

Social Capital & Brokerage Lab

For this lab we will use only one network and one attribute dataset:

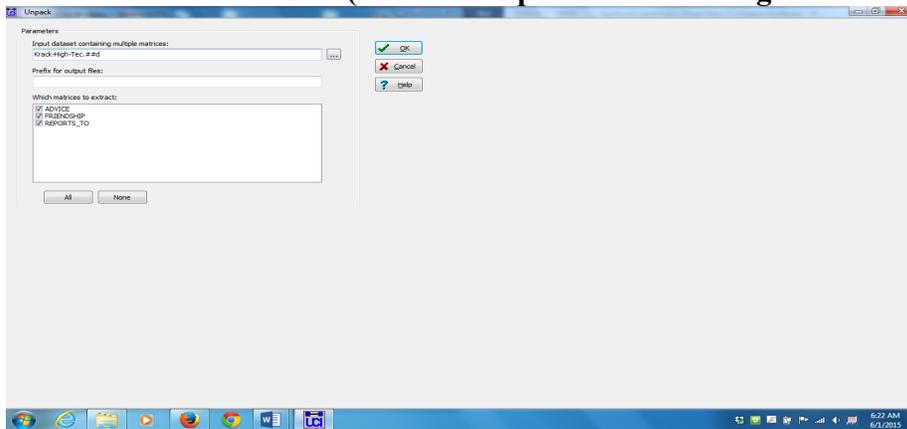
KRACK-HIGH-TEC (KHT) & HIGH-TEC-ATTRIBUTES (HTA)

This is a dataset collected by David Krackhardt from managers of a high tech company. KRACK-HIGH-TEC is a stacked dataset containing three directed, dichotomous matrices which represent ADVICE, FRIENDSHIP, and REPORTS_TO ties among 21 managers within the company. HIGH-TEC-ATTRIBUTES contains four attributes for each of the 21 actors, including each manager's age (in years), tenure with company, level in corporate hierarchy, and department.

NOTE: I use the abbreviations KHT and HTA to refer to the full filenames (KRACK-HIGH-TEC and HIGH-TEC-ATTRIBUTES) in the lab for brevity.

1) Structural Holes using UCINET and NetDraw with KHT and HTA

- a. If you have not already done so, unpack (Data | Unpack) the KHT dataset to get the three adjacency matrices **FRIENDSHIP**, **REPORTS_TO**, **ADVICE**. (Be sure the prefix is blank to get the filenames here.)



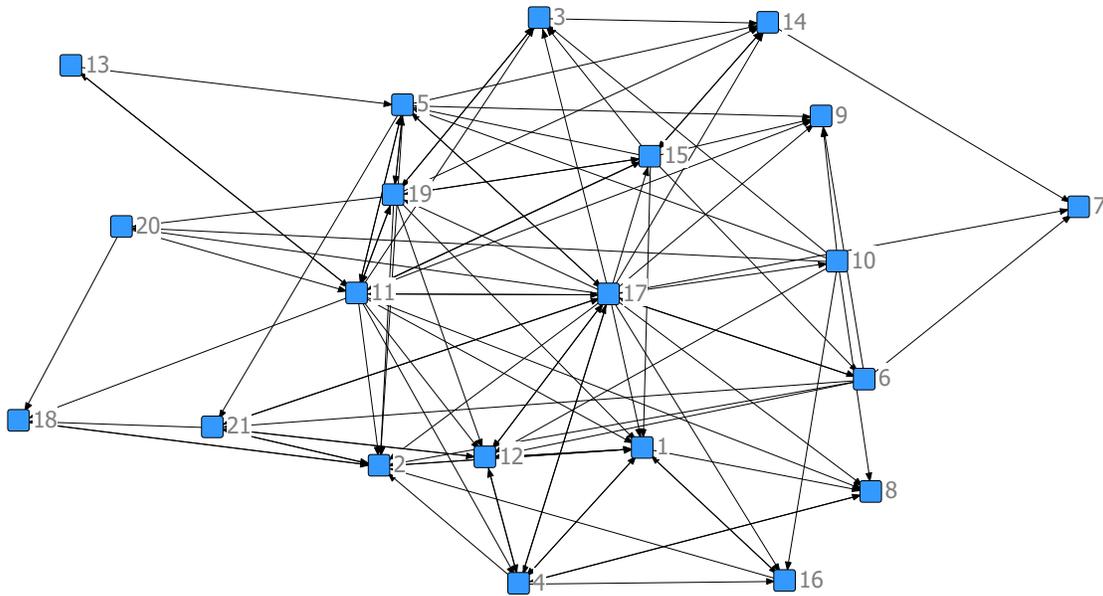
- b. Run Network | Ego Networks | Structural Holes on the **FRIENDSHIP** data. From the output, who appears to have the largest Effective Size? **Node 17 has the largest effective size with a score of 13.4375.**

