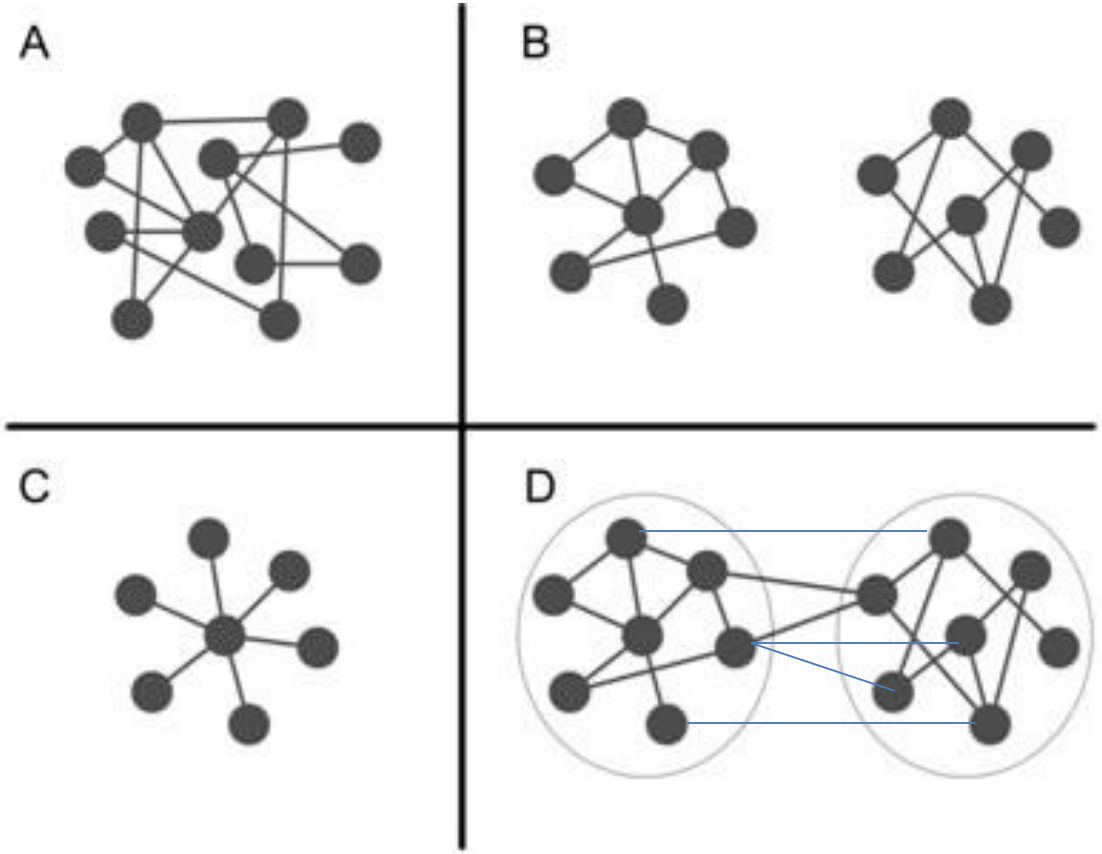
# SNA

Cohesion, geodesic distance

# Cohesion

- **Don't confuse the measure with the construct**
- Many ways to define or theoretically conceive of cohesion
- What is the mechanism that would relate cohesion to the outcome of interest?
- Define cohesion consistent with this mechanism
- For each way, we can then devise an operational measurement

# Cohesion



# Cohesions

- Network>Cohesion>Density
- Network>Cohesion>Reachability
- Network>Cohesion>No. of Geodesics
  - Geodesics?
- Network>Cohesion>Point Connectivity

# More functions…

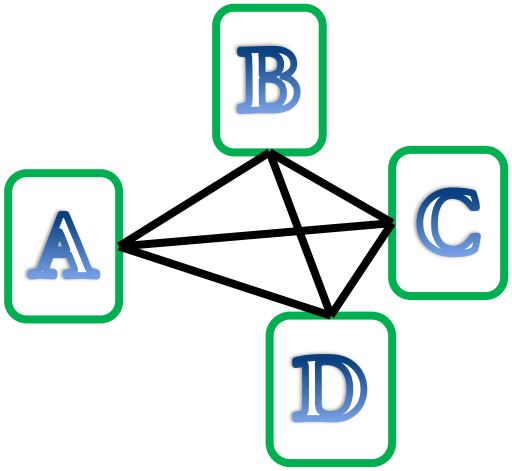
- Network>Cohesion>Density>Density of subgroups
- Network>Network properties>Reciprocity
  - For directional network
- Network>Cohesion>E-I index

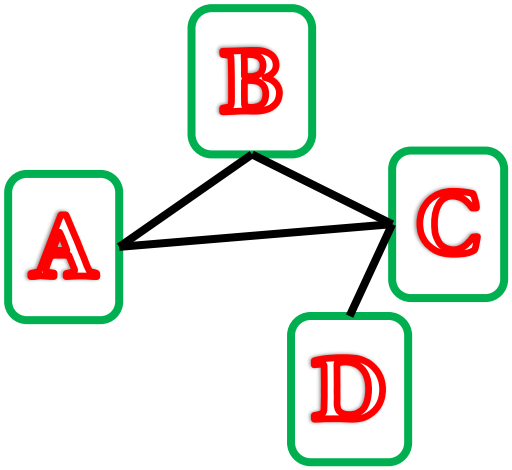
# Density

- An index of the degree of dyadic connection in a group
  - Binary data
    - The ration of the number of adjacencies that are present divided by the number of all possible pairs

# Density

- The density of a binary network is the total number of ties divided by the total number of possible ties



# Density

- Valued data
  - **The sum of the values of all ties divided by the number of possible ties**
  - The average strength of ties of across all possible (not all actual) ties

# Density by group

- Identify subgroups in a network
- Calculate the density, sum or average value **within** and **between** groups in a network or matrix.

# Density by groups

Open KNOKBUR

Create an attribute file (##d ##h)

- Group 1: governmental agencies: coun, educ, mayr  2 out of 3 links present
- Group 2: non-governmental agencies : comm, indu, news
- Group 3: welfare specialists: wro, uway, welf, west

