Lecture VI
Introduction to complex networks

Santo Fortunato

INSTITUTE
FOR SCIENTIFIC INTERCHANGE
FOUNDATION

Plan of the course

I. Networks: definitions, characteristics, basic concepts in graph theory
II. Real world networks: basic properties
III. Models
IV. Community structure I
V. Community structure II
VI. **Dynamic processes in networks**

Dynamic processes on networks

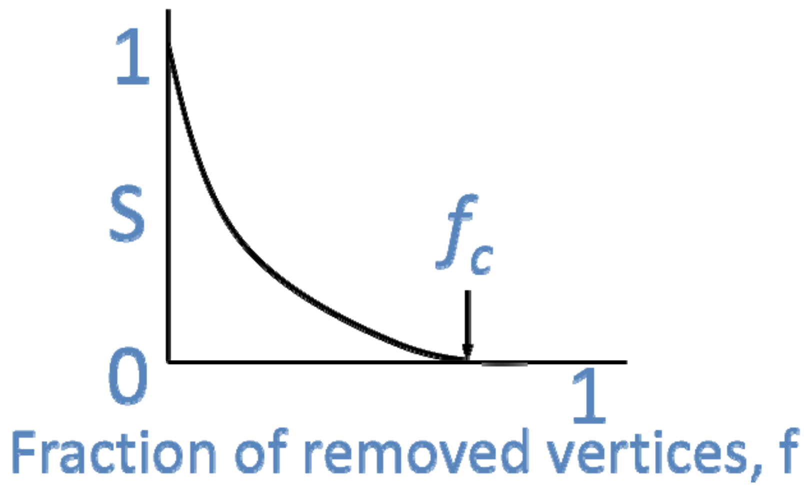
Many social networks are the support of some dynamical processes

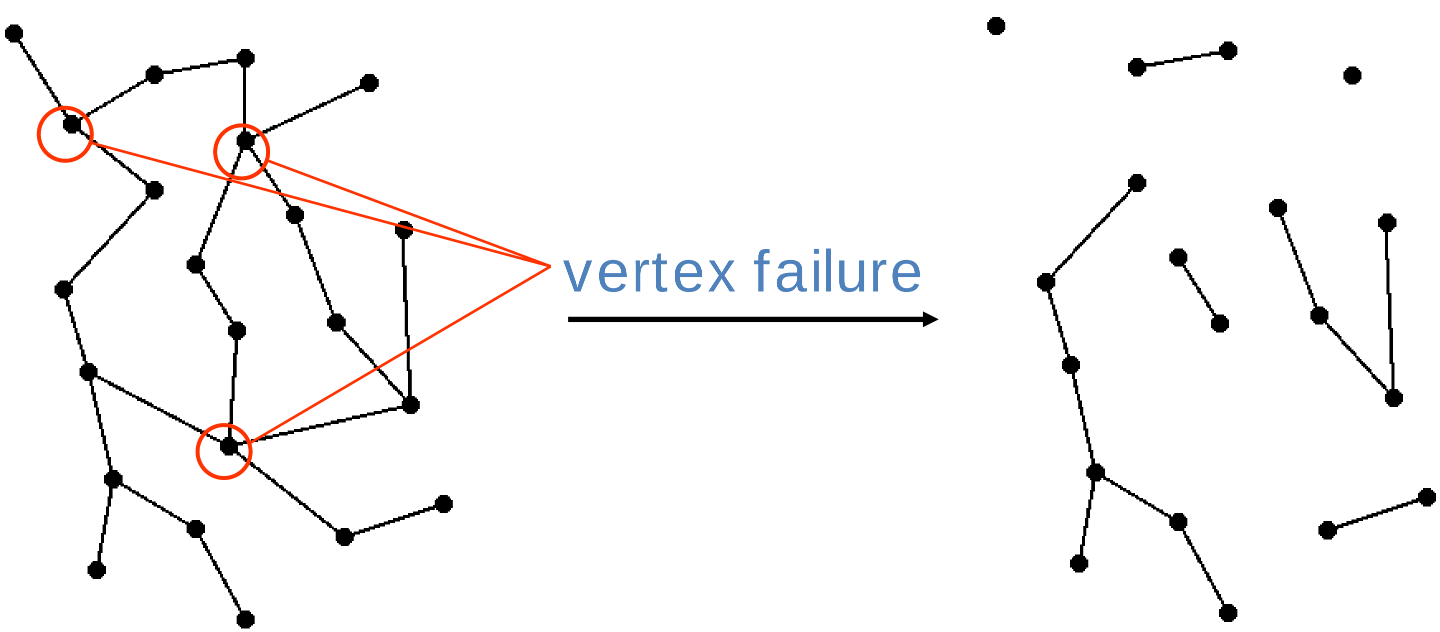
- Percolation -> Robustness/resilience
- Epidemics
- Opinion/consensus formation
- Search
- Navigation
- Cooperative phenomena…

Robustness

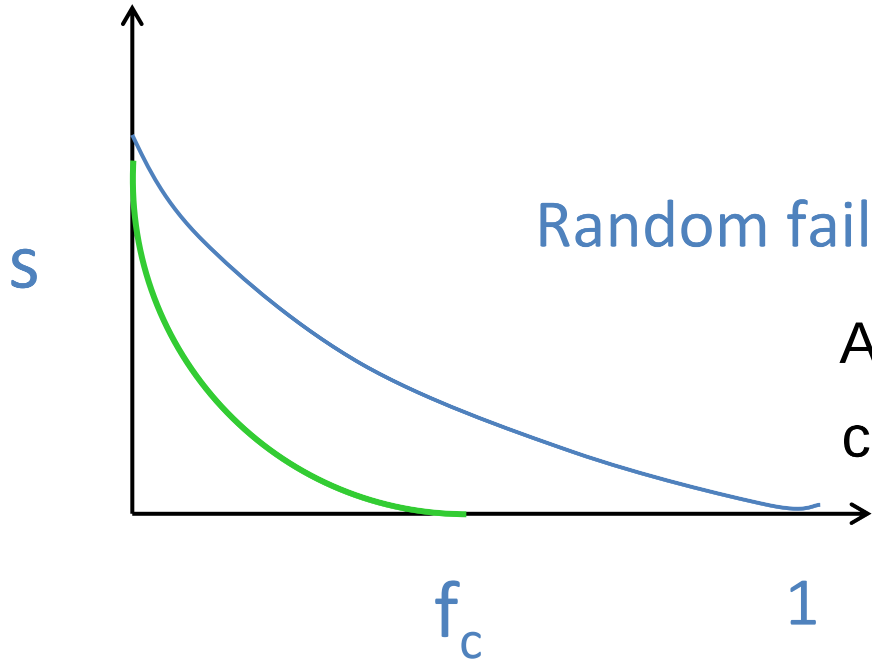
Complex systems maintain their basic functions
even under errors and failures
(cell $\rightarrow$ mutations; Internet $\rightarrow$ router breakdowns)

S: fraction of giant component





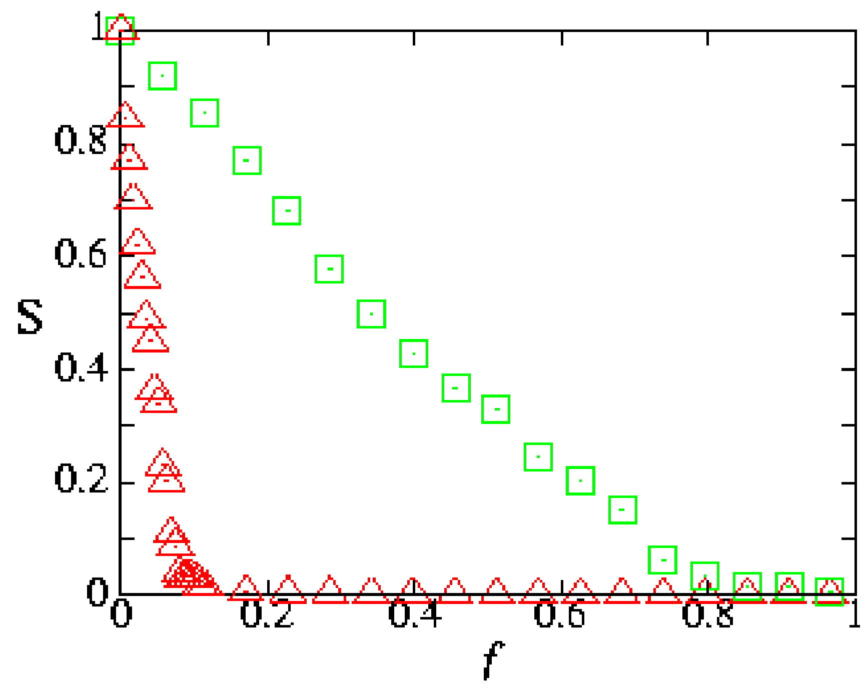
Case of Scale-free Networks



Random failure $f_c = 1$ $(2 < \gamma \leq 3)$

Attack = progressive failure of the most connected vertices $f_c < 1$

Internet



R. Albert, H. Jeong, A.L. Barabási, Nature **406**, 378 (2000)

Modelling resilience in networks: random attacks

Configuration model, degree distribution P(k)

A fraction 1-q of the vertices is removed at random

Probability that a vertex with degree k has k' neighbors left:

$$p_{kk'} = \binom{k}{k'} q^{k'} (1-q)^{k-k'}$$

Degree distribution of graph after random attack:

$$P(k') = \sum_{k=k'}^{\infty} P(k) p_{kk'} = \sum_{k=k'}^{\infty} P(k) \binom{k}{k'} q^{k'} (1-q)^{k-k'}$$

R. Cohen, D. ben-Avraham, S. Havlin, Phys. Rev. Lett. **85**, 4626 (2000)

Modelling resilience in networks: random attacks

$$P(k') = \sum_{k=k'}^{\infty} P(k) p_{kk'} = \sum_{k=k'}^{\infty} P(k) \binom{k}{k'} q^{k'} (1-q)^{k-k'}$$

Molloy-Reed criterion for the existence of a giant component:

$$\sum_{k} k(k-2)P(k) = 0$$

$$\frac{\langle k^2 \rangle}{\langle k \rangle} q_c + 1 - q_c = 2 \qquad \longrightarrow \qquad q_c = \frac{1}{\langle k^2 \rangle / \langle k \rangle - 1}$$

R. Cohen, D. ben-Avraham, S. Havlin, Phys. Rev. Lett. **85**, 4626 (2000)

Modelling resilience in networks: random attacks

Special case

Erdös-Rényi random graphs: $$P(k) = e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!}$$

$$\frac{\langle k^2 \rangle}{\langle k \rangle} = \langle k \rangle + 1 \quad \rightarrow \quad q_c = \frac{1}{\langle k \rangle}$$

Threshold for giant component: $\langle k \rangle \geq 1$

If <k> >1 one has to remove a macroscopic fraction of vertices to destroy the giant component!