You may have noticed that manually scanning the data on the UCINET output can be difficult. A solution is to open one of the output files created by UCINET in Excel and sort the data. Remember that for most procedures when you run an analysis in UCINET output files are created. When we run the structural holes routine we see that a separate data file on node-level measures is created by adding the suffix -SH to the Krack-High-Tec-Friendship dataset. We can use this file to easily sort through the output data. To do so:

Click on the Excel tab (the second icon on the toolbar) | Open UCINET Dataset | Select Krack-High-Tec-Friendship-SH

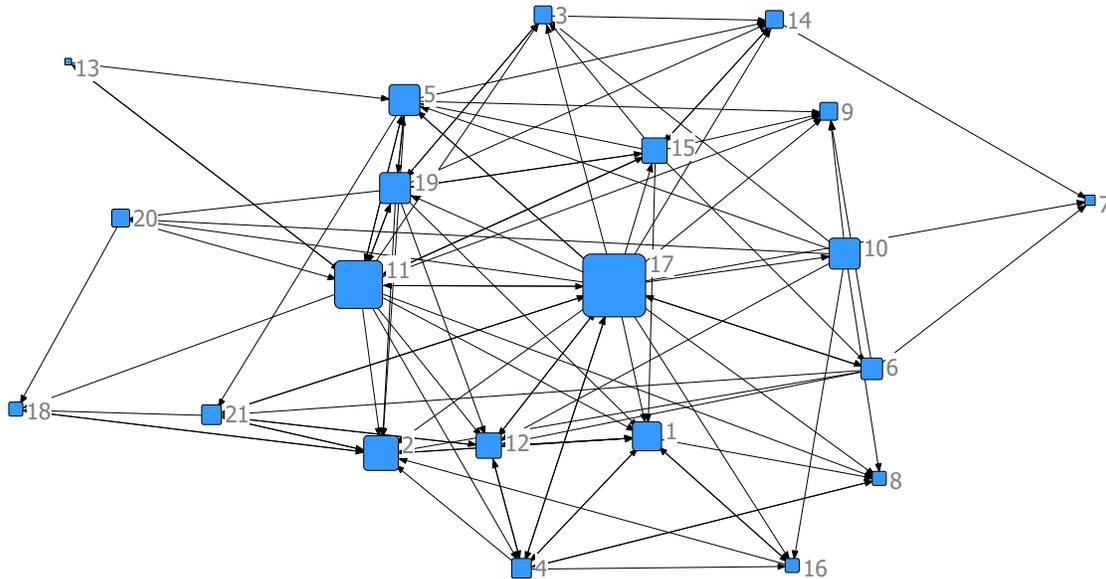
c. Load the **FRIENDSHIP** in Netdraw to visualize it.

