Social and Information Networks

CSCC46H, Fall 2025

Lecture 2

Prof. Ashton Anderson
ashton@cs.toronto.edu

Logistics

A1 out next week

Today

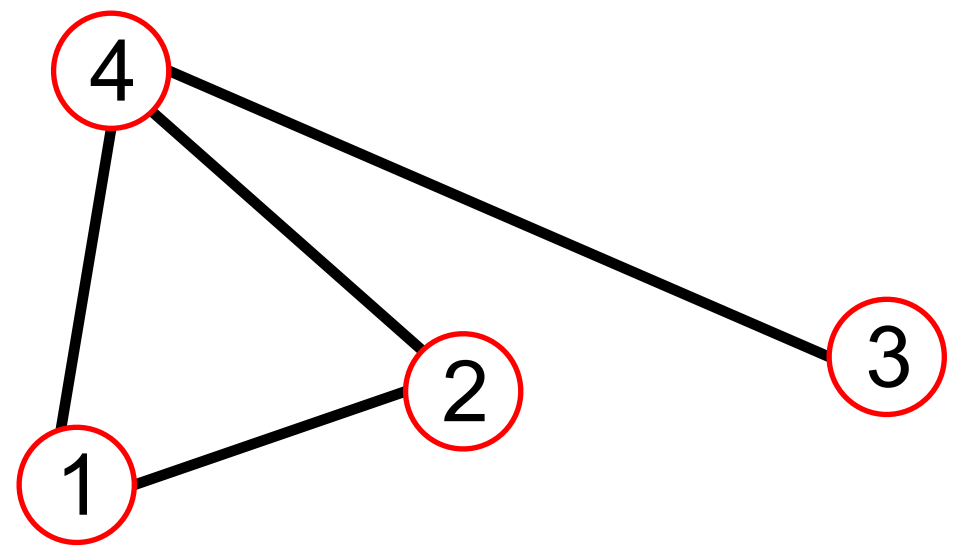
1) Building up our network vocabulary
2) Measuring networks; basic properties
3) Random graph model: G_{np}
4) Strong and weak ties (time willing)

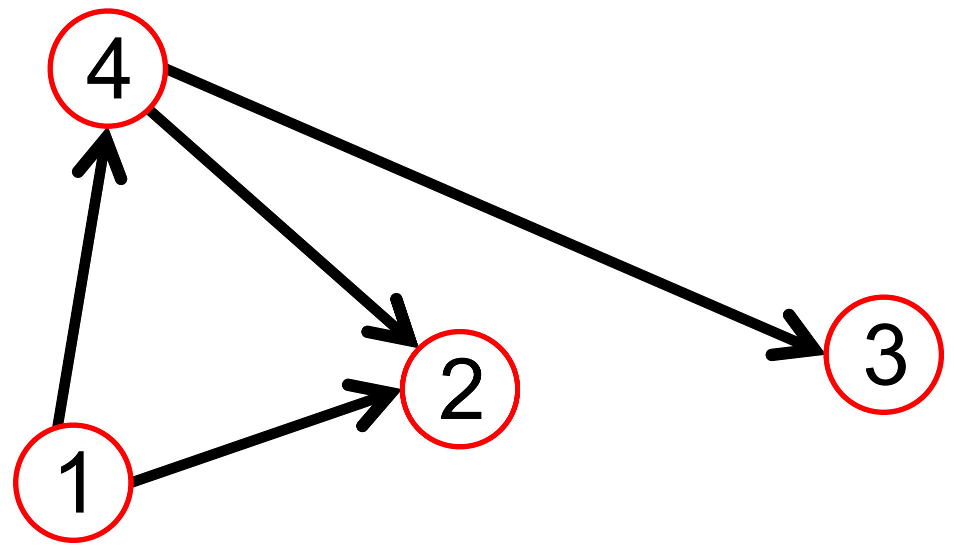
Network Representations

How do we represent graphs as mathematical objects?

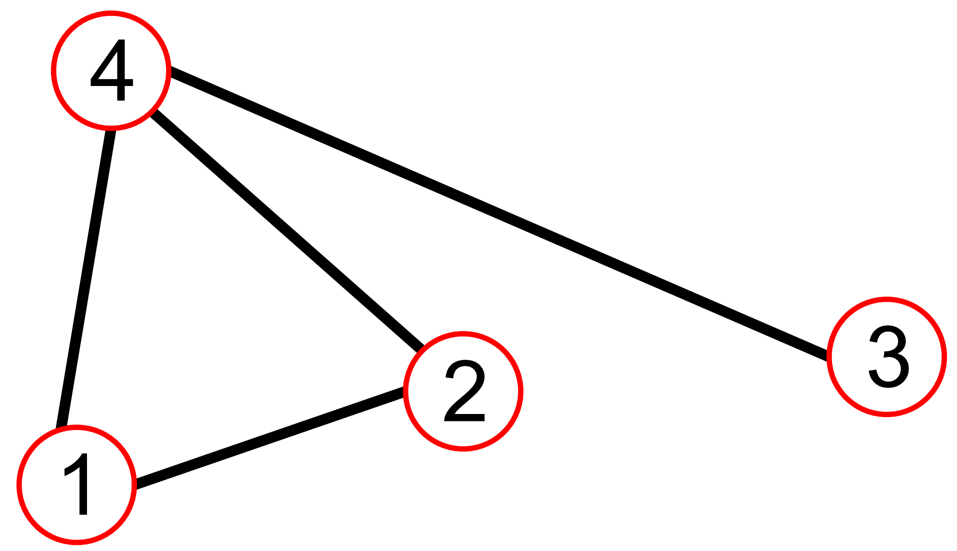
What are our choices when we're translating real-world networks into a graph representation?

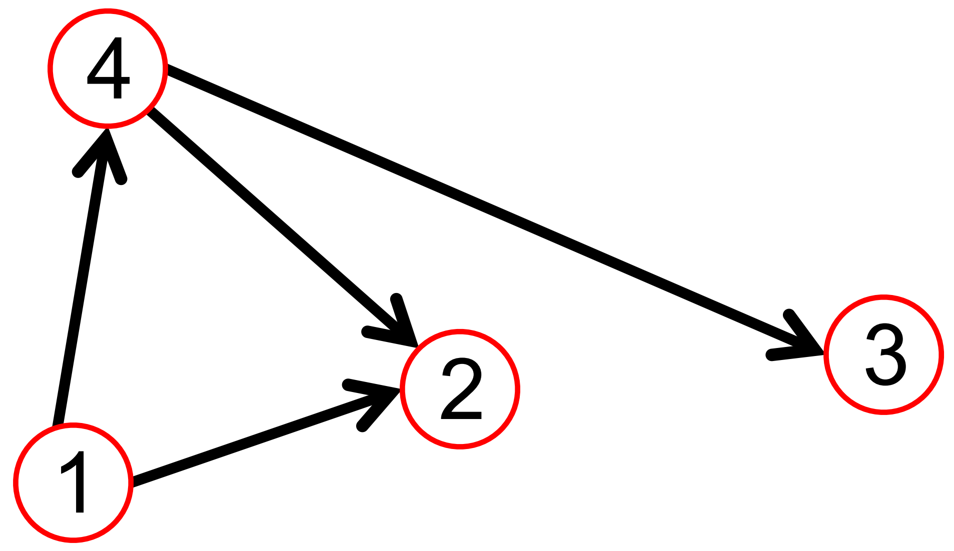
How to Store?





Edge List

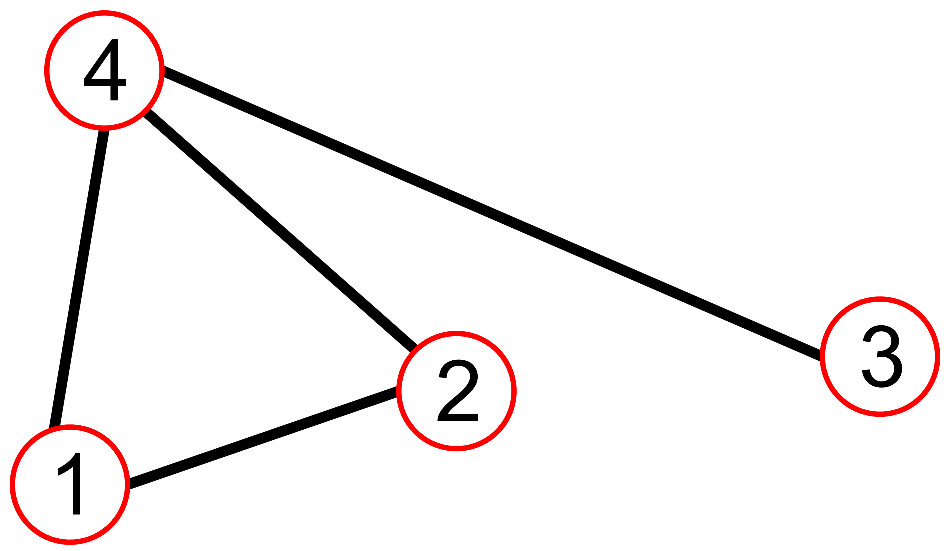


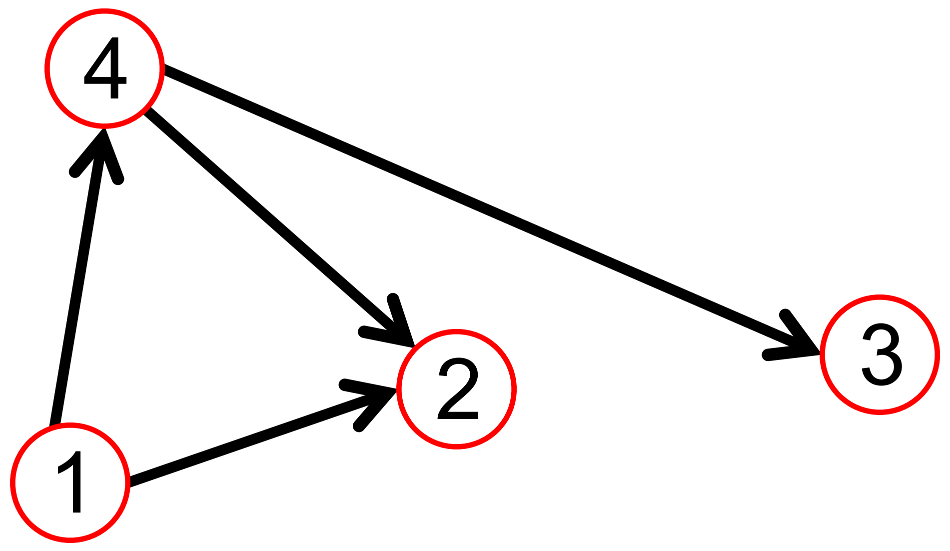


[(1,2),
(1,4),
(2,4),
(3,4)]

[(1,2),
(1,4),
(4,2),
(4,3)]

Adjacency List



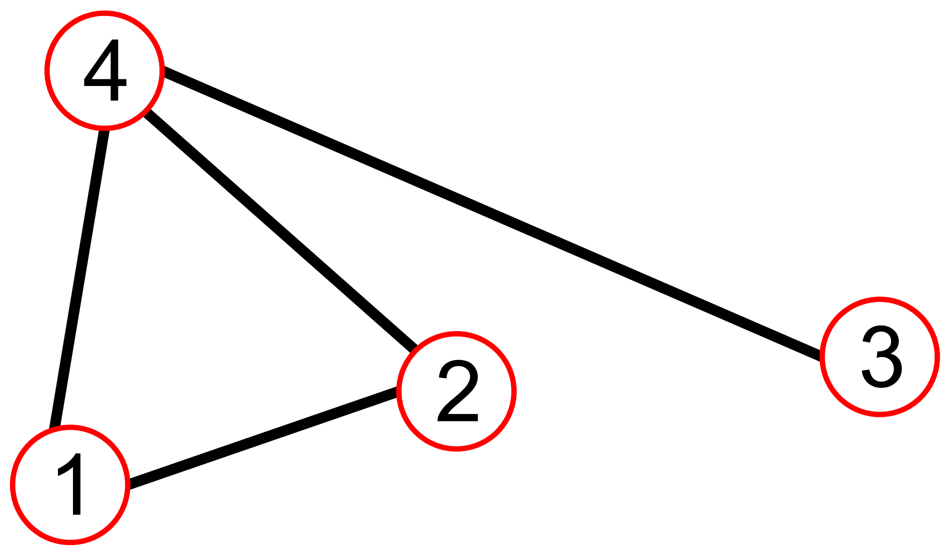


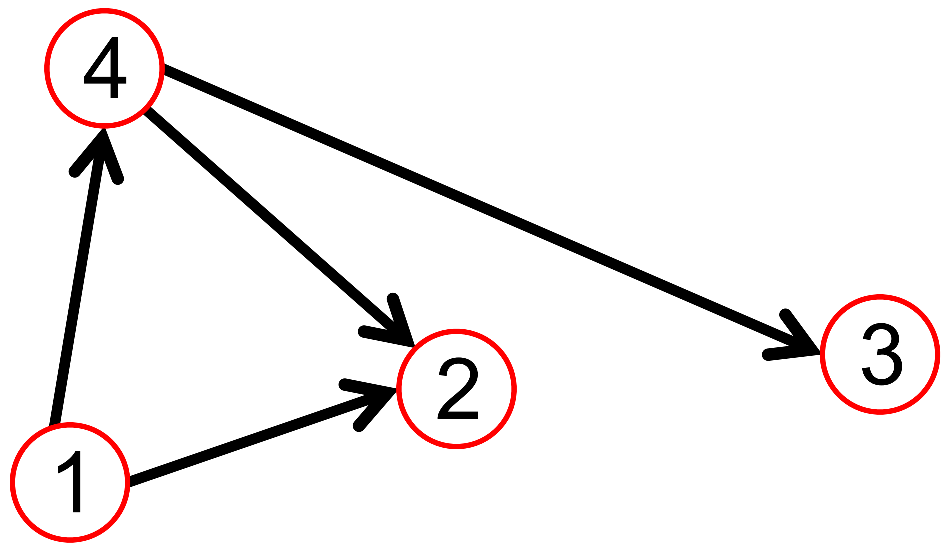
{1: [2,4],
 2: [1,4],
 3: [4],
 4: [1,2,3]}

{1: [2,4],
 4: [2,3]}

Total length of lists?

Adjacency Matrix





$A_{ij} = 1$ if there is a link from node *i* to node *j*

$A_{ij} = 0$ otherwise

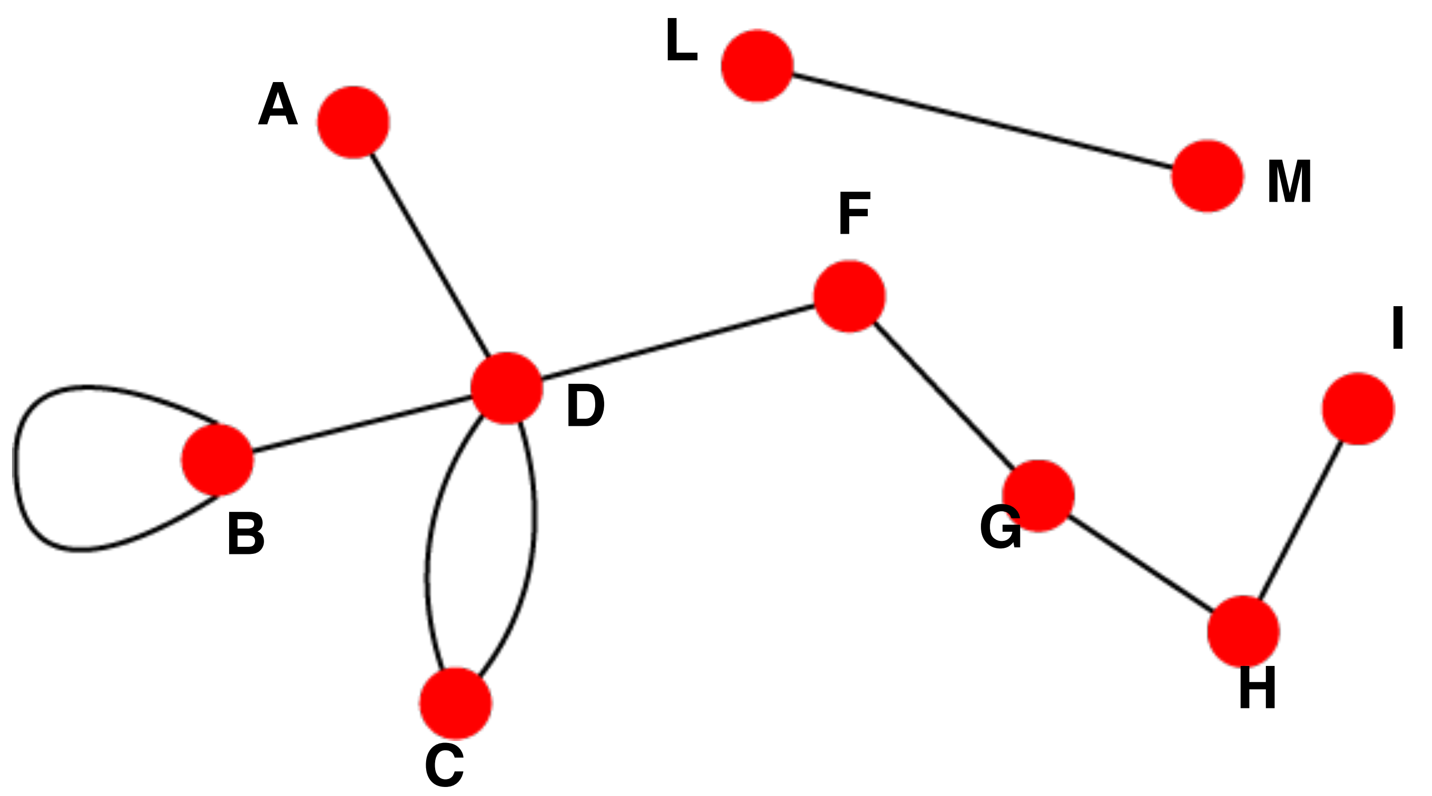
$$A = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix} \qquad A = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{pmatrix}$$

Note that for a directed graph (right) the matrix is not symmetric.

Undirected vs. Directed Networks

Undirected graphs

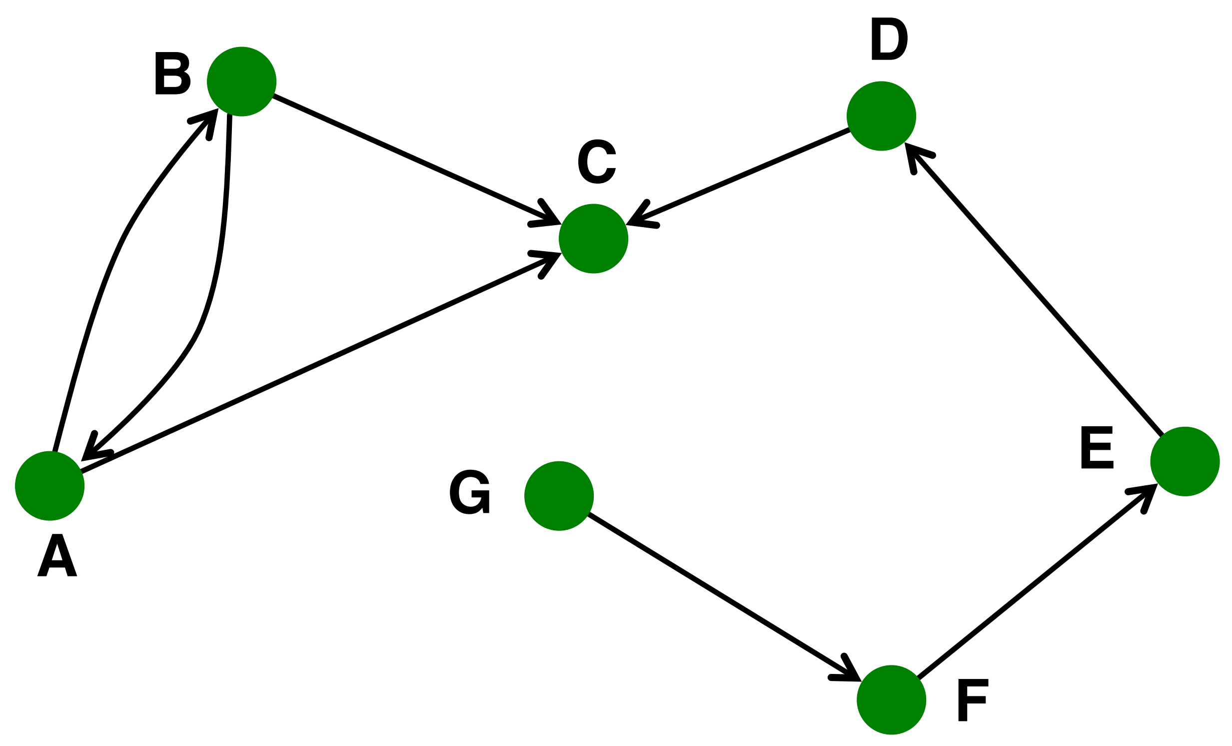
- **Links:** undirected (symmetrical, reciprocal relations)



- Undirected links:
  - Collaborations
  - Friendship on Facebook

Directed graphs

- **Links:** directed (asymmetrical relations)

