VE444: Networks

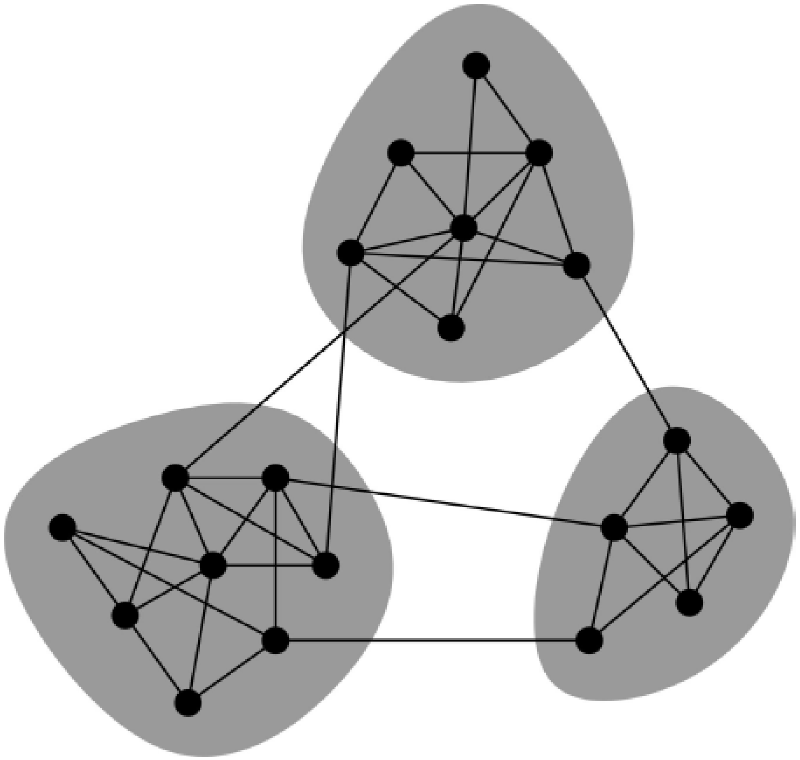
Yifei Zhu, assistant professor
University of Michigan-Shanghai Jiao Tong University

Strong and weak ties

Acknowledgment:
This lecture's slides are modified from the slides shared by Prof. Jure Leskovec

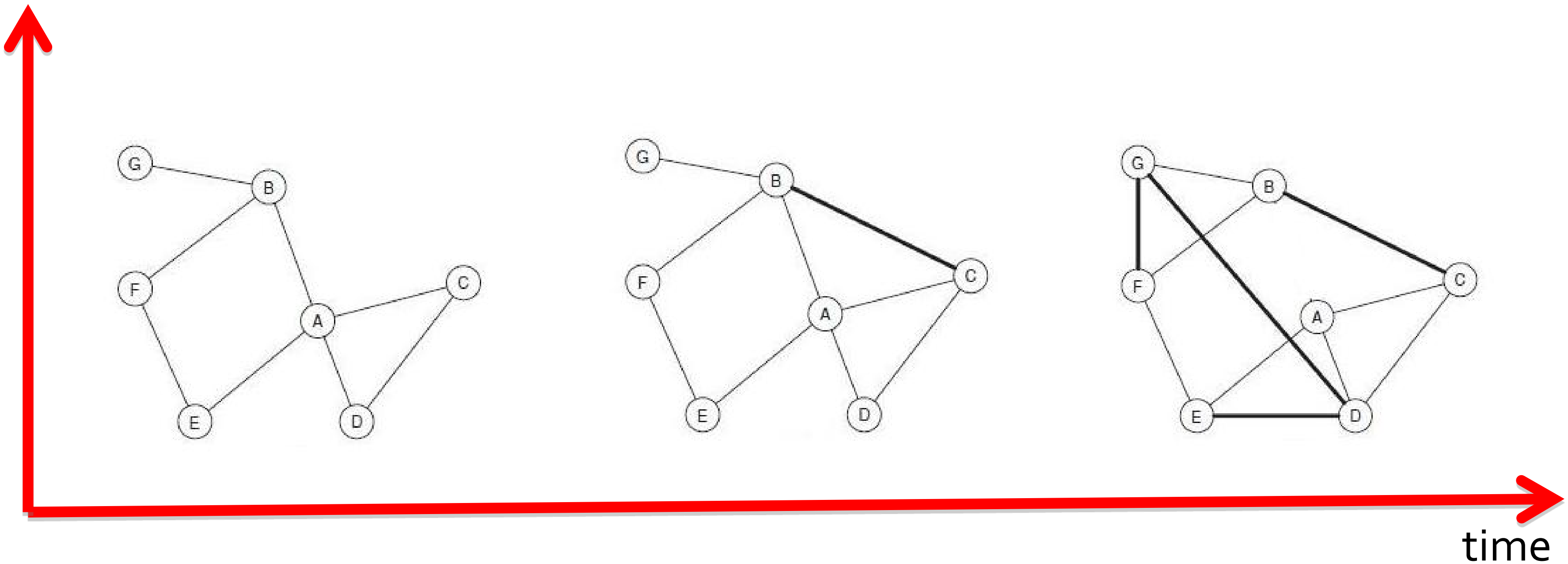
Networks & Communities

- We often think of networks "looking" like this:



- What led to such a conceptual picture?

