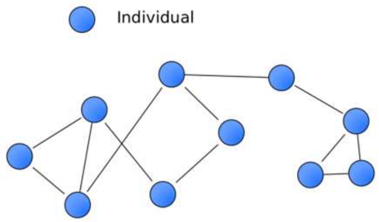
Strong and Weak Ties

We want to know

- How information flows through a social network



- How different <u>nodes</u> can play <u>structurally distinct roles</u> in this process

- How these <u>structural considerations</u> shape the <u>evolution</u> of the network itself over time

Networks structure

- Structure : composed of nodes and edges
  - They form groups of friends
  - There are close friends and acquaintances
  - How to bridge the local and the global ?

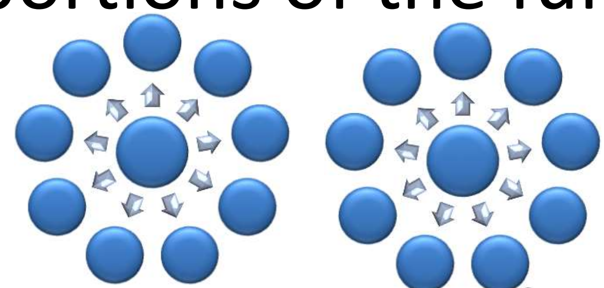
- Aim : to offer explanations for how simple processes at the level of individual nodes and links can have complex effects that ripple through a population as a whole

Network Structure

- Network structure evolves over time
- We take snapshot of nodes and edges at a particular time and form static structures.
- E.g. you are invited to be a friend of the friends of your friends.

Linking two perspectives on distant friendships

- Structural : the way these friendships span different portions of the full network



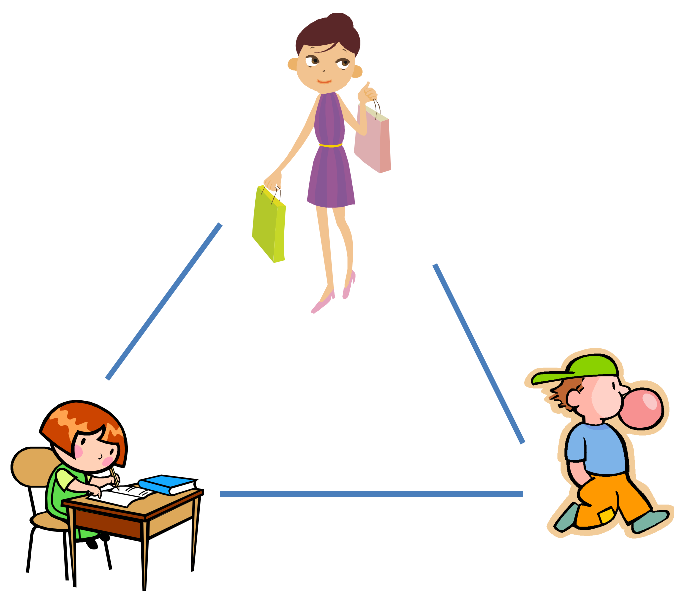
- The local consequences that following from a friendship between two people (strong vs weak ties)

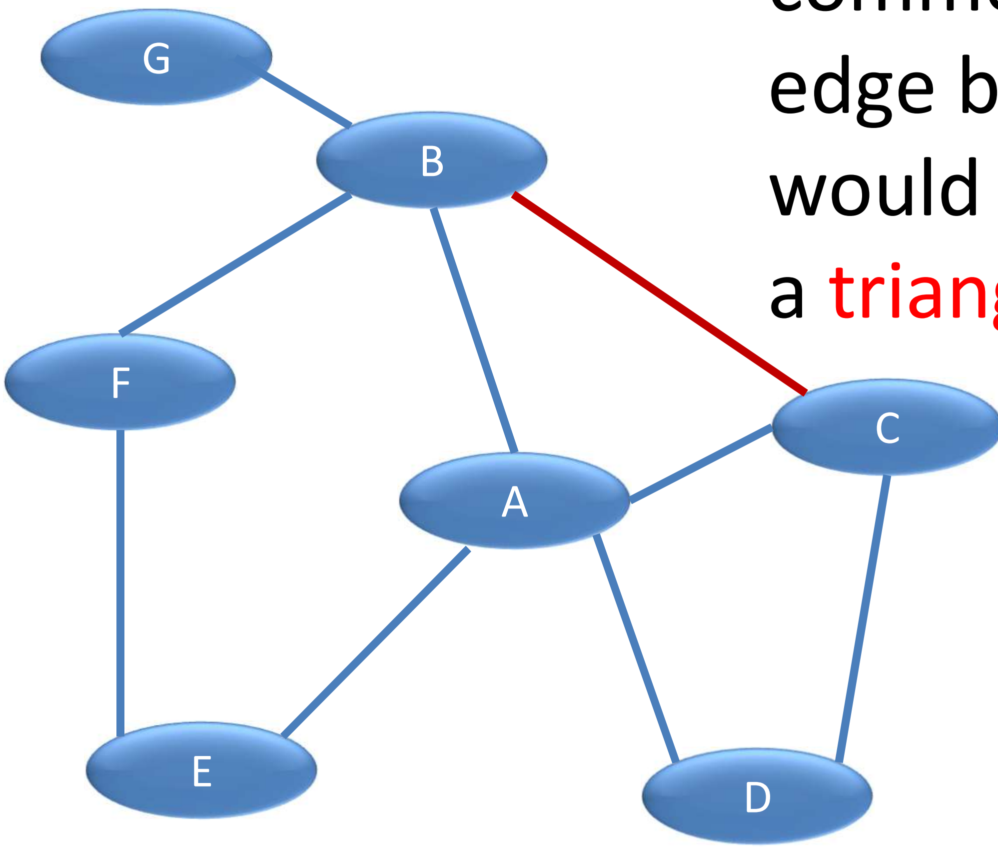
Outlines

1. Network structure : Triadic Closure
2. Strength of Weak Ties
3. Tie Strength and Network Structure : examples
4. Closure and Structure Holes
5. Graph Partitioning

1. Triadic closure

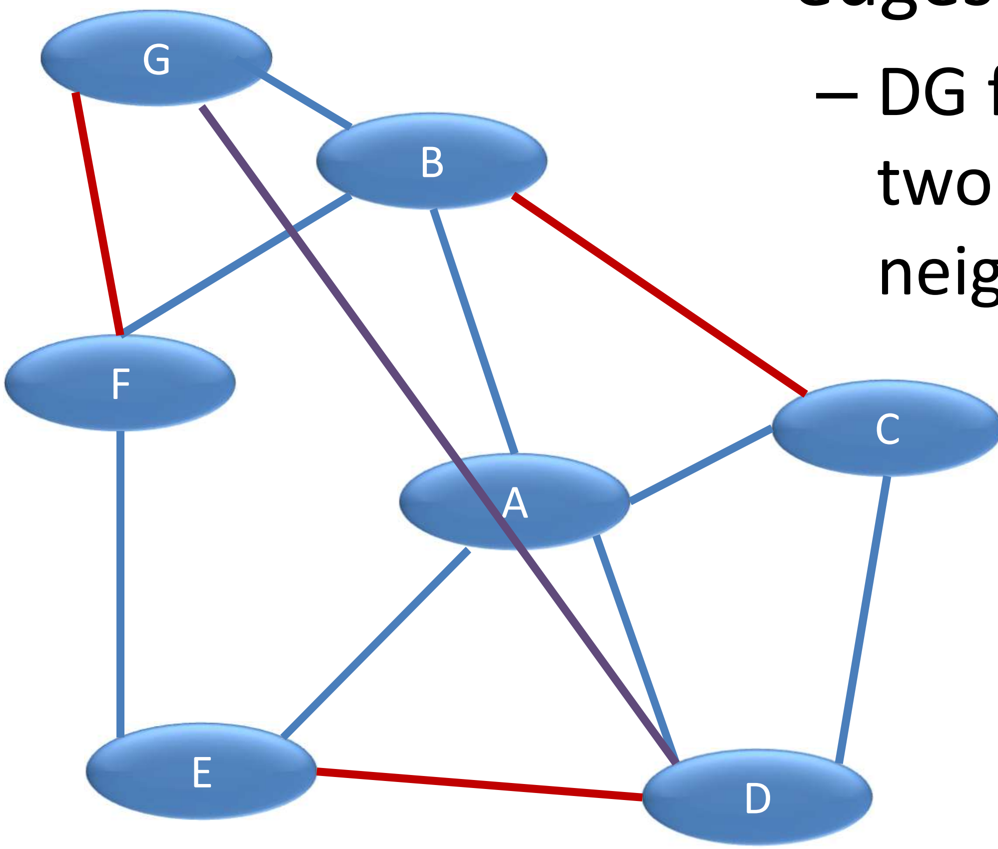
- If two people in a social network have a friend in common, then there is an increased likelihood that they will become friends themselves at some point in the future.





- B and C have a friend A in common, it is likely that an edge between B and C would be formed and form a triangle.

- Eventually, a number of new edges would be formed
  - DG forms even though the two endpoints have no neighbors in common



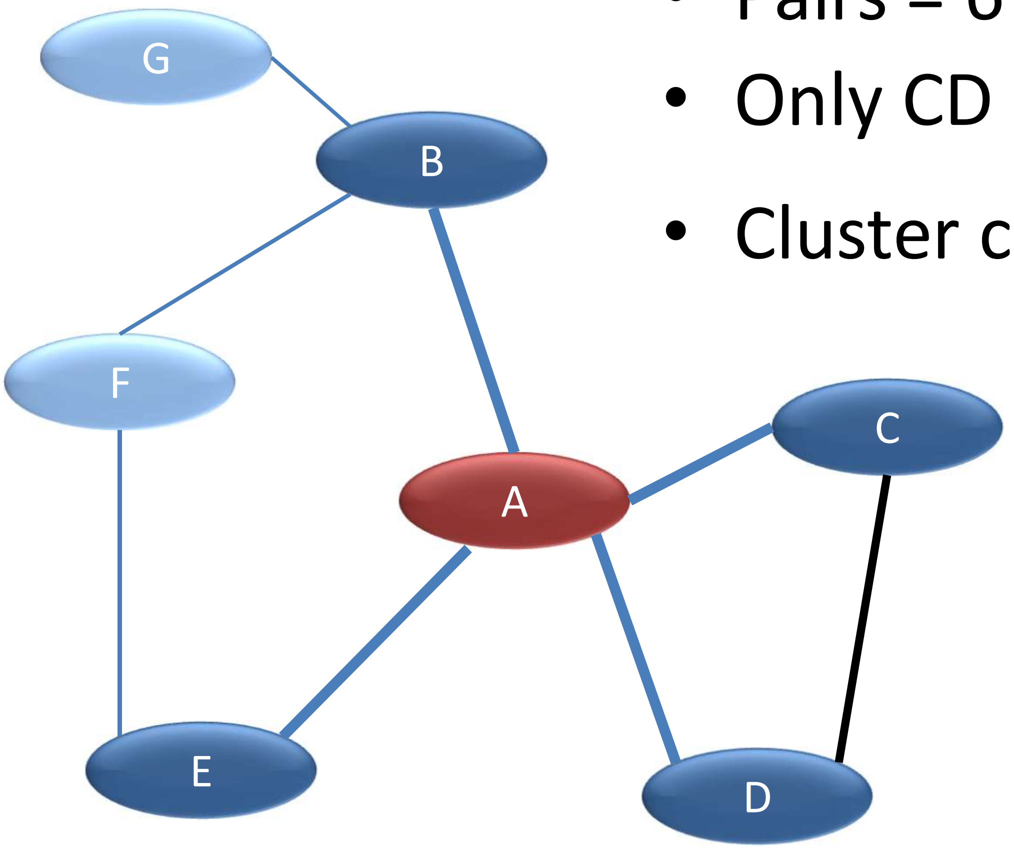
Cluster Coefficient

- The cluster coefficient of a node A is defined as the probability that two randomly selected friends of A are friends with each other.
- Cluster Coefficient (A) = CC(A)

$$= \frac{\text{pair of A}'\text{s friends that are connected}}{\text{total no. of pair of A}'\text{s friends}}$$

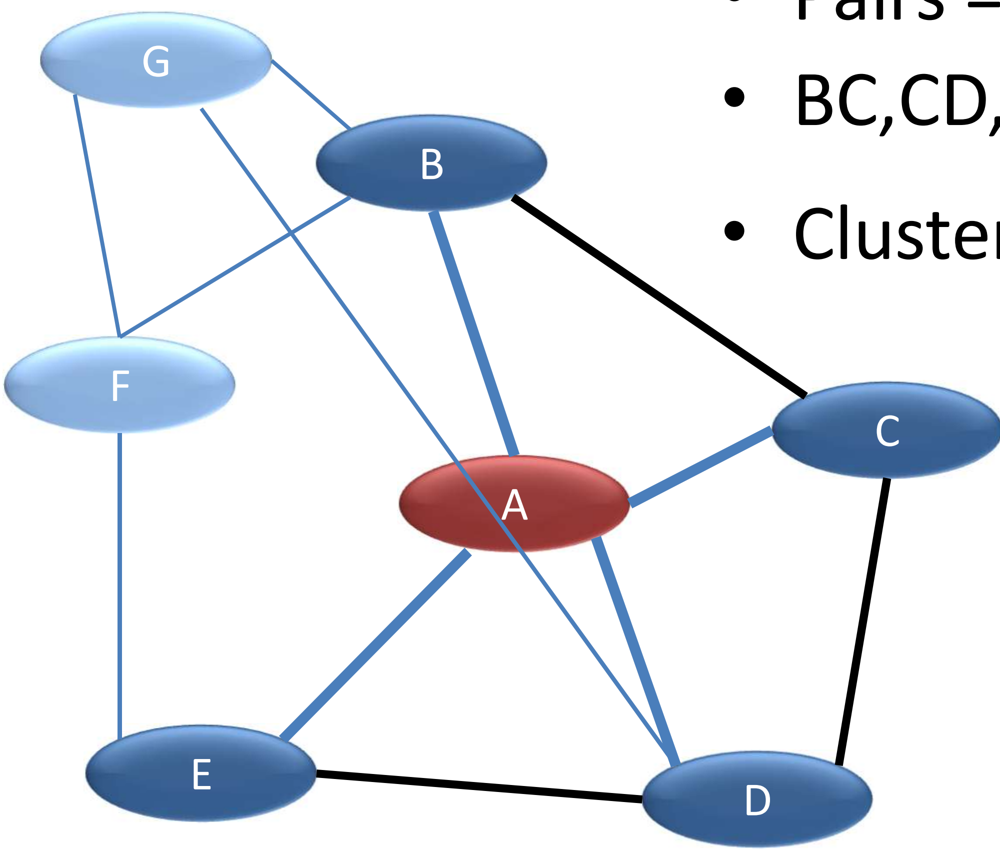
Fractions of your friends (in pairs) are friends themselves

Cluster Coefficient



- No. of friends of A = 4 (B,C,D,E)
- Pairs = 6 (BC,BD,BE,CD,CE,DE)
- Only CD are connected
- Cluster coefficient (A) $= \frac{1}{6}$

Cluster Coefficient



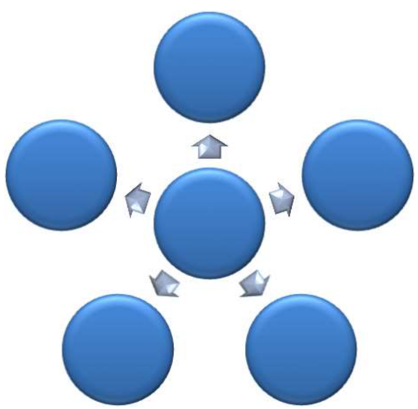
- No. of friends of A = 4 (B,C,D,E)
- Pairs = 6 (BC,BD,BE,CD,CE,DE)
- BC,CD,DE are connected
- Cluster coefficient (A) = $\frac{3}{6} = \frac{1}{2}$

Cluster Coefficient

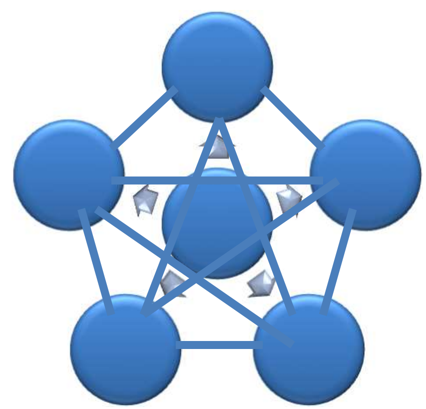
- Cluster Coefficient (A) = CC(A)

$$= \frac{\text{pair of A}'\text{s friends that are connected}}{\text{total no. of pair of A}'\text{s friends}}$$

CC(A) = 0

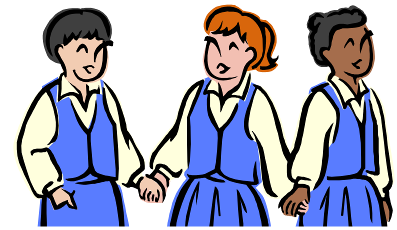


CC(A) = 1



Reasons for Triadic closure

When *B* and *C* have a common friend *A*



- More opportunity for *B* and *C* to meet
- A base for *B* and *C* to trust each other
- A source of latent stress for *A* if *B* and *C* are not friends
  – *Bearman and Moody* found that teenage girls with low clustering coefficient are more likely to contemplate suicide.

Outlines

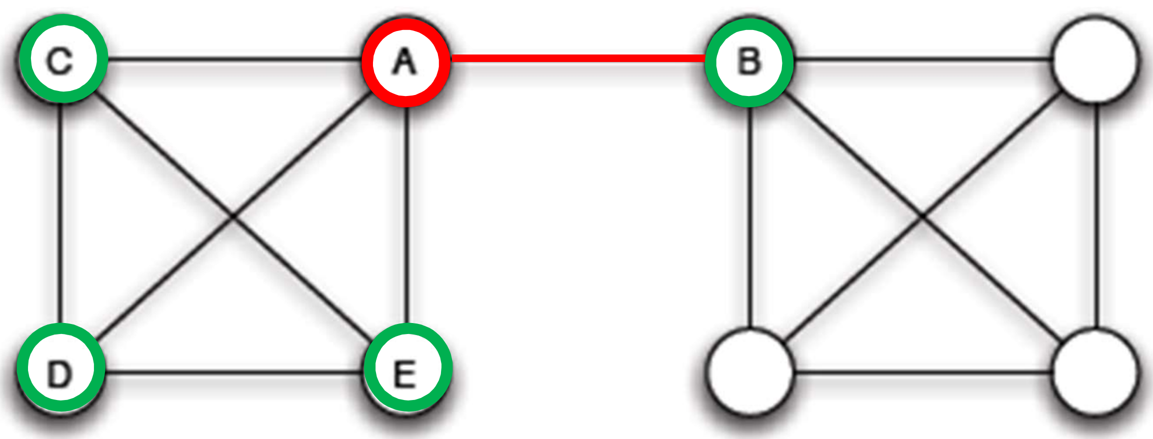
1. Network structure : Triadic Closure
2. Strength of Weak Ties
3. Tie Strength and Network Structure : examples
4. Closure and Structure Holes
5. Graph Partitioning

2. Strength of Weak Ties

- How does triadic closure influence the network ?
- Two more concepts
  - Links between two tightly-knit groups of friends : Bridges and Local bridges
  - Strength of the edges : strong ties and weak ties

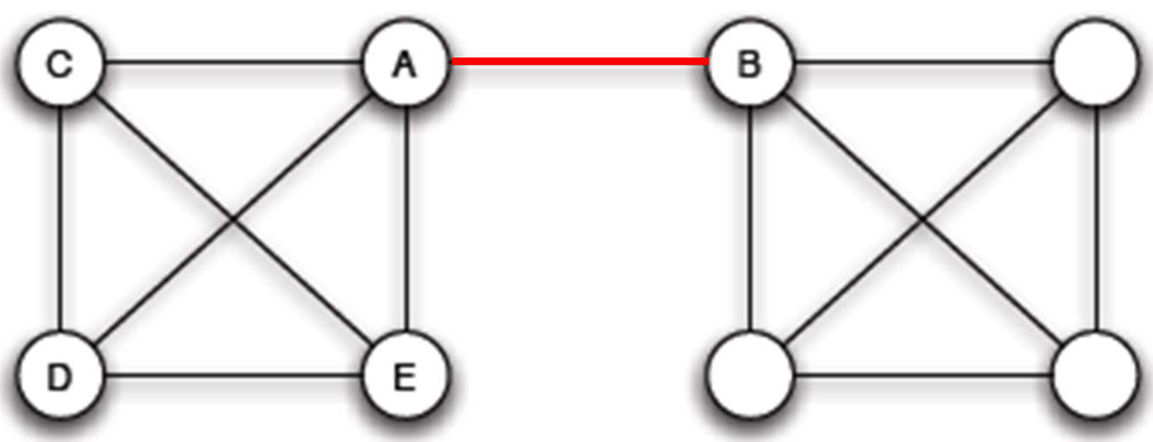
Bridge

- A has four friends
  - C, D, E come from a tightly-knit group of friends
    - A, C, D, E share similar opinions and source of information
  - B comes from a different part of the network
    - B offers A access to new information



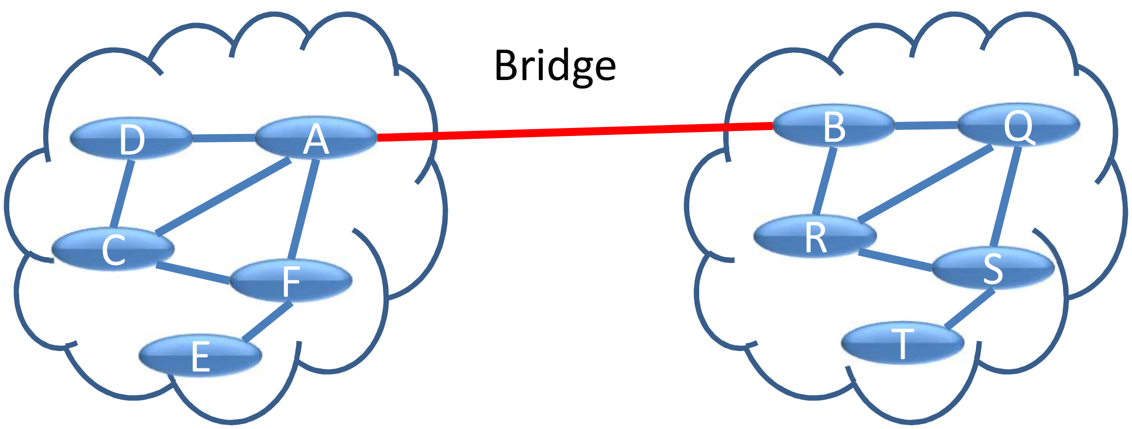
Bridge

- Bridge : the edge joining two nodes if deleting that edge would cause the two nodes in <u>two different components</u>.
- E.g. the edge A-B



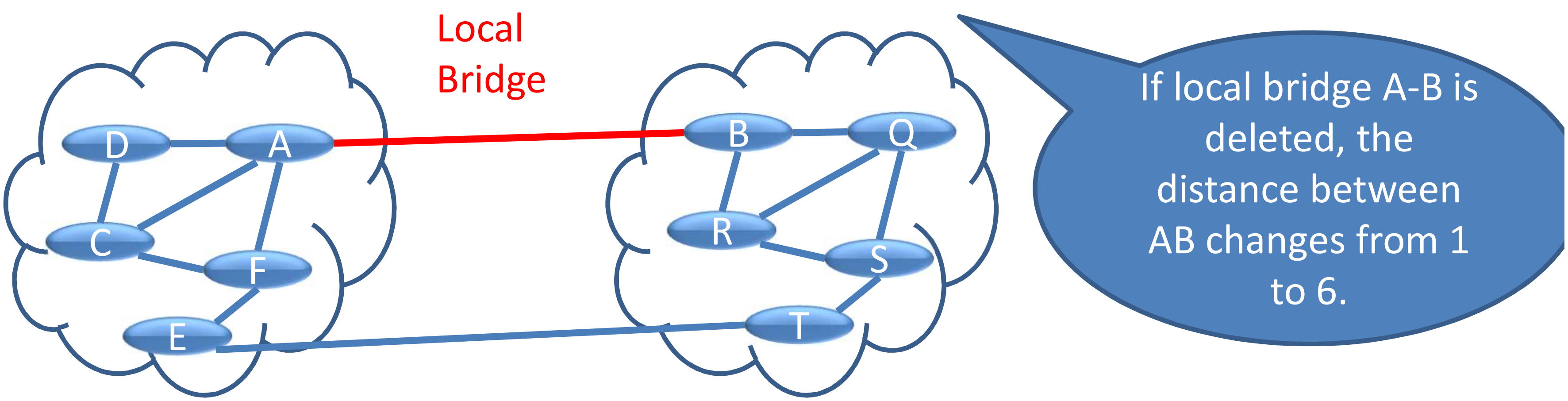
Local Bridge

- Bridges are rare in real social networks due to the giant components and small-world properties.
- Local bridge : the edge joining nodes A and B, if its endpoints A and B have no friends in common.



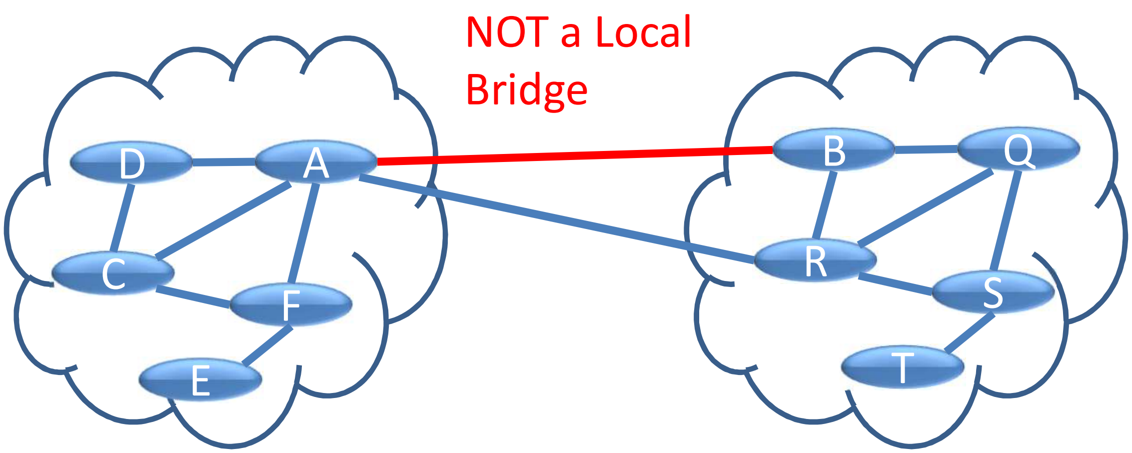
Local Bridge

- Bridges are rare in real social networks due to the giant components and small-world properties.
- Local bridge : the edge joining nodes A and B, if its endpoints A and B have no friends in common.
  - If deleting the edge would increase the distance between A and B to a value strictly more than two.

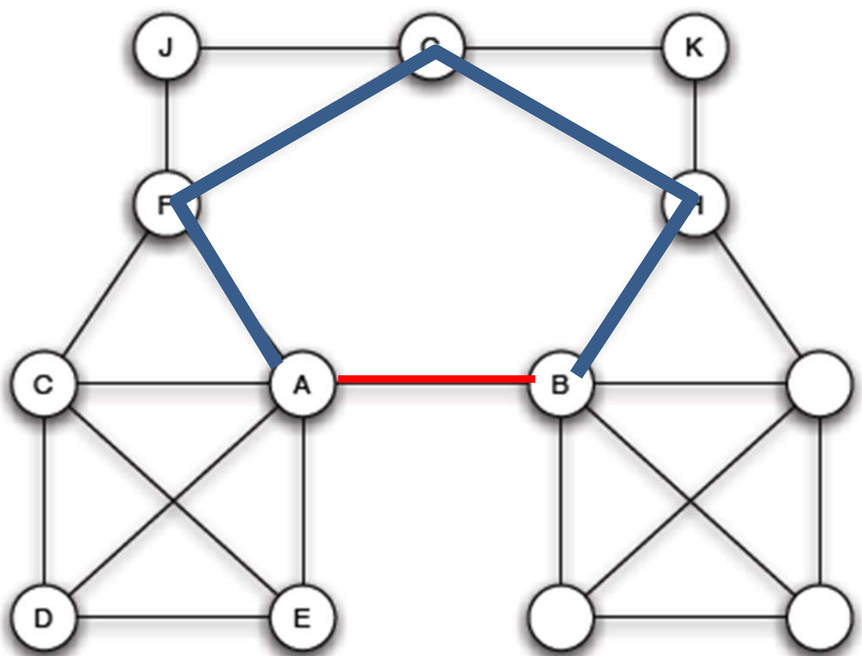


Local Bridge

- Bridges are rare in real social networks due to the giant components and small-world properties.
- Local bridge : the edge joining nodes A and B, if its endpoints A and B have no friends in common.
  - If deleting the edge would increase the distance between A and B to a value strictly more than two.



- The span of a local bridge = the distance between the end points if that edge were deleted.
- The span of a local bridge must be more than 2.