R. Cohen, D. ben-Avraham, S. Havlin, Phys. Rev. Lett. **85**, 4626 (2000)

Modelling resilience in networks: random attacks

Special case

Scale-free random graphs: $P(k) \sim k^{-\alpha}, \quad k \in [k_{min}, k_{max}]$

Threshold for giant component:

$$\frac{\langle k^2 \rangle}{\langle k \rangle} \to \left| \frac{2-\alpha}{3-\alpha} \right| \times \begin{cases} k_{min}, & \text{if } \alpha > 3; \\ k_{min}^{\alpha-2} k_{max}^{3-\alpha}, & \text{if } 2 < \alpha < 3; \\ k_{max}, & \text{if } 1 < \alpha < 2. \end{cases}$$

- If $\alpha > 3$, non-zero threshold!
- If $\alpha \leq 3$, zero threshold: there is always a giant component, no matter how many vertices are removed (super-robustness)

R. Cohen, D. ben-Avraham, S. Havlin, Phys. Rev. Lett. **85**, 4626 (2000)

Modelling resilience in networks: random attacks

Alternative approach

Configuration model, degree distribution P(k)

A fraction 1-q_k of the vertices with degree k is removed at random

Generating functions

$$F_0(x) = \sum_{k=0}^{\infty} P(k) q_k x^k, \qquad F_1(x) = \frac{\sum_k k P(k) q_k x^{k-1}}{\sum_k k P(k)}$$

Mean component size

$$\langle s \rangle = F_0(1) + \frac{F_0'(1) F_1(1)}{1 - F_1'(1)}$$

D. S. Callaway, M. E. J. Newman, S. H. Strogatz, D. J. Watts,
Phys. Rev. Lett. **85**, 5468 (2000)

Modelling resilience in networks: random attacks

Alternative approach

Relative size of giant component

$$S = F_0(1) - F_0(u), \qquad u = 1 - F_1(1) + F_1(u)$$

If q_k=q ➔ $q_c = \frac{1}{G_1'(1)}$ $\quad G_1(x) = \sum_k \frac{(k+1)P(k+1)x^k}{\langle k \rangle}$

D. S. Callaway, M. E. J. Newman, S. H. Strogatz, D. J. Watts, Phys. Rev. Lett. **85**, 5468 (2000)

Modelling resilience in networks: random attacks

Special case: scale-free networks

$$P(k) = \begin{cases} 0 & \text{for } k = 0 \\ k^{-\alpha}/\zeta(\alpha) & \text{for } k \geq 1 \end{cases}$$

$$q_c = \frac{\zeta(\alpha - 1)}{\zeta(\alpha - 2) - \zeta(\alpha - 1)}$$

- If $\alpha \leq 3$, q_c is zero or negative (unphysical!): there is always a giant component, no matter how many vertices are removed (super-robustness)
- If $3 \leq \alpha \leq 3.4788...$, $0 \leq q_c \leq 1$ ("normal" robustness)
- If $\alpha \geq 3.4788...$, $q_c \geq 1$ (unphysical) there is no giant component from the start

D. S. Callaway, M. E. J. Newman, S. H. Strogatz, D. J. Watts,
Phys. Rev. Lett. **85**, 5468 (2000)

Modelling resilience in networks: random attacks

Advantage of approach by Callaway et al.: vertices can be removed in the order of (decreasing or increasing) degree!

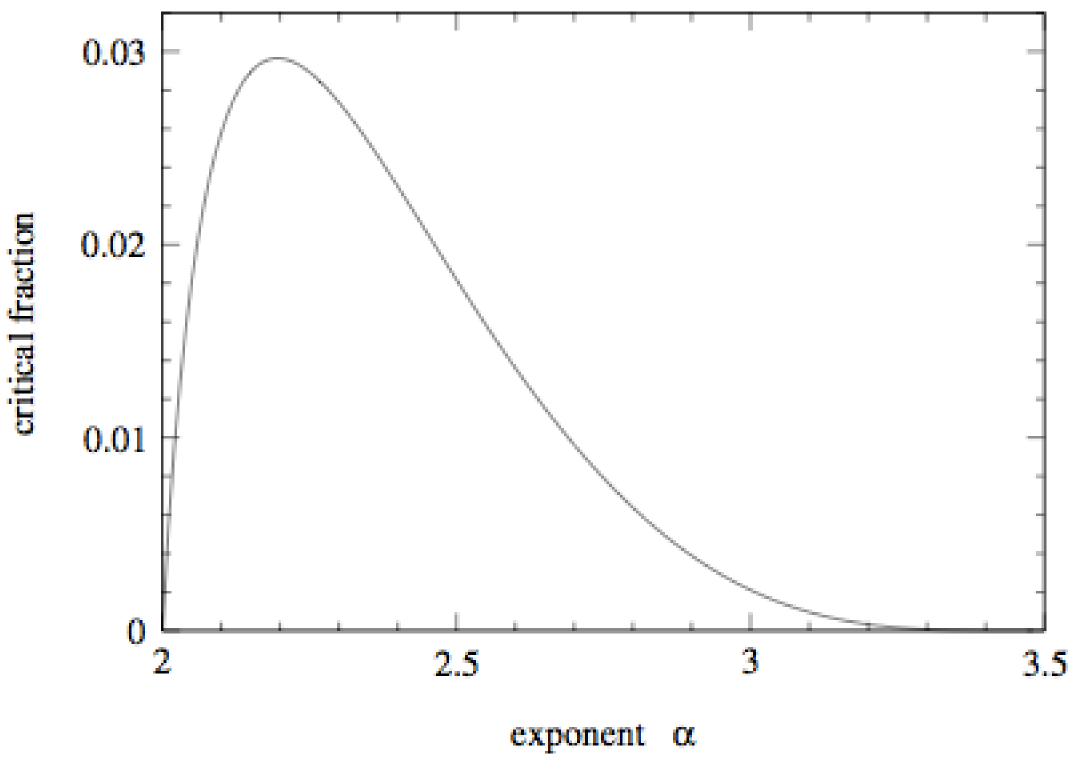
Example: $$q_k = \theta(k_{\mathbf{max}} - k)$$

All vertices with degree larger than k_{max} are removed!

Equations have to be solved numerically!

D. S. Callaway, M. E. J. Newman, S. H. Strogatz, D. J. Watts,
Phys. Rev. Lett. **85**, 5468 (2000)

Modelling resilience in networks: random attacks



D. S. Callaway, M. E. J. Newman, S. H. Strogatz, D. J. Watts, Phys. Rev. Lett. **85**, 5468 (2000)

Cascading failures in networks

Watts' model

Motivation: binary decisions with externalities

How it works:

- Random graph with given degree distribution (configuration model)
- A vertex i fails if a fraction Φ_i of its neighbors also fails
- The quantities $\{\Phi_i\}$ are taken from a distribution $f(\Phi)$
- Initially a fraction φ_0 of vertices, randomly selected, fail

D. J. Watts, Proc. Natl. Acad. Sci. USA **99**, 5766 (2002)

Cascading failures in networks

Watts' model

How to solve it:

- If $\Phi_0 \ll 1$ failed vertices are initially isolated (approximately)
- If a vertex i has only one failed neighbor, it will fail only if $\Phi_i<1/k_i$, where k_i is the degree of i (vulnerable vertices)
- The probability of a vertex with degree k being vulnerable is

$$q_k = \int_0^{1/k} f(\phi) d\phi$$

Through the generating function approach by Callaway et al. It is possible to find the condition for the existence of a giant component of vulnerable vertices

The computation of the full size of the cascades is done numerically

D. J. Watts, Proc. Natl. Acad. Sci. USA **99**, 5766 (2002)

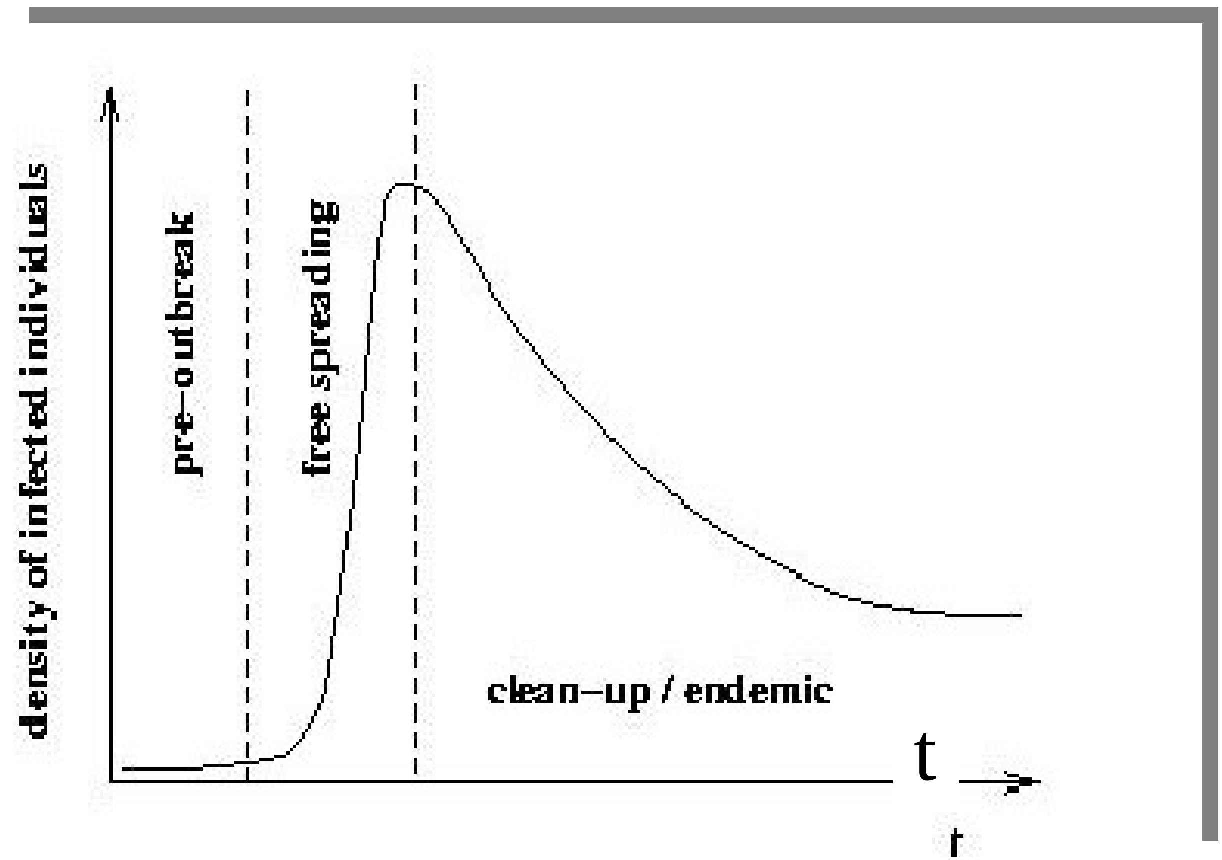
Epidemiology

Two levels:

- Microscopic: researchers try to disassemble and kill new viruses => quest for vaccines and medicines

- Macroscopic: statistical analysis and modeling of epidemiological data in order to find information and policies aimed at lowering epidemic outbreaks => macroscopic prophylaxis, vaccination campaigns...

