Diffusion on Complex Networks: A way of probing their large scale structures

Ingve Simonsen

Department of Physics

NTNU, Trondheim, NORWAY

Work done in collaboration with:

Kasper A. Eriksen, Lund

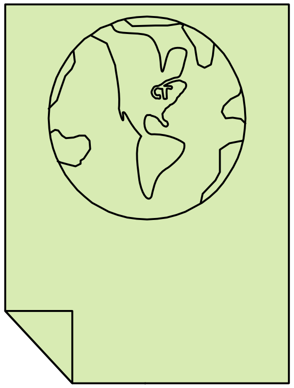
Sergei Maslov, Brookhaven

Kim Sneppen, Nordita/NBI

NTNU

NTNU

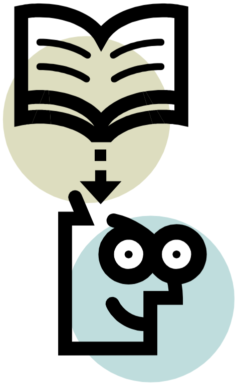
Motivation



- For many real-world networks it is important to:
  - have information on their modular structure
  - know how to increase their stability and robustness
  - identify critical links
- This is what we set out to do with diffusion!
- Physical Motivation:
  - Consider diffusion on complex networks (ensemble of Random Walkers)
  - Steady state will be reached first *locally*, then *globally*
- *Practical example*:
  - How easy is it by *random driving*, to reach South America if you start in North America ?

Outline of the talk

- Background Theory
  - Diffusion Equation for Complex Networks
  - Participation Ratio
- Real-World Examples
  - Zachary's friendship network
  - Internet network (Autonomous System)
- Application
- Conclusions



Diffusion on Complex Networks

- Conservation of walker density :

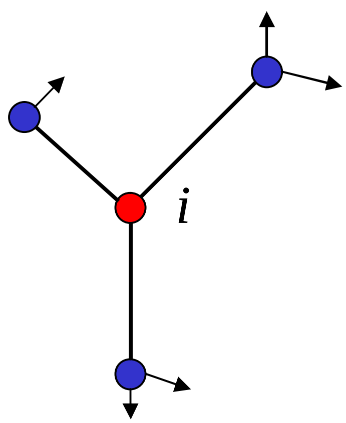
Adjacency matrix

Walker density

$$\rho_i(t+1) - \rho_i(t) = \sum_j A_{ij}\frac{\rho_j(t)}{k_j} - \sum_j A_{ji}\frac{\rho_i(t)}{k_i}$$

- In matrix form (T : transfer matrix, k_i: connectivity of node i)

$$\boldsymbol{\rho}(t+1) = \mathbf{T}\,\boldsymbol{\rho}(t) \qquad T_{ij} = \frac{A_{ij}}{k_j}$$



- A steady state exists, and is given by the connectivities :

$$\boldsymbol{\rho}(t) = \mathbf{T}^t\boldsymbol{\rho}(0) \qquad \Rightarrow \quad \rho_i(\infty) \propto k_i$$

Diffusion on Complex Networks, *cont.*

- The Transfer matrix is *not symmetric*, but it is similar to the symmetric matrix

$$\mathbf{S} = \mathbf{K}\mathbf{T}\mathbf{K}^{-1} \qquad K_{ij} = \frac{\delta_{ij}}{\sqrt{k_i}}$$

- Since S is symmetric, the transfer matrix has
  - real eigenvalues , $\lambda^{(\alpha)}$, (sorted in decreasing order)
  - due to conservation of walkers $|\lambda^{(\alpha)}| \leq 1$
  - and $\lambda^{(1)} = 1$ corresponds to the *stationary state*