Click on the Netdraw Icon | Open UCINET network dataset



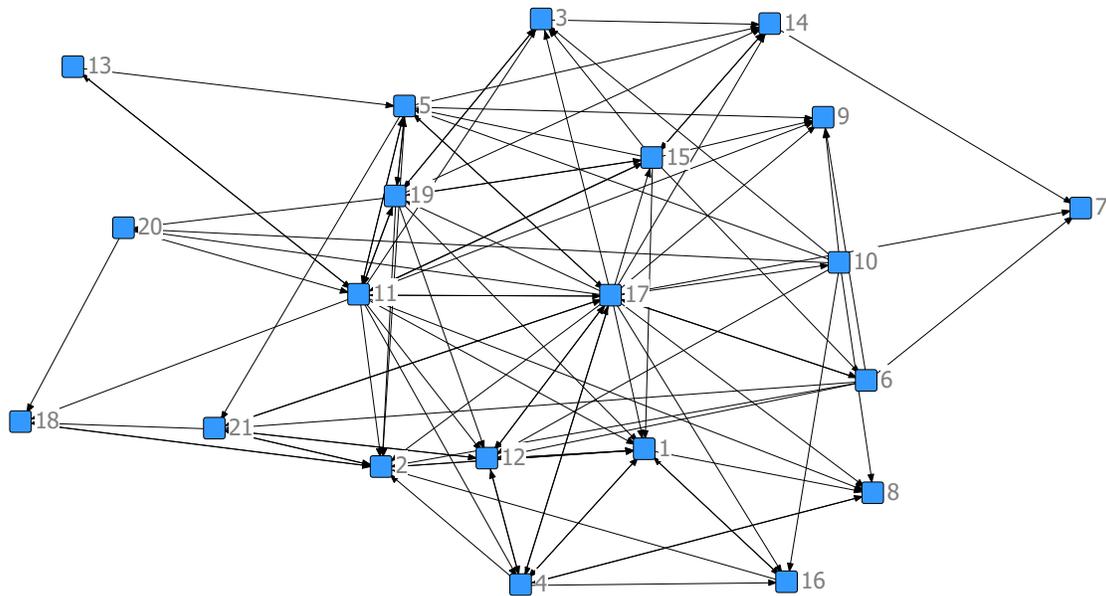
d. When you ran structural holes in UCINET it automatically saved the output in a dataset called “FREINDSHIP-SH” (unless you changed the name). Load **FRIENDSHIP-SH** as an attribute file in Netdraw and use the effective size attribute (EffSize) to size the nodes on the graph. What information does this convey in the graph?

We now see that that each node has been changed based on their effective size. We can also right click on each node to display attributes which will give us the score for each of our node attributes. We also see a tab to the right which displays a sorted list of each effective size value in our dataset.



- e. Again using the “Nodes” tab in the control region, select the density attribute (Density) and click on the “size” checkbox to resize the nodes based on Density. What happened? Why?

The size of the nodes reverse, this is because density is typically inversely proportional to effective size.



2) Clustering in UCINET using **KHT** (These were covered on Wednesday.)

- a. Run Network | Cohesion | Clustering Coefficient on the **FRIENDSHIP** data.

```

ucinetlog3 - Notepad
File Edit Format View Help
CLUSTERING COEFFICIENT
-----
Input dataset: Krack-High-Tec-FRIENDSHIP (C:\Users\csterling\Documents\UCINET data\Krack-High-Tec-FRIENDSHIP)
Overall graph clustering coefficient: 0.413
Weighted overall graph clustering coefficient: 0.324
Node Clustering Coefficients
-----
      1      2
Clus Coef nPairs
-----
1      0.417 36.000
2      0.344 45.000
3      0.500 15.000
4      0.476 21.000
5      0.311 45.000
6      0.310 21.000
7      0.500 3.000
8      0.500 10.000
9      0.500 15.000
10     0.179 28.000
11     0.253 91.000
12     0.339 28.000
13     1.000 1.000
14     0.467 15.000
15     0.389 36.000
16     0.500 10.000
17     0.229 153.000
18     0.333 6.000
19     0.367 45.000
20     0.500 10.000
21     0.400 15.000

Node-level clustering coefficients saved as dataset ClusteringCoefficients (C:\Users\csterling\Documents\UCINET data\ClusteringCoefficients)
Overall clustering coefficients saved as dataset OverallCC (C:\Users\csterling\Documents\UCINET data\OverallCC)
Please note that, for undirected graphs, the weighted clustering coefficient is simply transitivity.

Running time: 00:00:01
Output generated: 01 Jun 15 06:53:30
UCINET 6.573 Copyright (c) 1992-2015 Analytic Technologies

