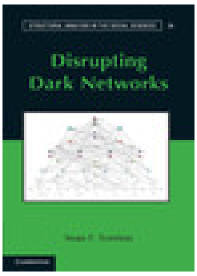


Disrupting Dark Networks

Sean F. Everton

Book DOI: http://dx.doi.org/10.1017/CBO9781139136877

Online ISBN: 9781139136877

Hardback ISBN: 9781107022591

Paperback ISBN: 9781107606685

Chapter

Appendix 3 - Multidimensional Scaling with UCINET pp. 404-416

Chapter DOI: http://dx.doi.org/10.1017/CBO9781139136877.021

Cambridge University Press

Appendix 3

Multidimensional Scaling with UCINET

This appendix illustrates how to estimate multidimensional scaling (MDS) coordinates in UCINET, which can then be visualized using NetDraw. The advantage of calculating MDS coordinates in UCINET is that UCINET calculates a stress statistic (while NetDraw currently does not – with one exception). Stress statistics are valuable because they indicate how well the network map fits the data. A stress statistic greater than .20 is generally considered intolerable. Thus, we would not want to rely too heavily on a network map with a high stress statistic. Another advantage of calculating coordinates in UCINET rather than in NetDraw is that with UCINET users can estimate metric and nonmetric MDS models, while NetDraw currently offers only metric MDS. Finally, UCINET can also estimate three-dimensional MDS models, which typically exhibit lower stress levels (i.e., they produce a better fit) than do two-dimensional models. While three-dimensional models cannot be visualized using NetDraw, they can be visualized using Mage, a program that is included as one of UCINET's helper programs.[1] In this appendix we will use a relatively small social network dataset, because it is easier to illustrate these techniques with smaller rather than with larger networks. That said, there is nothing we do in this appendix that you cannot do with larger datasets.

Multidimensional Scaling of Symmetric One-Mode Networks

We will begin with symmetric one-mode networks because it is easier to estimate MDS coordinates for symmetric one-mode networks than

[1] Mage was developed as a device to be used in molecular modeling (Richardson and Richardson 1992). For more information, see the article by Freeman, Webster, and Kirke (1998) or http://kinemage.biochem.duke.edu/software/mage.php. Pajek also exports images that can be visualized using Mage. Also, see de Nooy et al. (2011:388–389).

asymmetric ones. For this, we use the marital ties of Padgett's Florentine Families, which we have discussed before. Our first task is to use this network to calculate a set of related coordinates. We consider both metric and nonmetric MDS. We will then read these (and the related network) into NetDraw and Mage.

Metric Multidimensional Scaling

As previously noted, network analysts have long used sociograms to visualize social networks, and in recent years analysts have begun using a series of mathematical techniques to locate the points of a network in such a way that the distances between them are meaningful. MDS is one such technique. It uses the concepts of space and distance to represent a network's internal structure (Wasserman and Faust 1994). The typical input is a symmetric matrix consisting of measures of similarity or dissimilarity between pairs of actors. Output generally consists of a set of estimated distances between pairs of actors that we can represent in one-, two-, three-, or higher-dimensional space (Kruskal and Wish 1978; Wasserman and Faust 1994).

Metric MDS takes a given matrix of proximities that measure the similarities or dissimilarities among a set of actors and calculates a set of points in k-dimensional space, such that the distances between them correspond as closely as possible to the input proximities (Borgatti, Everett, and Freeman 2011).[2] Metric distance differs from distance in graph theory. In graph theory, the distance between two points is measured in terms of the number of lines in the path that connects the two points. In MDS the distance between two points is the most direct route between them. "It is a distance that follows a route 'as the crow flies,' and that may be across 'open space' and need not – indeed, it normally will not – follow a graph theoretical path" (Scott 2000:148–149).

Under UCINET's *Tools* menu, select the *Scaling/Decomposition> Metric MDS* command; this should bring up a dialog box similar to (Figure A3.1). There are a number of options available. In general, you will want to accept UCINET's defaults unless you have a good reason not to. Here, I have changed only one default setting: the name of the output dataset in order to make it easier to identify.

Scaling/ Decomposition> Metric MDS

Running this procedure produces both a scatter plot (not shown) and an output file that lists the MDS coordinates and a stress score (Figure A3.2). As you can see, the stress is .30, which tells us that the coordinates do not fit the data relatively well. The coordinates themselves are stored in the file `PadgetMetricMdsCoord2`, which we will use later use to read

[2] The Padgett data proximities represent similarities between the families. That is, a "1" in a matrix cell means that the two families represented by that cell share a marital tie.

