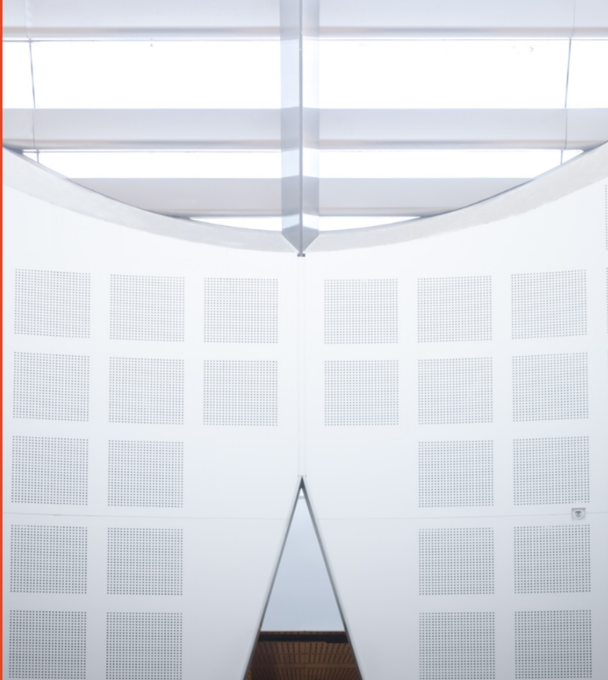
# COMP5313/COMP4313 - Large Scale Networks

## Week 2a: Graph Ties

**Lijun Chang**

March 6, 2025

THE UNIVERSITY OF SYDNEY



# Key Concepts Today

- Small world phenomenon, six degree of separation
- Strength of weak ties, local bridges, strong triadic closure
- Embeddedness, structural holes

- These concepts were firstly studied by social scientists on small-scale networks (in 1960s–1970s), then computer scientists verified these ideas on large-scale networks (in 1990s–2000s)

# Outline

# The Small World Phenomenon

- The small-world phenomenon is the idea that the world "looks" small
  - Also known as the "six degree of separation"

- In the 1960s, Stanley Milgram did an experiment to test the idea that people are really connected in the global friendship by short chains of friends, i.e., the distance between any pair of people is remarkably short. [1]
  - He asked 296 randomly chosen "starter" individuals to each try forwarding a letter to a designated "target" person living in Sharon, MA, USA.
  - He provided name, address, occupation, and personal information of the target person.
  - The participants could not mail the letter directly to the target
  - The participants could only forward the letter to an acquaintance that (s)he knew, with the goal of reaching the target as fast as possible

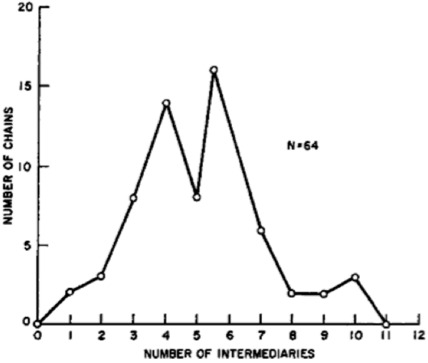
---

[1] S. Milgram. The small world problem. *Psychology today*, vol. 2, no. 1, 1967.

# The Small World Phenomenon

- $\frac{1}{5}$ of the letters eventually arrived at the target
- The median number of steps was 6
- This served as an experimental evidence for the existence of many short paths

Distribution of path lengths among the sixty-four chains that succeeded in reaching the target. The median length was 6.



# The Small World Phenomenon

- There are few caveats though:
  - The six degree of separation does not apply to all pairs of nodes in the network
  - The paths were to a single and fairly affluent (known) target
  - Many letters never arrived
  - Does short distance means that socially close?
- Nevertheless, many people consider social networks as small worlds
  - This is because the phenomenon was also confirmed in settings where we have the full network structure

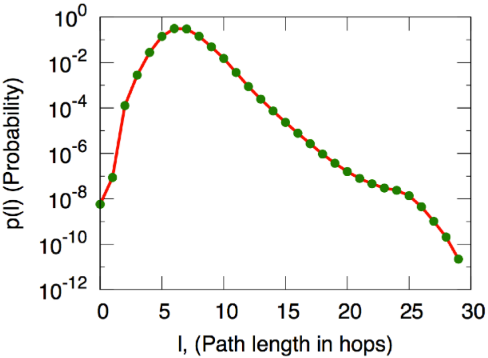
# The Small World Phenomenon

- One of the largest computational study analyzed 240M active Microsoft Instant Messaging users. [2]
  - A node represents a user
  - There is an edge between $u$ and $v$ if they engaged in a two-way conversation during a month-long observation period
- To get the distances of these given network structures
  - Code the Breadth-First Search algorithm and load the structure in a computer
  - Run the algorithm on the structure and extract the distances
- Conclusions:
  - The network contains a giant component containing almost all the nodes
  - The distances within this component were very small
    - Average distance of 6.6 and median distance of 7

