Lecture 4 - Social Network Growth and Triadic Closure

Thursday, January 29, 2026 8:32 AM

Plan for today:

- Ali Brooks presents: Emergence of Scaling in Random Networks
- Quick review (connectivity, network measures)
- Small-worldedness
- Shortest paths and diameter
- Social networks and triadic closure, homophily
- Dynamic and temporal graphs
- CODE MODE

Connectivity of a graph is a function of inherent network properties

Network measures:

Irregularity: degree distribution, cluster sizes, connectivity

Power-law: $P(k) \sim k^{-\gamma}$

Small-worldedness: low average shortest path lengths

→ How "small-world" is a graph?

→ we can explicitly measure

$$\frac{\text{sum(all pairs shortest path)}}{|\#\text{ of pairs}|}$$

Issue: $O(n^2)$ possible pairs
if $|V(G)| = n > 1$ million
↳ gets "tough" computationally

So → <u>approximation algorithms</u>

For average shortest paths:

$$\frac{\text{sum (some pairs shortests paths)}}{|\text{\# of pairs}|}$$

Main concern with such an approach:
→ sampling in a "smart" way to not bias our results

Another related problem: diameter

we expect small-world graphs to also have a small diameter

diameter = max(all pairs shortest paths)

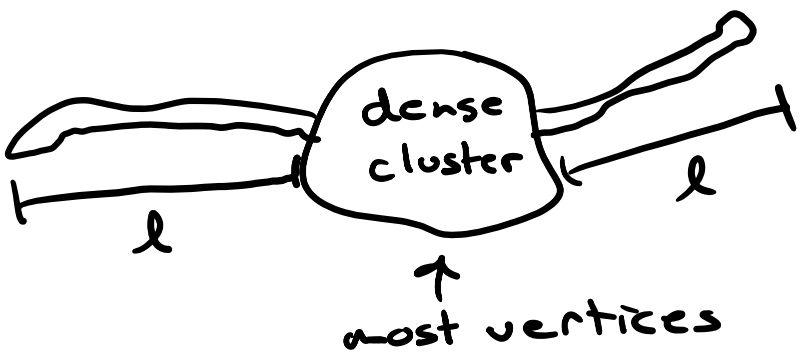
exact solution: compute APSP and

exact solution: compute APSP and take the max

(very slow)

approx solution: sampling-ish with BFS

Note: can't use naive sampling



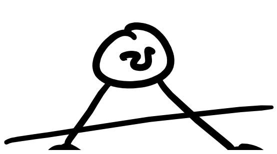
Issue: if we randomly sample, we are like to select vertices in dense core, which can result in error of 50%

Better Algorithm:

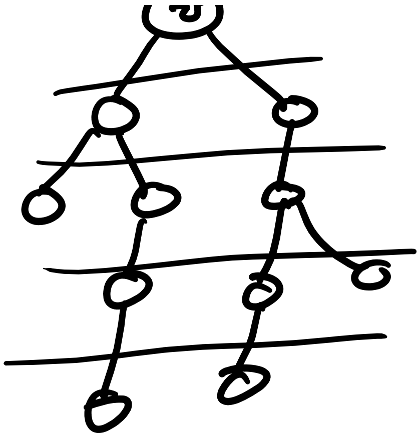
Select some root v randomly

For some # iter

$T = BFS(v)$

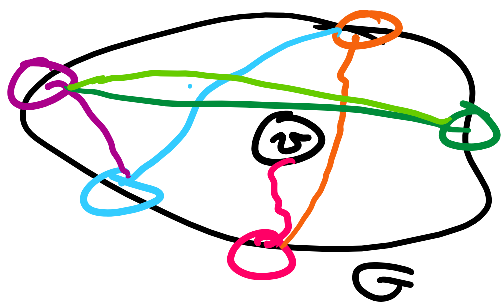


v = random selection from lowest



v = random selection from lowest level of T
(farthest from original root)

keep going until estimate stops increasing



Note: much better approximation, but it is still an approximation

---

As we talked about the emperical properties that arise in real networks $\longrightarrow$ why?

