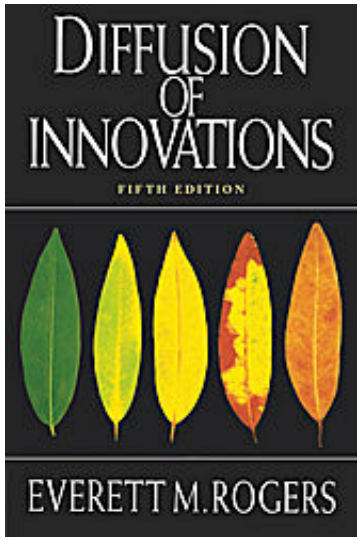
Social Network Analysis

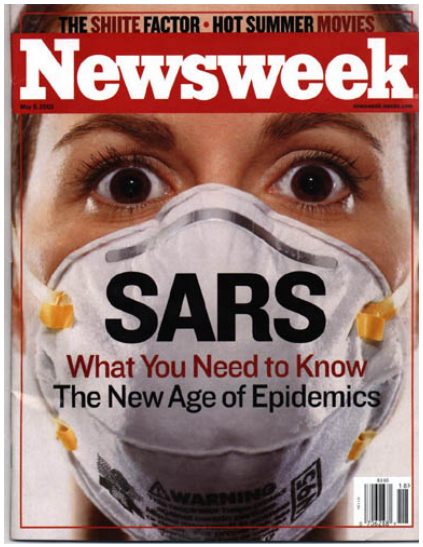
Dino Pedreschi
KDD LAB
Università di Pisa & ISTI-CNR

Diffusion, Spreading & Epidemics (1)

Epidemic spreading – Why?

Why is the spreading process important?





"Epidemic"

Epi + demos

upon *people*

Biological:

Airborne diseases (flu, SARS, …)

- Venereal diseases (HIV, …)
- Other infectious diseases including some cancers (HPV, …)
- Parasites (bedbugs, malaria, …)

Digital:

- Computer viruses, worms
- Mobile phone viruses

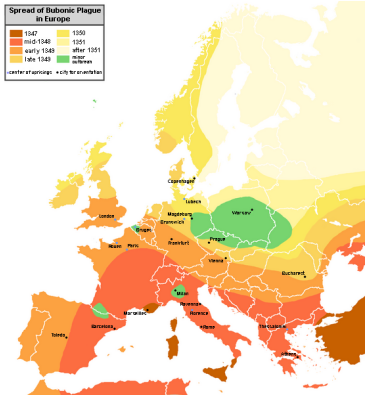
Conceptual/Intellectual:

- Diffusion of innovations
- Rumors
- Memes
- Business practices

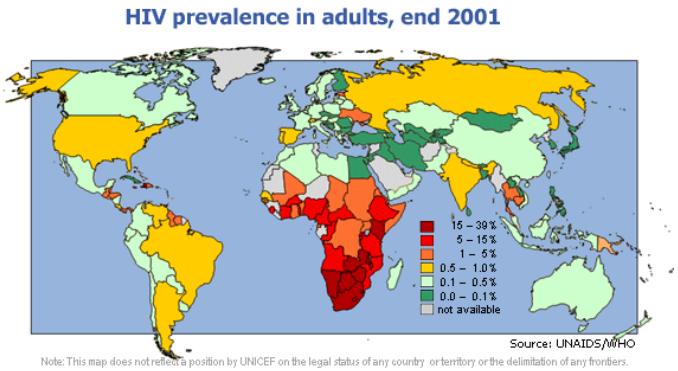
http://en.wikipedia.org/wiki/Epidemic

Biological: Notable Epidemic Outbreaks

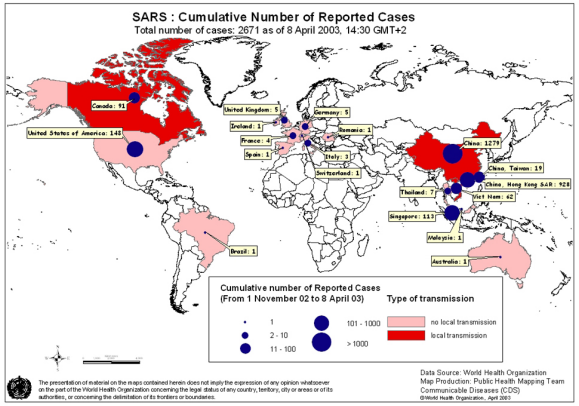
The Great Plague



HIV



SARS



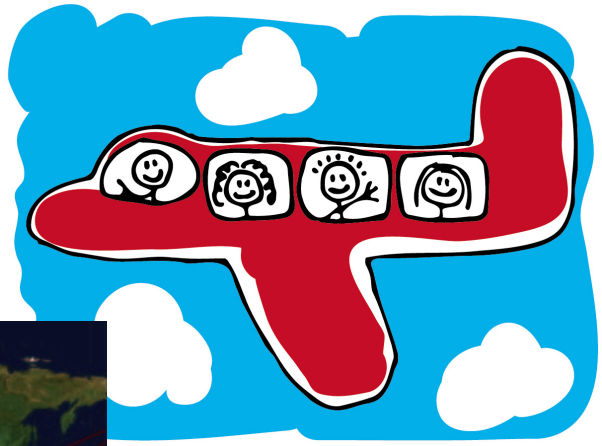
1918 Spanish flu

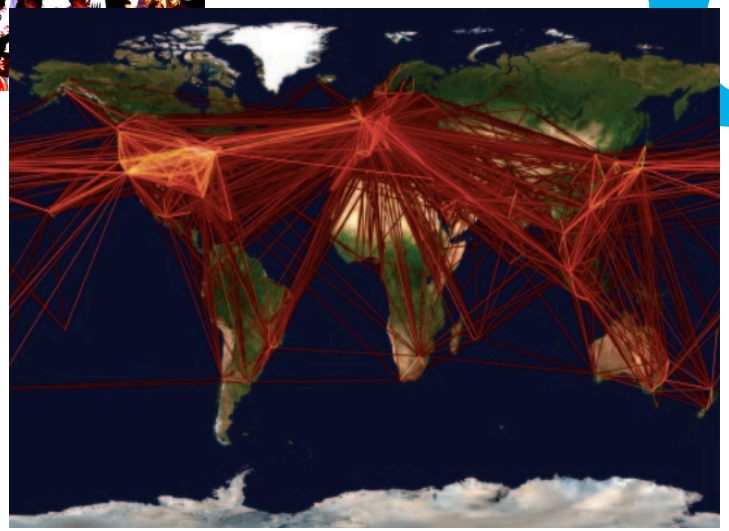
H1N1 flu

Epidemic spreading – Why does it matter now?

