Social Network Analysis

Strong & Weak Ties

Introduction

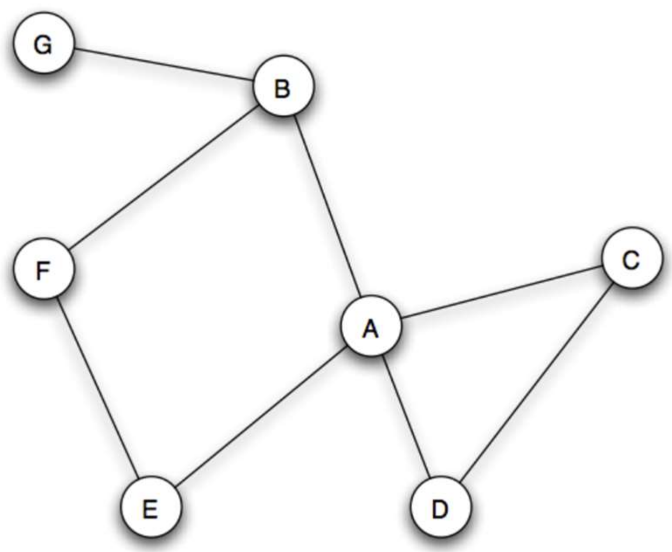
- Networks bridge the local to the global — to offer explanations for how simple processes at the level of individual nodes and links can have complex effects that ripple through a population as a whole.
- We will consider some fundamental social network issues that illustrate this theme:
  - how information flows through a social network,
  - how different nodes can play structurally distinct roles in this process,
  - how these structural considerations shape the evolution of the network itself over time.
- These themes all play central roles in different contexts as they arise.
- Our first focus is the famous "strength of weak ties" hypothesis from sociology

Overview

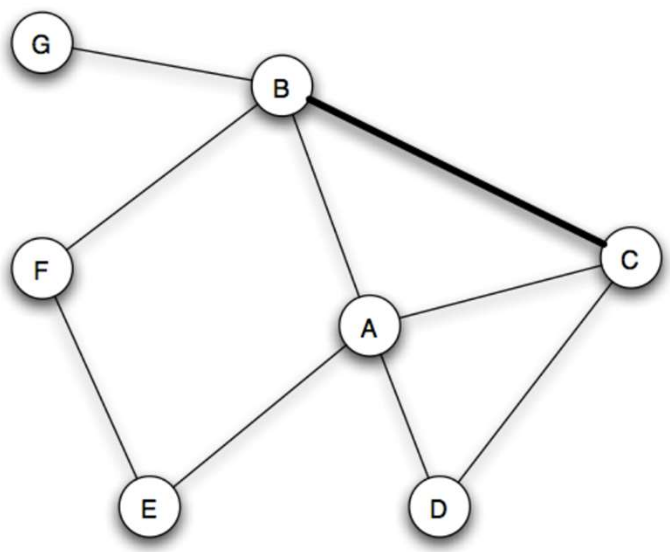
- Triadic Closure
- The Strength of Weak Ties
- Tie Strength and Network Structure in Large-Scale Data
- Tie Strength, Social Media, and Passive Engagement
- Closure, Structural Holes, and Social Capital

How are edges formed in a network?

- Instead of treating networks as static structures where we take a snapshot of the nodes and edges at a particular moment in time, and then ask about various measures.
- It is also useful to think about how a network evolves over time.



(a) *Before B-C edge forms.*



(b) *After B-C edge forms.*

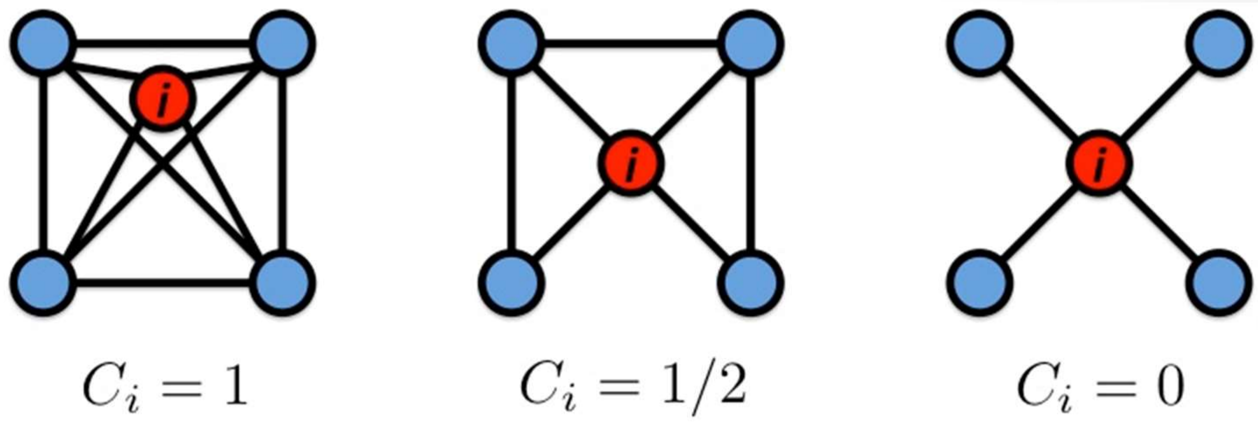
Triadic Closure

- 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
- The term Triadic Closure comes from the fact that the B-C edge has the effect of “closing” the third side of this triangle.
- If we observe snapshots of a social network at two distinct points in time, then in the later snapshot, we generally find a significant number of new edges that have formed through this triangle-closing operation, between two people who had a common neighbor in the earlier snapshot.

Clustering Coefficient

- **What fraction of your neighbors are connected?**
  - Node i with degree k_i
  - Clustering Coefficient C_i for a vertex i is in [0,1]

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



Clustering coefficient is a "local" property – each vertex has one.

Watts & Strogatz, Nature 1998.

