The Abdus Salam
International Centre for Theoretical Physics

United Nations
Educational, Scientific
and Cultural Organization

International Atomic
Energy Agency

SMR.1656 - 29

School and Workshop on Structure and Function of Complex Networks

16 - 28 May 2005

Covering and Traffic Handling in Complex Communication Networks

Yamir MORENO VEGA
Universidad de Zaragoza
Instituto de Biocomputation y Fisica de Sistemas Complejos
c/ Corona de Aragon 41
Edificio Cervantes
50009 Zaragoza
SPAIN

Covering and Traffic Handling in Complex Communication Networks

Yamir Moreno, Jesús Gómez-Gardeñes
Institute for Biocomputation and Physics of Complex Systems, BIFI, University of Zaragoza.

Also with:

Alexei Vázquez (Notre Dame, Indiana)
Pablo Echenique (BIFI, Zaragoza)

ICTP, May 2005.

OUTLINE

3. Complex Communication Networks

5. The Covering Problem: Heuristic

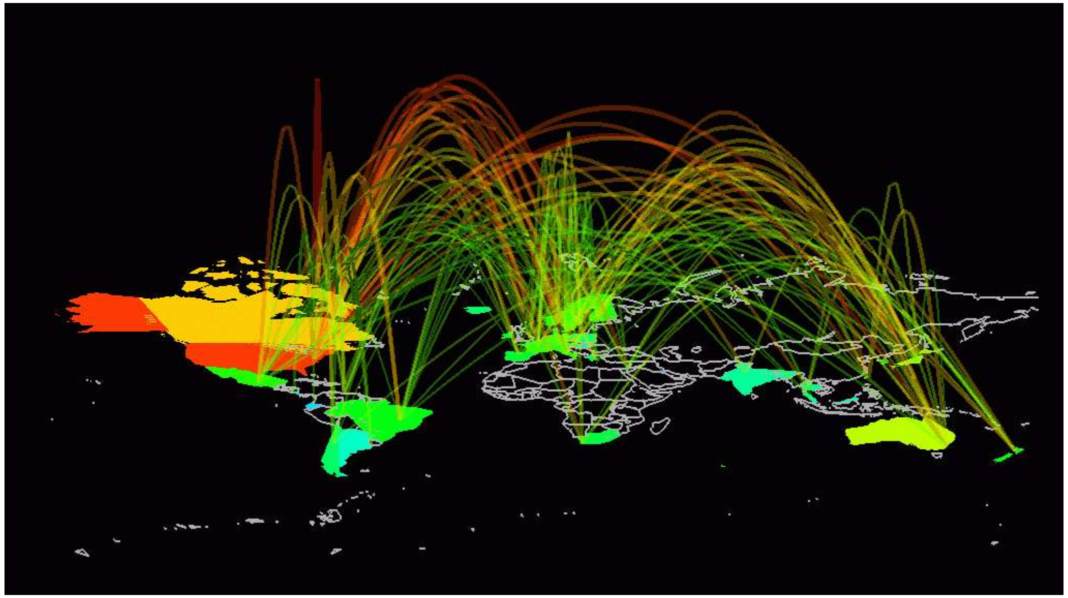
7. The Covering Problem: Immunization

9. Traffic-awareness algorithms.

11. Jamming Transitions.

13. Conclusions and Perspectives.

The Internet is a Complex Communication Network at both the Autonomous System and Router description levels.



From Stephen G. Eick

- It is a scale-free network with exponent around to 2.2
- It is large enough as to be considered statistically relevant. Besides, it is continuously growing.
- Our lives increasingly depend on its robust performance.

Two important problems:

- Efficient allocation of network resources.
- Alternative (and scalable) routing protocols for information transfer.

Efficient allocation of network resources.

Despite the many similarities between the AS and Router maps of the Internet, important differences arise in their statistical characterization. The most important is perhaps the degree correlations between adjacent nodes.

What are the effects of these differences? Do they determine the design of processes and/or protocols operating on the net?

The Distance-*d* Covering Problem

❑ **Motivation**:

1- Quest for the developing and deploying of a (digital) immune system.

2- Optimal placement of web mirror servers.

❑ **Aim**: Compute the minimum set of covered vertices (henceforth referred to as replicas, or mirrors) such that every vertex is covered or has at least one covered node at a distance at most *d*.

❑ **Optimization problem**: Trade-off between the number of replicas, $<x>$, and the number of other vertices, $<n>$, covered by each replica.

*P. Echenique, J. Gómez-Gardeñes, Y. Moreno and A. Vázquez , PRE **71**, 035102R (2005).*