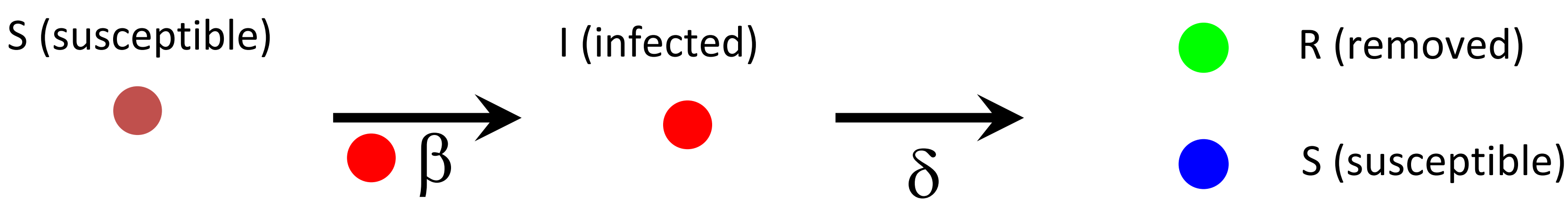
Stages of an epidemic outbreak



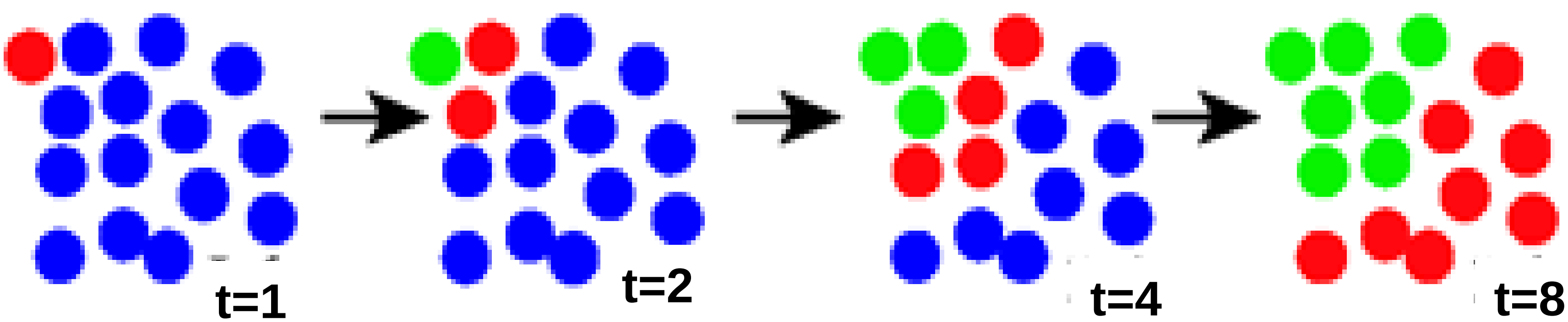
Infected individuals => prevalence/incidence

Standard epidemic modeling

Compartments: S, I, R... (for, e.g., influenza)



Homogeneous mixing assumption (mean-field)



The SIR model

W. O. Kermack, A. G. MacKendrick, Proc. Roy. Soc. Lond. A **115**, 700 (1927)

s = fraction of susceptible agents

i = fraction of infective agents

r = fraction of recovered agents

<k> = number of contacts in the unit time

Mean field equations:

$$\frac{ds}{dt} = -\beta < k > is, \quad \frac{di}{dt} = \beta < k > is - \gamma i, \quad \frac{dr}{dt} = \gamma i$$

There is an epidemic threshold!

Initial conditions: $s(0) \sim 1, \quad i(0) \sim 0, \quad r(0) \sim 0$

$$\frac{ds}{dr} = - < k > \frac{\beta}{\gamma} s = - < k > \lambda s \qquad \lambda = \frac{\beta}{\gamma}$$

The SIR model

Solution:

$$s(t) = e^{-<k>\lambda r(t)}$$

Large t limit
(i ➔ 0):

$$r_\infty = 1 - e^{-\lambda <k> r_\infty}$$

$$r_\infty = 0 \quad \text{always solution}$$

Condition for the existence of non-zero solution:

$$\frac{d}{dr_\infty}\left(1 - e^{-\lambda <k> r_\infty}\right)\Big|_{r_\infty = 0} > 1$$

$$\lambda > \lambda_c = \frac{1}{<k>}$$

The Susceptible-Infected-Susceptible (SIS) model

Model for non-immunizing diseases (e.g., tuberculosis, computer viruses)

- Each node is infected with rate β if connected to one or more infected nodes
- Infected nodes are recovered (cured) with rate γ
- Effective spreading rate: $\lambda=\beta/\gamma$

$$\frac{ds}{dt} = -\beta <k> is + \gamma i, \quad \frac{di}{dt} = \beta <k> is - \gamma i$$

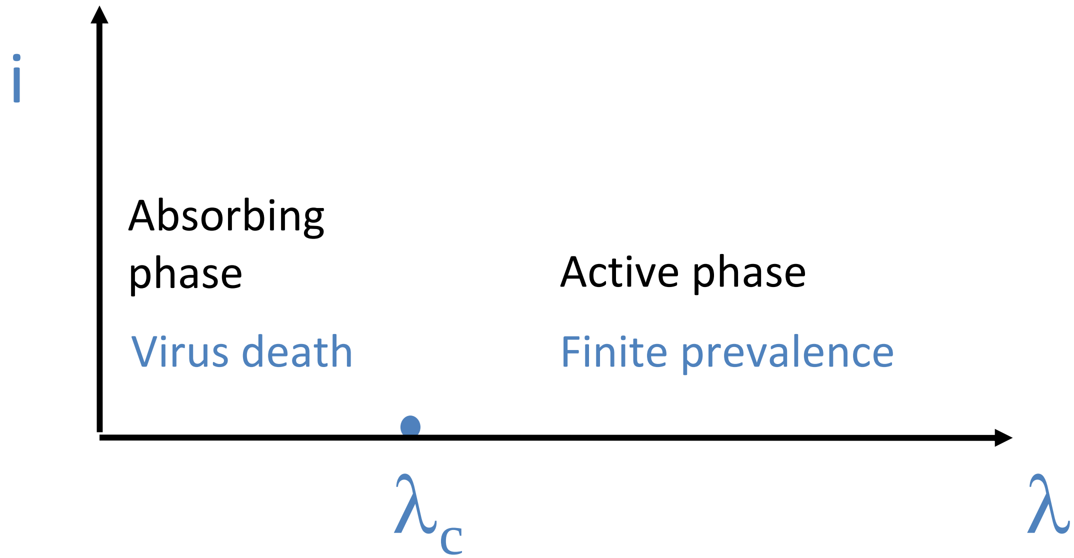
$$s + i = 1 \quad \rightarrow \frac{di}{dt} = (\beta <k> - \gamma) i - \beta <k> i^2$$

Stable states:

$$\frac{di}{dt} = 0 \quad i_f^1 = 0, \quad i_f^2 = 1 - \frac{\gamma}{\beta <k>} = 1 - \frac{1}{\lambda <k>}$$

$$\text{if } \lambda > \lambda_c = \frac{1}{<k>} \quad \rightarrow \quad i_f^2 > 0 \quad \text{and stable!}$$

Phase diagram



- Non-equilibrium phase transition
- λ_c=epidemic threshold = critical point
- Prevalence i =order parameter

NB: The question of thresholds in epidemics is central

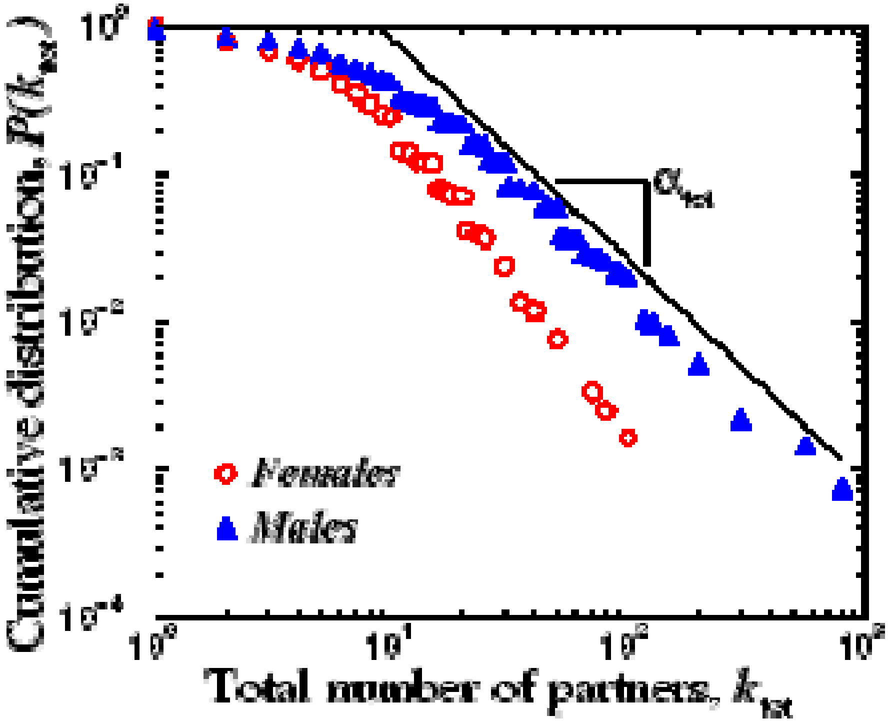
Complex networks

Viruses propagate on networks:

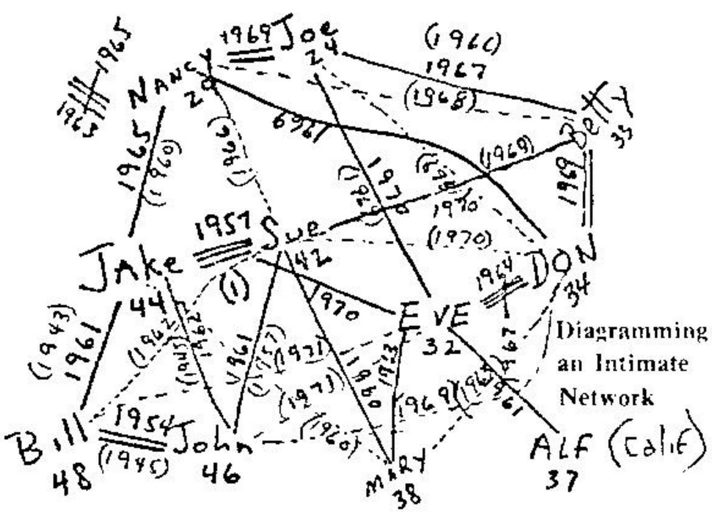
- Social (contact) networks
- Technological networks:
    - Internet, Web, P2P, e-mail...

