Online Social Networks and Media

Positive and Negative Edges

Strong and Weak Edges

POSITIVE AND NEGATIVE TIES

Structural Balance Theory

- originated in social psychology in the mid-20th-century, by Heider in the 1940s
- graph-theoretic approach by Cartwright and Harary in the 1960s

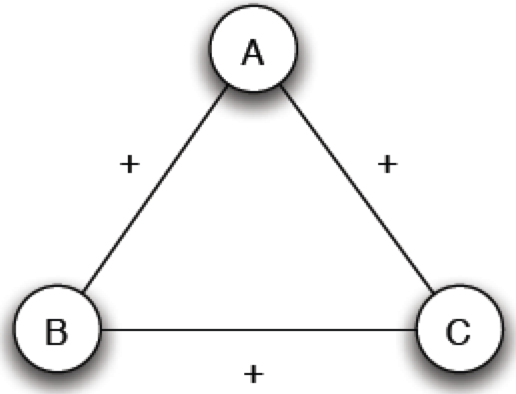
- considers the possible ways in which triangles on three individuals can be signed

- Lets look at all possible relationships *between 3 people* => 4 cases

- See if all are *equally possible* (local property)

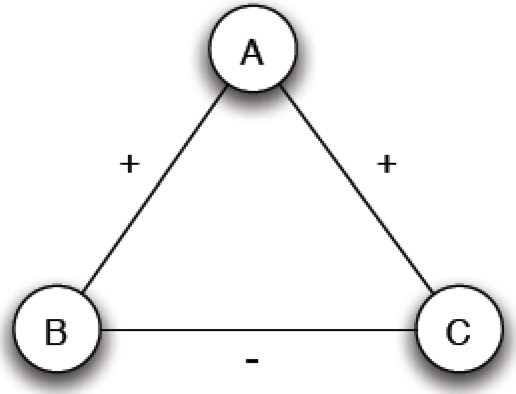
Structural Balance

Case (a): 3 +



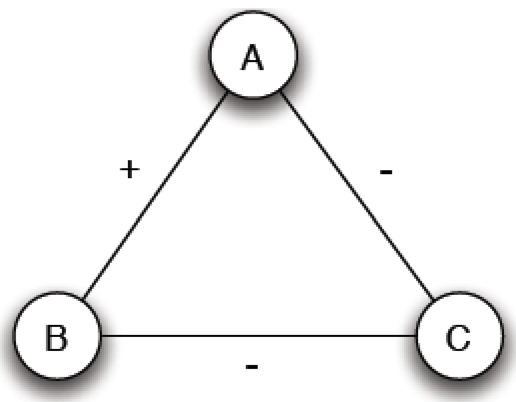
Mutual friends

Case (b): 2 +, 1 -



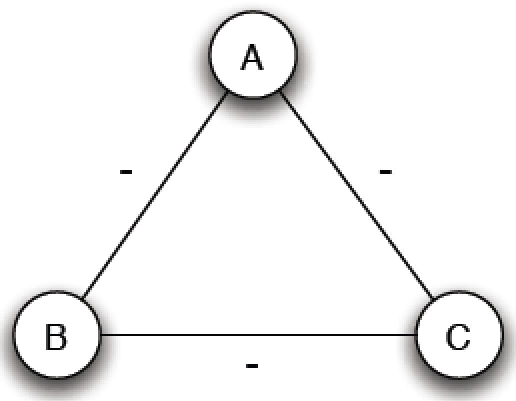
A is friend with B and C, but B and C do not get well together

Case (c): 1 +, 2 -



A and B are friends with a mutual enemy

Case (d): 3 -

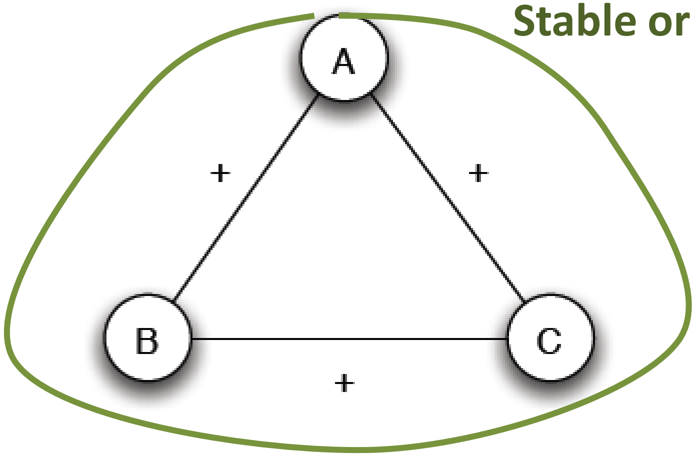


Mutual enemies

Structural Balance

Case (a): 3 +

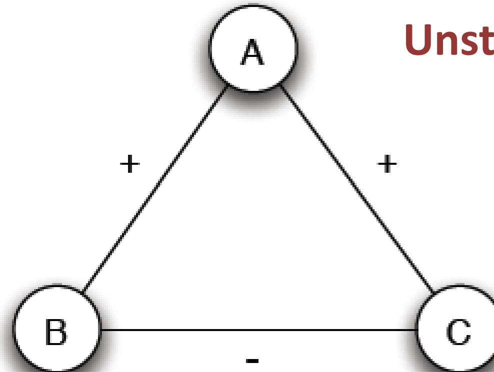
Stable or balanced



Mutual friends
“the friend of my friend is my friend,”

Case (b): 2 +, 1 -

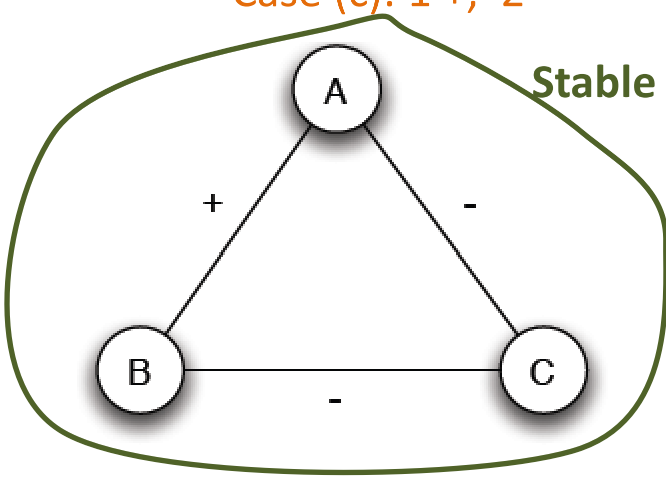
Unstable



A is friend with B and C, but B and C do not get well together
Implicit force to make B and C friends (- => +) or turn one of the + to -

Case (c): 1 +, 2 -

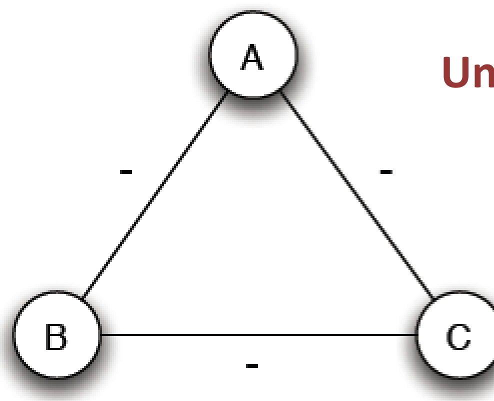
Stable or balanced



A and B are friends with a mutual enemy
“the enemy of my enemy is my friend”

Case (d): 3 -

Unstable

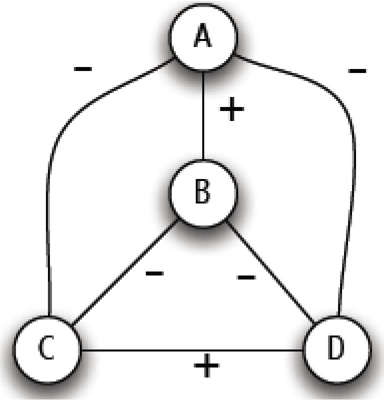


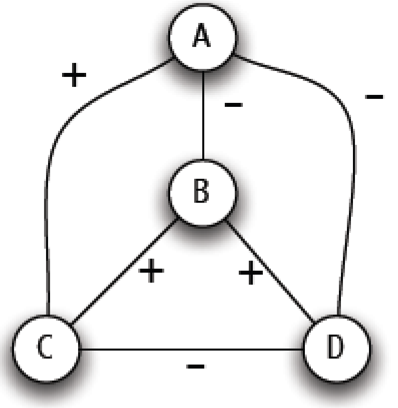
Mutual enemies
Forces to team up against the third (turn 1 – to +)

Structural Balance

A labeled complete graph is balanced if every one of its triangles is balanced

Structural Balance Property: For every set of three nodes, if we consider the three edges connecting them, either all three of these are labeled +, or else exactly one of them is labeled – (odd number of +)



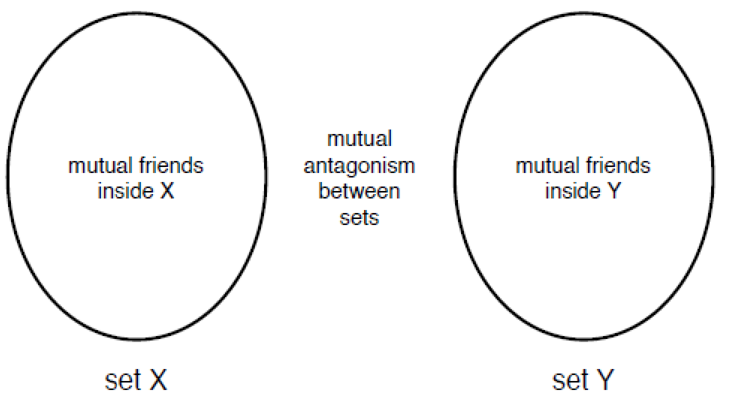


What does a balanced network look like?

The Structure of Balanced Networks

Balance Theorem: If a labeled *complete* graph is balanced,
(a) all pairs of nodes are friends, or
(b) the nodes can be divided *into two groups* X and Y, such that every pair of nodes in X like each other, every pair of nodes in Y like each other, and every one in X is the enemy of every one in Y.

From a local to a global property



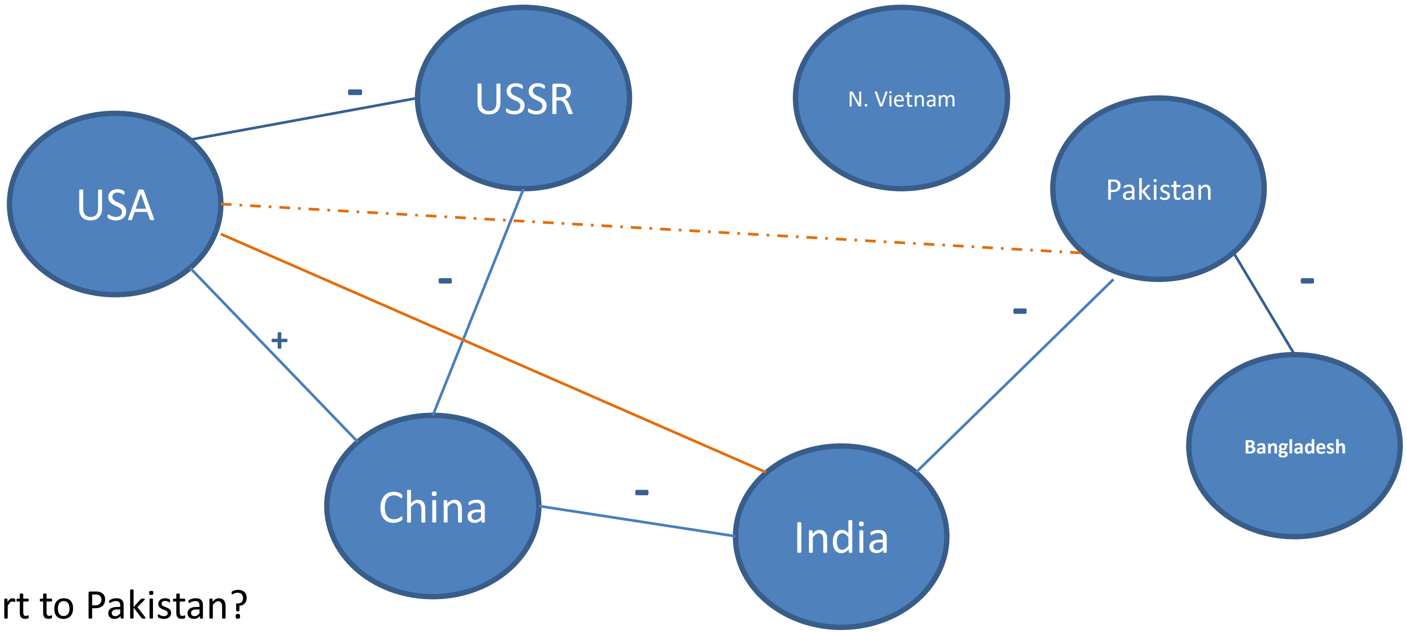
Proof ...

