

Modeling social networks with local mechanisms

R. Lambiotte
M. Ausloos
P. Krapivsky



Content

I. Some background

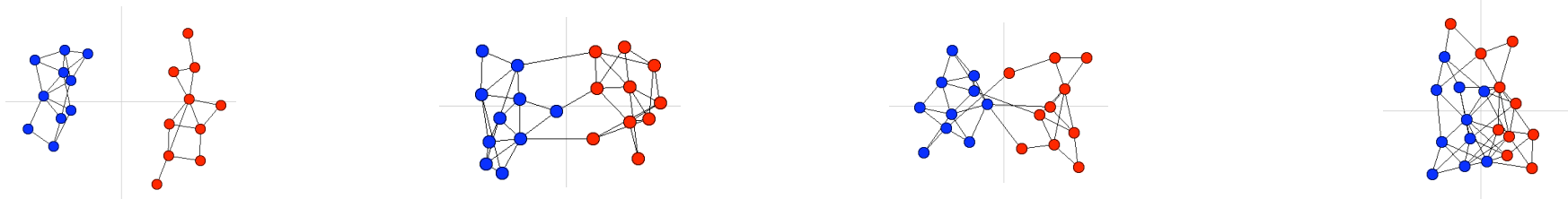
- a) What about networks?
- b) Barabasi-like networks

II. Models with local mechanisms

- a) Coupled networks
- b) Role of the localization in the network
- c) Copying each others' environment
- d) Some analytical predictions

III. Visualizing network dynamics

IV. Conclusion



Examples of complex networks:

- Internet
- Transport networks
- Protein interaction networks
- Food webs
- Social networks, e.g. collaboration networks, citation networks, friendship networks...

Chains of Affection: The Structure of Adolescent Romantic and Sexual Networks: Peter S. Bearman, James Moody and Katherine Stovel

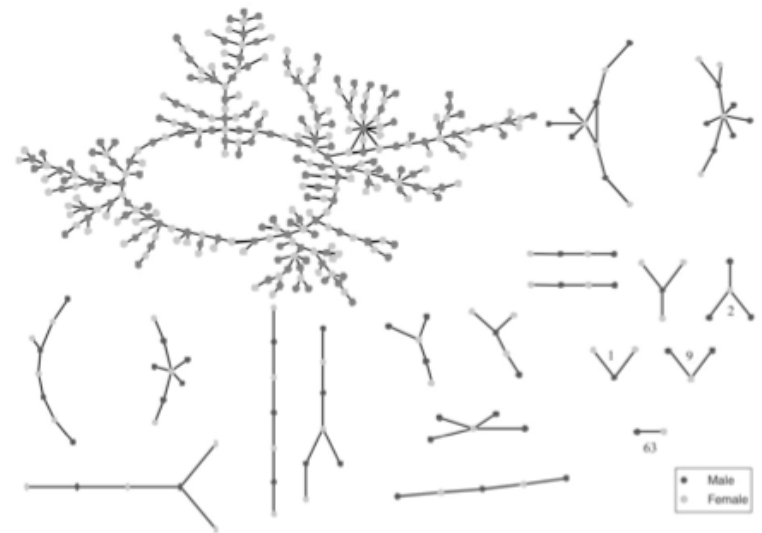


FIG. 2.—The direct relationship structure at Jefferson High



Hype of social networking

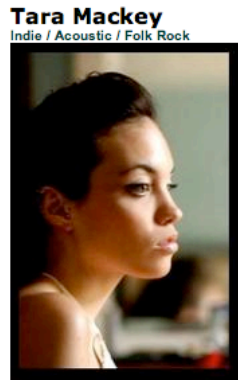
flicker, myspace, audioscrobbler, ...

MySpace.com | Home The Web MySpace Search Help | SignOut

INKT voor uw printer € 0*
* Of fotopapier (20 vel) voor € 0 als nieuwe klant vandaag!
www.inkclub.com

Home | Browse | Search | Invite | Film | Mail | Blog | Favorites | Forum | Groups | Events | Videos | Music | Comedy | Classifieds

MYSACE MUSIC Music Videos | Directory | Search | Top Artists | Shows | Music Forums | Music Classifieds | Artist Signup



Tara Mackey
Indie / Acoustic / Folk Rock

FLORAL PARK, New York
United States

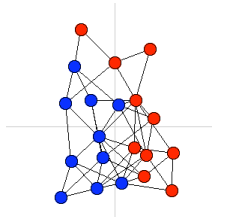
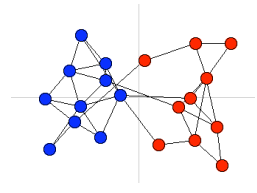
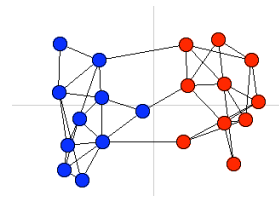
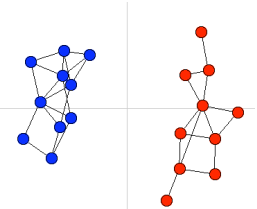
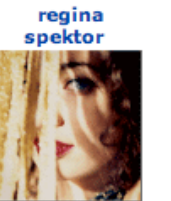
Profile Views: 17921

Last Login: 8/13/2006

View My: [Pics](#) | [Videos](#)

Contacting Tara Mackey

Total Plays: 26283



New marketing/sales strategies

Importance of the buzz of a product: discussions between bloggers, chatters... A modern “Bouche à oreille”

The huge amount of online databases helps to “measure” and predict the behaviour of users



The screenshot shows a music store interface with three main sections:

- Inside the Music Store:** Features the album "How to Dismantle an Atomic Bomb" by U2, released in 2004, with a 5-star rating. It includes links for "Reviews", "Gift This Music", and "Tell a friend".
- More from U2:** Lists two other albums: "Original of the Species (Single Version)" released in 2006, and "All Because of You - Single Version" released in 2005.
- Listeners Also Bought:** A list of 8 recommended albums, including "I'm Not Dead (International Version)" by Pink, "Ultimate Collection" by Eurythmics, "Playing the Angel" by Depeche Mode, "Nite Versions" by Soulwax, and "Pocket Revolution" by dEUS.



CREEN (Critical Events in Evolving Networks)

Data

Large databases available online, e.g. collaboration networks (ISI Web Of Knowledge, Arxiv), online communities (mySpace, musicmobs), mail networks, blog networks....

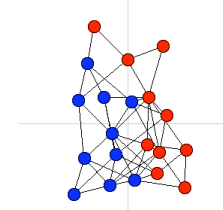
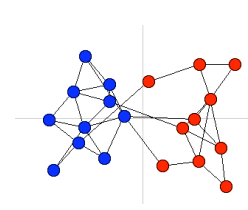
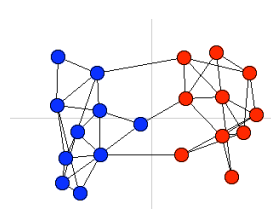
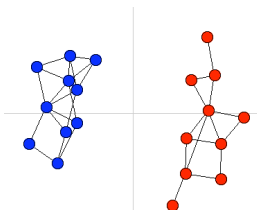
Visualisation

Visualise networks composed of thousands of links, with/without dynamical features

Modeling

Small World model, Barabasi-Albert like models

...



