Chapter 13

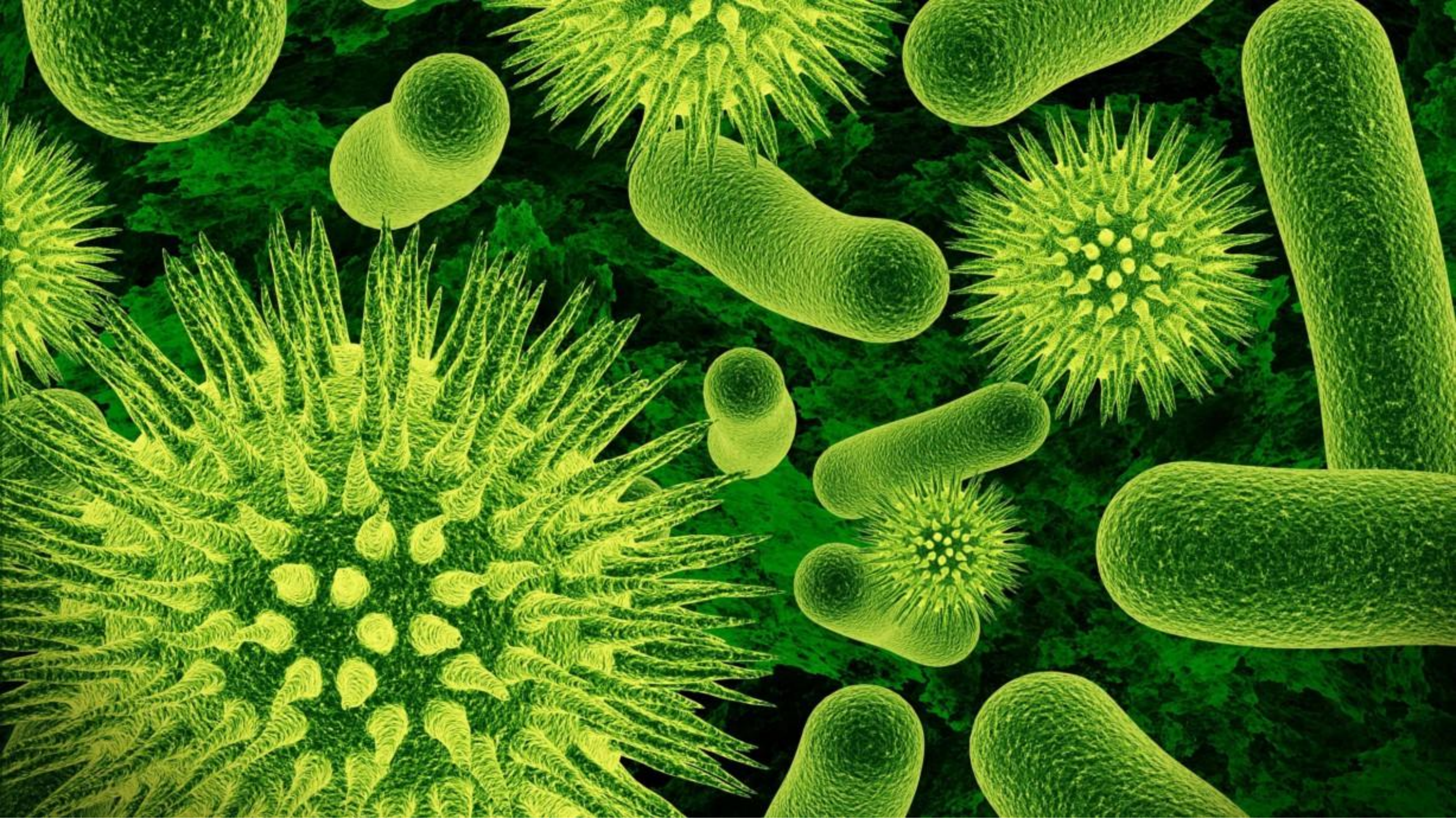
Diffusion: Decision-Based

Summary

- Diffusion phenomena
- Human Behaviours
- Threshold based models
- Diminishing returns
- Influence Maximization
- Advanced topics: Herding

Reading

- Chapter 19 of Kleinberg's book




Strategy Innovation
Vision
Creativity
Support

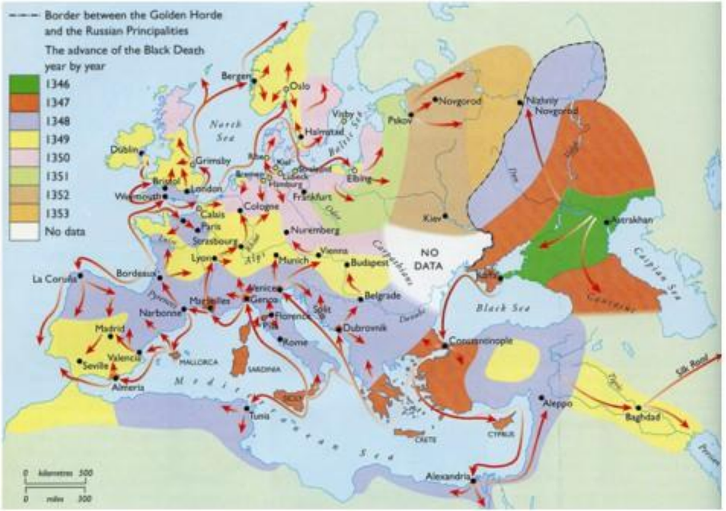

INNOVATION
BUSINESS
PLAN
IDEA

Why study diffusive processes matter?