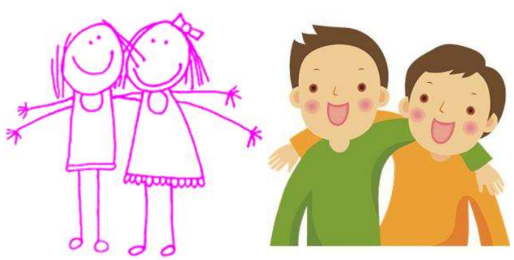
Span(A-B edge) = 4

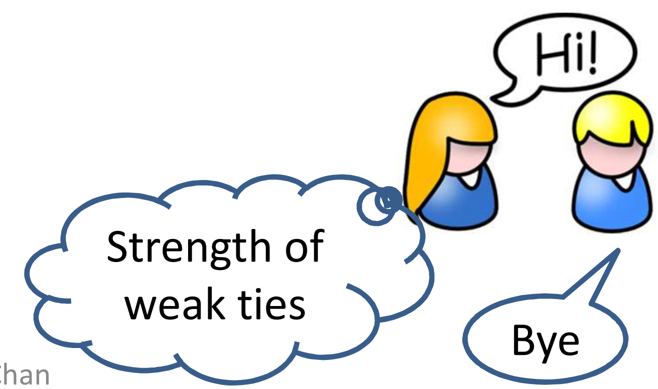
Local Bridges and triadic closure

- A local bridge cannot be part of a triangle by definition.
- Local bridges (especially those with large span) provide their endpoints with sources of information that would be far apart.
- Job seeking example : information of new jobs come from acquaintances.

Job hunting example by Mark Granovetter

- Mark Granovetter found in the late 60s that
  - many people learned information leading to their new jobs through personal contacts by interviewing people who had recently changed employers.
  - But these personal contacts were acquaintances rather than close friends.





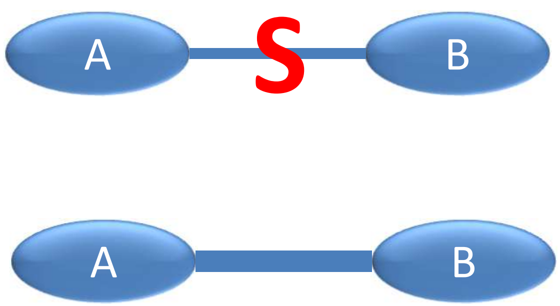
- What is the relationship between acquaintances and local bridges ?

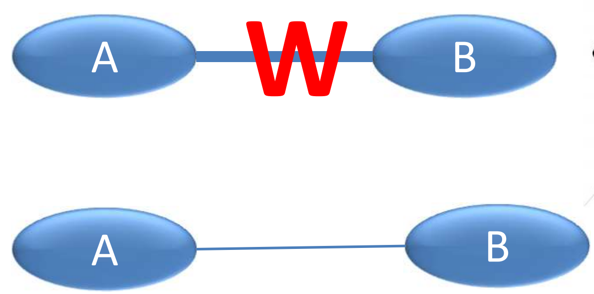
Strong/Weak ties and triadic closure

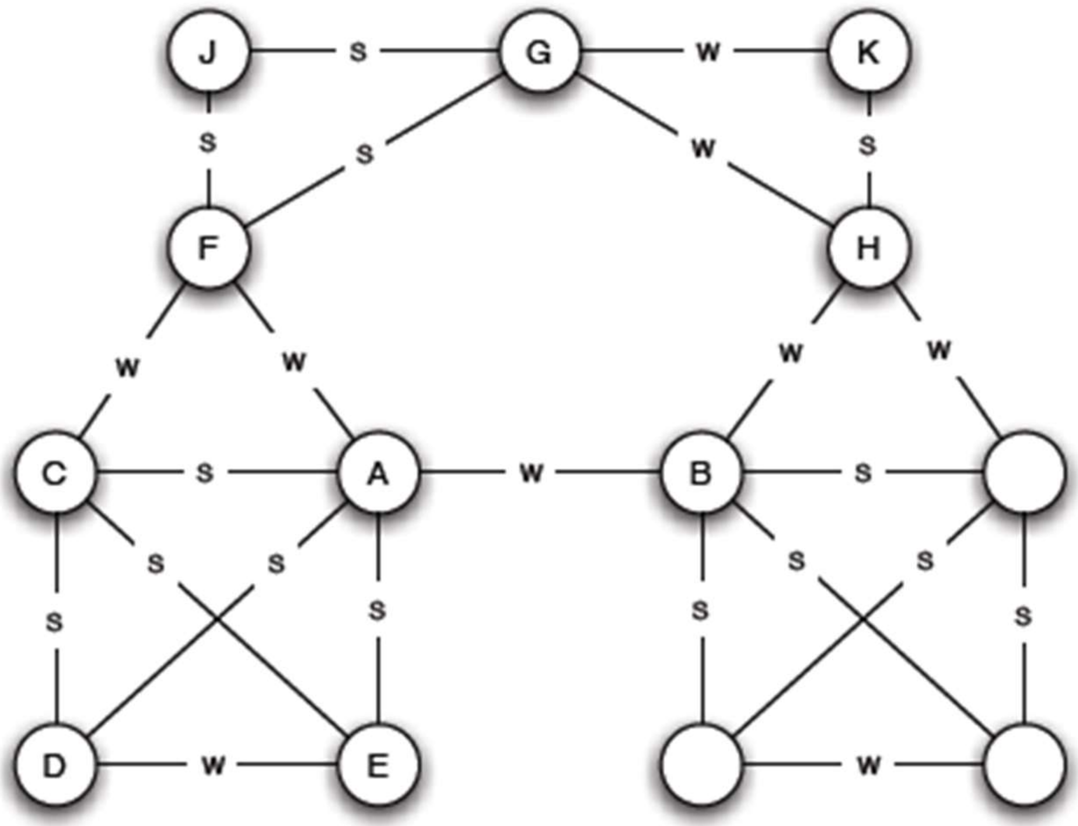
- The consequences that following from a friendship between two people (strong vs weak ties)

Close friend
Strong tie

Acquaintance
Weak tie

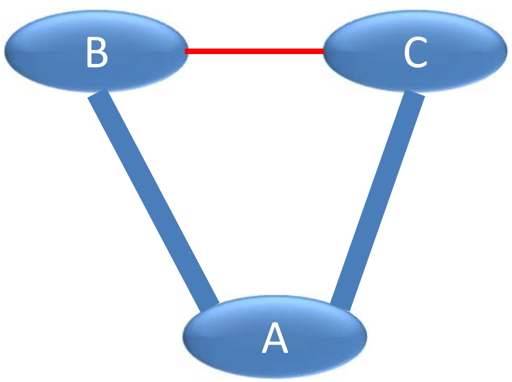





J
S
G
W
K
S
S
W
S
F
H
W
W
W
W
C
S
A
W
B
S
S
S
S
S
S
S
S
S
D
W
E
W

Strong/Weak ties and triadic closure

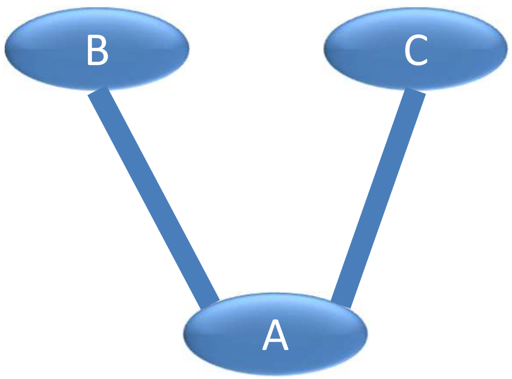
- If a node A has edges to nodes B and C, then B-C edge is especially likely to form if A’s edge to B and C are both strong ties.

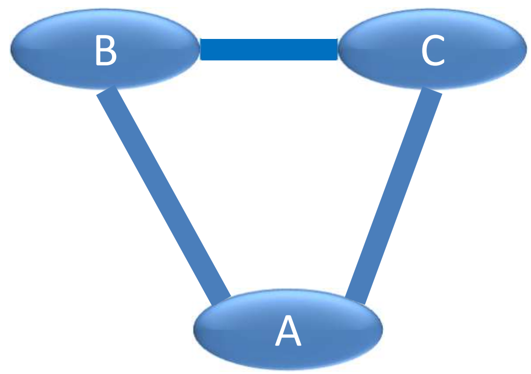


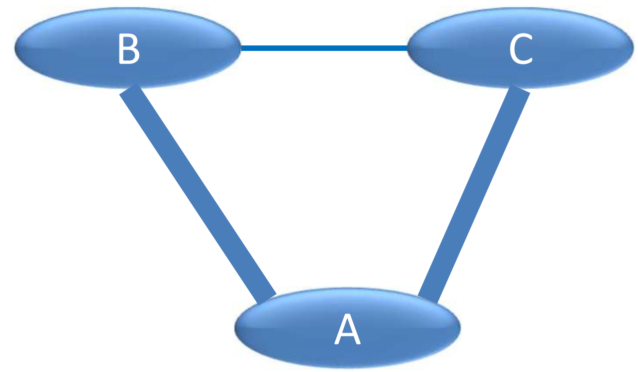
Strong Triadic Closure Property

- We say that a *node A violates the Strong Triadic Closure Property* if it has strong ties to two other nodes B and C, and there is no edge at all (either a strong or weak tie) between B and C.
- We say that a node A *satisfies* the Strong Triadic Closure Property if it does not violate it.

Strong Triadic Closure Property



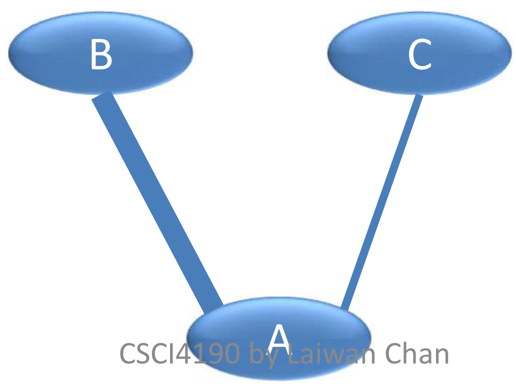


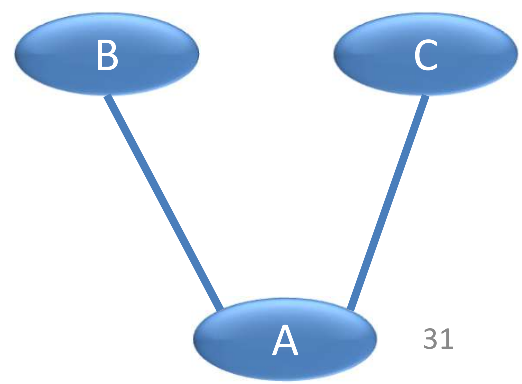


Violate !

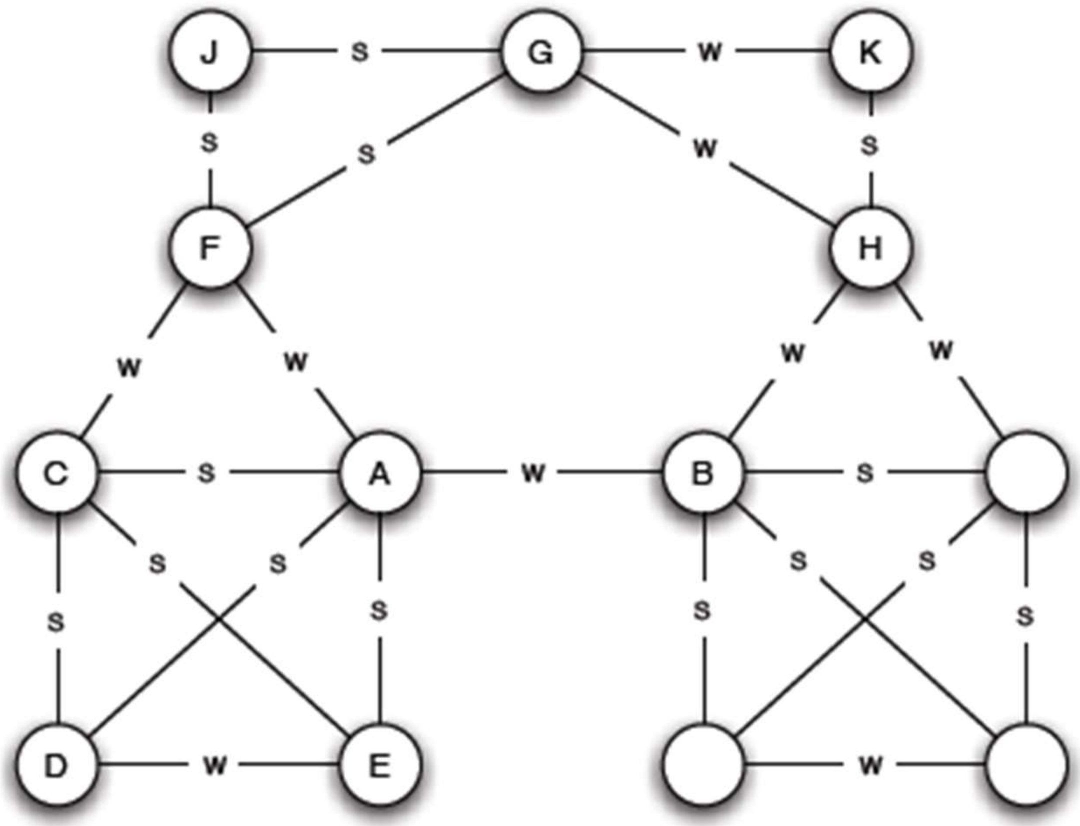
Satisfy !

Satisfy !

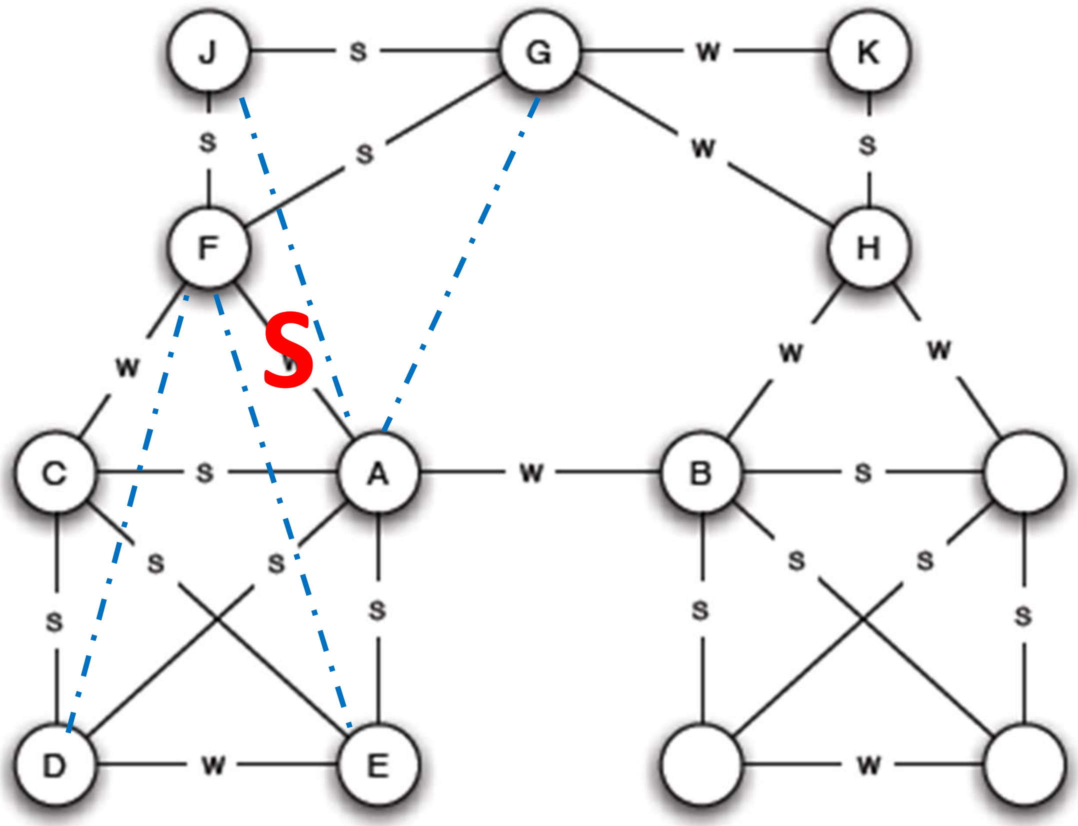




Satisfy !

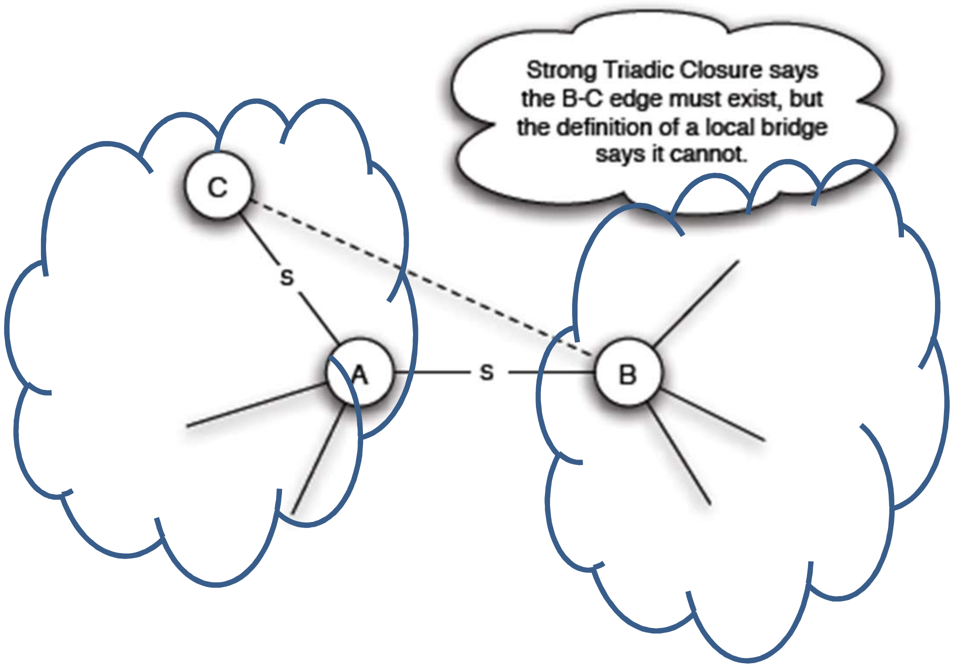


Violate !



Local Bridges and Weak Ties

- If a node A in a network satisfies the Strong Triadic Closure Property and is involved in at least two strong ties, then any local bridge it is involved in must be a weak tie.



The assumption of Strong Triadic Closure Property is too strong !

Outlines

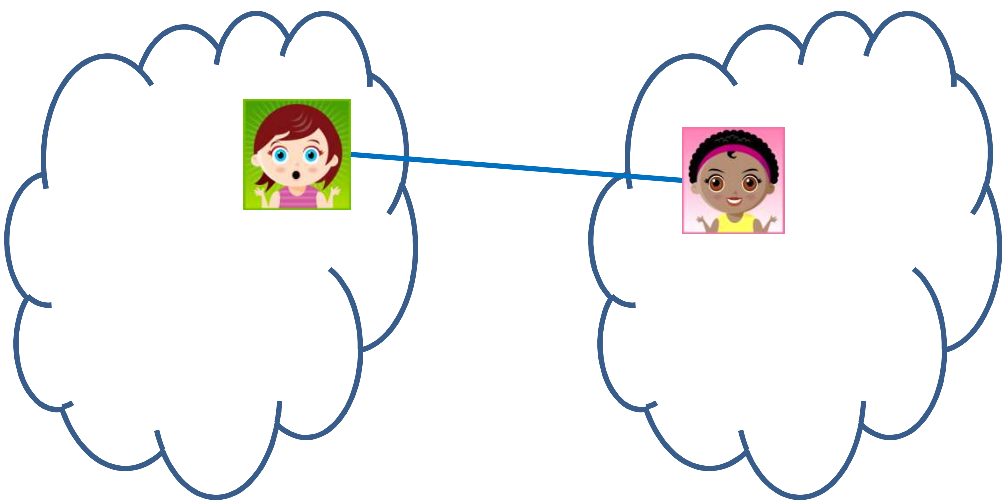
1. Network structure : Triadic Closure
2. Strength of Weak Ties
3. Tie Strength and Network Structure : examples
4. Closure and Structure Holes
5. Graph Partitioning

3. Tie strength and network structure in large scale data

a) Job hunting example by Mark Granovetter
b) Who-talks-to-whom example by Onnela
c) Facebook data analysis by Cameron Marlow
d) Twitter data analysis by Huberman, Romero, and Wu

a. Job hunting example by Mark Granovetter

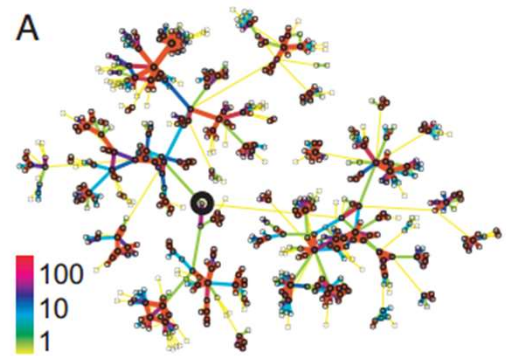
- New jobs are often rooted in contact with distant acquaintances.
- New sources of information
- Local bridge in the social network
- Weak ties



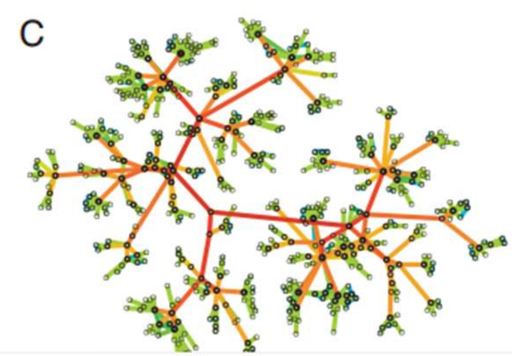
b. Who-talks-to-whom

- Study the relationship between tie strength and network structure
- Who-talks-to-whom by Onnela et al. published in 2007
  - Call records from a cell-phone provider
  - 20 % of a national population
  - Personal communication rather than business purpose

call duration



Betweenness
(high betweenness for links connecting communities and low values for links within the communities)

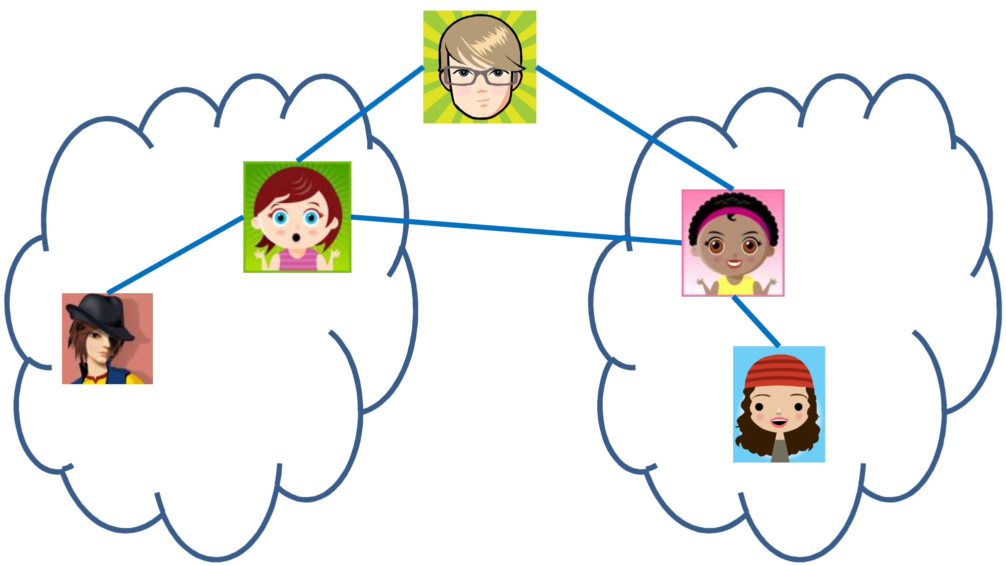


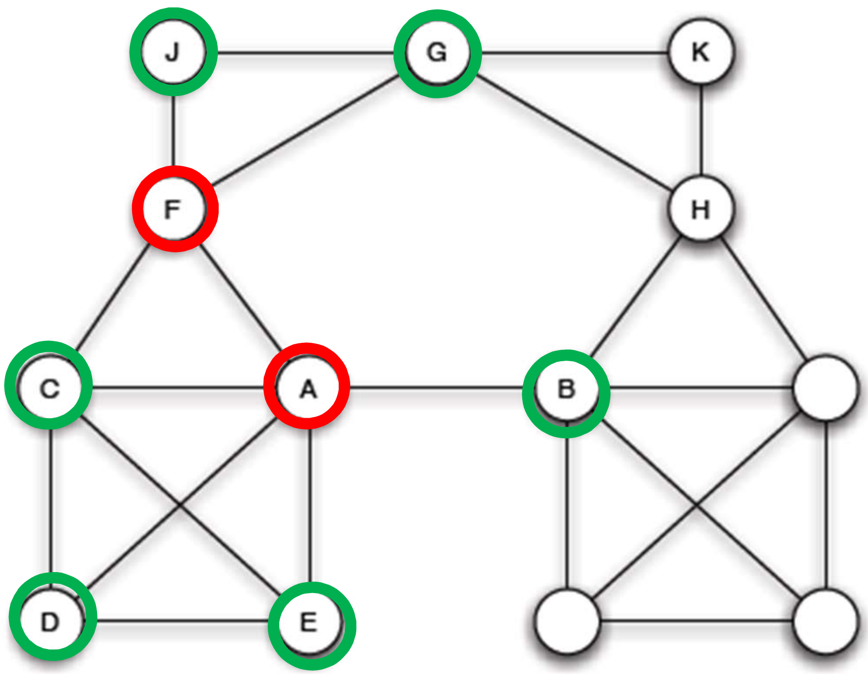
Who-talks-to-whom

- *Nodes* : cell-phone users
- *Edge* : if two users call each other in both directions over an 18 week observation period.
- *Strength* of an edge : total number of minutes spent on phone calls between the two users
- Local bridges are rare and so use *neighborhood overlap* instead

- Neighborhood overlap of an edge connecting A and B

$$= \frac{no.\ of\ nodes\ who\ are\ neighbors\ of\ both\ A\ and\ B}{no.\ of\ nodes\ who\ are\ neighbors\ of\ at\ least\ one\ of\ A\ or\ B}$$



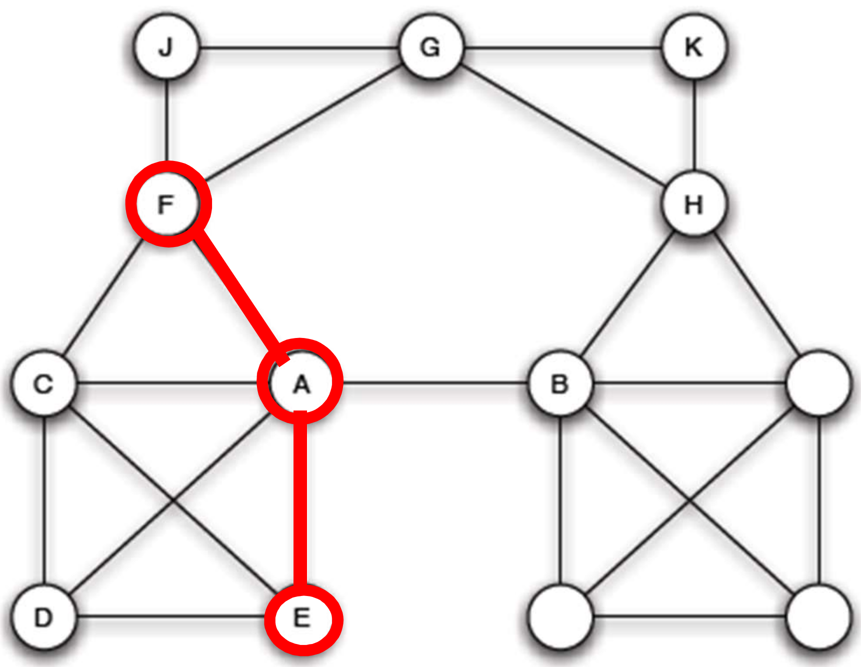


Neighbors of A : B, C, D, E, F
Neighbors of F : A, C, J, G

Total no. of A and F’s neighbors : 6
(B,C,D,E,J,G)
(exclude A and F)

No of common neighbors : 1 (C)

Neighborhood overlap = $\frac{1}{6}$



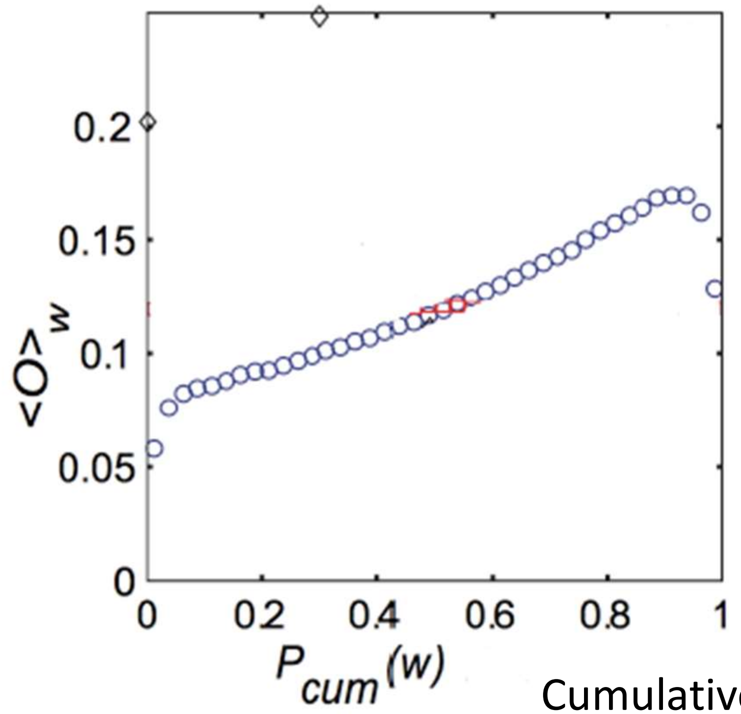
Neighbors of A : B, C, D, E, F
Neighbors of F : A, C, J, G
Total no. of A and F's neighbors : 6
(exclude A and F)
No of common neighbors : 1 (C)
Neighborhood overlap = O(A,F) = $\frac{1}{6}$

O(Local bridge) = 0
O(Almost Local bridge) = small

O(A,F) = $\frac{1}{6}$, O(A,E) = $\frac{1}{2}$
Edge A-F is much closer to being a local bridge than edge A-E

Tie strength and network structure

Neighborhood overlap



Cumulative tie strength

- Neighborhood overlap increases with the strength
- the stronger the tie between two users, the more their friends overlap
- a correlation that is valid for 95% of the links

Tie strength and network structure

- Local level : tie strength increases with neighborhood
- Global level : weak ties serve to link together different tightly-knit communities ?