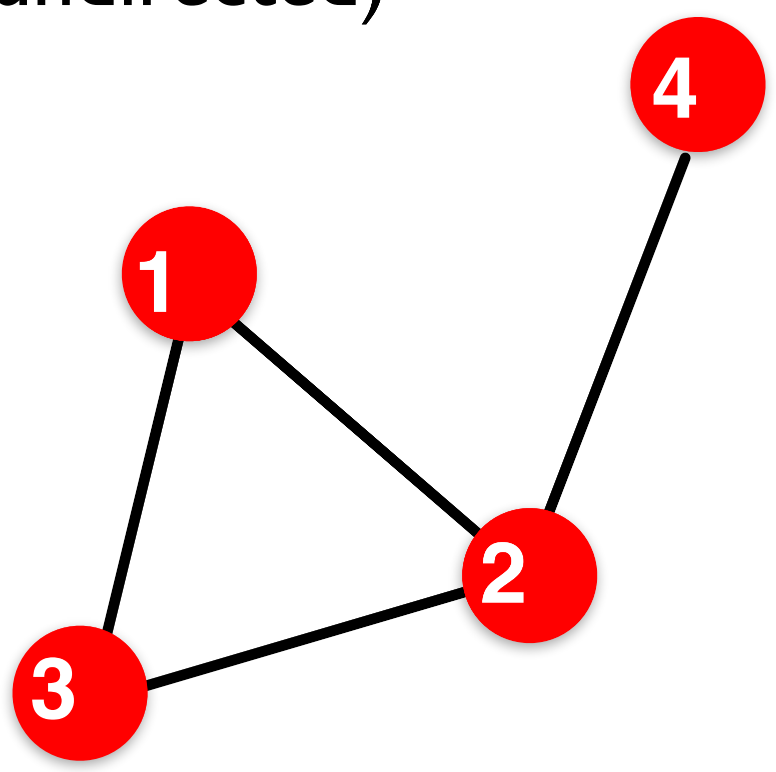


- Directed links:
  - Phone calls
  - Following on Twitter

Weighted Graphs

Unweighted

(undirected)



$$A_{ij} = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

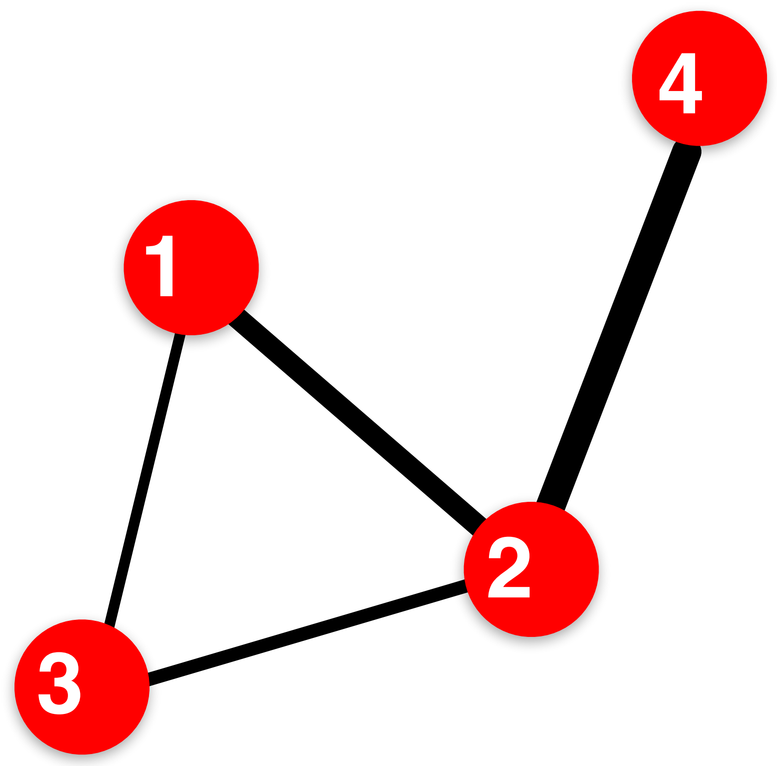
$$A_{ii} = 0 \qquad A_{ij} = A_{ji}$$

$$E = \frac{1}{2}\sum_{i,j=1}^{N} A_{ij} \qquad \bar{k} = \frac{2E}{N}$$

Examples: Friendship, Hyperlink

Weighted

(undirected)



$$A_{ij} = \begin{pmatrix} 0 & 2 & 0.5 & 0 \\ 2 & 0 & 1 & 4 \\ 0.5 & 1 & 0 & 0 \\ 0 & 4 & 0 & 0 \end{pmatrix}$$

$$A_{ii} = 0 \qquad A_{ij} = A_{ji}$$

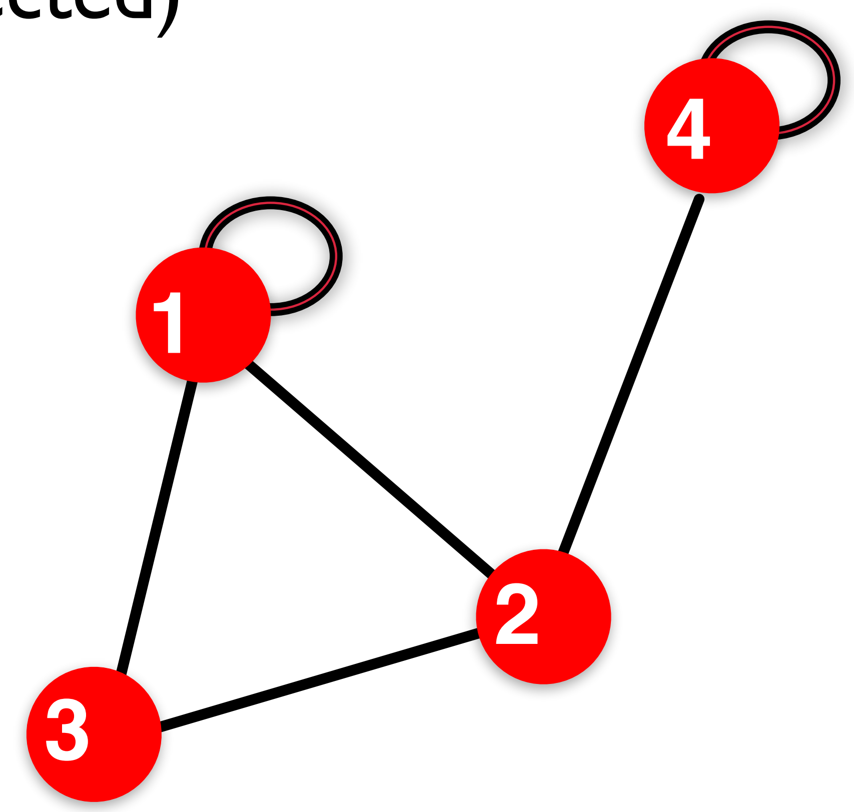
$$E = \frac{1}{2}\sum_{i,j=1}^{N} nonzero(A_{ij}) \qquad \bar{k} = \frac{2E}{N}$$

Examples: Collaboration, Internet, Roads

More Types of Graphs:

Graphs with self-edges

(undirected)



$$A_{ij} = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}$$

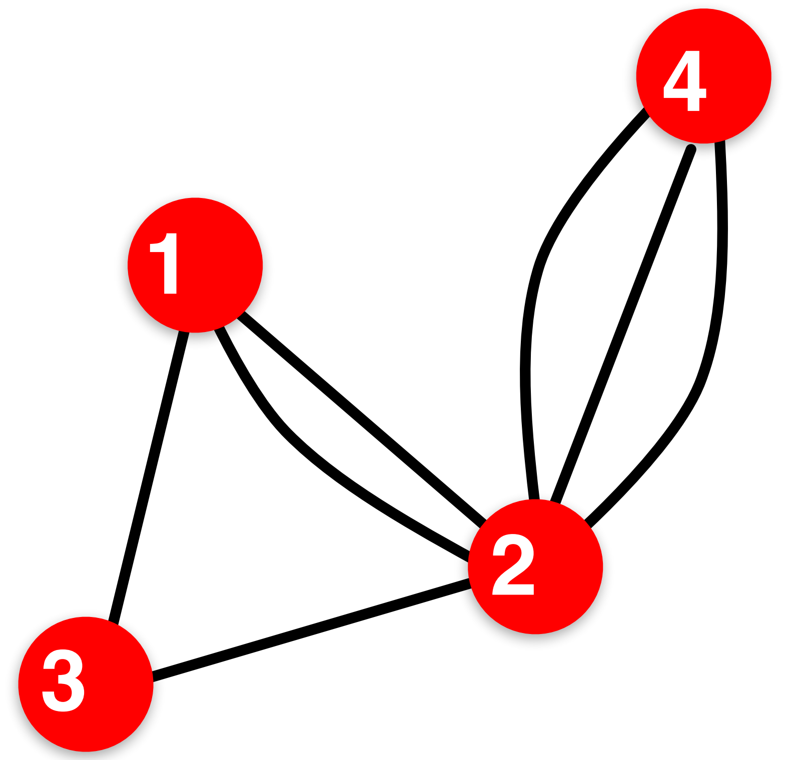
$$A_{ii} \neq 0 \qquad A_{ij} = A_{ji}$$

$$E = \frac{1}{2} \sum_{i,j=1, i \neq j}^{N} A_{ij} + \sum_{i=1}^{N} A_{ii}$$

Examples: Proteins, Hyperlinks

Multigraph

(undirected)



$$A_{ij} = \begin{pmatrix} 0 & 2 & 1 & 0 \\ 2 & 0 & 1 & 3 \\ 1 & 1 & 0 & 0 \\ 0 & 3 & 0 & 0 \end{pmatrix}$$

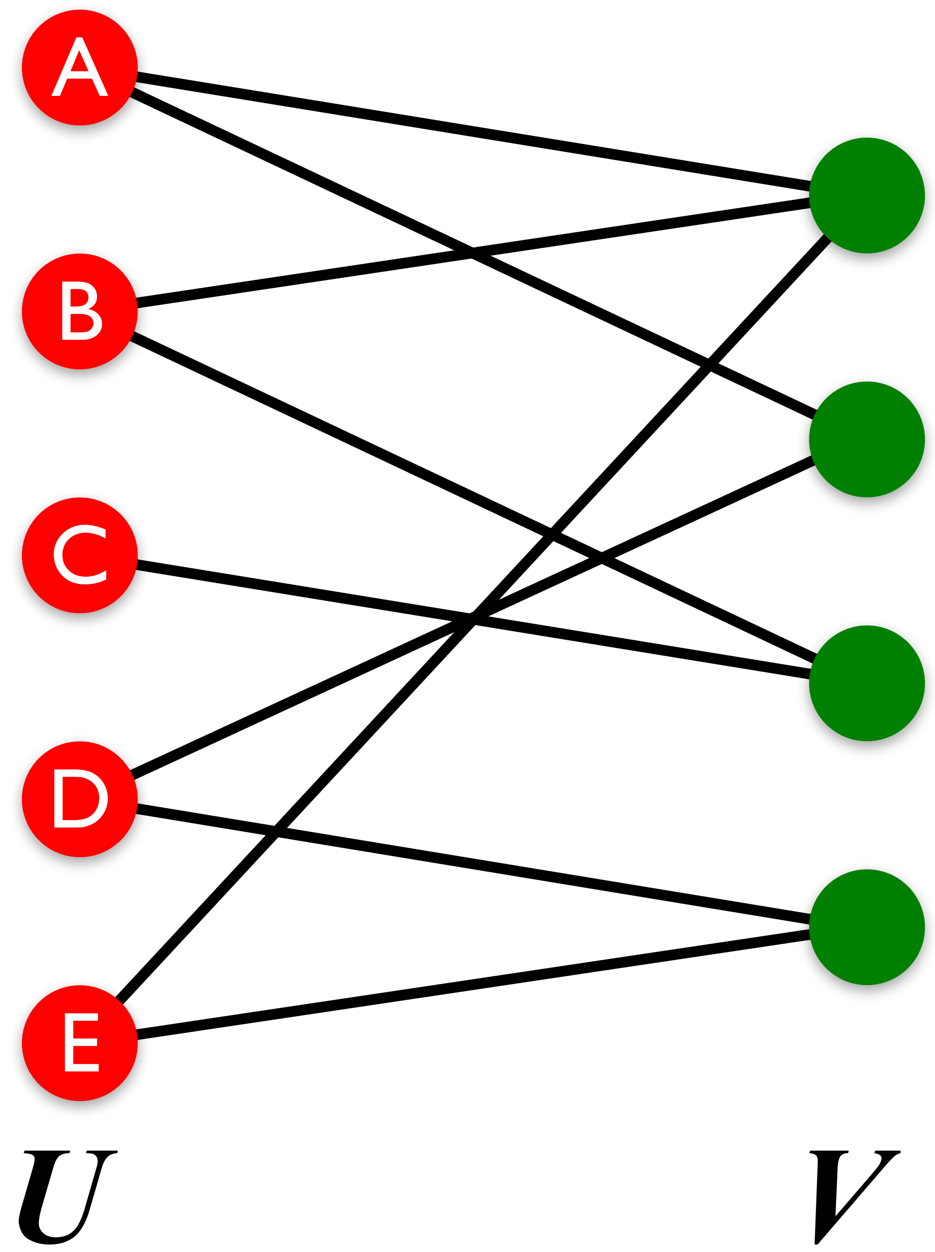
$$A_{ii} = 0 \qquad A_{ij} = A_{ji}$$

$$E = \frac{1}{2} \sum_{i,j=1}^{N} nonzero(A_{ij}) \qquad \bar{k} = \frac{2E}{N}$$

Examples: Communication, Collaboration

Bipartite Graph

Bipartite graph is a graph whose nodes can be divided into two disjoint sets U and V such that every link connects a node in U to one in V; that is, U and V are **independent sets**



Examples:

–Authors-to-papers (they authored)

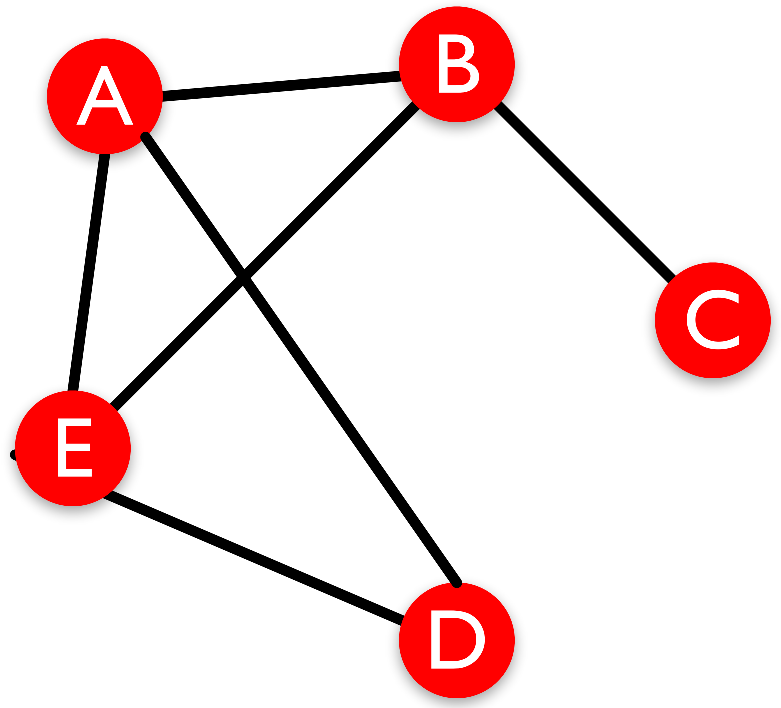
–Actors-to-Movies (they appeared in)

–Users-to-Movies (they rated)

"Folded" networks:

–Author collaboration networks

–Movie co-rating networks



Folded version of the graph above

Networks are Sparse Graphs

Most real-world networks are sparse

$E \ll E_{max}$ (or $\bar{k} \ll N-1$)

WWW (Stanford-Berkeley):	N=319,717	$\langle k \rangle$ =9.65
Social networks (LinkedIn):	N=6,946,668	$\langle k \rangle$ =8.87
Communication (MSN IM):	N=242,720,596	$\langle k \rangle$ =11.1
Coauthorships (DBLP):	N=317,080	$\langle k \rangle$ =6.62
Internet (AS-Skitter):	N=1,719,037	$\langle k \rangle$ =14.91
Roads (California):	N=1,957,027	$\langle k \rangle$ =2.82
Proteins (S. Cerevisiae):	N=1,870	$\langle k \rangle$ =2.39

(Source: *Leskovec et al., Internet Mathematics, 2009*)

Consequence: Adjacency matrix is filled with zeros!

(Density of the matrix (E/N^2): WWW=1.51×10^{-5}, MSN IM = 2.27×10^{-8})

Network Representations

WWW ➢

Facebook friendships ➢

Citation networks ➢

Collaboration networks ➢

Mobile phone calls ➢

Protein Interactions ➢

Network Representations

WWW ➢ directed multigraph with self-edges

Facebook friendships ➢ undirected, unweighted

Citation networks ➢ unweighted, directed, acyclic

Collaboration networks ➢ undirected multigraph or weighted graph

Mobile phone calls ➢ directed, (weighted?) multigraph

Protein Interactions ➢ undirected, unweighted with self-interactions

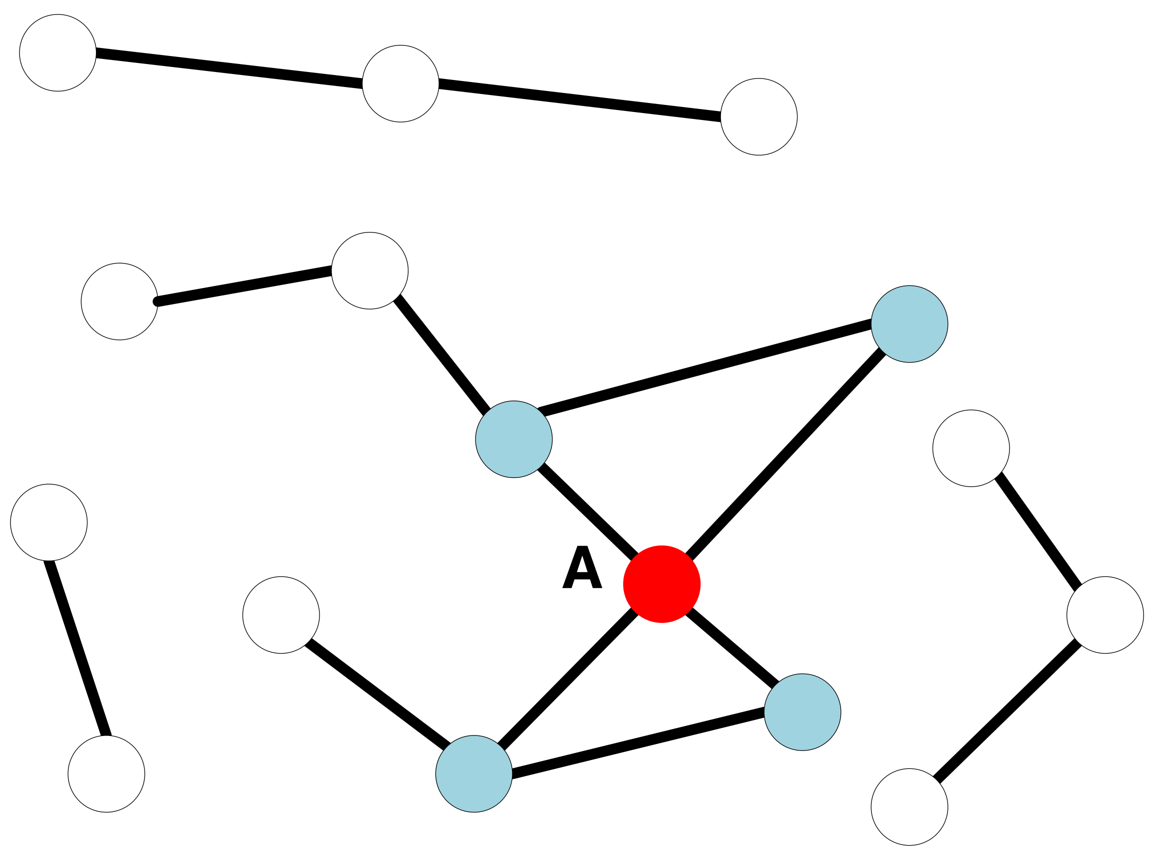
Network Properties:
How to Characterize/Measure a Network?

How do we measure properties in the graph representation of a network?

Focus on connectivity and distance

Connectivity: Node Degrees

Undirected

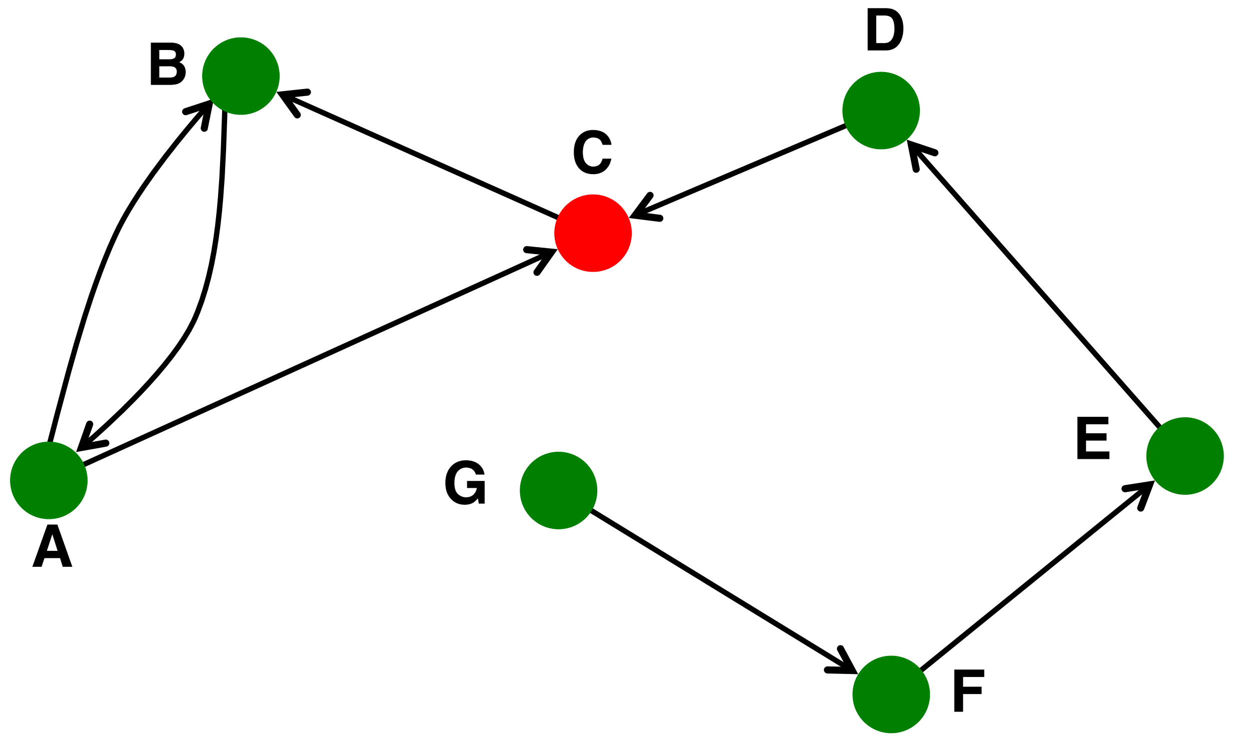


Node degree, k_i: the number of edges adjacent to node *i*

e.g. $k_A = 4$

Avg. degree: $$\bar{k} = \langle k \rangle = \frac{1}{N}\sum_{i=1}^{N} k_i = \frac{2E}{N}$$

Directed



In directed networks we define an in-degree and out-degree.
The (total) degree of a node is the sum of in- and out-degrees.

