Social Network Growth and Triadic Closure

Sunday, January 21, 2024 5:45 PM

A few people haven't selected papers yet
Submitty gradeable for presentation submissions
HW1 released this week, discuss on Thursday
Plan:

- Quick Review
- Small-world graphs
- Shortests paths and diameter
- Social networks and triadic closure
- Homophily
- Dynamic and Temporal graphs

Review

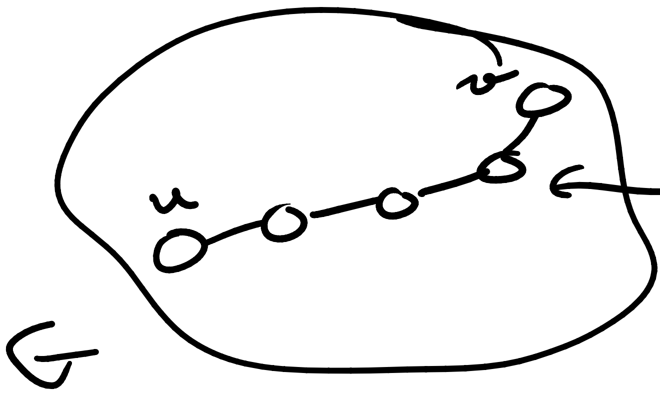
- connectivity: k-connectivity
  can vertex u reach vertex v?

- Network measures
  Irregularity: degree distributions
  cluster sizes, connected component sizes
  Pseudo-power-law: $P(k) \sim k^{-\gamma}$
  why?

  small worldedness

small worldedness

→ low avg. shortest path lengths



shortest u,v-path is fewest number of edges to traverse to get from u to v

How "small world" a graph:

Obvious measure: what's the average shortest path length

* over all possible u,v-pairs

$$\frac{\text{sum(all pairs shortest paths)}}{|\text{all pairs}|}$$

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

if $|V(G)| = n > 1$ million

→ gets "tough" computationally

So: we use approximation algos

SO: we use approximation algos
e.g.: sample a subset of
u, v pairs

Why do we care: small worldedness
of a graph affect things
like information diffusion,

→ idea on a social network
disease spread on
a contact network

Also: diameter
↳ largest possible shortest
path on same graph

How to solve: all pairs shortest
paths
(very slow)
However: we can approximate
using BFS

using BFS

* sampling u, v won't work here, since we'd need to select the u, v with longest shortest path

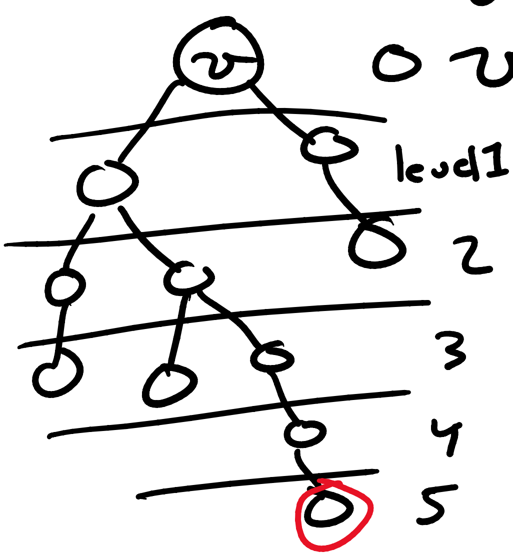
Algorithm:

select some root v

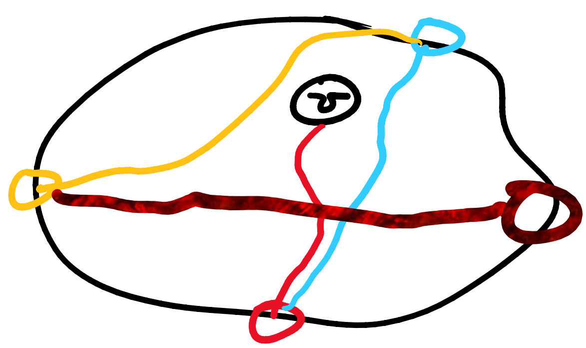
For some # iter:

BFS(v)

- v = some random vertex farthest from our root



keep going until our estimate stops increasing



Why?

why do the emperical properties

Why do the emperical properties we measure arise in real-world networks?