- *Outgoing* currents flowing from node i along each of its links
(Technically, $c^{(\alpha)}$ is the eigenvector for T^T):

$$c_i^{(\alpha)} \propto \frac{\rho_i^{(\alpha)}}{k_i}$$

Stationary state : $c_i^{(1)} = 1/\sqrt{N}$

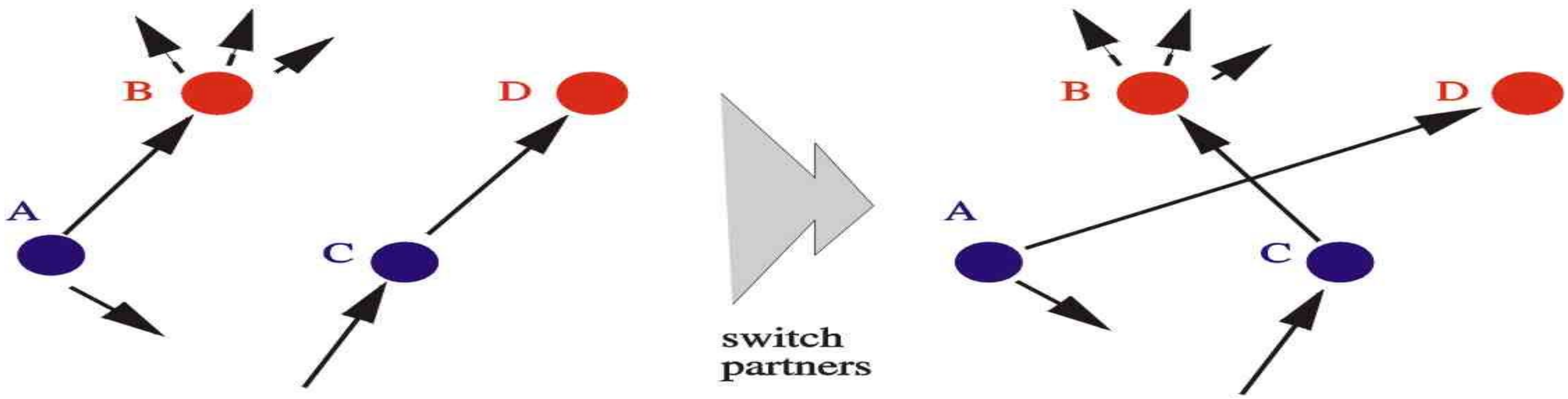
Participation Ratio

- For a normalized (link) current, $\sum_i (c_i^{(\alpha)})^2 = 1$, the participation ratio (PR) is defined as :

$$\chi_\alpha = \left[\sum_{i=1}^{N} \left(c_i^{(\alpha)} \right)^4 \right]^{-1}$$

- For the steady state : $\chi_1 = N$
- The PR gives a measure of the number of nodes participating significantly in a given eigenmode
- If the mode is related to a module, PR will roughly measure its size

The Null Model : A Randomized Network



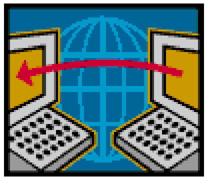
- Important : The randomization process should not alter the *degree distribution* of the network!
- The algorithm :
  - Choose two arbitrary nodes (A and C)
  - Choose one of their nearest-neighbors (or links) randomly (B and D)
  - Rewire : A→D and C→B if these links do not exist
- Reference : S. Maslov and K. Sneppen, Science **296**, 910 (2002).

Results for Real-World Networks

Two examples:

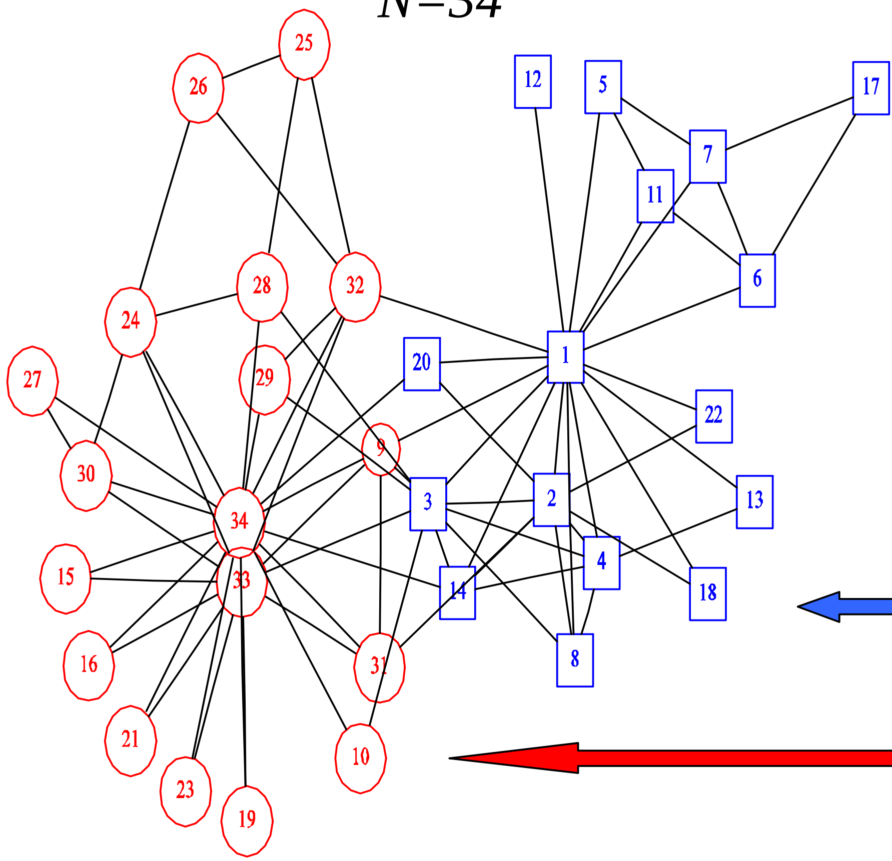
- Zachary's Karate club Network (friendship network)
  - The university karate club breaks up due to an internal conflict
  - A small network (N=34; L=72)
  - *The modular structure is known!*
  - Reference : W.W. Zachary, J. Antropol. Res. **33**, 452473 (1977).

- Course-grained Internet Network (Autonomous Systems)
  - Medium sized network (N=6,474; L=12,572)
  - *The modular structure is not well-known in advance*
  - Reference : National Laboratory of Applied Network Research http://moat.nlanr.net/AS/



Zachary's friendship Network

N=34



[From PNAS 99, 7821 (2002)]

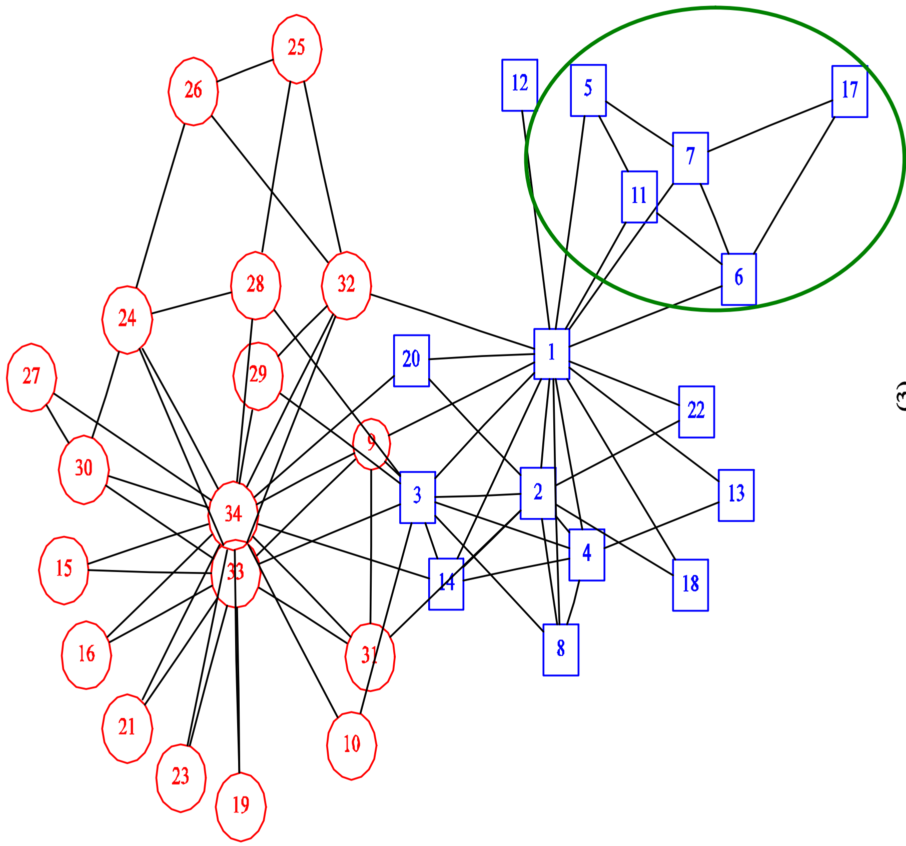
These are the assignment made by Zachary!