What if we consider the evolution of the a network



- Consider not only a snapshot, but evolution

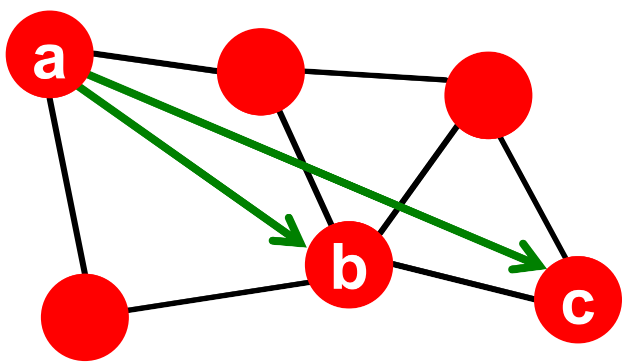
Motivating example

- **Network could facilitate the flow of information**
- **How do people find out about new jobs?**
  - Mark Granovetter, part of his PhD in 1960s
  - People find the information through personal contacts
- **But:** Contacts were often **acquaintances** rather than close friends
  - **This is surprising:** One would expect your friends to help you out more than casual acquaintances
- **Why is it that acquaintances are most helpful?**

[Granovetter '73]

Granovetter's Answer

- **Two perspectives on friendships:**
  - **Structural:** Friendships span different parts of the network
  - **Interpersonal:** Friendship between two people is either **strong** or **weak**
- **Structural role: Triadic Closure**

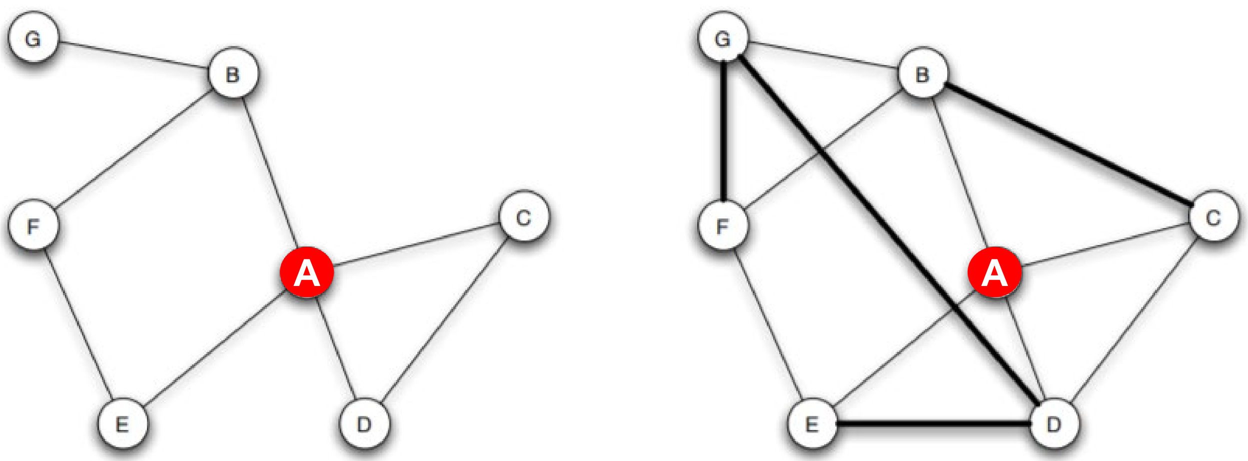


Which edge is more likely, a-b or a-c?

If two people in a network have a friend in common, then there is an increased likelihood they will become friends themselves.

Triadic Closure

- **Motivation of the clustering coefficient**



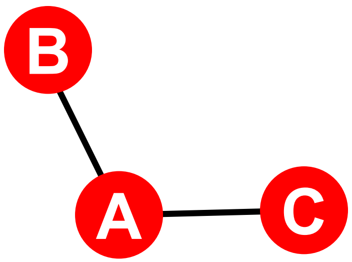
(a) *Before new edges form.* (b) *After new edges form.*

- **Triadic closure = High clustering coefficient**

Reasons for Triadic Closure

Reasons for triadic closure:

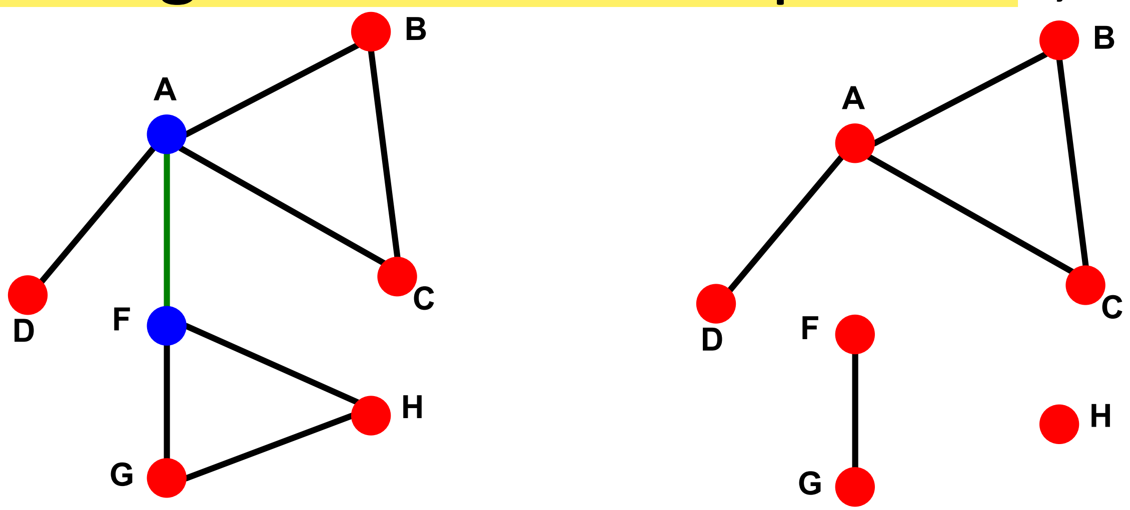
- If $\boldsymbol{B}$ and $\boldsymbol{C}$ have a friend $\boldsymbol{A}$ in common, then:
  - $\boldsymbol{B}$ is **more likely to meet** $\boldsymbol{C}$
    - (since they both spend time with $\boldsymbol{A}$)
  - $\boldsymbol{B}$ and $\boldsymbol{C}$ **trust** each other
    - (since they have a friend in common)
  - $\boldsymbol{A}$ has **incentive** to bring $\boldsymbol{B}$ and $\boldsymbol{C}$ together
    - (since it is hard for $\boldsymbol{A}$ to maintain two disjoint relationships)



