

Software and Data for Social Networks Seminar

Building a survey with Hansen's software	3
Accessing UCINET V	3
KliqueFinder	4
Other Visualization tools	4
P-star from Wasserman's web site:	5
p-star using SAS	6
p2)	7
influence	8
Conversion of data	9
Data sets available	10

Building a survey with Hansen's software

To build a survey, go to

<http://sndcadmin.tanglewood.net/>

I'll give you a log in name and password

(At one point, you are asked for a list of names, be sure to include no spaces in text file that has the names).

To take or complete a survey, go to

<http://sndc.tanglewood.net/>

Accessing UCINET V

Go to <http://www.analytictech.com/downloaduc6.htm> and follow instructions

Or,

We have contracted for 6 licenses with the company. Thus you may download it onto a lab, but I ask you not to download it at home.

To download:

- 1)ftp (binary) all files from ucinet2 to a directory you name on the c drive
- 2)go to the directory on the c drive and run setup.exe (associated with a dos icon that looks like a computer)
- 3)run UCINET V from the directory in which it was installed. The button may appear under the start-up icon

Along with the software, you will get several datasets in a folder called ucinetdata. You will also get a draft of a user guide.

If you want a copy for your home, go to the website:

<http://www.heinz.cmu.edu/project/INSNA/>

and go to software/ucinet

It's not expensive, and Steve Borgatti, the developer, deserves some compensation for his work.

ucinet has the capacity to run many types of analyses -- clustering, centrality, etc. Note that most do not involve both the set of ties as well as characteristics of actors -- the procedures operate primarily on the social network matrix.

KliqueFinder

In Windows:

Go to pikachu.harvard.edu/wkf and enter kf as the logon. Ask me personally for the password. Follow directions to install.

1) Under basic specs, do browse to choose the directory in which you want to work. Note that "Browse" cannot create a new directory.

2) setup file

choose "basic setup" and then "Run setup file"

3) Choose a data file (either your own or one of the examples)

4) choose options as described in this manual

5) click "run analysis"

6) look at clusters output, confirm evidence of clusters
to make the picture

7) click "make graph"

8) make modifications based on contents of this manual

9) save

10) click run sas (at the moment, this doesn't work perfectly. Sooo

Open sas

a) from the program editor, open the file socgramz.sas in the directory in which you were working above.

B) run

10) view file "socgramx.eps" using sas, word perfect, adobe, or ghostwriter.

You may have to edit socgramz.sas (as described in this manual) to modify plots.

Other Visualization tools

Netminer: http://www.netminer.com/NetMiner/home_01.jsp

Graphviz: http://www.research.att.com/sw/tools/reuse/packages/graphviz_1_0.html

MovieMol (chemistry based, longitudinal) <http://www.kvac.uu.se/~lars/moviemol.html>

Krackplot: <http://www.andrew.cmu.edu/~krack/>

Pajek: <http://vlado.fmf.uni-lj.si/pub/networks/pajek/default.htm>

P-star from Wasserman's web site:

<http://kentucky.psych.uiuc.edu/pstar/index.html>

to format data for prepstar:

in Windows

```
set printo(16)=3  
run mydata.list
```

in unix:

```
setup prepstar  
kcliqfind mydata.list
```

file mydata.pst is ready to be read by prepstar

[note that if the prefix for your original list is less than 6 characters, Kliquefinder will call the output file xxx.li.pst. This will have to be renamed xxx.pst to be read by prepstar].

In responding to prepstar, typically you will indicate

```
DIAGONAL IS INCLUDED (Y)  
DATA ARE DIRECTIONAL (D)  
BLOCK MODEL INFORMATION IS SUPPLIED (Y)  
COMPARATIVE MATRIX (N)
```

As an example:

I ran KliqueFinder on schelp.list (after setting print option 3 to 16)

This produced schelp.pst

(If you get a * in the output it is because some was more than 1 space – should be corrected in new version. If not, figure out what the value should be – either cluster size or a value from the network, and replace).

