- A simple model for collaboration networks in multidisciplinary fields

- Endo- vs Exo-genous shocks in sales and blog trends.

- Music/Fans as a paradigm for Bipartite Networks

A simple model for collaboration networks in multidisciplinary fields

- Model the interplay and interaction between scientists of different fields, like physics, informatics, sociology...
- Characterize the interface between the two *phases*.
- Model the influence of an external field (political decision).

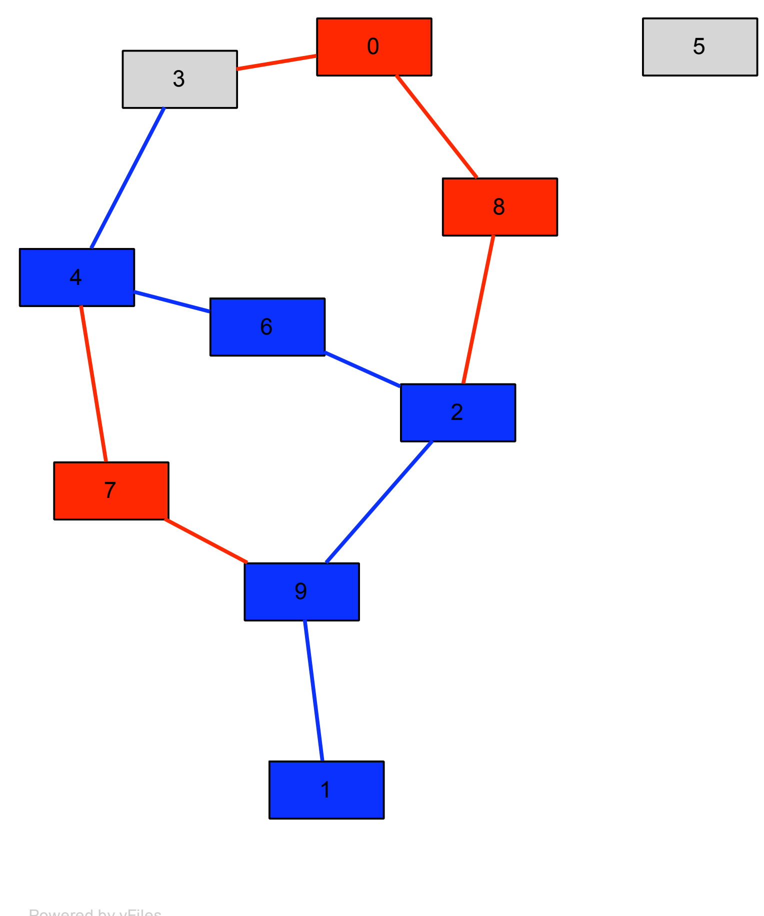
R. Lambiotte, M.Ausloos
(Supratecs, ULG, Belgium)

CREEN

Model ingredients:

- Agent-based model
- There are 2 possible kinds of collaboration, A and B, between **two** scientists (~ agents).

The collaboration network is a coupled network, i.e. a network where nodes (scientists) are related by two kinds of links (collaborations).



The state of the nodes is fixed by their links:
- majority of A collaborations => A scientist
- majority of B collaborations => B scientist
- equal number of A and B collaborations => A-B scientist

Model ingredients:

- As simple as possible: avoid complications due to non-stationary effects => constant number of nodes **and** of links.

- Monte-Carlo simulation:

  - At each time step, we remove one link, A or B, at random.

  - We add a new link between 2 randomly selected nodes, i and j

  - The kind of the added link, A or B, depends on the previous links of i and of j.

  - To do so, we calculate the proportions of links A for i and for j

$$p_A^i = \frac{N_A^i}{N^i} \qquad p_A^j = \frac{N_A^j}{N^j}$$

These quantities measure the *capacity* of i/j to work in the field A

The *capacity* of the pair is by definition the average:

$$p_A^{ij} = \frac{p_A^i + p_A^j}{2} \qquad 0 \le p_A^{ij} \le 1$$

Therefore, if: $p_A^{ij} > \frac{1}{2}$ → The selected pair should collaborate in the field A

$p_A^{ij} < \frac{1}{2}$ → The selected pair should collaborate in the field B

We implement this mechanism with the stochastic rule:

With probability $P = \frac{e^{\frac{(p_A^{ij} - p_D)}{T}}}{Z}$ the collaboration is A-type

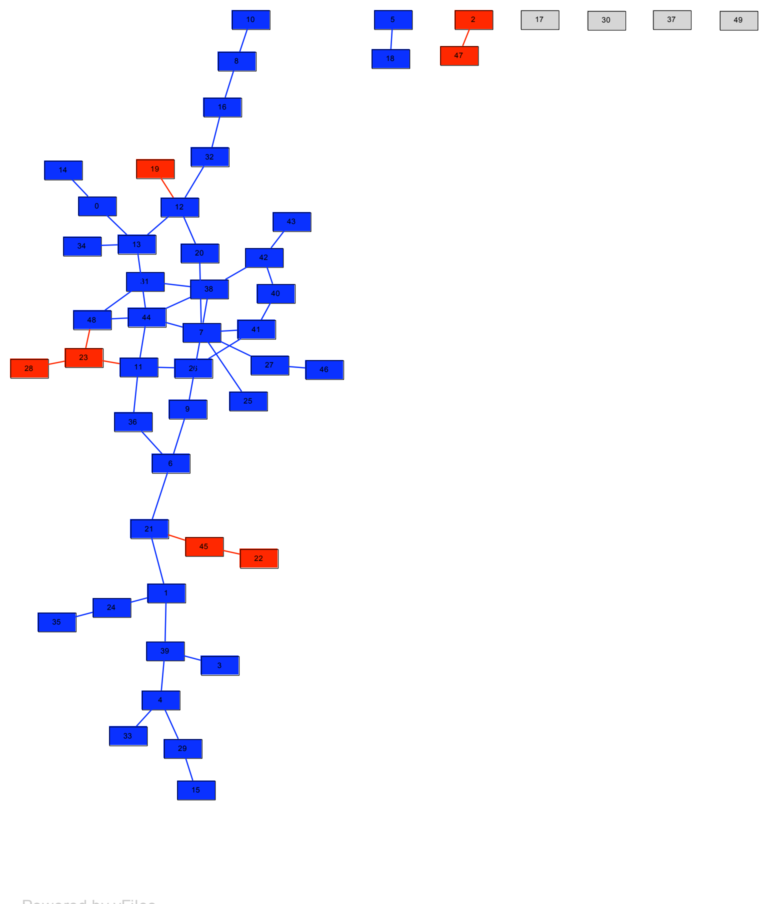
With probability $P = \frac{e^{\frac{(p_D - p_A^{ij})}{T}}}{Z}$ the collaboration is B-type

Where: $Z = e^{\frac{(p_A^{ij} - p_D)}{T}} + e^{\frac{(p_D - p_A^{ij})}{T}}$ is a normalizing constant

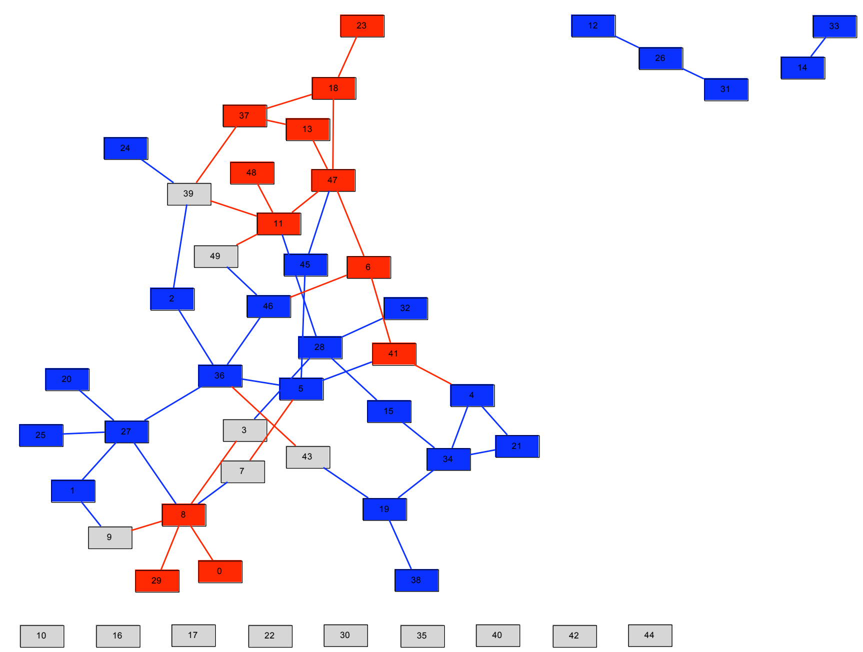
T plays the role of a temperature, ~ agitation, curiosity of the agents

p_D is a drift term, that breaks the internal symmetry => external field (political decision)

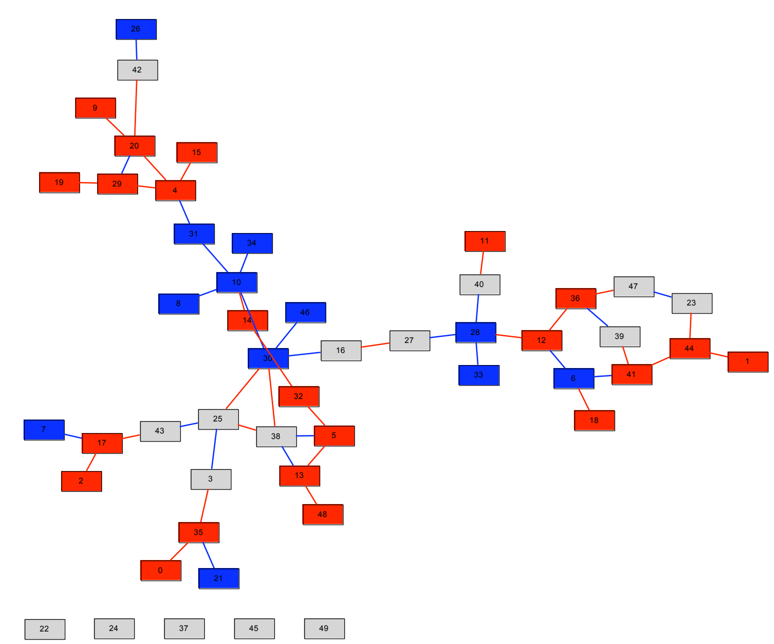
Typical asymptotic configurations, for small (50 agents) simulations