- **Empirical study by Bearman and Moody:**
  - Teenage girls with low clustering coefficient are more likely to contemplate suicide

Bridges and local bridges

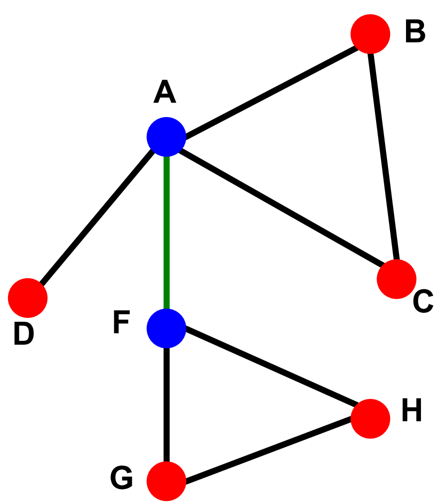
- How does triadic closure relates to the Granovetter's study?
- **Bridge** describes an edge's role where deleting this edge would cause its corresponding vertices falling in different components;

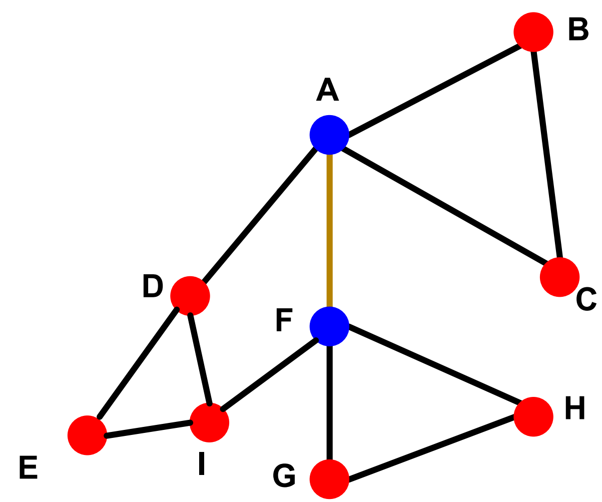


Articulation node: If we erase the **node**, the graph becomes disconnected

Bridges and local bridges

- An edge is a **local bridge** if its end points have no friends in common
  - (Alternatively, distance perspective?)
  - **Span** of a local bridge: the distance its endpoints would be from each other if the edge were deleted