Applications of Structural Balance

- How a network evolves over time
- Political science: International relationships

Applications of Structural Balance

The conflict of Bangladesh's separation from Pakistan in 1972 (1)

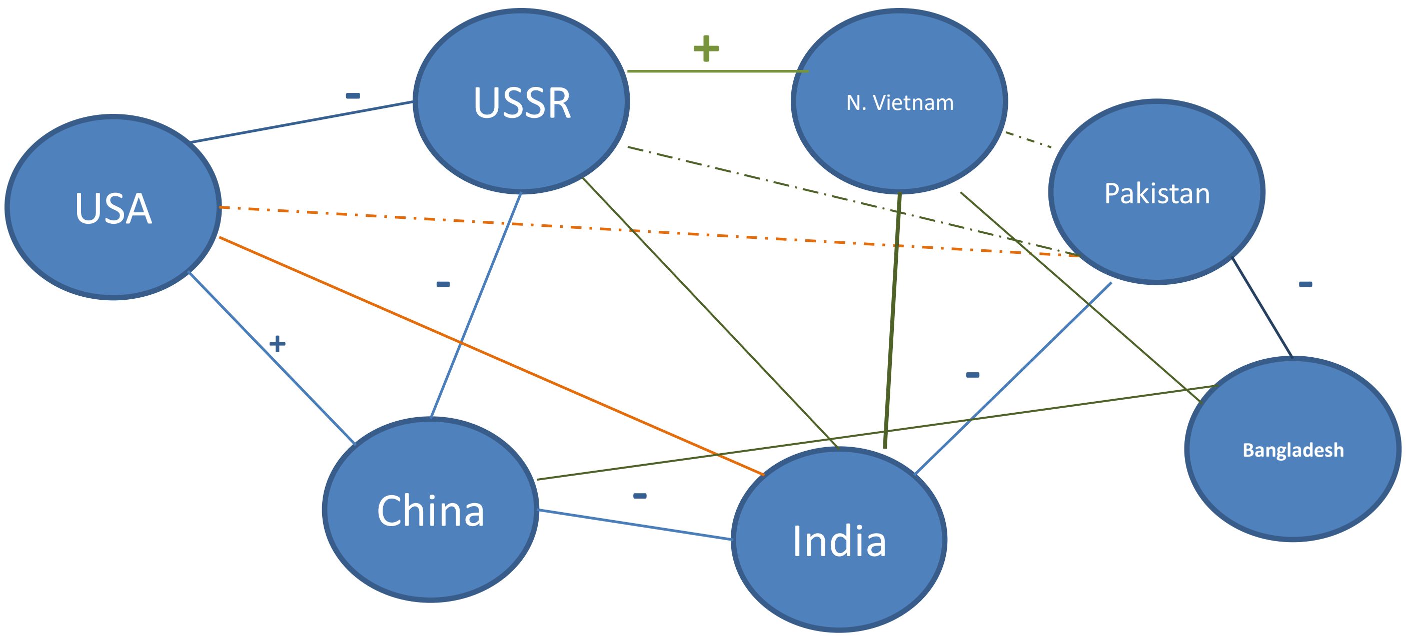


USA support to Pakistan?
"[T]he United States's somewhat surprising support of Pakistan ... becomes less surprising when one considers that the USSR was China's enemy, China was India's foe, and India had traditionally bad relations with Pakistan. Since the U.S. was at that time improving its relations with China, it **supported the enemies of China's enemies**.
Further reverberations of this strange political constellation became inevitable: North Vietnam made friendly gestures toward India, Pakistan severed diplomatic relations with those countries of the Eastern Bloc which recognized Bangladesh, and China vetoed the acceptance of Bangladesh into the U.N."

Applications of Structural Balance

✓ International relationships (I)

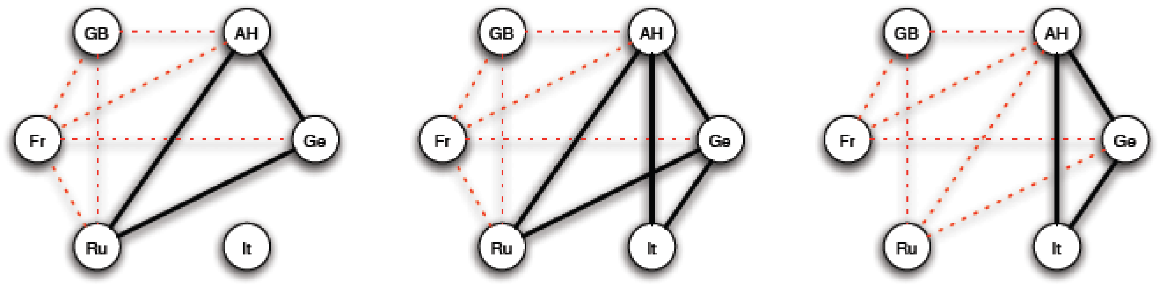
The conflict of Bangladesh's separation from Pakistan in 1972 (II)



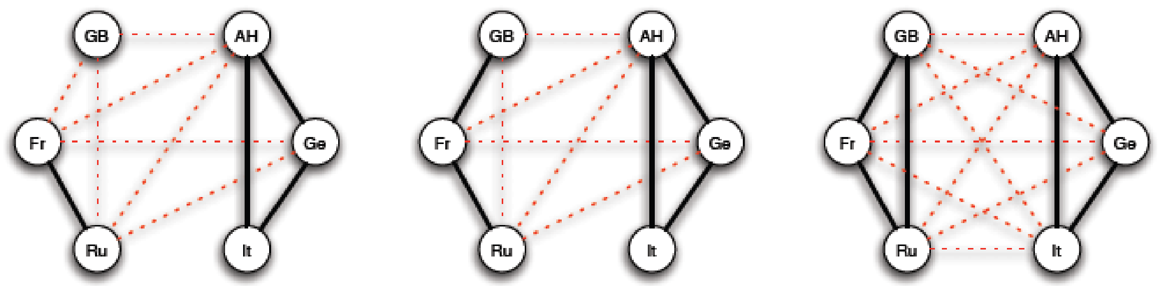
China?

Applications of Structural Balance

International relationships (II)



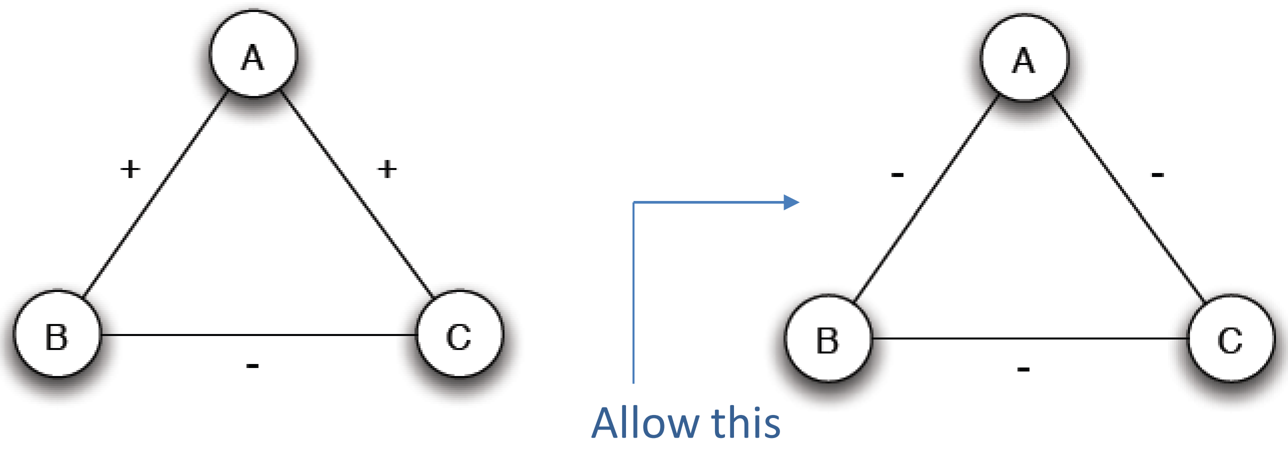
(a) *Three Emperors' League 1872–81* (b) *Triple Alliance 1882* (c) *German-Russian Lapse 1890*



(d) *French-Russian Alliance 1891–94* (e) *Entente Cordiale 1904* (f) *British Russian Alliance 1907*

Figure 5.5: The evolution of alliances in Europe, 1872-1907 (the nations GB, Fr, Ru, It, Ge, and AH are Great Britain, France, Russia, Italy, Germany, and Austria-Hungary respectively). Solid dark edges indicate friendship while dotted red edges indicate enmity. Note how the network slides into a balanced labeling — and into World War I. This figure and example are from Antal, Krapivsky, and Redner [20].

A Weaker Form of Structural Balance



Weak Structural Balance Property: There is no set of three nodes such that the edges among them consist of exactly two positive edges and one negative edge

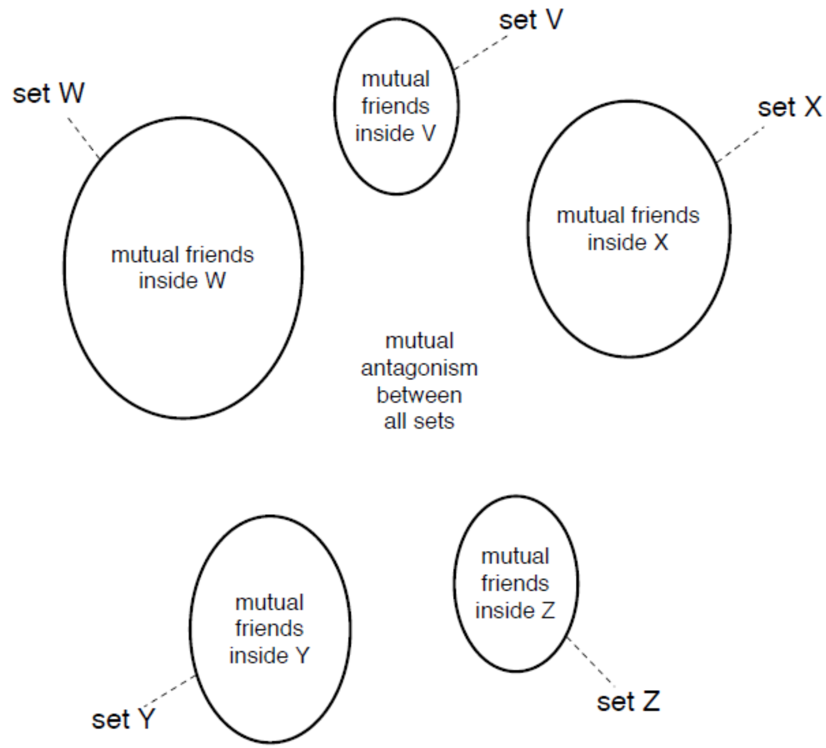
A Weaker Form of Structural Balance

Weakly Balance Theorem: If a labeled complete graph is weakly balanced, its nodes can be divided *into groups* in such a way that every two nodes belonging to the same group are friends, and every two nodes belonging to different groups are enemies.

From a local to a global property

Proof …

A Weaker Form of Structural Balance



Generalizing

1. Non-complete graphs

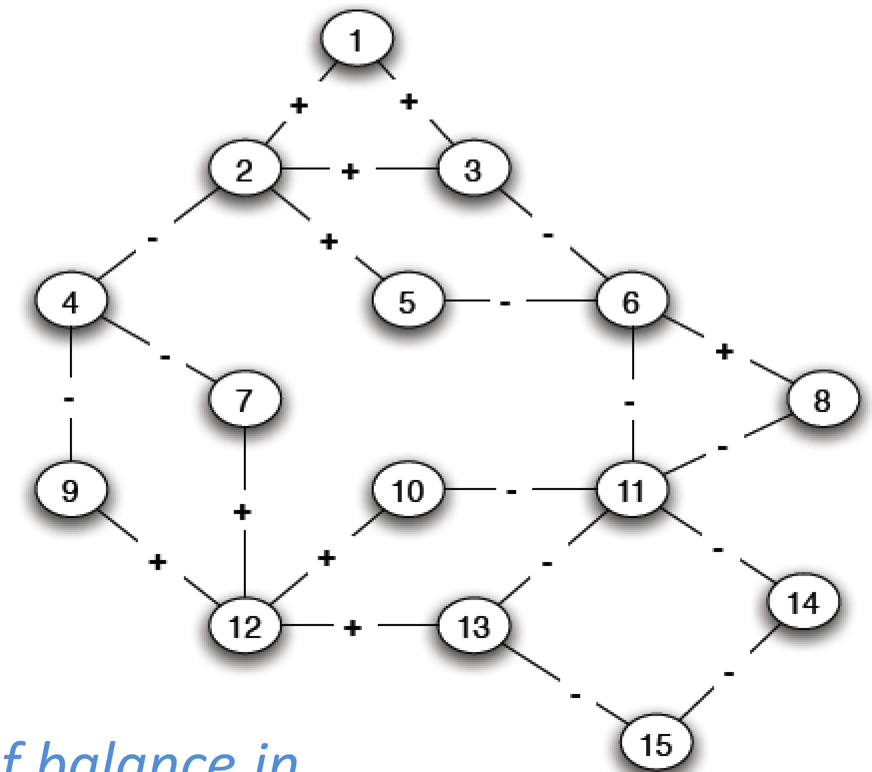
2. Instead of all triangles, “most” triangles, approximately divide the graph

We shall use the original (“non-weak” definition of structural balance)

Structural Balance in Arbitrary Graphs

Thee possible relations

- Positive edge
- Negative edge
- Absence of an edge

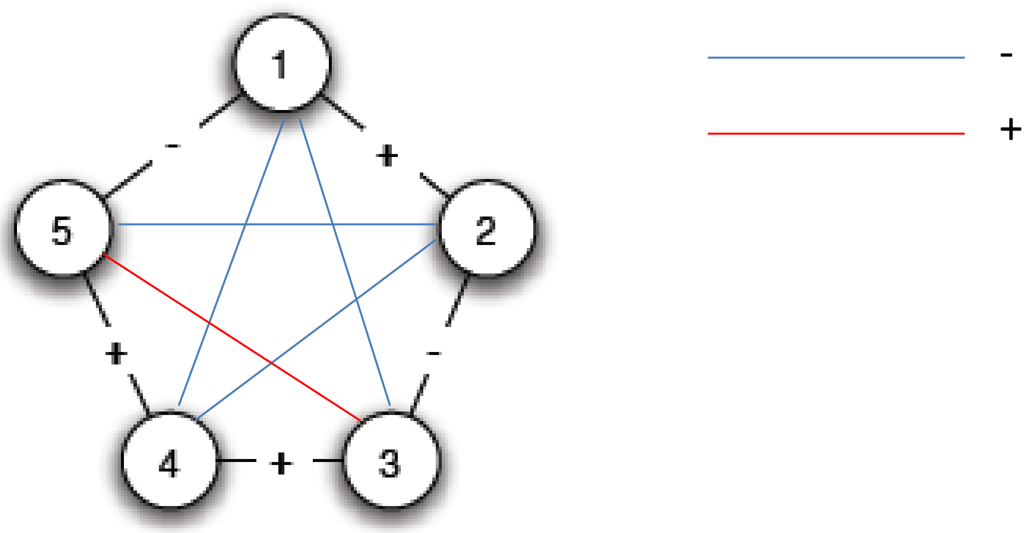


What is a good definition of balance in a non-complete graph?