We use unweigthed links!

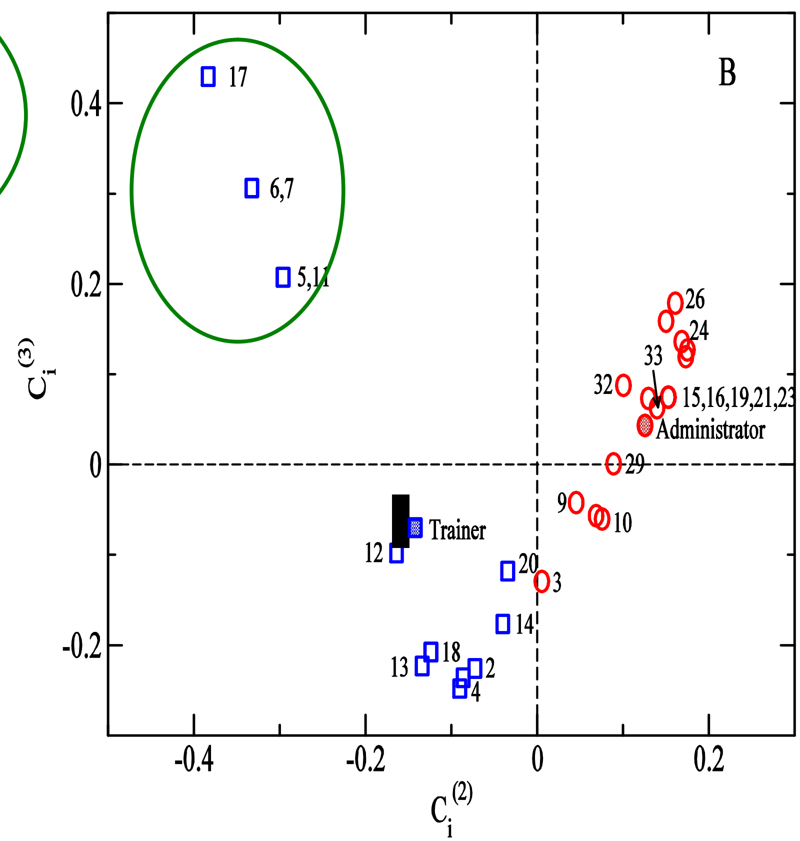
Trainer w/supporters

Administrator w/followers:

Zachary's friendship Network, *cont.*



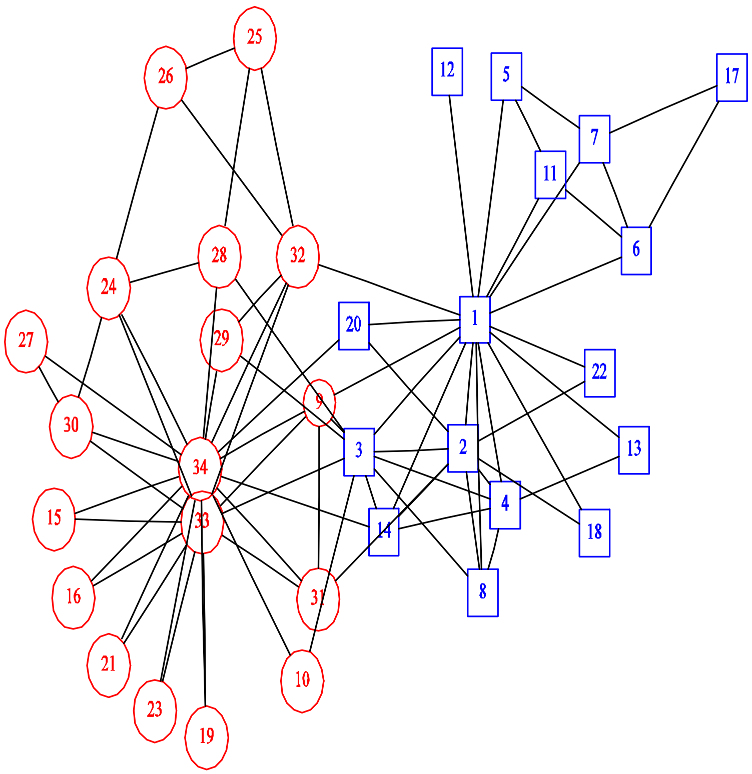
Current-current plot



[From PNAS 99, 7821 (2002)]

Node 3 is not correctly identified!

Summary: Zachary's friendship Network



Measured Cluster sizes :

$$\chi_2 = 14.0$$
$$\chi_3 = 13.3$$
$$\chi_2 \chi_3 = 5.4$$

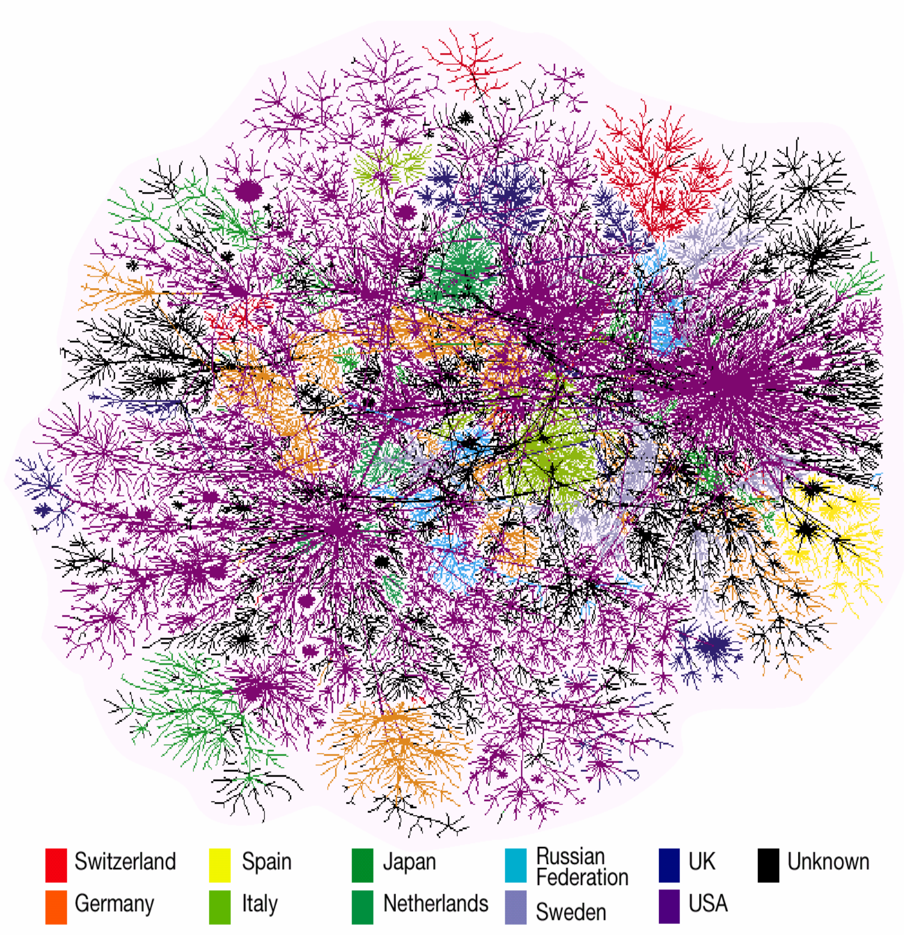
Theoretical sizes :

$$N_1 = 18$$
$$N_2 = 16$$
$$N_3 = 5$$

The known modular structure is well reproduced by the diffusion model !

