COSIN

IST–2001-33555

Coevolution and Self-Organization in Dynamical Networks

http://www.cosin.org

Final Report

Covering period 1.3.2002-31.8.2005

Report Version: 1

Report Preparation Date: 15 August 2005

Classification: Public

Contract Start Date: 1 March 2002 Duration: 42 Months

Project Co-ordinator: Guido Caldarelli (INFM) *Guido.Caldarelli@roma1.infn.it*

Partners:

CO1 (INFM) INFM *Italy*

CR2 (UDRLS) Università "La Sapienza" *Italy*,

CR3 (UB) Universitat de Barcelona *Spain*

CR5 (ENS) École Normale Supérieure, Paris *France*

CR7 (UNIKARL) Universität Karlsruhe *Germany*

CR8 (UPSUD) Université de Paris Sud *France*

CR9 (EPFL) École Polytechnique Federale de Lausanne *Switzerland*

Project funded by the European Community under the "Information Society Technologies" Programme (1998-2002)

Table of Contents

1. EXECUTIVE SUMMARY

- **Introduction**

During the years of the activity of the project (2002-2005) there has been a great interest in the field of scale-free networks in different communities. From Statistical Physics to Genomics and Biology different scientists started to consider the topology of graphs as a possible theoretical tool to describe their field of activity. In doing so they were forced to interact with the traditional experts of Graph Theory, that is to say Computer Scientists Mathematicians and for a little part Sociologists.
Even if the problems encountered in the respective fields were rather similar, still the lack of a common framework of terms, tools and ideas was the main issue in order to establish collaborations and expertise exchanges. We believe one of the main results of COSIN to have made a first step in the creation of this multidisciplinary discipline of growing networks. To give an example, the statistical approach adopted in the various cases of study has become the widespread methodology in the consortium. On the other hand the study of sociological quantities like communities distributions or betweenness is now a major topic of research in the community of Statistical Physicists. The activity of the consortium is witnessed by its scientific production. A complete list is available in the devoted section. Here, we only comment that both the quantity and the quality of the journals publishing COSIN research seem to be a rather objective measure of project success.

- **Objectives**

Given this situation at the time of the proposal the scientists involved in COSIN decided to put as a first objective of the collaboration the formation of a collective corpus of instruments in order to describe the real world. These instruments were visualization tools, simulations programs to produce numerical instances of different kinds of networks, real data sets to test theoretical hypothesis, new algorithms and models as well some extension of the notions of graph theory to deal with the experimental result. The ultimate goal with these tools was to provide a coherent and consistent description and understanding of Internet and the World Wide Web.
At the same time, when a discipline is starting the most important action that can be done is to have a reliable set of data in order to properly describe those systems.

- **Consortium**

The consortium consists of seven members and it is composed by three nodes active in the field of Statistical Physics (INFM, UB, UNIL-EPFL (formerly only UNIL), UPSUD), two nodes of Computer Scientists (UDRLS and UNIKARL (formerly UKON) and one node devoted to the study of social systems like that of Economics (ENS). The role of the various partners followed the plan presented in the proposal. In particular

1. **INFM** Istituto Nazionale Fisica per la Materia. Department of Physics, Rome Italy
(*G. Caldarelli, L. Pietronero, F. Coccetti, F. Colaiori, L. Buriol, Ramon Ferrer i Cancho V. Servedio*). This node had a supervision role, stimulating the activity of the various groups to follow the program of the project. It mantains the site of the project and produced a code for graph generation and visualization.
2. **UDRLS** Università di Roma “La Sapienza” Department of Computer Science, Rome Italy
(*S. Leonardi, D. Donato, L. Laura, S. Millozzi*). They worked mainly in the analyis of large portion of WWW. In collaboration with Node 1 produced some preliminar model of this system.
3. **UB** Universidad de Barcelona, Department of Physics, Barcelona Spain.
(*A. Díaz-Guilera, A. Arenas, M. Boguña, M. Catanzaro, M.-A. Muñoz, R. Pastor-Satorras).* They studied mainly the modelling of technological amd social networks. They have several collaboration with other nodes on this topic.

4. **UNIL-EPFL** École Polytechnique Federale de Lausanne (formerly Université de Lausanne), Lausanne Switzerland.
   (P. De Los Rios, D. Gfeller, T. Petermann). They mantained the data section of the project site and work on mathematical foundations of net exploration and formation.
5. **ENS** École Normale Supérieure, Paris France.
   (G. Weissbuch, S. Battiston, J.-P. Nadal). This node is mainly interested in application to Economics as the firm dynamics and board of directors analysis of the project results.
6. **UKON-UNIKARL** Universitaet Karlsruhe (formerly Universitaet Konstanz), Dept. of Computer Science, Karlsruhe Germany.
   (D. Wagner, U. Brandes, M. Gaertler). This node is mainly devoted to the production of visualization tools. Some of them have been inspired by the ideas of Renormalization Group in Statistical Phyisics.
7. **UPSUD** Université de Paris Sud, Paris France.
   (A. Vespignani, A. Barrat, M. Barthelemy, L. dall'Asta). This node is focussed in modelling visualizing and analysing the Internet.

- **Results**

The large scientific production of consortium established COSIN as one of the reference groups (together with Barabasi's group in Notre Dame, Mendes' group in Porto and Mark Newman group in Santa Fe) at least in the field of scale-free networks.
As a matter of facts COSIN site is one of the most used data repository by community. Many different people used the code produced by the consortium in order to create, analyse and visualize different networks. The consortium is just advertising some tools of visualization for large graphs and produced models for the Internet and the WWW.
While a complete crawl of the WWW was shown to be beyond the bandwidth accessible by COSIN, we still realised some smaller thematic crawls and analysed some geographical ones. In the case of the Internet also we have made several analysis of the system at AS level.

<u>Books</u>

- R. Pastor-Satorras and A. Vespignani *"Evolution and Structure of the Internet"* Cambridge University Press 2004.
- Statistical Mechanics of Complex Networks book of proceedings Springer 2003
- U. Brandes and T. Erlebach *"Network Analysis: Methodological Foundations"* Springer-Verlag 2005.
- G. Caldarelli and A. Vespignani *eds "Structure and Dynamics of Complex Networks"* to be published by World Scientific (2006).
- G. Caldarelli *"Scale-Free Networks"* to be published by Oxford University Press (2006).

<u>Conferences and Schools</u>

- School and Workshop Models and Algorithm for the WWW, Udine 2002
  http://www.dis.uniroma1.it/maw02/
- Sitges Conference in June 2002
  http://www.dis.uniroma1.it/maw02/
- Midterm conference Rome 2003 (about 100 participants)
  http://www.cosin.org/midterm.html
- *Sponsor* of Self-* International Workshop, Bertinoro 2004
  http://www.cs.unibo.it/self-star/
- *Sponsor* of European Conference Complex Systems, Turin 2004

http://www.isi.it/conference.html

- Final conference Salou 2004 (about 100 participants)
  http://complex.ffn.ub.es/cosin2005

Almost every site has a participant or more in charge of PhD courses on Networks, some other lectures have been given by G. Caldarelli and A. Vespignani at

Master in Complexity in Pavia (2004-2005)

School on Dynamics and Networks Trieste 2005

Papers

- 4 Nature
- 2 Proceedings of the National Acadamey of Science (USA) + 1 in press
- 13 Lectures Computer Science
- 8 Physical Review Letters
- 32 Physical Review E

2. PROJECT OBJECTIVES

The original objective of the project proposal were:

"The aim of this project is to develop a unified set of Complex Systems theoretical methodologies for the characterization of Complex Networks, helping addressing fundamental question about stability, efficiency and functionality of these networks.
We shall concentrate the research activity on the structures originated by the interplay of different agents in information society as the Internet network, the World Wide Web structure and the social and economic networks. In particular we intend to find a set of new tools for the analysis and the simulation of very large networks; devise efficient algorithms for measuring the relevant characteristics of such networks and for visualizing their evolution at different scales.
We shall also show that such tools can help in addressing the real-world problems faced in ITS technology as well as in the social studies."

Expanding a little bit this document with the extra information provided in the annex to the contract we can say that the main challenge we approach in this project is to give a quantitative description of complexity through the study of the networks with the goal of describing how complex behaviour could arise from microscopic interactions.
As for a non-rigorous characterization of complexity at least in Physics, we refer to the class of phenomena, like deposition, corrosion, cracking, growth of colonies and in general all the phenomena where the simple basic interactions between agents are such that to produce self-similar structures. These self-similar or fractal (i.e. they show the same shape at any level) structures are rather peculiar since they show correlation in shape at large distances both in space and in time. *A priori* this correlation is unpredictable from the microscopic dynamics. This mathematical property of self-similarity (i.e. to look the same at different scales) has been also called *criticality* in physics in relationship with the critical phenomena of phase transitions in thermodynamics.
The mathematical signature of complexity is the presence of power-law distributions for the quantities of interest.
Growing networks presents all the properties that we introduced in a non-rigorous way here. They effectively start from a small collections of vertices (or nodes) and arcs (or links) and through continuous growth, they develop some nontrivial complex features whose signatures are several power-law distributions as for example in the distributed frequency function of the number of links per node and of the in-degree distribution (the set of nodes ``uphill" in the spanning tree of the subgraph connected to a node in a oriented graph).
These are only few of the possible measure in order to quantitatively characterise networks. Others measures of interest that are computationally harder to model and to test are the clusterisation of the vertices, the distribution of cliques in order to evaluate size and distributions of communities or phenomena like the small-world (suitable shortcuts reduce the main distance between vertices to a characteristic small value) that arises as soon as the network reaches a stationary state and can be used to measure the efficiency of the net.

COSIN project decided to study complex networks focusing its attention on specific networks like Internet and WWW that are critical in the IST and on social networks that model relationships in companies.