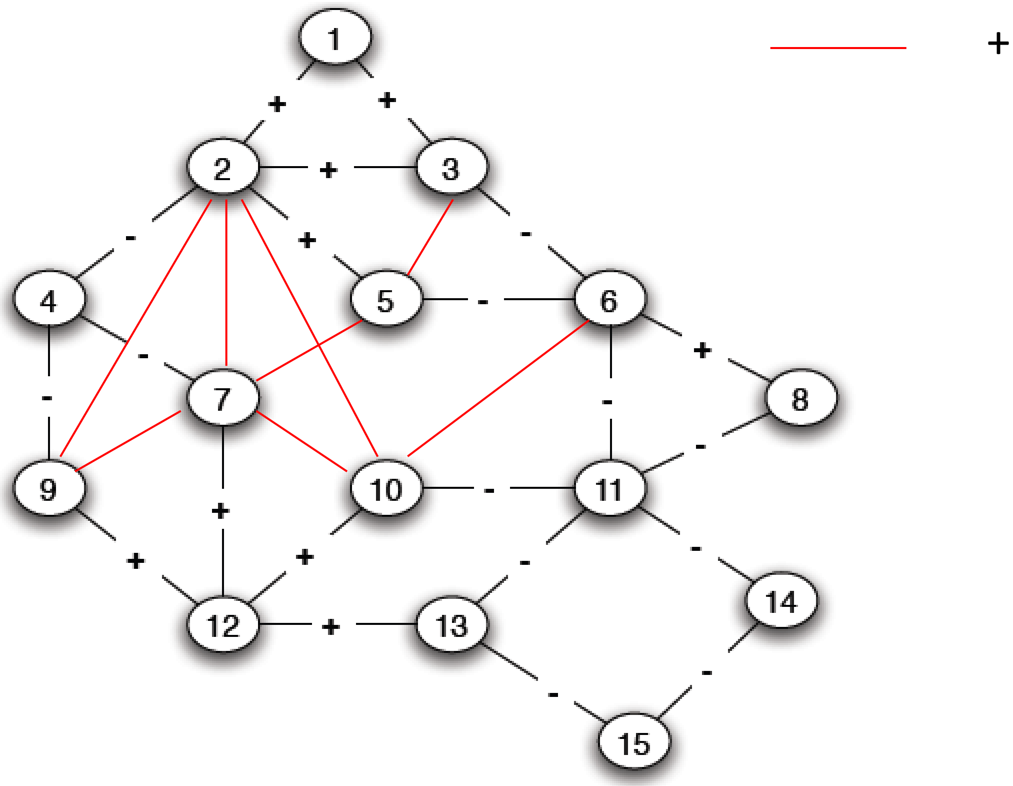
Balance Definition for General Graphs

1. Based on triangles (local view)
2. Division of the network (global view)

A (non-complete) graph is balanced if it can be completed by adding edges to form a signed complete graph that is balanced



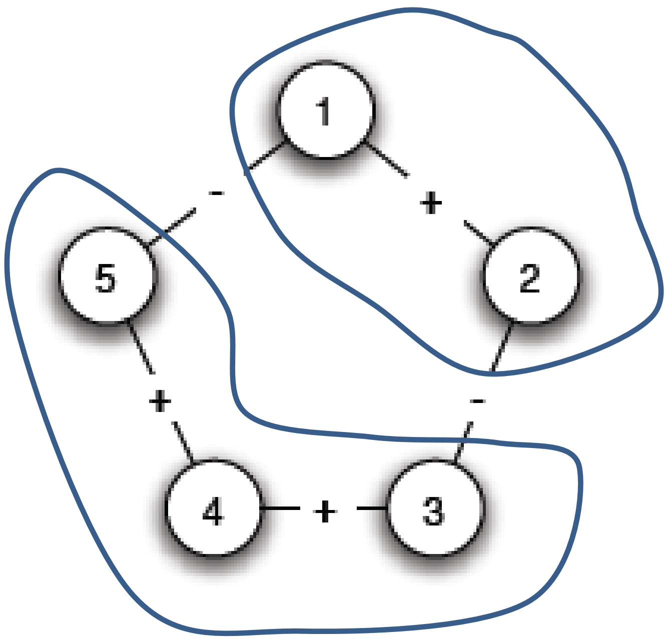
Balance Definition for General Graphs



Balance Definition for General Graphs

1. Based on triangles (local view)
2. Division of the network (global view)

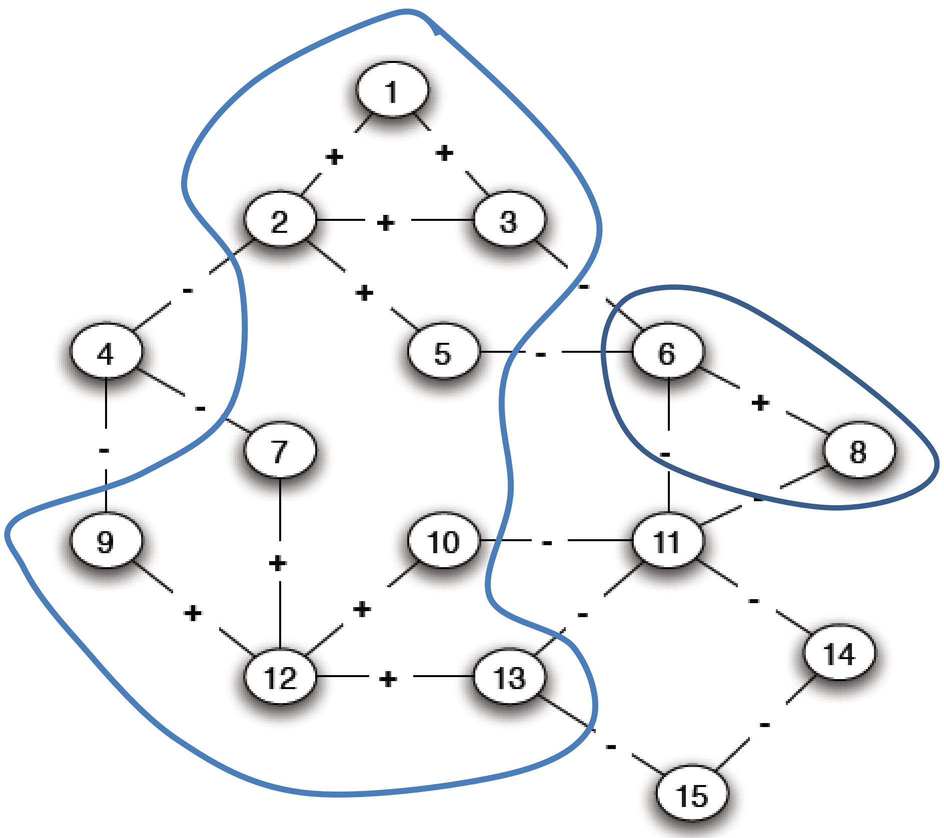
A (non-complete) graph is balanced if it possible to divide the nodes into two sets X and Y, such that any edge with both ends inside X or both ends inside Y is positive and any edge with one end in X and one end in Y is negative



The **two definition** are **equivalent**:
An arbitrary signed graph is balanced under the first definition, if and only if, it is balanced under the second definitions

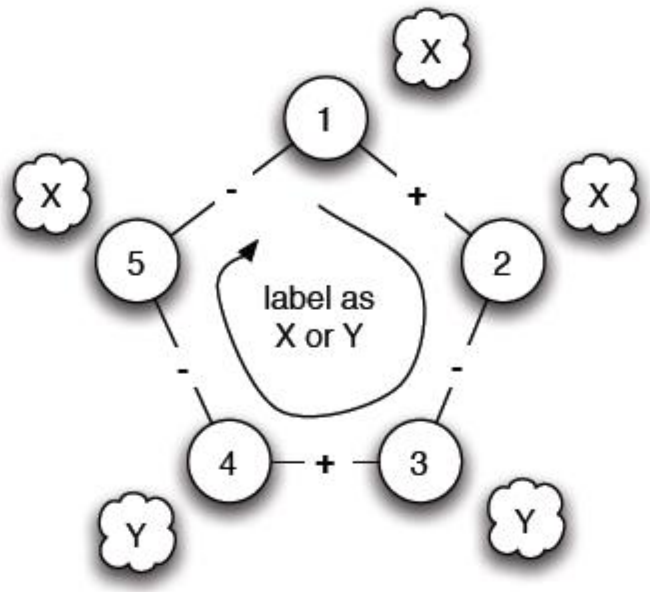
Balance Definition for General Graphs

Algorithm for dividing the nodes?



Balance Characterization

What prevents a network from being balanced?



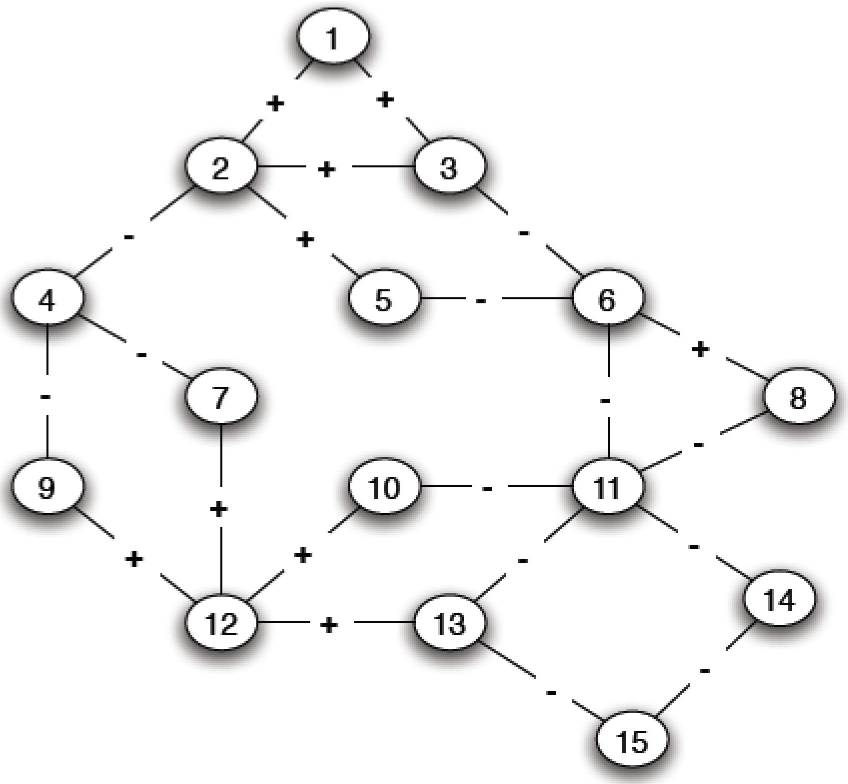
- Start from a node and place nodes in X or Y
- Every time we cross a negative edge, change the set

Cycle with odd number of negative edges

Balance Definition for General Graphs

Cycle with odd number of - => unbalanced

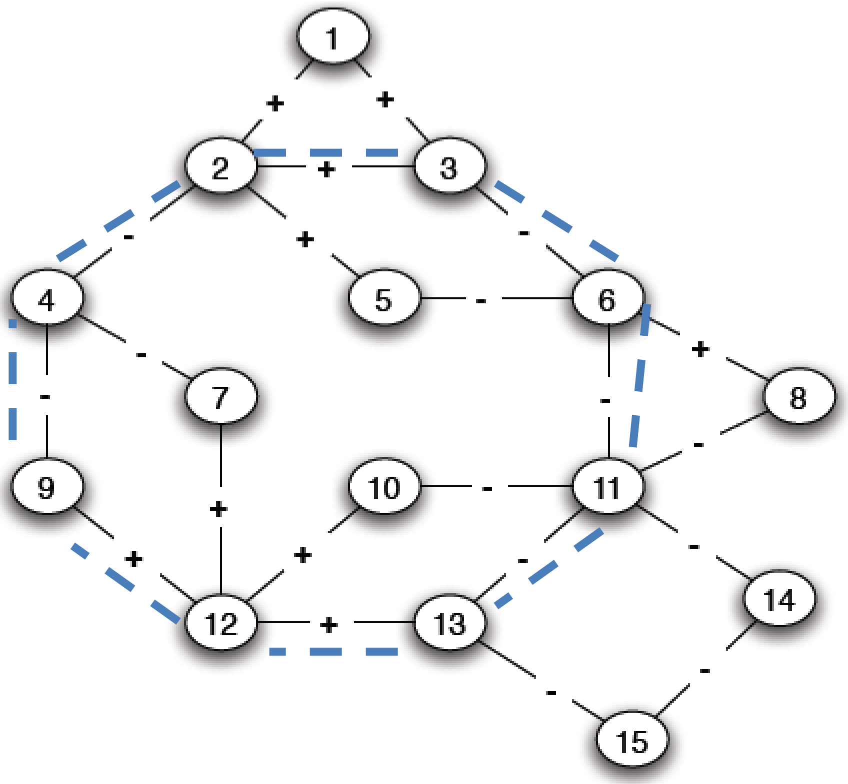
Is there such a cycle with an odd number of -?



