Reminders/Announcements

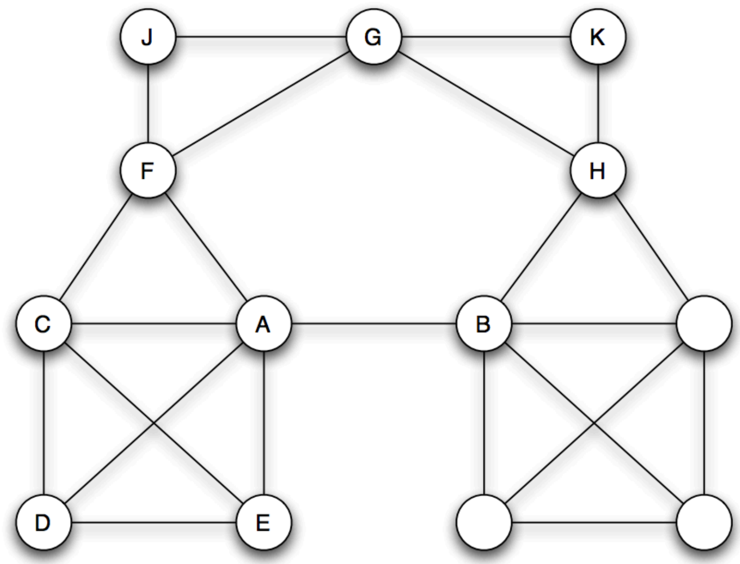
- When confused in lecture: ask right away!
- If still confused: ask during discussion, office hours, email, discussions on Canvas, etc.
- Bottom line: if confused: ask, ask ,ask!
- If you prefer to ask privately, you can email **si301.f2017@umich.edu** to reach me and the GSIs.
- Using my email address is fine, but response may be slower.
- Please register your iClicker on Canvas.
- When do I post lecture slides? Usually within 24 hours of the lecture, but if it has been longer and you want to see the slides, please email me.
- Please complete social network survey.

Last Time

- Why networks? Because they are everywhere.
- Connectivity:
  - Paths
  - Cycles
  - Connected graphs
  - Connected components
  - Giant component (informal)
  - Distance between two nodes
  - Breadth-First search

Warm up

1. Is the network connected?
2. How many connected components does it have?
3. Identify 3 cycles with more 3 edges.
4. Find the distance between E and G.
5. What the diameter of the network

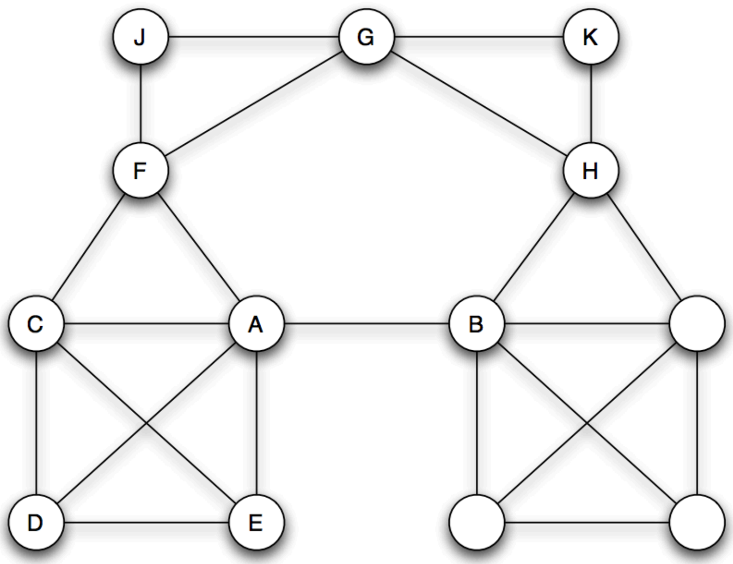


Warm up

1. Is the network connected?

A. Yes
B. No

Yes, there is a path from each node to every other node.

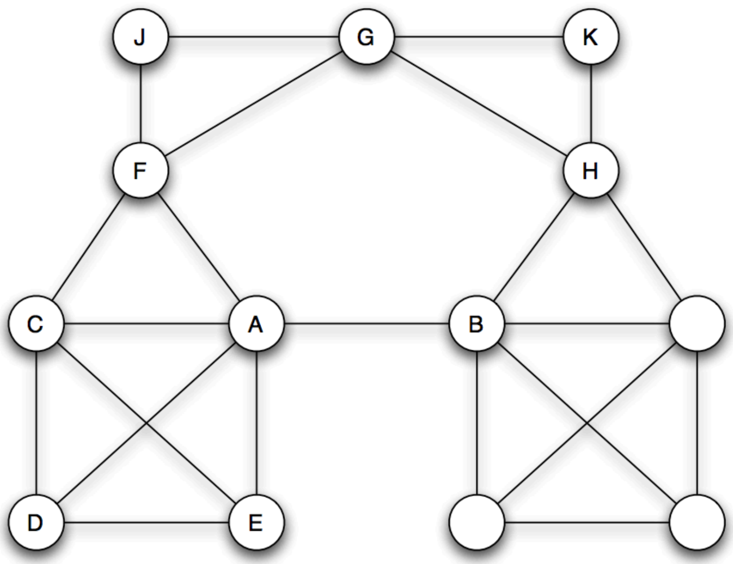


Warm up

2. How many connected components does it have?

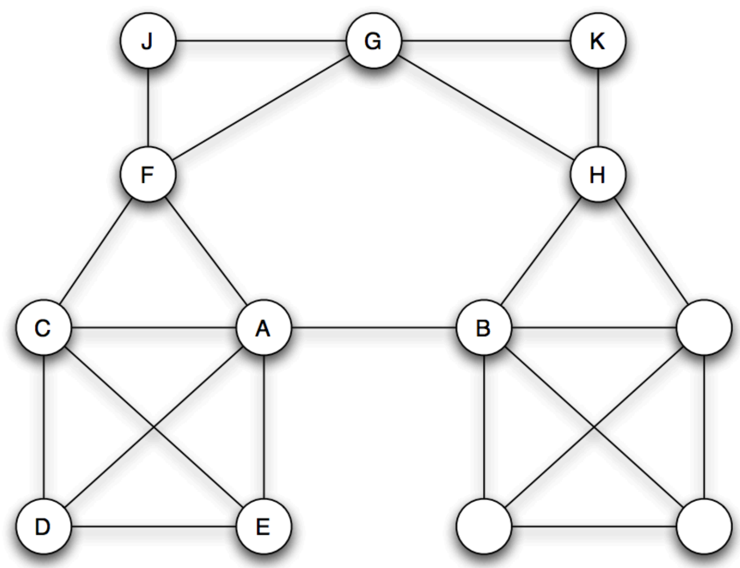
A. 1
B. 2
C. 3
D. 4

1, all connected networks have 1 connected component.