High mobility

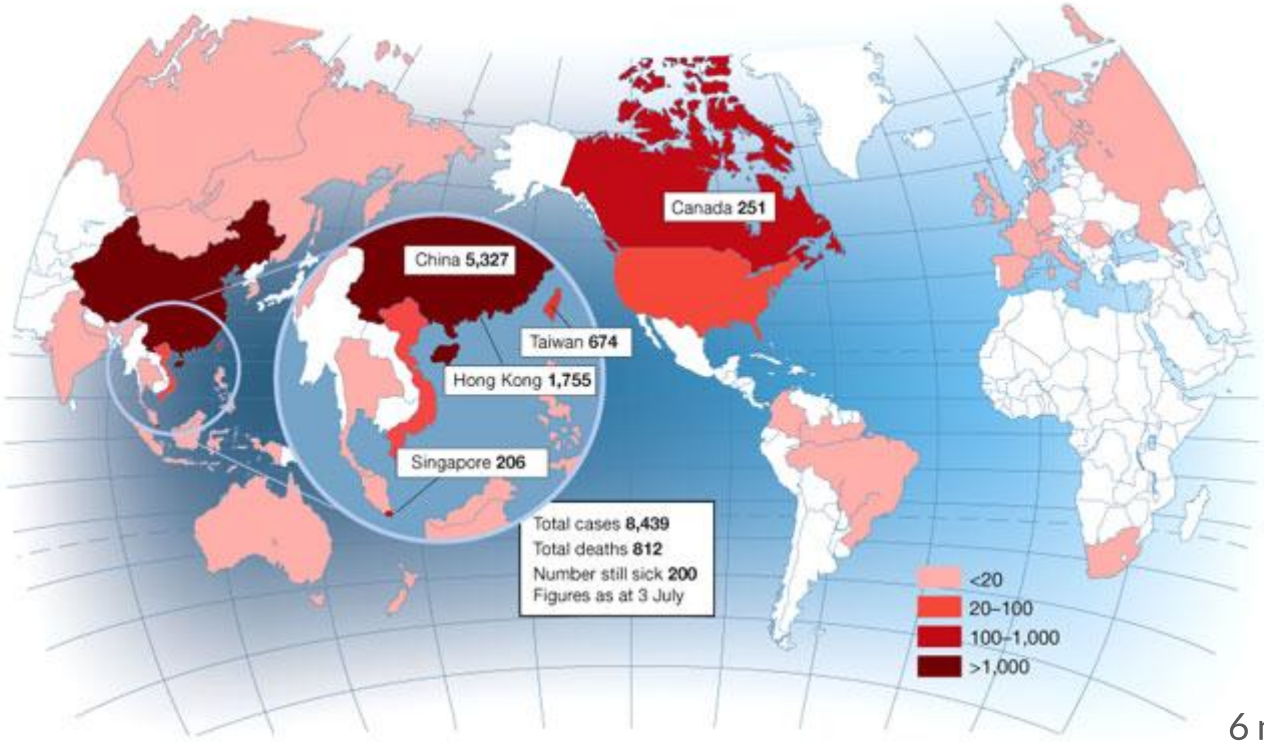
High population density

The great plague in 14th century



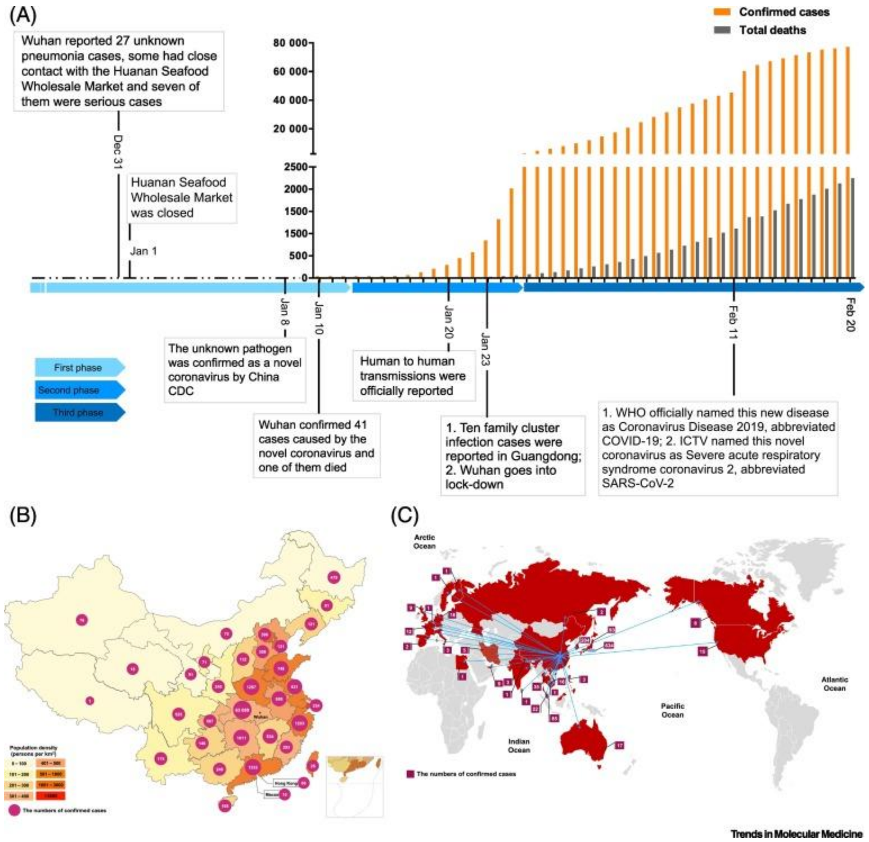
4 years from France to Sweden

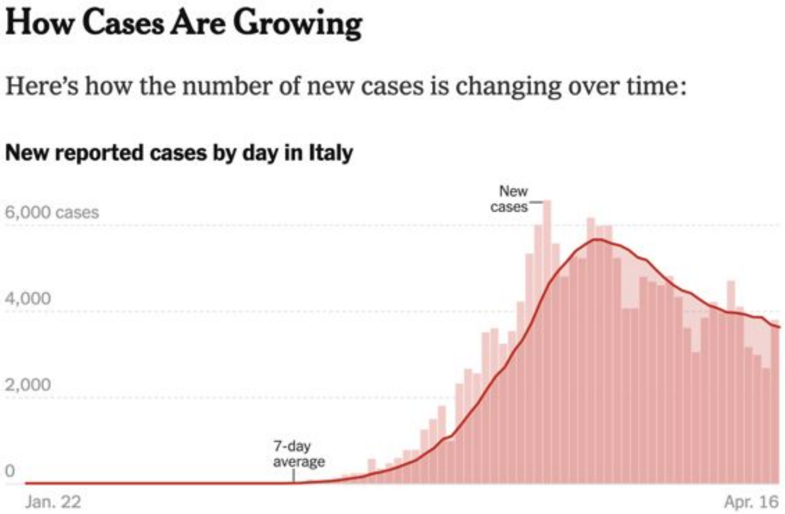
SARS in 21th century

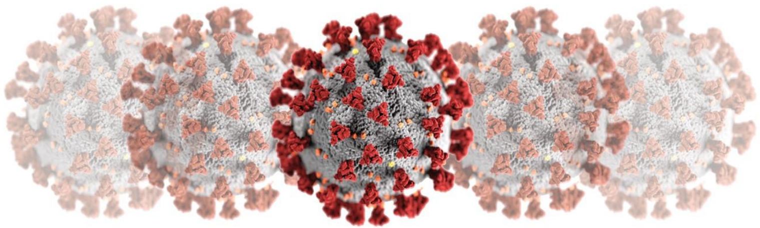


6 months...

Covid-19 today...







How can we model diffusive phenomena?

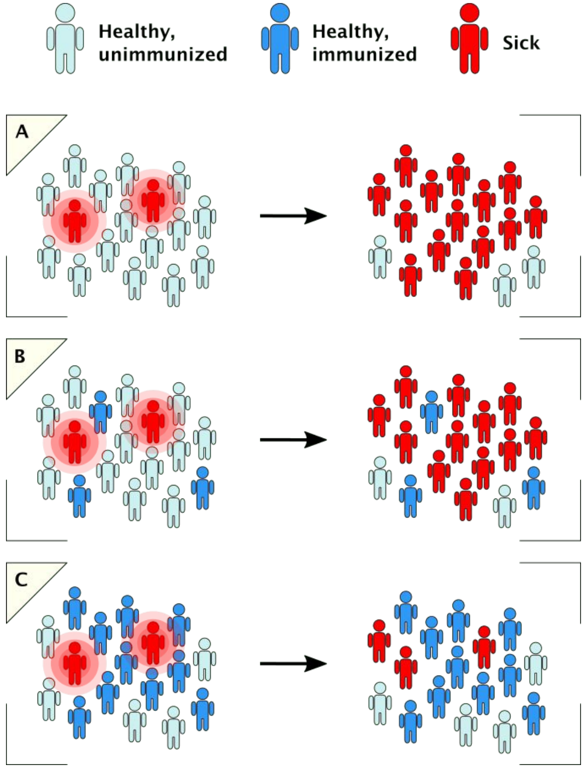
A First Approximation:

Mean Field modeling

MF Theory analyze high-dimensional models by studying a simpler models that approximate the original by averaging over degrees of freedom.

- Such models consider a large number of individual components that interact with each other.

- In MFT, the effect of all the other individuals on any given individual is approximated by a single averaged effect.



Refining MFT

Network effects

Diffusion happens only when the carries of the diseases/virus/idea are connected to susceptible nodes.

Diffusive phenomena can modeled describing:

- “node statuses”
- “transition rules”