Balance Definition for General Graphs

Cycle with odd number of - => unbalanced

Is there such a cycle with an odd number of -?



Balance Characterization

Claim: A signed graph is balanced, if and only if, it contains no cycles with an odd number of negative edges

(proof by construction)

Find *a balanced division:* partition into sets X and Y, all edges inside X and Y positive, crossing edges negative

Either succeeds or Stops with a cycle containing an odd number of -

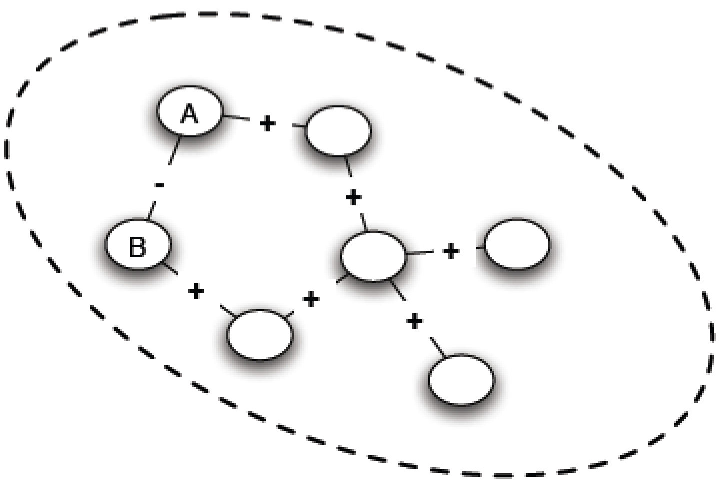
Two steps:

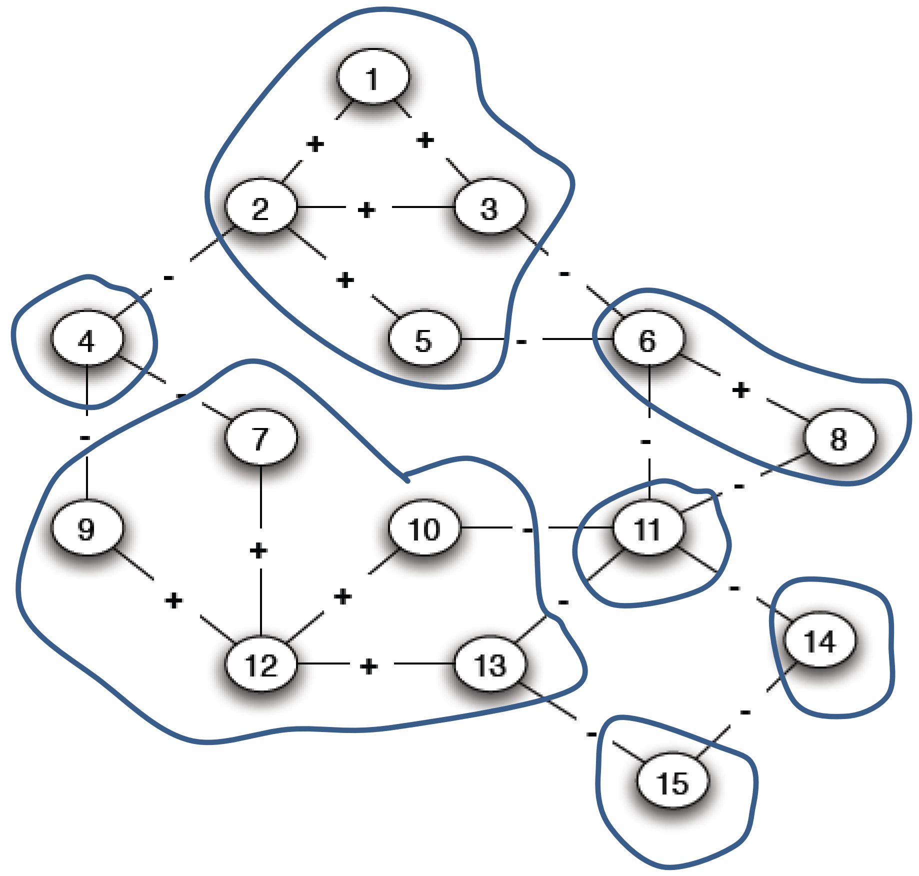
1. Convert the graph into a reduced one with only negative edges
2. Solve the problem in the reduced graph

Balance Characterization: Step 1

a. Find *connected components* (supernodes) by considering only positive edges

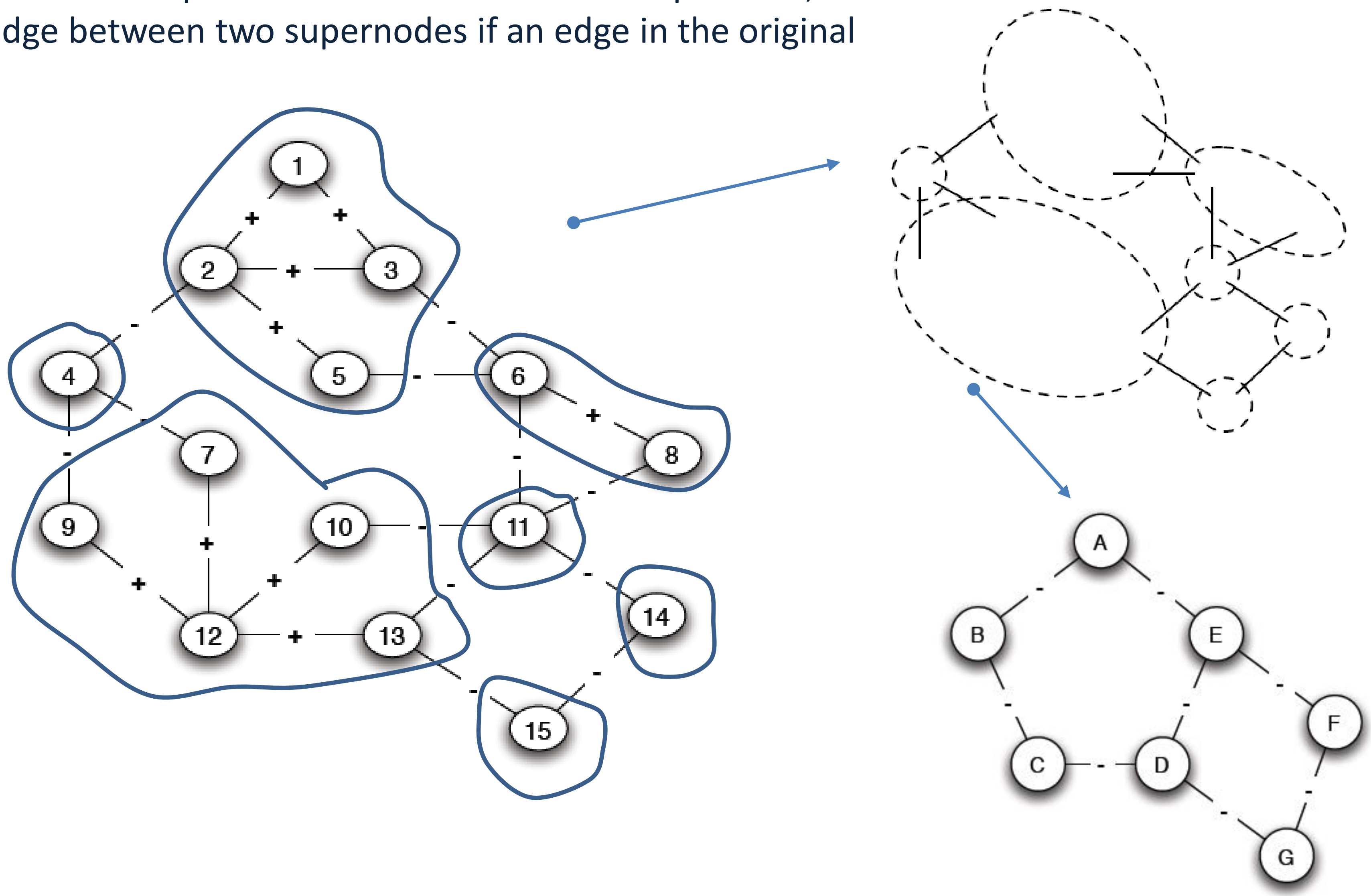
b. Check: Do supernodes contain a negative edge between any pair of their nodes
(i) Yes -> odd cycle
(ii) No -> each supernode either X or Y





Balance Characterization: Step 1

c. Reduced problem: a node for each supernode, an edge between two supernodes if an edge in the original



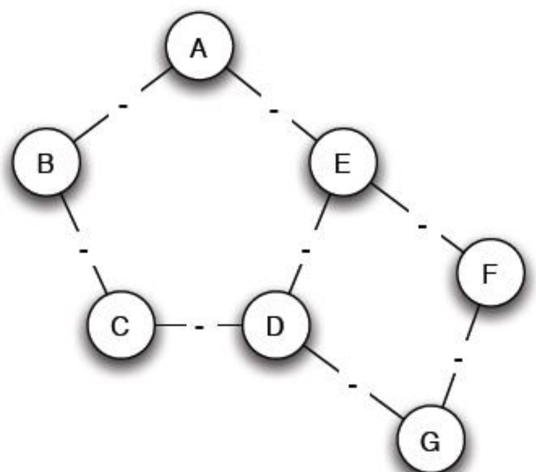
Balance Characterization: Step 2

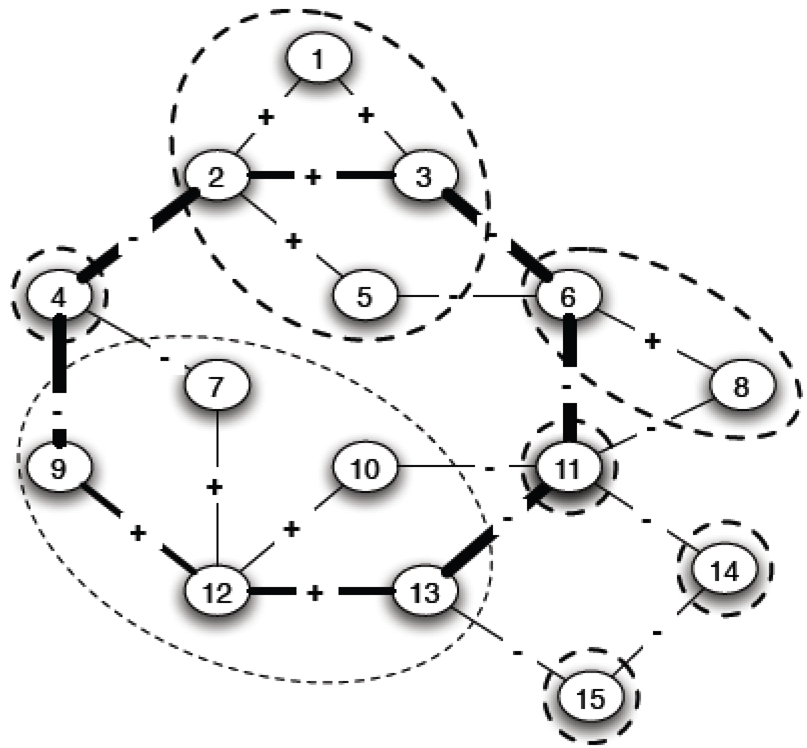
Note: Only negative edges among supernodes

Start labeling by either X and Y

If successful, then label the nodes of the supernode correspondingly

✓ A cycle with an odd number, corresponds to a (possibly larger) odd cycle in the original





Balance Characterization: Step 2

Determining whether the graph is bipartite (there is no edge between nodes in X or Y, *the only edges are from nodes in X to nodes in Y*)