# Density in UCINET and Gephi

Try Network>Cohesion>Overall Density

Try Network>Cohesion>Density by group

KNOKBUR

You gonna need attribute data

##D#, ##H

Also show the attributes using NetDraw

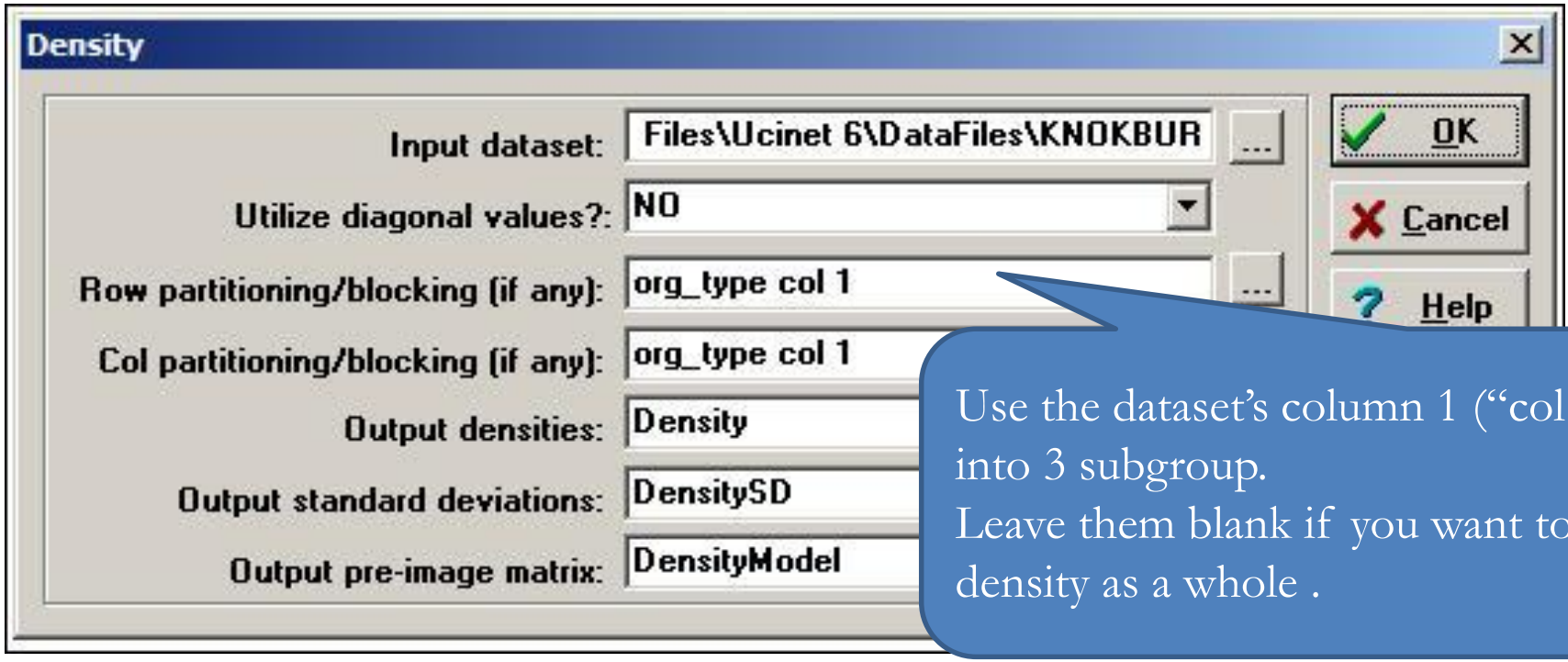
Do the same exercise with friendship data

# Density

- Exercise/unpack KRACK-HIGH-TECH
  - Open KRACKAD and KRACKFR from NetDraw
  - Use UCINET to compare the density between friendship and advice seeking density
  - Network>cohesion>density (overall)
    - * by group ? (Please use the old procedure, I have run into all kinds of problems with the "New")

# Within group density

- *Network>Cohesion>Density*
  - To calculate the density of whole populations, or of partitions
  - We can use an attribute or partition to divide the cases



Column
Block Old Code

| Block | Old Code | Members: |
|---|---|---|
| 1 | 1 | COUN EDUC MAYR |
| 2 | 2 | COMM INDU NEWS |
| 3 | 3 | WRO UWAY WELF WEST |

Relation: KNOKI

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

This shows the way you partitioned the cases. "Old Code" is from column 1.

The three sub-groups and the connections WITHIN them.

Density / average value within blocks

| | | 1 | 2 | 3 |
|---|---|---|---|---|
| | | 1 | 2 | 3 |
| 1 | 1 | 0.6667 | 0.8889 | 0.5000 |
| 2 | 2 | 0.6667 | 1.0000 | 0.1667 |
| 3 | 3 | 0.5833 | 0.6667 | 0.1667 |

Standard Deviations within blocks

| | | 1 | 2 | 3 |
|---|---|---|---|---|
| | | 1 | 2 | 3 |
| 1 | 1 | 0.4714 | 0.3143 | 0.5000 |
| 2 | 2 | 0.4714 | 0.0000 | 0.3727 |
| 3 | 3 | 0.4930 | 0.4714 | 0.3727 |

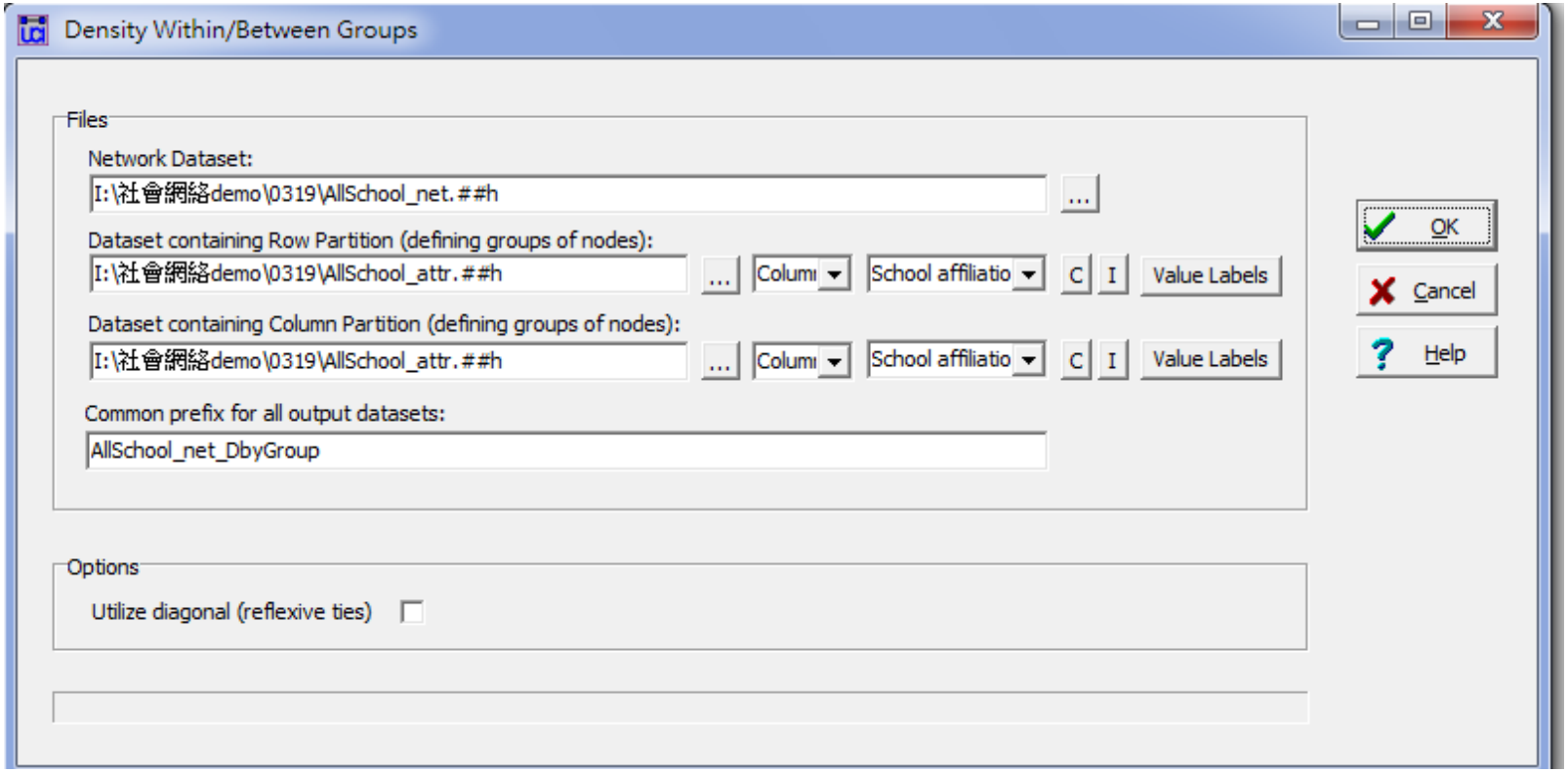
The density within- and between-blocks.

The SD within- and between-blocks.