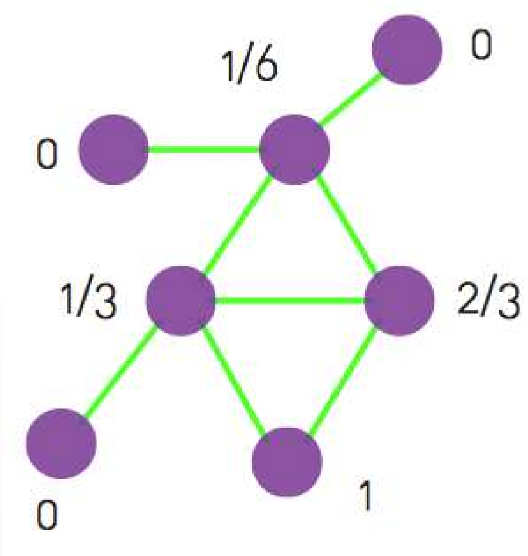
Clustering Coefficient

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

Clustering Coefficient of vertex i

Average Clustering Coefficient of the graph

$$\langle C \rangle = \frac{1}{N} \sum_{i=1}^{N} C_i \,.$$



<C> = (0+0+0+1+1/6+1/3+2/3)/7

<C> = 13/42 ≈ 0.310

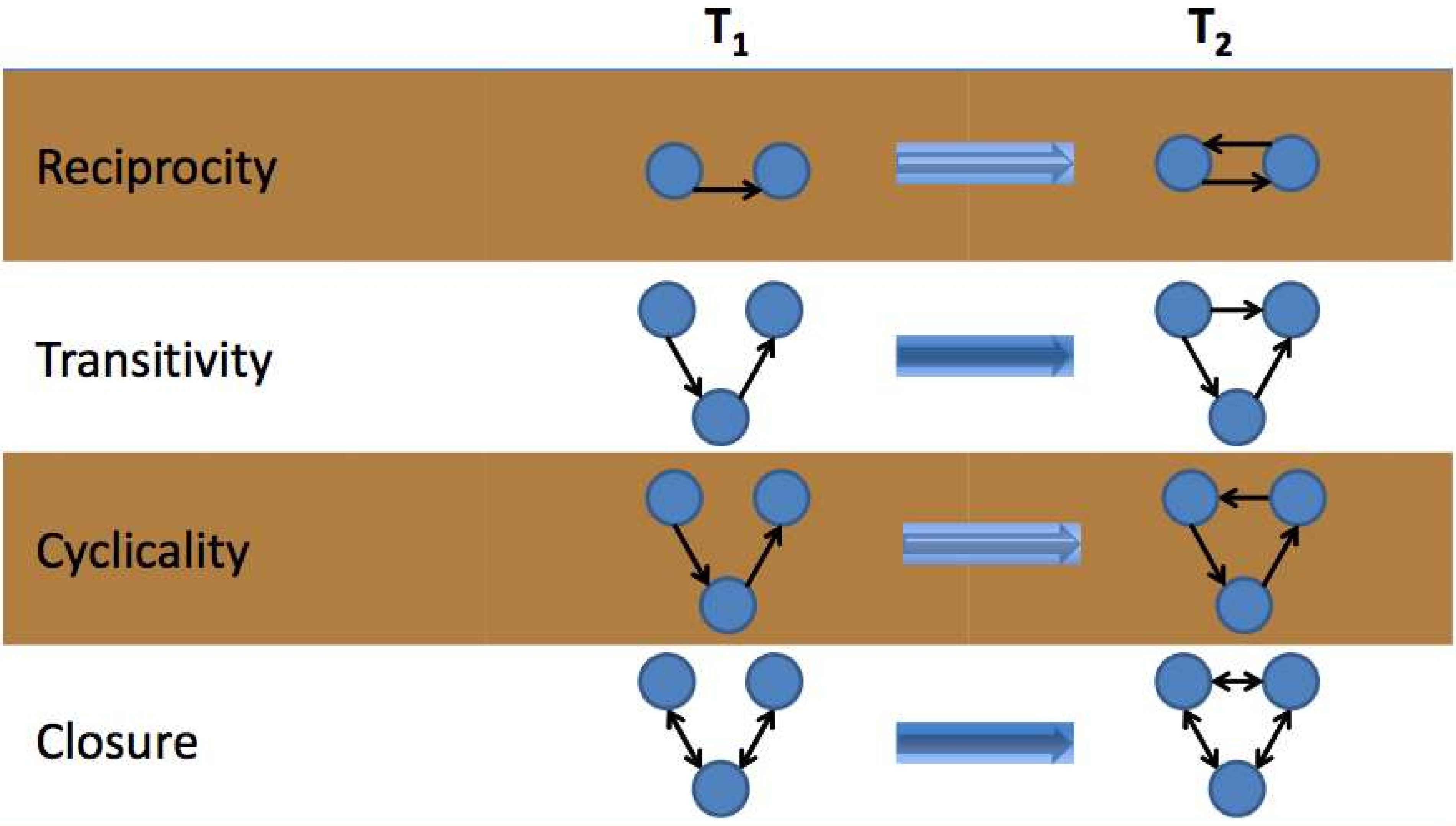
Clustering Coefficient

- ..is the fraction of pairs of A's friends that are connected to each other by edges.
- ..of a node A can also be defined as the probability that two randomly selected friends of A are friends with each other.
- The more the triadic closure (dynamic) is operating in the neighbourhood of a node, the higher the clustering coefficient (static) will be

Reasons for Triadic Closure

- When they have a common friend A, there is an increased opportunity for B and C to meet.
- Since each of B and C is friends with A (provided they are mutually aware of this) gives them a basis for trusting each other that an arbitrary pair of unconnected people might lack.
- A third reason is based on the incentive A may have to bring B and C together: if A is friends with B and C, then it becomes a source of latent stress in these relationships if B and C are not friends with each other.

Reciprocity, transitivity, & closure



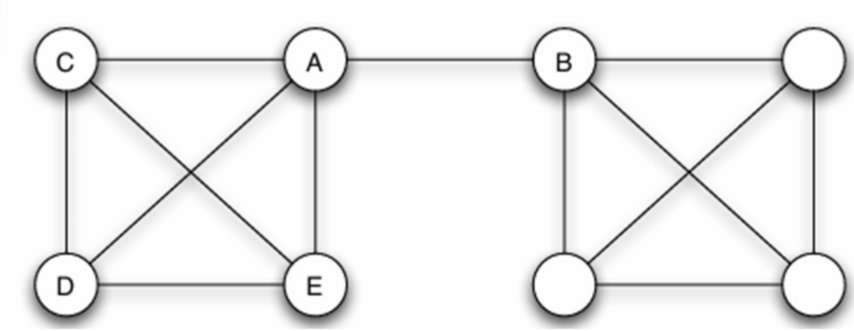
Overview

- Triadic Closure
- The Strength of Weak Ties
- Tie Strength and Network Structure in Large-Scale Data
- Tie Strength, Social Media, and Passive Engagement
- Closure, Structural Holes, and Social Capital

Granovetter's Research

- As part of his Ph.D. thesis research in the late 1960s, Mark Granovetter [1] interviewed people who had recently changed employers to learn how they discovered their new jobs?
  - Through personal contacts..
  - **Acquaintances**
  - not **Close friends**..

Strong and Weak Ties

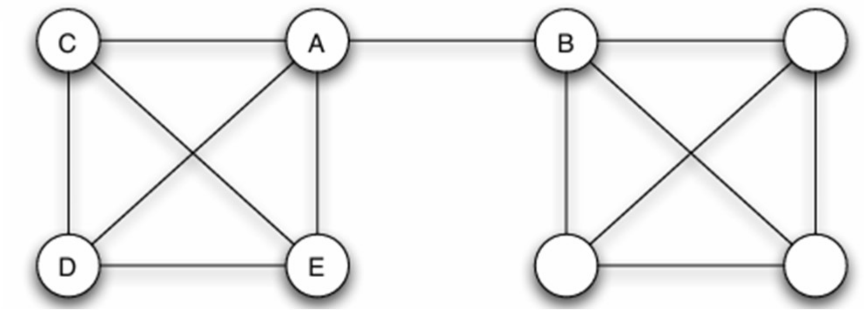


- A's links to C, D, and E connect her to a tightly-knit group of friends who all know each other,
- The link to B seems to reach into a different part of the network.
- We could speculate, that the structural peculiarity of the link to B will translate into differences in the role it plays in A's everyday life:
  - The tightly-knit group of nodes A, C, D, and E will all tend to be exposed to similar opinions and similar sources of information
  - A's link to B offers her access to things she otherwise wouldn't necessarily hear about.