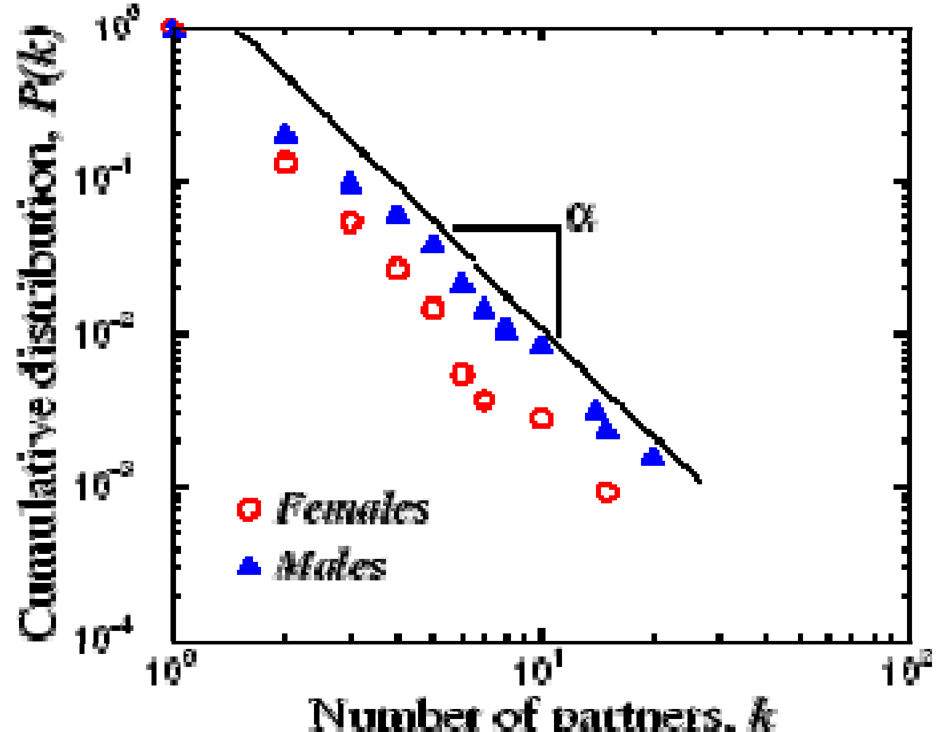
...which are complex, heterogeneous networks

Broad degree distributions

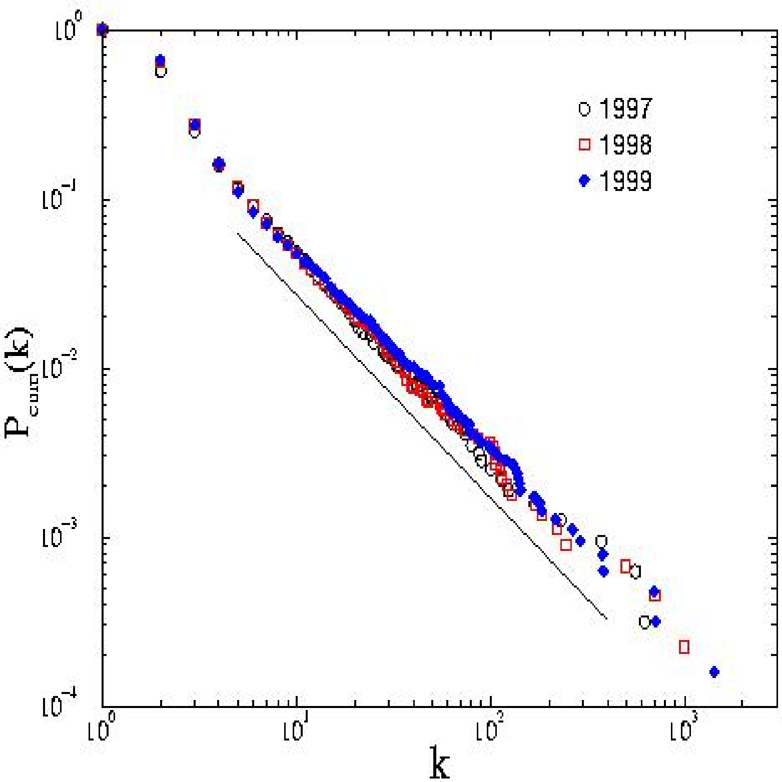


The web of human sexual contacts (Liljeros et al., Nature 411, 907-908, 2001)

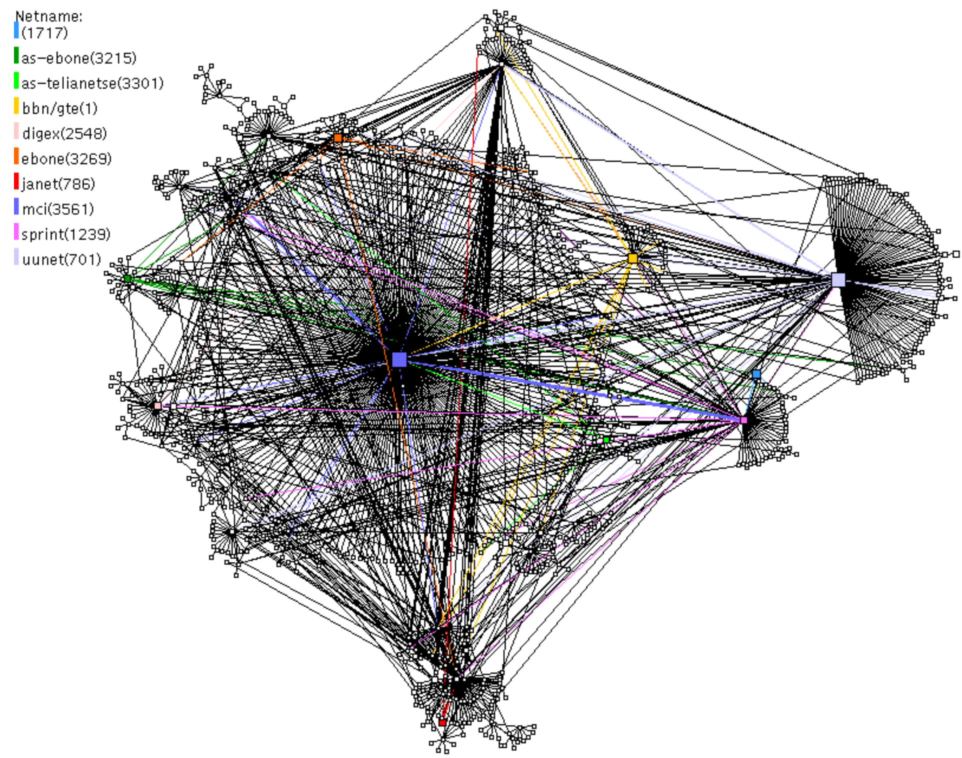


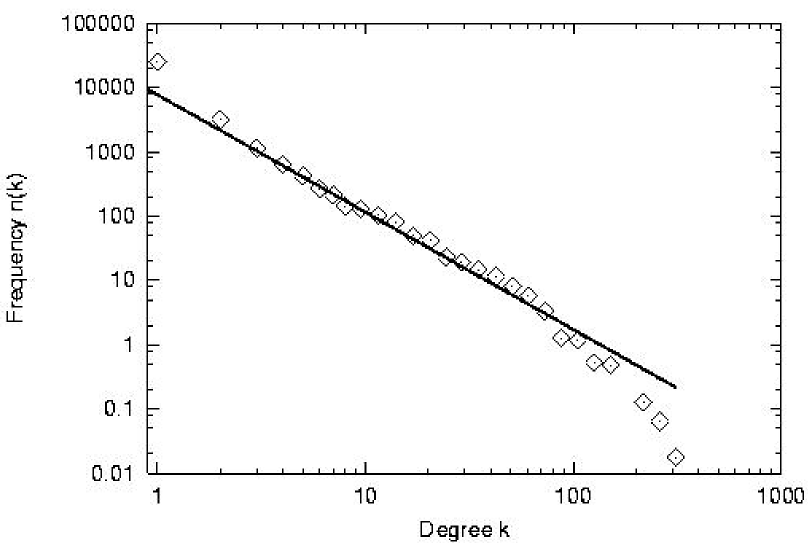


Broad degree distributions



Internet





E-mail network

H. Ebel, L.-I. Mielsch, S. Bornholdt, Phys. Rev. E **66**, 035103 (2002)

Epidemic spreading on heterogeneous networks

Number of contacts (degree) can vary a lot

huge fluctuations ($<k^2>/<k>$)

In networks without degree-degree correlations....:

$$\lambda_c = \frac{<k>}{<k^2>}$$

Epidemic threshold in heterogeneous networks

$$\lambda_c = \frac{\langle k \rangle}{\langle k^2 \rangle}$$

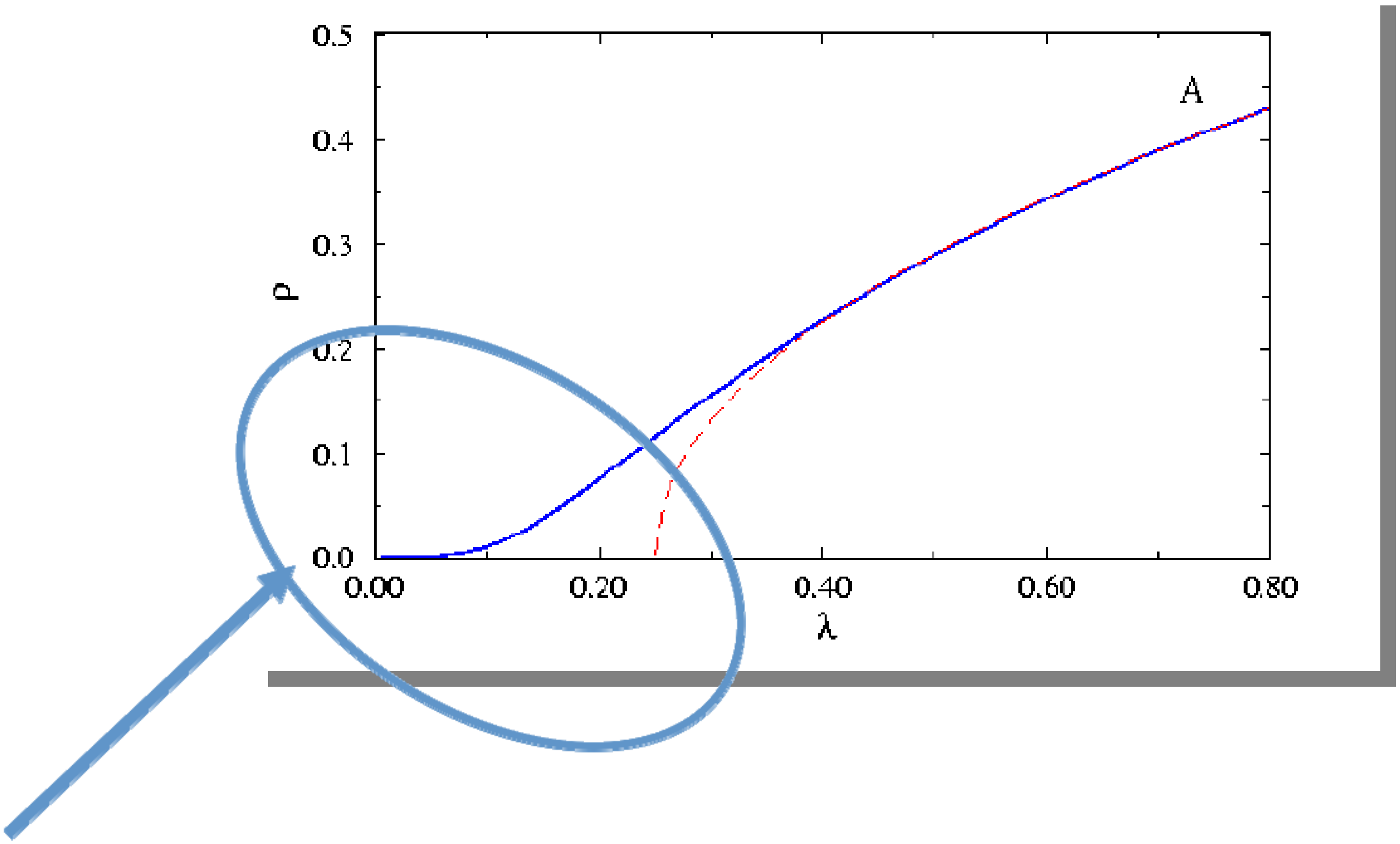
$$\langle k^2 \rangle \to \infty$$

$$\lambda_c \to 0$$

Order parameter behavior in an infinite system

$$i = 2e^{-1/m\lambda}$$

Epidemic threshold in heterogeneous networks



- Wide range of spreading rate with low prevalence
- Lack of healthy phase = standard immunization cannot drive the system below threshold!!!