Following the original project we therefore list the various particular objectives

1. we intend to develop a unified set of Complex Systems theoretical methodologies for the characterisation of all Complex Networks, helping addressing fundamental questions about

stability, efficiency and functionality of a network. Besides their theoretical interest there are, in particular two basic questions extremely relevant for the applications,:

How local interactions can be put in relation with a specified global behaviour?
What features of the local interaction makes a network stable against perturbations, thus preserving its functionality?

2. we intend to develop a set of tools for the analysis, the simulation and the visualization of very large networks. In particular we will devise efficient algorithms for measuring the relevant characteristics of such networks, and we will use these algorithms to collect data on Internet and WWW.
3. we want to study efficient visualization methods for such large networks.
4. we want to show how such tools can be applied to real-world problems faced in IST technology and economical and social networks. We will consider social networks that model the relationships in a firm and in firms in the same business area.
5. We will also consider the communication networks focusing our attention on Internet and WWW. We observe that the definition of suitable mathematical models that adequately represent Internet and WWW allows to simulate new applications and protocols whose efficiency is influenced by the structure of the network.

3. METHODOLOGIES

Being this project essentially focussed on basic research, the methodology in a strict sense has been very orthodox. That is to say from different datasets of a particular system we tried to select the features or better the ``physical quantities'' that describe it with enough accuracy. From this basic steps we passed to the modelling of the system and therefore whenever possible to the prediction of the future behaviour of other datasets. This elementary process known by everyone who is in charge of basic research, can be schematised as follows

A. Data Collection
B. Simple Modelling
C. General understanding (Theory)

While in general the third step is very difficult to achieve (a General Theory for fractal growth is still missing since the '70 when fractals have been introduced), still one can point out some ``active ingredients'' that can play a role in it. As stressed in many parts of the project documents we believe that this general theory corresponds to the science of complexity, that could clarify when and under which conditions simple entities interact to form something completely new. Put in this way the achievement of this third step is a formidable effort.

Our feeling is that the present state of art in the field is still somewhere between points 1 and 2. Not only we are still discovering new systems that can be described by means of graph theory but also the traditional ones display new features when other topological quantities are defined and checked. All this activity is in a loose sense related to the data collection or better to the creation of a repository of network data that can be used as a playground for scientists. On the other hand at the same time the first preliminary data have been interpreted and partly reproduced by a variety of toy models that can be clustered in relatively few different classes

- Random graphs and modifications (essentially the original model of Erdos-Renyi)
- Dynamical models with growth and preferential attachment
- Dynamical models with reciprocal rules for the establishment of a link
- Minimisation Models where the graph is chosen in such a way to minimise some cost function.

It is possible that some other ideas will be applied in the future to the modelling and it is very likely that some of them will play a role in the "General Theory" yet to come.

Therefore we can conclude that the methodology used until now world-wide can be defined as the research of reliable and good quality data on which test the hypothesis for new models. These models are then validated by new datasets in a feedback that should converge towards more and more refined models. Incidentally, data collection triggers further theoretical activity. As an example, if betweenness centrality becomes an important feature to monitor, then efficient algorithms must be produced in order to compute it on large graphs. Furthermore to pass from data analysis to modelling the step is to understand what is going on in the system and ultimately visualize it. A bad visualization can hidden the main features of a system as well as a good one can partly solve this problem. We tried to proceed along these lines in our activity alternating data analysis, modelling and testing the tools produced in the consortium.

STATE OF THE ART IN THE FIELD

Network research, in recent times, has been focusing on deeper analysis of networks structure, with new approaches in real data surveys and in mathematical models able to reproduce them. New models for complex networks have been introduced, in order to reproduce the features of real networks. A series of active ingredients from growth and preferential attachment, to fitnesses and minimisation of suitably chosen cost function have been proved to reproduce some of the properties of the systems observed.

On top of that a new interest has been paid on more refined representation of these systems, namely their dynamical evolution and the possible weight that can describe the nature of the edges. This new approach is the outcome of the realization that most complex networks are the result of a growth process. As a result, we currently view networks as dynamical systems that evolve through the subsequent addition and deletion of vertices and edges.
As already pointed out in the previous report and in the relative deliverable showing the state of the art in the field, the lines of research as perceived by the project members as well as the opinion of the major experts worldwide as collected during the most important conferences can be summarised as follows.

- **Beyond "Preferential Attachment".** Mechanisms generating complex features in networks (above all, the power-law degree distribution) have been deeply studied in the previous years, focussing mainly on evolving networks, whose size grows with time. This caused a very rapid development of the scientific field at its start. A seminal paper by A.-L. Barabási et al. introduced the idea that growth and preferential attachments were necessary to the occurrence of fat tails in the degree distribution. Nowadays, such approach is no more a priority: growing networks with various additional mechanisms have been introduced and reproduce almost any of the parameters measured in real networks. Nevertheless, the hypothesis of a dynamically evolving network is not always verified in reality.
- **Clustering and Communities** Measurements and exact results concerning the clustering patterns of networks mainly concern the occurrence of regular motifs and their correlations .However, many social and information networks, such as the World Wide Web, turn out to be approximately partitioned into communities of irregular shape: for example, web pages focusing on similar topics are strongly mutually connected and have a weaker linkage to the rest of the Web. The design of methods to partition a graph into several meaningful highly inter-connected components have then become a compelling application of graph theory to biological, social and information networks.
- **Weighted networks**. While complex networks are usually characterized by their topological complexity, they also often display a large heterogeneity in the capacity and intensity of the connections. In the Internet or in the Web, in ecosystems, or in the world-wide airport network, the strength of interactions varies greatly. This diversity in the weights of the interaction adds a complexity which cannot be overlooked in the study and description of these networks. Studies of this phenomenology as well as new models of complex networks explaining this heterogeneity are therefore necessary
- **Social Networks.** Recently, surveys carried on collaboration networks (authors linked by co-authorships), communication networks (Internet users exchanging e-mail messages) and markets (economic interactions connecting market agents), have confirmed that also social networks display complexity in the degree distribution $P(k)$ and other quantities of interest. Social networks represent then a paradigmatic example for the study of the onset of complexity.

ACTIVITY OF COSIN

As for the way in which basic research has been carried out in the consortium we tried to give our contribution in the world-wide process of research on scale-free networks. As already cited, a large number of the members of the COSIN project can be considered as leading figure in the area of complex dynamical networks. COSIN researchers have contributed to a large number of conferences and several papers with major impact in the community.

A) Data collection and analysis

Communities structure A major focus of the International community has concerned the developments of community detection algorithms with a hectic research activity in which even a large company such as HP have put substantial efforts. In this area the group of Rome and

Barcellona have without doubts contributed with major research papers. The application of spectral analysis in community detection is one of the basic strategies adopted world-wide. As well the practical application to organizational structure and chart analysis made by the group of Barcellona are among the most popular examples of community detection applications.
WebGraph The dataset analysis by the node in Rome La Sapienza represents the state of the art for the topological analysis of WWW data. It is worth to mention that the original paper by A.-L. Barabasi et al. who triggered the interest of statistical physics community in this area was made of nearly 300.000 pages all in the same domain (*.nd.edu). A real web graph is 10^3 times larger and display rather different statistical properties.

Financial and Ecological networks Ecole Normale and INFM has been among the precursors and standard setting event for the research directions in Economical networks. Even if this activity seems not in shape with research on technological networks. still it makes sense to consider the possible application that we could arise from this basic research.

Visualization and centrality measures The use of graph theory and statistical methods along with the skills and knowledge of the Karlsruhe group has opened some research directions that appears not matched at the moment. Noticeably the tools developed in this effort are already publicly available and used by the community at large.

B) Modelling

The project has stimulated and contributed to open the issue of the development of mathematical tools for a characterization of networks beyond the sole degree distribution. The consortium has been at the forefront of the analysis and modelling of high order correlations and their role in the understanding on network structures.

Fitness Models The project developed the idea that edges can be assigned to a set of vertices by making use of specific properties of the vertices themselves. This property, called fitness can be represented by one or more real numbers giving for example the appeal the vertices has in order to be linked by others. By studying different functional form for the linkage (i.e. draw a link when the sum of the fitnesses is larger than a certain threshold) in most of the cases scale-free networks with statistical properties similar to that of the real systems arise.

Multilayer models for the Web By extending the previous idea, one can think that in a multi-contents system like the WWW one page is characterised by a series of different features that could be conveniently described by a series of fitnesses. The system is then supposed to be organised according thematic layers where the different connection take place. The sum of the various ``thematic'' (in/out)-degrees over the different layers forms the total (in/out)-degree of a page.

Agent based models for Internet growth This model introduces a better *microscopic* description of the nodes of the Internet, by taking into account the hierarchical organization in users and Autonomous Systems. Moreover, also the dynamics of the vertices and edges is richer: both new users and AS join the network, although at different rates, users are allowed to change providers and AS can adapt the number of their edges so to satisfy the connectivity demands of their users. This model reproduce many statistical properties of the system among which also the fractal coverage of the network.

