Week 05

Local-level metrics

Tuesday, September 21

INFO 5613: Network Science

Brian Keegan
brian.keegan@colorado.edu

University of Colorado Boulder

Agenda

- The “neighborhood” around nodes encodes important information about network structure

- Clustering
- Triad census
- Degree assortativity
- Friendship paradox

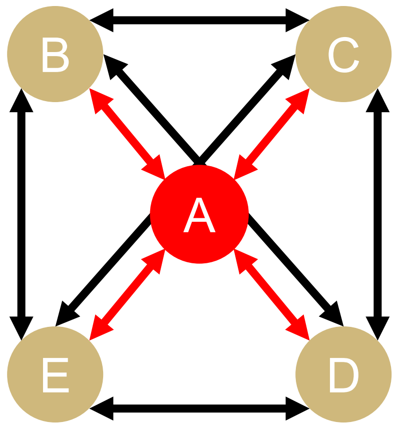
Review

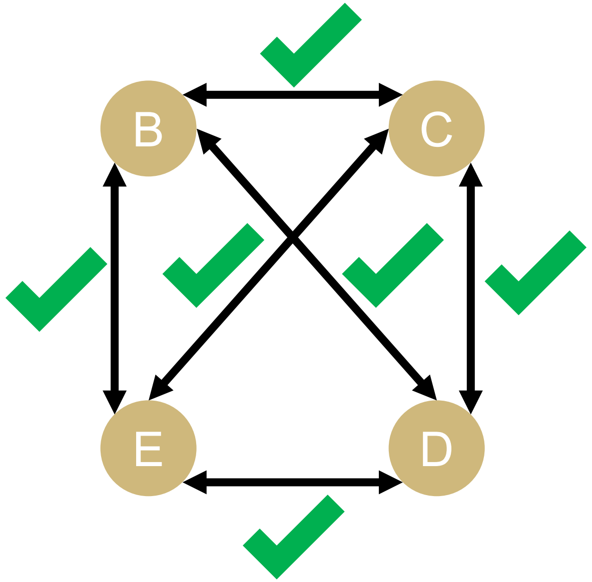
- **Week 2**: Network data structure, data collection, and validity threats
- **Week 3**: Visualization as an aesthetic and rhetorical practice
- **Week 4**: Node-level metrics like betweenness, closeness, degree, & eigenvector centrality

Clustering

Are a node's neighbors also neighbors?

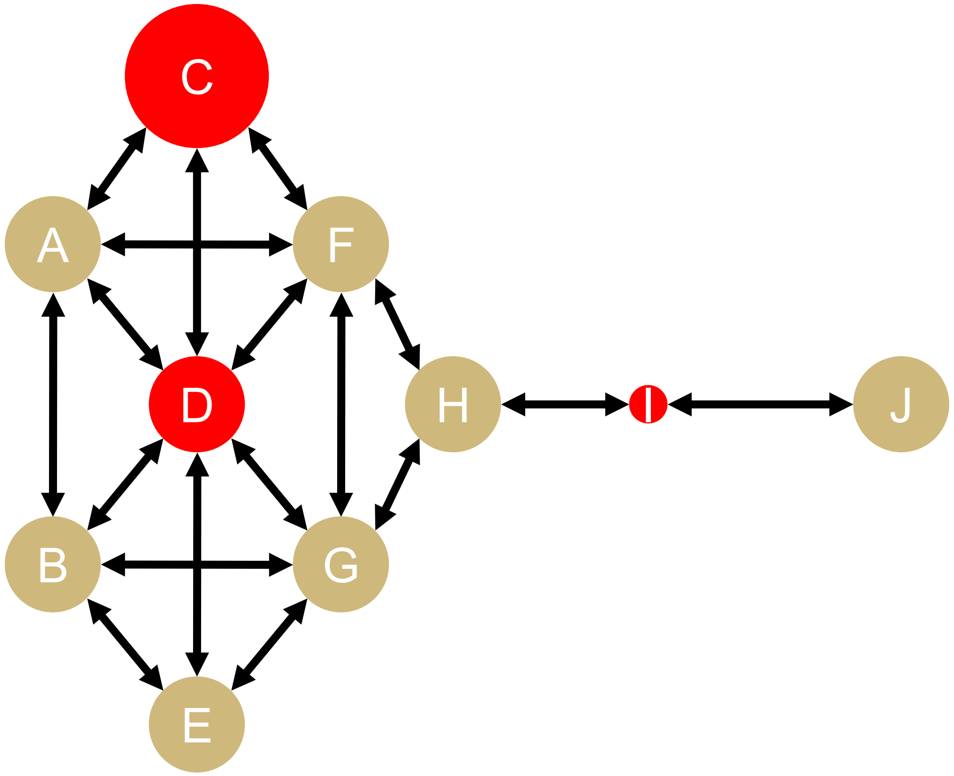
- Node A is connected to 4 nodes: B, C, D, E
- Are B, C, D, and E connected?
  - B,C
  - C,D
  - D,E
  - E,B
  - B,D
  - C,E
- All the connections among A's neighbors that could exist… do exist!