Dynamical behaviour

At short times: $i(t) \sim \exp(t/\tau)$, with

$$\tau = \frac{\langle k \rangle}{\lambda(\langle k^2 \rangle - \langle k \rangle)}$$

$\langle k^2 \rangle \to \infty$ → $\tau \to 0$

M. Barthélemy, A. Barrat, R. Pastor-Satorras, A. Vespignani
Phys. Rev. Lett. **92**, 178701(2004)

Summary: epidemic spreading

- Absence of an epidemic/immunization threshold
- The network is prone to infections (endemic state always possible)
- Small prevalence for a wide range of spreading rates
- Random immunization totally ineffective (targeted immunization instead!)
- Infinite propagation velocity

Huge consequences of the heterogeneous topology

R. Pastor-Satorras, A. Vespignani, Phys. Rev. Lett. **86**, 3200-3203 (2001)
R. Pastor-Satorras, A. Vespignani, Phys. Rev. E **65**, 036104 (2002)
Z. Dezso, A.-L. Barabási, Phys. Rev. E. **65**, 055103(R) (2002)
R. Cohen, D. ben Avraham, S. Havlin, Phys. Rev. E **66**, 036113 (2002)

Opinion formation models

Simplified models of interaction between agents

Questions:

* Convergence to consensus?
* How?
* In how much time?

Opinion formation models

Voter model:

n agents i=1,..n

Opinion s_i = 1 or -1

At each time step:

* Choose one agent i
* i chooses at random one of his neighbors j
* Agent i adopts the opinion of agent j

Voter model

- $n_A(t)$ = fraction of "active" links
(active= linking two non-agreeing agents)

- $\rho(t)$ = fraction of runs "surviving" at time t, i.e.
not having reached full agreement

- $\tau(n)$ = time for n agents to reach agreement

Voter model

Mean-field=agents on a complete graph

$$\tau(n) \sim n$$

No ordering in the infinite size limit!
(surviving runs keep a finite fraction of active links)

Voter model

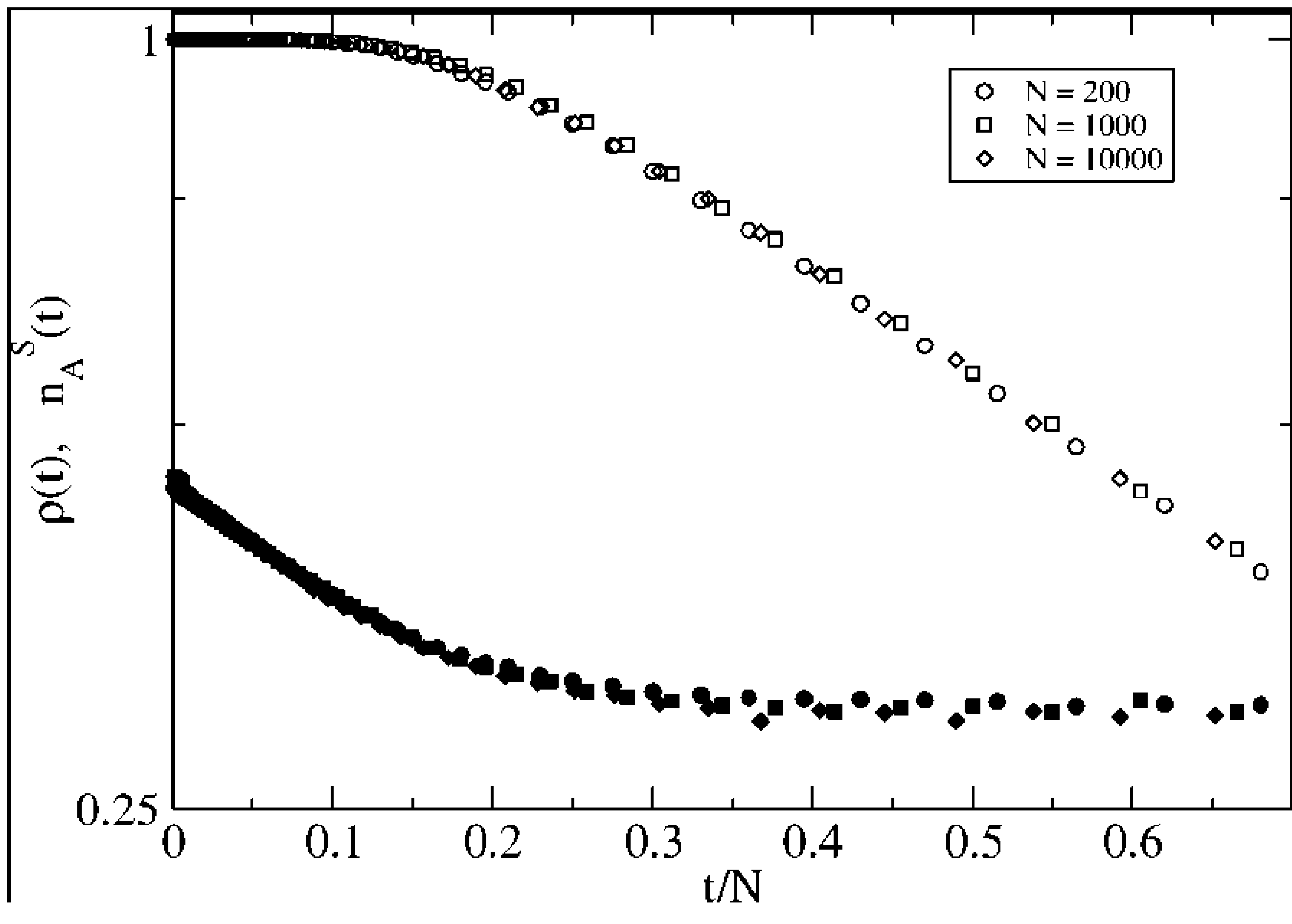
Agents forming a relationship network:
At each time step:
    choose an agent i
    choose a neighbor of j
    i adopts j's opinion
Or:
    choose an agent i
    choose a neighbor of j
    j adopts i's opinion

→ *Reverse* voter model

can be important in heterogeneous networks because:
- a randomly chosen node has typically small degree
- the neighbour of a randomly chosen vertex has typically large degree

Voter model

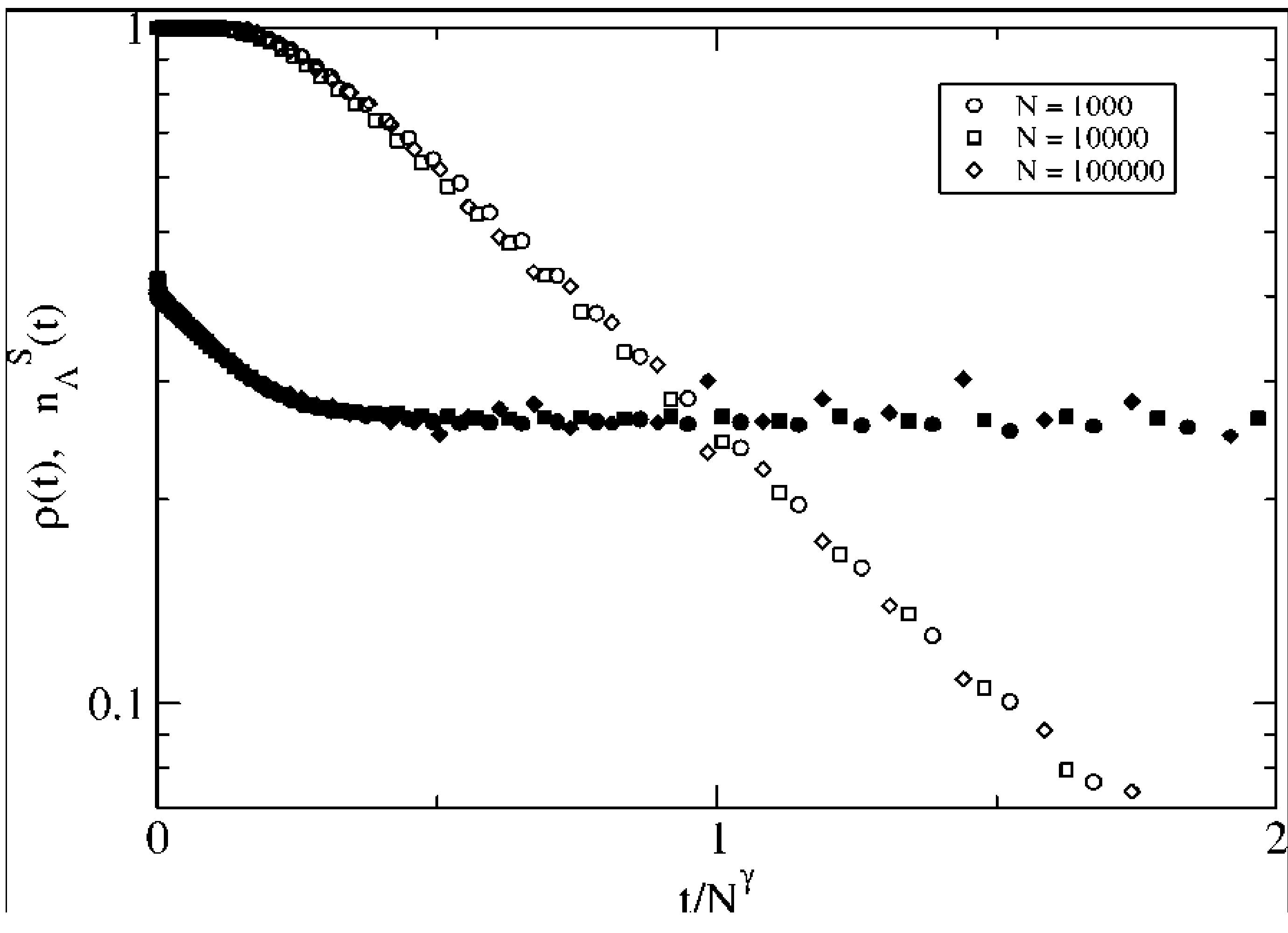
On random (homogeneous) graphs:
similar to mean-field
$\tau \sim n$
No ordering for surviving runs