High population density

High mobility

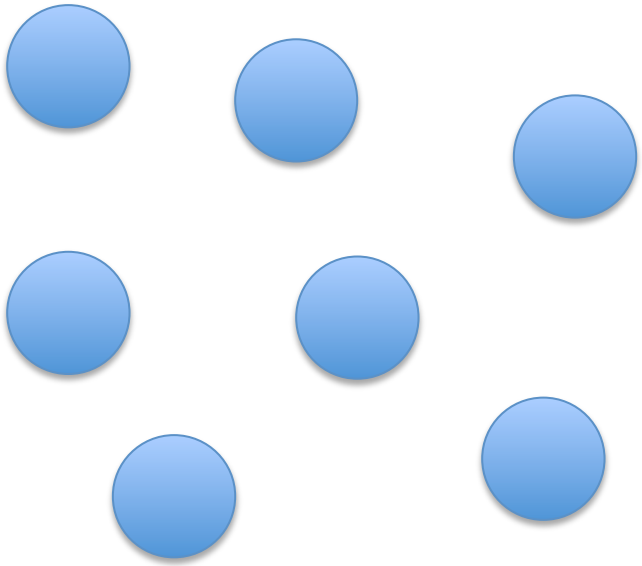




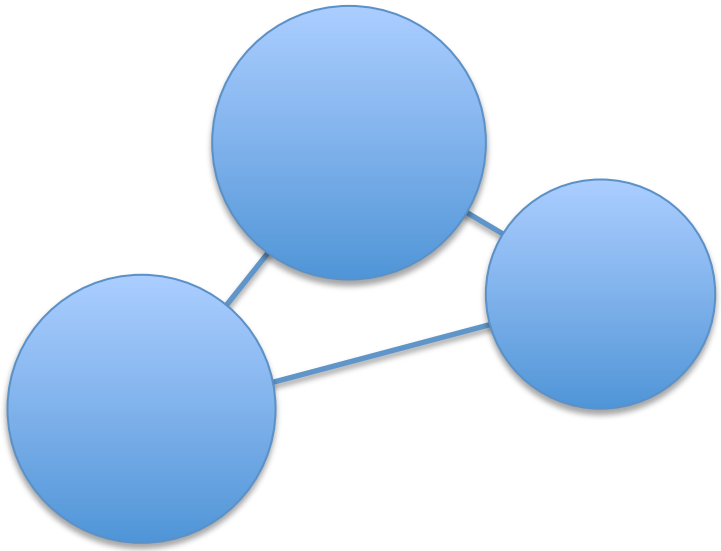
→ perfect conditions for epidemic spreading.

Airline figure: L. Hufnagel et al. *PNAS* **101,** 15124 (2004)

Large population can provide the “fuel”



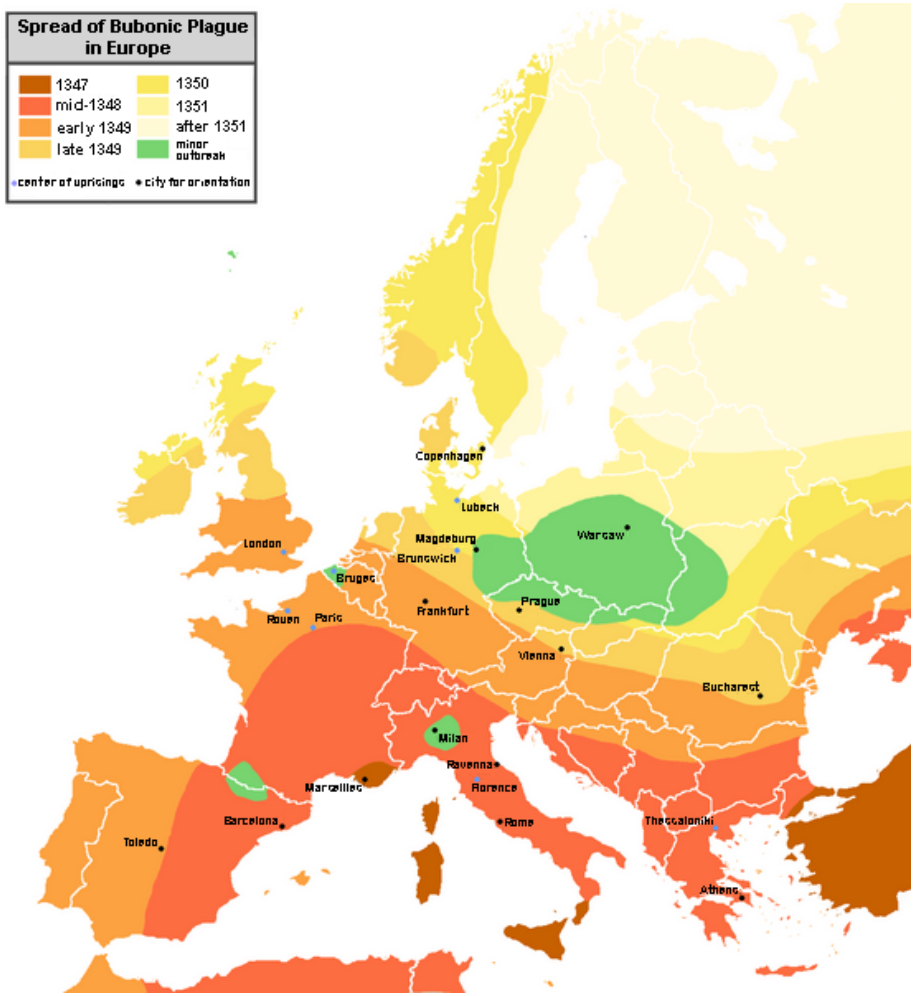
Separate, small population
(hunter-gatherer society, wild animals)



Connected, highly populated areas
(cities)

Human societies have “**crowd diseases**”, which are the consequences of large, interconnected populations (Measles, tuberculosis, smallpox, influenza, common cold, …)