# Exercise, apply density analysis on NTNU data

Simplified NTNU

Save into excel file then open with Gephi



# Density by groups

Number of ties/Sum of tie-strengths

| | 1<br>0 | 2<br>1 | 3<br>2 | 4<br>3 | 5<br>4 | 6<br>5 | 7<br>6 | 8<br>7 | 9<br>8 |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 0.000 | 1.000 | 1.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| 2 | 1 | 1.000 | 138.000 | 35.000 | 19.000 | 25.000 | 14.000 | 19.000 | 0.000 | 16.000 |
| 3 | 2 | 0.000 | 36.000 | 79.000 | 21.000 | 16.000 | 6.000 | 30.000 | 4.000 | 47.000 |
| 4 | 3 | 0.000 | 10.000 | 25.000 | 118.000 | 17.000 | 9.000 | 9.000 | 6.000 | 85.000 |
| 5 | 4 | 0.000 | 4.000 | 1.000 | 12.000 | 59.000 | 12.000 | 15.000 | 2.000 | 14.000 |
| 6 | 5 | 1.000 | 20.000 | 14.000 | 12.000 | 36.000 | 30.000 | 24.000 | 5.000 | 12.000 |
| 7 | 6 | 0.000 | 7.000 | 11.000 | 11.000 | 11.000 | 16.000 | 61.000 | 2.000 | 15.000 |
| 8 | 7 | 0.000 | 11.000 | 6.000 | 2.000 | 5.000 | 4.000 | 1.000 | 19.000 | 10.000 |
| 9 | 8 | 0.000 | 3.000 | 9.000 | 16.000 | 1.000 | 4.000 | 1.000 | 7.000 | 11.000 |

Density (prop of ties) / Average tie strength

| | 1<br>0 | 2<br>1 | 3<br>2 | 4<br>3 | 5<br>4 | 6<br>5 | 7<br>6 | 8<br>7 | 9<br>8 |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 0.000 | 0.009 | 0.010 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| 2 | 1 | 0.009 | 0.197 | 0.052 | 0.032 | 0.042 | 0.027 | 0.050 | 0.000 | 0.008 |
| 3 | 2 | 0.000 | 0.053 | 0.132 | 0.038 | 0.029 | 0.013 | 0.086 | 0.020 | 0.026 |
| 4 | 3 | 0.000 | 0.017 | 0.045 | 0.255 | 0.035 | 0.022 | 0.029 | 0.034 | 0.054 |
| 5 | 4 | 0.000 | 0.007 | 0.002 | 0.025 | 0.128 | 0.029 | 0.049 | 0.011 | 0.009 |
| 6 | 5 | 0.013 | 0.039 | 0.029 | 0.029 | 0.086 | 0.088 | 0.090 | 0.033 | 0.009 |
| 7 | 6 | 0.000 | 0.019 | 0.031 | 0.036 | 0.036 | 0.060 | 0.335 | 0.018 | 0.015 |
| 8 | 7 | 0.000 | 0.051 | 0.030 | 0.011 | 0.028 | 0.026 | 0.009 | 0.339 | 0.018 |
| 9 | 8 | 0.000 | 0.002 | 0.005 | 0.010 | 0.001 | 0.003 | 0.001 | 0.012 | 0.002 |

# Reachability

- An actor is reachable by another if there exists any set of connections by which we can trace from the source to the target actor, **regardless of how many others fall between them.**

# Reachability of Knoke "I" and "M" relations

Matrix #1

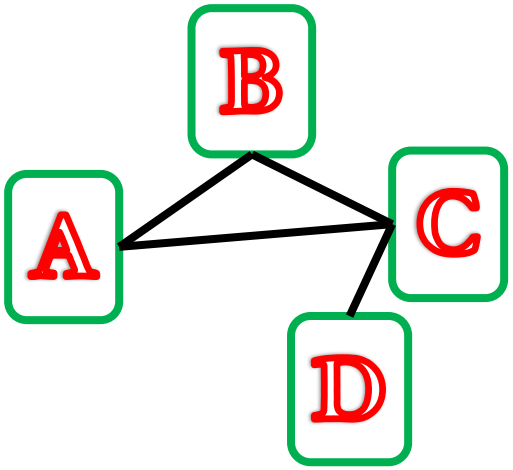
| | 1 C | 2 C | 3 E | 4 I | 5 M | 6 W | 7 N | 8 U | 9 W | 10 W |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 2 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 3 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 4 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 |
| 5 |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

Matrix #2

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

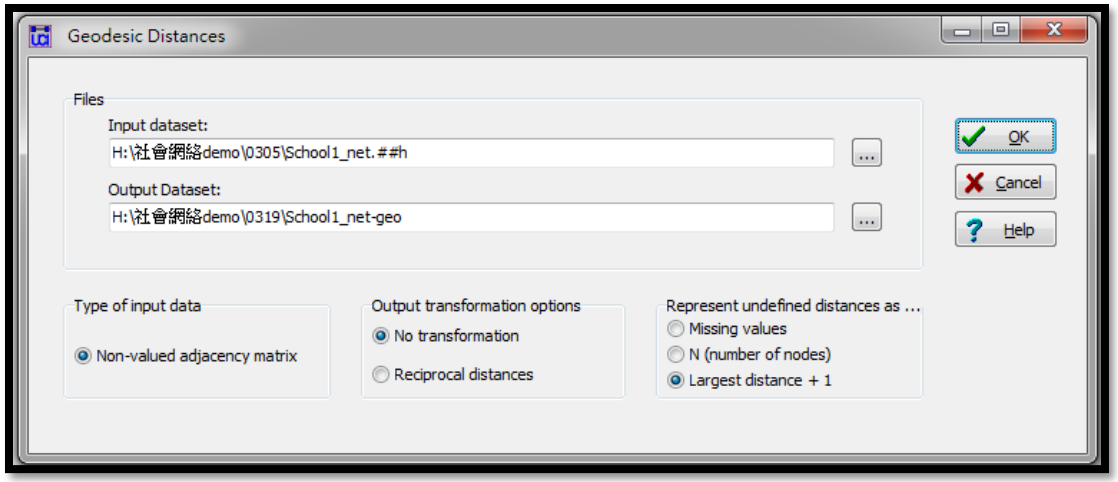
# (Geodesic) Distance

- The distance between two nodes is the length of the shortest path.
- Isolates?



# Distance

- Choose input, choose output.
- Use default settings.
- Suggestion: **save the text file** for average and frequencies.




| | Actor1 | Actor2 | Actor3 | Actor4 | Actor5 | Actor6 | Actor7 |
|---|---|---|---|---|---|---|---|
| Actor1 | 0 | 2 | 2 | 1 | 1 | 2 | 3 |
| Actor2 | 2 | 0 | 1 | 2 | 1 | 2 | 3 |
| Actor3 | 2 | 3 | 0 | 2 | 1 | 2 | 3 |
| Actor4 | 1 | 1 | 1 | 0 | 1 | 1 | 2 |
| Actor5 | 1 | 2 | 1 | 1 | 0 | 1 | 2 |
| Actor6 | 1 | 1 | 1 | 1 | 1 | 0 | 1 |
| Actor7 | 2 | 2 | 2 | 1 | 1 | 1 | 0 |

Geodesic distances for Knoke information exchange

Geodesic Distances

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

Compare 6 to 4 and 4 to 6, are they symmetrical

Save the output file and open it with netdraw
Reverse the value?

Who is close to the CEO ? Krack-high-tech

# Distance in a valued graph: tie strength

- How to measure distance with valued data?
- The distance between two actors is defined as the strength of the weakest link in the shortest path.
  - If A sends 6 units to B, B sends 4 units to C, the strength of the path from A to C (**assuming A to B to C is the shortest path**) is 4.
  - Wolf/fn (first save the data into separate file, then use the OLD procedure)

# No. of Geodesics

- Counts the number of different geodesic paths connecting each pair of nodes.
- Binary dataset only, or UCINET will warn you.
- What does the mean?

| | Actor1 | Actor2 | Actor3 | Actor4 | Actor5 | Actor6 | Actor7 | Actor8 |
|---|---|---|---|---|---|---|---|---|
| Actor1 | 1 | 1 | 2 | 1 | 1 | 2 | 2 | 1 |
| Actor2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 |
| Actor3 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 2 |
| Actor4 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 |
| Actor5 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 2 |
| Actor6 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| Actor7 | 3 | 2 | 3 | 1 | 1 | 1 | 1 | 1 |
| Actor8 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |

Who can influence the CEO ? Krack-high-tech

Number of geodesic paths for Knoke information exchange

# of Geodesic Paths

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

Next: Run on WOLFN data

# Flow

- Complementing geodesic paths
- Taking into account ALL connections between pairs of actors, **not just the most efficient ones**.
- Network>Cohesion>Maximum Flow
  - How many different actors in the NEIGHBORHOOD of a source leads to pathways to a target.

