Social Network Basics #8
Centrality & Prestige

When we use centrality to measure prestige, we typically do so with directed (rather than undirected) network data. Thus, for this exercise, we use directed network data collected by David Krackhardt from the managers of a high-tech company that manufactured on the West Coast. At the time, the firm had existed for 10 years, produced high-tech machinery for other companies, and employed approximately 100 people of whom 21 were managers.[1] The managers are the actors in the dataset. Krackhardt gave each manager a roster of the names of the other managers and asked to check the other managers to whom they would go for advice at work and with whom they were friends.[2] This is one of the standard datasets that comes with UCINET and other SNA software packages.

Part I – Centrality and Prestige in UCINET and NetDraw

Data>Unpack

1. Because Krackhardt's data is recorded in three stacked 21x21 matrices, we need to first extract the advice and friendship matrices with UCINET's *Data>Unpack* command, which brings up the Unpack dialog box (see Figure 1). Click on the "…" button to the right of the "Input dataset" drop box and select the Krack-High-Tec dataset. Next, select the "Advice" and "Friendship" networks. The "Unpack" dialog box should look similar to Figure 1. Click OK and you should have two "new" (i.e., unpacked) networks in your folder. You may want to also select the "Reports to" network in order to visualize it. It looks remarkably different from the other two and is a good illustration of how "official" lines of authority differ from informal ones.

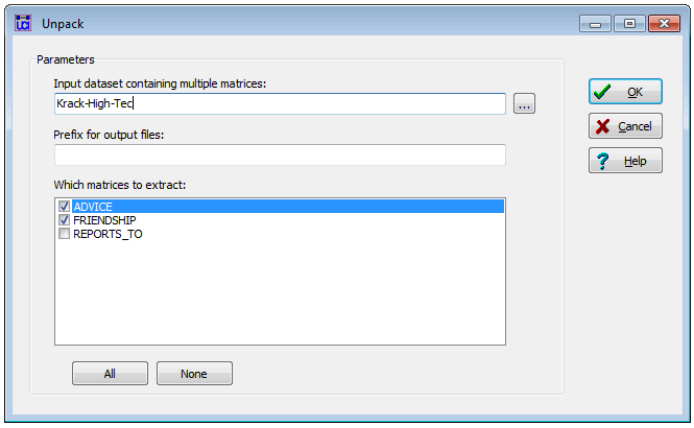

Figure 1: UCINET's Unpack Dialog Box

2. UCINET computes several different types of centrality, all of which are found under the *Network* menu. Since in this exercise we are interested in measures of prestige, we will first explore *indegree centrality*, which is the simplest measure of prestige or popularity. It assumes that nominations/choices on a positive social

[1] David Krackhardt. 1987. "Cognitive Social Structures." *Social Networks* 9:109-134.
[2] Krackhardt also collected data on "who reports to whom" for all 21 managers ("Reports to").

relation (e.g., working or living together) express prestige. For example, more nominations indicate higher prestige in an election or a popularity poll. It is important to emphasize that we can only calculate an actor's indegree or outdegree with a directed network. Put differently, with undirected networks, we cannot measure prestige.

Network >Centrality and Power >_Degree (legacy)

3. To calculate indegree centrality in UCINET, first select the *Degree (legacy)* option found under the *Network>Centrality and Power* submenu. This will bring up UCINET's Degree Centrality dialogue box. Make sure that the "Advice" network is highlighted in the "Input dataset" option box. ***Next, select the "No" option in the "Treat data as symmetric" option box***, which tells UCINET to calculate both the indegree and outdegree centrality scores of each actor in a network. (When you select "Yes," it calculates and then selects the higher of the indegree or outdegree centrality of each actor). ***In terms of the advice network, which actor enjoys the highest level of prestige?*** Repeat step #3 using the Friendship network. ***In terms of the friendship network, which actor enjoys the highest level of prestige?***

