Part IV - Ranking

Previous chapters did not pay much attention to the direction of social relations. In matters of cohesion or brokerage, it is more important to know that a relation exists than to know who initiates it. In this part, however, direction is central, especially asymmetry in social relations. Which choices are not reciprocated? Asymmetry in social relations points to social prestige and ranking.

9 Prestige

9.1 Introduction

In directed networks, people who receive many positive choices are considered to be prestigious. Prestige becomes salient especially if positive choices are not reciprocated, for instance, if everybody likes to play with the most popular girl or boy in a group but s/he does not play with all of them or, in the case of sentiments, if people tend to express positive sentiments towards prestigious persons but receive negative sentiments in return. In these cases, social prestige is connected to social power and the privilege of not having to reciprocate choices.

In social network analysis, prestige is conceptualized as a particular pattern of social relations. We will discuss techniques to calculate the **structural prestige** of a person from his or her social relations, notably sociometric choices. We will not compute a prestige score for an entire network.

Structural prestige is not identical to the concept of **social prestige** in the social sciences or in ordinary speech. For example, the medical profession is thought to be prestigious, but it is difficult to consider professions as a network in which many arcs point towards the medical profession. The prestige of an art museum may depend on the value and origins of its collection rather than on the number of art works it attracts ('receives') from other museums. However, social prestige is probably related to structural prestige. In community studies, for example, a physician is more often nominated in advice seeking relations than members of many other professions, and a prestigious art museum receives more attention from art critics than less prestigious ones.

In this chapter, we will compare the structural prestige of families within a network of visiting relations to their social prestige. As we will see, the two kinds of prestige are related but far from identical. Therefore, be careful not to equate structural prestige to social prestige. Instead, find out whether structural prestige scores on a social relation match indicators of social prestige which are measured by external variables. In a particular setting, which social relation is connected to social prestige?

9.2 Example

Let us have another look at the visiting relations between 75 *hacienda's* in San Juan Sur, a village in the Turrialba region of Costa Rica. In Chapter 3, we analyzed cohesive subgroups in this network. Now, we will concentrate on status and prestige. Members of the San Juan Sur community who were well informed about its population were asked to rank order all heads of households according to their importance to the community. Social status was computed for each family farm in this area as the average importance of its inhabitants and grouped into 14 classes. Prestige leaders were identified as those people who received more than 10 nominations within the community on the question: Which persons would you pick to represent you and the people of this place on a commission? These indicators of social prestige allow us to explore the relation between social and structural prestige.

