Elliott Brannon

Reminders

- Please make sure you register you iClicker on canvas.
- If you used your iClicker during the last lecture, but you did not get attendance or participation points, it probably means you have not registered your iClicker.

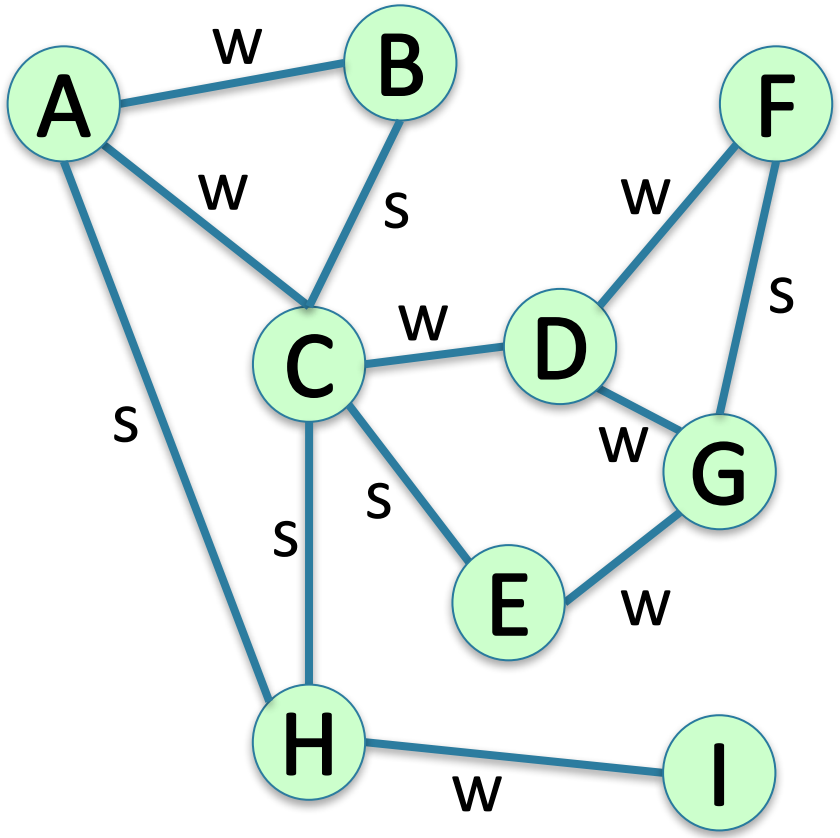
Last Time

- Bridges, local bridges, span
- Strong and weak ties
- Strong Triadic Closure Property
- Proof: Local bridges are weak ties under some assumptions.
- Neighborhood overlap.
- In large cell phone communication network:
  - Tie strength increased with neighborhood overlap -> Weak ties tend to be local bridges.
  - Removing weak ties decreased size of giant connected component faster than strong ties -> Weak ties crucial for connectivity.
- Types of connections in social media: Facebook and Twitter.

Warm up

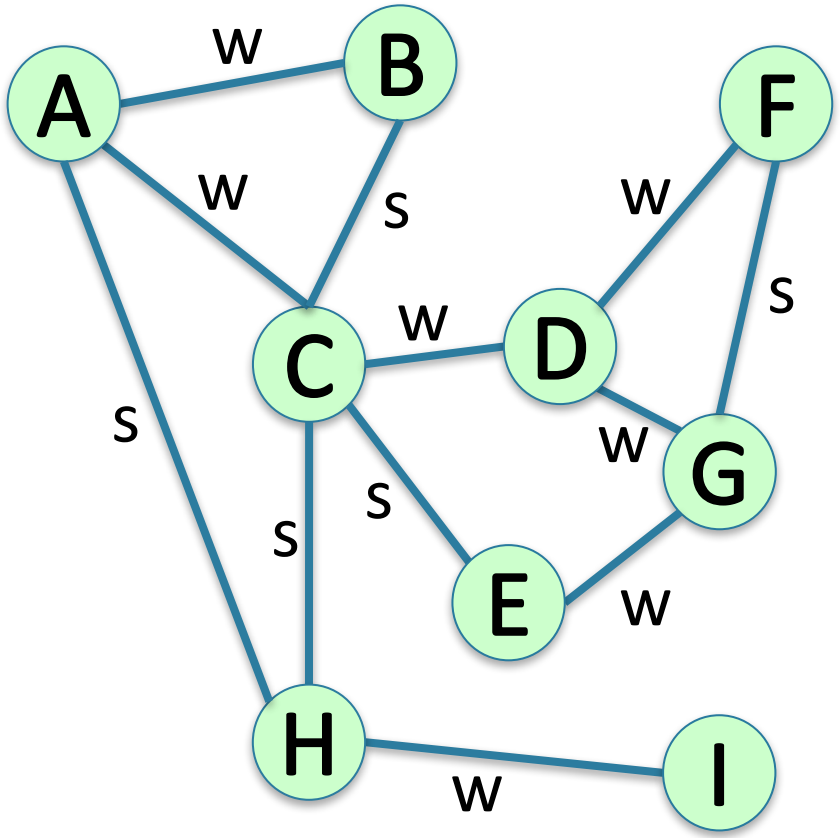
1. Find all bridges, local bridges, and their span.
2. Find the neighborhood overlap of all the edges.
3. Check which nodes satisfy the Strong Triadic Closure Property.

Note that you can ignore the weight of the edges for questions 1 and 2.



Warm up

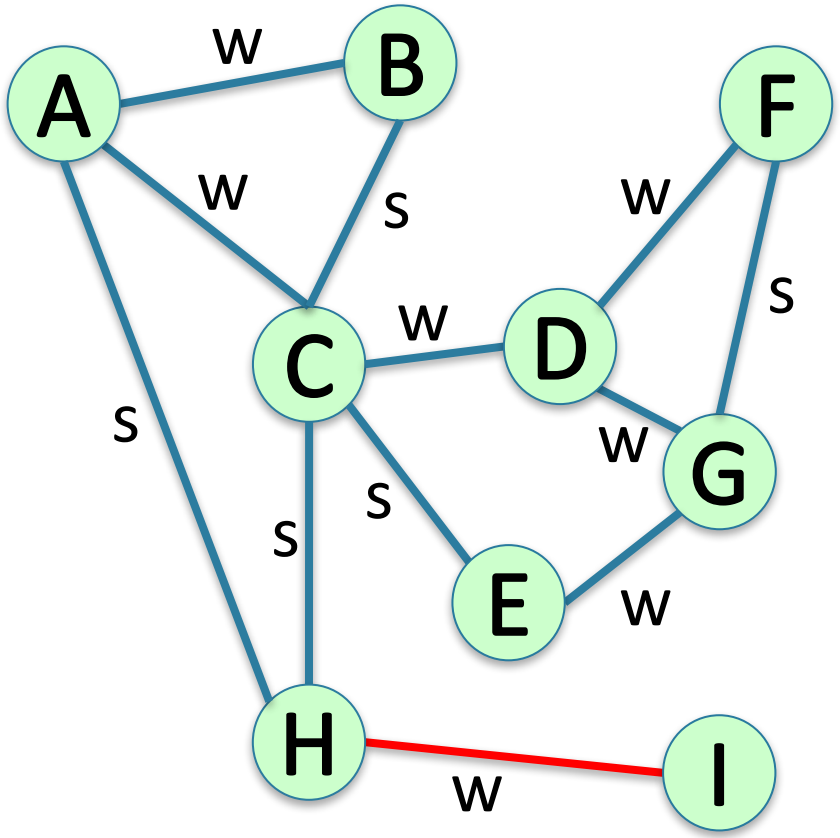
1. Which of the following edges are bridges:
   A. A—B
   B. C—D
   C. D—G
   D. H—I
   E. None of the above



Warm up

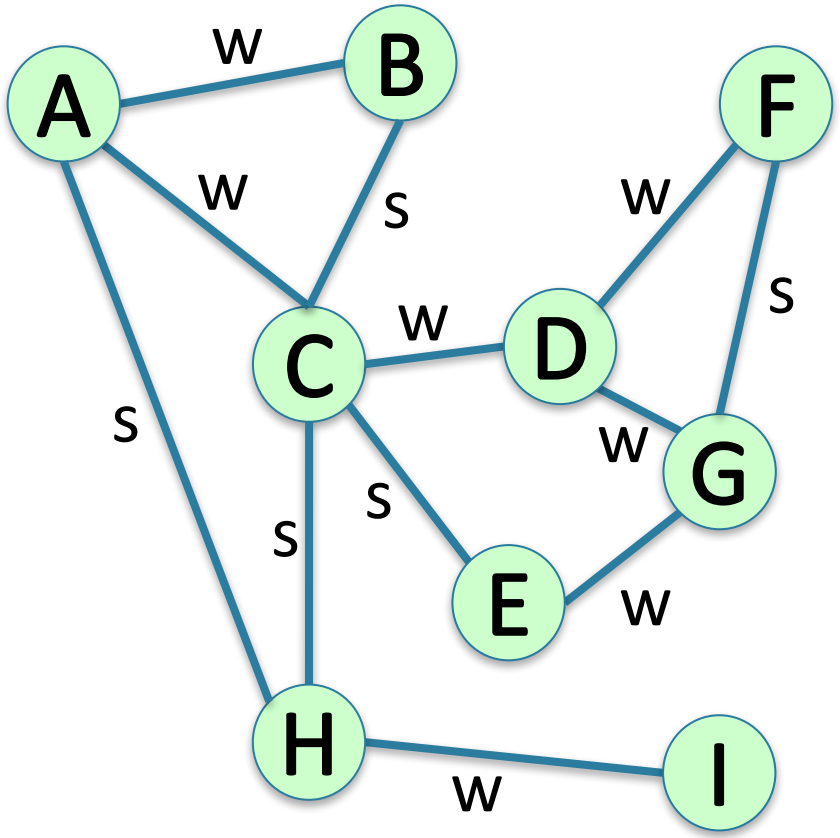
1. Which of the following edges are bridges:
   A. A—B
   B. C—D
   C. D—G
   D. H—I
   E. None of the above

If we remove edge H—I, nodes H and I belong to different components (no path from H to I).