Are weak ties linking different tightly-knit communities ?

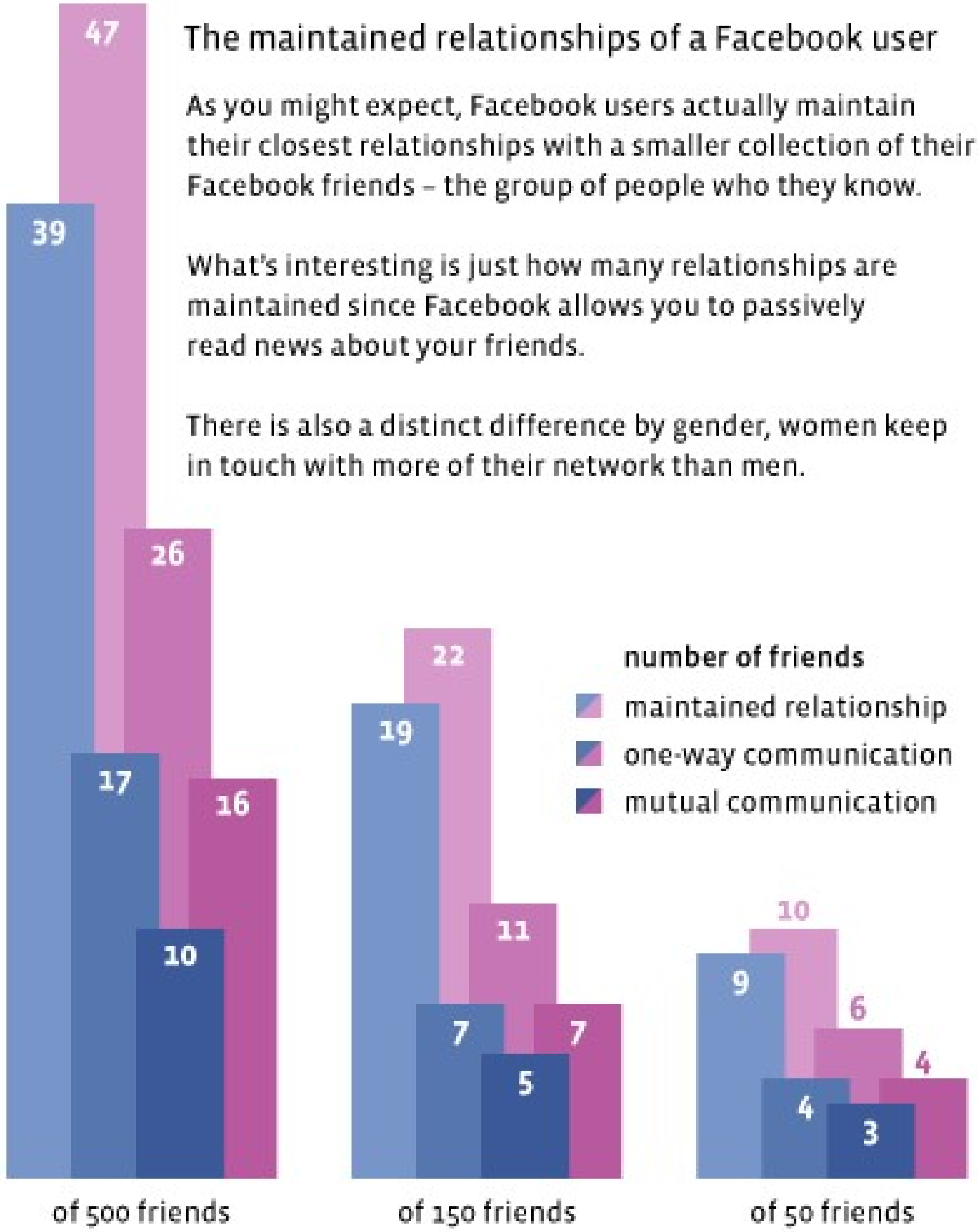
1. Delete edges one at a time, starting with strongest ties
   – the giant component shrank steadily
2. Deleted edges one at a time, starting with weakest ties
   – the giant component shrank more rapidly

Weak ties provide the more crucial connective structure
- *for holding together disparate communities and*
- *for keeping the global structure of the giant components intact.*

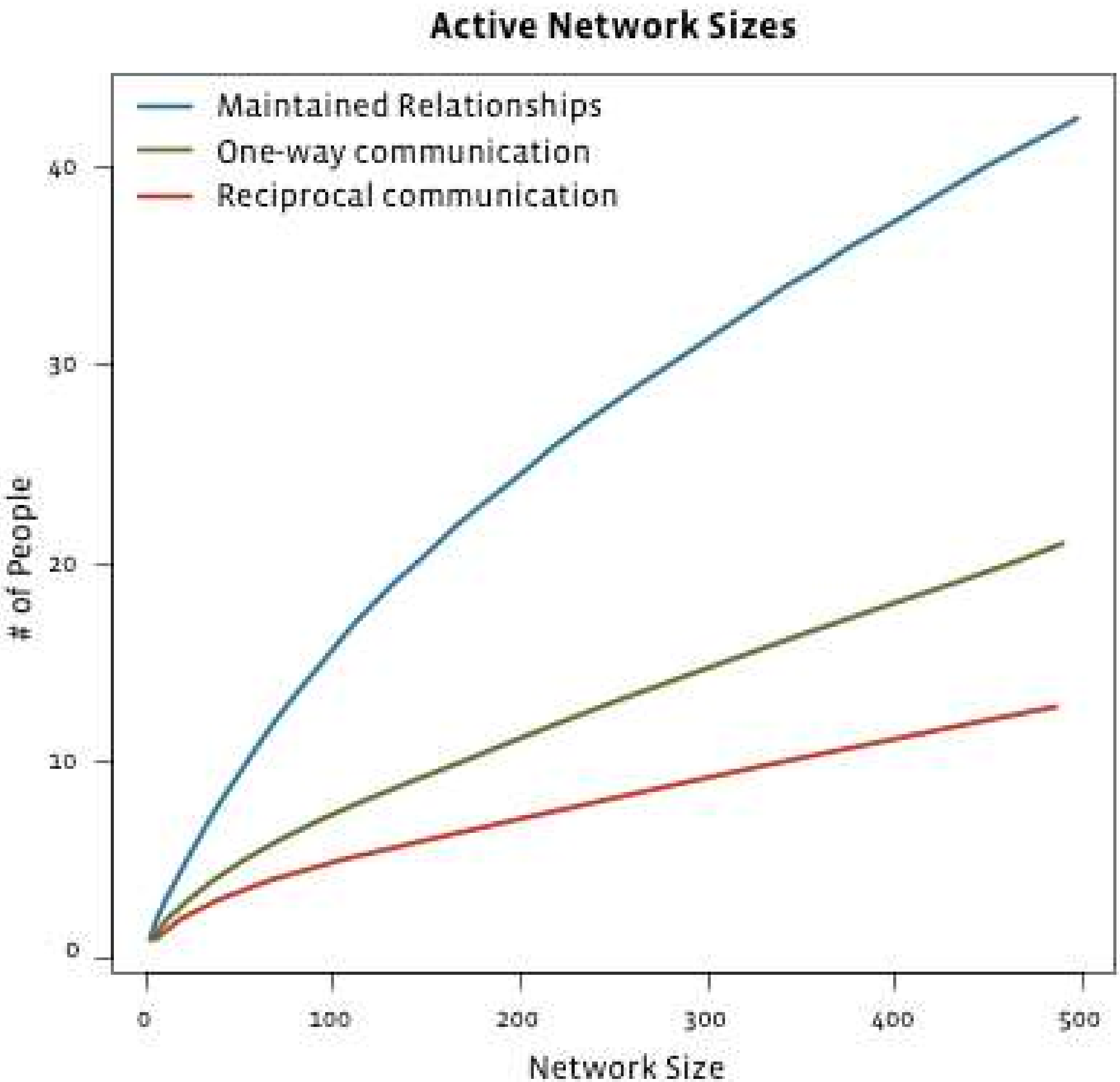
c. Tie strength on Facebook

- Cameron Marlow in his 2009 article "maintained relationships on Facebook" studied the types of relationships people maintain and the relative size of these groups.
  - They analyzed the friendship links reported in each user's profile, asking to what extent each link was actually used for social interaction.
  - Where are the strong ties among a user's friends?

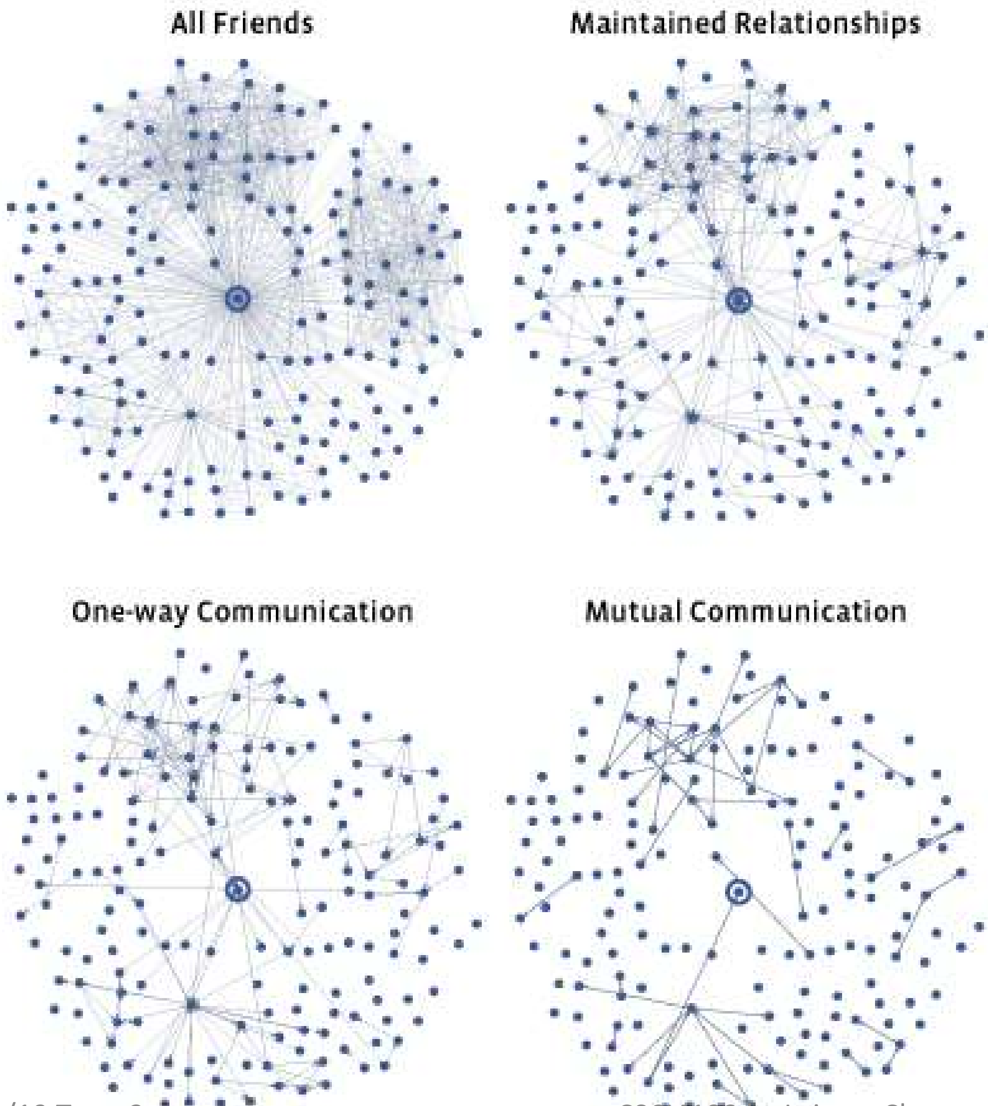
- Four categories of links
  - All friends
  - reciprocal (mutual) communication : if both users sent and received messages to each other during the observation period.
  - one-way communication : if the user sent one or more messages to the friend at the other end of the link (whether or not these messages were reciprocated).
  - a maintained relationship : if the user followed information about the friend at the other end of the link, whether or not actual communication took place;
    - “following information” here means either clicking on content via Facebook’s News Feed service or visiting the friend’s profile more than twice.



The maintained relationships of a Facebook user
As you might expect, Facebook users actually maintain their closest relationships with a smaller collection of their Facebook friends – the group of people who they know.
What's interesting is just how many relationships are maintained since Facebook allows you to passively read news about your friends.
There is also a distinct difference by gender, women keep in touch with more of their network than men.
number of friends
maintained relationship
one-way communication
mutual communication
39
47
17
26
10
16
19
22
7
11
5
7
9
10
4
6
3
4
of 500 friends
of 150 friends
of 50 friends


Active Network Sizes
Maintained Relationships
One-way communication
Reciprocal communication
# of People
0
10
20
30
40
0
100
200
300
400
500
Network Size

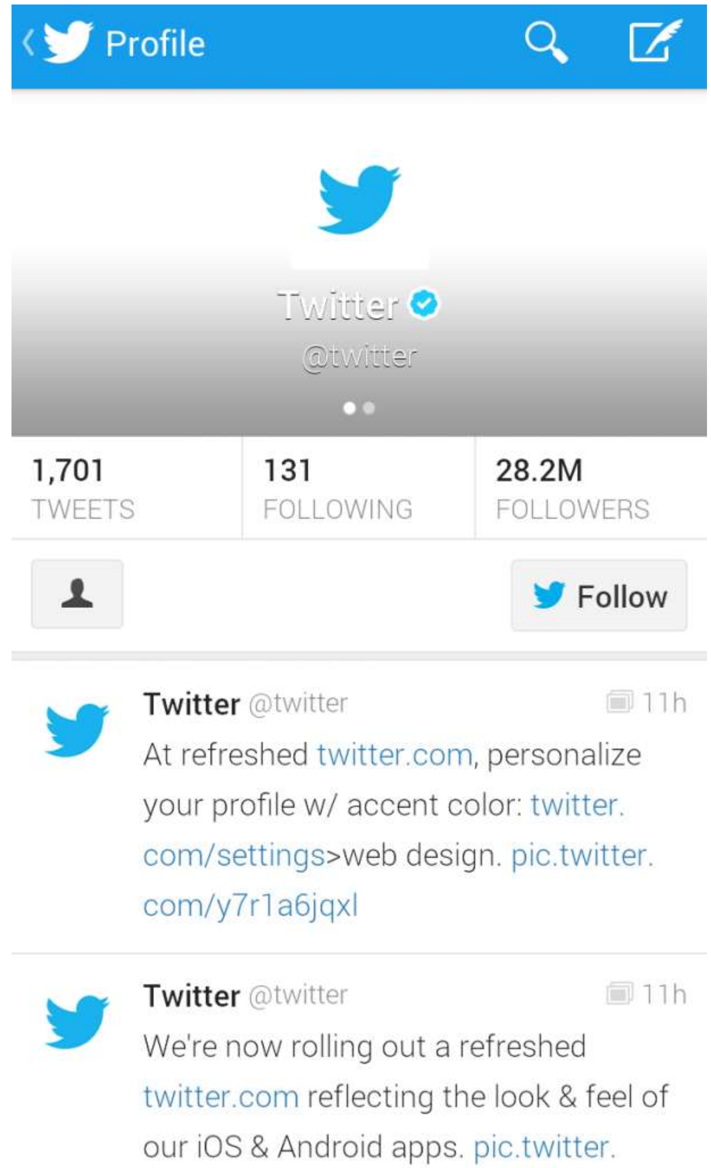
One's personal network



- Two clusters
  - One highly connected set of coworkers with frequent contacts
  - Another group without active contact.

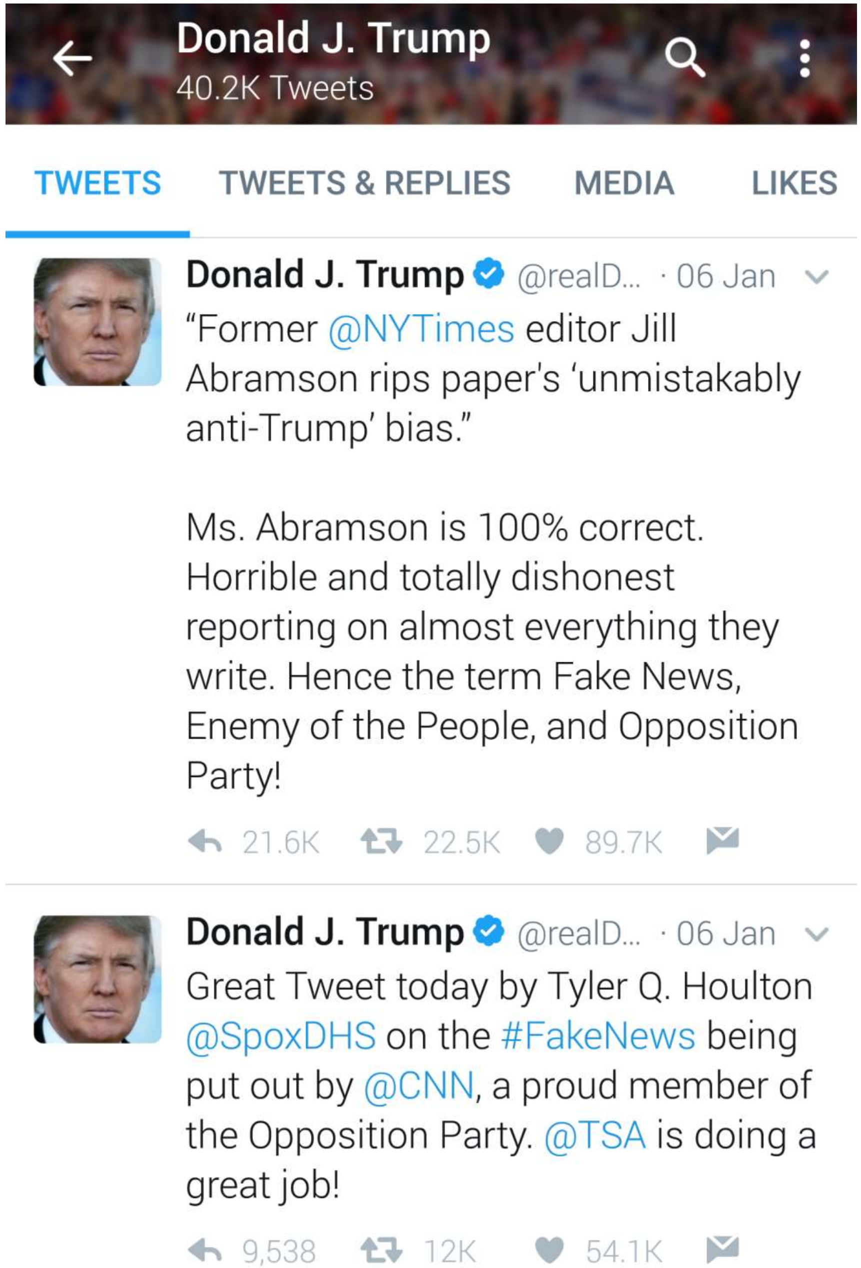
- A contrast between
  - Mutual communication
    - Strong ties
  - Passive engagement
    - Reading news via social media
    - News/events can propagate very quickly through this highly connected network

d. Tie strength on Twitter



- Twitter users
  - Followers (people who follow the person)
  - Followees (people followed by the person)
  - Friends (people who have received at least two direct messages)

- Twitter posts:
  - Direct public posts are used when a user aims her update to a specific person using an "@" symbol. Around 25.4 % of all posts are directed.
  - Indirect updates are used when the update is meant for anyone that cares to read it.

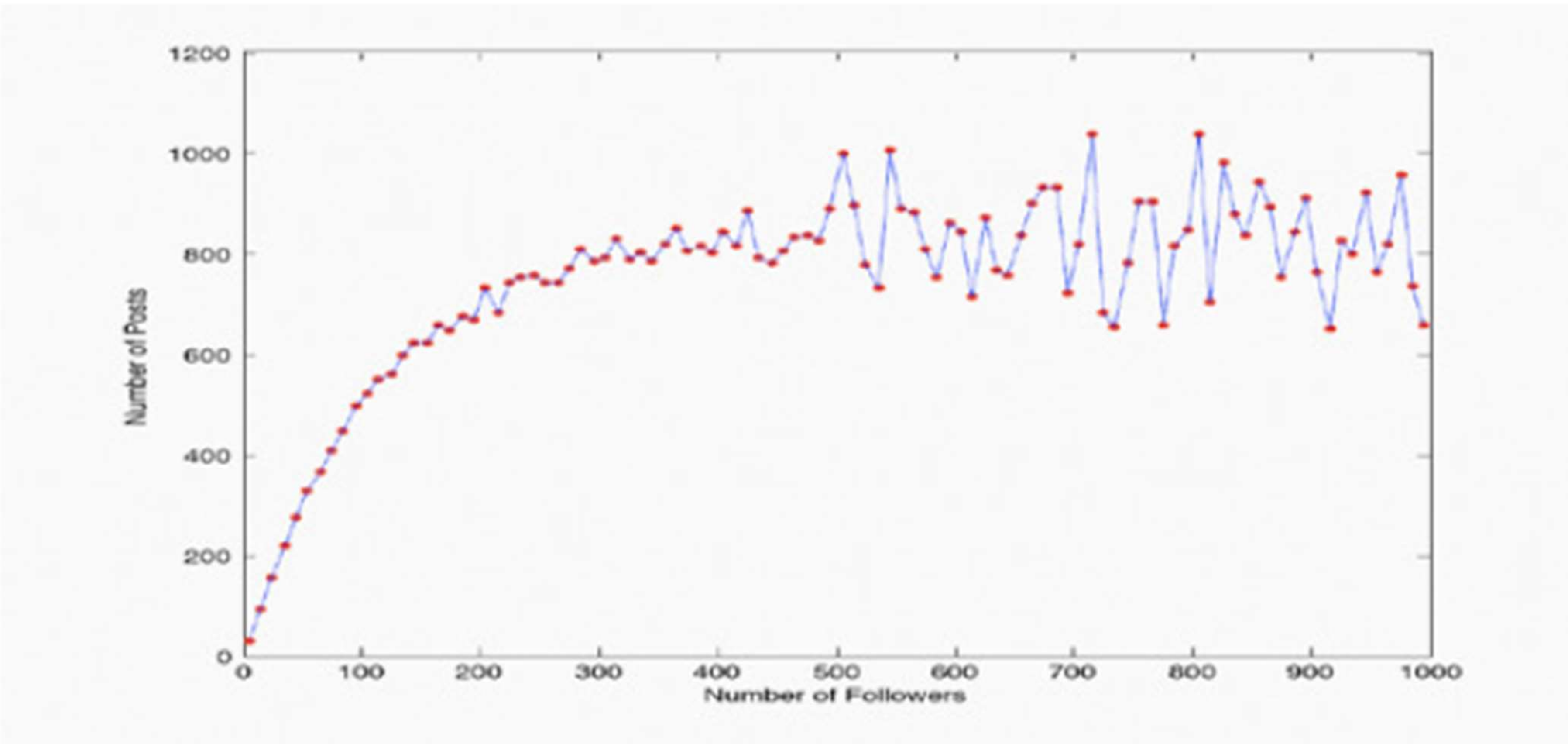

Donald J. Trump
40.2K Tweets
TWEETS
TWEETS & REPLIES
MEDIA
LIKES
Donald J. Trump @realD... · 06 Jan
"Former @NYTimes editor Jill Abramson rips paper's 'unmistakably anti-Trump' bias."
Ms. Abramson is 100% correct. Horrible and totally dishonest reporting on almost everything they write. Hence the term Fake News, Enemy of the People, and Opposition Party!
21.6K
22.5K
89.7K
Donald J. Trump @realD... · 06 Jan
Great Tweet today by Tyler Q. Houlton @SpoxDHS on the #FakeNews being put out by @CNN, a proud member of the Opposition Party. @TSA is doing a great job!
9,538
12K
54.1K

Tie strength on Twitter

- Huberman, Romero, and Wu in their 2009 article “social networks that matters : twitter under the microscope”
- a total of 309,740 users, who on average
  - posted 255 posts
  - 85 followers
  - followed 80 other users
- Among the 309,740 users only 211,024 posted at least twice (active users).
  - The active time of an active user is the time that has elapsed between his first and last post. On average, active users were active for 206 days.
- Strong ties : users with direct multiple messages to others
- Weak ties : users following others without direct communication

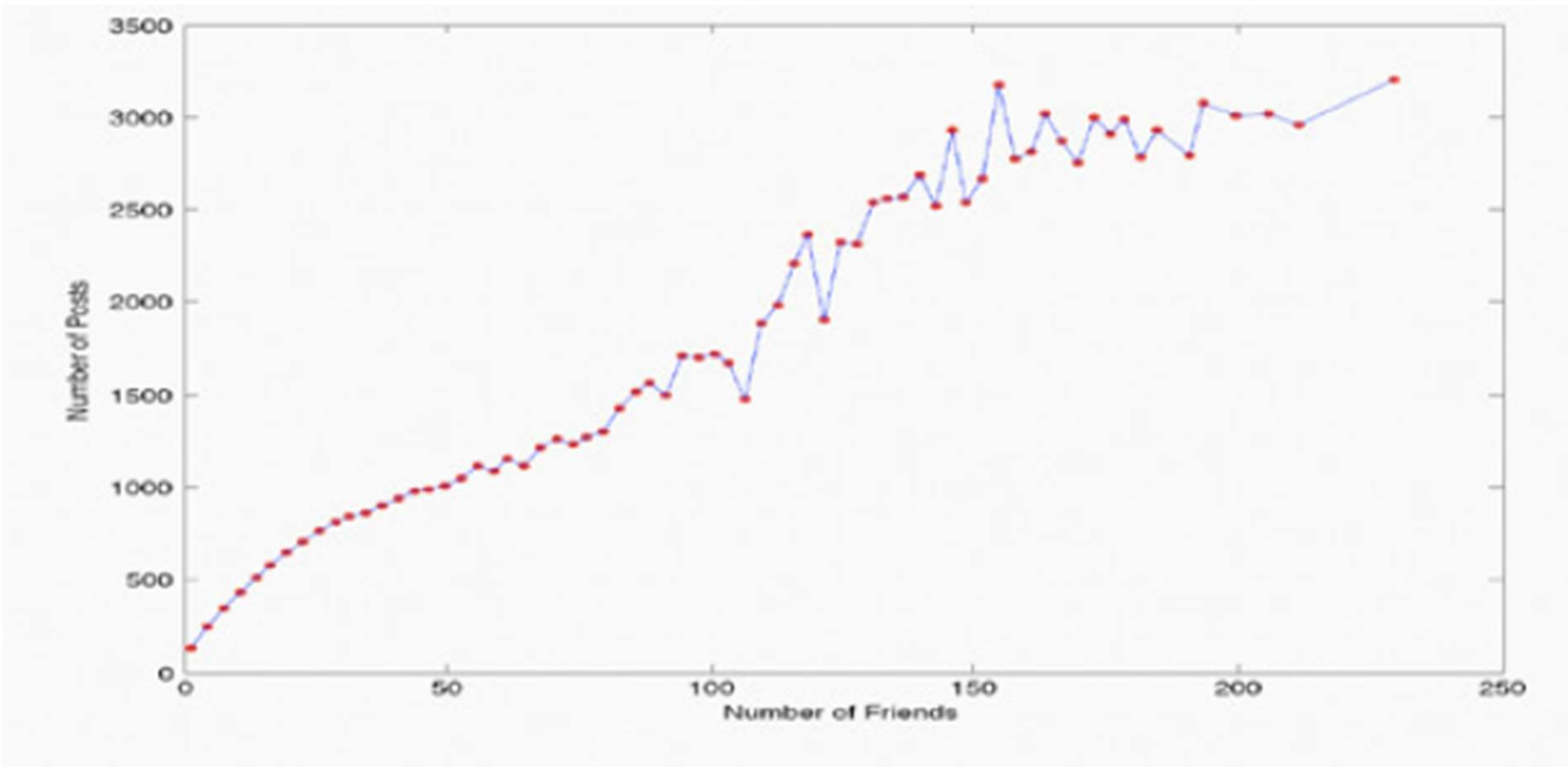
No. of posts verses no. of followers

saturation



No. of posts verses no. of friends

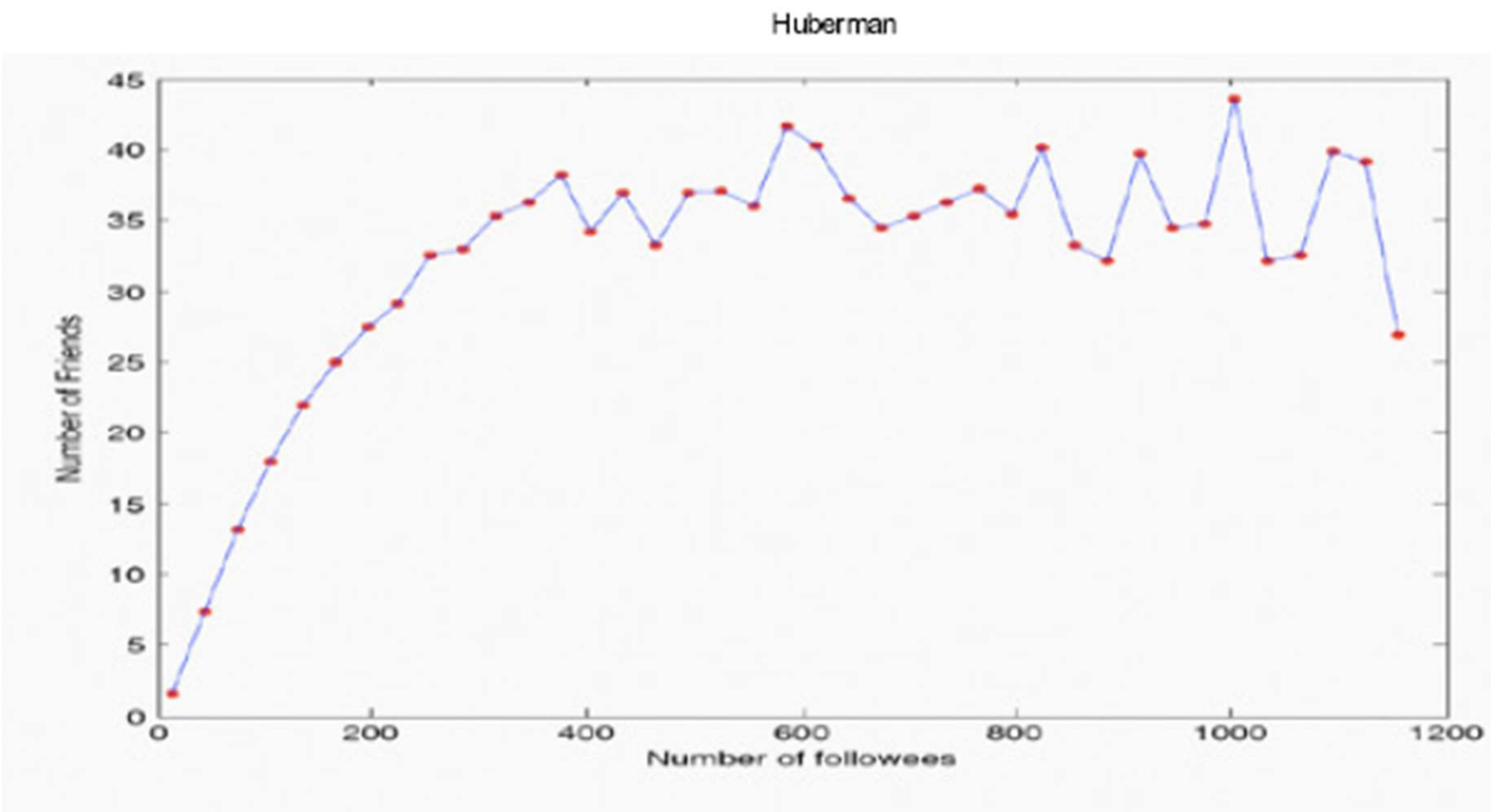
No saturation



Friends = those whom the user has directed at least two posts to.