4. Although the hubs and authorities algorithm was originally designed to rank web pages, it offers another approach to estimating prestige. A good *hub* is defined as one that points to many good *authorities*, and a good *authority* is one that is pointed to by many good *hubs*. In terms of prestige, it allows analysts to not only take into account the number of ties an actor receives (i.e., an authority) but to also weight those ties by whether the actor that is sending the tie (i.e., a hub) also sends ties to other prestigious actors. This should sound a lot like eigenvector centrality, and, in fact, when it's used with undirected networks, it is. Calculating hubs and authority centrality scores is straightforward in UCINET. All you do is select the *Hubs & Authorities* command under the *Network>Centrality and Power* submenu. Select the "Advice" network as your input network and click OK. ***In terms of the advice network, which actor enjoys the highest level of authority? How about the friendship network?***

Network >Centrality and Power >Hubs & Authorities

5. Indegree centrality is a somewhat limited measure of prestige because it only considers direct choices. Consequently, social network analysts have developed alternative measures that take into account direct and indirect choices. Reach centrality is one such measure. It counts the number of nodes each node can reach in k steps or less, and then weights this count by how far each node is from every other node (in terms of path distance). When k =1, reach centrality equals degree centrality, which means that if we are analyzing a directed network, the resulting scores will equal indegree and outdegree centrality. When $k = n - 1$ (i.e., the size of the network less one, which is its maximum value), a particular node's score equals the total number of actors that it can reach, weighted by how far, on

average they are. Thus, a maximum reach score is obtained when every other actor is one-step away, but it becomes less as actors are two-steps (weight = ½), three steps (weight = 1/3), four-steps (weight = ¼), and so on. Normalized reach score is calculated by dividing each score by the largest observed value. Reach is very similar to Pajek's proximity prestige measure (see below).

Network >Centrality and Power >Reach centrality

6. Calculating reach centrality in UCINET is straightforward. The command is located in the Centrality and Power submenu. It first calculates the weighted distance reach centrality of each node, which is the sum of the number of actors that can be reached in *k* steps divided by *k* (and is the same as ARD closeness centrality plus one). Figure 2 presents a portion of UCINET's output with regards to Krackhardt's Advice network. Here again, manager "2" ranks first in terms of this measure of prestige with "18" and "21" tied for second.

ucinetlog2.txt - Notepad

File Edit Format View Help

```
REACH CENTRALITY
-------------------------------------------------------------------------------

Input dataset:                          ADVICE (D:\Dropbox\Courses\SNA\Dark Networks\
Output dataset:                         ActorByDistanceReach (D:\Dropbox\Courses\SNA\

Note: Data not symmetric, therefore separate in-closeness & out-closeness computed.


Reach Centrality

                1         2         3         4
       OutdwReac IndwReach nOutdwRea nIndwReac
       --------- --------- --------- ---------
   15     21.000    12.167     1.000     0.579
   18     19.500    18.500     0.929     0.881
    3     18.500    13.333     0.881     0.635
    5     18.500    12.667     0.881     0.603
   10     18.000    14.833     0.857     0.706
    9     17.500    12.167     0.833     0.579
    4     17.000    15.000     0.810     0.714
   20     17.000    14.833     0.810     0.706
   21     16.500    18.500     0.786     0.881
   19     16.500    12.167     0.786     0.579
    8     15.000    16.000     0.714     0.762
    7     15.000    17.500     0.714     0.833
    1     14.000    17.000     0.667     0.810
   13     14.000    12.167     0.667     0.579
   14     13.000    16.000     0.619     0.762
   16     12.833    14.500     0.611     0.690
   17     12.667    15.500     0.603     0.738
    2     11.167    20.000     0.532     0.952
   11     11.167    16.333     0.532     0.778
   12     10.667    14.333     0.508     0.683
    6     10.000    16.000     0.476     0.762


Summary Statistics

                    1       2       3       4
                OutdwRe IndwRea nOutdwR nIndwRe
                ------- ------- ------- -------
   1      Mean   15.21   15.21    0.72    0.72
   2   Std Dev    3.06    2.23    0.15    0.11
   3       Sum  319.50  319.50   15.21   15.21
   4  Variance    9.37    5.00    0.02    0.01
   5       SSQ 5057.81 4965.86   11.47   11.26
   6     MCSSQ  196.84  104.90    0.45    0.24
   7  Euc Norm   71.12   70.47    3.39    3.36
   8   Minimum   10.00   12.17    0.48    0.58
   9   Maximum   21.00   20.00    1.00    0.95
  10  N of Obs   21.00   21.00   21.00   21.00
```