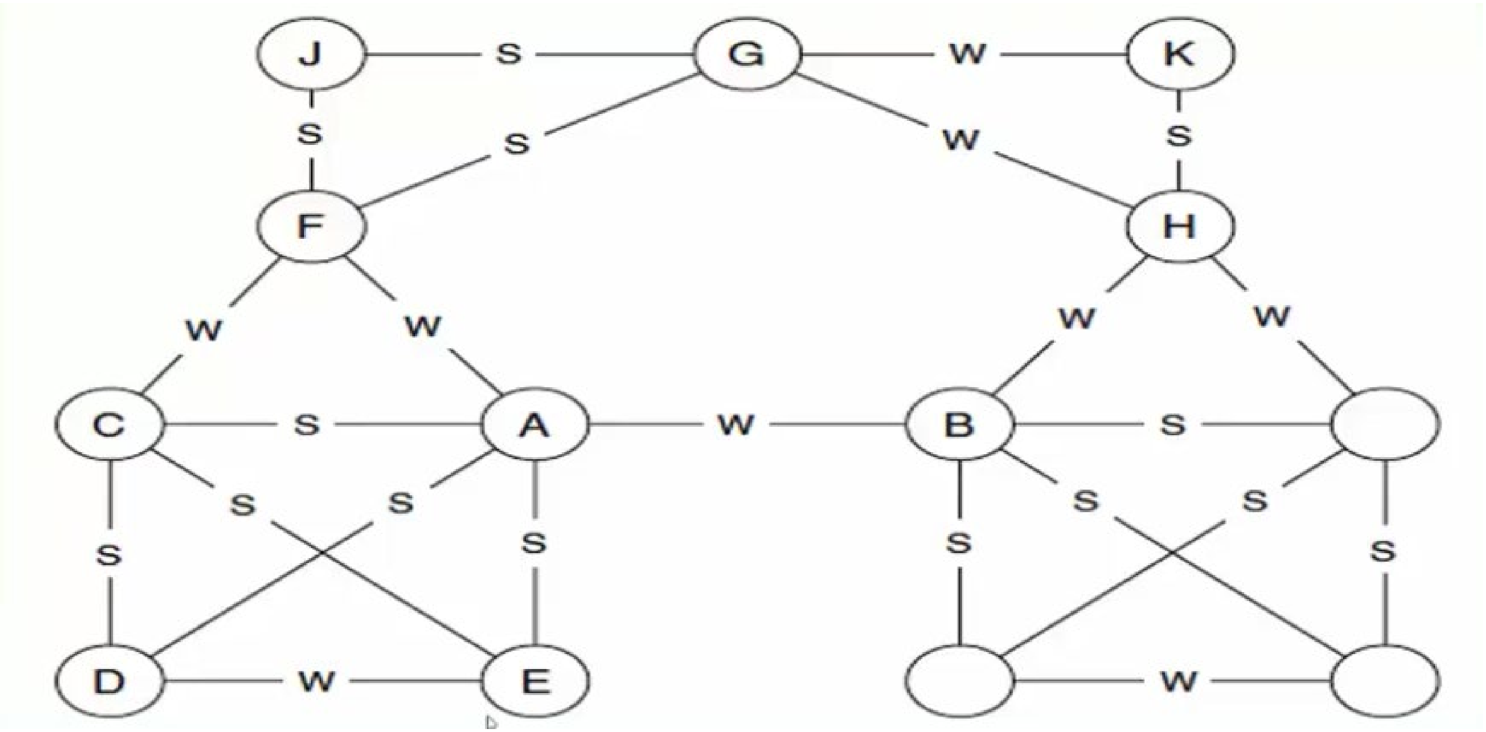


Strong and weak ties

- If we can categorize all links in the social networks as **strong ties** and **weak ties**
- Relating to the triadic closure, we make the following assumption:
  - 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.
  - More formally, as Granovetter suggested:
- 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 and weak tie) between B and C.

Strong triadic closure property

- Which node violates the strong triadic closure property?

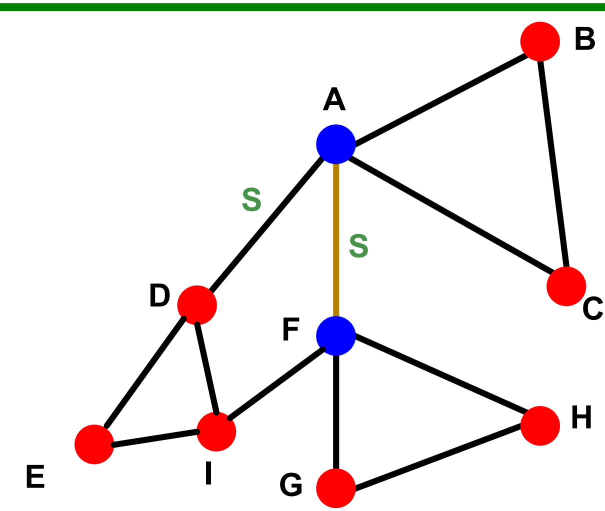


Structural level meets interpersonal level

- **Bridge**: global structural notation
- **Weak/Strong**: local interpersonal distinction
- How they link to each other? Granovetter's Theorem

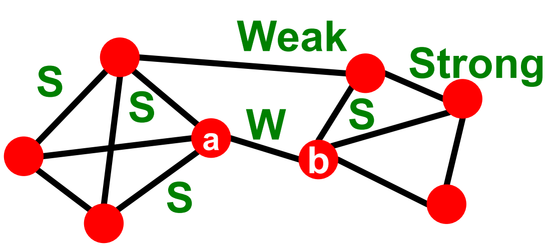
> if a node A in a network satisfies the STC property and is involved in at least two strong ties, then any local bridge it is involved in must be a weak tie

Local bridge



Granovetter's Explanation

- **Granovetter makes a connection between the social and structural role of an edge**
- **First point: Structure**
  - Structurally embedded edges are also socially strong
  - Long-range edges spanning different parts of the network are socially weak
- **Second point: Information**
  - Long-range edges allow you to gather information from different parts of the network and get a job
  - Structurally embedded edges are heavily redundant in terms of information access



From qualitative statement to quantitative statement

- **For many years Granovetter's theory was not tested**
- But, today we have large who-talks-to-whom graphs:
  - Email, Messenger, Cell phones, Facebook
- **Onnela et al. 2007:**
  - Cell-phone network of 20% of EU country's population
  - **Edge weight:** # phone calls