Warm up

3. Identify 3 cycles with more 3 edges.
   F – G – H – B – A – F
   C – A – E – D – C
   J – G – K – H – B – A – F – J

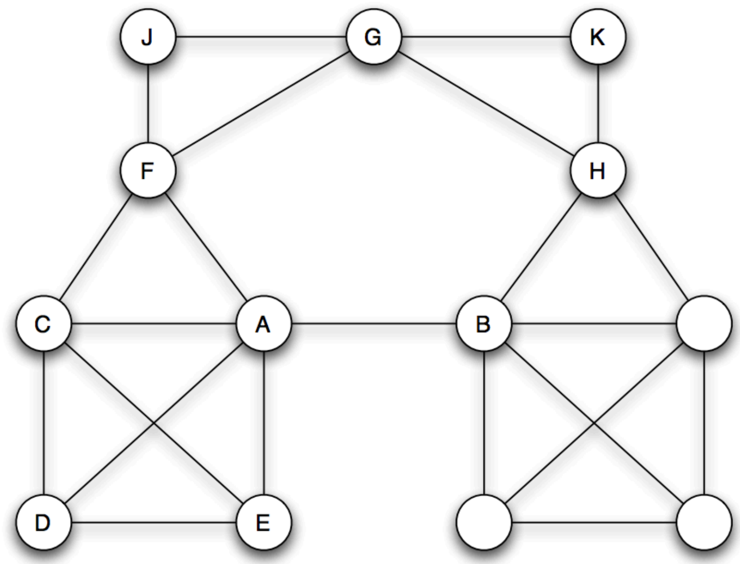


Warm up

4. Find the distance between E and G.

A. 2
B. 3
C. 4
D. 5

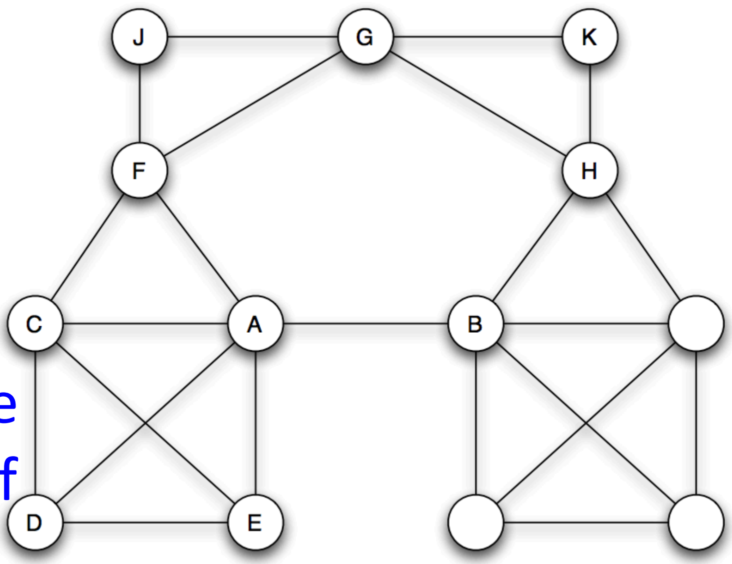
3, path E-A-F-G is the shortest path between E and F



Warm up

5. What the diameter of the network

A. 2
B. 3
C. 4
D. 5



4, D and K have distance 4 and this is the largest distance between any two pair of nodes.

The Small-World Phenomenon

The world is small in the sense that "short" paths exists between almost any two people.

The Small-World Phenomenon

The world is small in the sense that "short" paths exists between almost any two people.

How short are these paths?

How can we measure their length?

Milgram Small World Experiment

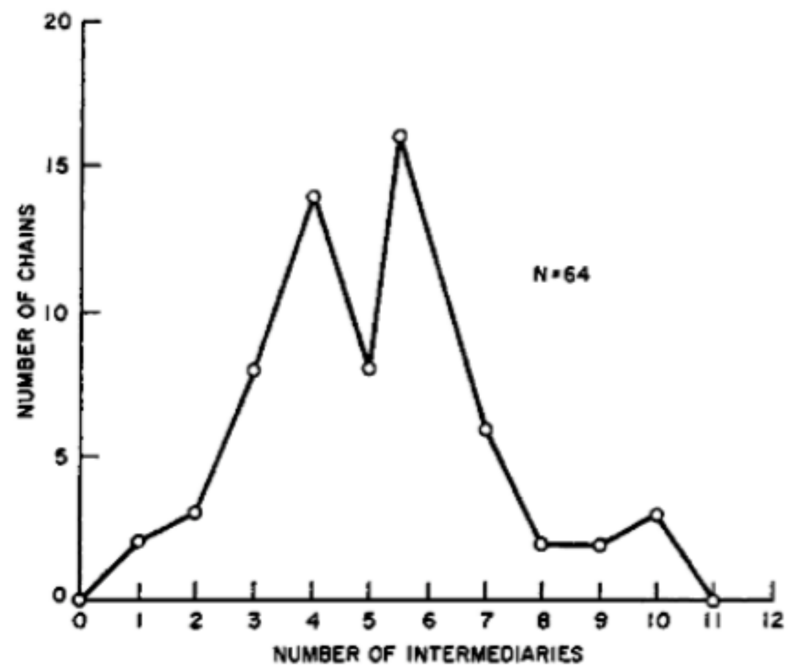
Set up:

- 296 randomly chosen “starters” asked to forward a letter to a “target” person.
- Target was a stockbroker in Boston.
- Instructions for starter:
  - Send letter to target if you know him on a first name basis.
  - If you do not know target, send letter (and instructions) to someone you know on a first name basis who is more likely to know the target.
- Some information about the target, such as city, and occupation, was provided.

Milgram Small World Experiment

Results:

- 64 out of the 296 letters reached the target.
- Median chain length was 6 (hence the phrase "six degrees of separation")



Key points:

- A relatively large percentage (>20%) of letters reached target.
- Paths were relatively short.
- People were able to find these short paths.

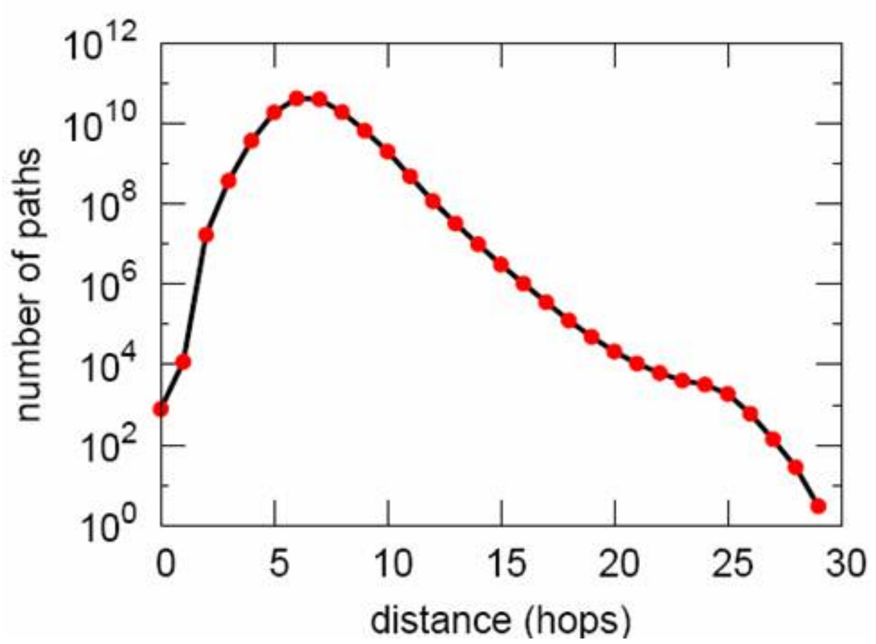
Caveats – see assignment 1.

Small World of Instant Message

Nodes: 240 million active users on Microsoft Instant Messenger.