```

- b. By default, this procedure creates a file called **ClusteringCoefficients**. We are going to add those columns to the FRIENDSHIP-SH file created in the previous item. Go to Data | Join | Join Columns and enter FRIENDSHIP-SH and ClusteringCoefficients in the Datasets to join box, and then set the Output dataset to **HolesAndClusters** and click okay. Display this new file to ensure you have all the columns from both files in one dataset.

```

ucinetlog4 - Notepad
File Edit Format View Help
JOIN COLUMNS
-----
Input files: Krack-High-Tec-FRIENDSHIP-SH (C:\Users\csterling\Documents\UCINET data\Krack-High-Tec-FRIENDSHIP-SH)
C:\Users\csterling\Documents\UCINET data\ClusteringCoefficients
Output dataset: Holes and Clusters (C:\Users\csterling\Documents\UCINET data\Holes and Clusters)
Holes and Clusters successfully saved.
Datasets joined as columns
-----
      1      2      3      4      5      6      7      8      9      10     11
Degree Effsize Efficiency constraint Hierarchy Ego Betw eerness Ln(Const raint) Indirect s Density Clus Coef nPairs
-----
1      9.000 5.808 0.645 0.378 0.063 20.833 -0.974 0.777 0.417 0.417 36.000
2      10.000 7.077 0.708 0.336 0.062 20.750 -1.092 0.766 0.344 0.344 45.000
3      6.000 3.357 0.560 0.560 0.071 5.250 -0.580 0.770 0.500 0.500 15.000
4      7.000 4.045 0.578 0.481 0.084 10.333 -0.731 0.760 0.476 0.476 21.000
5      10.000 6.462 0.646 0.379 0.134 19.583 -0.971 0.796 0.311 0.311 45.000
6      7.000 4.375 0.625 0.520 0.210 5.000 -0.653 0.730 0.310 0.310 21.000
7      3.000 1.833 0.611 0.945 0.098 0.000 -0.056 0.639 0.500 0.500 3.000
8      5.000 2.833 0.567 0.655 0.101 4.000 -0.423 0.731 0.500 0.500 10.000
9      6.000 3.333 0.556 0.576 0.060 0.000 -0.552 0.809 0.500 0.500 15.000
10     8.000 6.375 0.797 0.480 0.379 0.000 -0.735 0.623 0.179 0.179 28.000
11     14.000 10.211 0.729 0.761 0.110 42.033 -1.345 0.773 0.253 0.253 91.000
12     8.000 5.333 0.667 0.444 0.170 16.333 -0.812 0.721 0.339 0.339 28.000
13     2.000 1.000 0.500 1.235 0.057 0.000 0.211 0.556 1.000 1.000 1.000
14     6.000 3.571 0.595 0.551 0.071 4.333 -0.595 0.756 0.467 0.467 15.000
15     9.000 5.333 0.593 0.398 0.090 11.533 -0.920 0.795 0.389 0.389 36.000
16     5.000 2.833 0.567 0.645 0.081 4.000 -0.439 0.733 0.500 0.500 10.000
17     18.000 13.438 0.747 0.204 0.061 111.150 -1.587 0.829 0.229 0.229 153.000
18     4.000 2.500 0.625 0.705 0.127 1.000 -0.349 0.607 0.333 0.333 6.000
19     10.000 6.464 0.646 0.364 0.096 13.233 -1.012 0.796 0.367 0.367 45.000
20     5.000 3.400 0.680 0.610 0.098 5.000 -0.494 0.677 0.350 0.350 10.000
21     6.000 3.833 0.639 0.544 0.115 11.500 -0.610 0.711 0.400 0.400 15.000

21 rows, 11 columns, 1 levels.

Running time: 00:00:01 seconds.
Output generated: 01 Jun 15 07:00:19
UCINET 6.573 Copyright (c) 1992-2015 Analytic Technologies

```

- c. Now run Tools | Similarities on this new dataset to find correlations between the variables in the columns. Which of Burt's structural holes measures is the most like and the most opposite clustering coefficient?

Why?

Remember to change the default option from rows to columns as you run Similarities. It would also help to open the output file in excel to sort the data. Density and clustering coefficient are perfectly correlated, they are the same measure. Burt's measure of constraint is also highly correlated with the clustering coefficient. Efficiency and effective size are the most different from clustering coefficient. This is understandable as these measures count only non-redundant ties.

3) E-I Index with UCINET using **KHT & HTA**

- a. Run Network | Cohesion | E-I Index on the **FRIENDSHIP** data, partitioning the data based on department (which is in column 4 of the **HIGH-TEC-ATTRIBUTES** dataset, you must specify COL 4 after the filename, but the L button can look it up and fill it in for you). Looking at the individual E-I index statistics, who has the most homophilous ties (more concentrated within the same department), and who has the most heterophilous (most concentrated outside the same department) ones?