We will always start from a MF model, then extend it to a networked context.

Spreading speed & patterns will vary depending on:

- model’s parameter values (as in mean field)
- Initial infection seeds
- Topology that surrounds infected nodes (e.g., communities may act as barriers)
- Degree of homophily of connected nodes
- ...

Modeling Human Behaviours

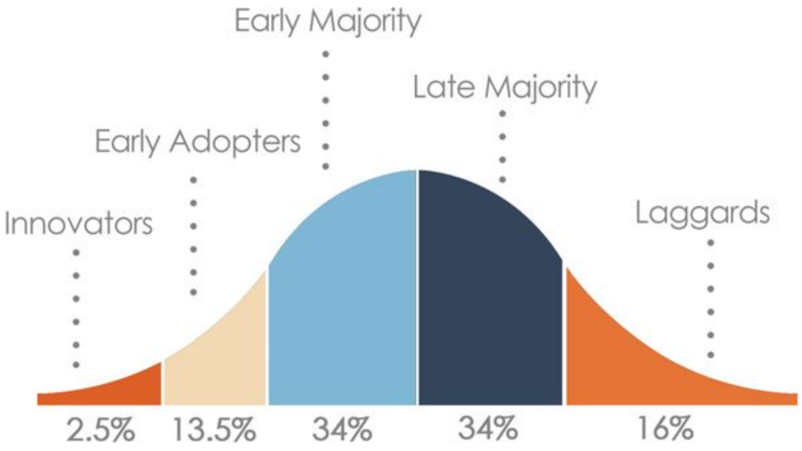
Diffusion of Innovations

Models of product adoption & decision making

An individual observes decisions of its peers and makes its own decision

Example:

You decide to buy an iPad only if only k% of the population has already done so



Rogers, Everett M. Diffusion of innovations. Simon and Schuster, 2010.

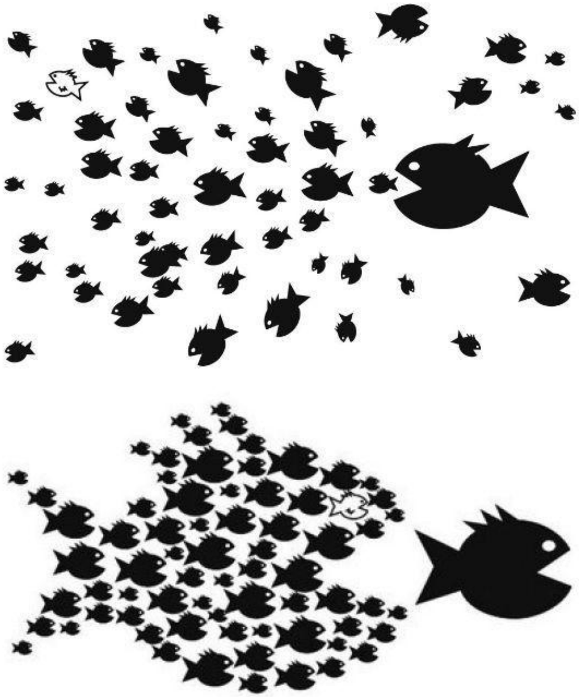
Collective Action

Everyone sees everyone else’s behaviour

- No network structure

Example:

- clapping, getting up and leaving in a theater
- keep your money or not in a stock market
- neighborhoods in cities changing their ethnic composition
- riots, protests, strikes...