14th Century – The Great Plague

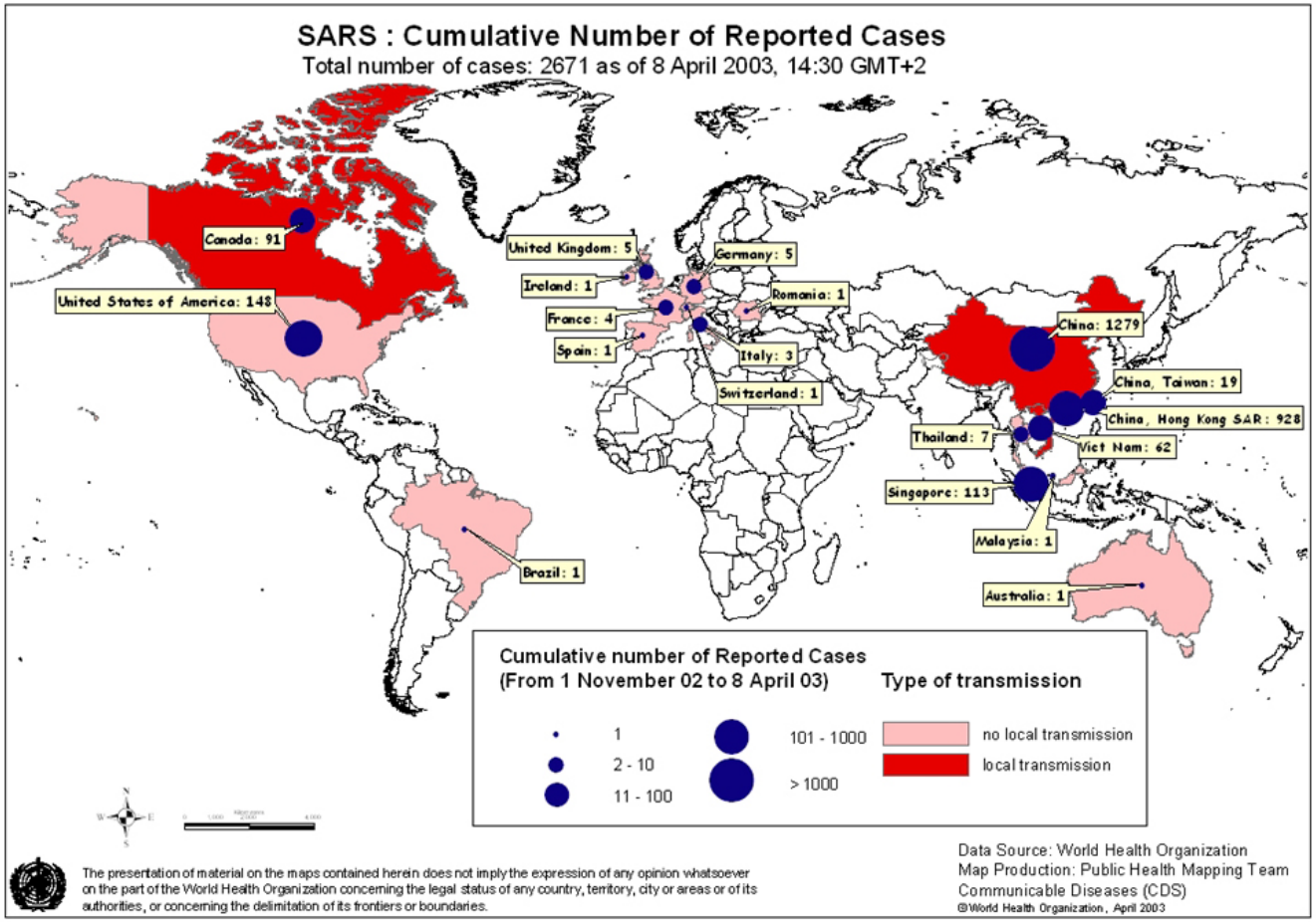


4 years from France to Sweden

Limited by the speed of human travel

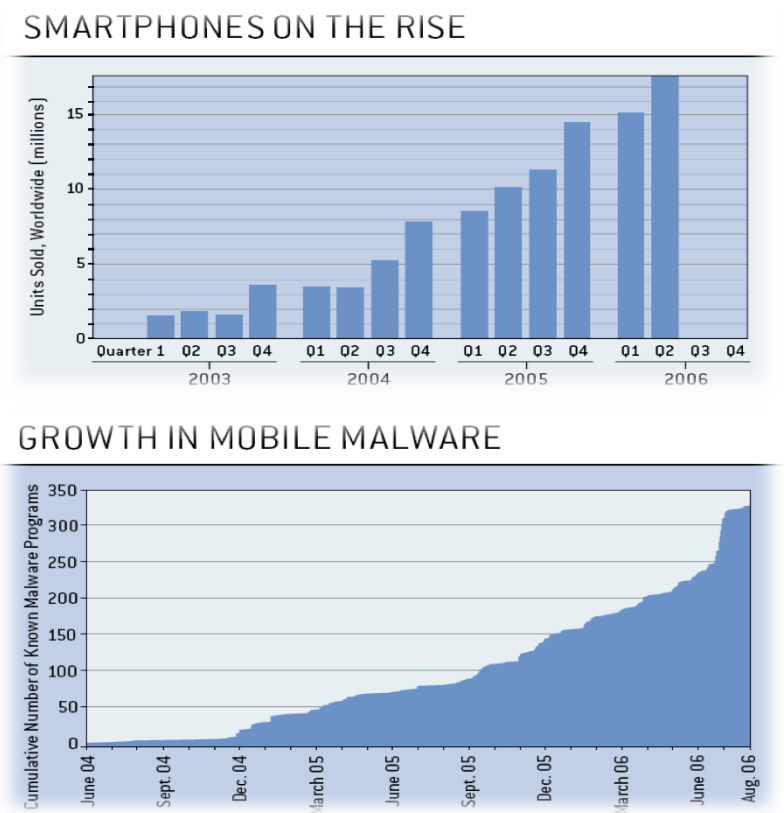
http://en.wikipedia.org/wiki/Black_Death
http://de.wikipedia.org/wiki/Schwarzer_Tod