Clustering

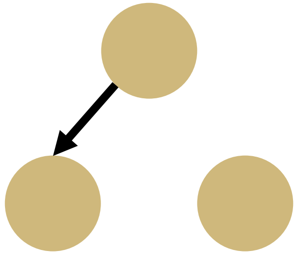
- How many of neighbors are also neighbors?
  - C's neighbors: A, D, F: A-D, D-F, A-F?
    - C's clustering = 3 existing ties / 3 possible ties = 1.0
  - I's neighbors: H, J: HJ
    - I's clustering = 0 existing ties / 1 possible ties = 0.0
  - D's neighbors: A, B, C, E, F, G
    - Possible: AB, AC, AE, AF, AG, BC, BE, BF, BG, CE, CF, CG, EF, EG, FG
    - Observed: AB, AC, AD, AF, BE, BG, CF, EG, FG
    - 9 observed ties / 15 possible ties = 0.6

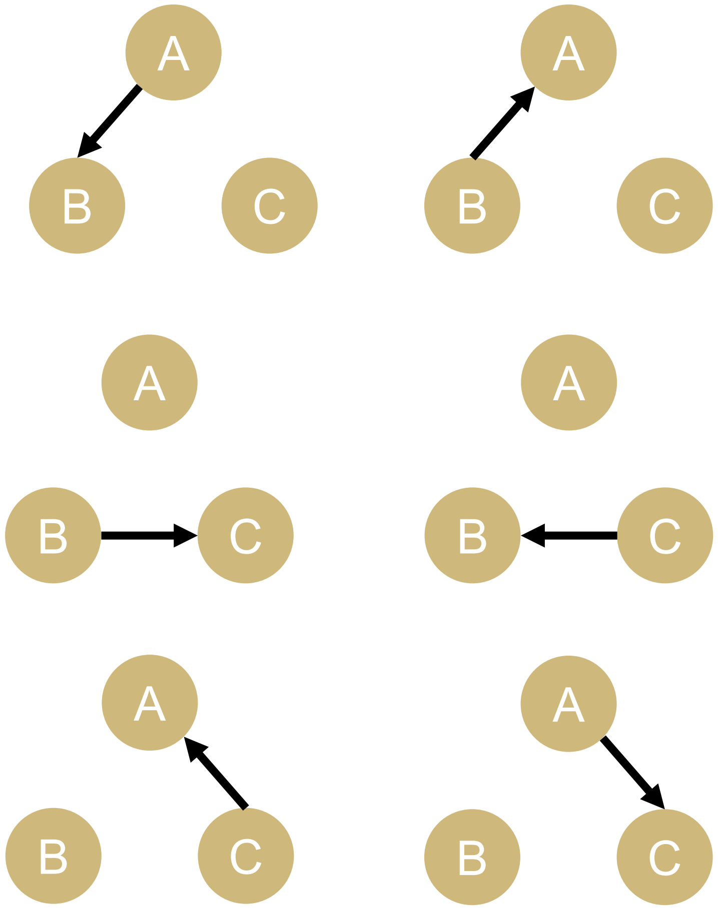


Triad census

Graph isomorphisms

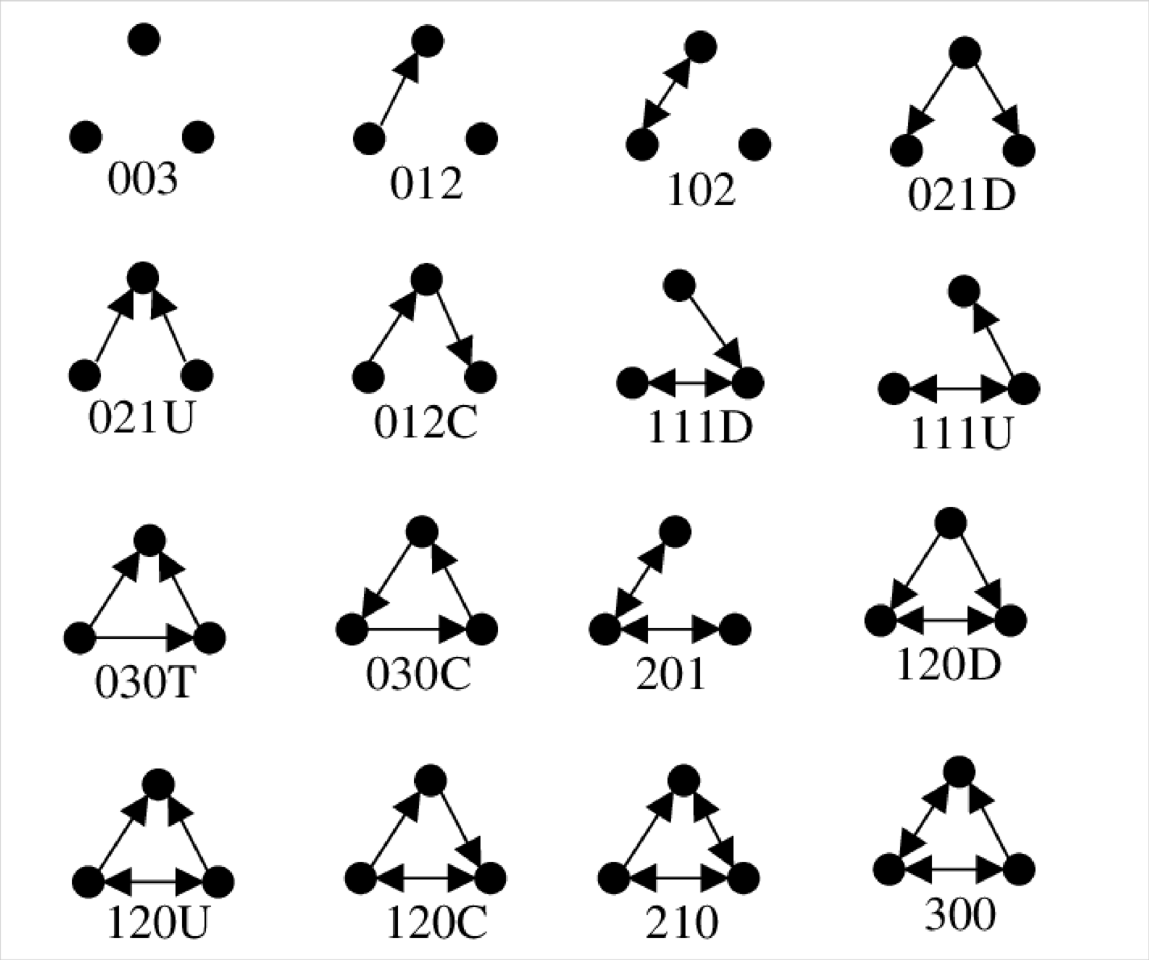