Low temperature

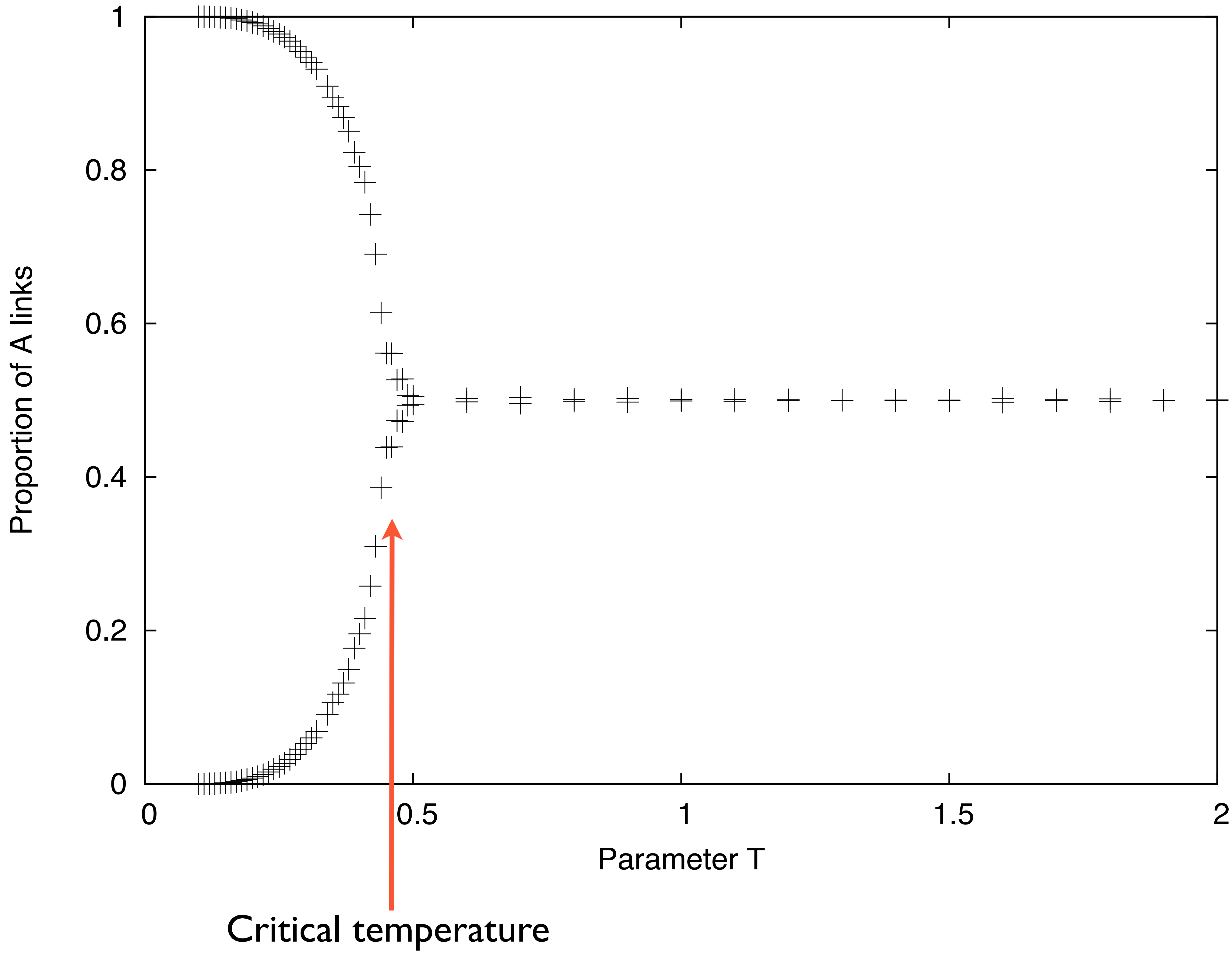


Medium temperature



High temperature

Bifurcation diagram (1000 agents, 10 links/agent), without external field ($p_D = \frac{1}{2}$).



In order to characterize the interface between the 2 networks, we calculate the *overlap coefficient*, defined by:

$$\Omega = \frac{< N_A^i N_B^i >_i}{< N_A^i >_i < N_B^i >_i} - 1$$

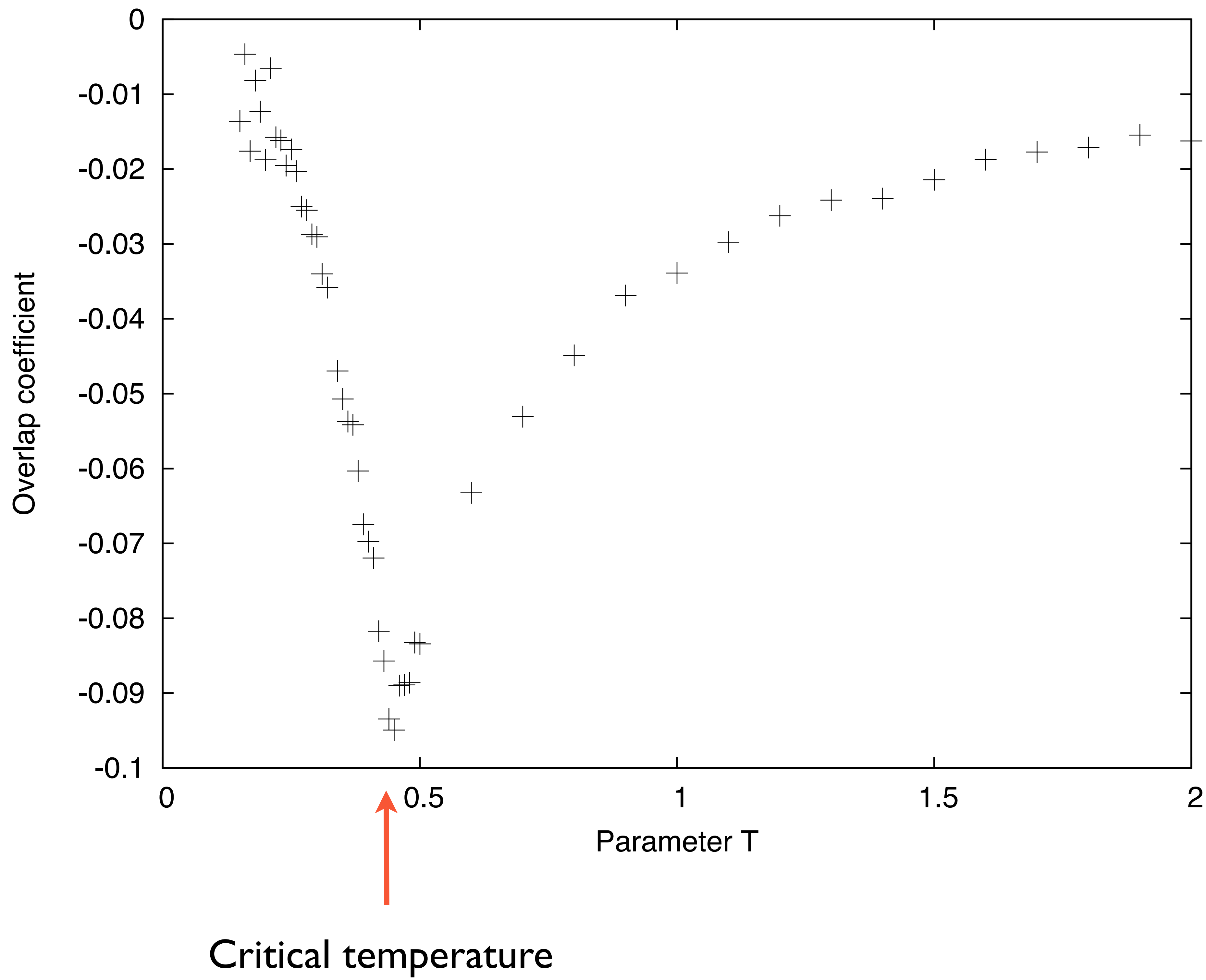
$\Omega = 0$ → A and B are independent networks

$\Omega < 0$ → Few actors work in A and B simultaneously

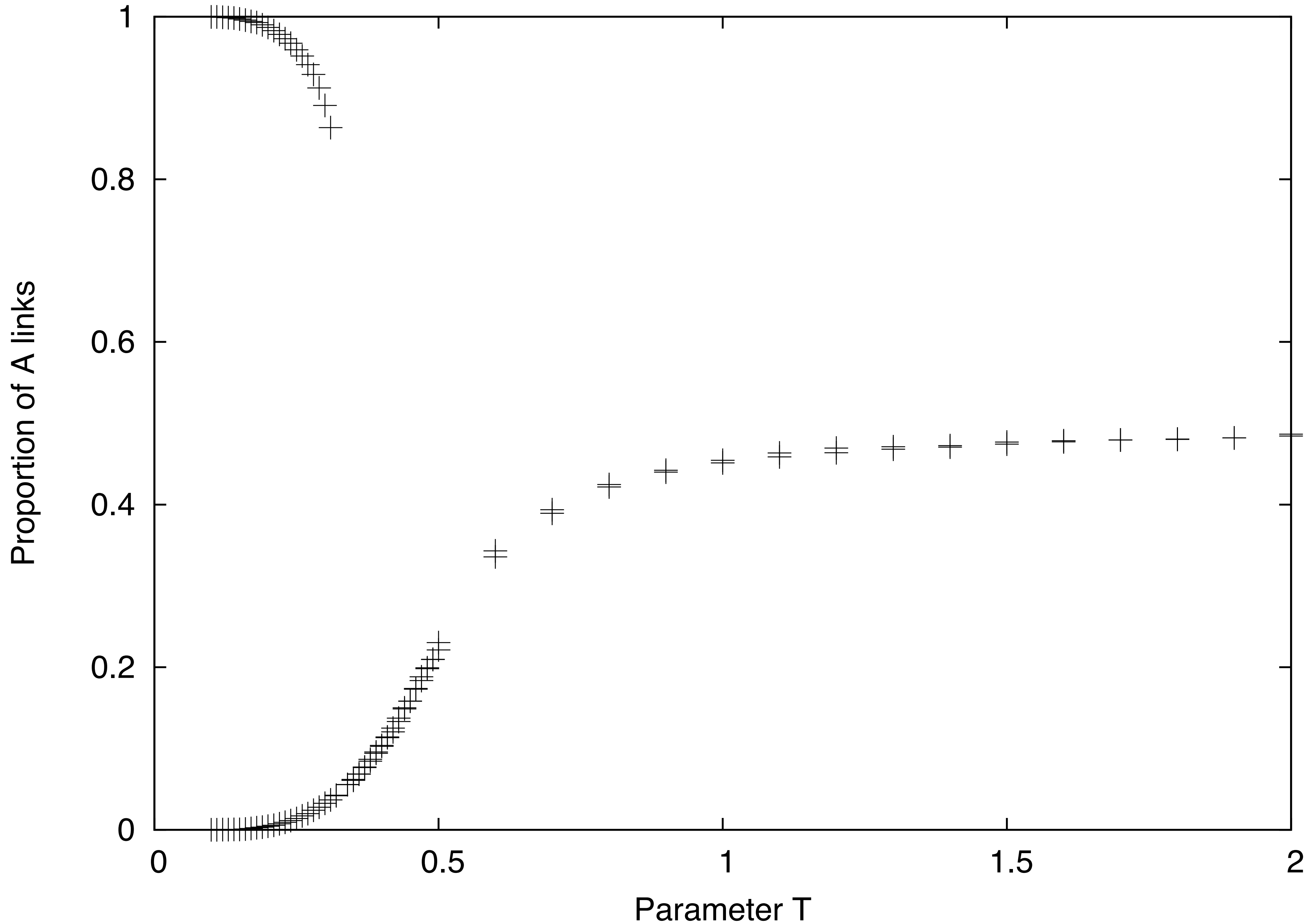
↓

2 well separated phases, where some scientists are an interplay between the two fields.

Bifurcation diagram for the overlap coefficient



Bifurcation diagram (1000 agents, 10 links/agent), with an external field ($p_D = 0.55$).



Metastability, hysteresis....

