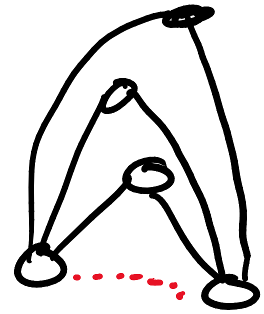
Lecture 5 - more social networks

Thursday, January 27, 2022 8:23 AM



Last class:

- Triadic closure
- Homophily → selection
  influence
- Temporal networks
  → changes over time

Today: strength of ties
diffusion

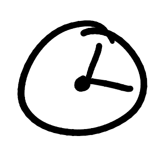
---

strong vs. weak ties

strong vs. weak => strength of links/
edge weight

- # of communications
- Time spent together
- etc.

Example: Story o'clock



Emperically: Job seekers found

Empirically: job seekers found open positions more often through acquaintences than close friends

why? Consider the structure of a social

→ dense clusters which are loosely coupled

→ information spreads through links connecting those clusters

strong tie: strong connection (usually internal to cluster)

weak tie: weaker connection (connecting clusters)

Connectivity

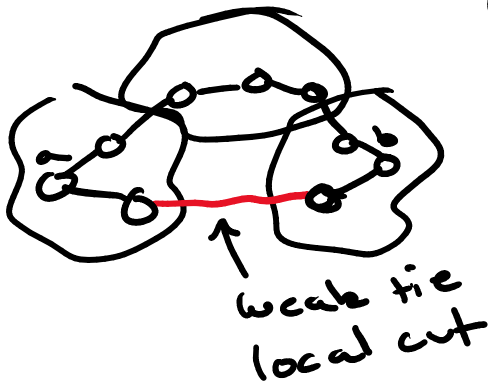
→ weak ties ≟ cut edges

not necessarily. more likely as

not necessarily, more likely as a whole

→ However, we have the notion of "local bridges" or a <u>local cut</u>

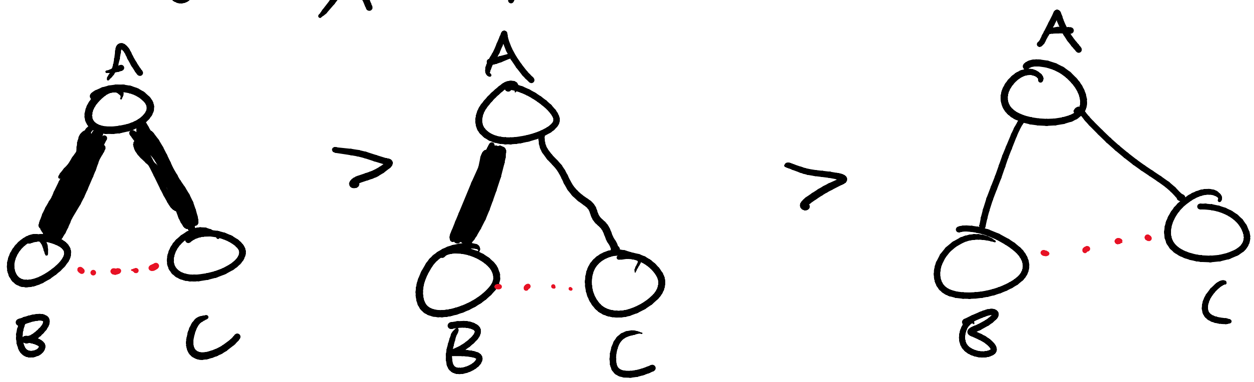
↳ removing a local cut increases shortest path distances for some vertex pair



How do ties relate to triadic closure?

