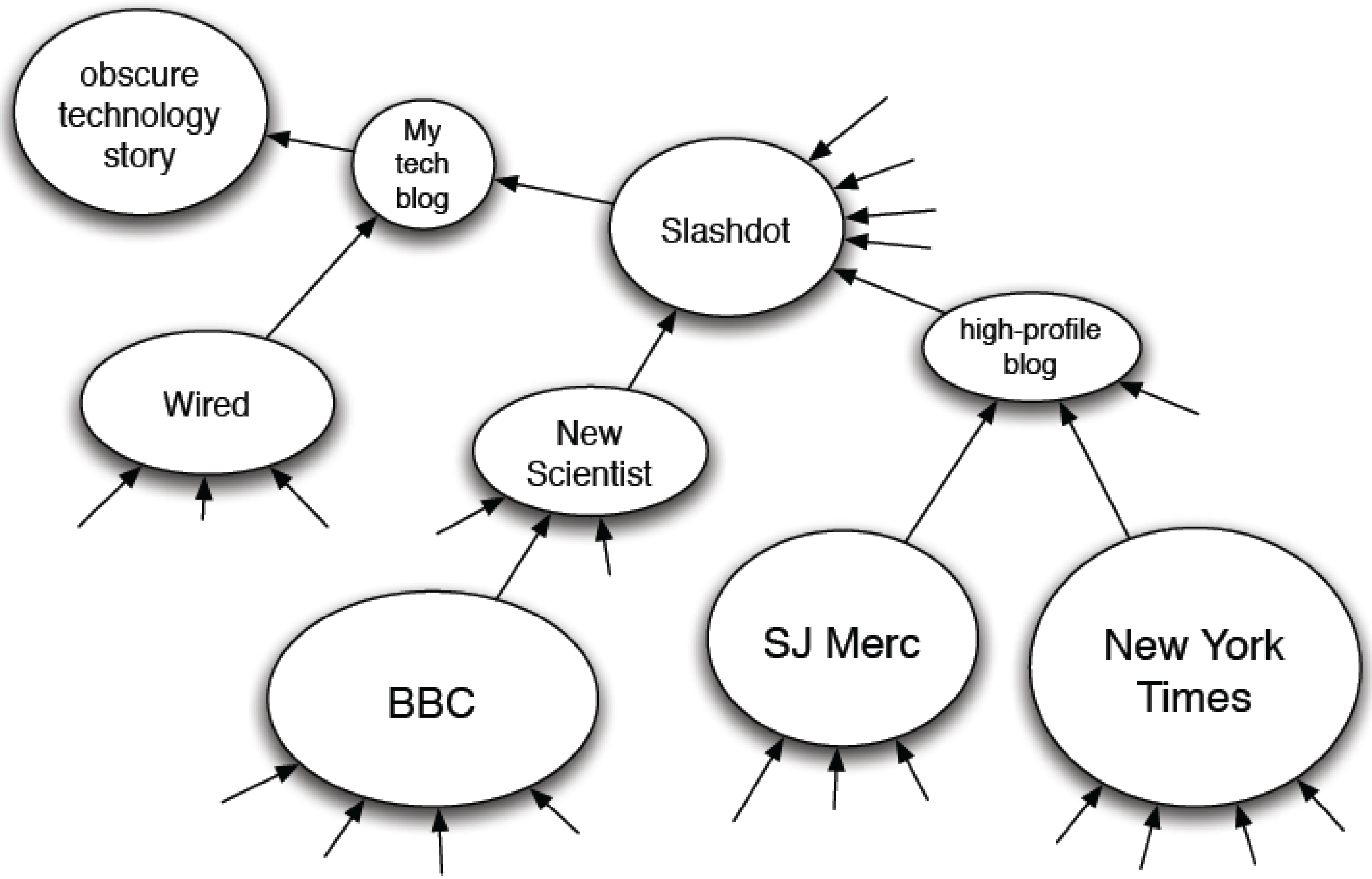
Cascading Behavior in Networks

CS 322: (Social and Information) Network Analysis
Jure Leskovec
Stanford University

