

SNA

Homophily, faction, and coreness

Grouping

- **Grouping/partition as owning an attribute**
 - **Pre-defined**
 - **Exogenous** (from the Greek words "exo" and "gignomi", meaning "outside" and "to come to be") refers to an action or object coming from outside a system.
- **Grouping/partition as interconnectedness**
 - **Emergent**
 - **Endogenous** substances are those that originate from within

Homophily

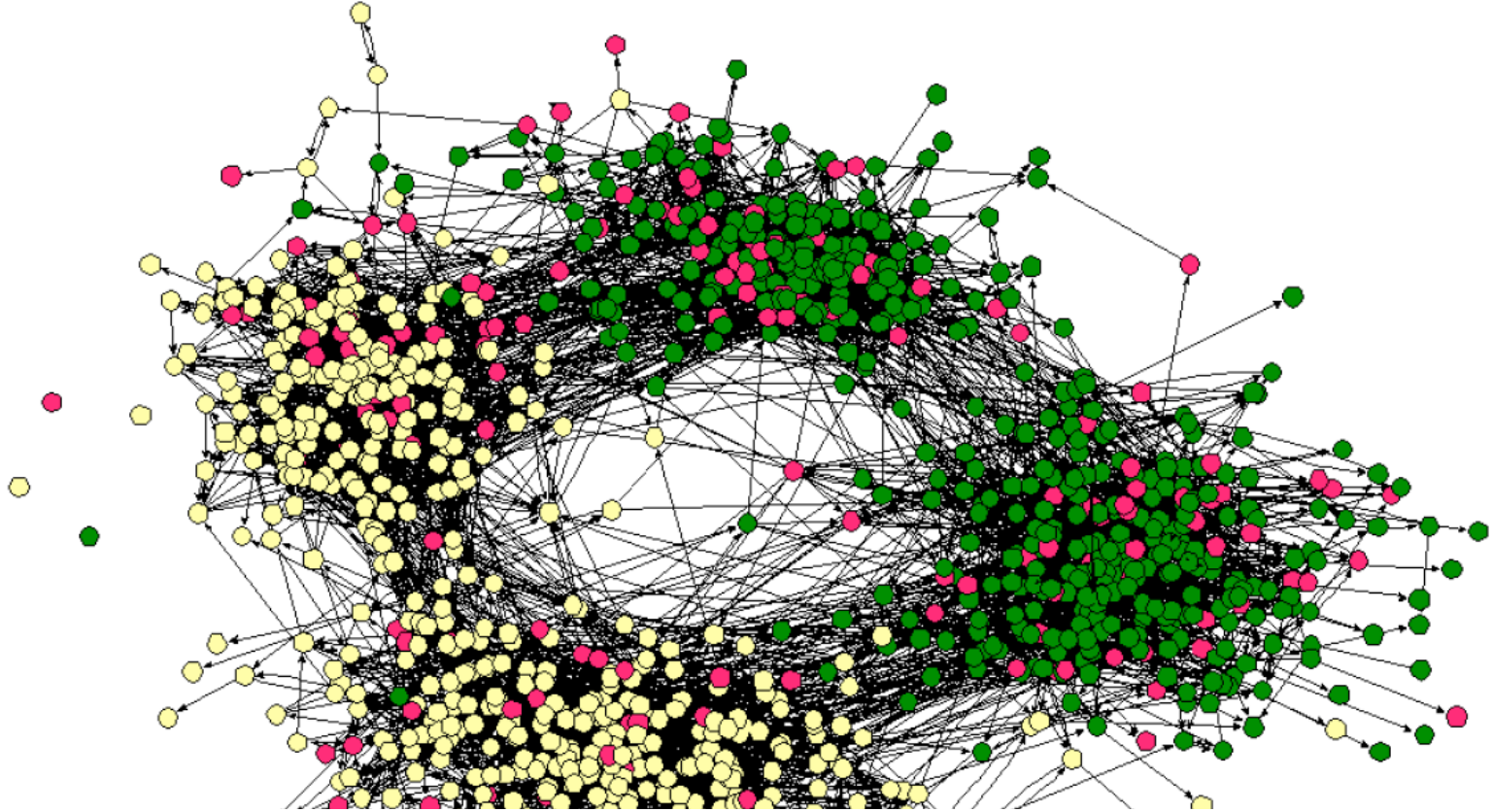
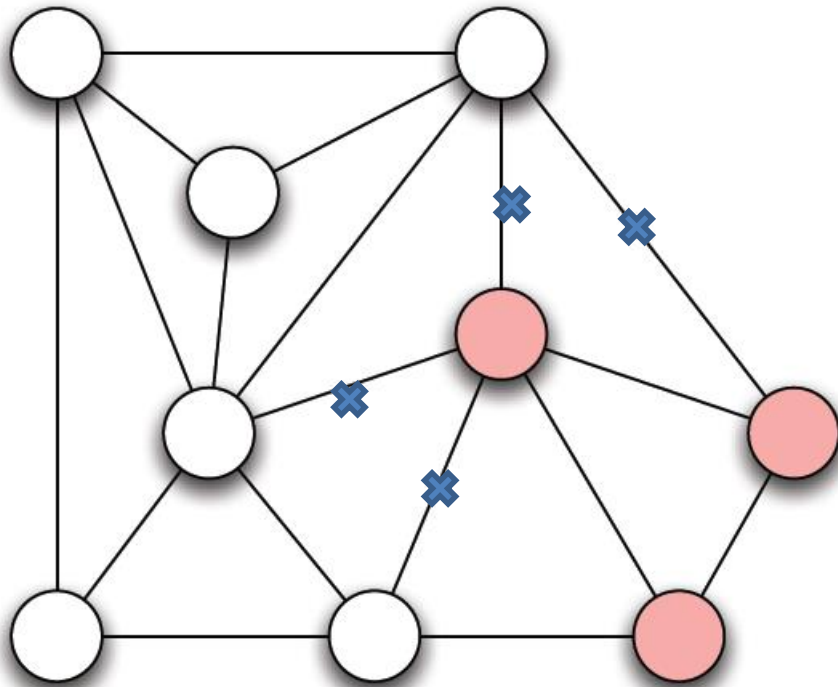


Figure 4.1: Homophily can produce a division of a social network into densely-connected, homogeneous parts that are weakly connected to each other. In this social network from a town's middle school and high school, two such divisions in the network are apparent: one based on race (with students of different races drawn as differently colored circles), and the other based on friendships in the middle and high schools respectively [304].

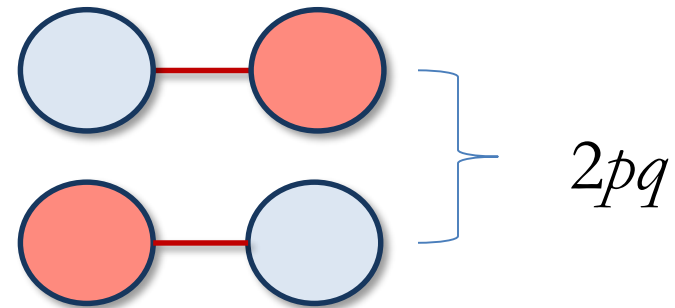
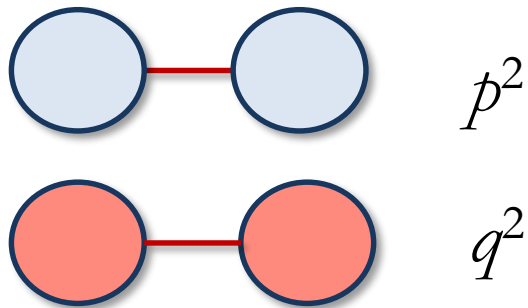
Homophily test

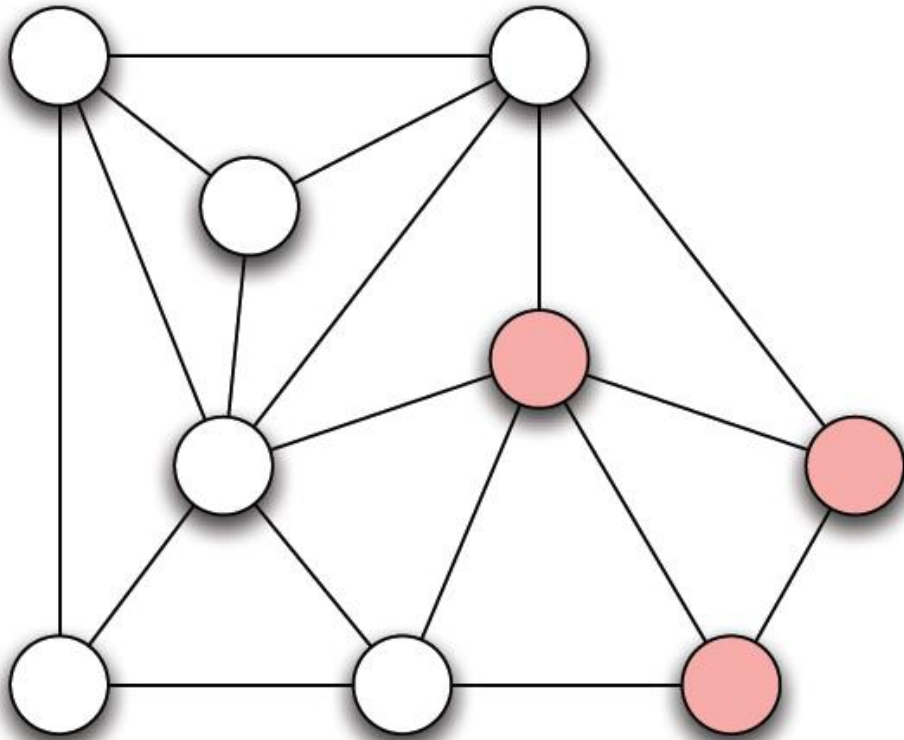


p fraction are male (white node),
q fraction are female (pink node)

Randomly Generated Network

- Probability of both ends of edge being male: p^2
female: q^2
- male on one end and female on other: $2pq$





Homophily Test: If the fraction of heterogeneous (cross-gender) edges is significantly less than $2pq$ then there is evidence for homophily