- Triad is a subgroup of three nodes and the links among them
- For a single tie among three nodes, there are six different combinations
- These are all a single “isomorphism”
- Remove the labels and its all the same pattern





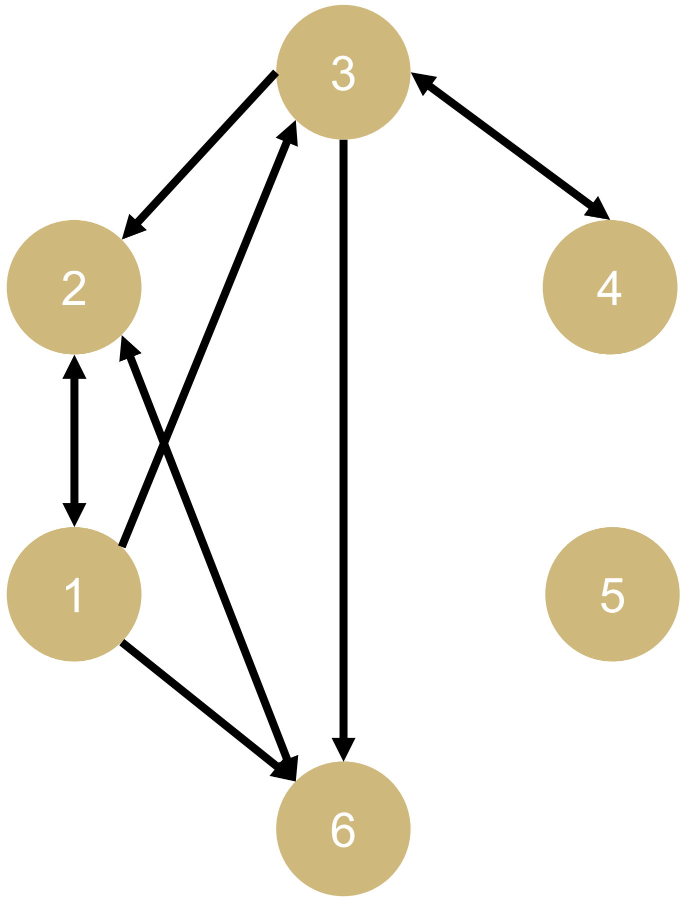
Classes of triads

- In a directed graph, there are 16 possible isomorphism classes
- "M-A-N" Nomenclature:
  - # Mutual, # Asymmetric, # Null
  - U = Up
  - D = Down
  - C = Cyclic
  - T = Transitive
- Each of these could be mapped to theoretical mechanisms around reciprocity, closure, *etc*.
  - More on statistical hypothesis testing in a bit