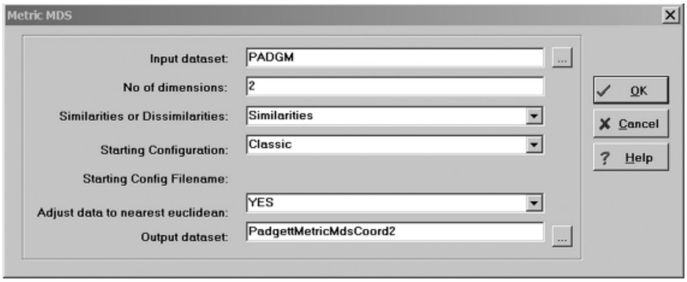
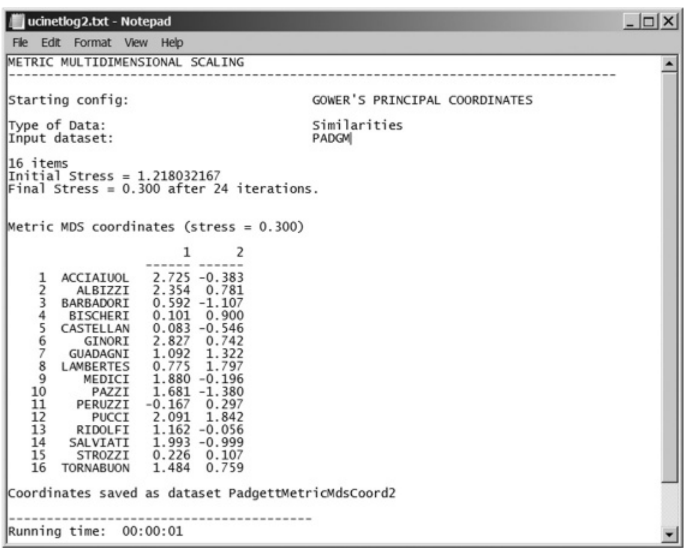


Figure A3.1. UCINET's Metric MDS Dialog Box

into NetDraw. Because the fit is rather poor, let's also estimate a three-dimensional metric MDS model, using the same command, except this time change the "No. of dimensions" option from 2 to 3 (not shown) and the name of the coordinate file that UCINET will generate to `PadgettMetricMdsCoord3` (or something similar – just so we do not overwrite the previous coordinate file). When we do this (the dialog box is not shown), we get a better-fitting result (Figure A3.3). The stress level is now below .30 (.195), which means we can be a little more confident with the resulting visualization. Before seeing how to use these two sets of coordinates to visualize the network in NetDraw and Mage, respectively, let's first turn our attention to nonmetric MDS, which is often a better choice when working with binary (i.e., dichotomous) data.



```
ucinetlog2.txt - Notepad
File  Edit  Format  View  Help
METRIC MULTIDIMENSIONAL SCALING
------------------------------------------------------------------------------

Starting config:                        GOWER'S PRINCIPAL COORDINATES

Type of Data:                           Similarities
Input dataset:                          PADGM

16 items
Initial Stress = 1.218032167
Final Stress = 0.300 after 24 iterations.


Metric MDS coordinates (stress = 0.300)

                          1      2
                     ------ ------
    1  ACCIAIUOL      2.725 -0.383
    2    ALBIZZI      2.354  0.781
    3  BARBADORI      0.592 -1.107
    4   BISCHERI      0.101  0.900
    5  CASTELLAN      0.083 -0.546
    6     GINORI      2.827  0.742
    7   GUADAGNI      1.092  1.322
    8  LAMBERTES      0.775  1.797
    9     MEDICI      1.880 -0.196
   10      PAZZI      1.681 -1.380
   11    PERUZZI     -0.167  0.297
   12      PUCCI      2.091  1.842
   13    RIDOLFI      1.162 -0.056
   14   SALVIATI      1.993 -0.999
   15    STROZZI      0.226  0.107
   16  TORNABUON      1.484  0.759

Coordinates saved as dataset PadgettMetricMdsCoord2

----------------------------------------
Running time:   00:00:01
```

Figure A3.2. UCINET's Metric MDS Output (2-D)

```
ucinetlog3.txt - Notepad
File  Edit  Format  View  Help
METRIC MULTIDIMENSIONAL SCALING
--------------------------------------------------------------------------------

Starting config:                       GOWER'S PRINCIPAL COORDINATES

Type of Data:                          Similarities
Input dataset:                         PADGM

16 items
Initial Stress = 0.896255808
Final Stress = 0.195 after 34 iterations.


Metric MDS coordinates (stress = 0.195)

                        1      2      3
                   ------ ------ ------
    1  ACCIAIUOL    1.987 -0.444  0.714
    2   ALBIZZI     1.814  1.004  0.868
    3  BARBADORI    0.556 -0.945  1.022
    4   BISCHERI   -0.233  1.061  0.658
    5  CASTELLAN   -0.178 -0.387  1.225
    6    GINORI     1.869  1.102  1.604
    7   GUADAGNI    0.910  1.393  0.410
    8  LAMBERTES    0.736  1.864  1.014
    9    MEDICI     1.404 -0.193  0.581
   10     PAZZI     0.976  0.010  2.338
   11   PERUZZI    -0.430  0.419  1.213
   12     PUCCI     0.360  1.113  2.270
   13   RIDOLFI     0.662 -0.041 -0.195
   14  SALVIATI     1.391 -0.273  1.934
   15   STROZZI    -0.270  0.208  0.540
   16  TORNABUON    1.151  0.538 -0.209

Coordinates saved as dataset PadgettMetricMdsCoord2

----------------------------------------
Running time:  00:00:01
```