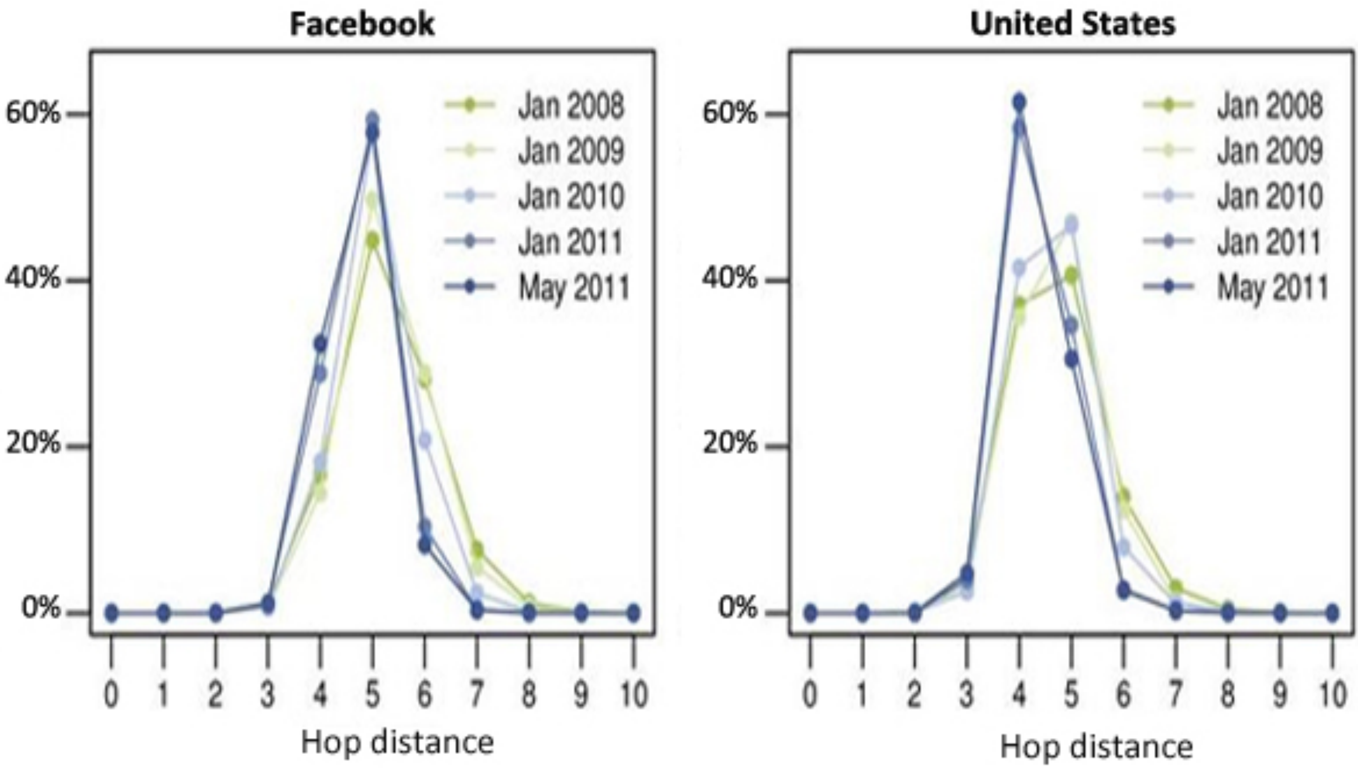
Edges: Users engaged in two-way communication over a one-month period.

Estimated median path length of 7.



[Leskovec and Horvitz, 2008]

Small World of Facebook



[Backstrom, Boldi, Rosa, Ugander, Vigna, 2012]

- Global network: average path length in 2008 was 5.28 and in 2011 it was 4.74.
- Path are even shorter if network is restricted to US only.

The Strength of Weak Ties

The **strength** of **weak ties**
MS Granovetter - American journal of sociology, 1973 - JSTOR
Analysis of social networks is suggested as a tool for linking micro and macro levels of sociological theory. The procedure is illustrated by elaboration of the macro implications of one aspect of small-scale interaction: the strength of dyadic ties. It is argued that the ...
Cited by 34761 Related articles All 108 versions Cite Save More

- Sample of 54 people who has recently found a job through a personal contact.
- Respondents were asked how often they saw the contact who helped find the job
- Results:
  - **Often** (twice/week or more): **16.7%**
  - **Occasionally** (between once/year and twice/week): **55.6%**
  - **Rarely** (once/year or less): **27.8%**

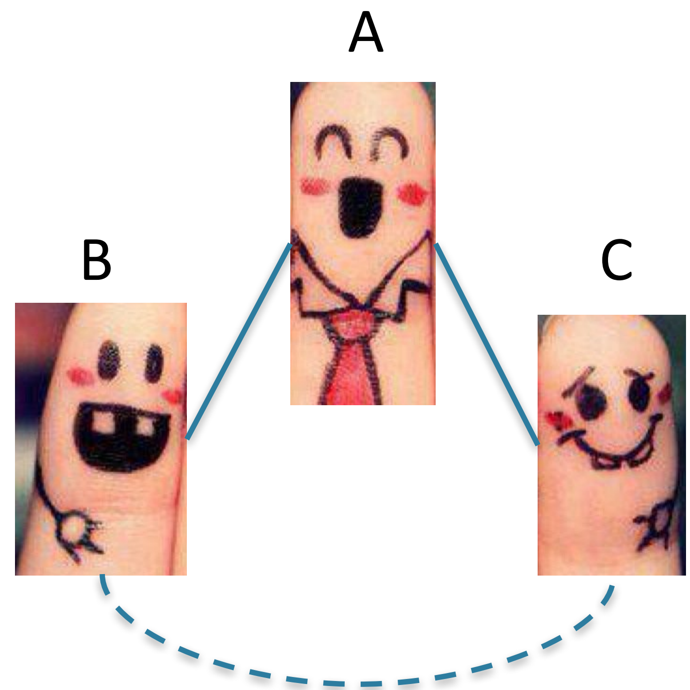
Weak ties were more important than strong ties in job search

Triadic Closure

In a social network, when two people have a connection in common, they are more likely to become connected.

Triadic Closure

In a social network, when two people have a connection in common, they are more likely to become connected.



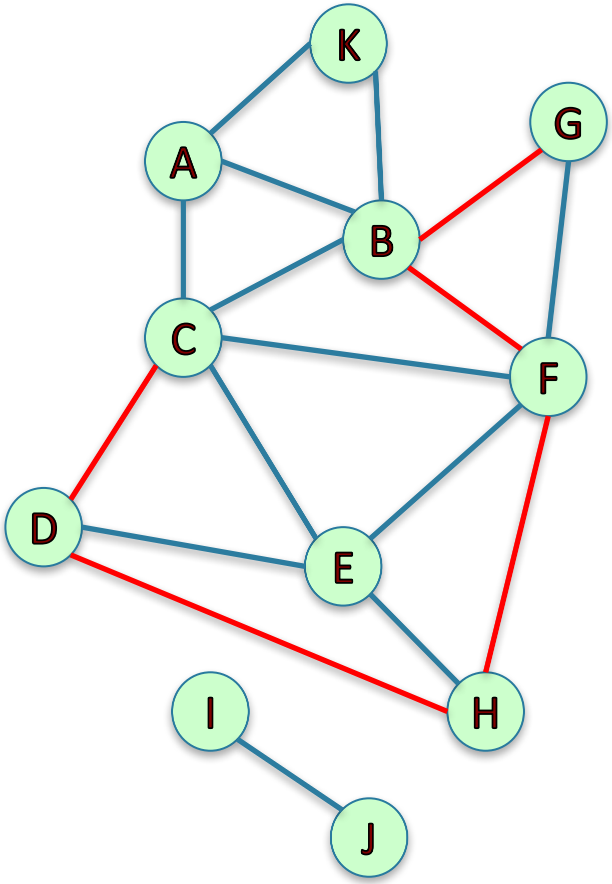
Why?

