Graph Theory and Social Networks - part 2

EE599: Social Network Systems

Keith M. Chugg
Fall 2014

USC University of Southern California

Overview

- Continuation of graph theory for social networks
  - Typical social network properties
  - Graph measures to quantify
  - Examples

Last Time: Basic Graph Defs/Props

- Paths, walks, cycles
- Connectedness and components
    - *Giant component*
- Node degree, Node degree statistics
    - *Sparseness & heavy-tailed node degree distribution*
- Adjacency matrix
- Distance and diameter
    - *Small World Phenomena*

References

- Easley & Kleinberg, Ch 3
  - Focus on relationship to social nets with little math
- Barabasi, Ch 2
  - General networks with some math
- Jackson, Ch 2-3
  - Social network focus with more formal math
    - Next time

Motivation

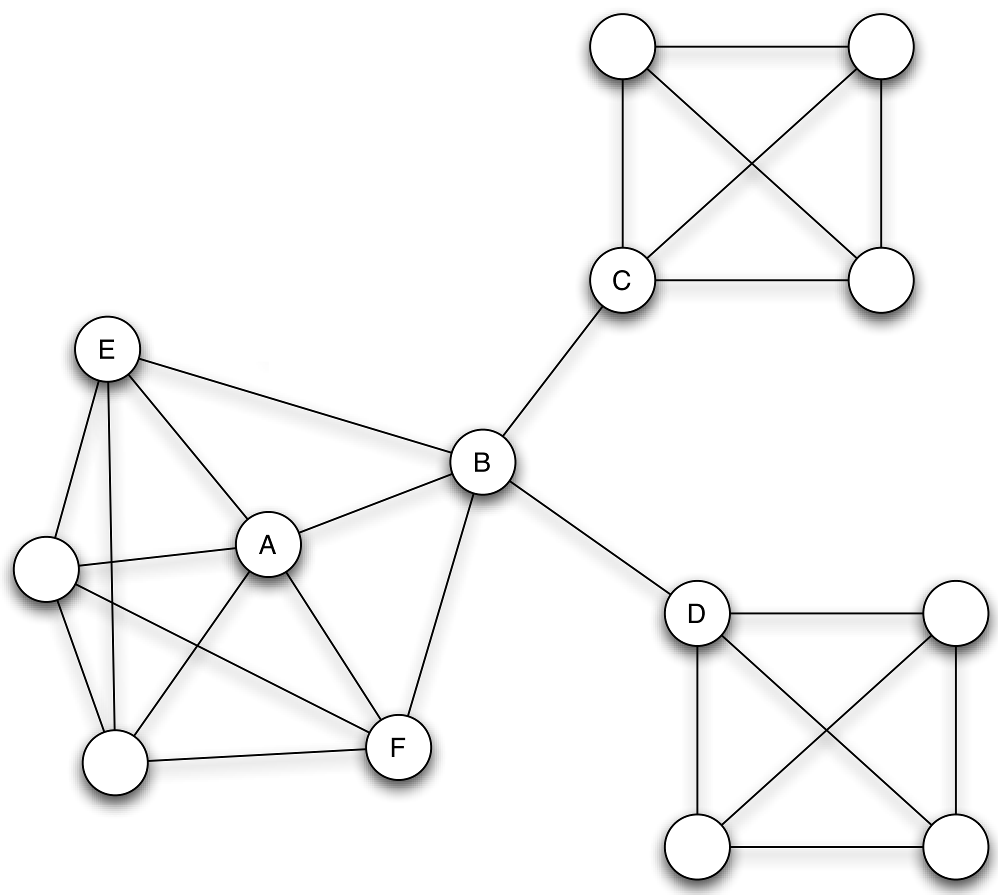


Figure 3.11: The contrast between densely-knit groups and boundary-spanning links is reflected in the different positions of nodes A and B in the underyling social network.
Easley & Kleinberg

- Not all nodes and edges are "equal"
  - A vs. B

Motivation

Our discussion thus far suggests 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 — with a few edges spanning different groups while most are surrounded by dense patterns of connections.