4. PROJECT RESULTS AND ACHIEVEMENTS

As regards *the relations and synergies with other relevant projects* the activity of COSIN will largely merge in the activity of the IP DELIS
As regards *any implications for EU policies and standards*, we can comment that the basic research can play an active role in this topic by designing more efficient and reliable information networks. In general any scientific progress in the knowledge of the technological networks could result in more efficient immunization against computer viruses more efficient algorithms for data retrieval and more robust and reliable systems. Another possibility is given by the fact that understanding Internet and WWW growth even at the level of simple statistical models is the first step in the construction of control tools.
For the rest we tried to work as much as possible on the questions related to the original objectives. For such reason, amongst the various activity we present here those that more closely can be put in contact with the previously listed series of objectives (see section 2)

- **Objective 1) A New Theory for Networks Growth**

For a long time all these systems have been considered as haphazard set of points and connections, mathematically framed in the random graph paradigm. This situation has radically changed in the last five years, during which the study of complex networks has received a boost from the ever-increasing availability of large data sets and the increasing computer power for storage and manipulation. In particular, mapping projects of the WWW and the physical Internet represented the the first chance to study the topology of large complex networks. Gradually, other maps followed describing many networks of practical interest in social science, critical infrastructures and biology. Researchers thus have started to have a systematic look at these large data sets, searching for hidden regularities and patterns that can be considered as manifestations of underlying laws governing the dynamics and evolution of these complex systems.
Amongst the regularities that have been found we list the following

- Many of these systems show the small-world property, which implies that in the network the average topological distance between the various nodes increases very slowly with the number of nodes (logarithmically or even slower).
- A particularly important finding is the realization that many networks are characterized by the statistical abundance of "hubs"; i.e. nodes with a large number of connections to other elements. This feature has its mathematical roots in the observation that the number of elements with k links follows a power-law distribution, indicating the lack of any characteristic scale. This has allowed the identification of the class of scale-free networks whose topological features turn out to be extremely relevant in assessing the physical properties of the system as a whole, such as its robustness to damages or vulnerability to malicious attack.

The attempt to model and understand the origin of the observed topological properties of real networks has led to a radical change of perspective, shifting the focus from static graphs, aiming to reproduce the structure of the network in a certain moment, to modeling network evolution. This new approach is the outcome of the realization that most complex networks are the result of a growth process. As a result, we currently view networks as dynamical systems that evolve through the subsequent addition and deletion of vertices and edges.
<u>The contribution of COSIN to this research has been to</u>

a) provide a series of statistical quantities in order to classify the different instances of networks. On top of the degree distribution and the small-world phenomenon we studied the correlation on AS systems between vertices measured by plotting the average degree of vertices neighbors with respect to the degree of the central vertex. This allow to classify the various networks as assortative or disassortative according if this function of degree k is an increasing one or not. Another measure that is currently under investigation by the community is the characteristic density of cycles of particular length. Internet at the autonomous system level is characterized by values not yet reproduced by all the models. (See the book by Vespignani and Pastor-Satorras, Bianconi G, and Capocci A, Phys. Rev. Lett. 90 078701 (2003). Bianconi G, Caldarelli G, A. Capocci, Phys. Rev. E (2003).

b) Shift the interest of the community on the weighted nature of these network, thereby using a generalized degree (defined as sum of the weight incident on the same vertex) and finding new scaling relation describing these systems (see Barrat A, Barthelemy M, Vespignani A, Phys. Rev. Lett. 92 228701 (2004) Barrat A, Barthelemy M, Pastor-Satorras R, and Vespignani A, PNAS, 101, 3747-3752 (2004).

c) produce one of the new ingredients for the modeling of network growth, namely the possibility for a network to increase its edges and vertices not only following the preferential attachment rule. In this new model vertices are attached a characteristic feature that decides their success in the future evolution of the system. In a loose sense, the larger this fitness, the larger the degree this vertex will obtain, thereby linking the microscopic properties of the system with the macroscopic ones (see Caldarelli G, A. Capocci, De Los Rios P, and M. A. Muñoz Phys. Rev. Lett. 89, 258702 (2002)).

d) To study the modification given by congestion and load on network topologies (see R. Guimerà, A. Díaz-Guilera, F. Vega-Redondo, A. Cabrales, and Arenas A, Phys. Rev. Lett. 89, 248701 (2002).

- **Objective 2) A Tool for the analysis of large Networks**

While as tools we intended both theoretical and algorithmic tools as described in the previous objective, but also specific piece of code. One that we believe particularly useful is a piece of software running under Linux that could allow the researchers to produce immediately the plot and the results they needed for basic models of graphs.

The program NetIni is used to produce networks according to the various models of growth and it can be compiled under Linux with the use of specified libraries downloadable from the site. While other codes as for example Pajek (that is nevertheless available only under Microsoft Windows) do a similar job, NetIni has the added value to direct automatically the files of the results to a program for plotting and fitting function. Thereby allowing to compute instantaneously the slope of a power-law. The main features of the code are the following

FEATURES OF NetIni

- Text based initialization file.
- Graphical interface (if you do not like to edit text and if you have Qt 3.0 installed).

- Import of user network structure data (undirected, directed, weighted networks).
- Creation of complex networks using the fitness model of "Caldarelli G. et al., PRL 89, 258702 (2002)".
- Creation of complex networks using the Barabasi-Albert model of preferential attachment.

- Graph reduction with the minimum betweenness criterium.
- Calculation of eigenvalues and eigenvectors of graph associated matrices (requires Lapack and GSL)

- Analysis of most relevant statistical quantities:
    - degree distribution,

- transitivity,
- degree correlation,
- site and edge betweenness,
- pair distances,
- cluster dimensions.

- Direct interface with XmGrace (available under GNU General Public License) at http://plasma-gate.weizmann.ac.il/Grace
- Plot networks with Graphviz and/or Grip.

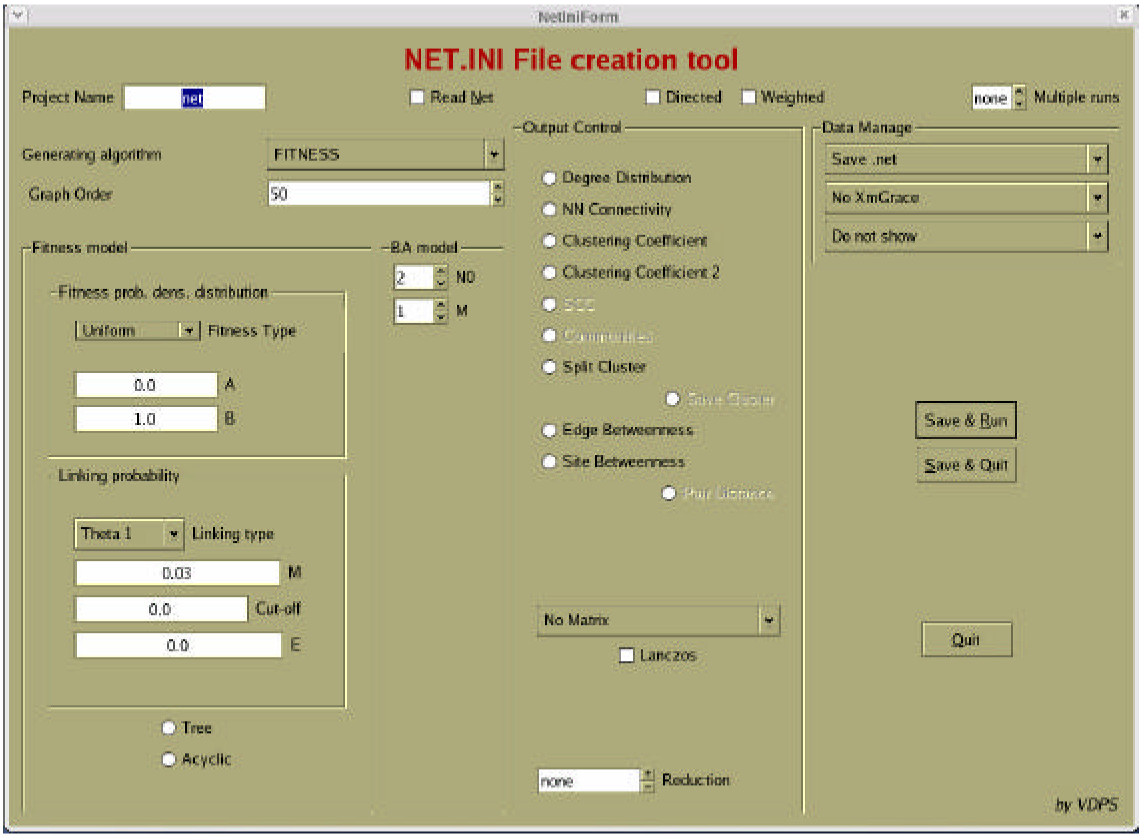


Figure The Graphical Interface of the program NetIni under Linux

- **Objective 3 Efficient visualising methods**

We came up with three different possibilities that are already usable and will be probably refined during the activity in the European project DELIS.

- One form of clustering and grouping are nested decompositions, i.e., sequences of subsets, where every subset contains all following ones. They are a commonly used tool to model hierarchical structures and express structural importance or core-periphery compositions. Unfortunately, in most real networks, this kind of information is only implicitly given. Since manually classifying the elements of these instances is not possible, one needs to find proper graph-theoretic decompositions that approximate or synthesize the original information. Such an instance is the physical Internet at the Autonomous System level. ASes can be roughly classified into: back-bone, national, regional, and local providers as well as customers. Node CR7 UNIKARL verified that this structure is well related to an already known and efficiently computable decomposition called k-cores (see below). Based on this decomposition, it has been established one of the first non-trivial and readable drawings for the AS graph that showed all nodes and edges.