1. *Opportunity* – Because B and C know A, they are likely to meet through A.
2. *Trust* – Once they meet, trust is easier to establish. A serves as a reference.
3. *Incentive* – A may get more out of her friendships if B and C are also friends.

Triadic Closure

In a social network, when two people have a connection in common, they are more likely to become connected.

How can we measure the prevalence of triadic closure in a network?

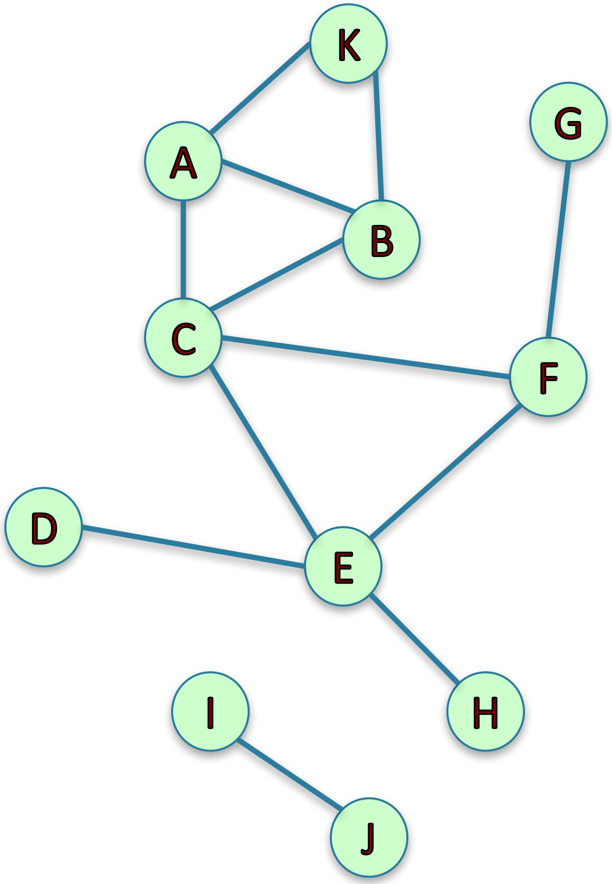


Clustering Coefficient

Local clustering coefficient of a node:
Fraction of pairs of the node's friends that are friends with each other.

Compute the local clustering coefficient of node C:

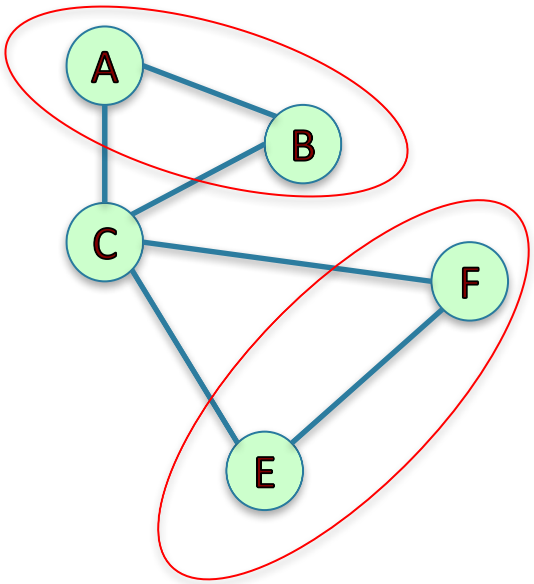
$$\frac{\text{\# of pairs of C's friends who are friends}}{\text{\# of pairs of C's friends}}$$



Clustering Coefficient

Local clustering coefficient of a node:
Fraction of pairs of the node's friends that are friends with each other.

Compute the local clustering coefficient of node C:

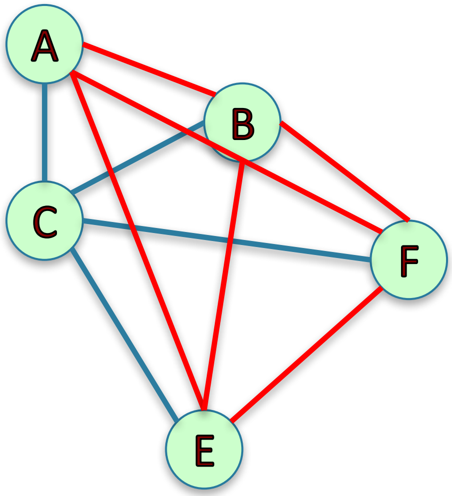


$$\frac{\text{\# of pairs of C's friends who are friends}}{\text{\# of pairs of C's friends}} = \frac{2}{}$$

Clustering Coefficient

Local clustering coefficient of a node:
Fraction of pairs of the node's friends that are friends with each other.

Compute the local clustering coefficient of node C:

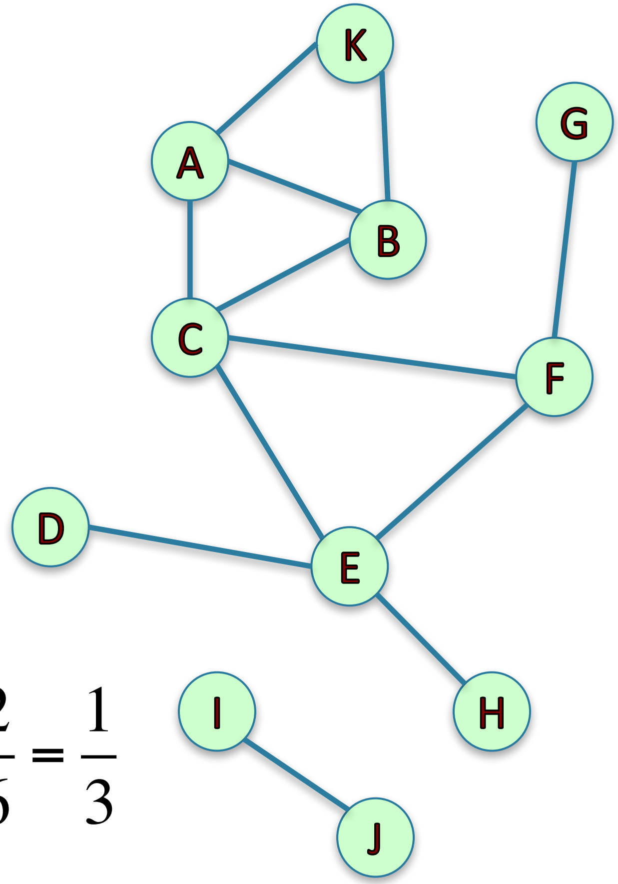


$$\frac{\text{\# of pairs of C's friends who are friends}}{\text{\# of pairs of C's friends}} = \frac{2}{6}$$

Clustering Coefficient

Local clustering coefficient of a node:
Fraction of pairs of the node's friends that are friends with each other.