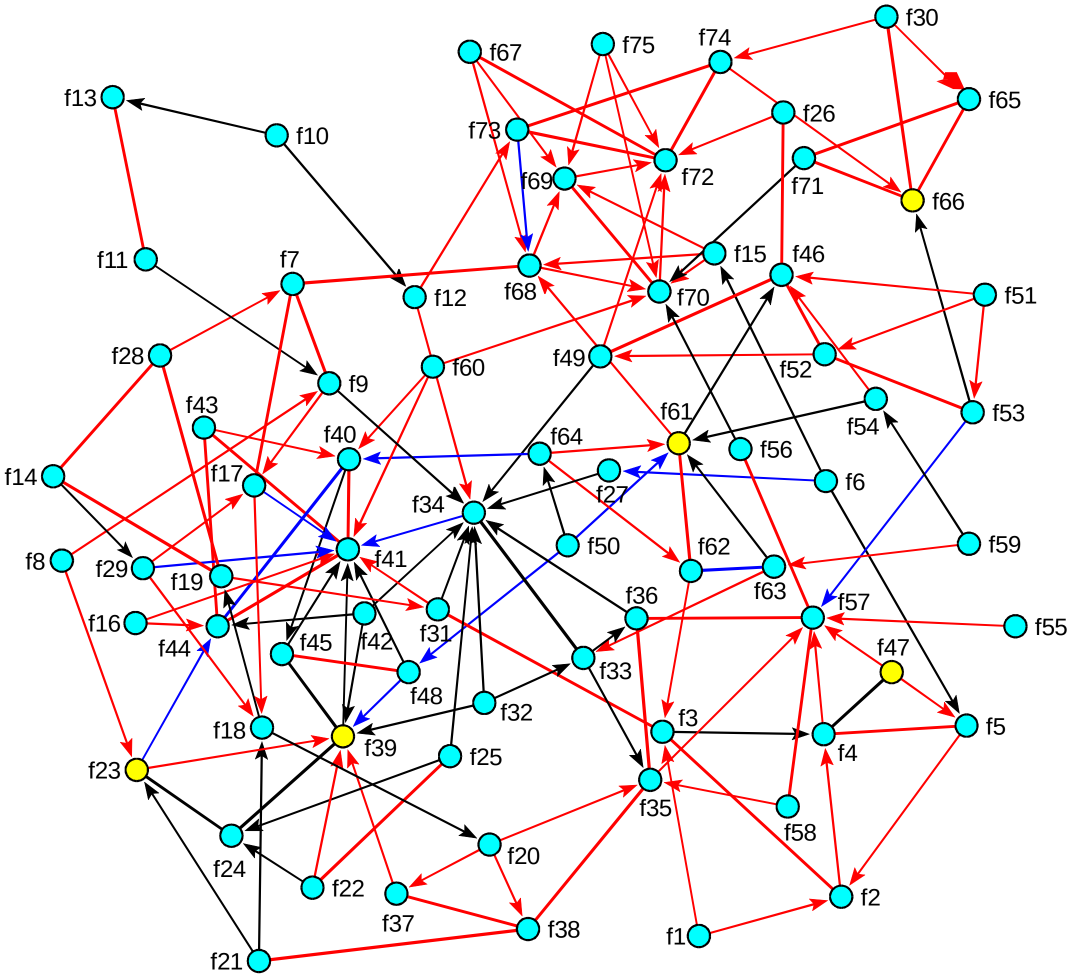


Figure 1 - Visiting relations and prestige leaders in San Juan Sur.

Figure 1 depicts the simple network of visiting relations. Note that bi-directional arcs are replaced by lines. The data are available in the project file `SanJuanSur2.paj`. Note that the color of arcs as well as line values indicate the type of relation between two families. Red arcs (line value 2) represent visits among kin, blue arcs (line value 3) are visits among families bound by god-parent or god-child ties ('church relations'), and other types of ties are drawn in black arcs (line value 1). The prestige leaders are yellow (see partition `SanJuanSur_leaders.clu`).

9.3 Popularity and indegree

At a first glance, this sociogram tells us little about the structural positions of prestige leaders. The leaders are dispersed over the network. They are situated in dense areas (e.g., family f39) as well as in the margins (families f23, f49, and f66). We will need some calculations to get a better view of structural prestige.

The **popularity** or indegree of a vertex is the number of arcs it receives in a directed network

The simplest measure of structural prestige is called popularity and it is measured by the number of choices a vertex receives: its indegree. Nominations on a positive social relation, e.g., working or living together, express prestige; more nominations indicate higher structural prestige, for example, in an election or a popularity poll. In this example, receiving more visitors indicates higher structural prestige. Note that the indegree of a vertex can only be determined in a directed network. In undirected networks, we can not measure prestige; instead, we use degree as a simple measure of centrality (see Chapter 6). In fact, several centrality measures are equal or similar to prestige measures applied to undirected networks.

Of course, a high indegree on a relation such as ‘lend money to someone’ does not reflect the popularity of an actor: it merely identifies someone who owes money to many persons. We should note, that indegree does reflect prestige if we **transpose** the arcs in such a network, that is, if we reverse the direction of arcs. In the transposed network, arcs represent the ‘owe money to’ relation and someone with a large indegree has lend money to many other people. Probably, this actor is quite rich compared to the other actors and more prestigious.

In the original network, the direction of the arcs depends on the way the researcher has defined the relation and worded the sociometric question. In the analysis, it is sometimes better to change the direction of the arcs. You are allowed to do this, because no information is lost in the transposed network: just transpose it again and you obtain the original network. It is interesting to note that several structural properties of a network do not change when the arcs are transposed, e.g., the components remain unchanged, and other properties are just swapped, e.g., outdegree becomes indegree and vice versa.

Application

Net>Partitions>Degree>Input

In Chapter 3, you have learned to compute the indegree of vertices in a directed network by means of the *Input* command in the *Net>Partitions>Degree* submenu. This command creates a new partition which can be displayed with *Info>Partition*. Table 1 shows the frequency count of the indegree of family farms in San Juan Sur. Thirteen families were not visited, so their indegree is zero. They have minimal structural prestige. Family number f41 is most popular because it is visited by twelve families (see entry of class 12 in Table 1). Note that the indegree is equal to the number of visiting families because there are no multiple arcs. In Figure 1, we can see the high number of visits which family f41 receives. This simple frequency tabulation summarizes the distribution of popularity better than the sociogram. The table shows that half of the families receive two visits at most. No more than a fifth of all families receive five or more visits (see column CumFreq%).

Table 1 - Indegree listing in Pajek.

Cluster	Freq	Freq%	CumFreq	CumFreq%	Representative
0	13	17.3333	13	17.3333	f1
1	17	22.6667	30	40.0000	f11
2	11	14.6667	41	54.6667	f13
3	15	20.0000	56	74.6667	f2
4	5	6.6667	61	81.3333	f3
5	6	8.0000	67	89.3333	f35
6	2	2.6667	69	92.0000	f44
7	1	1.3333	70	93.3333	f70
8	3	4.0000	73	97.3333	f39
10	1	1.3333	74	98.6667	f34
12	1	1.3333	75	100.0000	f41
Sum	75	100.0000			

File>Partition>Edit

How about the prestige leaders? May we conclude that families containing prestige leaders are structurally prestigious? Inspecting the sociogram or the indegree partition (use *File>Partition>Edit*), we note that prestige leaders f23, f39, f47, f61, and f66 have indegree 3, 8, 1, 5, and 5 respectively. All prestige leaders except for family f47 have indegree above average and three out of five families belong to the top 20 percent because they receive five or more visits. Therefore, we conclude that the prestige leaders are visited quite often, but there are other families which receive even more visits. Structural prestige measured by indegree does not distinguish between prestige leaders and other frequently visited families in this example.