Our focus: social networks

skewed, irregular, small-world, small diameter

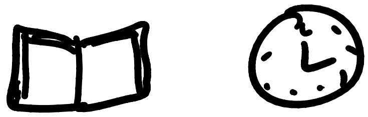
Social Networks:

Human interaction networks

- vertices: humans (or animals)
- edges: friendships, interactions, communication

Triadic closure

Story time w/ Slota



For sharing photos on my phone, Google gave me 4 default options:

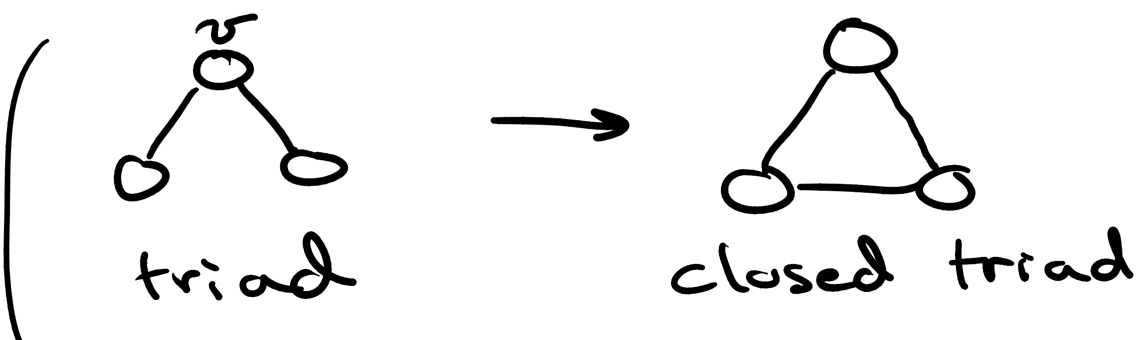
Google gave me 4 default options:

1. my wife → communicate regularly
2. my friend's wife → I don't communicate with at all
3. my ex-girlfriend → haven't communicated with in years
4. my mom → same as 1

Speculation:

why are options 2, 3 shown?

Triadic closure



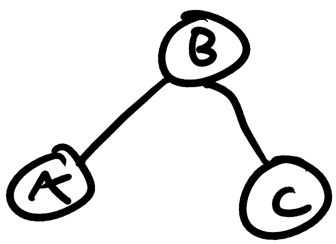
→ triads close at a frequency higher than randomly expected

Why do triads close?

Why do triads close:

Note: triadic closure is a main driver of social network growth over time

Consider: A and B are friends



B and C are friends

A and C are not friends

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

Why:

Opportunity: B spends time with A and C separately, but likely could spend time with A, C together

Trust: B trusts A and C

A and C trust B

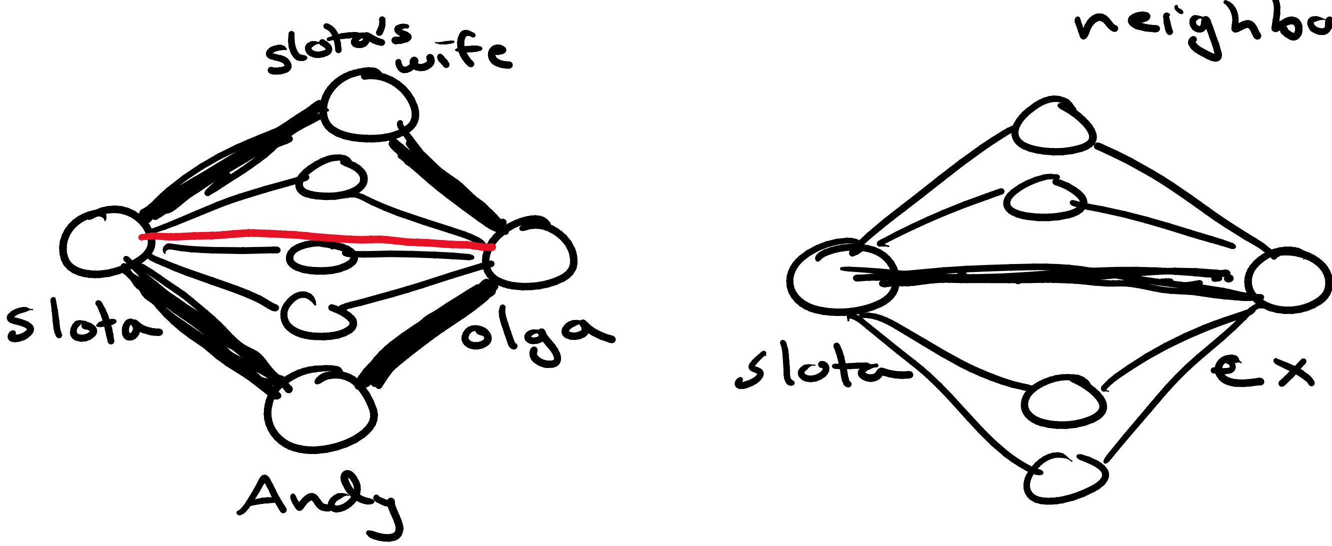
over time: A and C might trust each other trust each other