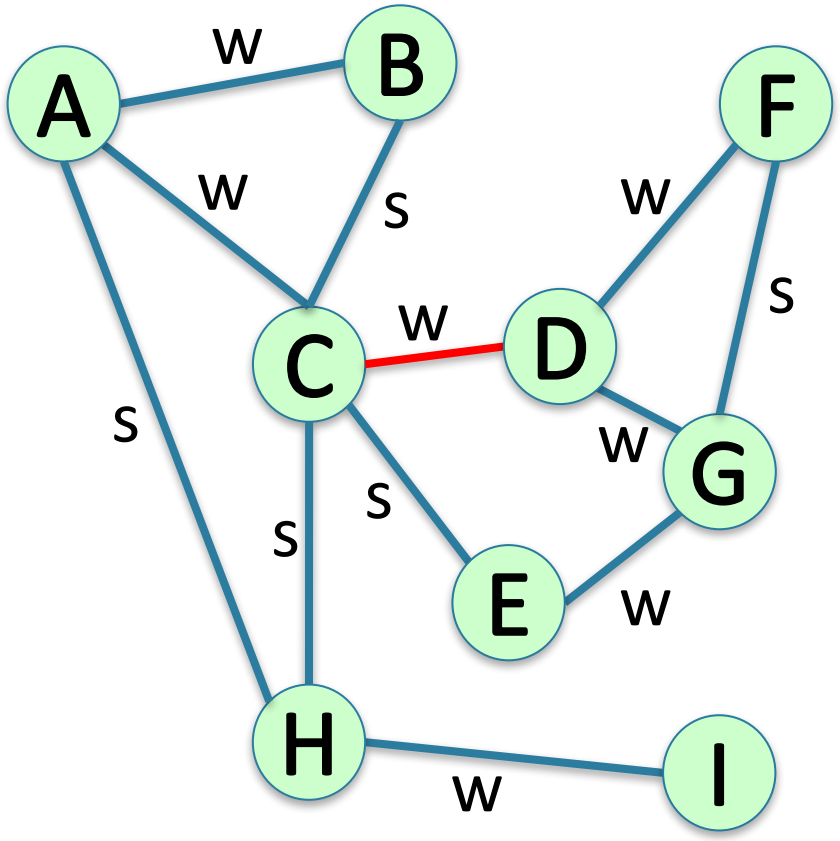
Warm up

2. Which of the following edges are local bridges:
   A. A—B
   B. C—D
   C. D—G
   D. A—H
   E. None of the above



Warm up

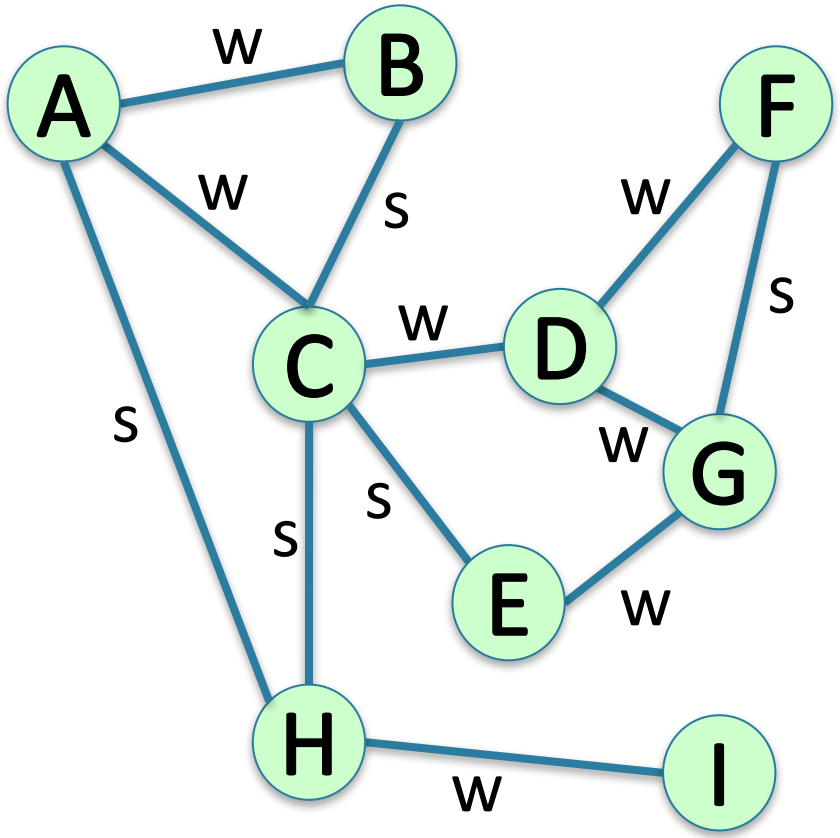
2. Which of the following edges are local bridges:
   A. A—B
   B. C—D
   C. D—G
   D. A—H
   E. None of the above



Nodes C and D have no friends in common.

Warm up

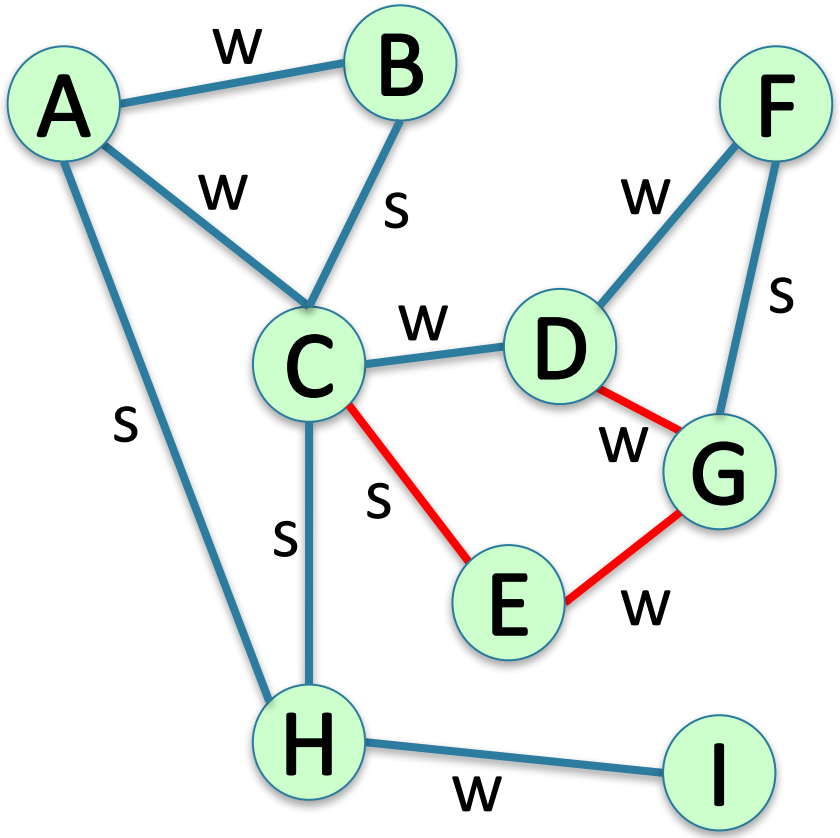
3. What is the span of local bridge C—D?
   A. 1
   B. 2
   C. 3
   D. 4
   E. None of the above



Warm up

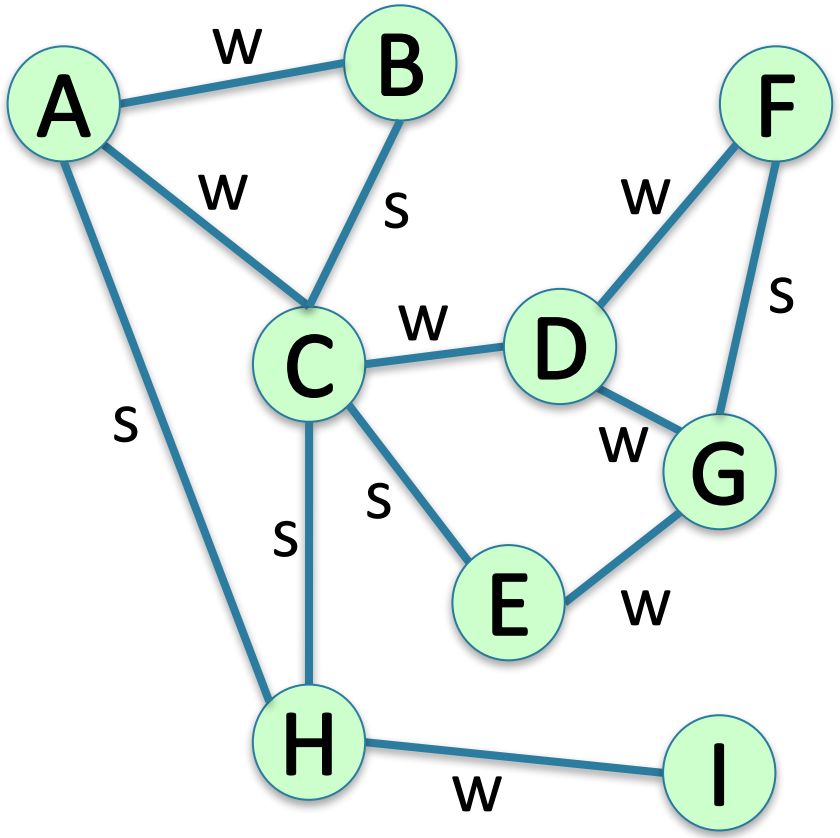
3. What is the span of local bridge C—D?
   A. 1
   B. 2
   C. 3
   D. 4
   E. None of the above

If we remove edge C—D, the new distance between nodes C and D is 3.