9.4 Correlation

Does structural prestige indicated by indegree match social status as it was rated by experts within the community? In order to answer this question, we have to use standard statistical analysis to the results from our network analysis, which are the structural prestige scores. Since this is not a course in statistics, we will keep it as simple as possible. It is our primary goal to show that social network analysis and statistical analysis are two sets of techniques that work very well together in social research.

In statistics, the association between two phenomena is usually measured by **correlation coefficients**. Correlation coefficients range from 1 to -1. A positive coefficient indicates that a high score on one feature is associated with a high score on the other, e.g., high structural prestige occurs in families with high social status. A negative coefficient points towards a negative or inverse relation: a high score on one characteristic combines with a low score on the other, e.g., high structural prestige is found predominantly with low social status families. As a rule of thumb, we may say that there is no correlation if the absolute value of the coefficient is less than .05. If the absolute value of a coefficient is between .05 and .25, association is weak, coefficients from .25 to .60 (and from -.25 to -.60) indicate moderate association, and .60 to 1.00 (or -.60 to -1.00) is interpreted as strong association. Usually, a coefficient of 1 or -1 is said to display perfect association, but it is very unlikely that you will find this unless you correlate a characteristic to itself.

In Pajek, two kinds of correlation coefficients can be computed: Spearman's rank correlation and Pearson's correlation. **Spearman's rank correlation** determines whether the ranking of vertices on one characteristic, e.g., indegree, matches the ranking on another characteristic, e.g., status. The magnitude of differences between ranks is unimportant. Of course, both characteristics must have scores that can be ranked. Spearman's rank correlation is a robust measure of association provided that few cases have equal ranks.

Pearson's correlation coefficient uses the exact numerical scores on both characteristics. It assumes a linear association between two characteristics, which means that a unit increase in one characteristic will be associated with a fixed increase (or decrease) in the other. In our example, Pearson's correlation assumes that one extra indegree of structural prestige is accompanied by a fixed amount of additional social status, e.g., 2.4 extra points of social prestige.

Pearson is more precise and more sensitive than Spearman. This can be an advantage as well as a disadvantage. If a linear association exists among two features of vertices in the network, Pearson's correlation coefficient describes it more accurately than that of Spearman. However, the assumption that unit change on one feature is associated with a fixed change in another is very strict and often not met. For example, one extra indegree may involve substantial extra social status for families in the lower classes, whereas it may be associated with little extra status for families in the middle or upper classes. In this case, Pearson's coefficient underestimates the actual association, whereas that of Spearman does not. Therefore, it is important to use Pearson's correlation coefficient only if its results do not diverge too much from Spearman's coefficient. If results are very different, the data contain irregularities.

Application

In order to compute a correlation coefficient, we need two characteristics of each vertex in the network. As we learned in Chapter 2, features of vertices are stored in partitions and vectors. A partition contains integers, a vector is a list of numbers with decimals. Since Spearman's rank correlation coefficient only takes the rank order of scores into account, it operates on partitions. To calculate Spearman, you need two partitions. Hence, Spearman can be found in the *Partitions* menu. Pearson's correlation coefficient, however, uses the exact magnitude of scores. In Pajek, Pearson needs two vectors as input data and the procedure is to be found in the *Vectors* menu.

Partitions>First Partition , Second Partition

Partitions>Info >Spearman Rank

Social status scores are available as a partition (`SanJuanSur_status.clu`) which must be opened in Pajek in order to compute its correlation with the indegree partition. Load both partitions in the *Partitions* menu by selecting the partition in the drop list and clicking on the commands *First Partition* and *Second Partition* respectively. It does not matter which partition is first. When both partitions are selected, choose the command *Spearman Rank* on the *Info* submenu (see Figure 2) and Pajek will compute the rank correlation coefficient. In this case, it is .40, meaning that there is a moderate positive rank correlation between indegree and social status. Families with larger indegree tend to be families with higher status. Hence, we may conclude that structural prestige is moderately associated with status in this example.

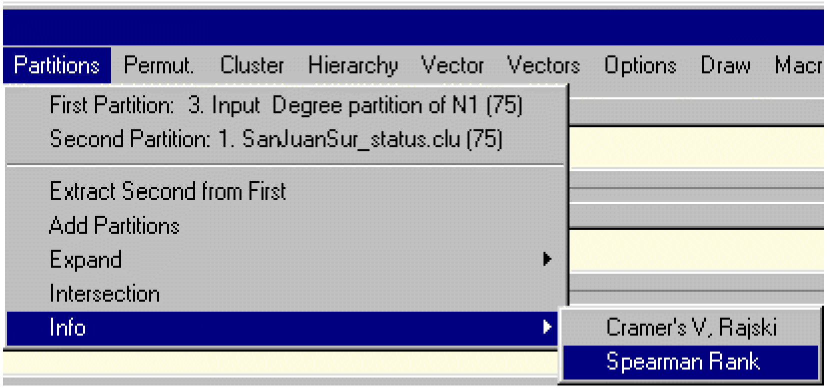


Figure 2 - Partitions menu in Pajek.