Autonomous System Network

We now consider a larger network!



$$N = 6{,}474$$

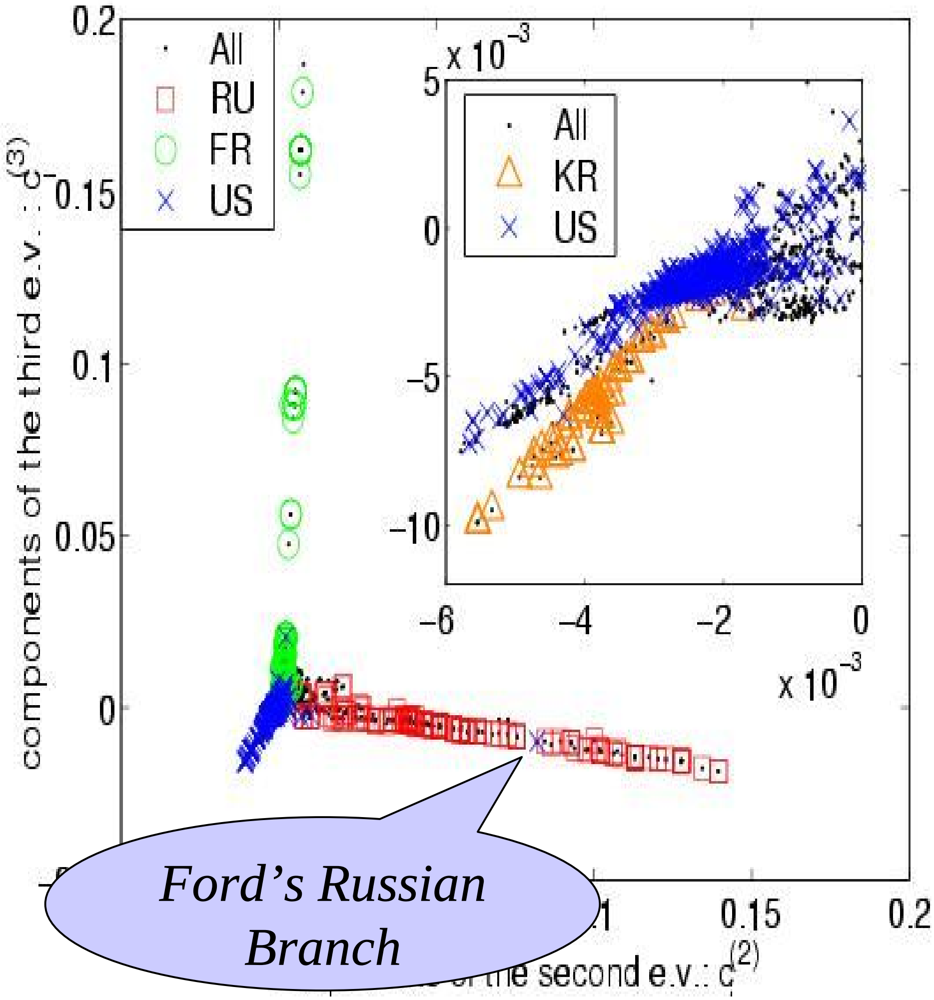
$$L = 12{,}572$$

Definition: An Autonomous System (AS) is a connected segment of a network consisting of a collection of subnetworks interconnected by a set of routers.