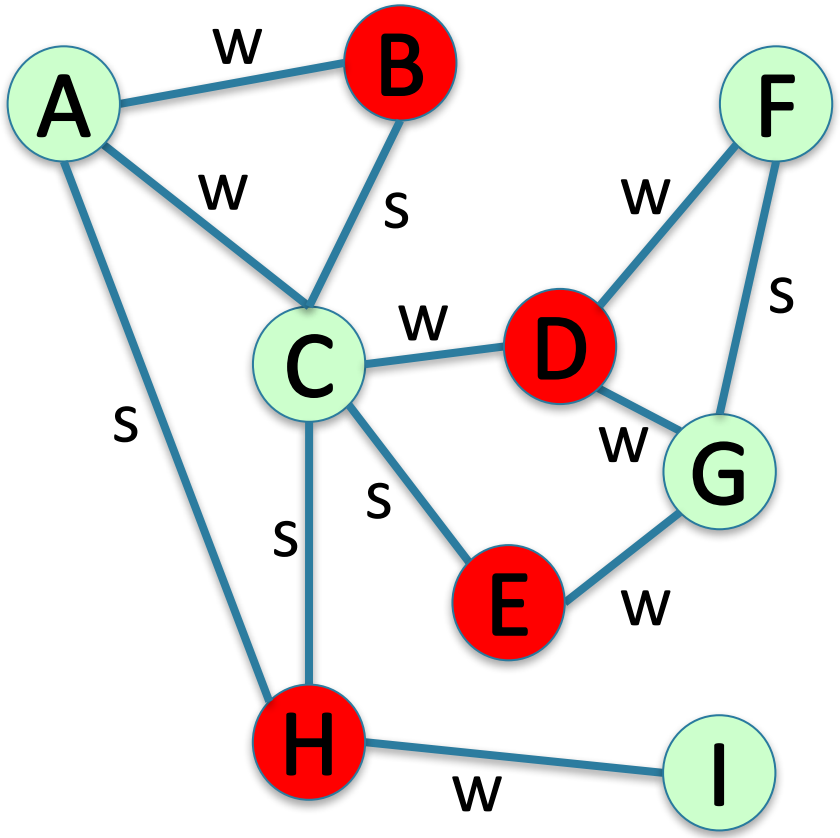
Warm up

4. What is the neighborhood overlap of edge C—A?
   A. 1/3
   B. 2/3
   C. 1/2
   D. 1
   E. None of the above



Warm up

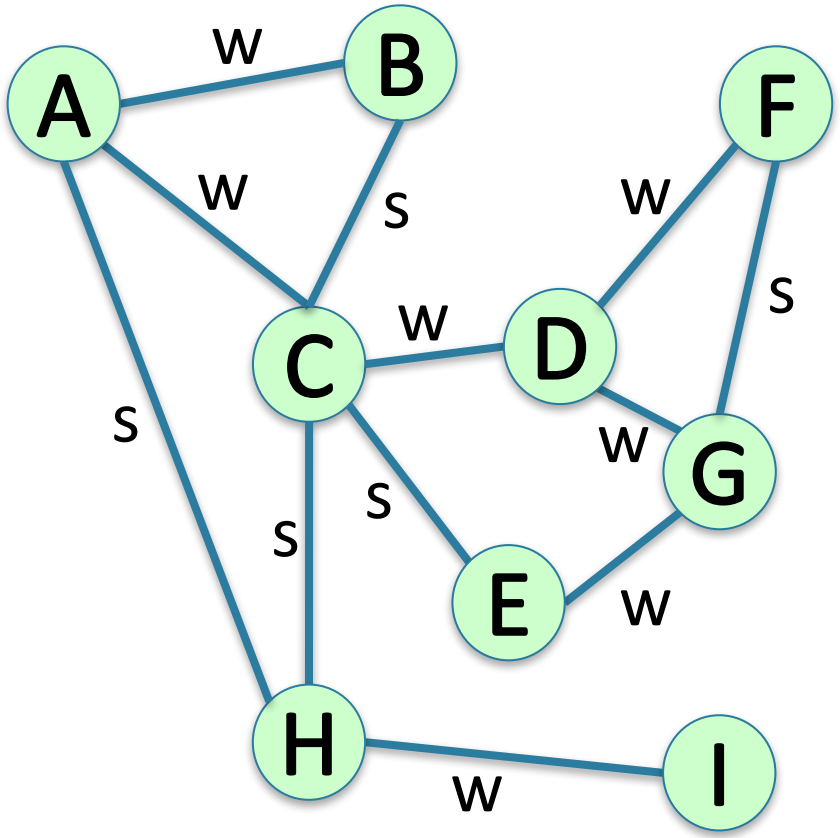
4. What is the neighborhood overlap of edge C—A?
   A. 1/3
   B. 2/3
   C. 1/2
   D. 1
   E. None of the above



2 nodes (B and H) are connected to both A and C.
A and C have 4 friends in total.
Neighborhood overlap of C—A is 2/4 = 1/2.

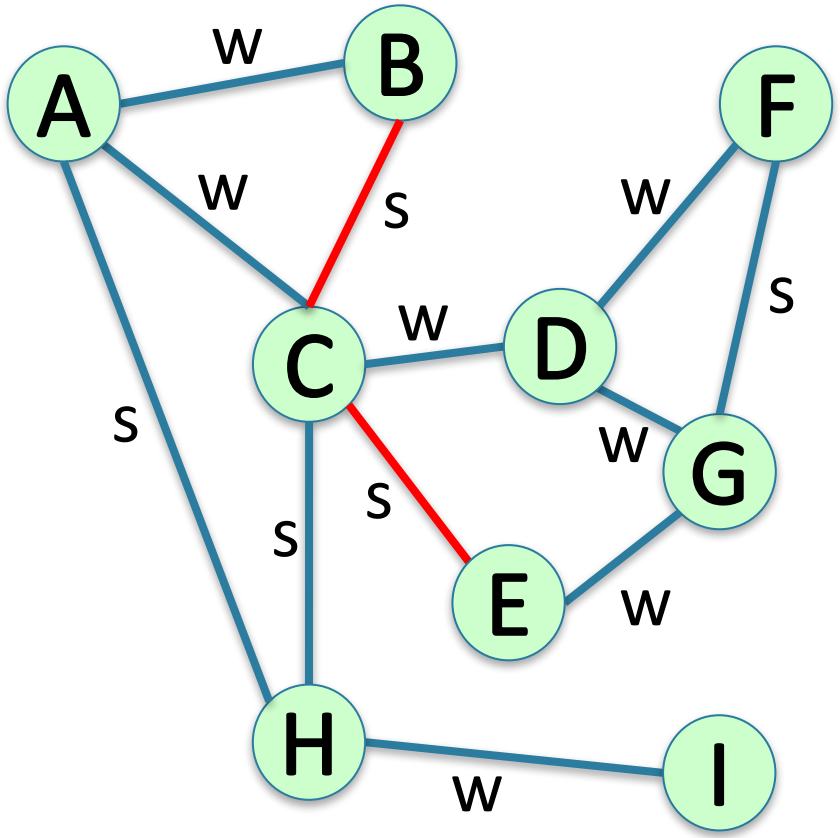
Warm up

4. Which node violates the Strong Triadic Closure Property?
   A. A
   B. B
   C. C
   D. D
   E. Node of the above



Warm up

4. Which node violates the Strong Triadic Closure Property?
   A. A
   B. B
   C. C
   D. D
   E. Node of the above

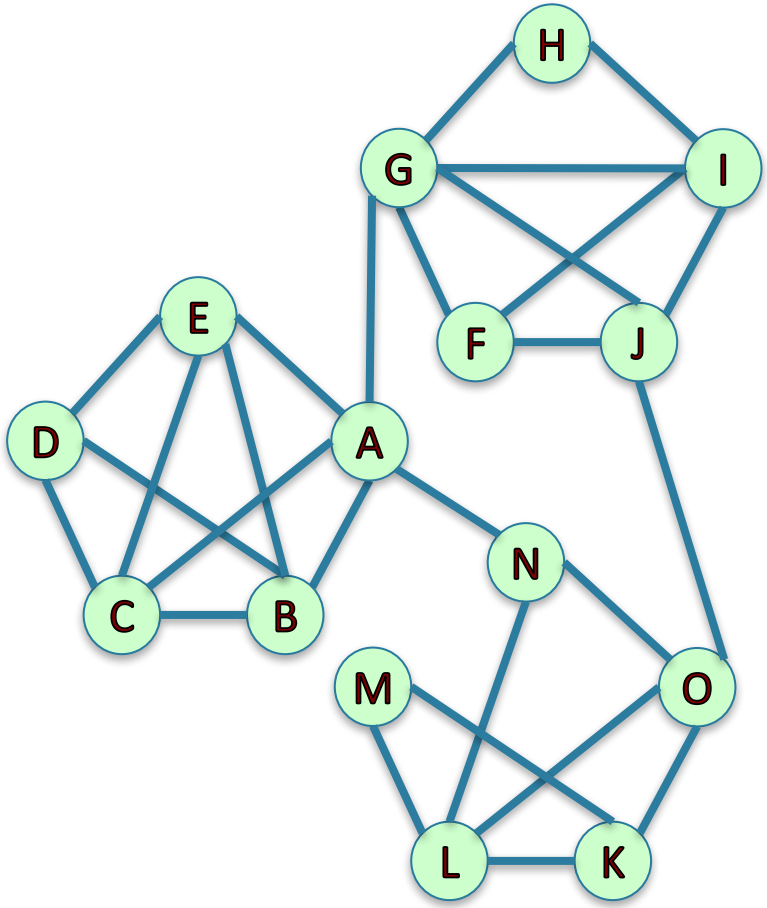


C has strong connection to B and E, but edge B—E does not exist.

Embeddedness

Embeddedness of an edge A – B is the number of friends A and B have in common.

What's the embeddedness of A—B?

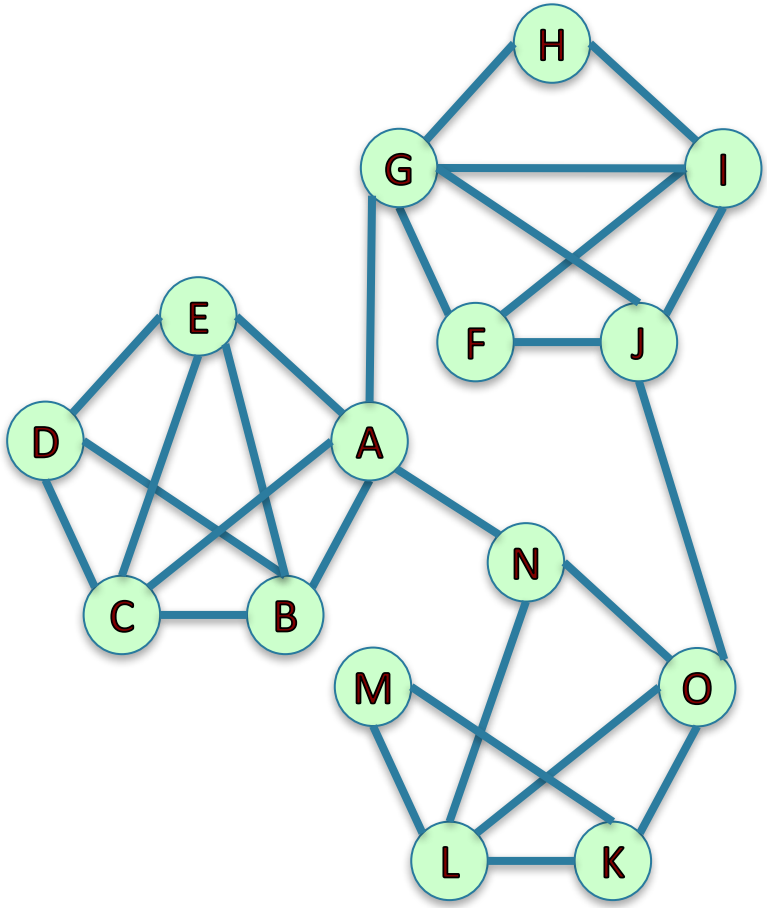


Embeddedness

Embeddedness of an edge A – B is the number of friends A and B have in common.

What's the embeddedness of A—B?

2. A and B are both friends of E and C



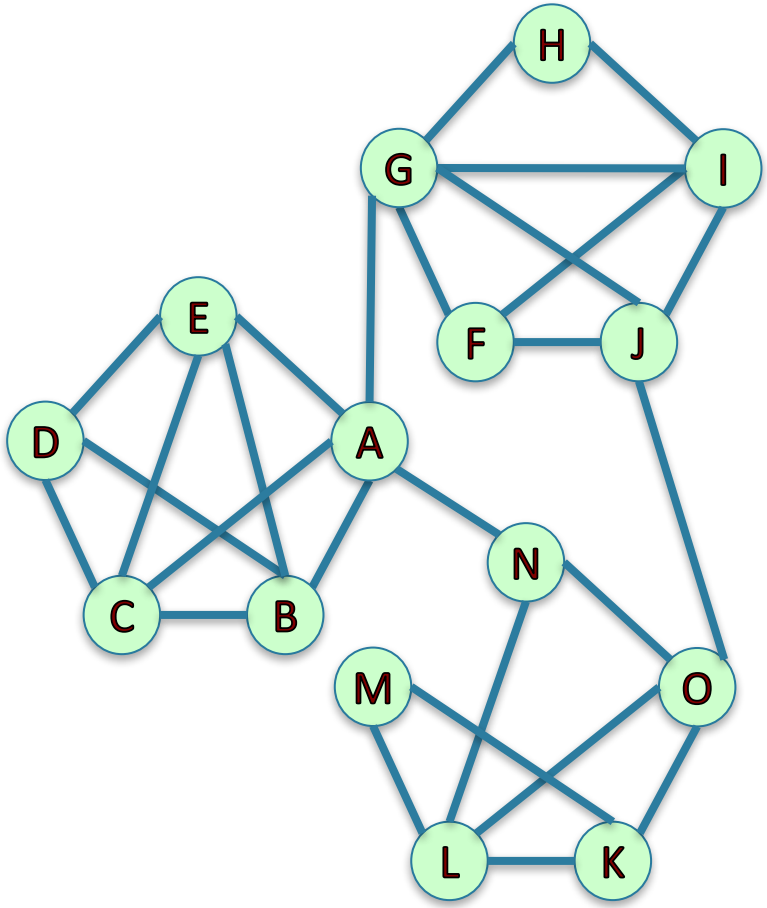
Embeddedness

Embeddedness of an edge A – B is the number of friends A and B have in common.