Figure 2: UCINET's Reach Centrality Output

[NetDraw] File>Open>Ucinet dataset >Network

File>Open>Ucinet dataset >Attribute data

Properties>Nodes>Size >Attribute-based

7. Open NetDraw and read in the ***original*** `Krack-High-Tech` dataset, using the *File>Open>Ucinet dataset>Network* command. One way to visualize the various measures of prestige is to read the files you created in UCINET as attribute files, using the *File>Open>Ucinet dataset>Attribute data* command (see Figure 3 below where I read in the indegree and outdegree centrality file; I could have just as easily read in the hubs and authorities scores or the reach centrality scores) and then vary the size of the nodes, using NetDraw's *Properties>Nodes>Symbols>Size>Attribute-based* command. This will produce a network map similar to Figure 5 (next page). **Note**: if you vary node size based on the indegree centrality calculated using the Advice network, then make sure that the Advice network is selected in the far-right panel (see Figure 5).

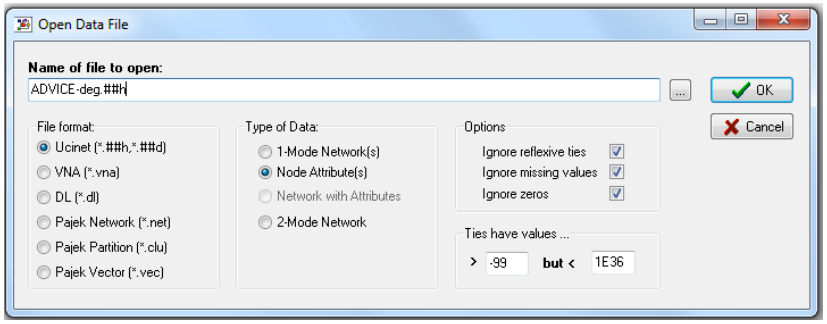

Figure 3: Import Indegree and Outdegree Attribute Data

Analysis >Centrality measures

Properties>Nodes>Size >Attribute-based

8. Another way is to use the centrality command implemented in NetDraw. Here, after loading the network data, choose which network you want to visualize and calculate indegree centrality (see Figure 5, next page). Then select the *Analysis>Centrality measures* command, which brings up the following dialog box (Figure 4). Select degree centrality and indicate that you want the "Directed versions" of the calculation. Then click "OK" and vary the node size based on indegree centrality, using the *Properties>Nodes>Symbols>Size>Attribute-based* command, which should produce a network map similar to Figure 5 (next page).