Strong and Weak Ties

- We won't define "strength" precisely, but we mean it to align with the idea that stronger links represent closer friendship and greater frequency of interaction.
- In general, links can have a wide range of possible strengths, but for conceptual simplicity — and to match the friend/acquaintance dichotomy
- We'll categorize all links in the social network as belonging to one of two types:
  - strong ties (the stronger links, corresponding to friends)
  - weak ties (the weaker links, corresponding to acquaintances)

Bridges



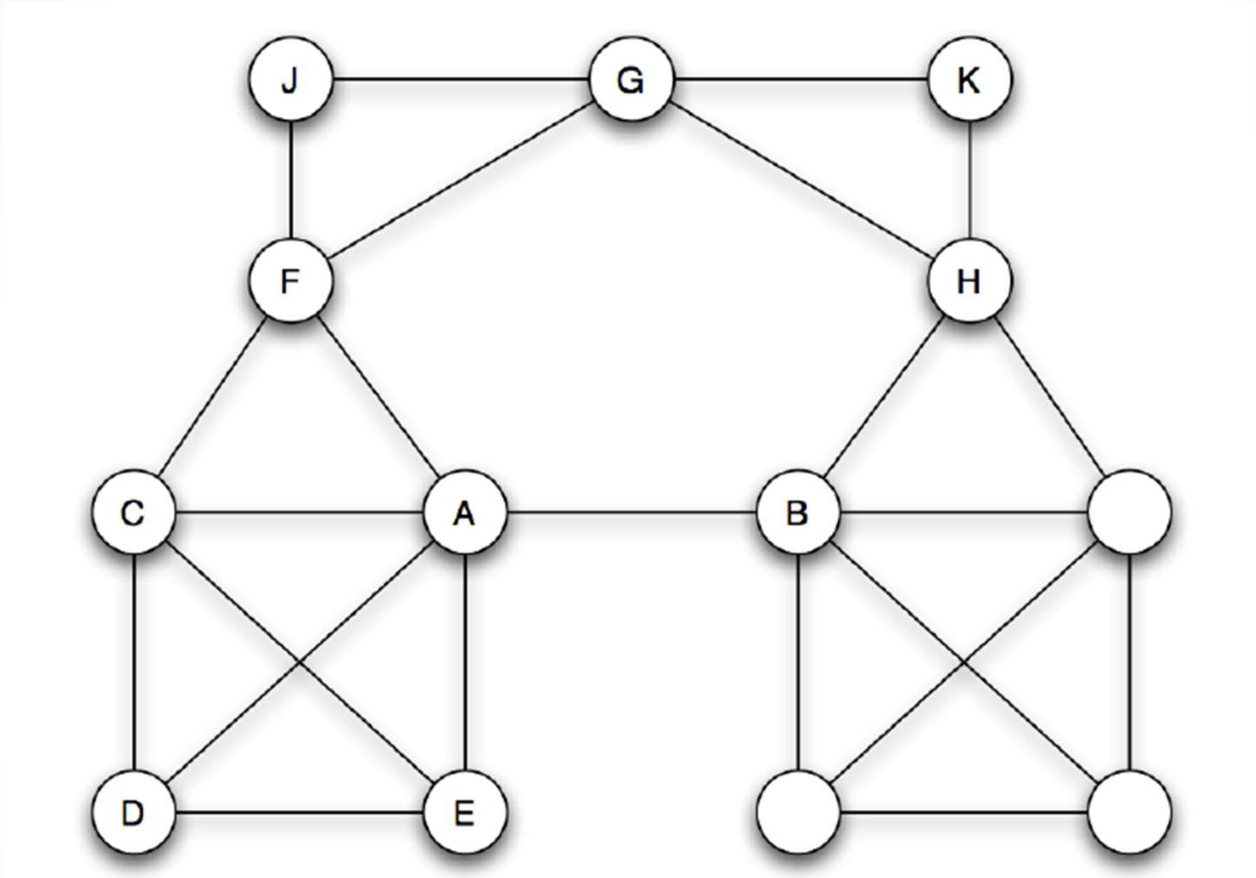
- A-B is a bridge; it's removal disconnects the graph into 2 components, with A and B lying in different components.
- Bridges provide nodes with access to parts of the network that are unreachable by other means.
- If a node like A is going to get truly new information, the kind that leads to a new job, it might come unusually often (though certainly not always) from a friend connected by a local bridge.
- The closely-knit groups that you belong to, though they are filled with people eager to help, are also filled with people who know roughly the same things that you do

- *In real social networks, bridges may be rare, so…?*

Local Bridge

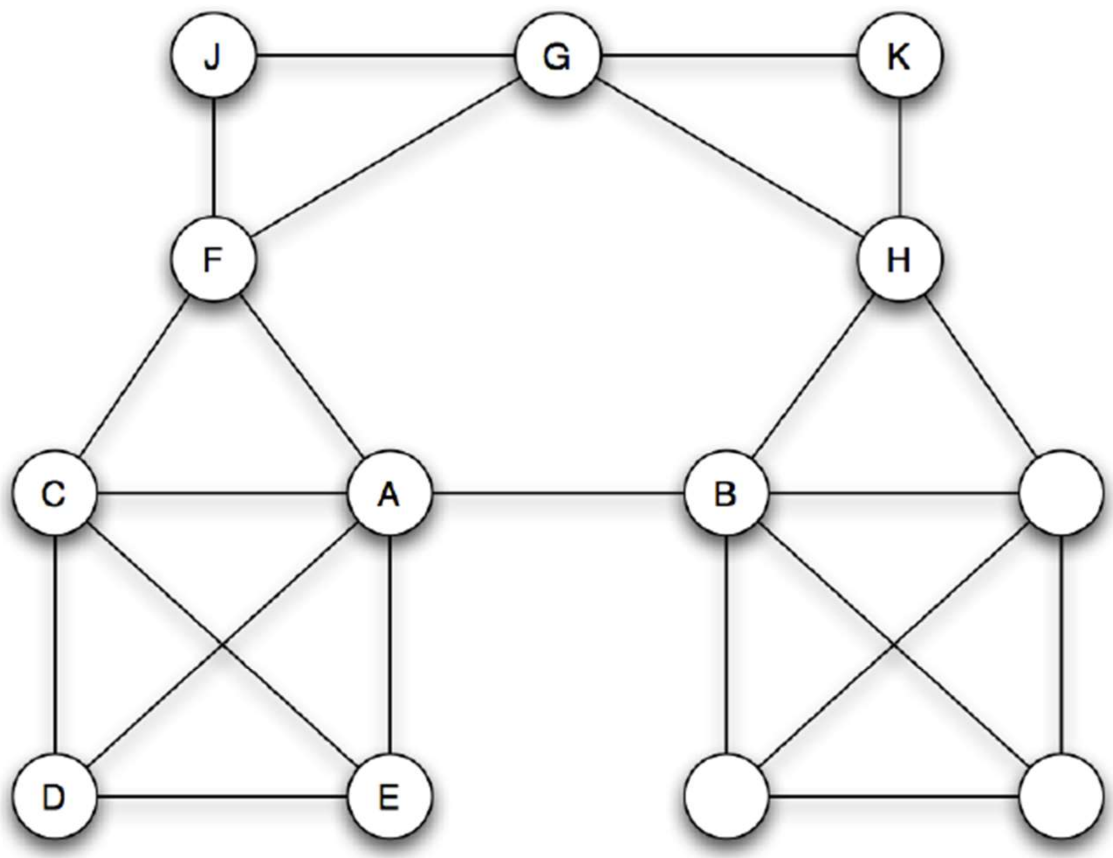
- An edge joining two nodes A and B in a graph is a local bridge if its endpoints A and B have no friends in common
  - in other words, if deleting the edge would increase the distance between A and B to a value strictly more than two.
- We say that the span of a local bridge is the distance its endpoints would be from each other if the edge were deleted

Local Bridges



We say that an edge joining two nodes A and B in a graph is a local bridge if its endpoints A and B have no friends in common —
in other words, if deleting the edge would increase the distance between A and B to a value strictly more than two.