Our focus for now: <u>social networks</u>

$\hookrightarrow$ Human interaction networks

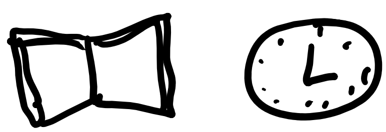
vertices: humans

vertices: humans
edge: friendships, other forms of interaction

Biggest driver of social network growth in general

→ <u>Triadic closure</u>

Story time w/ Slota

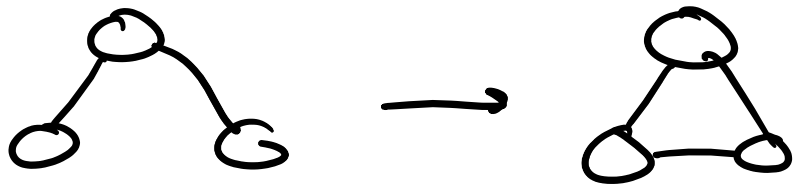


For sharing photos on my phone:
Google gave me 4 options for sharing photos:

1. my wife → strongest connection
2. my friend's wife → I don't communicate with at all
3. my ex → haven't communicated with in some time
4. mom → same as 1

Speculation: why were options 2,3 given?

Triadic closure

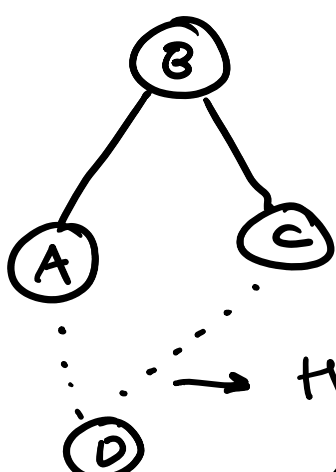


triad → closed triad

Triads close at a frequency <u>higher than expected</u> on social networks

↳ relative to random edges

Consider:



A and B are friends

B and C are friends

A and C are NOT friends

→ there is a good chance A and C become friends over time

Why:

<u>Opportunity</u>: B spends time with

<u>Opportunity</u>: B spends time with both A and C separately, but will likely spend time with both together

<u>Trust</u>: B trusts A and C

A and C trust B

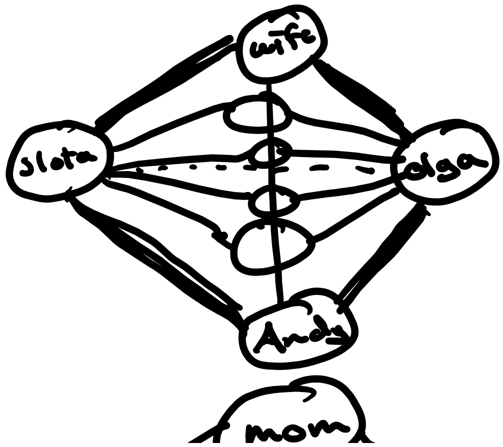
over time: A and C trust each other

<u>Incentive</u>: B might want to close the triad

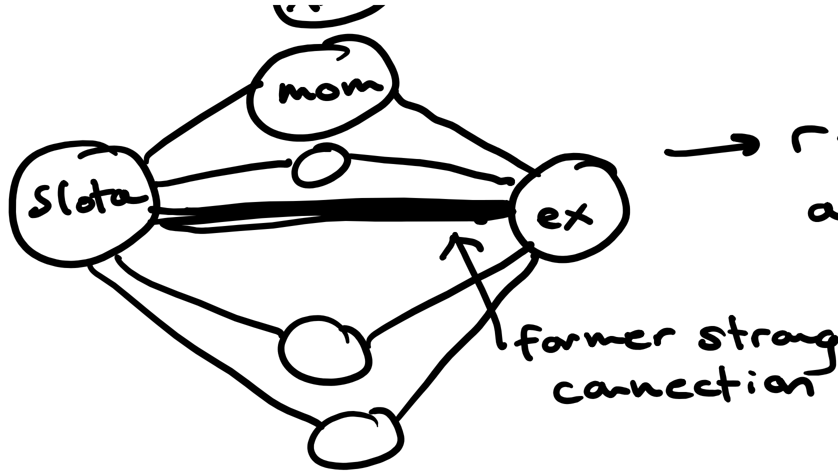
→ Result: triads close faster than random