Vectors>First Vector, Second Vector

Vectors>Info

Pearson's correlation coefficient is computed in a similar way. Select a first and second vector in the *Vectors* menu and choose the *Info* submenu, which has no options other than Pearson's coefficient. In this example, you may use the normalized input degree vector created by the *Network>Partitions>Degree>Input* command but you have to translate the status partition (`SanJuanSur_status.clu`) to a vector first with the *Partition>Make Vector* command. Pearson's correlation coefficient is .35, which is slightly lower than Spearman's correlation indicating that the association is not linear. Using the rule of thumb specified above, however, we reach the same conclusion about the association between indegree and social status.

9.5 Domains

Popularity is a very restricted measure of prestige because it only takes direct choices into account. With popularity it does not matter whether choices are received from people who are not chosen themselves or from popular people. The overall structure of the network is disregarded.

Several efforts have been made to extend prestige to indirect choices. The first idea which comes to mind is to count all people by whom someone is nominated directly or indirectly, that is, without or with go-betweens. This is the input domain of an actor, which has been called the influence domain because structurally prestigious people are thought to influence people who regard them as their leaders. The larger the input domain of a person, the higher his or her structural prestige.

The **input domain** of a vertex in a directed network is the number or percentage of all other vertices which are connected by a path to this vertex.

Note that the output domain is more likely to reflect prestige in the case of a relation such as 'lend money to'. It is easy to define the output domain of a vertex and we guess that you understand that the output domain of a vertex is identical to the input domain of the vertex in the transposed network. In fact, we may distinguish between three domains: input domain, output domain, and (overall) domain, which is the union of the input and the output domain.

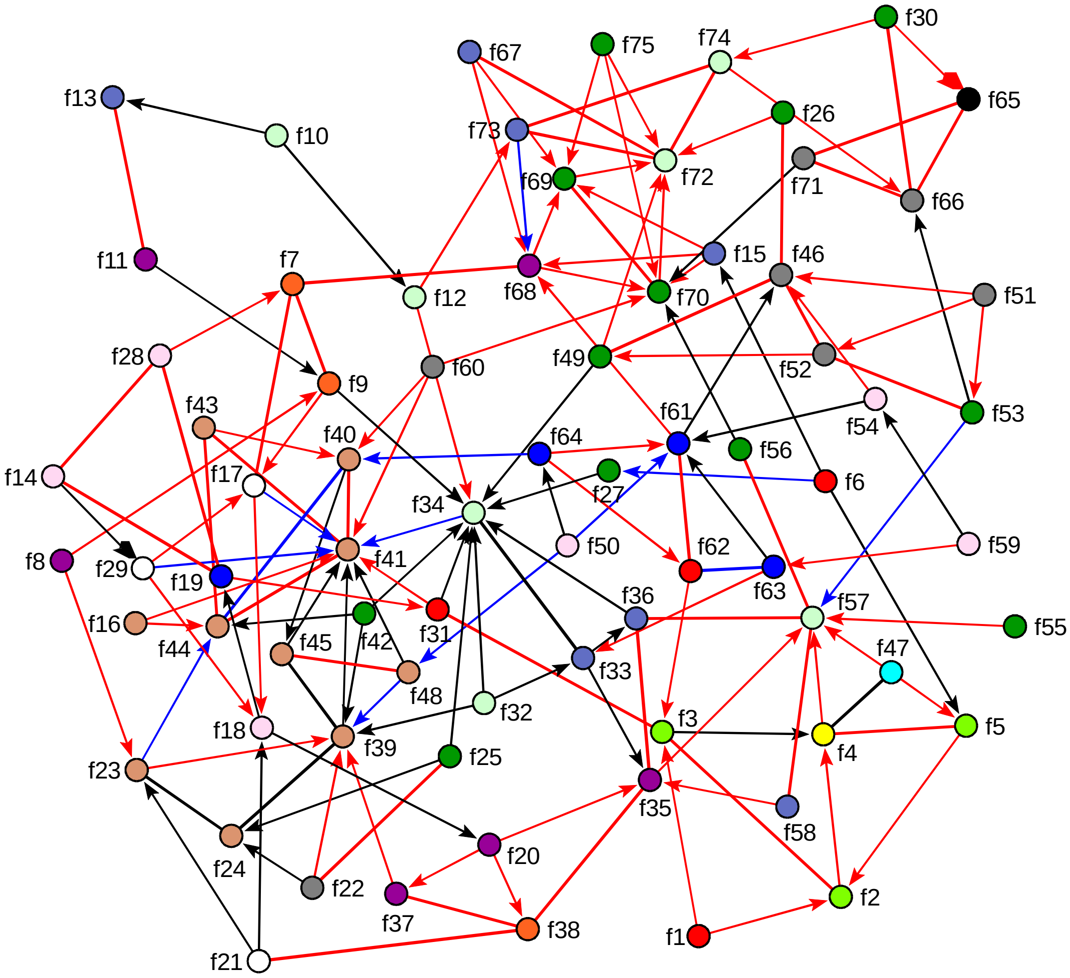


Figure 3 - Distances to family 47.