Granovetter, Mark. "Threshold models of collective behavior." American journal of sociology 83.6 (1978): 1420-1443.

Collective Action

Idea

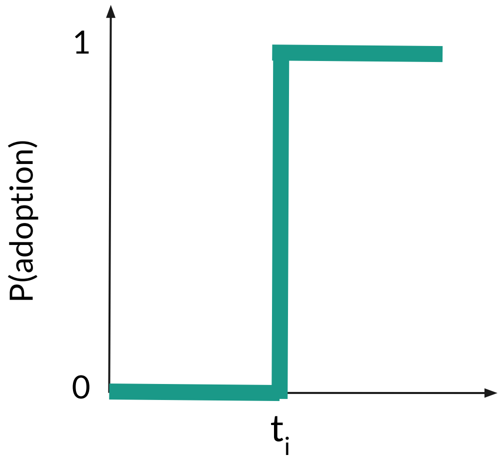
N people: everyone observes all actions

Each person i has a threshold t_i

- Node i will adopt the behaviour iff at least t_i other people are adopters
  - Small t_i: early adopter
  - Large t_i: late adopter

The population is described $\{t_1, \ldots, t_n\}$

- F(x) is the fraction of people with threshold $t_i \leq x$



Collective Action

Description

Think of the step-by-step change in number of people adopting the behaviour:

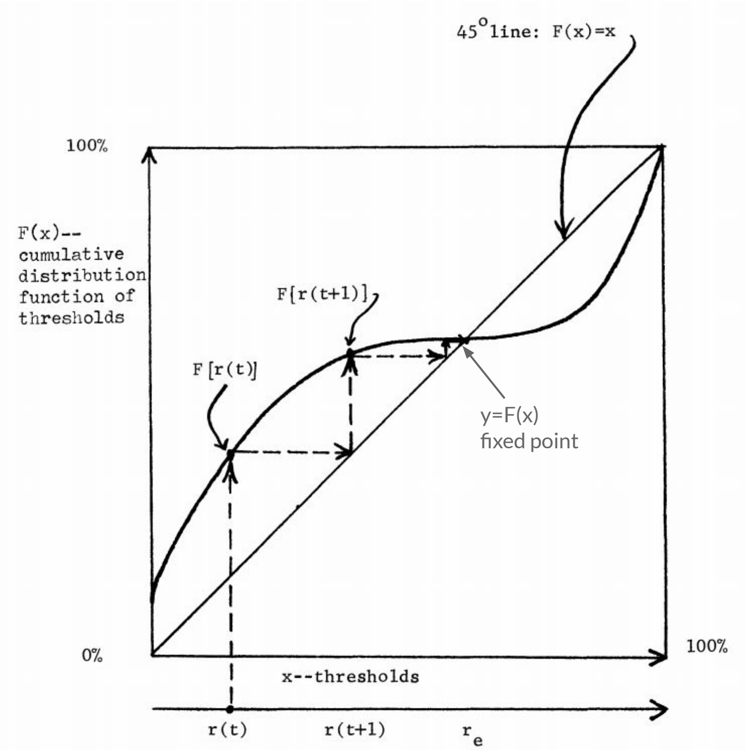
- **F(x):** fraction of the people with threshold ≤x
- **s(t) :** number of participants at time t

Easy to simulate:

- s(0) = 0
- s(1) = F(0)
- s(2) = F(s(1)) = F(F(0))
- s(t+1) = F(s(t)) = $F^{t+1}(0)$

Fixed point: F(x) = x

There could be other fixed points but starting from 0 we never reach them



Collective Action

Simulations

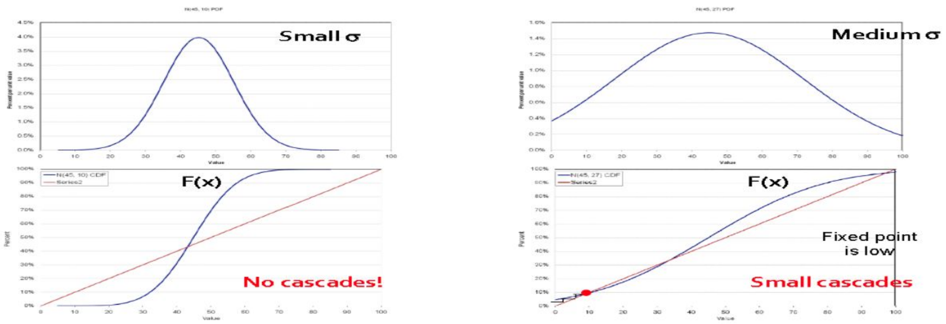
Each threshold t_i is drawn independently from some distribution:

- $F(x) = \Pr[\text{threshold} \leq x]$

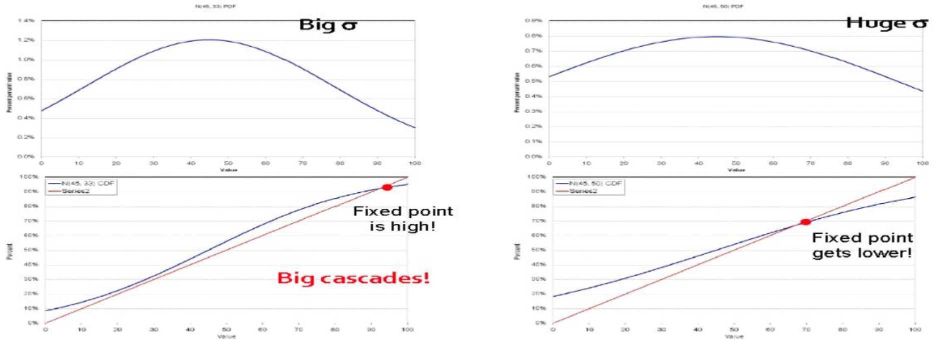
Let's assume a normal distribution with

- mean $\mu > n/2$
- variance σ

Increasing variance allow bridges from early adopters to mainstream...