Figure A3.3. UCINET's Metric MDS Output (3-D)

Nonmetric Multidimensional Scaling

There are some limitations to using metric MDS for visualizing social networks. Many relational datasets, such as the Padgett data, are dichotomous (i.e., they simply indicate either the presence or absence of a tie, not its strength). Consequently, UCINET cannot directly use such data to measure proximities. Instead, it first needs to convert it into an alternative (and scaled) metric, such as correlation coefficients, before calculating its metric properties. However, because dichotomous data are not metric (i.e., they only consist of 1's and 0's), it is possible for UCINET to estimate the coordinates incorrectly. Even when the data are valued, metric assumptions may be inappropriate if the data do not represent a definitive scale. For example, a family with four marital ties may not be twice as central a family with only two.

While nonmetric MDS procedures (like metric MDS procedures) use symmetrical adjacency matrices to calculate similarities or dissimilarities between actors, unlike metric MDS, they do not directly convert these values into Euclidean distances. Instead, they consider only rank order. They treat the data, in other words, as ordinal and "seek a solution in which the rank ordering of the distances is the same as the rank ordering of the original values" (Scott 2000:157). Nonmetric MDS is often preferred because it tends to provide a better "goodness-of-fit" (stress) statistic.

To estimate nonmetric MDS coordinates, select the *Nonmetric MDS* option found under the *Tools>Scaling/Decomposition* submenu. This brings up a dialog box similar to the one we used for calculating metric

Tools>Scaling/Decomposition >Nonmetric MDS

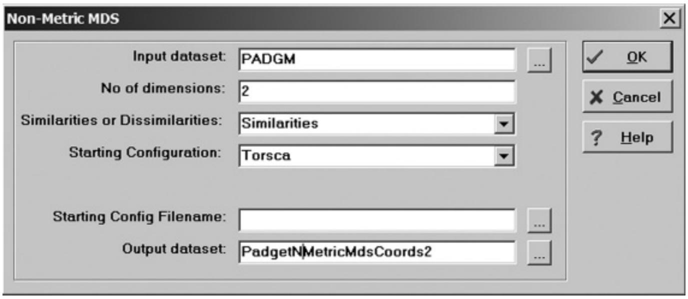


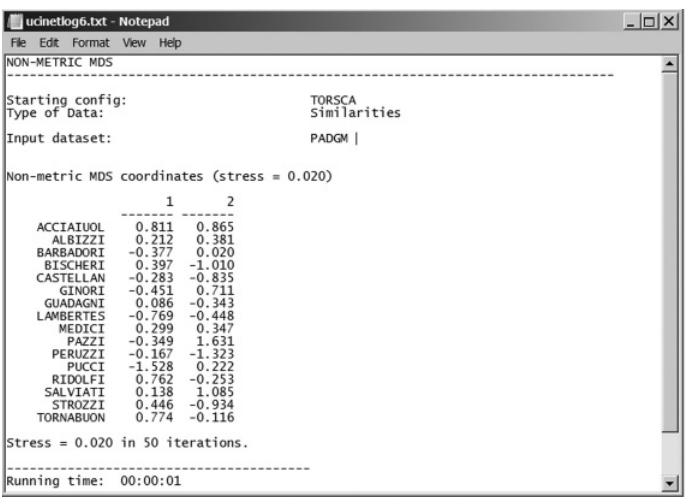
Figure A3.4. UCINET's Nonmetric MDS Scaling Dialog Box

MDS (Figure A3.4). As before, I accepted all of UCINET's defaults except that I changed the name of the output dataset in order to make it easier to identify. Moreover, I estimated both two- and three-dimensional models.

As with metric MDS, this procedure produces scatter plots (not shown), output files (Figures A3.5 and A3.6) that list the MDS coordinates, and a stress score. As you can see, the stress statistic for the two-dimensional model is .02, and the stress statistic for the three-dimensional model is .00, both of which fit the data far better than did the coordinates estimated using metric MDS. Thus, we will use the nonmetric coordinates for visualizing in NetDraw and Mage.

Using UCINET Coordinates with NetDraw and Mage

Using MDS coordinates generated by UCINET to visualize a network in NetDraw is straightforward. First open the Padgett data in NetDraw



```
ucinetlog6.txt - Notepad
File  Edit  Format  View  Help
NON-METRIC MDS
------------------------------------------------------------------------------

Starting config:                        TORSCA
Type of Data:                           Similarities

Input dataset:                          PADGM

Non-metric MDS coordinates (stress = 0.020)

                      1       2
                -------  -------
     ACCIAIUOL    0.811   0.865
       ALBIZZI    0.212   0.381
     BARBADORI   -0.377   0.020
      BISCHERI    0.397  -1.010
     CASTELLAN   -0.283  -0.835
        GINORI   -0.451   0.711
      GUADAGNI    0.086  -0.343
     LAMBERTES   -0.769  -0.448
        MEDICI    0.299   0.347
        PAZZI    -0.349   1.631
       PERUZZI   -0.167  -1.323
         PUCCI   -1.528   0.222
       RIDOLFI    0.762  -0.253
      SALVIATI    0.138   1.085
       STROZZI    0.446  -0.934
     TORNABUON    0.774  -0.116

Stress = 0.020 in 50 iterations.

----------------------------------------
Running time:  00:00:01
```