Remember that E- Index is calculated by $(E-I)/(E+I)$. Nodes 7 and 10 have networks that are completely heterophilous and nodes 14 and 16 are the most homophilous.

- b. Rerun E-I index using the same partitioning, but instead of using the **FRIENDSHIP** dataset, use the stacked **KHT (KRACK-HIGH-TEC)** dataset. Bearing in mind that you ran this on a stacked dataset, what do you think these results tell you? How could you find out?

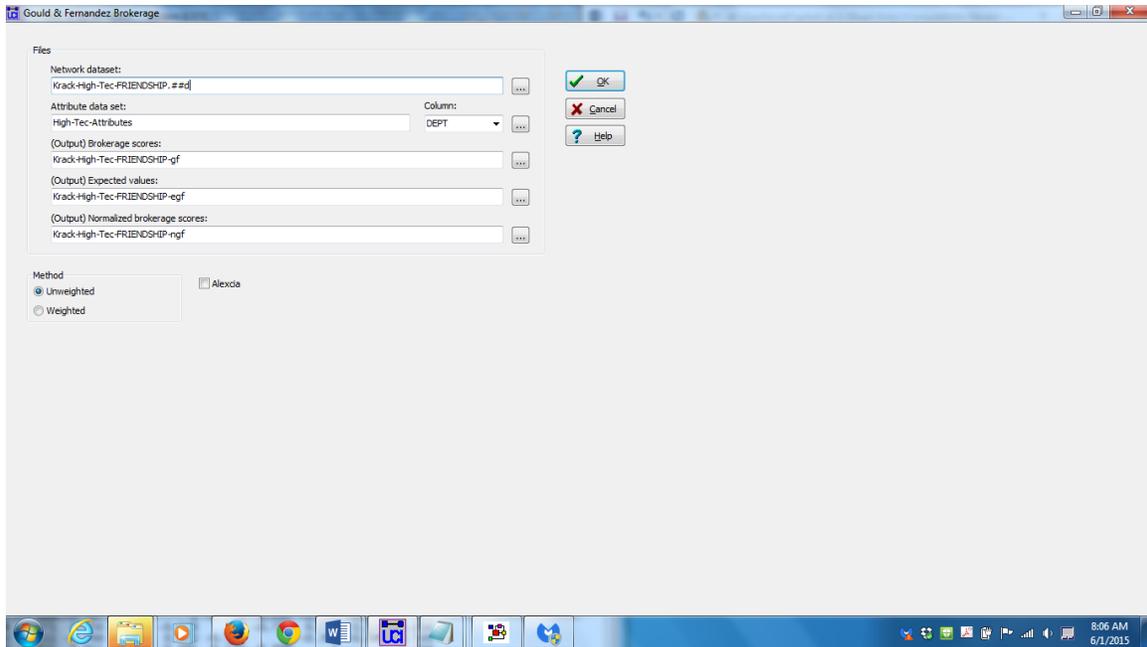
Sometimes UCINET measures will work on each relationship in a stacked dataset, other times UCINET measures will only work on the first relationship in the stacked dataset

- c. Display (using the "D" icon) the **KHT** stacked dataset. Rerun the E-I index on the individual dataset that is displayed first from this command and compare the results to the results from step b. Is this what you thought was happening?