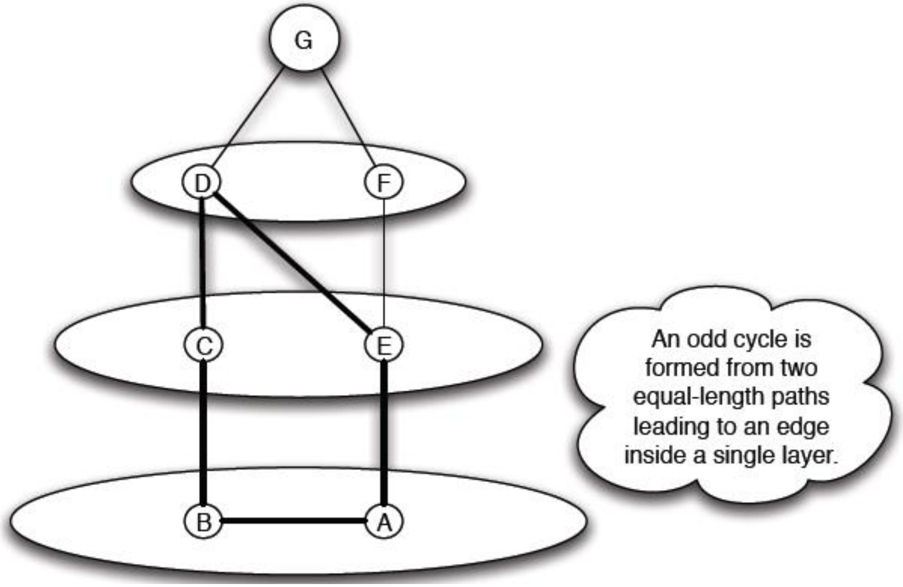
Use Breadth-First-Search (BFS)

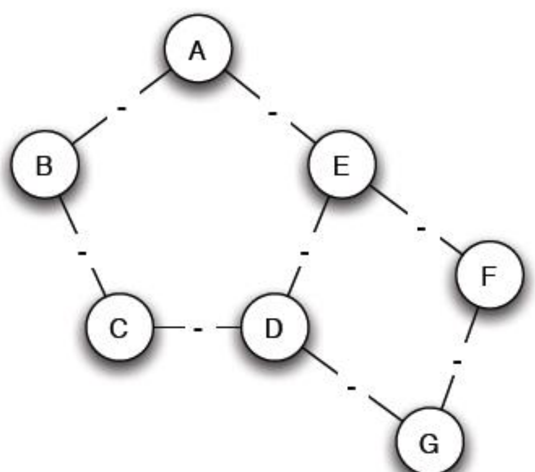
Two type of edges: (1) between nodes in adjacent levels (2) between nodes in the same level

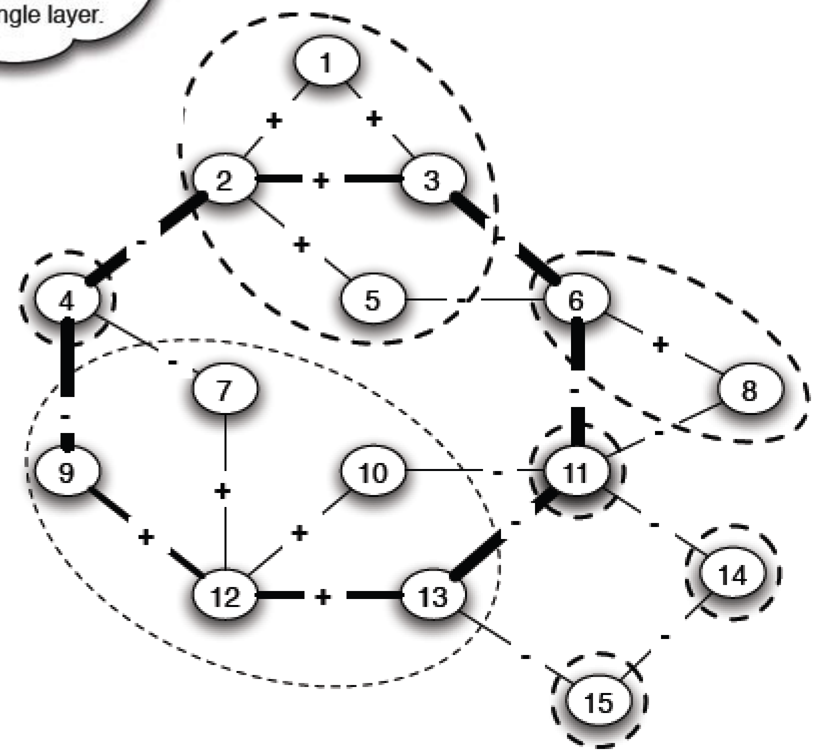
If only type (1), alternate X and Y labels at each level

If type (2), then odd cycle

Balance Characterization





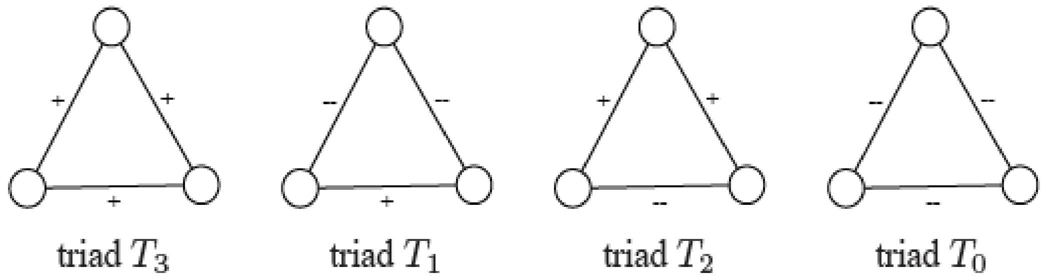


Status theory in practice

	Epinions	Slashdot	Wikipedia
Nodes	119,217	82,144	7,118
Edges	841,200	549,202	103,747
+ edges	85.0%	77.4%	78.7%
− edges	15.0%	22.6%	21.2%
Triads	13,375,407	1,508,105	790,532

- **Epinions:** product review Web site, where users can indicate their *trust* or *distrust* of the reviews
- **Slashdot:** the social network of the blog where a signed link indicates that one user *likes* or *dislikes* the *comments*
- **Wikipedia:** its voting network where a signed link indicates a positive or negative *vote* by one user *on the promotion* to admin status of another.

Structural balance theory in practice



Triad T_i		$\|T_i\|$	$p(T_i)$	$p_0(T_i)$	$s(T_i)$
Epinions					
T_3	$+++$	11,640,257	0.870	0.621	1881.1
T_1	$+--$	947,855	0.071	0.055	249.4
T_2	$++-$	698,023	0.052	0.321	-2104.8
T_0	$---$	89,272	0.007	0.003	227.5
Slashdot					
T_3	$+++$	1,266,646	0.840	0.464	926.5
T_1	$+--$	109,303	0.072	0.119	-175.2
T_2	$++-$	115,884	0.077	0.406	-823.5
T_0	$---$	16,272	0.011	0.012	-8.7
Wikipedia					
T_3	$+++$	555,300	0.702	0.489	379.6
T_1	$+--$	163,328	0.207	0.106	289.1
T_2	$++-$	63,425	0.080	0.395	-572.6
T_0	$---$	8,479	0.011	0.010	10.8

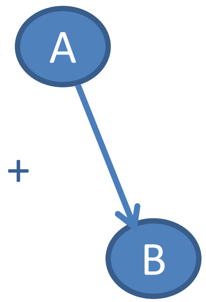
Symbol	Meaning
T_i	Signed triad, also the number of triads of type T_i
Δ	Total number of triads in the network
p	Fraction of positive edges in the network
$p(T_i)$	Fraction of triads T_i, $p(T_i) = T_i/\Delta$
$p_0(T_i)$	A priori prob. of T_i (based on sign distribution)
$E[T_i]$	Expected number of triads T_i, $E[T_i] = p_0(T_i)\Delta$
$s(T_i)$	Surprise, $s(T_i) = (T_i - E[T_i])/\sqrt{\Delta p_0(T_i)(1-p_0(T_i))}$

- All-positive triad T_3 is *heavily overrepresented* in all three datasets. T_3 tends to be over represented by about 40% in all three datasets
- Triad T_2 consisting of two enemies with a common friend is *heavily underrepresented*. T_2 is underrepresented by about 75%in Epinions and Slashdot and 50% in Wikipedia
- More consistent with weak structural balance

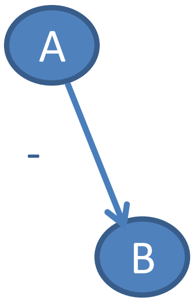
A theory of status

Directed networks

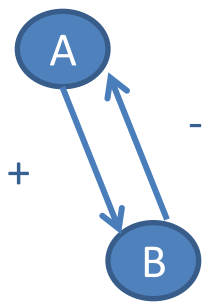
A *positive edge* (*A, B*) means that *A* regards *B* as having *higher status* than herself



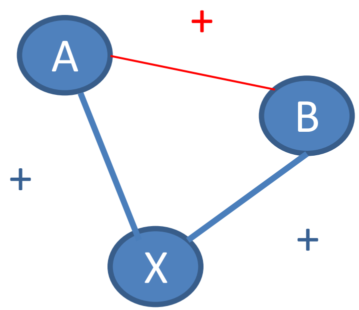
A *negative edge* (*A, B*) means that *A* regards *B* as having lower status than herself



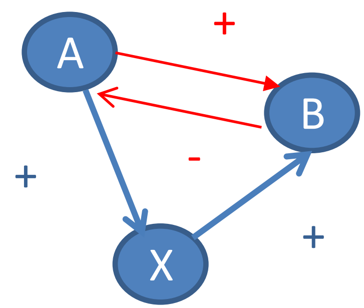
Assuming that all participants agree on status ordering, status theory predicts that when the direction of an edge is flipped, its sign should flip as well.

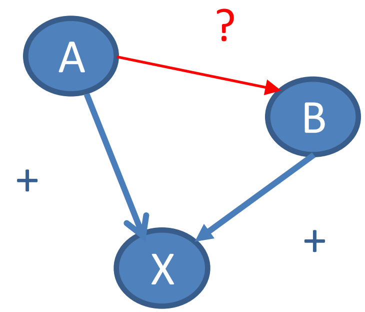


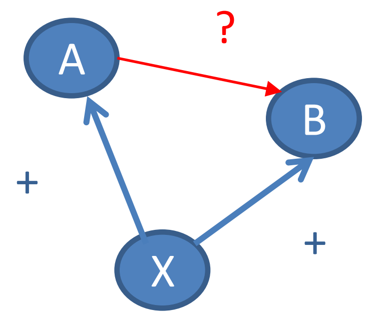
A theory of status

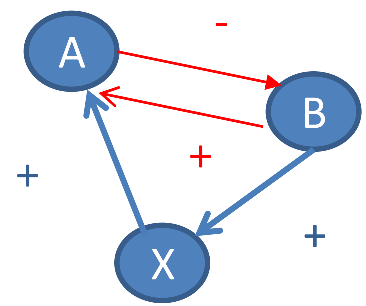


Structural balance









A theory of status: local property

For any edge (*u, v*), and any third node *w*, possible to assign distinct numerical "status values" to *u*, *v*, and *w* in such a way that the *positive edges* among them (if any) go from nodes of lower status to nodes of higher status, and the *negative edges* among them (if any) go from nodes of higher status to nodes of lower status.

Three nodes *u, v,* and *w* are *status-consistent* if this condition holds.

A theory of status: global property

Let *G* be a signed, directed graph, and suppose that all sets of three nodes in G are status-consistent.
Then it possible to order the nodes of G as $v_1, v_2, \ldots, v_n$ in such a way that each positive edge (v_i, v_j) satisfies $i < j$, and each negative edge (v_i, v_j) satisfies $i > j$.

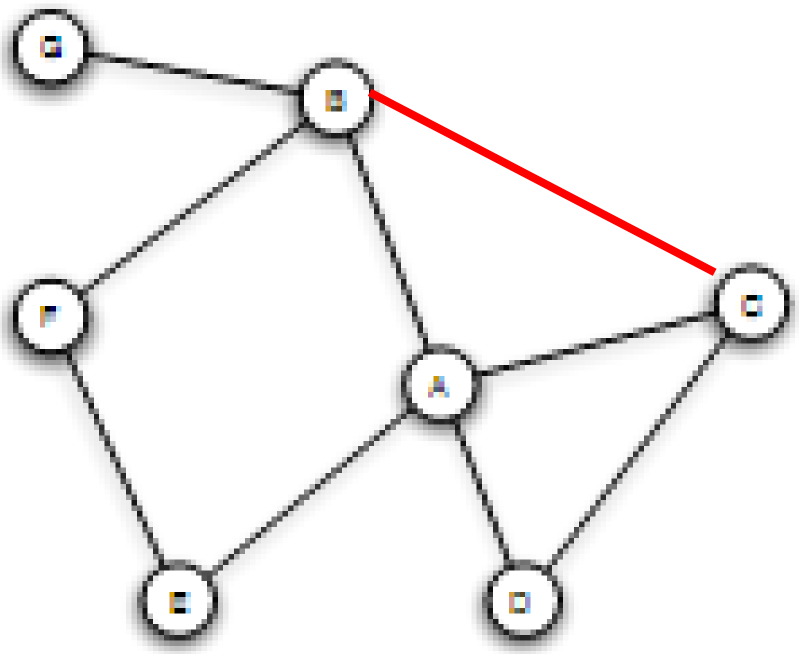
References

Networks, Crowds, and Markets (Chapter 5)