P. Echenique et al, PRE (RC) 2005.

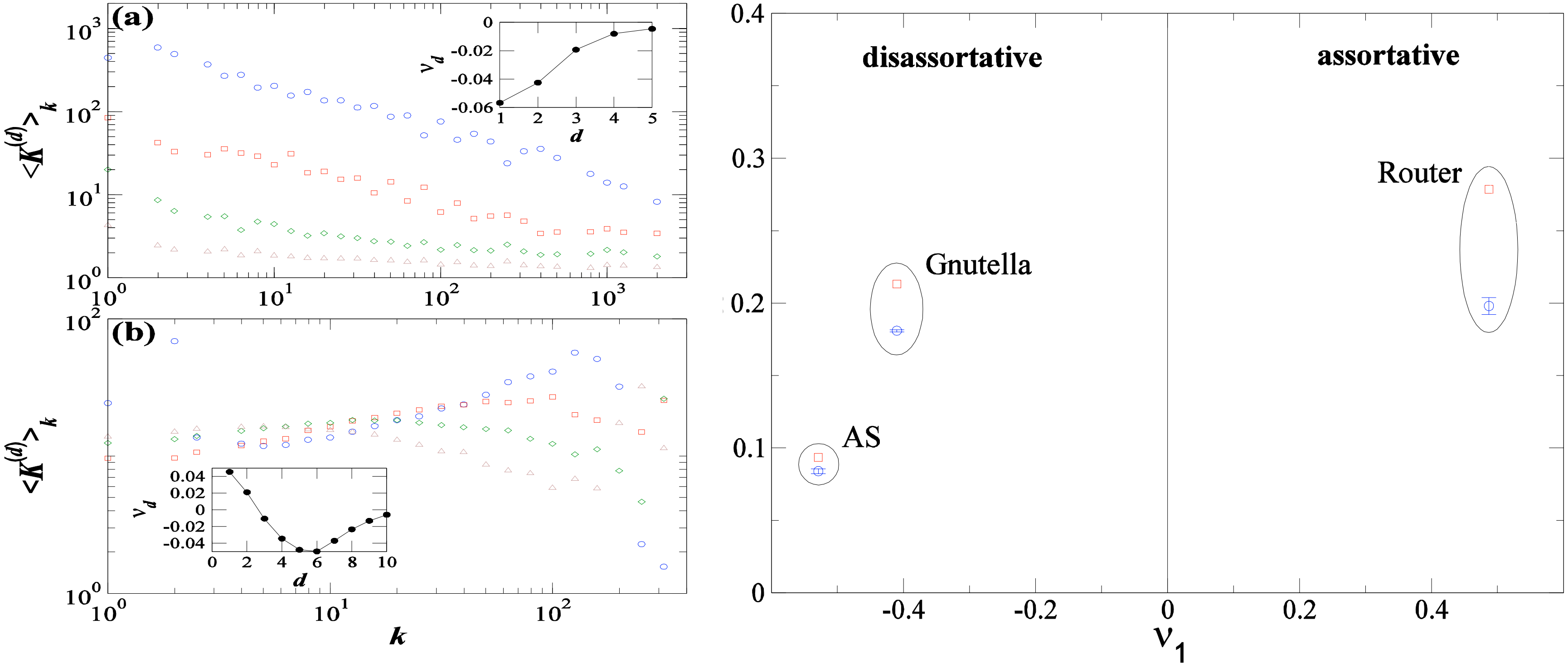
☐ Heuristic

- For each node, look for its neighbors at a distance at most *d* and cover the one with the highest connectivity,
- If there is a replica in the neighborhood, keep it and cover the node with maximal connectivity as well.

☐ The algorithm is not optimal, but it is *fast* and *accurate*, specially for disassortative networks.

- Our primary intent is not to develop an optimal algorithm. Instead, our main focus is in assessing the impact of correlations on the design of networked systems, and hence provide motivations, or lack thereof, for moving to more complex heuristics in the context of covering problems in real nets.

Numerical simulations of the covering algorithm on several communication networks.



At the router level representation of the Internet, degree-degree correlations depend on the distance. At relative large distances, correlations change.

<x>: set of covered nodes. Give the number of "special" nodes that you have in your system.

<n>: average number of nodes covered by a server (or average number of nodes to which a covered vertex gives service) . Directly related to the servers´ capacities.

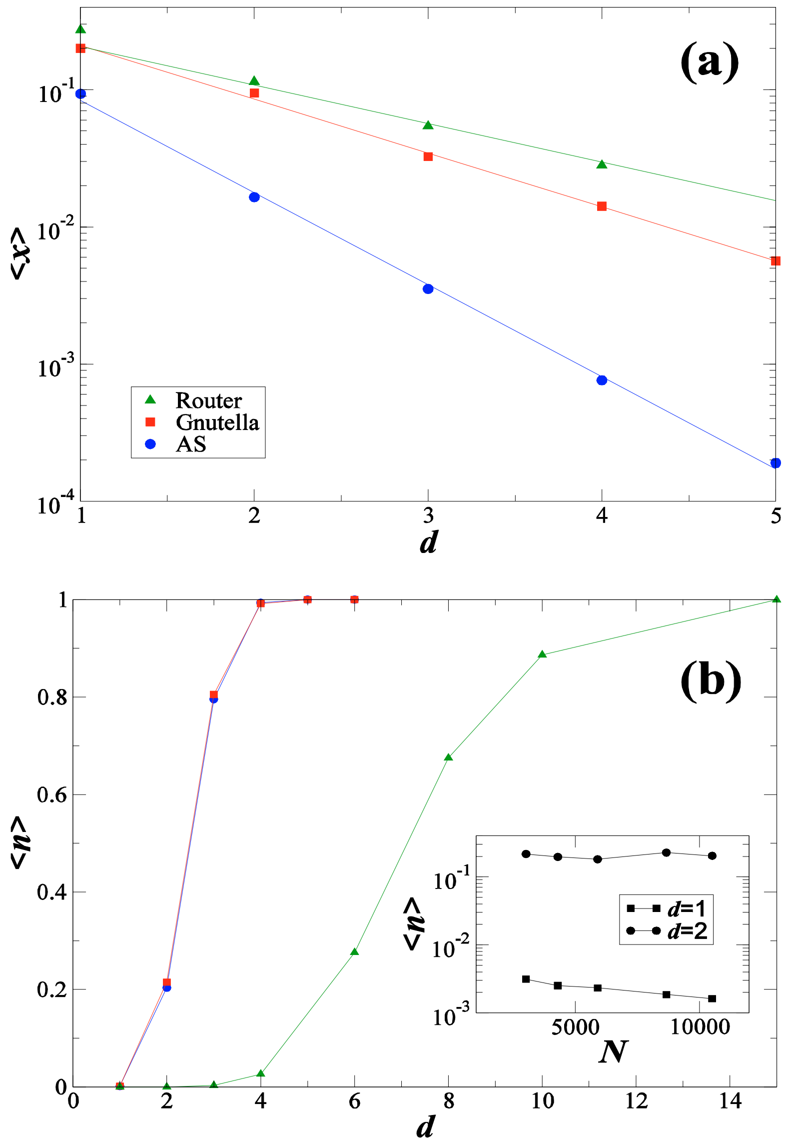
The trade-off between these two magnitudes points to the optimal design, i.e., a distributed or a centralized one!!!