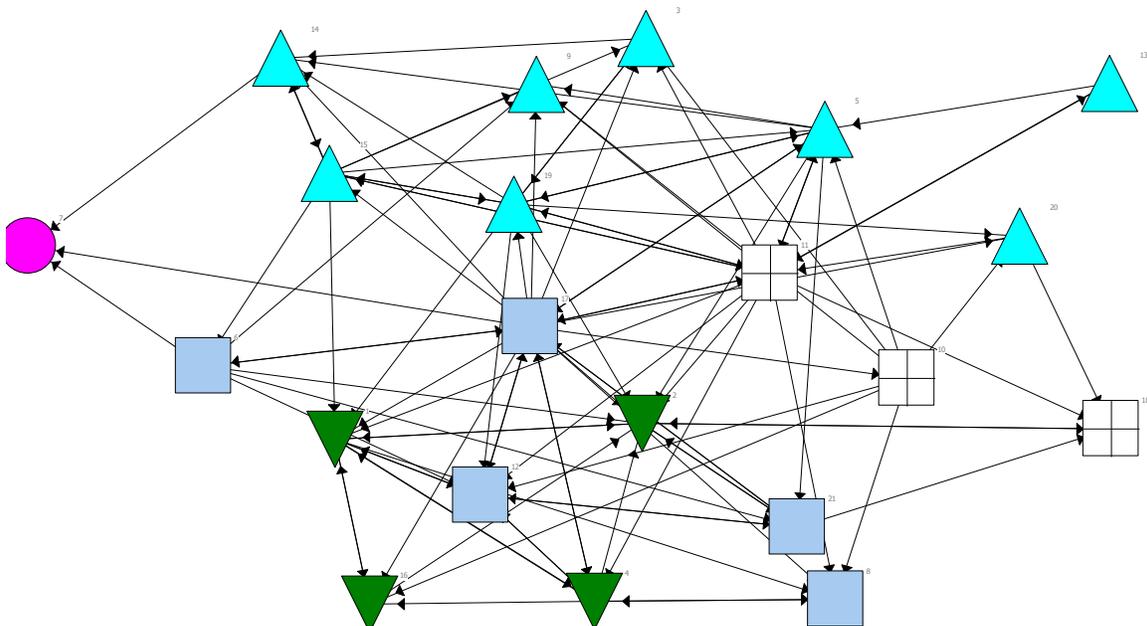
The output from 3b is identical to 3c. This means that the E-I index only recognizes the first matrix in the stacked dataset, in this case that is Krack-High-Tec- Advice

4) Brokerage with UCINET using **KHT & HTA**

- a. Run Network | Ego Networks | G&F Brokerage roles on the **FRIENDSHIP** data, again using the department attribute. You will notice this time, when you put in the name of the attribute file, it populates the drop down list with the attributes in that file and you can just select DEPT.



- b. Open **KRACK-HIGH-TEC** in Netdraw and, using the “Rels” tab in the control region, display only the **FRIENDSHIP** relation. Compare this visualization with the results from step a. It may help to color or shape the nodes by department. Can you find at least one example of each kind of brokerage for Actor 5? (Remember, direction counts in brokerage, so make sure you have the arrows on and visible.)



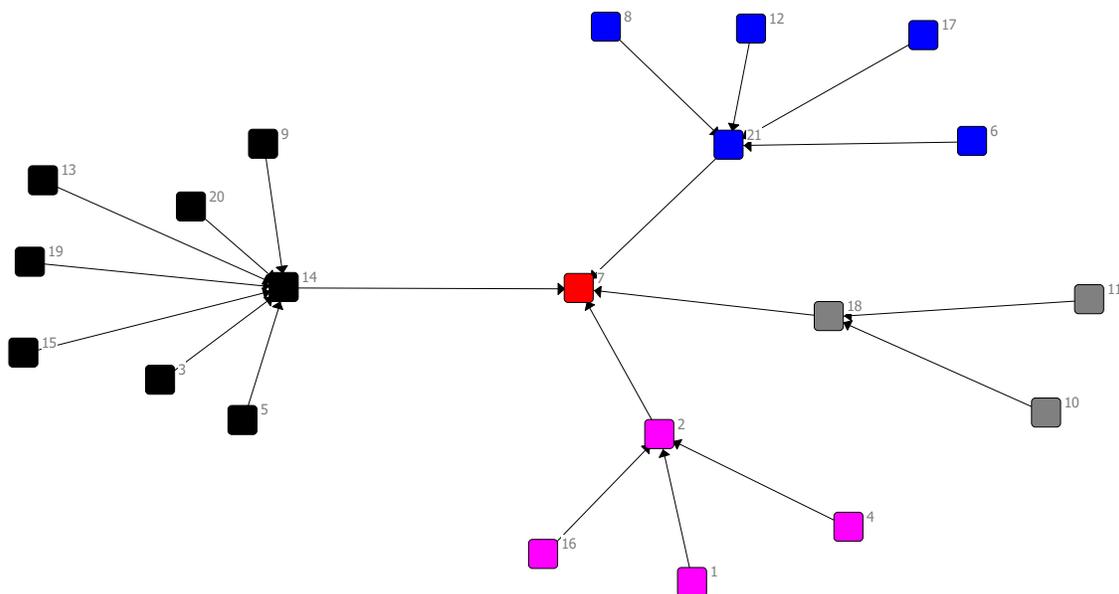
It is very difficult to discern brokerage role simply by scanning the network diagram. Remember that both direction of the relationship and department affiliation of each node determines brokerage roles. You may want to increase the size of arrows and color nodes by department in Netdraw before answering this question. Also increasing the symbol size really helps in visualizing the arrows. By looking at Actor 5 we can see several types of brokerage A) Coordinator: 13→05→09 B) Gatekeeper: 10→05→19 C) Representative: 13→05→17 D) Consultant: 10→05→11 E) Liaison 10→05→17