STRONG AND WEAK TIES

Triadic Closure

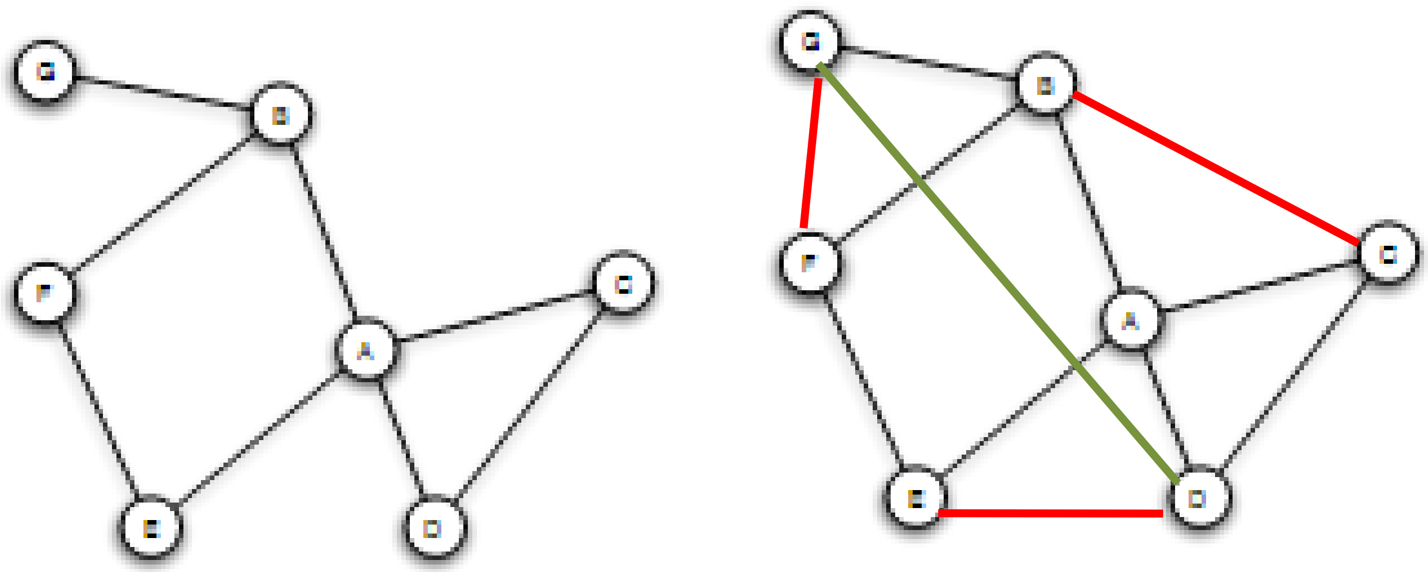
If two people in a social network have *a friend in common*, then there is an increased likelihood that they will become *friends themselves* at some point in the future



Triangle

Triadic Closure

Snapshots over time:



Clustering Coefficient

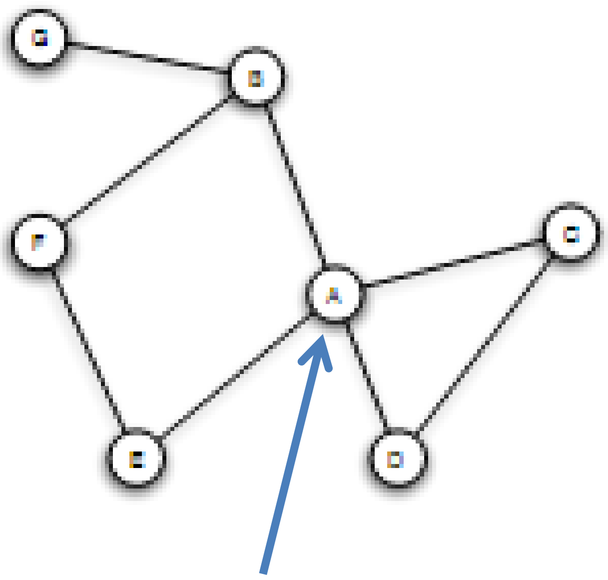
(Local) clustering coefficient for a node is the probability that two randomly selected friends of the node are friends with each other *(i.e., form a triangle)*

$$C_i = \frac{2|\{e_{jk}\}|}{k_i(k_i - 1)} \qquad e_{jk} \in E, u_i, u_j \in Ni, k \text{ size of } \mathrm{N_i}, \mathrm{N_i} \text{ neigborhood of } u_i$$

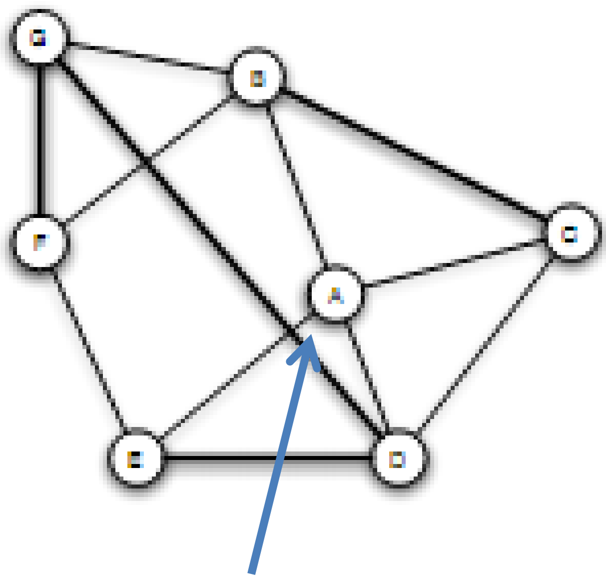
Fraction of the friends of a node that are friends with each other (i.e., connected)

$$\mathrm{C}^{(1)} = \frac{\sum_i \text{triangles centered at node i}}{\sum_i \text{triples centered at node i}}$$

Clustering Coefficient



1/6

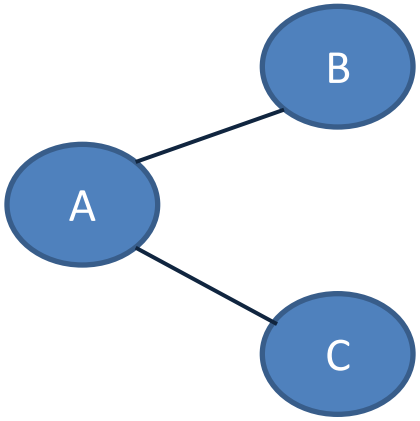


1/2

Ranges from 0 to 1

Triadic Closure

If A knows B and C, B and C are likely to become friends, *but WHY?*



1. Opportunity
2. Trust
3. Incentive of A (latent stress for A, if B and C are not friends, dating back to social psychology, e.g., relating low clustering coefficient to suicides)

The Strength of Weak Ties Hypothesis

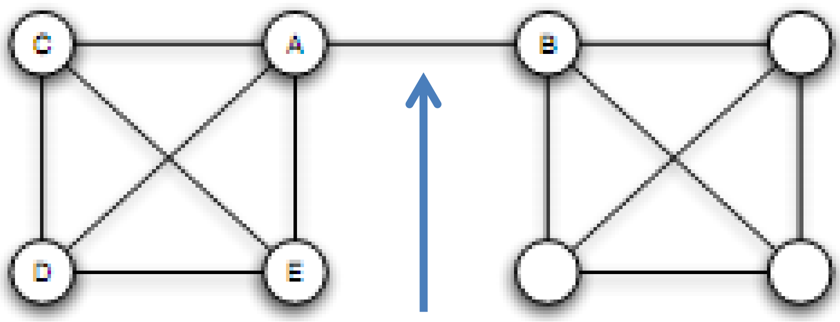
Mark Granovetter, in the late 1960s

Many people learned information leading to their current job *through personal contacts*, often described as *acquaintances* rather than *closed friends*

Two aspects

- Structural
- Local (interpersonal)

Bridges and Local Bridges

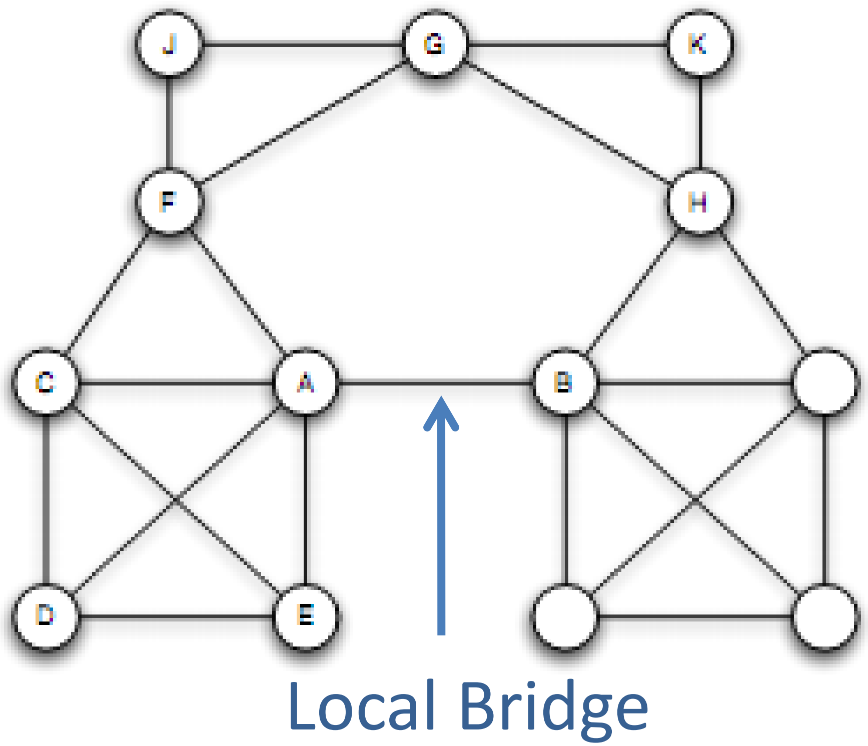


Bridge
(aka cut-edge)

- An edge between A and B is a *bridge* if deleting that edge would cause A and B to lie in two different components
  AB the only "route" between A and B

- Extremely rare in social networks

Bridges and Local Bridges

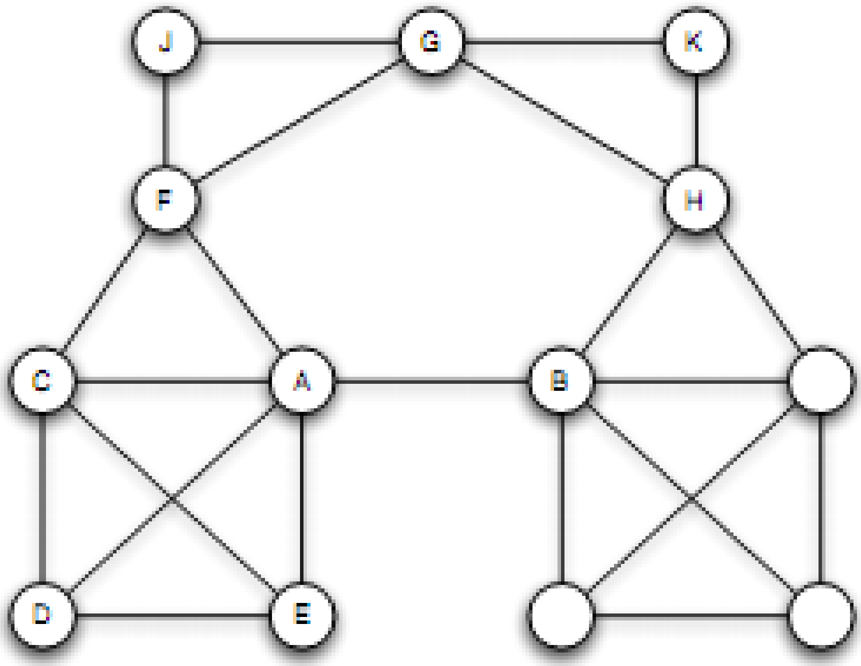


- An edge between A and B is a *local bridge* if deleting that edge would increase the distance between A and B to a value strictly more than 2

  A and B have no friends in common

- *Span of a local bridge:* distance of the its endpoints if the edge is deleted

Bridges and Local Bridges

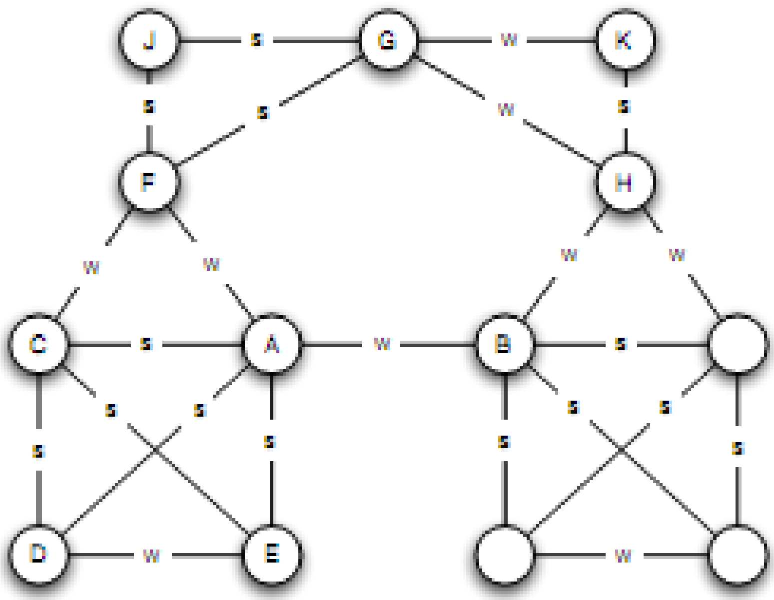


An edge is a local bridge, if an only if, it is not part of any triangle in the graph

The Strong Triadic Closure Property

- Not all links the same
- Levels of strength of a link: Strong and Weak ties
- May vary across times and situations