Cross-gender edges:
5 of 18

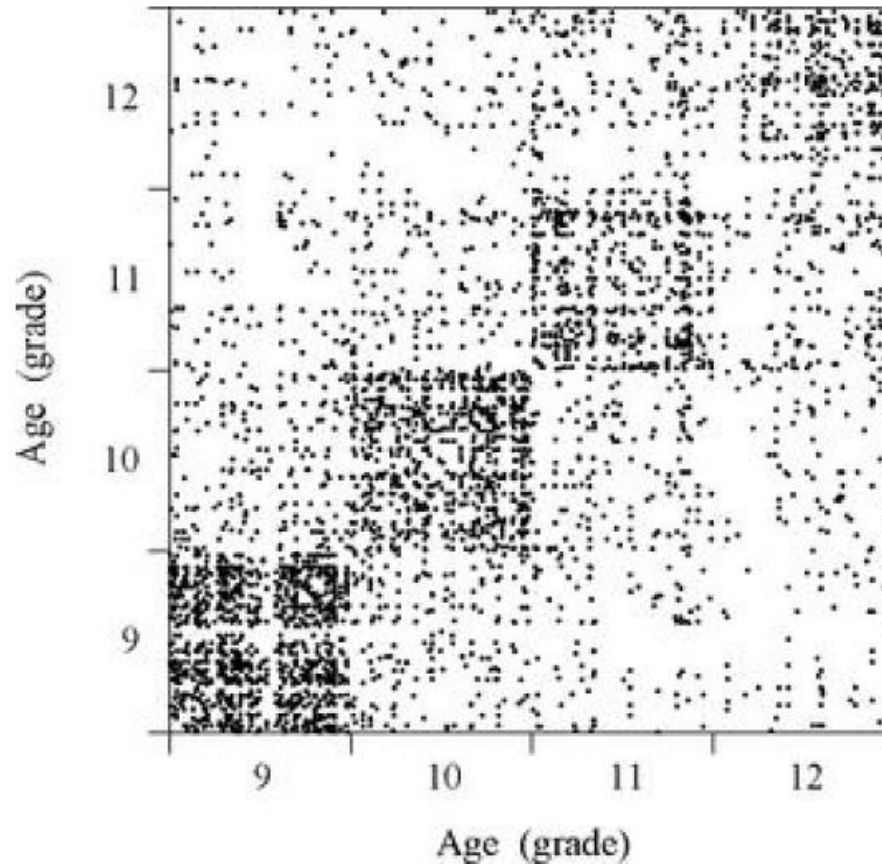
$$p = 6/9 = 2/3$$

$$q = 3/9 = 1/3$$

If no homophily, # of cross-gender edges should be $2pq = 4/9 = 8$ out of 18

\therefore Evidence of homophily

Homophily: correlation between network and attribute variables



Homophily test with UCINET

- **NETWORK>COHESION>HOMOPHILY**

1. Make a partition of a network into a number of mutually exclusive groups
- 2. The E-I index , which ranges from 1 to -1 and can be seen as a measure of the extent a group chooses themselves a value of -1 showing homophily and a value of +1 showing heterophily.
- 3. For valued data it is the sum of the tie strengths instead of the number of ties.
- 4. Find the correlation between a network and an attribute variable by correlating the dichotomous variable "is in the same group as" with the corresponding entry in the data matrix.

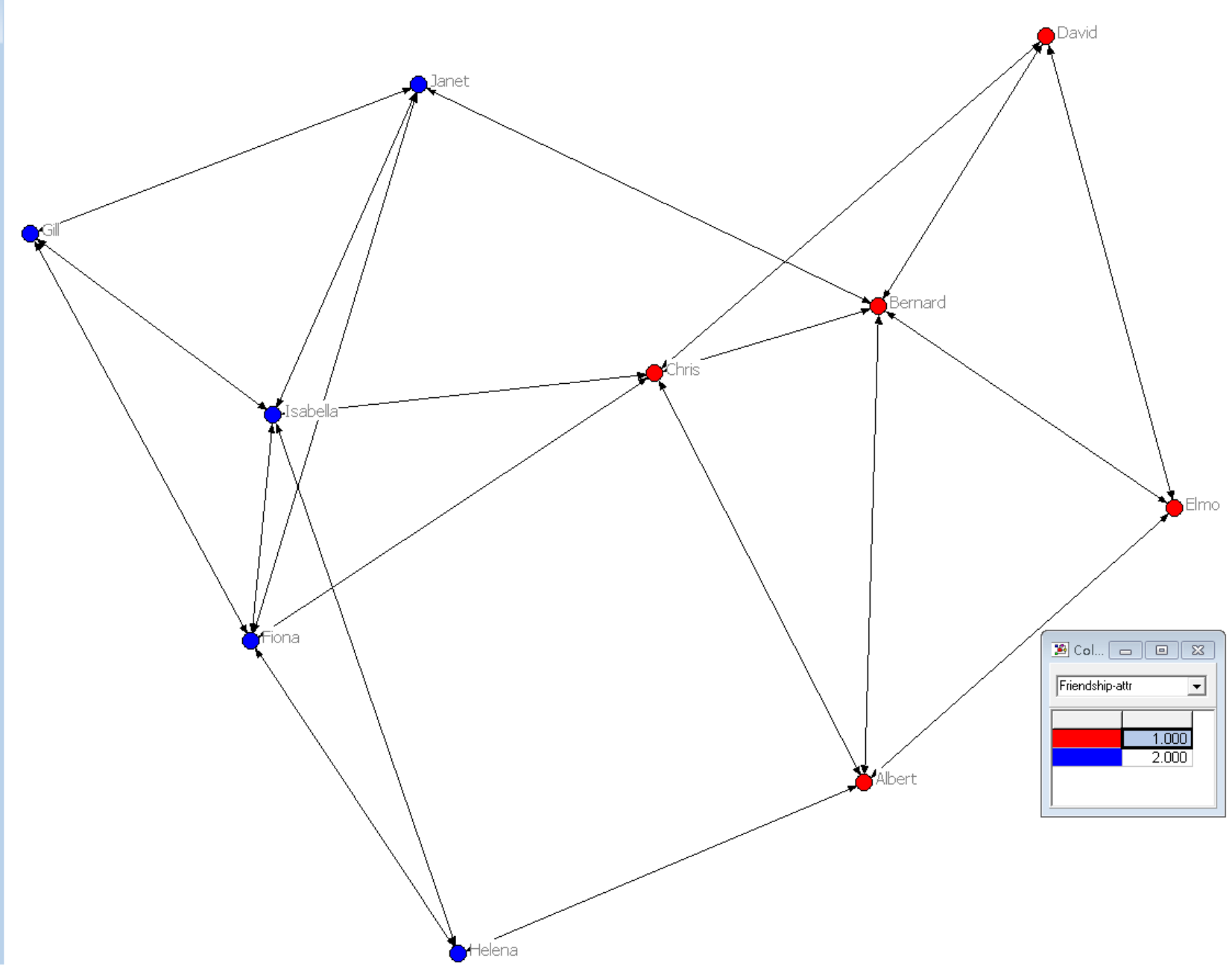
Exercises

- Use the excel file friendship and friend-attr files to run both

NETWORK>COHESION>HOMOPHILY
AND

Network>Cohesion>E-I Index

Then compare their results



Relations:

- ☒ Friendship

Col...

Friendship-attr

	1.000
	2.000

Dn Up Cl All R

Size:

> 0.0 + -

Save As New Relation

Options:

☐ AND ☒ OR Inc:

☒ Self-Loops Dec:

☒ Link wts ->

☐ Width ☐ Color

NODE PARTITION

Block	Value	Freq	Members:
1	1	5	Albert Bernard Chris David Elmo
2	2	5	Fiona Gill Helena Isabella Janet

No. of ties within and between groups

	1	2
1	16	4
2	4	16

Whole Network Homophily Measures

	1	2
E-I	Ind	PBSC
Matrix 1	-0.6000	0.6400

PBSC = point biserial correlation.

Density matrix

		1	2
		1	2
1	1	0.800	0.160
2	2	0.160	0.800

40 ties.

Whole Network Results

		1 Freq	2 Pct	3 Possibl	4 Density
1	Internal	32.000	0.800	40.000	0.800
2	External	8.000	0.200	50.000	0.160
3	E-I	-24.000	-0.600	10.000	0.111

Max possible external ties: 50.000

Max possible internal ties: 40.000

E-I Index: -0.600

Expected value for E-I index is: 0.111

Max possible E-I given density & group sizes: 1.000

Min possible E-I given density & group sizes: -1.000

Re-scaled E-I index: -0.600

Permutation Test

Number of iterations = 5000

		1 Obs	2 Min	3 Avg	4 Max	5 SD	6 P >= Ob	7 P <= Ob
1	Internal	0.800	0.200	0.446	0.800	0.091	0.009	1.000
2	External	0.200	0.200	0.554	0.800	0.091	1.000	0.009
3	E-I	-0.600	-0.600	0.108	0.600	0.181	1.000	0.009

E-I Index is significant (p < 0.05)

Group level E-I Index

		1	2	3	4
		Intern	Extern	Total	E-I
1	1	16.000	4.000	20.000	-0.600
2	2	16.000	4.000	20.000	-0.600

Individual Level E-I Index

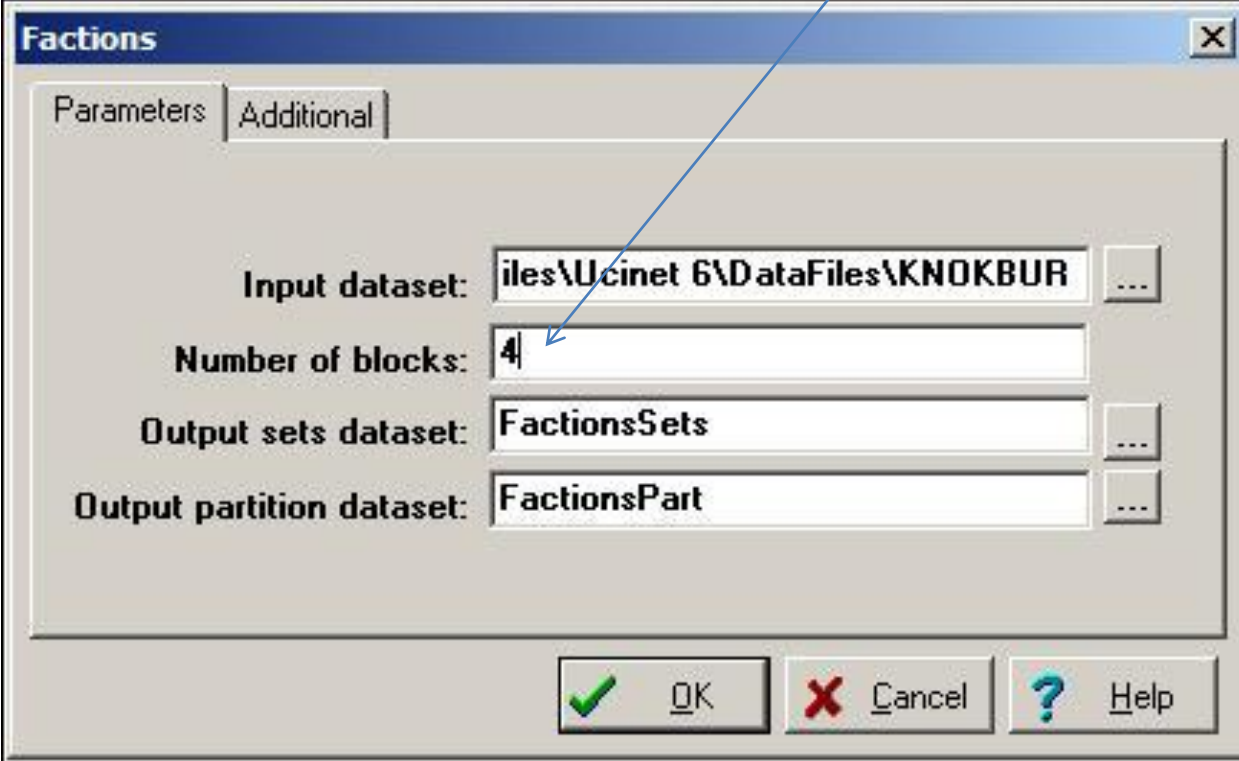
		1	2	3	4
		Intern	Extern	Total	E-I
1	Albert	3.000	1.000	4.000	-0.500
2	Bernard	4.000	1.000	5.000	-0.600
3	Chris	3.000	2.000	5.000	-0.200
4	David	3.000	0.000	3.000	-1.000
5	Elmo	3.000	0.000	3.000	-1.000
6	Fiona	4.000	1.000	5.000	-0.600
7	Gill	3.000	0.000	3.000	-1.000
8	Helena	2.000	1.000	3.000	-0.333
9	Isabella	4.000	1.000	5.000	-0.600
10	Janet	3.000	1.000	4.000	-0.500