[2] J. Leskovec and E. Horvitz, "Planetary-scale views on a large instant-messaging network," in *Proceedings of the 17th international conference on World Wide Web*, pp. 915–924, ACM, 2008.

# The Small World Phenomenon

Distribution of distances averaged over a random sample of 1000 users of the graph of all active Microsoft Instant Messenger user accounts during one month.




- Remarks:
  - The experiment is quite different from Milgram's
  - Users are technologically endowed and conversation does not imply friendship

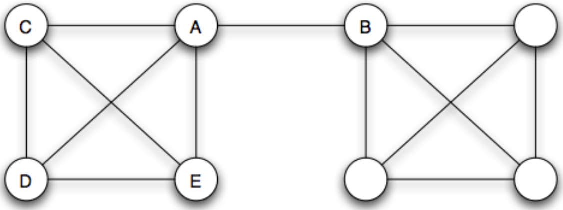
# Outline

## On the Power of Acquaintances

- In the 1960s, Mark Granovetter interviewed people who had recently changed employers to learn how these people had discovered their new jobs. [3]
- Personal contacts led many people to their current job
  - More acquaintances rather than close friends
- While your close friends have more motivation to help you, why are your acquaintances so much more helpful?
  - Information about jobs is intuitively relatively scarce
  - Hearing about jobs from others means they have access to info you don't
  - Acquaintances span different regions of the network that have access to different information

[3] M. Granovetter, Getting a job: A study of contacts and careers. *University of Chicago press*, 1974.

# On the Power of Acquaintances

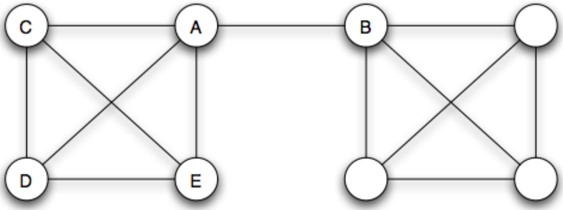



- A's close friends C, D, E connect her to a tightly-knit group of friends, while A's acquaintance B reaches into a different part of the network

  $\Longrightarrow$ A, C, D, E will be exposed to the same info

  $\Longrightarrow$ B offers A access to info she wouldn't be able to access otherwise

## Bridges and Local Bridges



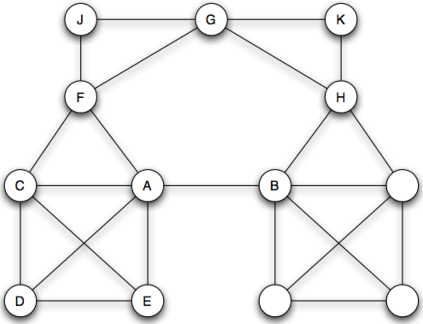

- Edge (A,B) is a bridge meaning that its removal would place A and B in distinct connected components.
- Bridges provide nodes with access to parts of the network that would be unreachable by other means.

## Bridges and Local Bridges

- Remember: giant connected components tell us that bridges are presumably extremely rare in real social networks
- You may have a friend from a very different background, and it may seem that your friendship is the only thing that bridges your world and his world
- But one expects in reality that there will be other multi-step paths that also span these worlds
- Hence the previous graph is unlikely to standalone and he is more likely embedded into a larger graph

# Bridges and Local Bridges

- Let's have a look at such a larger graph.