Theoretical treatment

Three assumptions

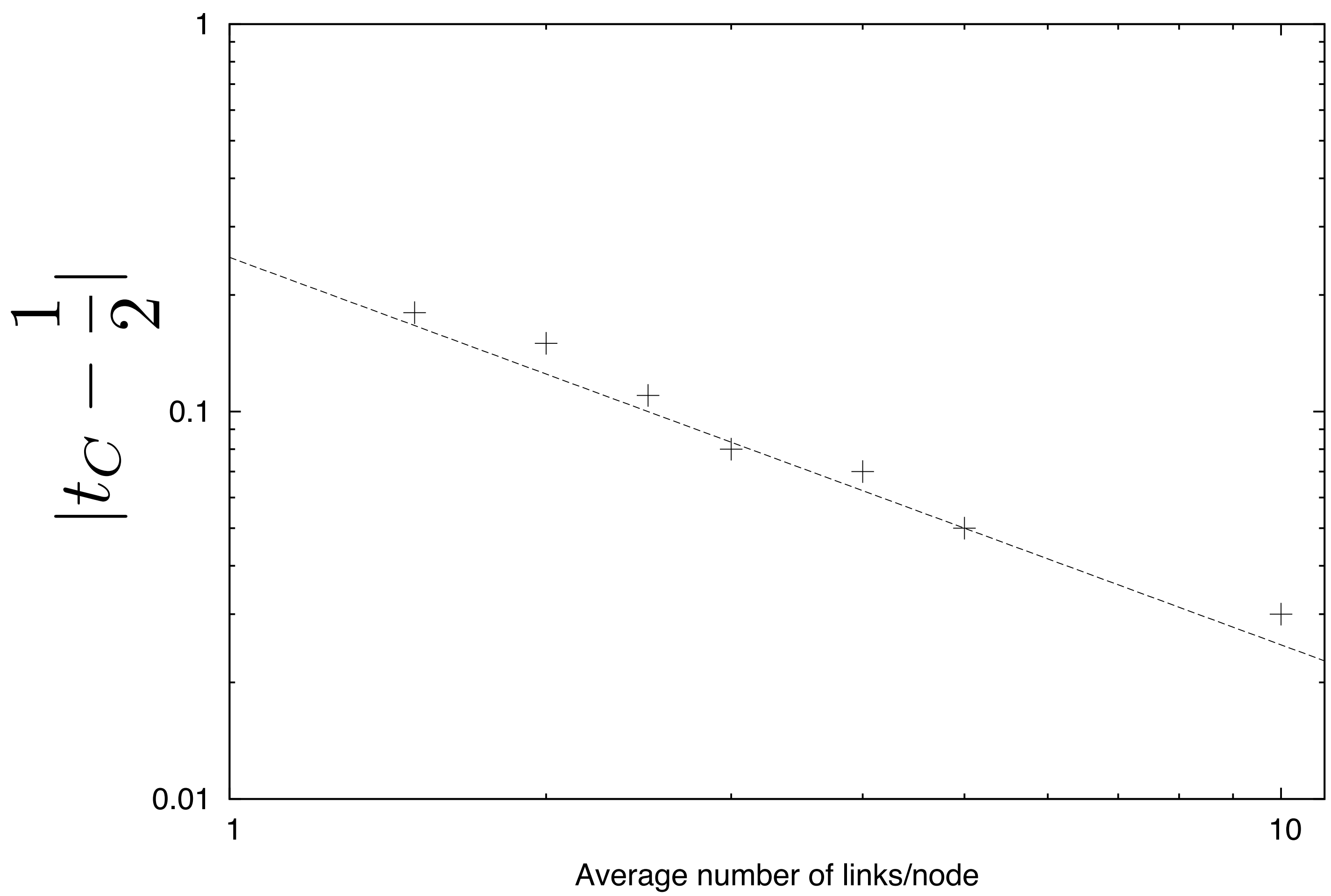
Fluctuations of the number of links/node are negligible

Mean field approximation

Stationary solution satisfies detailed balance

Then, the stability of symmetric state derives straightforwardly from *Free Energy* principles.

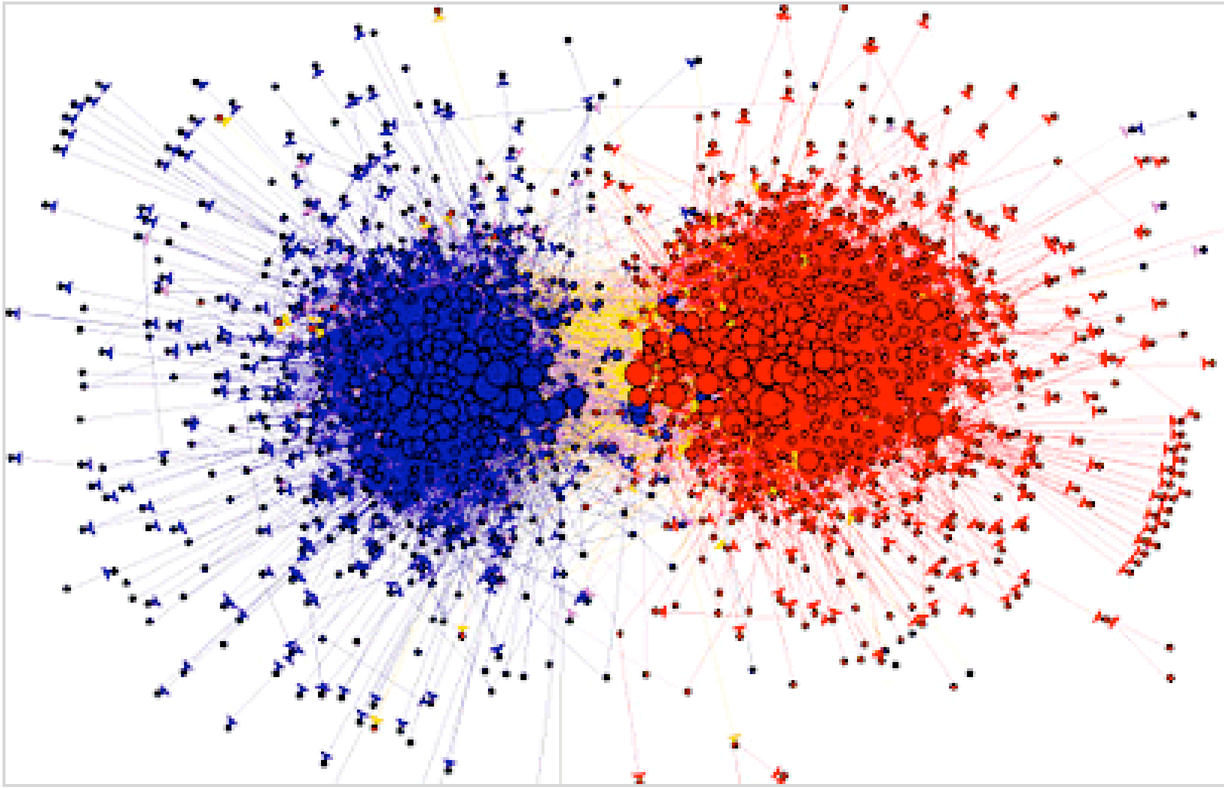
$$\boxed{T_C = \frac{1}{2}}$$



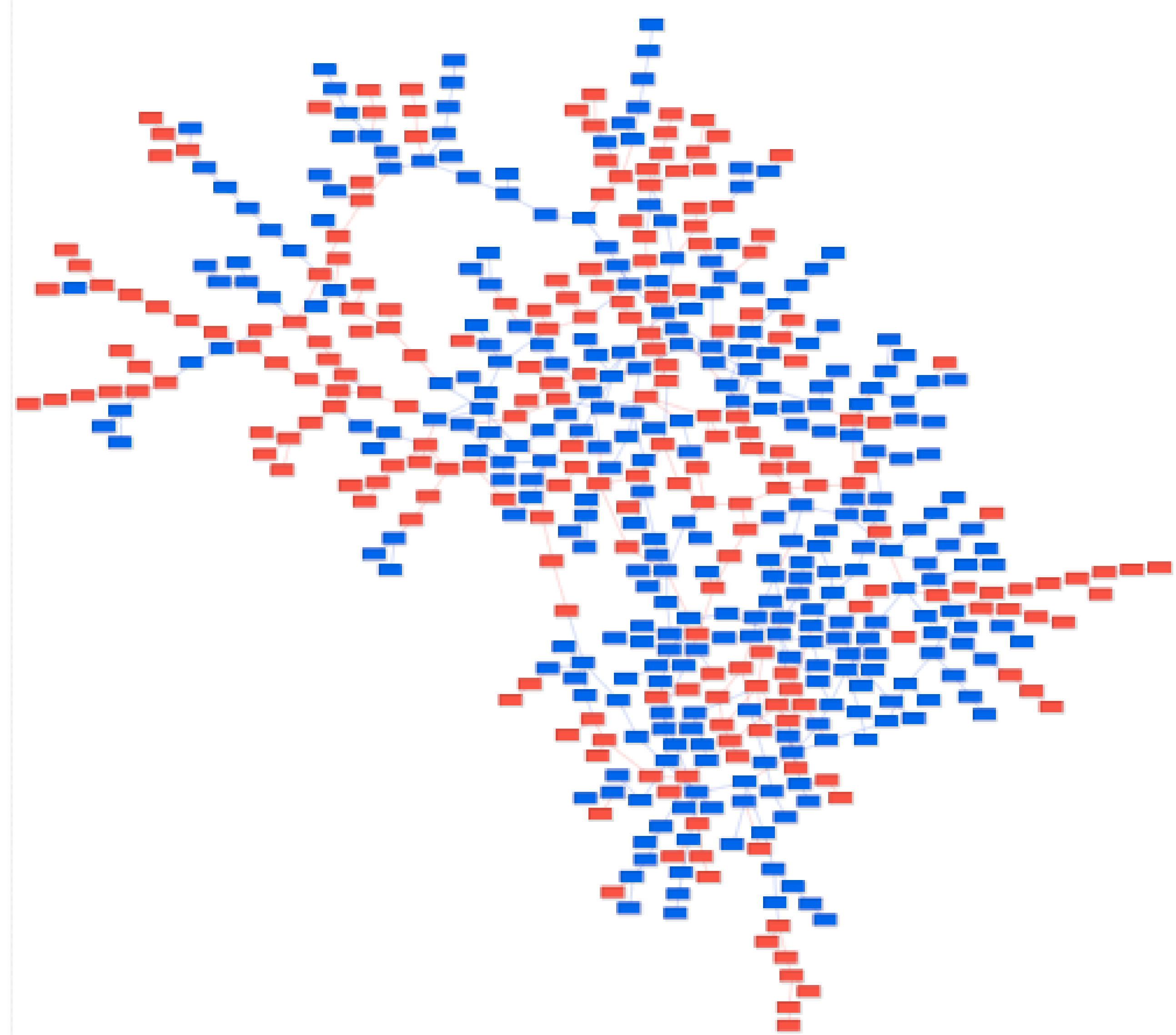
Conclusion...

- Very simple *stationary* model for connected multidisciplinary scientists.
- Qualitatively, its features are those of an Ising model for magnetic systems, i.e. hysteresis, Curie temperature...
- Favouring one of the fields, A or B, breaks the internal symmetry, and leads to metastability.
- Decreasing the temperature leads to the formation of two distinct phases.
- Mean field theoretical predictions
- What's next: theoretical framework (canonical formalism, phase transition, in progress), extension to *k* scientific fields, to open systems, growing networks...

... and comparison with experimental data!!



Links between conservative and liberal blogs (L. Adamic and N. Glance, *blogpulse.com*)



Asymptotic simulation configuration, *T=0.51*

Endo- vs Exo-genous shocks in sales and blog trends.

- Applicability of the fluctuation-dissipation theorem to sociological systems.

- Characterize after-shocks relaxations.

- Discriminate endogenous shocks from exogenous ones.

R. Lambiotte, M.Ausloos
(Supratecs, ULG, Belgium)

CREEN

Critical events in evolving networks

Project of 6th EU Framework Programme

Fluctuation-Dissipation theorem:

Key tool of Statistical Mechanics,
that relates 2 classes of dynamical
features.

Fluctuation phenomena,
i.e. stochastic deviations
from the Equilibrium State.