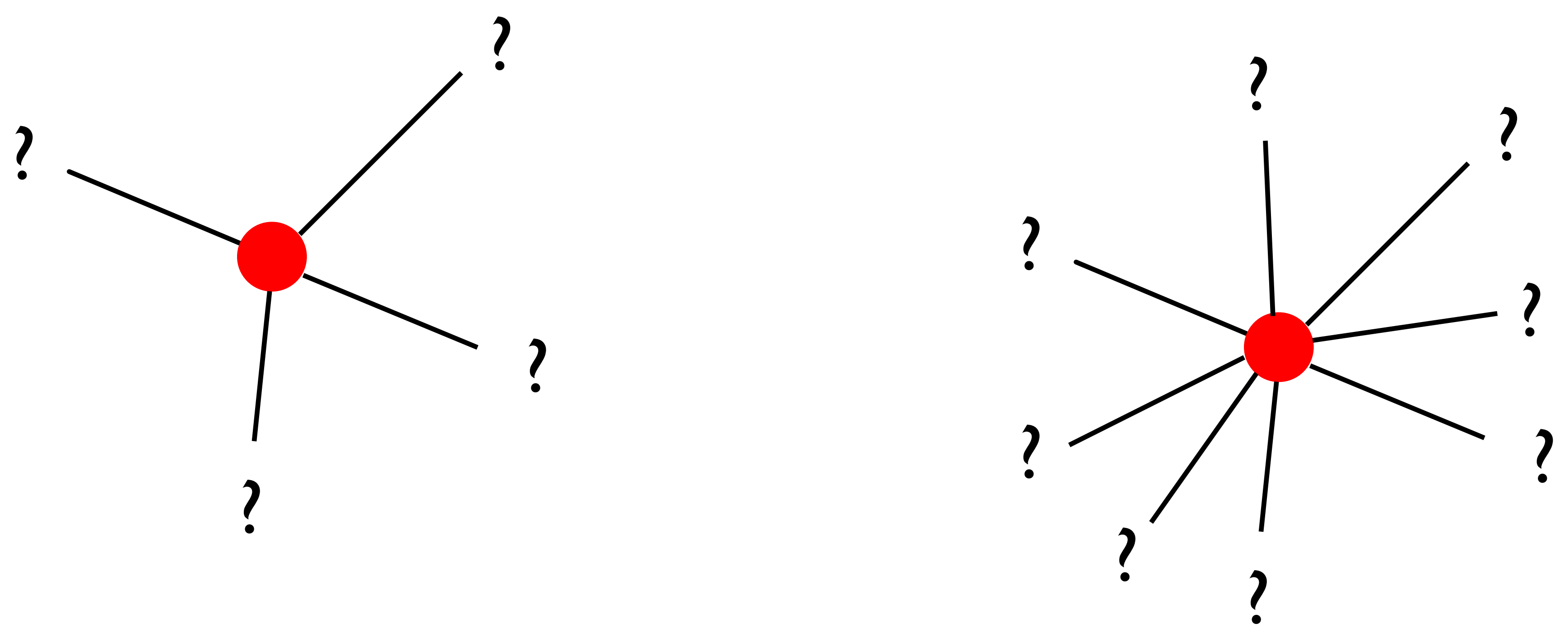
$$k_C^{in} = 2 \quad k_C^{out} = 1 \quad k_C = 3$$

Source: Node with $k^{in} = 0$
Sink: Node with $k^{out} = 0$

$$\overline{k^{in}} = \overline{k^{out}}$$

Connectivity: How Connected Are Nodes?

How many neighbours do nodes tend to have in your graph?



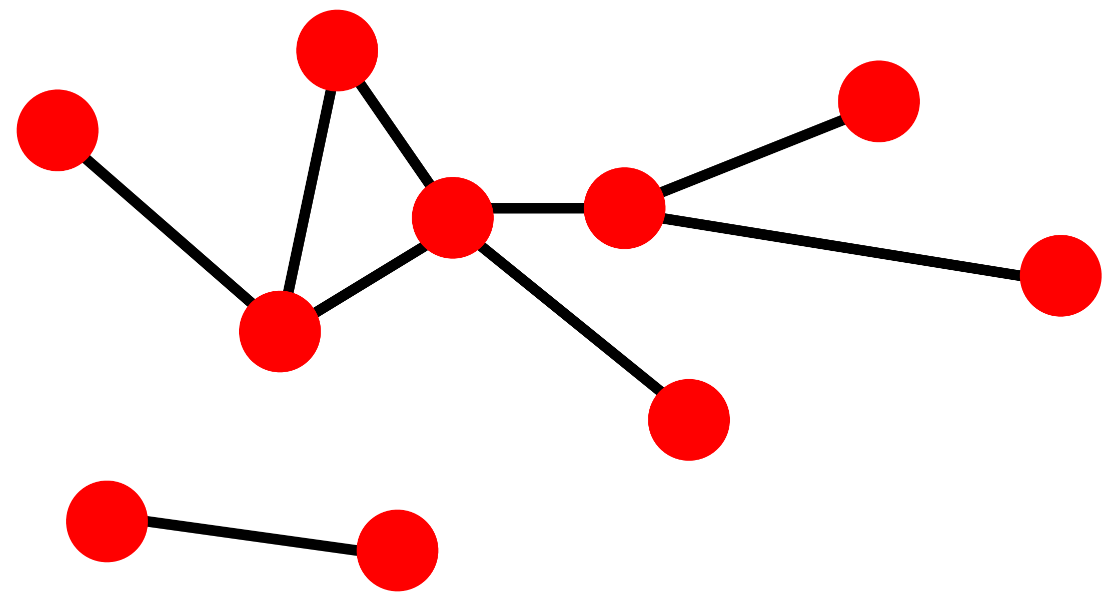
Connectivity: Degree Distribution

Degree distribution *P(k)*: Probability that a randomly chosen node has degree *k*

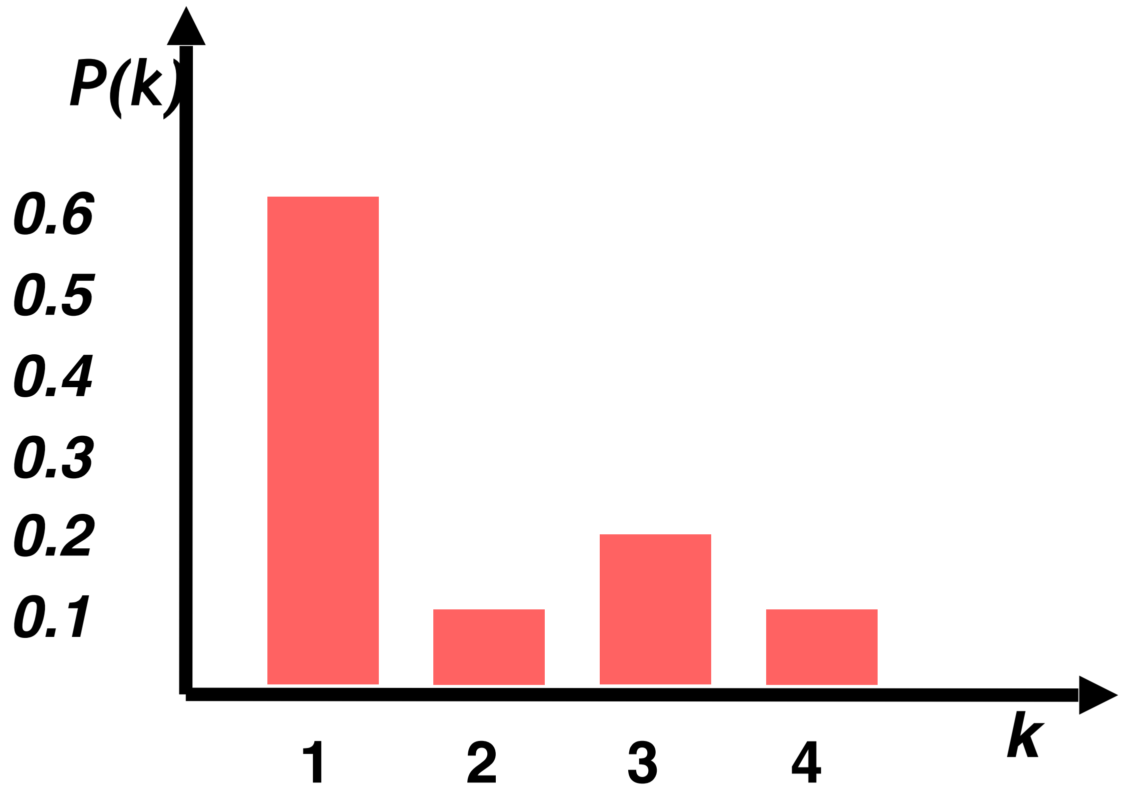
N_k = # nodes with degree *k*

Normalized histogram:

$P(k) = N_k / N$ → plot



→



Connectivity: Local Clustering

Are the nodes “clustered" in the graph? Do nodes with common neighbours tend to know each other?

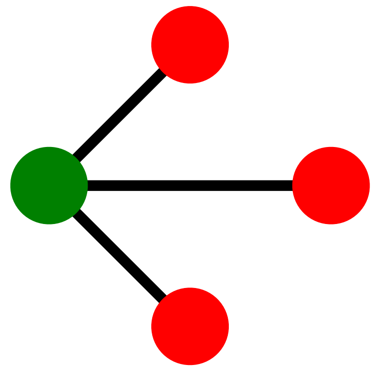
Connectivity: Clustering Coefficient

What's the probability that a random pair of your friends are connected?

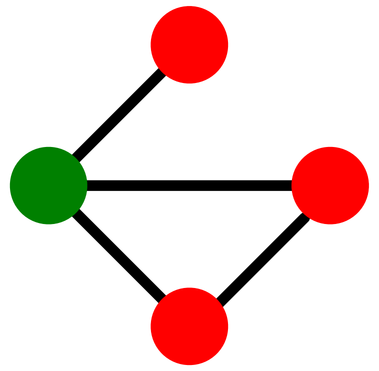
$$C_i \in [0, 1]$$

$$C_i = \frac{e_i}{\binom{k_i}{2}} = \frac{e_i}{k_i(k_i - 1)/2} = \frac{2e_i}{k_i(k_i - 1)}$$

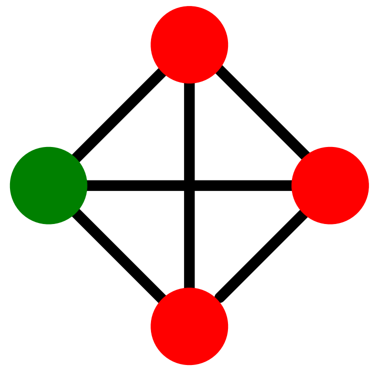
where e_i is the number of edges between the neighbours of node i and k_i is the degree of node i



$C_i=?$



$C_i=?$



$C_i=?$

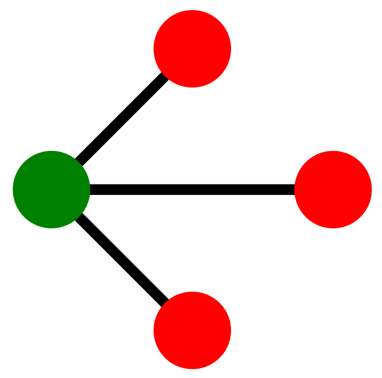
Connectivity: Clustering Coefficient

What's the probability that a random pair of your friends are connected?

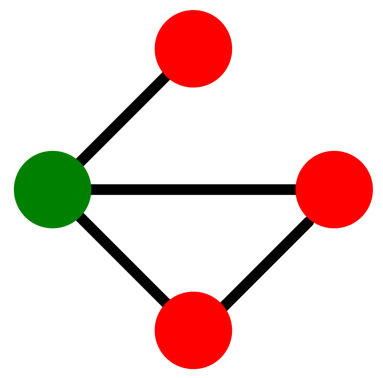
$$C_i \in [0, 1]$$

$$C_i = \frac{e_i}{\binom{k_i}{2}} = \frac{e_i}{k_i(k_i-1)/2} = \frac{2e_i}{k_i(k_i-1)}$$

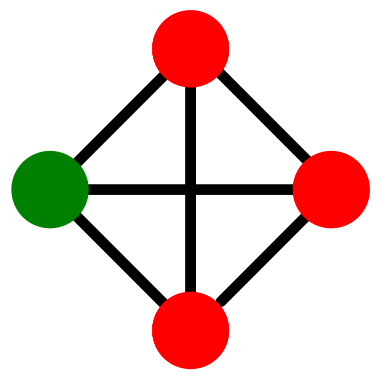
where e_i is the number of edges between the neighbors of node I and k_i is the degree of node I



$C_i=0$



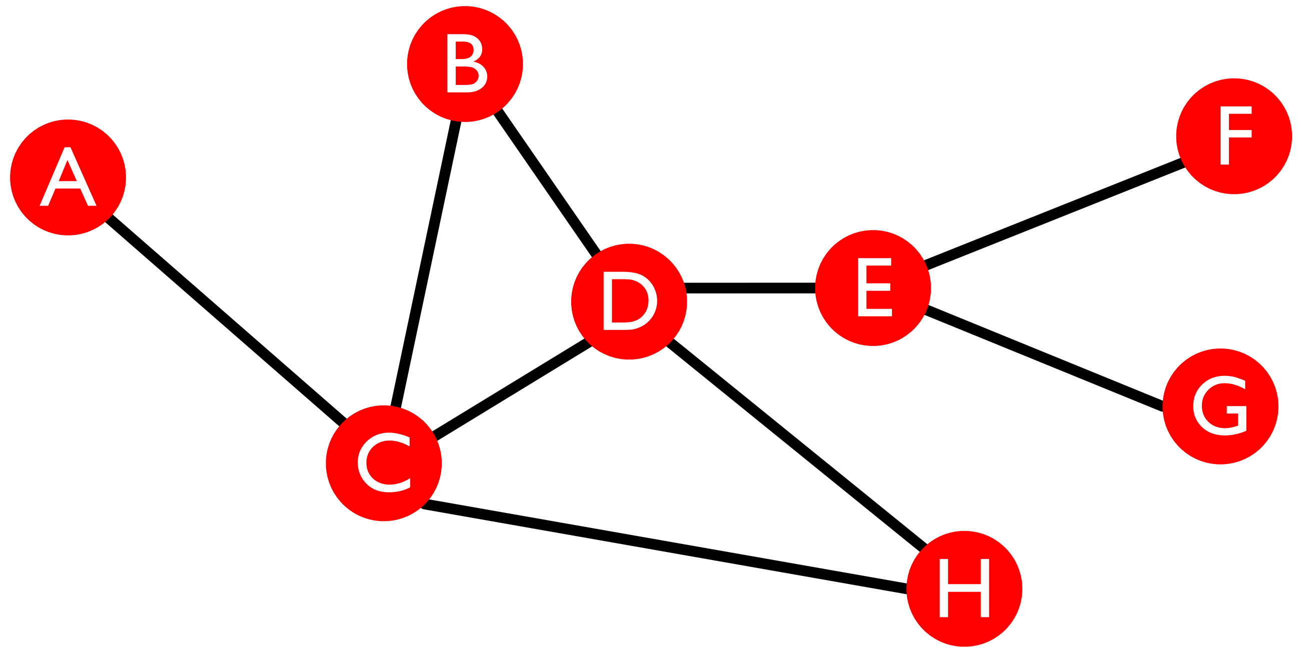
$C_i=1/3$



$C_i=1$

Average clustering coefficient: $C = \frac{1}{N}\sum_{i}^{N} C_i$

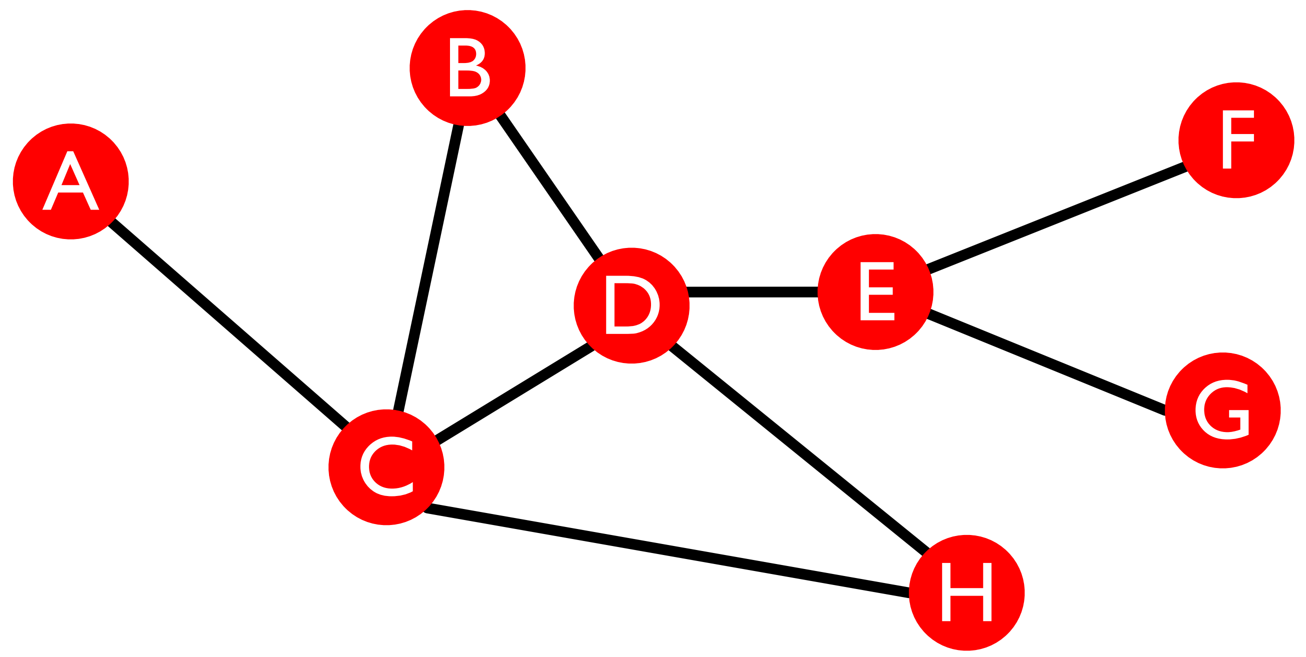
Connectivity: Clustering Coefficient



k_B=?, e_B=?, C_B=? = ?

k_D=?, e_D=?, C_D=? = ?

Connectivity: Clustering Coefficient



$k_B=2, \; e_B=1, \; C_B=2/2 = 1$

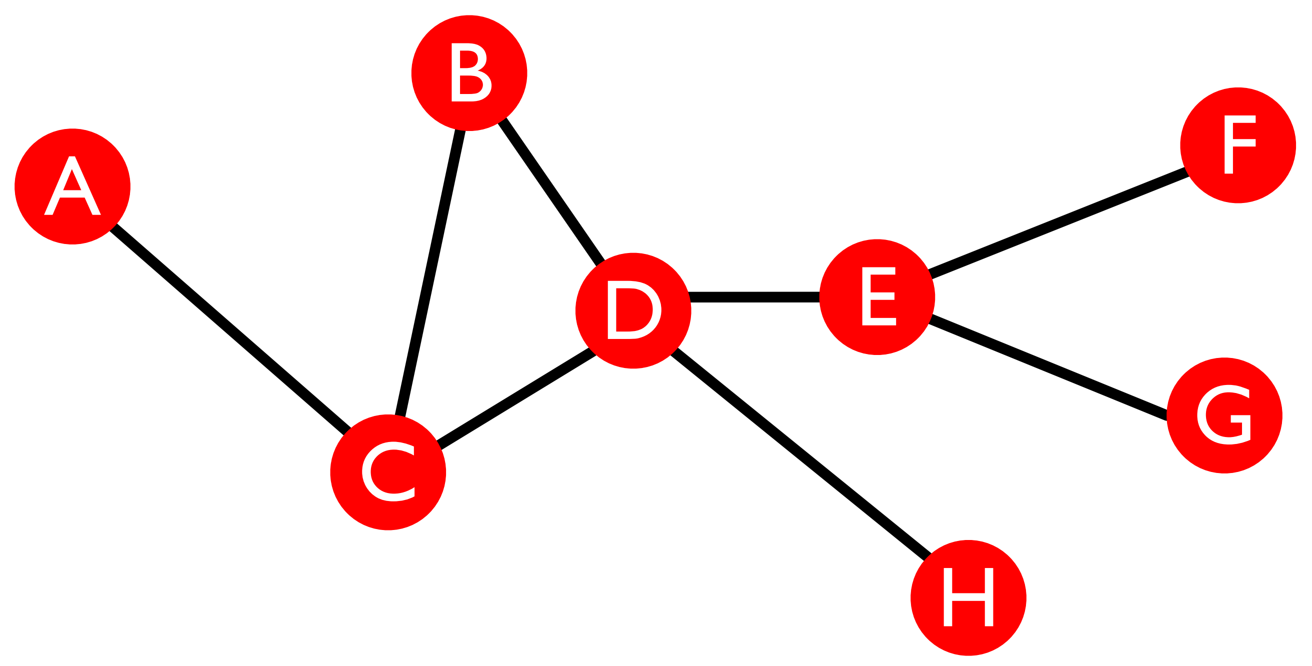
$k_D=4, \; e_D=2, \; C_D=(2*2)/(4*3) = 4/12 = 1/3$

Distance: Paths in a Graph

- A ***path*** is a sequence of nodes in which each node is linked to the next one

$$P_n = \{i_0, i_1, i_2, \ldots, i_n\} \qquad P_n = \{(i_0, i_1), (i_1, i_2), (i_2, i_3), \ldots, (i_{n-1}, i_n)\}$$

- Path can intersect itself and pass through the same edge multiple times
  - E.g.: ACBDCDEG
  - In a directed graph a path can only follow the direction of the "arrow"



Distance: Number of Paths

Number of paths between nodes *u* and *v*:

Length $h=1$: If there is a link between u and v, $A_{uv}=1$ else $A_{uv}=0$

Length $h=2$: If there is a path of length two between *u* and *v* then $A_{uk}A_{kv}=1$ else $A_{uk}A_{kv}=0$

$$H_{uv}^{(2)} = \sum_{k=1}^{N} A_{uk}A_{kv} = [A^2]_{uv}$$

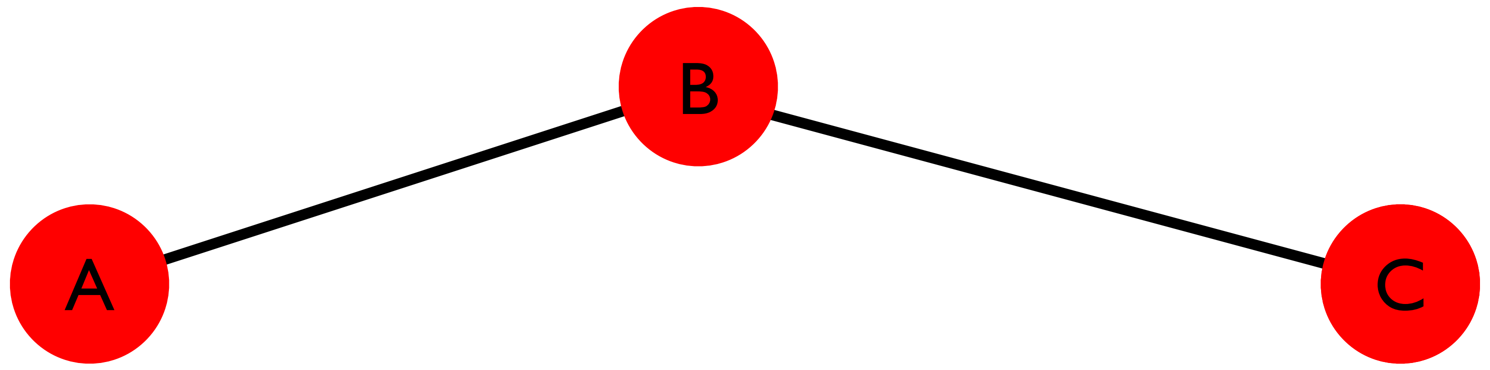
Length h: If there is a path of length *h* between *u* and *v* then $A_{uk} \ldots A_{kv}=1$ else $A_{uk} \ldots A_{kv}=0$