Figure A3.5. UCINET's Nonmetric MDS Output (2-D)

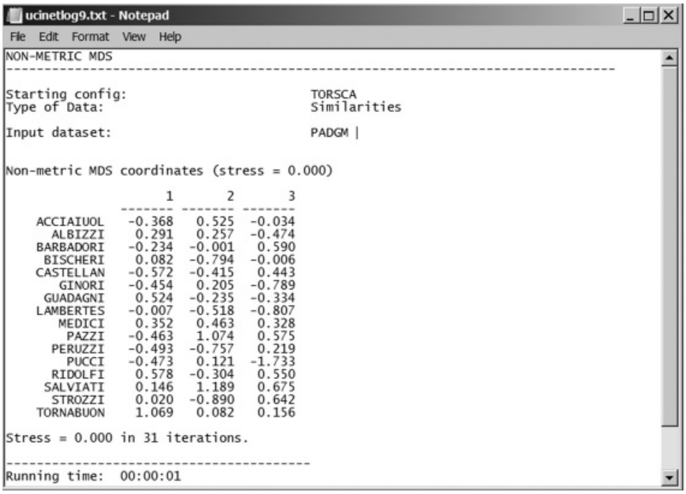

```
ucinetlog9.txt - Notepad
File Edit Format View Help
NON-METRIC MDS
--------------------------------------------------------------------------------

Starting config:                        TORSCA
Type of Data:                           Similarities

Input dataset:                          PADGM |


Non-metric MDS coordinates (stress = 0.000)

                   1       2       3
             ------- ------- -------
  ACCIAIUOL  -0.368   0.525  -0.034
    ALBIZZI   0.291   0.257  -0.474
  BARBADORI  -0.234  -0.001   0.590
   BISCHERI   0.082  -0.794  -0.006
  CASTELLAN  -0.572  -0.415   0.443
     GINORI  -0.454   0.205  -0.789
   GUADAGNI   0.524  -0.235  -0.334
  LAMBERTES  -0.007  -0.518  -0.807
     MEDICI   0.352   0.463   0.328
      PAZZI  -0.463   1.074   0.575
    PERUZZI  -0.493  -0.757   0.219
      PUCCI  -0.473   0.121  -1.733
    RIDOLFI   0.578  -0.304   0.550
   SALVIATI   0.146   1.189   0.675
    STROZZI   0.020  -0.890   0.642
  TORNABUON   1.069   0.082   0.156

Stress = 0.000 in 31 iterations.

----------------------------------------
Running time:  00:00:01
```

Figure A3.6. UCINET's Nonmetric MDS Output (3-D)

with its *File>Open>Ucinet dataset>Network* command. Next use the *File>Open>Ucinet dataset>Coordinates* command to open the related coordinate file (`PadgetNMetricMdsCoords2.##h`). NetDraw automatically assigns these coordinates to the respective nodes. You should get a network map similar to the one displayed in Figure A3.7. If you choose a mapping algorithm resident in NetDraw, then in order to recall the coordinates calculated in UCINET, you will need to reissue the *File>Open>Ucinet dataset>Coordinates* command.

[NetDraw] File>Open>Ucinet dataset>Network

File>Open >Ucinet dataset >Coordinates

To visualize a network with Mage using MDS coordinates generated by UCINET takes a few more steps than it does with NetDraw. We need to first export both the network and coordinate data in a format that Mage can read. To do this, we use UCINET's *Data>Export>Mage* command, which calls up a dialog box similar to Figure A3.8.

[UCINET] Data>Export> Mage

As you can see, using the radio buttons on the right side of the dialog box, both the network (`PADGM.##h`) and coordinate data (`PadgetNMetricMdsCoords2.##h`) have been loaded into the dialog box. Initially, you will probably want to accept UCINET's default settings. If, later, you want to adjust the size of nodes, the width of ties, and so on, you can play with the various options and see how they alter the resulting sociogram. Clicking "OK" brings up a dialog box (Figure A3.9) that allows you to open the newly formatted network (`PADGM.kin`) in Mage (Mage can also be accessed with the *Visualize>Mage* command).[3]

Visualize>Mage

Clicking "OK" opens Mage along with a three-dimensional visualization of the network that appears as an interactive computer display

[3] If the Mage program did not come with UCINET, you can download it at http://kinemage.biochem.duke.edu/software/mage.php.