No. of friends verses no. of followees

saturation

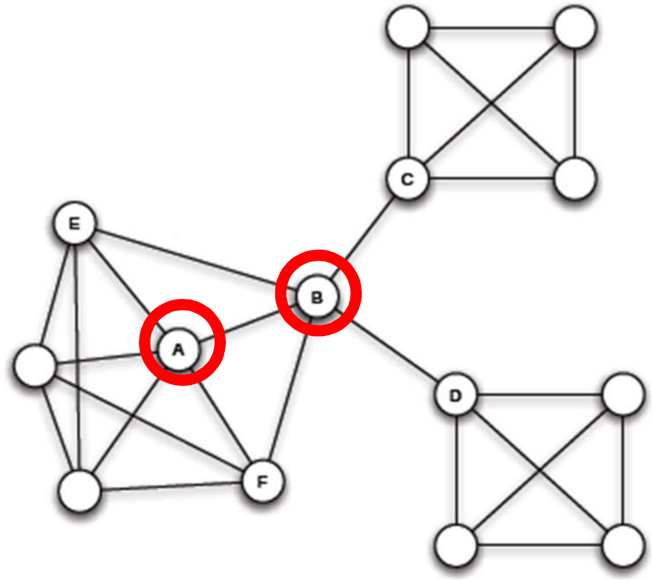


- Even for users who maintain very large numbers of weak ties on-line, the number of strong ties remains relatively modest.
- Each strong tie requires investment of time and effort.
- The formation of weak ties is governed by much milder constraints
  - they need to be established at their outset but not necessarily maintained continuously

Outlines

1. Network structure : Triadic Closure
2. Strength of Weak Ties
3. Tie Strength and Network Structure : examples
4. Closure and Structure Holes
5. Graph Partitioning

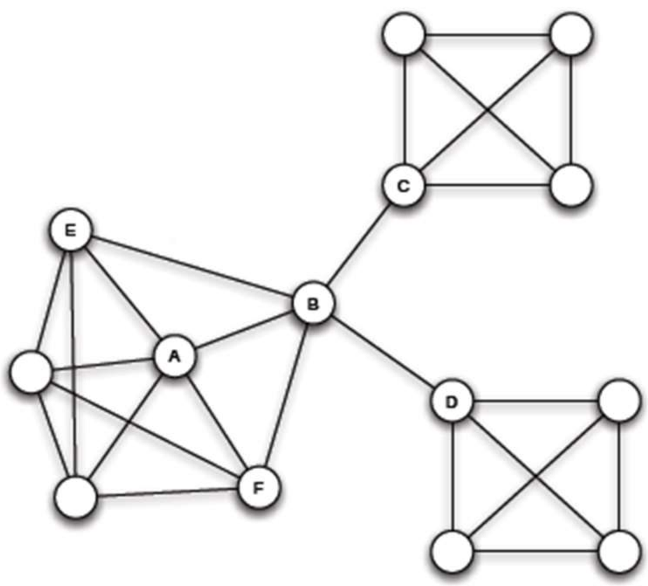
4. Closure and Structural Holes



- Edges
  - bridges and local bridges
  - Strong ties and weak ties
- Nodes
  - What are the roles of different nodes ?
  - Some nodes are positioned at the interface between multiple groups (B)
  - Others are positioned in the middle of a single group (A)

Embeddedness

- *Embeddedness* of an edge in a network is the number of common neighbors the two endpoints have.



Embeddedness(A-B)=2
(common neighbors are E and F)

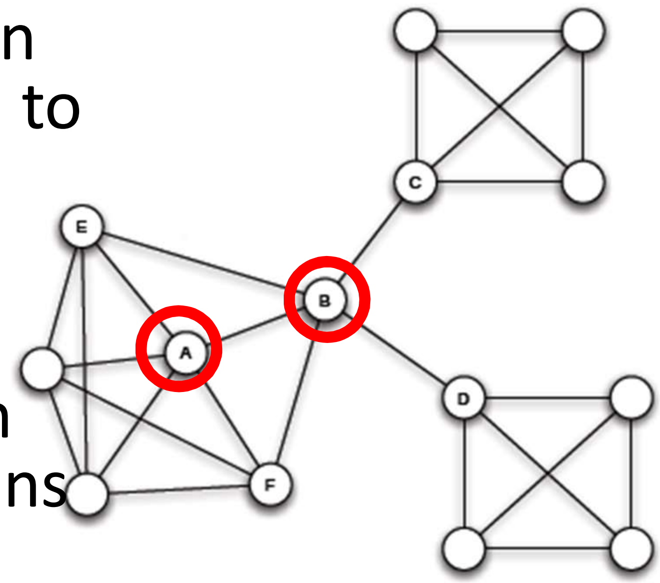
Neighborhood overlap

$$= \frac{no.of\ common\ neighbors}{total\ no.of\ neighbors}$$

Embeddedness(Local bridge) =0

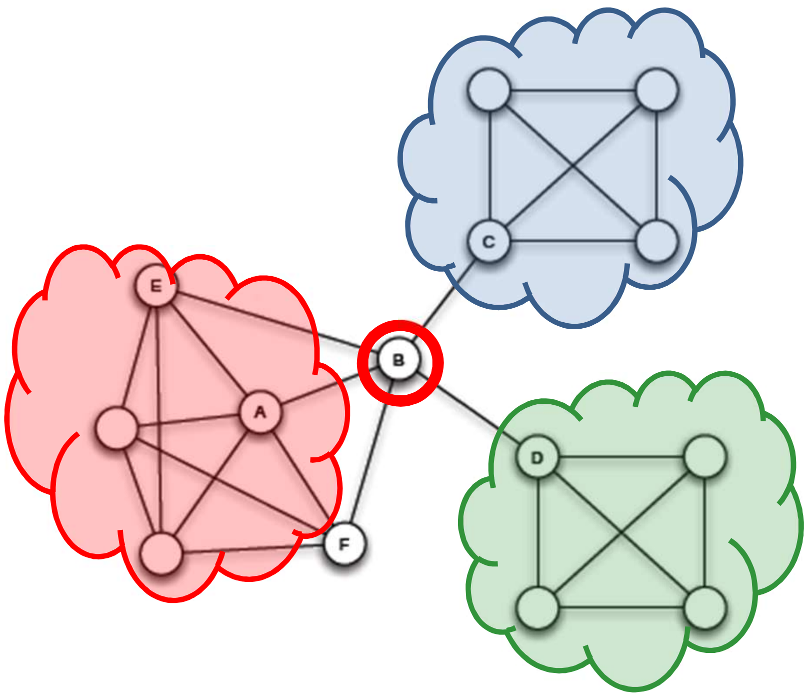
Embeddedness

- If two individuals are connected by an embedded edge, it is easier for them to trust one another
  - lots of mutual friends
  - E.g. A and his friends
- If two are connected by an edge with zero embeddedness, their interactions are risky.
  - e.g. B and C, B and D
- If one is involved in different groups (e.g. B), one might have contradictory norms and expectations from the groups.



Structure Hole

- Structure Hole : separation between non-redundant contacts



Three classes of structure holes in the network

- Holes between contacts and contacts
- Holes between contacts and contacts
- Holes between contacts and contacts

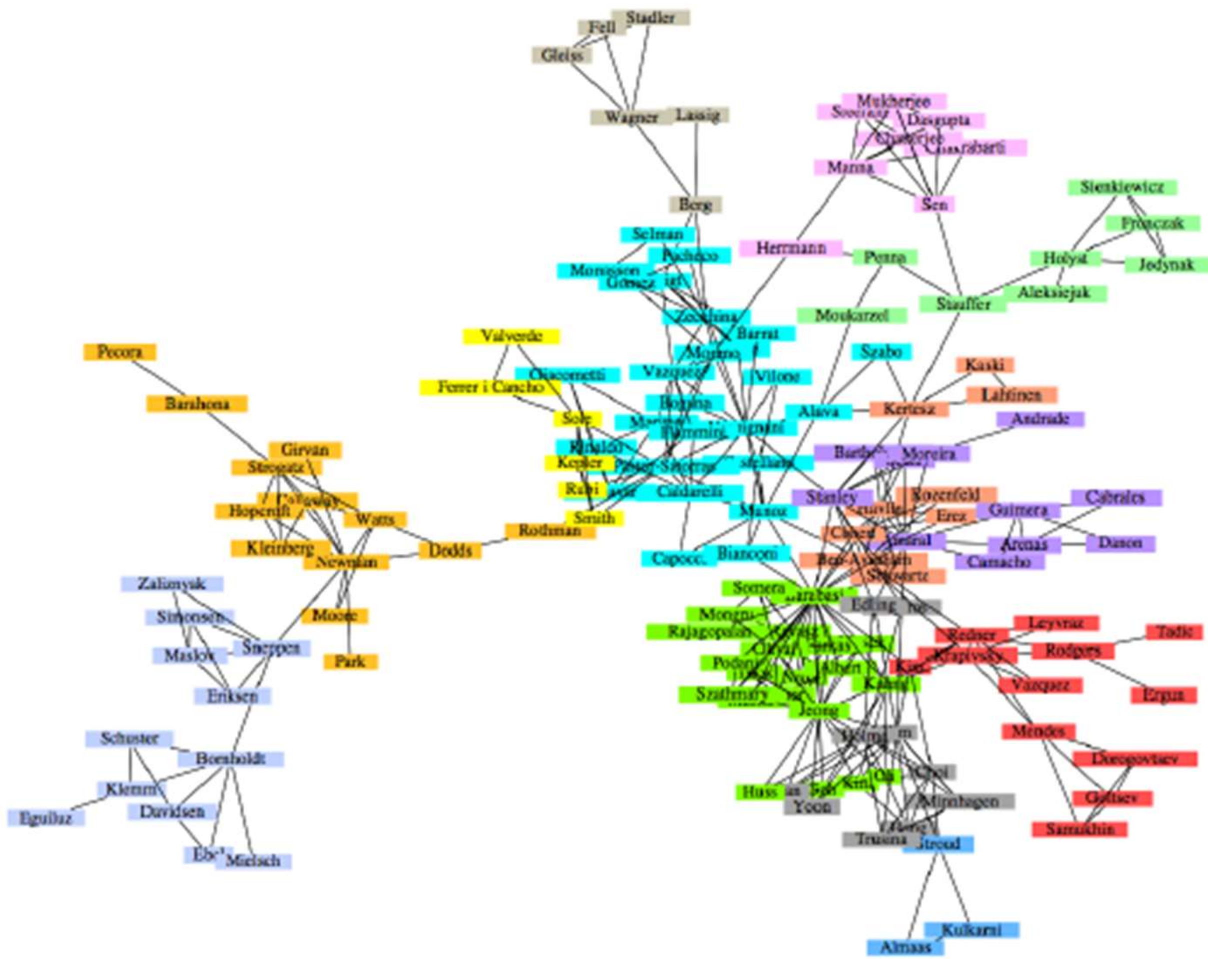
Structure Hole

- Networks rich in structural holes are a form of social capital in that they offer information benefits.
- The main player (Node B) in a network that bridges structural holes enables him to
  - access information originating in multiple, non-interacting parts of the network.
  - have opportunities for innovations arose from the unexpected synthesis of multiple ideas, each of them on their own perhaps well-known, but well-known in distinct and unrelated bodies of expertise.
  - have an opportunity for a kind of social “gate-keeping”; a source of power in the organization.

5. Betweenness measures and Graph Partitioning

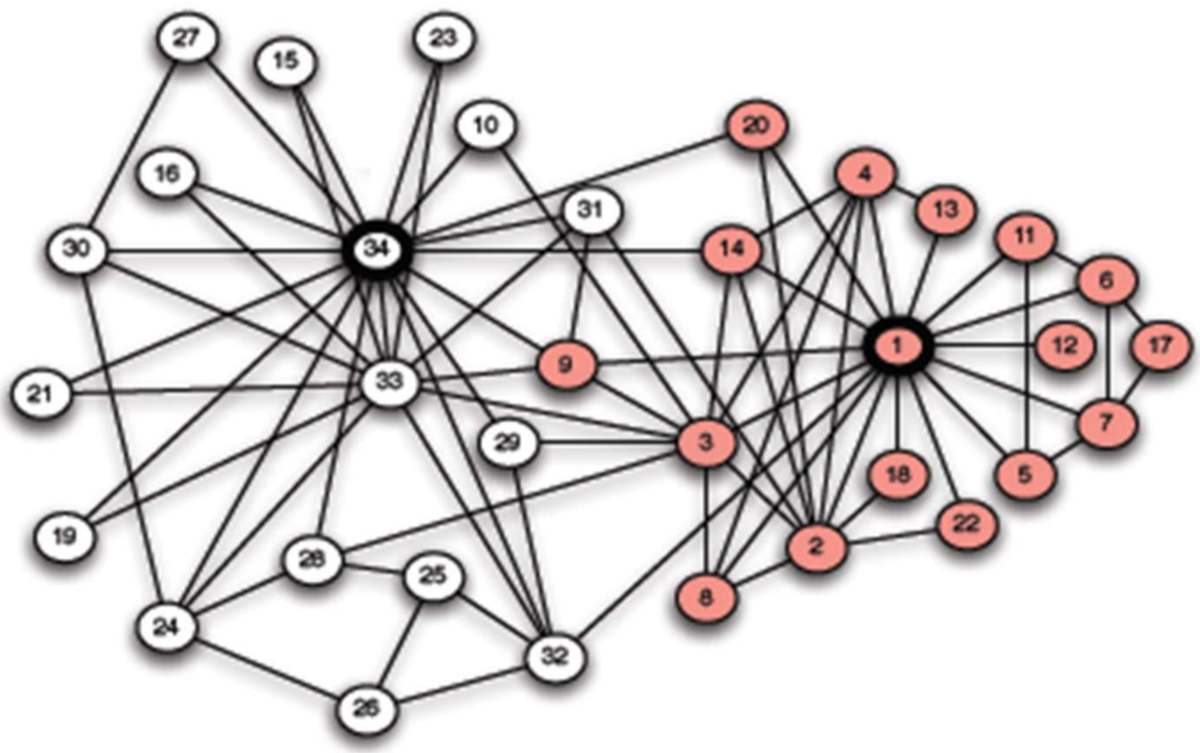
- Theme :
  - Tightly-knit regions
  - Graph partitioning : algorithm that break a network down into a set of tightly-knit regions, with sparser interconnections between the regions.

Co-authorship network of physicists and applied mathematicians



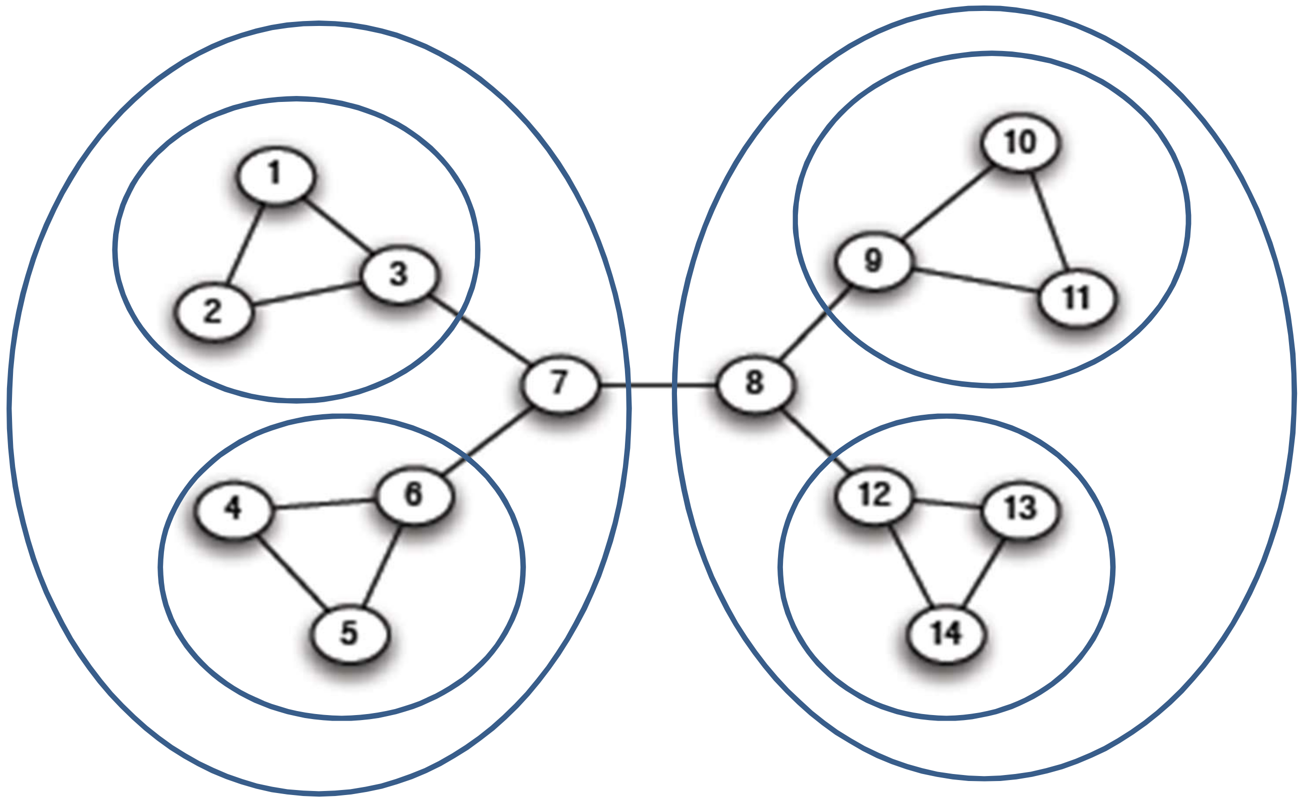
Karate Club

- A dispute between the club president (node 34) and the instructor (node 1)
- Two conflicting groups are still heavily inter-connected.



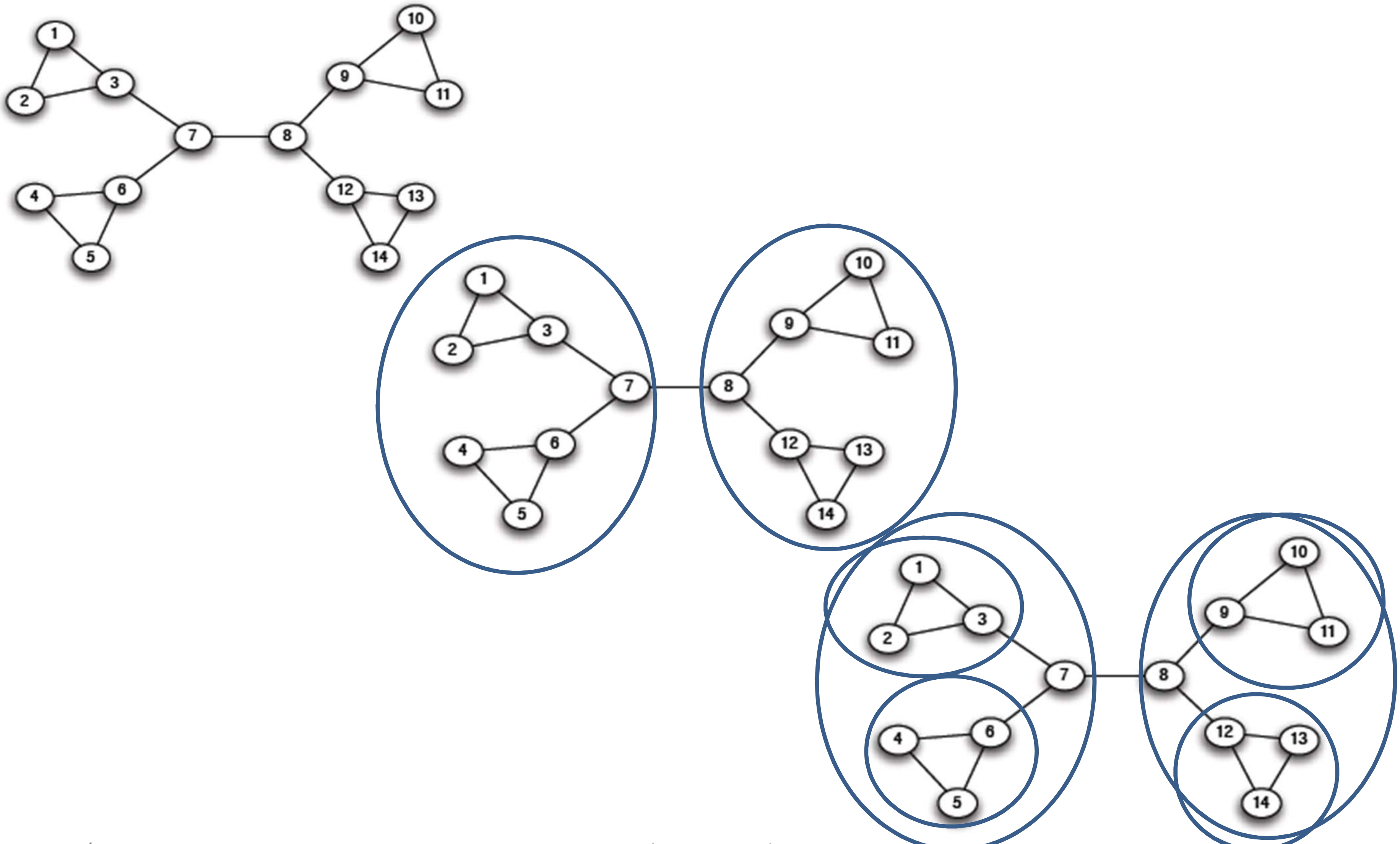
- Graph partitioning : Describing a method that can take a network and **break** it into a set of tightly-knit regions, with sparser interconnections between the regions.

Nested structure of a network

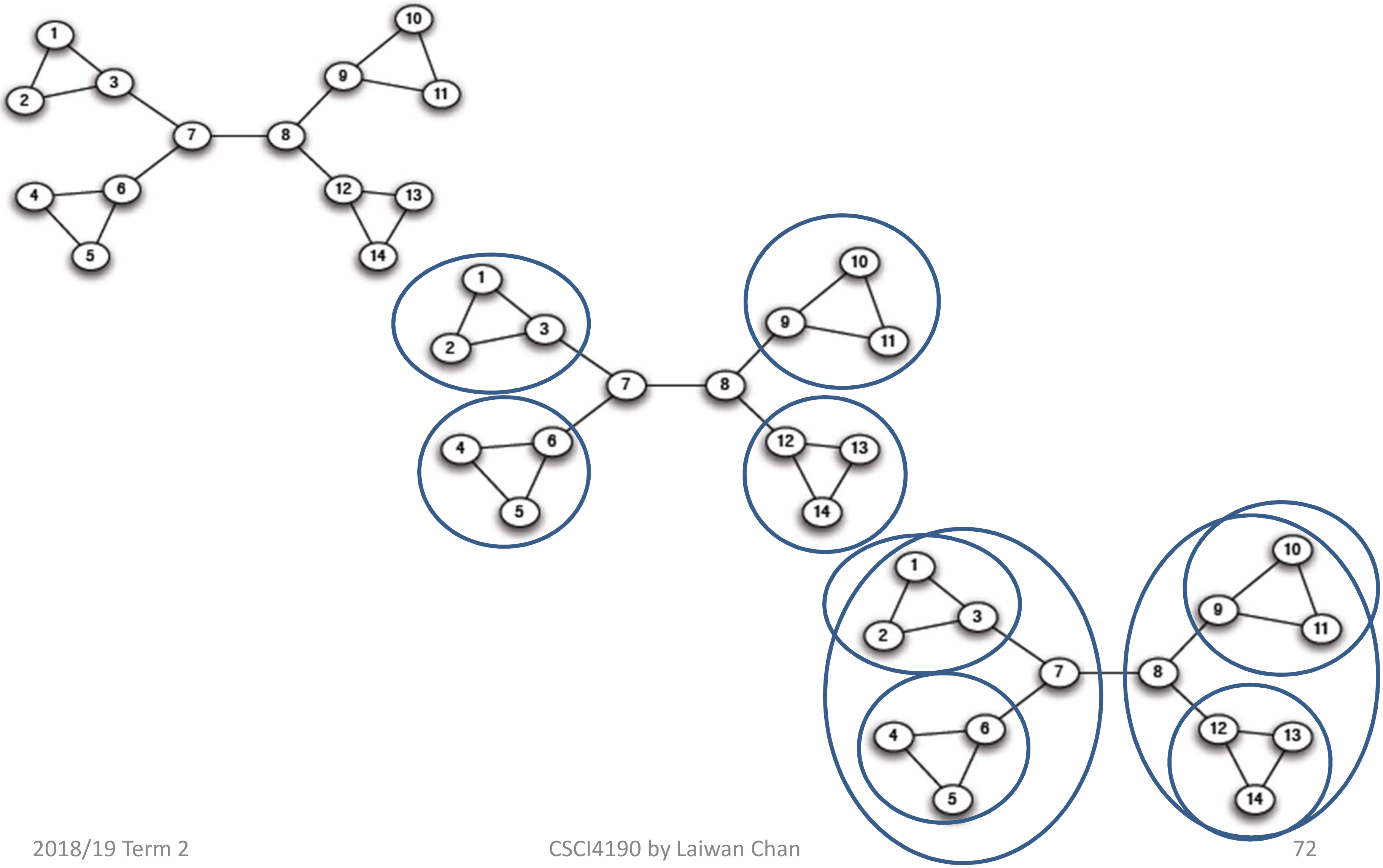


- General Approaches to Graph Partitioning
  - divisive methods : identifying and removing the “spanning links” between densely-connected regions. Once these links are removed, the network begins to fall apart into large pieces;
  - agglomerative methods : focusing on the most tightly-knit parts of the network. find nodes that are likely to belong to the same region and merge them together

Divisive methods

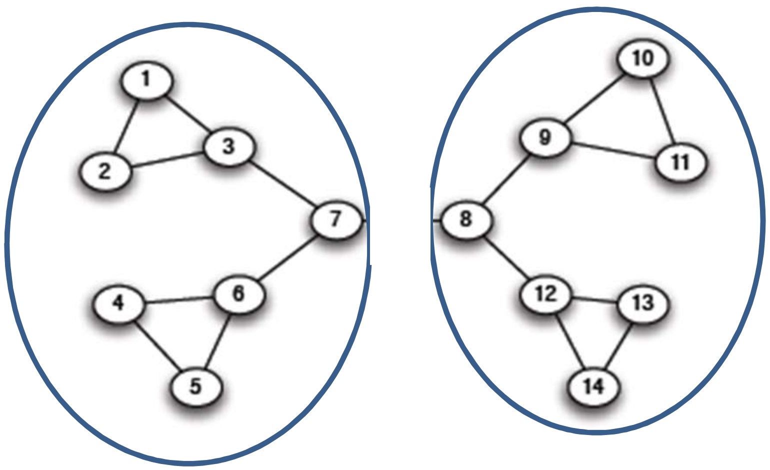


Agglomerative methods



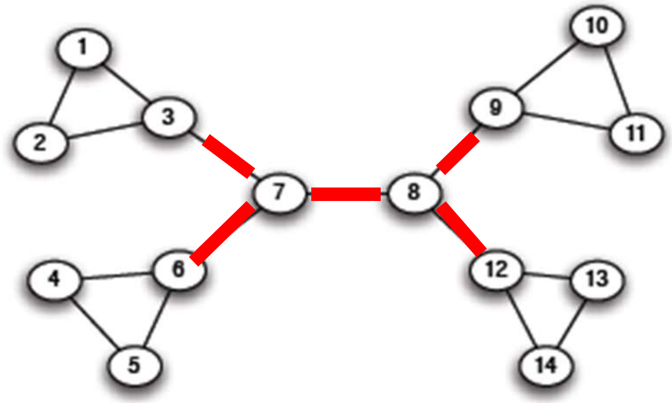
Divisive methods

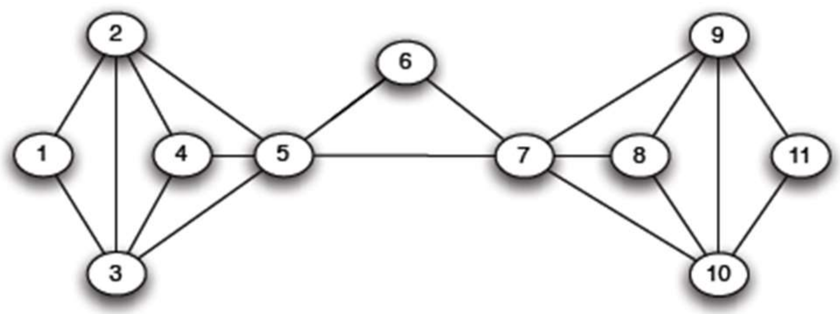
- For divisive methods, bridges are good candidates for removal.