Local Bridges

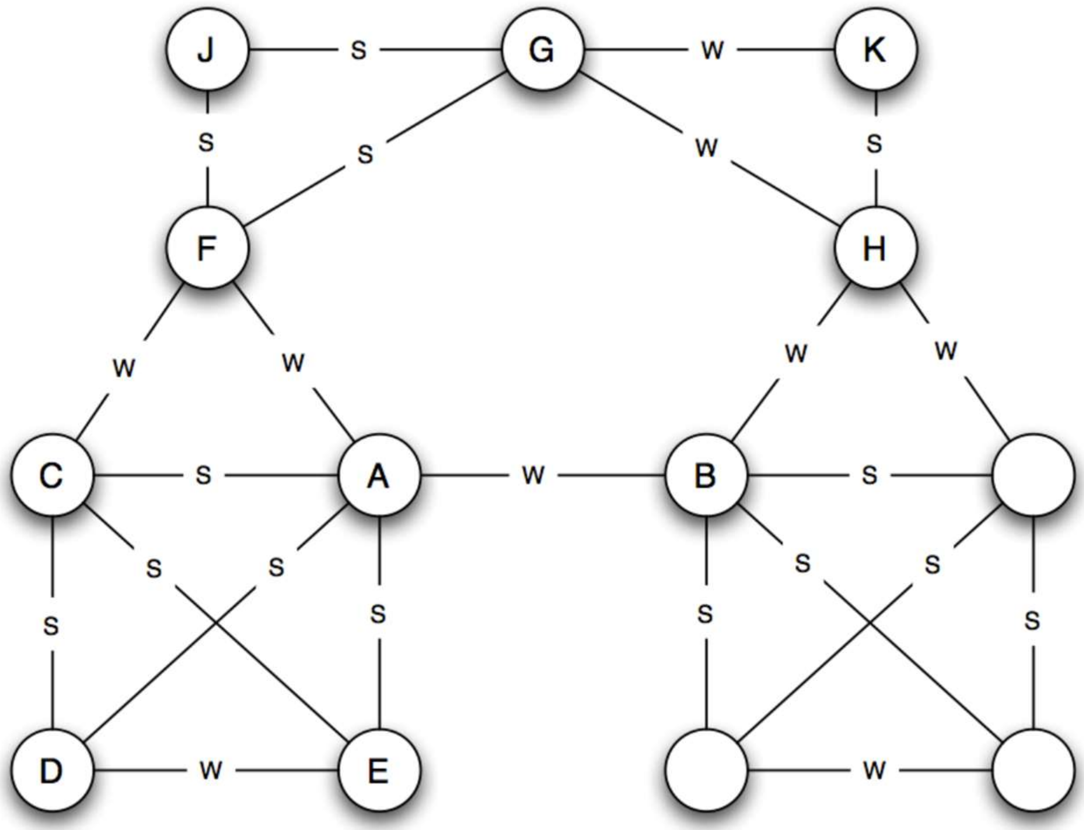


The A-B edge is a local bridge of span 4, since the removal of this edge would increase the distance between A and B to 4.

Strong Triadic Closure Property

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

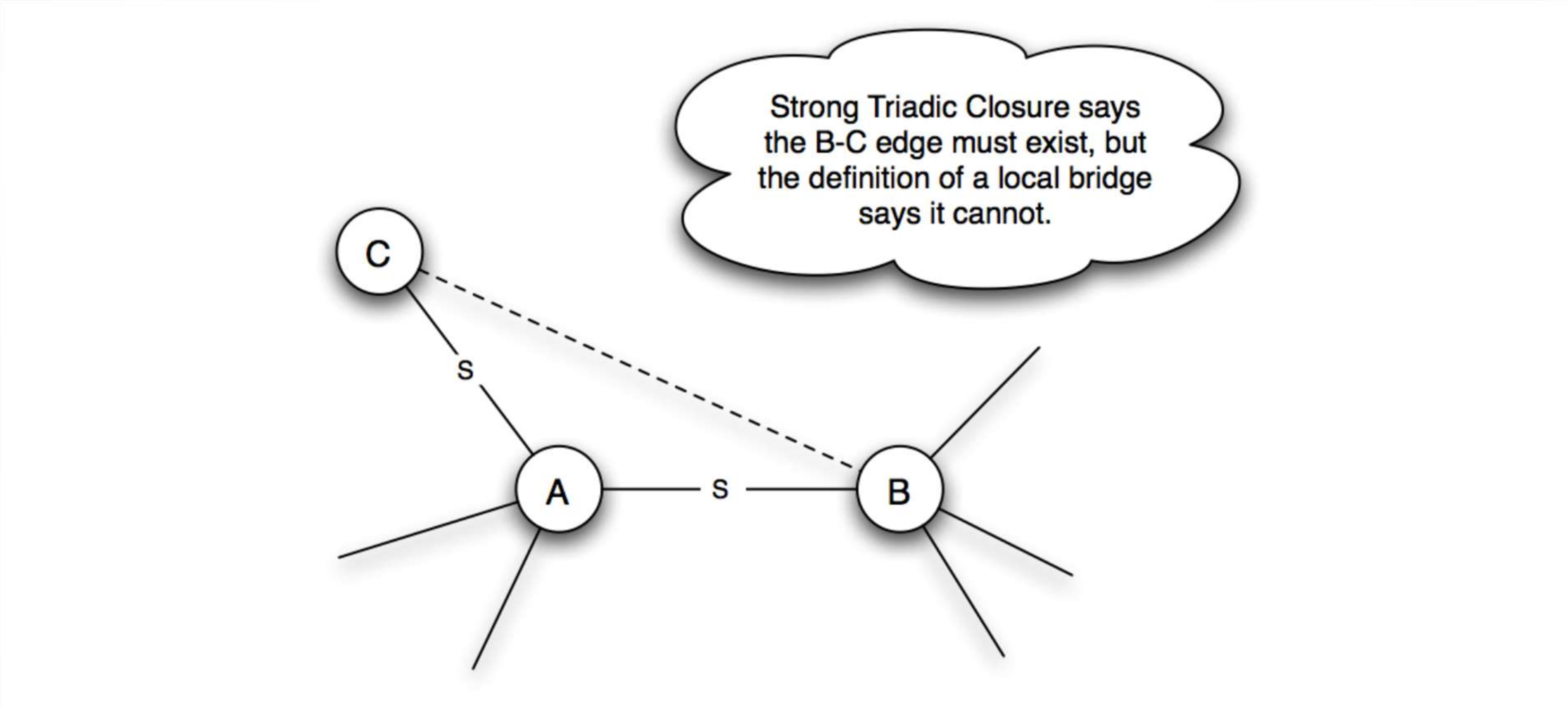


- No node violates the Strong Triadic Closure Property, and hence all nodes satisfy the Property.
- On the other hand, if the A-F edge were to be a strong tie rather than a weak tie, then nodes A and F would both violate the Strong Triadic Closure Property. *Why?*