*P. Echenique, J. Gómez-Gardeñes, Y. Moreno and A. Vázquez , PRE **71**, 035102R (2005)*
J. Gómez-Gardeñes, P. Echenique, Y. Moreno, and A. Vázquez, in preparation..

Results

➢ ***Two possible solutions***

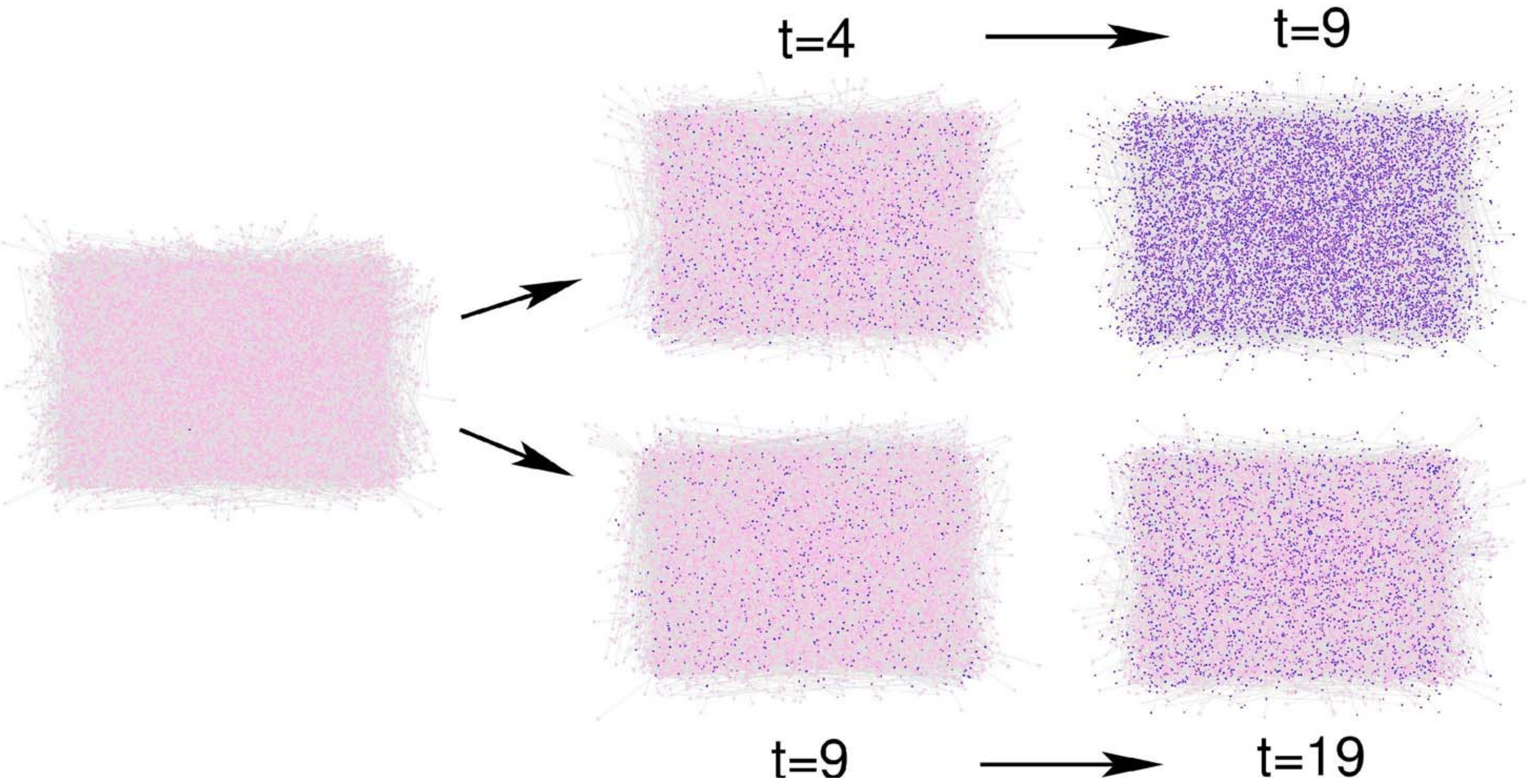
- Centralized Design: Small number of covered nodes, but with large capacities. Best suited for disassortative networks such as the Internet at the *Autonomous System* level and Gnutella.
- Distributed Design: More covered nodes, but with limited capacities. Best suited for assortative networks such as the Internet at the *Router* level and Social Networks.