So, the no. of paths of length *h* between *u* and *v* is

$$H_{uv}^{(h)} = [A^h]_{uv}$$

(holds for both directed and undirected graphs)

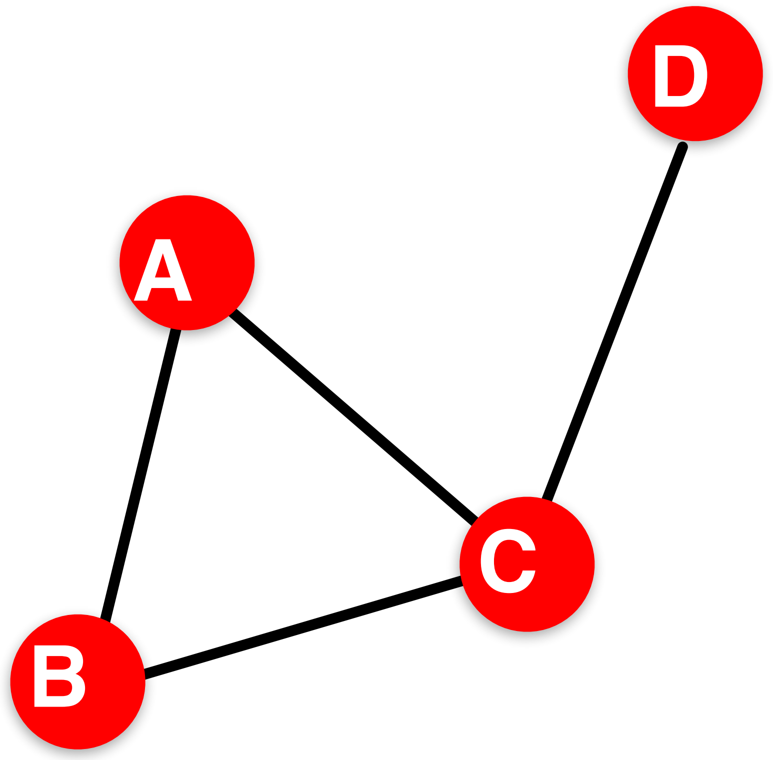
Distance: Number of Paths



$$H^{(1)} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

$$H^{(2)} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

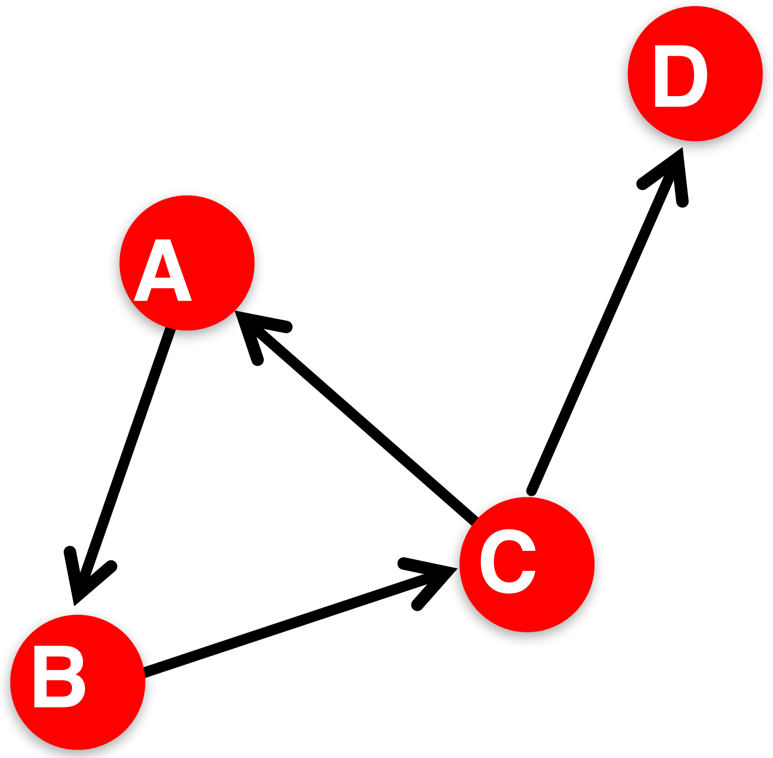
Distance: definition



$h_{B,D} = 2$

Distance (shortest path, geodesic) between a pair of nodes is defined as the number of edges along the shortest path connecting the nodes

*If the two nodes are disconnected, the distance is usually defined as infinite



$h_{B,C} = 1$, $h_{C,B} = 2$

In **directed graphs** paths need to follow the direction of the arrows

Consequence: Distance is **not symmetric**: $h_{A,C} \neq h_{C,A}$

Distance: Graph-level measures

- **Diameter:** the maximum (shortest path) distance between any pair of nodes in a graph

- **Average path length** for a connected graph (component) or a strongly connected (component of a) directed graph

$$\bar{h} = \frac{1}{2E_{max}} \sum_{i, j \neq i} h_{ij}$$

where h_{ij} is the distance from node i to node j,
And Emax is the maximum number of edges (=n*(n-1)/2)

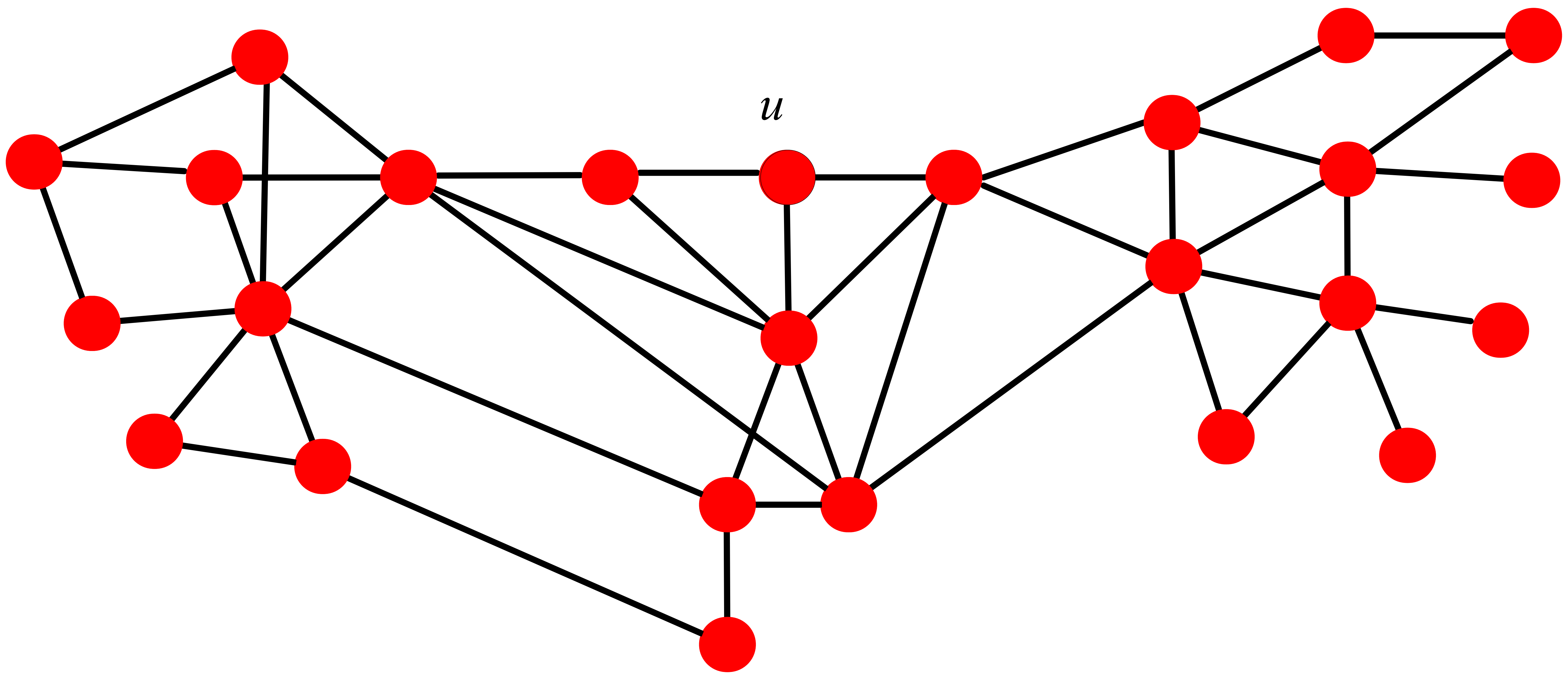
  - Many times we compute the average only over the connected pairs of nodes (that is, we ignore "infinite" length paths)

Key Network Properties

Degree distribution:	$P(k)$
Clustering coefficients:	C
Path lengths:	L
Diameter:	D

Finding Shortest Paths

- **Breadth First Search:**
  - Start with node u, mark it to be at distance $h_u(u)=0$, add u to the queue
  - While the queue not empty:
    - Take node v off the queue, put its unmarked neighbors w into the queue and mark $h_u(w)=h_u(v)+1$



Let’s measure these properties in a real network!

Key Network Properties

Degree distribution: $P(k)$

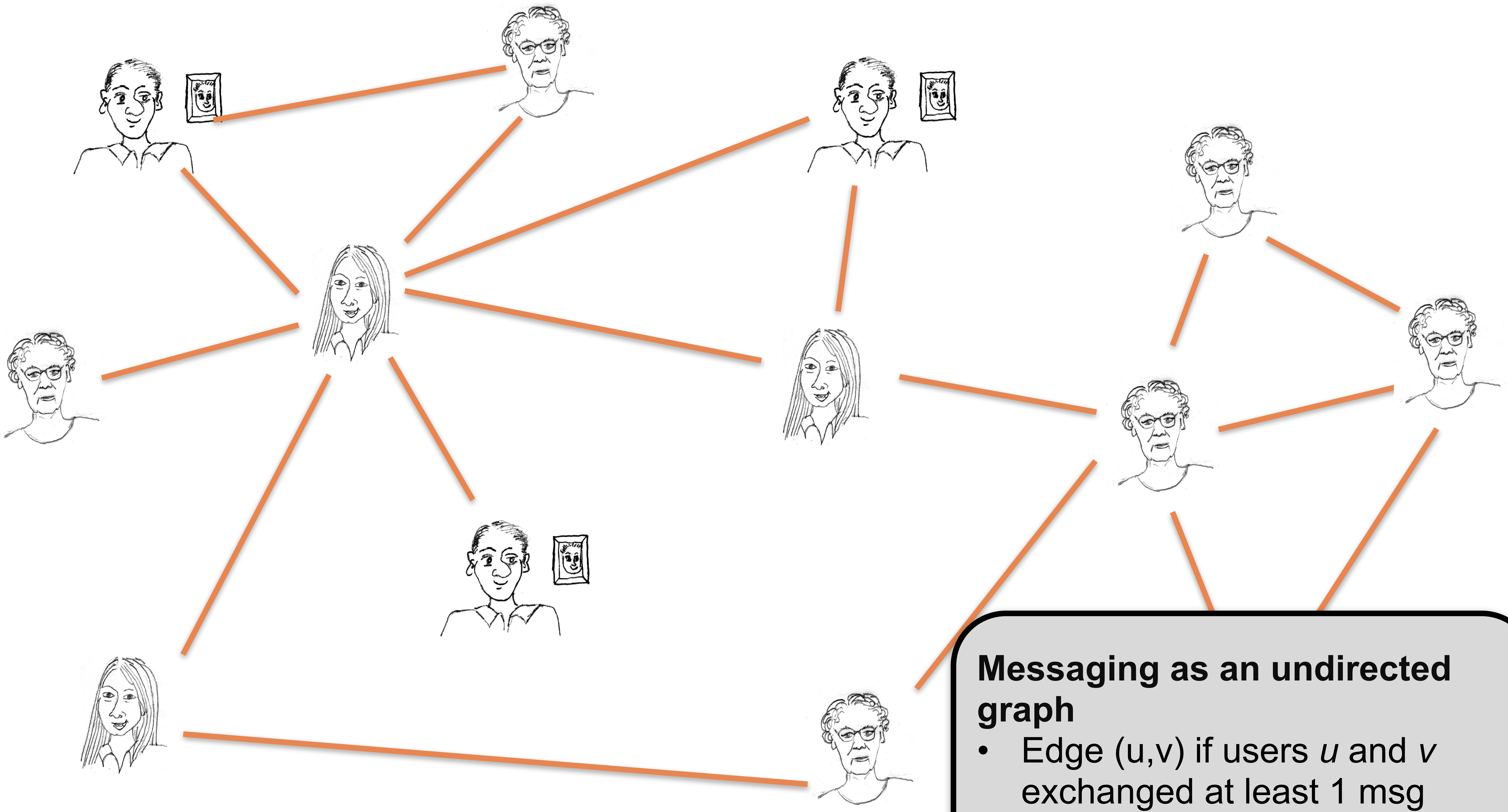
Clustering coefficient: C

(we'll look at distance later in the course)

MSN Messenger

- **MSN Messenger activity in June 2006:**
  - 245 million users logged in
  - 180 million users engaged in conversations
  - More than 30 billion conversations
  - More than 255 billion exchanged messages

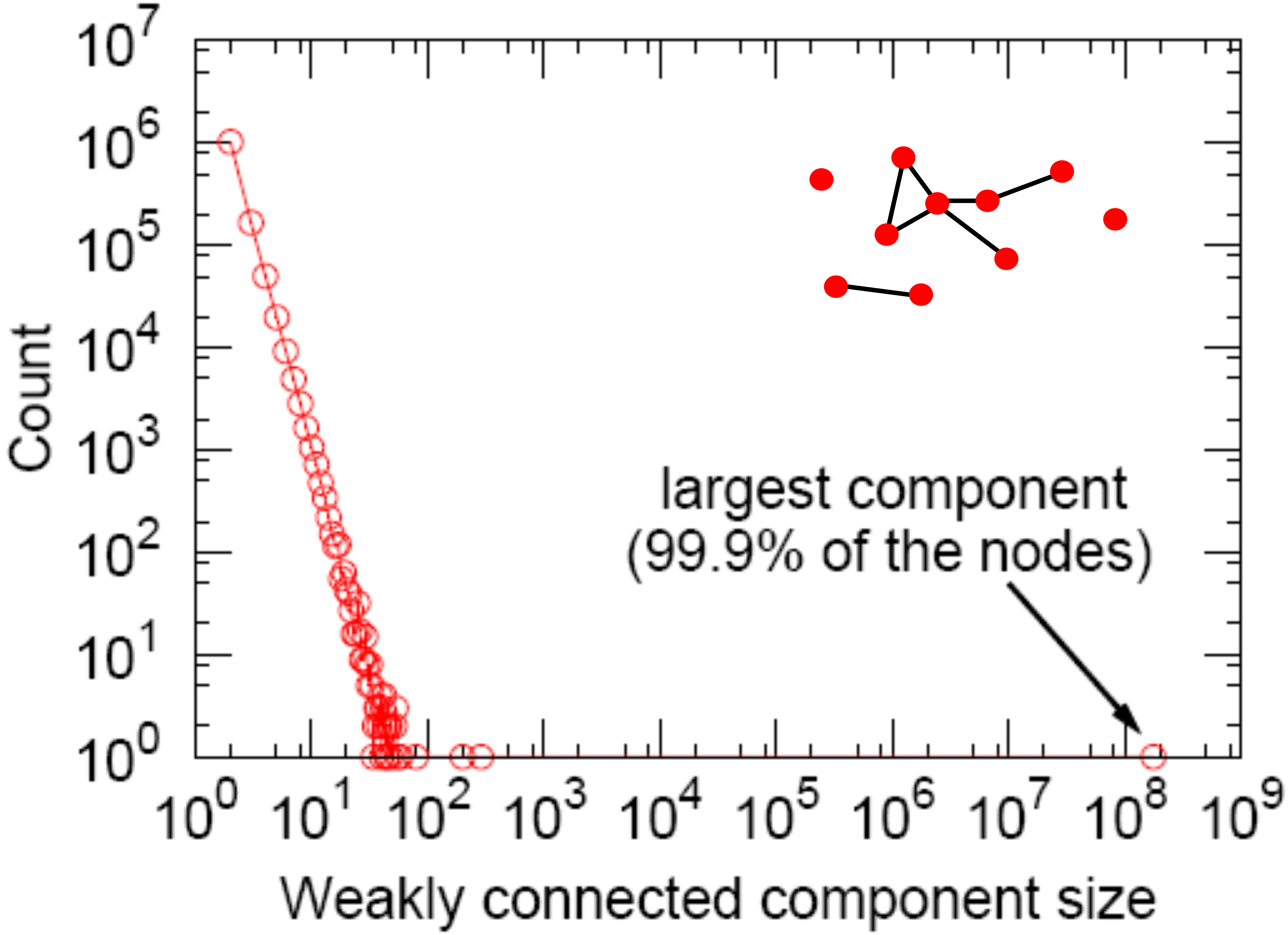
Messaging as a simple graph



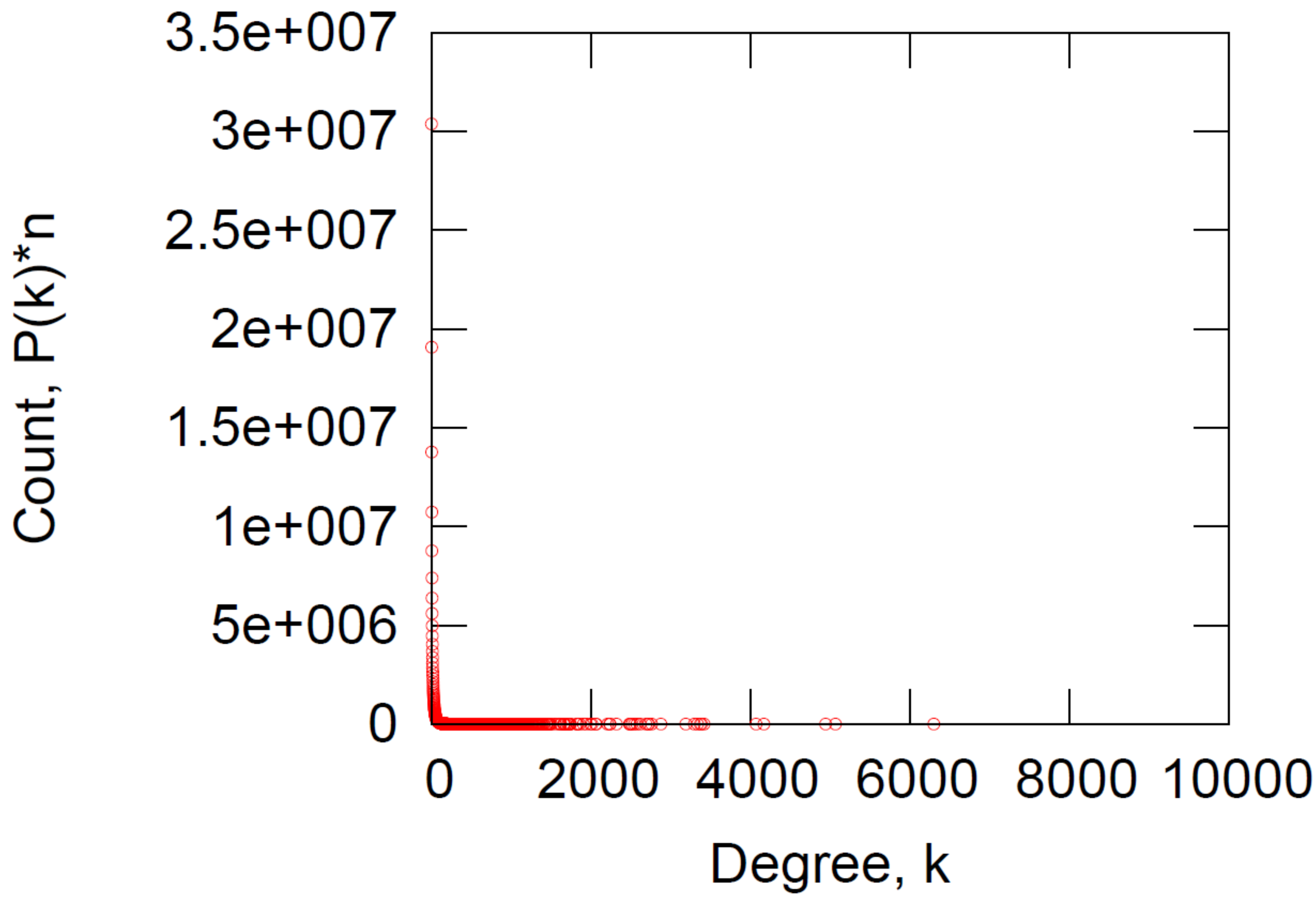
Messaging as an undirected graph

- Edge (u,v) if users *u* and *v* exchanged at least 1 msg
- N=180 million people
- E=1.3 billion edges