... until the fixed point lowers!



Collective Action

Limitations

Limitations:

- No notion of social structure
- Leverages only volume of early adoptions, not their characteristics
- Captures only number of participants, not their awareness
- Simplified threshold:

  Richer distributions?
  Derivating threshold from more basic assumptions (e.g., game theoretic models)

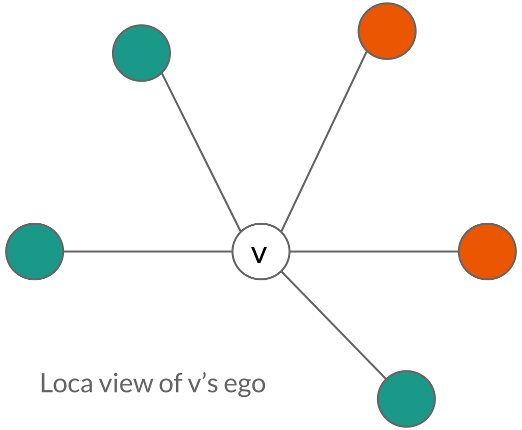
Cascades and Threshold based models

Cascades

Game theoretic models

Based on 2 players coordination games

- Each player chooses a technology
- Each person can adopt only one "behaviour" (e.g., A or B)
- Your gain increase if your friend has adopted the same behaviour as you



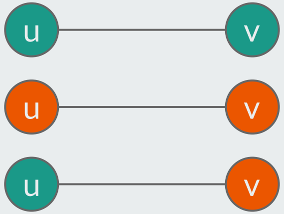
Loca view of v's ego

Payoff

Game theoretic models

Payoff matrix:

- if u and v adopt the same behaviour A they each get payoff a>0
- if u and v adopt the same behaviour B they each get payoff b>0
- if u and v opposite behaviours they each get 0



Each node v plays a copy of the game with each of its neighbors

Payoff:

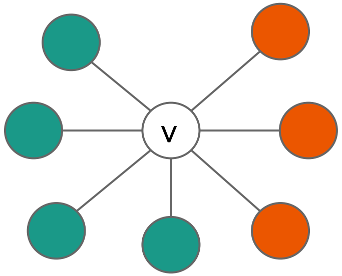
Sum of nodes payoff per game

Decision rule

- let v have d neighbors
- assume fraction p of v's neighbor adopt A

$$\text{Payoff}_v = a*p*d \quad \text{if v chooses A}$$
$$= b*(1-p)*d \quad \text{if v chooses B}$$

- v chooses A if: a*p*d > b*(1-p)*d



v chooses A if p>q $\quad q = \frac{b}{a+b}$

Cascades

Example

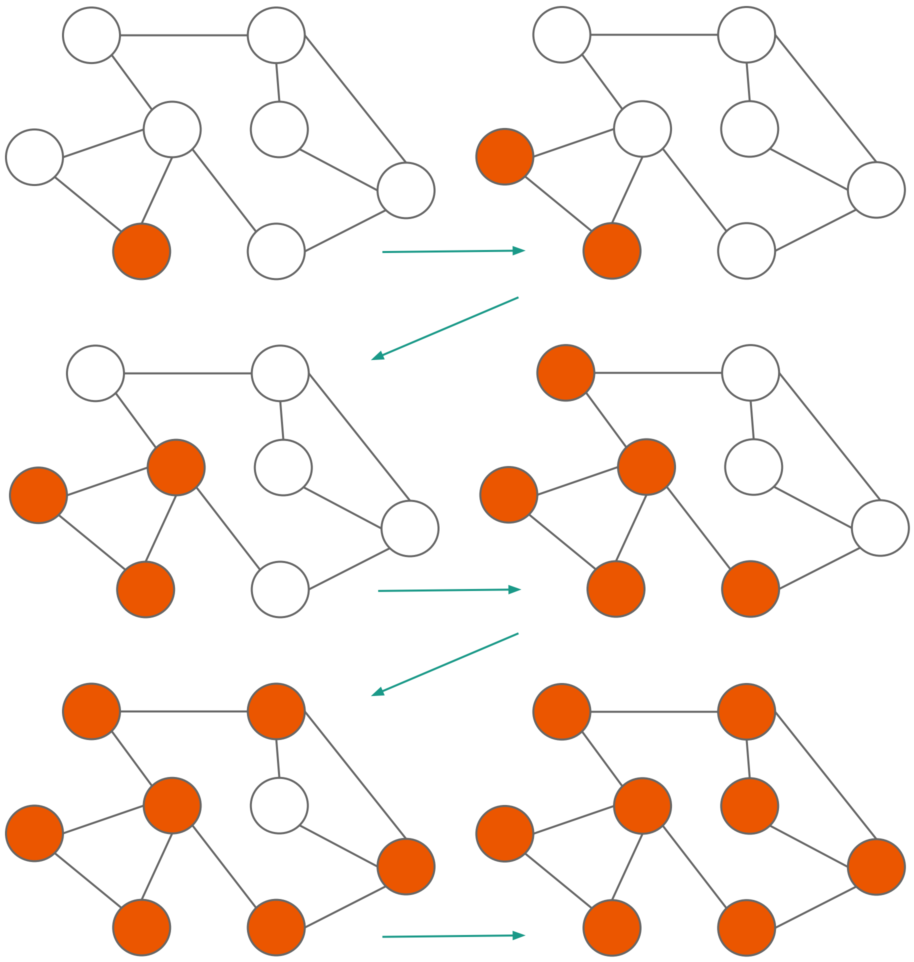
Graph where everyone starts with B.
Small set S of early adopters of A

- Hard wire S: they keep using A no matter what payoffs tell them to do

Payoff are set in such a way that nodes say:

- If at least 30% of my friends are red I'll be red

NB: the cascade process (fixed S, the threshold and, the network topology) is deterministic.