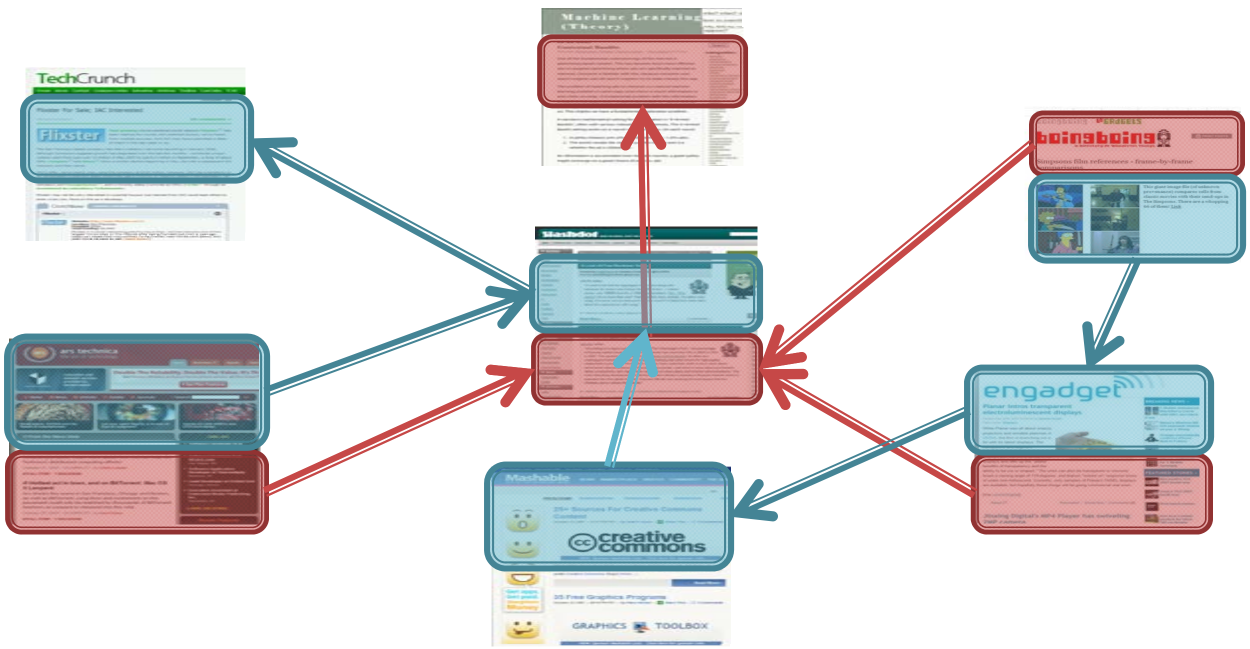
Announcements

- Homework 2 is out
- It is due in a week
  - 27 October in class

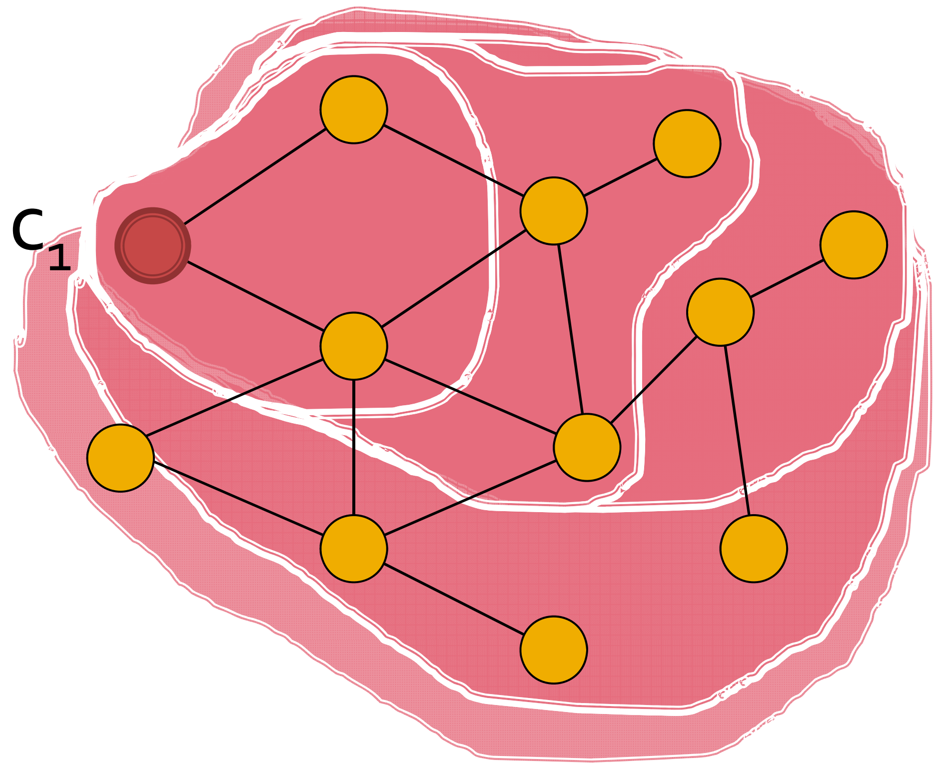
Information difussion



Information diffusion

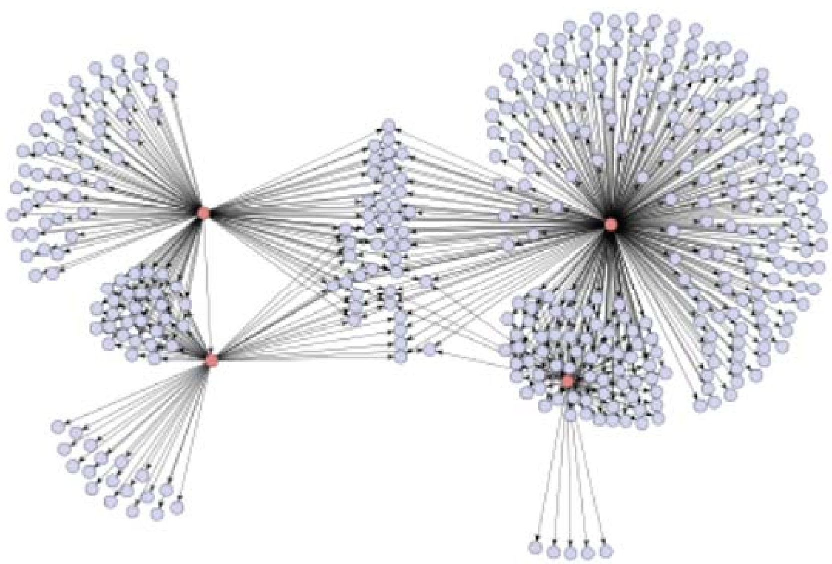


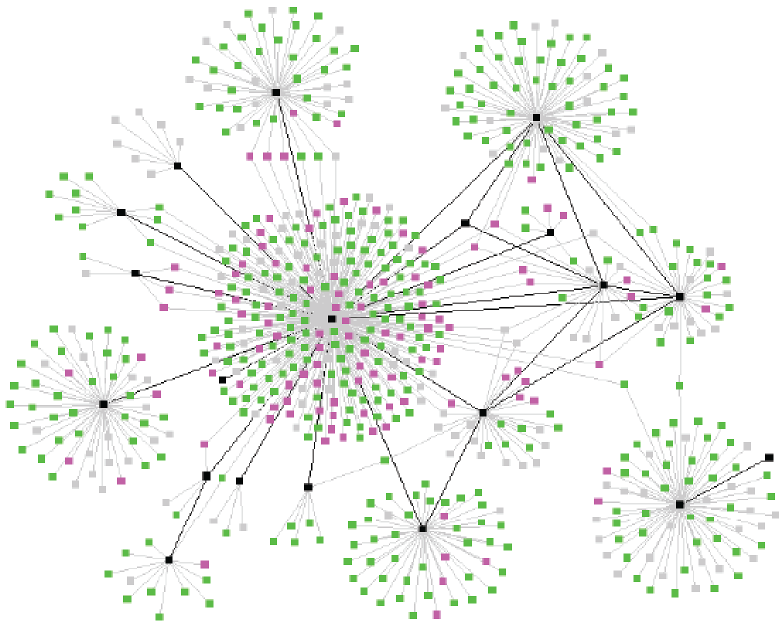
Spread of diseases



Diffusion in Social Networks

- One of the networks is a spread of a disease, the other one is product recommendations
- Which is which? ☺