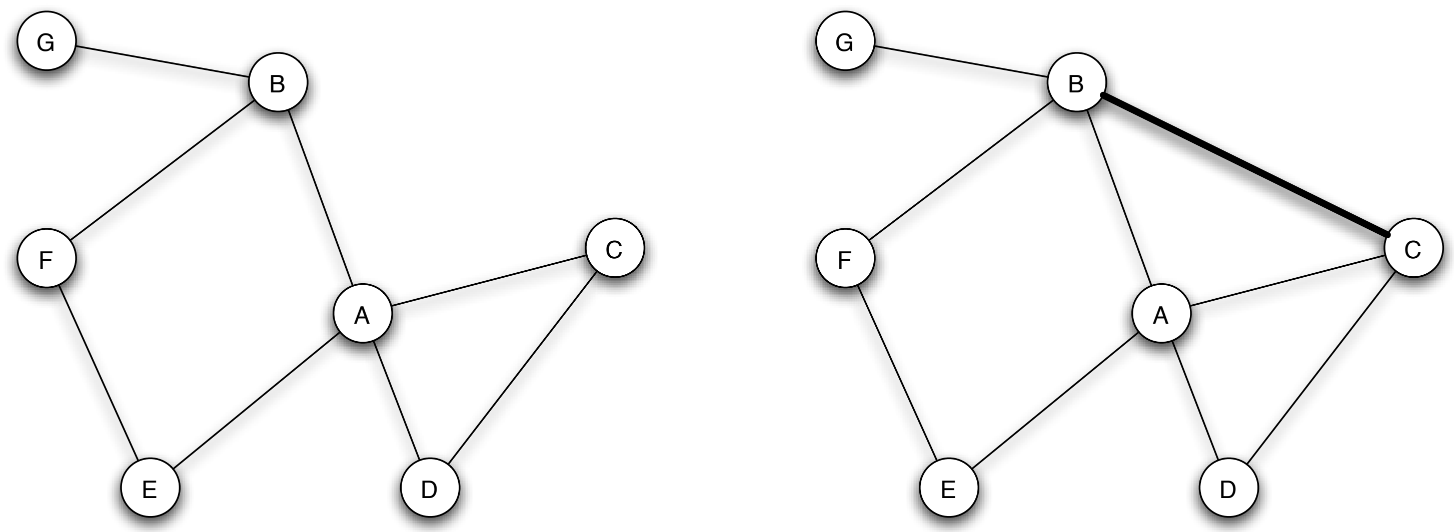
Easley & Kleinberg, pg. 64

- Our goal
  - Motivate this from a few social rules - simple, yet reasonable
  - Define qualitative terms and quantitative measures to capture these properties

Overview

- Triadic Closure & Cluster Coefficients
  - *Our friends usually become friends*
- Strong & Weak Ties
  - *Most people get jobs from acquaintances rather than close friends*
- Centrality and Prestige Measures
  - *Some people (or connections) are more critical than others*

Triadic Closure



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

Figure 3.1: The formation of the edge between B and C illustrates the effects of triadic closure, since they have a common neighbor A.

- Our friends tend to be (or become) friends
  - *Opportunity - likely to meet*
  - *Trust - implicitly trust a friend-of-friend*
  - *Incentive - latent stress if triangle is not closed*

Cluster Coefficient

- Measures the degree of triadic closure in a network
  - *High cluster coefficient = dense local connectivity (many friends are friends)*
- Two variations
  - *Local or individual cluster coefficient*
  - *Global or overall cluster coefficient*

Local Cluster Coefficient

- Cluster coefficient of node i
  - The fraction of node i's neighbors that are neighbors of each other

$$C_i = \frac{\sum_{j,k:(i,j,k)\ \text{distinct}} a_{ij} a_{ik} a_{jk}}{\sum_{j,k:(i,j,k)\ \text{distinct}} a_{ij} a_{ik}} = \frac{L_i}{\binom{k_i}{2}}$$

C_i = 0 if degree is 0 or 1

$$L_i = \text{Number of links between the } k_i \text{ neighbors of node } i$$

$$C_{\text{ave}} = \frac{1}{N} \sum_{i=1}^{N} C_i = \text{Average cluster coefficient}$$

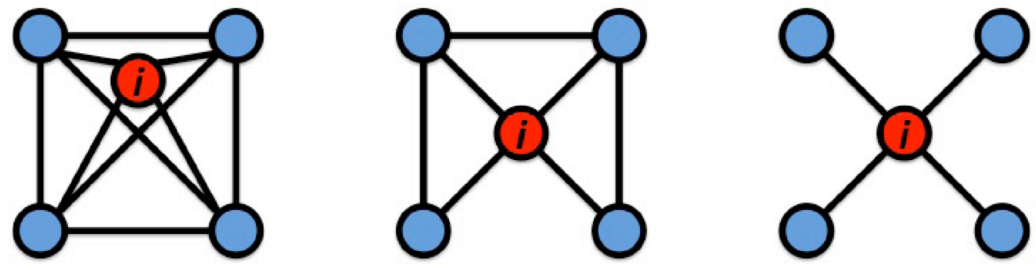
Global Cluster Coefficient

- Property of entire network
  - Average triadic closure over all (possible) triangles

$$C_{\text{global}} = \frac{\sum_{(i,j,k)\text{ distinct}} a_{ij}a_{ik}a_{jk}}{\sum_{(i,j,k)\text{ distinct}} a_{ij}a_{ik}} = \frac{\sum_{i=1}^{N} L_i}{\sum_{i=1}^{N} \begin{pmatrix} k_i \\ 2 \end{pmatrix}}$$

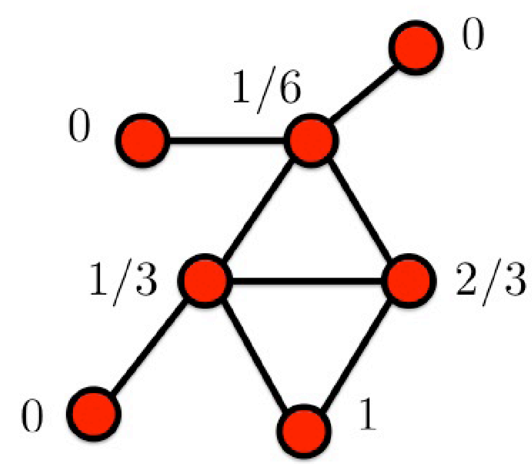
- Average cluster coefficient more heavily weights low degree nodes (relative to global cluster coefficient)

Local Cluster Coefficient



$C_i = 1$ $\quad$ $C_i = 1/2$ $\quad$ $C_i = 0$

$C = 1$ $\quad$ $C = 9/14$ $\quad$ $C = 0$



$$\langle C \rangle = \frac{13}{42} \approx 0.310$$ average

$$C = \frac{3}{8} = 0.375$$ global

Image 2.15
Clustering Coefficient.