Voter model

On scale-free (heterogenous) graphs

- $\tau \sim n^{\gamma}$
- No ordering for surviving runs



Axelrod model

n agents i=1,..,n on a lattice

- Each agent has F attributes
- Each attribute can take q values

i	j
1	4
2	3
5	5
3	1
1	2
1	3
4	2

Axelrod model

Dynamical interaction:

- If i and j have no common attribute:

  No interaction possible

- If i and j have at least one common attribute:

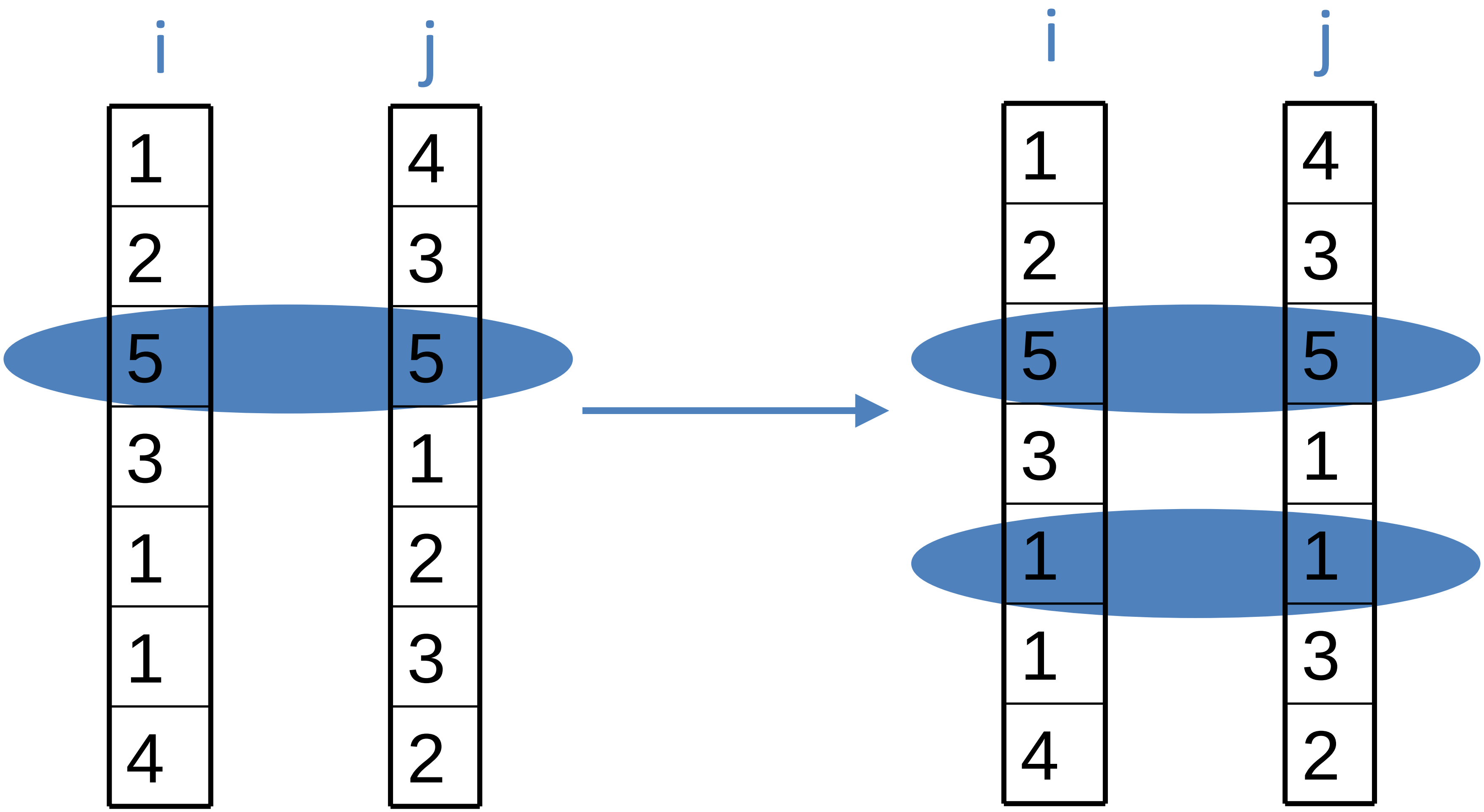
  i chooses one of the other attributes

  and adopt j's value

Bounded confidence

Axelrod model

Dynamical interaction: example



Axelrod model

Dynamical interaction: favors convergence

Large F (number of attributes): large probability to have at least one common attribute

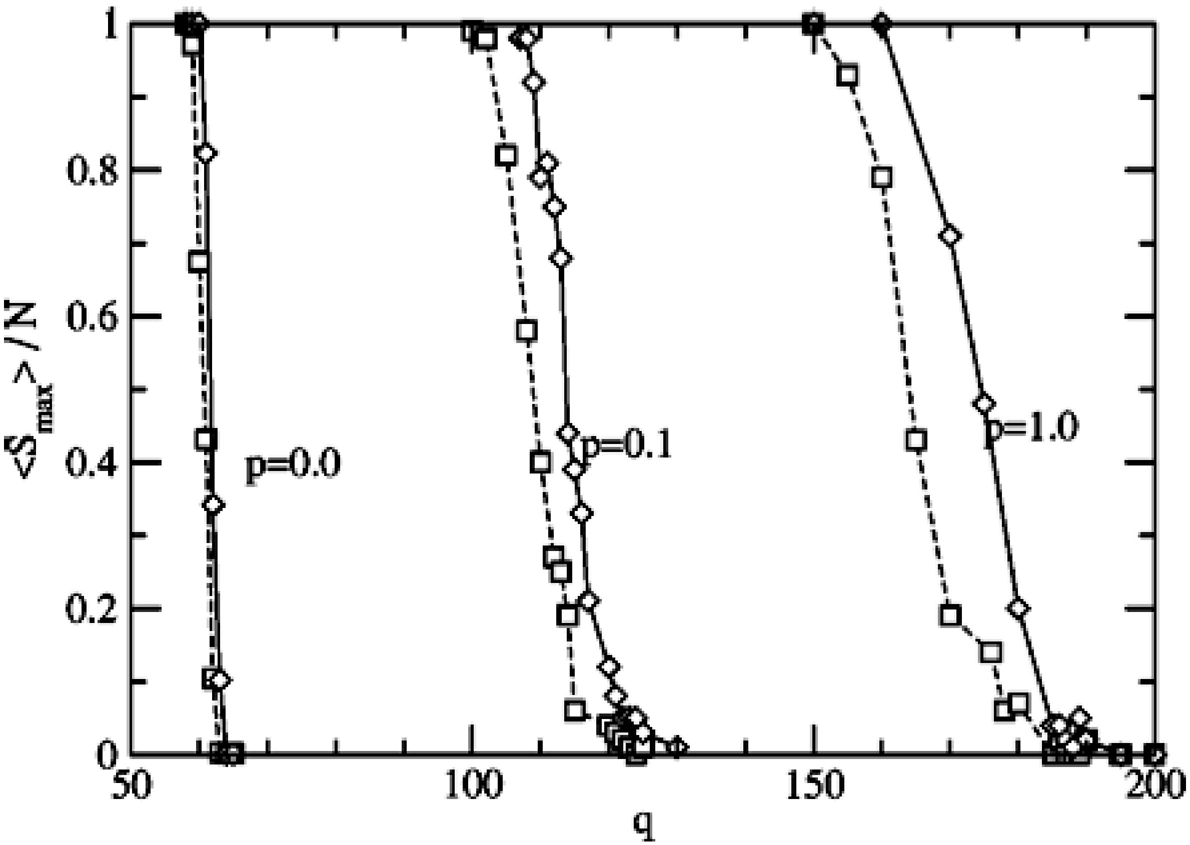
Large q: small probability to have at least one common attribute

Transition from consensus to fragmented state as q increases

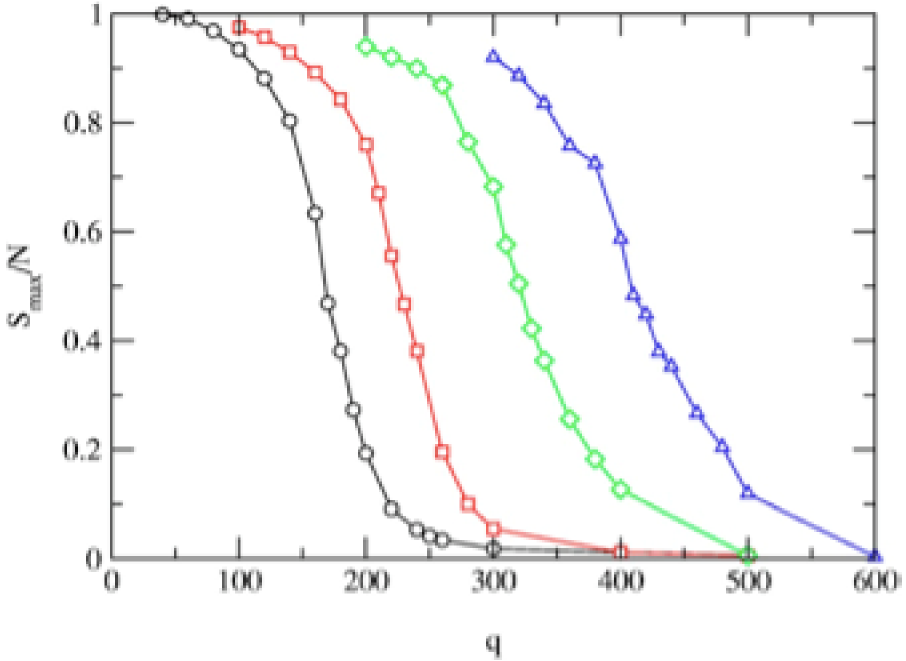
Axelrod model

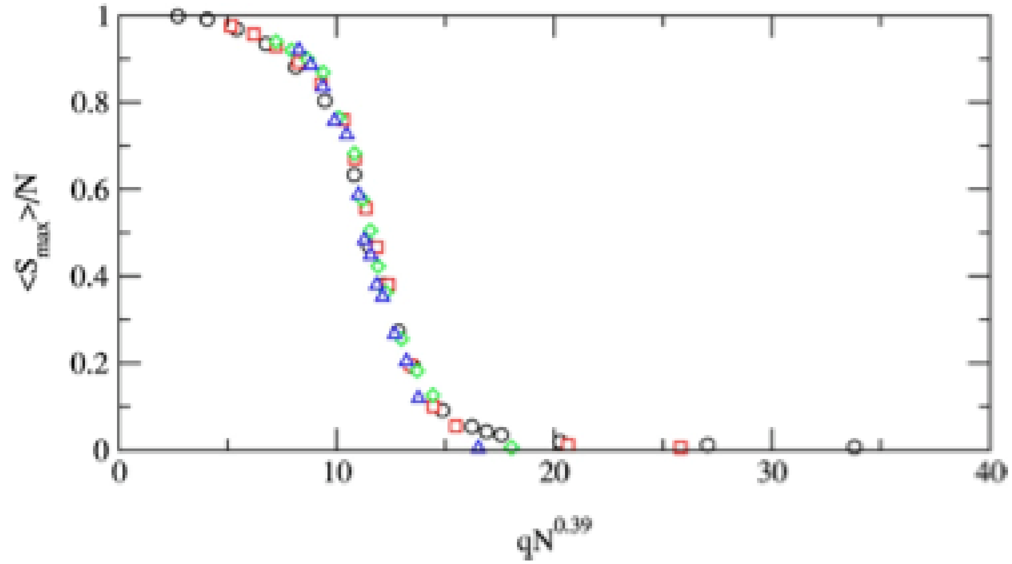
On a Watts-Strogatz network

Order parameter: Size of the largest cluster with agreeing agents



Axelrod model





q_c grows as $n^{0.39}$: the transition disappears in the limit of infinite network size, only one cultural domain survives for any value of q