run prepstar on schelp.pst and specified output as SCHELP.PDA

I then edited SCHELP.PDA to remove the beginning informational output

I then used readhps.sas to read in the data from SCHELP.PDA

readhps also merges the data with another data file (or see influence.sas).

p-star using SAS

In Windows

in sas,

- 1) run: newlib1 be sure to designate the correct path name for your data files
- 2) [you must change the variable *numpeop* to correspond to the number of people in your network;
- 3) run pstar

In Unix

```
% setup pstar
```

```
% emacs pstar.sas
```

[you must change the variable *numpeop* to correspond to the number of people in your network;
ctrl-x, ctrl-c to exit emacs]

```
% sas pstar
```

pstar reads a file called *pstar.list* which contains data in chooser, chosen, weight of tie format. This is the same type of file that KliquesFinder reads. Remember that 99999 indicates group membership.

Pstar produces *pstar.log* containing errors, warnings, etc., *pstar.lst* containing output including a full pstar model, and *pstar.dat* containing data to be read by another package. For code used to produce this, as well as format of output file, see the file called pstar.sas.

Making dyadic attributes

The program diffattb.sas makes dyadic level attributes, such as difference in a belief, or whether or not two actors have the same attribute.

p2)

1) Run HGLM2 (awkward interface, but ask me)

or

2) maybe p2explore + glimmix (from Sam Field)

wo programs, the glimmix macro (in case you don't have it) and the program that generates p2 data from a random draw and runs the glimmix macro.

make sure you alter the include statement to reflect the location of the macro on your computer before running the program. There may be other references to my directory structure in the program as well (not many) that you will have to debug if you are going to run the program.

By the way, I have been running some simulations and it appears that failing to account for correlated sender and receiver effects (glimmix models do not) results in biased estimates of the reciprocity parameter and, possibly, dyad level covariates.

3) **Preferred?**

Maritje Van Duijn's version at:

<http://stat.gamma.rug.nl/stocnet/>

Stocnet reads a standard social network matrix

Peter Hoff in R

SOFTWARE: <http://www.stat.washington.edu/hoff/Code/GBME/>

influence

sas program

influenceb.sas creates influence variables using proc means and merges, may be easier to follow

Or, using IML:

In windows

- 1) copy stor6.sas7bcat or use sas to run newlib1 in your data directory
- 2) put network data in *influence.list*
- 3) put individual "y" data in *yvar.list*
- 4) in influence.sas you must change the variable *numpeop* to correspond to the number of people in your network
- 5) in sas, run influence.sas

In unix

- % setup influence
- % emacs influence.sas

[you must change the variable *numpeop* to correspond to the number of people in your network; ctrl-x, ctrl-c to exit emacs]

% sas influence

influence.sas reads from two files: *influence.list* and *yvar.list*.

It produces *influence.log* which contains errors, warnings, etc., *influence.lst* which contains output, and *influence.dat* which contains a list of individuals and their level of exposure in three contexts:

- 1) exposure within the subgroup -- the beliefs of the people a person talks to within the subgroup
- 2) exposure outside of the subgroup -- the beliefs of the people a person talks to outside the subgroup
- 3) total exposure -- sum of all beliefs of the people a person talks to

The variables in *influence.dat* are: id, influence from within the subgroup, influence from outside the subgroup, total influence.

Influence.sas also produces *attdiff.dat*, which is a pair level data set consisting of absolute values of differences on y-var.

You can see how the code for this works by looking at *newlib1.sas*. If you want to change something about the way these variables are computed, you would have to modify *newlib1.sas*, then run *newlib1.sas* and rerun *influence.sas* by typing:

```
% sas newlib1
% sas influence
```

Recombining with existing data

At the end of *Influence.sas*,
influence.dat is read in and merged with *yvar.list* to create the permanent data set *withinfl*.

attdiff.dat is read in and merged with *influence.list* to create the permanent data set *withdiff*.

Conversion of data

A. KliqueFinder ÷ UCINET
% setup ucinet
% kliqfind mydata.list

There is now a file called *ucwn.DAT* that contains information to be read by UCINET. To read this, open with the spreadsheet with ucinet (VI) and delete the first two lines and then save.