Dissipative response of
the system to an external
field

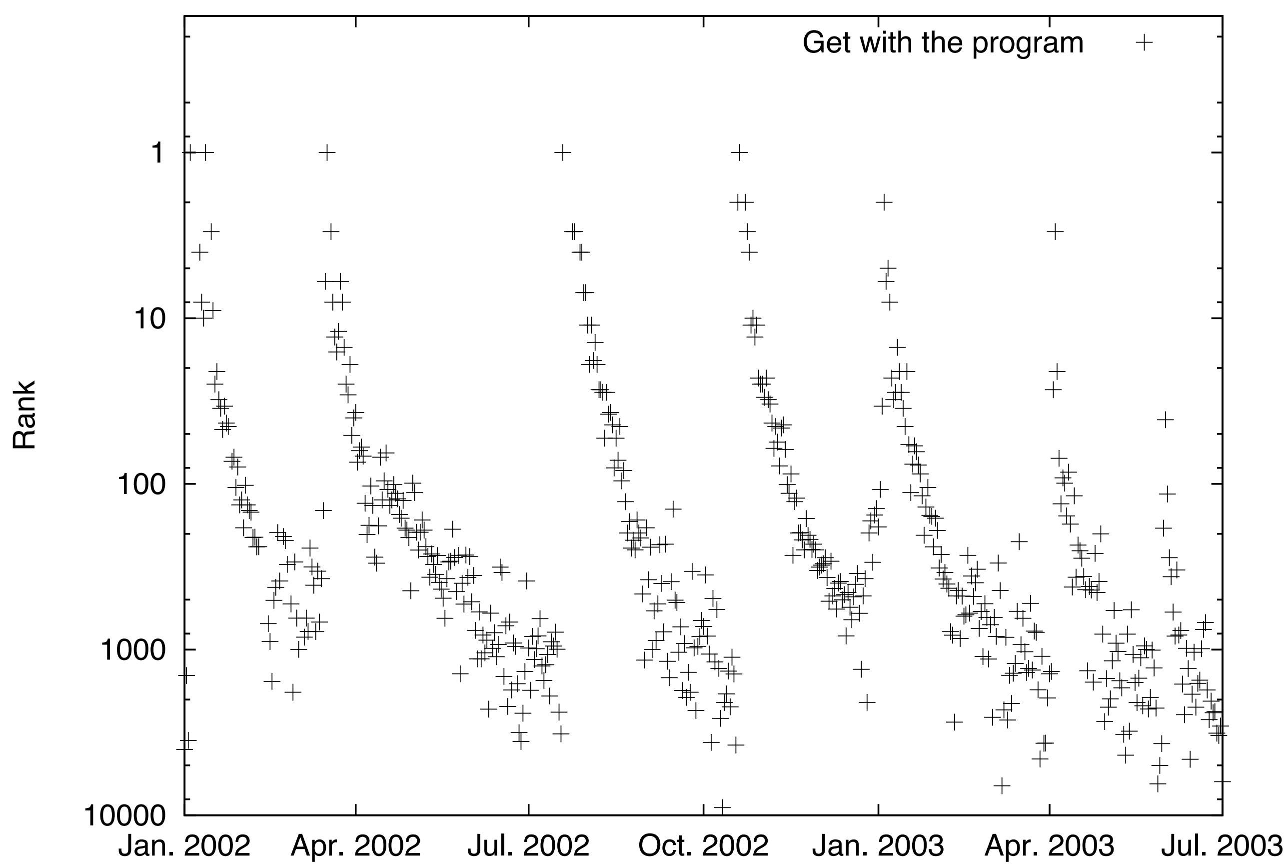
Sociological and economical systems are out-of-equilibrium

→ Subjects to outliers, bubble formation, self-emergence of trends...

Two kinds of shocks:

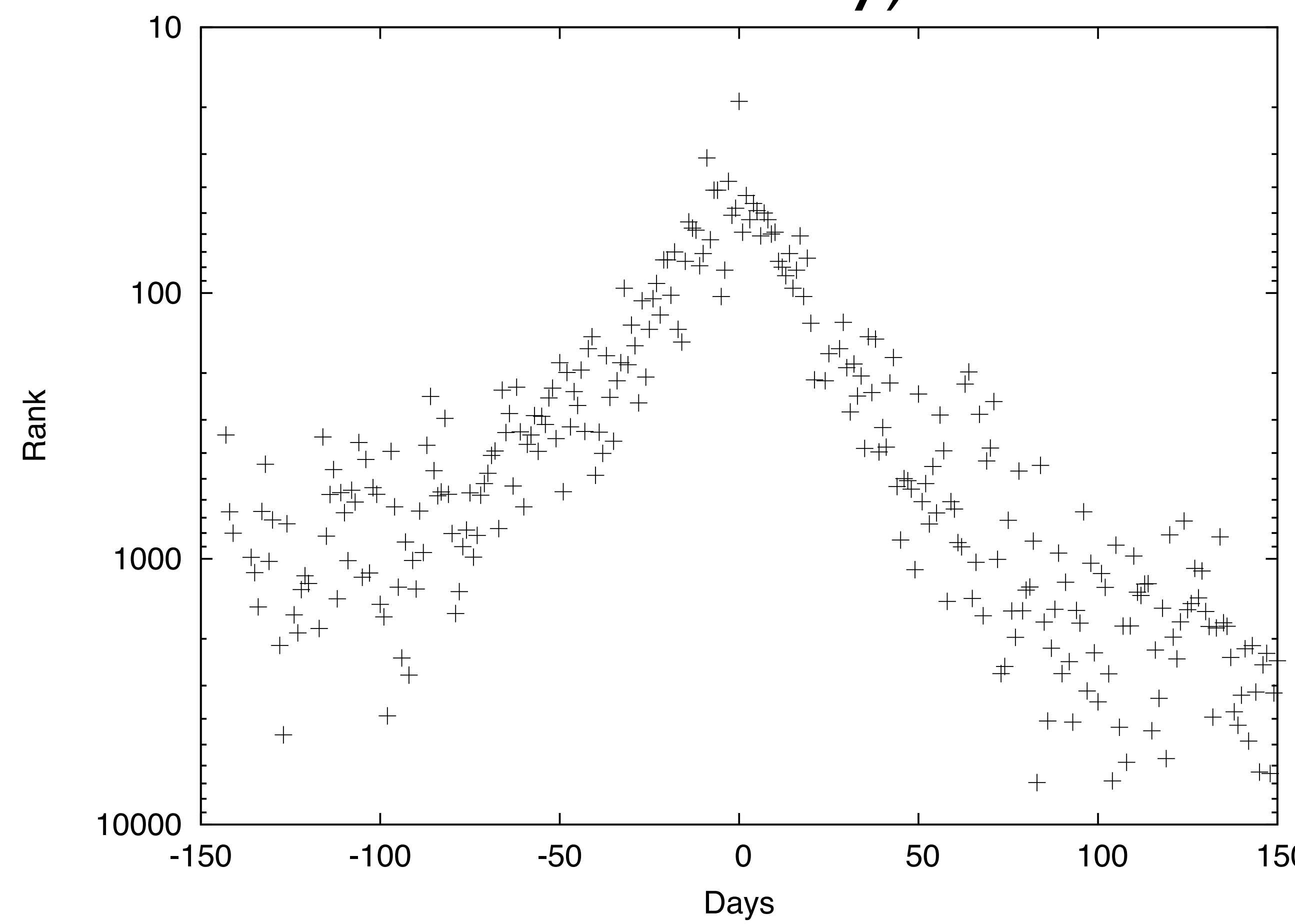
Exogenous shock

Response to some external field.



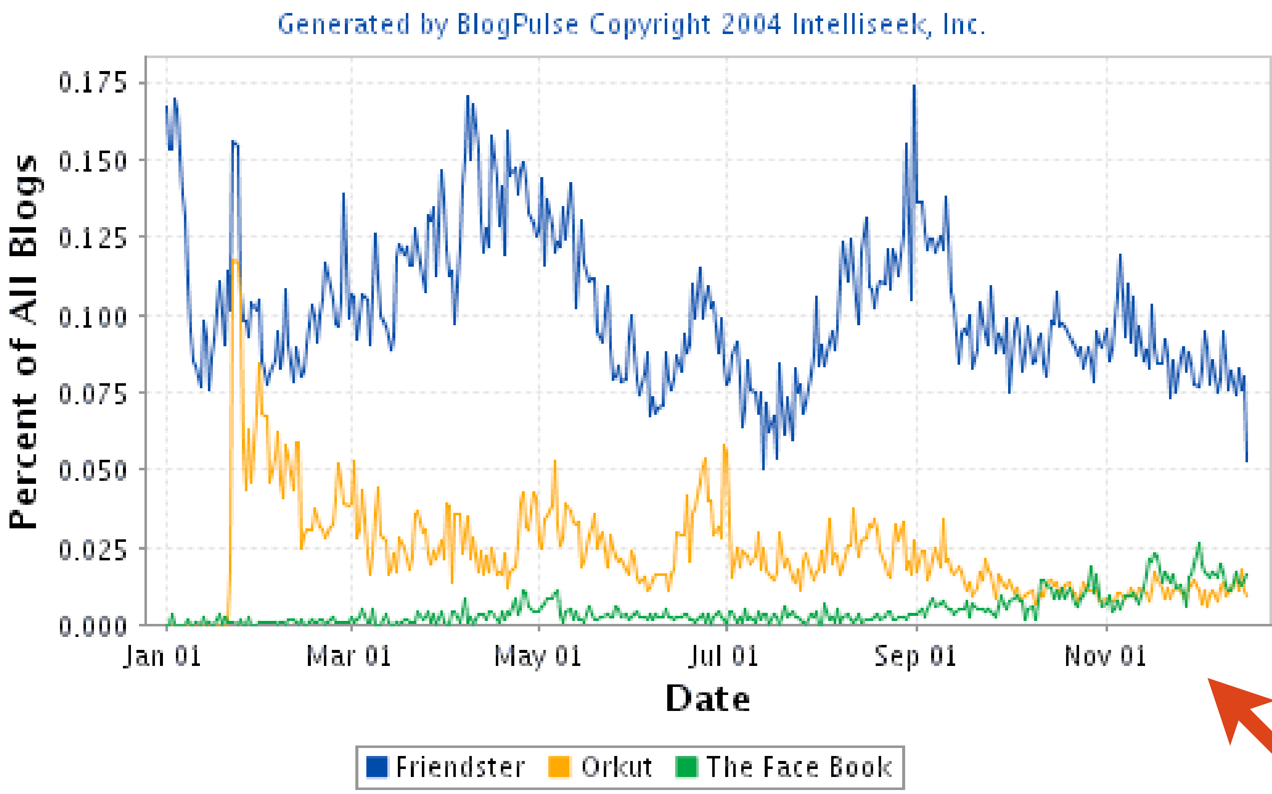
Endogenous shock

Spontaneous evolution of the system (Self-Organized Criticality)



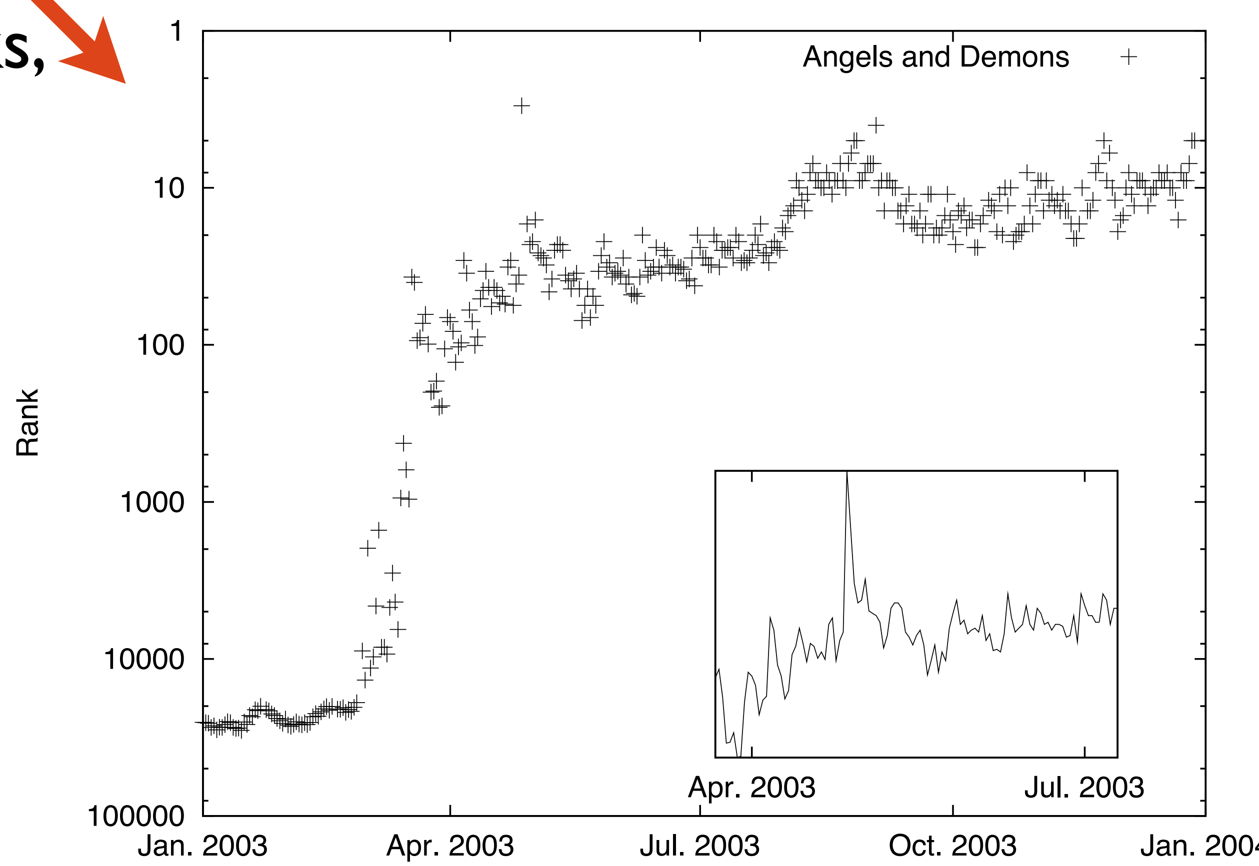
Most systems are driven by an interplay of the two mechanisms

Sales vs blog trends.