# Flow

- "If I need to send a message to you, and there is only one other person to whom I can send this for retransmission, my connection is weak– **even if the person I send it to may have many ways of reaching you**.
- If, on the other hand, there are four people to whom I can send my message, **each of whom has one or more ways of retransmitting my message to you**, then my connection is stronger."

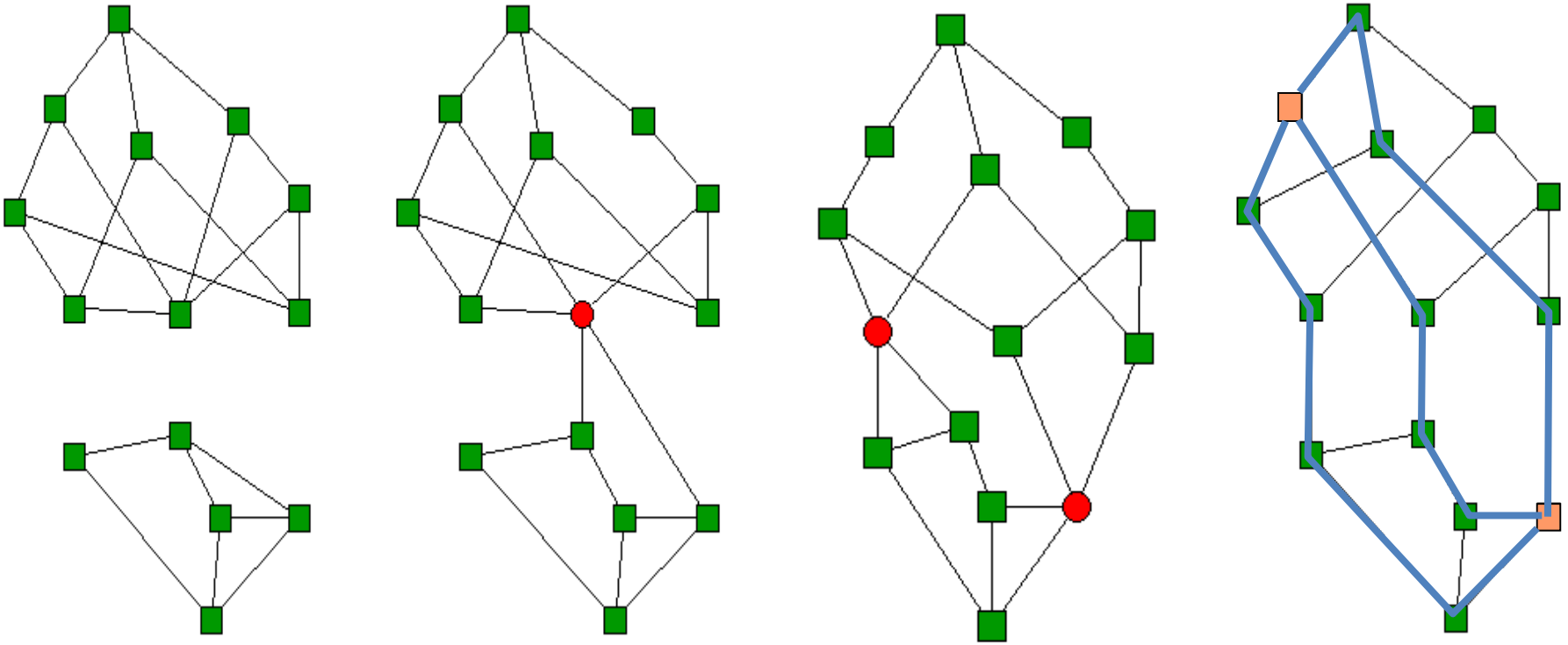
Maximum flow for Knoke information network

| | 1<br>C | 2<br>C | 3<br>E | 4<br>I | 5<br>M | 6<br>W | 7<br>N | 8<br>U | 9<br>W | 10<br>W |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 4 | 3 | 4 | 4 | 1 | 4 | 2 | 4 | 2 |
| 2 | 5 | 0 | 3 | 5 | 7 | 1 | 7 | 2 | 5 | 2 |
| 3 | 5 | 6 | 0 | 5 | 6 | 1 | 6 | 2 | 5 | 2 |
| 4 | 4 | 4 | 3 | 0 | 4 | 1 | 4 | 2 | 4 | 2 |
| 5 | 5 | 8 | 3 | 5 | 0 | 1 | 8 | 2 | 5 | 2 |
| 6 | 3 | 3 | 3 | 3 | 3 | 0 | 3 | 2 | 3 | 2 |
| 7 | 3 | 3 | 3 | 3 | 3 | 1 | 0 | 2 | 3 | 2 |
| 8 | 5 | 6 | 3 | 5 | 6 | 1 | 6 | 0 | 5 | 2 |
| 9 | 3 | 3 | 3 | 3 | 3 | 1 | 3 | 2 | 0 | 2 |
| 10 | 5 | 5 | 3 | 5 | 5 | 1 | 5 | 2 | 5 | 0 |

Who can influence the CEO ? Krack-high-tech

# Global Connectivity

•Networks are structurally cohesive if they remain connected even when nodes are removed



0 1 2 3

# Node/point connectivity

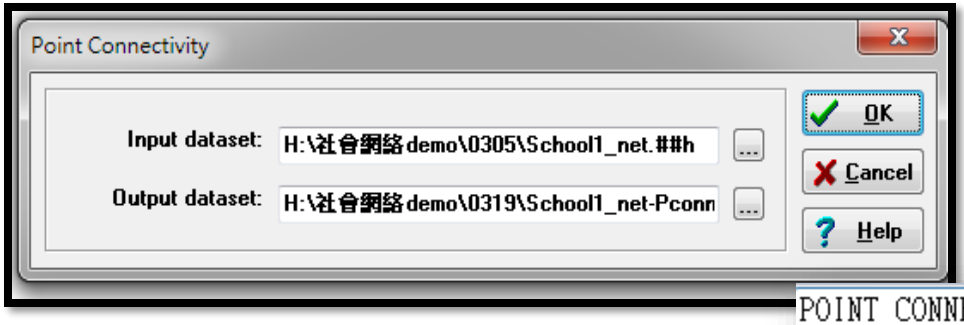
- The minimum number of nodes that need to be removed to disrupt the path **between any two pair of nodes.**
- Equal to the maximum number of **node** independent path between any two pair of nodes.

# Point/local connectivity

- **NETWORK>COHESION>POINT CONNECTIVITY**
  **PURPOSE** Compute the local point connectivity between all pairs of nodes in a network.
  **DESCRIPTION** The local (point) connectivity of two non-adjacent vertices is the number of vertices that need to be deleted so that no path connects them, this is equal to the maximum number of vertex disjoint paths connecting them.
- NOTE: This procedure only operates on the first matrix in a dataset.
- Hint: data > unpack

# Point Connectivity

- Compute the local point connectivity between all pairs of nodes in a network




POINT CONNECTIVITY

Input dataset: School1_net (H:\
Output connectivity: School1_net-Pconne

| | | 1 Actor | 2 Actor | 3 Actor | 4 Actor | 5 Actor | 6 Actor | 7 Actor |
|---|---|---|---|---|---|---|---|---|
| 1 | Actor1 | 0.000 | 2.000 | 2.000 | 2.000 | 2.000 | 2.000 | 1.000 |
| 2 | Actor2 | 1.000 | 0.000 | 2.000 | 1.000 | 2.000 | 1.000 | 1.000 |
| 3 | Actor3 | 1.000 | 1.000 | 0.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| 4 | Actor4 | 3.000 | 2.000 | 4.000 | 0.000 | 5.000 | 2.000 | 1.000 |
| 5 | Actor5 | 3.000 | 2.000 | 3.000 | 3.000 | 0.000 | 2.000 | 1.000 |
| 6 | Actor6 | 3.000 | 2.000 | 4.000 | 4.000 | 6.000 | 0.000 | 1.000 |
| 7 | Actor7 | 3.000 | 2.000 | 3.000 | 3.000 | 3.000 | 3.000 | 0.000 |