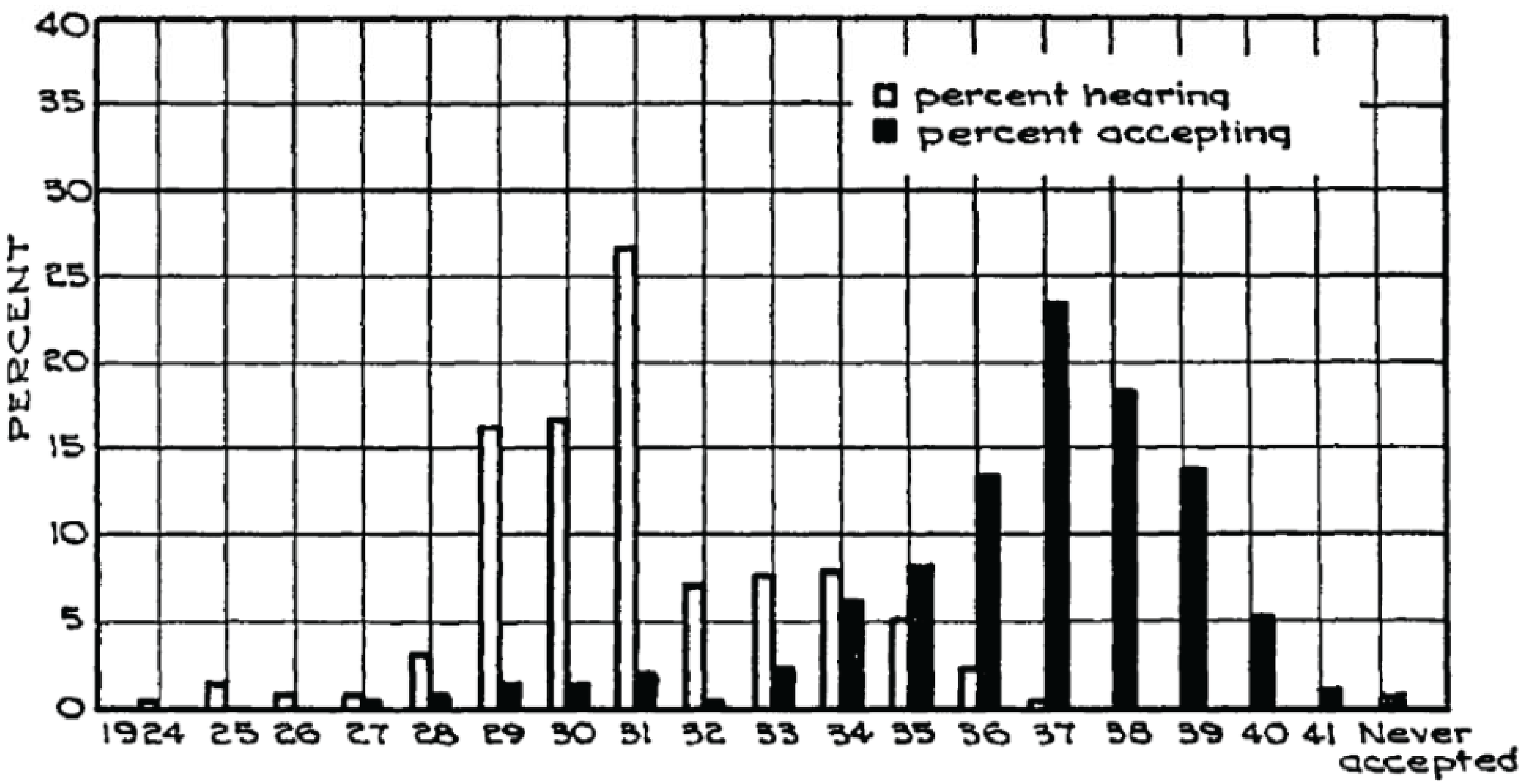
Diffusion in Networks

- A fundamental process in social networks:
  Behaviors that cascade from node to node like an epidemic
  - News, opinions, rumors, fads, urban legends, …
  - Word-of-mouth effects in marketing: rise of new websites, free web based services
  - Virus, disease propagation
  - Change in social priorities: smoking, recycling
  - Saturation news coverage: topic diffusion among bloggers
  - Internet-energized political campaigns
  - Cascading failures in financial markets
  - Localized effects: riots, people walking out of a lecture

Empirical Studies of Diffusion

- Experimental studies of diffusion:
  - Spread of new agricultural practices [Ryan-Gross 1943]
    - Adoption of a new hybrid-corn between the 259 farmers in Iowa
    - Classical study of diffusion
    - Interpersonal network plays important role in adoption → Diffusion is a social process
  - Spread of new medical practices [Coleman et al. 1966]
    - Studied the adoption of a new drug between doctors in Illinois
    - Clinical studies and scientific evaluations were not sufficient to convince the doctors
    - It was the social power of peers that led to adoption

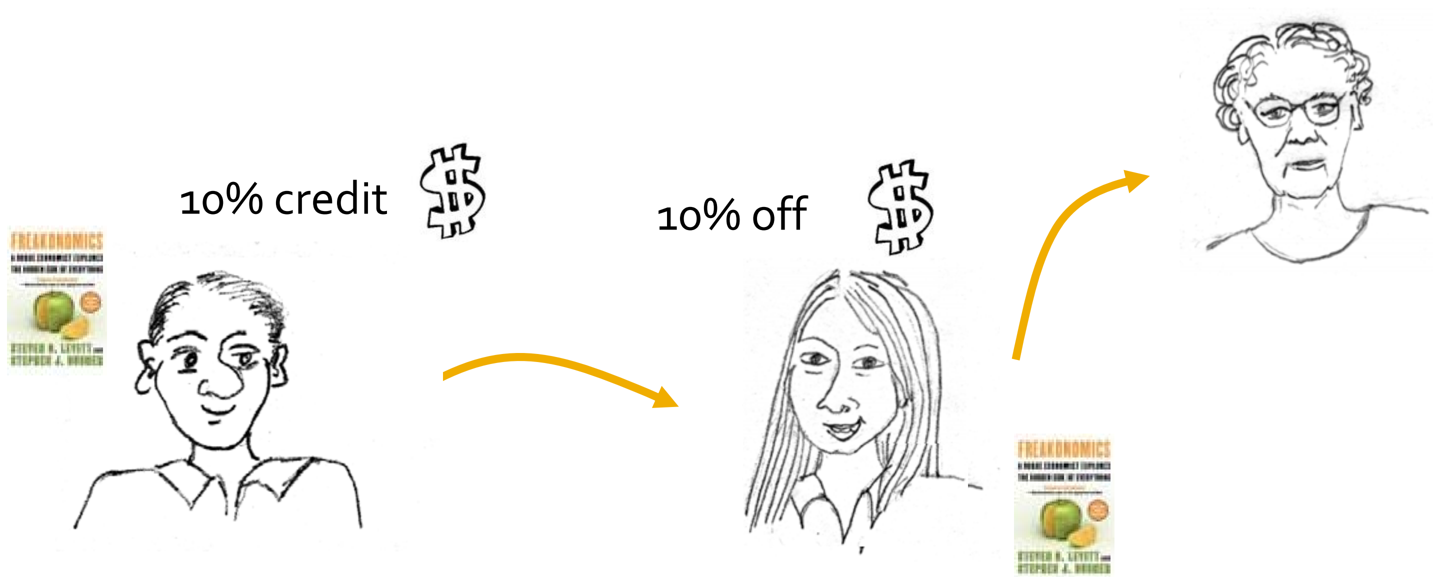
Hybrid Corn [Ryan-Gross 1966]