Individual E-I values saved as dataset C:\Users\assistant\Desktop\0326Matrix\IndE-I

Factions (Endogenous grouping)

- Factions: a subgroup in which each person is closely tied to all others within the group and no or few connections at all to other sub-groups
- Network > Subgroups > Factions
 - Maximize **within-group** connections/density, minimize **between-group** connections
 - Errors (sum of **zeros within**, and **ones between**, i.e., non-diagonal blocks)
 - **Use in an exploratory manner** (i.e. trail and errors)

Need to specify the # of factions



The screenshot shows a Windows-style dialog box titled "Factions". It has two tabs: "Parameters" (selected) and "Additional". The "Parameters" tab contains four labeled input fields, each with a browse button (three dots) to its right:

- Input dataset:** The text box contains the path "iles\Ucinet 6\DataFiles\KNOKBUR".
- Number of blocks:** The text box contains the number "4". A blue arrow from the text "Need to specify the # of factions" points to this field.
- Output sets dataset:** The text box contains the text "FactionsSets".
- Output partition dataset:** The text box contains the text "FactionsPart".

At the bottom of the dialog box are three buttons: "OK" (with a green checkmark icon), "Cancel" (with a red X icon), and "Help" (with a blue question mark icon).

Number of factions: 4
 Measure of fit: Hamming
 Input dataset: KNOKBUR

Initial proportion correct: 0.567

...Badness of fit: 27.000
 ...Badness of fit: 27.000
 ...Badness of fit: 29.000

Final proportion correct: 0.700

Group Assignments:

1: 1 2 4 5 7 8
 2: 9
 3: 10
 4: 3 6

Density Table				
	1	2	3	4
1	0.83	0.67	0.17	0.17
2	0.50		0.00	0.00
3	0.67	0.00		0.50
4	0.42	0.50	0.50	1.00

The assignment of actors into 4 groups

Grouped Adjacency Matrix

Show within and between density

How to count classification
 “errors”/badness of fit?

(sum of zeros within, and
 ones between, i.e., non-diagonal blocks)

	1	2	8	4	5	7	9	10	6	3
	C	C	U	I	M	N	W	W	W	E
1	1				1	1	1			
2	1		1	1	1	1	1			1
8	1	1		1	1	1	1			
4	1	1			1	1				
5	1	1	1	1		1	1	1		1
7	1			1	1					
9		1			1	1				
10	1	1			1	1				1
6						1	1		1	1
3		1		1	1	1		1	1	

Number of factions: 3
 Measure of fit: Hamming
 Input dataset: KNOKBUR

Initial proportion correct: 0.656

...Badness of fit: 25.000
 ...Badness of fit: 25.000
 ...Badness of fit: 25.000

Final proportion correct: 0.722

Group Assignments:

1: 3 10
 2: 6
 3: 1 2 4 5 7 8 9

Grouped Adjacency Matrix

Density Table

	1	2	3
1	1.00	0.50	0.57
2	0.50		0.29
3	0.21	0.00	0.76

	3	0	6	2	4	1	7	8	9	5
E	W	W	C	I	C	N	U	W	M	
3	1	1	1	1	1	1	1	1	1	1
10	1		1	1	1	1	1	1	1	1
6	1					1	1			
2	1			1	1	1	1	1	1	1
4				1	1	1	1			1
1				1		1		1	1	
7				1	1					1
8				1	1	1	1		1	1
9				1		1				1
5	1	1		1	1	1	1	1	1	

Compare "goodness" of fit

Number of factions: 4
 Measure of fit: Hamming
 Input dataset: KNOKBUR

Initial proportion correct: 0.567

...Badness of fit: 27.000
 ...Badness of fit: 27.000
 ...Badness of fit: 29.000

Final proportion correct: 0.700

Group Assignments:

1: 1 2 4 5 7 8
 2: 9
 3: 10
 4: 3 6

Grouped Adjacency Matrix

Density Table

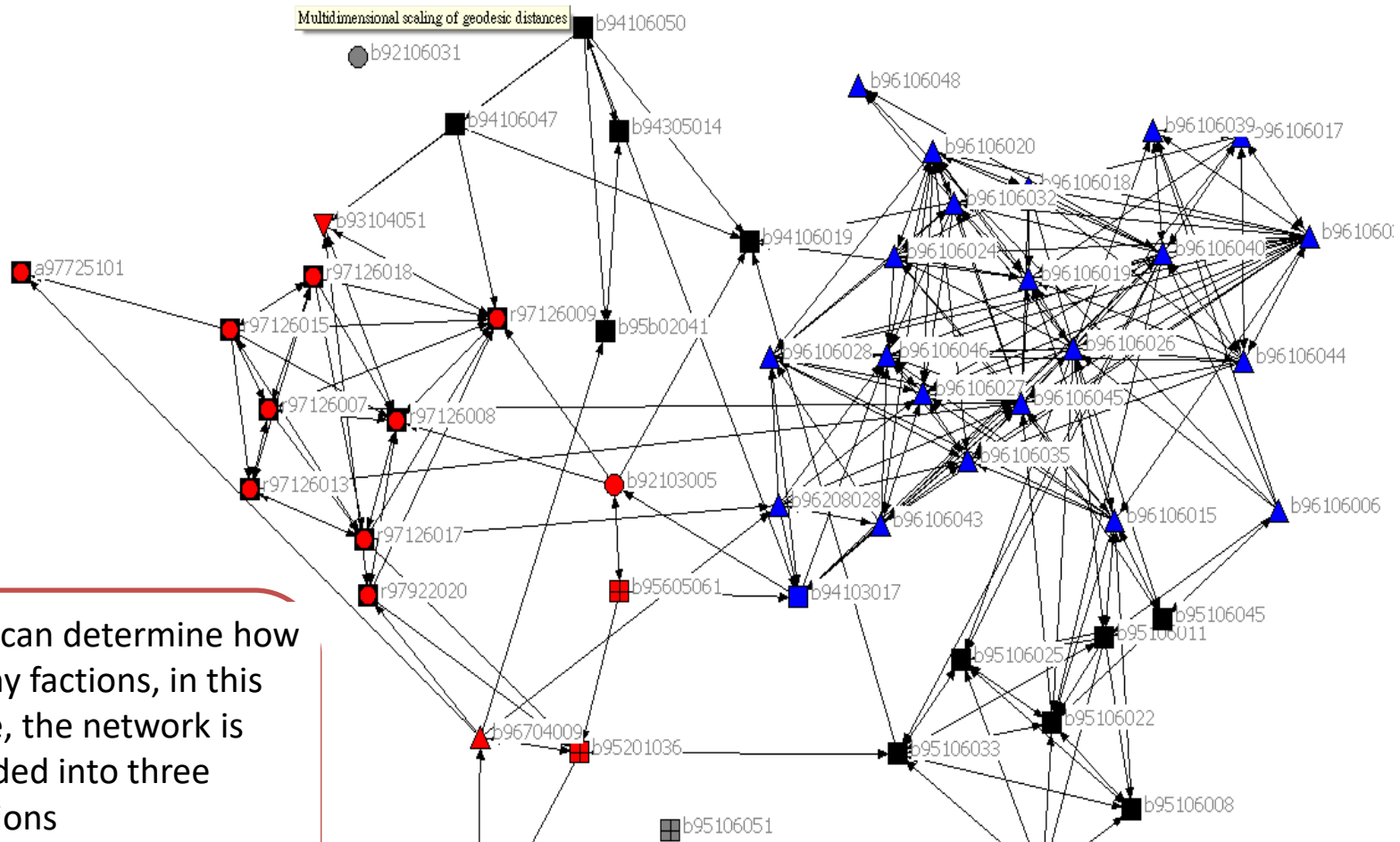
	1	2	3	4
1	0.83	0.67	0.17	0.17
2	0.50		0.00	0.00
3	0.67	0.00		0.50
4	0.42	0.50	0.50	1.00

	1	2	8	4	5	7	9	0	6	3
C	C	U	I	M	N	W	W	W	E	
1	1		1	1		1				
2	1	1	1	1	1	1			1	
8	1	1		1	1	1	1			
4	1	1		1	1					
5	1	1	1	1	1	1	1	1		1
7		1		1	1					
9		1		1	1					
10	1	1		1	1					1
6					1	1	1		1	
3		1	1	1	1		1	1	1	

Factions with Netdraw

- *Analysis>Subgroups>Factions* (select number).
- A "faction" is a part of a graph in which the nodes are more tightly connected to one another than they are to members of other "factions."

Subgroups: *Analysis* > *Subgroups* > *Factions*



You can determine how many factions, in this case, the network is divided into three factions

Exercises

- Try setting faction number at 2 and see the correction rate
- Try assign different faction number to the NTNU faculty data
- Try visualization of the NTNU data with Netdraw
 - *Analysis>Subgroups>Factions* (select number).