Correlations (topology) strongly influence design principles!!!

IMMUNIZATION

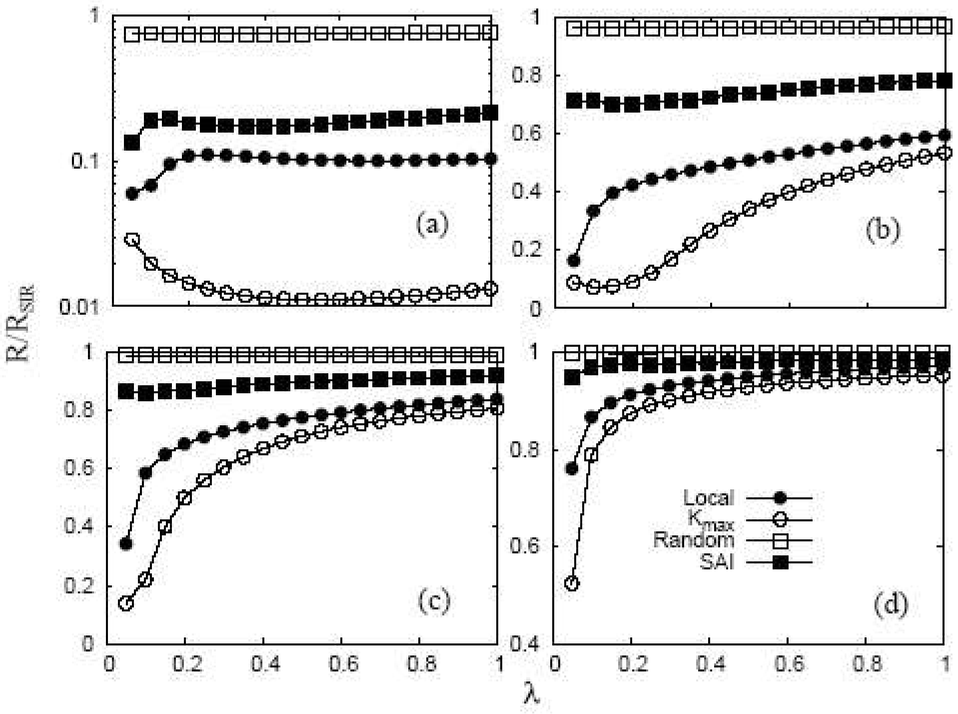
We consider a SIR model and assume that covered nodes correspond to immune sites (cannot be infected).



AS, Targetted immunization, 1% of immune nodes.

IMMUNIZATION

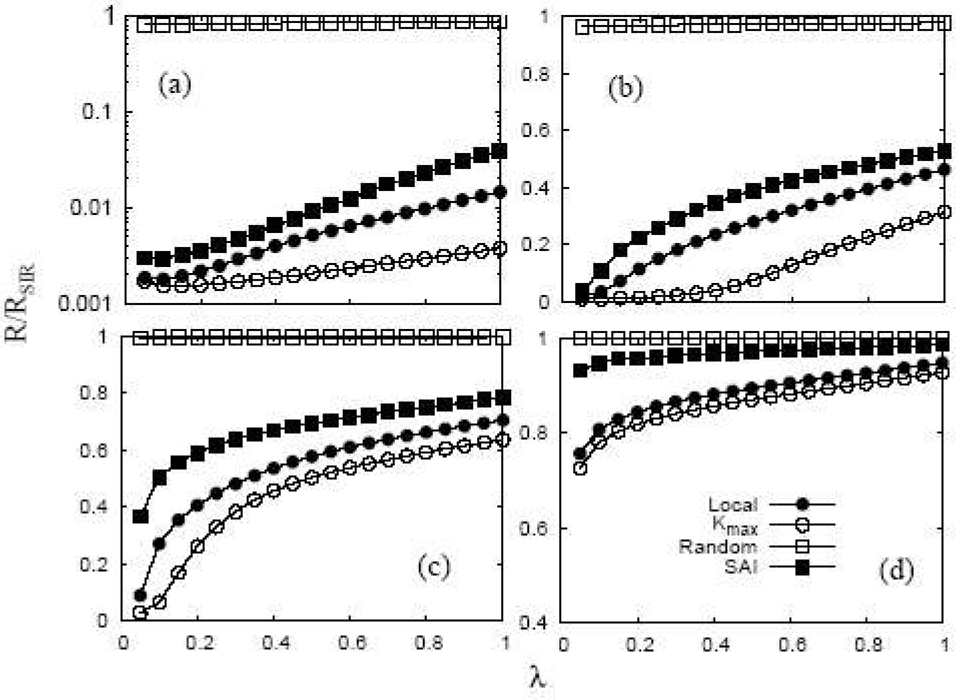
We consider a SIR model and assume that replicas correspond to immunized nodes (cannot be infected)



Legend: Local (covering), K_{max} (targetted), Random, SAI (Cohen et al., PRL 91, 247901 (2003)). SAI: Single Acquaintance Immunization.