Edge Strength in Real Data

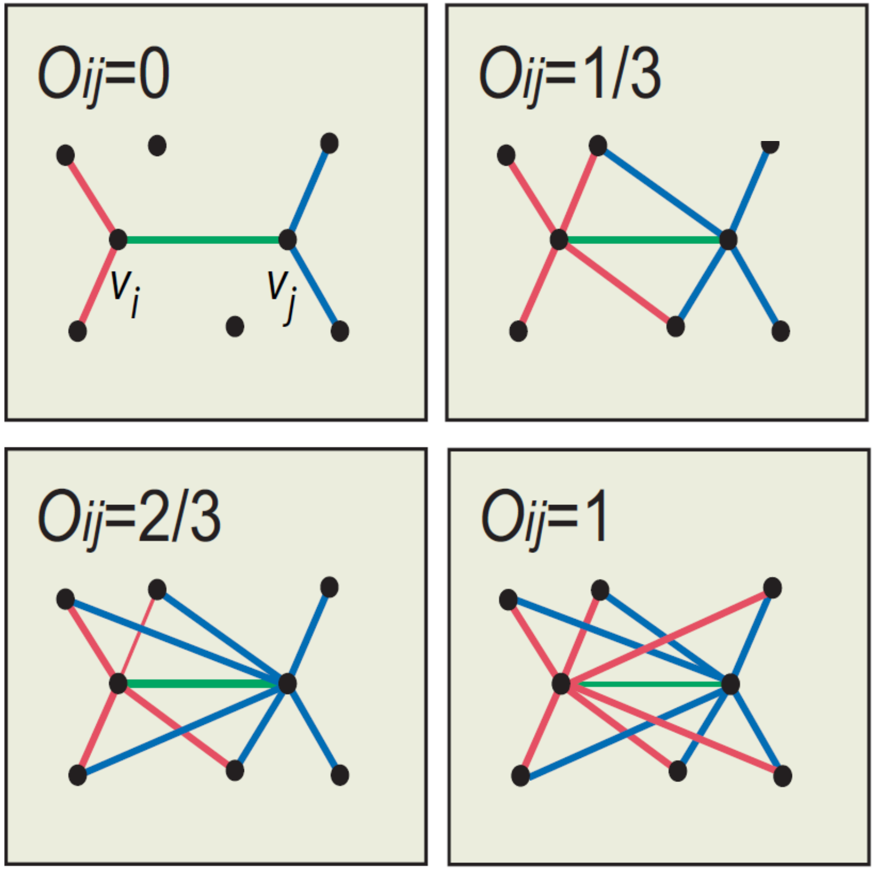
- **For many years Granovetter's theory was not tested**
- But, today we have large who-talks-to-whom graphs:
  - Email, Messenger, Cell phones, Facebook
- **Onnela et al. 2007:**
  - Cell-phone network of 20% of EU country's population
  - **Edge weight:** # phone calls

Edge Overlap

- **Edge overlap:**

$$O_{ij} = \frac{|(N(i) \cap N(j)) \setminus \{i,j\}|}{|(N(i) \cup N(j)) \setminus \{i,j\}|}$$

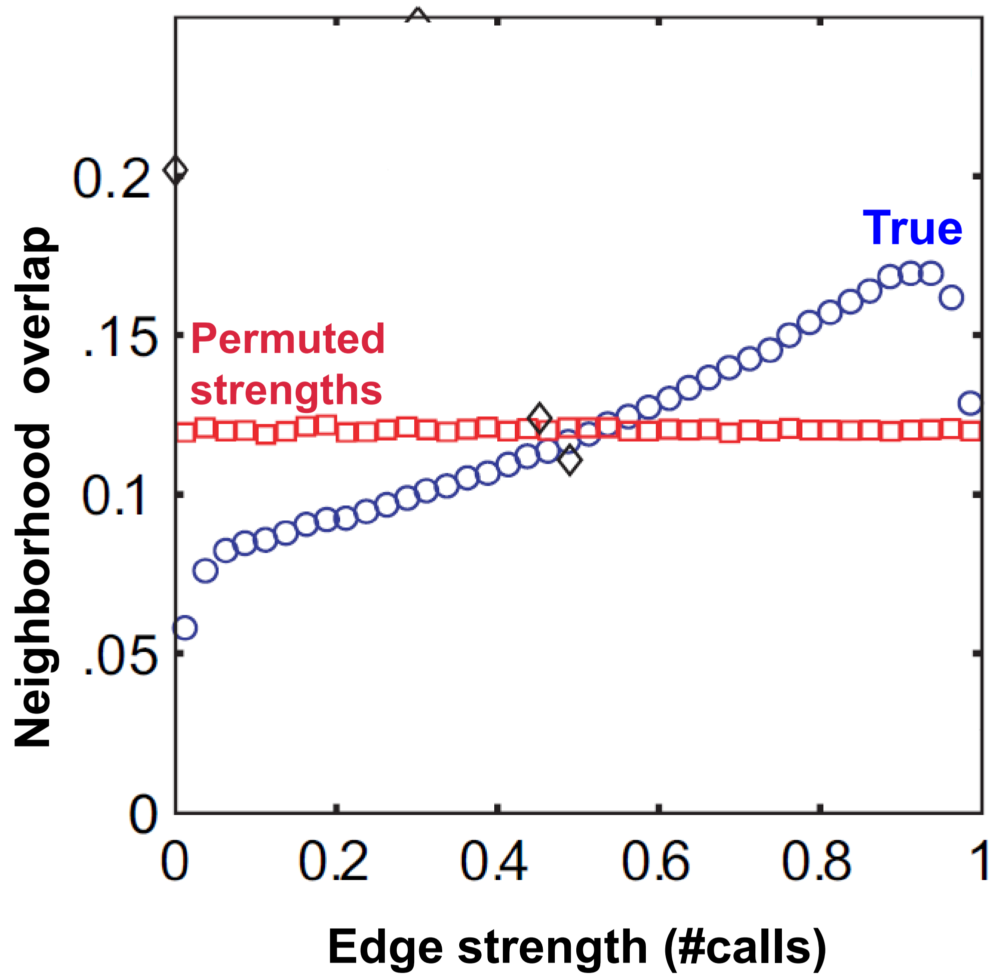
  - $N(i)$ … the set of neighbors of node i