Core/periphery structures

- Classification of community members by connectivity
 - A dense, cohesive **core**, and
 - A parse, unconnected **periphery**, only connect to the core, but not to each other

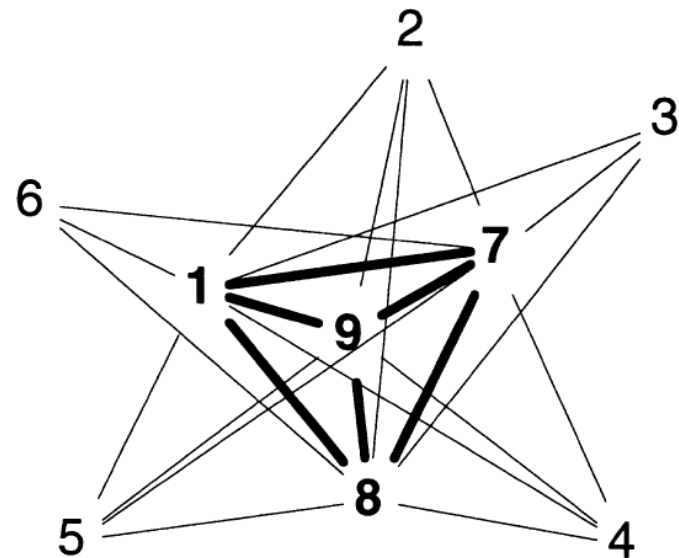
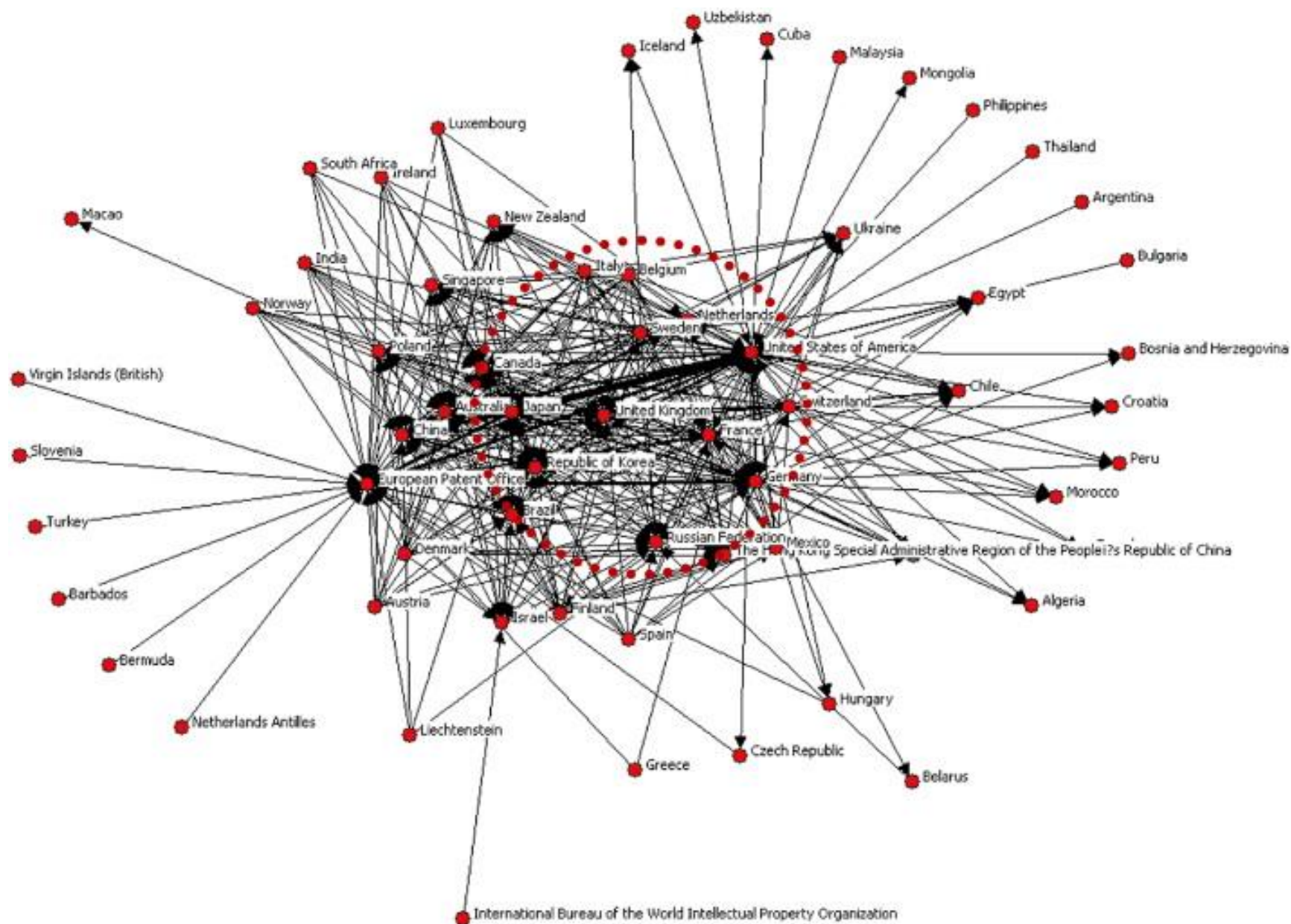


Fig. 3. Core/periphery structure.



Core/periphery

- To identify a core-periphery structure, we compare an **observed** block structure to an **ideal** block structure

Idealized core/periphery structure

	1	2	3	4	5	6	7	8	9	10
1		1	1	1	1	1	1	1	1	1
2	1		1	1	1	1	1	1	1	1
3	1	1		1	1	1	1	1	1	1
4	1	1	1		1	1	1	1	1	1
5	1	1	1	1		0	0	0	0	0
6	1	1	1	1	0		0	0	0	0
7	1	1	1	1	0	0		0	0	0
8	1	1	1	1	0	0	0		0	0
9	1	1	1	1	0	0	0	0		0
10	1	1	1	1	0	0	0	0	0	

Compare an observed block structure to an ideal block structure in order to identify the core/periphery structure

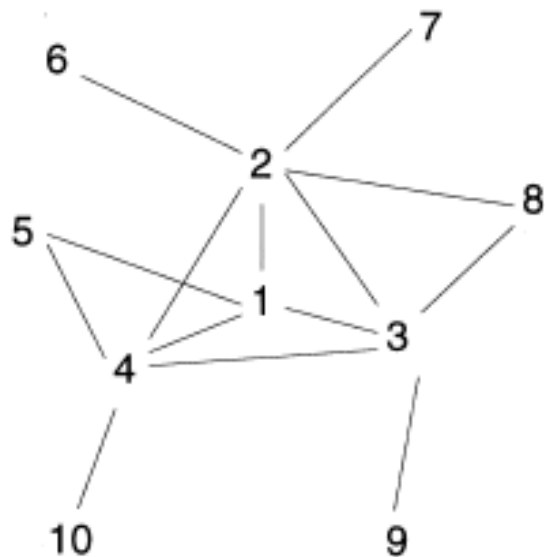


Fig. 1. A network with a core/periphery structure.

	1	2	3	4	5	6	7	8	9	10
1		1	1	1	1	0	0	0	0	0
2	1		1	1	0	1	1	1	0	0
3	1	1		1	0	0	0	1	1	0
4	1	1	1		1	0	0	0	0	1
5	1	0	0	1		0	0	0	0	0
6	0	1	0	0	0		0	0	0	0
7	0	1	0	0	0	0		0	0	0
8	0	1	1	0	0	0	0		0	0
9	0	0	1	0	0	0	0	0		0
10	0	0	0	1	0	0	0	0	0	

(observed blocked network)

(Ideal CP blocked network)

	1	2	3	4	5	6	7	8	9	10
1		1	1	1	1	1	1	1	1	1
2	1			1	1	1	1	1	1	1
3	1	1		1	1	1	1	1	1	1
4	1	1	1		1	1	1	1	1	1
5	1	1	1	1		0	0	0	0	0
6	1	1	1	1	0		0	0	0	0
7	1	1	1	1	0	0		0	0	0
8	1	1	1	1	0	0	0		0	0
9	1	1	1	1	0	0	0	0		0
10	1	1	1	1	0	0	0	0	0	

(observed blocked network)

	1	2	3	4	5	6	7	8	9	10
1		1	1	1	1	0	0	0	0	0
2	1			1	1	0	1	1	1	0
3	1	1			1	0	0	0	1	1
4	1	1	1			1	0	0	0	0
5	1	0	0	1		0	0	0	0	0
6	0	1	0	0	0		0	0	0	0
7	0	1	0	0	0	0		0	0	0
8	0	1	1	0	0	0	0		0	0
9	0	0	1	0	0	0	0	0		0
10	0	0	0	1	0	0	0	0	0	

A core periphery structure exists to the extent that **the correlation between the ideal structure and the observed structure is high**. We can search for cores by simply proposing a partition (many times) and then selecting the best fitting partition.

Core/periphery structures

- When a network cannot be subdivided into exclusive cohesive subgroups (such as components, cliques etc.)
- The assumption:
 - The network consists of just one group to which all actors belong to a greater or lesser extent

Core/periphery with UCINET

- **NETWORK > CORE/PERIPHERY > CATEGORICAL**

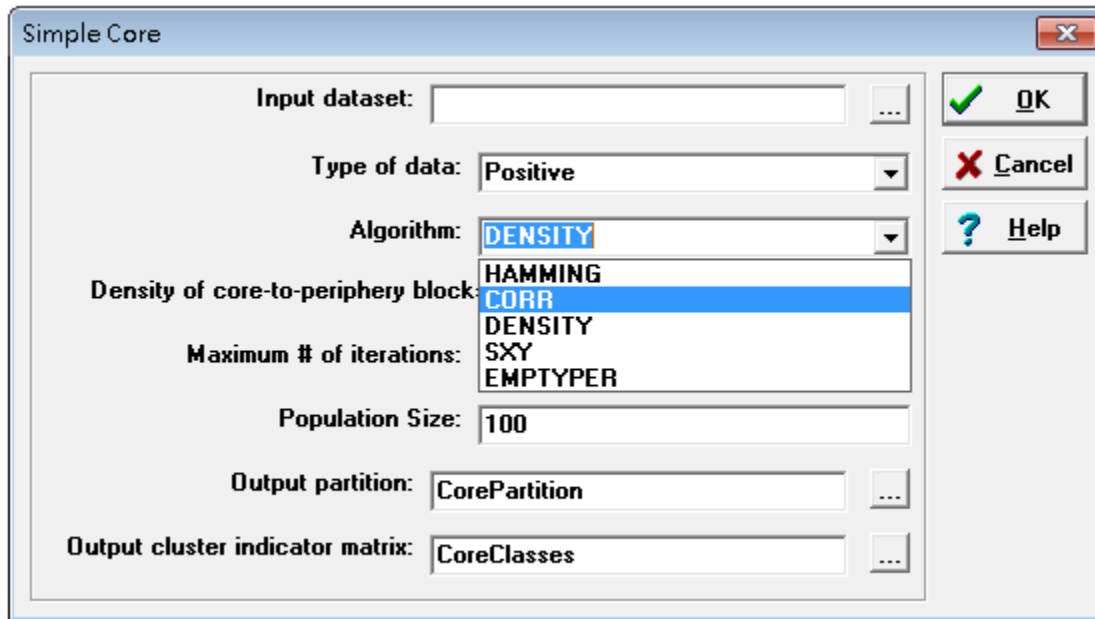
PURPOSE Uses a genetic algorithm to fit a core/periphery model to the data.