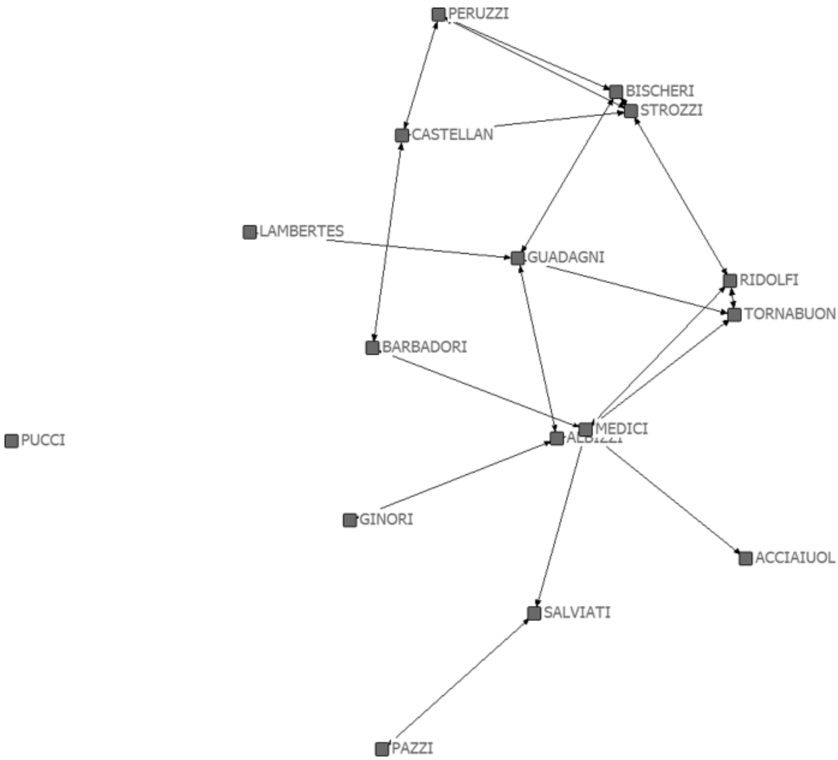


Figure A3.7. Nonmetric MDS Map of Padgett Marriage Network (NetDraw)

(Figure A3.10). Researchers can rotate Mage images, turn parts of the displays on or off, use the mouse to select and identify various points of the network, and animate changes between different arrangements of objects. Here, the background has been changed from black to white and the node

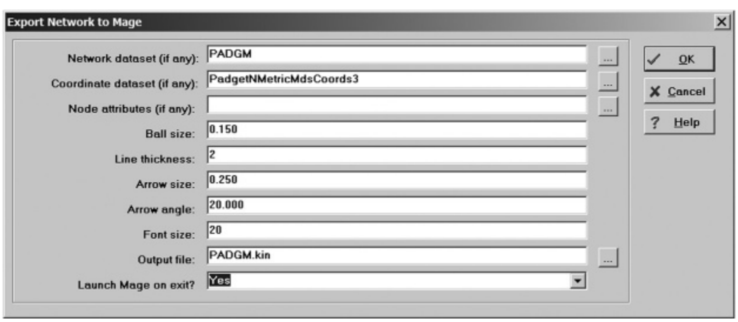


Figure A3.8. UCINET's Export to Mage Dialog Box

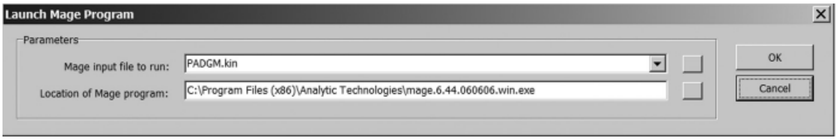


Figure A3.9. UCINET's Launch Mage Program Dialog Box

color to gray scale using the using the *Display>White Background* and *Display>Gray Scale* commands, respectively.

Display>White Background

Display>Gray Scale

Mage also allows users to easily change the color of the vertices. For example, let's highlight the Medici family because of its centrality. To change the color of the Medici vertex, first select the "Change Color" option under the "Edit" menu. This calls up a new "Changecolor" option along the right hand side of the display (not shown). If you check it and click on the Medici (or any) node, this will bring up a color selection dialog box (not shown). This, in turn, gives you a choice between changing the color of the entire network (list) or a single node (point) in the network. To change the color of the Medici node, select the "Point" option and then the color of your choice. Mage files actually allow for a considerable amount of sophisticated editing. Although that is beyond the scope of this appendix, see Appendix A of *A Guide to the Visually Perplexed* (Everton 2004).