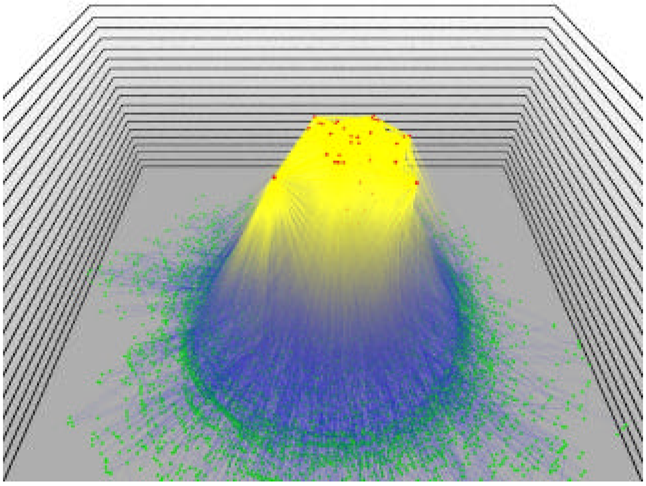


Figure *Hierarchical visualization of the Autonomous System network in 2.5D*

- An alternative way is made by considering a k-core decomposition. This decomposition, based on a recursive pruning of the least connected vertices, allows to disentangle the hierarchical structure of networks by progressively focusing on their central cores. By using this strategy we develop a general visualization algorithm that can be used to compare the structure of various networks and highlight their hierarchical structure. The low computational complexity of the algorithm O(n), where n is the size of the network, makes it suitable for the visualization of very large networks. We apply the proposed visualization tool to several real and synthetic graphs, showing its utility in finding specific structural fingerprints of computer generated and real world networks. layout several topological and hierarchical properties of large scale networks. The k-core decomposition consists in identifying particular subsets of the graph, called k-cores, each one obtained by recursively removing all the vertices of degree smaller than k, until the degree of all remaining vertices is larger than or equal to k. Larger values of core-ness clearly correspond to vertices with larger degree and more central position in the network's structure. When applied to the graphical analysis of real and computer-generated networks, this visualization tool allows the identification of networks' fingerprints, according to properties such as hierarchical arrangement, degree correlations and centrality, etc

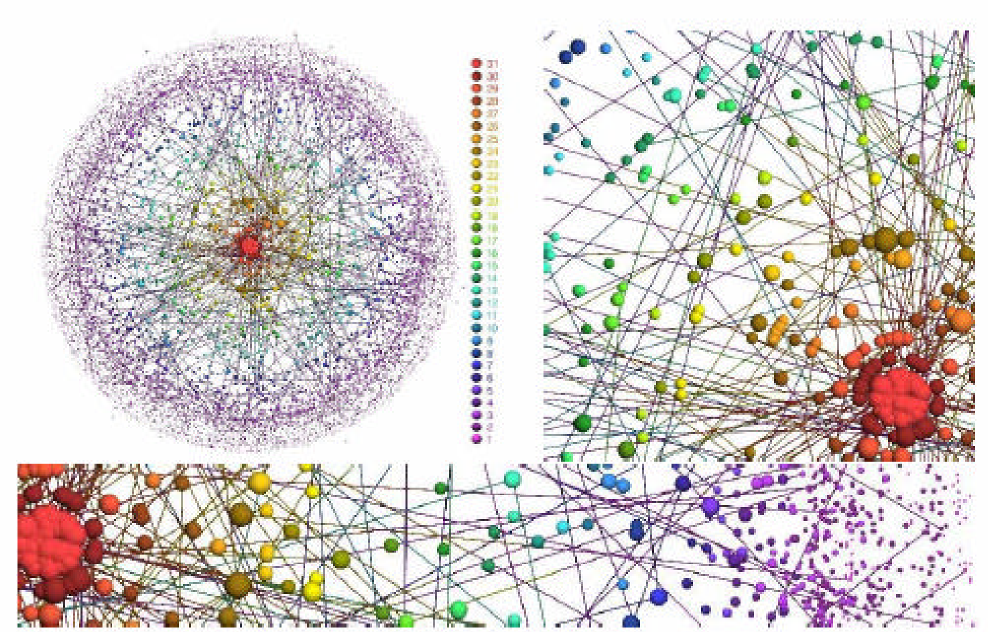


***Figure** A k-core decomposition representing the AS system*

- Another possibility is that to decompose the graph in order to keep only the "relevant" edges as defined by a centrality measure as that of the betweenness. The procedure works as in the case of the Renormalization Group in physics alternating the two process of *decimation* where the edges are removed according to their value of betweenness (low value means deletion) and *rescaling*, that is to say a new computation of the various betweennesses in order to have a ``Rescaled Graph''. From analytical computation and computer simulation this method seem to preserve the main statistical properties of graphs, as degree distribution, clustering coefficient and correlation. The practical drawback is the rescaling procedure where for large graphs we cannot repeat the computation at every time step. In all the cases of interest it seems from preliminary analysis that this condition of rescaling can be relaxed allowing to re-compute this quantity only after a certain number of steps.

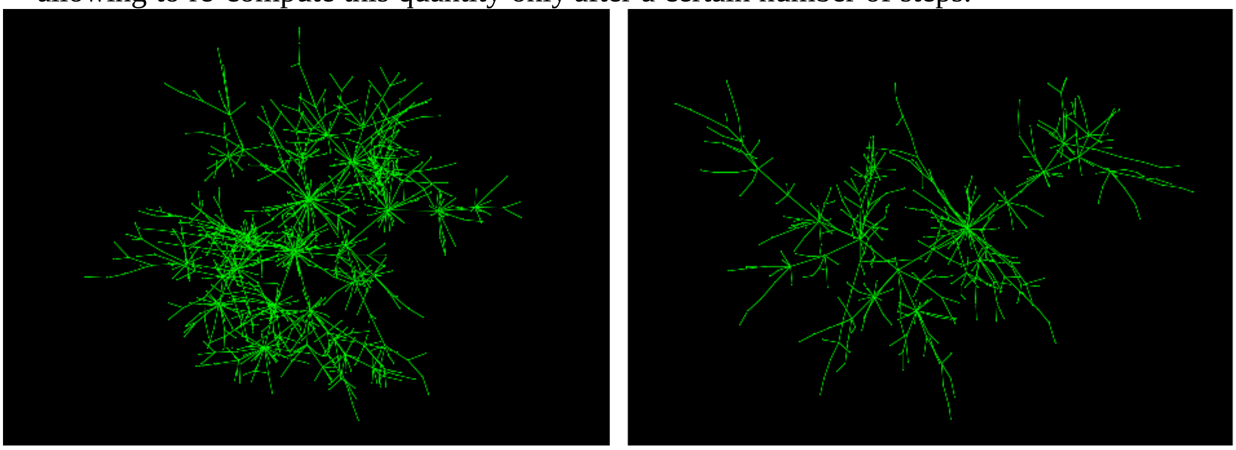


***Figure:** A network of 3000 vertices reduced to 500 vertices, by means of betweenness decimation and rescaling. Networks, rescaling and Images are obtained by NetIni code produced by COSIN. Algorithm is described in this report and is presented in Ref [2]*

- **Objective 4 Analysis of dynamics inside firms**

As for the technological networks the real socio-economical networks are often poorly described by means of random graphs and many dynamical processes display on realistic graphs a behavior dramatically different than on random graphs (ad example the spread of viruses and the Ising model). Moreover, the process of organization of human activities in global world-wide interwoven structures, results in networks the size of which goes beyond the descriptive power of traditional network analysis tools. For this reason we decided to start an analysis of these systems hoping to find similarities/differences with Internet and WWW. Concerning this objective we have achieved several scientific results in the understanding of socio economic networks, from a statistical physics perspective. We were not planning to produce results immediately applicable to business or policy making. However, the results suggest several ideas prompty applicable in further investigations in economics, management and policy making. Indeed, our findings have inspired already several collaboration projects with economists and, remarkably three consulting projects for the French government (such consulting activity has dealt with identification of most influencial decision makers among foreign investors and patterns of investment behavior among largest foreign firms). We have achieved novel understanding mainly in the following topics:

1) Decision making process in the board of an organisation: we have provided novel topological quantities that measure the role of well connected minorities and lobbies

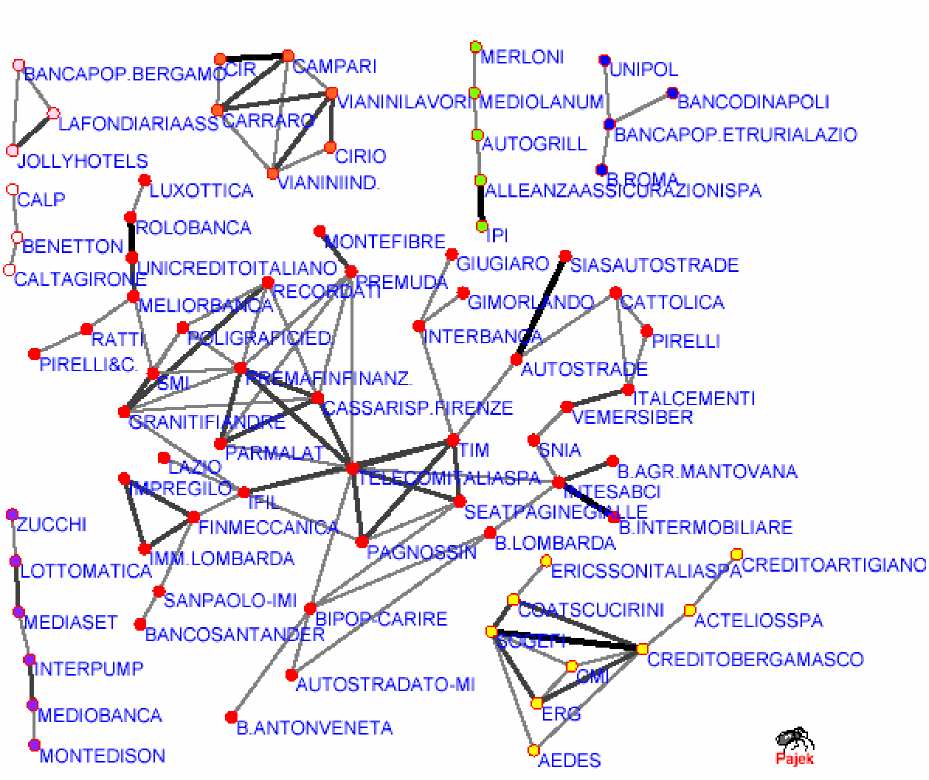
2) Decision making across organizations sharing some of their members : we have characterized two regimes of spreading of decisions, practices and visions through the network of interlocking relationships. Both in topic 1) and 2) we have developed models that will be validated as soon as data about decision processes will be available to us. This is already a big progress with respect to the previous state of the art in the field, which was based on qualitative models, surveys or correlation studies.