21st Century – SARS



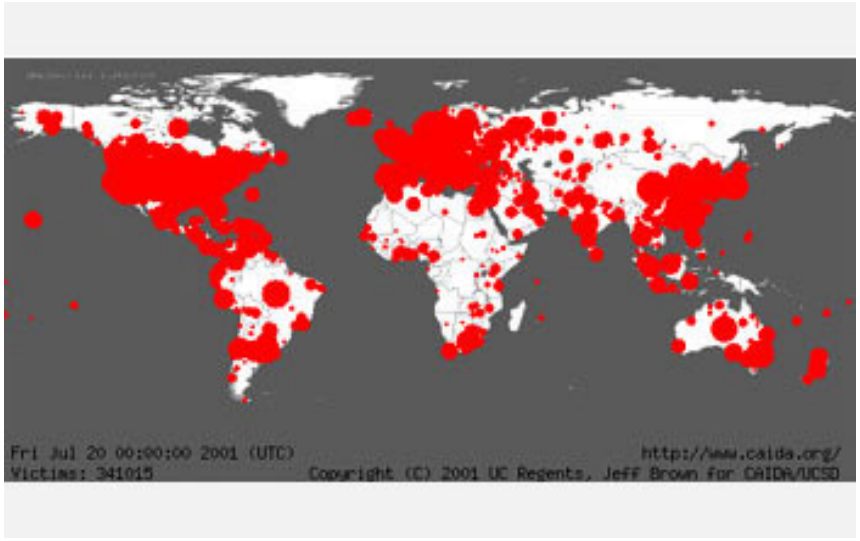
Source: World Health Organization

Computer Viruses, Worms, Mobile Phone Viruses



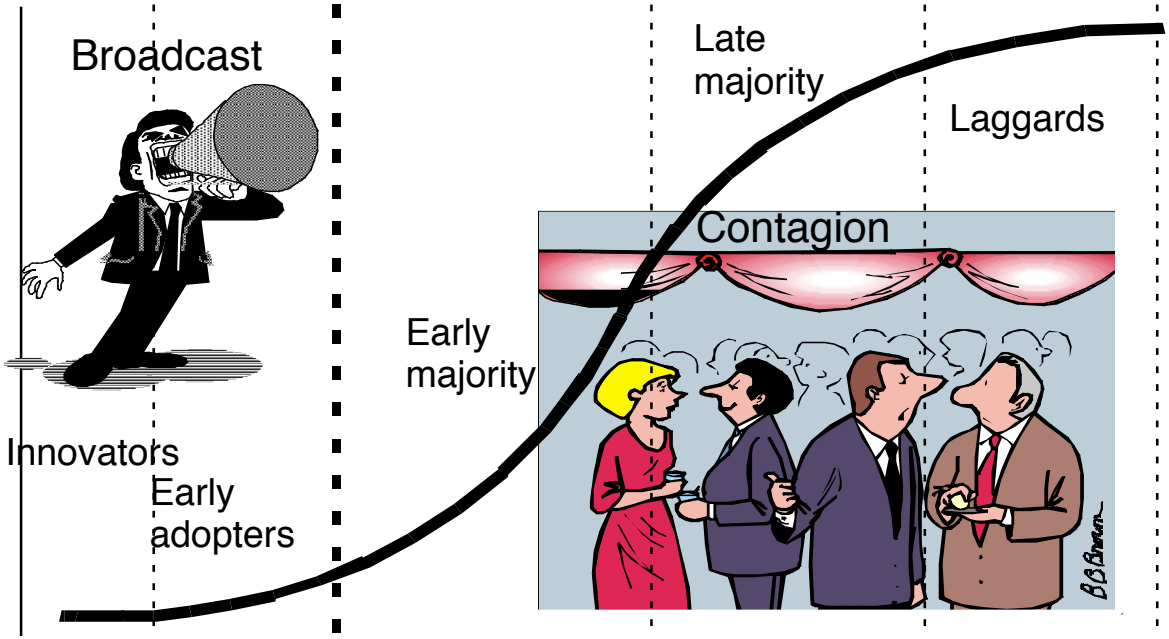
Hypponen M. *Scientific American* Nov. 70-77 (2006).