Cascades

Stopping cascades

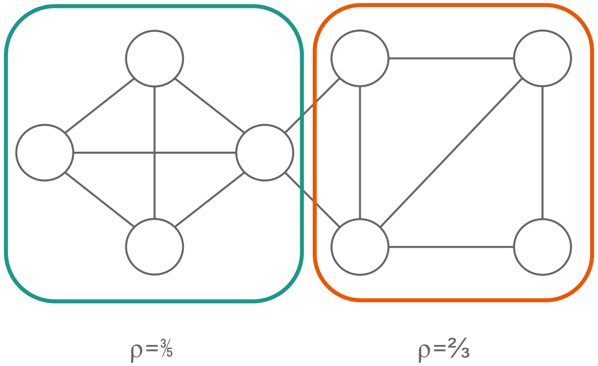
What prevents cascades from spreading?

Assuming an homogeneous threshold **q**:

- if G/S contains a cluster of density > (1-q) than S can not cause a cascade
- if S fails to create a cascade then there is a cluster of density >(1-q) in G/S

Definition:

A cluster of density ρ is a set of nodes C where each node in the set has at least ρ fraction of edges in C



Diminishing Returns

Case Study

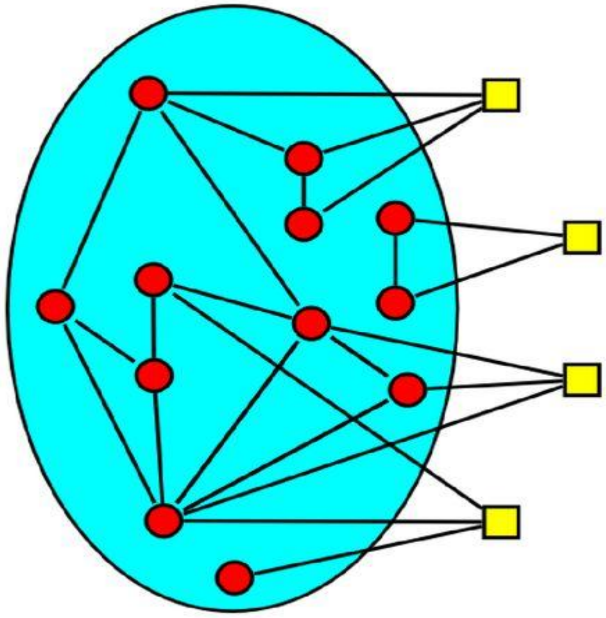
Adoption Curve: LiveJournal

Group memberships spread over the network:

- **Red** circles represent existing group members
- **Yellow** squares may join

Question

How does prob. of joining a group depend on the number of friends already in the group?



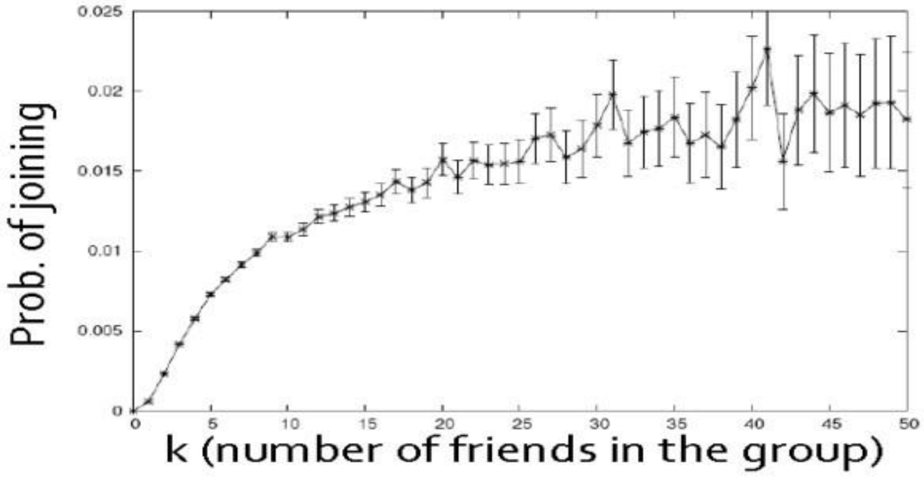
Backstrom, Lars, et al. "Group formation in large social networks: membership, growth, and evolution." ACM SIGKDD (2006).

Case Study

Adoption Curve: LiveJournal (cont'd)

Probability of joining increases with the number of friends in the group…

… but increases get smaller and smaller...