Local properties

node degree: numbers of links arriving at one node

in(out) degree: number of ingoing (outgoing) links for one node (directed networks)

clustering coefficient: number of triangles

local correlations: assortativity...



Global properties

percolation: formation of gigantic islands

distance: shortest path between nodes

time evolution of the total number of links, nodes...

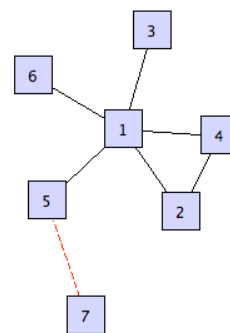
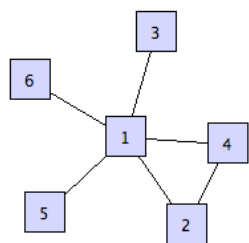
large scale structures, communities...



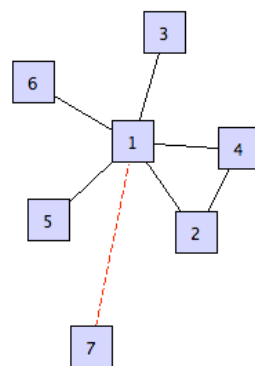
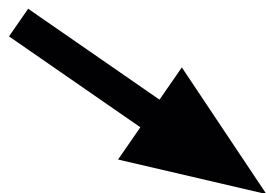
Barabasi-like networks

Many social systems exhibit what is called cumulated advantage, preferential attachment...

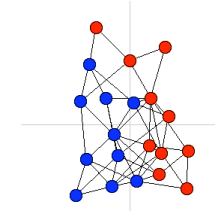
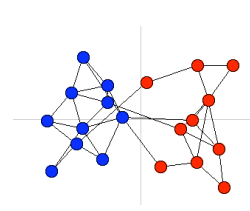
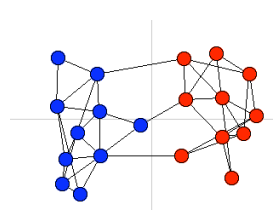
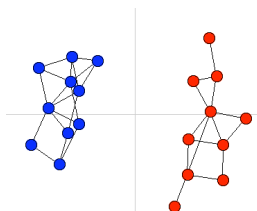
→ multiplicative process



small probability

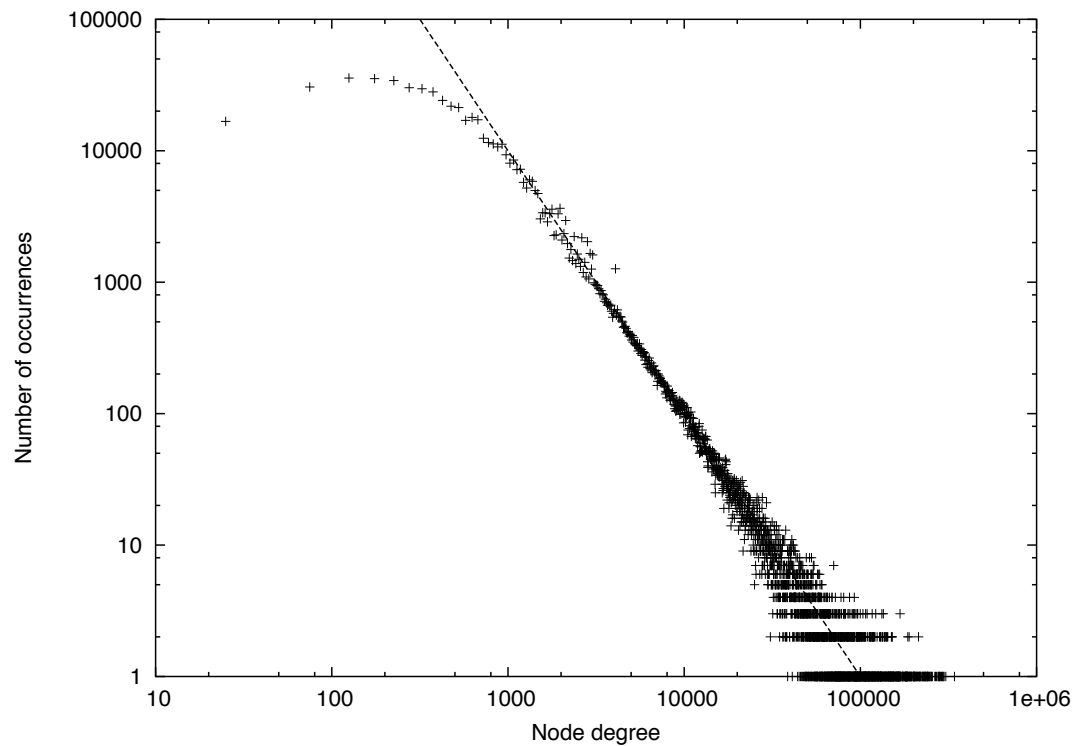


large probability



This model accounts for the importance of the degree of the nodes...