Diffusion is a social process

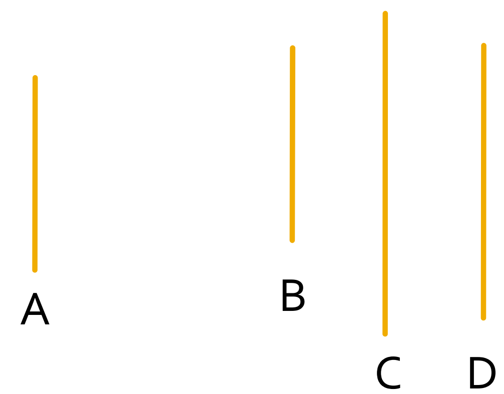
Diffusion in Viral Marketing

- Senders and followers of recommendations receive discounts on products



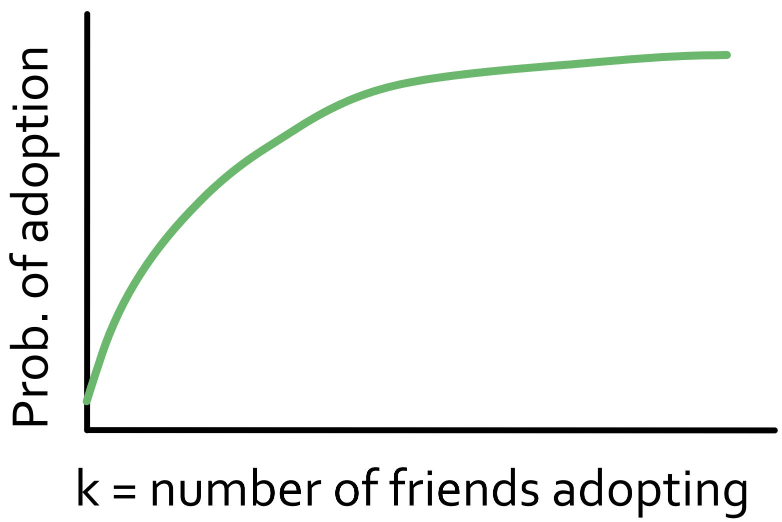
Empirical Studies of Diffusion (2)

- Diffusion has many (very interesting) ☺ flavors, *e.g.*:
  - The contagion of obesity [Christakis et al. 2007]
    - If you have an overweight friend your chances of becoming obese increases by 57%
  - Psychological effects of others' opinions, *e.g.*: Which line is closest in length to A? [Asch 1958]

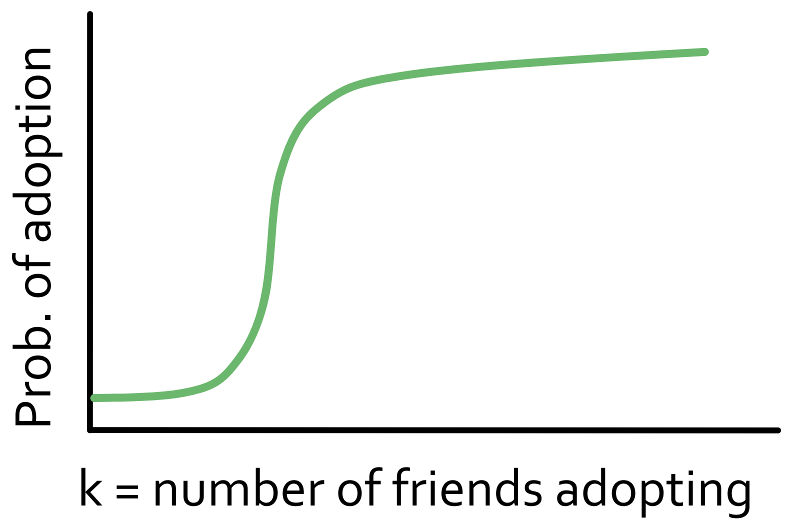


Diffusion Curves (1)

- Basis for models:
  - Probability of adopting new behavior depends on the number of friends who have adopted [Bass '69, Granovetter '78, Shelling '78]
- What's the dependence?

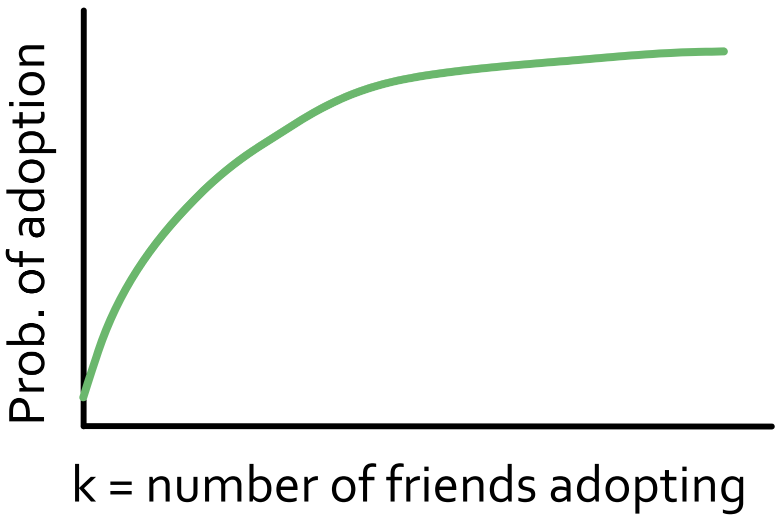


Diminishing returns?

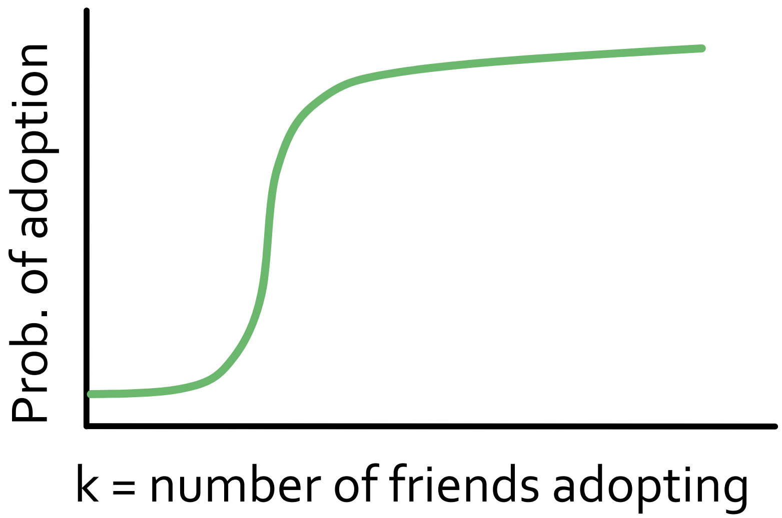


Critical mass?

Diffusion Curves (2)



Diminishing returns?

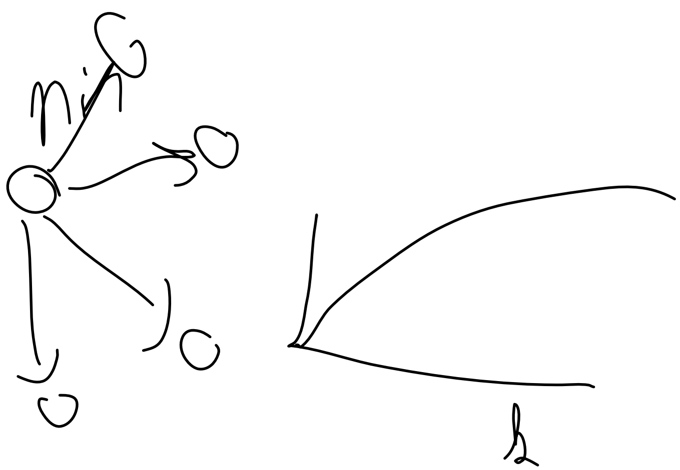


Critical mass?