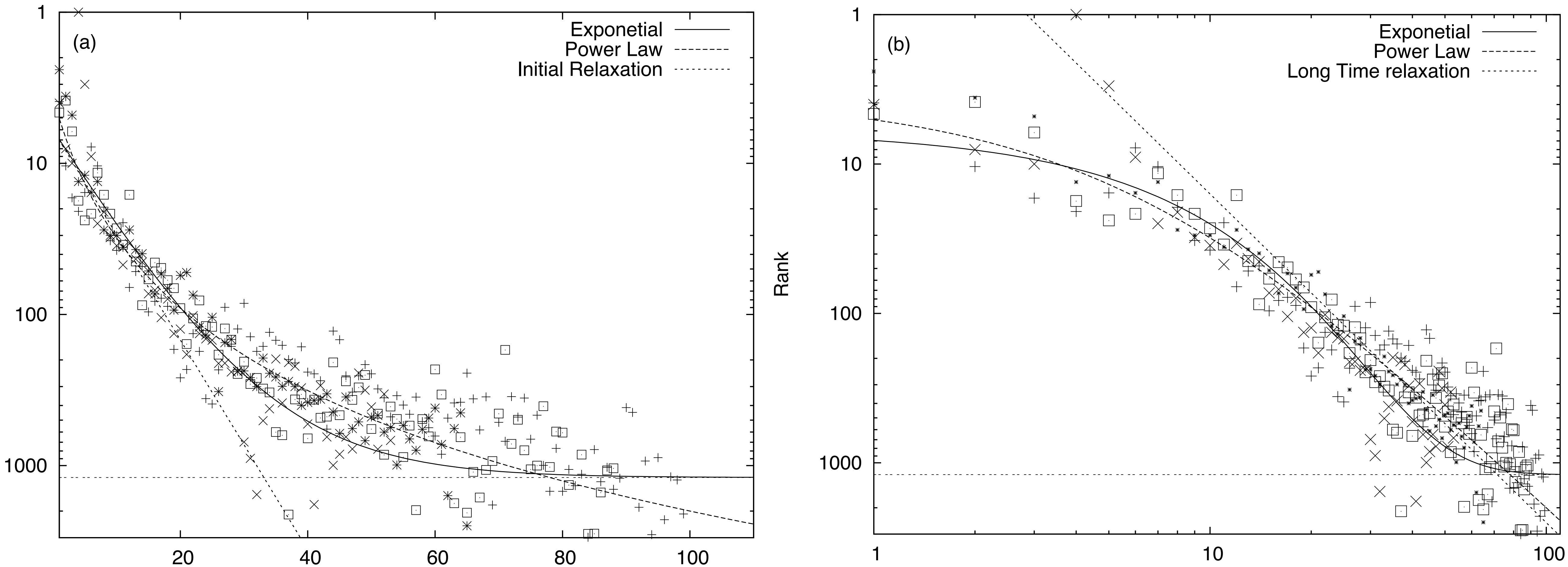
blogpulse, frequency of words

Endo-/Exo-genous shocks, correlations between signals...



junglescan, rank from Amazon sales

First requirement: Experiment reproducibility



Two possible descriptions

Lambiotte & Ausloos

Exponential relaxation + saturation

$$R = (R_\infty^{-\frac{1}{2}} + (R_0^{-\frac{1}{2}} - R_\infty^{-\frac{1}{2}})e^{-\frac{\lambda}{2}t})^{-2}$$

Friction parameter, λ

Sornette et al.

μ

Power-law relaxation

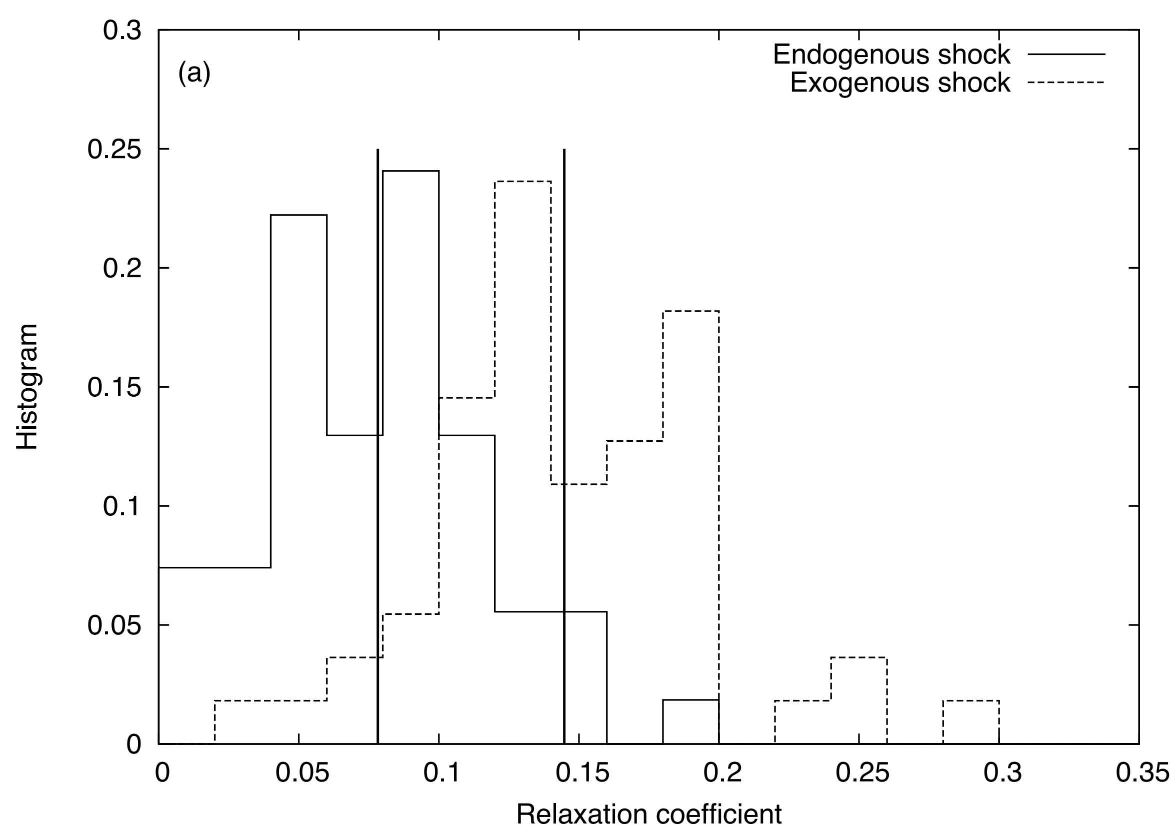
$$R = (t_c + t)^\mu$$

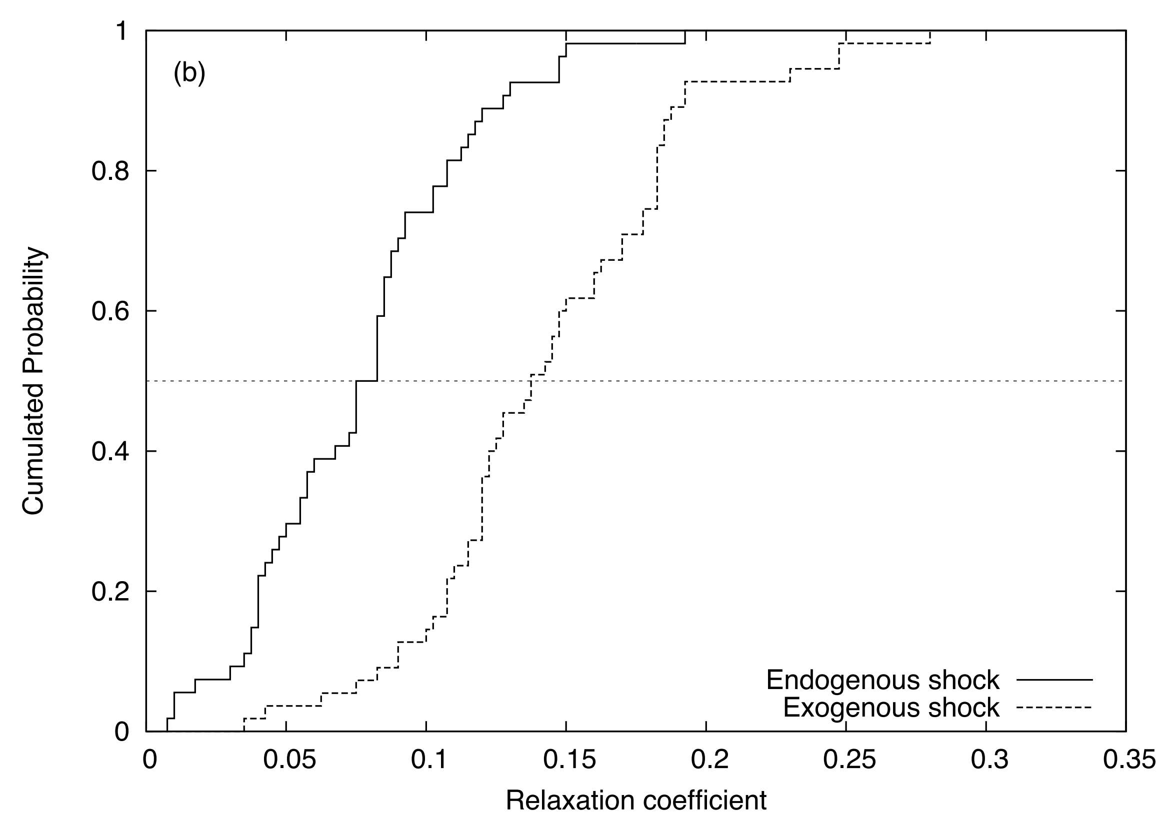
Epidemic-like model, with long memory effetcs

Universal features

Sornette et al. show that exo- and endo-genous relaxations differ on the long time scale, i.e. different exponents μ

In contrast, we discriminate shocks by their *short-time* behaviour: the relaxation time $t_R = \frac{1}{\lambda}$ seems to be twice shorter in exogenous shocks than in endogenous ones.





What is next?