DESCRIPTION Simultaneously fits a core/periphery model to the data network, and identifies which actors belong in the core and which belong in the periphery.

Apply core/periphery in bibliometrics

[illegible]

Core/periphery with categorical data



A screenshot of a software dialog box titled "Simple Core". The dialog box contains several input fields and a list box. The "Input dataset:" field is empty. The "Type of data:" dropdown is set to "Positive". The "Algorithm:" dropdown is set to "DENSITY", and a list box below it shows "HAMMING", "CORR", "DENSITY", "SXY", and "EMPTYPER", with "CORR" selected. The "Density of core-to-periphery block:" field is empty. The "Maximum # of iterations:" field is set to "100". The "Population Size:" field is set to "100". The "Output partition:" field is set to "CorePartition". The "Output cluster indicator matrix:" field is set to "CoreClasses". On the right side, there are three buttons: "OK" (with a green checkmark icon), "Cancel" (with a red X icon), and "Help" (with a blue question mark icon).

Simple Core

Input dataset:

Type of data: Positive

Algorithm: DENSITY

Density of core-to-periphery block: HAMMING
CORR
DENSITY
SXY
EMPTYPER

Maximum # of iterations: 100

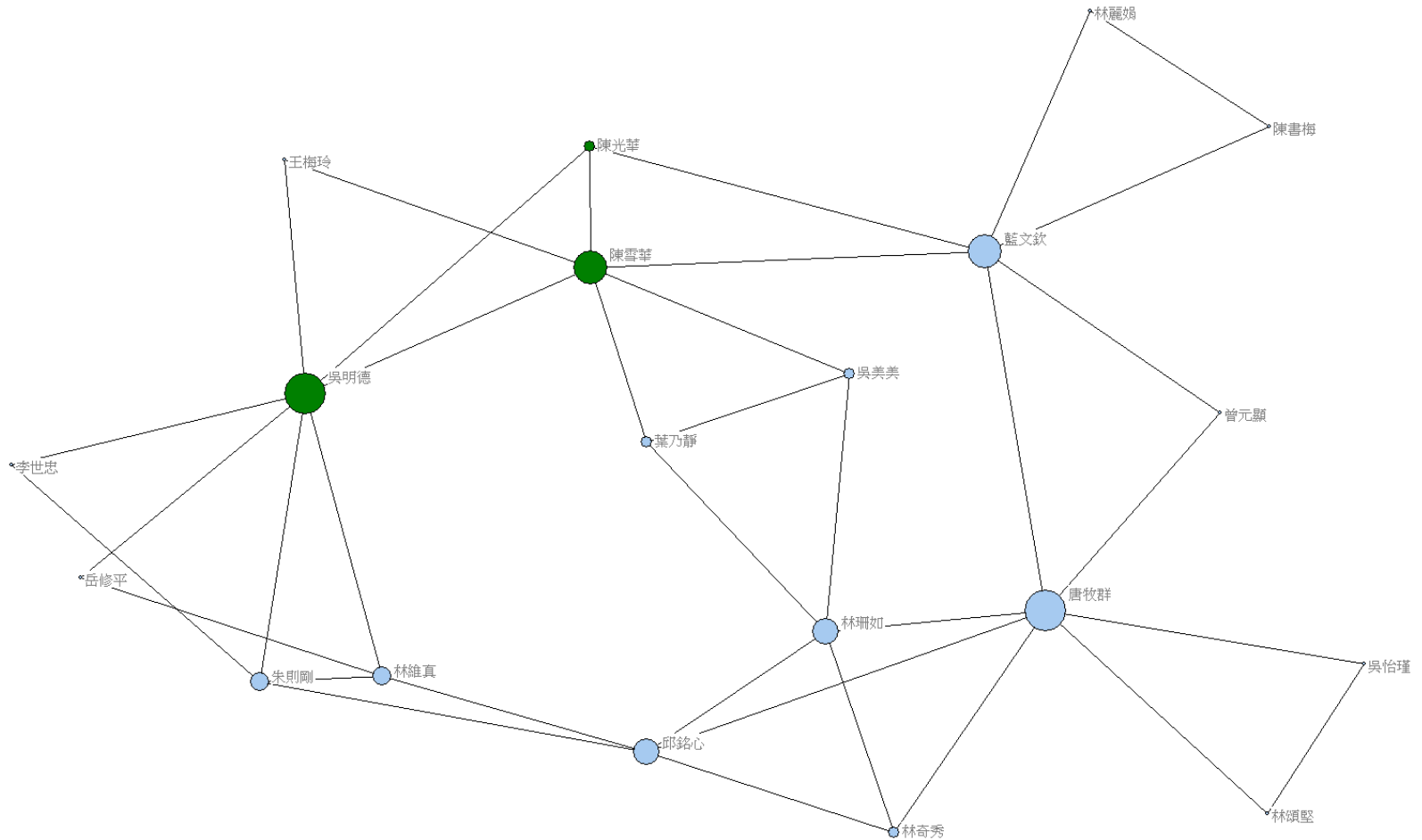
Population Size: 100

Output partition: CorePartition

Output cluster indicator matrix: CoreClasses

OK Cancel Help

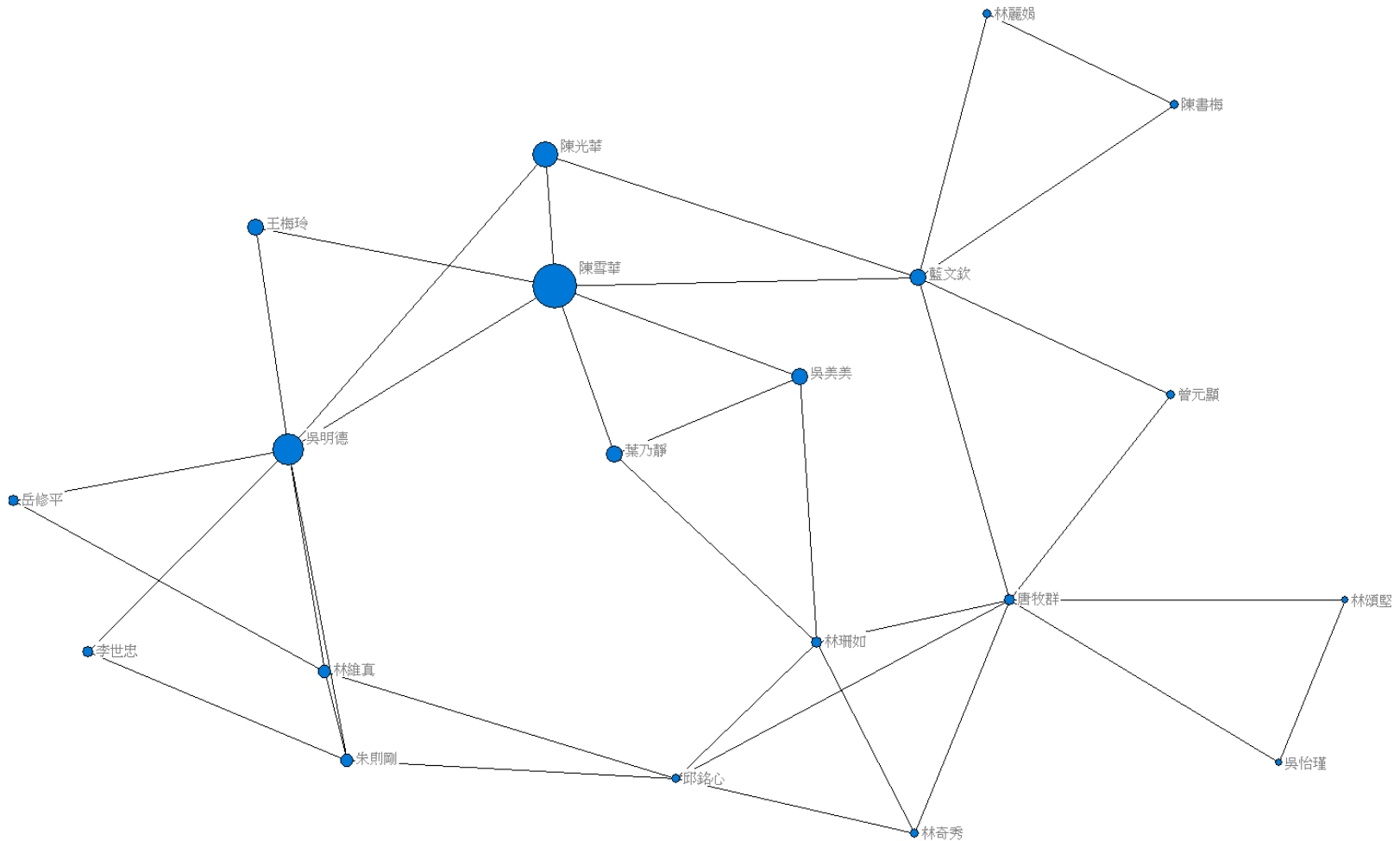
Try `ntu_TwomodRows ##d`



Performed with categorical core-periphery procedure

Size = degree

Color = core



Performed with continuous core-periphery procedure

Size = Coreness

More credits to the library director

UCINET Spreadsheet - C:\Users\assistant\Documents\UCINET da

File Edit Transform Fill Labels Options Help

Fill +.0 -.0 Ren

	PARTITION
黃慕萱	1
朱則剛	1
吳明德	1
林奇秀	2
藍文欽	1
林維真	2
陳雪華	1
謝寶媛	2
唐牧群	1
林珊如	1
陳光華	2
陳書梅	1
潘美月	2
盧秀菊	2
吳可久	2
鄭雪玫	2
卜小蝶	2
吳美美	2
陳昭珍	2
王梅玲	2
宋雪芳	2
林頌堅	2
莊道明	2
林呈濱	2
黃乾綱	2
吳玲玲	2
邱銘心	2