The local clustering coefficient, C_i, of the central node with degree k_i=4 for three different configurations of its neighborhood. The clustering coefficient measures the local density of links in a node's vicinity. The bottom figure shows a small network, with the local clustering coefficient of a node shown next to each node. Next to the figure we also list the network's average clustering coefficient <C>, according to Eq. (20), and its global clustering coefficient C, declined in Appendix A, Eq. (21). Note that for nodes with degrees k_i=0,1, the clustering coefficient is taken to be zero.

Baraba'si

Clustering Patterns

- Many social networks exhibit:
  - *Less triadic closure as the node degree increases*
- This is captured by: $C_{\mathrm{global}} < C_{\mathrm{ave}}$

Cluster Coeff.	Biology Collaboration Network	Math Collaboration Network	Physics Collaboration Network
Global	0.09	0.15	0.45
Average	0.60	0.34	0.56

Jackson, 3.2.5

Overview

- Triadic Closure & Cluster Coefficients
    - *Our friends usually become friends*
- Strong & Weak Ties
    - *Most people get jobs from acquaintances rather than close friends*
- Centrality and Prestige Measures
    - *Some people (or connections) are more critical than others*

Getting a Job

- Many studies in the social sciences indicate that people often get new jobs through acquaintances rather than close friends
  - Why?
    - *Triadic closure & weak/strong tie properties*

Strong vs. Weak Ties

- Strong Tie: connection between friends
  - Generally requires some degree of regular interaction and active participation in the relationship
- Weak Tie: connection between acquaintances
  - Infrequent and/or passive interaction
- Note: this is a 2-value weighting of the graph

Bridges

- Bridge: edge that, if removed, break the network into two or more connected components

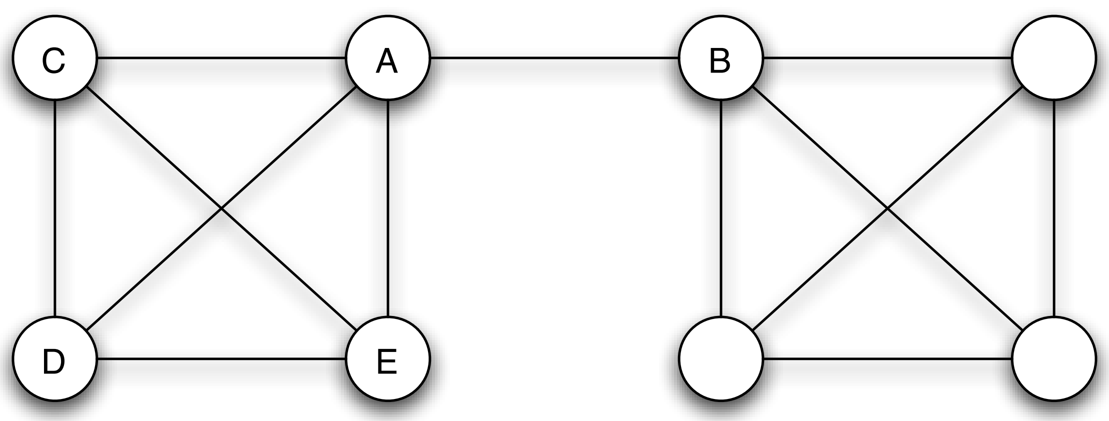


Figure 3.3: The A-B edge is a *bridge*, meaning that its removal would place A and B in distinct connected components. Bridges provide nodes with access to parts of the network that are unreachable by other means.

Easley & Kleinberg

Bridges like this are rare in real social networks, why?

Bridges connect people to new opportunities and information

Local Bridges

- Local Bridge: edge connecting two nodes with no common friends

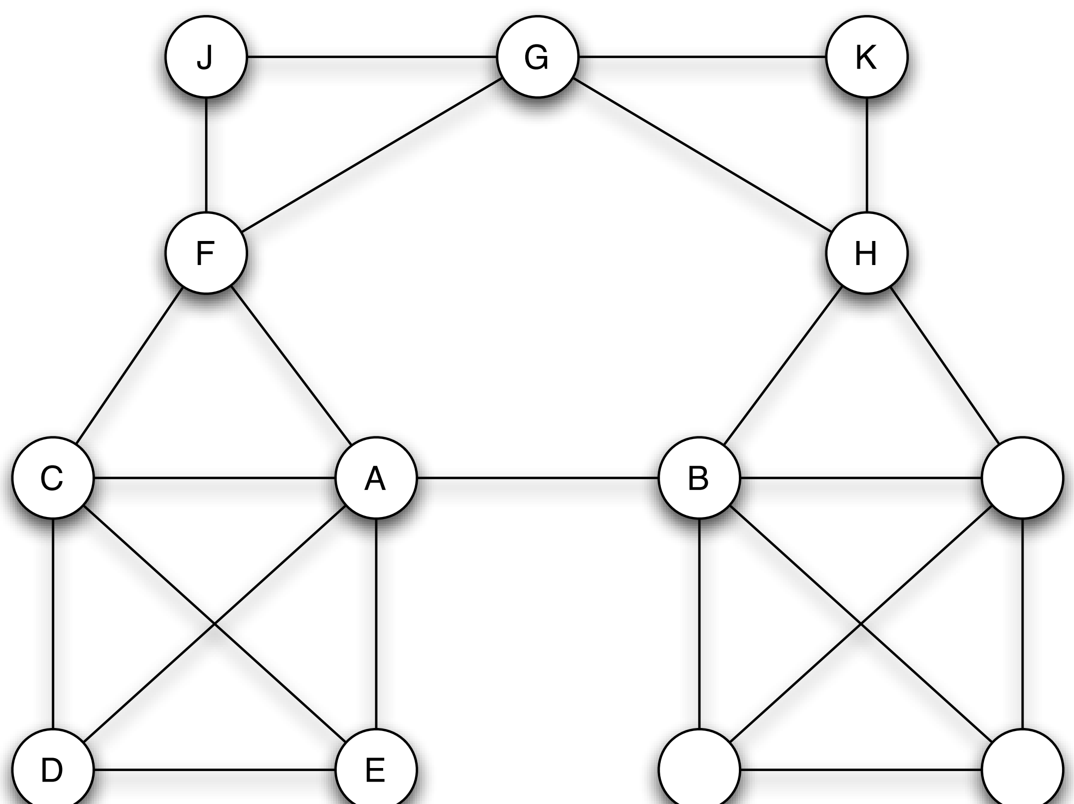


Figure 3.4: The A-B edge is a local bridge of span 4, since the removal of this edge would increase the distance between A and B to 4.