Let us have a look at the visiting relations network again to understand the concept of an input domain. In Figure 3, vertex color indicates the distance to family f47. Clearly, family f47 has zero distance to itself, hence its color is light blue. This family is visited by family f4 only: its distance to family f47 is one so its color is yellow. Families f2, f3, and f5 (lime green) visit family f4, so they can reach family f47 via family f4. The 'green' families are visited by three 'red' families (distance three) and, ultimately, family f47 can be reached by 64 out of the remaining 74 families (86%) in San Juan Sur. The input domain of family f47 equals 64 vertices or 86 percent.

The ten families outside the input domain of family f47 are colored brown in Figure 3. Note that the 'brown' families occupy the densest part of the network and that they include family f41 with highest indegree as well as prestige leaders f23 and f39. Family f47, which is also a prestige leader, turns out to be unreachable for the prestige leaders in the center of the network. This family was probably nominated as a representative by a relatively isolated group of families, including families f2, f3, f4, and f5. In this case, prestige leadership does not necessarily imply high overall social or structural prestige. The prestige leader is probably just a little more prestigious than the subgroup he or she represents.

In a well-connected network with many reciprocal relations, vertices are reachable from most other vertices. Hence, input domain scores display little variation. In this case, it is more interesting to capture the network structure in a prestige index which does not consider the entire input domain. For example, we can count the vertices which are able to reach a person in one or two steps: direct choices and indirect choices with one go-between. This **restricted input domain** only takes into account the direct popularity of the people by whom one is nominated. The input domain of family f47 restricted to two steps (distance two) is four (or 5%): one (yellow) family at distance one, and three (lime green) families at distance two.

Application

Net>k-Neighbours>Input

Info>Partition

The input domain of a particular vertex can be found with the *Net>k-Neighbours>Input* command, which is discussed in Chapter 6. In the first dialog box, enter the number or label of a vertex, e.g., f47, and in the second dialog box accept the default value (zero) to compute all distances. Then, the command creates a partition specifying the distances of all vertices to the selected vertex. From a frequency tabulation, created with the *Info>Partition* command (Table 2), you can calculate the number of vertices (CumFreq) in the input domain of the selected vertex at a particular maximum distance, for instance, the input domain at maximum distance two contains four vertices: the five vertices at maximum distance two minus family f47 itself. The entry identified by 'Unknown' in the table shows the number of vertices which are not connected by a path to the selected vertex: they do not belong to its input domain. In our example, 10 out of 74 vertices (do not count the selected vertex itself!) are outside the input domain of family f47, which is 14 percent; the remaining 86 percent of the vertices are inside its input domain. Note that you can not find these percentages in the table because all percentages there include family f47.

Table 2 - Input domain of family f47.

Cluster	Freq	Freq%	Valid%	CumFreq	CumFreq%	CumValid%	Representative
0	1	1.3333	1.5385	1	1.3333	1.5385	f47
1	1	1.3333	1.5385	2	2.6667	3.0769	f4
2	3	4.0000	4.6154	5	6.6667	7.6923	f2
3	4	5.3333	6.1538	9	12.0000	13.8462	f1
4	4	5.3333	6.1538	13	17.3333	20.0000	f19
5	6	8.0000	9.2308	19	25.3333	29.2308	f14
6	3	4.0000	4.6154	22	29.3333	33.8462	f17
7	3	4.0000	4.6154	25	33.3333	38.4615	f7
8	6	8.0000	9.2308	31	41.3333	47.6923	f8
9	7	9.3333	10.7692	38	50.6667	58.4615	f13
10	7	9.3333	10.7692	45	60.0000	69.2308	f10
11	12	16.0000	18.4615	57	76.0000	87.6923	f25
12	7	9.3333	10.7692	64	85.3333	98.4615	f22
13	1	1.3333	1.5385	65	86.6667	100.0000	f65
Sum	65	86.6667	100.0000				
Unknown	10	13.3333					
Total	75	100.0000					

Net>Partitions>Domain >Input

It is quite cumbersome to repeat this command for each vertex in a network, so Pajek contains a command which calculates the size of the input domains of all vertices in one go: *Net>Partitions>Domain>Input*. Use the command *Input* to restrict the analysis to incoming arcs only. A dialog box, which is similar to the one displayed by *k-Neighbours*, allows you to specify a maximum distance for the input domain.

The *Domain>Input* command produces three new data objects: one partition and two vectors. The partition specifies the number of vertices within the input domain of each vertex. The vector labeled 'Normalized Size of input domain' lists the size of input domains as a proportion of all vertices (minus the vertex itself) and the second vector gives the average distance to a vertex from all vertices in its input domain. Of course, it is impossible to compute average distance in the case of a vertex with an empty input domain, that is, a vertex which is not chosen at all. In this case, average distance is set to 999998, which represents infinity. The average distances vector is very useful, as we will see in the following section.

Table 3 - Size of input domains in the visiting relations network.