Agents on a scale-free network: no transition as $n \to \infty$

The hubs "polarize" the system $\to$ convergence

Other models

- Deffuant model: continuous opinions, bounded confidence
- Naming Game
- …
- Many possible variants: zealots, external fields, noise…
- Dynamically evolving networks

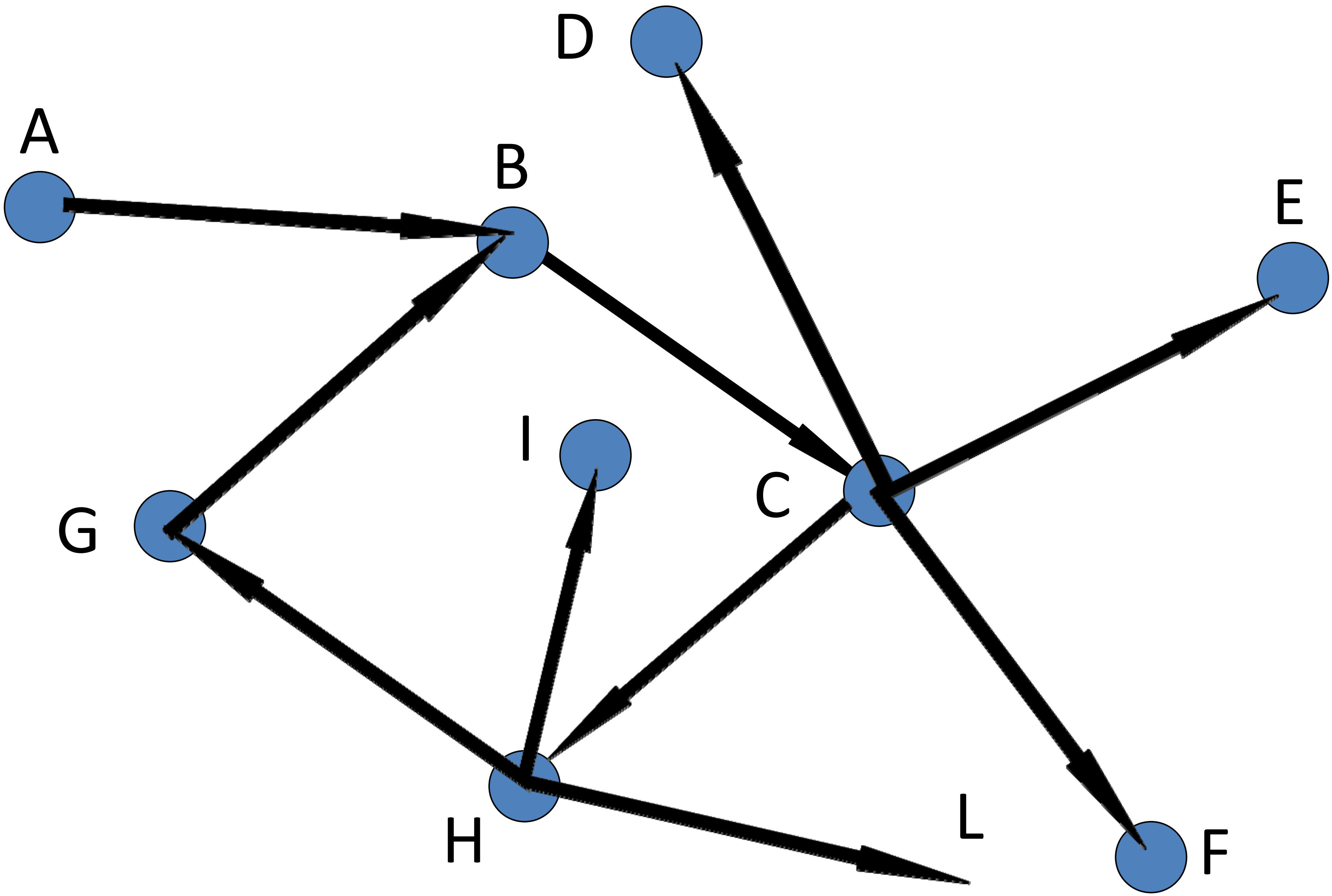
Search and Navigation on the Web

The two processes are intertwined: what is the best strategy to look for interesting sites on the Web?

Before the search engines era, people used to look for sites by surfing on the Web following hyperlinks.

Problems with Web surfing:

- Pages without incoming links (in-degree zero) are unreachable
- If one reaches a page without outgoing links (out-degree zero, dangling end), one gets stuck in there

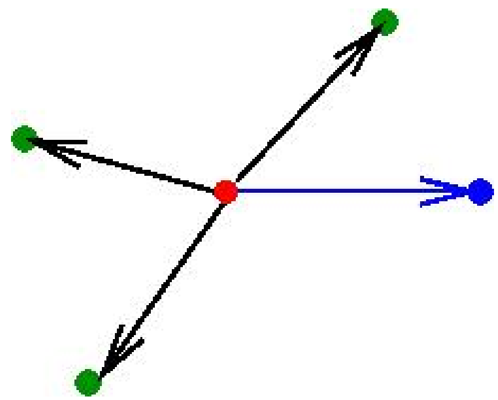

D
A
B
E
I
C
G
L
H
F

What is then the best way to navigate on a directed network?

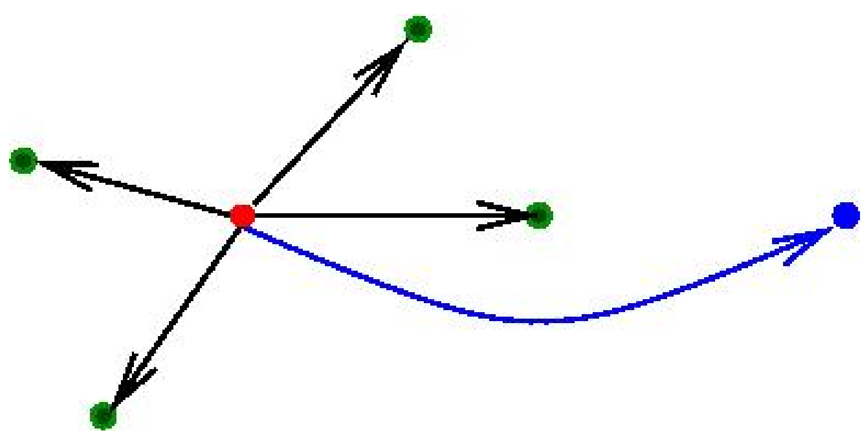
Ideally, one should have the chance to visit all pages and to move on from dangling ends: the simplest way is to introduce a probability to leap from a page to any other!

The resulting navigation is a mix of two processes

a) with probability 1-q, a user moves from a page O to a neighboring page by following any of the outlinks of O (random walk);



b) with probability q, it jumps to any other page of the Web (random jump).



PageRank

$$p(i) = \frac{q}{n} + (1-q)\sum_{j \to i} \frac{p(j)}{k^j_{out}}$$

The navigation can be simulated by placing an agent on an arbitrarily chosen node and making it move

After a sufficient number of iterations, for a fixed probability q, each node will have a well defined probability to be visited by the agent.

Problem: users cannot jump from a page to another, because they would need to know all Web pages

Who can do that?

All search engines need to rank Web pages according to their supposed importance

PageRank is the prestige measure of a Web page according to Google

Google stores a (large) sample of the Web graph in its database; the full information contained in the sampled Web pages is stored as well

The PageRank value of all pages/nodes of the graph is calculated

When a user submits a query, Google selects all pages which contain the input string(s) and return them listed in decreasing order of PageRank dress http://www.google.com/ Go

ks TOSHIBA Access Customize Links Free Hotmail RealPlayer Windows Windows Marketplace Windows Media

Personalized Home

Google™

Web Images Groups News Froogle Local Scholar **more »**

maria sharapova

Google Search I'm Feeling Lucky

Advanced Search
Preferences
Language Tools

Get less spam. Use all 2.5GB for your real email instead. Gmail.

Advertising Programs - Business Solutions - Hurricane Katrina Resources - About Google

Make Google Your Homepage!

http://www.google.com - maria sharapova - Google Search - Microsoft Internet Explorer provided by Insight Broadband

File Edit View Favorites Tools Help

Back Search Favorites

Address http://www.google.com/search?hl=en&q=maria+sharapova&btnG=Google+Search Go

Links TOSHIBA Access Customize Links Free Hotmail RealPlayer Windows Windows Marketplace Windows Media

Web Images Groups News Froogle Local Scholar **more »**

maria sharapova [Search] Advanced Search Preferences

Web Results **1** - **10** of about **2,040,000** for maria **sharapova**. (**0.07** seconds)

Image results for **maria sharapova**

http://english.pravda.ru/science/19/95/380/13530_Sharapova.html

Sponsored Links

Bob Larson's Tennis News
The Official Tennis Newswire
Tennis News, Results, Photos, Info
www.tennisnews.com

Maria Sharapova
everything about **Maria Sharapova**: photos, interviews, articles, statistics, results and much more!
www.**maria**world.net/ - 27k - Sep 17, 2005 - Cached - Similar pages
PHOTOS - UPDATES - MULTIMEDIA - BIOGRAPHY
More results from www.mariaworld.net »

Get news, photos and more about tennis player **Maria Sharapova**
This Martina Hingis provides photos, news, infos and results about the Swiss tennis superstar.
www.**maria-sharapova**.net/ - 33k - Sep 17, 2005 - Cached - Similar pages

Maria Sharapova WTA Tour Profile
Includes a biography, match record, results, activity and career highlights.
www.wtatour.com/players/playerprofiles/ PlayerBio.asp?PlayerID=310137 - 38k - Cached - Similar pages

Maria Sharapova - Tennis Princess
Everything about teen queen **Maria Sharapova**: News, pictures, photos, wallpapers, interviews, articles, statistics, results and more!
www.**maria-sharapova**.org/ - 43k - Sep 17, 2005 - Cached - Similar pages

http://www.google.com - maria sharapova - Google Search - Microsoft Internet Explorer provided by Insight Broadband

File Edit View Favorites Tools Help

Back Search Favorites

Address http://www.google.com/search?hl=en&q=maria+sharapova&btnG=Google+Search Go

Links TOSHIBA Access Customize Links Free Hotmail RealPlayer Windows Windows Marketplace Windows Media

Latest News, **Maria Sharapova** Matches, Statistics, Rankings.
www.sexy**sharapova**.com/ - 89k - Sep 17, 2005 - Cached - Similar pages