Compute the local clustering coefficient of node C:

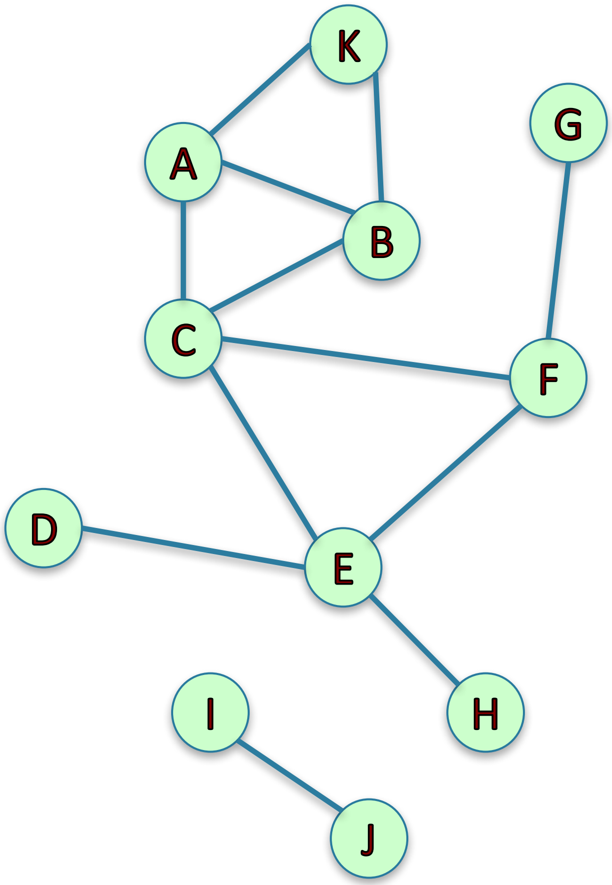


$$\frac{\text{\# of pairs of C's friends who are friends}}{\text{\# of pairs of C's friends}} = \frac{2}{6} = \frac{1}{3}$$

Clustering Coefficient

Compute the local clustering coefficient of node E:

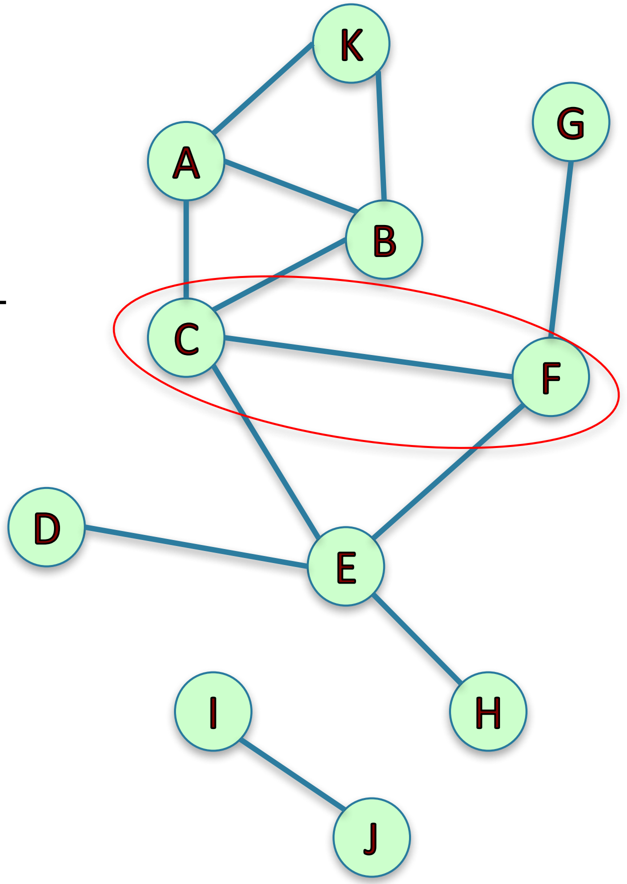
$$\frac{\text{\# of pairs of E's friends who are friends}}{\text{\# of pairs of E's friends}} =$$



Clustering Coefficient

Compute the local clustering coefficient of node E:

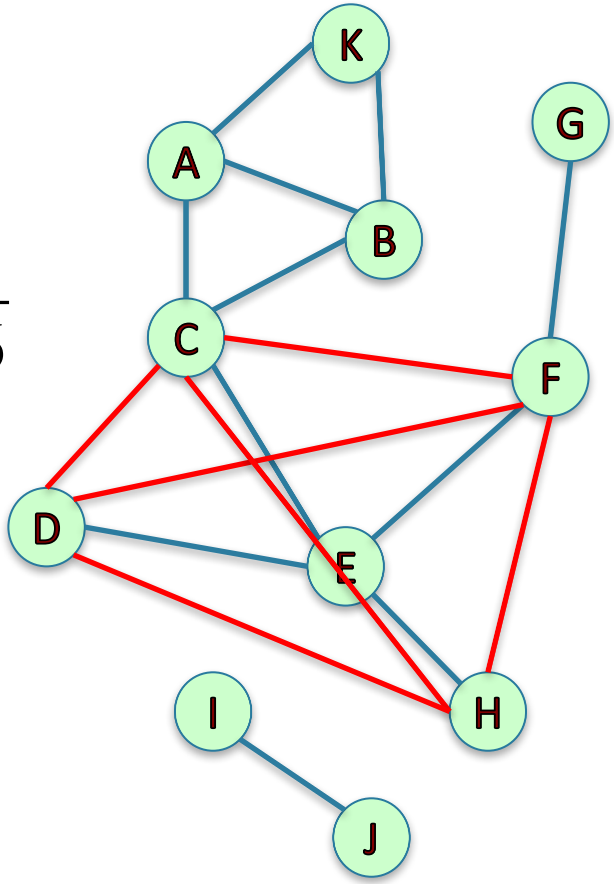
$$\frac{\text{\# of pairs of E's friends who are friends}}{\text{\# of pairs of E's friends}} = \frac{1}{\;}$$



Clustering Coefficient

Compute the local clustering coefficient of node E:

$$\frac{\text{\# of pairs of E's friends who are friends}}{\text{\# of pairs of E's friends}} = \frac{1}{6}$$



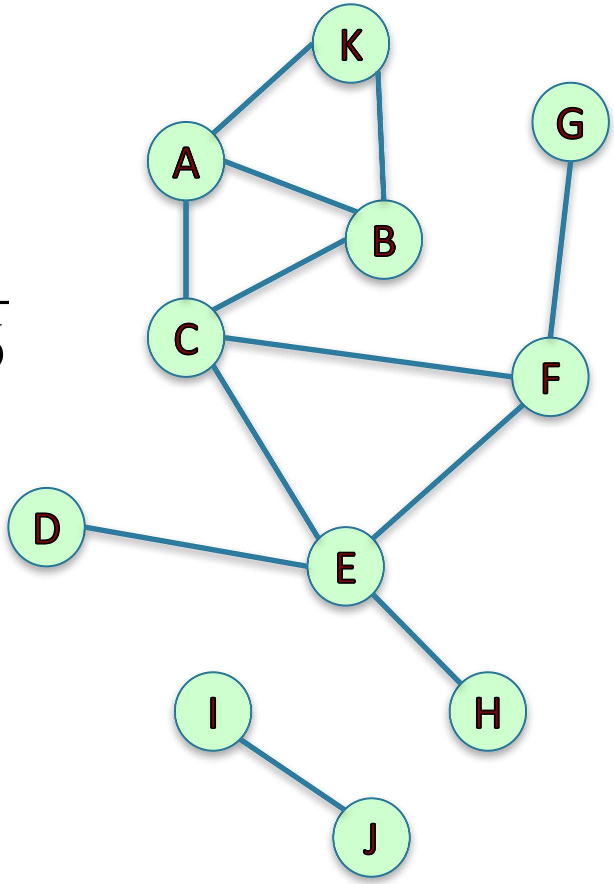
Clustering Coefficient

Compute the local clustering coefficient of node E:

$$\frac{\text{\# of pairs of E's friends who are friends}}{\text{\# of pairs of E's friends}} = \frac{1}{6}$$