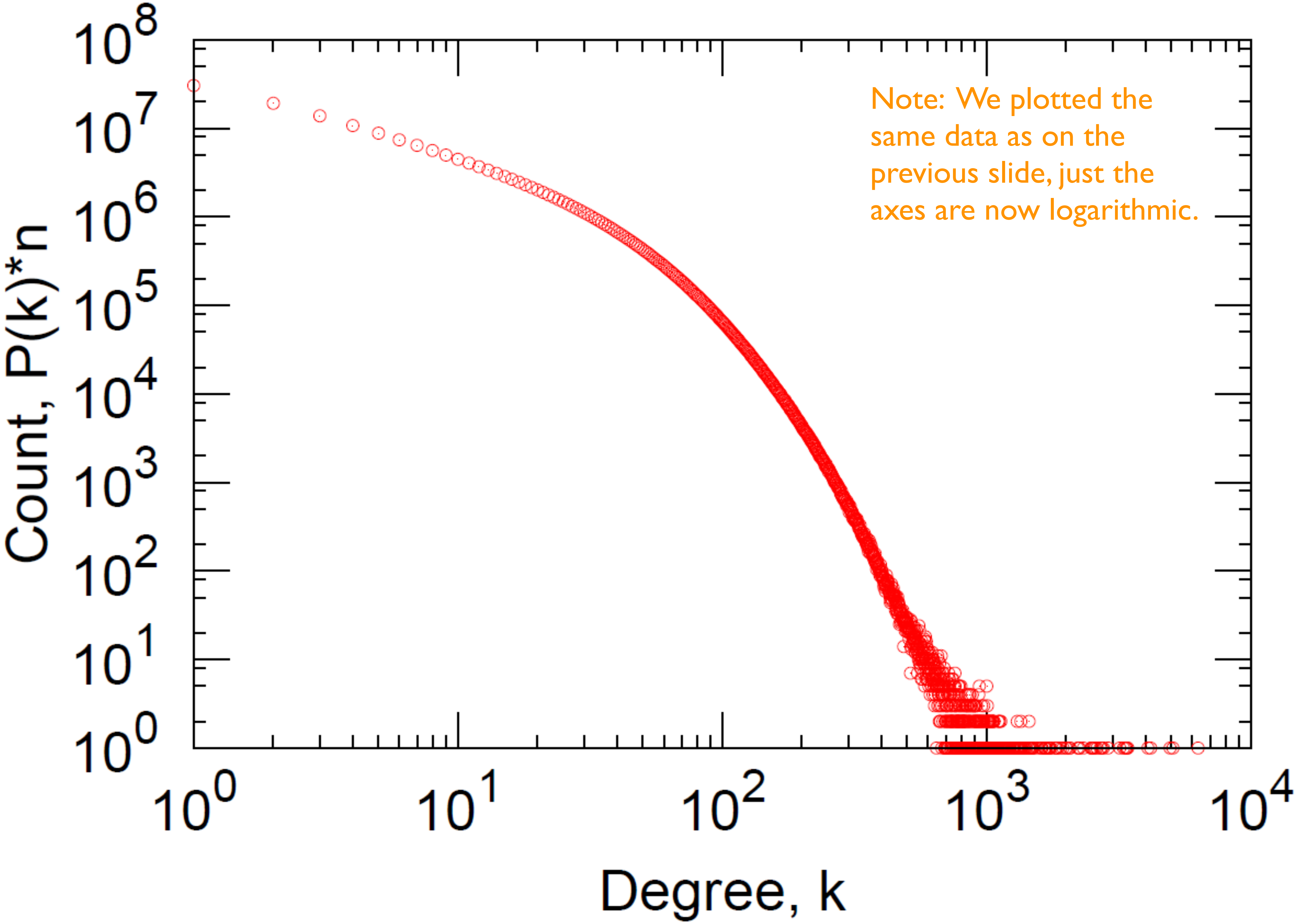
MSN Network: Connectivity



MSN: Degree Distribution



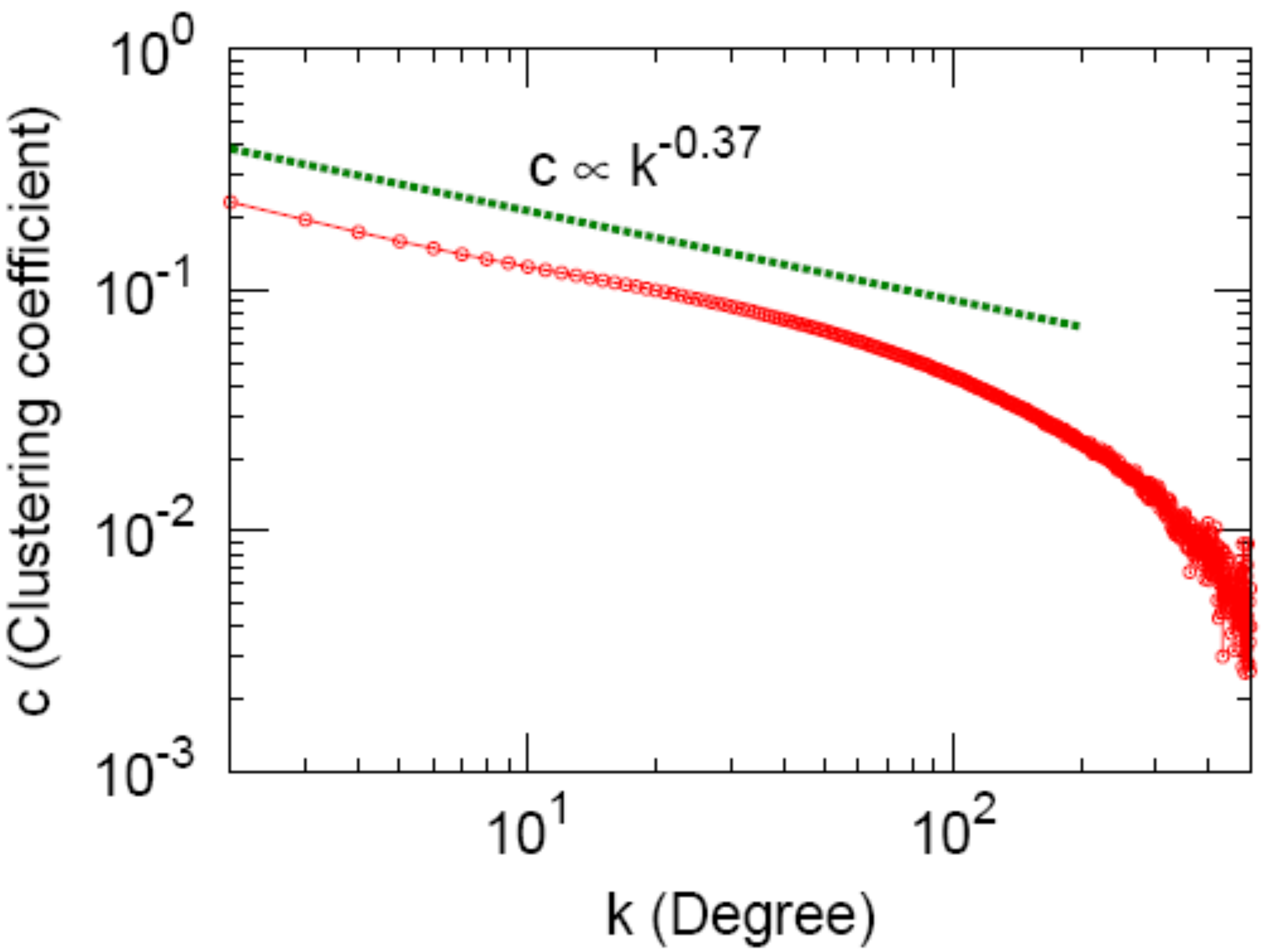
MSN: Log-Log Degree Distribution



MSN: Clustering

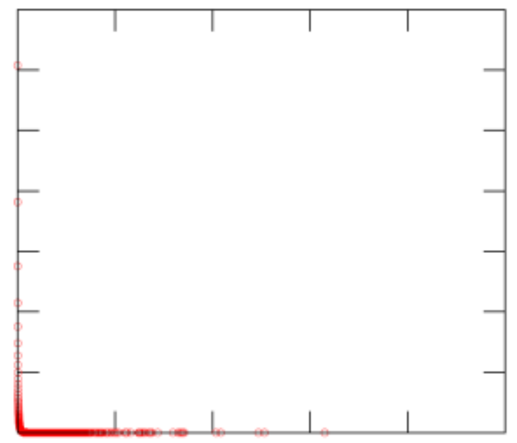
C_k: average C_i of nodes i of degree k:

$$C_k = \frac{1}{N_k} \sum_{i:k_i=k} C_i$$



Avg. clustering coefficient of the MSN graph:
C = 0.1140

MSN: Recap

Degree distribution: 

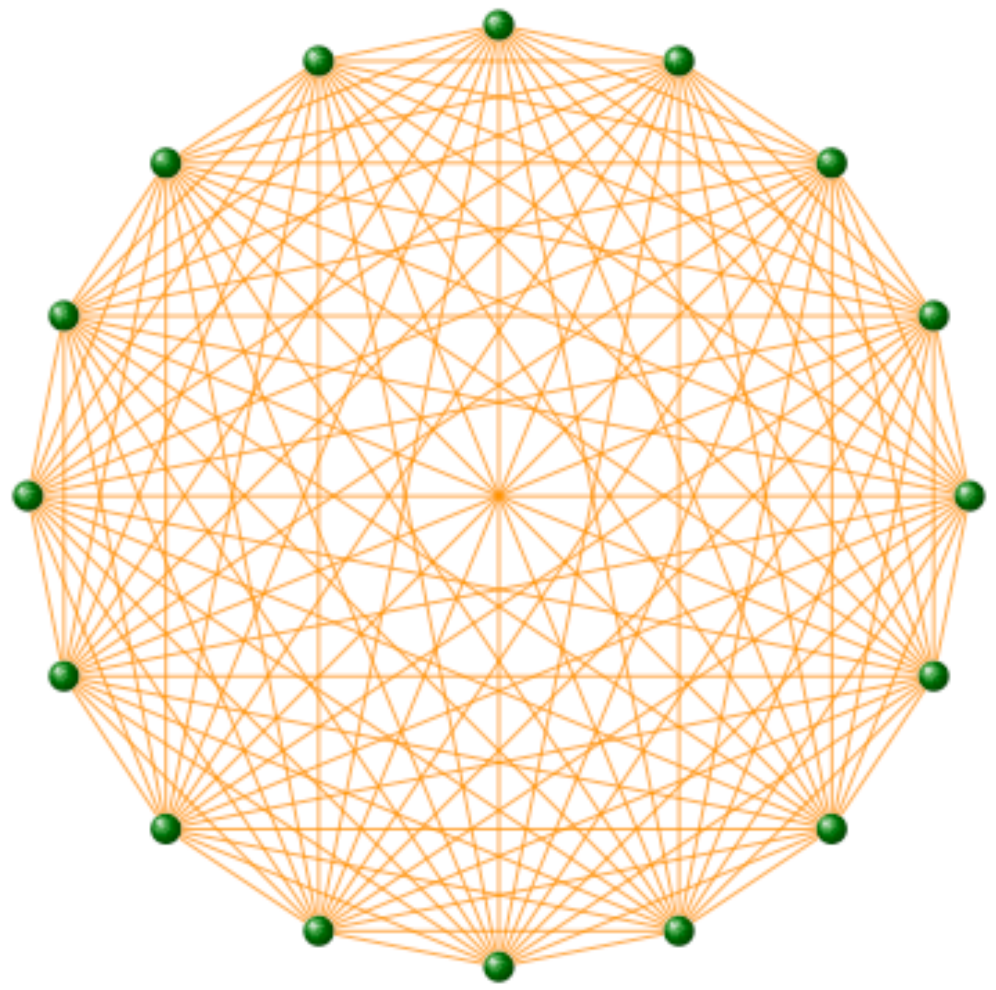
Clustering coefficient: *0.11*

Is this behaviour expected or surprising?

Let’s compare with simple models

Complete Graph

The maximum number of edges in an undirected graph on *N* nodes is



$$E_{\max} = \binom{N}{2} = \frac{N(N-1)}{2}$$

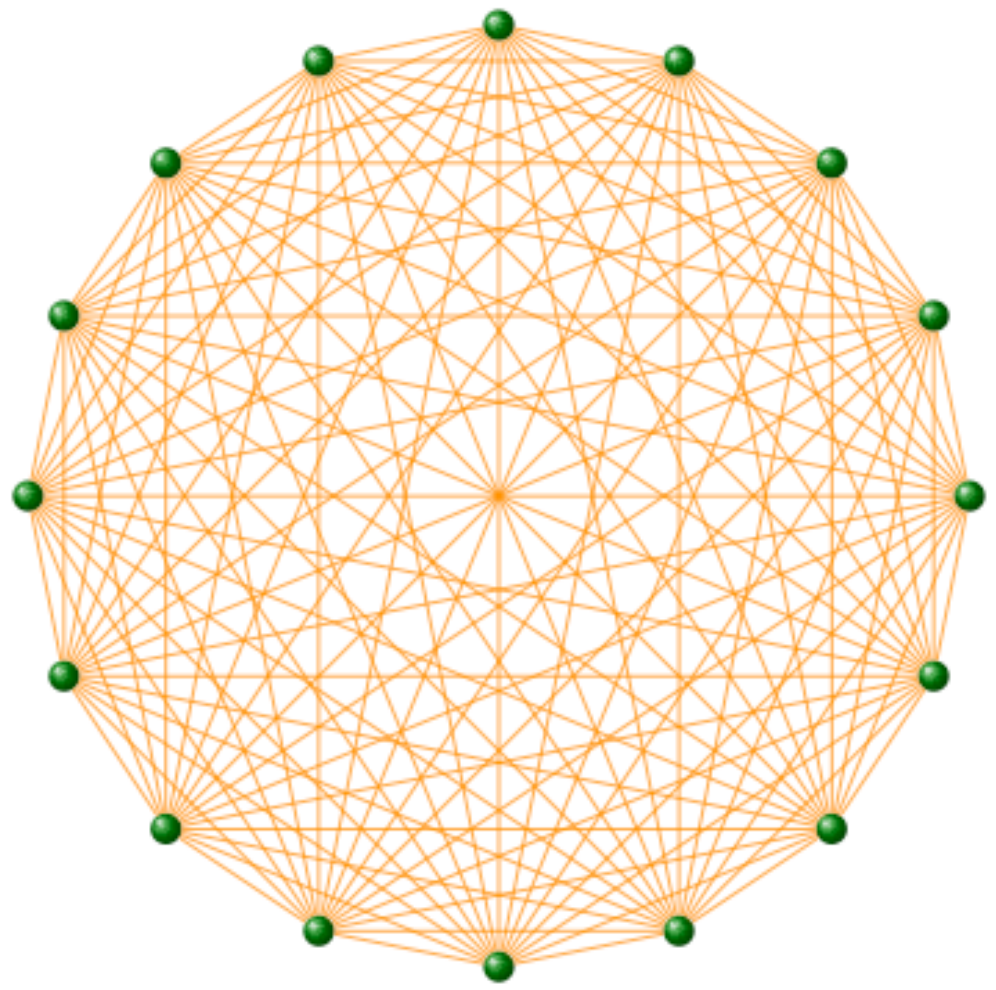
An undirected graph with the number of edges *E* = E_{max} is called a complete graph

Every node has degree ?

Every node has clustering coefficient ?

Complete Graph

The maximum number of edges in an undirected graph on N nodes is

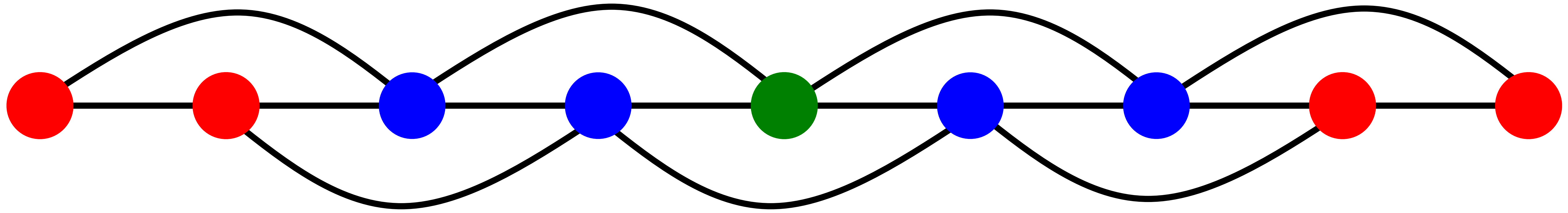


$$E_{\max} = \binom{N}{2} = \frac{N(N-1)}{2}$$

An undirected graph with the number of edges $E = E_{max}$ is called a complete graph

Every node has degree n-1
Every node has clustering coefficient 1

Is MSN Network like a "chain"?



Degree distribution: $P(k) = \delta(k-4)$

Clustering: $C = \frac{1}{N}\left(\frac{1}{2}(N-4) + 2 + 2\frac{2}{3}\right) \to \frac{1}{2}$ as $N \to \infty$

Constant degree,
high average clustering coefficient

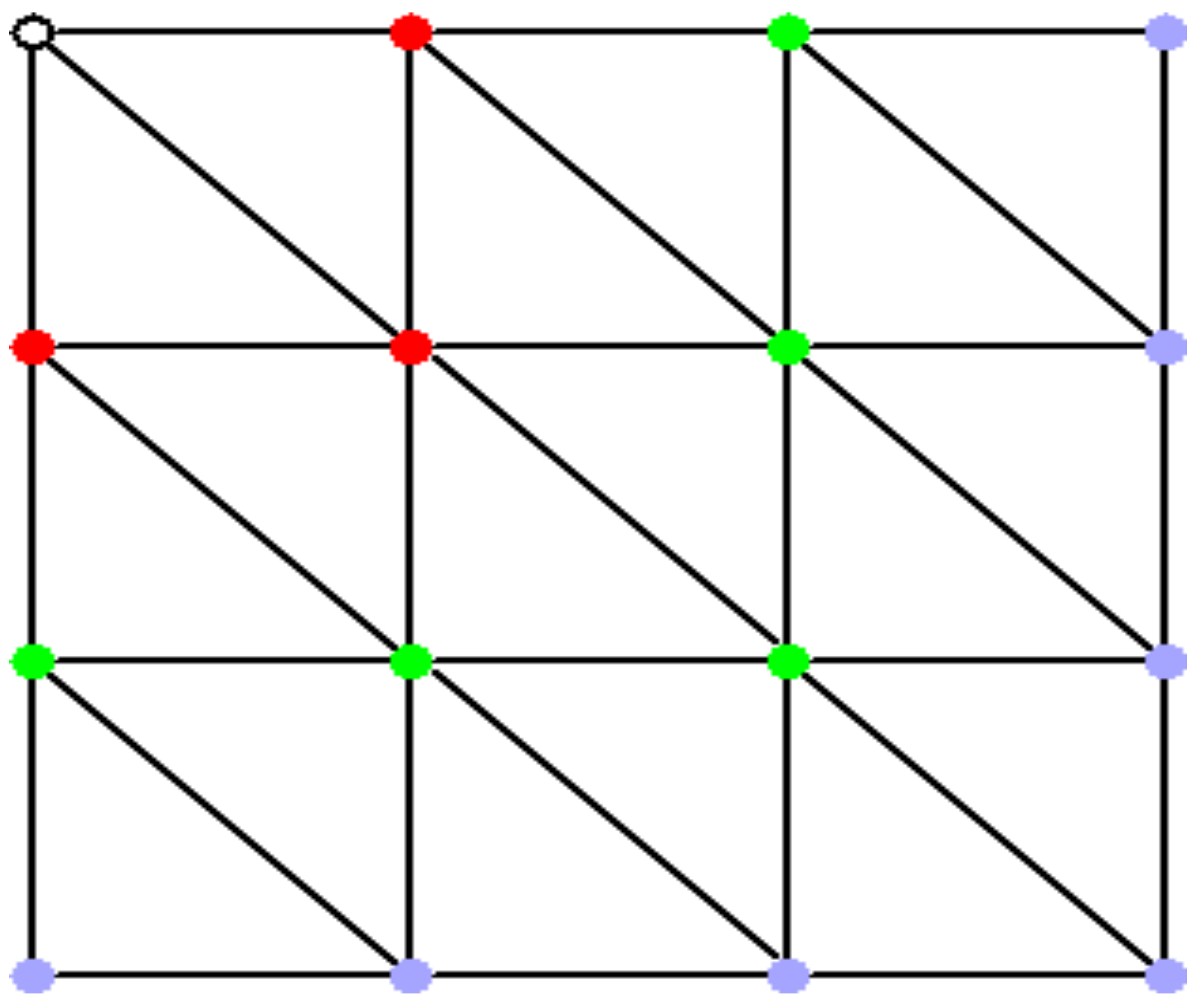
Note about calculations:
We are interested in quantities as graphs get large ($N \to \infty$)

We will use big-O:
$f(x) = O(g(x))$ as $x \to \infty$
if $f(x) < g(x)*c$ for all $x > x_0$ and some constant c.

Is MSN Network like a "grid"?

- $P(k) = \delta(k-6)$
  - $k = 6$ for each inside node
- $C = 6/15$ for inside nodes
- **Path length:**

$$h_{max} = O(\sqrt{N})$$



- **In general, for lattices:**
  - Average path-length is $\bar{h} \approx N^{1/D}$ (D… lattice dimensionality)
  - Constant degree, constant clustering coefficient

MSN Network is neither a chain nor a grid

Need a model to compare:

Erdös-Renyi
Random Graph Model

Simplest Model of Graphs

Erdös-Renyi Random Graphs [Erdös-Renyi, '60]

$G_{n,p}$: undirected graph on n nodes and each edge (u,v) appears i.i.d. with probability p

Simplest random model you can think of

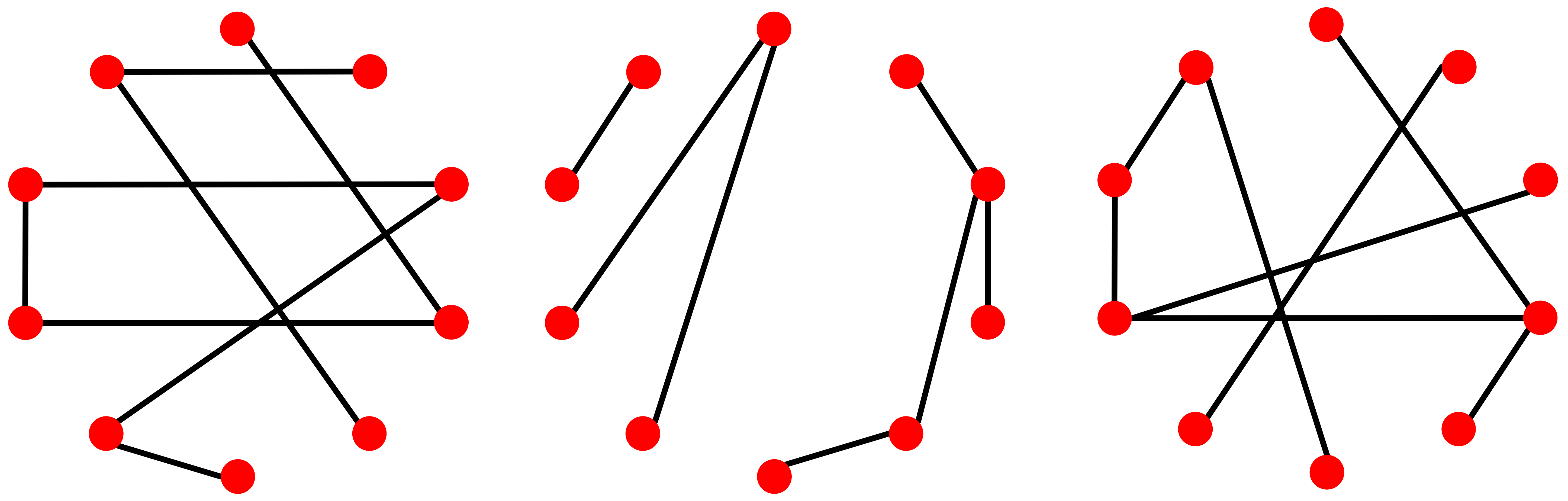
What kinds of networks does such a model produce?

Random Graph Model

n and *p* do not uniquely determine the graph!