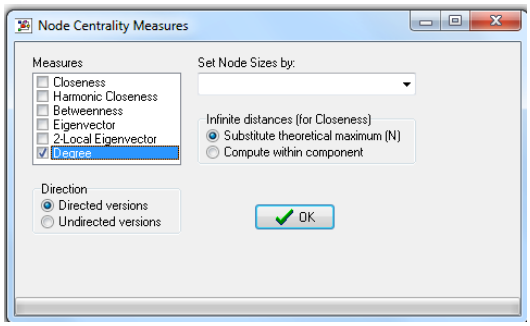

Figure 4: Node Centrality Measures Dialog Box

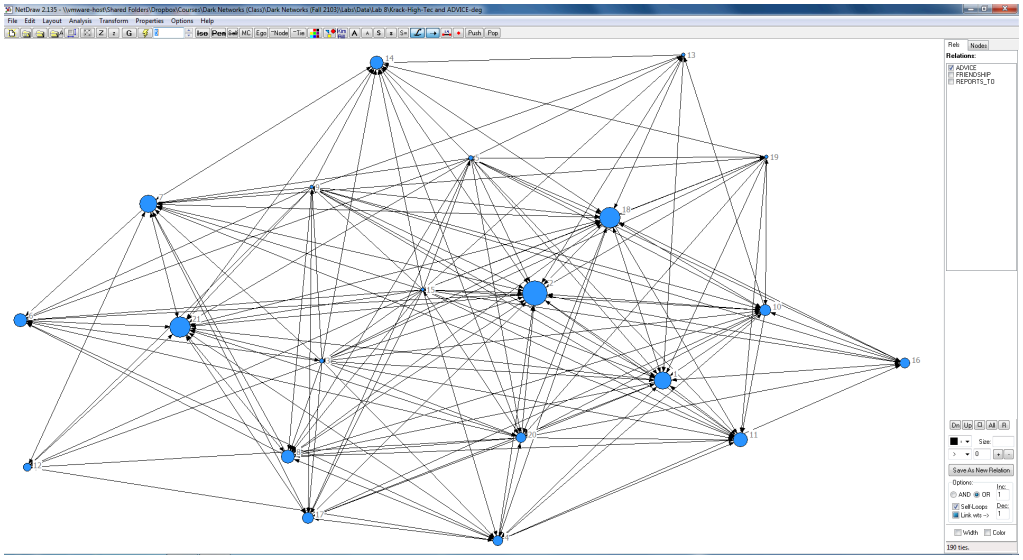

Figure 5: Advice Network with Node Size Varying by Indegree Centrality

Part II – Centrality and Prestige in Pajek

File>Network>Read

Network>Create Partition >Degree>Input

Partition>Info

1. Read the Advice network into Pajek with the *File>Network>Read* command. Next, select the *Network>Create Partition>Degree>Input* command, which creates a new partition, which we can info with Pajek's *Partition>Info* command, which calls up a dialog box similar to Figure 6. Type "1" in the top box and 21 in the bottom (so you can see the scores of all managers), and click OK.

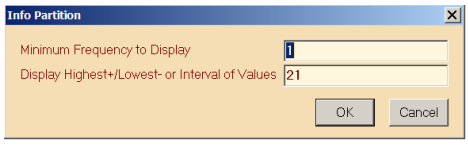

Figure 6: Pajek's *Partition>Info* Dialog Box

2. This should create a report that looks something like Figure 7. The upper part ("highest clusters values") ranks the actors in terms of indegree centrality while the lower ("frequency distribution of cluster numbers") indicates the number of actors in each cluster/class (cluster number equals indegree centrality). Thus, actor 2 has an indegree centrality of 18 and four actors have an indegree centrality of 4, while only one actor has an indegree centrality of 18. ***Which actor has an indegree centrality of 5? How many actors have an indegree centrality of 7?*** Close the report window and select the *File>Partition>Edit* command. This displays the indegree centrality measures of each actor.