Compute the local clustering coefficient of node B:

$$\frac{\text{\# of pairs of B's friends who are friends}}{\text{\# of pairs of B's friends}}$$



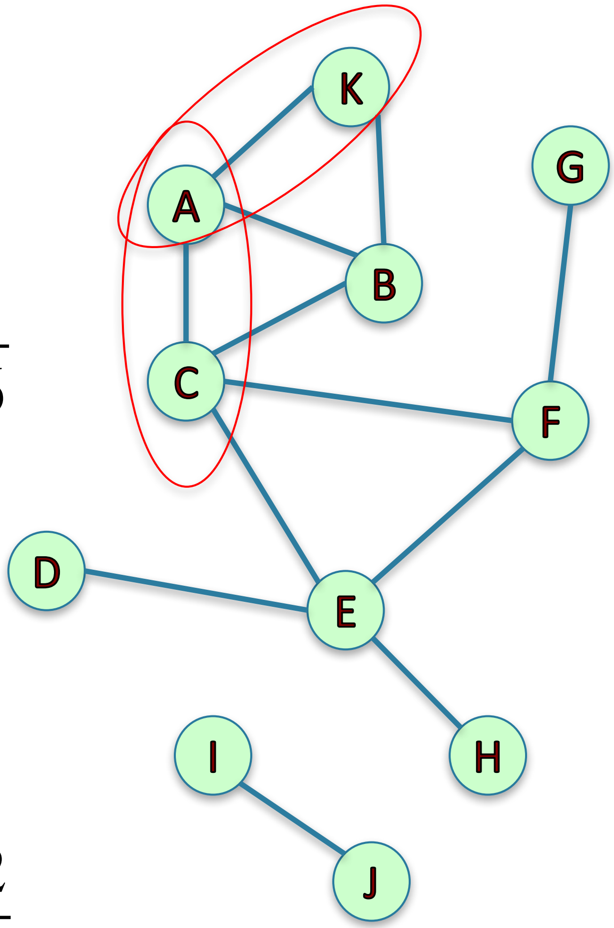
Clustering Coefficient

Compute the local clustering coefficient of node E:

$$\frac{\text{\# of pairs of E's friends who are friends}}{\text{\# of pairs of E's friends}} = \frac{1}{6}$$

Compute the local clustering coefficient of node B:

$$\frac{\text{\# of pairs of B's friends who are friends}}{\text{\# of pairs of B's friends}} = \frac{2}{}$$



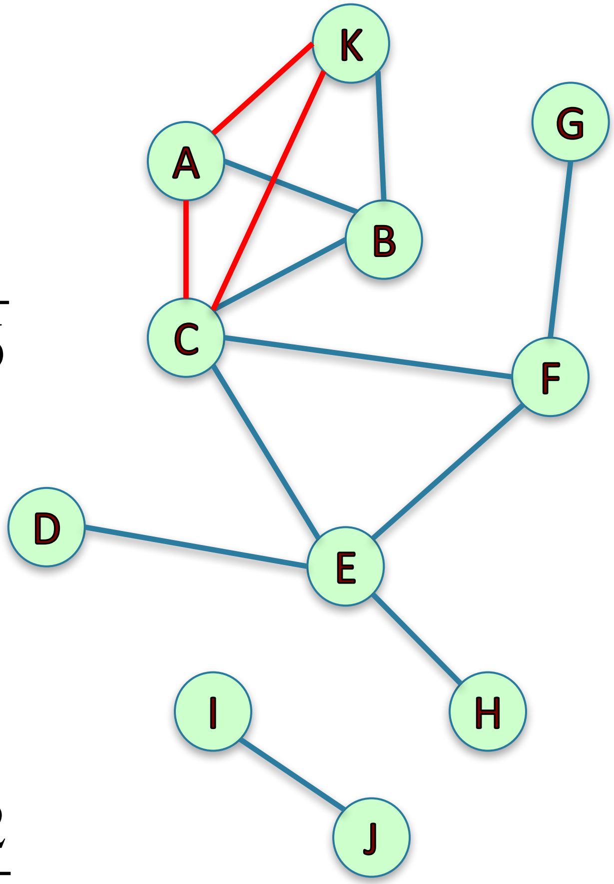
Clustering Coefficient

Compute the local clustering coefficient of node E:

$$\frac{\text{\# of pairs of E's friends who are friends}}{\text{\# of pairs of E's friends}} = \frac{1}{6}$$

Compute the local clustering coefficient of node B:

$$\frac{\text{\# of pairs of B's friends who are friends}}{\text{\# of pairs of B's friends}} = \frac{2}{3}$$



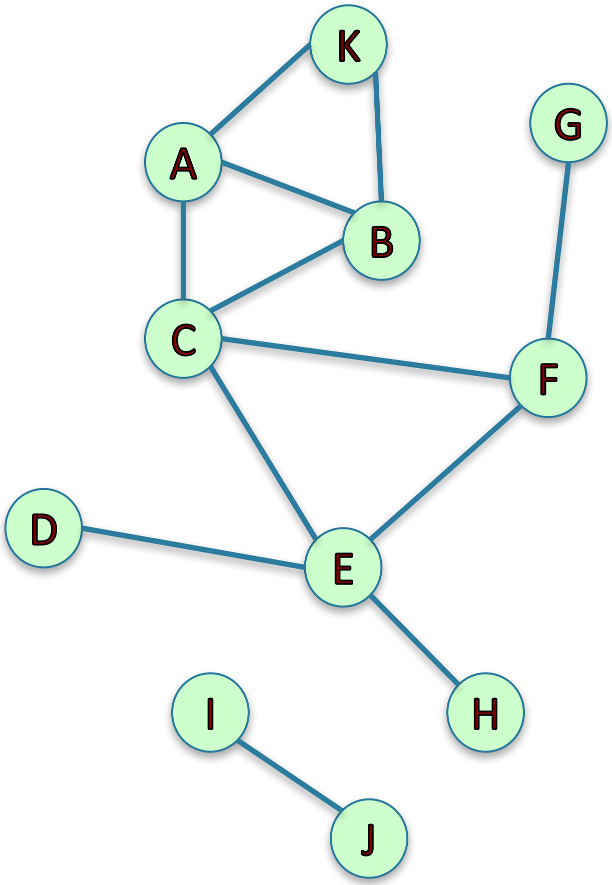
Clustering Coefficient

Compute the local clustering coefficient of node I:

# of pairs of I's friends = 0

Can't divide by 0!

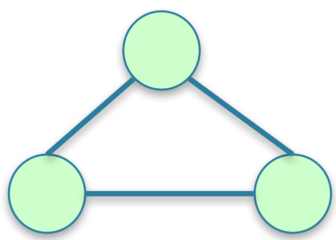
If a node has less than 2 connections, we say it has local clustering coefficient of 0.