IMMUNIZATION

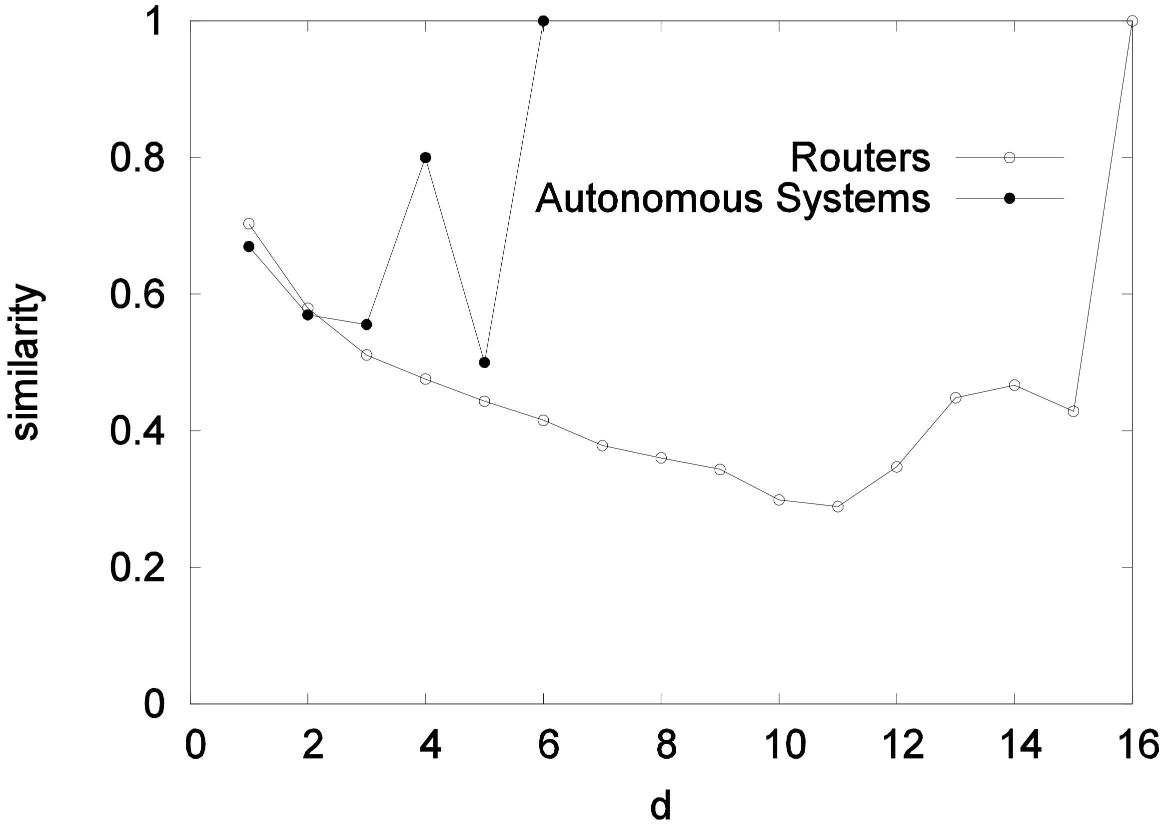
We consider a SIR model and assume that replicas correspond to immunized nodes (cannot be infected)



Legend: Local (covering), K$_{max}$ (targetted), Random, SAI (Cohen et al., PRL 91, 247901 (2003)).
SAI: Single Acquaintance Immunization.

IMMUNIZATION

We consider a SIR model and assume that replicas correspond to immunized nodes (cannot be infected)



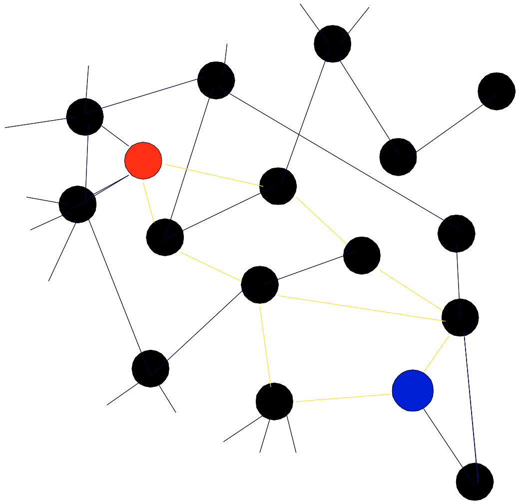
Traffic (information) flow and Jamming Transitions

Goal: improve actual routing strategies by incorporating traffic awareness to current models, i.e, knowledge of the load status in the network.