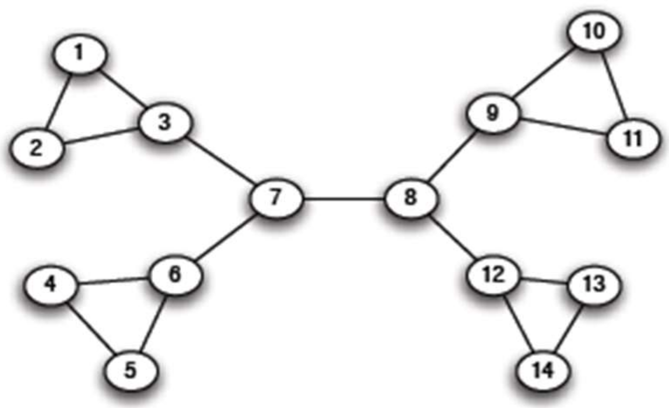
Divisive methods

- Can we simply remove the bridges and local bridges ?
  - Which one first if we have a number of bridges ?
  - What if there is no bridges nor local bridges ?





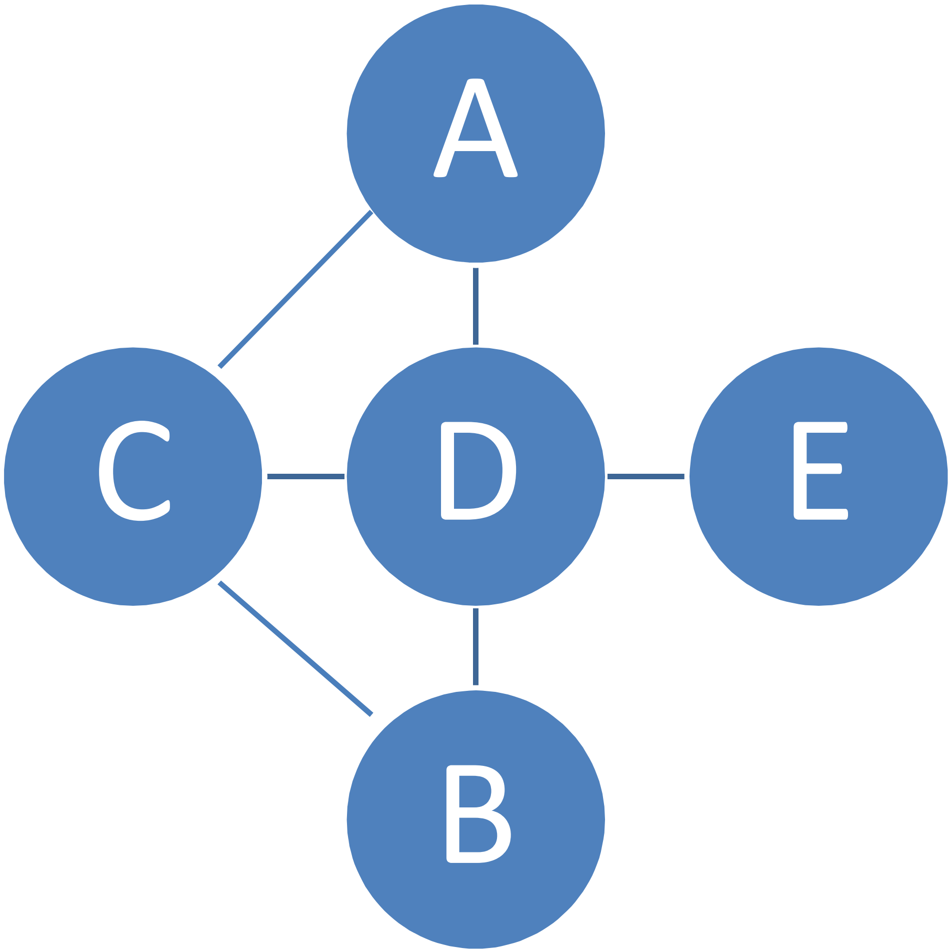
- Bridges and local bridges form part of the shortest path between pairs of nodes in different parts of the network.
- Define "traffic" on the network and look for the edges that carry most of the traffic.



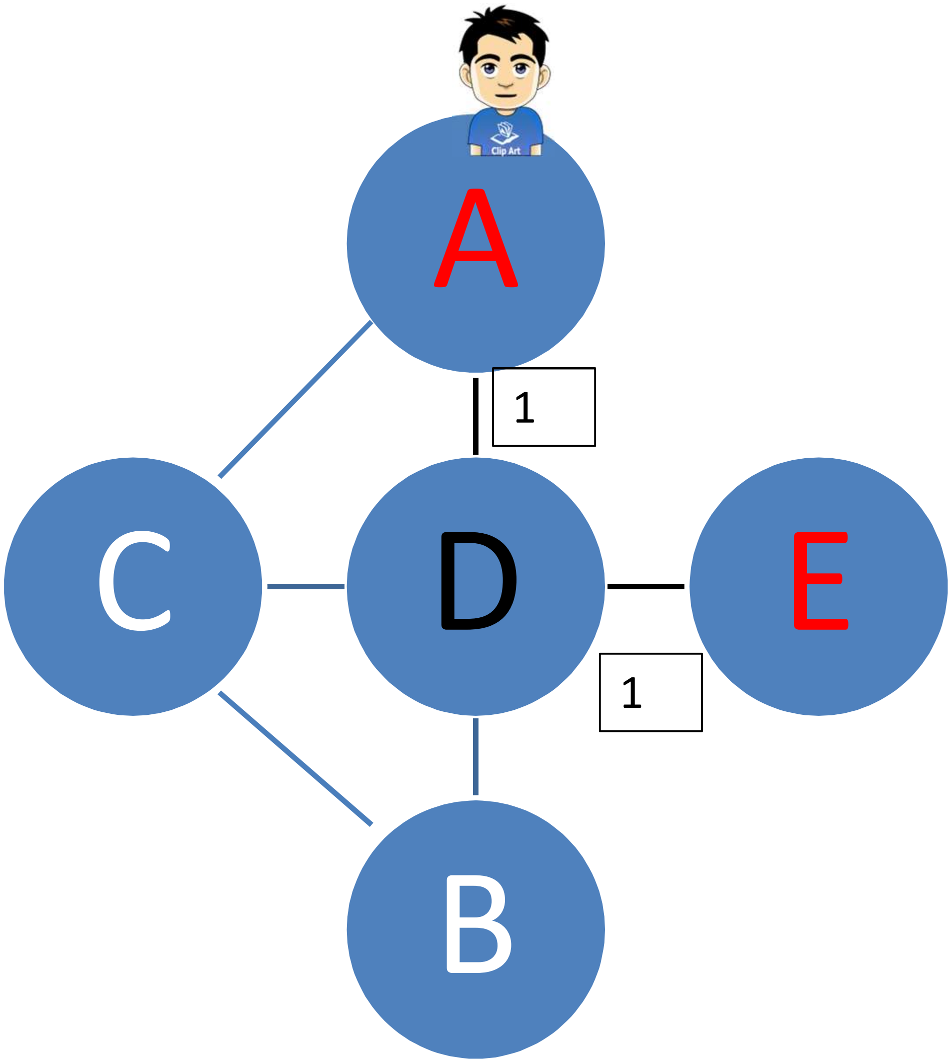
Traffic in a Network

For each pair of nodes A and B

- 0 unit of traffic “flow” if A and B belong to different component
- Otherwise, 1 unit of traffic “flow” from A to B
  - the flow divides itself evenly along all the possible *shortest* paths from A to B
  - if there are k shortest paths between A and B, then 1/k units of flow along each one.


A
C
D
E
B

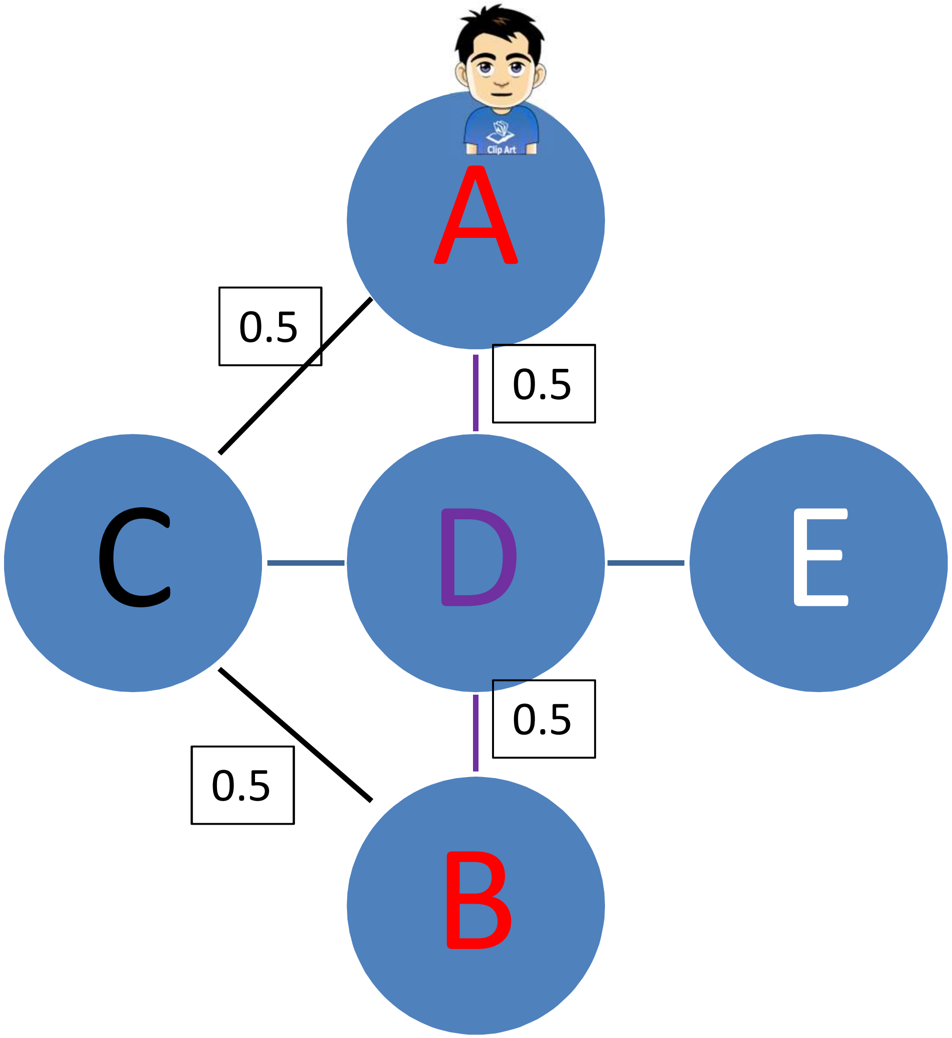
Traffic flow from A to E



Shortest path between A and E

- A– D – E (1 unit)

Traffic flow from A to B

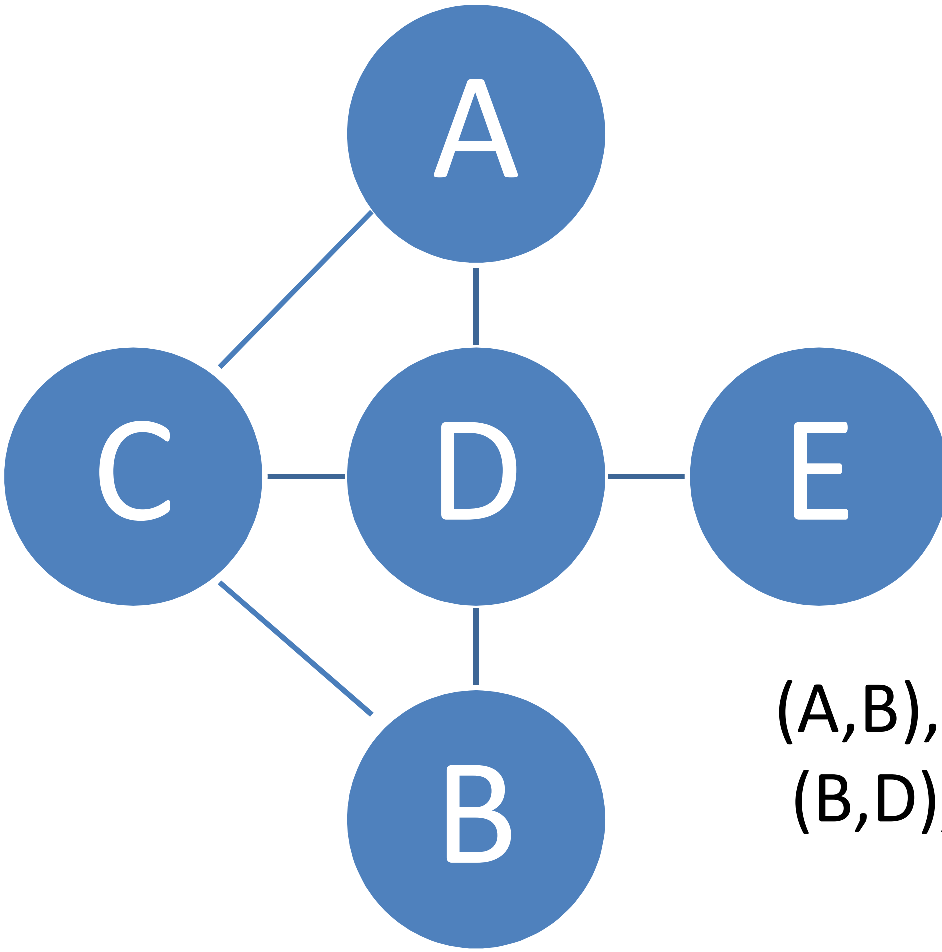


Shortest path between
A and B

- A– C – B (1/2 units)
- A –D – B (1/2 units)

Betweenness

- Betweenness of an edge : total amount of flow that the edge carries, counting flow between all pairs of nodes using this edge.
- Betweenness of a node : total amount of flow that the node carries, when a unit of flow between each pair of nodes is divided up evenly over shortest paths (same as for edges)
  - nodes of high betweenness occupy critical roles in the network, as it interfaces between tightly-knit groups.



5 nodes

→

${}_nC_2 = {}_5C_2 = 10$ pairs of nodes

(A,B),(A,C),(A,D),(A,E),(B,C),
(B,D),(B,E),(C,D),(C,E),(D,E)

1 Shortest path between

A – C

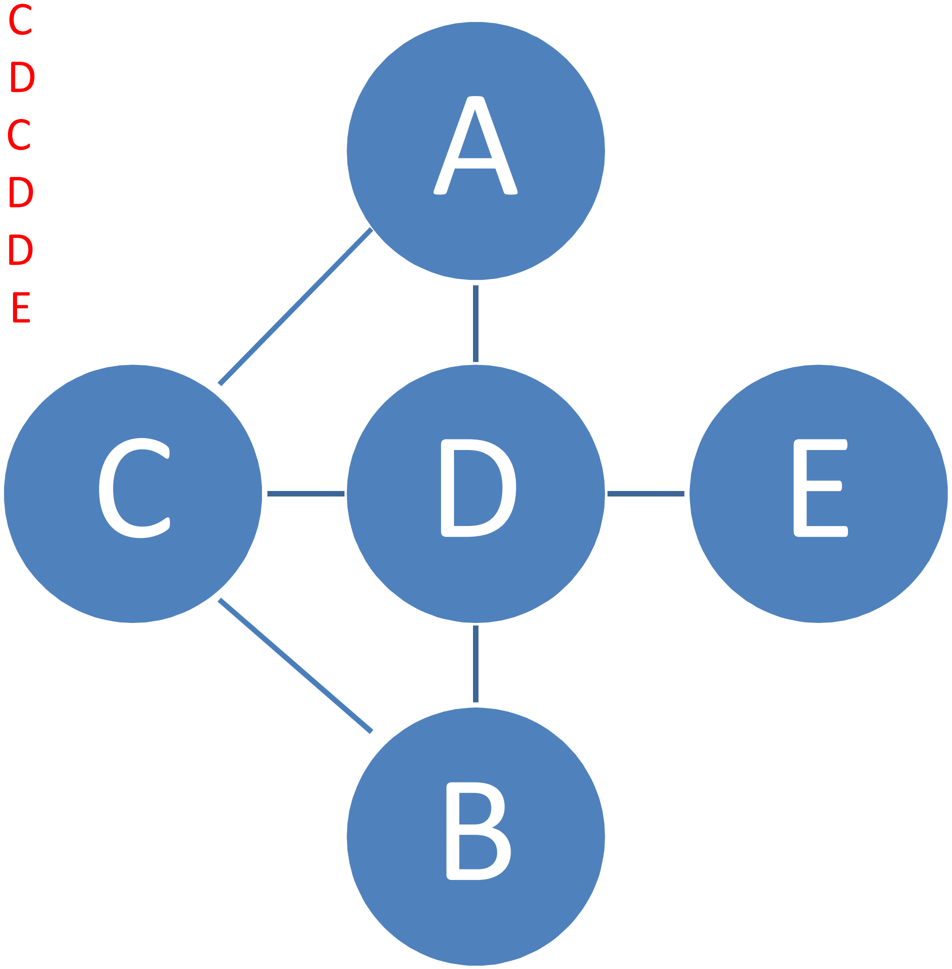
A – D

B – C

B – D

C – D

D – E



2 Shortest paths between

A and B

- A– C – B (1/2 units)
- A –D – B (1/2 units)

1 Shortest path between

A and E

- A– D – E (1 unit)

1 Shortest path between

B and E

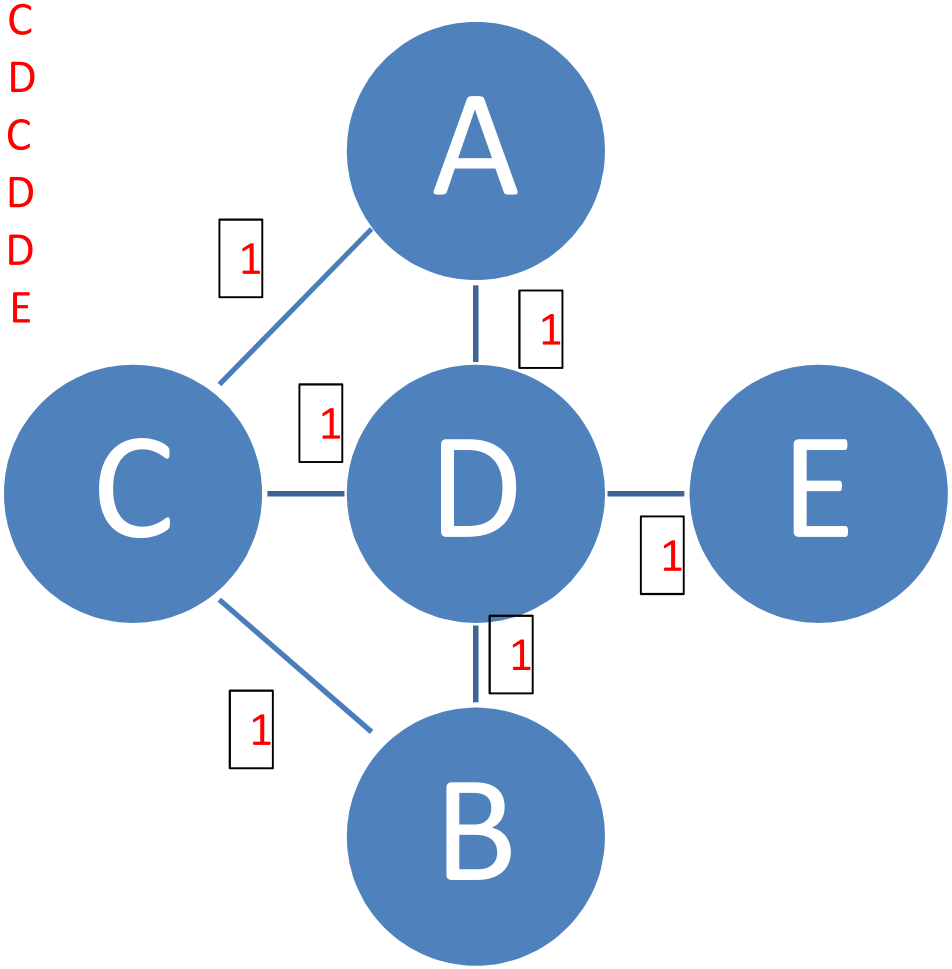
- B– D – E (1 unit)

1 Shortest path between

C and E

- C– D – E (1 unit)

1 Shortest path between
A – C
A – D
B – C
B – D
C – D
D – E



1 Shortest path between

A – C

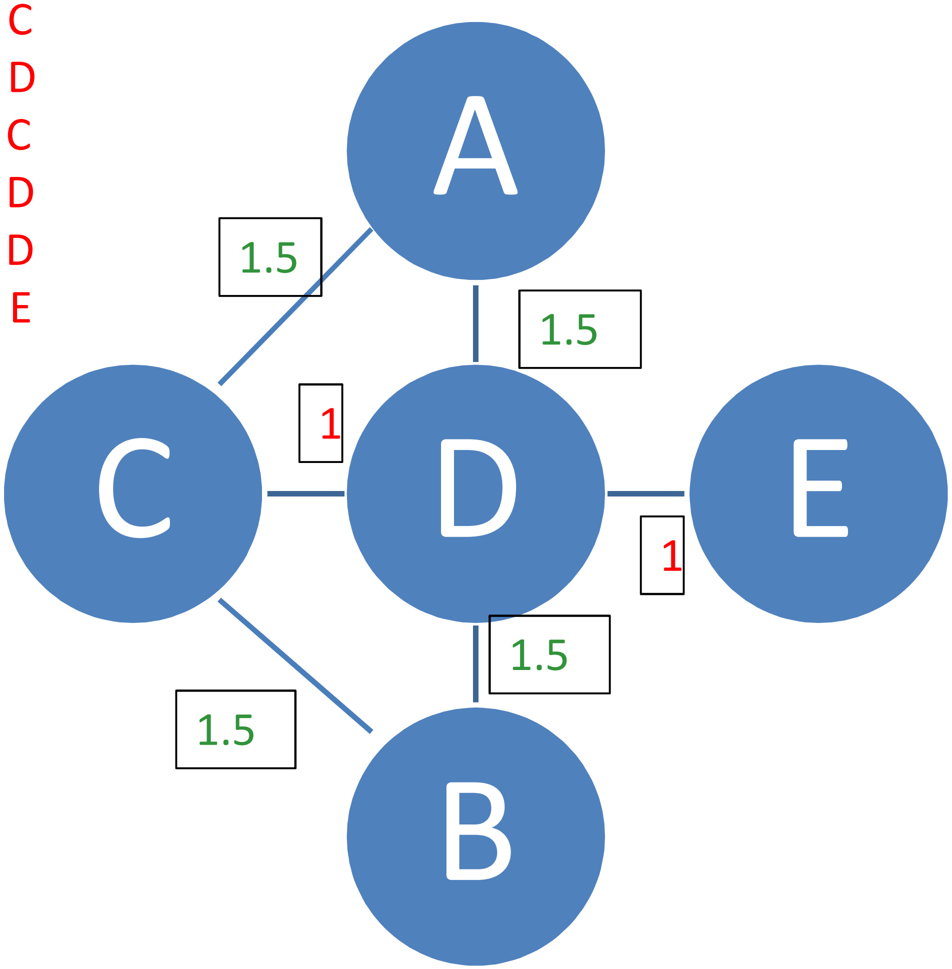
A – D

B – C

B – D

C – D

D – E



2 Shortest paths between

A and B

- A– C – B (1/2 units)
- A –D – B (1/2 units)

1 Shortest path between

A – C

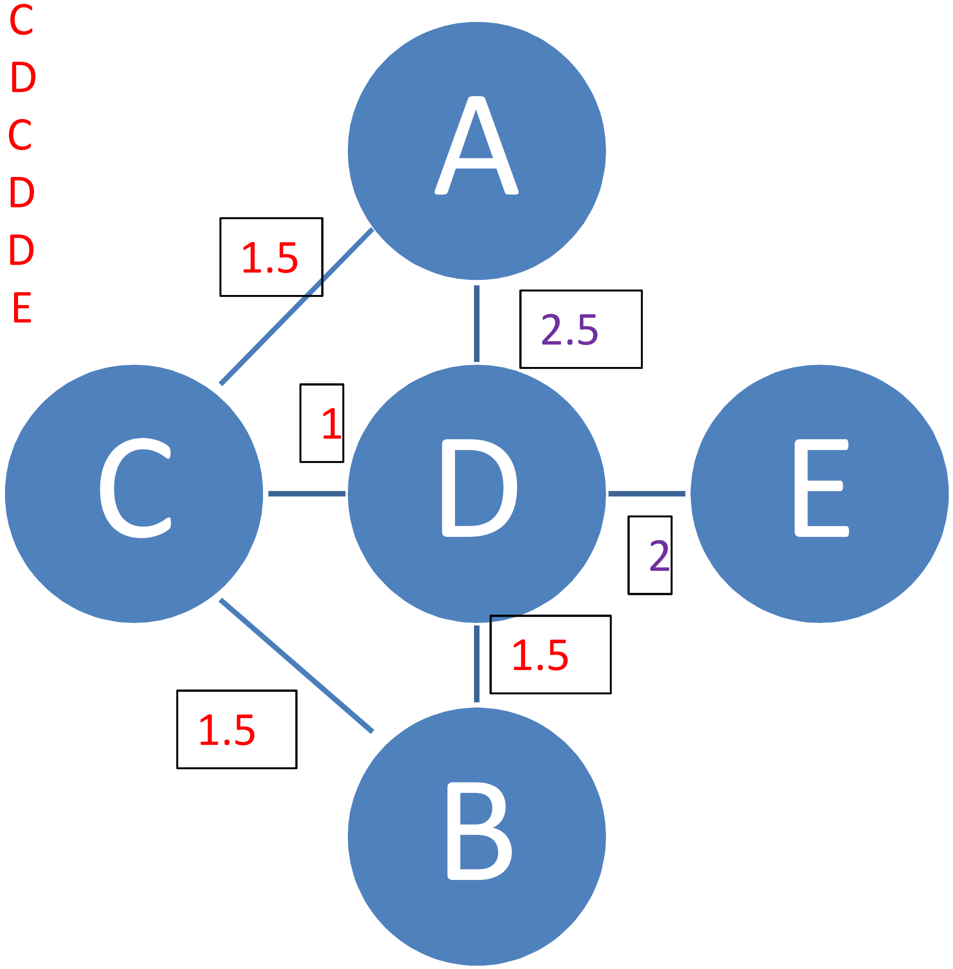
A – D

B – C

B – D

C – D

D – E



2 Shortest paths between

A and B

- A– C – B (1/2 units)
- A –D – B (1/2 units)