Class	Freq	Freq%	CumFreq	CumFreq%	Representative
0	13	17.3333	13	17.3333	f1
1	7	9.3333	20	26.6667	f12
2	2	2.6667	22	29.3333	f11
6	3	4.0000	25	33.3333	f61
12	5	6.6667	30	40.0000	f26
64	36	48.0000	66	88.0000	f2
74	9	12.0000	75	100.0000	f23
Sum	75	100.0000			

Table 3 lists the size of input domains in the visiting relations network. Nine families have maximal input domains; they are reachable from all 74 other vertices. Prestige leaders f23 and f39 are among them. The third prestige leader, family f47, is situated in the class of families with an input domain of size 64 as we noted before. Inspecting the partition with influence sizes with the *File>Partition>Edit* procedure, we find that prestige leader f66 belongs to this class too. Family f61 is the only prestige leader which has a small input domain of size 6. We may conclude that most prestige leaders have large input domains, but many families with equally large input domains are not prestige leaders.

The rank correlation between structural prestige measured as the size of the input domain and social prestige indicated by social status scores can easily be computed (see Section 9.4). Spearman's rank correlation coefficient is .36, which is a little less than the rank correlation between popularity (indegree) and social status. Nevertheless, it points to positive, moderate association between input domain and social status: larger input domains occur among families of higher social status.

9.6 Proximity prestige

In the previous section, we noted that the input domain of a vertex is not a perfect measure of prestige. In a well-connected network, the input domain of a vertex often contains all or almost all other vertices, so it does not distinguish very well between vertices. In this case, we proposed to limit the input domain to direct neighbors or to neighbors at maximum distance two on the assumption that nominations by close neighbors are more important than nominations by distant neighbors. An indirect choice contributes less to prestige if it is mediated by a longer chain of intermediaries.

Of course, the choice of a maximum distance from neighbors within a restricted input domain is quite arbitrary. The concept of proximity prestige overcomes this problem. This index of prestige considers all vertices within the input domain of a vertex but it attaches more importance to a nomination if it is expressed by a closer neighbor. In other words, a nomination by a close neighbor contributes more to the proximity prestige of an actor than a nomination by a distant neighbor, but many 'distant nominations' may contribute as much as one 'close nomination'.

In order to allow direct choices to contribute more to the prestige of a vertex than indirect choices, proximity prestige weighs each choice by its path distance to the vertex. A higher distance yields a lower contribution to the proximity prestige of a vertex, but each choice contributes something. In the calculation of proximity prestige, this is accomplished by dividing the input domain of a vertex (expressed as a proportion of all vertices which may be part of the input domain) by the average distance from all vertices in the input domain. A larger input domain (larger numerator) yields a higher proximity prestige because more vertices are choosing an actor directly or indirectly. In addition, a smaller average distance (smaller denominator) yields a higher proximity prestige score because there are more nominations by close neighbors.

Maximum proximity prestige is achieved if a vertex is directly chosen by all other vertices. This is the case, for example, in a star network in which all choices are directed to the central vertex. Then, the proportion of vertices in the input domain is one and the mean distance from these vertices is one, so proximity prestige is one divided by one. Vertices without input domain get minimum proximity prestige by definition, which is zero.

The **proximity prestige** of a vertex is the proportion of all vertices (except itself) in its input domain divided by the mean distance from all vertices in its input domain.

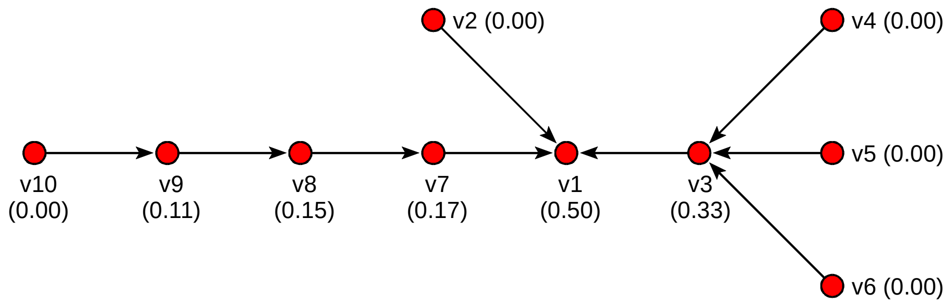


Figure 4 - Proximity prestige in a small network.

In Figure 4, all vertices at the extremes of the network (v2, v4, v5, v6, and v10) have empty input domains, hence they have proximity score zero. The input domain of vertex v9 contains vertex v10 only, so its size is one out of nine (.11). Average distance within the input domain of vertex v9 is one, so the proximity prestige of vertex 9 is .11 divided by one. You can see that the proximity prestige of vertices increase if they have a longer ‘tail’ from vertex v10 to v1. Vertex v1 has a maximal input domain, since it can be reached by all nine vertices (a proportion of 1.00). Average distance is 2.0, so proximity prestige amounts to 1.00 divided by 2.0, which is .5.

Application

Net>Partitions>Domain >Input

In the previous section, we have learned how to compute the size of input domains and average distance from all vertices within the input domain (command *Net>Partitions>Domain>Input*). Thus, we obtain the two vectors which we need to compute proximity prestige: the size of the input domain expressed as a proportion (‘Normalized Size of input domain’) and the average distance from vertices within the input domain (‘Average distance from input domain’).