<u>Strong triadic closure property</u>

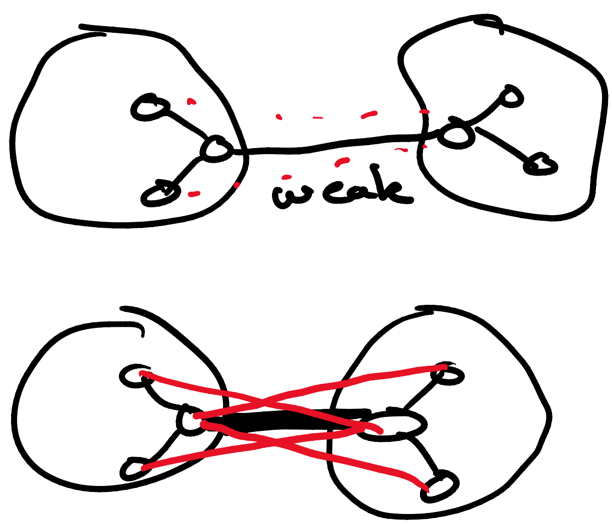
If A is connected to B, C via strong ties and B, C aren't connected, there's a higher likelihood edge (B, C) is created than if at least one of A's ties was weak of A's ties was weak



Think of

weak ties → local bridges

If a local bridge is a strong ties → the surrounding neighborhood is likely to become more tightly coupled



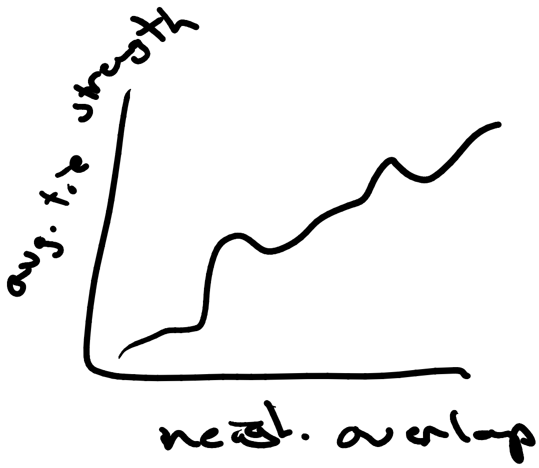
Let's quantify this:

Let's go into this.

- Consider a weighted network

First: correlate neighborhood overlaps with tie strength

Second: remove ties from $w \to s$ and $s \to w$ and observe connectivity



Diffusive processes

Generally: we've considering how information, data, etc. spread through some networks

Basic models:

Vertex-centric behavior

$\to v$ updates its state based off of that state of $N(v)$

→ v updates its state based off of that state of $N(v)$

Complexity of dynamics depends on complexity of each vertex's updating

→ small local change can have large global change

Simple example: Label propagation

Our algorithm

```
For v ∈ V(G):          ← vertex id 1....n
    State[v] = vid(v)

while updates happen:
    For v ∈ V(G):
        counts = {}
        For u ∈ N(v):
            counts[state[u]] += 1
        State[v] = argmax(counts)
```

Q: How are diffusion and network ties related?

→ cutting weak ties increases

→ cutting weak ties increases avg. shortest paths length

→ increase steps for a diffusive process to spread

→ disconnecting graph stops process

Think: network resilience

* We can strategically cut a network if that is our goal (future: epidemiology)

Growth models

Note: the impact of all our observations is dependent on an underlying growth process

Many networks grow via "preferential attachment" (P.A.)

- P. A. = higher likelihood of attachment to higher degree verts

aka 'Rich get richer'

aka rich get

- Barabasi-Albert model

start with no vertices

attach new vertex v to some existing vertex u with probability

$$p_{uv} = \frac{d(u)}{\sum_{i \in V(G)} d(i)}$$

all of this explains degree skew

To summarize:

Triadic closure:

→ clustering coefficient increases

→ existence of communities

useful for link prediction

Homophily:

→ clusters become more homogeneous

→ selection and influence

useful for vertex classification

Preferential attachment:

Preferential attachment:

→ existence of hubs

→ degree skew

→ rich get richer

useful for link prediction

Weak vs. strong ties:

→ strong triadic closure

→ diffusive processes "jump" between clusters via weak ties