Easley & Kleinberg

- Span of a local bridge is the distance between its connected nodes when the edge is removed
- This is always >2
- Exception of triadic closure: bridges are not in triangles

Local Bridges & Weak Ties

- Local bridges are (typically) weak ties

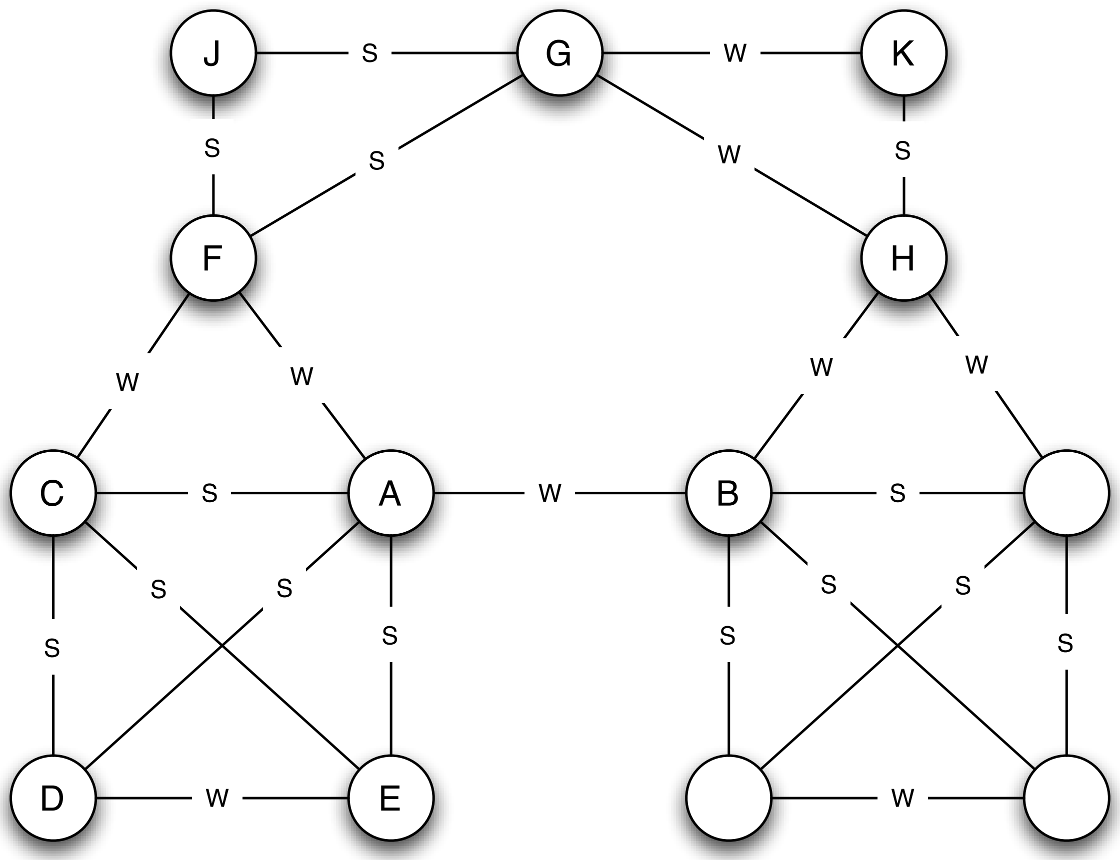


Figure 3.5: Each edge of the social network from Figure 3.4 is labeled here as either a *strong tie* (S) or a *weak tie* (W), to indicate the strength of the relationship. The labeling in the figure satisfies the Strong Triadic Closure Property at each node: if the node has strong ties to two neighbors, then these neighbors must have at least a weak tie between them.