Nodes : AS numbers

Links : Information sharing

Autonomous System Network, *cont.*



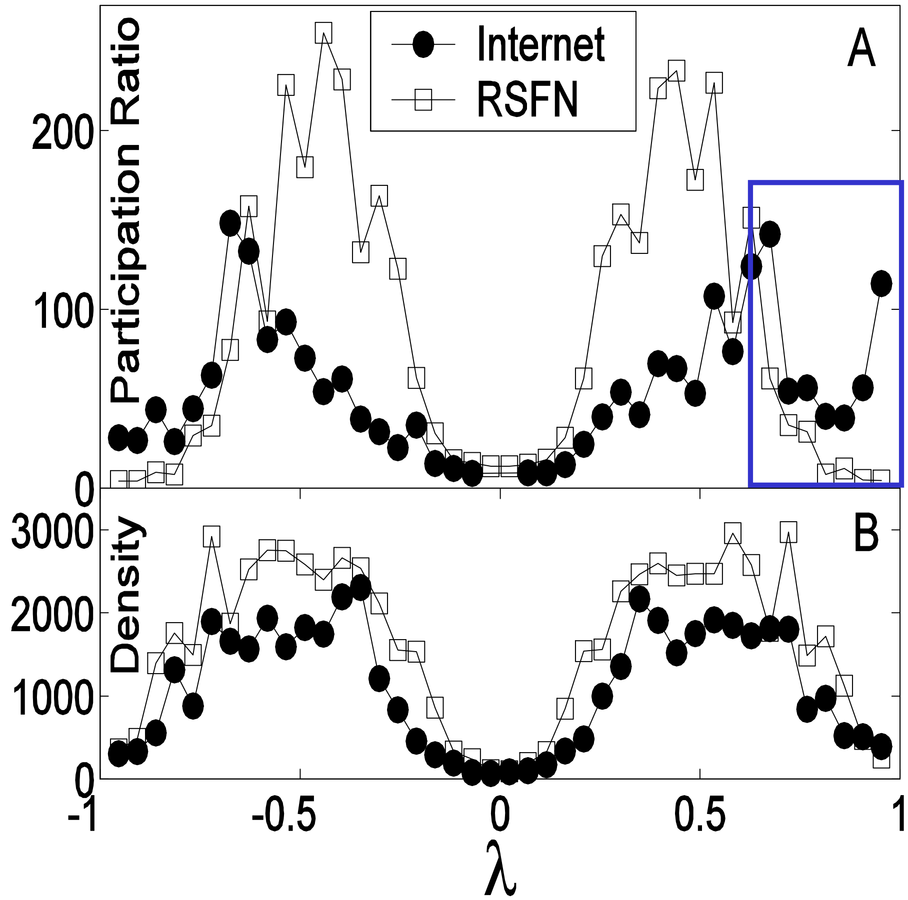
- The Internet is *indeed* modular
- The modules corresponds roughly to the national structure

The *extreme edges* of the Internet are represented by *Russian and US military* sites !

These country modules *could not* have been detected using spectral analysis of A

Ref : PRE **64**, 026704 (2001)

Autonomous System Network, *cont.*



$$\textbf{\textit{Binning}}\,\Delta\lambda = 0.05$$

- A Module is defined by a PR much larger then its random counterpart (RSFN)
- Number of Modules M=100
- Modularity

$$\sum_{\alpha=1}^{M} \chi_\alpha = 5400$$

and avoiding *double counting* gives:

$$1800/N = 30\%$$

Applications

the way to search the WED..

The advantage of Google:

Popular web pages are listed first among the search results!

- This is achieved by associating a PageRatio with each page related to the number of links pointing to this page.
- In our diffusion language this means

$$\boldsymbol{PageRatio(i) = \rho_i(\infty)}$$

Ref : S. Brin and L. Page (1998)

Conclusions

- ✓ Generalized diffusion to (discreet) complex networks
- ✓ Diffusion goes beyond next-nearest neighbor interaction

- ✓ Shown that diffusion may assist in probing the network's (large scale) topology
- ✓ Participation ratio is an important quantity that contains information about :
  - ✓ the number of significant modules (relative to a randomized network)
  - ✓ the size of the modules

The End

Reference :

K. Eriksen, I. Simonsen, S. Maslov, and K. Sneppen
Modularity and extreme edges of the Internet
Phys. Rev. Lett. **90,** 148701 (2003)

NTNU

The currents