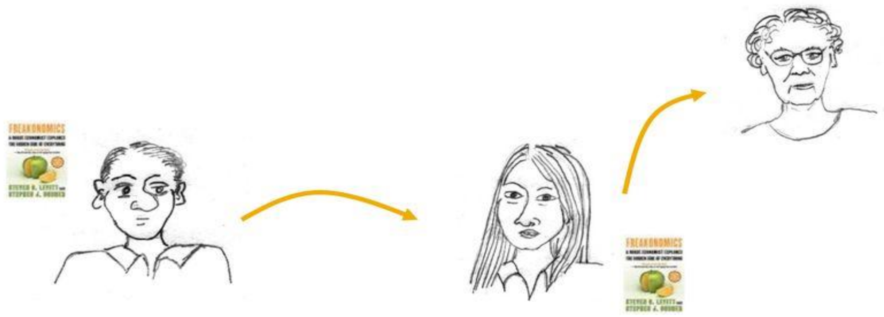
Case Study

Diffusion and Viral Marketing

Senders and followers of recommendations receive discounts on products

Data: Incentivated Viral Marketing program

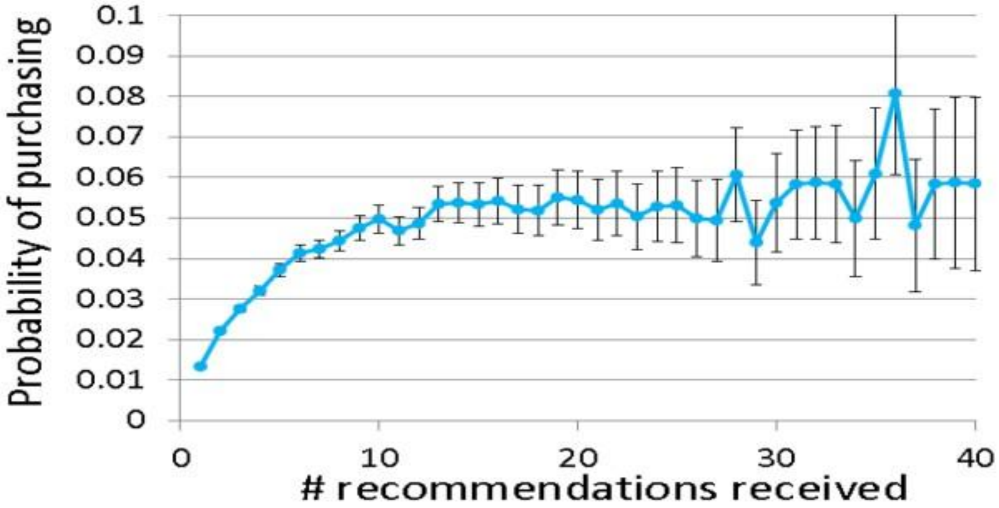
- 16 million recommendations
- 4 million people, 500k products



Leskovec, Jure, Lada A. Adamic, and Bernardo A. Huberman. "The dynamics of viral marketing." ACM Transactions on the Web (TWEB) 1.1 (2007): 5-es.

Case Study

Diffusion and Viral Marketing (cont'd)



Let's abstract a little bit...

Diminishing Returns' Law

(network effect)

There is a point where an increased level of inputs does not equal to an equal increase level of outputs.

In other words, after a certain point each input will not increase outputs at the same rate.

Influence Maximization

How to create big cascades?

Blogs - Information epidemics

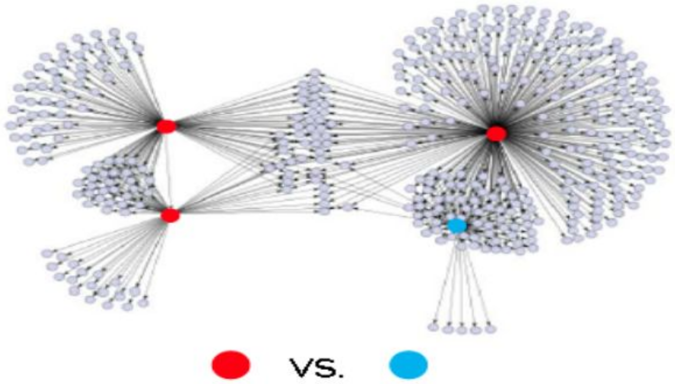
- which are the influential/infectious blogs?
- Which blogs create big cascades?

Viral Marketing

- Who are the influencers?
- Where should I advertise?

Disease spreading

- Where to place monitoring stations to detect epidemics?



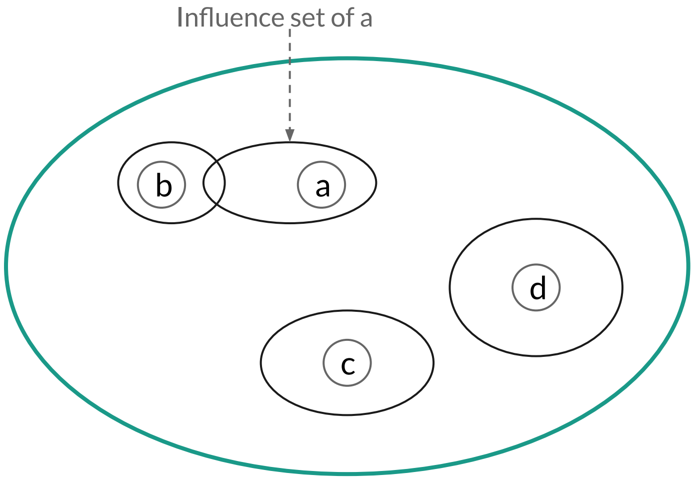
Most influential sets of nodes

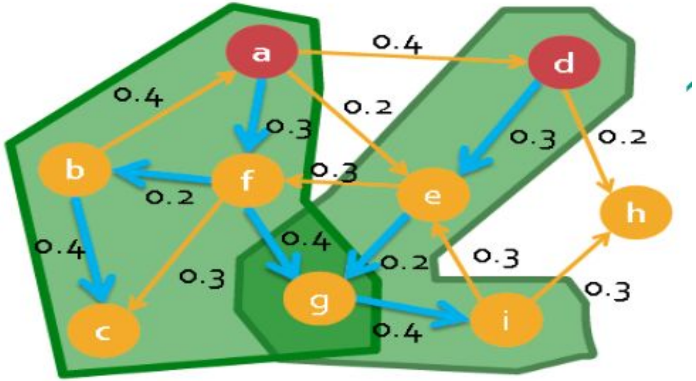
S: initial active set
f(S): expected size of final active set

S is more influential if f(S) is larger
e.g., f({a,b}) < f({a,c}) < f({a,d})

Problem:
Find the most influential set of size k

Namely, the set S of k nodes producing the largest expected cascade size f(S) if activated





How hard is the problem?

NP-HARD!

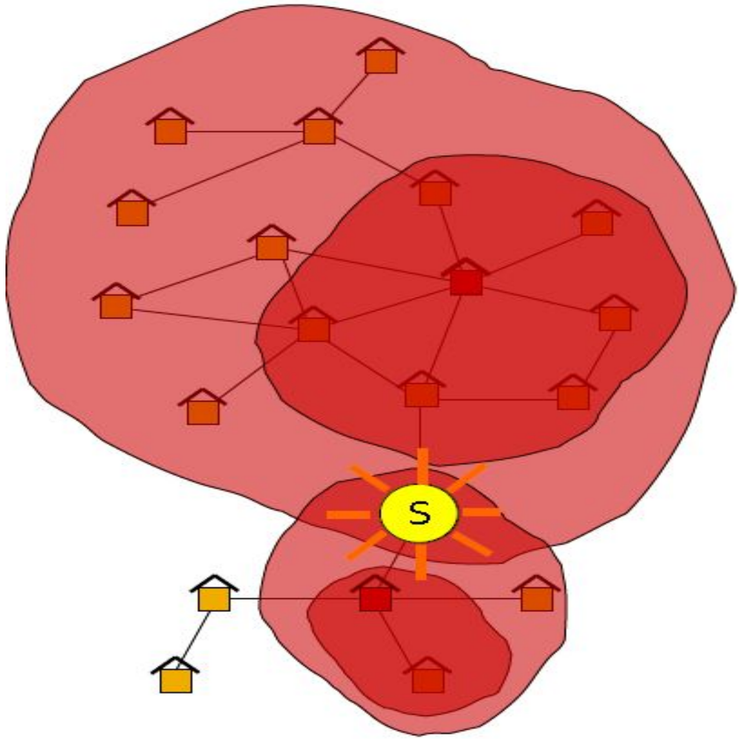
But if f(S) is «diminishing returns»

- Monotonic and submodular
- Then the approximated solution computed with a greedy algorithm (hill climbing) has a bounded distance with the global optimum!