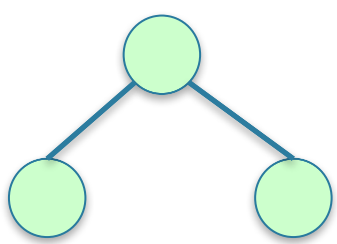


Clustering Coefficient

Global clustering coefficient (GCC) of a network:

Definition 1: $$\textbf{Transitivity} = \frac{3 * \text{Number of closed triads}}{\text{Number of open triads}}$$

Closed triads: 

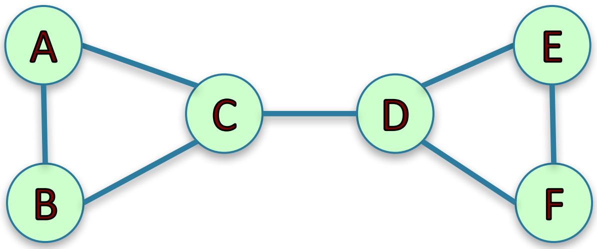
Open triads: 

Clustering Coefficient

Global clustering coefficient (GCC) of a network:

$$\textbf{Transitivity} = \frac{3{*}\text{Number of closed triads}}{\text{Number of open triads}}$$

Compute the transitivity of the network:



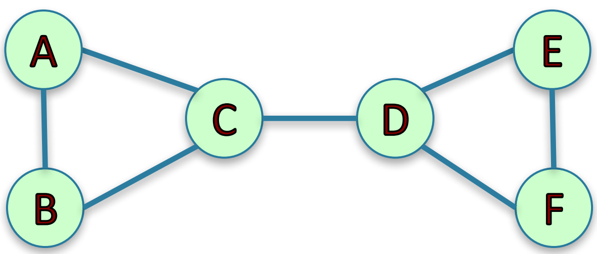
Clustering Coefficient

Global clustering coefficient (GCC) of a network:

$$\textbf{Transitivity} = \frac{3 * \text{Number of closed triads}}{\text{Number of open triads}}$$

Compute the transitivity of the network:

Number of closed triads:



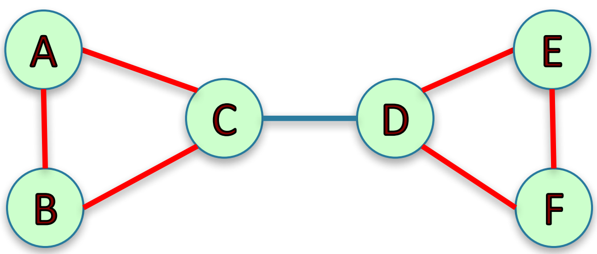
Clustering Coefficient

Global clustering coefficient (GCC) of a network:

$$\textbf{Transitivity} = \frac{3 * \text{Number of closed triads}}{\text{Number of open triads}}$$

Compute the transitivity of the network:

Number of closed triads: 2



Clustering Coefficient

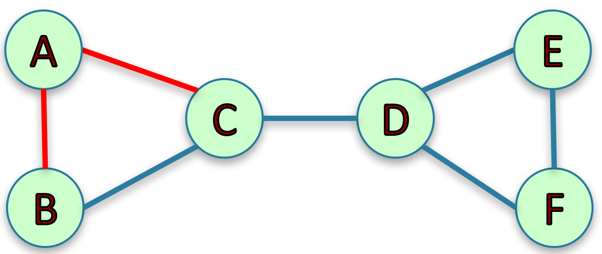
Global clustering coefficient (GCC) of a network:

$$\textbf{Transitivity} = \frac{3 * \text{Number of closed triads}}{\text{Number of open triads}}$$

Compute the transitivity of the network:

Number of closed triads: 2

Number of open triads: 1



Clustering Coefficient

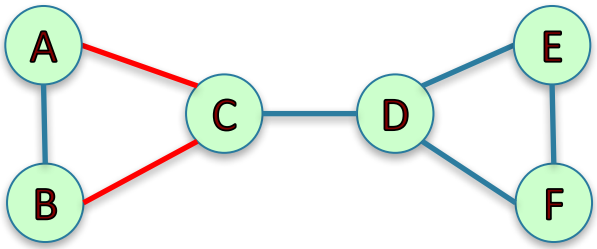
Global clustering coefficient (GCC) of a network:

$$\textbf{Transitivity} = \frac{3 * \text{Number of closed triads}}{\text{Number of open triads}}$$

Compute the transitivity of the network:

Number of closed triads: 2

Number of open triads: 2



Clustering Coefficient

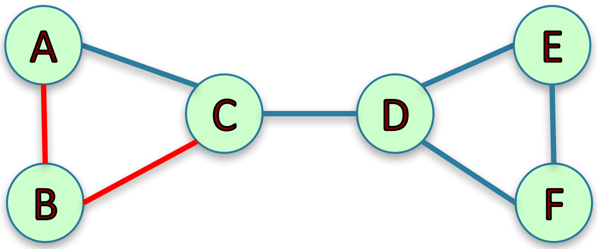
Global clustering coefficient (GCC) of a network:

$$\textbf{Transitivity} = \frac{3 * \text{Number of closed triads}}{\text{Number of open triads}}$$

Compute the transitivity of the network:

Number of closed triads: 2

Number of open triads: 3



Clustering Coefficient

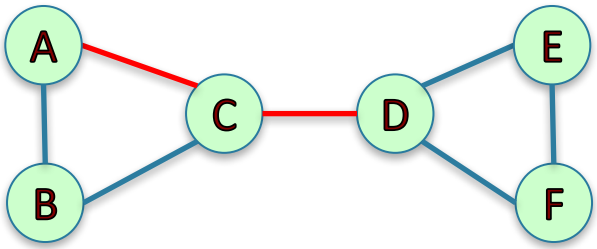
Global clustering coefficient (GCC) of a network:

$$\textbf{Transitivity} = \frac{3 * \text{Number of closed triads}}{\text{Number of open triads}}$$

Compute the transitivity of the network:

Number of closed triads: 2

Number of open triads: 4



Clustering Coefficient

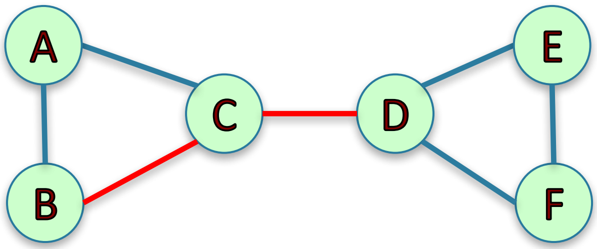
Global clustering coefficient (GCC) of a network:

$$\textbf{Transitivity} = \frac{3*\text{Number of closed triads}}{\text{Number of open triads}}$$

Compute the transitivity of the network:

Number of closed triads: 2

Number of open triads: 5



Clustering Coefficient

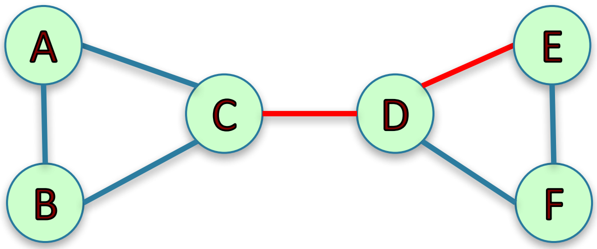
Global clustering coefficient (GCC) of a network:

$$\textbf{Transitivity} = \frac{3 * \text{Number of closed triads}}{\text{Number of open triads}}$$