...and reproduces quite well the emergence of power-law degree distributions

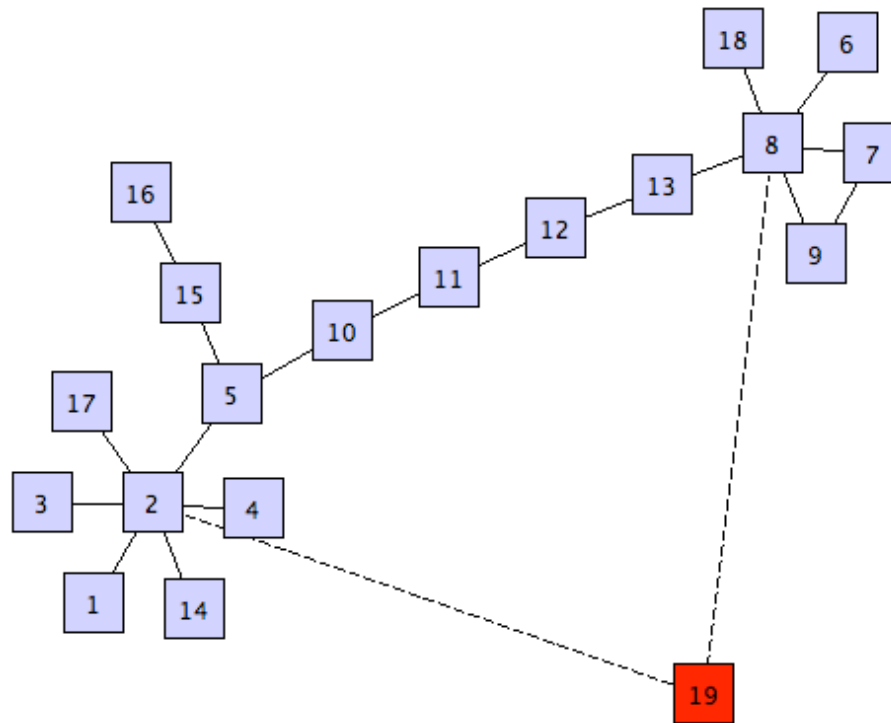


Distribution of the audience of music groups



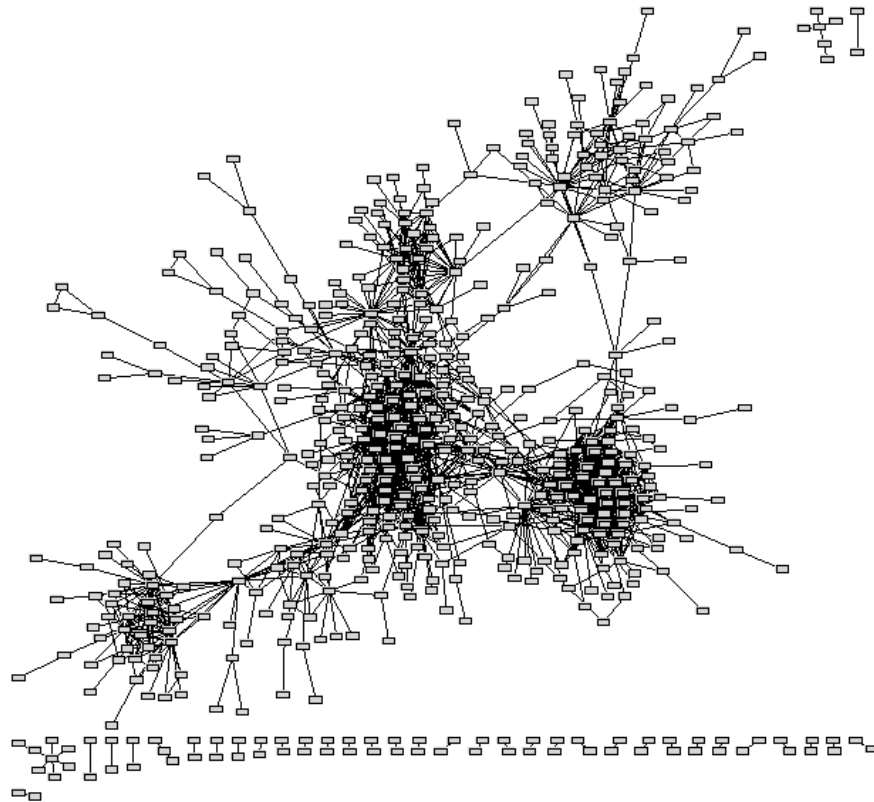
This model accounts for the importance of the degree of the nodes...

... but it is a “mean-field” model where the localization of the nodes is not taken in account (taste, kind, genre.. \)

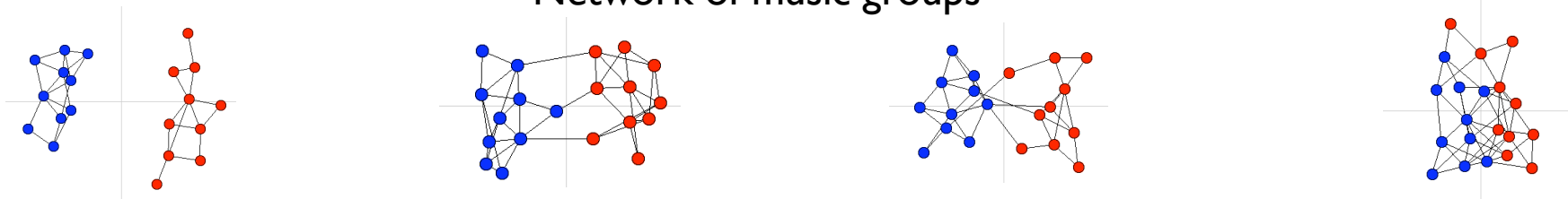


This model accounts for the importance of the degree of the nodes...

...and does not reproduce the structuring of the network (clustering coefficient, communities...)



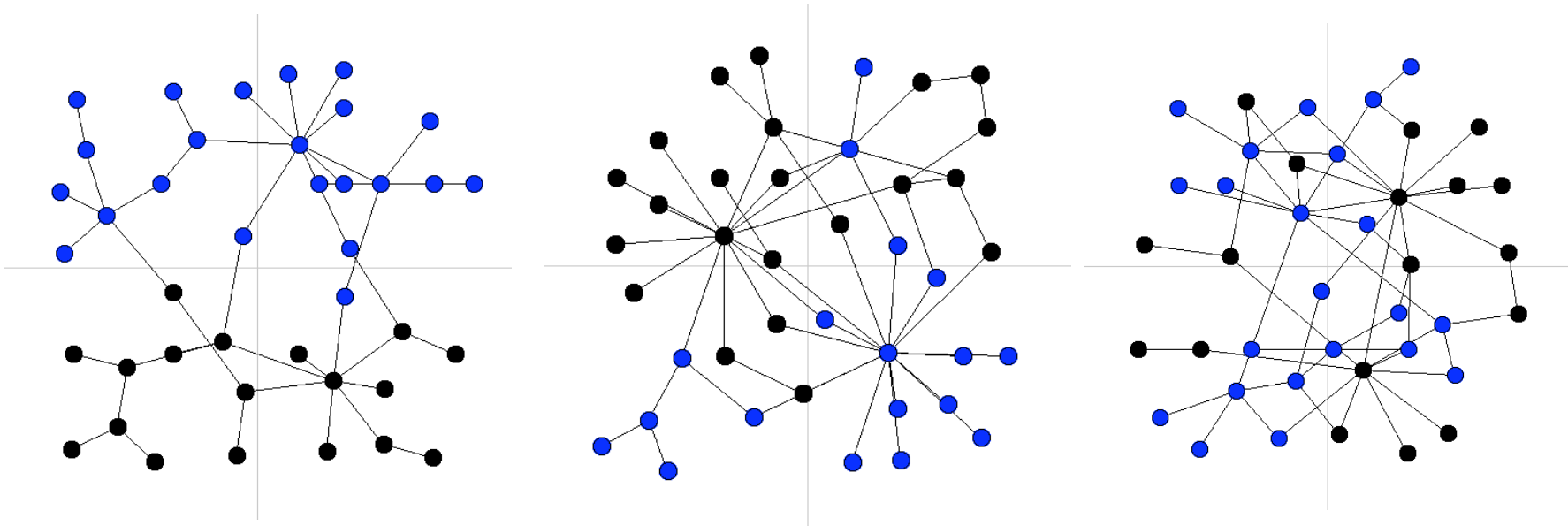
Network of music groups



Coupled networks

One possibility is to attribute “tags” to the nodes and model the system by coupled networks

Alike nodes have a tendency to link, while different nodes have a bigger probability to avoid each others



E.g. At each time step, a new node enters the system. With probability w , it relates to one existing node chosen with preferential attachment, and receives the kind of his parent node. With probability $(1-w)$, a pair of nodes of any kinds is chosen with PA.



The method is easy to implement, analytically and theoretically

But...

It is very difficult to characterise human beings by a small number of variables or tags, and the access to these characteristics is practically very difficult in most of the cases.

To avoid the use of these uncertain hidden-like variables for the nodes, is it possible to use the localisation in the network itself as the central role for the node evolution?