Make sure you know the location of these two output files

Coreness
Partition

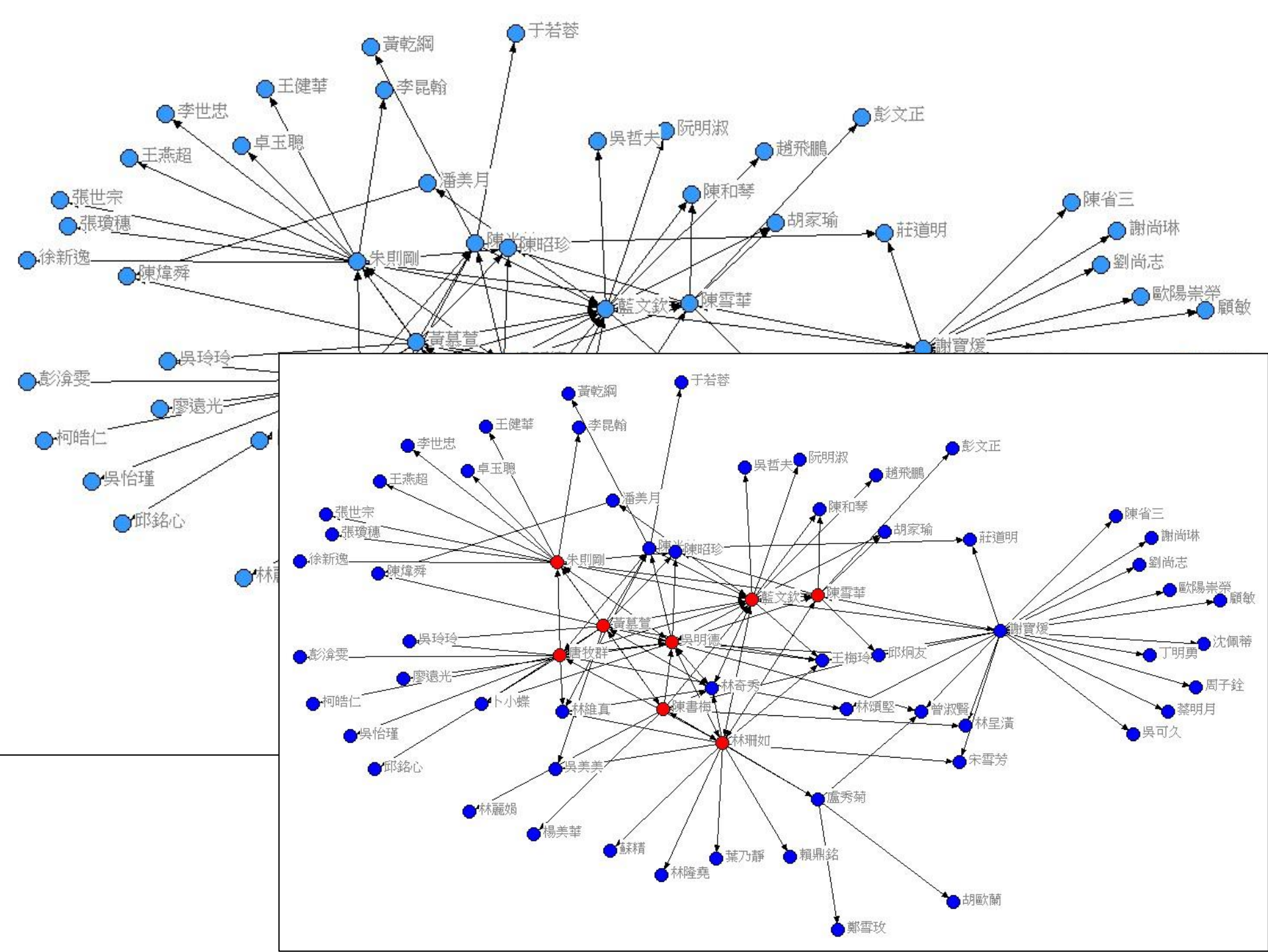
Coreness
Classes

UCINET Spreadsheet - C:\Users\assistant\Documents\UCINET data\CoreClasses.##h

File Edit Transform Fill Labels Options Help

Fill +.0 -.0 Ren

	黃慕萱	朱則剛	吳明德	林奇秀	藍文欽	林維真	陳雪華	謝寶媛	唐牧群	林珊如	陳光華
1	1	1	1	0	1	0	1	0	1	1	0
2	0	0	0	1	0	1	0	1	0	0	1



Categorical

[illegible]

Coreness

- A continuous version of “coreness” can be had by generalizing the ideal image seen above. Instead of just 0/1, pairs of “high core” nodes have a very strong tie connecting them, and core-periphery nodes have a very low score.
- **Nodes with high coreness are more likely to be at the center of a core-periphery structure.**

Coreness

- Coreness can thus be defined as a type of centrality, **but one that assumes a particular underlying structure to the network.**
- Coreness is essentially Eigenvector centrality, and **UCINET sorts nodes by eigenvector centrality and build the “core” until the correlation between ideal/observed drops.**

Again, correlation between network and attribute variables

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
		M	M	M	M	M	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F
1	M		2	10	4	5	5	9	7	4	3	3	7	3	2	5	1	4	1	0	1
2	M	2		5	1	3	1	4	2	6	2	5	4	3	2	2	6	3	1	1	1
3	M	10	5		8	9	5	11	7	8	8	14	17	9	11	11	5	9	4	6	5
4	M	4	1	8		4	0	3	4	2	3	5	3	11	4	7	0	4	3	3	0
5	M	5	3	9	4		3	5	7	4	3	5	6	3	4	4	1	2	1	3	3
6	F	5	1	5	0	3		5	2	3	2	2	4	4	3	1	1	2	0	1	2
7	F	9	4	11	3	5	5		5	4	6	3	9	5	5	4	2	6	3	2	2
8	F	7	2	7	4	7	2	5		3	0	3	4	2	1	3	0	1	1	1	0
9	F	4	6	8	2	4	3	4	3		1	3	2	4	5	4	3	4	1	3	2
10	F	3	2	8	3	3	2	6	0	1		4	7	5	5	7	2	2	3	3	2
11	F	3	5	14	5	5	2	3	3	3	4		9	3	4	4	2	4	2	3	1
12	F	7	4	17	3	6	4	9	4	2	7	9		7	7	8	3	7	2	4	3
13	F	3	3	9	11	3	4	5	2	4	5	3	7		8	11	3	8	2	5	3
14	F	2	2	11	4	4	3	5	1	5	5	4	7	8		8	1	5	4	4	1
15	F	5	2	11	7	4	1	4	3	4	7	4	8	11	8		2	5	2	2	1
16	F	1	6	5	0	1	1	2	0	3	2	2	3	3	1	2		6	1	0	1
17	F	4	3	9	4	2	2	6	1	4	2	4	7	8	5	5	6		4	3	3
18	F	1	1	4	3	1	0	3	1	1	3	2	2	2	4	2	1	4		2	1
19	F	0	1	6	3	3	1	2	1	3	3	3	4	5	4	2	0	3	2		6
20	F	1	1	5	0	3	2	2	0	2	2	1	3	3	1	1	1	3	1	6	

Interactions among a troop of monkeys

Network > core/periphery > continuous

Coreness

Input dataset: 2013_NTU

Positive or Negative Data: POSITIVE

Algorithm: Corr

Prevent Negatives: Distance
Minres

Max # of iterations:

Diagonal values valid: NO

Output dataset: 2013_NTU-Coreness

Output partition: 2013_NTU-CorenessPart

Output concentration: Concentration

OK

Cancel

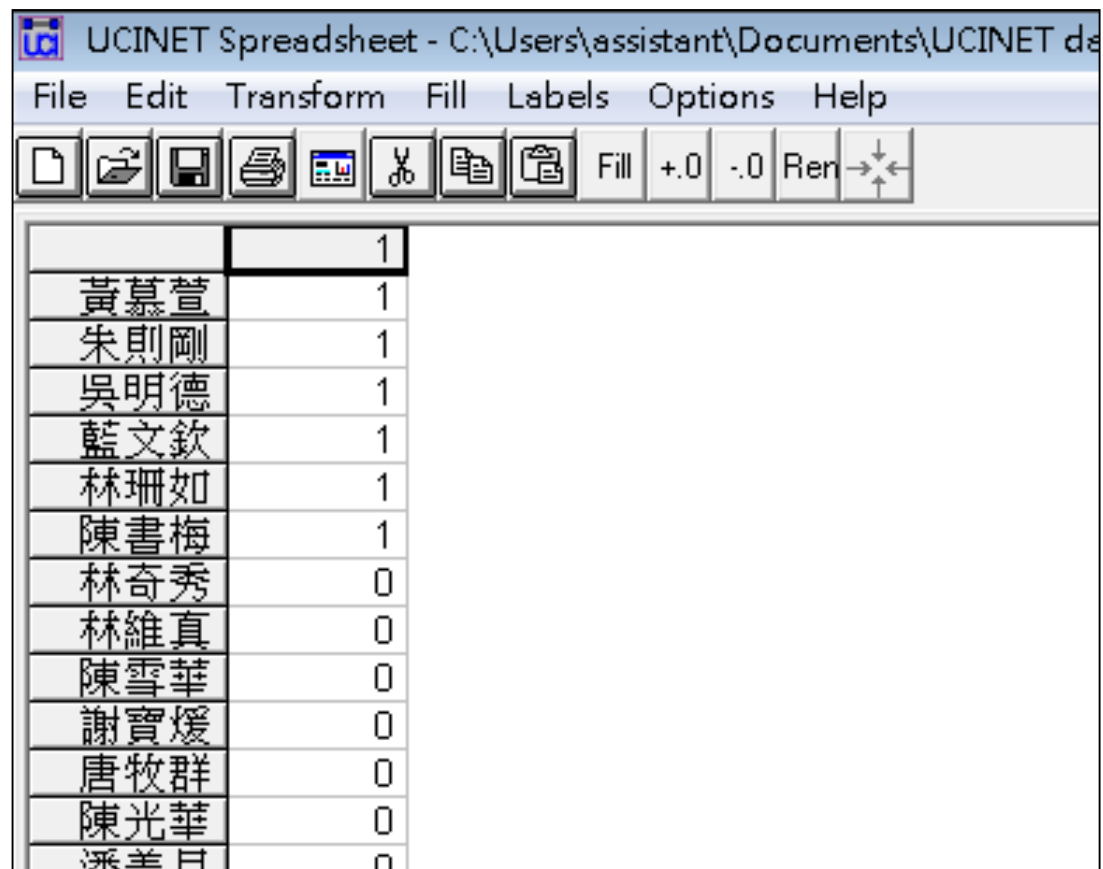
Help

Coreness exercise

Output :