Code Red Worm paralyzed many countries' Internet

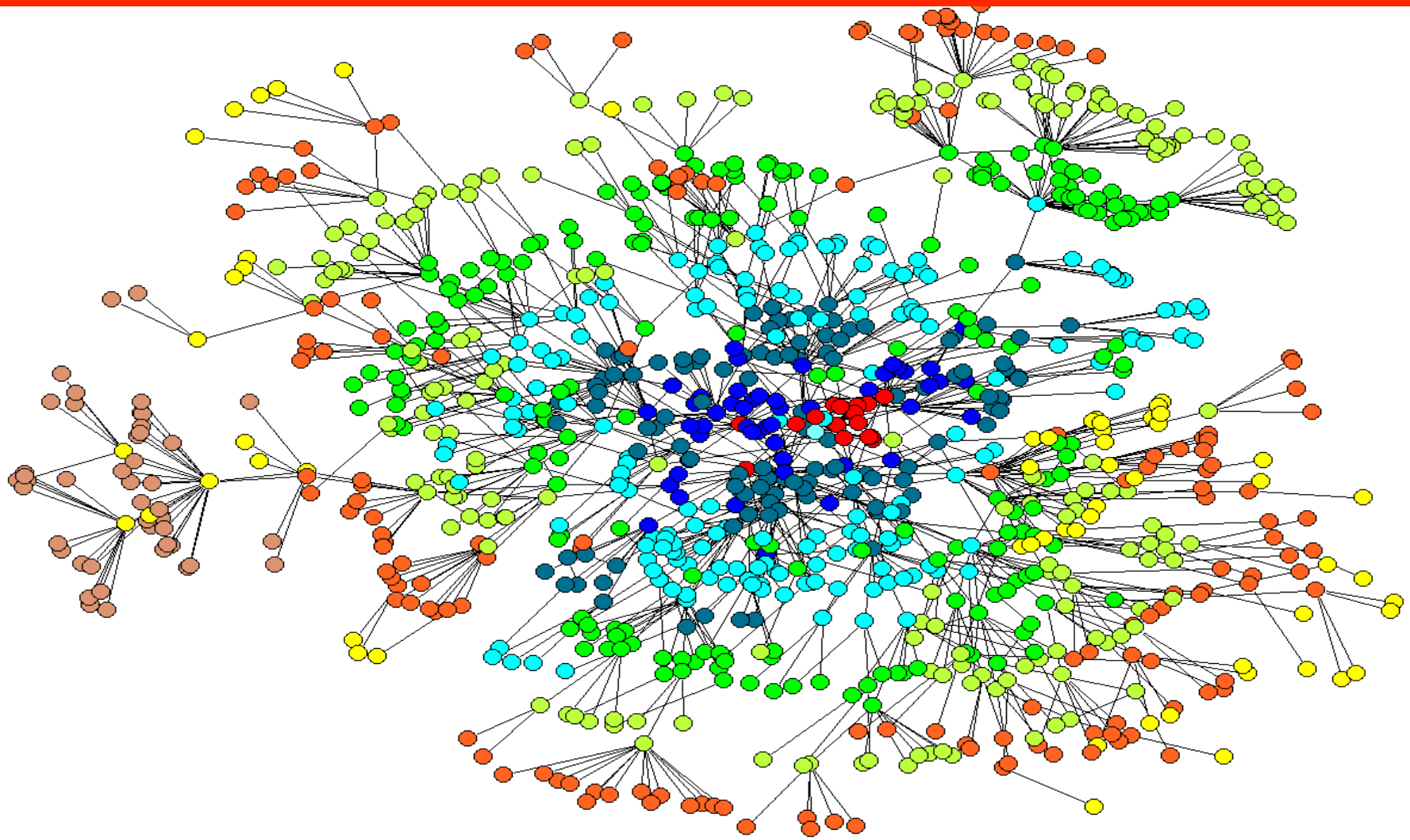


http://www.caida.org/publications/visualizations/

Diffusion of Innovation – The Adoption Curve



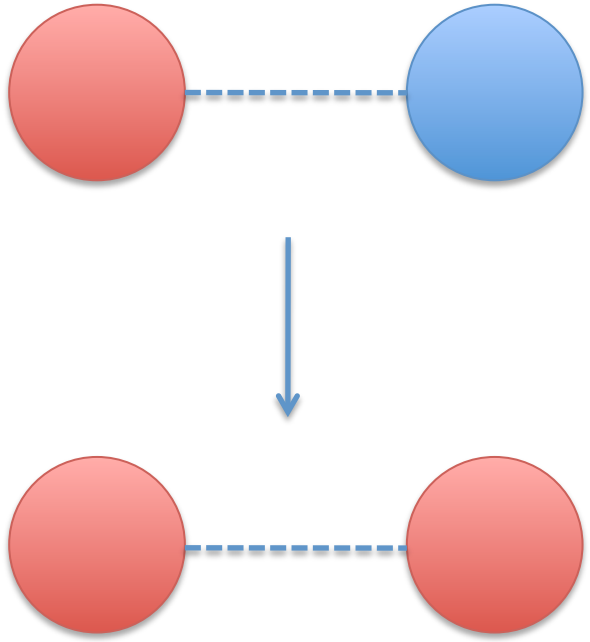
Information Spreading



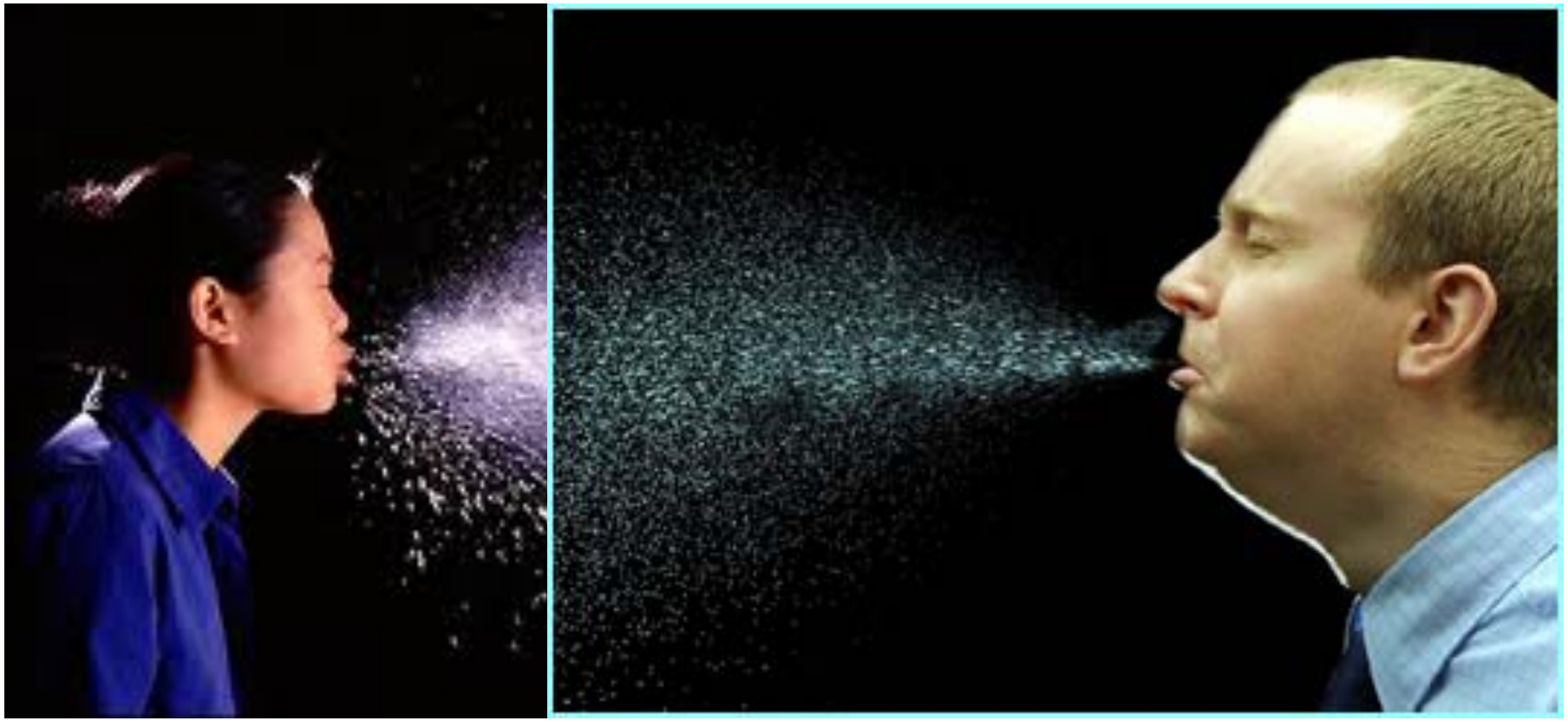
Epidemic Spreading – Network

- Epidemic spreading always implies network structure!