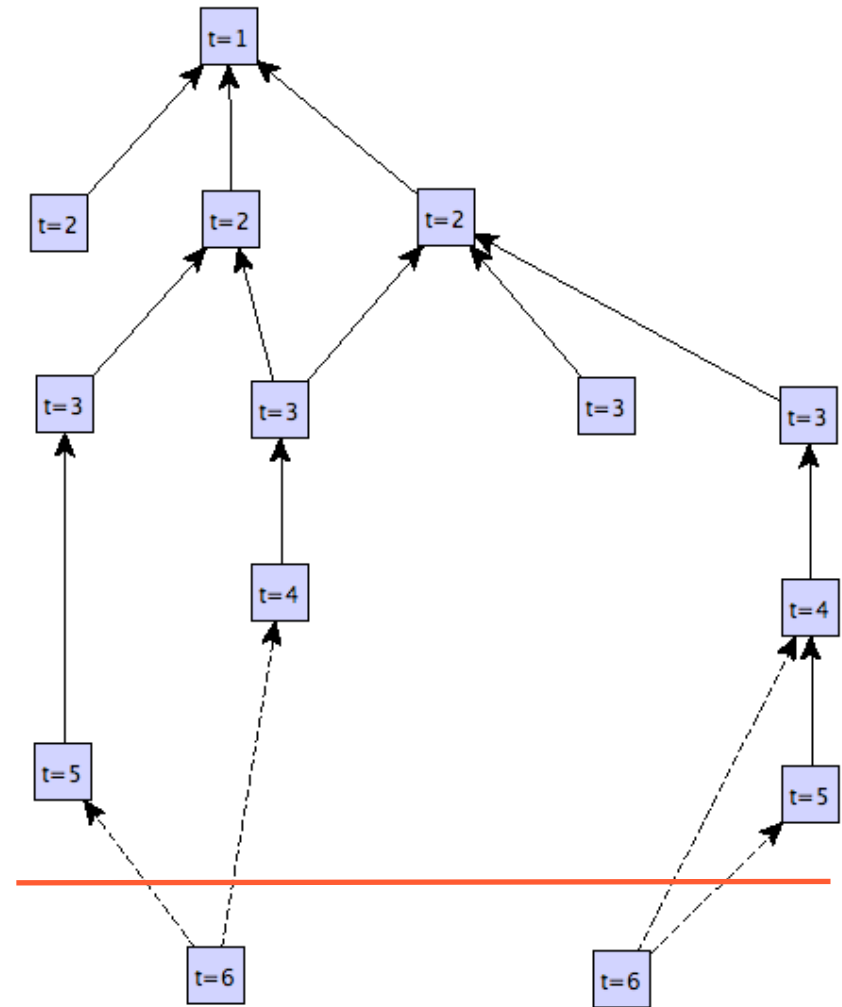


Is it possible to measure this importance of locality in *real* social networks?

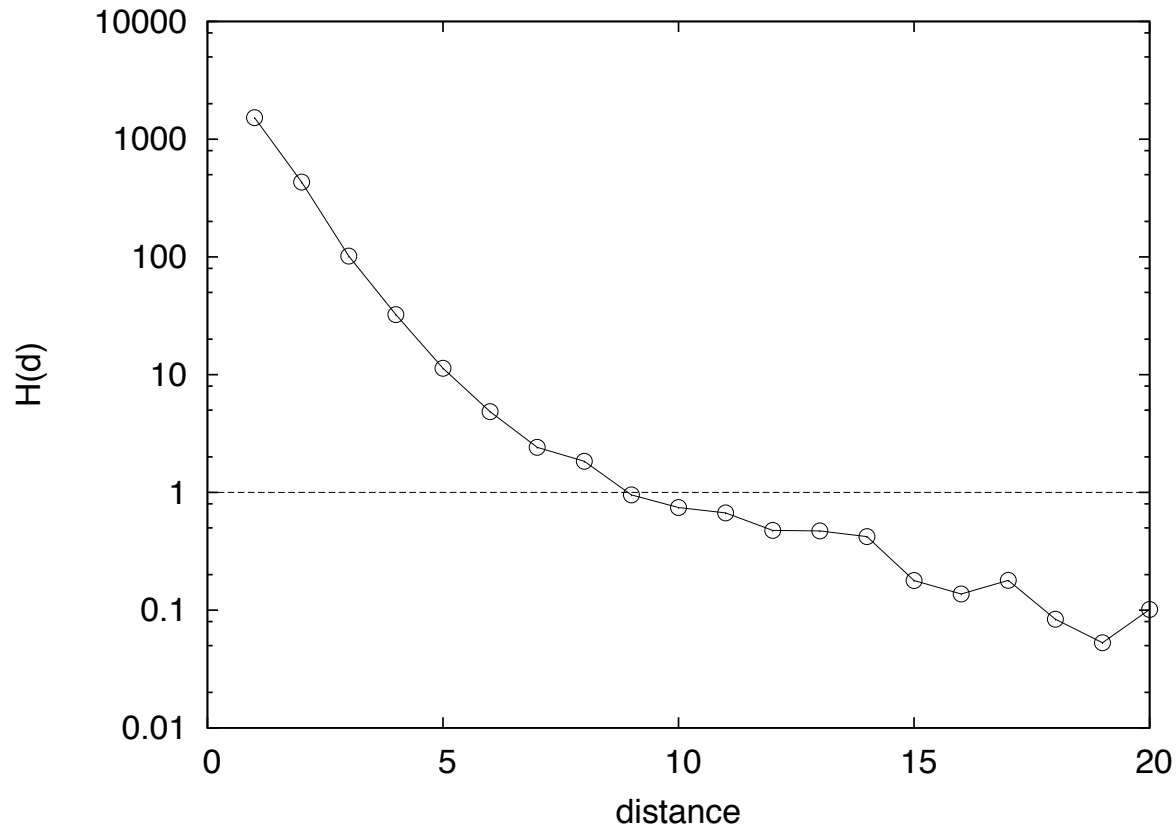
A listing of all articles published in *PRL*, together with their references has been downloaded [1996,2005].

We built the citation network of these ~ 32000 articles

When a new article cites $k=2,3,\dots$ articles, what are the distances between these k old articles



Distance between cited articles

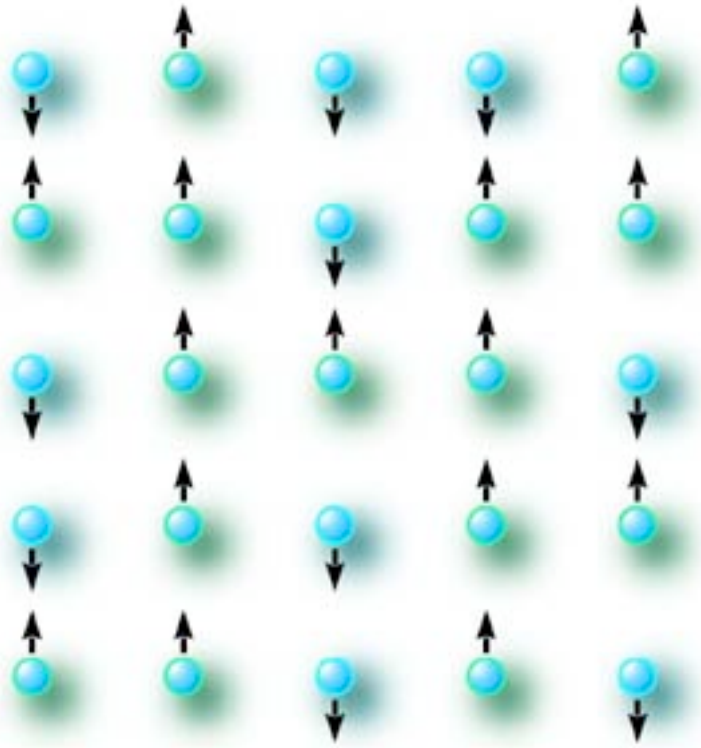


Measurement in
collaboration
networks, mail
networks, GSM
networks...