Easley & Kleinberg

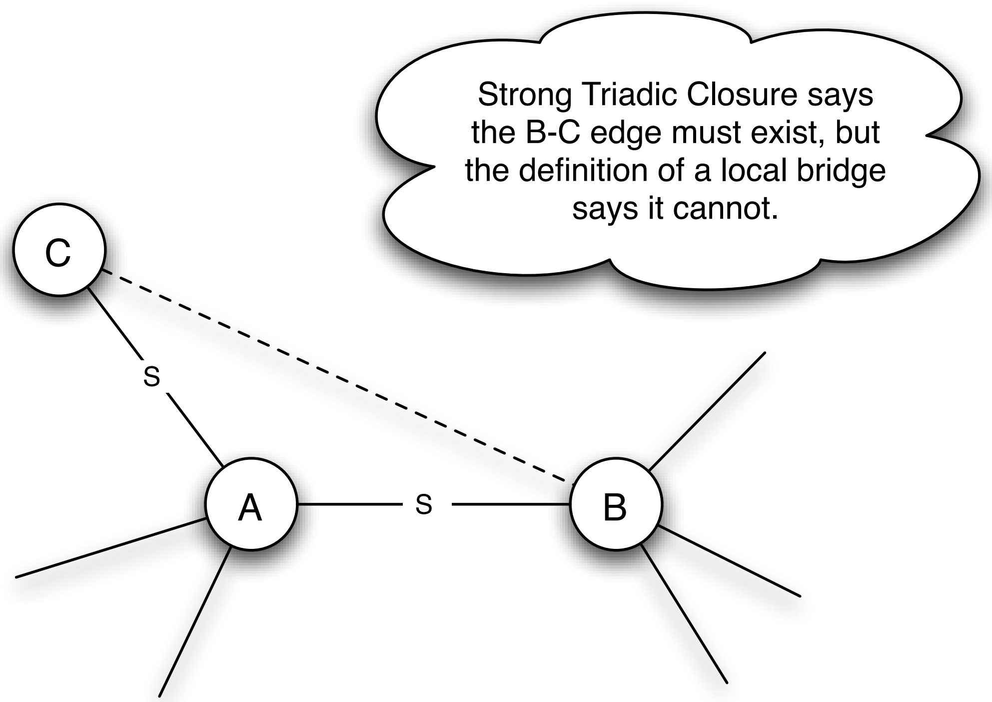
Local Bridges are Weak Ties

- Only need two assume two properties to prove this
  - *Node has 2 or more strong ties*
  - *Strong Triadic Closure Property*
    - *If A has a strong tie with B and a strong tie with C*
      - *then C and B must have a tie*

Easley & Kleinberg

Local Bridges are Weak Ties

Claim: If a node A in a network satifies 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

Figure 3.6: If a node satifies Strong Triadic Closure and is involved in at least two strong ties, then any local bridge it is involved in must be a weak tie. The figure illustrates the reason why: if the A-B edge is a strong tie, then there must also be an edge between B and C, meaning that the A-B edge cannot be a local bridge.

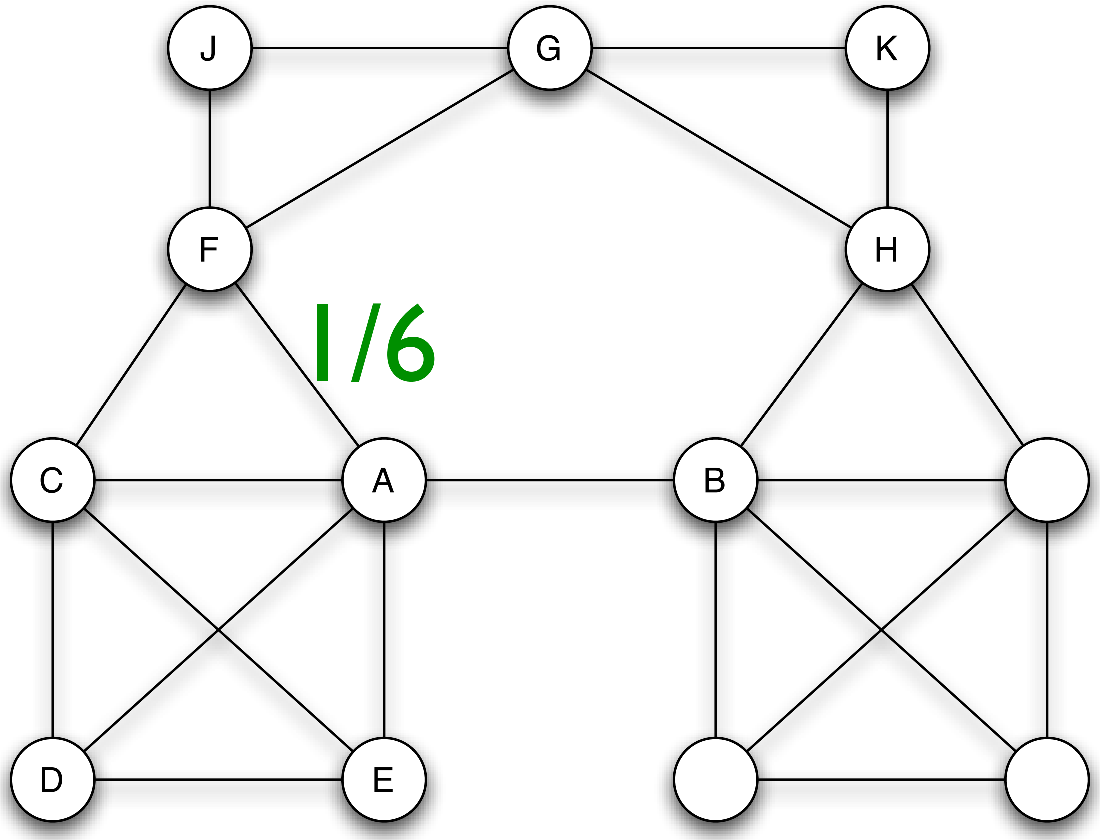
Easley & Kleinberg

“Soft” Measure of Bridges

- Neighborhood overlap of edge connecting A and B
    - *(number of common neighbors)/ (number of total neighbors !=A,B)*

$$\frac{|\mathcal{N}(A) \bigcap \mathcal{N}(B)|}{|[\mathcal{N}(A) \bigcup \mathcal{N}(B)]/\{A, B\}|}$$