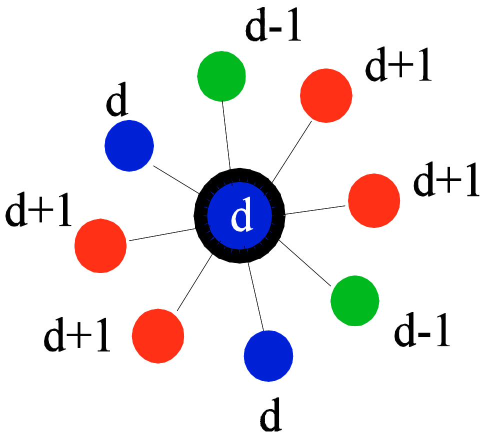
The Model

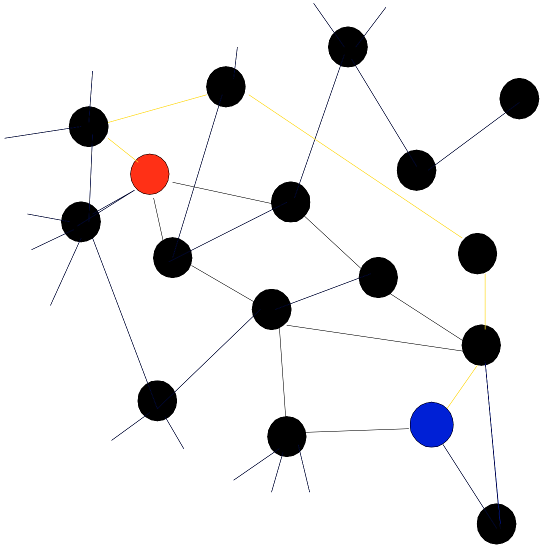
- A transport process is defined such that at every time step packets are sent through the network. All packets are associated with a source and a destination when they are created.
- Every node has a (not limited) queue in which packets are stored. They are sent on a FCFS basis.

*P. Echenique, J. Gómez-Gardeñes, Y. Moreno PRE **70**, 056105 (2004)*
P. Echenique, J. Gómez-Gardeñes, Y. Moreno Europhysics Lett., in press (2005).



Standard protocol: packets are delivered following the shortest paths to their destinations.





Traffic-awareness: packets are delivered taking into account both the topology and the actual congestion status in the network according to:

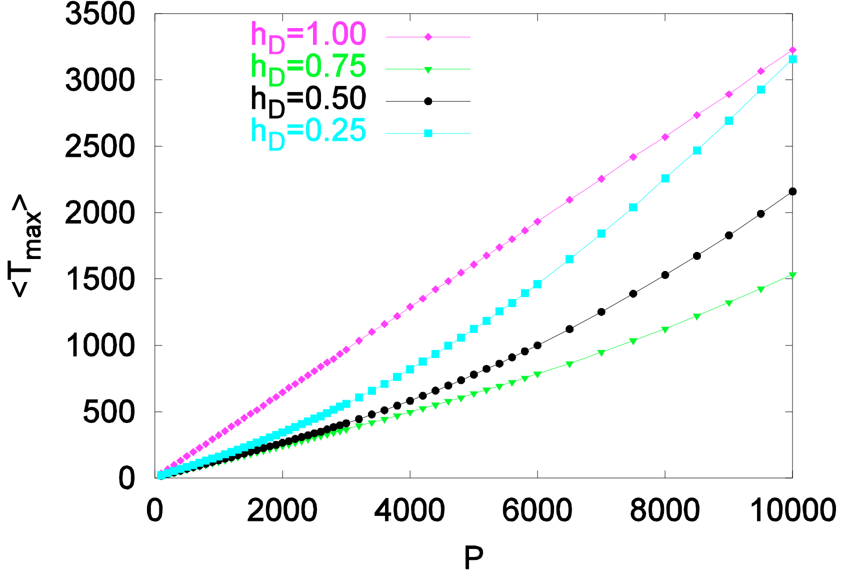
$$D_{ij} = h\, d_{ij} + (1-h)\, c_i$$

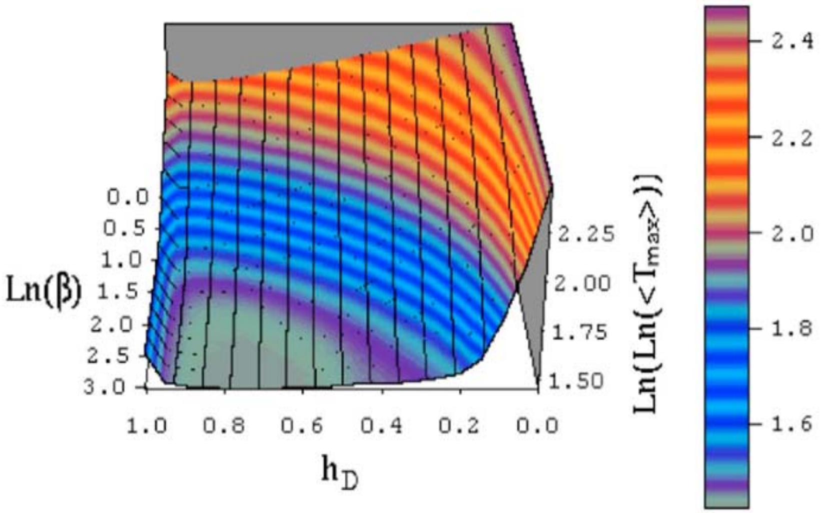
P. Echenique, J. Gómez-Gardeñes, Y. Moreno PRE ***70****, 056105 (2004)*

P. Echenique, J. Gómez-Gardeñes, Y. Moreno Europhysics Lett., in press (2005).

SINGLE INPUT

- At the beginning of the process, p packets are created. To each of them, one assigns a randomly chosen source and a destination. This is the control parameter of the model.
- We vary h and compute the maximum time it takes for a packet to reach its destination, averaged over many realizations.