- c. Because this dataset is actors by brokerage roles, it (like most output from UCINET routines) can be used as an attribute file. Load **BROKERAGE** as an attribute file in NetDraw (making sure the **FRIENDSHIP** relation is open and displayed first). Now, size the nodes by the various brokerage roles (Consultant, Representative, etc.). Does doing this help identify the different kinds of brokerage roles people play in the network?

Doing this allows us to compare each brokerage role among people, but does not allow an easy comparison of the different brokerage roles played by each person.

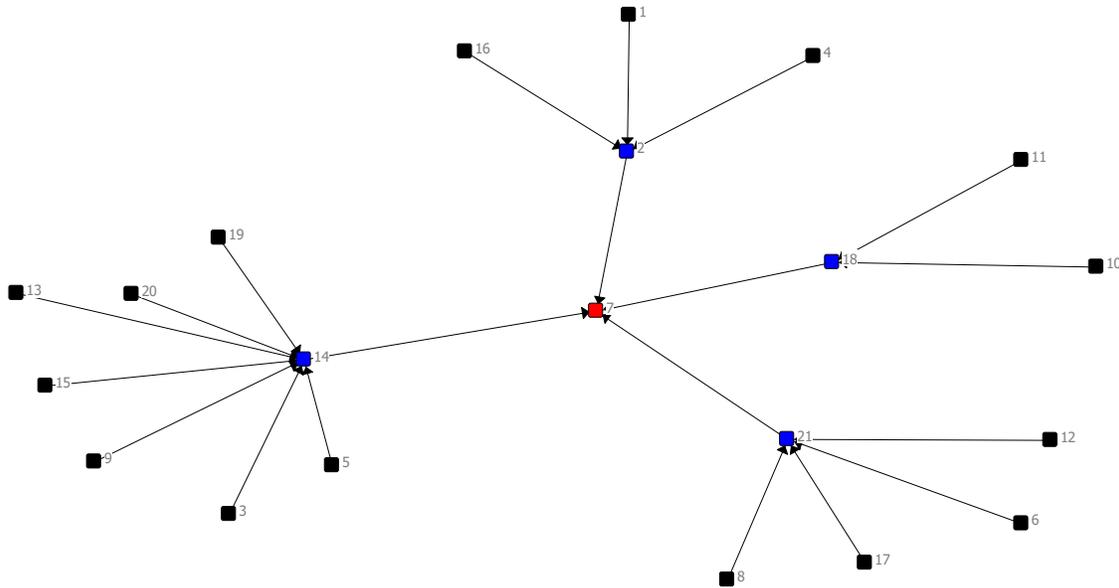
- d. Rerun the brokerage routine using the **REPORTS_TO** data instead of the **FRIENDSHIP** data, still partitioning based on the department attribute. Thinking about the nature of this relationship, what can you tell about the actors based on this output.

This will represent the organizational structure of the organization. Managers are representatives of their respective departments to the CEO.



- e. Re-run brokerage on the **REPORTS_TO** data, but this time use the Level attribute to partition that data. How does this data compare to the previous output? Why is it different?

