Lecture 4 - Social networks and triadic closure

Monday, January 24, 2022 7:50 AM

Plan for today:

- Topics:
  - This week: evolution of social networks
  - Next week: (general) link prediction
- Homework released tomorrow (probably)
- Review last class
- Shortest paths and small-worldness
- Social networks
- Triadic closure
- Homophily
- Dynamic and temporal networks
- Code mode

Review

- Connectivity: k-connectivity

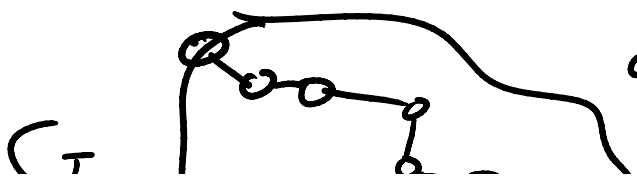
→ *reachability*
can u reach v?

- Network measures

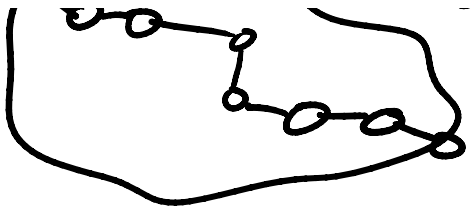
Irregularity: degree distributions, etc.

- Power law: $P(d) \sim d^{-\alpha}$
- $\alpha = 1 + n \sum_{v \in V(G)} \frac{1}{\ln(d(v))}$

Small-worldness property



diameter: longest shortest path

G  shortest path

How "small-world" a graph is:

- What's the average shortest path distance?
  * considering all u,v pairs $\in V(G)$
  * $\frac{\text{sum (all-pairs shortest paths)}}{|\text{all pairs}|}$

Issue: $O(n^2)$ possible pairs

- $|V(G)| > 1$ million
  ↳ computationally gets tough

So: we can do an approximate algo
- sample some subset of u,v-pairs

What about diameter?

- We'd have to solve APSP
  all pairs shortest paths
- Again, not ideal - $O(n^2)$

However, we can approximate

However, we can approximate

- Sampling u, v pairs "not gonna work"
- But, we can use BFS
  breadth-first search

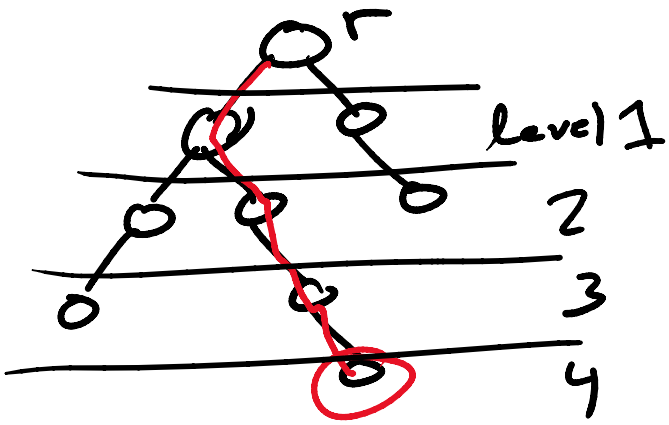
Algorithm:

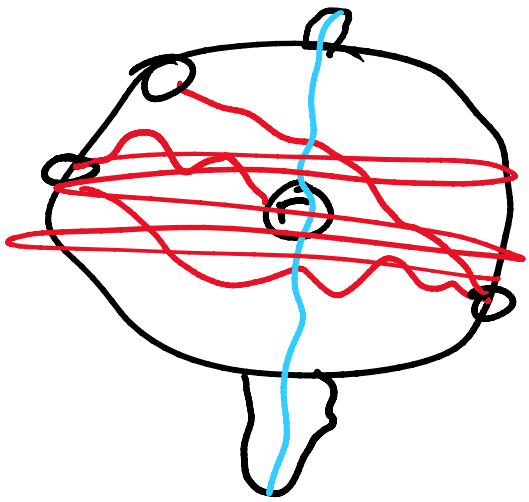
select some root v

For some # iter:

BFS(v)

v = some random vertex farthest from initial root





Why?

Why do these emperical properties occur or arise in social networks?

Social Networks

Really: human interaction networks

vertices: humans

edges: friendships, communication

Properties above exist tremendously in social networks

Triadic closure

Story time (1)

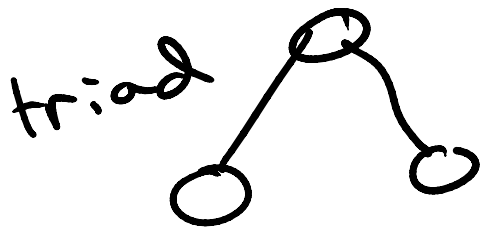
For sharing photos, I get options:

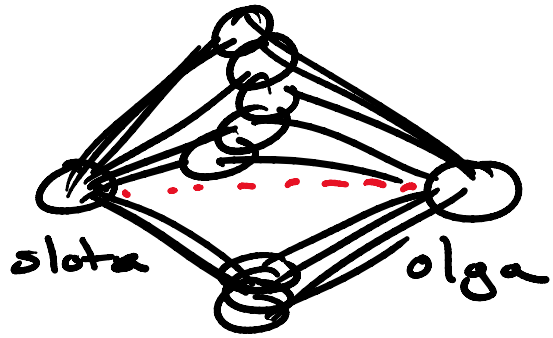
#1 current girlfriend

#2 my friend's wife

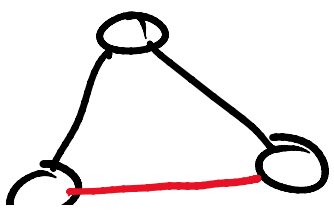
#3 ex-girlfriend

#4 mom





↳ closure

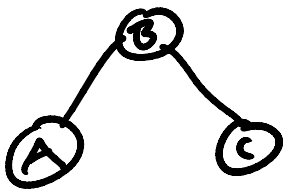


↳ closure 

Why do triads tend to close?

Note: triadic closure is a driving force behind social network evolution

Consider: A and B are friends
B and C are friends
A and C are not friend



→ there is a higher-than-random chance A, C will become friends

Why is this observed:

Opportunity: B hangs out with A, C separately, but may get A, C to hang out together

Trust: B trusts A and C
A and C trust B

A and C trust B

so A and C are likely to trust each other

Incentive: B might want to close the triad

Over time, as network evolves:

Triads close at a higher rate than non-triads

- proportional to # common neighbors

→ clustering coefficient will increase over time

$$\text{Clustering coefficient} = \frac{\text{closed triads}}{\text{all possible triads}}$$

Dynamic and temporal networks

dynamic: changes over time

temporal: we have time data for our vertices/edges

temporal —

for our vertices / edges

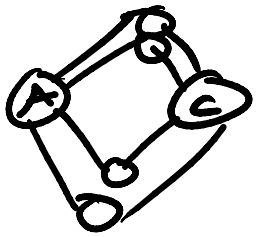
Experiment: (aka gettin' minin')

- We expect triads to close over time
- We measure at some to some number of open triads

→ what fraction closed after same Δt

→ is there a relation to "strength" of triad

↳ how many shared neighbors



Similar concept: Homophily

Homophily: "birds of a feather flock together"

"like attracts like"

Or: similar people tend to be friends

- Selection: we seek out similar people
- Influence: we become similar to people we spend time with

people we spend time with

Expanding triadic closure:

consider affiliation networks