$$\Pi_i = e^{-\beta H_i} / \sum e^{-\beta H_i}$$

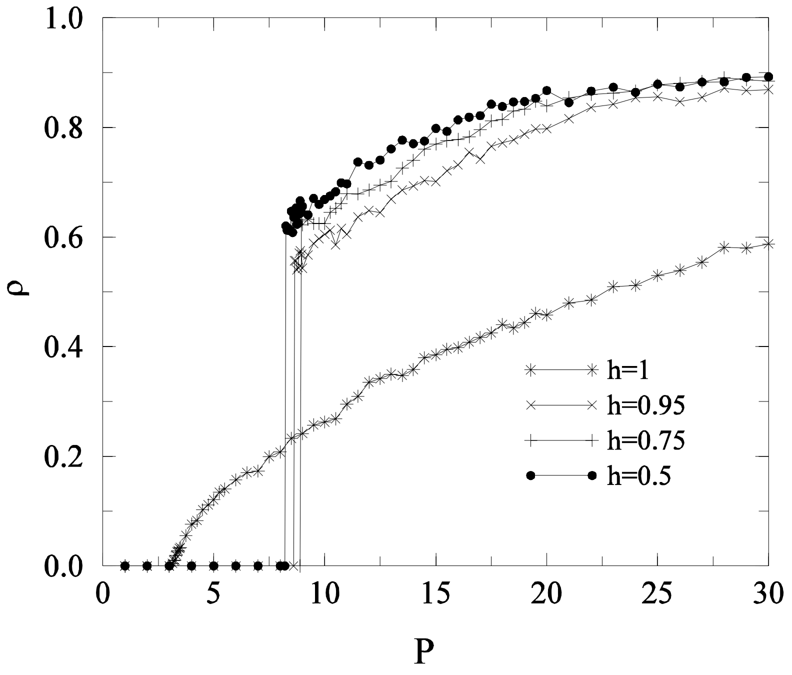
*P. Echenique, J. Gómez-Gardeñes, Y. Moreno PRE **70**, 056105 (2004)*

STEADY INPUT

- At each time step, *p* packets are created. To each of them, one assigns a randomly chosen source and a destination. This is, as before, the control parameter of the model.

- We vary *h* and compute a new order parameter given by:

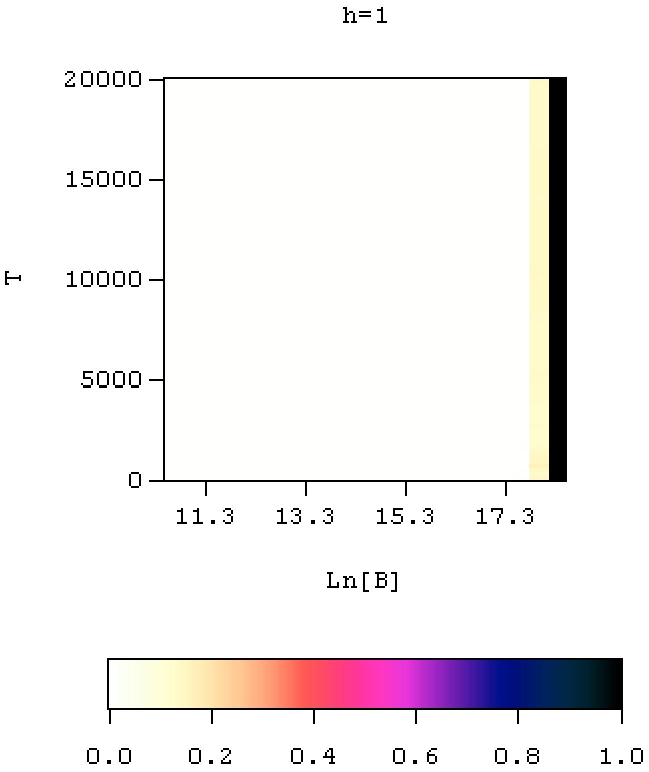
$$\rho = \lim_{t\to\infty} \frac{A(t+\tau)-A(t)}{p\tau}$$

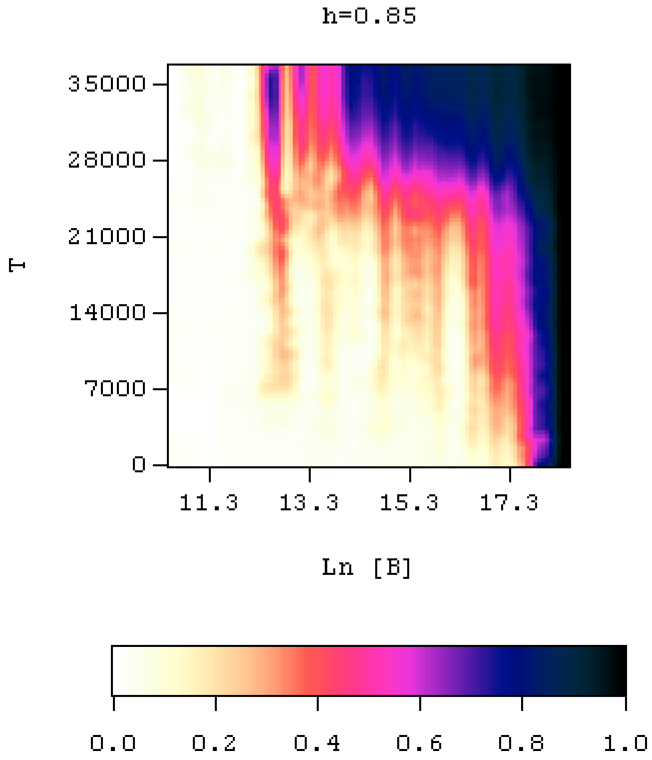


P. Echenique, J. Gómez-Gardeñes, Y. Moreno Europhysics Lett., in press (2005).

MICROSCOPIC DYNAMICS

Betweenness of node *i*: The number of shortest paths that passes through it. It is in general proportional to the connectivity of the node; the larger the node connectivity, the larger its betweenness.