1 Shortest path between

A and E

- A– D – E (1 unit)

1 Shortest path between

A – C

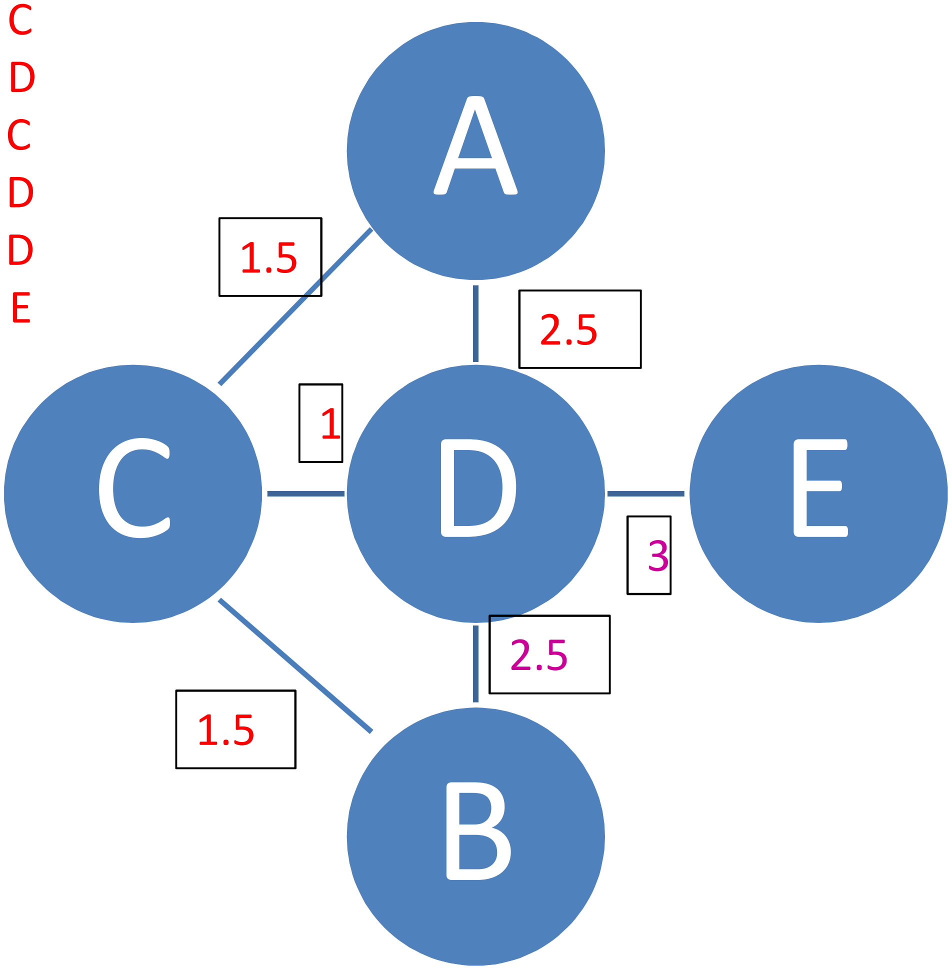
A – D

B – C

B – D

C – D

D – E



2 Shortest paths between

A and B

- A– C – B (1/2 units)
- A –D – B (1/2 units)

1 Shortest path between

A and E

- A– D – E (1 unit)

1 Shortest path between

B and E

- B– D – E (1 unit)

1 Shortest path between

A – C

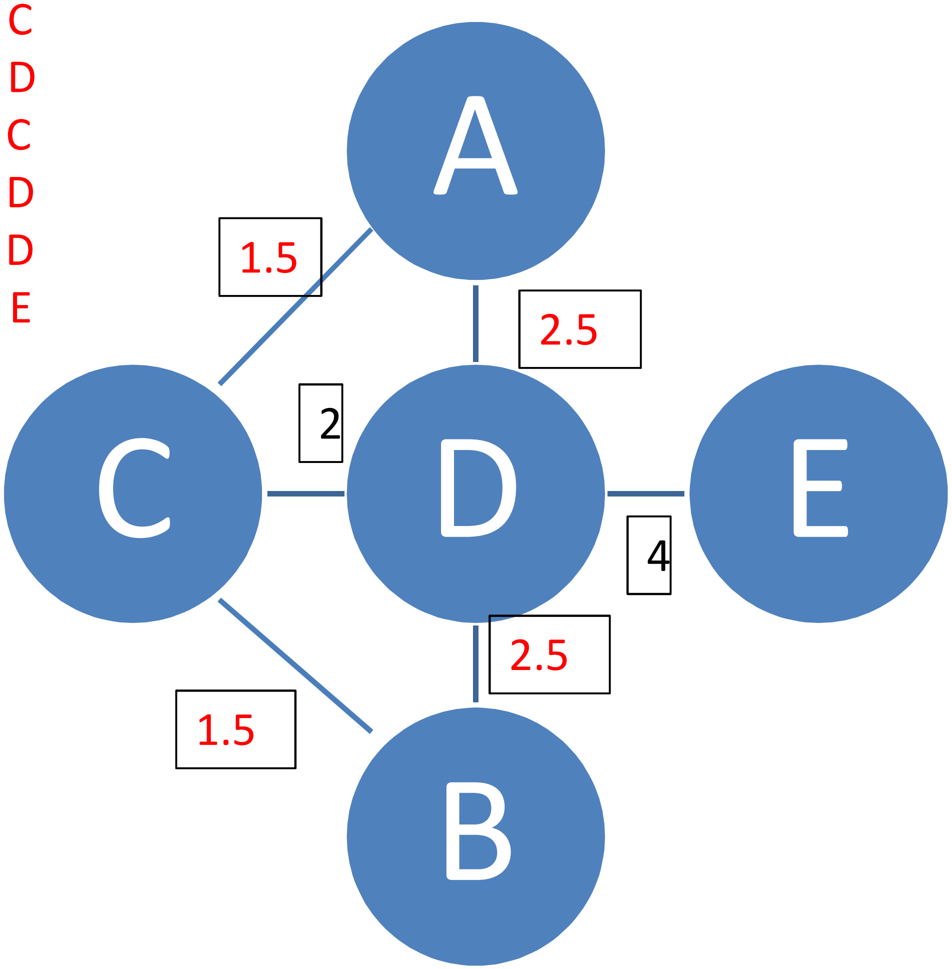
A – D

B – C

B – D

C – D

D – E



2 Shortest paths between

A and B

- A– C – B (1/2 units)
- A –D – B (1/2 units)

1 Shortest path between

A and E

- A– D – E (1 unit)

1 Shortest path between

B and E

- B– D – E (1 unit)

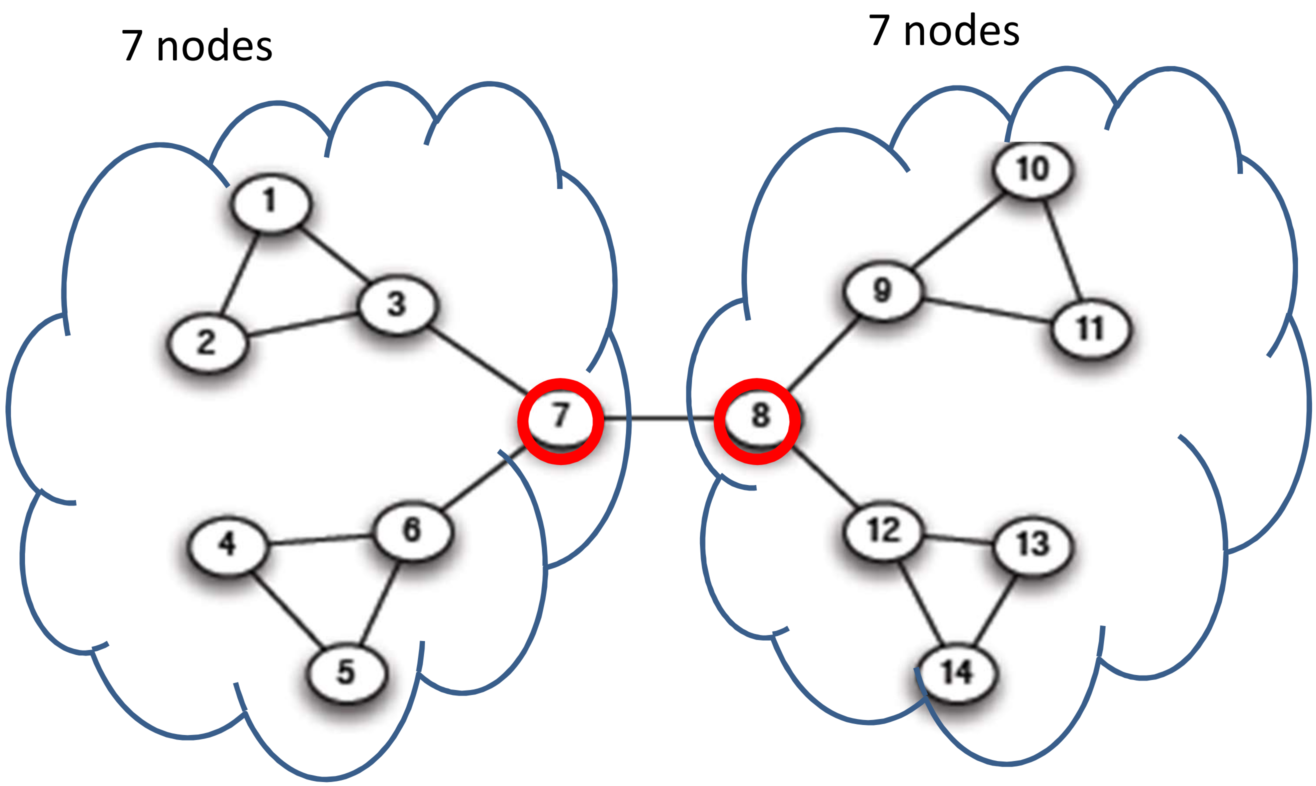
1 Shortest path between

C and E

- C– D – E (1 unit)

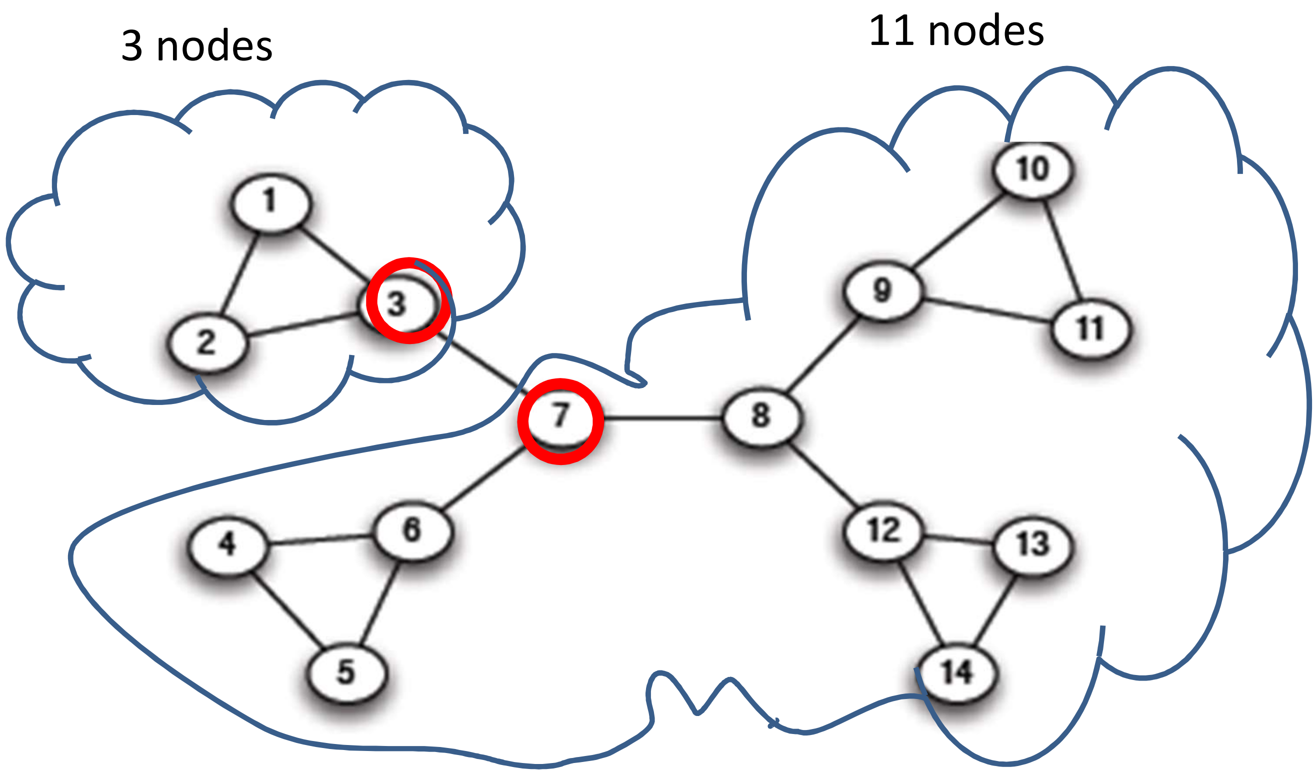
Edge Betweenness

- Betweenness(edge 7-8) = 7x7=49



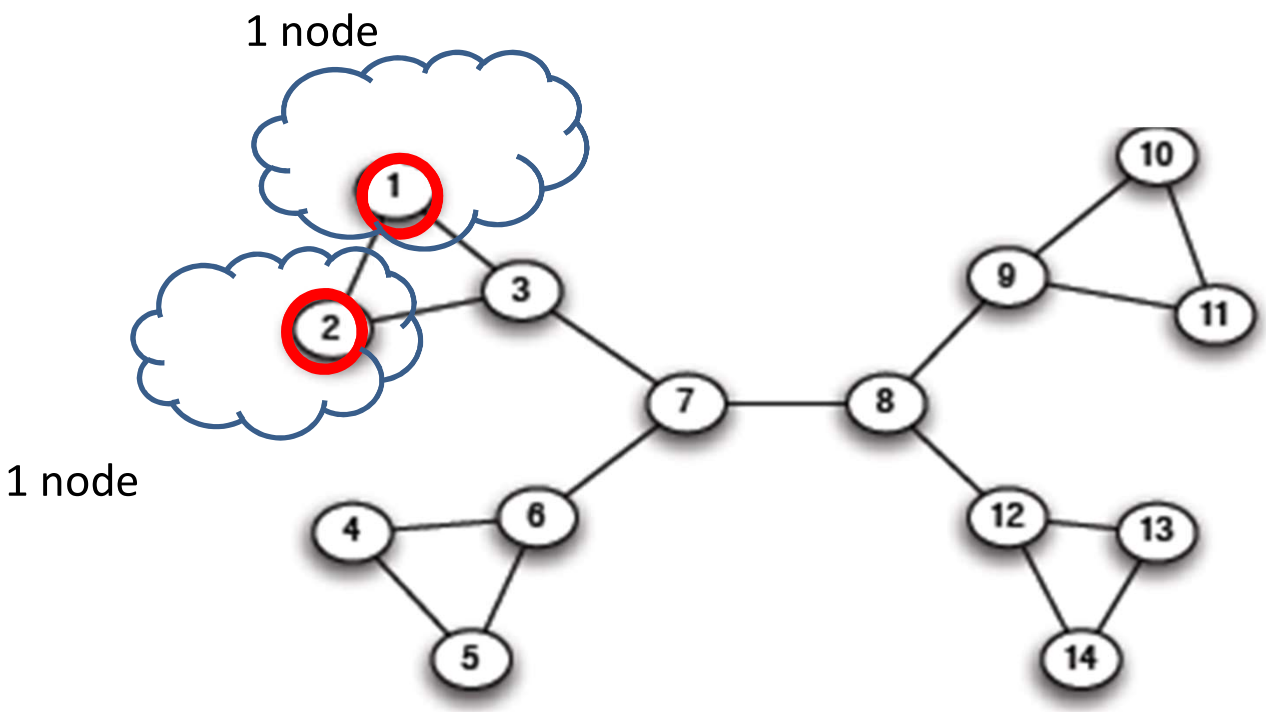
Edge Betweenness

- Betweenness(edge 3-7) = 3x11=33



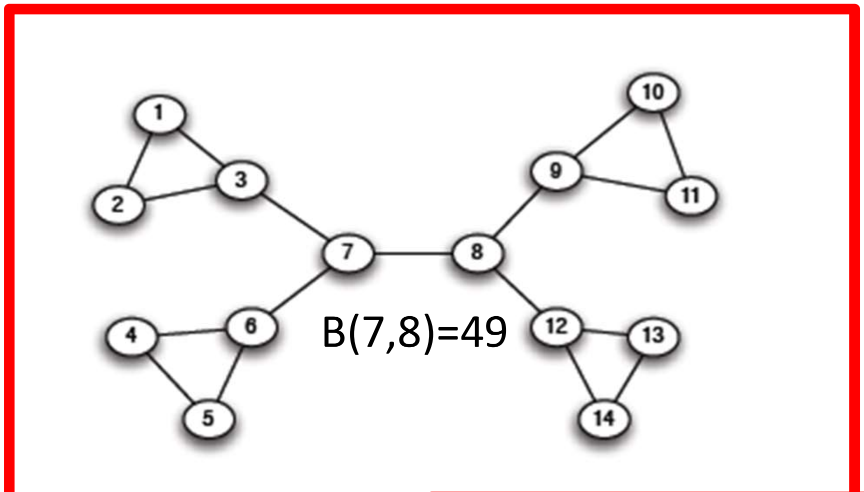
Edge Betweenness

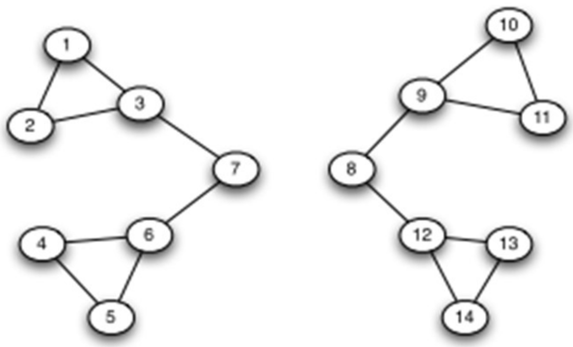
- Betweenness(edge 1-2) = 1x1=1



Girvan-Newman Method

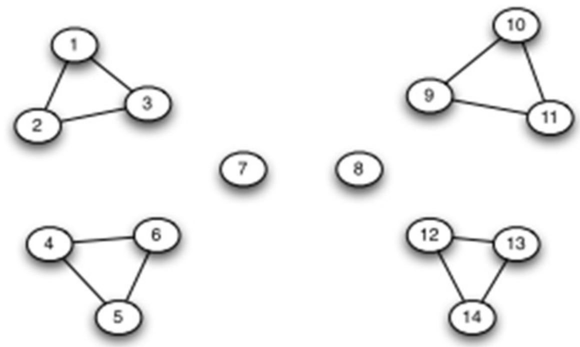
- Find the edge of highest betweenness (or multiple edges of highest betweenness, if there is a tie) and remove these edges from the graph.
  - This may cause the graph to separate into multiple components.
- Recalculate all betweennesses, and again remove the edge or edges of highest betweenness.
  - This may break some of the existing components into smaller components; if so, these are regions nested within the larger regions.
- (...) Proceed in this way as long as edges remain in graph, in each step recalculating all betweennesses and removing the edge or edges of highest betweenness.



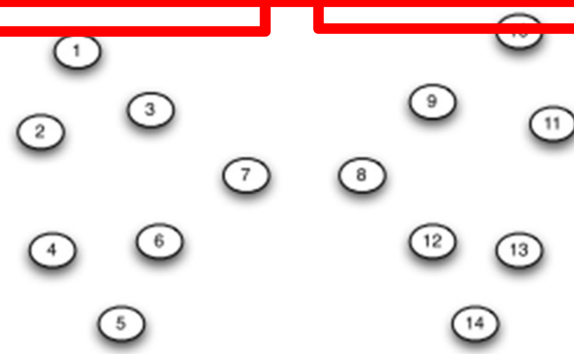


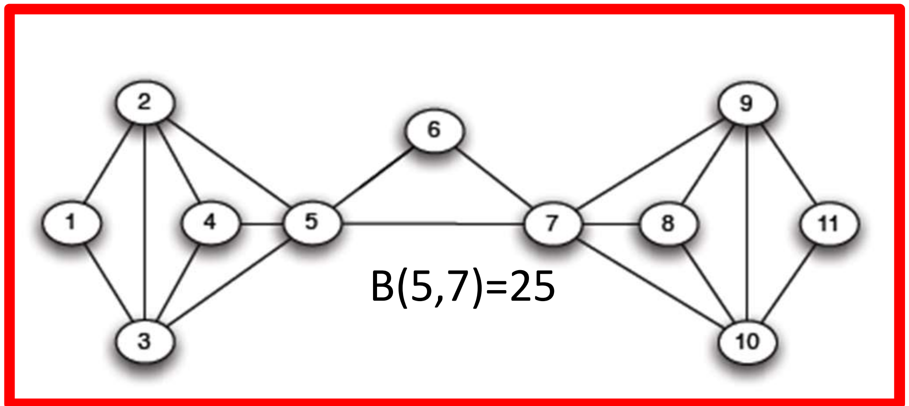
B(3,7)=B(6,7)=B(8,9)
=B(8,12)=12

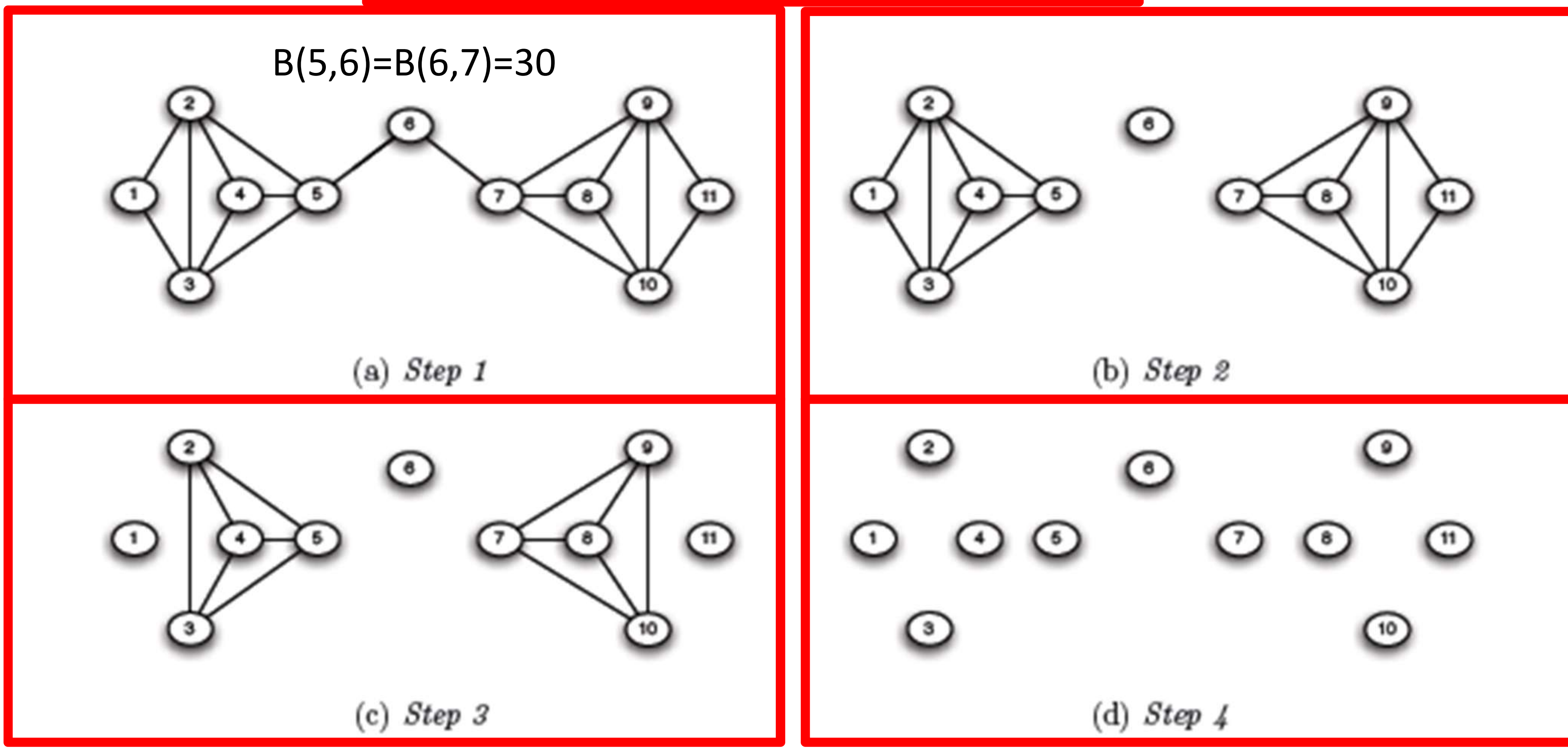
(a) *Step 1*



(b) *Step 2*



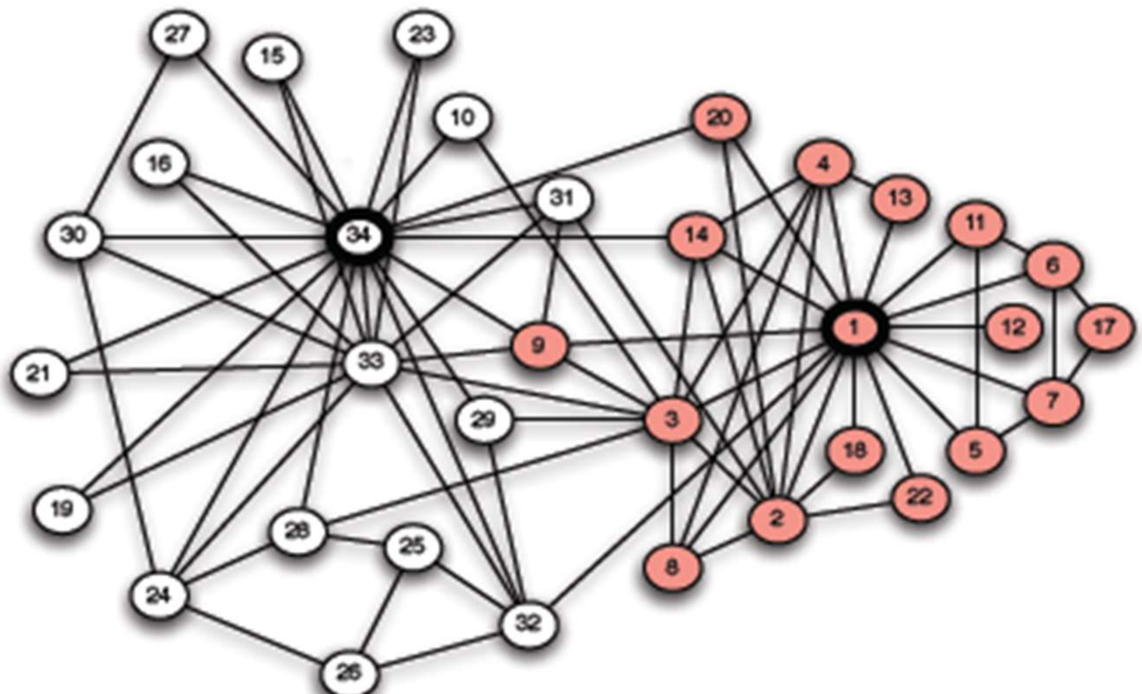




(a) *Step 1*
(b) *Step 2*
(c) *Step 3*
(d) *Step 4*

Karate Club Example

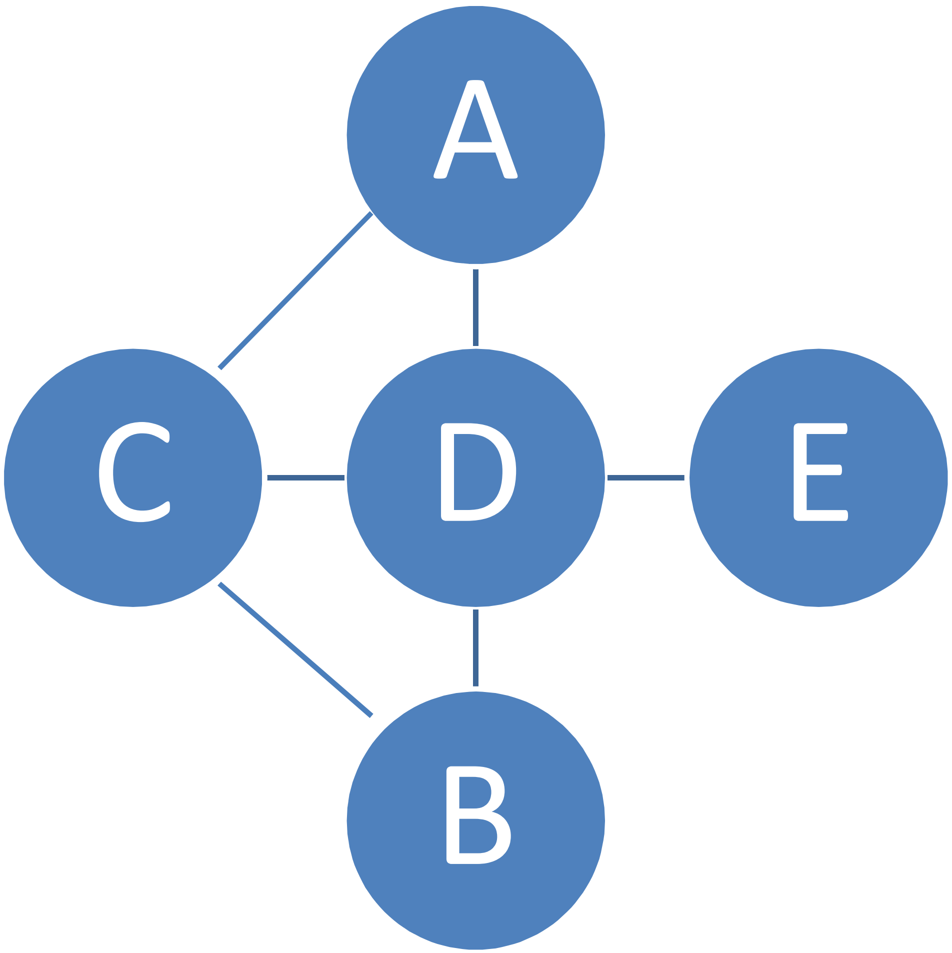
- Dispute between the club president (node 34) and the instructor (node 1)
- Use Girvan-Newman Method to remove edges until the graph separates into two pieces.
- The result agrees with the actual split except node 9, who has to follow the instructor when he was only three weeks away from completing a four year quest to obtain a black belt