The graph is a result of a random process

We can have many different realizations given the same *n* and *p*



n = 10
p = 1/6

Random Graph Model: Edges

How likely is a graph with *E* edges?

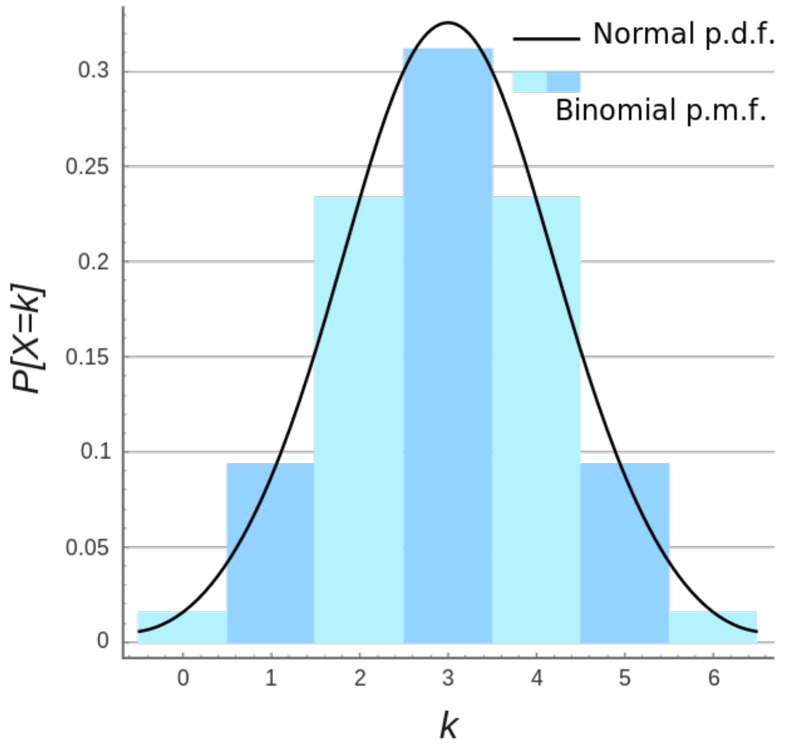
P(E): the probability that a given G_{np} generates a graph with exactly *E* edges:

$$P(E) = \binom{E_{\max}}{E} p^{E} (1-p)^{E_{\max}-E}$$

where $E_{max}=n(n-1)/2$ is the maximum possible number of edges in an undirected graph of *n* nodes

P(E) is exactly the Binomial distribution >>>

Number of successes in a sequence of E_{max} independent yes/no experiments



Node Degrees in a Random Graph

What is expected degree of a node?

Let X_v be a rnd. var. measuring the degree of node v

We want to know: $$E[X_v] = \sum_{j=0}^{n-1} j P(X_v = j)$$

An easier way:

Recall linearity of expectation

For any random variables $Y_1, Y_2, \dots, Y_k$

If $Y = Y_1 + Y_2 + \dots + Y_k$, then $E[Y] = E[Y_1 + Y_2 + \dots + Y_k] = \sum_i E[Y_i]$

Decompose X_v to $X_v = X_{v,1} + X_{v,2} + \dots + X_{v,n-1}$

where $X_{v,u}$ is a $\{0,1\}$-random variable which tells if edge (v,u) exists or not

$$E[X_v] = \sum_{u=1}^{n-1} E[X_{vu}] = (n-1) \cdot p$$

How to think about this?

- Prob. of node u linking to node v is p
- u can link (flips a coin) to all other $(n-1)$ nodes
- Thus, the expected degree of node u is: $p(n-1)$

Properties of G_{np}

Degree distribution: $P(k)$

Clustering coefficient: C

What are values of these properties for G_{np}?

Degree Distribution

Fact: **Degree distribution of G_{np} is Binomial.**

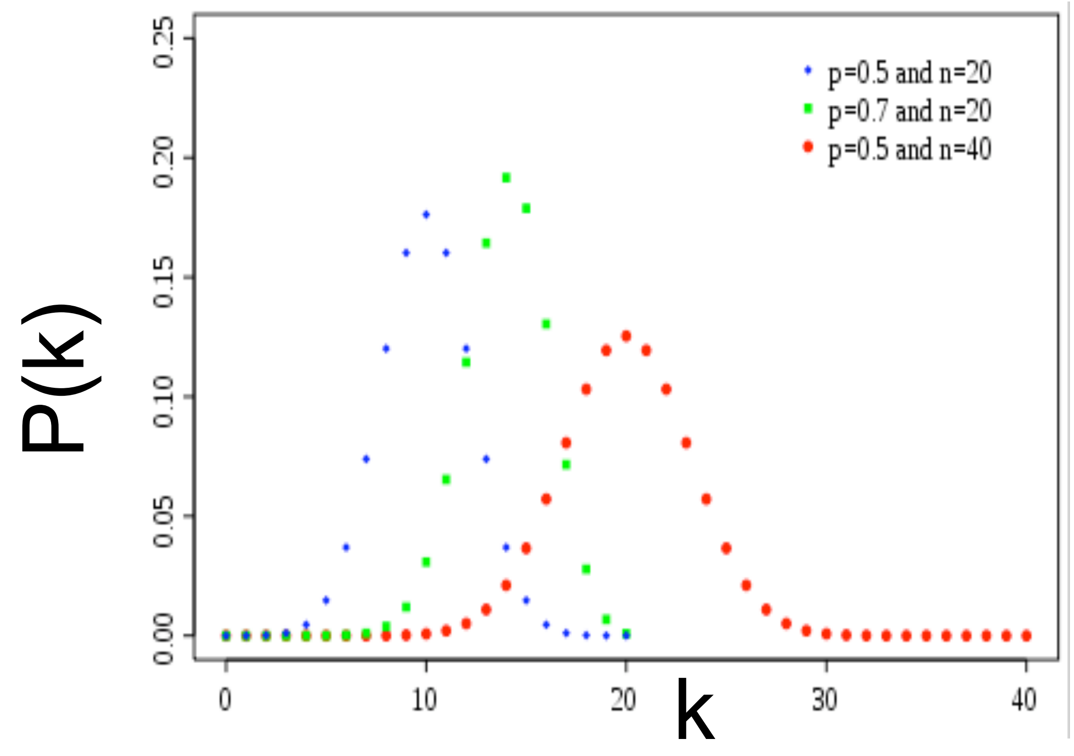
Let ***P(k)*** denote a fraction of nodes with degree ***k***:

$$P(k) = \binom{n-1}{k} p^k (1-p)^{n-1-k}$$

Select *k* nodes out of *n-1*

Probability of having *k* edges

Probability of missing the rest of the *n-1-k* edges



$$\bar{k} = p(n-1)$$

Clustering Coefficient of G_{np}

Remember: $$C_i = \frac{2e_i}{k_i(k_i - 1)}$$

Where e_i is the number of edges between i's neighbours

Edges in G_{np} appear i.i.d. with prob. p

So: $$e_i = p\frac{k_i(k_i - 1)}{2}$$

Each pair is connected with prob. p

Number of distinct pairs of neighbors of node i of degree k_i

Then: $$C = \frac{p \cdot k_i(k_i - 1)}{k_i(k_i - 1)} = p$$

And: $$C = p = \frac{\bar{k}}{n - 1}$$

Since $\bar{k} = p(n-1)$

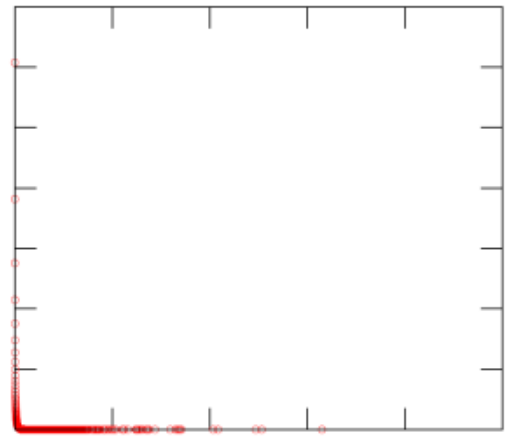
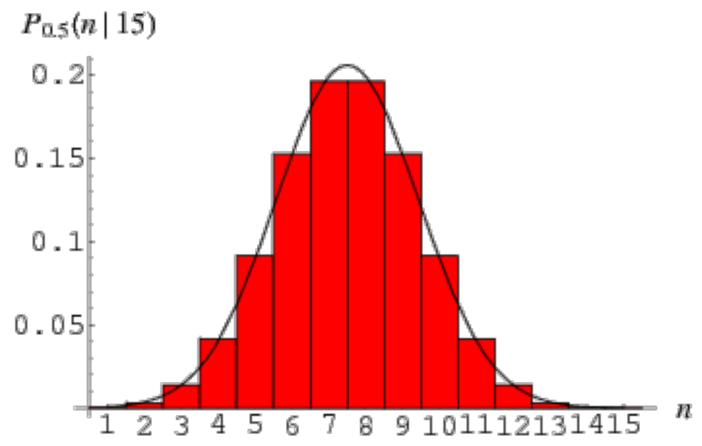
Clustering coefficient is very small (if you fix average degree, decreases with graph size)

Network Properties of G_{np}

Degree distribution: $$P(k) = \binom{n-1}{k} p^k (1-p)^{n-1-k}$$

Clustering coefficient: $$C = p = k / n$$

MSN vs. G_{np}

	MSN	G_{np}
Degree distribution:		
Clustering coefficient:	*0.11*	k / n ≈ $8 \cdot 10^{-8}$

Real Networks vs. G_{np}

Are real networks like random graphs?

Clustering Coefficient: ☹

Degree Distribution: ☹

Problems with the random networks model:

Degree distribution differs from that of real networks

No local structure – clustering coefficient is too low

Most important: Are real networks random?

NOPE!

Real Networks vs. G_{np}

If G_{np} is wrong, why did we spend time on it?

—It will help us calculate many quantities that can then be compared to the real data

—If the quantity you just calculated also shows up in G_{np}, it's probably not that interesting ("this also happens if you assume complete randomness")

—It is the reference model for the rest of the class.

So, while G_{np} is not realistic, it is extremely useful!

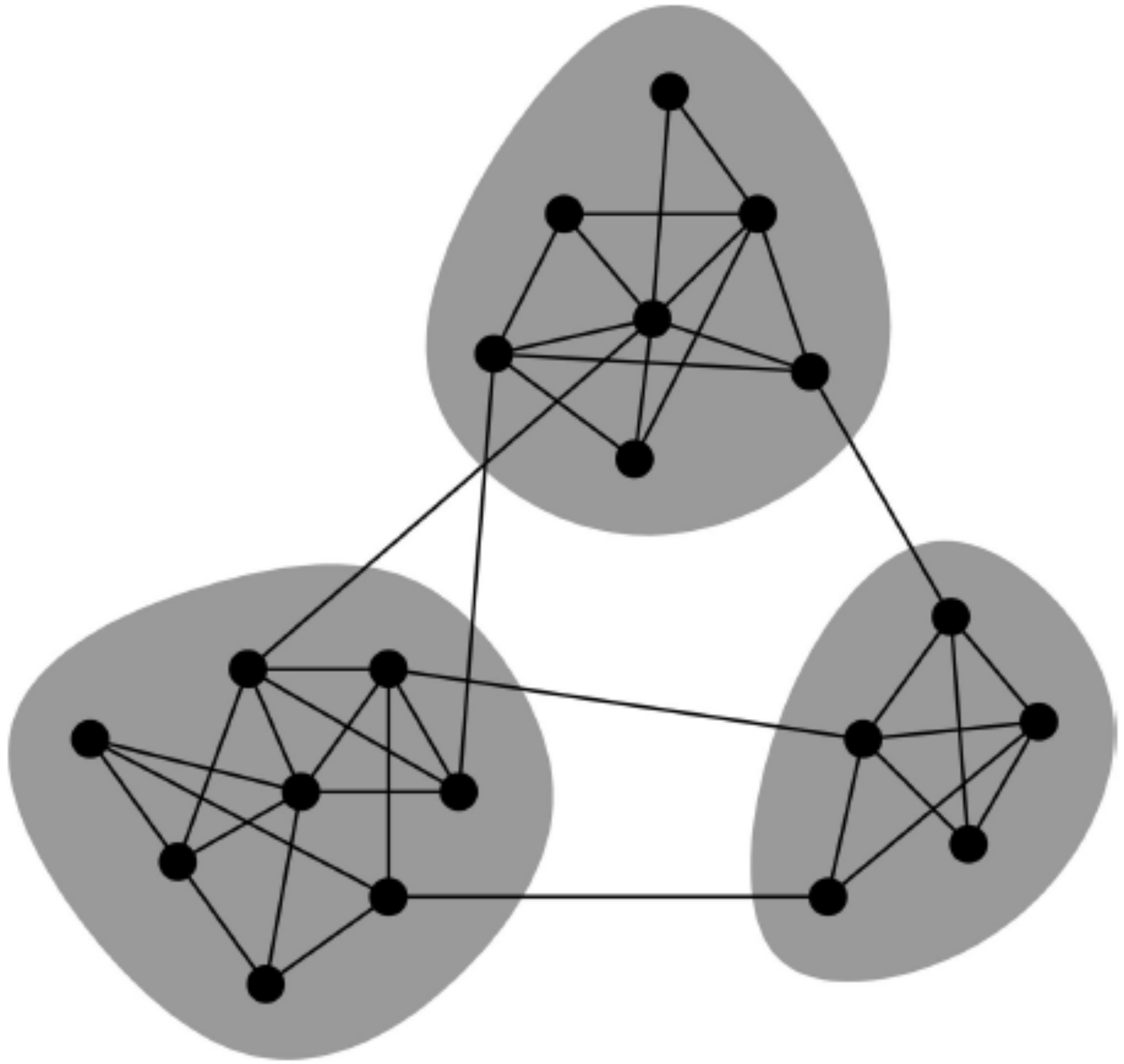
(Because if the phenomenon you observed also happens in Gnp, it's probably not that interesting)

Strong and weak ties

Modeling relationships of varying strength

Networks & Communities

We often think of networks "looking" like this:



What can lead to such a conceptual picture?

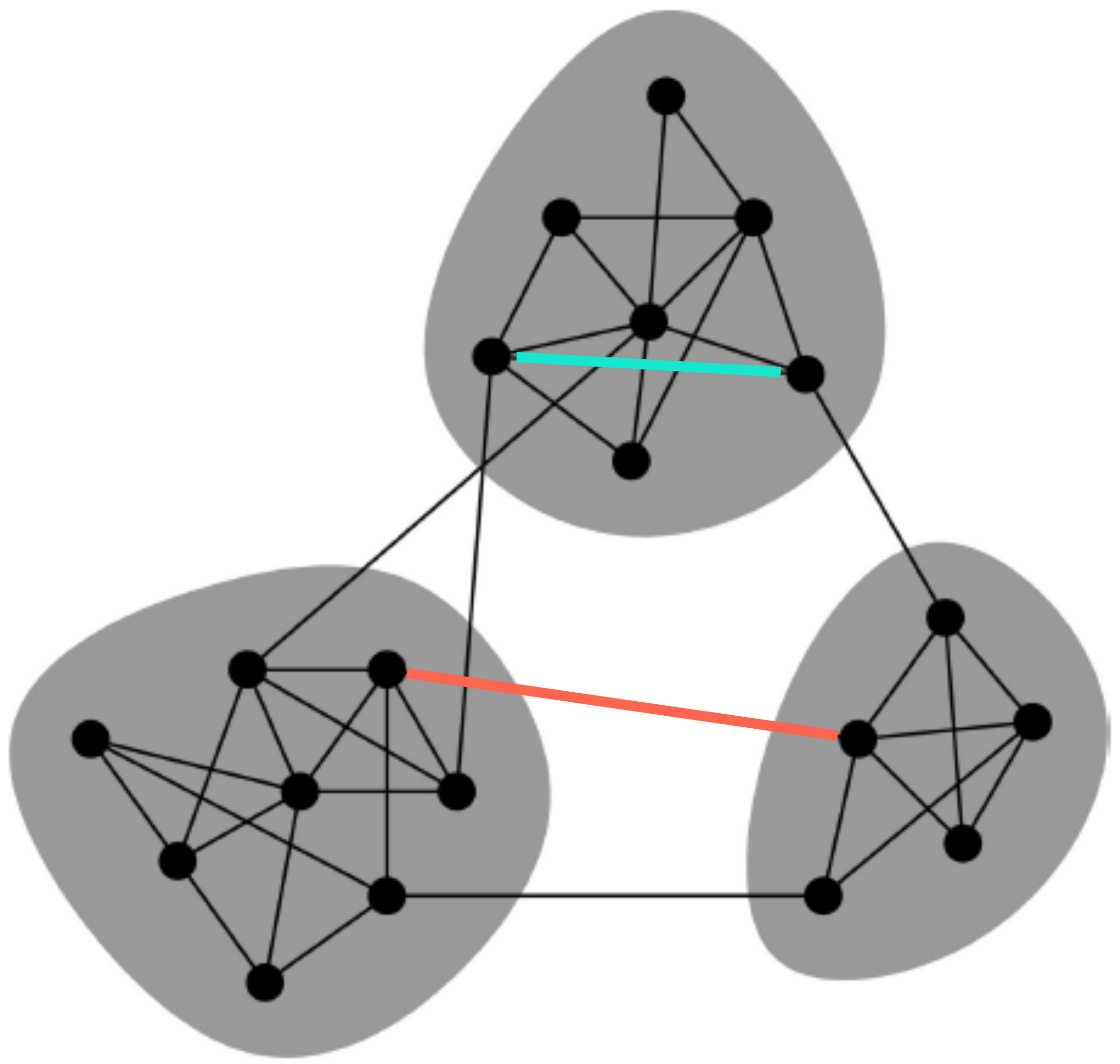
Networks: Flow of Information

- How information flows through the network?
  - What structurally distinct roles do nodes play?
  - What roles do different links (short vs. long) play?
- How people find out about new jobs?
  - Mark Granovetter, part of his PhD in 1960s
  - People find the information through personal contacts
- But: Contacts were often acquaintances rather than close friends
  - This is surprising: One would expect your friends to help you out more than casual acquaintances
- Why is it that acquaintances are most helpful?

Granovetter's Answer

Two perspectives on friendships:

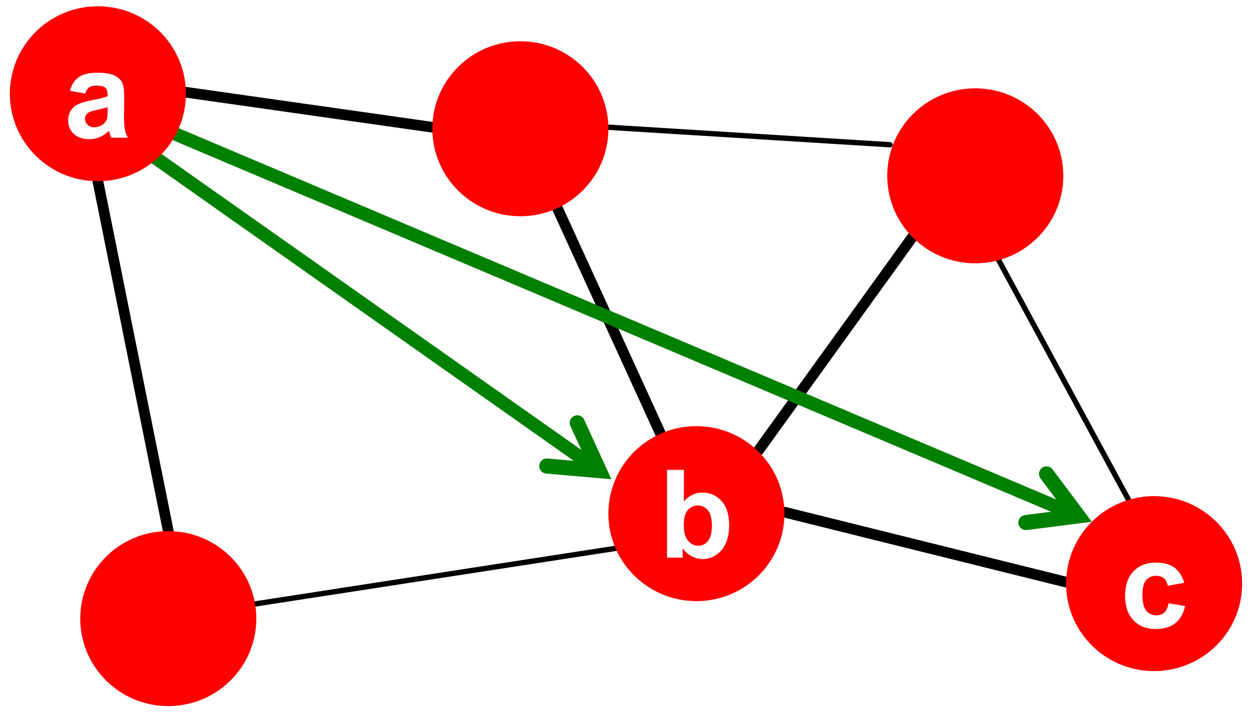
Structural: Friendships span different parts of the network



The two highlighted edges are structurally different: one spans two different "communities" and the other is inside a community

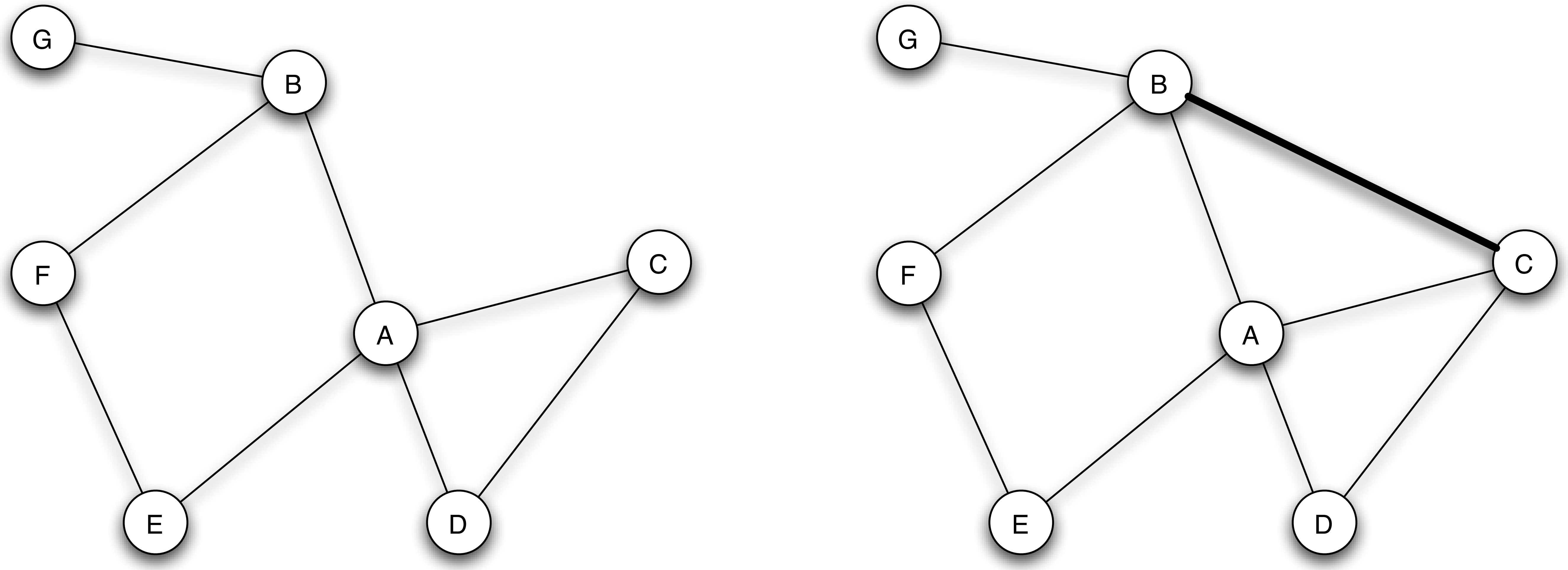
Interpersonal: Friendship between two people vary in strength, you can be close or not so close to someone

Structural force: Triadic closure



Which edge is more likely:
a–b or a–c?