What's the embeddedness of A—B?

2. A and B are both friends of E and C

What's the embeddedness of A—G?



Embeddedness

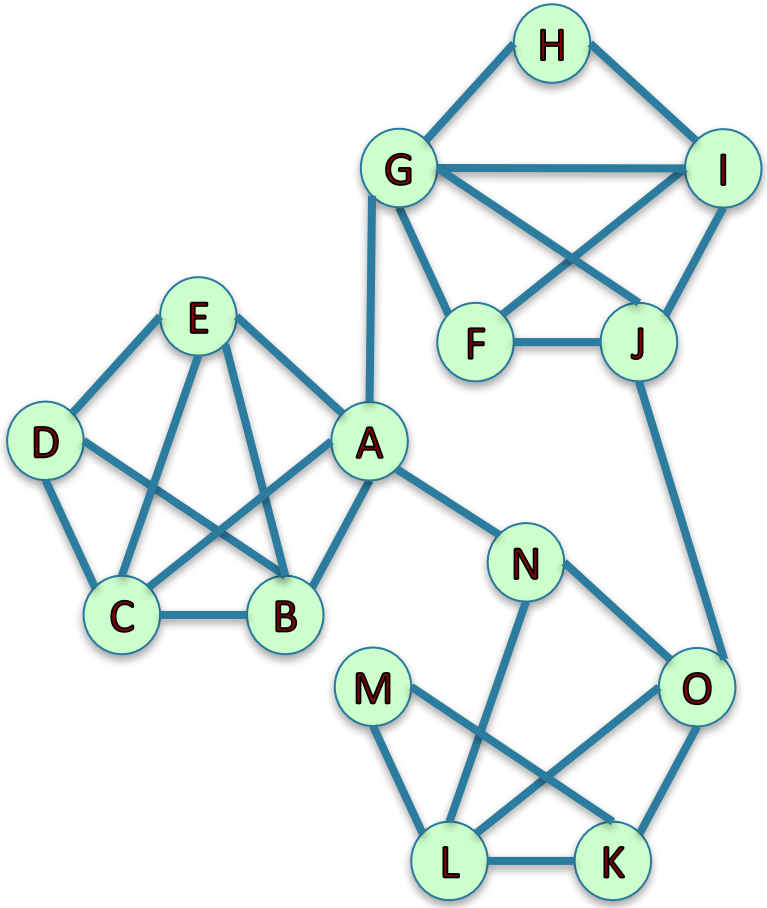
Embeddedness of an edge A – B is the number of friends A and B have in common.

What's the embeddedness of A—B?

2. A and B are both friends of E and C

What's the embeddedness of A—G?

0. A and G have no friends in common

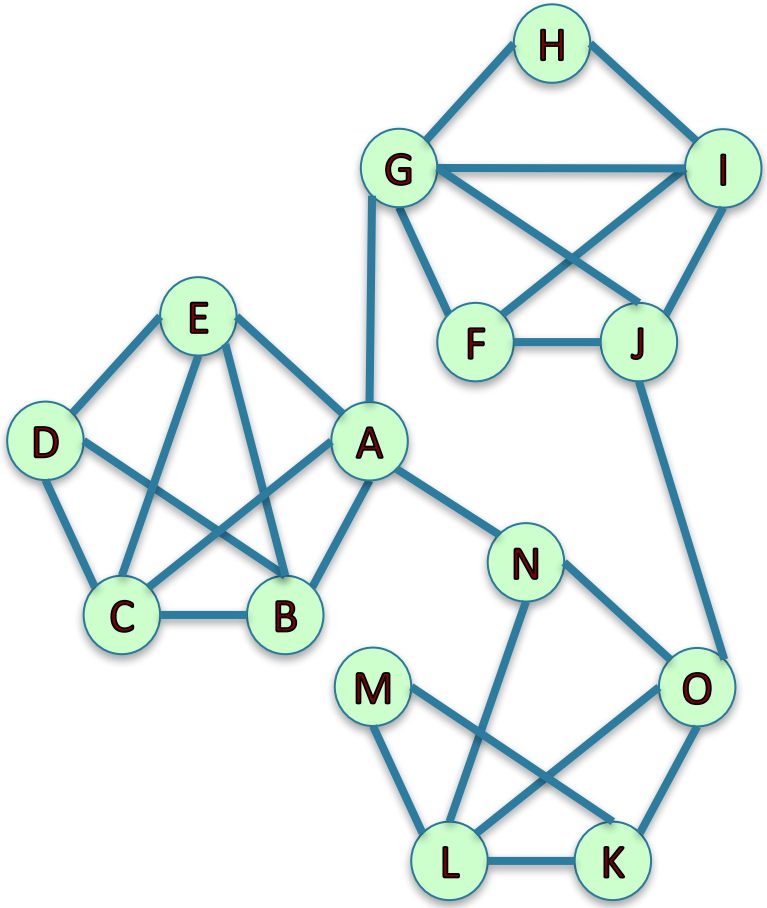


Embeddedness

Embeddedness of an edge A – B is the number of friends A and B have in common.

People connected by edges of high embeddedness can more easily trust each other.

Why?



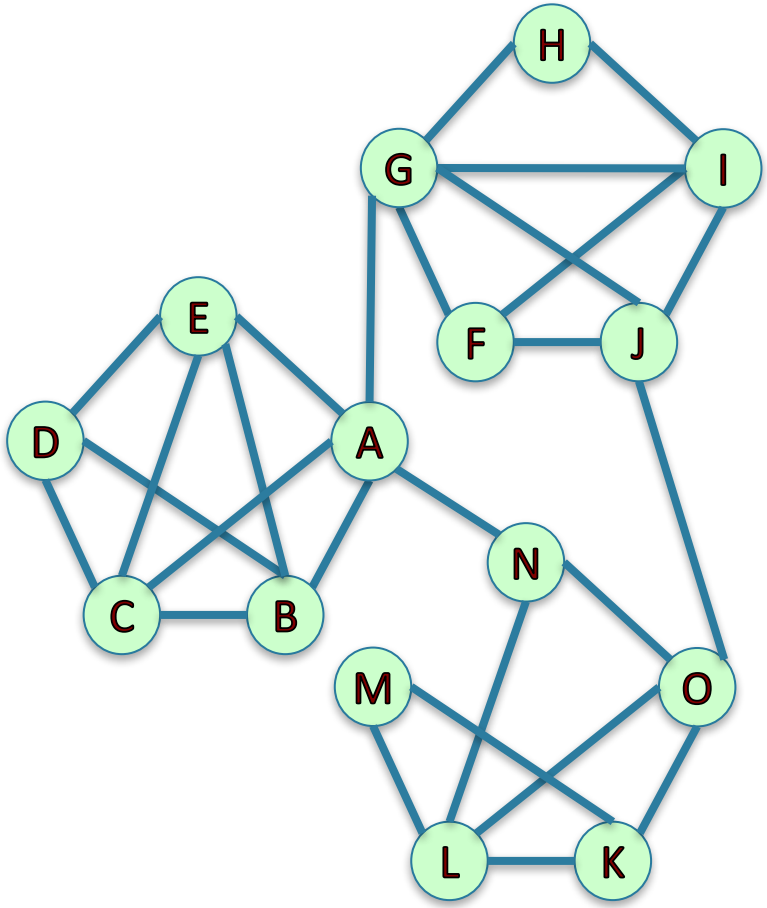
Embeddedness

Embeddedness of an edge A – B is the number of friends A and B have in common.

People connected by edges of high embeddedness can more easily trust each other.

Why?

"Misbehavour" risks, not only the pair's relationship, but also the relationship with common friends.

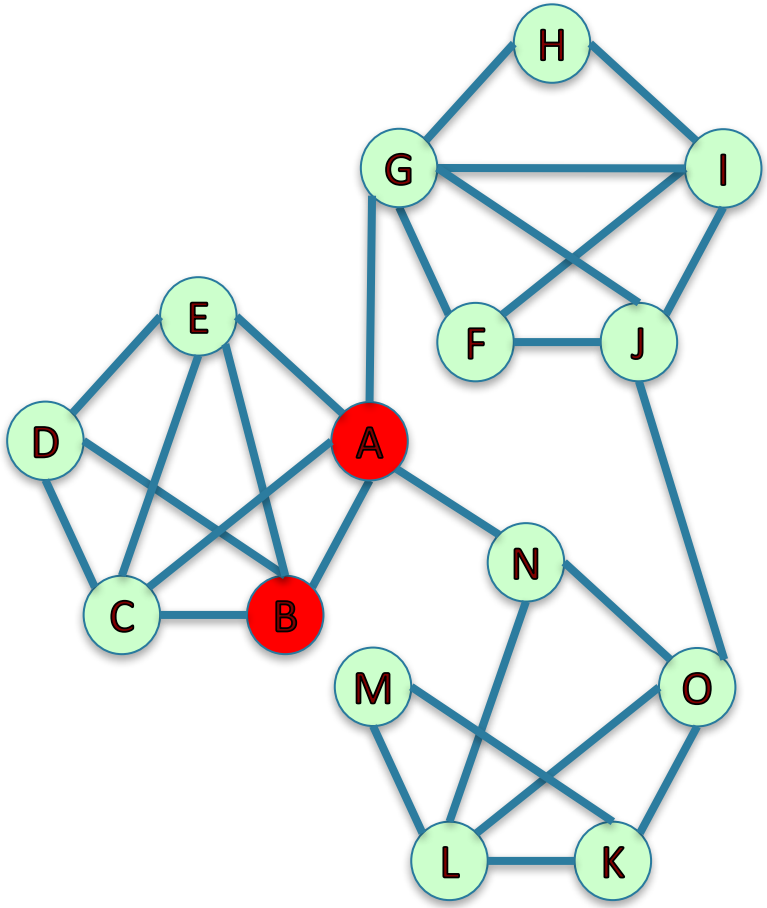


Embeddedness and Local Bridges in Organizations

Would you rather be node A or node B in this network?