- Local bridge has overlap 0, small overlap means ~ almost a bridge



Local Bridges in Real Data

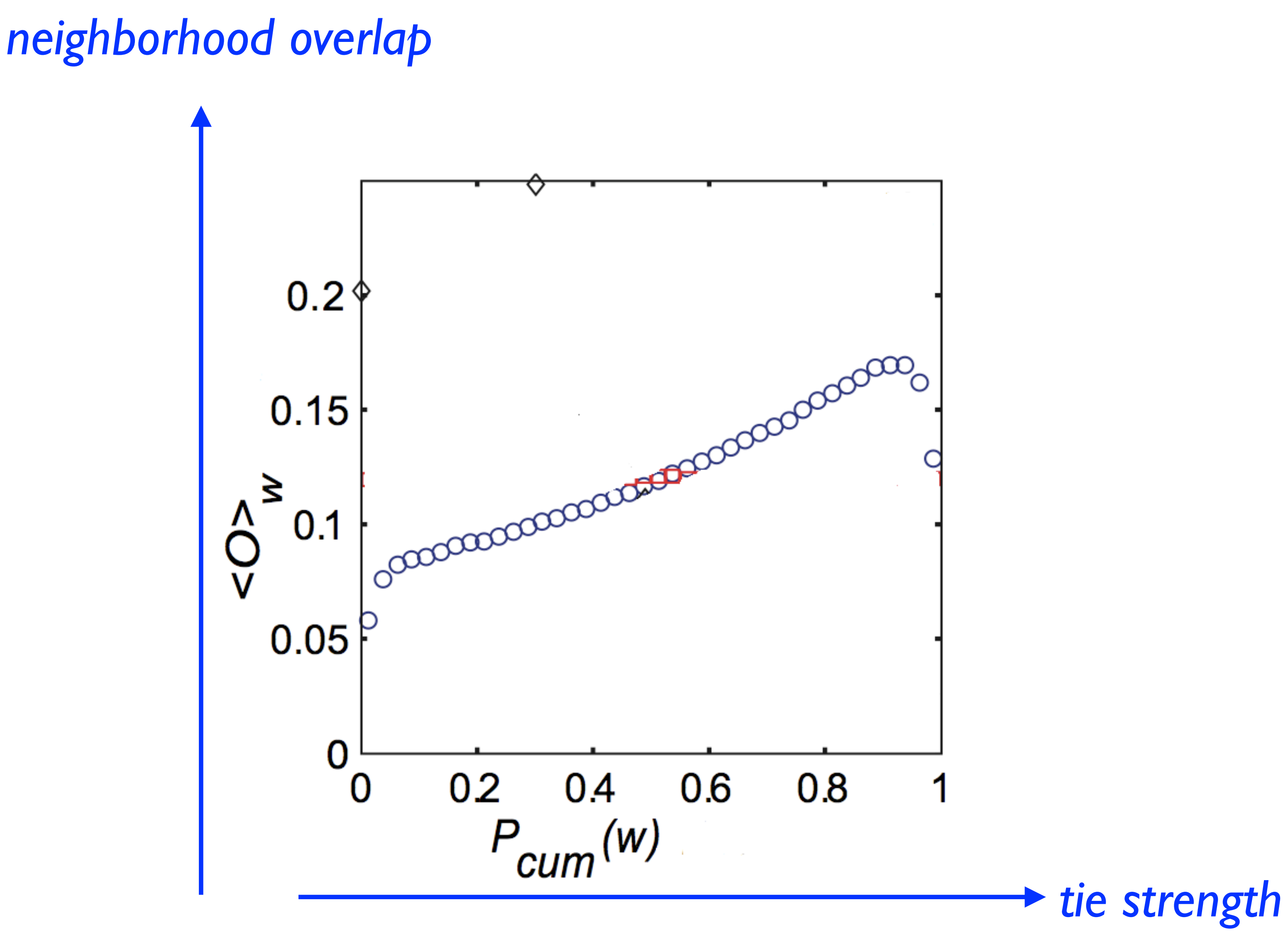


Figure 3.7: A plot of the neighborhood overlap of edges as a function of their percentile in the sorted order of all edges by tie strength. The fact that overlap increases with increasing tie strength is consistent with the theoretical predictions from Section 3.2. (Image from [334].)

Easley & Kleinberg

weaker ties ~ more like bridges

Tie Strength in Real Data

community from past?