檔名-Coreness 、 檔名-CorenessPart

檔名-CorenessPart



	1
黃慕萱	1
朱則剛	1
吳明德	1
藍文欽	1
林珊如	1
陳書梅	1
林奇秀	0
林維真	0
陳雪華	0
謝寶煖	0
唐牧群	0
陳光華	0
潘美月	0

檔名-Coreness

UCINET Spreadsheet - C:\Users\assistant\Documents\UCINET data\20

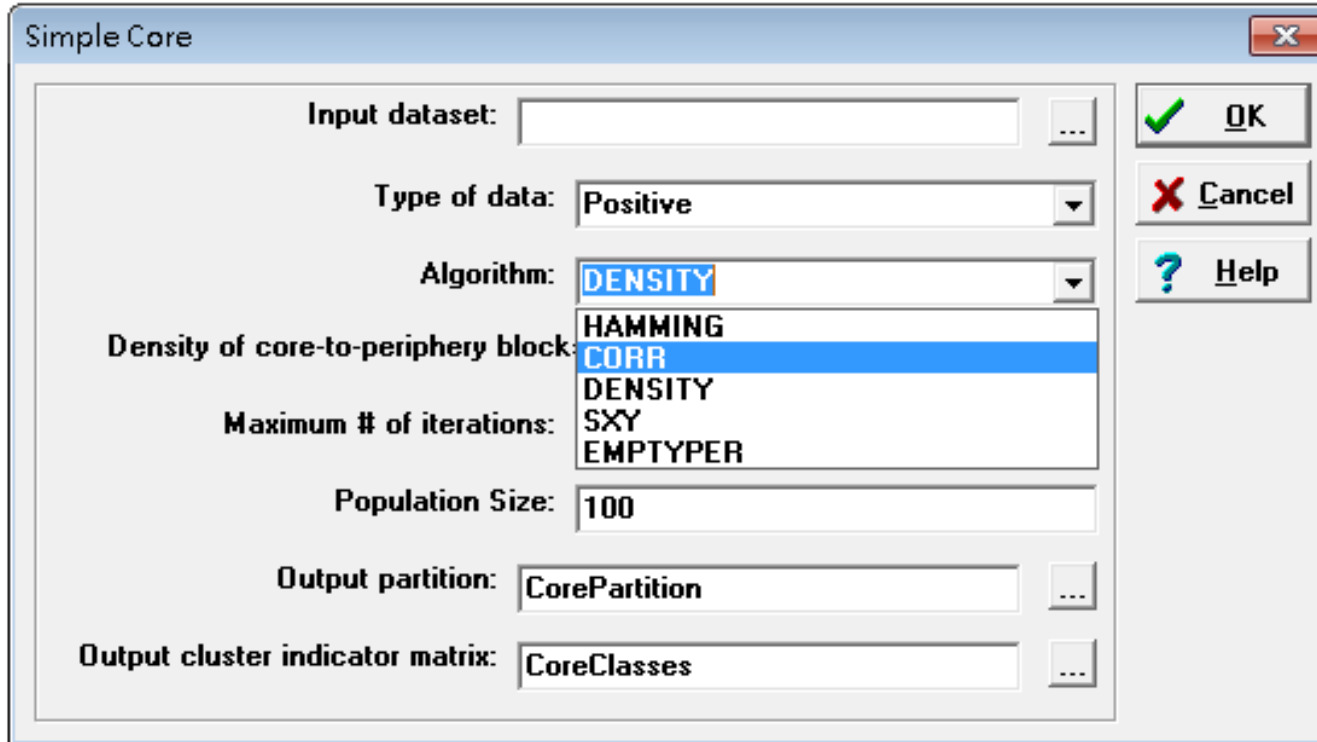
File Edit Transform Fill Labels Options Help

File Edit Transform Fill Labels Options Help

	Coreness
吳明德	0.607710 95752716 1
黃慕萱	0.489119 52972412 1
陳書梅	0.321688 05599212 6
朱則剛	0.268220 48425674 4
林珊如	0.255740 07630348 2
藍文欽	0.230224 37095642 1
唐牧群	0.163089 97571468 4
陳雪華	0.142371 13296985 6
林奇秀	0.115417 80829429 6
陳光華	0.094239 86077308 65
陳昭玲	0.077988

Use this file as attribute

Network > core/periphery > categorical



A screenshot of a software dialog box titled "Simple Core". The dialog box contains several input fields and a list box. The "Input dataset:" field is empty with a browse button (...). The "Type of data:" dropdown is set to "Positive". The "Algorithm:" dropdown is set to "DENSITY", and a list box below it shows the following options: HAMMING, CORR (highlighted), DENSITY, SXY, and EMPTYPER. The "Density of core-to-periphery block:" label is positioned to the left of the list box. The "Maximum # of iterations:" field is empty. The "Population Size:" field is set to "100". The "Output partition:" field is set to "CorePartition" with a browse button (...). The "Output cluster indicator matrix:" field is set to "CoreClasses" with a browse button (...). On the right side of the dialog box, there are three buttons: "OK" (with a green checkmark icon), "Cancel" (with a red X icon), and "Help" (with a blue question mark icon).

Simple Core

Input dataset: ...

Type of data: Positive

Algorithm: DENSITY

Density of core-to-periphery block: HAMMING
CORR
DENSITY
SXY
EMPTYPER

Maximum # of iterations:

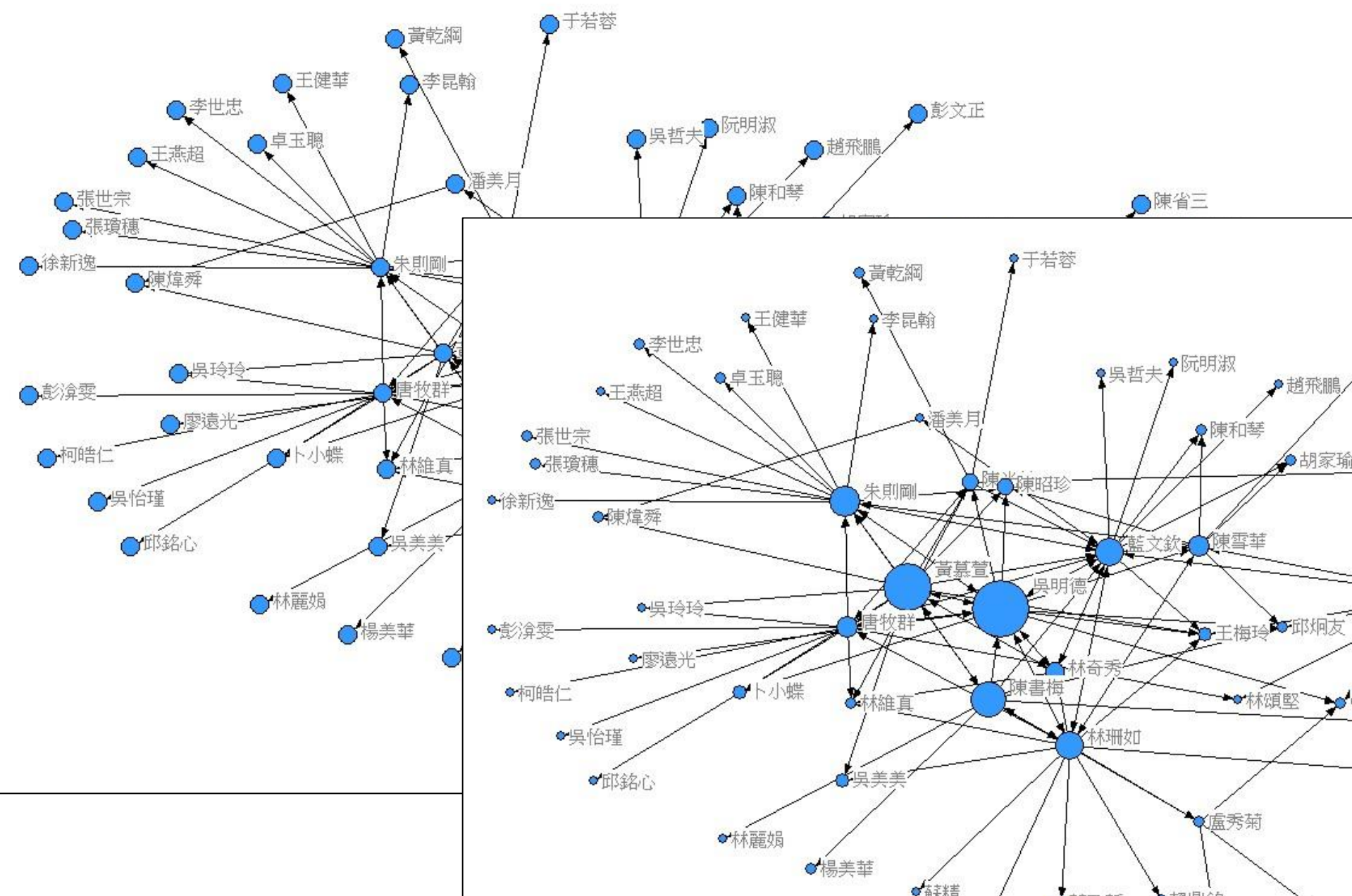
Population Size: 100

Output partition: CorePartition ...

Output cluster indicator matrix: CoreClasses ...