- **Note: Overlap $=$ 0** when an edge is a **local bridge**

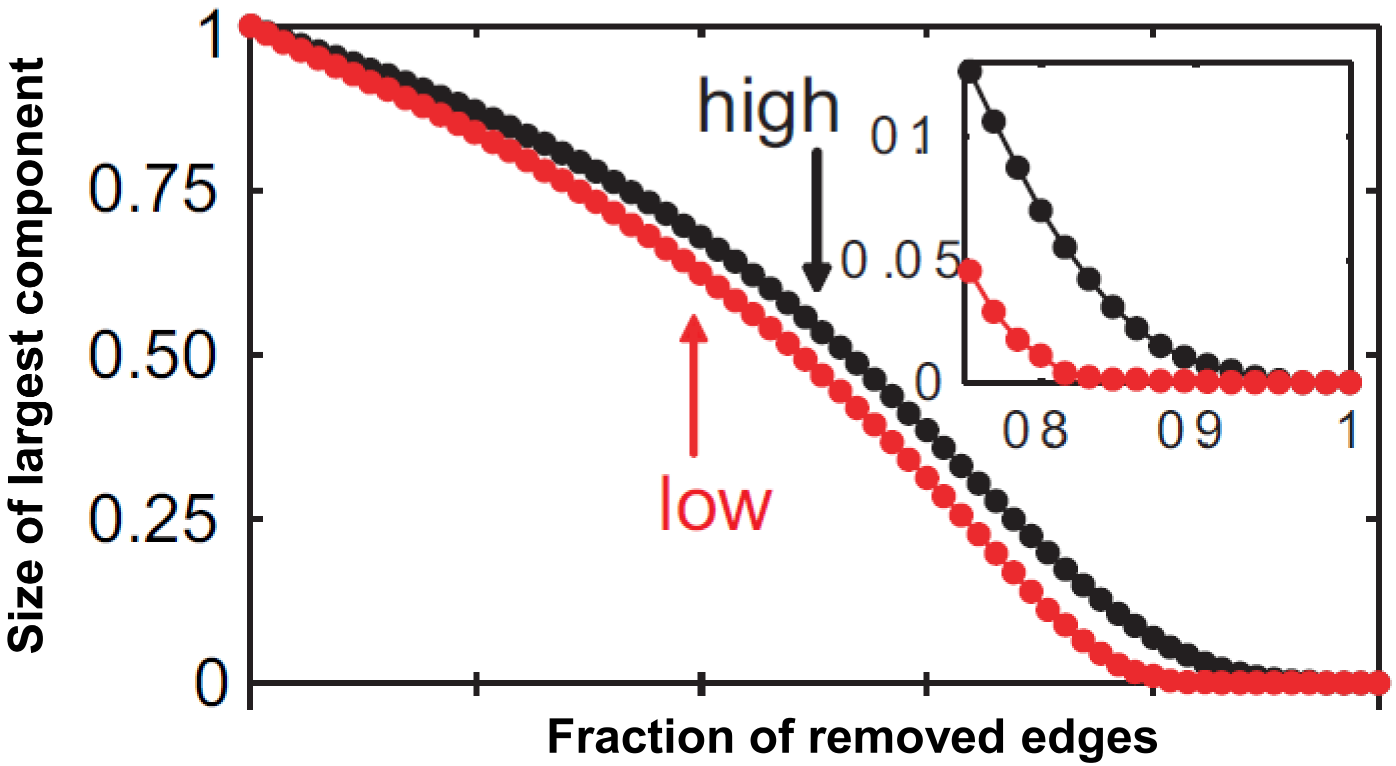
Phones: Edge Overlap vs. Strength

- Cell-phone network
- **Observation:**
  - Highly used links have high overlap!

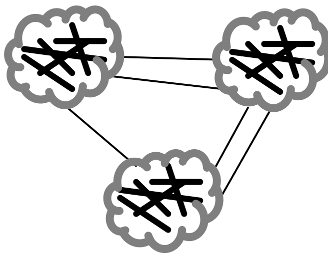


- Legend:
  - **True:** The data
  - **Permuted strengths:** Keep the network structure but randomly reassign edge strengths