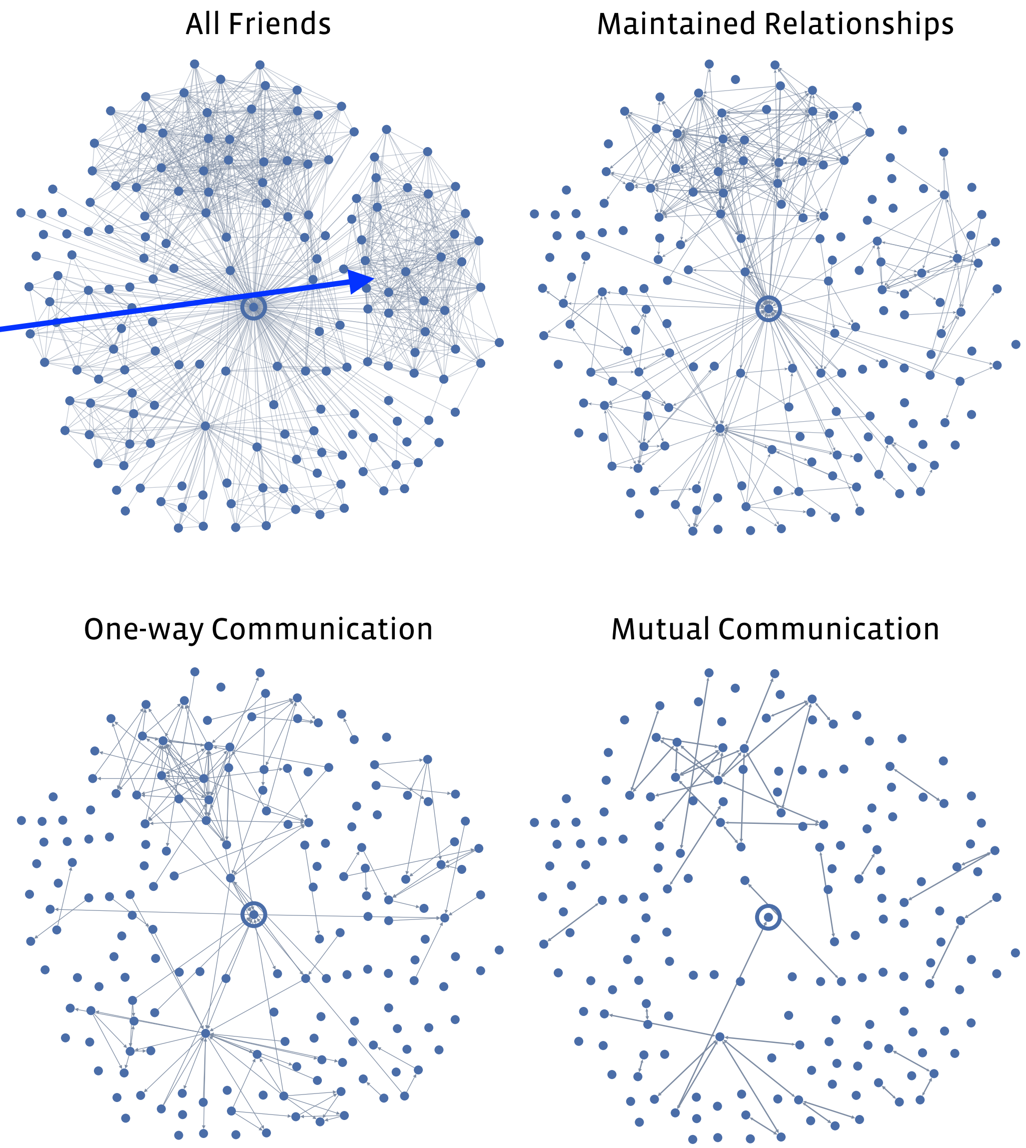


Figure 3.8: Four different views of a Facebook user's network neighborhood, showing the structure of links coresponding respectively to all declared friendships, maintained relationships, one-way communication, and reciprocal (i.e. mutual) communication. (Image from [286].)