Triadic closure



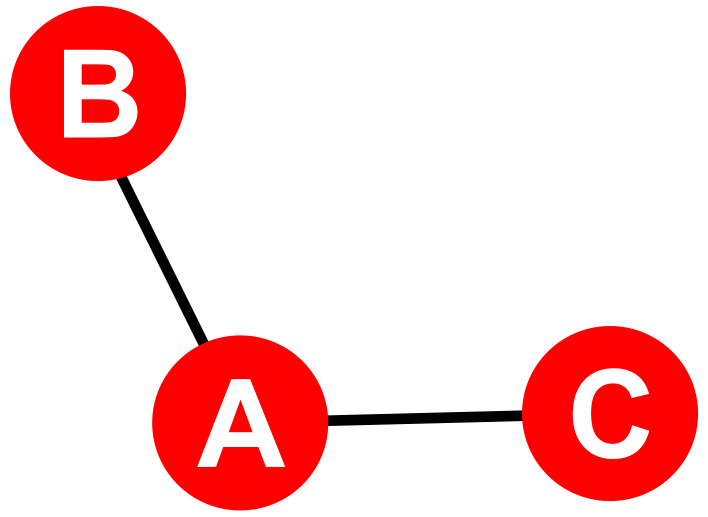
Informally: If two people in a social network have a friend in common, then there is an increased likelihood that they will become friends themselves at some point in the future.

Triadic Closure

- **Triadic closure == High clustering coefficient**

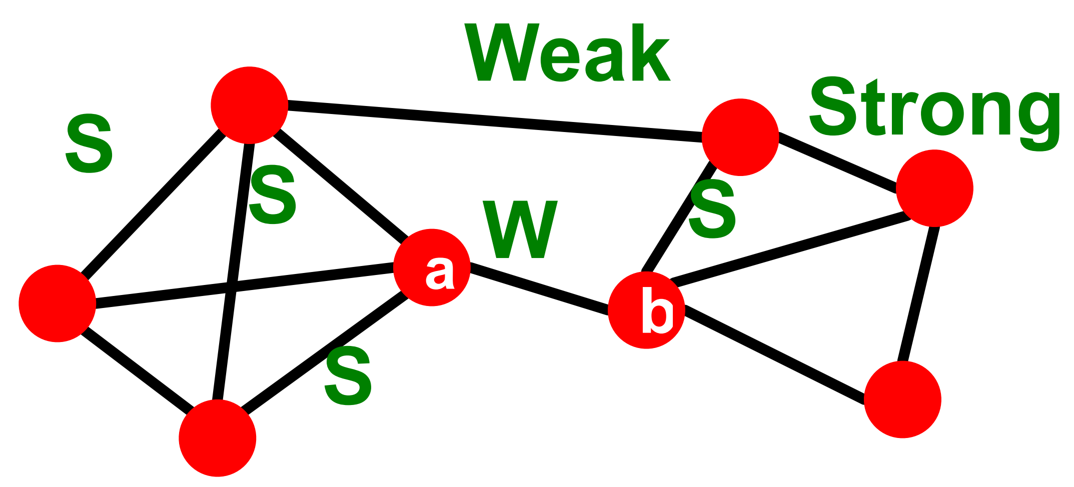
Reasons for triadic closure:

- If $\boldsymbol{B}$ and $\boldsymbol{C}$ have a friend $\boldsymbol{A}$ in common, then:
  - $\boldsymbol{B}$ is **more likely to meet** $\boldsymbol{C}$
    - (since they both spend time with $\boldsymbol{A}$)
  - $\boldsymbol{B}$ and $\boldsymbol{C}$ **trust** each other
    - (since they have a friend in common)
  - $\boldsymbol{A}$ has **incentive** to bring $\boldsymbol{B}$ and $\boldsymbol{C}$ together
    - (as it is hard for $\boldsymbol{A}$ to maintain two disjoint relationships)



Granovetter's Explanation

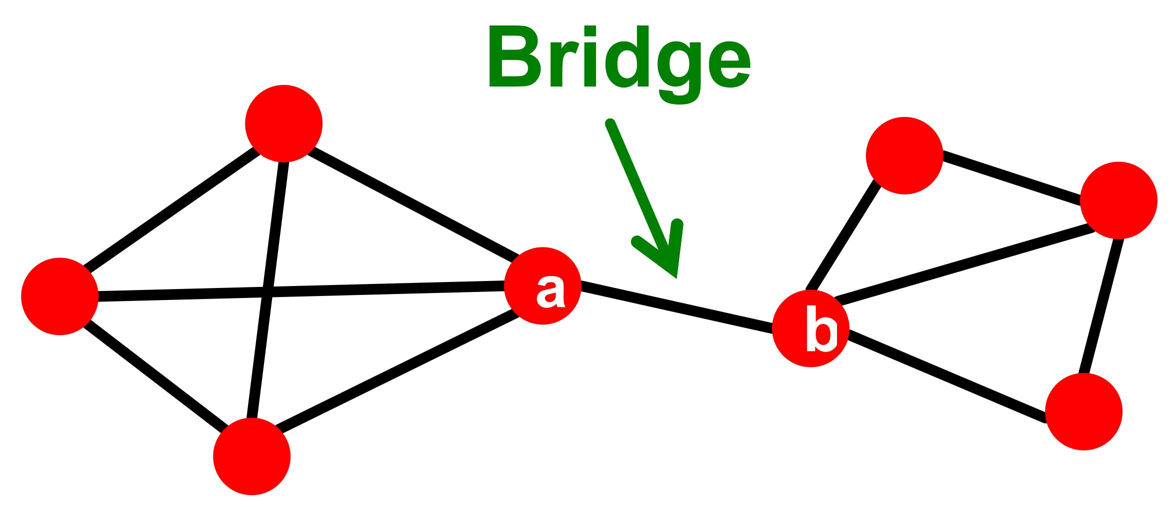
- Granovetter makes a connection between social and structural role of an edge
- **First point: Structure**
  - Structurally embedded edges are also socially strong
  - Long-range edges spanning different parts of the network are socially weak
- **Second point: Information**
  - Long-range edges allow you to gather information from different parts of the network and get a job
  - Structurally embedded edges are heavily redundant in terms of information access



Network Vocabulary: Span and Bridges

Define: **Bridge edge**

If removed, it disconnects the graph



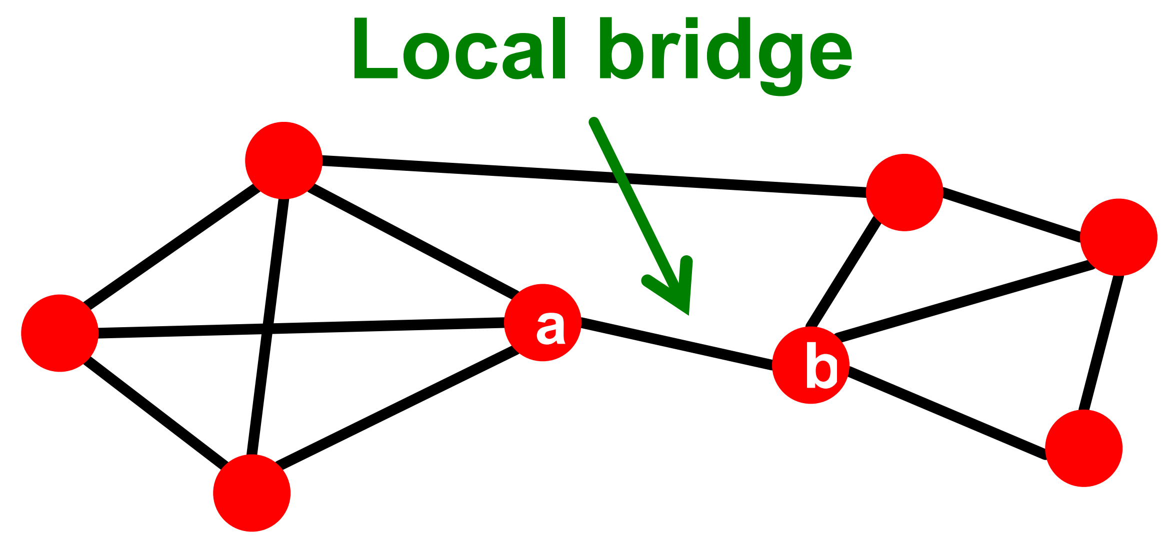
Define: **Span**

The **Span** of an edge is the distance of the edge endpoints if the edge is deleted.

Define: **Local bridge**

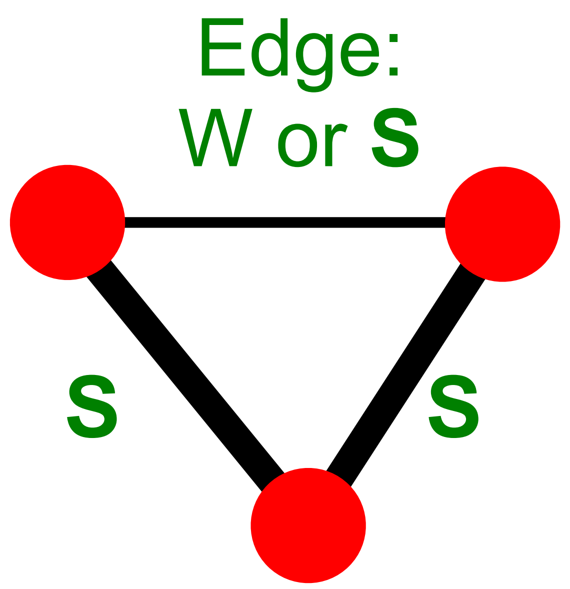
Edge of **Span > 2**

(any edge that doesn't close a triangle)



Idea: Local bridges with long span are like real bridges

Granovetter's Explanation



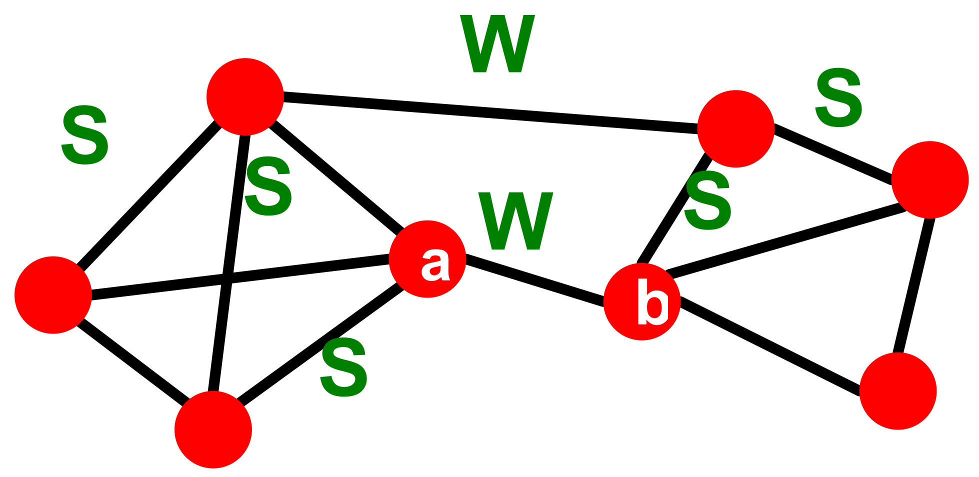
Model: Two types of edges:

Strong (friend), **Weak** (acquaintance)

Model: Strong Triadic Closure property:

Two strong ties imply a third edge

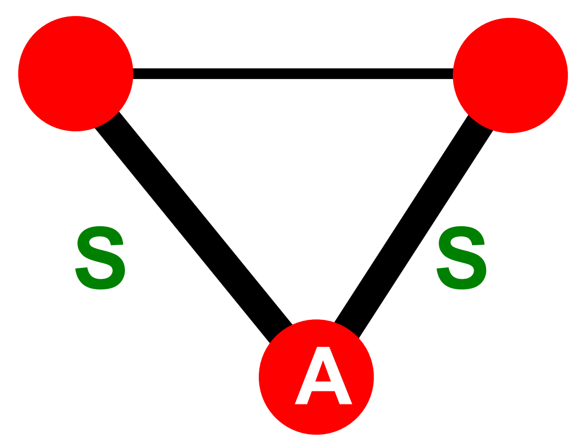
Fact: If strong triadic closure is satisfied then **local bridges are weak ties!**

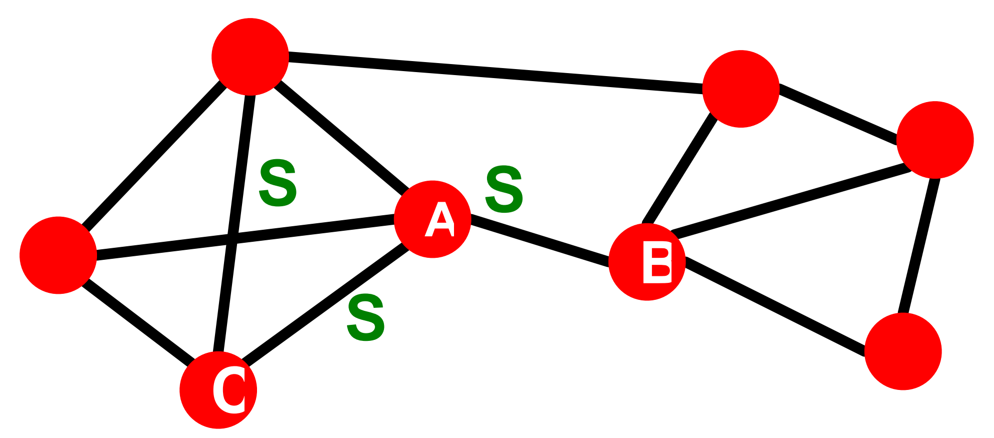


Local Bridges and Weak ties

- **Claim:** If node A satisfies **Strong Triadic Closure** and is involved in at least **two strong ties**, then any **local bridge** adjacent to A must be a **weak tie**.

- **Proof by contradiction:**
  - Assume A satisfies **Strong Triadic Closure** and has **2 strong ties**
  - Let $A - B$ be **local bridge** and a **strong** tie
  - Then $B - C$ must exist because of **Strong Triadic Closure**
  - But then $A - B$ is **not a bridge!**
    (since **B-C** must be connected due to Strong Triadic Closure property)





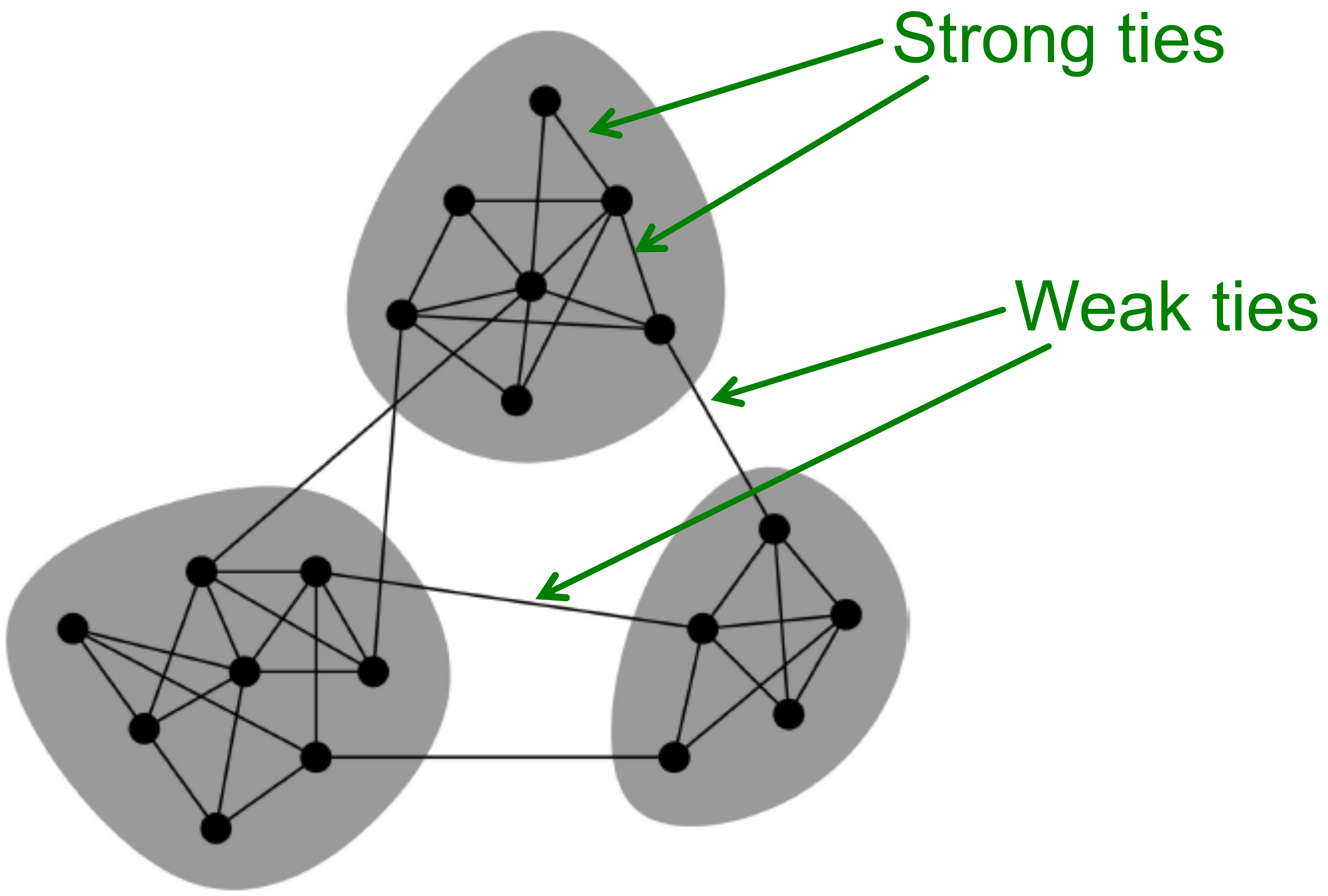
Granovetter's Explanation

Weak ties have access to different parts of the network! Access to other sources and other kinds of information

Strong ties have redundant information

Conceptual Picture of Networks

Granovetter's theory leads to the following conceptual picture of networks



Tie strength in real data

- **For many years Granovetter's theory was not tested**
- But, today we have large who-talks-to-whom graphs:
  - Email, Messenger, Cell phones, Facebook
- **Onnela et al. 2007:**
  - Cell-phone network of 20% of country's population
  - **Edge strength:** # phone calls