- Similar study of trend formation in scientific fields (data from Ruby?)
  - Rank ~ Frequency of selected words
  - Automatic location of maxima => short-time (1-20 days) and long-time (1 month, 2 month) relaxations
- Characterization of the random signals
  - Fractal dimension, Hurst exponent, Noise Intensity
  - Link with the friction parameter? The power-law exponent?
  - Time correlations between different signals

Linear response theory

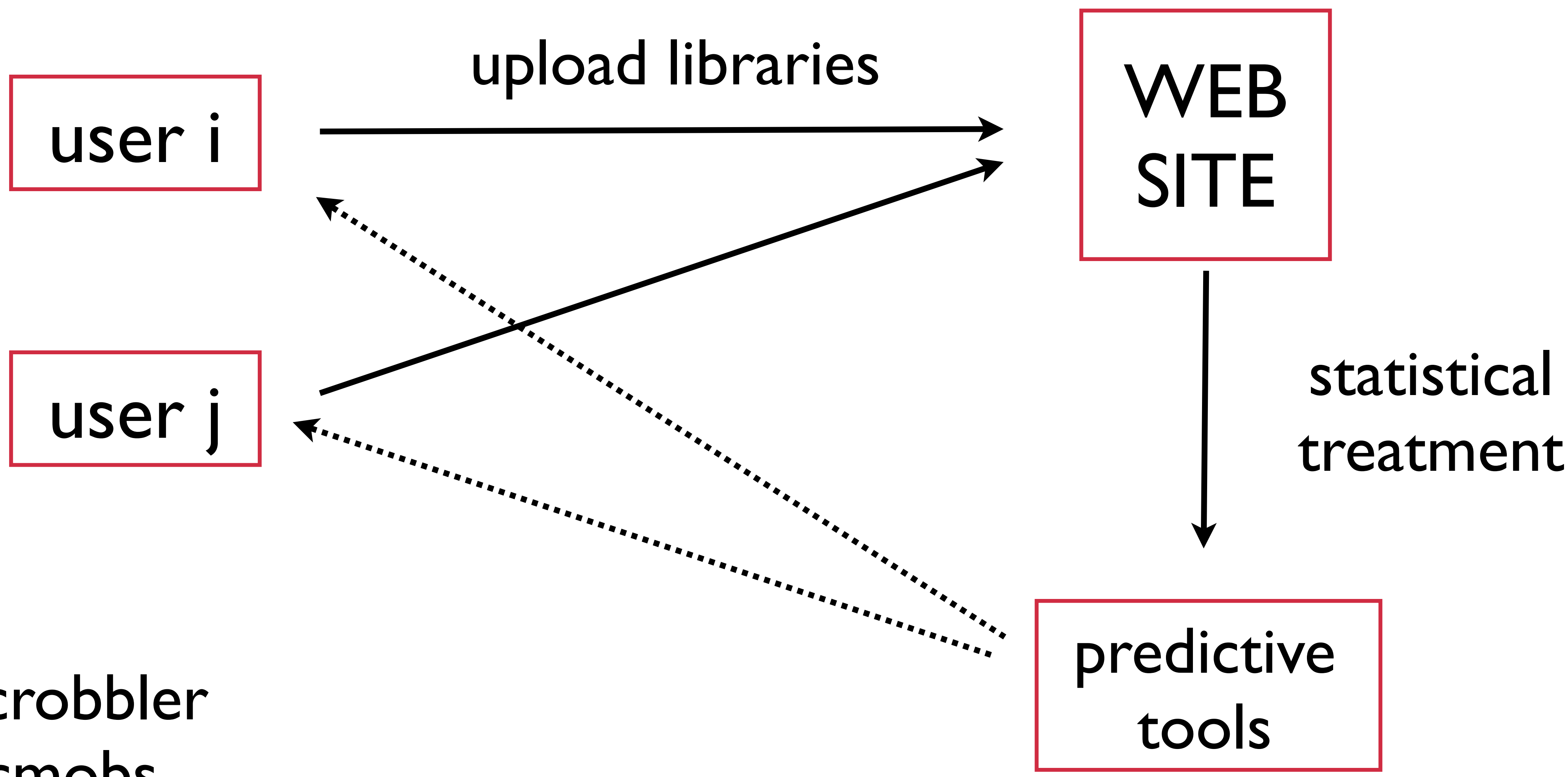
Music/Fans as a paradigm for Bipartite Networks

- Network with strong sociological behaviour
- Bipartite structure (scientist/article)
- Large available databases
- Evolving structures, trends, avalanches

R. Lambiotte, M.Ausloos
(Supratecs, ULG, Belgium)

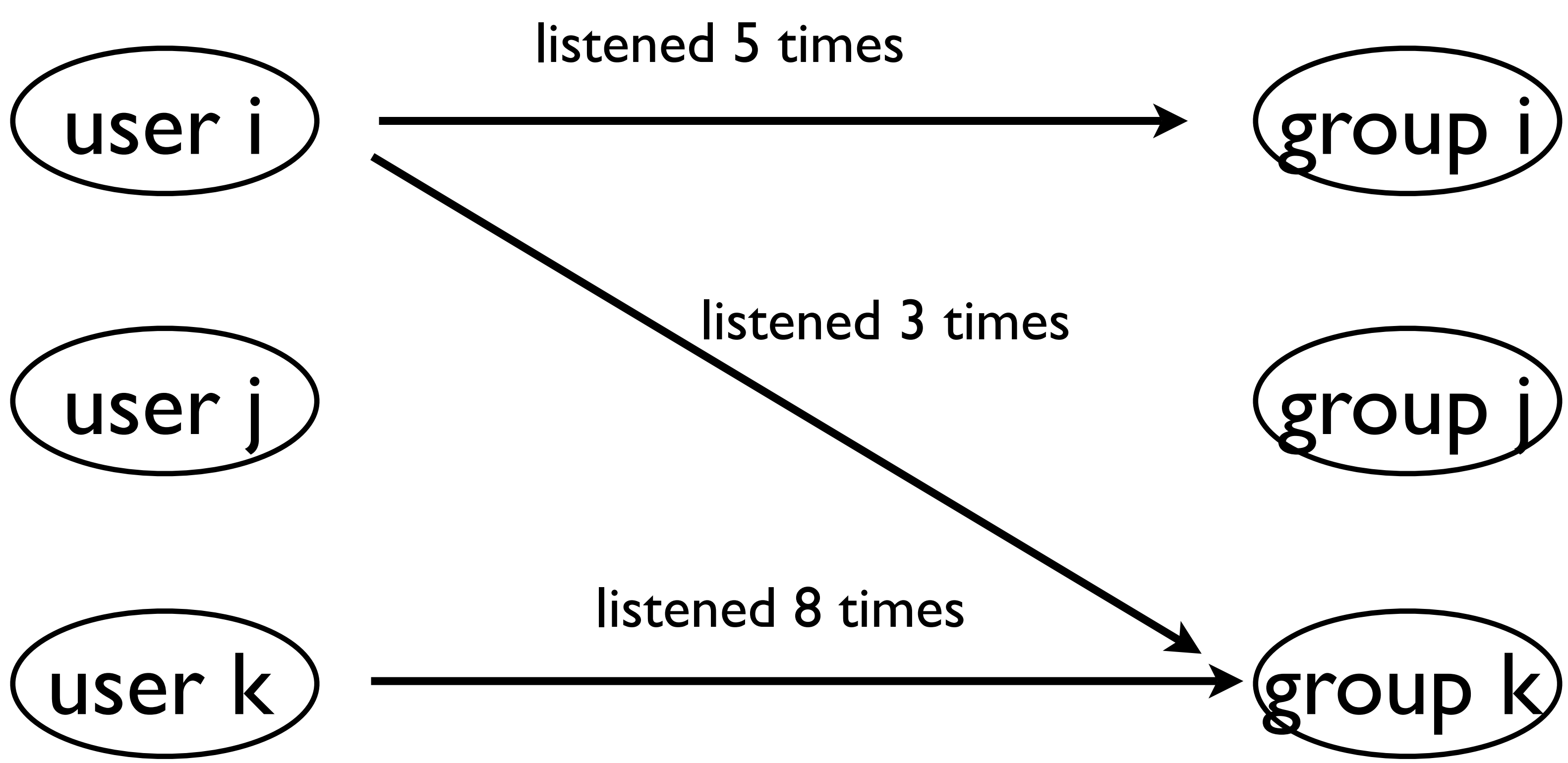
Cooperative filtering

Recently (1 year), new free services on the web:



audioscrobbler
musicmobs
...