Easley & Kleinberg

Limited Strong Ties in Real Data

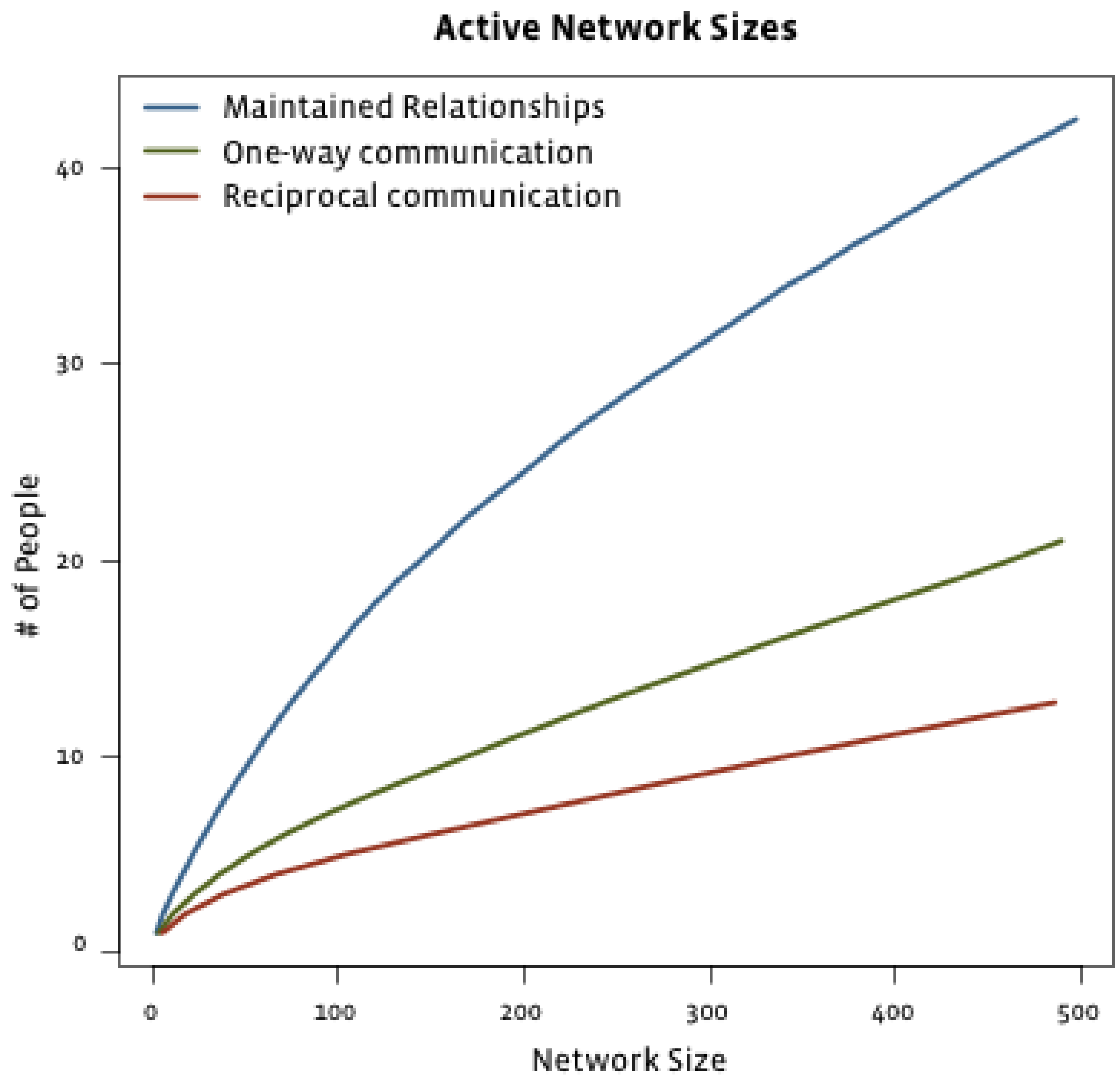


Figure 3.9: The number of links corresponding to maintained relationships, one-way communication, and reciprocal communication as a function of the total neighborhood size for users on Facebook. (Image from [286].)

Easley & Kleinberg

Limited Strong Ties in Real Data

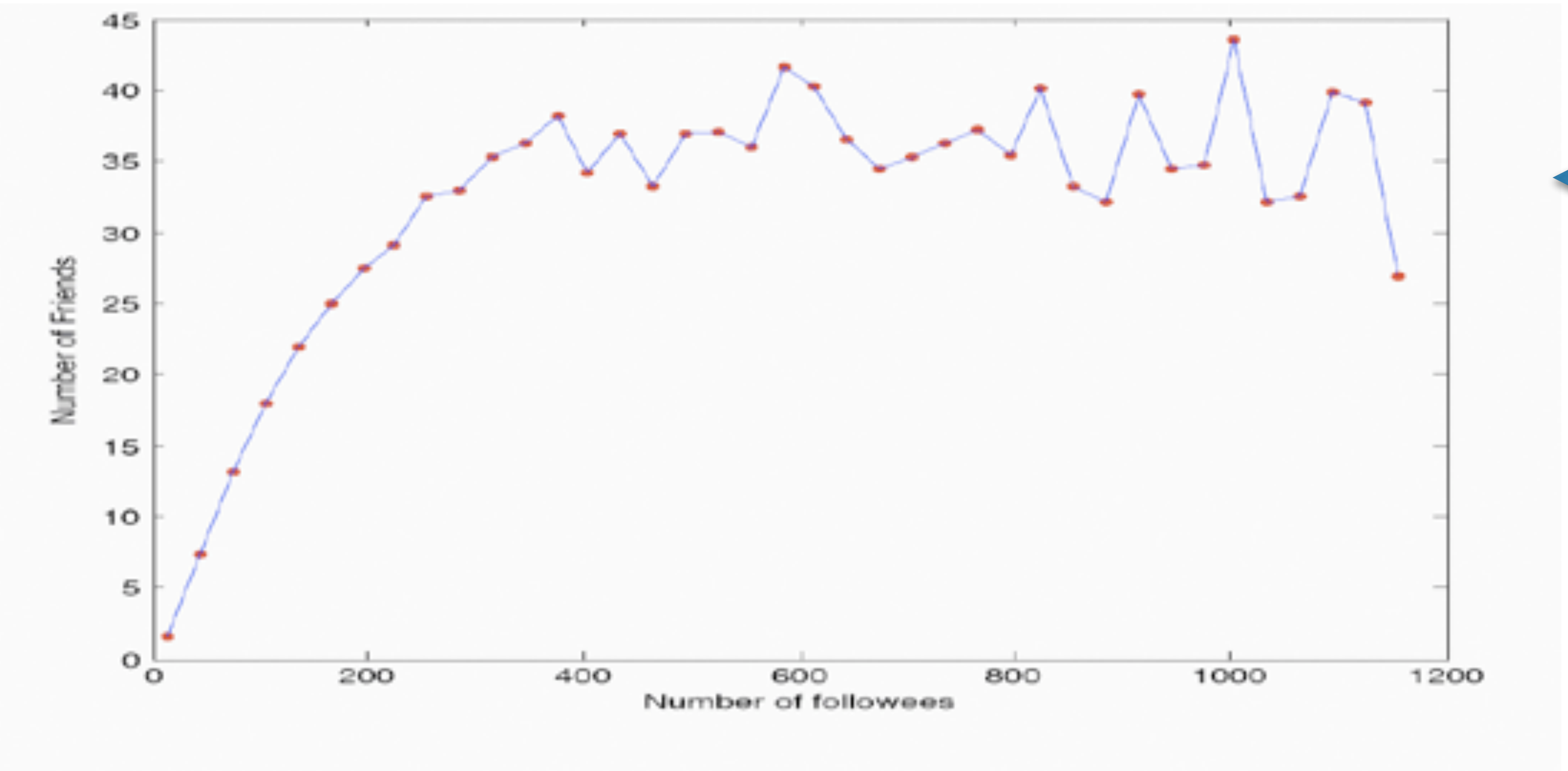


Figure 3.10: The total number of a user's strong ties (defined by multiple directed messages) as a function of the number of followees he or she has on Twitter. (Image from [222].)

Easley & Kleinberg

limited number of strong ties are maintained regardless of total number of ties

Overview

- Triadic Closure & Cluster Coefficients
  - *Our friends usually become friends*
- Strong & Weak Ties
  - *Most people get jobs from acquaintances rather than close friends*
- Centrality and Prestige Measures
  - *Some people (or connections) are more critical than others*
    - Next time