Triad census

- 003 = {145, 245, 456}
- 012 = {135, 146, 156, 235, 356}
- 102 = {124, 125, 246, 256, 345}
- 111U = {234, 346}
- 111D = {134}
- 030T = {136}
- 120D = {236}
- 120C = {123}
- 210 = {126}



Hypothesis testing

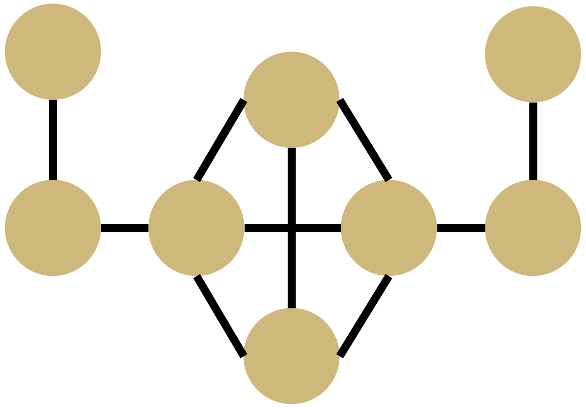
- Calculate the triad census on observed network
- Simulate a bunch of similar random networks
  - Similar on size, density, etc.
  - More on random networks in week 8!
- Calculate triad censuses on each random network
- Compare the observed graph's statistics to the distribution of random networks

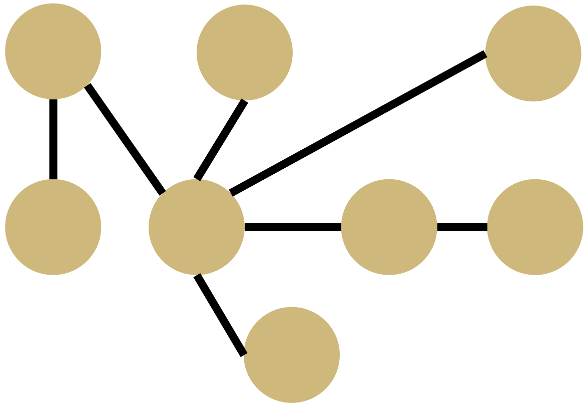
- This will be clearer in the notebook!

Degree assortativity

Neighbor degree varies with node degree?

- **Degree centrality**: a node's importance depends its number of connections
- **Eigenvector centrality**: a node's importance depends on its neighbors' importance
- A folk sociological theory
  - "Well-connected people are connected to other well-connected people"
  - Check back in on "homophily" in week 11!
- Does the degree of a node's neighbors vary with the degree of the node?
  - High/low degree nodes connected to other high/low degree nodes?
  - Or is degree of neighbors independent from degree of nodes?





Friendship paradox

Your friends have more friends than you

- Neighbors of a random node will have a higher average degree than the node itself – on average
  - This isn't a trivial statistical anomaly like a single outlier neighbor brings up the average
- People with many friends are more likely to be your friend – we are introverts orbiting extroverts

- Similar to the class size paradox (Feld & Grofman 1977):
  - Most classes are small, but most students experience large classes
  - Professors experience average class sizes, students do not
- Can be generalized to other contexts:
  - Traffic/crowding paradox: Spaces are mostly empty, but most people experience them crowded
  - Settlement size paradox: Most human settlements are small, but most people live in large cities

Feld, S. L., & Grofman, B. (1977). Variation in class size, the class size paradox, and some consequences for students. *Research in Higher Education*, *6*(3), 215-222.

Next class

Next class

- Readings
  - Feld, S. L. (1991). Why your friends have more friends than you do. *American Journal of Sociology*, 96(6)
  - Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D., and Alon, U. (2002). Network motifs: Simple building blocks of complex networks. *Science*, 298(5594)
  - Newman, M. E. J. and Park, J. (2003). Why social networks are different from other types of networks. *Physical Review E*, 68(3):036122
    - Just skim through equations, focus on prose and “Section V: Examples”
- Prompts
  - Think of some other examples of “friendship paradoxes”, outline what causes the mis-match between subjective and objective perspectives, and their implications for analyzing social behavior
  - Each network motif corresponds with distinctive social behaviors. What are some kinds of networks you might expect to see high frequencies of motifs compared to typical or random networks and why?
  - Newman and Park (2003) argue that social networks differ from other kinds of networks because group structure induces clustering and positive degree correlations. If you don’t find this persuasive, what could be some counter-examples and arguments?