Models with local mechanisms: beyond the mean-field approach

Simplest interaction possible: you copy what you friend is doing. E.g. listen to the same music, have the same friends, cite articles in the same scientific field...



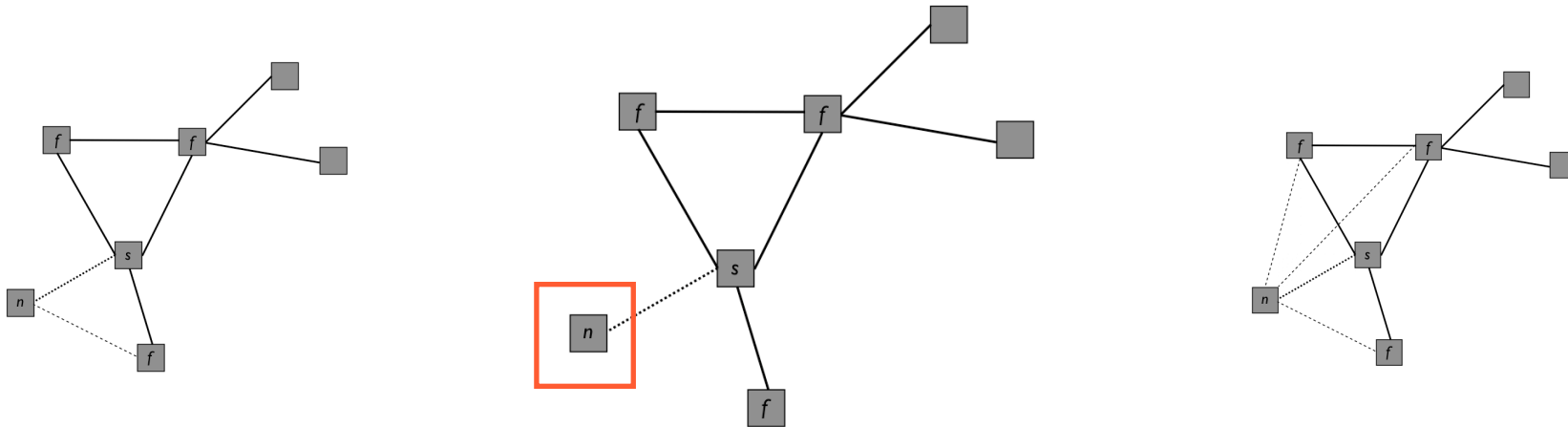
In the Ising model, spins of neighbours have a tendency to align. Similar model for the network structure?



Undirected network with copying processes

At each time step, a new node randomly connects to a target node. With probability p , it also connects to the neighbours of the target node.

For instance, the target node already has 3 neighbours:

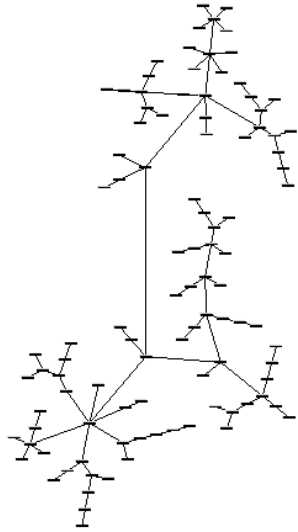


$$P = p(1 - p)^2$$

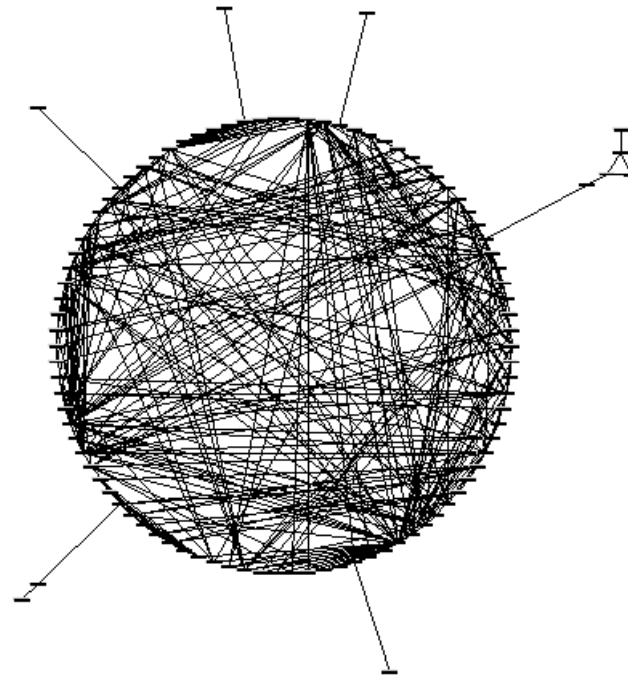
$$P = p^3 \dots\dots\dots$$



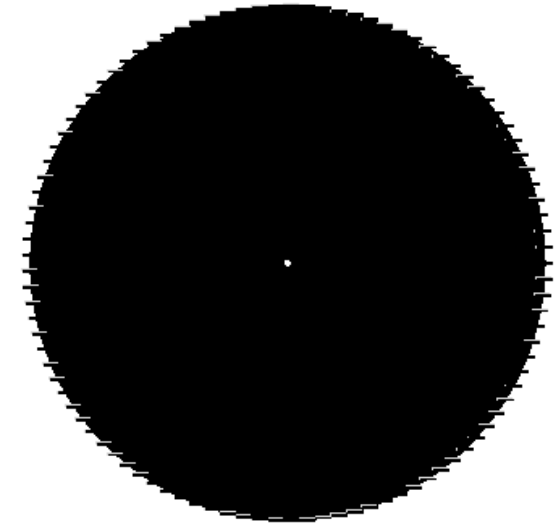
By construction, starting from 1 root node, the system grows ($N+1$ nodes after N steps) and behaves like:



Purely random network when $p=0$



What are its properties when $p \in]0, 1[$?