- Key issue: qualitative shape of diffusion curves
  - Diminishing returns? Critical mass?
  - Distinction has consequences for models of diffusion at population level

How to model diffusion?

- Probabilistic models:
  - "catch" a disease with some prob. From neighbors in the network
- Decision based models:
  - Adopt new behaviors if *k* of your friends do



Models

- Two flavors, two types of questions:
  - A) Probabilistic models: Virus Propagation:
    - SIS: Susceptible – Infective – Susceptible (*e.g.*, flu)
    - SIR: Susceptible – Infective – Recovered (*e.g.*, chicken-pox)
    - Question: Will the virus take over the network?
  - B) Decision based models: Diffusion of Innovation
    - Independent contagion model
    - Threshold model
    - Questions:
      - Finding influential nodes
      - Detecting cascades

Decision based models

- Model where everyone sees everyone else's behavior
- Collective action [Granovetter, '78]
  - Example:
    - Neighborhoods in cities changing ethnic composition
    - Riots, protests, strikes
    - People leaving a lecture that runs over

A simple model

- n people – everyone can observe all actions
- Each person i has threshold t_i:
    - Node i will take part in the behavior iff at least t_i other people are already doing it:
        - Small t_i: early adopter
        - Large t_i: late adopter

Collective action

- Easy to simulate:
  - Given $t_1, \ldots, t_n$
  - $F(x)$ = fraction of people whose threshold $\leq x$ $\in (0,1)$
  - $s(t)$ = fraction of participants at time t $\in (0,1)$
- Then:

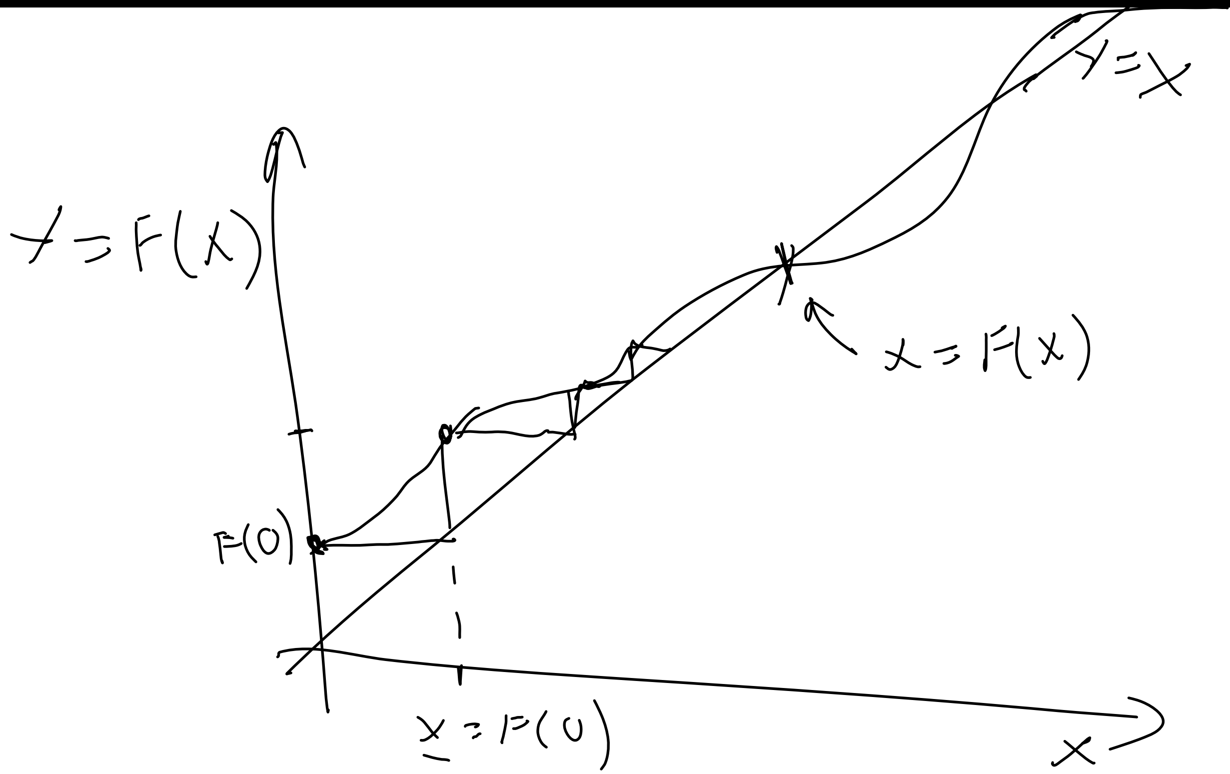
$$s(0) = 0$$

$$s(1) = F(0)$$

$$s(2) = F(s(1))$$

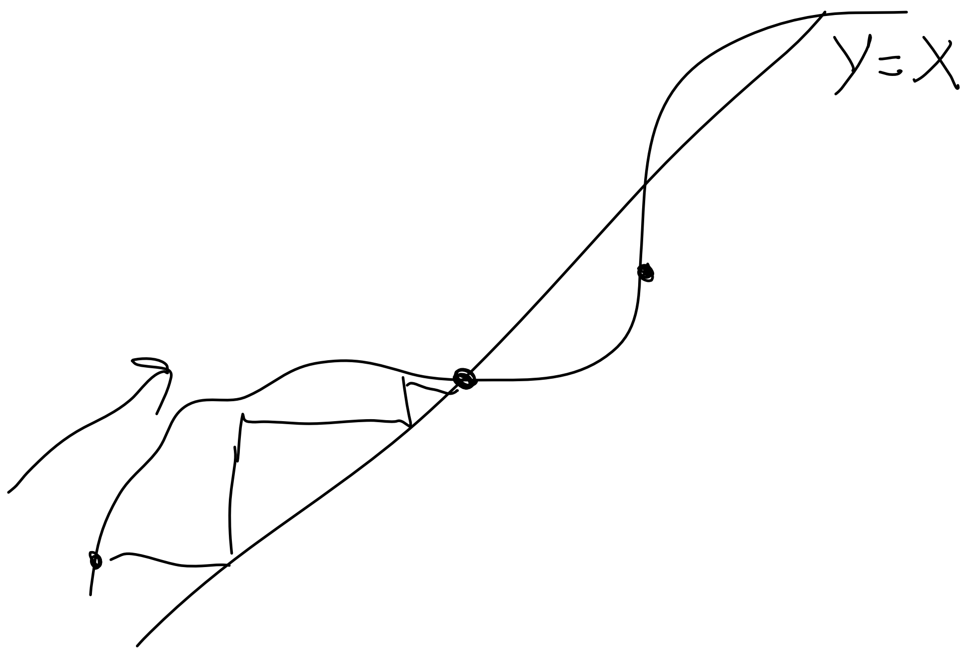
$$s(t+1) = F(s(t))$$

Collective action

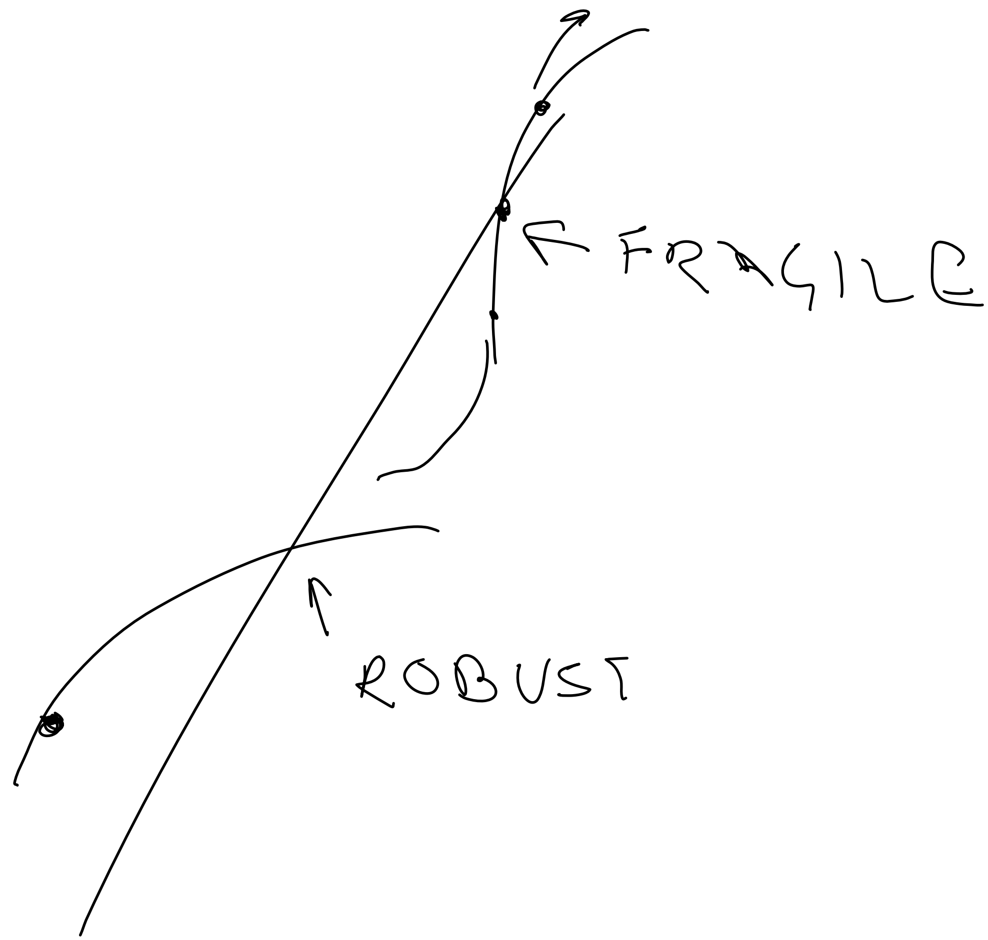


Starting elsewhere

- What if we start the process somewhere else?
  - We move up/down to the next fixed point



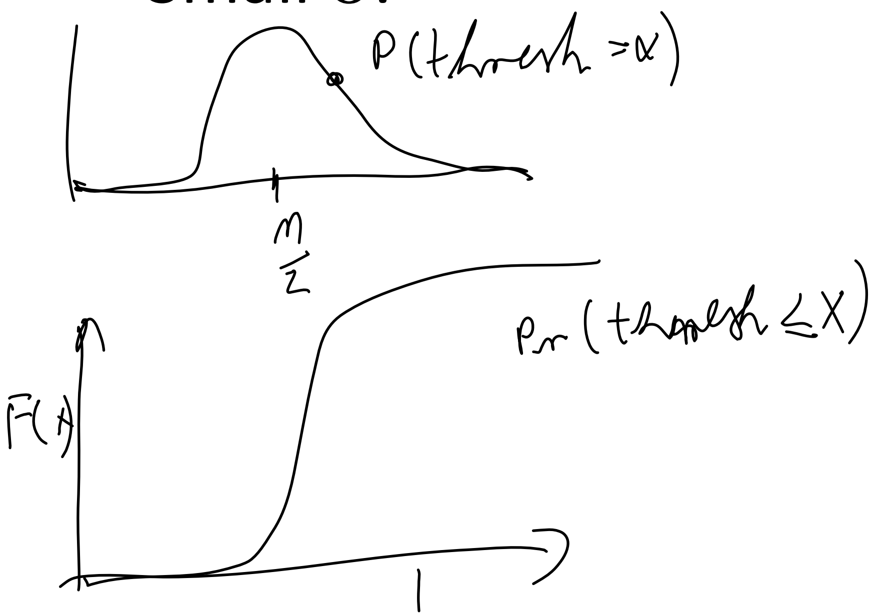
Fragile vs. robust fixed point



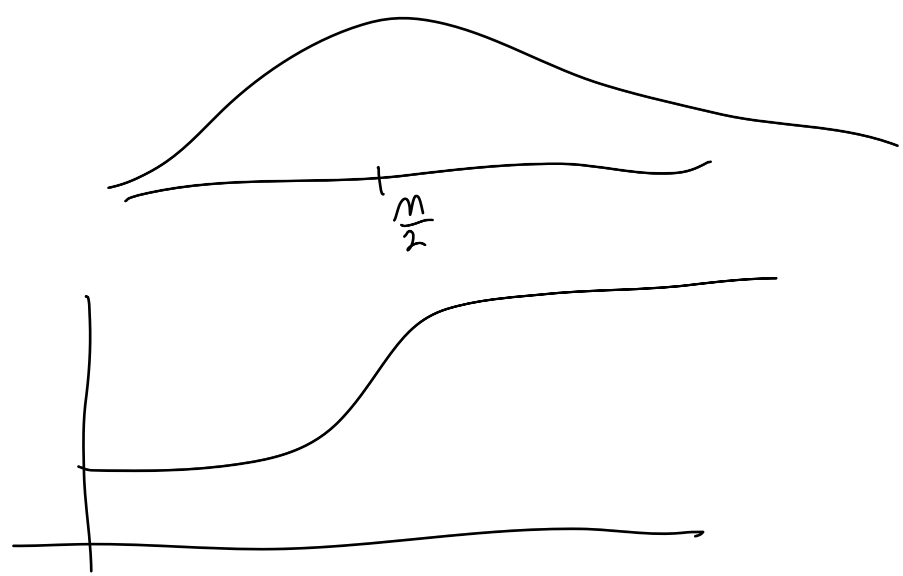
Discontinuous transition

- Each threshold t_i is drawn independently from some distribution $F(x) = Pr[thresh \leq x]$
  - Suppose: Normal with $\mu=n/2$, variance σ