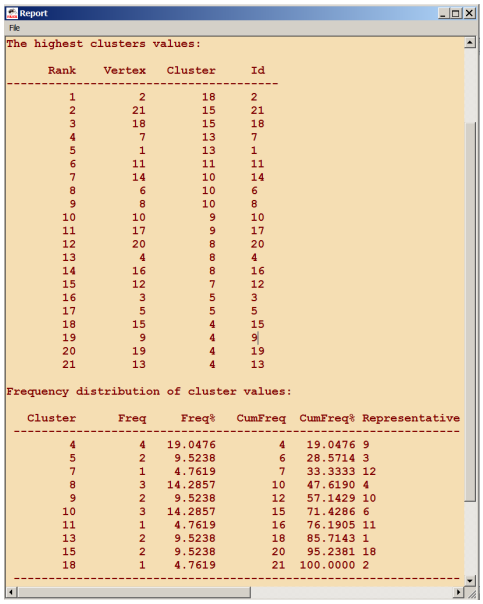

Figure 7: Pajek's Report Window: Partition Information

Network>Create Vector >Centrality >Hubs-Authorities

Vector>Info

3. We learned how Pajek estimates Hub and Authority centrality scores in the previous lab with Pajek's *Network>Create Vector>Centrality>Hubs-Authorities* command. Using Pajek's *Vector>Info* feature, compare the Authority scores for the Advice network estimated by Pajek with those estimated by UCINET (remember that in the dialog box indicate that you want to see the scores for all 21 managers).

4. As noted above indegree centrality is a somewhat limited measure of prestige because it only considers direct choices. ***Input domain*** is a measure that counts all people by whom someone is nominated whether directly or indirectly. However, input domain is an imperfect prestige measure because in a well-connected network, such as Krackhardt's advice and friendship networks, an actor's input domain often contains all or almost all other actors, so it does not always distinguish between actors very well.

Network>Create Vector>Centrality >Proximity Prestige>Input

5. An alternative approach is to weight each nomination by its path distance to the actor. This is what the ***proximity prestige*** does. With proximity prestige, a close neighbor's nomination contributes more to an actor's proximity prestige than does a distant neighbor's nomination. Pajek now calculates both proximity prestige and input domain together with its *Network>Create Vector>Centrality> Proximity Prestige>Input* command (accept Pajek's default—Distance (0 - No limit). This creates a new partition (Input Domain) and three new vectors (*Normalized Input Domain*,[3] *Average Distance from Input Domain*,[4] and *Input*

[3] Normalized input domain is the size of an actor's input domain as a proportion of all actors (minus the focal actor itself)

Proximity Prestige). Input proximity prestige is calculated by dividing the normalized input domain by average distance. In fact, earlier versions of Pajek required users to divide the two vectors themselves. Now Pajek does it for us.

Partition>Info

Vector>Info

6. Use the *Partition>Info* command to examine the input domain partition (Note: be sure that the new partition is showing in the partition drop list), and the *Vector>Info* command to examine the input prestige proximity scores. ***Which actors have the highest input domain? Which actors have the highest input proximity prestige scores?***

7. As noted above, input domain scores in a well-connected network often display little or no variation, and one solution is to calculate input proximity prestige. Another is to limit the input domain to direct neighbors or to neighbors at a maximum distance. For example, Christakis and Fowler argue that a person's influence ceases to have a noticeable effect on others beyond three degrees of separation.[5] In other words, Christakis and Fowler might argue that the influence that one has in terms of their advice doesn't extend beyond a path length of three (i.e., an advisee of an advisee of an advisee). Thus, repeat the above steps for computing proximity prestige, except this time change Pajek's distance default to 3 (see Figure 8), and then examine the input domain partition. ***Is there more variation in input domain than before?*** Try it again with the distance set to 2. ***Now, which actors have the highest input domain? What input domain score do you think you will get if you set the distance to 1?***

Network>Create Vector >Proximity Prestige>Input

Partition>Info

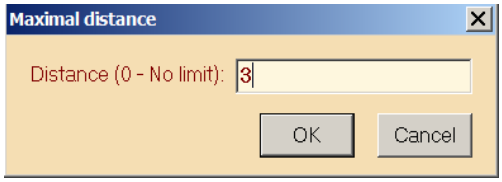

Figure 8: Proximity Prestige Dialog Box

[4] Average distance gives the average distance to an actor from all actors in its input domain. It is impossible to compute average distance in the case of an actor with an empty input domain. In this case, average distance is set to 999998, which represents infinity.

[5] Nicholas A. Christakis and James H. Fowler, *Connected* (NY: Little, Brown and Company: 2009). **Note**: Others have argued that their method for disentangling influence from other factors is flawed, so we should be cautious in applying their theory of influence.