Local Bridges and Weak Ties

- A purely local, interpersonal distinction between kinds of links — whether they are weak ties or strong ties — as well as a global, structural notion — whether they are local bridges or not.
- What is the connection?
- Claim: 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.
- Assuming the Strong Triadic Closure Property and a sufficient number of strong ties, the local bridges in a network are necessarily weak ties.
- *How to prove this?*

Local Bridges and Weak Ties



- Since A is involved in at least two strong ties, and the edge to B is only one of them, it must have a strong tie to some other node, which we'll call C.
- Since the edge from A to B is a local bridge, A and B must have no friends in common, and so the B-C edge must not exist.
- But this contradicts Strong Triadic Closure, which says that since the A-B and A-C edges are both strong ties, the B-C edge must exist.
- This contradiction shows that our initial premise, the existence of a local bridge that is a strong tie, cannot hold

Insights

- We found the connection between the local property of tie strength and the global property of serving as a local bridge.
- A way to think about the way in which interpersonal properties of social-network links are related to broader considerations about the network's structure.

Strength of Weak Ties

- This analysis provides a concrete framework for thinking about the initially surprising fact that life transitions such as a new jobs are often rooted in contact with distant acquaintances.
- The argument is that these are the social ties that connect us to new sources of information and new opportunities.
- Their conceptual "span" in the social network (the local bridge property) is directly related to their weakness as social ties.
- This dual role as weak connections but also valuable conduits to hard-to-reach parts of the network — this is the surprising strength of weak ties.

Overview

- Triadic Closure
- The Strength of Weak Ties
- Tie Strength and Network Structure in Large-Scale Data
- Tie Strength, Social Media, and Passive Engagement
- Closure, Structural Holes, and Social Capital

Does the results hold true for large social networks?

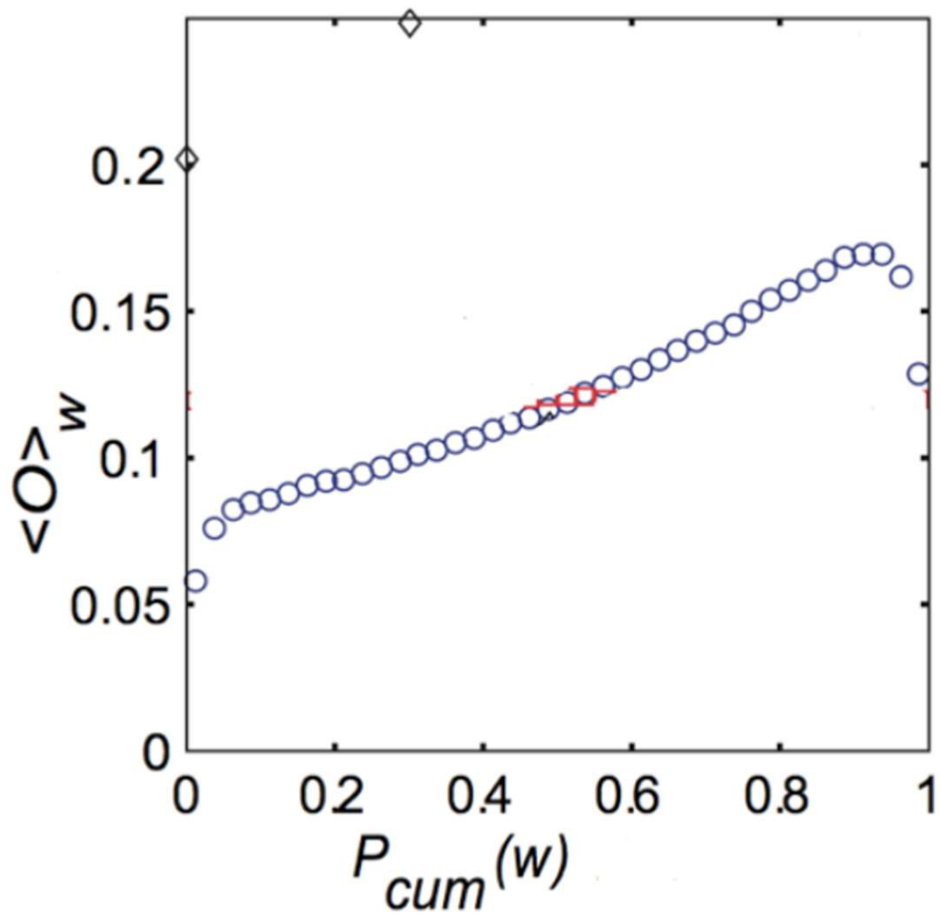
- For many years after Granovetter's work, these predictions remained untested on large social networks..
- Difficulty of finding data that would reliably capture these parameters.
- Enter digital communication data.
- Onnela et al. [2] studied the who-talks-to-whom network maintained by a cell-phone provider that covered roughly 20% of a national population.
  - Note that this network has a Giant Component — a single connected component containing most (in this case 84%) of the individuals in the network

New Metrics

- Tie strength:
  - We can make the strength of an edge a numerical quantity
  - Defining it to be the total number of minutes spent on phone calls between the two ends of the edge.
- It is also useful to sort all the edges by tie strength, so that for a given edge we can ask what percentile it occupies this ordering of edges sorted by strength.
- *What about local bridges and "almost" local bridges?*
- Define the **neighbourhood overlap** of an edge connecting A and B as…

$$\frac{\text{number of nodes who are neighbors of } \textit{both } A \text{ and } B}{\text{number of nodes who are neighbors of } \textit{at least one of } A \text{ or } B}$$

Tie-strength & Network Structure in Large-Scale Data



A plot of the neighbourhood overlap (O) of edges as a function of their percentile (P) in the sorted order of all edges by tie strength. The fact that overlap increases with increasing tie strength is consistent with the theoretical predictions.

Empirical Results on Tie Strength and Neighborhood Overlap

- First, we can ask how the neighborhood overlap of an edge depends on its strength
  - The strength of weak ties predicts that neighbourhood overlap should grow as tie strength grows.
  - This is shown in the Figure
- How can this type of data can be used to evaluate the more global picture suggested by the theoretical framework, that weak ties serve to link together different tightly-knit communities that each contain a large number of stronger ties?
  - Onnela et al., first deleted edges from the network one at a time.
    - Starting with the strongest ties and working downward in order of tie strength, the giant component shrank steadily and gradually.
  - Starting from the weakest ties and working upward in order of tie strength.
    - In this case, they found that the giant component shrank more rapidly, and moreover that its remnants broke apart abruptly once a critical number of weak ties had been removed.