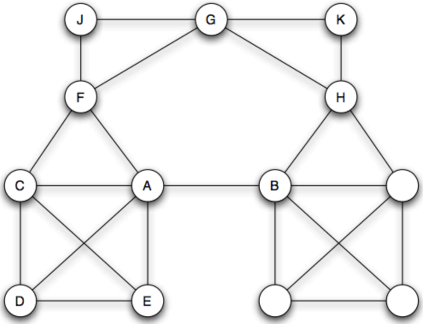



- The edge (A,B) is not the only path that connects its two endpoints
- There is a longer path through F, G and H

# Bridges and Local Bridges

- A local bridge is an edge joining two nodes A and B in a graph if its endpoints A and B have no friends in common
- In other words, deleting a local bridge between A and B would increase the distance between A and B to a value strictly larger than 2
- We say that the span of a local bridge is the distance its endpoints would be from each other if the edge were deleted

# Bridges and Local Bridges

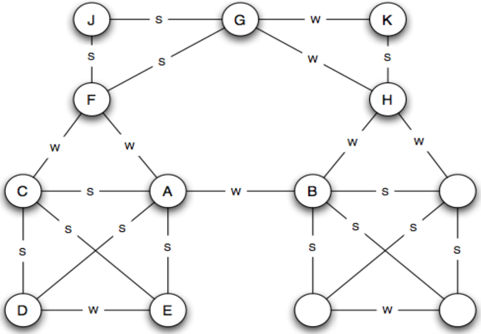



- Edge (A,B) is a local bridge and has a span of 4.
- Local bridges are more common than bridges
- Local bridges that have a large span play a role similar to bridges

$\Longrightarrow$ Local bridges span different regions of a network

# Strong and Weak Ties

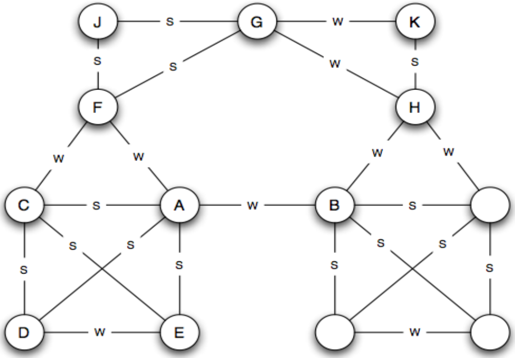
- To distinguish close friends and acquaintances, edges are classified into strong ties and weak ties.



# Strong Triadic Closure Property

- We say that 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 (be it strong or weak) between B and C. We say that node A satisfies the Strong Triadic Closure Property if it does not violate it.

All nodes in this figure satisfy the Strong Triadic Closure Property



# Weak Ties and Local Bridges

- Using triadic closure we can establish a connection between a purely local differentiation of edges and the global structural notion of bridges
  - Strong/weak ties
  - Local bridges or not

**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

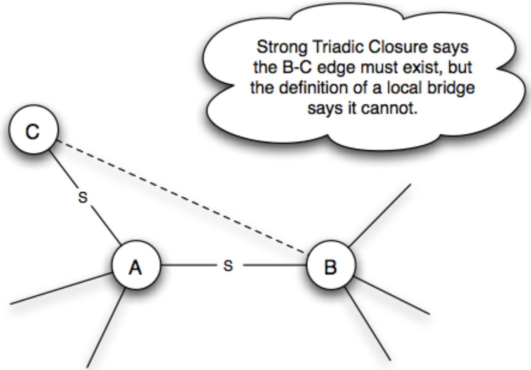
- Assuming the Strong Triadic Closure Property and a sufficient number of strong ties, the local bridges in a network are necessarily weak ties.

# Weak Ties and Local Bridges

**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

- **Proof by contradiction:**



# Empirical Validation on Large Scale Networks

- To empirically validate the idea that local bridges are weak ties, Onnela et al. studied the who-talks-to-whom network maintained by a cell phone provider. [4]
  - Nodes are cell phone users
  - Edges represent a phone call to each other in both directions over an 18 week observation period
    - Phone calls were personal (not business phone calls)
    - Cell phone numbers were generally exchanged between people who know each other

  $\Longrightarrow$ Could be viewed as conversations occurring in a social network

- Network statistics
  - It covered 20% of a national population
  - A single giant connected component included 84% of the nodes

[4] J.-P. Onnela, J. Saramäki, J. Hyvönen, G. Szabó, D. Lazer, K. Kaski, J. Kertész, and A.-L. Barabási, "Structure and tie strengths in mobile communication networks," *Proceedings of the national academy of sciences*, vol. 104, no. 18, pp. 7332–7336, 2007.