In the old ucinet (V) , go to datasets/import/ucinet3.0 . Enter the filename containing the data (probably *ucwn.DAT* unless you renamed it), indicate the data are in integer format, and designate an output file name.

B. KliqueFinder ÷ STOCNET
same procedure as above, but open with STOCNET instead of UCINET

C. UCINET ÷ KliqueFinder

- 1)run UCINET
 - 2)go to datasets/export/raw
 - 3)indicate data set to export. Consider these defaults:
 - diagonals: absent
 - decimal places: 0
 - fieldwidth: 9
 - Guarantee space: yes
 - page width: 255
 - embed row labels: no
 - embed column labels: no
 - embed level variables: no
 - indicate output data set
 - 4)ftp output data set to edstat
- This data set should be in a format ready to read directly into KliqueFinder

Data sets available

A. French financial elite (look in data/ffe directory)

Jeff Yasumoto and I analyzed these data in the paper starting on page 335 of the coursepack. They consist of friendships and transactions among the French financial elite. Basic list files are in: *ffe.list* for friendship ties, *hostile.list* for hostile actions, and *suppor.list* for supportive actions

Sas background data are in a sas data set:
ffe.ssd01

A table linking the ID numbers in your data and those used in the paper is in *nicetab3.lst*, along with other characteristics of the people.

UCINET data files are in **.###h* and *#.###d* -- you need both the *h* and the *d* extension to call up in UCINET.

B. "Our Hamilton High" --ourham (look in data/ourham)

These are the data I analyzed in Frank (1995), and Frank (1996) -- see pages 187-189, 193-22 in the coursepack. They consist of professional discussions among teachers in a high school.

The basic list file is in *ourham.xlist*.

Characteristics of individuals are in a sas file called *background.ssd01*.

A table linking ID's in the data to those reported in the article is in *nicetab3.lst*.

UCINET data sets are in *magent.###h*, *magent.###d* (moral agency data) and *ourham.###h* and *ourham.###d* (professional discussions among teachers).

C) Academic awareness of participants in a social capital conference

At a conference at MSU in April, 1998, we had many internationally known researchers in the field of social capital (Etzioni, Putnam, Fukuyama), as well as in education (Valerie Lee, Charles Bidwell, Bo Bileau, Aage Sorensen, Ralph McNeal, Barbara Schneider, Steve Morgan). The basic list consists of who was aware of whom academically. We also developed measures of each person's definition of social capital. We have limited background information. I have a 5 page write up on this if you are interested.

Basic list file is in *sammy1.list* and *sammy2.list*. People's ideas about the definition of social capital are in *socdef.ssd01*.

D) Interaction & influence among those participating in and studying the non-profit sector

The Kellogg foundation has funded several centers for the study, development, and implementation of programs in the non-profit sector. One of their goals is to bring academic and practitioners closer together. Kellogg has funded conferences, web pages, and other mechanisms for bringing people together. We have obtained information at 2 (soon to be 3) time points regarding who interacts with whom, as well as a measure of each person's sense of the extent to which practitioners and academics must coordinate their actions. We also have background information.

E) Coordination among social service agencies

Large changes in the social service and mental health sectors have had an impact on how social service agencies coordinate care for patients. The Robert Wood Johnson Foundation and other foundations have funded several studies of how these agencies are reaction. We have access to two sets of data indicating the extent to which agencies coordinate action, give referrals to one another, share information, etc.

F) Relationships among test items

Mark Reckase has data on relationships between pairs of items based on graphical distances. Kyle Fahrback has conducted preliminary analyses to identify subgroups, but no picture has been made (250 items), and other info about the items has not been linked in.

G) Schools like “Our Hamilton High”

In addition to the data in “Our Hamilton High” we have data from up to 50 other schools, 14 of which I am relatively familiar with. There are also some case study materials available, and student achievement for some of the schools. Might take a little to coordinate use of these data, but it can be done.

H) UCINET data sets

a variety of data sets in UCINET format from some famous analyses.

I) Other data on the web

Check out: <http://www.heinz.cmu.edu/project/INSNA/>
and go to information about social network data sets.