Small σ:

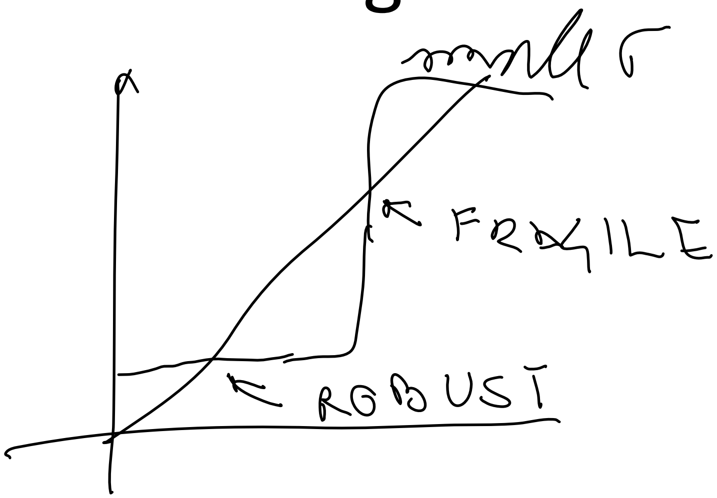


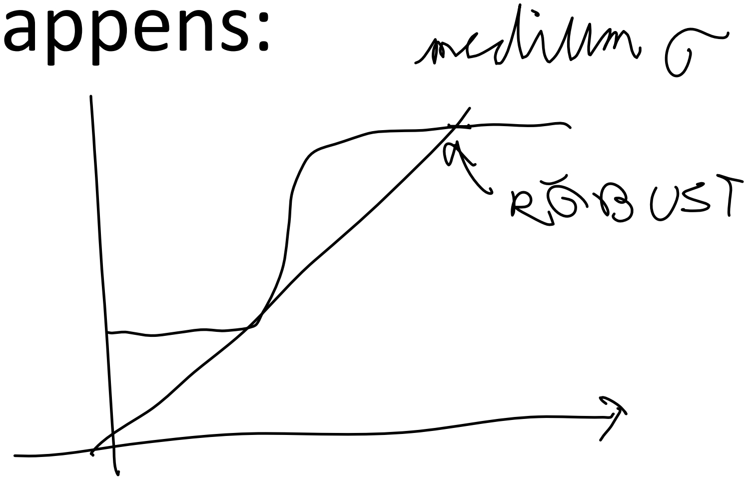
Large σ:



Discontinuous transition

- As we continuously vary the parameter σ, something discontinuous happens:

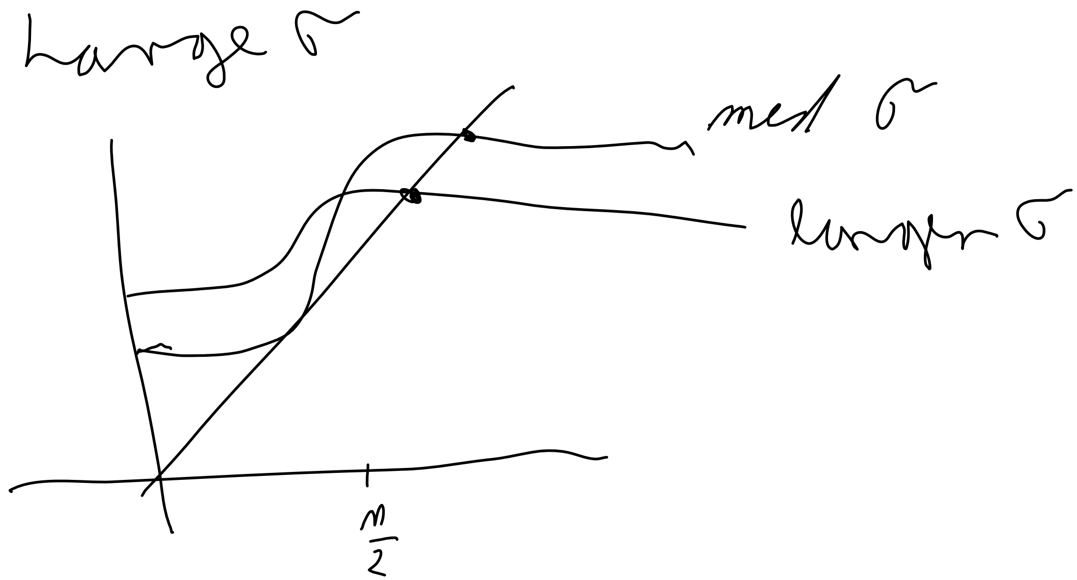




- Bigger variance let's you build a bridge from early adopters to mainstream

Discontinuous transition

- But if we increase the variance even further we move one higher fixed point lower



Caveat: Pluralistic ignorance

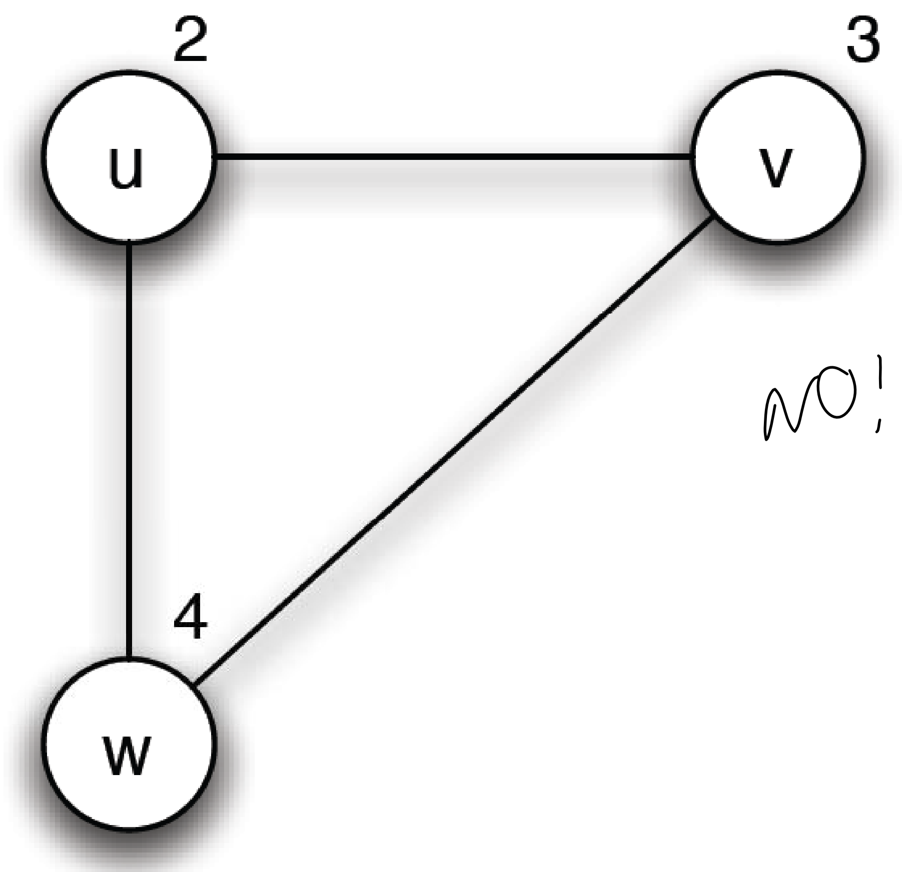
- Pluralistic ignorance – erroneous estimates about the prevalence of certain opinions in the population
- Survey conducted in the U.S. in 1970 showed that while a clear minority of white Americans at that point favored racial segregation, significantly more than 50% believed that it was favored by a majority of white Americans in their region of the country