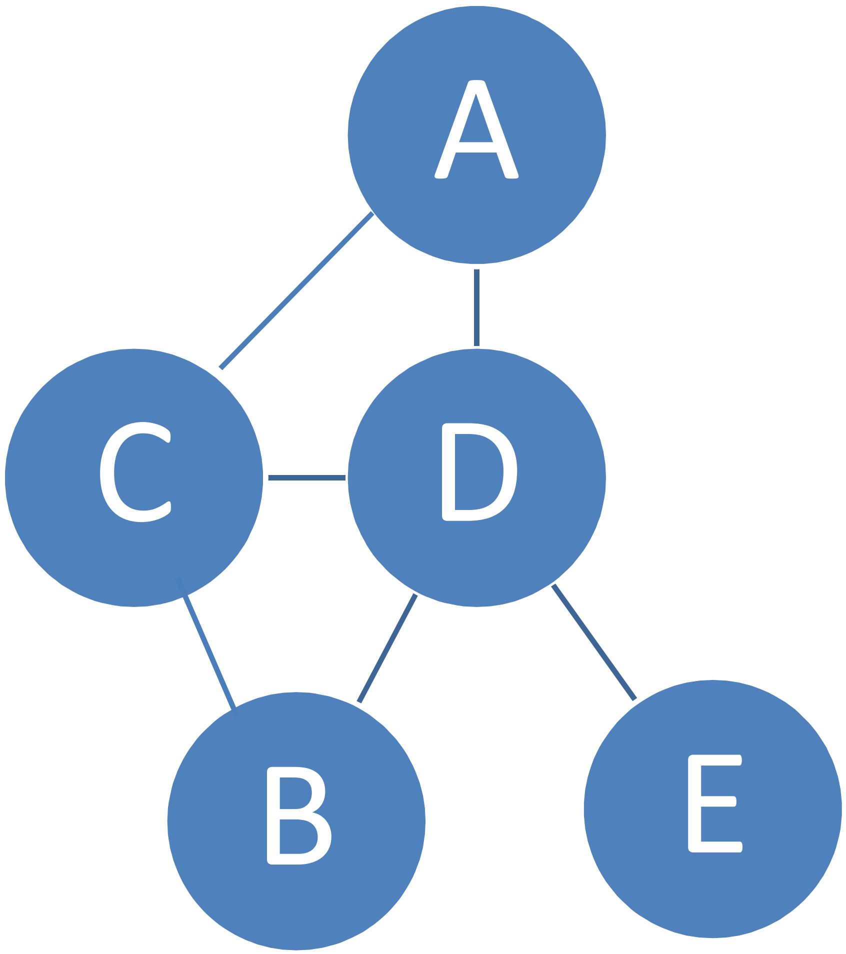
How to Computing the Betweenness values for the Girvan-Newman method ?

- It involves the set of all shortest paths and there are too many shortest paths when the graph is large.
- We need a systematical approach to compute the numbers of shortest paths and the flow.
- Determine the flow from A to other nodes
  - Perform a breadth-first search of the graph, starting at A.
  - Determine the number of shortest paths from A to each other node.
  - Based on these numbers, determine the amount of flow from A to all other nodes that uses each edge.

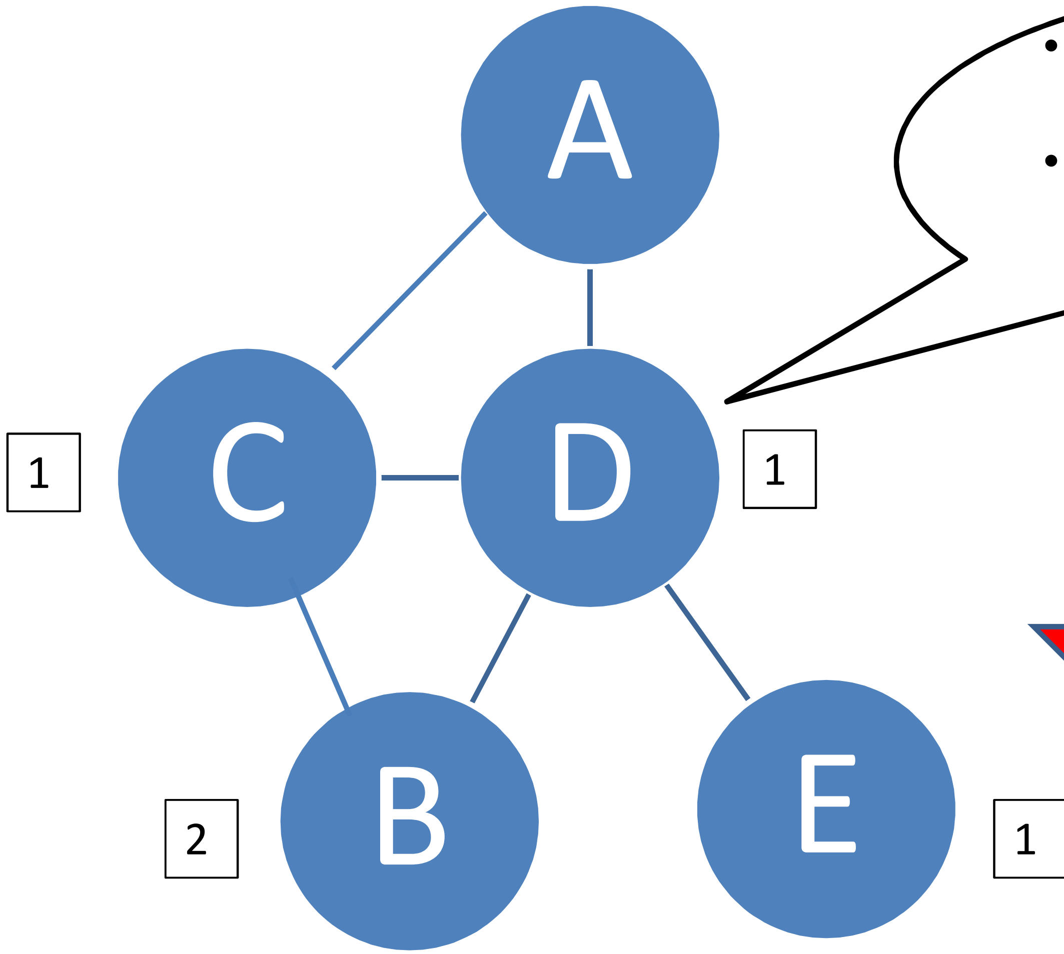
Example



Step 1 : BFS from A



Step 2 : determine the number of shortest paths from A to other nodes

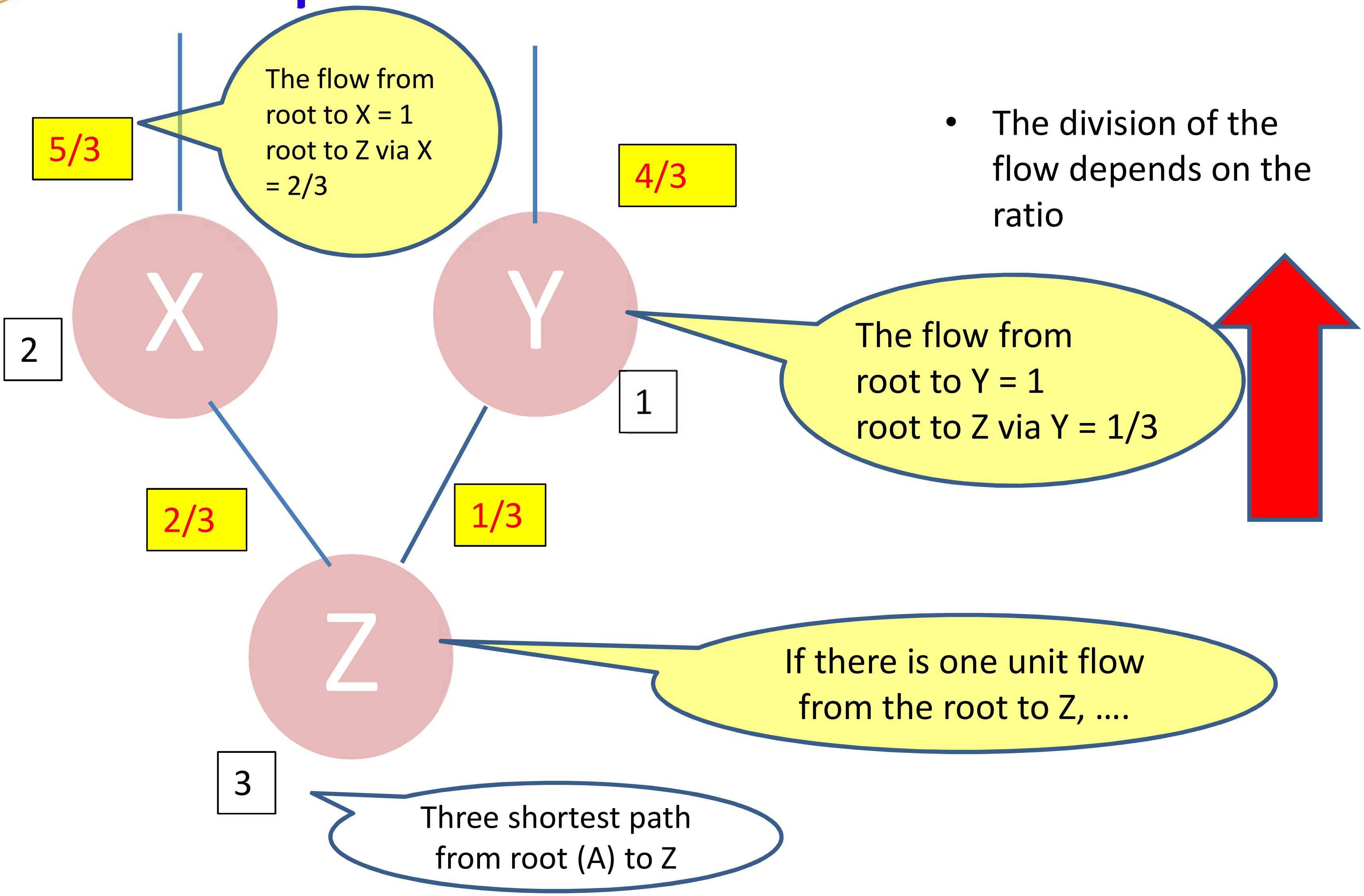


- 1 unit each for nodes in the first level
- The numbers in the next level are the sum of those in the previous level

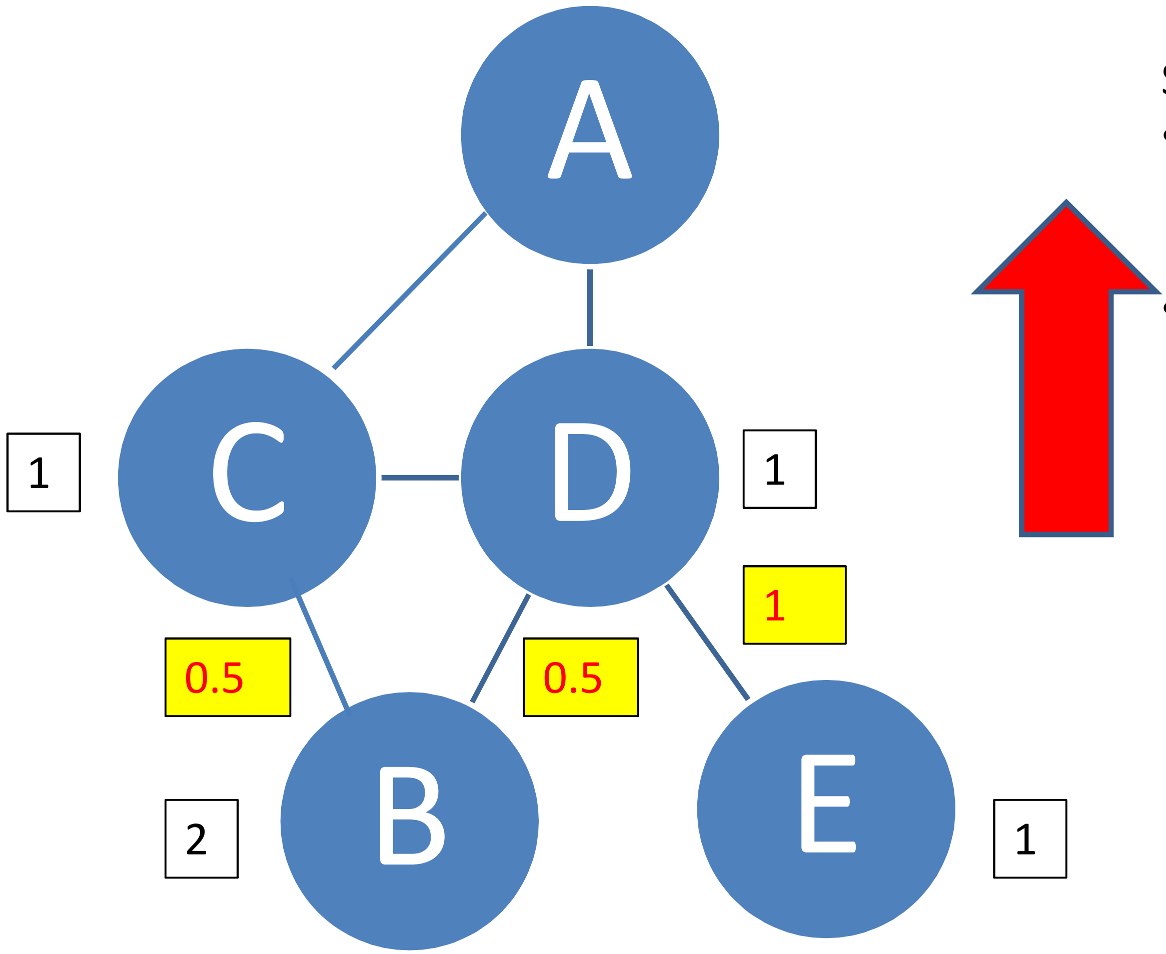
No. of shortest path
- From A to C is 1
- From A to D is 1
- From A to B is 2
- From A to E is 1

Step 3 : determine the flow

- The division of the flow depends on the ratio



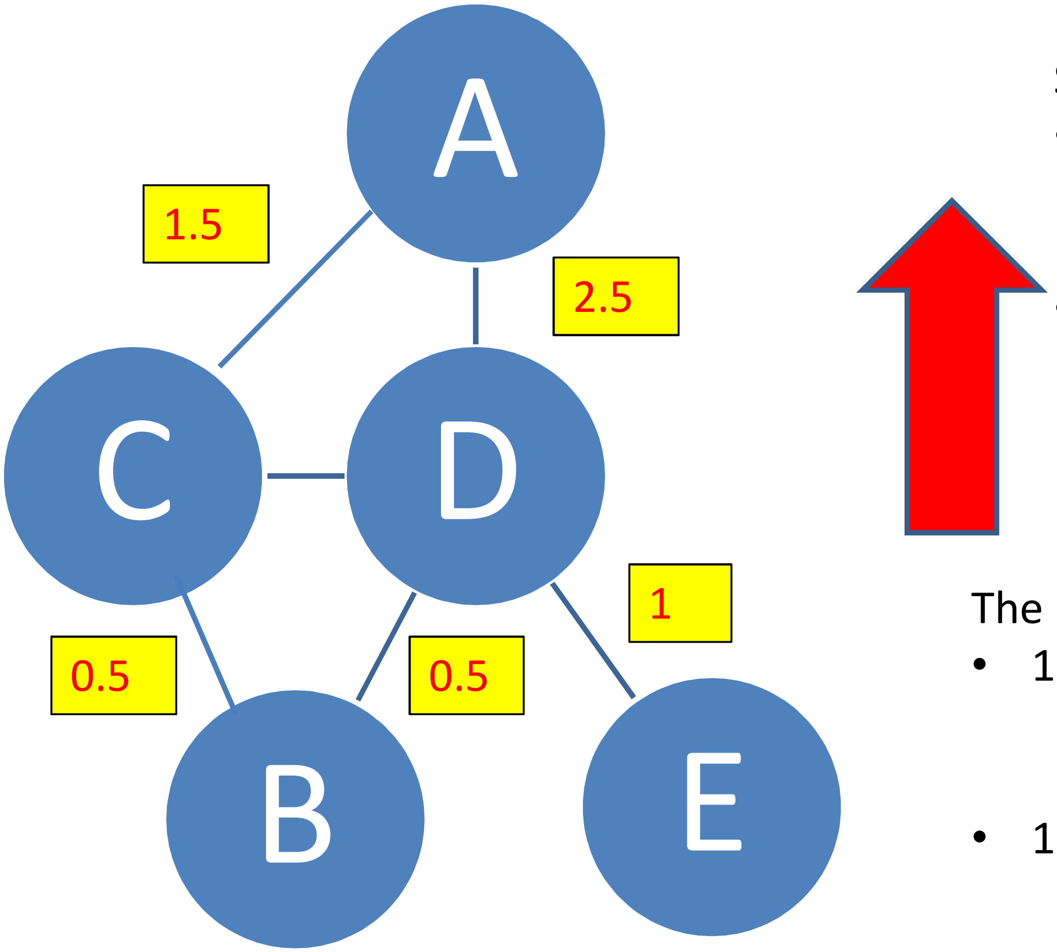
Step 3 : determine the flow



Start from the lowest layer

- 1 unit of flow from B (to A)
  - ½ goes via C and ½ goes via D
- 1 unit of flow from E (to A)
  - 1 goes via D

Step 3 : determine the flow



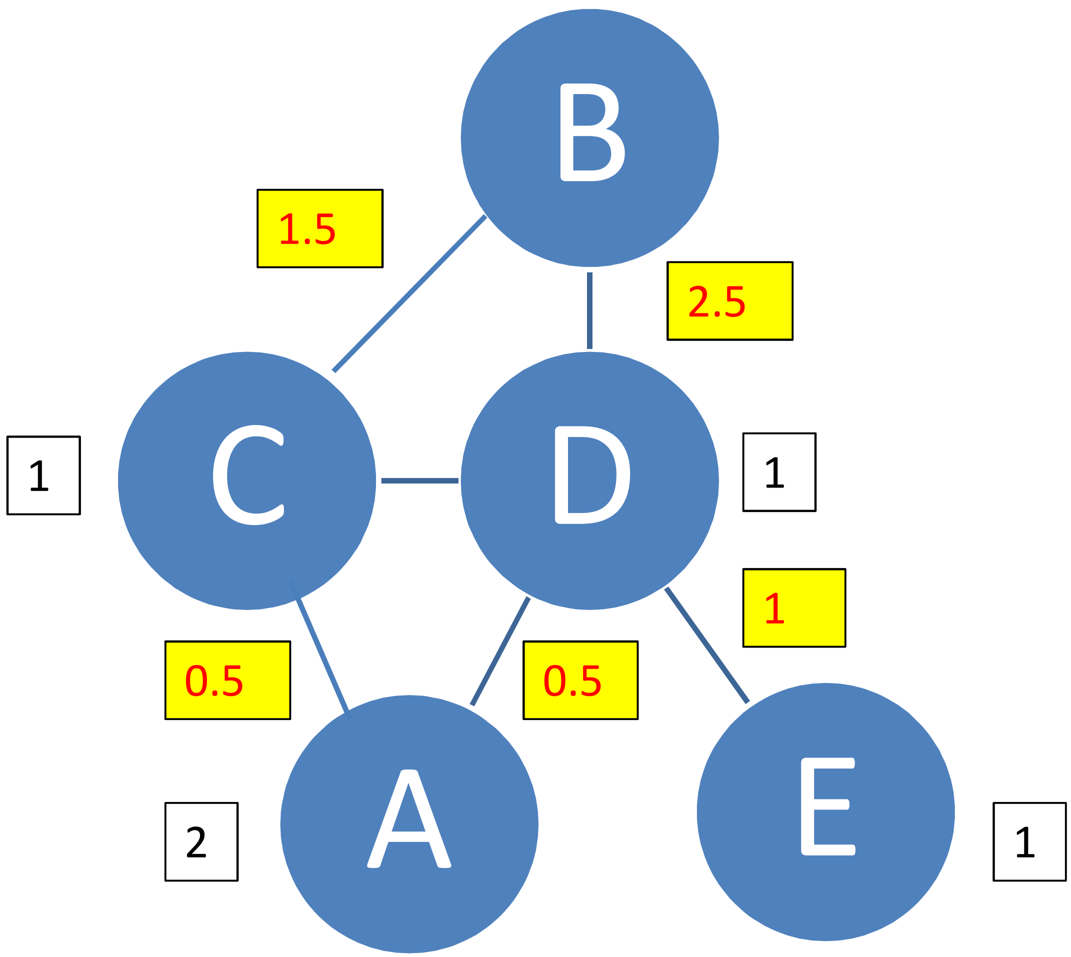
Start from the lowest layer

- 1 unit of flow from B
  - ½ goes to C and ½ goes to D
- 1 unit of flow from E
  - 1 goes to D

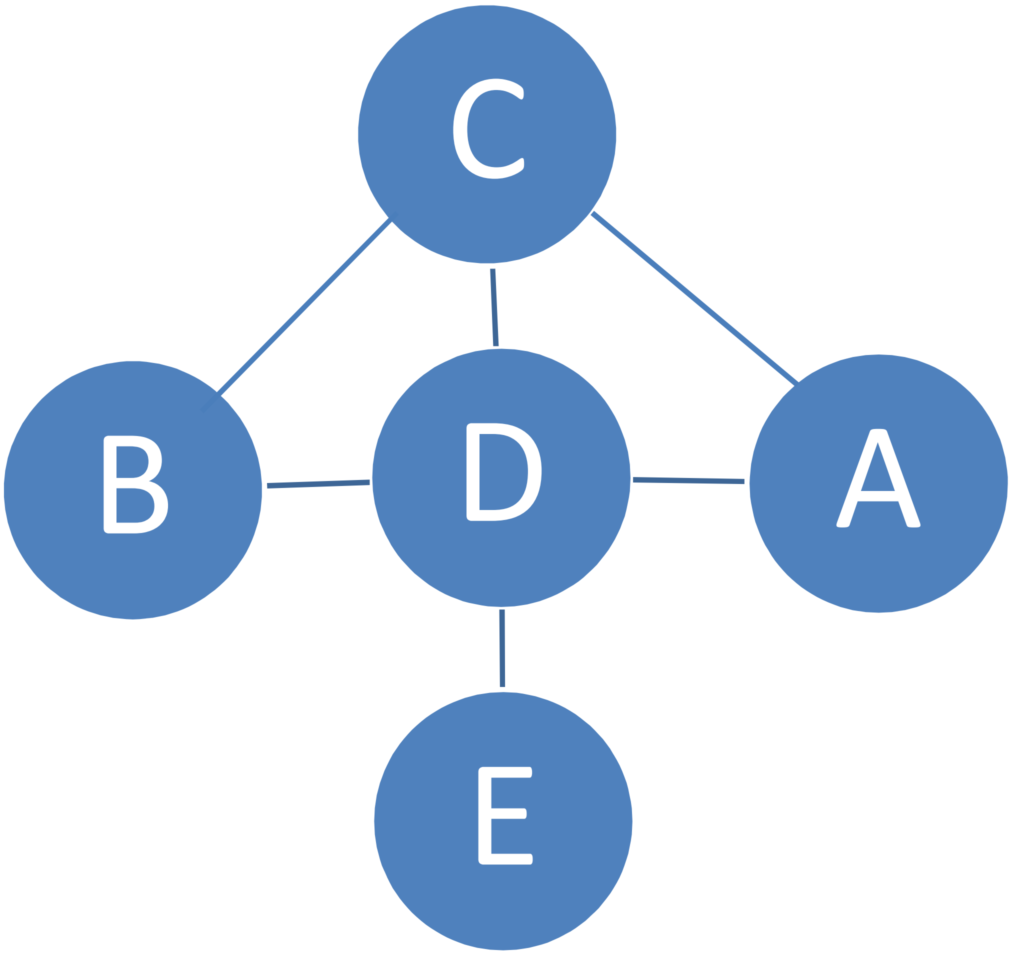
The next layer

- 1 unit of flow from C
  - Add to ½ units from lower layers
- 1 unit of flow from D
  - Add to ½ units from B and 1 unit from E