# Empirical Validation on Large Scale Networks

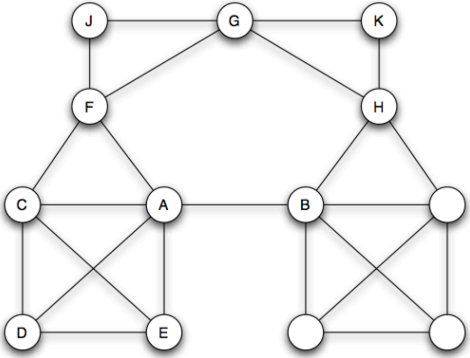
- Constraints of our definitions:
  - An edge is either a weak tie or a strong tie
  - An edge is either a local bridge or not
- Generalize tie strength and local bridge to exhibit smoother gradation
  - A tie strength can be a numerical quantity
    - total number of minutes spent on phone between the two end-points of an edge
  - Since a very small fraction of the edges are local bridges, let's define the
    - Neighborhood overlap:

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

## Empirical Validation on Large Scale Networks

**Example of neighborhood overlap:**

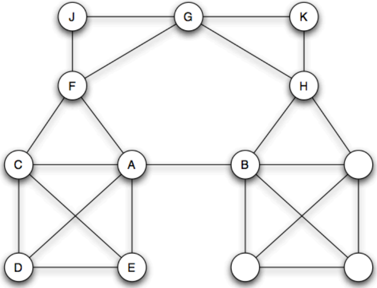
- There are 6 neighbors of A or F: B, C, D, E, G and J.
- There is only 1 of these that is common to both A's neighbors and F's neighbors: C
- Hence the neighborhood overlap of edge (A,F) is $\frac{1}{6}$



# Empirical Validation on Large Scale Networks

- When does the neighborhood overlap nullify?
  - When the numerator is 0
  - This happens when the edge is a local bridge

    $\Longrightarrow$ Edges with low neighborhood overlap are "almost" local bridges

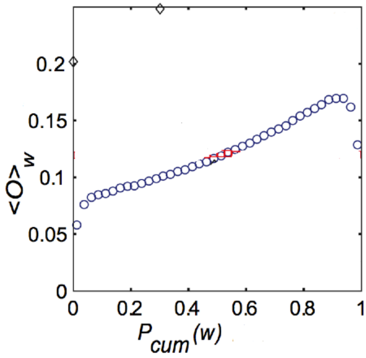



  - (A,F) is closer to be a local bridge than (A,E) for example

# Empirical Validation on Large Scale Networks

- The figure shows neighborhood overlap of edges as a function of their percentile in the sorted order of all edges by tie strength

- The relationship between these quantities aligns well with the theoretical predictions
  - Neighborhood overlap grows as tie strength grows



# Empirical Validation on Large Scale Networks

- How these data can show that weak ties connect tightly-knit communities?
  - Onnela et al. first deleted edges in descending order of their strength [5]
    - The giant component shrank steadily
  - They then re-started but by deleting edges in ascending order of their strength
    - The giant component shrank more rapidly
    - Its remnants broke abruptly once a critical number of weak ties were removed
      $\Longrightarrow$ This is consistent with a picture in which the weak ties provide the more crucial connective structure for holding together disparate communities

[5] J.-P. Onnela, J. Saramäki, J. Hyvönen, G. Szabó, D. Lazer, K. Kaski, J. Kertész, and A.-L. Barabási, "Structure and tie strengths in mobile communication networks," *Proceedings of the national academy of sciences*, vol. 104, no. 18, pp. 7332–7336, 2007.

# Outline

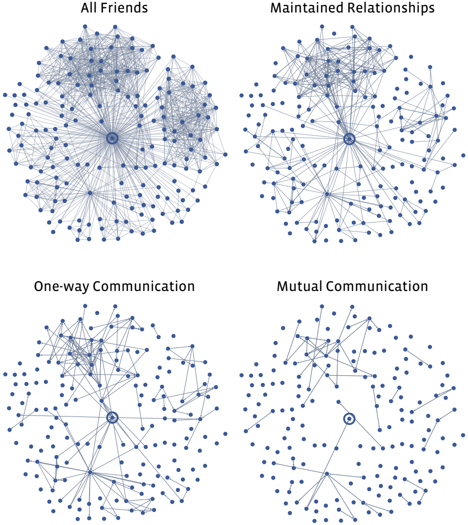
# Tie Strength on Facebook

Facebook

- Friendship is symmetric: if A is friend with B, then B is friend with A
- Users post direct public messages on other users' wall
- Users can send private messages to other users through (instant) messages
- Users post indirect messages, photos, videos, links that other users may see
  - A user's view aggregates friends' posts
  - Following passively