Edge Removal by Strength



Low disconnects the network sooner

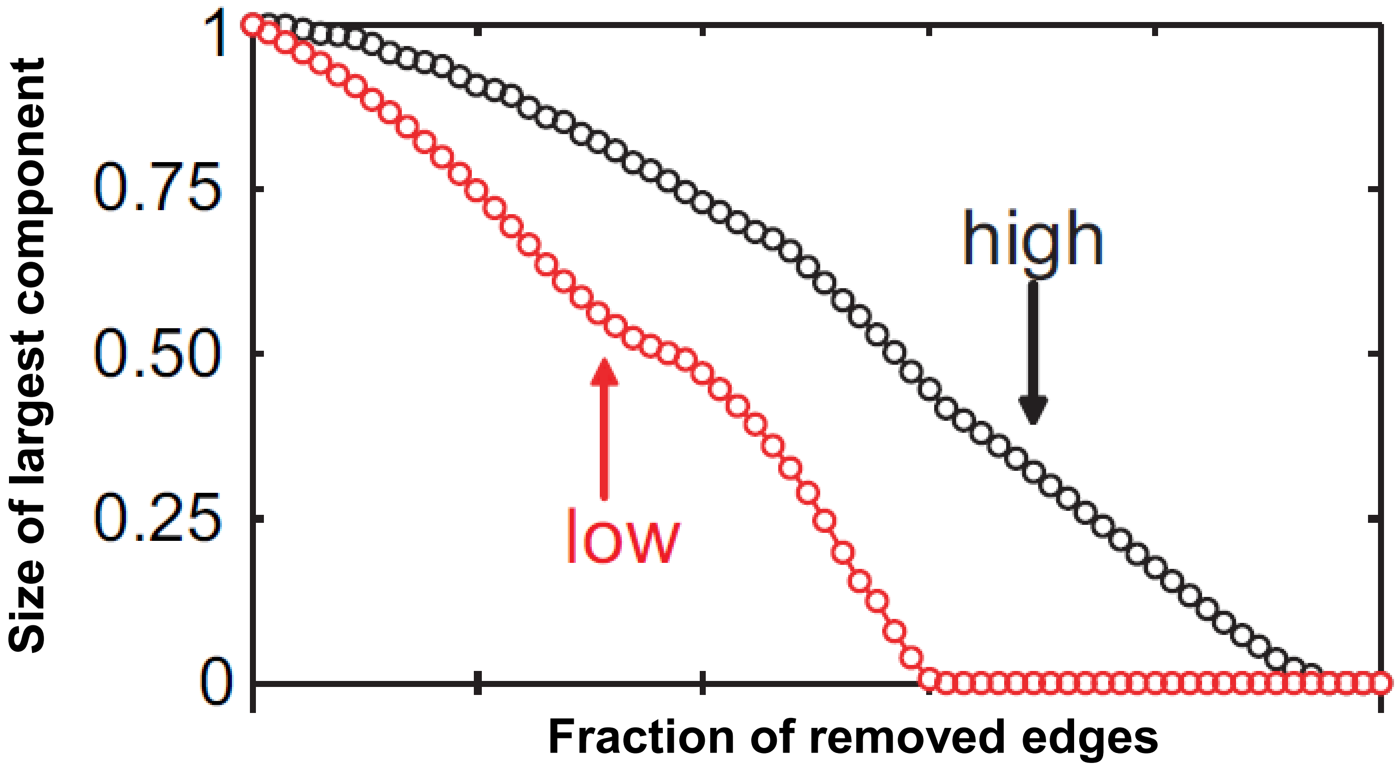


Conceptual picture of network structure

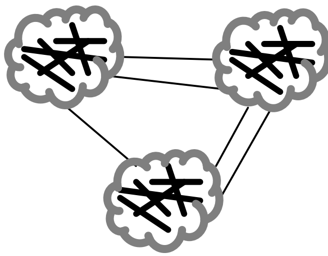
Removing edges based on **strength (#calls)**