Typical bipartite graph



→ Projected network

→ Structure of music trends, genres

→ Sociological structure of listeners

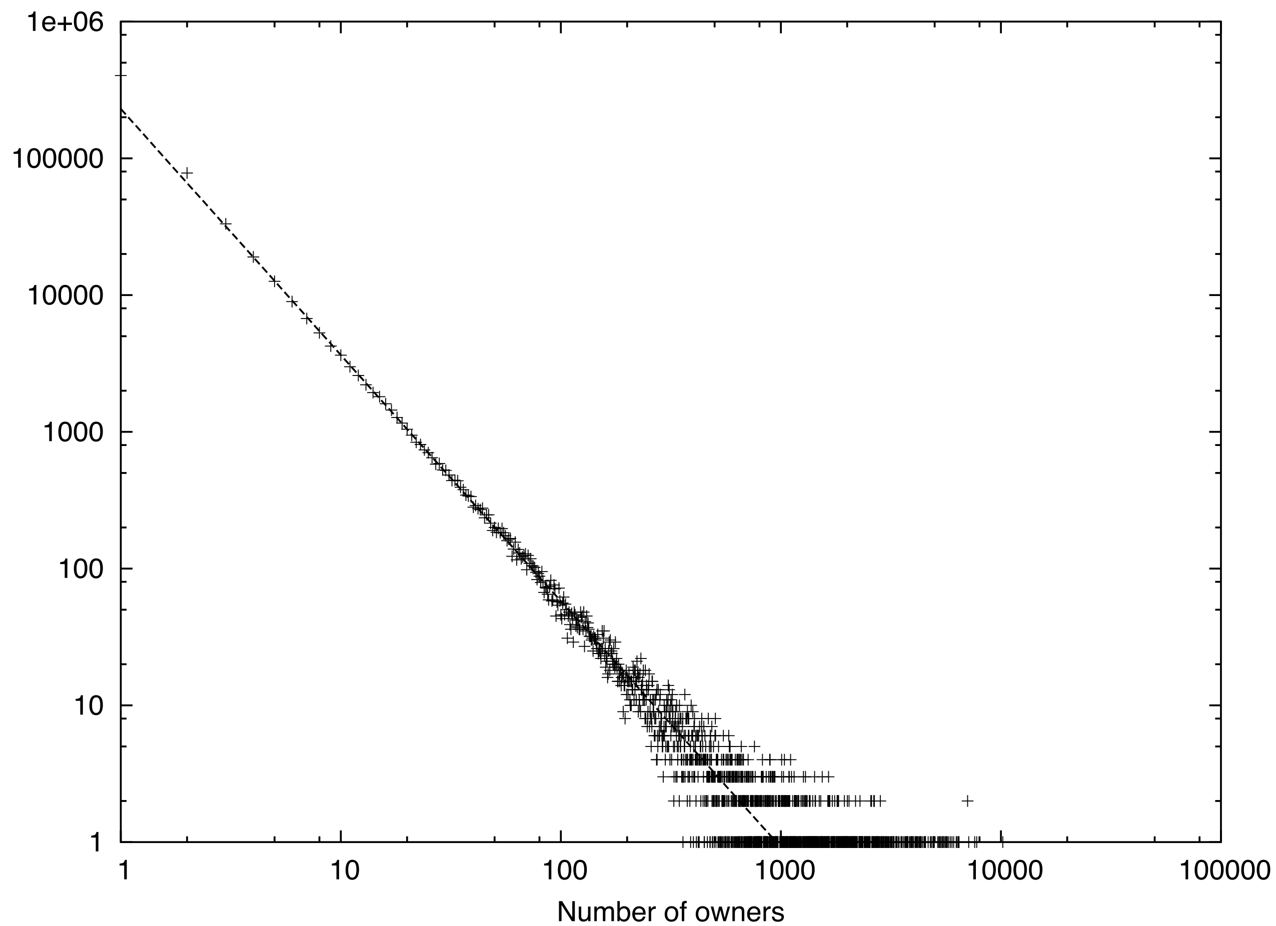
Analysis of bipartite graph

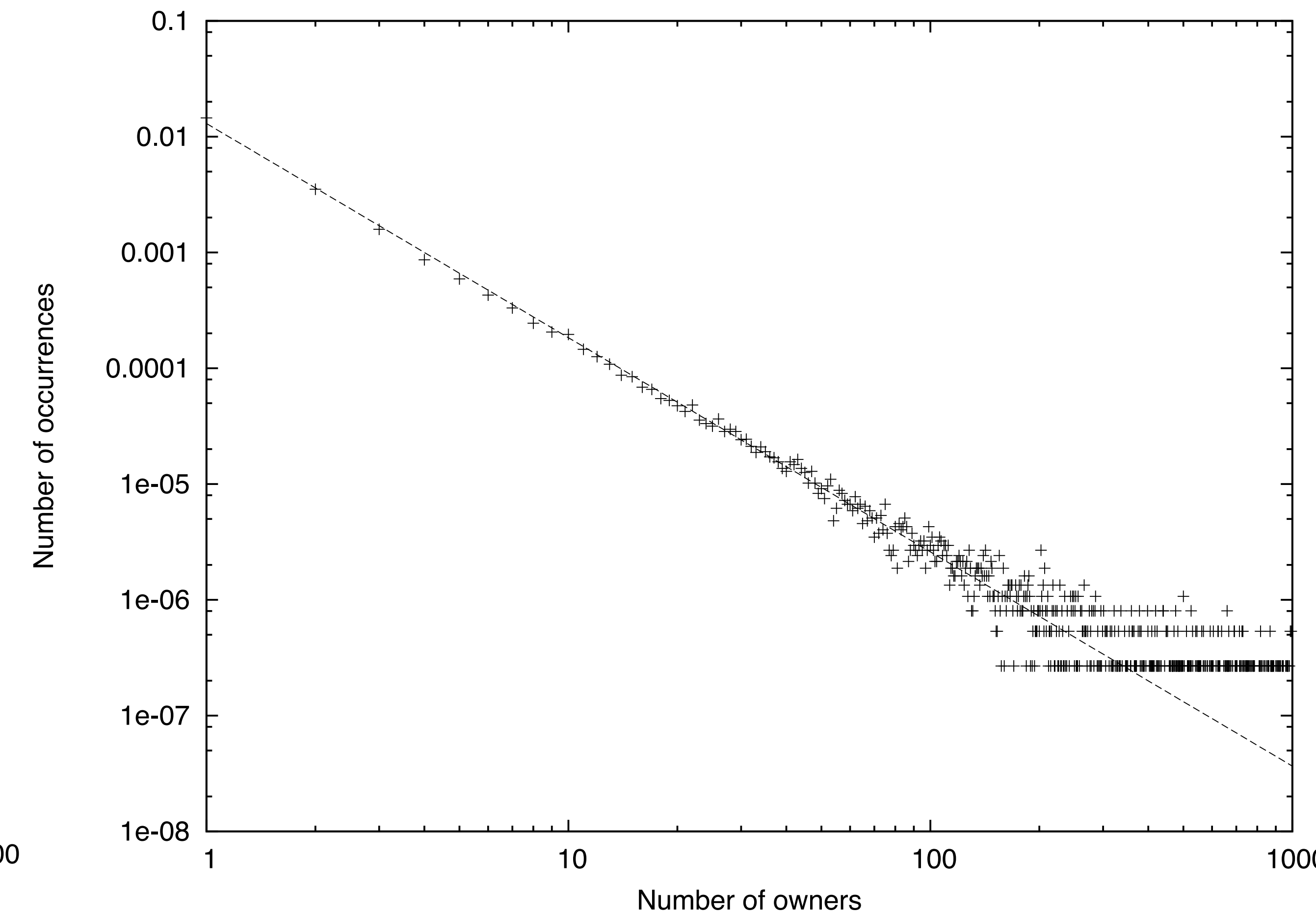
From audioscrobbler, a data base with:
- 35916 users
- the music library of each user + the number of times they listen to each group

There is a total of 617900 *different* music groups

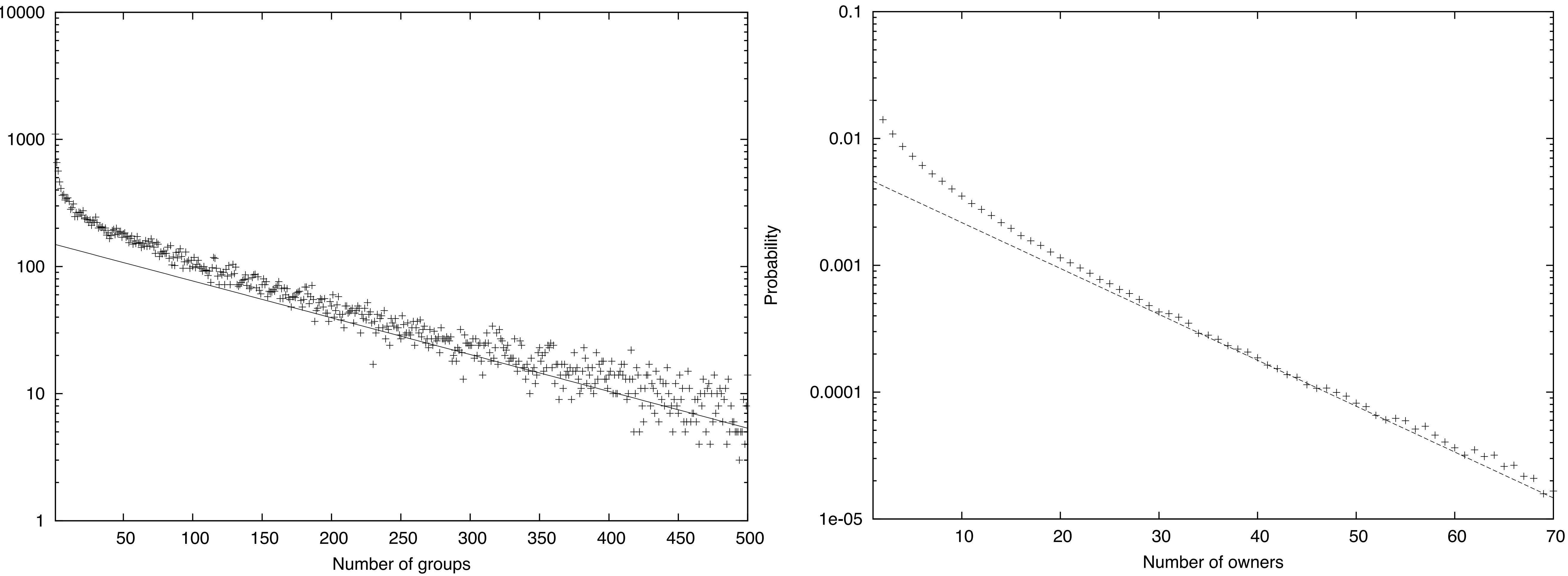
Mozart: 1468 users
Wolfgang Amadeus Mozart: 539 users
Amadeus Mozart: 17 users
Mozart Wolfgang Amadeus: 7 users
Wolfang Amadeus Mozart: 8 users
...

In the bipartite graph, there are 5028580 links, and the total number of playcounts is 54386834

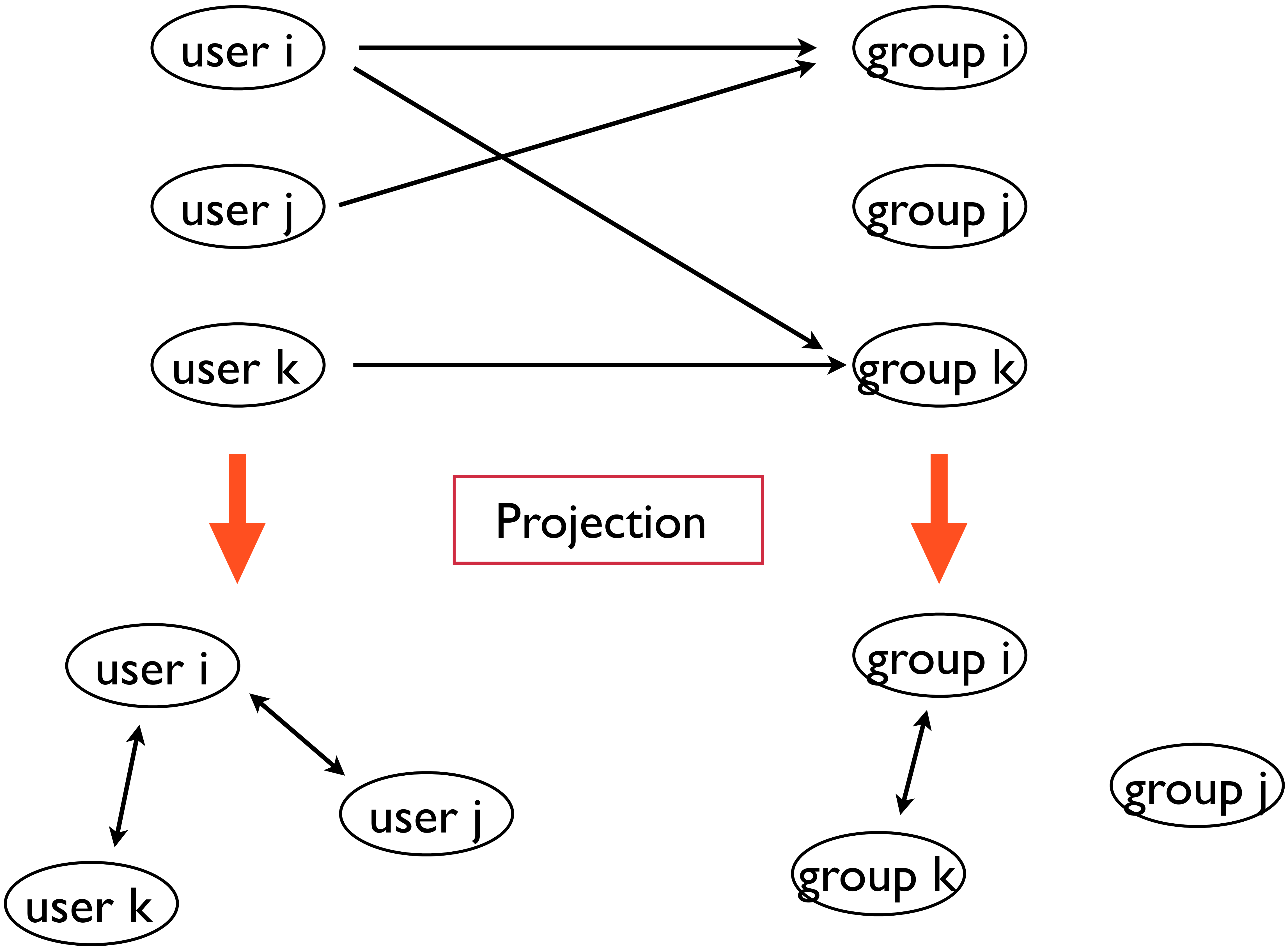




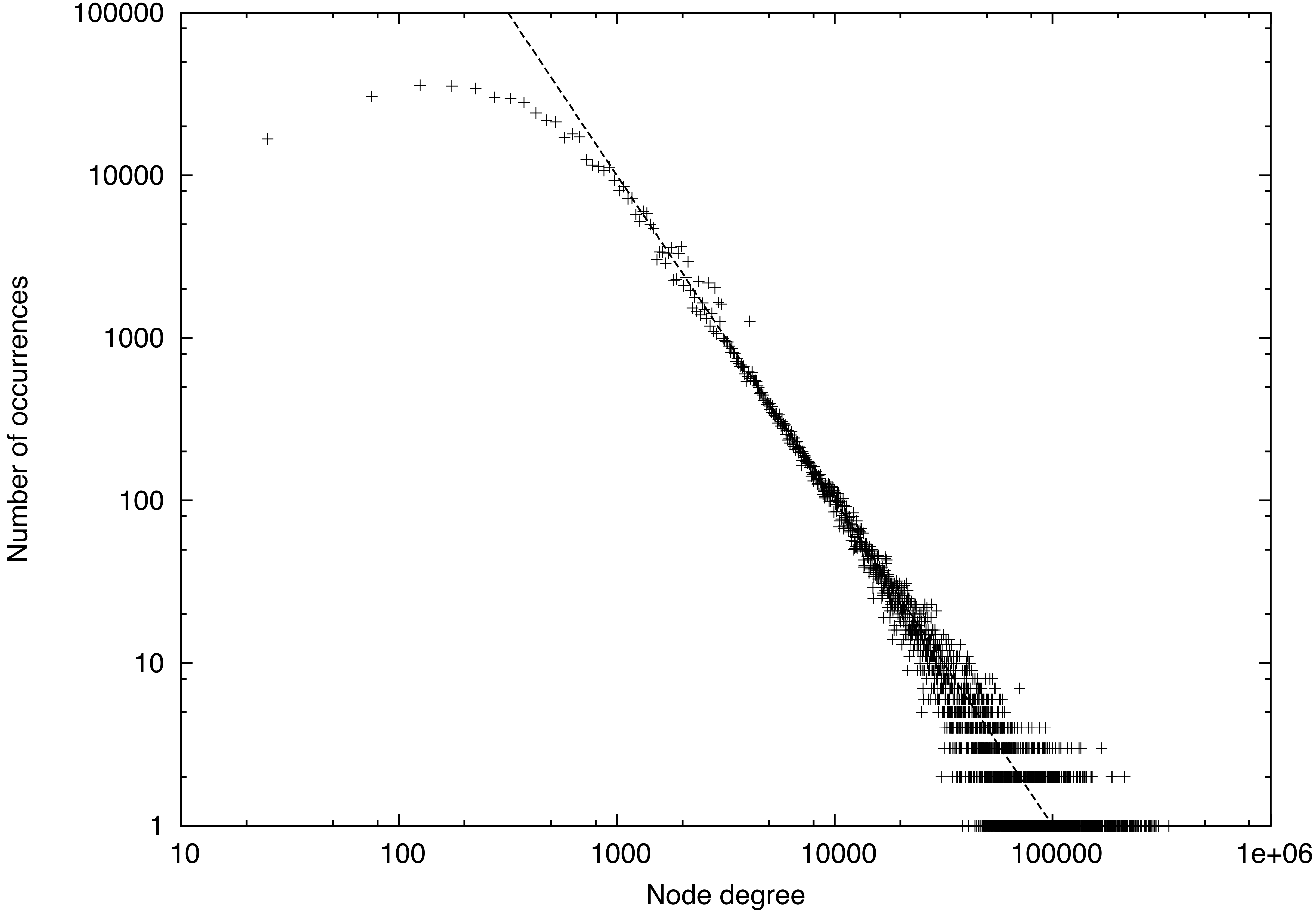
In the left figure, distribution of the number of listeners per group. This distribution is shown to behave like the power-law $\sim n^{-1.8}$.
In the right h.s. figure, same distribution from a simple growing bipartite network.