Overview

- Triadic Closure
- The Strength of Weak Ties
- Tie Strength and Network Structure in Large-Scale Data
- **Tie Strength, Social Media, and Passive Engagement**
- Closure, Structural Holes, and Social Capital

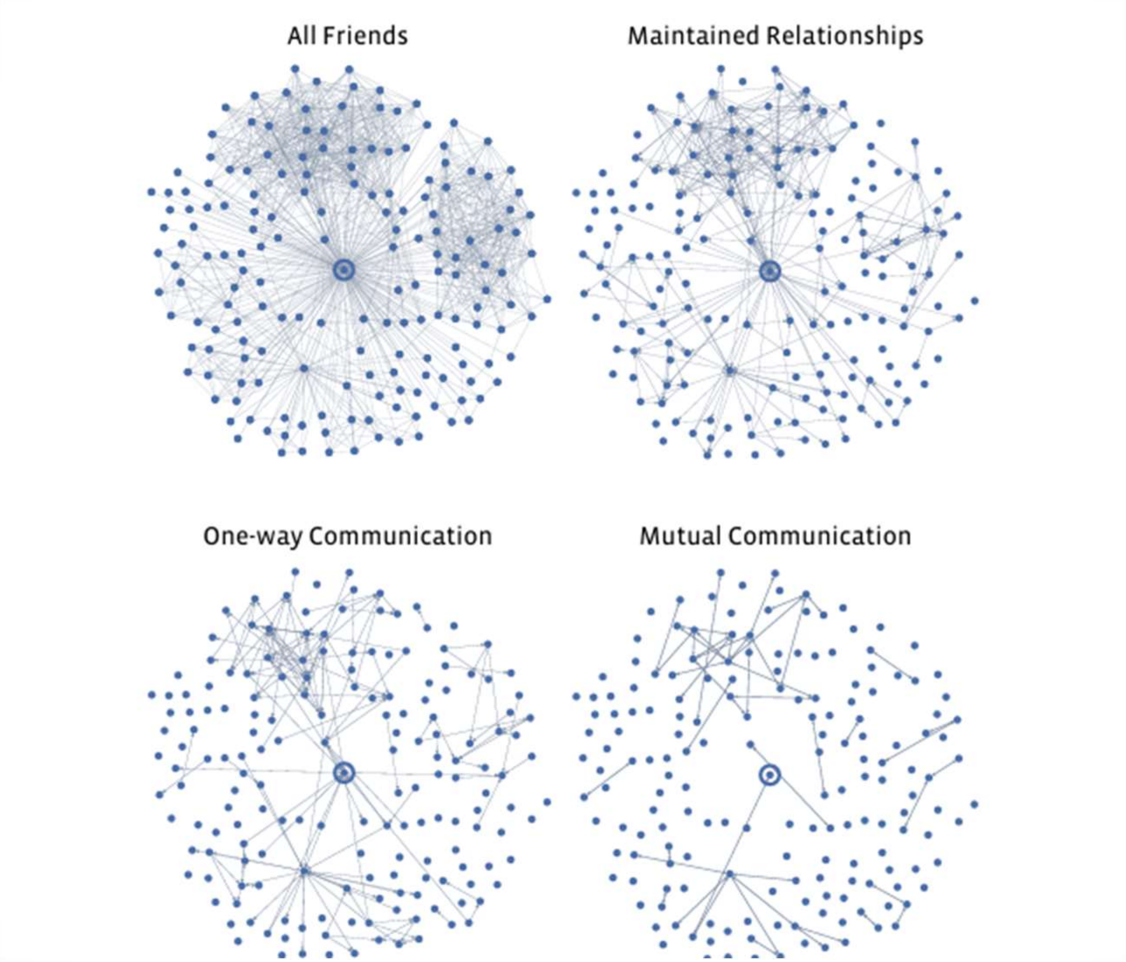
Results from Social Media

- As is well-known to users of social networking tools, people maintain large explicit lists of friends.
- In the past, these lists were much more implicit.
- What effect does this have on social network structure more broadly?
- Tie strength can provide an important perspective on such questions
  - Providing a language for asking how on-line social activity is distributed across different kinds of links
  - In particular, how it is distributed across links of different strengths.

Tie Strength on Facebook

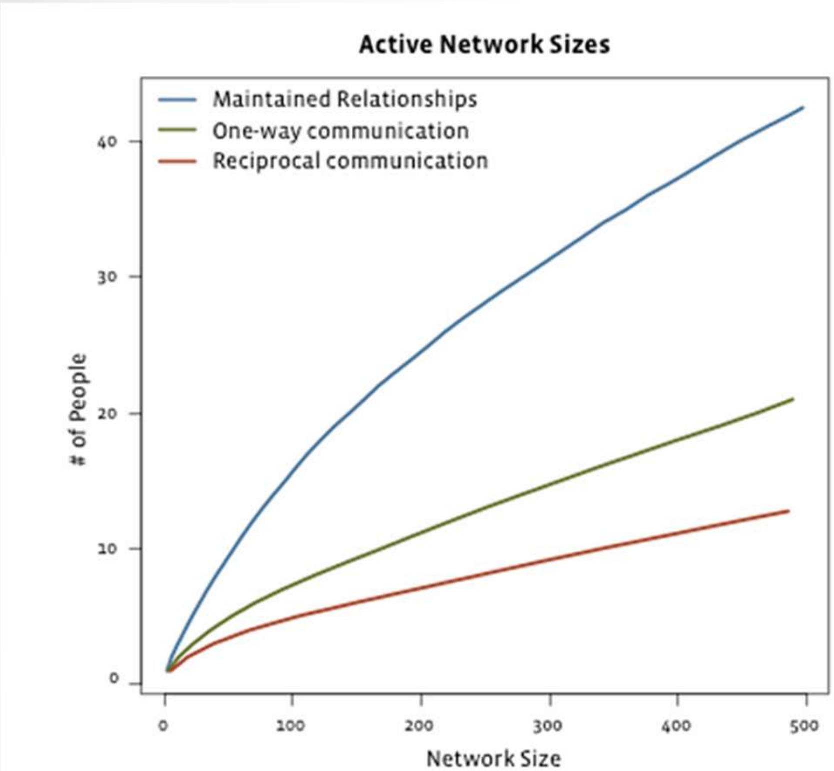
- At Facebook, Cameron Marlow and his colleagues analyzed the friendship links reported in each user's profile, asking to what extent each link was actually used for social interaction, beyond simply being reported in the profile [3].
- In other words, where are the strong ties among a user's friends?
- To make this precise using the data they had available, they defined three categories of links based on usage over a one-month observation period.
  - A link represents **reciprocal (mutual)** communication, if the user both sent messages to the friend at the other end of the link, and also received messages from them during the observation period.
  - A link represents **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 link represents 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 (providing information about the friend) or visiting the friend's profile more than once

Facebook user's network neighborhood



The set of links becomes sparser once we consider only maintained relationships, one-way communication, or reciprocal communication