Fully connected network when $p=1$



The system exhibits transitions for different characteristics.
 For instance, the total number of links behaves like:

$$L(N) = \begin{cases} (1 - 2p)^{-1} N & \text{for } p < 1/2; \quad \longrightarrow \text{ Sparse phase} \\ N \ln N & \text{for } p = 1/2; \\ A(p) N^{2p} & \text{for } 1/2 < p \leq 1. \quad \longrightarrow \text{ Dense phase} \end{cases}$$

The degree distribution satisfies the non-trivial set of equations:

$$\frac{dN_k}{dN} = \frac{N_{k-1} - N_k}{N} + p \frac{(k-1)N_{k-1} - kN_k}{N} + m_k$$

$$m_k = \sum_{s \geq k-1} n_s \binom{s}{k-1} p^{k-1} (1-p)^{s-k+1}$$

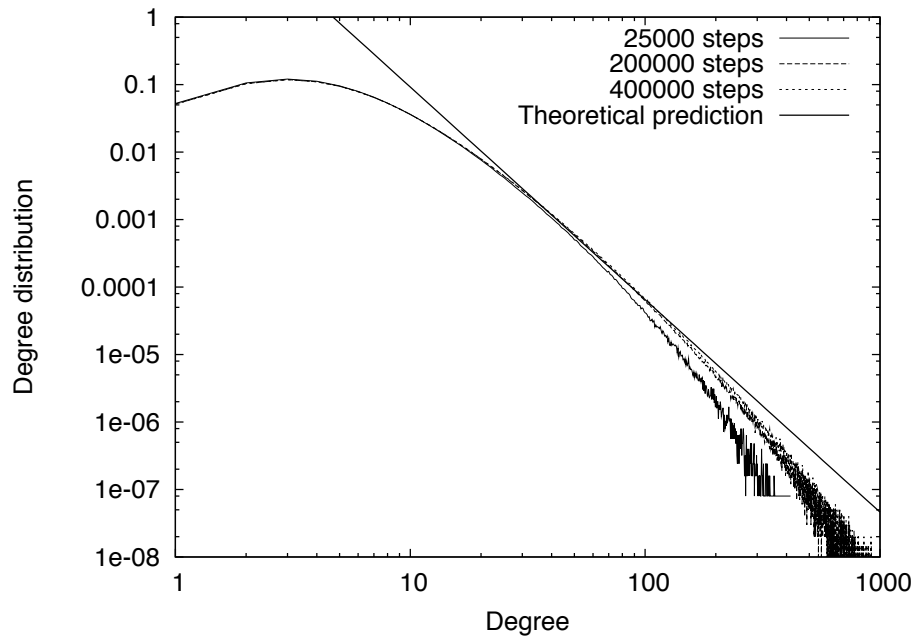
When $p < 1/2$, the asymptotic solution is stationary:

$$n_k \sim k^{-\gamma} \quad \text{as } k \gg 1$$

$$\gamma = 1 + p^{-1} - p^{\gamma-2}$$



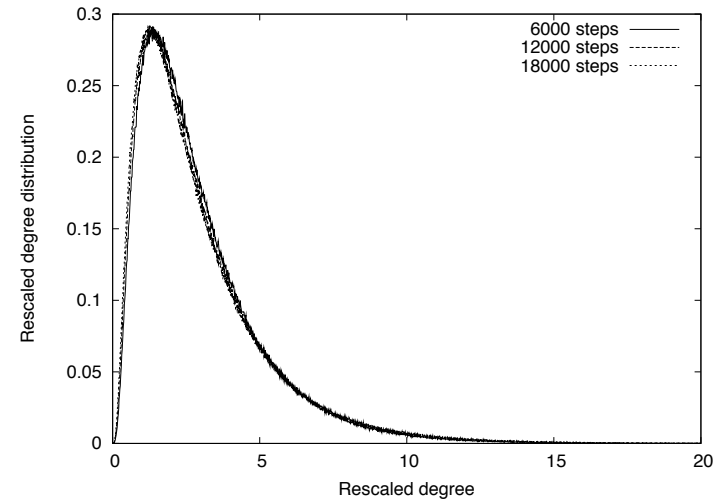
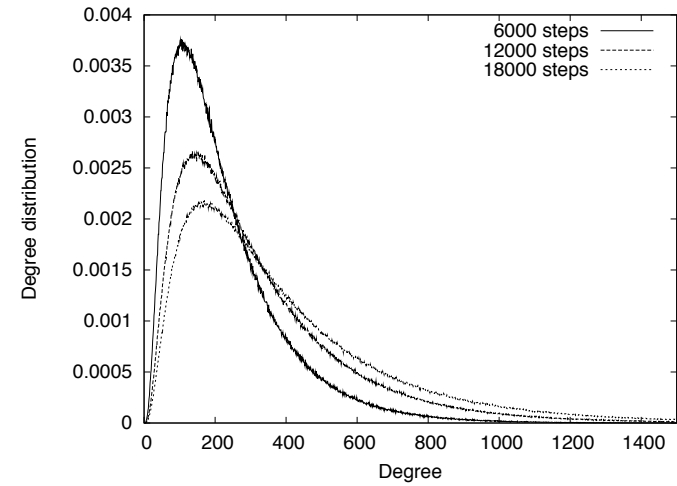
$$p < 1/2$$



Asymptotic stationary solution



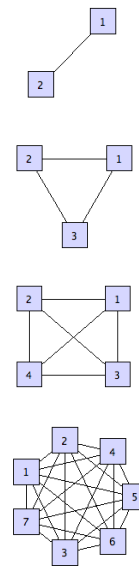
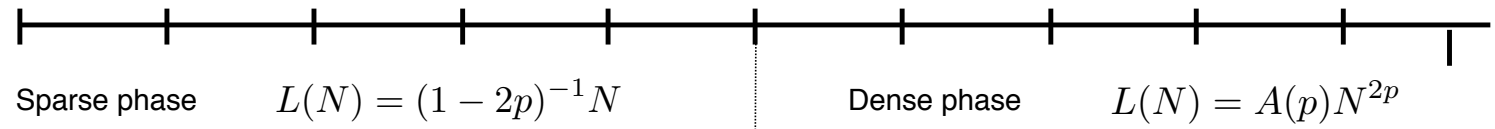
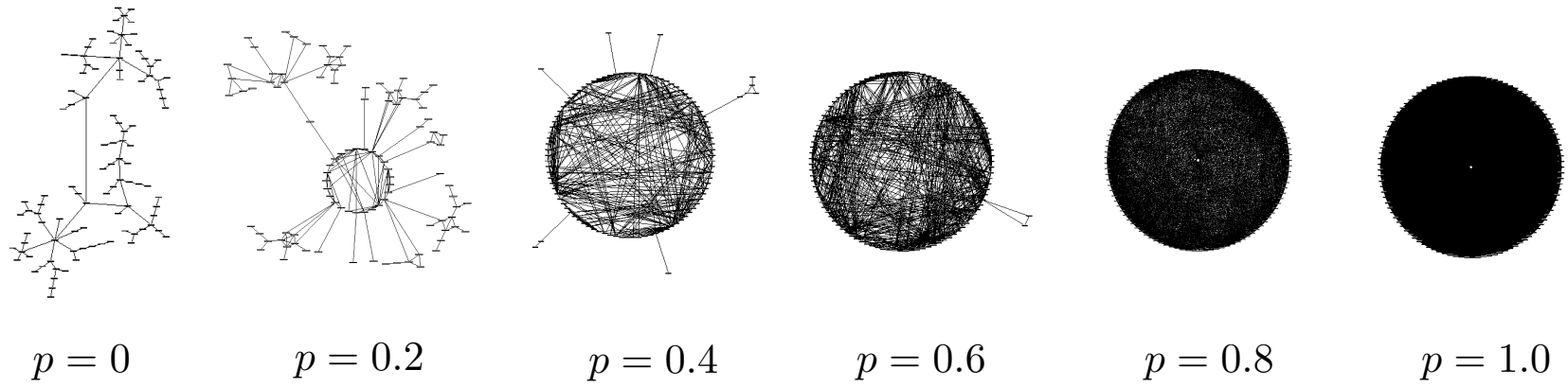
$$p > 1/2$$