Distribution of the number of music groups per user, exponential tail.



Newman, Watts, Strogatz, Physical Review E, 64, 026118 (2001).



Degree distribution of the number of links per groups. For n>1000, the distribution behaves like n^{-2}.

Alternative way to project the network

In the following, we focus on the k largest groups, and try to get a structure for these groups.

For each group, we define a 35916 vector, with 1 if the the user i owns it, and 0 if not.

→ (1,0,0,...,0,1,...,1)

For each pair of group, we calculate the cosine between their 2 vectors:

$$\boxed{c_{ij} = \frac{\mathbf{v}_i . \mathbf{v}_j}{|v_i||v_j|}}$$

→ Symmetric measure of correlations, in [0,1]

→ Symmetric $k \times k$ matrix

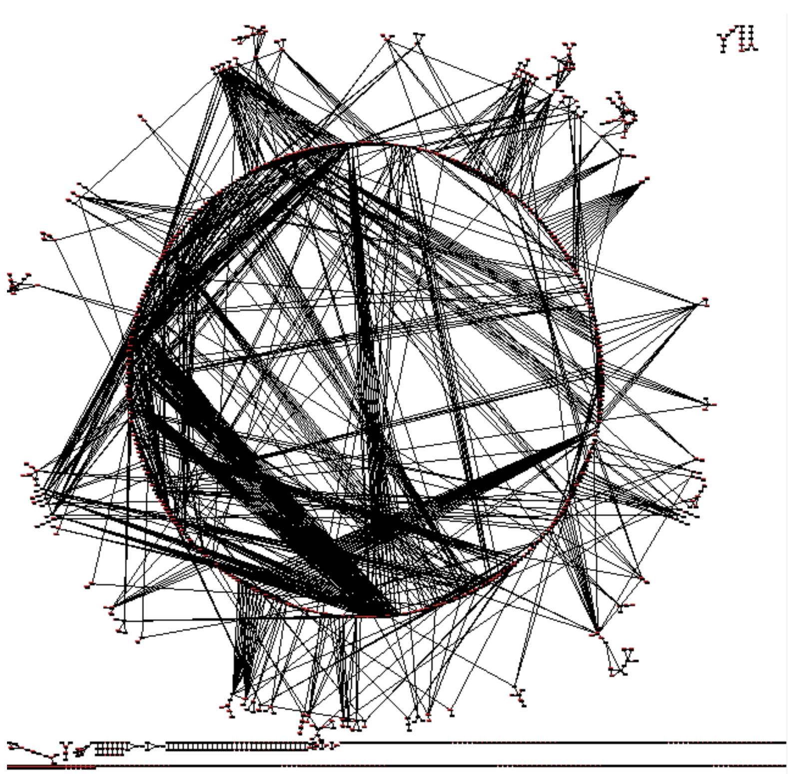
One applies the same scheme for the users, with a 617900 dimensional vector.

We construct a graph by filtering the matrix.

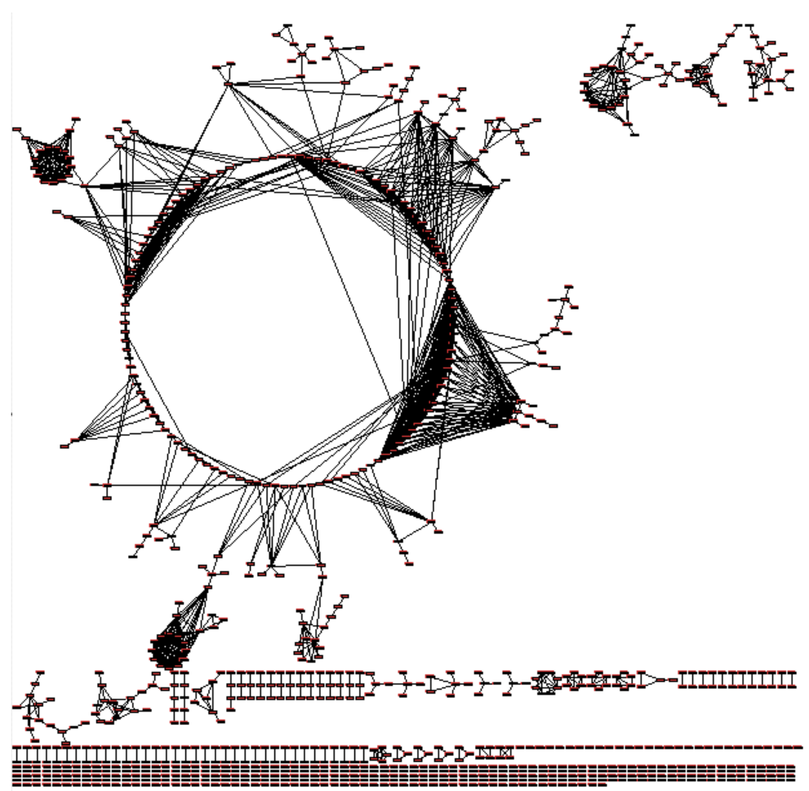
If $c_{ij} > h$	→	link between i and j
Else	→	no link between i and j

NB: when h=0, we recover the precedent projective method.

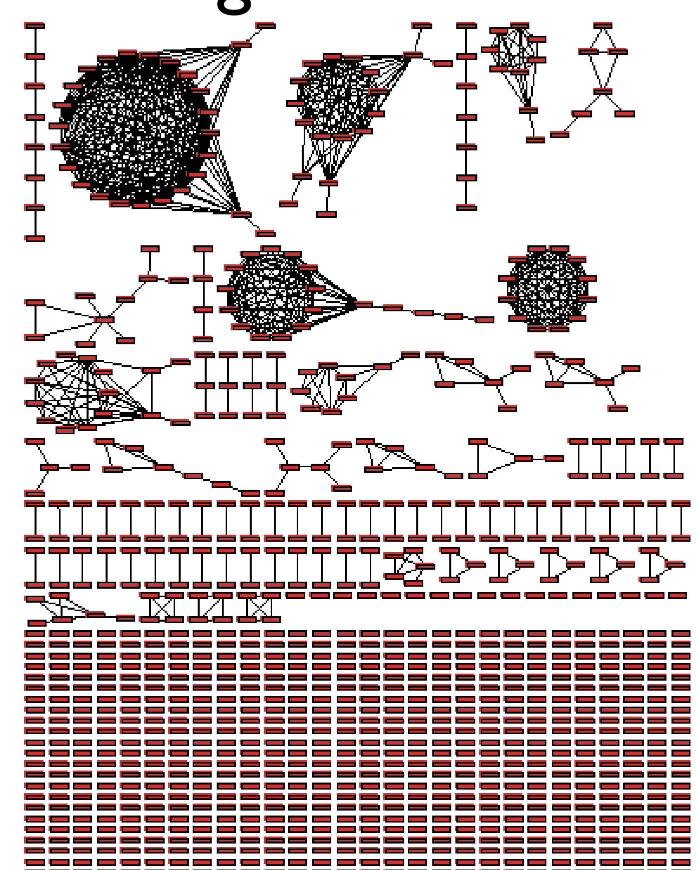
small h

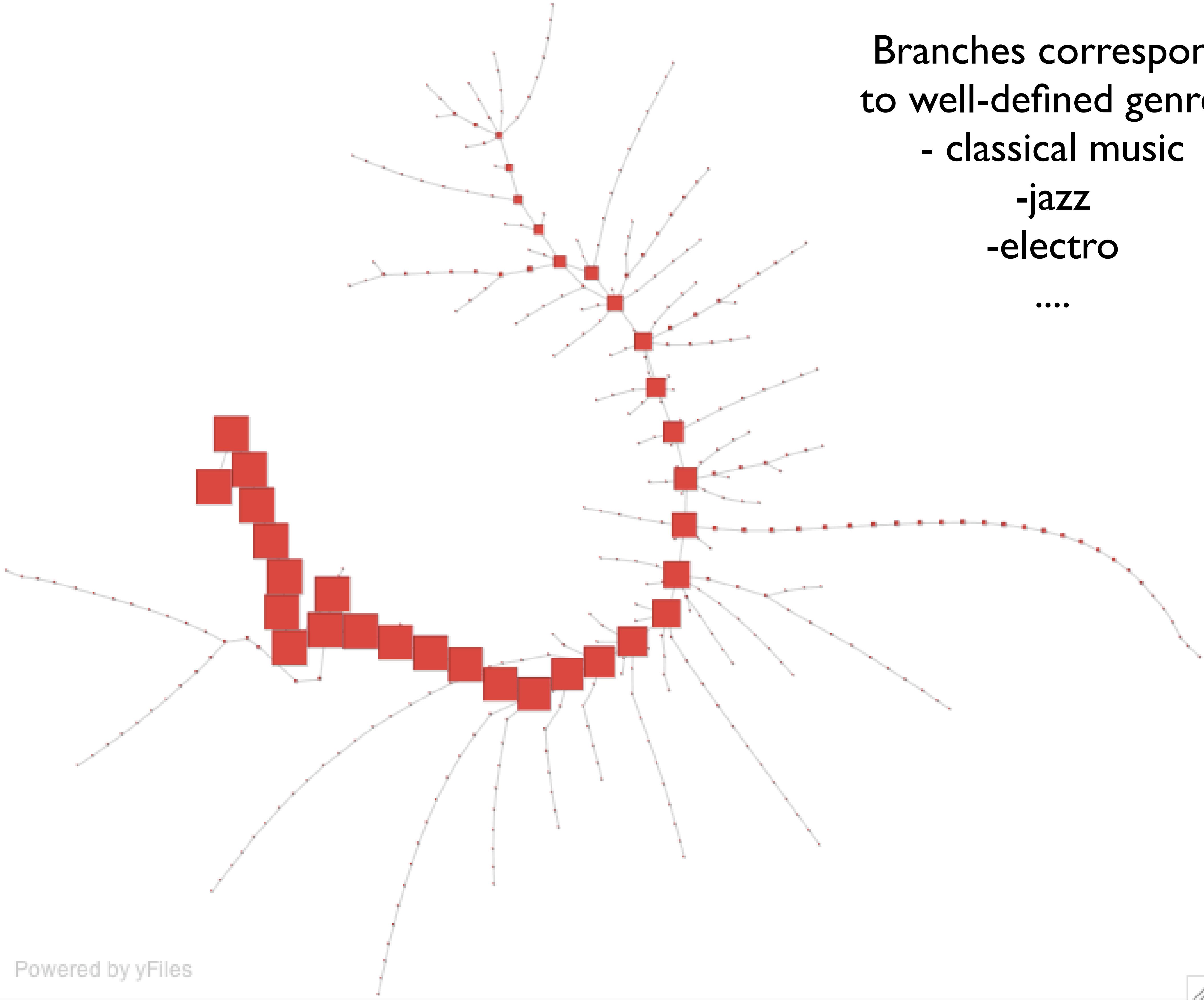


medium h



high h





Branches correspond
to well-defined genres:
- classical music
-jazz
-electro
....

What is next?

- improve the dynamical model
- characterization of the branching process
- application to scientist/articles bipartite networks