3) Structure of the membership and ownership networks in stock markets: we were the first to investigate this topic and we have found that ownership networks are scale free but in some cases the topology reflects an organization in several groups of interest. Specific quantities to detect such groups were introduced.

4) Propagation of failures in production networks: we have developed reasonably realistic models of productions networks that yet can be investigated in a statistical physics perspective. We have found that production networks can exhibit spontaneous macroscopic fluctuations both in space and time.

Details are provided in the descriptions of the deliverables.

Two consulting projects with the French Agency of Foreign Investments were based on results on topic 1, 2, 3. Collaborations with Prof. Gallegati (Univ. Politecnica delle Marche), Prof. Delli Gatti (Univ. Cattolica of Milan) and Prof. Stiglitz (Columbia Univ., Nobel Prize for Economics) were also a consequence of results on topic 1, 2, 3 . This collaboration contributes to our findings on topic 4 and have resulted now in a larger on going project.



- **Objective 5 A model for the Internet and the WWW**

Internet

Finding a model that could reliably describe the Internet's structure and the principles shaping it would be a precious result, since it would open the possibility to play with the Internet, to see the effects of different kind of perturbations and, ultimately, to try designing a better network. Yet, we would like to know the microscopic mechanisms at work in shaping the Internet's structure. At the moment it seems difficult to aim for a *first principles* description of the Internet. Rather, the analysis of Internet data provides a benchmark against which any model that is proposed has to be tested. This has been our approach in this deliverable, and we have found that a purely topological description is very likely insufficient to describe the Internet, and that some further, finer level of details, capturing the intrinsic qualities of nodes and edges, should be included in the models to go beyond simplistic self-referential topological mechanisms. Before the seminal paper of Faloutsos Faloutsos and Faloutsos, none paid any attention on the fact that the Internet was displaying power-law distributions for the degree and a reasonable model might have been the Random Graph one. In order to describe this new statistical feature, the Barabasi-Albert model was introduced. At this point some other quantities as the cycle density or the disassortativity of the system were not reproduced by the model. In this frantic feedback between data analysis and modelling, COSIN played the following role. . Capocci *et al (*Phys. Rev. E **68**, 047101 (2003)) proposed a mechanism where the above mentioned fitness modified PA rule takes place only over nodes whose fitness is larger than the one of the new entering node, mimicking a choice where only intrinsically more authoritative nodes are considered as viable partners. The main result of this paper is that this model is able to reproduce, with a good approximation, the values of the correlations measured for the Internet. A complete description of these approaches together with their own scientific results is available in the book of R. Pastor-Satorras and A. Vespignani.

A better quantitative agreement is recovered by a new model introduced by Ángeles Serrano *et al.* (Phys. Rev. Lett **94**, 038701 (2005)). This model introduces a better *microscopic* description of the nodes of the Internet, by taking into account the hierarchical organization in users and Autonomous Systems. Moreover, also the dynamics of the vertices and edges is richer: both new users and AS join the network, although at different rates, users are allowed to change providers and AS can adapt the number of their edges so to satisfy the connectivity demands of their users. Each part of the dynamics being characterized by specific rates, partly fitted to the empirical growth rates of the real Internet, the network of Autonomous Systems turns out to be scale-free and the power-law degree distribution decays with an exponent close to 2, as measured in real data. Starting from these premises, Ángeles Serrano and co-workers then refine their model placing AS in two-dimensional space in such that they cover a fractal set, as measured for the Internet [13]. Connections between AS take then place according to the same rules as above, but also taking into account the costs of long-distance connections and the bandwidth needs of different AS (that is, only AS that need to increase their bandwidth because of an increase of users can link to each other). The final result is a model where very many microscopic details close to the ones of the Internet have been introduced and that is able to closely reproduce many of its correlation properties. Far from being an arrival point, the model of Ángeles Serrano and co-workers shoes that more and more details have to be plugged back in the models to obtain detailed a reproduction of the real data

WWW

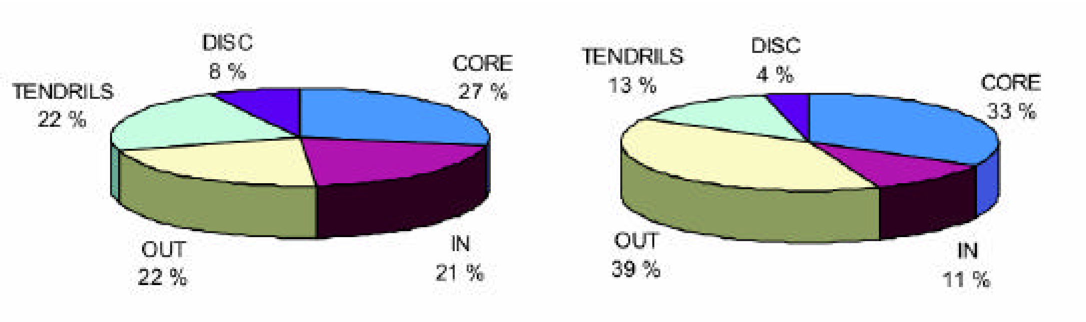
The most noticeable activity of the Web graph, has been made by node CR2 UDRLS repeated the experiments of Broder et al. [3] on the macroscopic analysis of the graph computing the in-degree, out-degree and WCC size distributions of a large Webgraph of more than 300 million pages. In order to crawl this large data sets new algorithms were needed, in particolar we implemented algorithms that achieve remarkable performance improvements when processing data that are stored on external memory. We implemented *semi-external* algorithms, that use only a small constant amount of memory for each node of the graph, as well as *fully-external* algorithms that use an amount of main memory that is independent of the graph size.

We implemented the following algorithms.

- External versions of Breadth and Depth First search, based on random accesses to the disk, in order to avoid maintaining the data in main memory.
- The traditional traversal algorithms that work in main memory.
- A semi-external graph traversal that allows to determine the verteces reachability for determining vertex reachability using only 2 bits per node. The one bit is set when the node is first visited, and the other when all its neighbors have been visited (we say that the node is ``completed'')The algorithm operates on the principle that the order in which the vertices are visited is not important. Starting from an initial set of nodes, it performs multiple passes over the data, each time visiting the neighbors of the non-completed nodes. A semi-external Breadth First Search that computes blocks of reachable nodes and splits them up in layers according to their distance from the root. In a second step, these layers are sorted to produce the standard BFS traversal of the graph
- A semi-external Depth First Search (DFS) that needs 12 bytes plus one bit for each node in the graph. This traversal has been developed following that approach suggested by Sibeyn et al. [6] .
- A semi-external algorithm for computing all SCCs of the graph based on the semi-external DFS.
- An algorithm for computing the largest SCC of the Web graph. The algorithm adopts a heuristic approach that exploits the structural properties of the Web graph to compute the biggest SCC, using a simple reachability algorithm. As a result of the algorithm The algorithm exploits the fact that the largest SCC is a sizable fraction of the Web graph. Thus, by sampling a few nodes of the graph, we can obtain a node of the largest SCC with high probability.

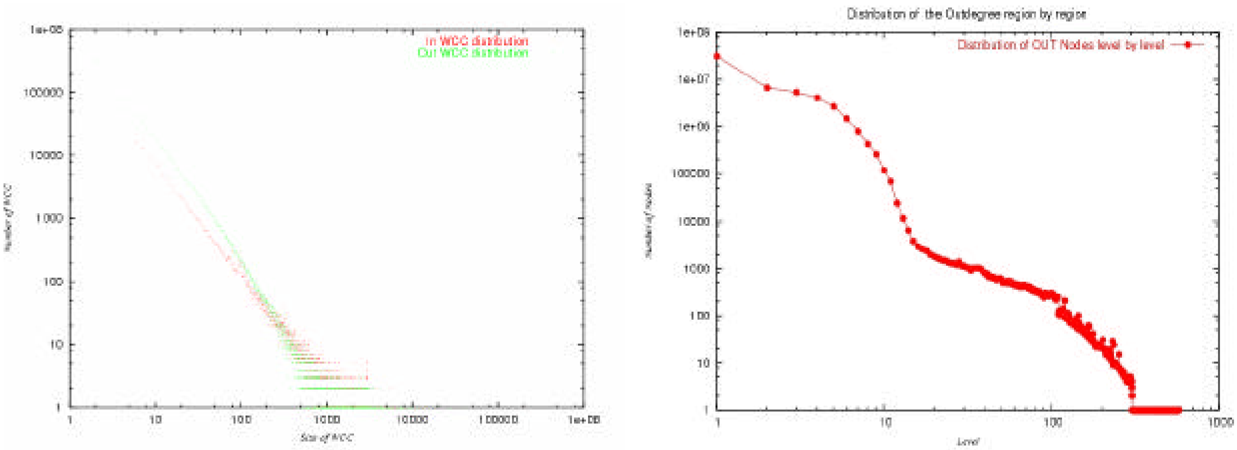
We can then identify the nodes of the SCC using the reachability algorithm. As an end product we obtain the bow-tie regions of the Web graph, and we are able to compute all the remaining SCCs of the graph efficiently using the semi-external DFS algorithm.

A software library containing a suite of algorithms for generating and processing massive Web graphs is available online **http://www.dis.uniroma1.it/~cosin/**.
A detailed presentation of some of these algorithms and a study of their efficiency has been presented in [7]. A complete description of these algorithms is available in the extended version of this work [8].