Asymptotic non-stationary
self-similar solution



The network structure has been studied by looking at the average number of complete sub-graphs in the network



Sparse phase	$L(N) = (1 - 2p)^{-1}N$	Dense phase	$L(N) = A(p)N^{2p}$
$T(N) \sim L(N)$		$T(N) \sim L(N)^{\frac{3p^2}{2p}}$	
$Q(N) \sim T(N)$		$Q(N) \sim T(N)^{\frac{4p^3}{3p^2}}$	

$M_m \sim M_{m-1}$... with transitions at $p_C = 1 - \frac{1}{m}$



Distance distribution in the “*all-or-none*” version of the model: with probability p , all the neighbours of the target node are linked; otherwise, none is linked.

$$P(d, N + 1) = \left[1 + \frac{2p}{N} \right] P(d, N) + \frac{2(1 - p)}{N} P(d - 1, N)$$

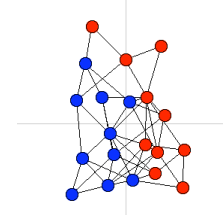
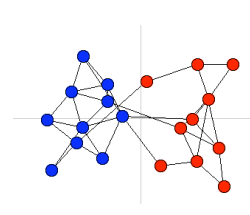
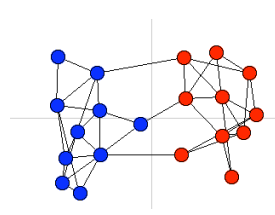
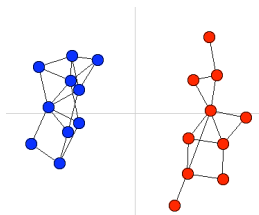
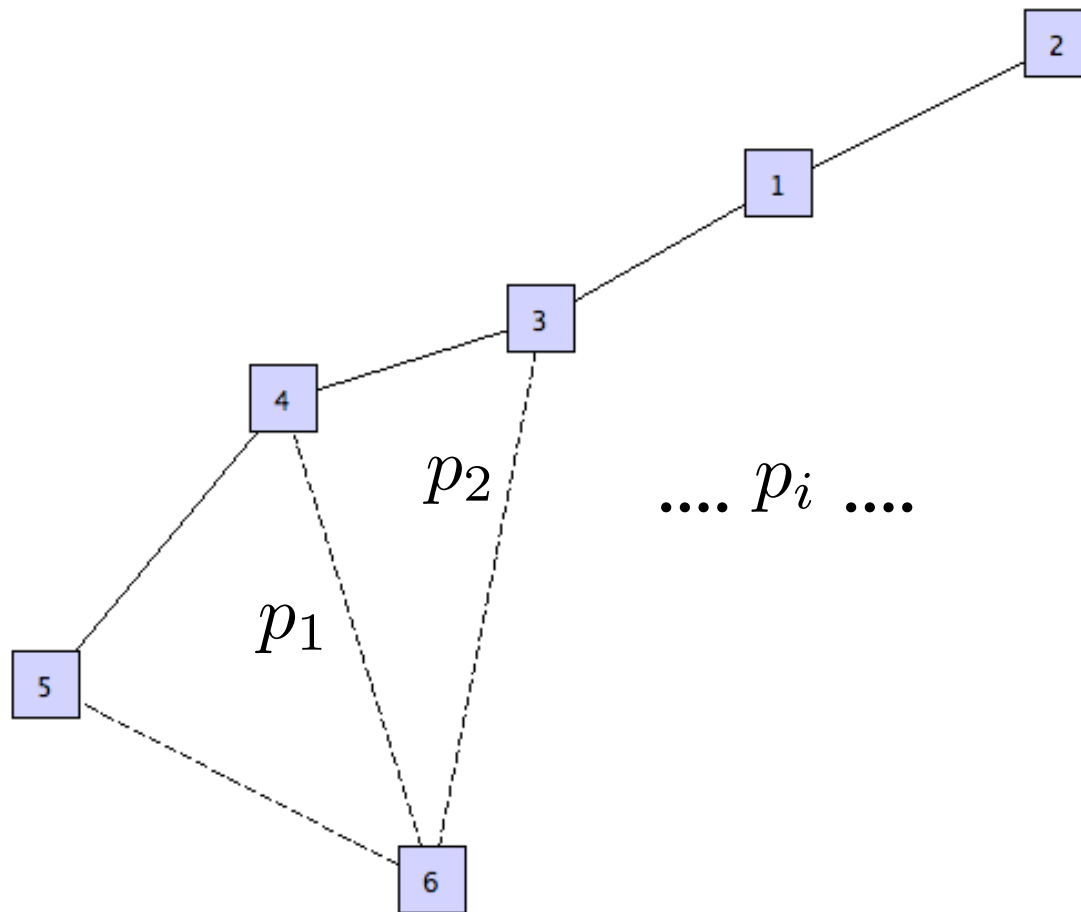
From which the distance moments can be found.
The average distance asymptotically behaves like:

$$d \sim 2(1 - p) \ln N$$

while the fluctuations are gaussian and asymptotically vanish.

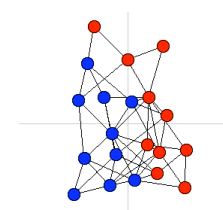
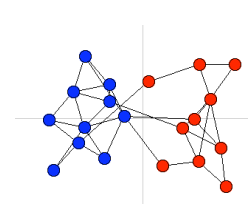
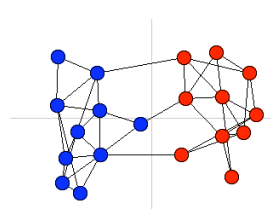
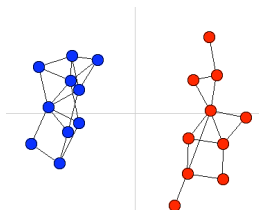
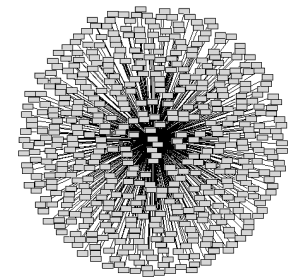
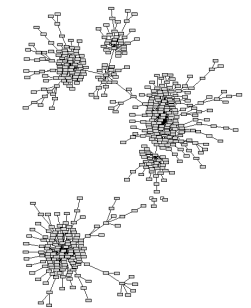
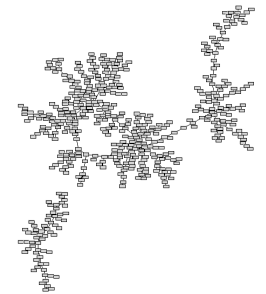
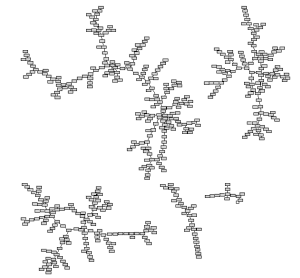


Interactions at larger distances



Perspectives

- New results on the distance statistics in the network: application to random recursive trees
- Generalization to longer distance interactions
- Model with redirection: condensates
- Comparison with empirical data: citation networks, friendship networks
- ...