Related problem: outbreak detection

Which node(s) initiated a cascade?

- Given a real city water distribution, and data on how contaminants spread in the network
- Detect the contaminant as quickly as possible



Advanced Topic:

Herding

Herding

Influence of actions of others

- Everyone sees everyone else’s behaviour
- Sequential decision making

Example: Picking a restaurant

- Consider you are choosing a restaurant in an unfamiliar town
- Based on Yelp reviews you intend to go to restaurant A
- When you arrive there is no one eating at A buth the next door restaurant B is nearly full

What will you do?

- Information that you can infer from other’s choices may be more powerful than your own

Herding

- There is a decision to be made
- People make the decision sequentially
- Each person has some private information that helps guide the decision
- You can't directly observe private information of the others but can see what they do

You can make inferences about the private information of others

Herding experiment

Consider an urn with 3 marbles. It can be either:

- Majority-blue: 2 blue, 1 red, or
- Majority-red: 1 blue, 2 red

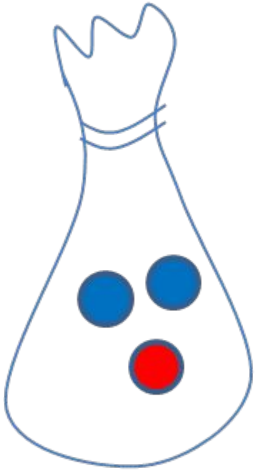
Each person wants to best guess whether the urn is majority-blue or red.

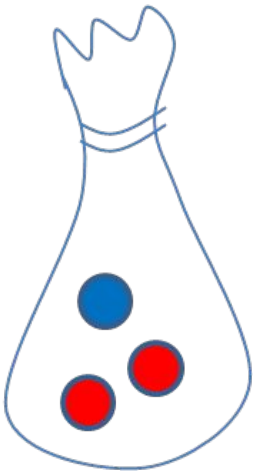
- Guess red if P(majority-red | what seen or heard) > ½

Experiment: one by one each person:

1. draw a marble
2. Privately looks at the color and puts the marble back
3. Publicly guesses whether the urn is majority-red or blue

You see all the guesses beforehand:
how should you make your guess?





Herding experiment

State of the world:

- whether the urn is MR or MB

Payoffs:

- utility of making a correct guess

Signals:

- Models private information
  (color you draw)
- Models public information
  (MR & MB guesses of people before you)

Decision: Guess MR if P(MR|past actions)>½

Analysis (Bayes rule):

- **#1** Follow her own color (private signal)

$$P(MR|\mathbf{r}] = \frac{P(MR)P(r|MR)}{P(r)} = \frac{1/2 \cdot 2/3}{1/2} = 2/3$$

$$P(r) = P(r|MB)P(MB) + P(r|MR)P(MR) = \frac{1}{2}\frac{1}{3} + \frac{1}{2}\frac{2}{3} = 1/2$$

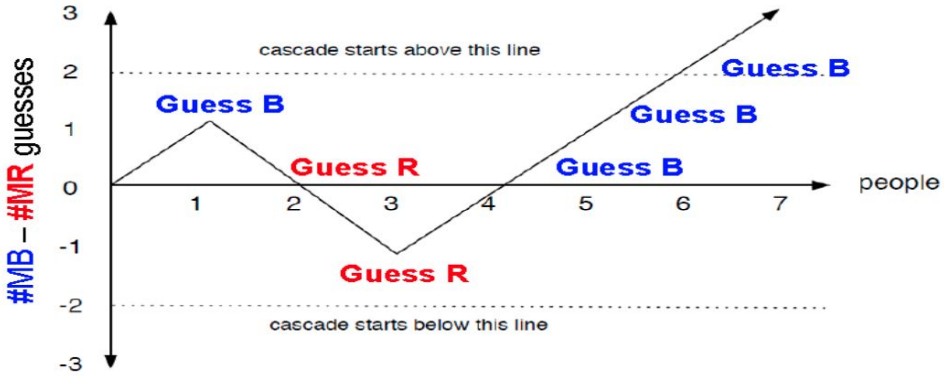
- **#2** Guesses her own color (private signal)
  #2 knows #1 revealed her color, so #2 gets 2 colors
  If they are the same, easy.
  If not, break the tie in favor of her own color
- **#3** gets 3 signals.
  If #1 and #2 are opposite, #3 goes with her own color.
  If #1 and #2 are equal, #3 decision conveyed no info.
  Cascade starts.

Herding experiment

How does this unfold?

You are the Nth person

- #MB = #MR: you guess your color
- |#MB - #MR| = 1: your color makes you indifferent, or reinforces your guess
- |#MB - #MR| ≥2: ignore your signal. Go with majority.



Cascade begins when the difference between the number of blue and red guesses reaches 2

Chapter 12

Conclusion

Take Away Messages

1. Diffusive processes can be modeled both in Mean Field and on network topologies
2. Social Contagion can be described using adoption Thresholds
3. Diminishing returns' law play a relevant role in social contagion
4. Influence maximization is a key issue

Suggested Readings

- Chapter 19 of Kleinberg's book
- Rogers, E. M. "*Diffusion of innovations*"
- Granovetter, M. "*Threshold models of collective behavior.*"

What's Next

Chapter 14:
Diffusion: Epidemics