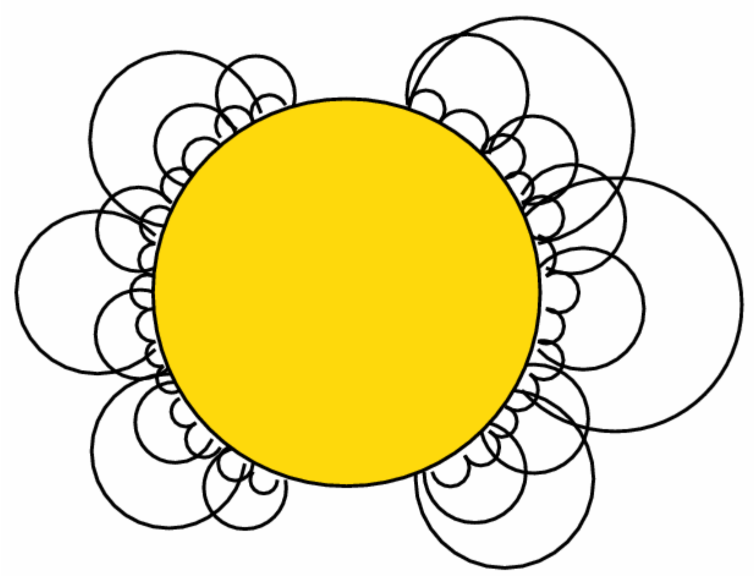
The proportion of the various component of the Web Graph
On the left the AltaVista Graph on the right the analysis on the WebBase Graph

As expected, the in-degrees, and the sizes of SCCs follow a power-law distribution, while the out-degree distribution follows an imperfect power-law.



- The inner structure as regards degree distributions and sizes of WCC and SCC is the same for the IN and OUT component.
- The CORE has *entry points* that is to say nodes connected with at least one node in the IN region and has also *exit points* that point at least to a node in the OUT region. *Bridges* are told those vertices that are both entry and exit points. Surprising the vast majority of vertices (~72 %) are EXIT points. This means that about 80% of the core is connected to the external regions, while only 20% belong to inner core region.
- One can also define connectors those vertices of the core that have a single in-coming and out-coming link. A connector forms a petal if if they coincide. Surprisingly only 6% of the CORE vertices are connectors (47% of which are petals).

Putting together the various feature it seems that a different macroscopic picture of the WWW must be done. Instead of Bow-tie a Daisy seem to represent better the shape of the graph,



The proposed Daisy shape for the structure of the WWW

5. DELIVERABLES AND REFERENCES

5.1 List of Deliverables

First Year	D1: Power Point Presentation
	D2-D3: Dissemination and Use Plan + Setup Advisory Board
	D4: Universality in Networks
	D5: Preliminary Analysis of Collected data
	D6: Algorithms for Network centrality
	D7: Centrality and Groups in Social networks
	D8: Modelling WWW
	D9: First Year report
Second Year	D10: Check of the state of the art
	D11: Self-organized criticality in network formation
	D12: Database describing complex networks, internet and www
	D13: A library of software tools for performing measures on large networks
	D14: Customization and usability study of general purpose software tools for visualization of large networks
	D15: Modelling dynamics and interactions in firms structures
	D16: Inter-firm network dynamics: risk propagation fusion/outsourcing dynamics
	D17: Algorithms for network traffic analysis
	D18: Second Year report
Third Year	D19: Optimization and network shaping
	D20: Statistical properties of collected data
	D21: Web interface for datasets
	D22: Algorithms to find paths and connections from local information
	D23: Cyber-communities in World Wide Web
	D24: Modelling of the Internet graph
	D25: Organization of workshops, schools and conferences
	D26: Papers and book on growing networks and complexity
	D27: WWW site with COSIN results
	D28: Technology implementation plan
	D29: Third Year report

5.2 List of References

2005 publications

- Tuning clustering in random networks with arbitrary degree distributions
  Angeles Serrano M, and Boguna M,
  cond-mat/0507535

- Weighted Configuration Model
  Angeles Serrano M, and Boguna M,
  cond-mat/0501750

- Competition and Adaptation in an Internet Evolution Model
  Angeles Serrano M, Boguna M, and Diaz-Guilera A,
  Phys. Rev. Lett. **94**, 038701 (2005)

- Rate equation approach for correlations in growing network models
  Barrat A, and Pastor-Satorras R,
  Phys. Rev. E **71**, 036127 (2005)

- The effects of spatial constraints on the evolution of weighted complex networks
  Barrat A, Barthelemy M, Vespignani A
  J. Stat. Mech. P05003 (2005)

- Dynamical patterns of epidemic outbreaks in complex heterogeneous networks
  Barthelemy M, Barrat A, Pastor-Satorras R, Vespignani A,
  J. of Theor. Biol. **235**, 275-288 (2005)

- Characterization and modeling of weighted networks
  Barthelemy M, Barrat A, Pastor-Satorras R, Vespignani A,
  Physica A **346**, 34-43 (2005)

- Efficient generation of large random networks
  Batagelj V, Brandes U
  Phys. Rev. E **71** 036113 (2005)

- Parallel scheduling problems in next generation wireless networks
  Becchetti L, Leonardi S, Marchetti-Spaccamela A, et al.
  Networks **45** 9-22 (2005)

- Generalized percolation in random directed networks
  Boguna M, and Angeles Serrano M,
  Phys. Rev. E **72**, 016106 (2005)

- Studying the emerging global brain Analyzing and visualizing the impact of co-authorship teams
  Borner K, Dall'Asta L, Ke WM, et al.
  Complexity **10** 57-67 (2005)

- Centrality measures based on current flow
  Brandes U, Fleischer D
  Lect. Comp. Sci. **3404** 533-544 (2005)

- Network analysis - Methodological foundations - Introduction
  Brandes U, Erlebach T
  Lect. Comp. Sci. **3418** 1 (2005)

- Fundamentals
  Brandes U, Erlebach T
  Lect. Comp. Sci. **3418** 7-15 (2005)

- Universal scaling in food-web structure
  Camacho J and A. Arenas A
  Nature **435**, E3-4 (2005)

- Comparison of voter and Glauber ordering dynamics on networks
  Castellano C, Loreto V, Barrat A, et al.
  Phys. Rev. E **71** 066107 (2005)

- Generation of uncorrelated random scale-free networks
  Catanzaro M, Boguna M, and Pastor-Satorras R,
  Phys. Rev. E **71**, 027103 (2005)

- Analytic solution of a static scale-free network model
  Catanzaro M, and Pastor-Satorras R,
  Eur. Phys. J. B **44**, 241-248 (2005)

- Characterization and modeling of protein-protein interaction networks
  Colizza V, Flammini A, Maritan A, et al.
  Phys. A **352** 1-27 (2005)

- Statistical theory of Internet exploration
  Dall'Asta L, Alvarez-Hamelin I, Barrat A, et al.
  Phys. Rev. E **71** 036135 (2005)

- Community structure identification
  Danon L, Duch J, Arenas A, and Diaz-Guilera A,
  cond-mat/0505245

- Improved spectral algorithm for the detection of communities
  Donetti L, and A Munoz M,
  physics/0504059

- Entangled networks, super-homogeneity and optimal network topology
  Donetti L, I Hurtado P, and A Munoz M,
  cond-mat/0502230

- Community detection in complex networks using Extremal Optimization
  Duch J, and Arenas A,
  Phys. Rev. E **72**, 027104 (2005)

- Spectral methods cluster words of the same class in a syntactic dependency network
  Ferrer i Cancho R, Capocci A and Caldarelli G
  cond-mat/0504165

- The variation of Zipf's law in human language
  Ferrer i Cancho R

Eur. Phys. J. B **44** 249-257 (2005)

- The consequences of Zipf's law for syntax and symbolic reference
  Ferrer i Cancho R, Bollobas B and Riordan O
  Proc. R. Soc. Lond. Series B 272, 561-565 (2005)

- Universal scaling in food-web structure
  Garlaschelli D, Caldarelli G and Pietronero L
  Nature **435**, E4 (2005)

- The Scale-free topology of market investments.
  Garlaschelli D, Battiston S, Castri M, Servedio VDP, Caldarelli G,
  Physica A **350** 491 (2005)

- Efficiency of informational transfer in regular and complex networks
  Vragovic I, Louis E, and Diaz-Guilera A,
  Phys. Rev. E **71**, 036122 2005

- Persuasion dynamics
  Weisbuch G, Deffuant G, Amblard F
  Phys. A **353** 555-575 (2005)

2004 publications

- Virtual Round Table on ten leading questions for network research
  Amaral LAN, Barrat A, Barabasi AL, et al.
  Eur. Phys. J. B **38** 143-145 (2004)
- Emergence of complex dynamics in a simple model of signaling networks
  Amaral LAN, Diaz-Guilera A, Moreira AA, Goldberger AL, and Lipsitz LA
  PNAS **101**, 15551-15555 (2004)

- Community analysis in social networks
  Arenas A, Danon L, Diaz-Guilera A, Gleiser PM, Guimera R,
  Eur. Phys. J. B, **38**, 373-380 (2004)

- Local search with congestion in complex communication networks
  Arenas A, Danon L, Diaz-Guilera A, et al.
  Lect. Comp. Sci. **3038**, 1078-1085 (2004)
- Weighted evolving networks: coupling topology and weights dynamics
  Barrat A, Barthelemy M, Vespignani A,
  Phys. Rev. Lett. **92** 228701 (2004)

- The architecture of complex weighted networks
  Barrat A, Barthelemy M, Pastor-Satorras R, and Vespignani A,
  PNAS, **101**, 3747-3752 (2004)