AskMen.com - **Maria Sharapova**
AskMen.com feature on **Maria Sharapova** that includes pics, pictures, biography, ... A brilliant tennis player with a deadly backhand, **Maria Sharapova** was the ...
www.askmen.com/women/models_200/ 223_**maria_sharapova**.html -
Similar pages

Maria Sharapova hot Pics
Maria Sharapova hot sexy modeling Pics, **Maria Sharapova** non nude Pictures, **Maria Sharapova** topless Photos, **Maria Sharapova** sexy Images, hot Gallery, ...
maria-sharapova.boom.ru/ - 10k - Cached - Similar pages

Maria Sharapova wins Wimbledon, **Maria Sharapova** Wallpaper, **Maria** ...
Maria Sharapova wins Wimbledon, **Maria Sharapova** Wallpaper, **Maria Sharapova** Pics, **Maria Sharapova** Biography, **Maria Sharapova** pictures.
www.**maria-sharapova**.cc/ - 9k - Cached - Similar pages

Gooooooooooogle

Result Page: 1 2 3 4 5 6 7 8 9 10 **Next**

Google Desktop Search 9:30 AM

Free! Instantly find your email, files, media and web history. Download now.

maria sharapova Search

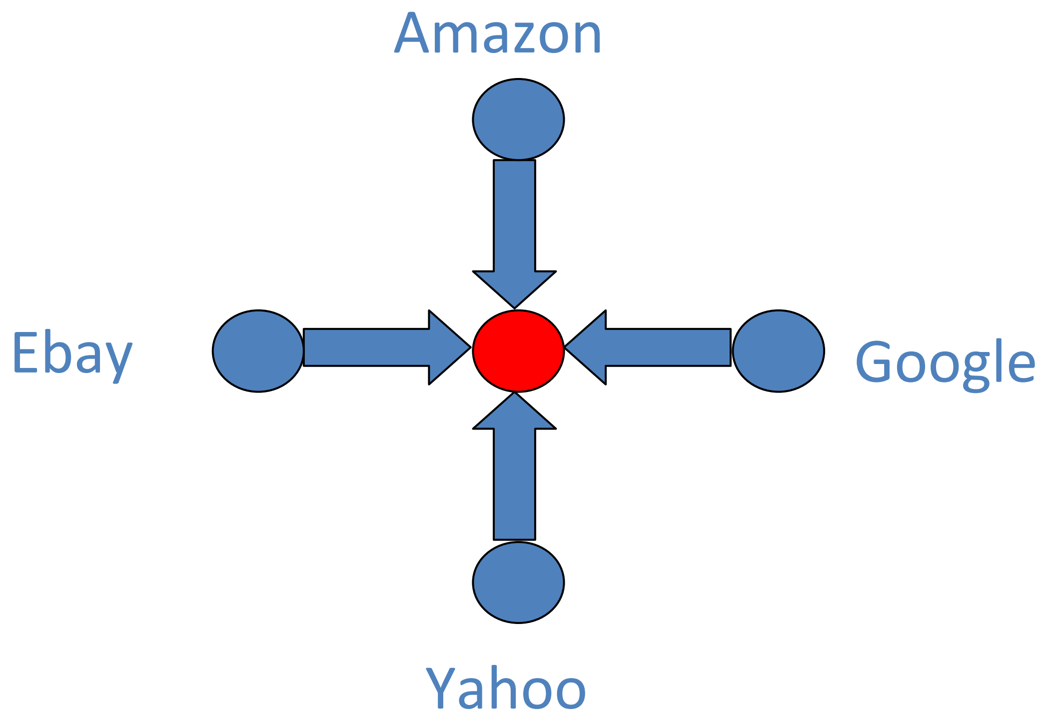
Search within results | Language Tools | Search Tips | Dissatisfied? Help us improve

Google Home - Advertising Programs - Business Solutions - Hurricane Katrina Resources - About Google

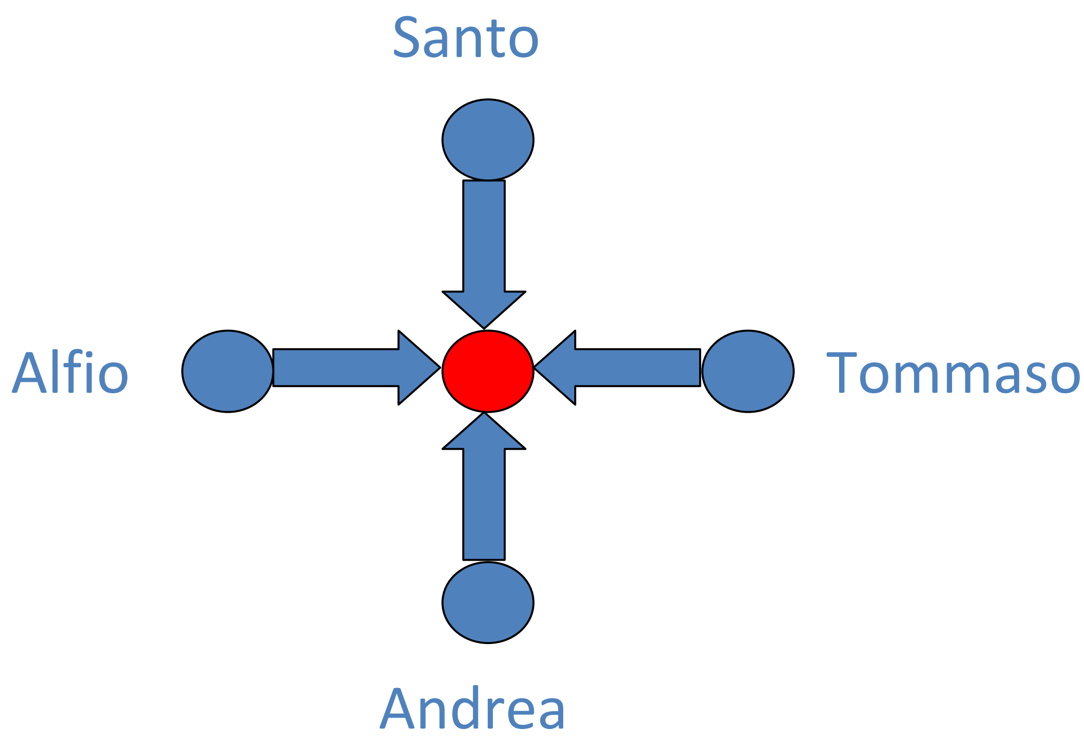
A simple alternative to PageRank would be just to count the number of incoming links to a page (link popularity)

A page with many in-links is usually more important than a page with just a few in-links

PageRank is better than link popularity because it not only takes into account the in-degree of a page but also how important the in-neighbors of the page are

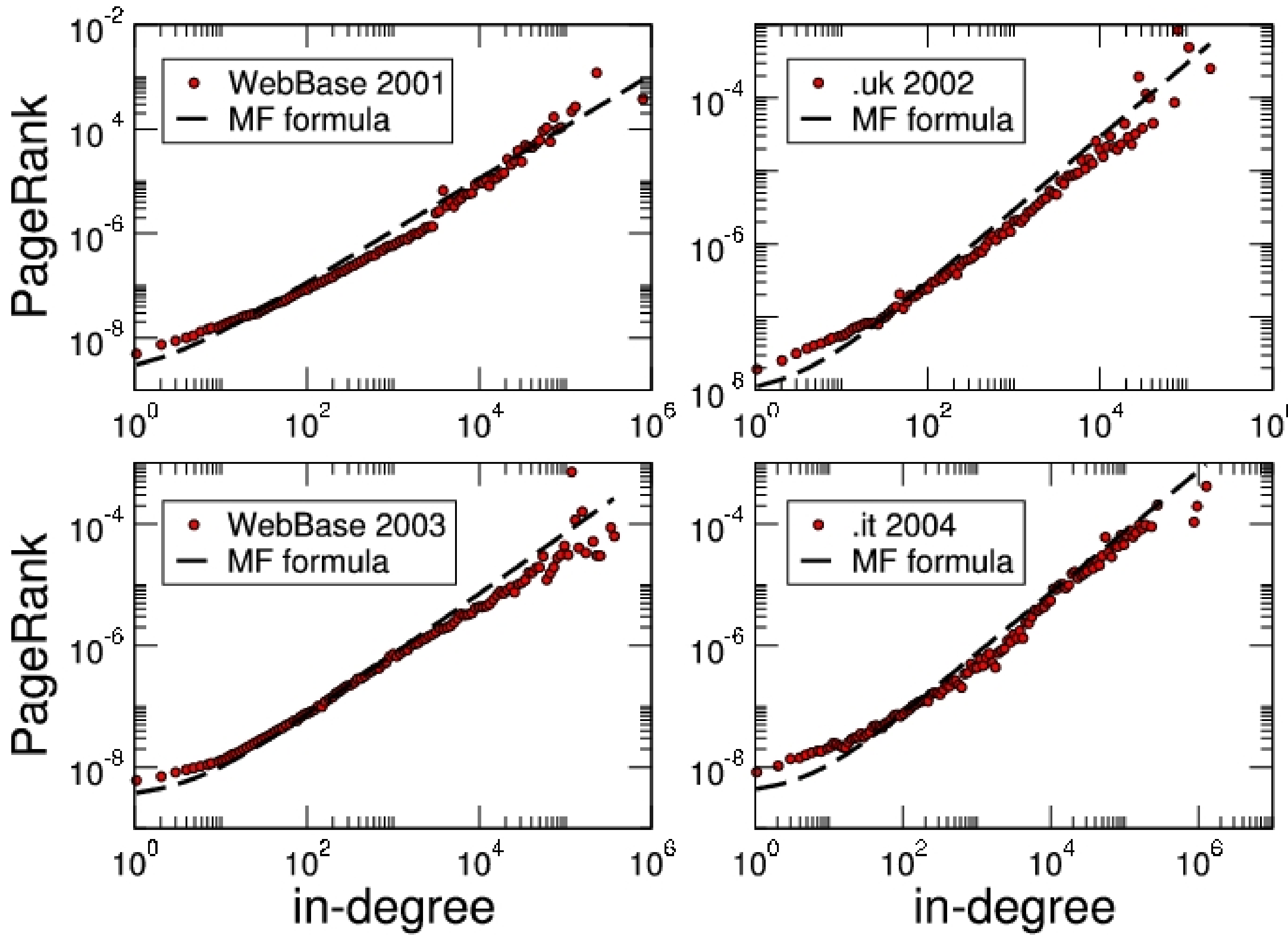


Link popularity: 4. PageRank p



Link popularity: 4. PageRank q«p

Relation between PageRank and in-degree



References

* **Evolution of networks**
  S.N. Dorogovtsev, J.F.F. Mendes, Adv. Phys. 51, 1079 (2002), cond-mat/0106144
* **Statistical mechanics of complex networks**
  R. Albert, A-L Barabasi,
  Reviews of Modern Physics 74, 47 (2002),
  cond-mat/0106096
* **The structure and function of complex networks**
  M. E. J. Newman, SIAM Review 45, 167-256 (2003), cond-mat/0303516
* **Complex networks: structure and dynamics**
  S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, D.-U. Hwang, Physics Reports 424, 175-308 (2006)
* **Community detection in graphs**
  S. Fortunato, Physics Reports 486, 75-174 (2010)
  arXiv: 0906.0612