Local connectivity: the minimum # of nodes need to be removed to disconnect a pair of nodes

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

# Flow

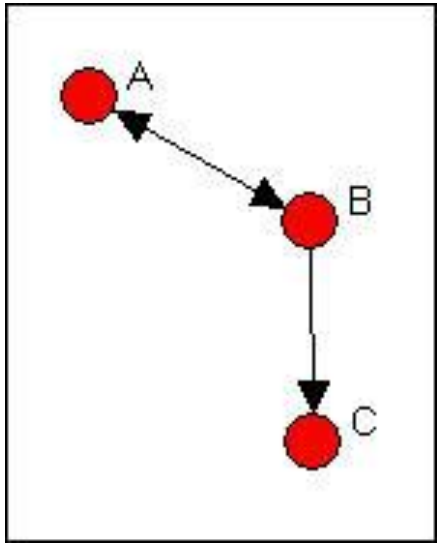
- Complementing geodesic paths
- Taking into account ALL connections between pairs of actors, **not just the most efficient ones**.
- Network>Cohesion>Maximum Flow
  - How many different actors in the NEIGHBORHOOD of a source leads to pathways to a target.

# Flow

- "If I need to send a message to you, and there is only one other person to whom I can send this for retransmission, my connection is weak– **even if the person I send it to may have many ways of reaching you**.
- If, on the other hand, there are four people to whom I can send my message, **each of whom has one or more ways of retransmitting my message to you**, then my connection is stronger."

# Reciprocity in directional network

- Four possible dyadic relationships with directed data
- A network that has a predominance of **null or reciprocated ties** over asymmetric connections may be a more "equal" or "stable" network than **one with a predominance of asymmetric connections**



# Reciprocity

- *Network>Cohesion>Reciprocity*
- Method
  - Arc-based:
    - "Hybrid"

      (A, B), (B, C) (C, A) > 1/3
  - Dyad-based

    **All actual pairs**

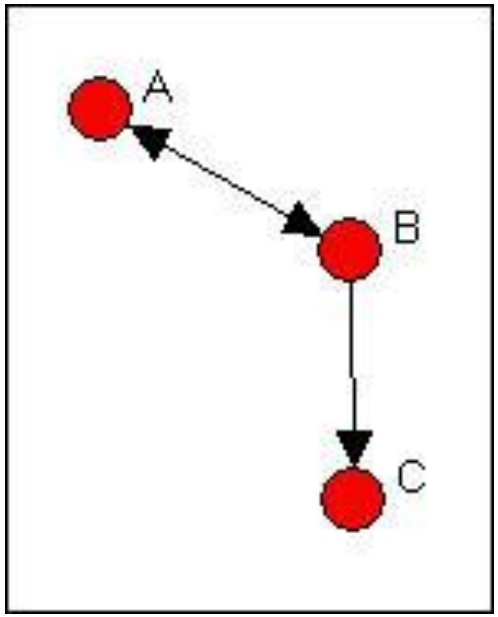
    (A, B), (B, C) > 1/2

    AB, BA, BC, CB, CA, AC > 1/3

    **All actual ties**

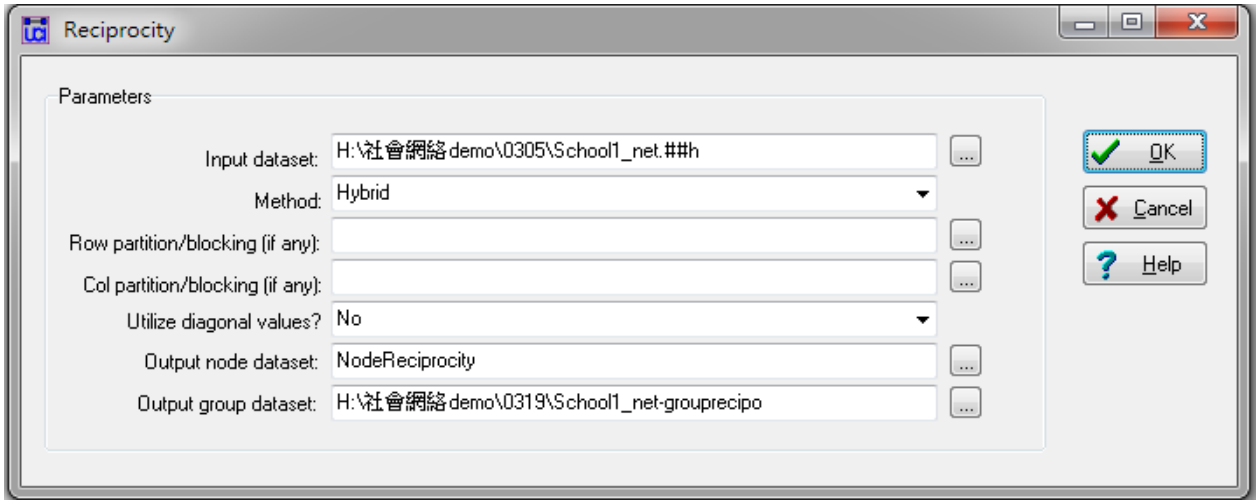
    AB, BA, BC, CB, CA, AC > 2/3

Consider all possible pairs
Is less sensible in a large network, why?

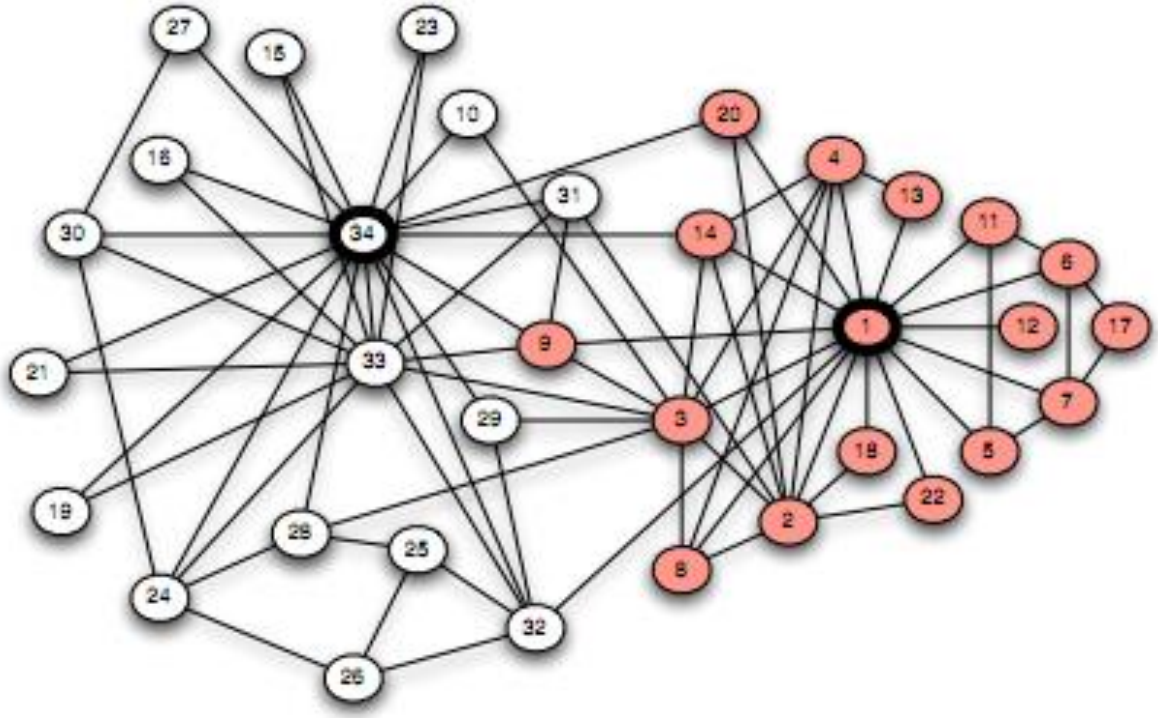


# Reciprocity

- Calculate the amount of ties that are reciprocated in a network.
- Use default settings.
- Suggestion:
  save the text file for detail info.