P. Echenique, J. Gómez-Gardeñes, Y. Moreno Europhysics Lett., in press (2005).

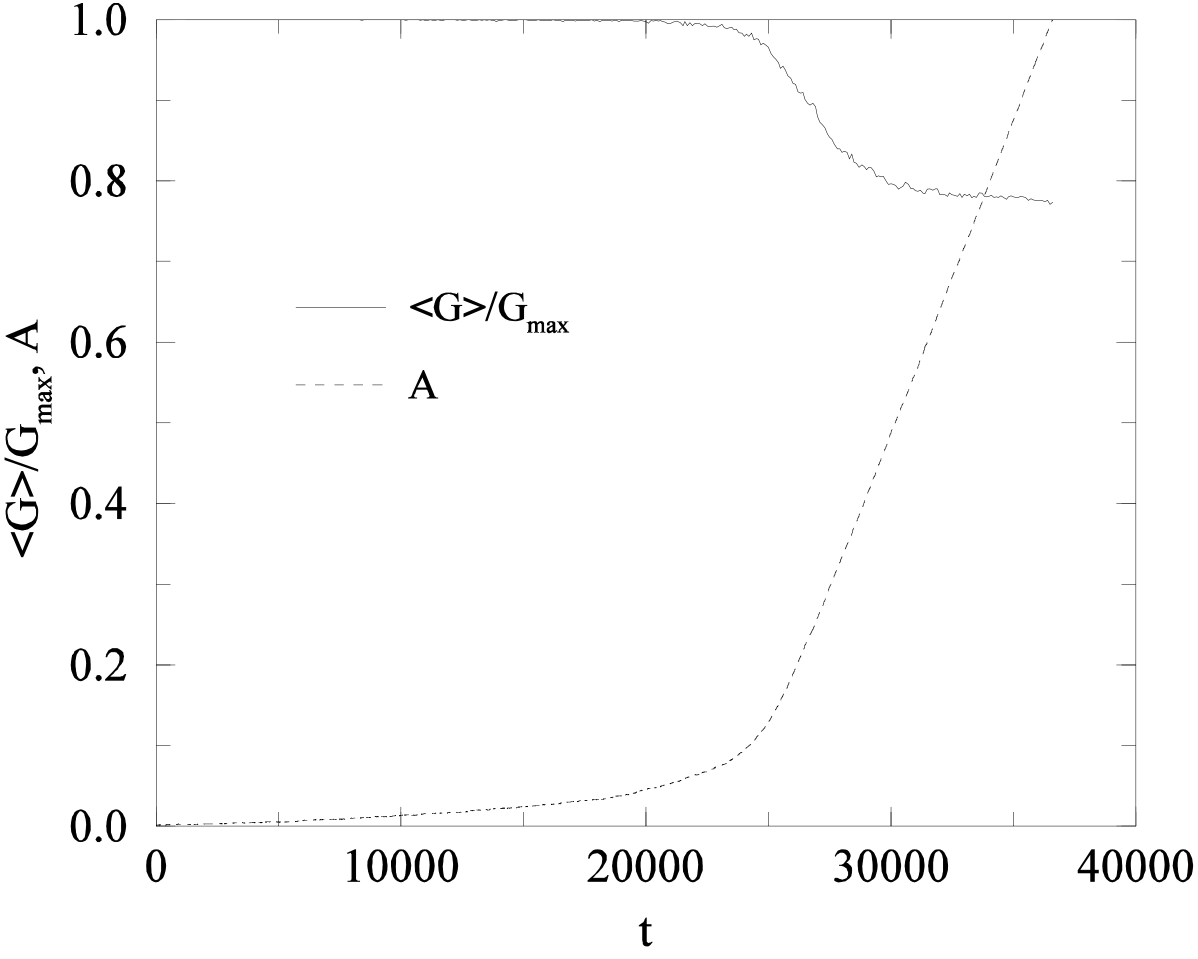
- For $h<1$, the system self-organizes the distribution of jammed nodes. As $h \to 1$ the time needed for self-organization eventually diverges.

Once again, the choice of what routing protocol is best suited in practice depends on a delicate trade-off between two factors:

- Smaller critical load, but second order phase transition, i.e., more “warnings”.

- Larger critical point, but first order like phase transition, i.e., congestion arises suddenly.

P. Echenique, J. Gómez-Gardeñes, Y. Moreno Europhysics Lett., in press (2005).



1.0
0.8
0.6
0.4
0.2
0.0
0
10000
20000
30000
40000
t
$\langle G\rangle/G_{max}$, A
$\langle G\rangle/G_{max}$
A