- Low to high
- High to low

Link Removal by Overlap



Low disconnects the network sooner



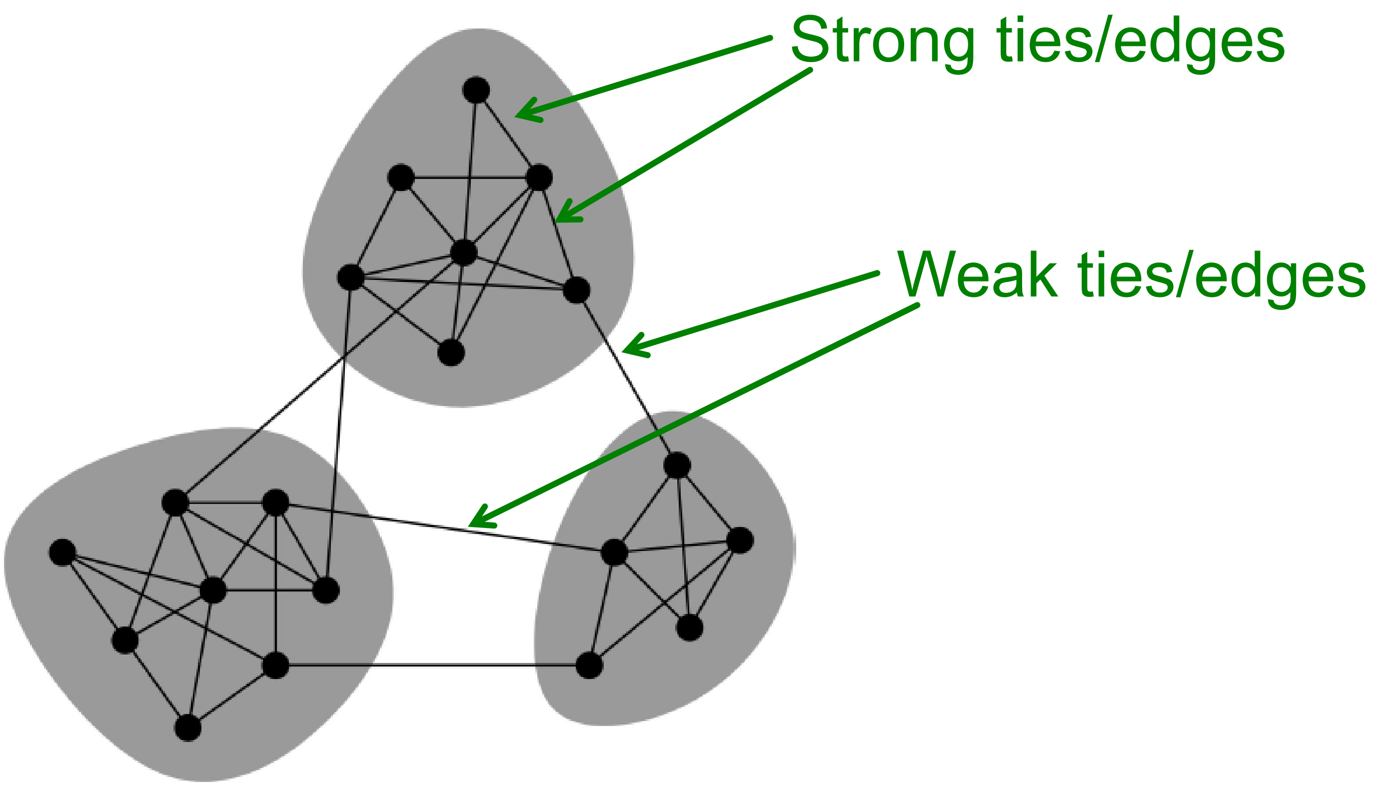
Conceptual picture of network structure

Removing edges based on **edge overlap**

- Low to high
- High to low

Conceptual Picture of Networks

- **Granovetter's theory leads to the following conceptual picture of networks**



Social capital

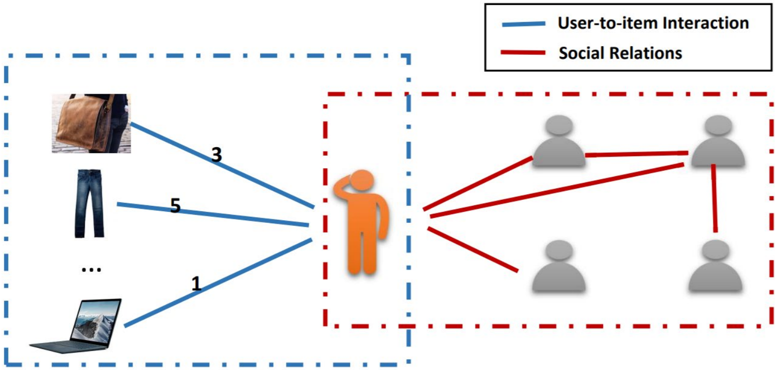
- We studied properties in graph theory, and sociology
- We saw different properties can be linked together, i.e., from one property to another property
- We studied from social science problems → abstract into graph and → study its properties (with new definitions) → validation from big data

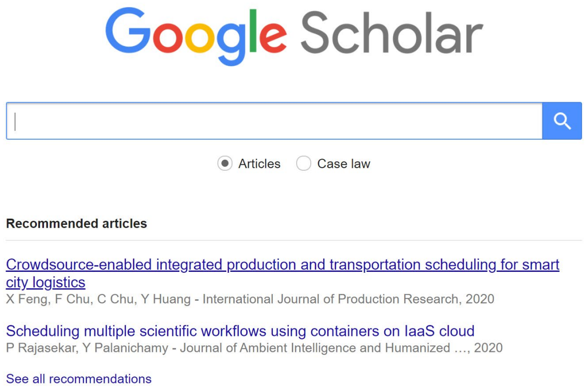
A summary: research process of network analysis

- We studied properties in graph theory, and sociology
- We saw different properties can be linked together, i.e., from one property to another property
- We studied from social science problems → abstract into graph → study its properties (with new definitions) → validation from big data

Weak ties for recommendations

- **Social recommendations:** leverage social network information to help mitigate the "cold start" problem





Homophily

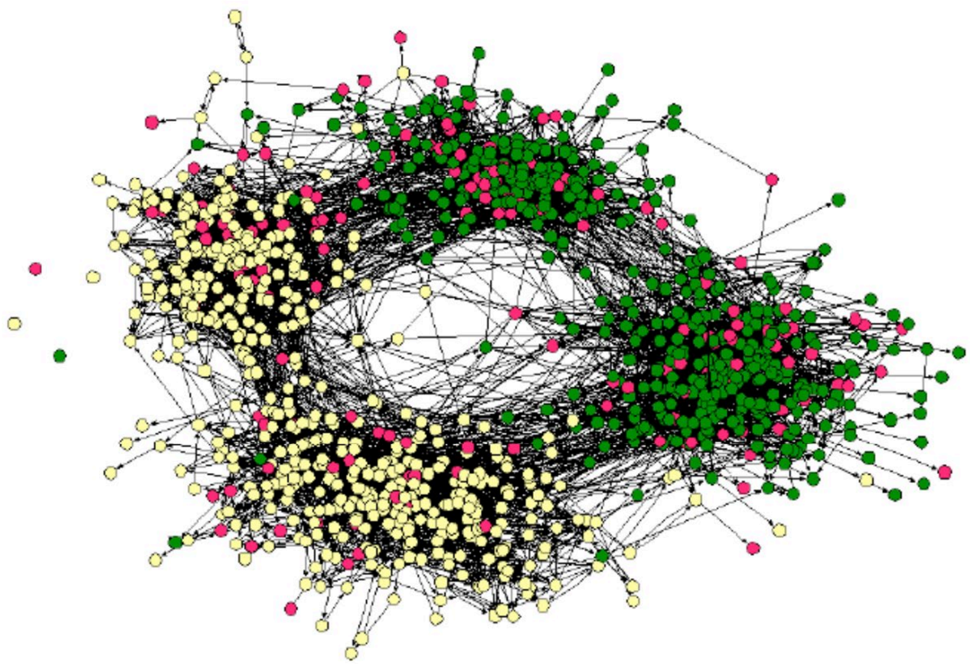
Homophily

- **Homophily**: the tendency of individuals to associate and bond with similar others
  - *"Birds of a feather flock together"*
  - It has been observed in a vast array of network studies, based on a variety of attributes (e.g., age, gender, organizational role, etc.)
  - **Example**: people who like the same music genre are more likely to establish a social connection (meeting at concerts, interacting in music forums, etc.)

Correlations Exists in Networks

Example:

- Real social network
  - Nodes = people
  - Edges = friendship
  - Node color = race
- People are segregated by race due to homophily



(Easley and Kleinberg, 2010)