Edit>Change color

Multidimensional Scaling of Asymmetric One-Mode Networks

Visualizing asymmetric (directed) one-mode networks using UCINET is somewhat different because the MDS routines require symmetric

[UCINET]

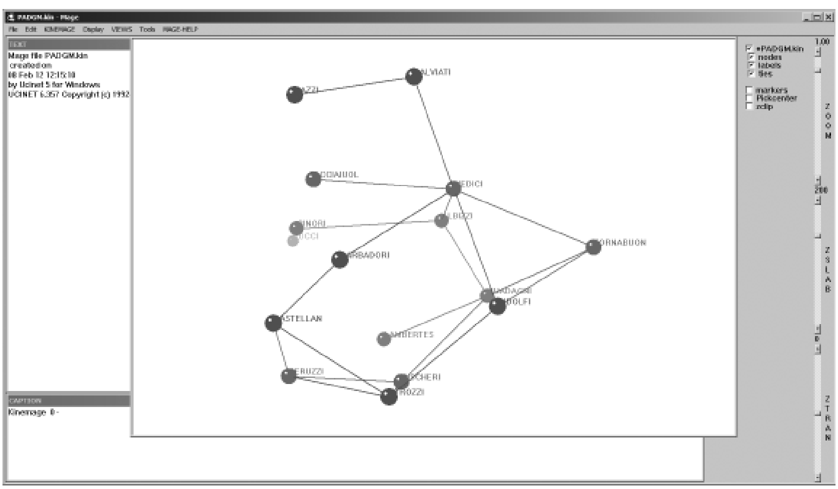


Figure A3.10. Mage's Visualization of Padgett Marriage Data

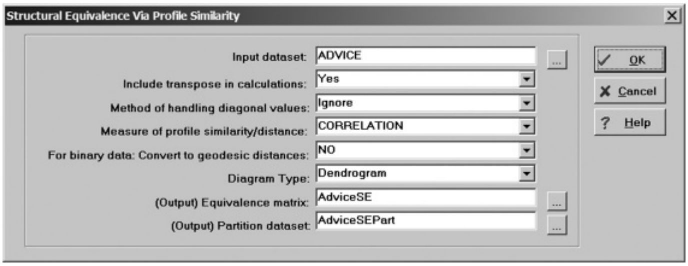


Figure A3.11. UCINET's Structural Equivalence Profile Similarity Dialog Box

matrices.[4] Thus, we need to first calculate an equivalence matrix, based either on the distances (e.g., Euclidean) or the correlations between the nodes of the directed matrix. We then submit this equivalence matrix, which is symmetric, to MDS algorithms. For our purposes here we will use the advice network of Krackhardt's High Tech Managers (1987a), which we have used previously. To calculate an equivalence matrix, under the *Network* menu, choose the *Roles & Positions>Structural>Profile* command. This brings up UCINET's profile similarity dialog box (see Figure A3.11). As is generally the case, accept UCINET's defaults, but change the "Measure of profile similarity/distance" option to "Correlation" because in this case a tie between two individuals indicates similarity.[5] You may also want to change the names of the output files in order to make them easier to identify. This produces both a dendogram (not shown) and a structural equivalence matrix (not shown).

Network>Roles & Positions >Structural >Profile

The next step in the process is to submit the structural equivalence matrix (`AdviceSE.##h`) to the MDS techniques previously discussed (not shown). The stress statistics for the two-dimensional metric and nonmetric MDS were .196 and .129, respectively, while the statistics for the three-dimensional metric and nonmetric MDS were .126 and .088, respectively. This illustrates that nonmetric MDS generally provides a better fit with binary data than metric MDS, and three-dimensional models typically fit the data better than two-dimensional models. Figure A3.12 presents the nonmetric MDS two-dimensional sociogram as

[4] You can submit asymmetric networks to UCINET's *Metric MDS* and *Nonmetric MDS* commands and UCINET will output coordinate files; however, these will most likely be incorrect.

[5] If we had chosen to calculate the matrix using the Euclidean distance option, then in the resulting matrix the larger the number would indicate the greater the distance of one actor from another. For example, in Figure A3.7 the correlation coefficients along the diagonal are all 1.00 (because each actor is perfectly correlated with itself); if we had chosen the Euclidean distance option, the coefficients along the diagonal would be 0.00. Thus, if we had chosen the Euclidean option, when we instructed UCINET to perform MDS on the structural equivalence matrix, we would need to choose the "Dissimilarities" option rather than the "Similarities" option (see Figure A3.8).

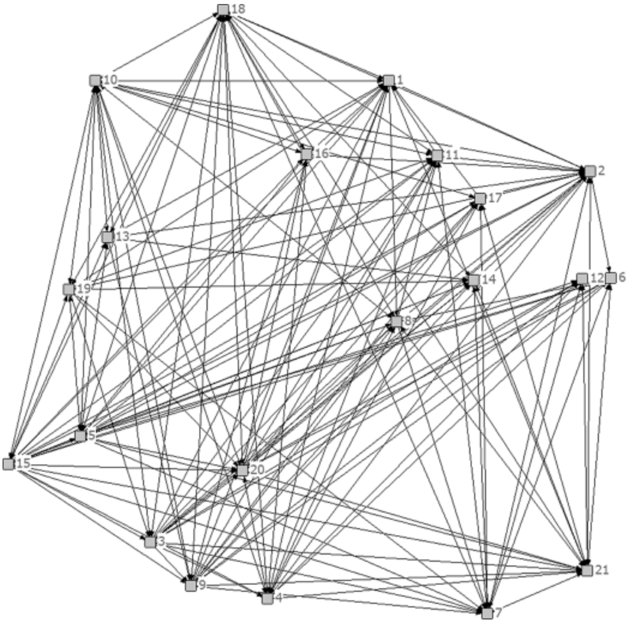


Figure A3.12. NetDraw's Nonmetric MDS Graph of Krackhardt Advice Network (2-D)