# Tie Strength on Facebook

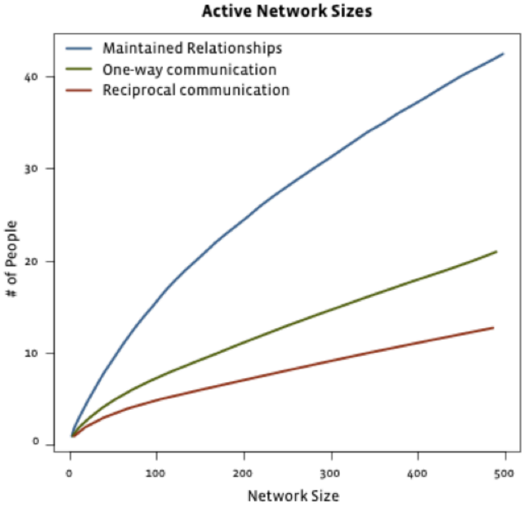
- A link represents reciprocal (mutual) communication, if the user both sent messages to the friend at the other end of the link, and also received messages from them
- A link represents a one-way communication if the user sent one or more messages to the friend at the other end of the link
- A link 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
- These three categories are not mutually exclusive. [a]




[a] https://overstated.net/2009/03/09/maintained-relationships-on-facebook/

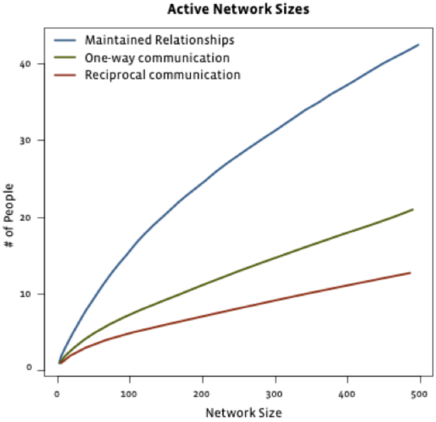
# Tie Strength on Facebook

- 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.



# Tie Strength on Facebook

- Even for users with large set of friends ($\sim$ 500)
  - They actually communicate with 10-20 (these are the strong ties), and
  - They follow passively only 50 other users
  - The remaining links are weak ties
- The set of users followed passively is interesting
  - It is a middle ground between strong and weak ties
  - It allows important news (e.g., new born baby) to disseminate very rapidly through this highly connected network



# Tie Strength on Twitter

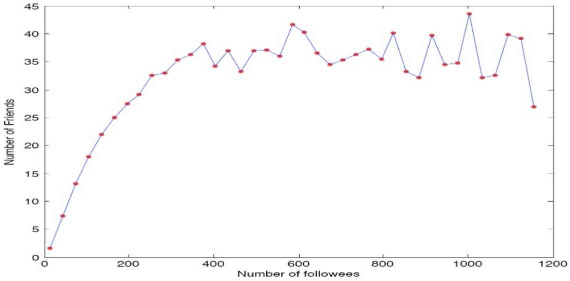
- Twitter
  - Online micro-blogging service
  - Users post tweets that are very short (280 characters)
  - Users specify users they want to follow, i.e., from which they want to receive public messages

    $\Longrightarrow$ A follower of $u$ is someone who follows $u$

    $\Longrightarrow$ A followee of $u$ is someone that $u$ follows
  - Users can direct messages to another user (even though it is public, it is marked)
- Followers define weak ties of a social network
- Directing messages define a stronger kind of direct interaction

# Tie Strength on Twitter

- A study on Twitter shows that users with 1000 followees have less than 50 strong ties. [6]
  - A strong tie between $u$ and $v$ if $u$ sent two direct messages to $v$ during the observation period




[6] B. A. Huberman, D. M. Romero, and F. Wu, "Social networks that matter: Twitter under the microscope," *CoRR*, vol. abs/0812.1045, 2008.

## Tie Strength in Online Social Networks

- It seems that network of strong ties can still be sparse where weak ties abound
  - Strong tie requires a continuous investment of time and effort to maintain
  - Weak tie does not need to be maintained continuously, so it is easier for someone to accumulate them in large numbers.