Vectors >Divide First by Second

To calculate proximity prestige, we just divide the input domain size by the average distance. Select the vector with the normalized size of the input domain in the vectors drop list and click command *First Vector* in the *Vectors* menu to use it as the numerator in the division operation. Select the vector with average distances as the second vector in a similar manner and click on the command

Divide First by Second in the *Vectors* menu. This will create a new vector containing the proximity prestige scores of all vertices. Inspect them with the command *Info>Vector* or browse with *File>Vector>Edit*. Proximity prestige scores must range from zero to one. If they do not, you probably specified the wrong vectors in the *Vectors* menu.

In the network of visiting relations at San Juan Sur, proximity prestige ranges from 0.0 to .33. Family f41 has the highest proximity prestige. Three out of five prestige leaders have a proximity prestige above average (.12). However, the proximity prestige of families f47 (.11) and f61 (.07) is below average. We must conclude that prestige leaders are not characterized by high proximity prestige. In Section 9.5, we already noted that family f47 occupies a special position in the network. Inspection of the average distances confirms this: family f47 has the largest average distance (8.03). This family is difficult to reach in the network.

Vector>Make Partition >by Intervals >First Threshold and Step

Finally, let us see whether proximity prestige is associated with social status in San Juan Sur. Before we can compute Spearman's rank correlation, we must turn the vector with proximity prestige scores into a partition. As you have learned in Chapter 2, this can be done in several ways. In this case, the easiest way to convert the vector into a partition is to create classes of equal width with the procedure *Vector>Make Partition>by Intervals>First Threshold and Step*. Specify 0.01 as the first threshold (the upper limit of the lowest class) and enter this number also as the step (the class width) to obtain a partition with classes between 0 and 100.

Partitions>Info >Spearman Rank

The newly created partition with proximity prestige scores can be correlated to the existing partition with social status (`SanJuanSur_status.clu`) in the manner described in Section 9.4. Spearman's correlation coefficient is .26, indicating a low or moderate association between proximity prestige within the network and social status rated separately by members of the community. In this example, social status is related less to proximity prestige than to popularity (indegree), which has a rank correlation of .40 (see Section 9.4).

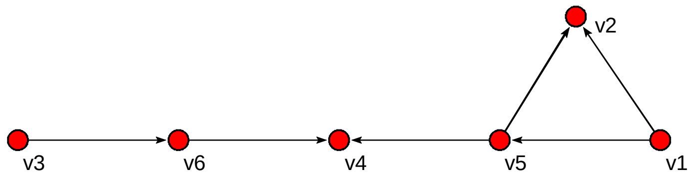
9.7 Summary

This is the first chapter of the book to deal with asymmetry in social networks. We present the simplest way to take the direction of relations into account, which is to pay attention to incoming relations only. Structural indices which do this are called measures of prestige. Actors who receive a lot of choices are popular provided, of course, that the choices express a positive social relation. Popularity, which is measured as the indegree of a vertex, is the first index of prestige we discuss. More advanced measures of prestige also take indirect choices into account. We present two advanced measures: the input domain of a vertex and proximity prestige.

It is important to distinguish between structural prestige and social prestige. The indices introduced in this chapter assess structural prestige, that is, a pattern of relations which network analysts call prestige. They are called prestige because actors in prestigious network positions often enjoy high social prestige. However, the example which we use shows that structural prestige and social prestige do not match perfectly; we find moderate association only. We use correlation coefficients to establish the strength of the association between structural prestige and social status scores measured independently from the network. This is an example of an important research strategy, namely using structural indices such as prestige scores in statistical analysis.

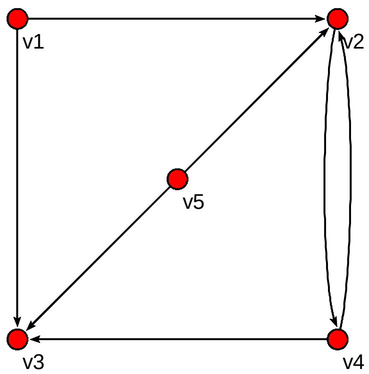
9.8 Exercises

1 In the network below, which vertex or vertices have the highest proximity prestige?



a vertex v2
b vertex v4
c vertices v2 and v4 have equal proximity prestige
d it is impossible to tell from this sociogram

2 In the network below, which vertex or vertices have minimal structural prestige?



a vertex v5
b vertices v5 and v1
c all vertices have equal structural prestige
d it is impossible to tell from this sociogram

3 In the network presented in exercise 2, which vertex or vertices have the highest social prestige?

a vertex v3
b vertices v2 and v3
c vertices v2, v3, and v4 have approximately equal prestige
d it is impossible to tell from this sociogram

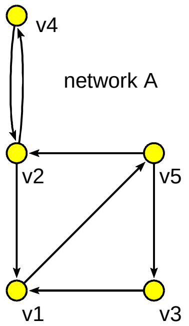
4 What is the correct interpretation of a correlation coefficient of size -.20?