Annotated graph

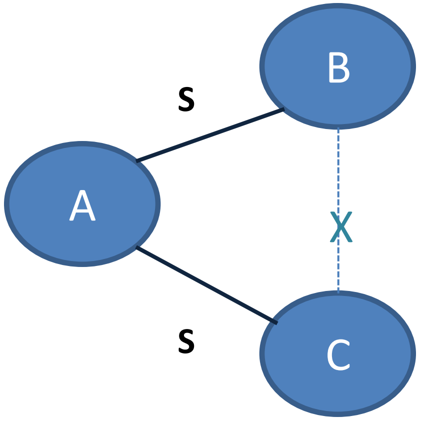


The Strong Triadic Closure Property

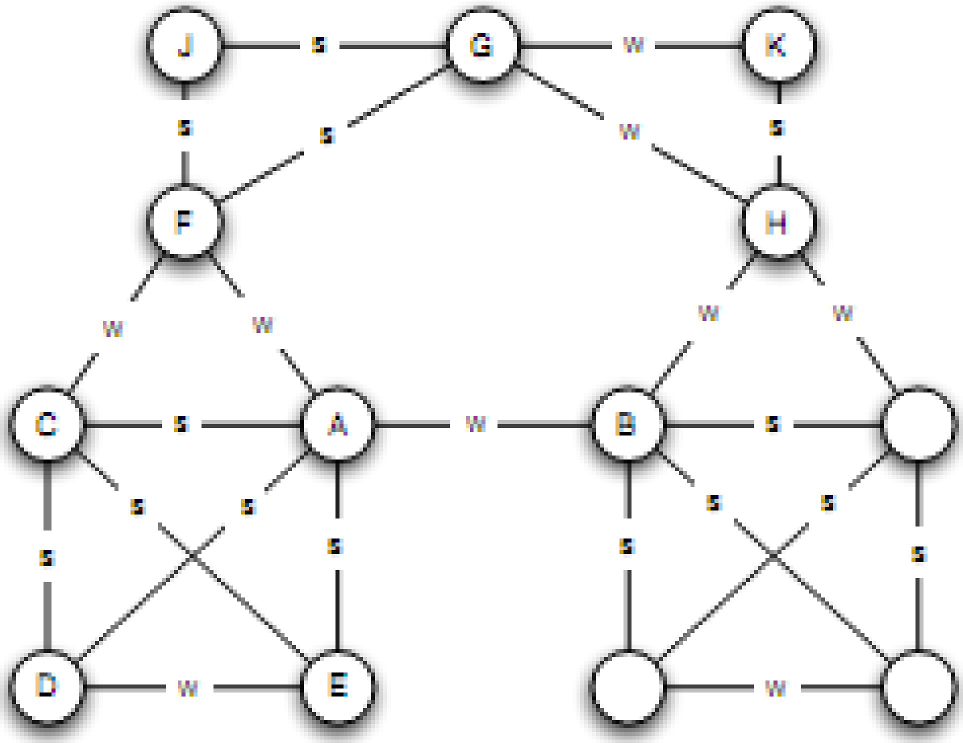
If a node A has edges to nodes B and C, then the B-C edge is <u>especially likely</u> to form if both A-B and A-C are <u>strong ties</u>

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 (strong or weak tie) between B and C.

A node A satisfies the Strong Triadic Property if it does not violate it



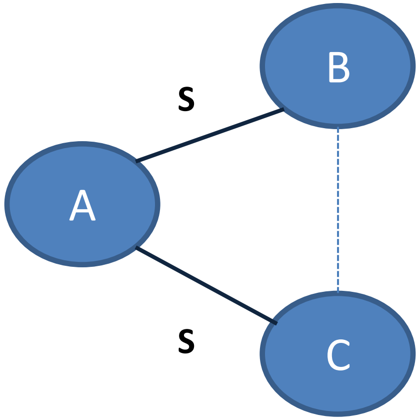
The Strong Triadic Closure Property



Local Bridges and Weak Ties

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

Proof: by contradiction



local distinction (weak and strong ties) related to *global structural* distinction (local bridges or not)

Relation to job seeking?

Tie Strength and Network Structure in Large-Scale Data

How to test these prediction on large social networks?

Tie Strength and Network Structure in Large-Scale Data

Cell-phone study [Omnela et. al., 2007]

"who-talks-to-whom network", covering 20% of the national population

- *Nodes*: cell phone users
- *Edge*: if they make phone calls to each other in both directions over 18-week observation periods
- *Strength of the tie*: time spent talking during an observation period

Is it a social network?

- Cells generally used for personal communication + no central directory, thus cell-phone numbers exchanged among people who already know each other
- Broad structural features of large social networks (*giant component*, 84% of nodes)

Generalizing Weak Ties and Local Bridges

So far:

- Either weak or strong
- Local bridge or not

Tie Strength: Numerical quantity (= number of min spent on the phone)

How to quantify "local bridges"?

Generalizing Weak Ties and Local Bridges

Bridges

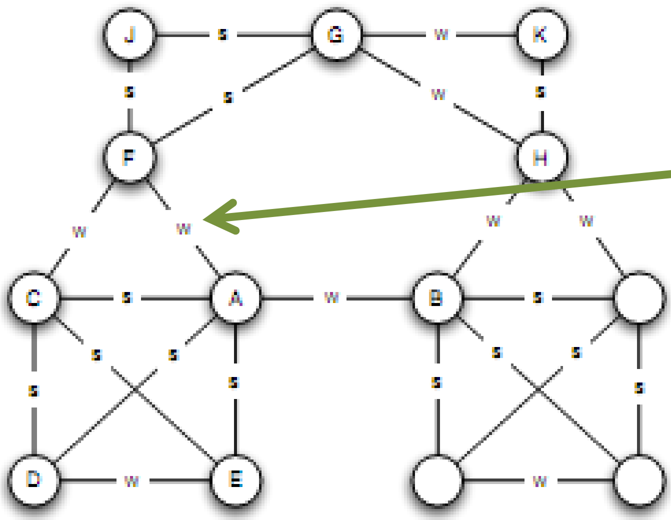
"almost" local bridges

Neighborhood overlap of an edge e_{ij}

(*) In the denominator we do not count A or B themselves

$$\frac{|N_i \cap N_j|}{|N_i \cup N_j|}$$

Jaccard coefficient



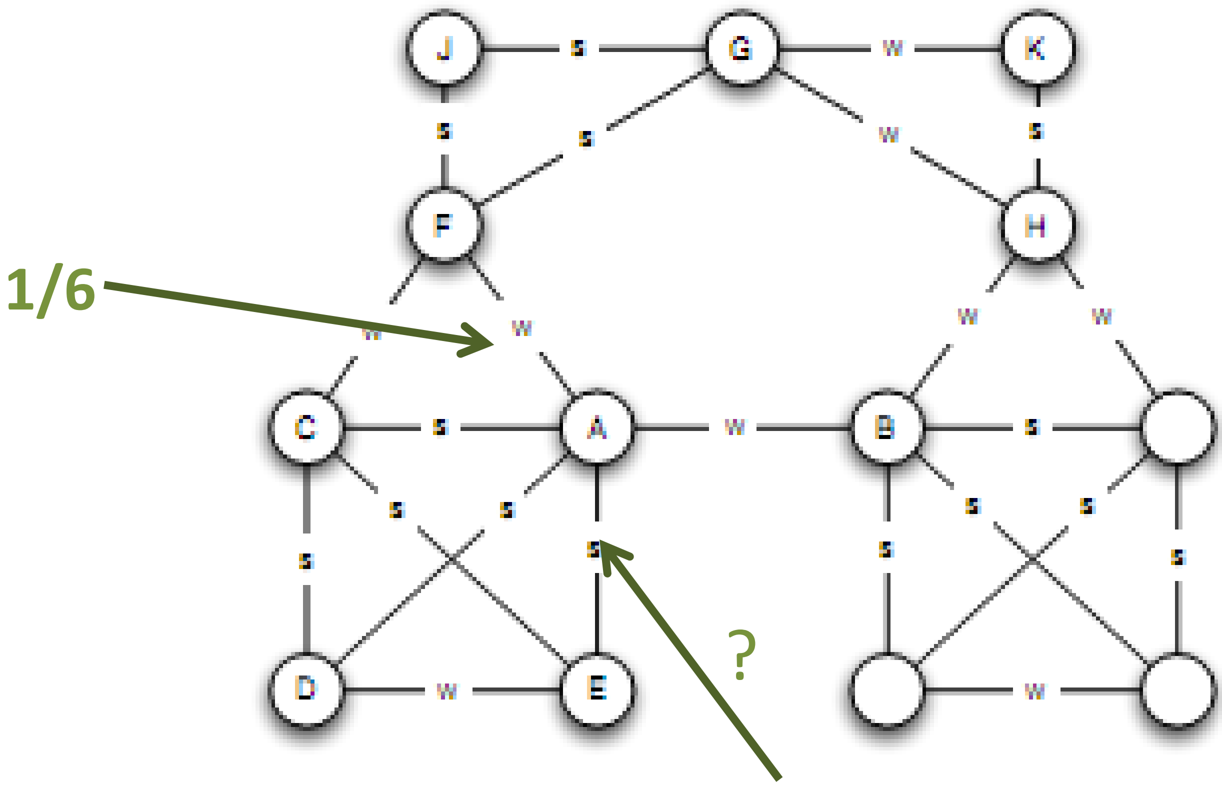
A: B, E, D, C
F: C, J, G

1/6

When is this value 0?

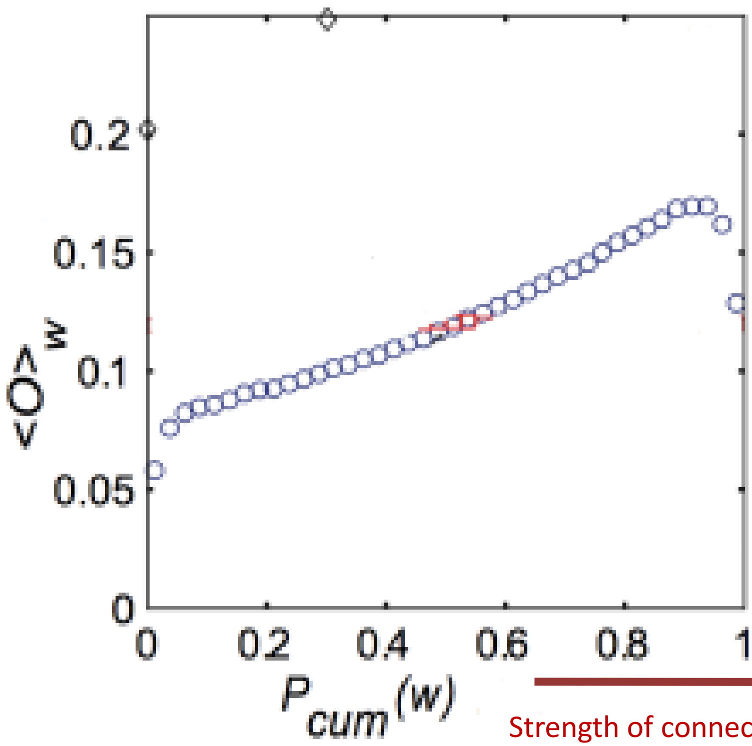
Generalizing Weak Ties and Local Bridges

Neighborhood overlap = 0: edge is a local bridge
Small value: “almost” local bridges



Generalizing Weak Ties and Local Bridges:
Empirical Results

Hypothesis: the strength of weak ties predicts that *neighborhood overlap* should *grow* as *tie strength grows*



(*) Some deviation at the right-hand edge of the plot

sort the edges -> for each edge at which percentile

Strength of connection (function of the percentile in the sorted order)

Generalizing Weak Ties and Local Bridges:
Empirical Results

How to test the following global (macroscopic) level hypothesis:

Hypothesis: weak ties serve to link different tightly-knit communities that each contain a large number of stronger ties

Generalizing Weak Ties and Local Bridges: Empirical Results

Delete edges from the network one at a time

- Starting with the strongest ties and working downwards in order of tie strength
    - giant component shrank steadily

-Starting with the weakest ties and upwards in order of tie strength
    - giant component shrank more rapidly, broke apart abruptly as a critical number of weak ties were removed