The Model

- Personal threshold k: “I will show up to the protest if I am sure at least *k* people in total (including myself) will show up”
- Each node in the network knows the thresholds of all their friends

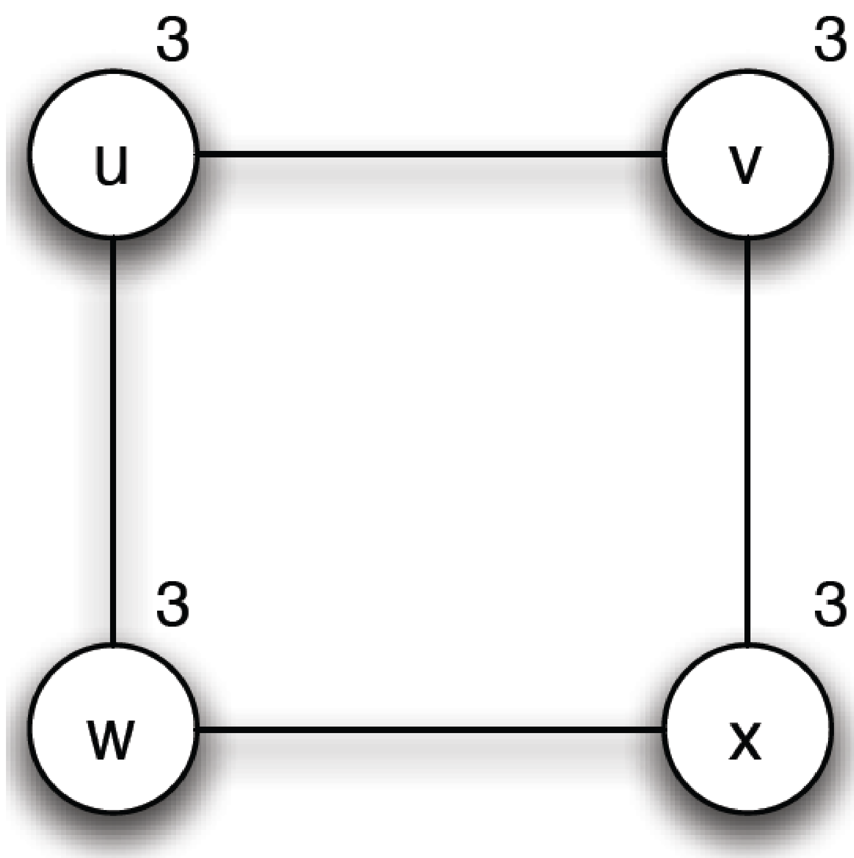
Subtle issues

- Will uprising occur?



Subtle issues

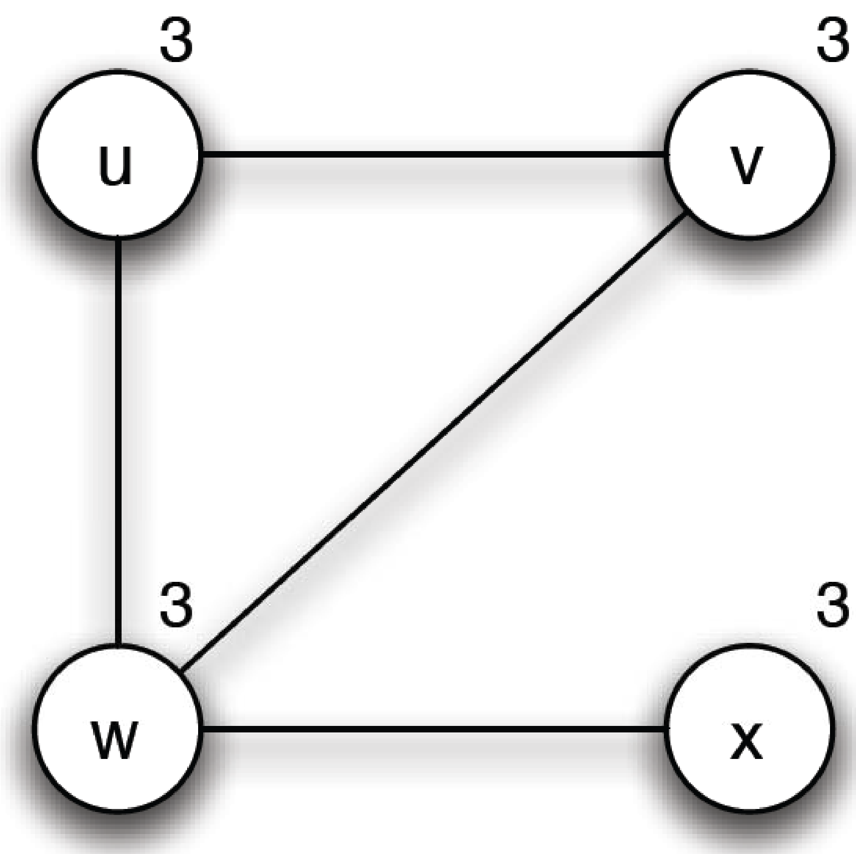
- Will uprising occur?



NO?

Subtle issues

- Will uprising occur?



YES!

Weaknesses of the model

- It does not take into account:
  - No notion of social network – more influential users
  - It matters who the early adopters are, not just how many
  - Models people's awareness of size of participation not just actual number of people participating
  - Modeling thresholds
    - Richer distributions
    - Deriving thresholds from mode basic assumptions
      - game theoretic models

Weaknesses of the model

- It does not take into account:
  - Modeling perceptions of who is adopting the behavior/ who you believe is adopting
  - Non monotone behavior – dropping out if too many people adopt
  - Similarity – thresholds not based only on numbers
  - People get “locked in” to certain choice over a period of time

Game theoretic model of diffusion

- Based on 2 player coordination game
- 2 players – each chooses technology A or B