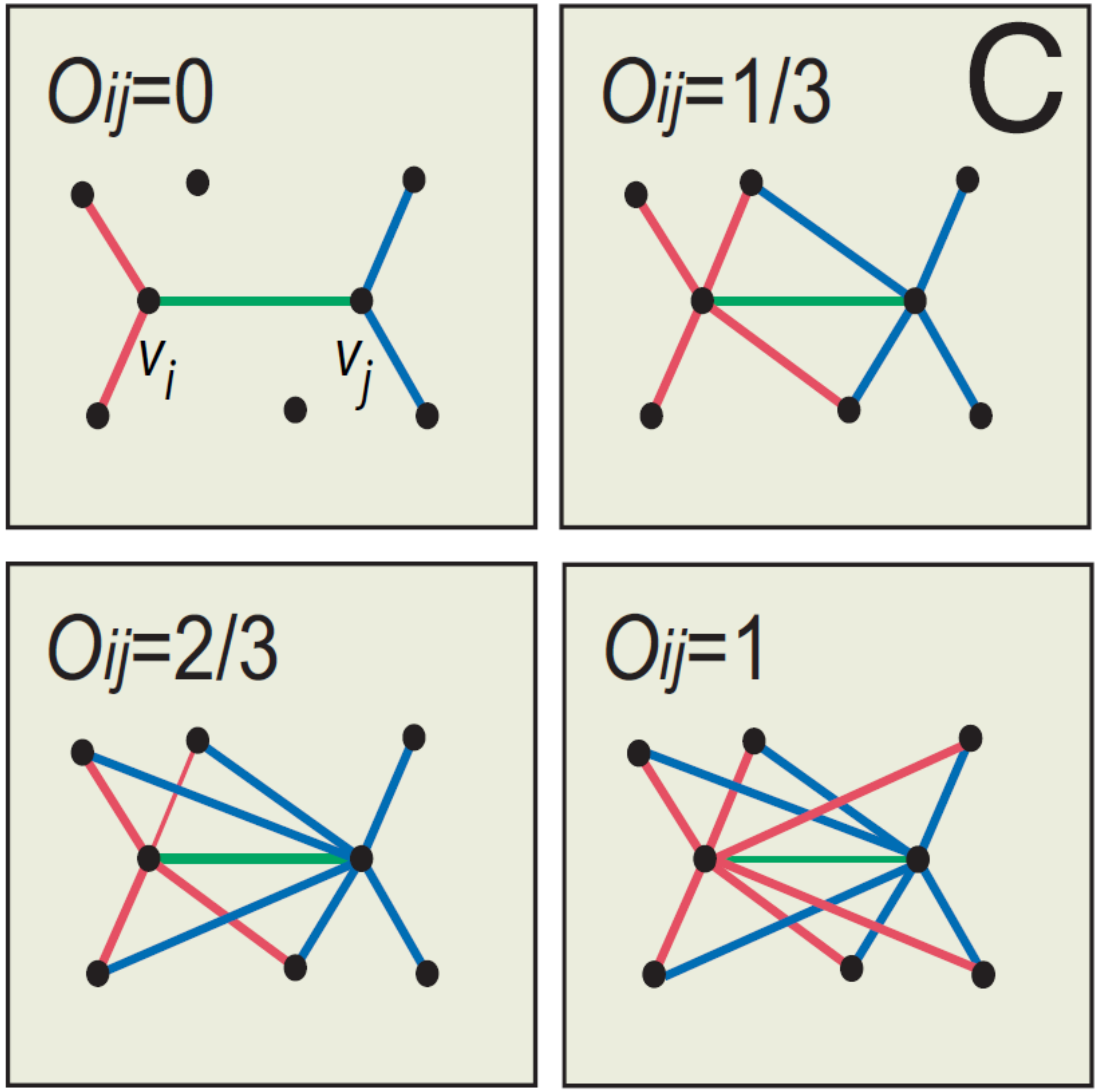
Neighborhood Overlap

- **Edge overlap:**

$$O_{ij} = \frac{N(i) \cap N(j)}{N(i) \cup N(j)}$$

  - $N(i)$ … a set of neighbors of node i

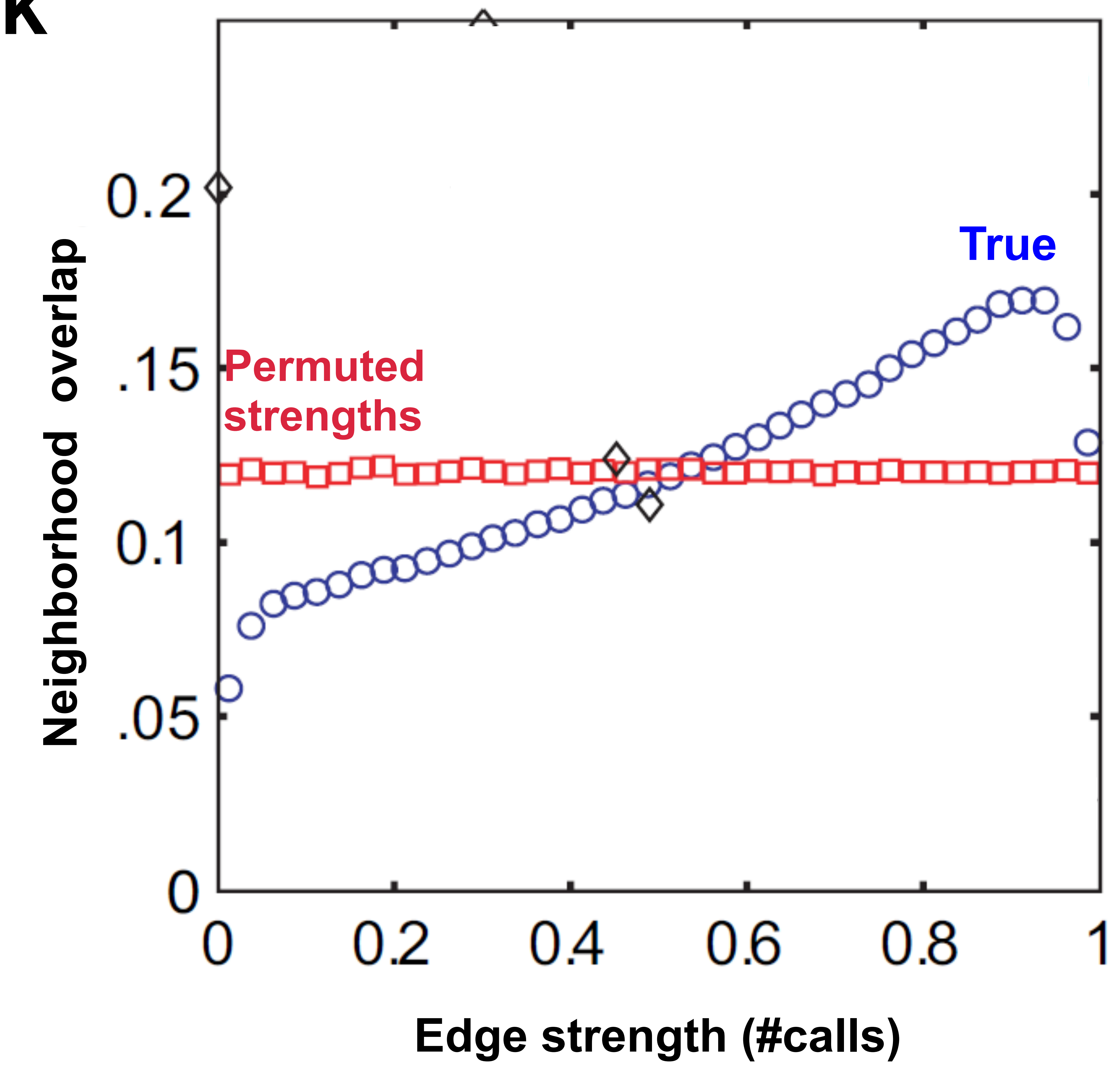
- **Overlap = 0** when an edge is a **local bridge**



Phones: Edge Overlap vs. Strength

- Cell-phone network
- **Observation:**
  - Highly used links have high overlap!

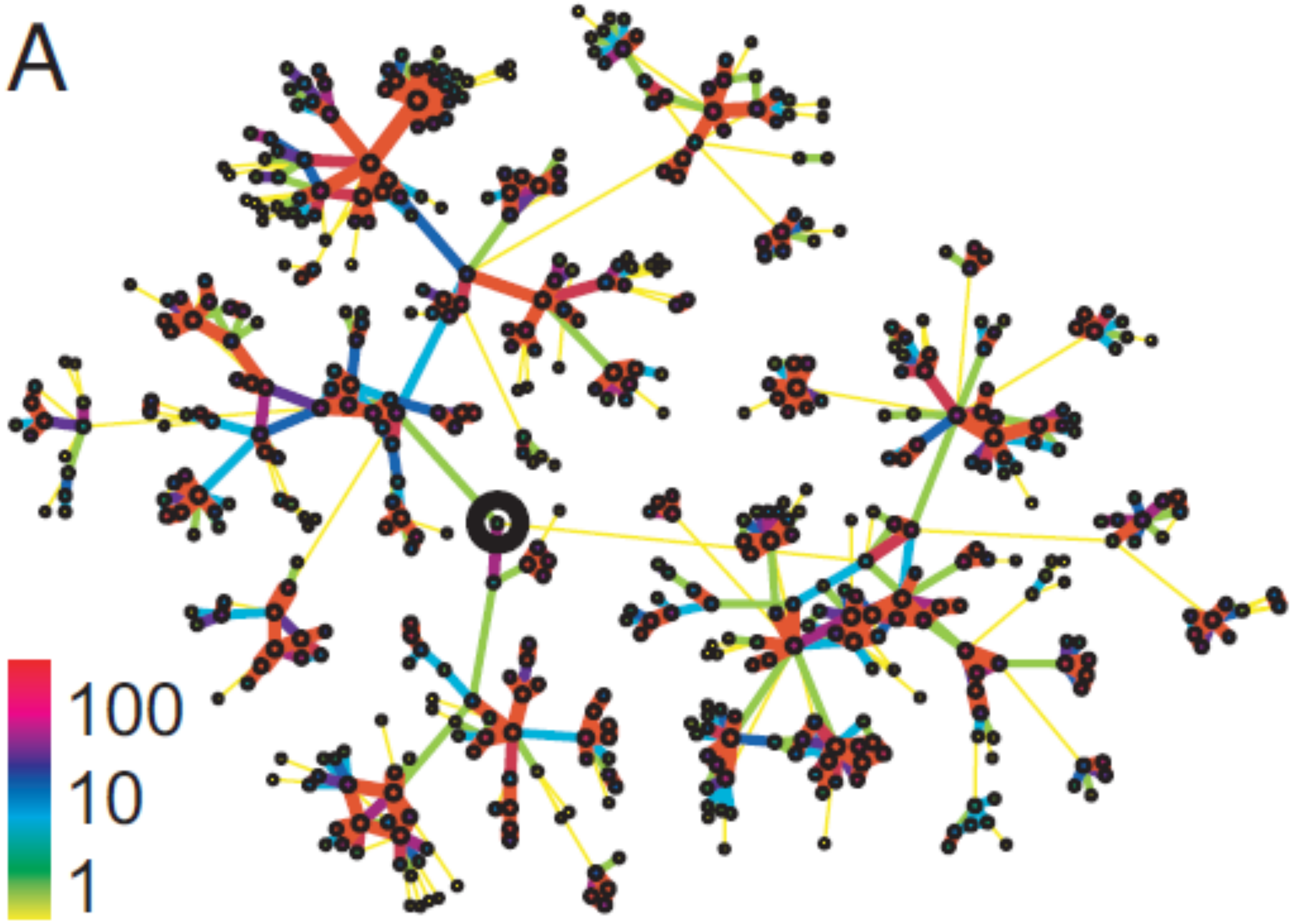
- Legend:
  - **True:** The data
  - **Permuted strengths:** Keep the network structure but randomly reassign edge strengths



Real Network, Real Tie Strengths

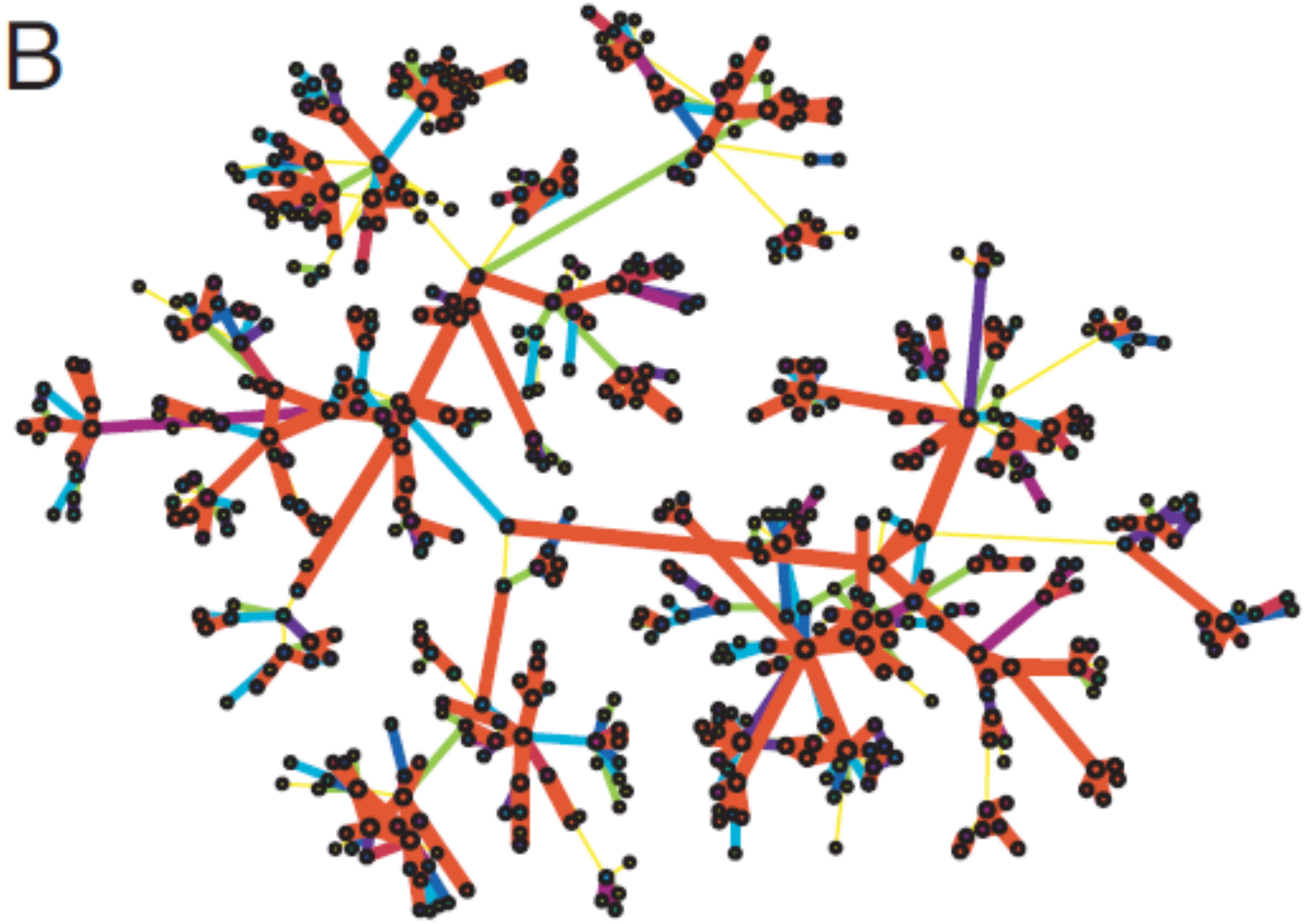
Real edge strengths in mobile call graph

Strong ties are more embedded (have higher overlap)



Real Net, Permuted Tie Strengths

Same network, same set of edge strengths
but now **strengths are randomly shuffled**

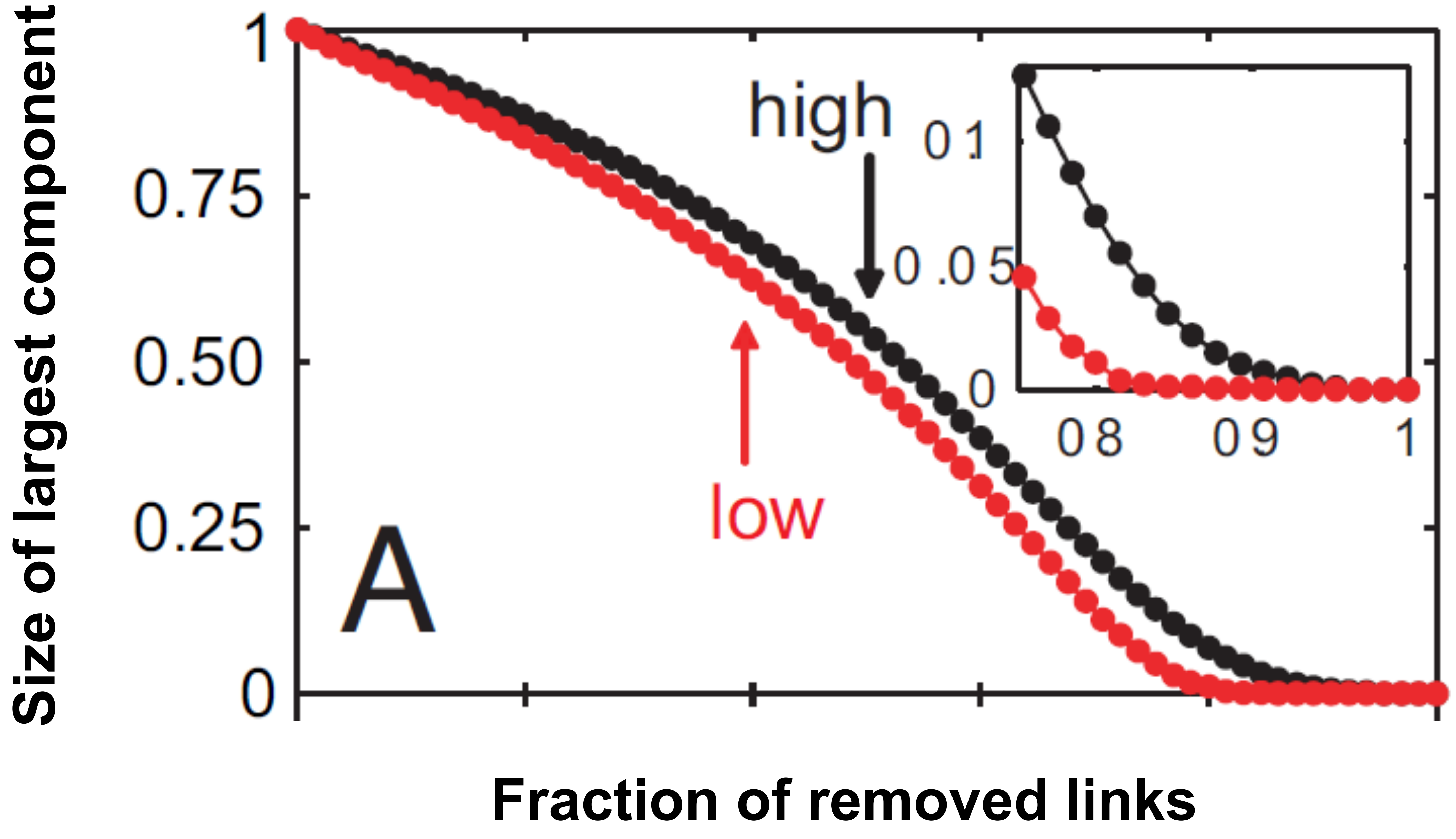


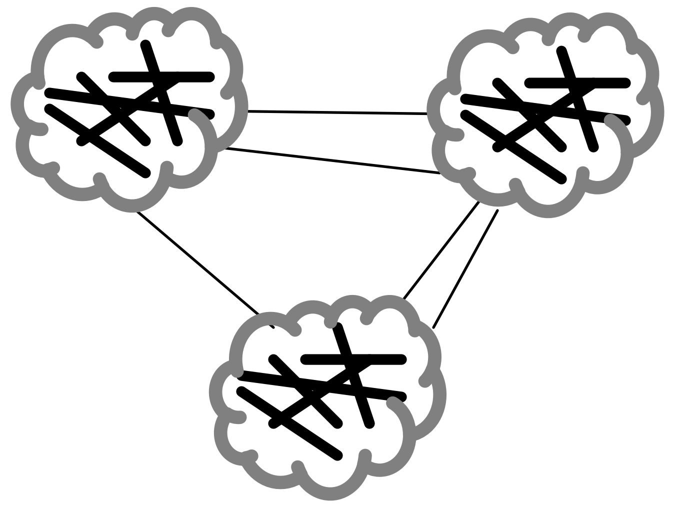
Link Removal by Strength

Removing links by **strength (#calls)**

- Low to high
- High to low

Low disconnects the network sooner





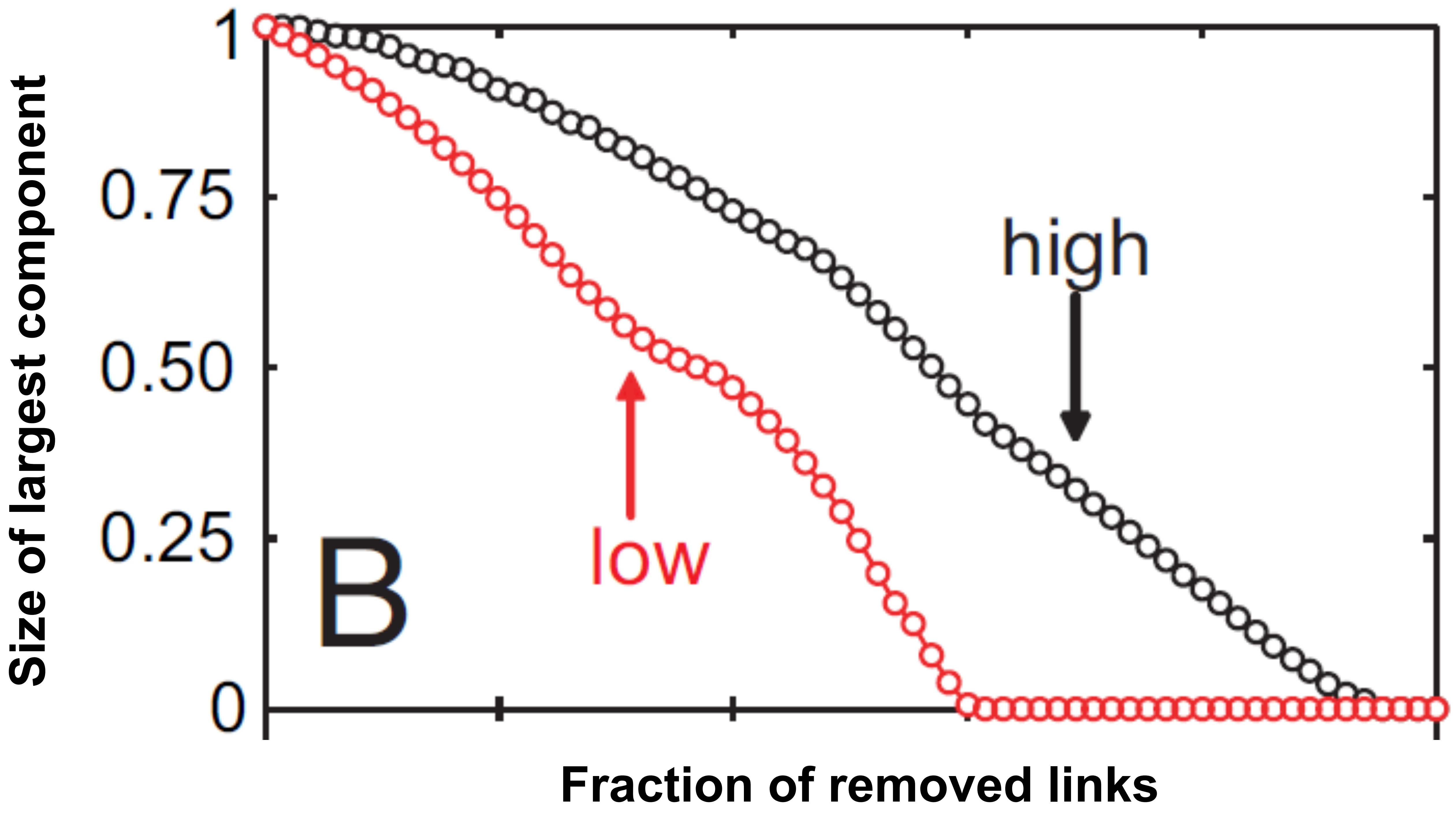
Conceptual picture of network structure

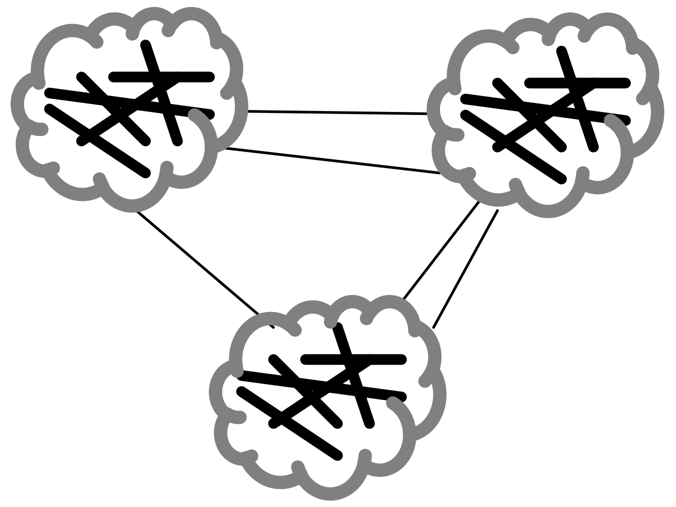
Link Removal by Overlap

Removing links based on **overlap**

- Low to high
- High to low

Low disconnects the network sooner





Conceptual picture of network structure

Course progress

Assignment 1 out next week

- Due 2 weeks later
- Get started early!