Compute the transitivity of the network:

Number of closed triads: 2

Number of open triads: 6



Clustering Coefficient

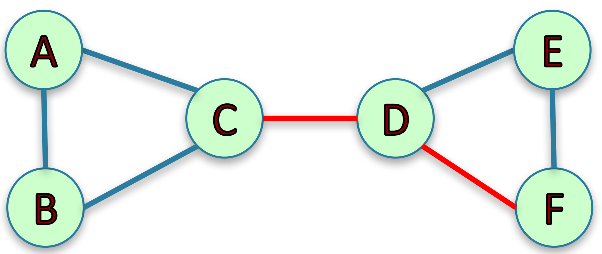
Global clustering coefficient (GCC) of a network:

$$\textbf{Transitivity} = \frac{3 * \text{Number of closed triads}}{\text{Number of open triads}}$$

Compute the transitivity of the network:

Number of closed triads: 2

Number of open triads: 7



Clustering Coefficient

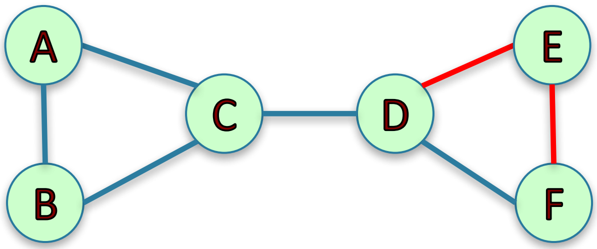
Global clustering coefficient (GCC) of a network:

$$\textbf{Transitivity} = \frac{3*\text{Number of closed triads}}{\text{Number of open triads}}$$

Compute the transitivity of the network:

Number of closed triads: 2

Number of open triads: 8



Clustering Coefficient

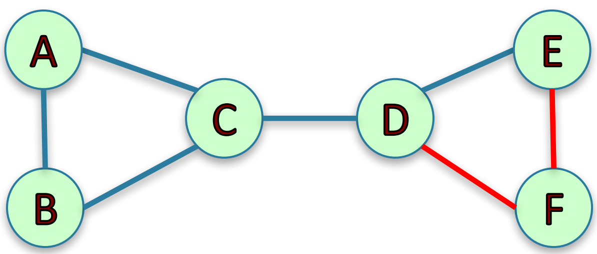
Global clustering coefficient (GCC) of a network:

$$\textbf{Transitivity} = \frac{3*\text{Number of closed triads}}{\text{Number of open triads}}$$

Compute the transitivity of the network:

Number of closed triads: 2

Number of open triads: 9



Clustering Coefficient

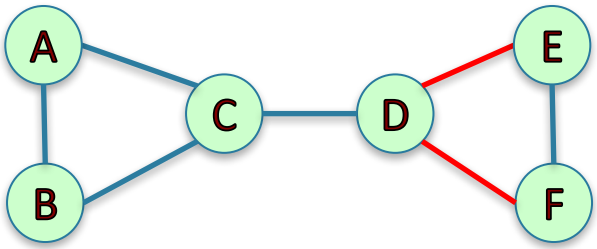
Global clustering coefficient (GCC) of a network:

$$\textbf{Transitivity} = \frac{3 * \text{Number of closed triads}}{\text{Number of open triads}}$$

Compute the transitivity of the network:

Number of closed triads: 2

Number of open triads: 10



$$\text{Transitivity} = \frac{3 * 2}{10} = \frac{3}{5}$$

Clustering Coefficient

Global clustering coefficient (GCC) of a network:

Definition 2: Mean Local Clustering Coefficient (MLCC) over all nodes

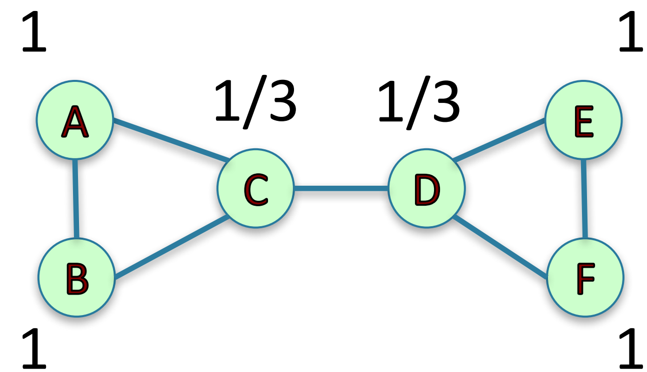
Compute the MLCC of the network:

First, compute LCC of all nodes

Now take the mean

MLCC = Mean of {1, 1, 1/3, 1/3, 1, 1}

$$\text{MLCC}=\frac{7}{9}$$



Clustering Coefficient

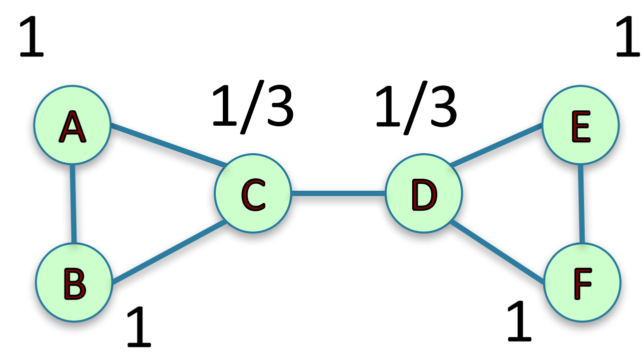
Global clustering coefficient (GCC) of a network:

$$\textbf{Transitivity} = \frac{3 * \text{Number of closed triads}}{\text{Number of open triads}}$$

MLCC = Mean Local Clustering Coefficient (MLCC) over all nodes.

$$\text{Transitivity} = 3/5 = 0.6$$

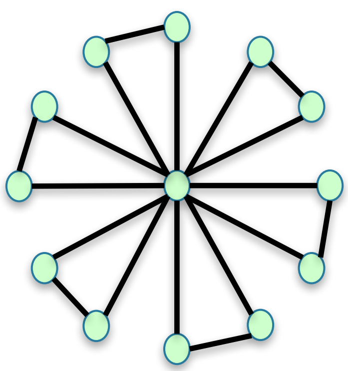
$$\text{MLCC} = 7/9 = 0.77$$



- Transitivity and MLCC can be very different on the same graph.
- MLCC weights every node equally
- Transitivity gives higher weight to nodes with many connections

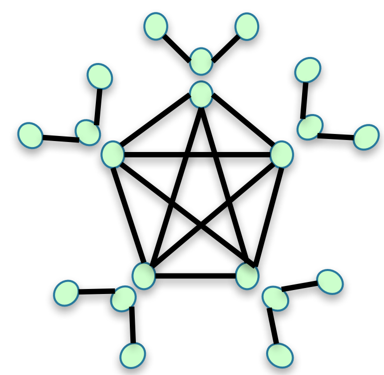
Transitivity vs. Average Clustering Coefficient

Both measure the tendency for edges to form triangles.
Transitivity weights nodes with large degree higher.



- Most nodes have high LCC
- The high degree node has low LCC

Ave. clustering coeff. = 0.93
Transitivity = 0.23



- Most nodes have low LCC
- High degree nodes have high LCC

Ave. clustering coeff. = 0.25
Transitivity = 0.86