Passive Network



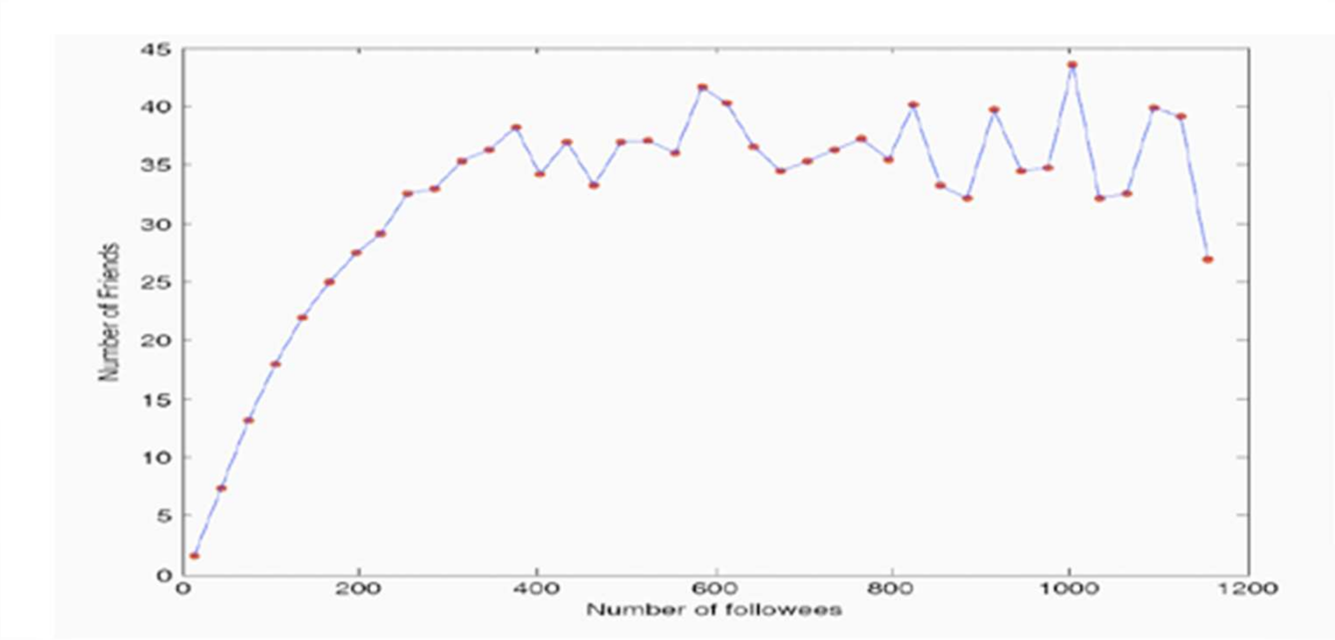
The number of links corresponding to maintained relationships, one-way communication, and reciprocal communication as a function of the total neighborhood size for users on Facebook

- Even for users who report very large numbers of friends on their profile pages the number with whom they actually communicate is generally between 10 and 20
- The number they follow even passively (e.g. by reading about them) is under 50.
- But beyond this observation, Marlow and his colleagues draw a further conclusion about the power of media like Facebook to enable this kind of passive engagement, in which one keeps up with friends by reading news about them even in the absence of communication.
- They argue that this **passive network** occupies an interesting middle ground between the strongest ties maintained by regular communication and the weakest ties from one's distant past, preserved only in lists on social-networking profile pages.

Twitter Tie Strength

- Twitter also includes social-network features, and these enable one to distinguish between stronger and weaker ties:
  - Each user can specify a set of other users whose messages he or she will follow.
    - defines a social network based on more passive, weak ties
    - it is very easy for a user to follow many people's messages without ever directly communicating with any of them
  - Each user can also direct messages specifically to another user. (In the latter case, the remains public for everyone to read, but it is marked with a notation indicating that it is intended for a particular user.)
    - corresponds to a stronger kind of direct interaction.
- Huberman, Romero, and Wu analyzed the relative abundance of these two kinds of links on Twitter [4].
  - Weak Ties: For each user they considered the number of users whose messages she followed (her "followees")
  - Strong Ties: Consist of the users to whom she had directed at least two messages over the course of an observation period.

Twitter Tie Strength



- Shows how the number of strong ties varies as a function of the number of followees.
- Even for users who maintain very large numbers of weak ties on-line, the number of strong ties remains relatively modest.

Strong vs Weak Ties in Social Media

- In Social Media like Facebook and Twitter there is contrast between the ease of forming links and the relative scarcity of strong ties.
- By definition, each strong tie requires the continuous investment of time and effort to maintain, and so even people who devote a lot of their energy to building strong ties will eventually reach a limit — imposed simply by the hours available in a day — on the number of ties that they can maintain in this way.
- The formation of weak ties is governed by much milder constraints — they need to be established at their outset but not necessarily maintained continuously — and so it is easier for someone to accumulate them in large numbers.

- *What about WhatsApp, Instagram?*

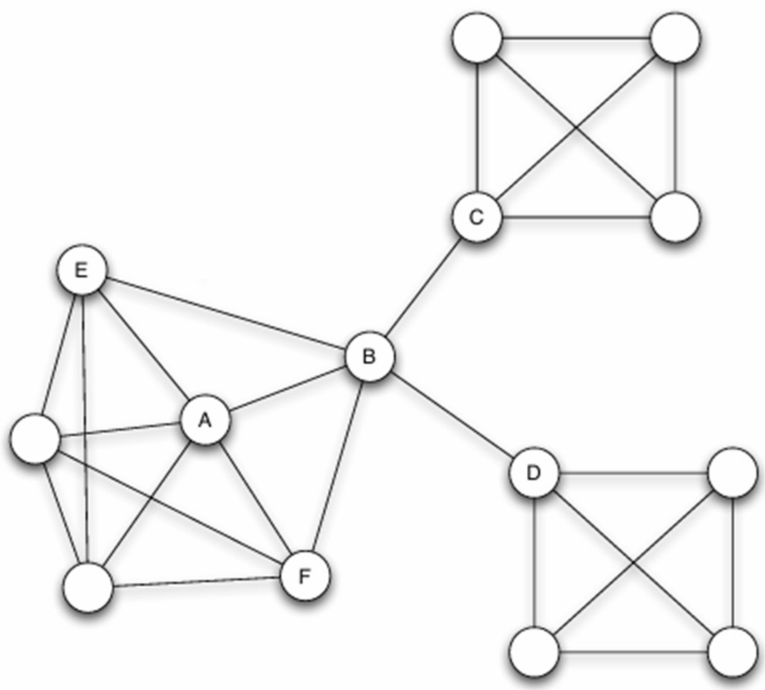
Overview

- Triadic Closure
- The Strength of Weak Ties
- Tie Strength and Network Structure in Large-Scale Data
- Tie Strength, Social Media, and Passive Engagement
- **Closure, Structural Holes, and Social Capital**

Roles of Nodes

- So far, we have a general view of social networks in terms of tightly-knit groups and the weak ties that link them.
- The analysis has focused primarily on the roles that different kinds of **edges** of a network play in this structure —
  - Some edges spanning different groups
  - Some others are surrounded by dense patterns of connections.
- There is a lot of further insight to be gained by asking about the roles that different **vertices** play in this structure as well.
- In social networks, access to edges that span different groups is not equally distributed across all nodes:
  - Some nodes are positioned at the interface between multiple groups, with access to boundary-spanning edges,
  - Others are positioned in the middle of a single group.