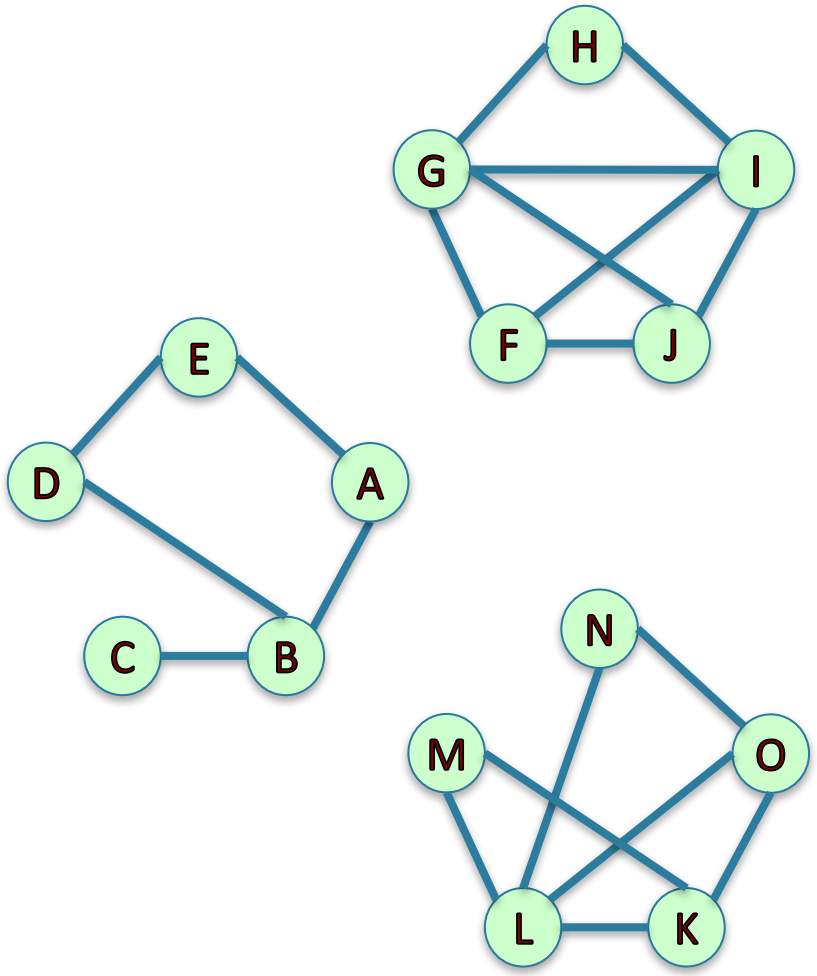
Node B: Involved in edges with high embeddedness. High clustering coefficient → **Many strong ties and high trust among her friends**.

Node A: Involved in local bridges (edges with low emdeddedness). Low clustering coefficient → **Many weak ties spanning different communities. Access to richer and more diverse information**.



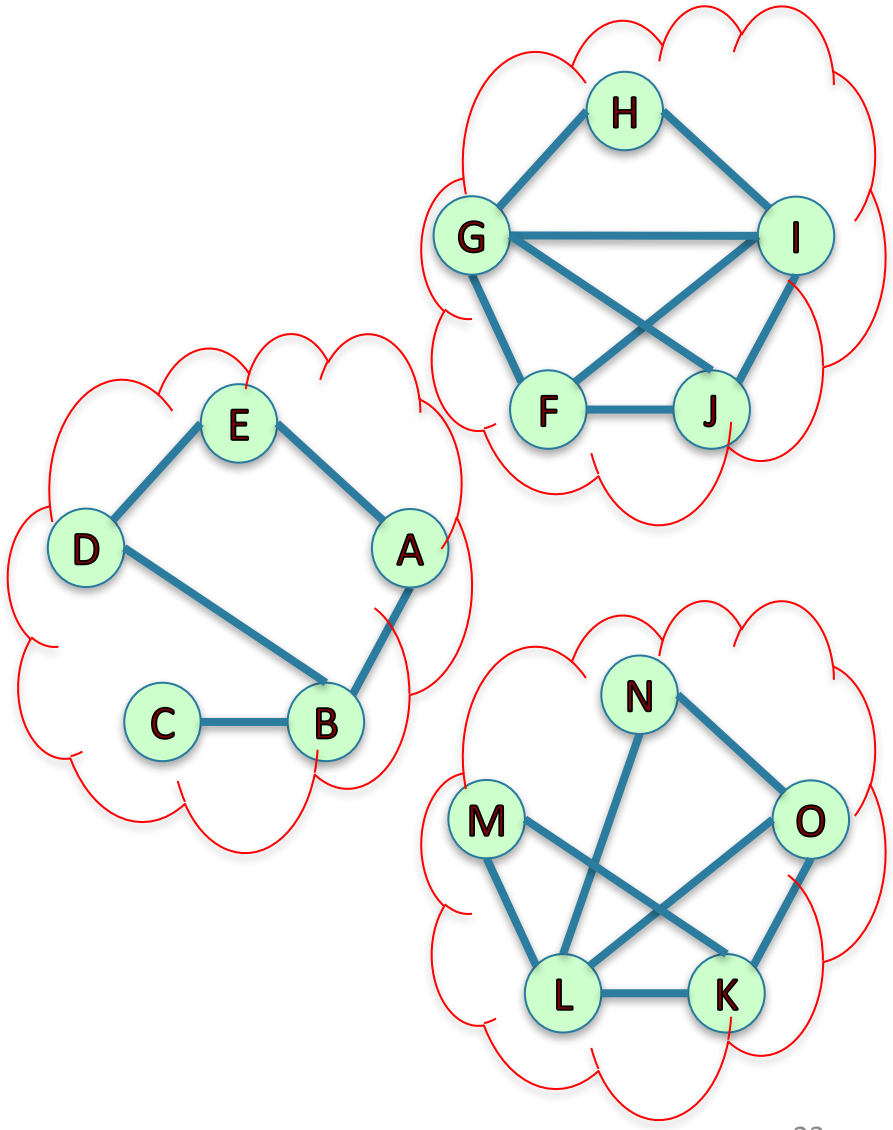
Structural Holes

Structural holes (informal): Gaps in a network that do not allow two sets of people to communicate with each other.



Structural Holes

Structural holes (informal): Gaps in a network that do not allow two sets of people to communicate with each other.



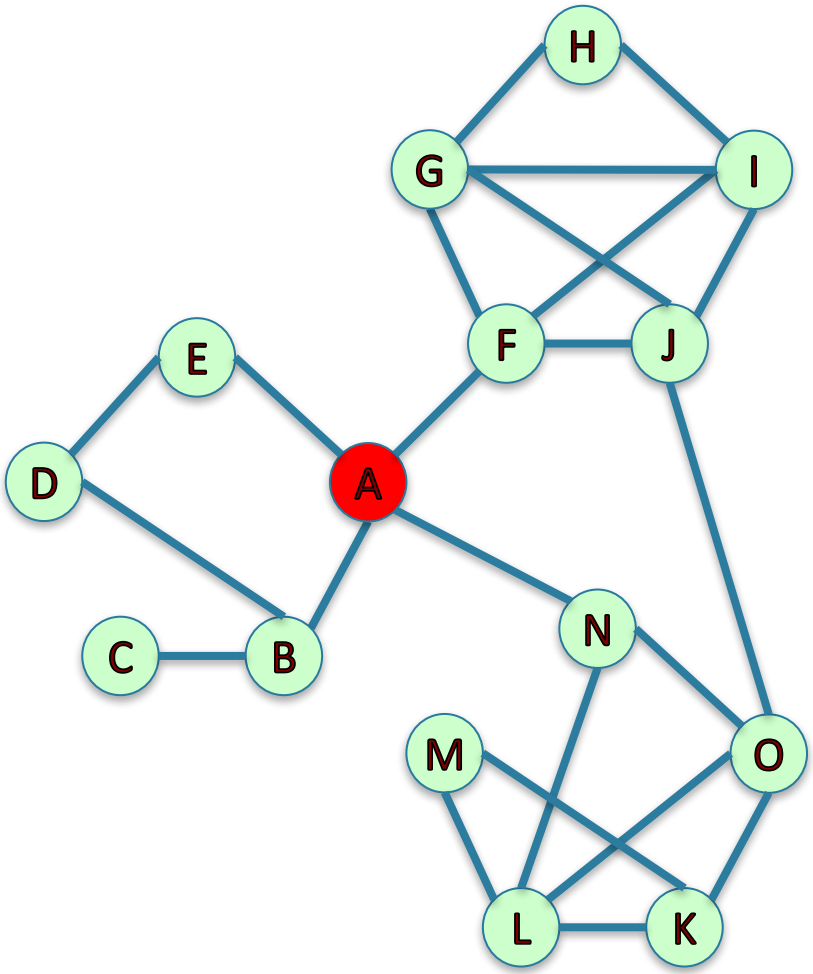
Structural Holes

Structural holes (informal): Gaps in a network that do not allow two sets of people to communicate with each other.

Being part of local bridges (like node A) has been found to correlate with:

- likelihood of promotion
- wages
- quality of ideas
- evaluation ratings

[Burt 1992, 2000, 2004]

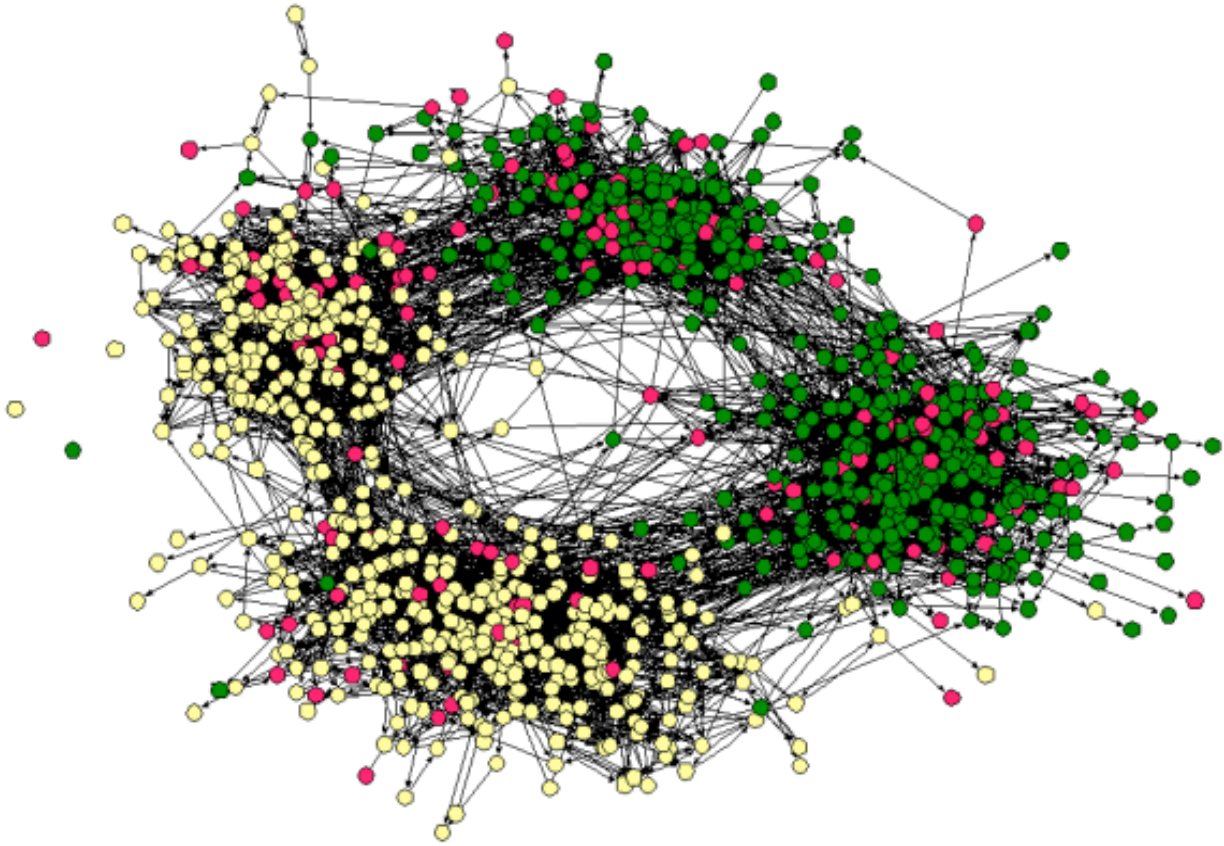


Homophily

Homophily: The principle that friends tend to be similar.