- Modeling the evolution of weighted networks
  Barrat A, Barthelemy M, Vespignani A
  Phys. Rev. E **70** 066149 (2004)
- Traffic-driven model of the World Wide Web graph
  Barrat A, Barthelemy M, Vespignani A
  Lect. Comp. Sci. **3243** 56-67 (2004)
- Velocity and Hierarchical Spread of Epidemic Outbreaks in Scale-Free Networks
  Barthelemy M, Barrat A, Pastor-Satorras R, and Vespignani A,
  Phys. Rev. Lett. **92**, 178701 (2004)

- Betweenness centrality in large complex networks
  Barthelemy M
  Eur. Phys. J. B **38** 163-168 (2004)
- Statistical properties of board and director networks
  Battiston S, and Catanzaro M,
  Eur. Phys. J. B **38** 245 (2004)
- Inner structure of capital control networks
  Battiston S
  Physica A **338** 107-112 (2004)
- Drawing the AS graph in 2.5 dimensions
  Baur M, Brandes U, Gaertler M, et al.
  Lect. Comp. Sci. **3383** 43-48 (2004)
- Crossing reduction in circular layouts
  Baur M, Brandes U
  Lect. Comp. Sci. **3353** 332-343 (2004)
- Models of social networks based on social distance attachment
  Boguna M, Pastor-Satorras R, Diaz-Guilera A, and Arenas A
  Phys. Rev. E, vol. 70, 056122 (2004)
- Conditional dynamics driving financial markets
  Boguna M, Masoliver J
  Eur. Phys. J. B **40** 347-352 (2004)
- Cut-offs and finite size effects in scale-free networks
  Boguna M, Pastor-Satorras R, and Vespignani A,
  Eur. Phys. J. B **38**, 205-209 (2004)
- GraphML transformation
  Brandes U, Pich C
  Lect. Comp. Sci. **3383** 89-99 (2004)
- Structural similarity in graphs - A relaxation approach for role assignment
  Brandes U, Lerner J
  Lect. Comp. Sci. **3341** 184-195 (2004)
- How to draw the minimum cuts of a planar graph
  Brandes U, Cornelsen S, Fiess C, et al.
  Comp. Geom. **29** 117-133 (2004)
- Visualizing related metabolic pathways in two and a half dimensions
  Brandes U, Dwyer T, Schreiber F
  Lect. Comp. Sci. **2912** 111-122 (2004)
- Characterizing families of cuts that can be represented by axis-parallel rectangles
  Brandes U, Cornelsen S, Wagner D
  Lect. Comp. Sci. **2912** 357-368 (2004)
- Generating Node Coordinates for Shortest-Path Computations in Transportation Networks
  Brandes U, Schulz F, Wagner D and Willhalm T
  J. Exp. Alg. **9** 1.1 (2004)
- Visual Understanding of Metabolic Pathways Across Organisms Using Layout in Two and a Half Dimensions
  Brandes U, Dwyer T and Schreiber F

J. Int. Bionf. **0002** (2004)

- Graph drawing contest report
  Brandenburg FJ, Brandes U, Eades P, et al.
  Lect. Comp. Sci. **2912** 504-508 (2004)

- Structure of cycles and local ordering in complex networks
  Caldarelli G, Pastor-Satorras R, and Vespignani A,
  Eur. Phys. J. B **38**, 183-186 (2004)

- Preferential exchange Strengthening connections in complex networks
  Caldarelli G, Coccetti F, De Los Rios P
  Phys. Rev. E **70** 027102 (2004)

- Statistical features of drainage basins in mars channel networks - Can one guess from the landscape the past presence of water?
  Caldarelli G, De los Rios P, Montuori M, et al.
  Eur. Phys. J. B **38** 387-391 (2004)

- Widespread occurrence of the inverse square distribution in social sciences and taxonomy
  Caldarelli G, Cartozo CC, De Los Rios P, et al.
  Phys. Rev. E **69** 035101 (2004)

- Emergence of Complexity in Financial Networks
  Caldarelli G, Battiston S, Garlaschelli D, Catanzaro M,
  Lect. Notes Phys. "Complex Networks", Springer-Verlag (2004)

- Structure of cycles and local ordering in complex networks.
  Caldarelli G, Pastor-Satorras R, and Vespignani A,
  Eur. Phys. J. B **38** 183-186 (2004)

- Preface on "Applications of Networks"
  Caldarelli G, Erzan A, Vespignani A
  Eur. Phys. J. B **38** 141 (2004)

- Archive of BGP Updates: Integration and Visualization.
  Di Battista G,Mariani F,Patrignani M and Pizzonia M
  Graph Drawing **2192** 295-306 (2004)

- Large scale properties of the Webgraph
  Donato D, Laura L, Leonardi S, et al.
  Eur. Phys. J. B **38** 239-243 (2004)

- Simulating the webgraph. A comparative analysis of models
  Donato D, Laura L, Leonardi S, et al.
  Comp. in Sci. & Eng. **6** 84-89 (2004)

- Detecting network communities
  Donetti L, and A Munoz M,
  J. Stat. Mech. P10012 (2004)

- Euclidean distance between syntactically linked words
  Ferrer i Cancho R
  Phys. Rev. E **70** 056135 (2004)

- Model of experience for public organisations with staff mobility
  Kitowski J, Krawczyk K, Majewska M, et al.
  Lect. Comp. Sci. **3025**, 75-84 (2004)

- Efficiency and reliability of epidemic data dissemination in complex networks
  Moreno Y, Nekovee M, Vespignani A
  Phys. Rev. E **69** 055101 (2004)

- Role of clustering and gridlike ordering in epidemic spreading
  Petermann T, De Los Rios P
  Phys. Rev. E **69** 066116 (2004)

- Cluster approximations for epidemic processes a systematic description of correlations beyond the pair level
  Petermann T, De Los Rios P
  J. Theor. Biol. **229** 1-11 (2004)

- Exploration of scale-free networks - Do we measure the real exponents?
  Petermann T, De los Rios P
  Eur. Phys. J. B **38** 201-204 (2004)

- Social Interactions in Economic Theory: an Insight from Statistical Mechanics
  Phan D, M. B. Gordon M. B. and J.-P. Nadal
  (2004)

- Self-organization of collaboration networks
  Ramasco JJ, Dorogovtsev SN and Pastor-Satorras R,
  Phys. Rev. E **70**, 036106 (2004)

- Generation of Scale Free Random Networks within the Formalism of Hidden Variables
  Servedio VDP, Caldarelli G, Butta P,
  Phys. Rev. E (2004)

- Influence of topology on the performance of a neural network
  Torres JJ, Munoz MA, Marro J, et al.
  Neurocomp. **58** 229-234 (2004)

- Bounded confidence and social networks
  Weisbuch G
  Eur. Phys. J. B **38** 339-343 (2004)

2003 publications

- Decision spread in the corporate board network
  Battiston S, Weisbuch G, Bonabeau E
  Adv. Com. Sys. **6** 631-644 (2003)

- Decision making dynamics in corporate boards.
  Battiston S, Bonabeau E and Weisbuch G,
  Phys. A **322** 567 (2003).
- Size of Quantum Networks.
  Bianconi G,
  Phys. Rev. E **67**, 056119 (2003)
- Number of Loops of size *h* in growing scale-free networks.
  Bianconi G, and Capocci A,
  Phys. Rev. Lett. **90** 078701 (2003).
- Number of h-cycles in the Internet at the Autonomous System Level.
  Bianconi G, Caldarelli G, A. Capocci,

Phys. Rev. E (2003).

- Absence of Epidemic threshold in scale-free networks with degree correlations.
  Boguna M, Pastor-Satorras R, and Vespignani A,
  Phys. Rev. Lett. **90**, 028701 (2003).
- Emergence of clustering, correlations, and communities in a social network model
  Boguna, Pastor-Satorras R, Diaz-Guilera A, and Arenas A,
  cond-mat/0309263
- Class of correlated random networks with hidden variables
  Boguna M, Pastor-Satorras R
  Phys. Rev. E **68** 036112 (2003).
- Topology of correlation based minimal spanning trees in real and model markets.
  Bonanno G, Caldarelli G, Lillo F and Mantegna R
  Phys. Rev. E **68** 046130 (2003).
- Special Issue - Algorithmic Aspects of Communication - University of Bologna,- Foreword
  Bonuccelli M, Marchetti-Spaccamela A
  Discrete Applied Mathematics **129** 1 (2003).
- Centrality in Policy Network Drawings.
  Brandes U, Kenis P, and Wagner D
  IEEE transactions on Visualization and Computer Graphics. **9**, 241 (2003).
- Experiment on Graph Clustering Algorithms
  Brandes U, Gaertler M and Wagner D
  Lect. Comp. Sci. **2832** 568 (2003).
- Visual Ranking of Link Structures
  Brandes U, S. Cornelsen
  J. Graph Algor. and Appl. **7** 181-201 (2003).
- Visual Unrolling of Network Evolution and the Analysis of Dynamic Discourse
  Brandes U and Corman SR
  Information Visualization **2** 40 (2003).
- Communicating Centrality in Policy Network Drawings
  Brandes U, Kenis P, and Wagner D
  IEEE Transactions on Visualization and Computer Graphics **9** 241-253 (2003).
- Analysis and Visualization of Social Networks
  Brandes U and Wagner D
  Graph Drawing Software Michael Jünger and Petra Mutzel (eds.) 321-340 Springer-Verlag 2003.
- Generalized Network Growth: Microscopic Strategies and Real Internet Properties.
  Caldarelli G, De Los Rios P, L. Pietronero,
  ArXiv:cond-mat/0307610 (2003).
- Quantitative description and modelling of real networks.
  Capocci A, Caldarelli G, and De Los Rios P,
  Phys. Rev. E **68** 047101 (2003).
- **C. Castellano, D. Vilone, Vespignani A,**
  Incomplete ordering of the voter model on small-world networks
  Europhys. Lett. **63** 153-158 (2003).
- A System for Building Animated Presentations over the Web,
  B.A. Colombo, C. Demetrescu, I. Finocchi, L. Laura
  Proceedings of the AICCSA'03 Workshop on Practice and Experience with Java Programming in Education, Tunis, (2003).
- A Portable Virtual Machine for Program Debugging and Directing,
  C. Demetrescu and I. Finocchi
  Technical Report ALCOM-FT, **ALCOMFT-TR-03-1**,(2003).
- Dynamic Graph Algorithms
  C. Demetrescu, I. Finocchi, G.F. Italiano
  Handbook of Graph Theory , **10.2**, J. Yellen, J.L. Gross eds.,
  CRC Press, Discrete Mathematics and its Applications, ISBN 1-58488-090-2, (2003).
- Combinatorial Algorithms for Feedback Problems in Directed Graphs,
  C. Demetrescu and I. Finocchi

Inf. Proc. Lett. **86** 129-136 (2003).