# E-I: Group External vs. Internal ties



# E-I: Group External vs. Internal ties

- Takes the number of external group ties, subtracts the number of internal ties, and divides by the total number of ties
- The lower the E-I, the higher the degree of **closure**
- Range from -1 (all ties are internal; homophile or closure) to +1 (all ties are external)
- Directions of ties are ignored

# E-I index

- Units of analysis
  - Entire population
  - Each sub-group
  - Individual

# External and internal ties

- Network>Cohesion>E-I index
  - Given **a partition** of a network into a number of **mutually exclusive groups**, then
  - the E-I index is the number of ties external to the groups minus the number of ties that are internal to the group divided by the total number of ties.

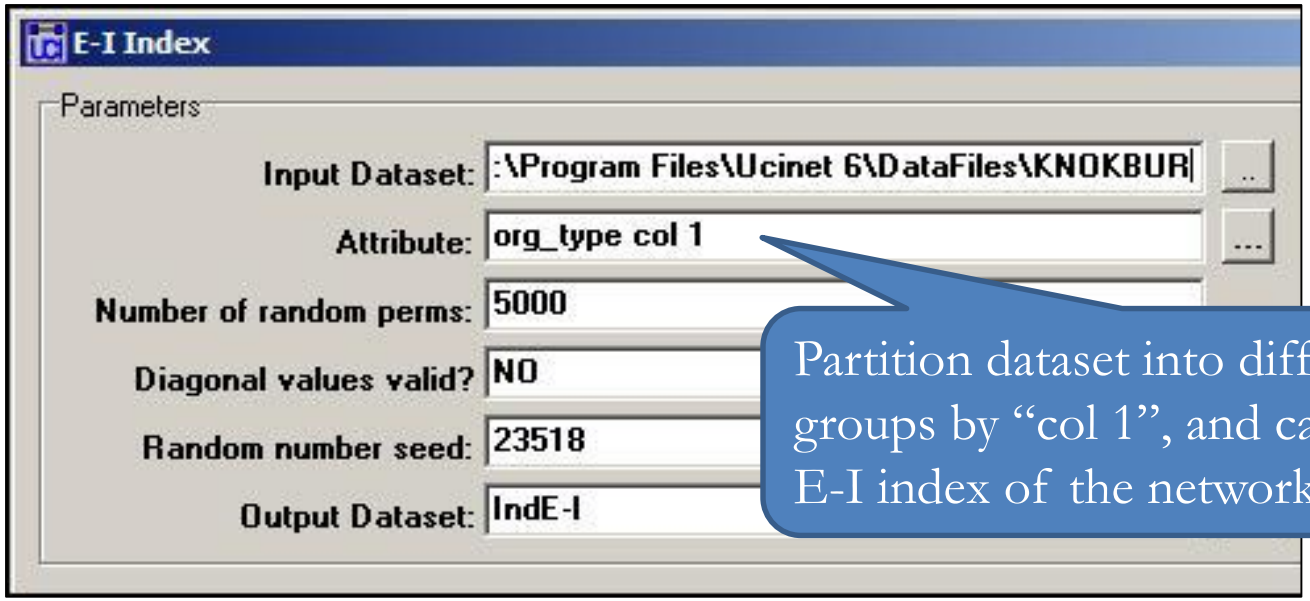
You can also try

Network>Cohesion>Homophily

more about this later

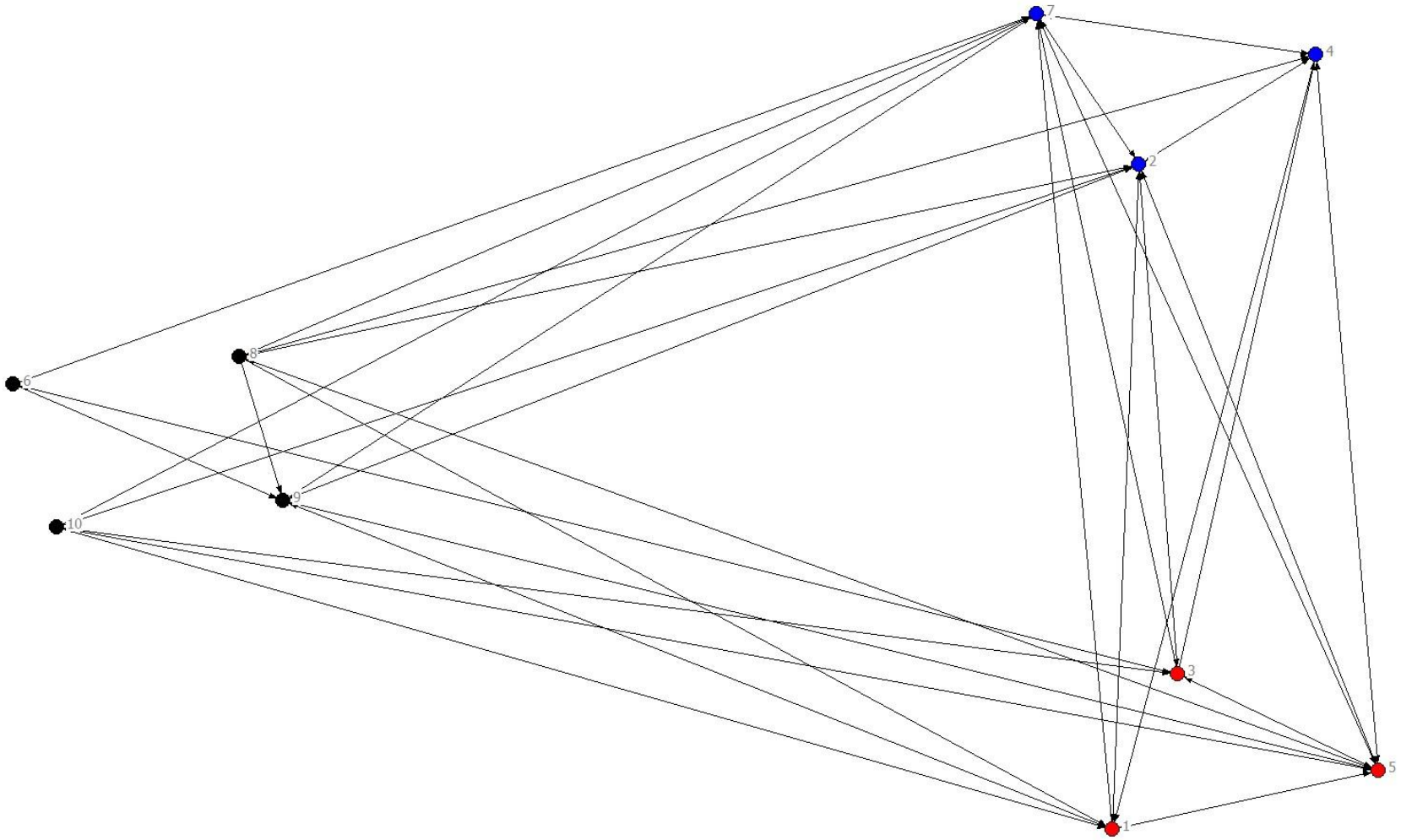
# E-I Index

- *Network>Cohesion>E-I Index*
- Compare the numbers of ties within groups and between groups.
  - The directions of ties are ignored




nologies\datafiles\High-Tec-Attributes.##h" col 1

# E-I



| | 1<br>Freq | 2<br>Pct | 3<br>Possib | 4<br>Densit |
|---|---|---|---|---|
| 1 Internal | 14.000 | 0.219 | 24.000 | 0.583 |
| 2 External | 50.000 | 0.781 | 66.000 | 0.758 |
| 3 E-I | 36.000 | 0.563 | 42.000 | 0.467 |

Max possible external ties: 66.000
Max possible internal ties: 24.000

E-I Index: 0.563
Expected value for E-I index is: 0.467

the expected value under a random distribution (given the size of the group and the overall density)

Max possible E-I given density & group sizes: 1.000
Min possible E-I given density & group sizes: 0.250

Re-scaled E-I index: -0.167

Re-scaled E-I Index by the demographic constraints and overall density.