Network of a town's middle school and high school students

Color of node indicates race



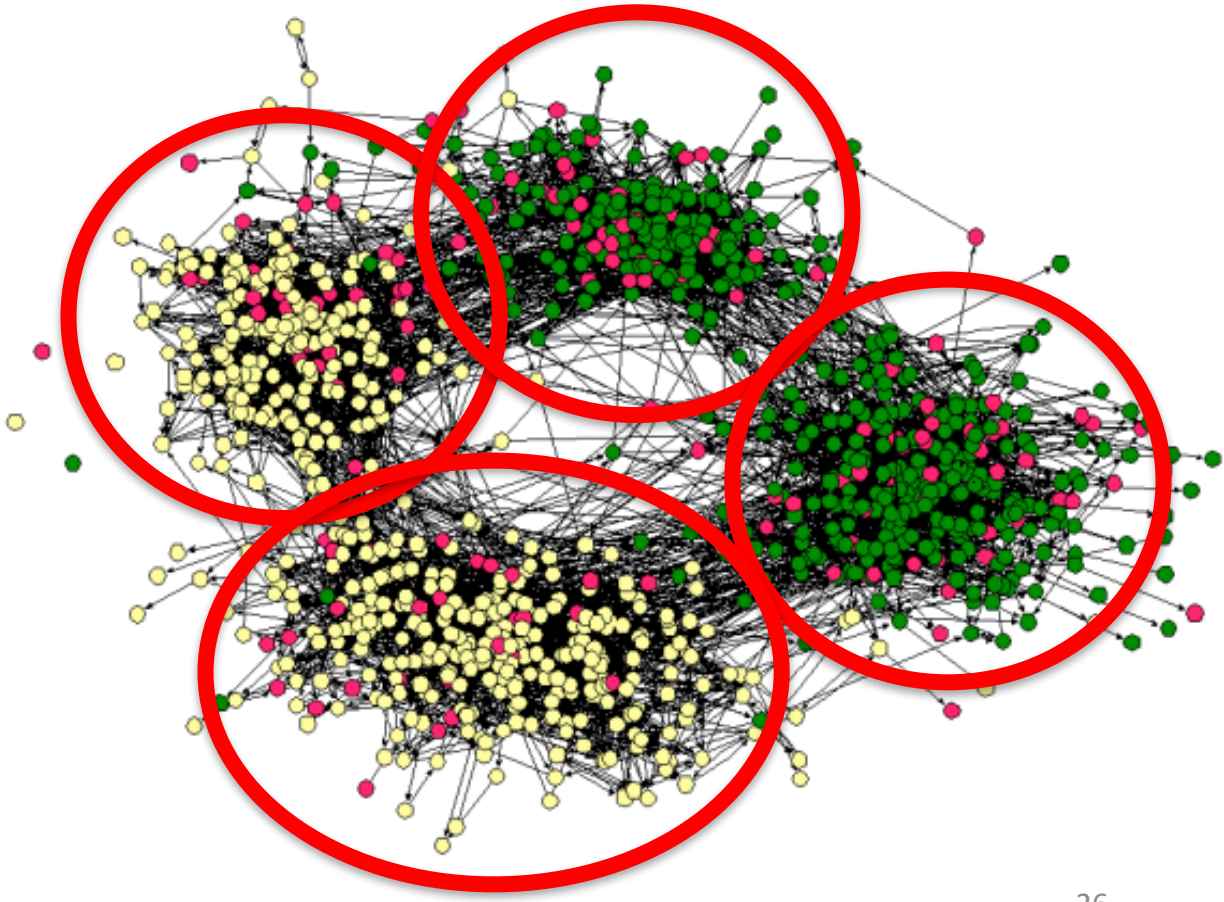
Homophily

Homophily: The principle that friends tend to be similar.

Network of a town's middle school and high school students

Color of node indicates race

Nodes cluster by race and school level



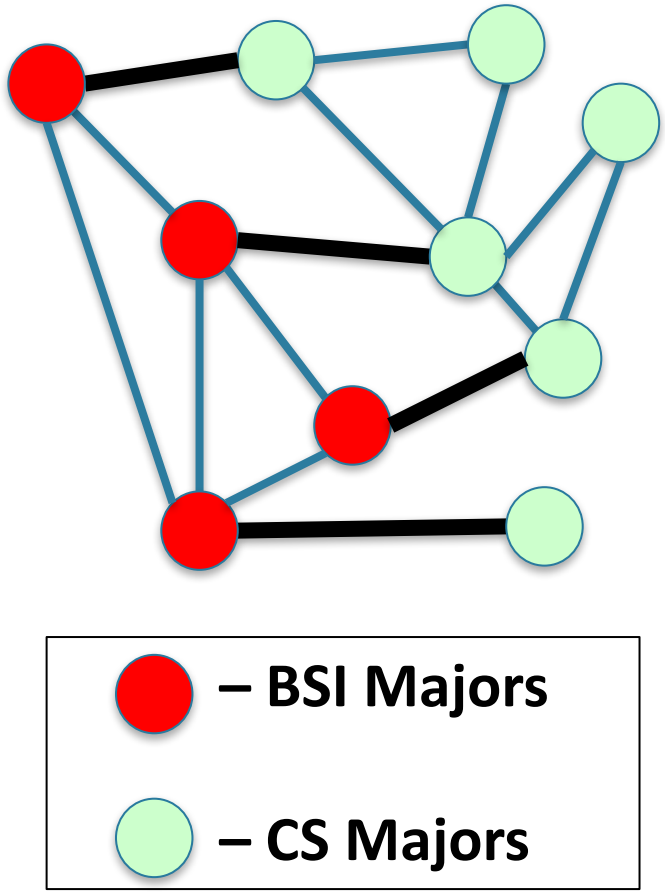
Measuring Homophily

Is homophily present in this network?

4/15 = 26.65% of edges connect nodes of different majors.

Without homophily, what percentage of edges should be cross major?

Compare with random major assignment.

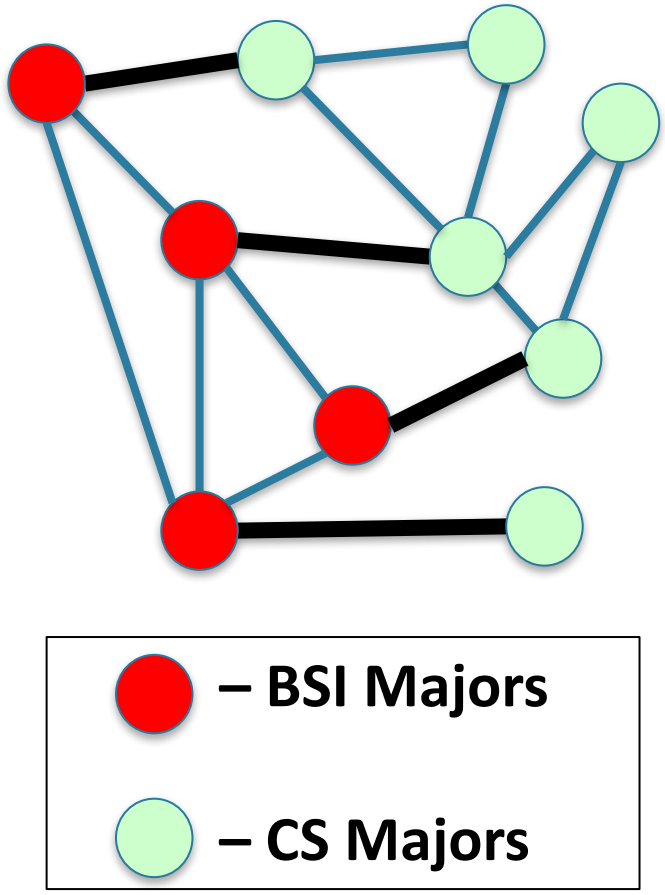


Measuring Homophily

Thought experiment:

- A population has 40% BSI majors and 60% CS majors
- If we randomly assign a major to the endpoints of an edge (with 0.6 prob. of CS and 0.4 prob. of BSI), what's the probability that they will both be CS majors?

$0.6*0.6 = 0.6^2 = 0.36$

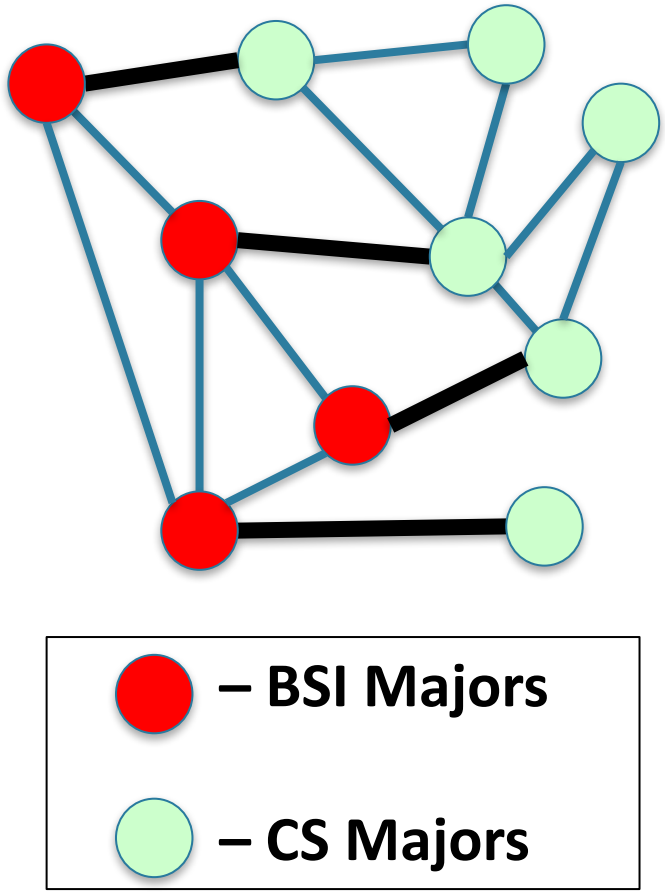


Measuring Homophily

Thought experiment:

- A population has 40% BSI majors and 60% CS majors
- If we randomly assign a major to the endpoints of an edge (with 0.6 prob. of CS and 0.4 prob. of BSI), what's the probability that they will both be BSI majors?

$0.4*0.4 = 0.4^2 = 0.16$

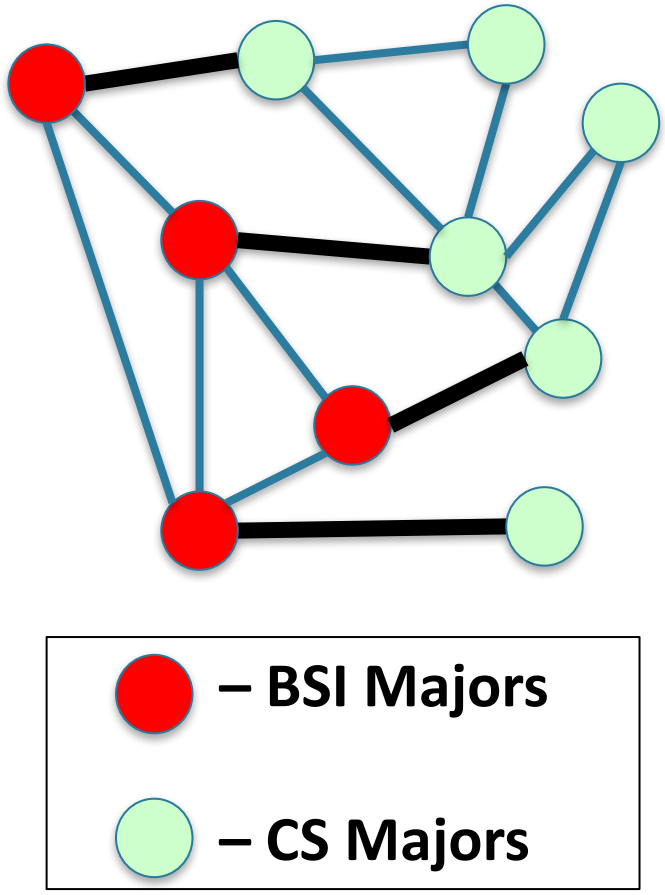


Measuring Homophily

Thought experiment:

- A population has 40% BSI majors and 60% CS majors
- If we randomly assign a major to the endpoints of an edge (with 0.6 prob. of CS and 0.4 prob. of BSI), what's the probability that one will be a BSI major and the other a CS major?