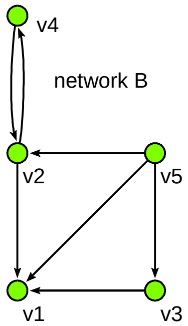
a weak negative association
b medium association

- c medium positive association
- d no association

5 Which prestige indices take indirect relations into account?
- a proximity prestige only
- b proximity prestige and input domain
- c proximity prestige and popularity
- d input domain and popularity

6 For which of the networks below it is useless to compute structural prestige as the full input domain of a vertex?





- a A
- b B
- c A and B
- d A nor B

7 In an undirected network, is proximity prestige equal to closeness centrality?
- a yes, because proximity prestige is equal to mean distance to all other vertices in an undirected network
- b yes, because a prestige index applied to an undirected network is equal to a centrality index
- c no, unless the network is connected
- d no, because the calculation of an input domain is meaningless in an undirected network

9.9 Assignment

In Chapter 8, we learned that the diffusion of innovations resembles a contagion process because social contacts are needed to persuade people to adopt innovations. It is hypothesized, therefore, that prestige is associated with adoption time: less prestigious actors adopt later because they wait for more prestigious opinion leaders to adopt first.

The file `Galesburg_discussion.net` contains a network of discussion relations between 31 physicians in Galesburg (Illinois) in the 1950s. The researchers asked each physician to name three doctors with whom they would choose to discuss medical matters. For 17 physicians, the date that they first prescribed a new drug (gammanym) was recorded. The partition `Galesburg_adoptiontime.clu` measures the adoption time as the number of months since the introduction of the drug. Note that the adoption time is unknown (code 999999) for 14 physicians. For most of them, the new drug was not relevant.

Investigate whether adoption time is associated with the prestige rather than the centrality of doctors in the discussion network. Compute the indices of prestige presented in this chapter (indegree, restricted input domain with maximum distance two, and proximity prestige) as well as the corresponding centrality measures in the undirected network. Use rank correlation and note that adoption time is higher when a doctor adopts later.

Another hypothesis states that friendship relations are more important than discussion relations for the adoption of a new drug because it is easier to persuade friends than people you only know professionally. Physicians with many direct or indirect friends would adopt sooner than physicians with less central positions in the friendship network. The file `Galesburg_friends.net` contains the friendship network between the doctors. Is the adoption time of the new drug related to prestige or centrality in the friendship network rather than in the discussion network?

9.10 Further reading

- The data on San Juan Sur are taken from Charles P. Loomis, Julio O. Morales, Roy A. Clifford, and Olen E. Leonard, *Turrialba. Social Systems and the Introduction of Change* (Glencoe (Ill.): The Free Press, 1953). Consult this book to learn more about the research project.
- The medical innovation project is taken from James S. Coleman, Elihu Katz, and Herbert Menzel, *Medical Innovation. A Diffusion Study* (Indianapolis: Bobbs-Merrill, 1966). David Knoke and Ronald S. Burt reanalyzed the data in a chapter on prominence in R. S. Burt and M. J. Minor (Eds.), *Applied Network Analysis. A Methodological Introduction* (Beverly Hills: Sage Publications, 1983, pp. 195-222). This paper contains the basic argument to distinguish between directed prestige and undirected centrality.
- To learn more about prestige indices, use Chapter 5 of Stanley Wasserman and Katherine Faust, *Social Network Analysis. Methods and Applications* (Cambridge: Cambridge University Press, 1994).

9.11 Answers

1 Answer b is correct. Vertex v4 has the largest input domain, which contains all vertices except for v2, and paths to v4 are quite short for all vertices in its input domain. A large numerator (size of input domain) and a relatively small denominator (average distance) yields high proximity prestige. Vertex v2 is second best because it has an even lower average distance from vertices in its input domain (it is directly chosen by both vertices in its input domain), but its input domain is a lot smaller.

2 Answer b is correct. Vertices v1 and v5 are not chosen. They have zero indegree, hence no input domain and minimum proximity prestige. Both vertices have minimal scores on all prestige indices presented in this chapter.

3 Answer d is correct. Since structural prestige is not necessarily equal to social prestige, we cannot tell which actor has the most social prestige from this sociogram.

4 Answer a is correct. According to the rule of thumb presented, the association is weak. The sign of the coefficient tells us that there is a negative or inverse relation between the two characteristics.

5 Answer b is correct. Input domain counts direct choices as well as indirect choices of vertices at distance two or higher, so it definitely takes indirect relations into account. Proximity prestige uses the input domain, so it uses indirect relations too. Popularity is just the indegree of a vertex, the direct choices it receives. Clearly, it does not use indirect relations.

6 Answer a is correct. In network A, each vertex is reachable for all other vertices, so each vertex has an input domain of size 4. In other words, the network is one strong component. Because there are no differences between vertices with respect to the size of their input domain, this prestige index is useless. Network B differs from network A in the relation between v1 and v5. Changing the direction of this relation ‘breaks’ the strong component: v5 is no longer reachable for any other vertex and as a consequence v1 can no longer reach v2 and v3. Now, the size of the input domain varies between vertices; it is a useful prestige index.

7 Answer c is correct. Proximity prestige is calculated as the average distance from all vertices in the input domain of a vertex divided by the size of the input domain as a proportion of the maximum number of vertices it can hold. Closeness centrality is similar to the numerator of this fraction: average distance. In an undirected network, proximity prestige is equal to closeness prestige only if the denominator of the fraction is 1, which means that all other vertices are part of the input domain. This is the case if the network is connected because each vertex is reachable from all other vertices in a connected undirected network.