# Outline

# Roles of Nodes

- We have analysed the roles that different kinds of edges of a network play. Now we are interested in the roles that different nodes play in a network structure.
- Access to edges that span different groups is not equally distributed across all nodes
  - Some nodes are at the interface between multiple groups
  - Others are placed at the middle of a single group
- What is the effect of this heterogeneity?
  - We can answer this question with a story about node experience in a network

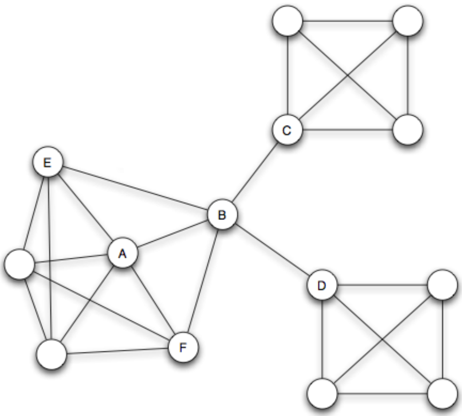
# Embeddedness

- The embeddedness of an edge in a network is the number of common neighbors its two endpoints have
- A long line of research in sociology has argued that, if two individuals are connected by an embedded edge, it is easier for them to trust each another
  - The presence of mutual friends puts the interactions between the two people "on display" in a social sense
  - In case of misbehavior by one of the two parties, there is potential for social sanctions and reputational consequences from their mutual friends
- Granovetter said "My mortification at cheating a friend of long standing may be substantial even when undiscovered. It may increase when a friend becomes aware of it. But it may become even more unbearable when our mutual friends uncover the deceit and tell one another"

# Roles of Nodes

A and B have different roles

- All of A's edges have significant embeddedness
- Interactions that B has with C and D are much more riskier than the embedded interactions experienced by A.
- Moreover B is subject to potentially contradictory norms from its different groups



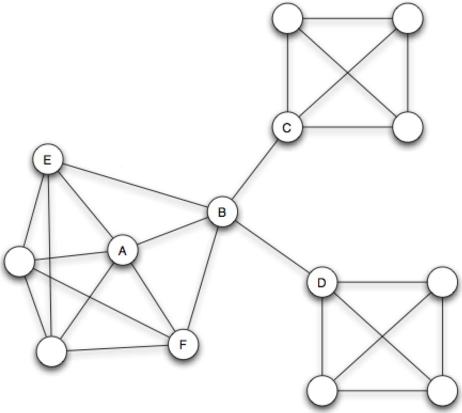
# Structural Holes

- In the previous example, the advantage of A is that A can trust all friends.
  - All of A's edges have significant embeddedness
- But, what about the advantages of B, i.e, being at the end of local bridges?
- Empirical studies of managers in large corporations have correlated an individual's success within a company with the access to local bridges. [7]
  - Managers are nodes
    - Collaborate on common objectives
    - Implicitly compete for career advancement
  - Relations are edges
    - Managers are linked if they know each other and talk to each other
    - The edges do not represent hierarchical reporting relations

[7] V. Bush and J. Wang, "As we may think," *Atlantic Monthly*, vol. 176, pp. 101–108, 1945.

# Structural Holes

- 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 (Informal definition).
- B's position offers advantages relative to A
  - B has access to information originating in multiple, non-interacting parts of the network
  - Standing at one end of a local bridge can amplify creativity
  - B has a gatekeeper role by regulating accesses of C and D to her tightly knit groups, and by controlling the ways in which her own group learns about information coming from C and D's groups.



# Structural Holes

- B is powerful as a gatekeeper.
  - B may not want to loose the gatekeeping role
  - B may want to prevent triangles from forming around their local bridges
    - Another edge from C or D into B's group would reduce B's gatekeeping role.

- Her interest may not be aligned with the one of the company
  - The organization may want to accelerate the flow of information between groups

# Conclusions

- The median/average distances between nodes is often surprisingly small
- Strength of ties is heterogeneous
  - Intuitively, the strength translates into the frequency of interactions
  - The number of strong ties is generally low compared to the number of weak ties
  - Weak ties can be extremely useful in disseminating rare/important information in social networks
- Tightly-knit regions and local bridges are natural concepts
  - Tightly-knit regions induce implicit trust in social networks
  - Local bridges (in holes) often bring importance to its end points
- Identifying these tightly knit regions and holes
  - Is useful (to maximize information flow, detect powerful nodes)
  - But requires a lightweight algorithm to apply to large scale

# Reading

- Reading for this week
  - Chapter 3 of the textbook, excluding the advanced material
- Reading for next week
  - Chapters 4 and 5 of the textbook