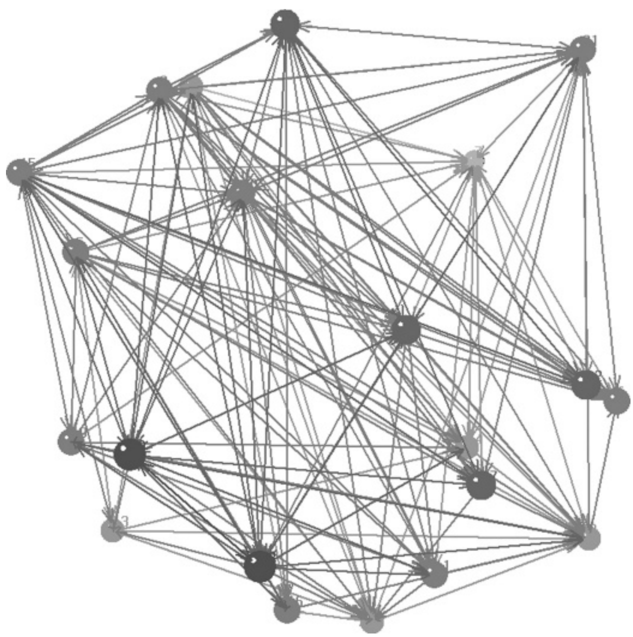


Figure A3.13. Mage's Nonmetric MDS Graph of Krackhardt Advice Network (3-D)

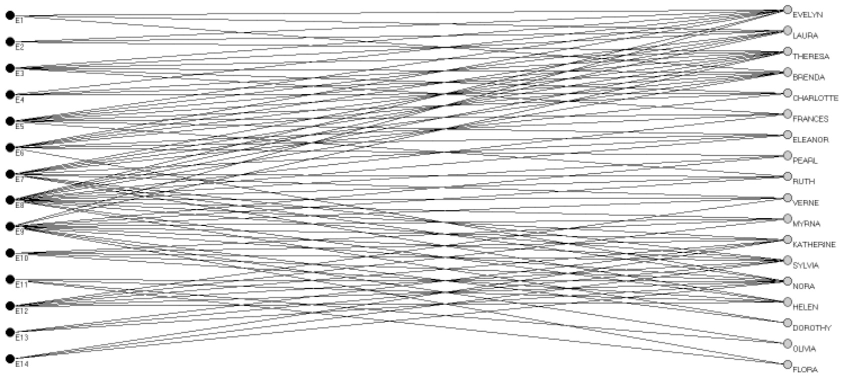


Figure A3.14. Bipartite Graph of Davis's Southern Women

visualized in NetDraw, and Figure A3.13 presents the nonmetric MDS three-dimensional sociogram as visualized in Mage.

Visual Representation of Two-Mode Networks

Two-mode data present additional visualization complexities. To illustrate these, we will use Davis's Southern Club Women (Breiger 1974; Davis et al. 1941), which we examined in Chapter 4. Depending on how we manipulate the data, we can use UCINET to visualize two-mode networks in a variety of ways. We can, of course, convert the data to one-mode (actors or events) data and visualize them using the techniques we discussed. Alternatively, we can visualize the original two-mode network: There are a number of approaches (Borgatti and Everett 1997; Everton 2004). A common approach is to use correspondence analysis, but Borgatti and Everett (1997:247) argue that in correspondence analysis representations of two-mode data that depict the distances between nodes are not Euclidean (i.e., the distances do not necessarily reflect social distance; see, however, Roberts 2000). As such, they recommend that we first convert the two-mode network to a bipartite graph, from which we compute the geodesic distances between all pairs of nodes, which we then submit to MDS techniques (Borgatti and Everett 1997: 249–251).

What is a bipartite graph? "A graph is bipartite if the vertices may be partitioned in exactly two mutually exclusive sets such that there are no ties wholly within either set – i.e., the endpoints of every tie come from different sets" (Borgatti and Everett 1997:247–248). See Figure A3.14. To create a bipartite graph from a two-mode network, we use UCINET's *Transform>Bipartite* command. This brings up a dialog box (Figure A3.15) that requires us to indicate which two-mode network we want to transform. It is important to note that you will want to tell

[UCINET] Transform >Bipartite

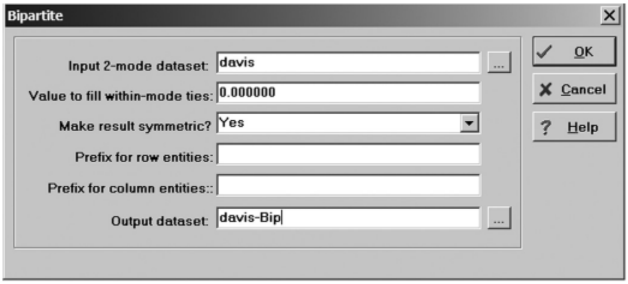


Figure A3.15. UCINET's Bipartite Dialog Box

UCINET to make the resulting graph *symmetric* – this is not UCINET's default option. If you do not change the option to symmetric, you will not be able to calculate the geodesic distance between nodes in the following step. Clicking "OK" will generate a bipartite graph/matrix (not shown) should be a symmetric, one-mode graph with thirty-two rows and columns (eighteen women + fourteen events).