Incentive: B might want to close the triad

Result: triads close at a higher than random rate

AND

they close faster proportionally to the number of common neighbors



Consequences of triad closure:

- Large clustering
- small world

- small world
- small diameter

→ clustering coefficient will increase over time

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

Similar concept: Homophily

Homophily: "birds of a feather flock together"
"like attracts like"

OR: similar people become friends

- Selection: we seek out similar people as friends
- Influence: we become more similar to our friends over time

How this drives social

How this drives social network growth:

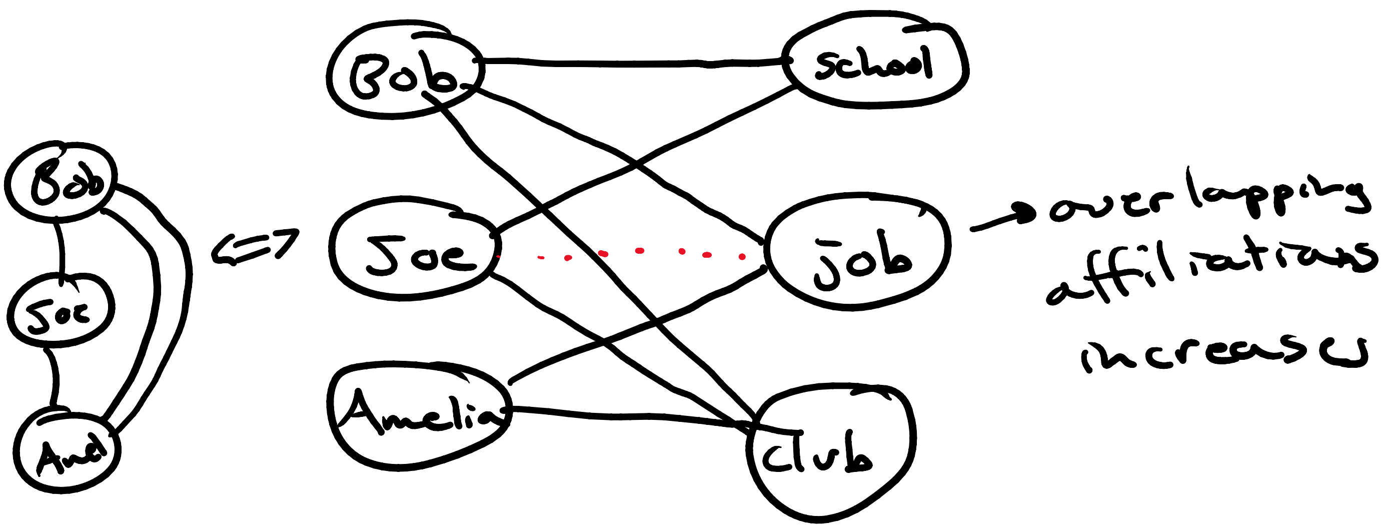
* results in dense clusters with people who are "similar"

Why might online social networks care about this:

Easy to infer (unknown) properties of individuals within a topological cluster

Expanding triadic closure:

Affiliation networks



Same basic idea, if not "triads"

same basic idea, if not ...

→ mutual affiliations will grow over time, proportional to existing mutual affiliations

Takeaway: we also observe an increase in strength of connections over

---

Dynamic and temporal networks

dynamic: changes over time

temporal: we have timestamps associated with creation of vertices/edges

Experiment: (aka gettin' minin')

- we expect triads to close over time
- First, we measure at some t_0 the number of open triads

→ we then after some Δt

... closed

→ we then after some [illegible]

we measure how many have closed

→ then we look at this measure in terms of the number of mutual neighbors

i.e., strength of triad

Q: is there the emperical evidence of triadic closure?