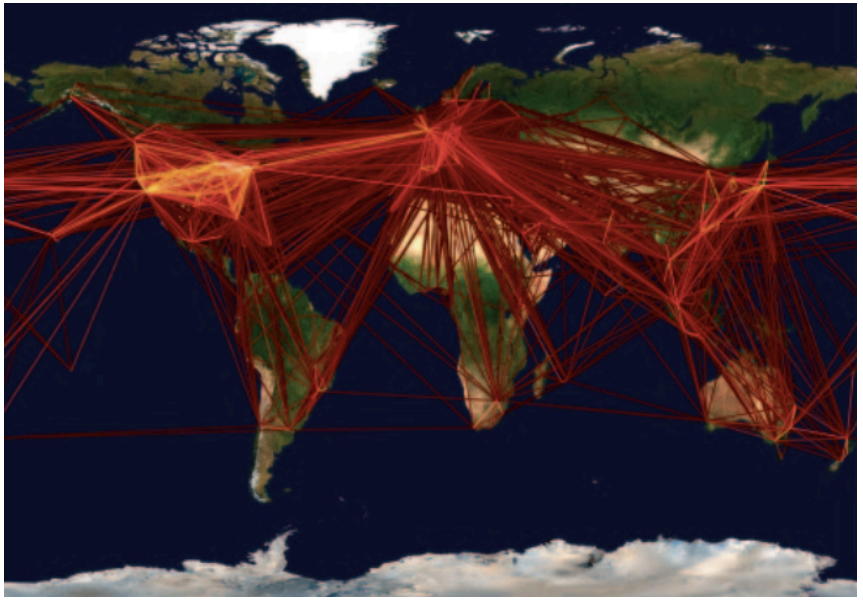
Spreading happens only when the carries of the diseases/virus/idea are **connected to each other**.



Epidemic Spreading – Network

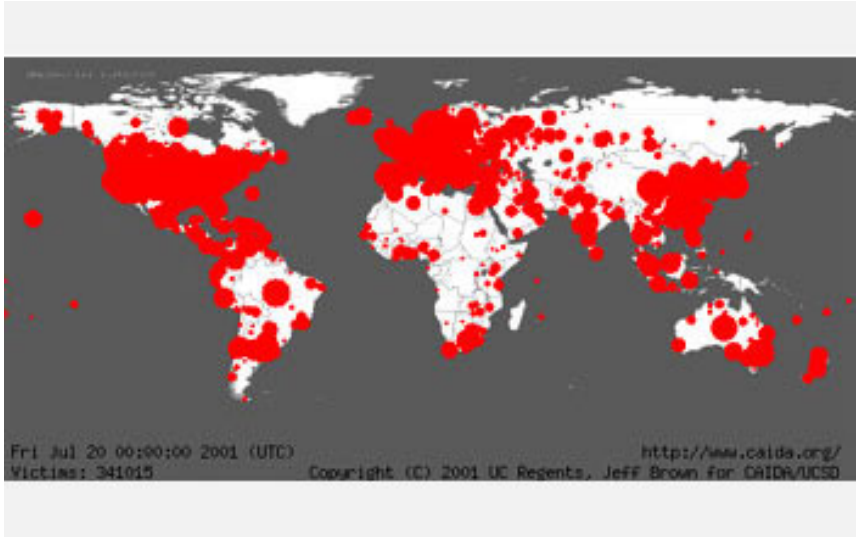


Epidemic Spreading – Network



The transportation network

L. Hufnagel et al. *PNAS* **101,** 15124 (2004)



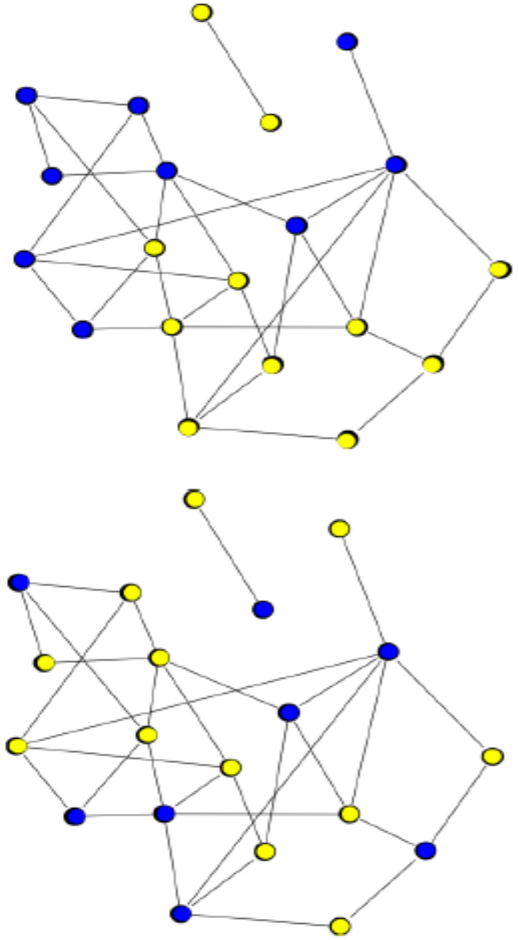
Internet

http://www.caida.org/publications/visualizations/

Group formation dynamics

Group formation in networks

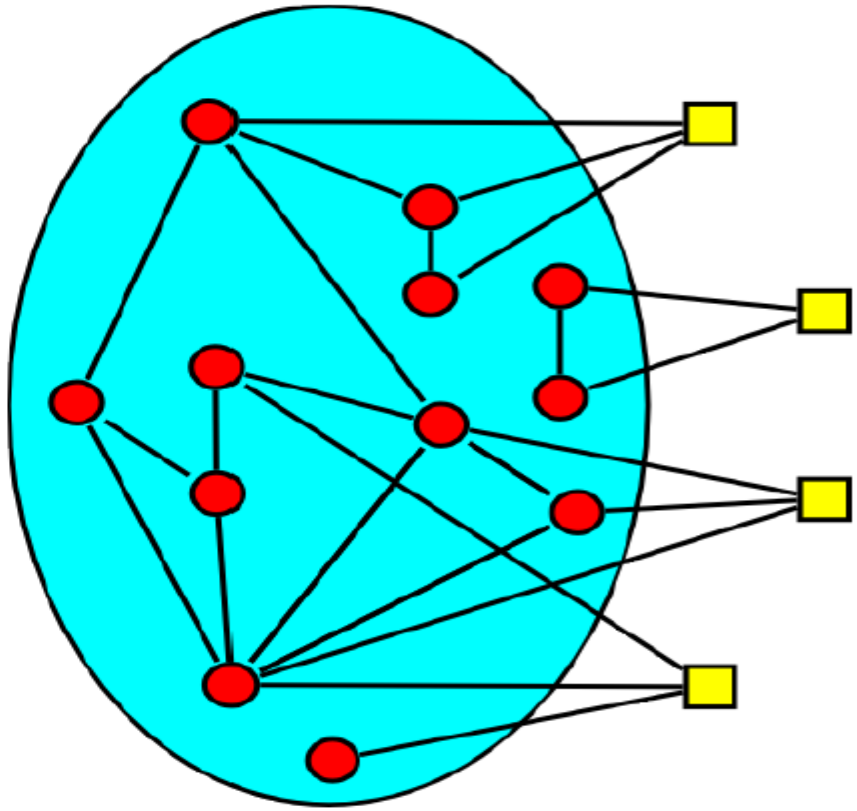
- In a social network **nodes explicitly declare group membership:**
  - Facebook groups, Publication venue
- Can think of groups as **node colors**
- Gives **insights into social dynamics:**
  - Recruits friends? Memberships spread along edges
  - Doesn't recruit? Spread randomly
- **What factors influence a person's decision to join a group?**