Roles of Nodes



The contrast between densely-knit groups and boundary-spanning links is reflected in the different positions of nodes A and B in the underyling social network

Embeddedness

- We define the *embeddedness of an edge* in a network to be the number of common neighbors the two endpoints have.
- embed(A,B) = 2
- This is the numerator in the neighborhood overlap definition
- A local bridge has embeddedness = 0.

Advantages of A

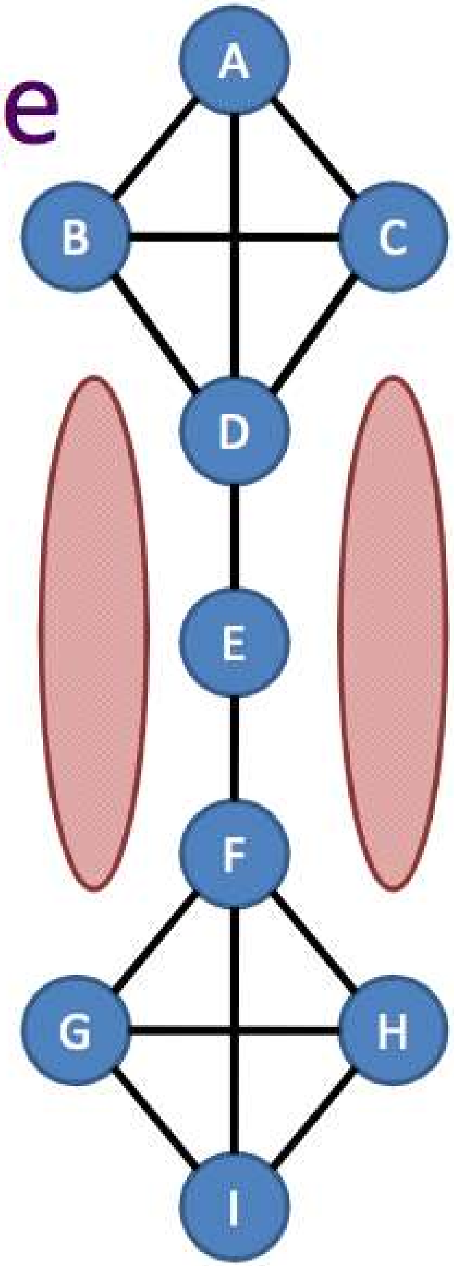
- Node A's set of network neighbors has been subject to considerable triadic closure
- A has a high clustering coefficient.
- All of A's edges have significant embeddedness
- Sociological research has argued that if two individuals are connected by an embedded edge, then this makes it easier for them to trust one another, and to have confidence in the integrity of the transactions
- The presence of mutual friends puts the interactions between two people "on display" in a social sense, even when they are carried out in private;
- In the event of misbehaviour by one of the two parties to the interaction, there is the potential for social sanctions and reputational consequences from their mutual friends.

Structural Hole

- For edges with zero embeddedness, since there is no one who knows both people involved in the interaction.
- Moreover, the constraints on B's behavior are made complicated by the fact that she is subject to potentially contradictory norms and expectations from the different groups she associates with.
- Node B, with her multiple local bridges, spans a **structural hole** in the organization — the "empty space" in the network between two sets of nodes that do not otherwise interact closely [5]
- Unlike the term "local bridge," which has a precise mathematical definition in terms of the underlying graph, definition of the term "structural hole" is somewhat informal

Structural holes & Brokerage

- **Structural holes**
  - Places where people are unconnected in a network
- **Brokers**
  - Actors who exploit structural holes
  - Gain access to information, power to filter, timing for competitive advantage, and ability to refer other actors
  - Difficult entrée, requires accurate maps of relationships in each groups, costly to maintain, high potential to be undercut



Advantages of B

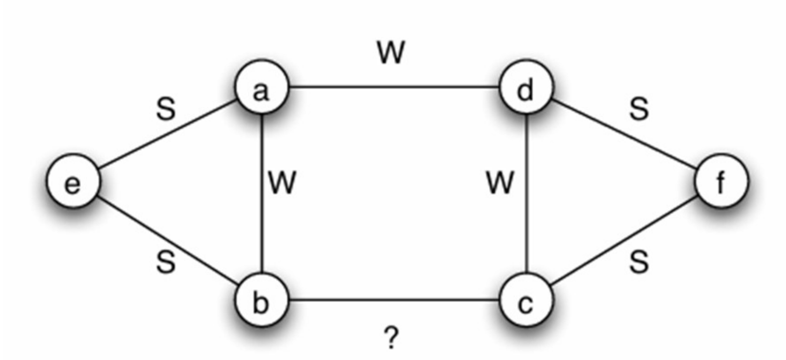
1. Informational: B has early access to information originating in multiple, non-interacting parts of the network.
2. Creativity: standing at one end of a local bridge can be an amplifier for creativity.
   - Experience from many domains suggests that innovations often arise from the unexpected synthesis of multiple ideas, each of them on their own perhaps well-known, but well-known in distinct and unrelated bodies of expertise
3. Gate- keeping: B’s position in the network provides an opportunity for a kind of social “gate- keeping”
   - She regulates the access of both C and D to the tightly-knit group she belongs to
   - She controls the ways in which her own group learns about information coming from C’s and D’s groups.
   - This becomes a source of *power* for B.

Social Capital

- All of these arguments are framed in terms of individuals and groups deriving benefits from an underlying social structure or social network; as such, they are naturally related to the notion of *social capital*.
- Social capital is sometimes viewed as a property of a group, with some groups functioning more effectively than others because of favorable properties of their social structures or networks.
- Alternately, it has also been considered as a property of an individual; a person can have more or less social capital depending on his or her position in the underlying social structure or network.

Exercise

- According to the theory of strong and weak ties, with the strong triadic closure assumption, how would you expect the edge connecting *b* and *c* to be labeled?



References

1. Mark Granovetter. Getting a Job: A Study of Contacts and Careers. University of Chicago Press, 1974.
2. J.-P. Onnela, J. Saramaki, J. Hyvonen, G. Szabo, D. Lazer, K. Kaski, J. Kertesz, and A.-L. Barabasi. Structure and tie strengths in mobile communication networks. Proc. Natl. Acad. Sci. USA, 104:7332{7336, 2007
3. Cameron Marlow, Lee Byron, Tom Lento, and Itamar Rosenn. Maintained relationships on Facebook, 2009. On-line at http://overstated.net/2009/03/09/maintainedrelationships-on-facebook.
4. Bernardo A. Huberman, Daniel M. Romero, and Fang Wu. Social networks that matter: Twitter under the microscope. First Monday, 14(1), January 2009.
5. Ronald S. Burt. Structural Holes: The Social Structure of Competition. Harvard University Press, 1992

Readings

- Networks, Crowds, and Markets: Reasoning About a Highly Connected World https://www.cs.cornell.edu/home/kleinber/networks-book/
  - Chapter 3.1-3.5