Recall that geodesic distance refers to the length of the shortest path between two nodes. To calculate this in UCINET we use the *Network>Cohesion>Distance* command. In the resulting dialog box (Figure A3.16) indicate that the input dataset is the bipartite network calculated earlier, accept UCINET's defaults, and click "OK." This produces a distance matrix (not shown). If you examine the matrix closely you will note that the geodesic distances between any two women or between any two events is never less than two (Borgatti and Everett 1997:249; Faust 1997). This is because the women are only connected to one another through events and the events are only connected to one another through women, so it always takes at least two steps to get from one woman to another or from one event to another.

[UCINET] Network >Cohesion >Distance

The next step is submitting this distance matrix to the MDS routines discussed earlier. Nonmetric MDS models yielded the best measures of fit (two-dimensional = .213; three-dimensional = .153) and are displayed in Figures A3.17 and A3.18.

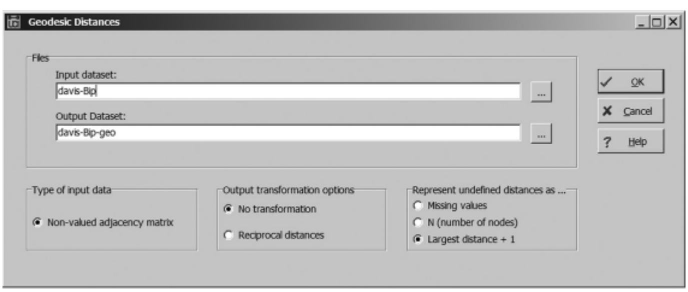


Figure A3.16. UCINET's Geodesic Distance Dialog Box

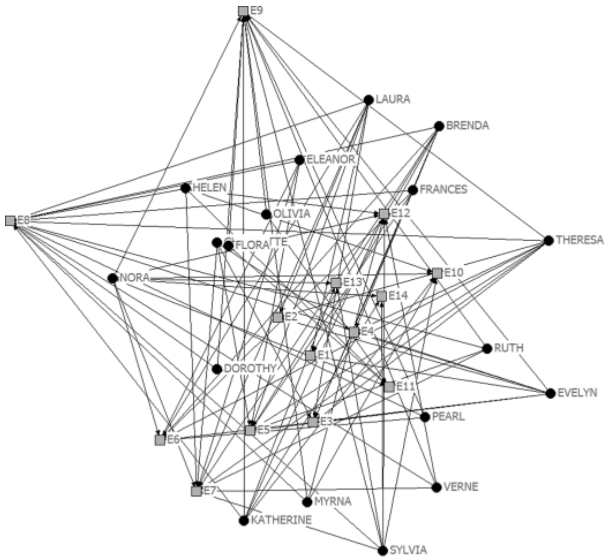


Figure A3.17. NetDraw's Geodesic MDS Graph of Davis's Southern Women (2-D)

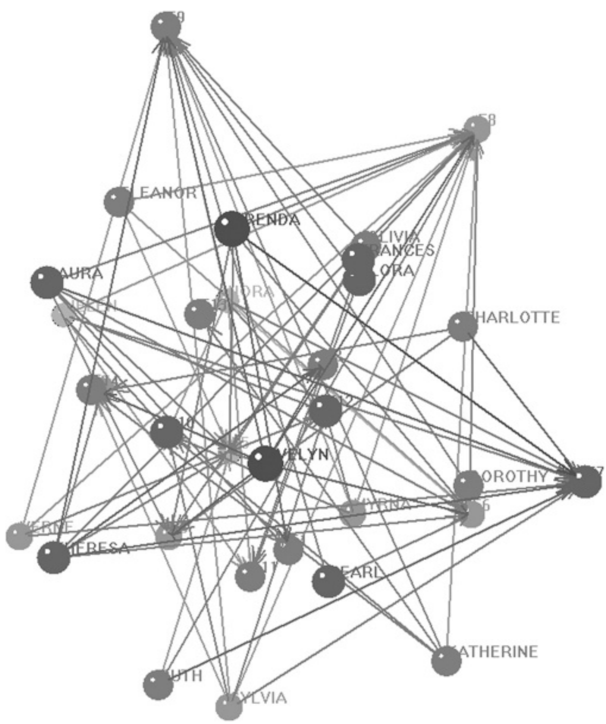


Figure A3.18. Mage's Geodesic MDS Graph of Davis's Southern Women (3-D)