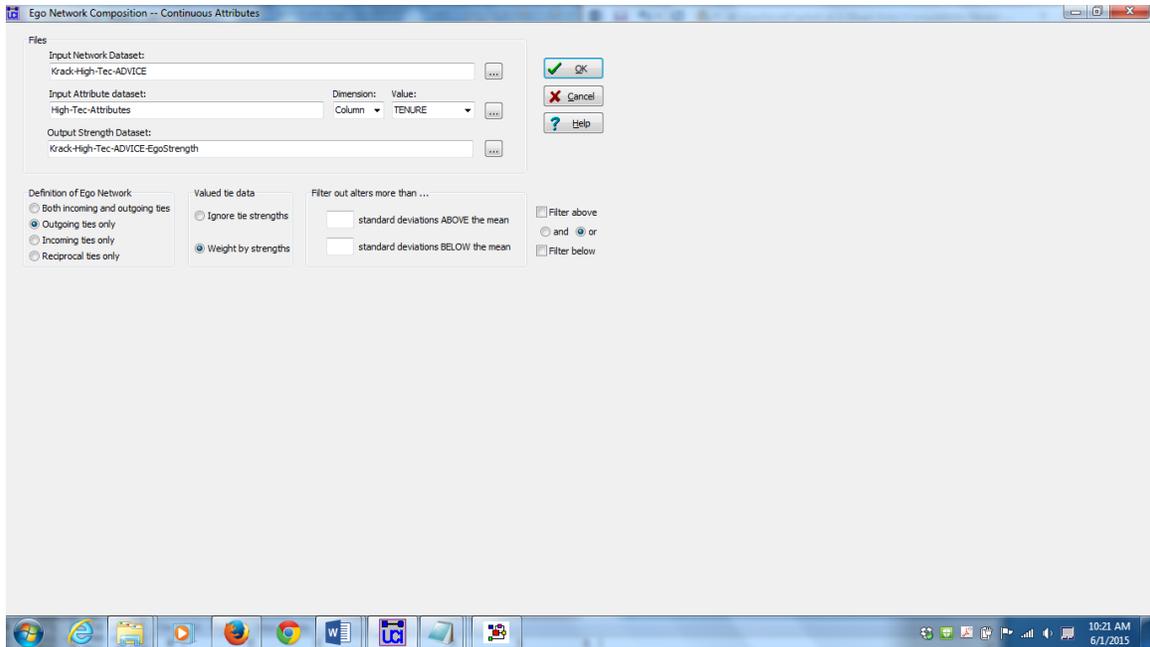
Now managers are liaisons between staff and managers.



5) Ego-Net Strength with UCINET and NetDraw using **KHT**

- a. Run Network | Ego Networks | Egonet composition | Continuous alter attributes specifying the **ADVICE** dataset you previously unpacked from **KHT** for the Input Network Dataset. For the Input Attribute dataset, specify **HIGH-TEC-ATTRIBUTES** and select the column labeled “Tenure.”

In running this procedure we will have to specify how the ego network is defined, the default option is to include both incoming and outgoing ties. This would be fine for relationships that have been symmetrized or for relationships where directionality may not be important in certain instances (i.e. basic communication ties) but this should matter for advice ties. Let’s assume that the advice ties are *advice seeking* ties thus outgoing advice ties would measure “getting advice from other people” In this case we should measure the ego network using outgoing ties only.



- b. This procedure gives information about each actors' egonet with respect to their access to "Tenure". If we say that getting advice from people who have worked for the company longer is more likely to lead to success, which people are most likely to succeed based on these results? What other measures reported might be indicators of success based on the composition of an actor's advice network, and why?

If the tenure of advice givers matters most for success the ego net composition measure Avg. should be a good measure to use. When testing this hypothesis we would probably want to control for the size of the advice network however. According to this approach we might expect that nodes 2, 12, 8, 14 would be highly successful individuals.

An alternative theory could be related to the diversity of the advice we receive. Therefore the StdDev in tenure of advice givers may represent the diversity of perspectives from individuals that give advice. According to this approach we might expect that nodes 8, 5, 17, 15

- c. Go back to NetDraw, load **KHT** and ensure that the **ADVICE** relation is being displayed.
- d. Now load the dataset just created (by default it was called **ADVICE-EgoStrength**) as an Attribute file in NetDraw. Size the nodes based on the various measures (Sum, Avg, StdDev, etc.). How do the results differ?

Visualizing these different attributes show us that there are different patterns based on which attribute is chosen, standard deviation looks to be more uniform, and sum and avg. do not appear to be very strongly related in this case.