Permutation Test
Number of iterations = 5000

P=.203, not significant

| | 1<br>Obs | 2<br>Min | 3<br>Avg | 4<br>Max | 5<br>SD | 6<br>P >= Ob | 7<br>P <= Ob |
|---|---|---|---|---|---|---|---|
| 1 Internal | 0.219 | 0.625 | 0.733 | 0.844 | 0.039 | 1.000 | 0.000 |
| 2 External | 0.781 | 0.156 | 0.267 | 0.375 | 0.039 | 0.000 | 1.000 |
| 3 E-I | 0.563 | 0.250 | 0.467 | 0.688 | 0.078 | 0.203 | 0.953 |

Group level E-I Index

| | | 1 Intern | 2 Extern | 3 Total | 4 E-I |
|---|---|---|---|---|---|
| 1 | 1 | 4.000 | 17.000 | 21.000 | 0.619 |
| 2 | 2 | 6.000 | 17.000 | 23.000 | 0.478 |
| 3 | 3 | 4.000 | 16.000 | 20.000 | 0.600 |

Individual Level E-I Index

| | 1 Inter | 2 Exter | 3 Total | 4 E-I |
|---|---|---|---|---|
| 1 | 1.000 | 6.000 | 7.000 | 0.714 |
| 2 | 2.000 | 6.000 | 8.000 | 0.500 |
| 3 | 1.000 | 5.000 | 6.000 | 0.667 |
| 4 | 2.000 | 4.000 | 6.000 | 0.333 |
| 5 | 2.000 | 6.000 | 8.000 | 0.500 |
| 6 | 1.000 | 2.000 | 3.000 | 0.333 |
| 7 | 2.000 | 7.000 | 9.000 | 0.556 |
| 8 | 1.000 | 5.000 | 6.000 | 0.667 |
| 9 | 2.000 | 4.000 | 6.000 | 0.333 |
| 10 | 0.000 | 5.000 | 5.000 | 1.000 |

Raw counts of ties within and between each group.

E-I Index for each group.

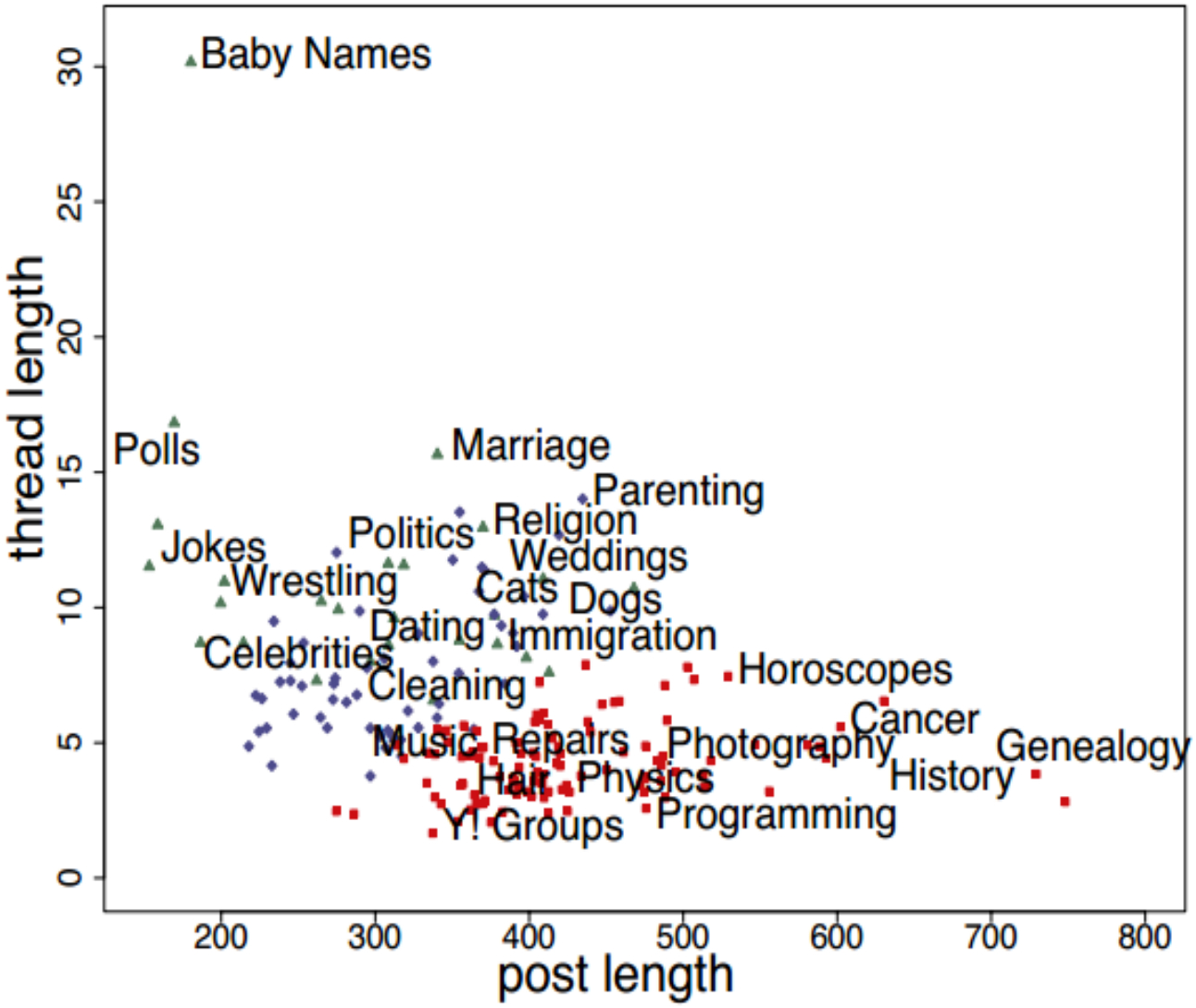
# Exercise I

- Use Friendship data and do the following analyses
- 1. Run density within Gender and Club
- 2. Run E-I Index on both attributes and see if there exists homophily by either attribute
- 3. Identify the most outward/Inward - looking group and individuals
- 4. Which gender is more outward looking?

# Exercise II

- Apply E-I index analysis on Krack High Tech, and Try both level (column 3) and department (column 4)
  - The procedure only runs on the first spreadsheet
    - Run on report-to network and check the E-I value