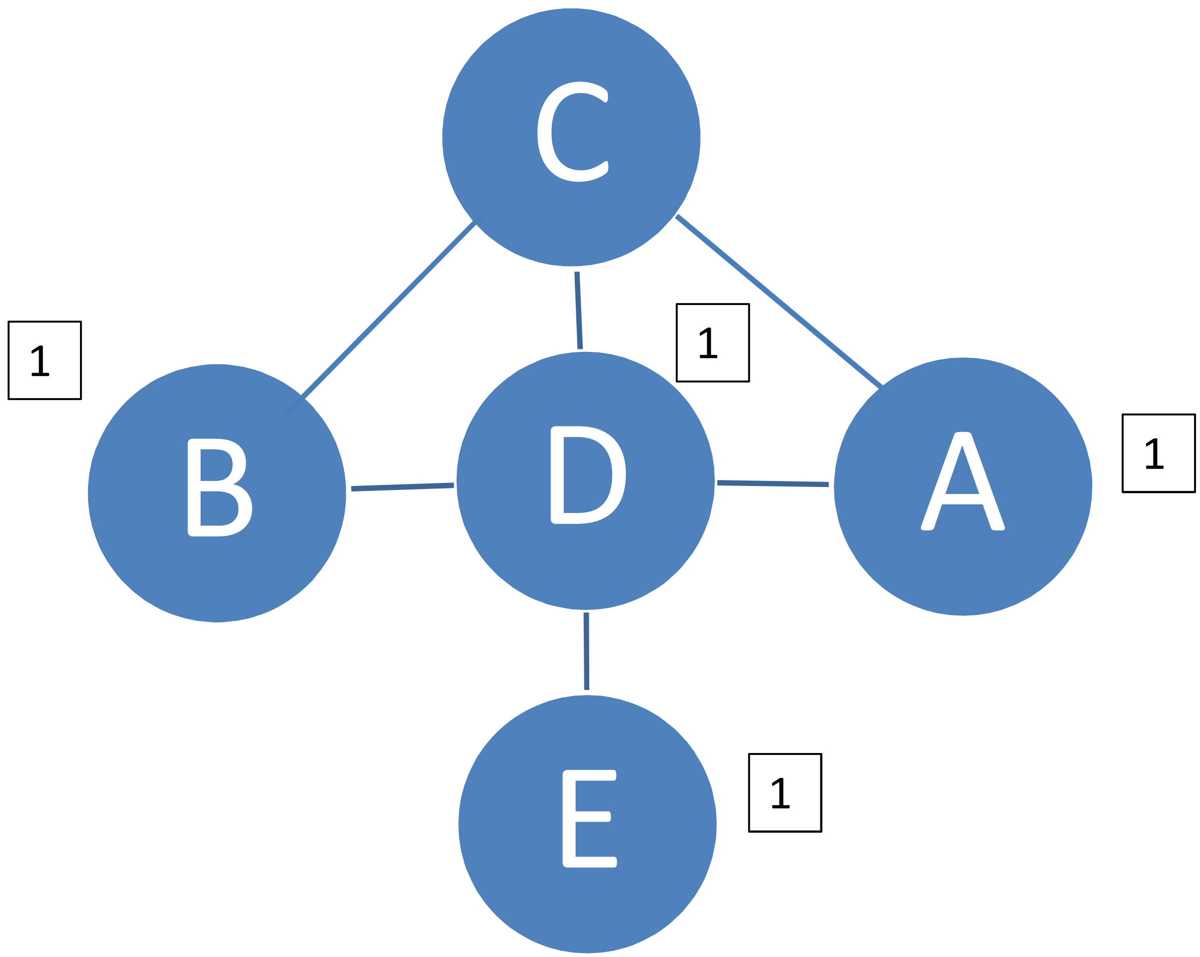
Repeat the steps starting from B



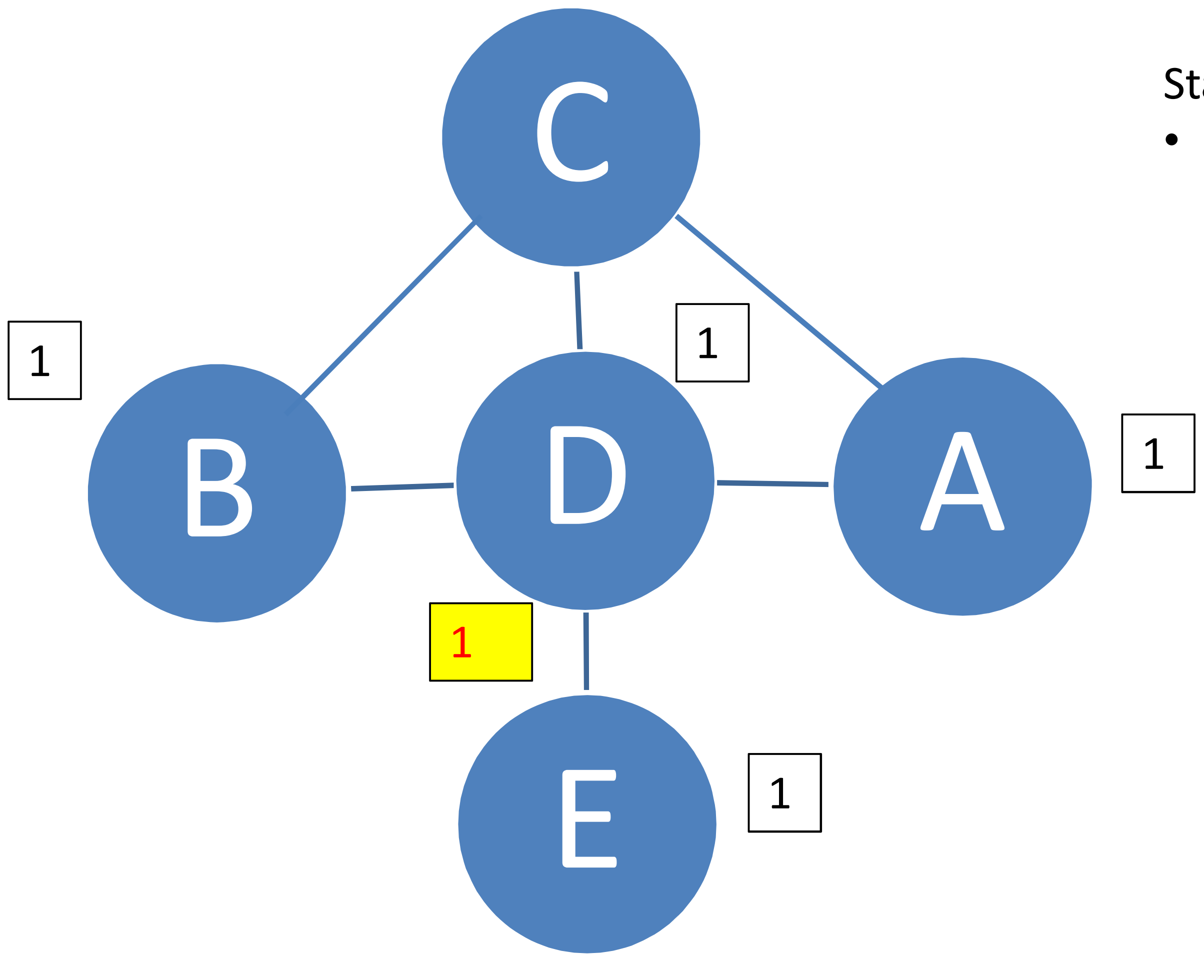
Starting from C



Starting from C



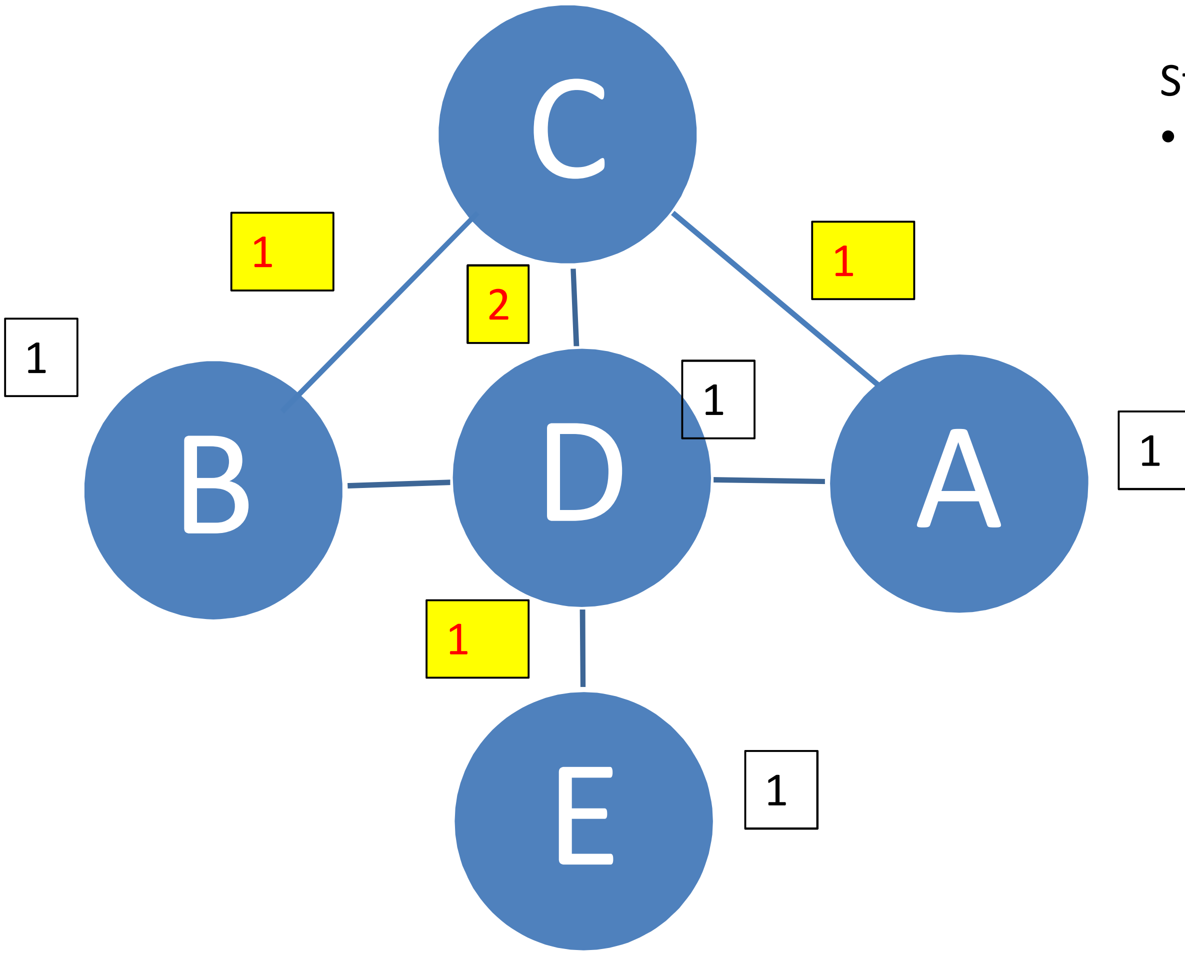
Starting from C



Start from the lowest layer

- 1 unit of flow from E
  - 1 goes to D

Starting from C



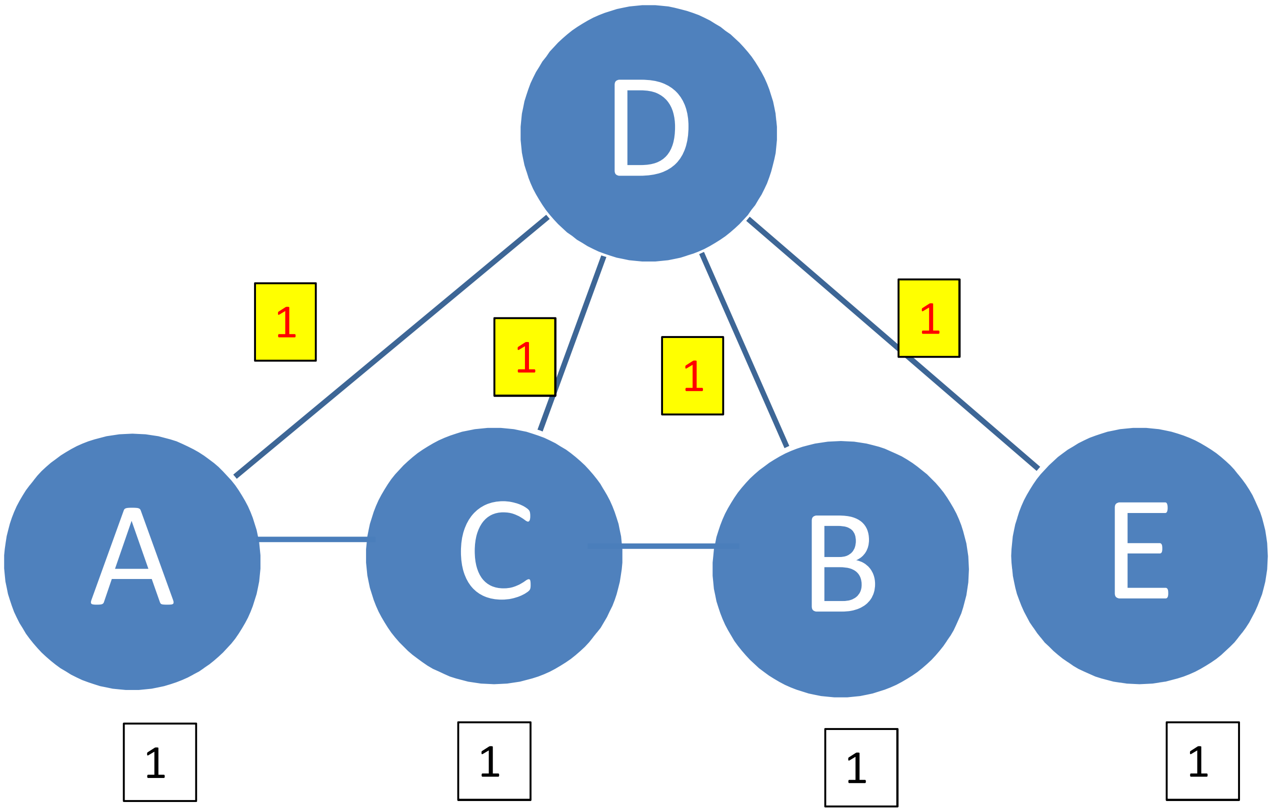
Start from the lowest layer

- 1 unit of flow from E
  - 1 goes to D

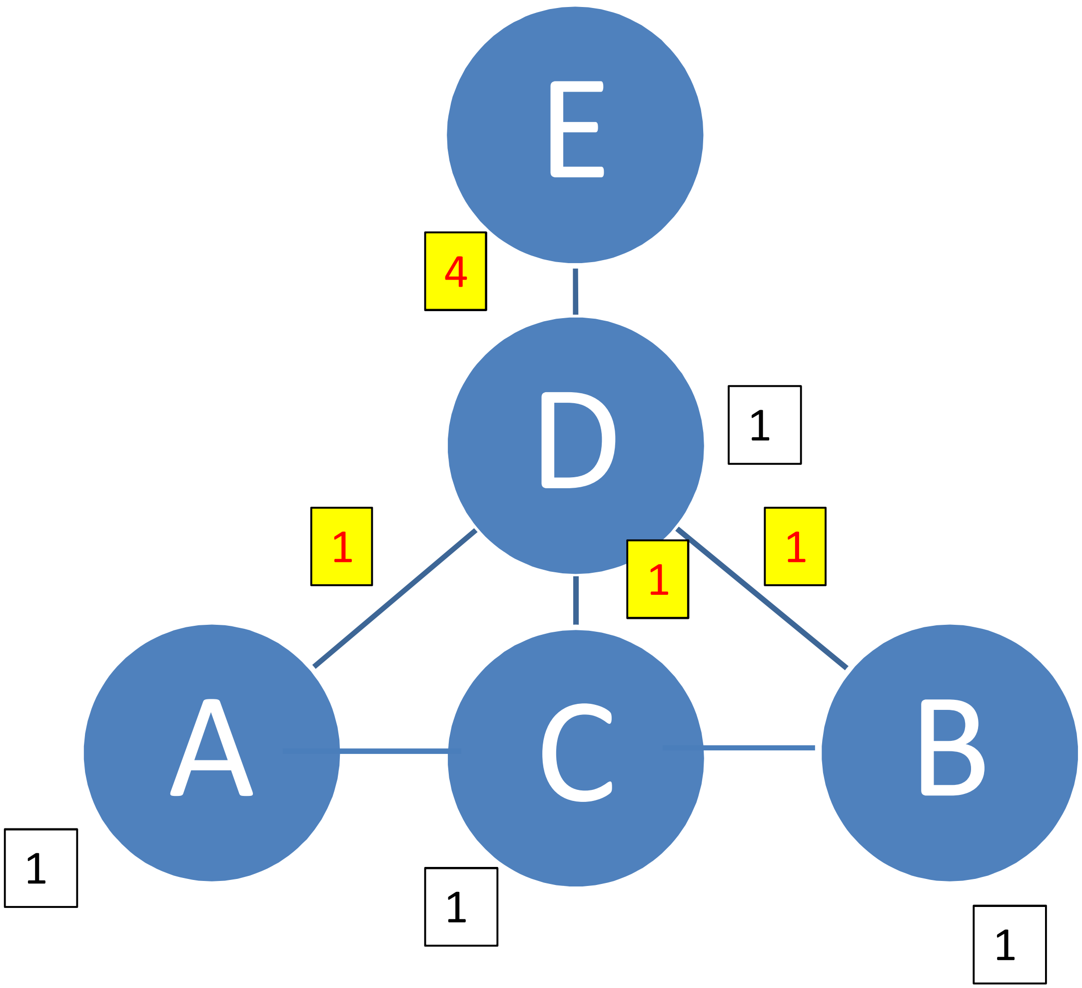
the next layer

- 1 unit of flow from B
  - 1 goes to C
- 1 unit of flow from A
  - 1 goes to C
- 1 unit of flow from D
  - Add 1 unit from E and 2 units go to C

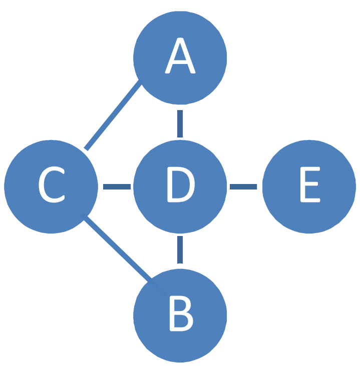
Start from D



Start from E



	Edge AC	Edge AD	Edge CD	Edge CB	Edge BD	Edge DE
BFS from A	1.5	2.5	0	0.5	0.5	1
BFS from B	0.5	0.5	0	1.5	2.5	1
BFS from C	1	0	2	1	0	1
BFS from D	0	1	1	0	1	1
BFS from E	0	1	1	0	1	4
Total	3	5	4	3	5	8
Total/2	1.5	2.5	2	1.5	2.5	4



We have to divide the results by 2 as we have counted the flow twice using from X to Y and from Y to X

The edge betweenness

1 Shortest path between

A – C

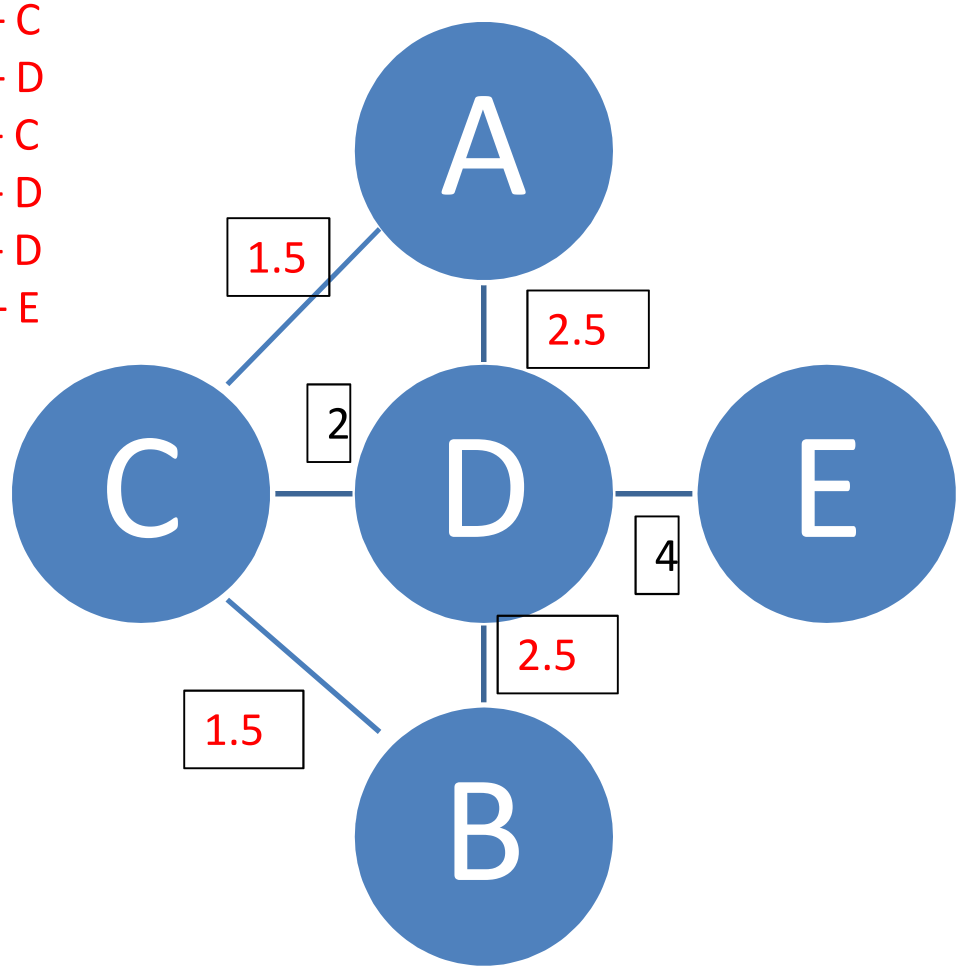
A – D

B – C

B – D

C – D

D – E



2 Shortest paths between

A and B

- A– C – B (1/2 units)
- A –D – B (1/2 units)

1 Shortest path between

A and E

- A– D – E (1 unit)

1 Shortest path between

B and E

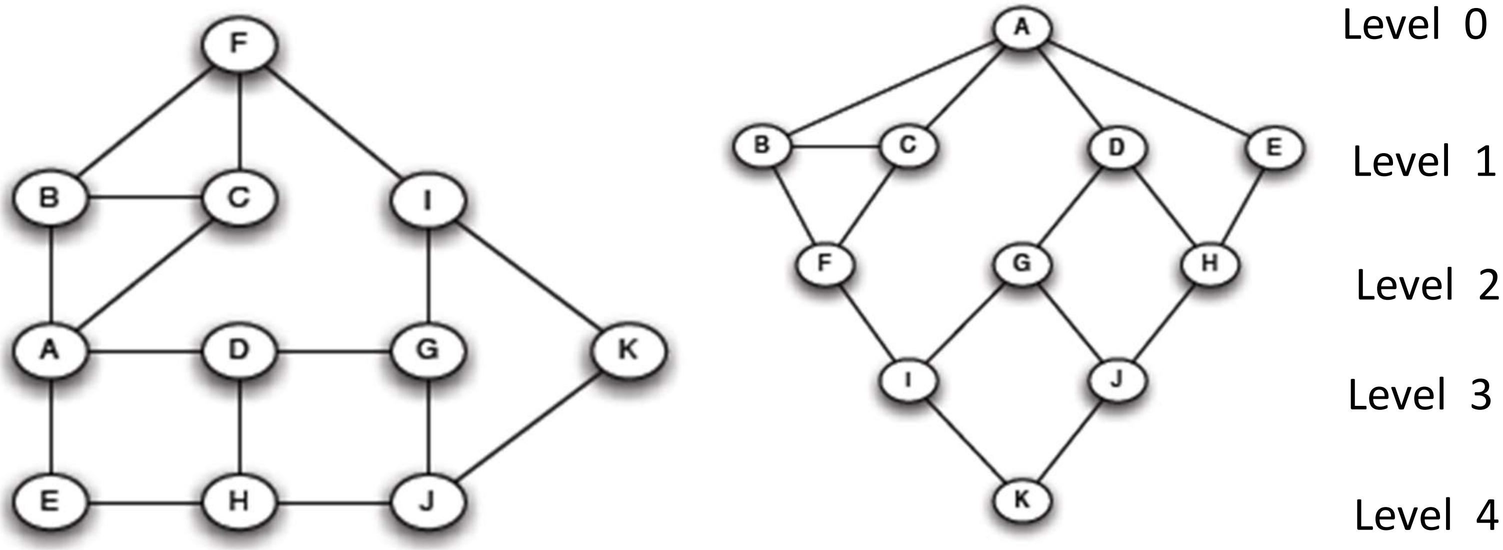
- B– D – E (1 unit)

1 Shortest path between

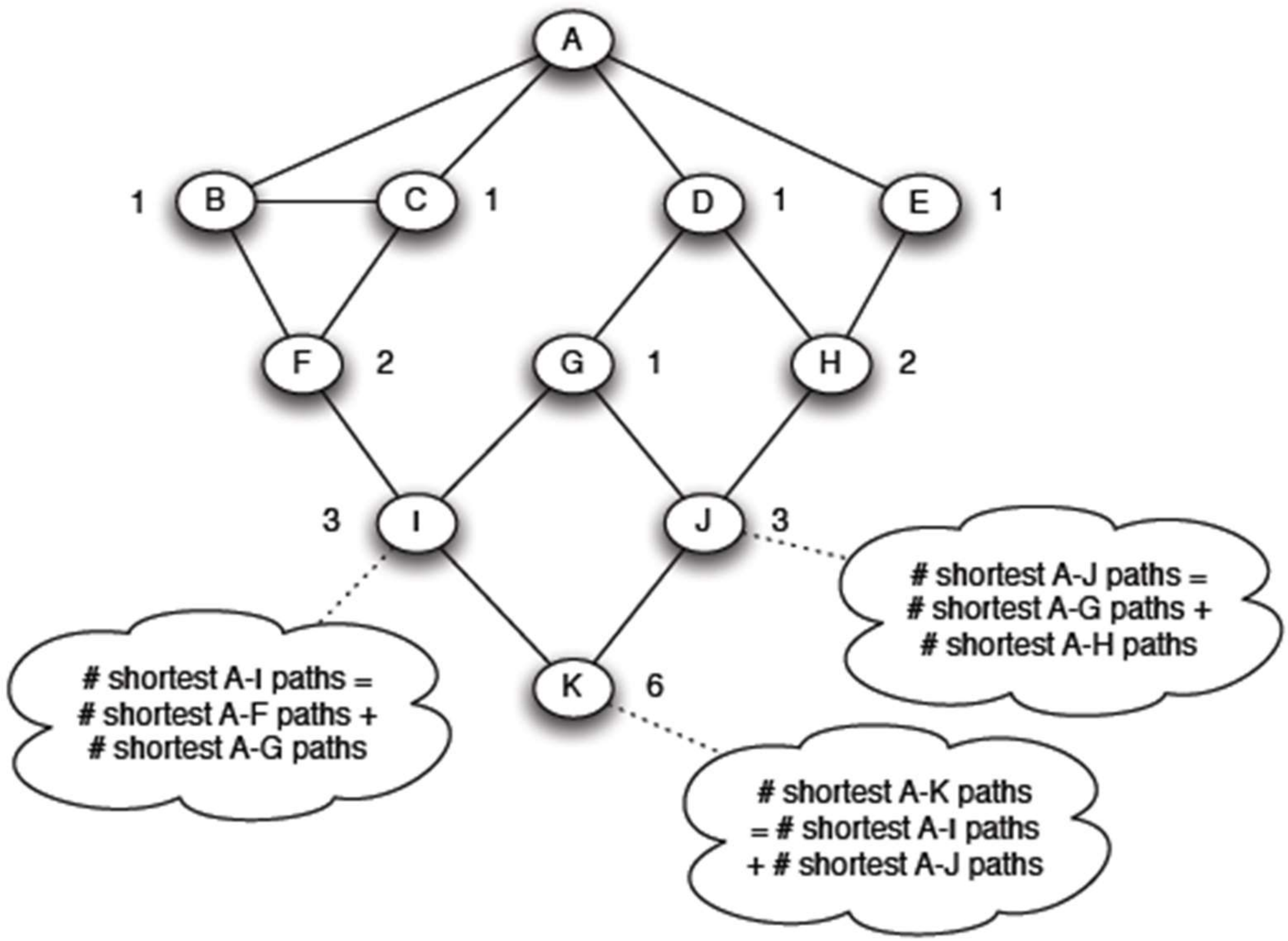
C and E

- C– D – E (1 unit)

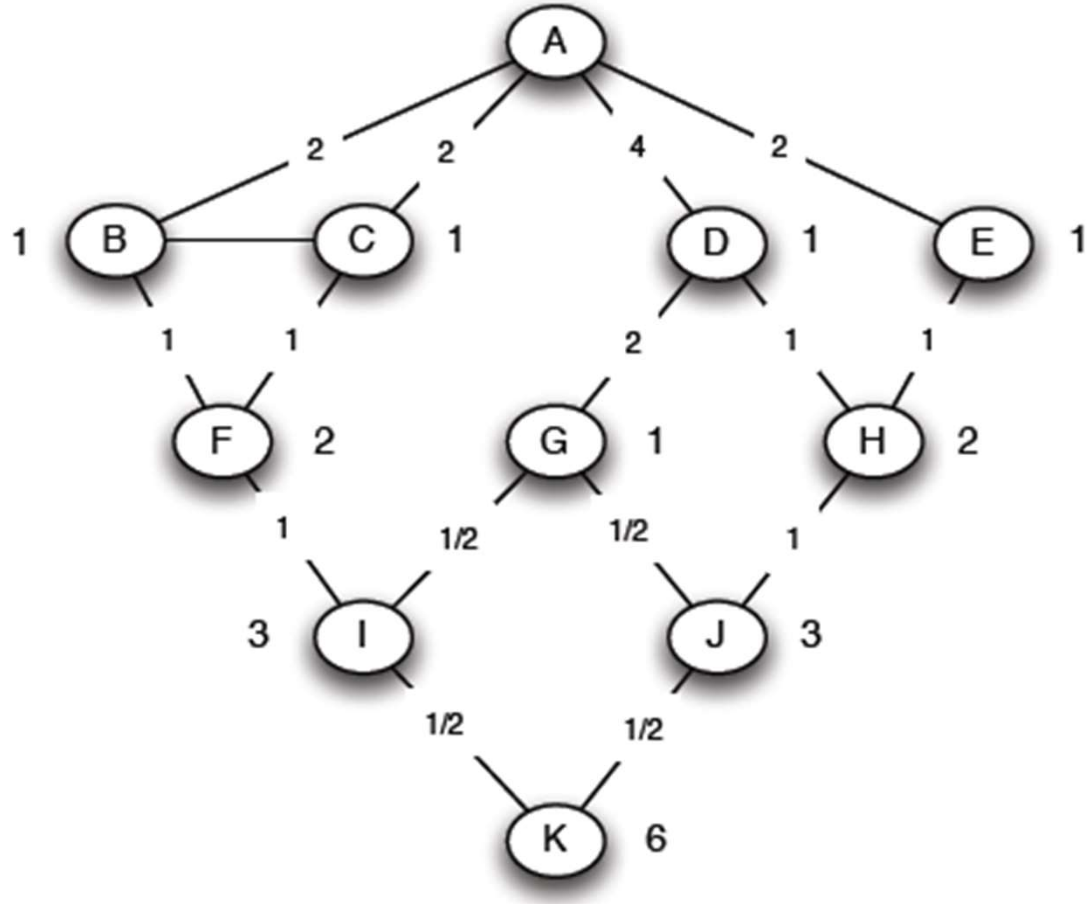
Another example
Step 1 : BFS



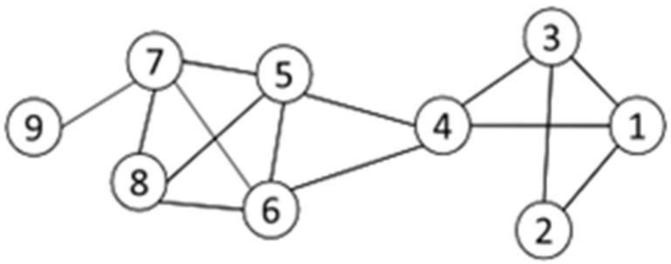
Step 2 : determine the number of shortest paths from A to other nodes



Step 3 : Determine Flow Values



Divisive clustering based on edge betweenness



Initial betweenness value

Table 3.3: Edge Betweenness

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

After remove e(4,5), the betweenness of e(4, 6) becomes 20, which is the highest;

After remove e(4,6), the edge e(7,9) has the highest betweenness value 4, and should be removed.

33

Idea: progressively removing edges with the highest betweenness