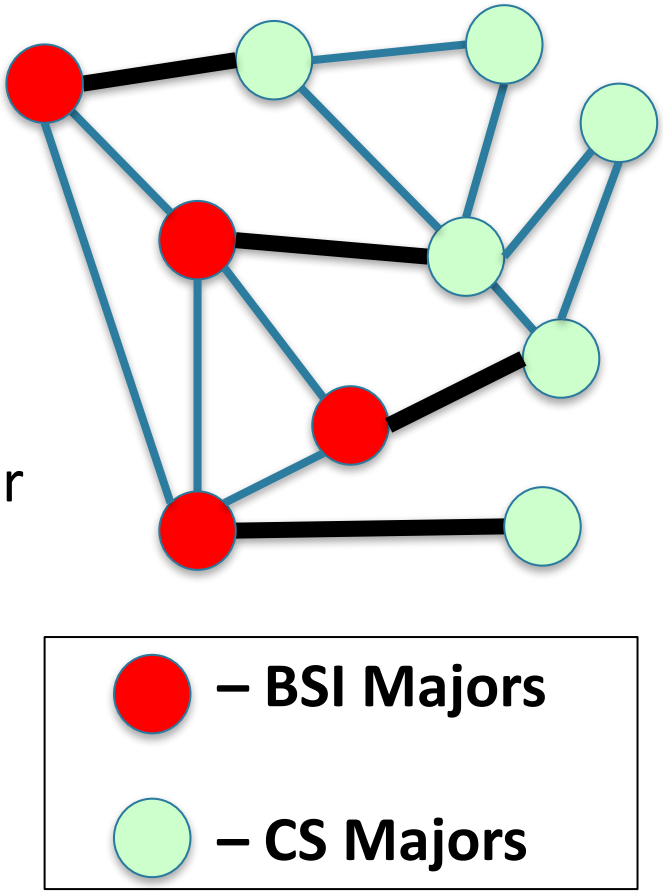
2*0.6*0.4 = 0.48



Measuring Homophily

Thought experiment:

- A population has 40% BSI majors and 60% CS majors
- If we randomly assign a major to the endpoints of an edge (with 0.6 prob. of CS and 0.4 prob. of BSI), what's the probability that one will be a BSI major and the other a CS major?



2*0.6*0.4 = 0.48

Without homophily, we would expect 48% of all edges in this network to be cross-major. However, only 4/15 = 26.6% are. **Homphily seems to be present.**

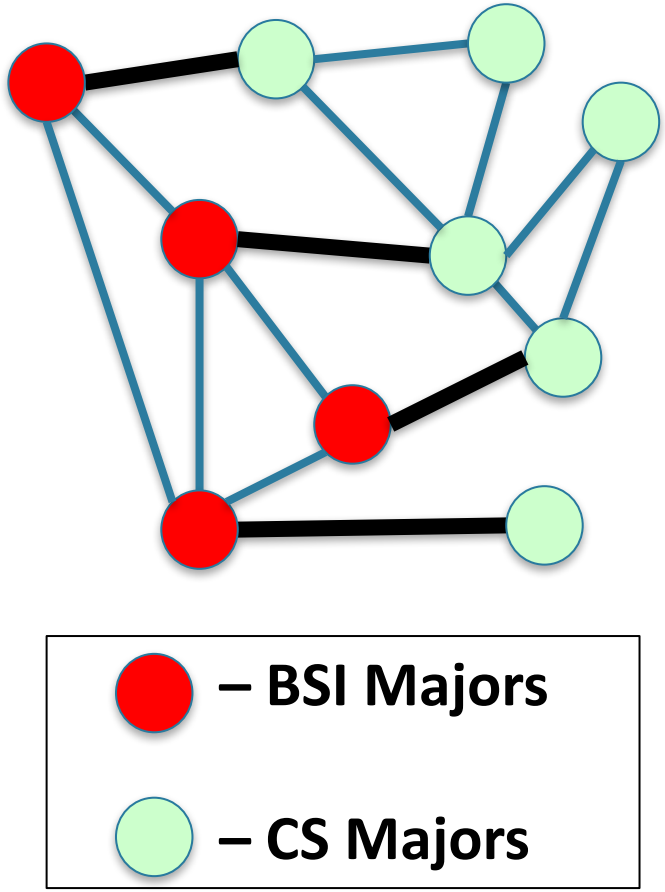
Measuring Homophily

Homophily test:

Given network with two types of nodes such that:

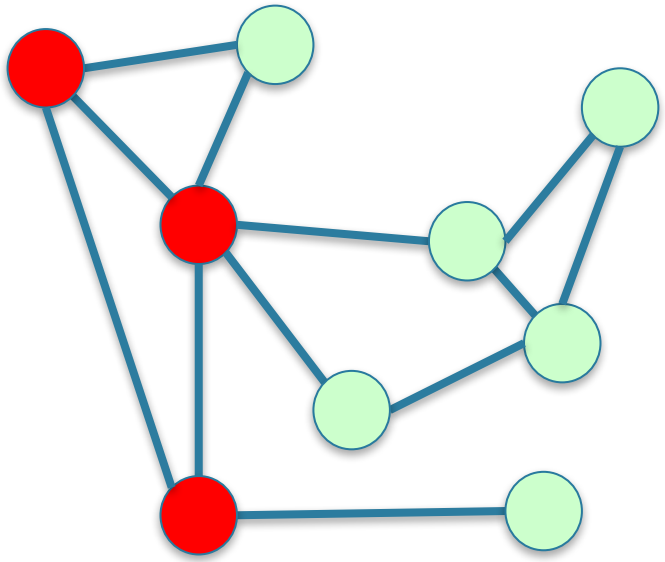
- A *p* fraction of nodes are of type 1
- A *q* fraction of nodes are of type 2.
- A *c* fraction of the edges are cross-type (the two endpoints are of different type)

Then we say there is evidence for homophily if $c < 2pq$.



Measuring Homophily

Run the homophily test on the network:



Measuring Homophily

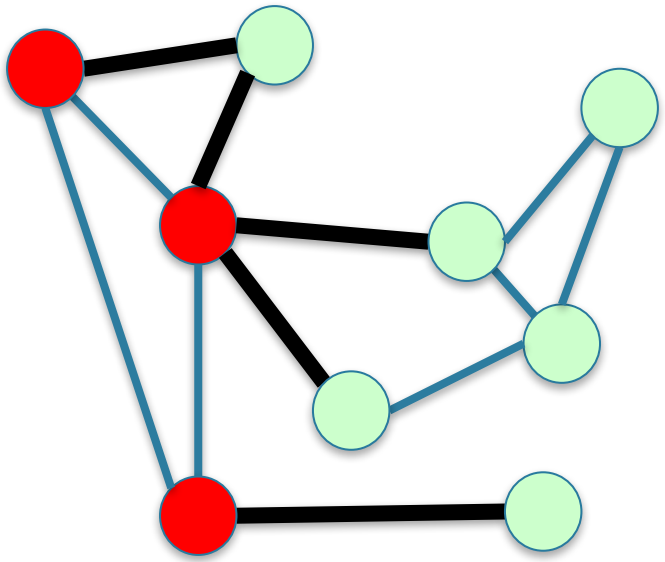
Run the homophily test on the network:

p =

q =

$2*p*q$ =

Frac. cross-type edges =



Measuring Homophily

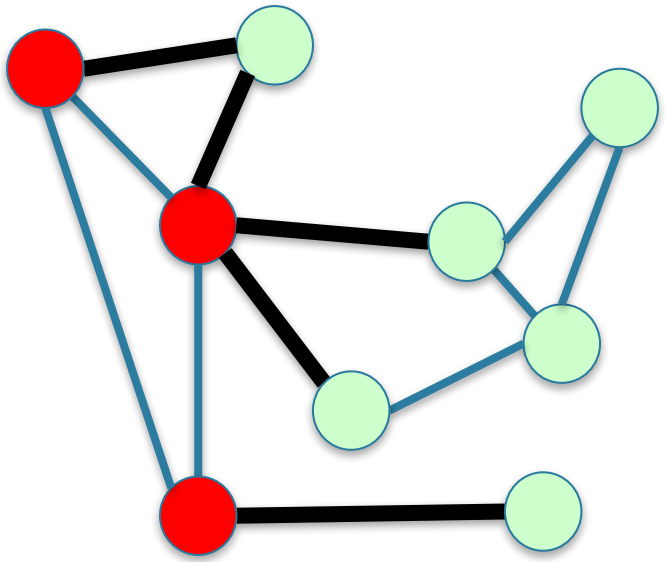
Run the homophily test on the network:

$p = 3/9$

$q =$

$2*p*q =$

Frac. cross-type edges =



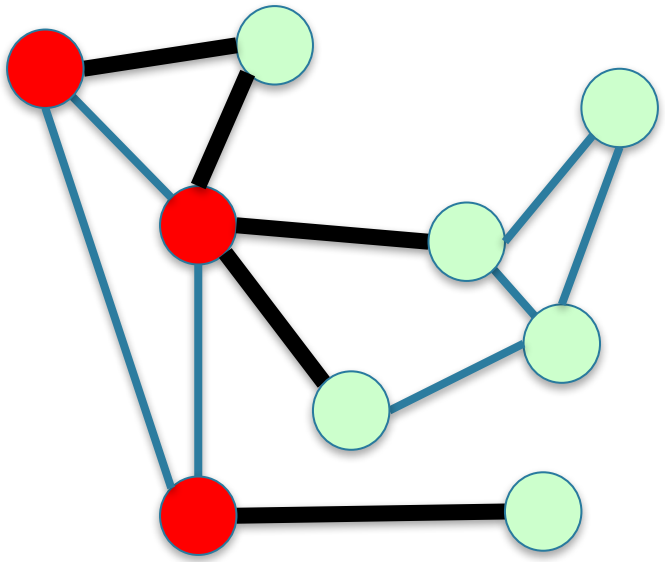
Measuring Homophily

Run the homophily test on the network:

$p = 3/9$
$q = 6/9$

$2*p*q =$

Frac. cross-type edges =



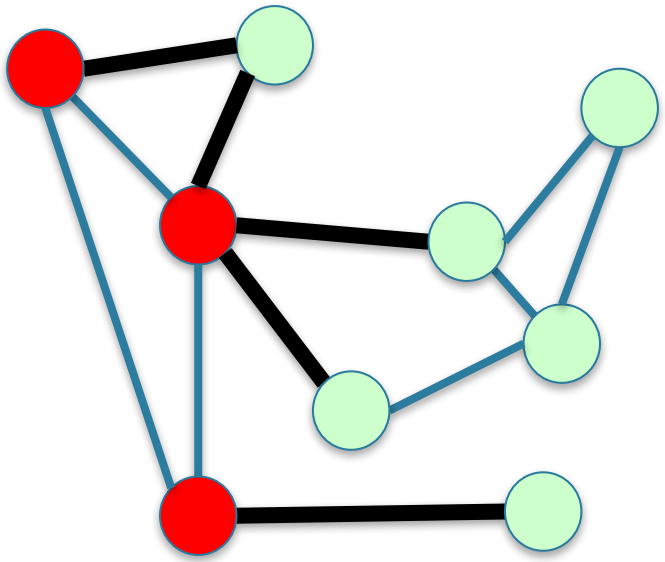
Measuring Homophily

Run the homophily test on the network:

$p = 3/9$
$q = 6/9$

$2*p*q = 0.44$

Frac. cross-type edges =



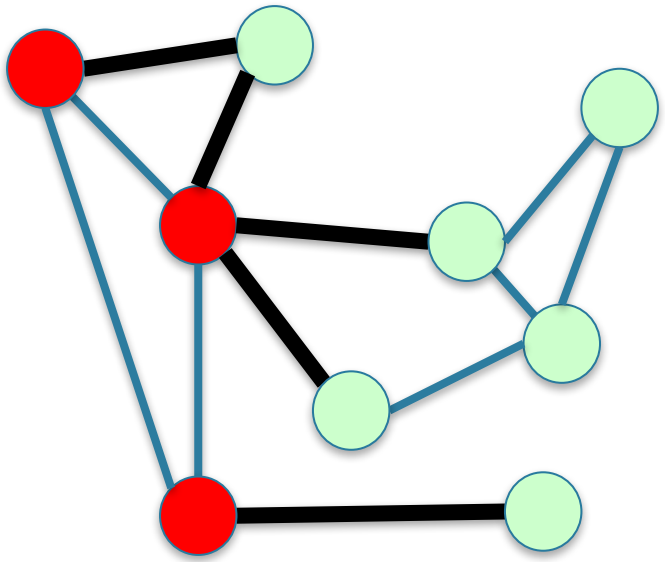
Measuring Homophily

Run the homophily test on the network:

$p = 3/9$
$q = 6/9$

$2*p*q = 0.44$

Frac. cross-type edges = 5/12 = 0.416



Measuring Homophily

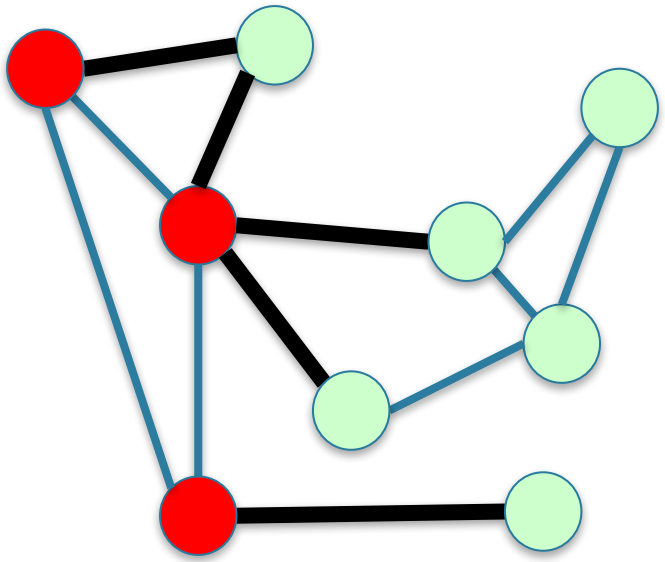
Run the homophily test on the network:

$p = 3/9$
$q = 6/9$

$2*p*q = 0.44$

Frac. cross-type edges = 5/12 = 0.416

Since frac. cross-type edges is less than 2*p*q, then we find that there is evidence of homophily.



Drivers of Homophily

Drivers of Homophily

Selection: A tendency of people to form friendships with others who are like them (A and B are similar ***then*** they become friends).

Drivers of Homophily

Selection: A tendency of people to form friendships with others who are like them (A and B are similar ***then*** they become friends).

Examples:

- Elementary school students often prefer to form friendship with those of the same gender (intentional).

Drivers of Homophily

Selection: A tendency of people to form friendships with others who are like them (A and B are similar ***then*** they become friends).

Examples:

- Elementary school students often prefer to form friendship with those of the same gender (intentional).
- Most people live in the same country as most of their friends (unintentional).

Drivers of Homophily

Selection: A tendency of people to form friendships with others who are like them (A and B are similar ***then*** they become friends).

Examples:

- Elementary school students often prefer to form friendship with those of the same gender (intentional).
- Most people live in the same country as most of their friends (unintentional).

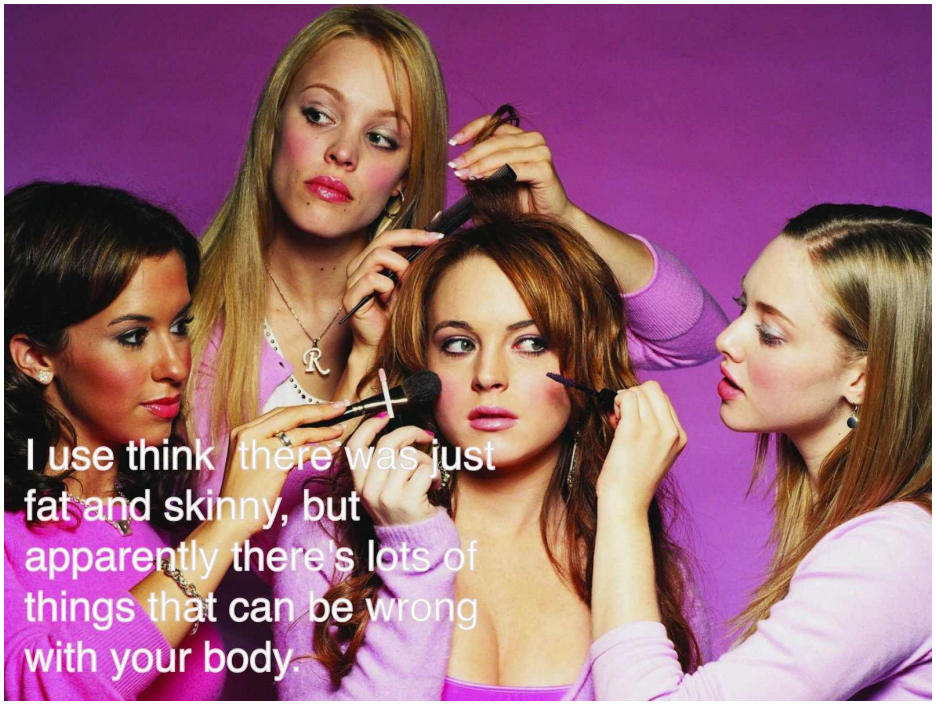
Note: Selection doesn't necessarily imply intentionality.

Drivers of Homophily

Socialization or Social Influence: A tendency for people to modify their behavior to match their friends' (A and B are friends ***then*** they become similar).

Examples:

- People may adopt a similar taste in music as their friends.
- Political opinions can often change due to social interactions/pressure.

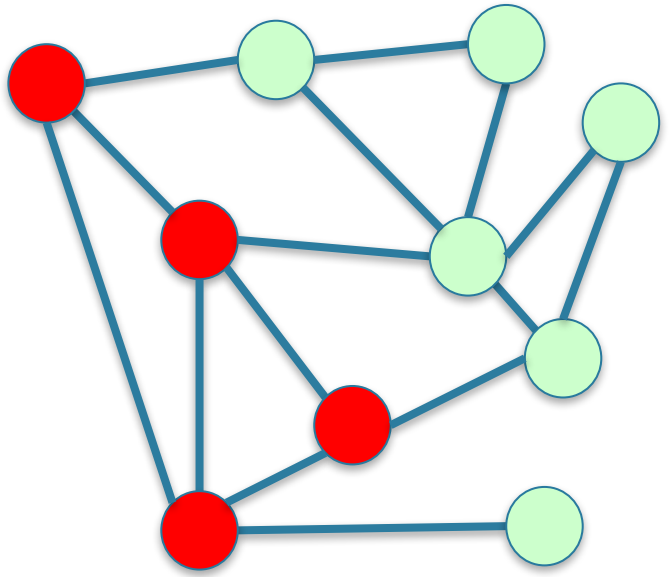


Distinguishing Between Selection and Social Influence

Is the homophily in this network due to selection, social influence, or something else?

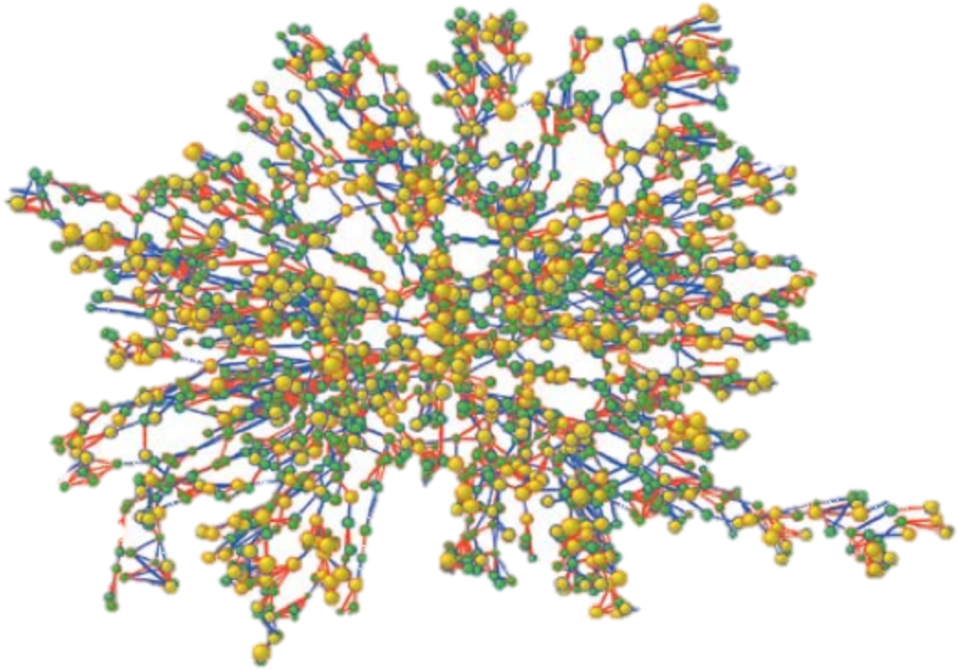
Impossible to tell.
We would need information about the temporal dynamics of the network.

Specific mechanisms are important to identify if we want to make interventions to encourage or discourage a certain behavior.



The Spread of Obesity in Social Networks

- Social network of 12,067 people assessed repeatedly for 32 years (1971 to 2003).
- Yellow node denotes an obese person (body-mass index ≥ 30) and green denotes a nonobese person (sample network from 2000).



[Christakis – Fowler 2007]

- Homophily was observed in all 32 years.
- Longitudinal statistical models suggest that **social influence** and **not selection** was the driver of homophily.