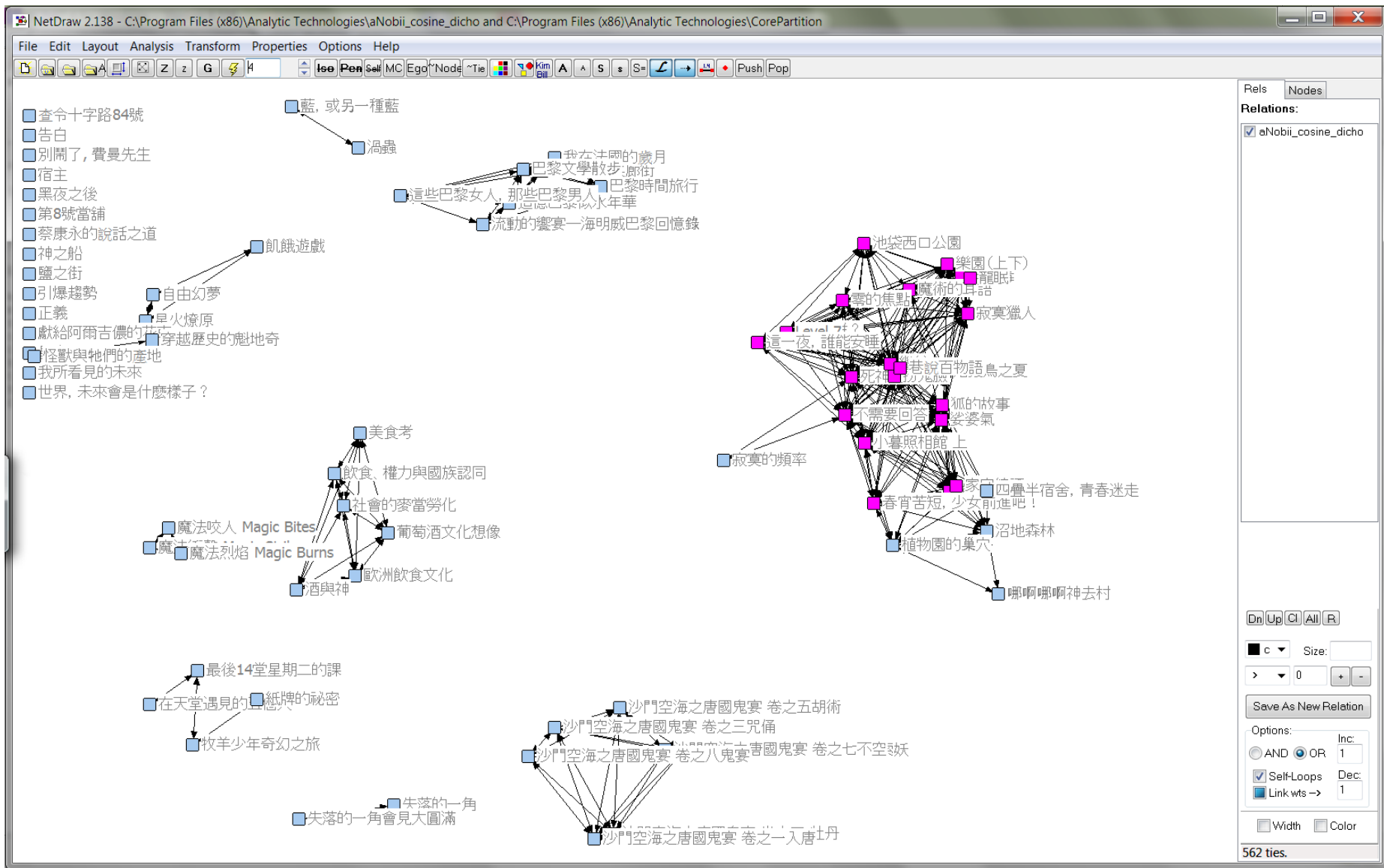
OK Cancel Help

Output : CorePartition 、 CoreClasses

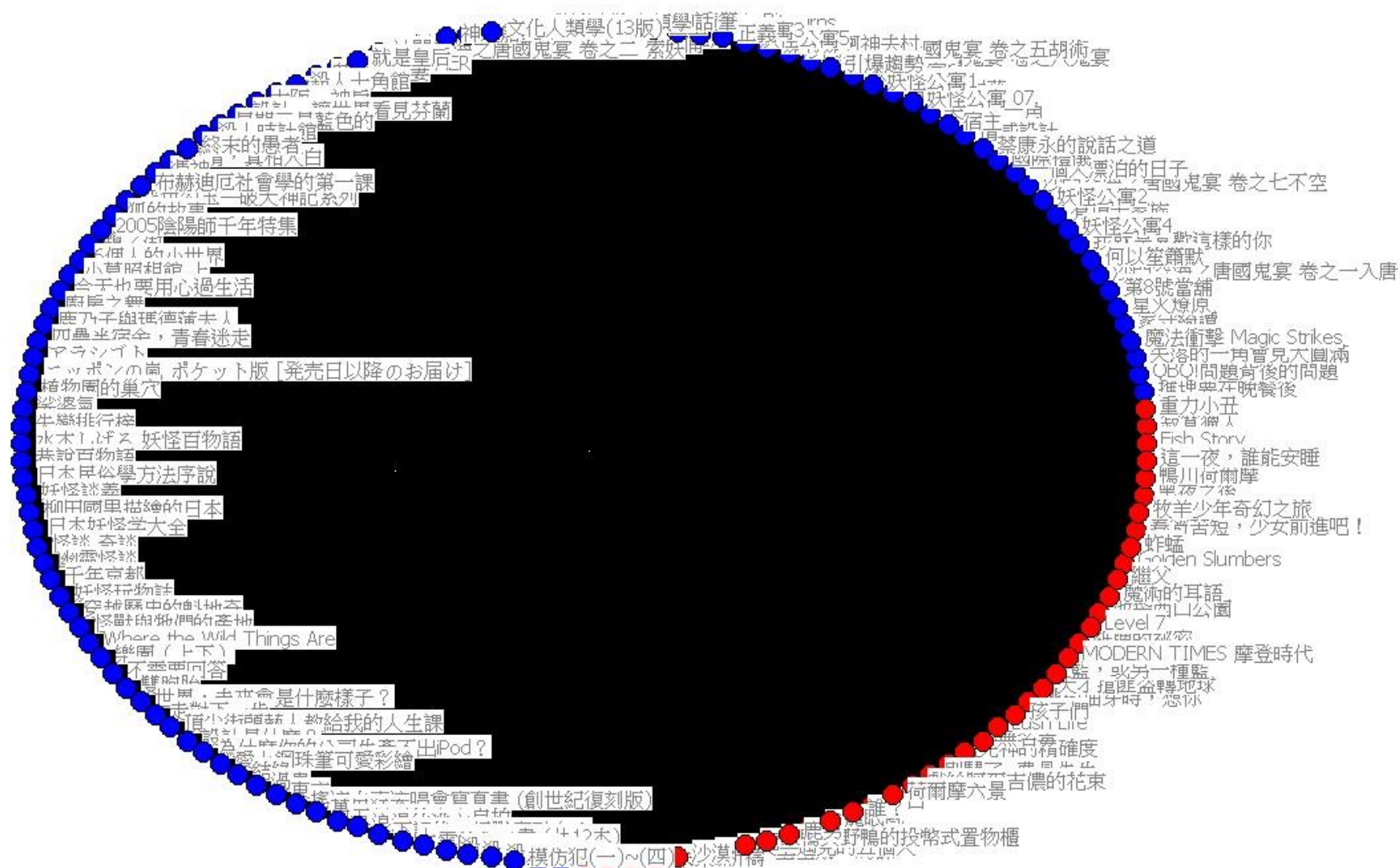


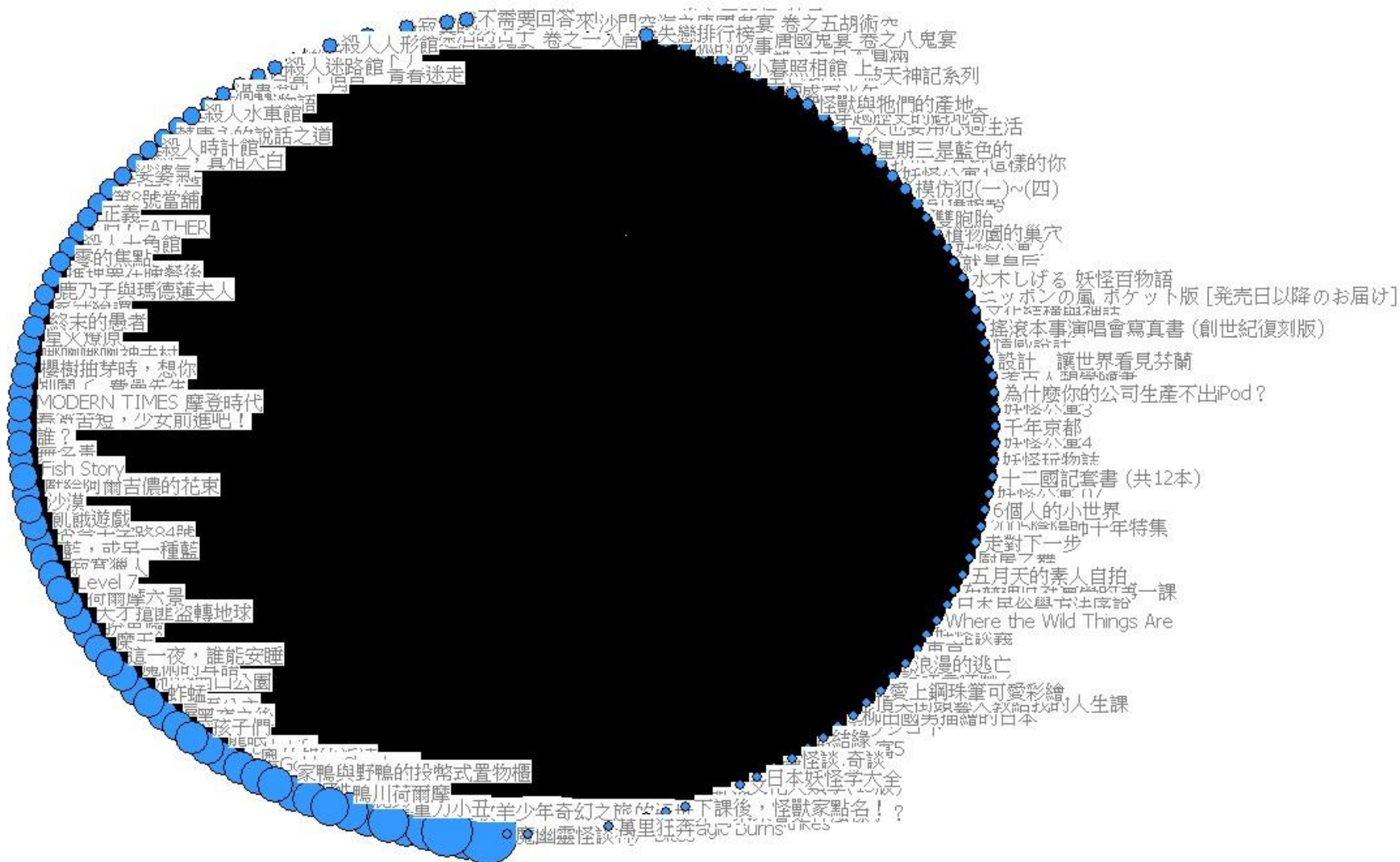
Exercise

- 1. Use aNobii cosine ([aNobii cosin ###d](#) ; [aNobii cosin ###h](#)) to
 - A. Open the file in UCINET and NetDraw to examine the cells
 - B. Dichotomize by the threshold of 0.7, then run categorical core/periphery procedure, note where CorePartition file is saved
 - B. Import CorePartition into NetDraw
 - C. Visualize the raw data with NetDraw then set edge > 0.7
 - E. Run continuous core/periphery procedures and save the coreness output
 - F. Input partition and coreness as attributes and redraw the book-book network, again set edge > 0.7









我所看見的未來
世界，未來會是什麼樣子？

怪獸與牠們的產地
穿越歷史的魁地奇