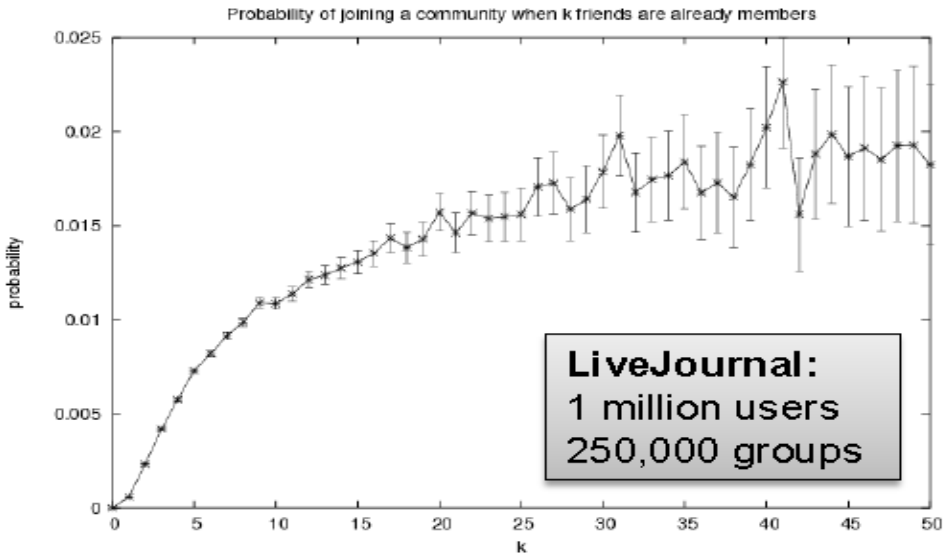


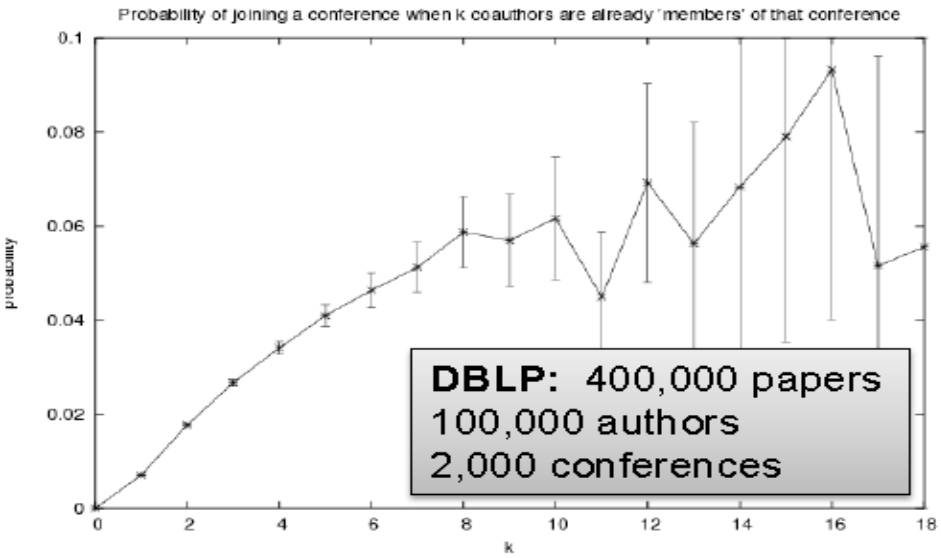
Group memberships spread over the network:

- 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?



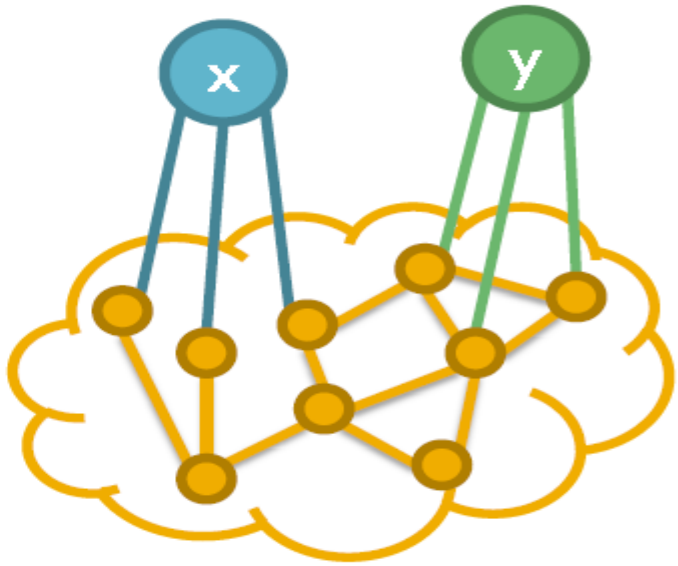




- **Diminishing returns:**
  - Probability of joining increases with the number of friends in the group
  - But increases get smaller and smaller

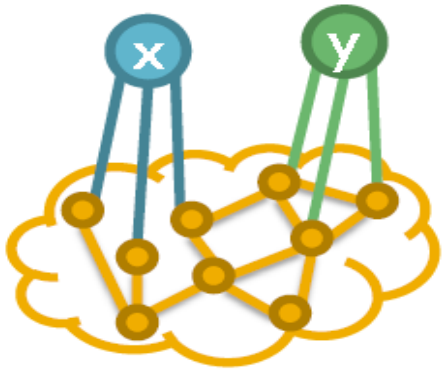
Connectedness of friends and group membership

- *x* and *y* have three friends in the group
- *x*'s friends are independent
- *y*'s friends are all connected



Who is more likely to join?