# Yahoo Answers (YA)



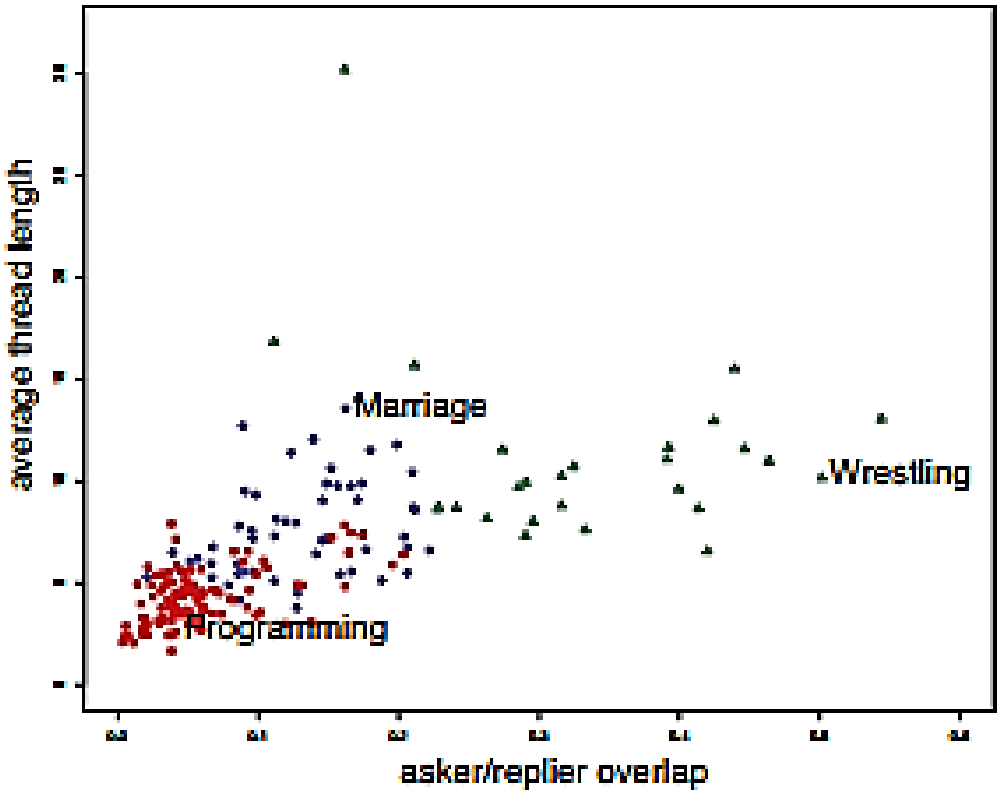



**Figure 2: Clustering of categories by thread length and overlap between askers and repliers**

Adamic, L. A., Zhang, J., Bakshy, E., & Ackerman, M. S. (2008, April). Knowledge sharing and yahoo answers: everyone knows something. In*Proceedings of the 17th international conference on World Wide Web* (pp. 665-674). ACM.

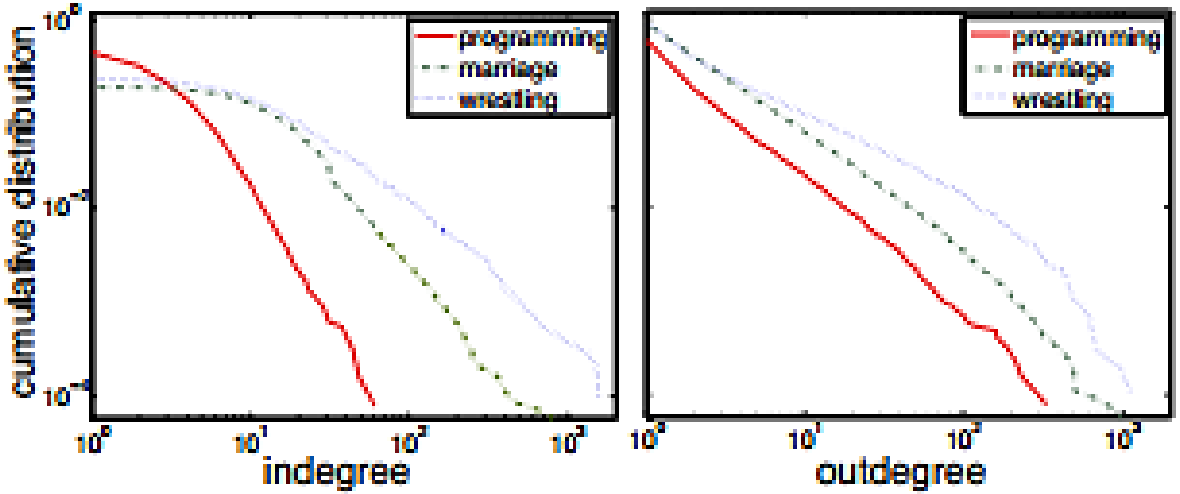



**Figure 3: Distributions of indegree (number of users one has received answers from) and outdegree (number of users one has answered)**

Adamic, L. A., Zhang, J., Bakshy, E., & Ackerman, M. S. (2008, April). Knowledge sharing and yahoo answers: everyone knows something. In*Proceedings of the 17th international conference on World Wide Web* (pp. 665-674). ACM.

**Table 1: Summary statistics for selected QA networks**

| Category | Nodes | Edges | Avg. deg. | Mutual edges | SCC |
|---|---|---|---|---|---|
| Wrestling | 9,959 | 56,859 | 7.02 | 1,898 | 13.5% |
| Program. | 12,538 | 18,311 | 1.48 | 0 | 0.01% |
| Marriage | 45,090 | 164,887 | 3.37 | 179 | 4.73% |

Adamic, L. A., Zhang, J., Bakshy, E., & Ackerman, M. S. (2008, April). Knowledge sharing and yahoo answers: everyone knows something. In *Proceedings of the 17th international conference on World Wide Web* (pp. 665-674). ACM.

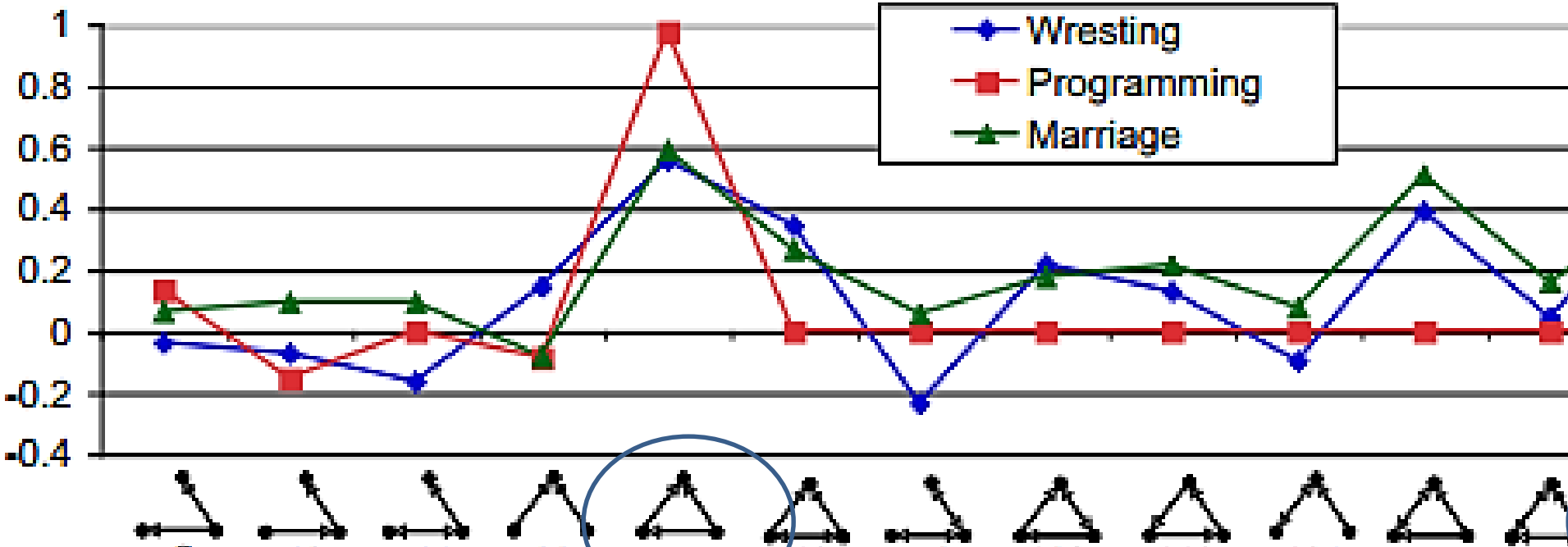



**Figure 5: Motif profiles of selected categories**

Adamic, L. A., Zhang, J., Bakshy, E., & Ackerman, M. S. (2008, April). Knowledge sharing and yahoo answers: everyone knows something. In*Proceedings of the 17th international conference on World Wide Web* (pp. 665-674). ACM.