Social Media and Passive Engagement

People maintain large explicit lists of friends

Test:
How *online activity* is distributed across *links of different strengths*

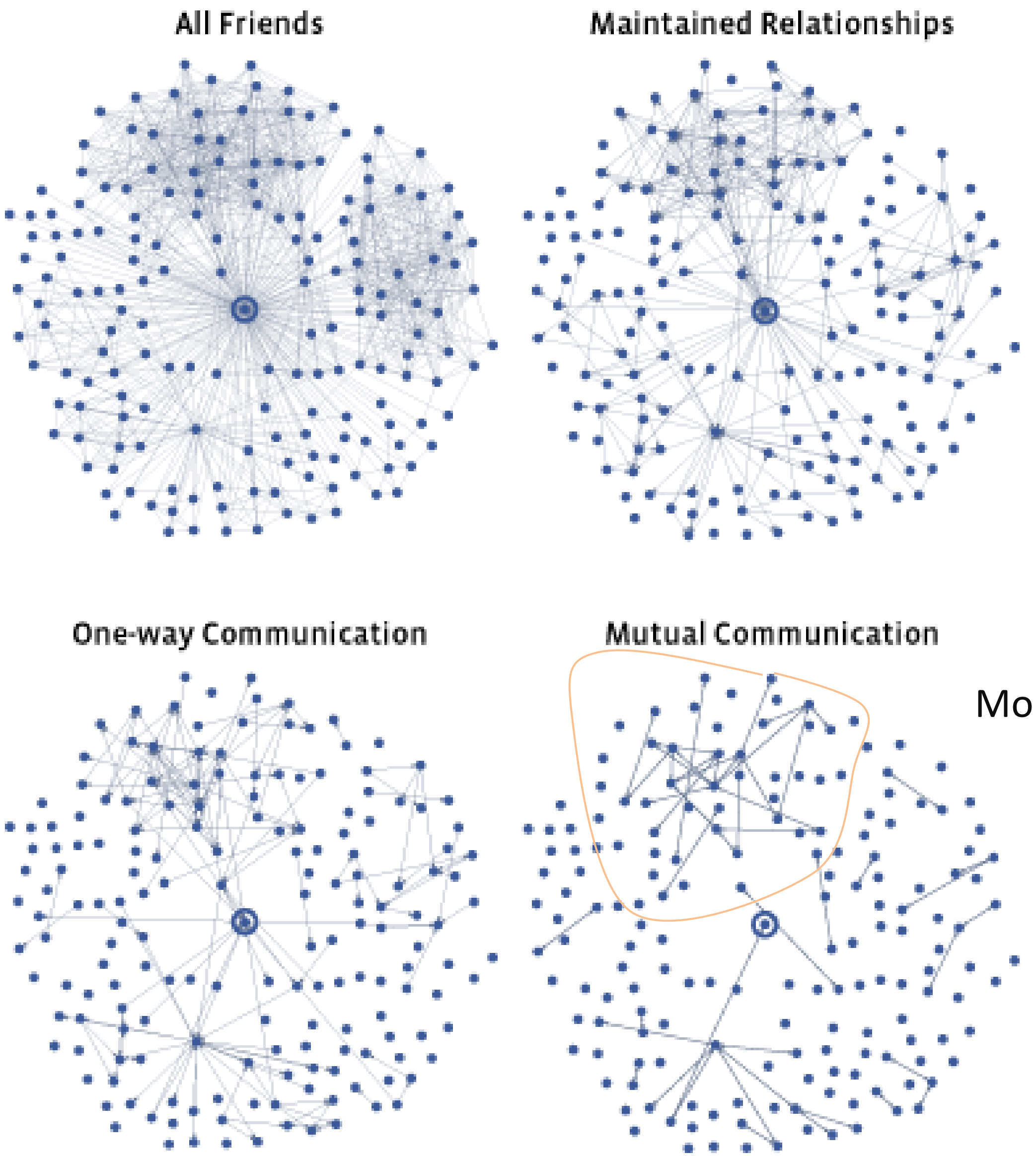
Tie Strength on Facebook

Cameron Marlow, et al, 2009
At what extent each link was used for social interactions

Three (not exclusive) kinds of ties (links)

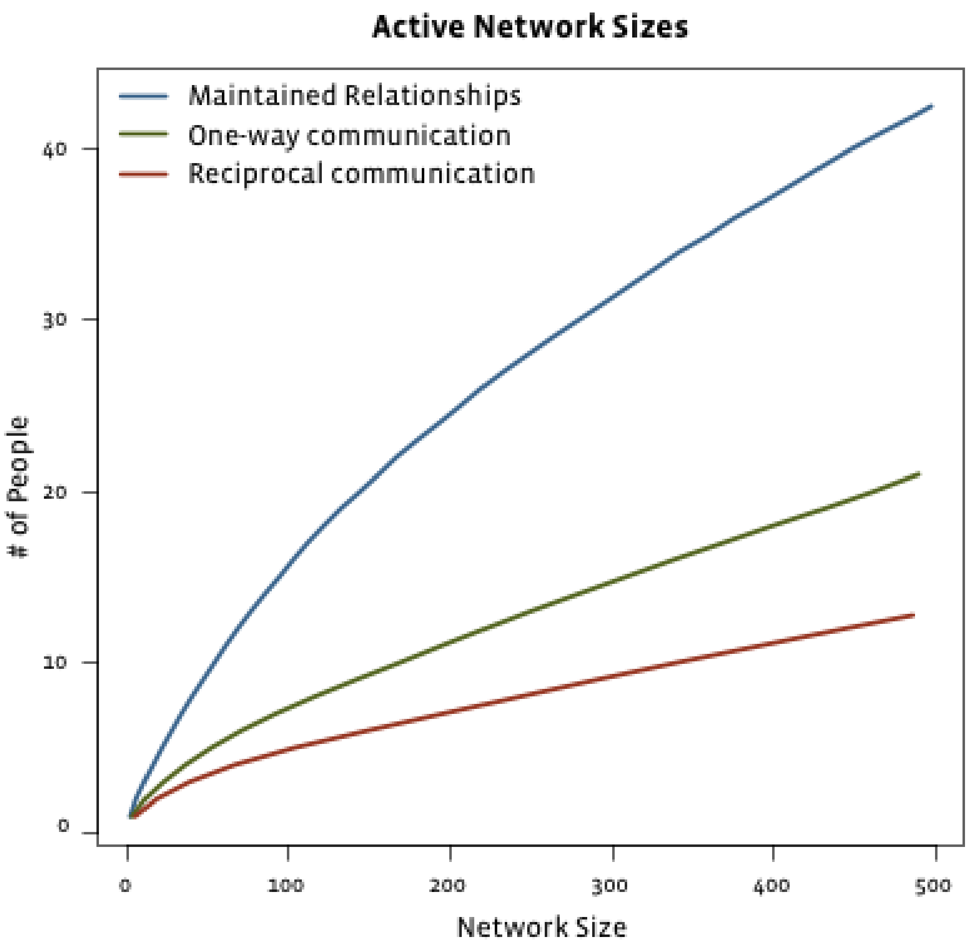
1. Reciprocal (mutual) communication: both *send and received messages* to friends at the other end of the link
2. One-way communication: the user *send* one or more message to the friend at the other end of the link (including reciprocal)
3. Maintained relationship: the user *followed information* about the friend at the other end of the link (click on content via News feed or visit the friend profile more than once)

Tie Strength on Facebook



More recent connections

Tie Strength on Facebook



Total number of friends

Even for users with very large number of friends

- actually communicate : 10-20
- number of friends follow even passively <50

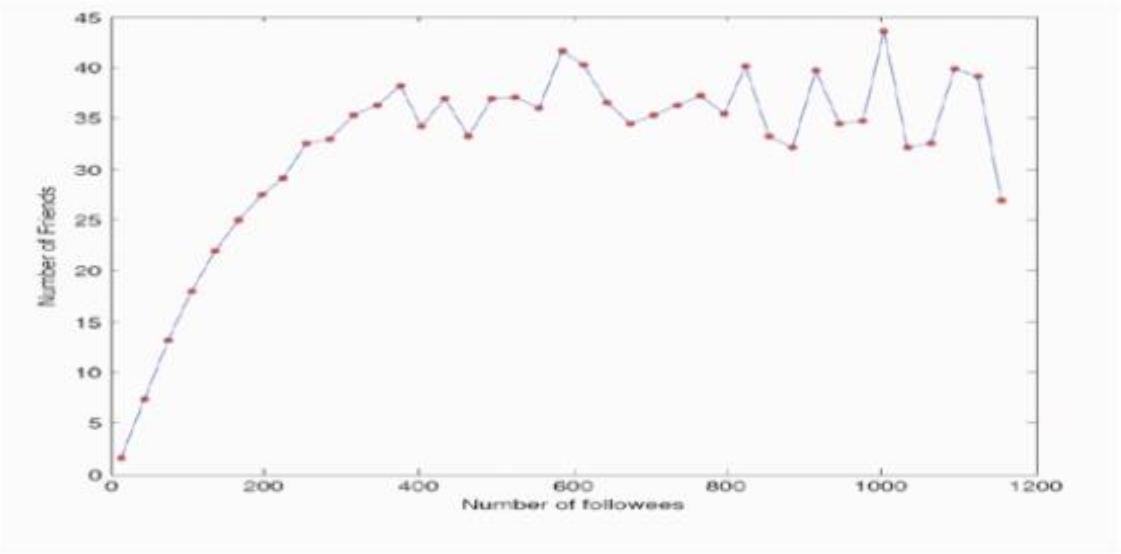
Passive engagement: keep up with friends by reading about them even in the absence of communication

Tie Strength on Twitter

Huberman, Romero and Wu, 2009

Two kinds of links

- Follow
- Strong ties (friends): users to whom the user has *directed at least two messages* over the course if the observation period



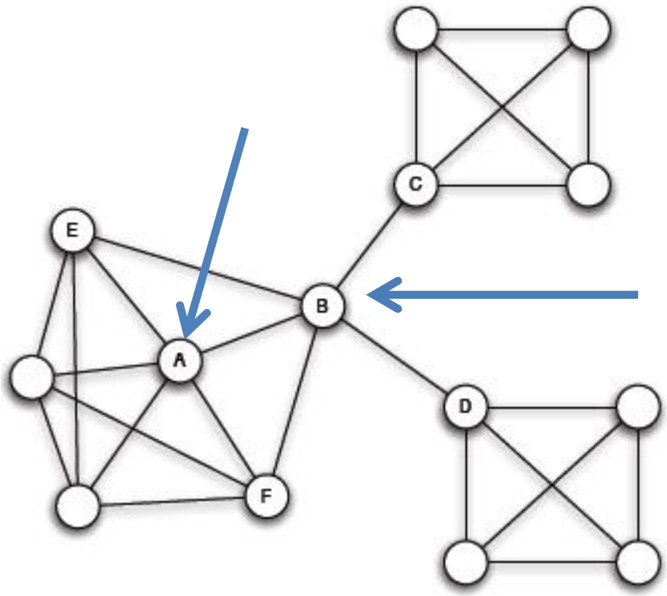
Social Media and Passive Engagement

- Strong ties require continuous investment of time and effort to maintain (as opposed to weak ties)
- Network of strong ties still remain sparse
- How different links are used to convey information

Closure, Structural Holes and Social Capital

Different roles that *nodes* play in this structure

Access to edges that span different groups is not equally distributed across all nodes

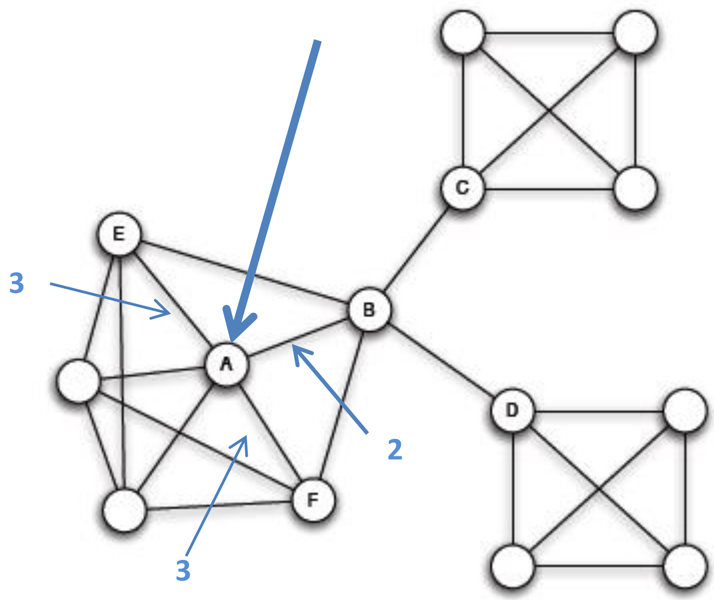


Embeddedness

A has a large clustering coefficient

- Embeddedness of an edge: number of common neighbors of its endpoints (neighborhood overlap, local bridge if 0)

For A, all its edges have significant embeddedness



(sociology) if two individuals are connected by an embedded edge => trust

- "Put the interactions between two people on display"

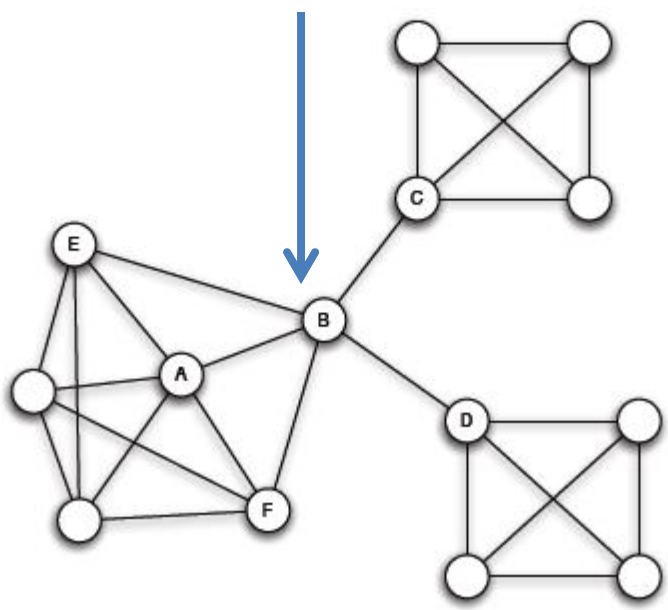
Structural Holes

(sociology) B-C, B-D much riskier, also, possible contradictory constraints
Success in a large cooperation correlated to access to local bridges

B spans a structural hole

- B has access to information originating in multiple, non interacting parts of the network
- An amplifier for creativity
- Source of power as a social “gate-keeping”

Social capital



References

Networks, Crowds, and Markets (Chapter 3)