AND

they close faster proportional to the number of common neighbors



→ Google wanted me to close the triad



→ re-close triad in a temporal aspect

Consequences of triadic closure:

* Large clustering
  (clustering coefficient will increase over time)

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

Note: can be local or global

local → for a single vertex

global → over all vertices

* Small diameter: note how triadic closure decreases a "local diameter"

* And decrease avg. shortest path lengths

Related concept: Homophily

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

OR: similar people are more likely to become friends

- Selection: we seek out people similar to ourselves or are inherently close to them

- Influence: we become more similar to people close to us

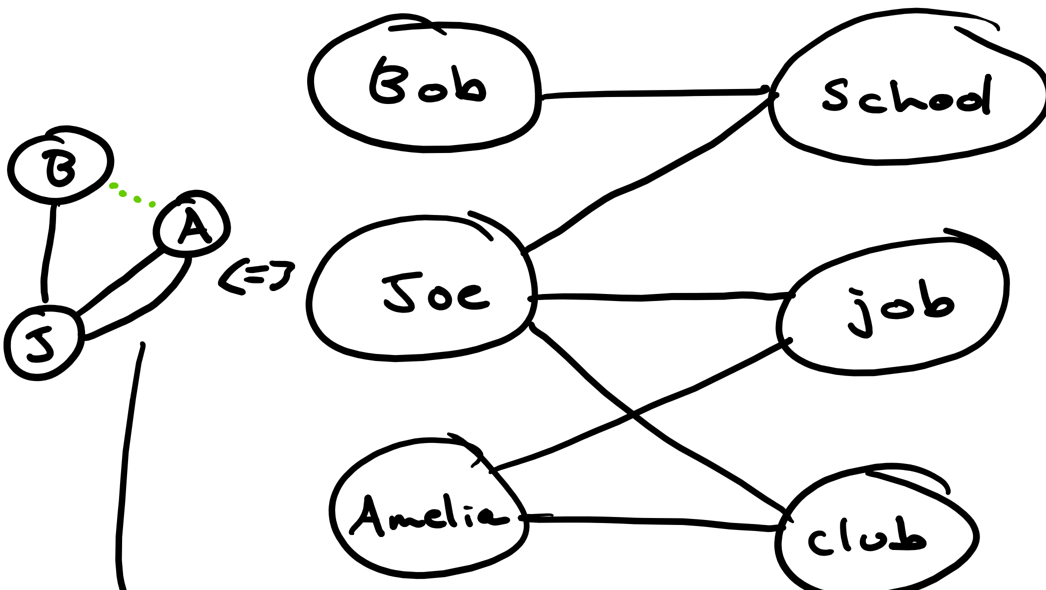
How homophily drives social network growth:

↳ dense clusters of people who are "similar" in some way

Relevance to graph mining:
We can use this idea to:

1. Predict network growth
2. Infer properties of individuals or groups

Expanding Triadic Closure to affiliation networks



→ overlapping affiliations facilates triadic closure

→ Bob and Amelia might overlap in affiliation over time

Same basic idea and consequences of standard triadic closure

→ mutual affiliations increase over time

One more takeaway:

One more takeaway:
connection strength increases over time

Dynamic and Temporal Networks

dynamic: changes over time

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

* Useful for training or evaluating growth predictors

Experiment:

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

→ after Δt we measure how many have closed

→ We can also look at strength of triadic closure based on number of mutual neighbors

of triadic closure based on
number of mutual neighbors

Q: can we observe evidence
of triadic closure over time?