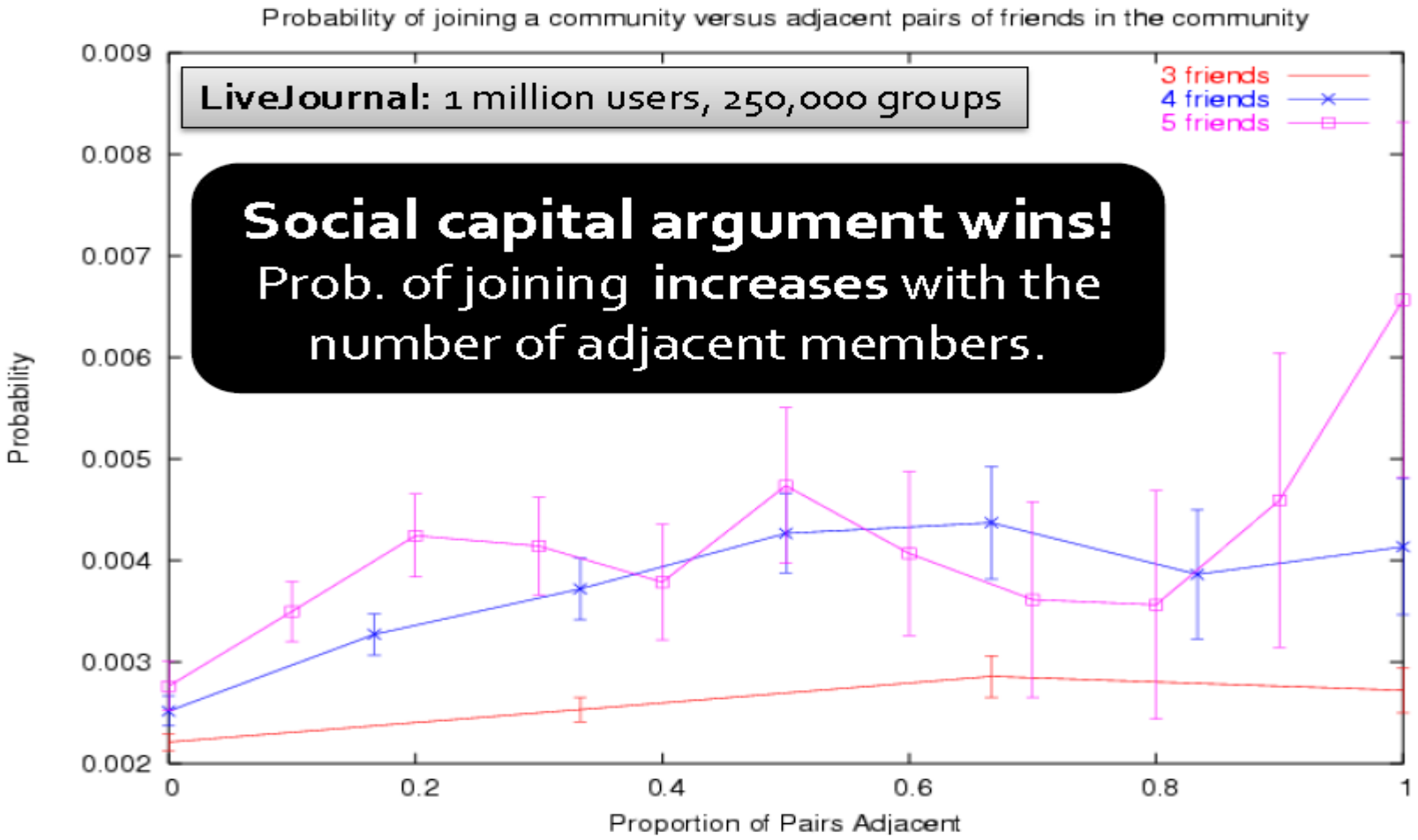
- **Competing sociological theories:**
  - Information argument [Granovetter '73]
  - Social capital argument [Coleman '88]

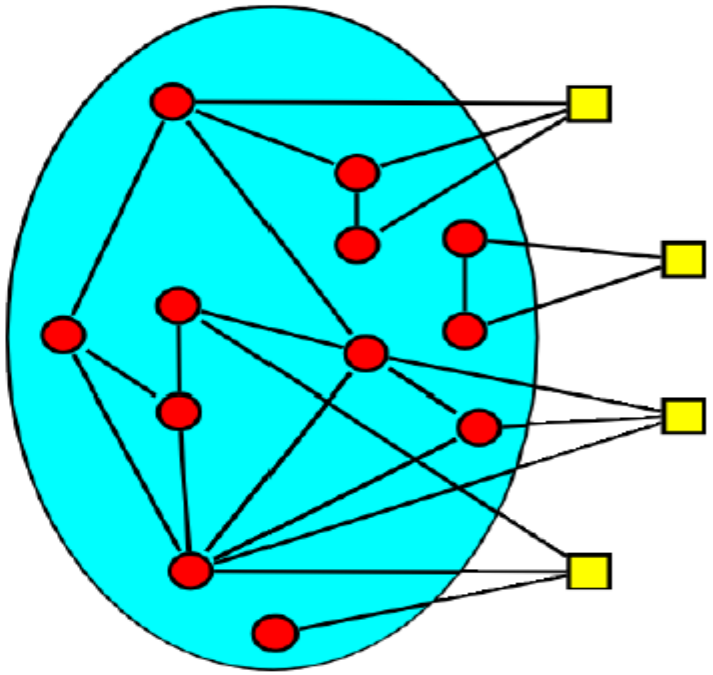


- **Information argument:**
  - Unconnected friends give independent support
- **Social capital argument:**
  - Safety/trust advantage in having friends who know each other

… and the winner is …

[Backstrom et al., KDD 2006]



The strength of **strong** ties

- **A person is more likely to join a group if**
  - she has more friends who are already in the group
  - friends have more connections between themselves
- **So, groups form clusters of tightly connected nodes**