Visualizing network dynamics

A good visualizing tool should aggregate nodes which are linked, and separate nodes that are not linked.

Similarities with physical system:

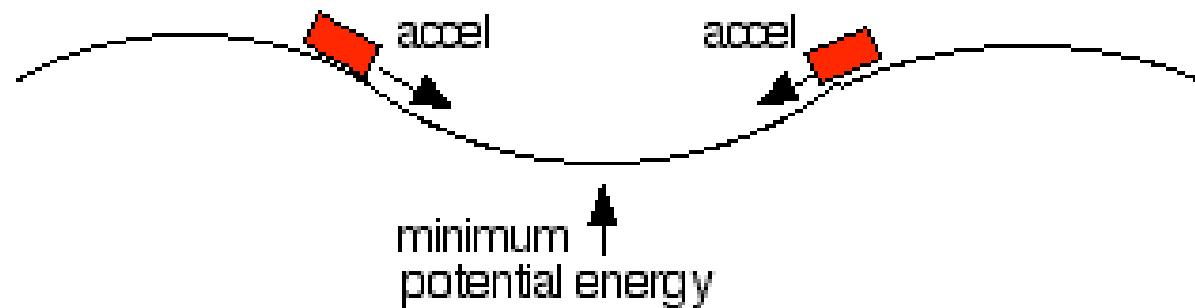
node \sim atoms

links \sim attractive forces, no links \sim repulsive forces

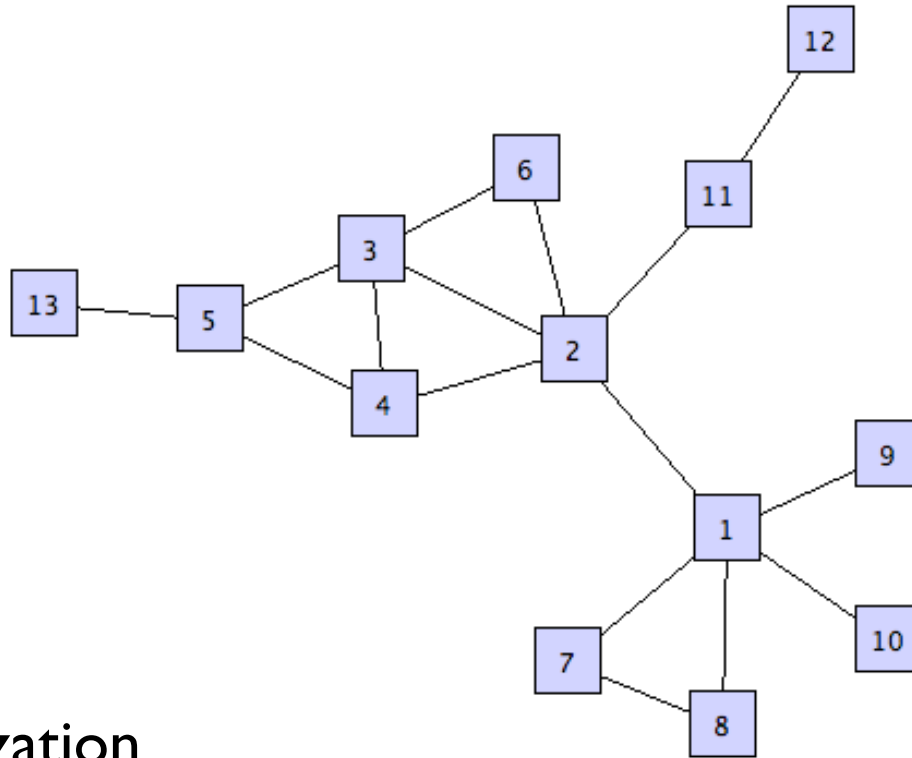
2 kinds of simulations

Looking for the minima of
some energy

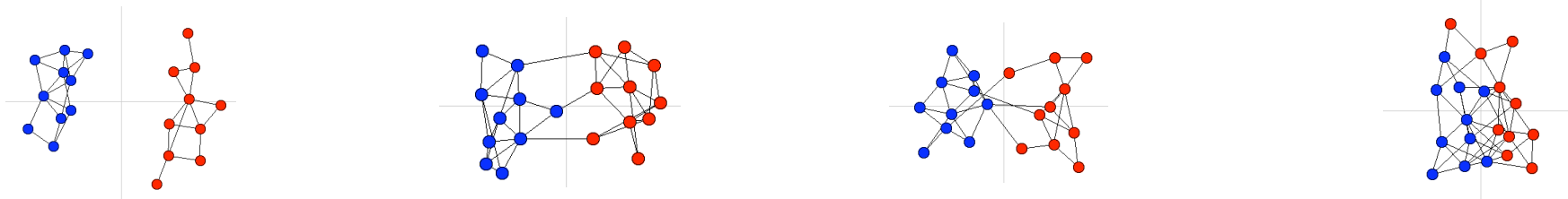
Solve the equation of
motion of the atoms



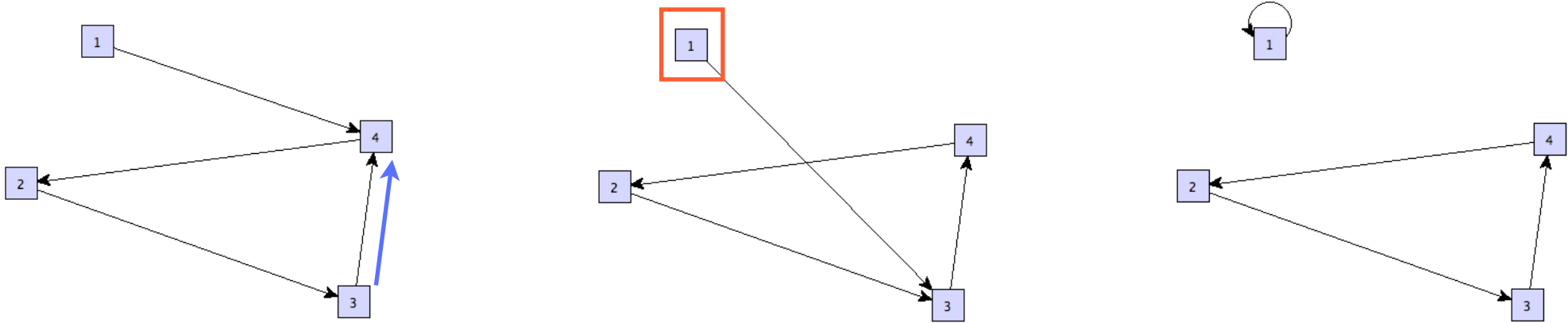
Contrary to classical systems, the interactions between nodes may evolve in time, if links evolve!



Static visualization
using Visone



Application to a model with redirection, with fixed number of nodes and links.



p go to the “grand-father”

$(1 - p)$ connects to a node chosen randomly

small p , random structures

high p , condensation around 1 node