- Experimental Analysis of Dynamic All Pairs Shortest Path Algorithms,
C. Demetrescu, S. Emiliozzi, and G.F. Italiano
Technical Report ALCOM-FT, **ALCOMFT-TR-03-9** (2003).
- Comentarios a "El problema del mundo pequeño" de S. Milgram,
A. Díaz-Guilera, Arenas A, R. Guimerà,
Araucaria **10** 29 (2003).
- Fully dynamic shortest paths in digraphs with arbitrary arc weights
D. Frigioni, A. Marchetti-Spaccamela, U. Nanni
Algorithm **49** 86-113 (2003).
- Universality in Food webs
Garlaschelli D, Caldarelli G, L. Pietronero
Nature **423** 165 (2003).
- Macro and micro structures of trust networks
X. Guardiola, R. Guimerà, Arenas A, Diaz-Guilera A, D. Streib, L.A.N. Amaral
ArXiv:cond-mat/0206240
- Self-similar community structure in a network of human interactions
R. Guimerà, Danon L, Diaz-Guilera A, F. Giralt, Arenas A,
Phys. Rev. E **68** 065103 (2003).
- A Multi-Layer Model for the Web Graph
L. Laura, S. Leonardi, Caldarelli G, De Los Rios P,
Proceedings of the 2nd International Workshop on Web Dynamics (2003).
- Optimization as a result of the interplay between dynamics and structure,
M. Llas, P. M. Gleiser, Diaz-Guilera A, C. J. Perez,
Phys. A **326** 567 (2003).
- Nonequilibrium phase transition in a model for the propagation of innovations among economic agents,
M. Llas, P. M. Gleiser, J. M. Lopez, Diaz-Guilera A,
Phys. Rev. E **68** 066101 (2003).
- Degree stability of a minimum spanning tree of price return and volatility
S. Micciché, G. Bonanno, F. Lillo, R.N. Mantegna
Phys. A **324** 66-73 (2003).
- Critical load and congestion instabilities in scale-free networks
Y. Moreno, Pastor-Satorras R, A. Vazquez et al.
Europhys. Lett. **62** 292-298 (2003).
- Stochastic theory of synchronization transitions in extended systems
M.A. Muñoz, Pastor-Satorras R,
Phys. Rev. Lett. **90** 204101 (2003).
- Monopoly Market with Externality: an Analysis with Statistical Physics and Agent-based Computational Economics
J.-P. Nadal, Phan D, M. B. Gordon and J. Vannimenus,
WEHIA 2003, Kiel (2003).
- Evolving Protein Interaction Networks through gene duplication.
Pastor-Satorras R, E. Smith, R. V. Sole
J. Theor. Biol. **222** 199 (2003).
- Search and Congestion in Complex Networks,
Pastor-Satorras R, M. Rubi, and Diaz-Guilera A,
Proceedings of the Conference "Statistical Mechanics of Complex Networks", Sitges, Spain, June 2002, eds., Springer (2003)
- Scale-free behavior of the Internet global performance
R. Percacci, Vespignani A,
Eur. Phys. J. B **32** 411-414 (2003).
- The Monopolist's Market with Discrete Choices and Network Externality Revisited: Small-Worlds, Phase Transition and Avalanches in an ACE Framework
Phan D, S. Pajot, J.-P. Nadal
Ninth annual meeting of the Society of Computational Economics, University of Washington, Seattle, USA,

- Is there something out there? Inferring space from sensorimotor dependencies
  D. Philipona, K. O'Regan and J.-P. Nadal
  Neural Computation **15** 2029-2049 (2003).
- Topology of the World Trade Web.
  M. A. Serrano, and M. Boguñá
  Phys. Rev. E **68** 151011-151014 (2003).
- Modeling traffic of information packets on graphs with complex topology
  B. Tadic
  Lect. Comp. Sci. **2657** 136-143 (2003).
- Orientational minimal redundancy wavelets: from edge detection to perception
  A. Turiel, J.-P. Nadal and N. Parga
  Vision Research **43** 1061-1079 (2003).
- Topology and correlations in structured scale-free networks
  A. Vazquez, M. Boguñá, Y. Moreno, et al.
  Phys. Rev. E **67** 046111 (2003).
- **Vespignani A,**
  Evolution thinks modular
  Nature Genetics **35** 118-119 (2003).
- Analysis and Visualization of Social Networks.
  Wagner D
  Lect. Comp. Sci. **2647** 261 (2003).
- **Weisbuch G, Stauffer D,**
  Adjustment and social choice
  Phys. A **323** 651-662 (2003)
- A market of inhomogeneous threshold cellular automata
  Stauffer D, Weisbuch G
  Int. J. Mod. Phys. B **17** 5495-5501 (2003)

2002 publications

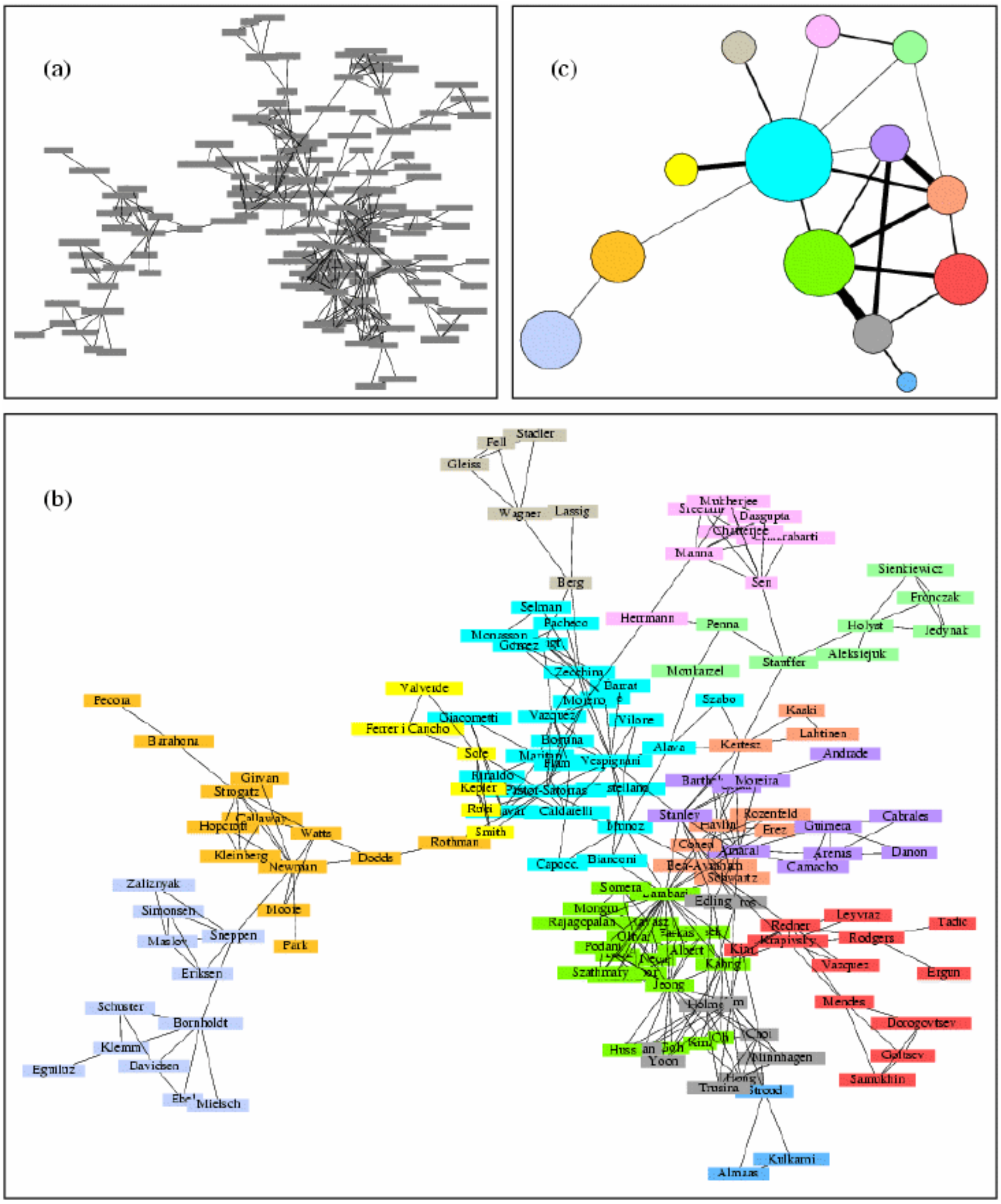
- Epidemic spreading in complex networks with degree correlations.
  M. Boguñá Pastor-Satorras R, and Vespignani A,
  Phys. Rev. E **66**, 047104 (2002).
- GraphML Progress Report: Structural Layer Proposal.
  Brandes U, M. Eiglsperger, I. Herman, M. Himsolt, M. Scott Marshall
  Lect. Comp. Sci. **2265** 501-512 (2002).
- Guide - Changes
  V. Maag, T. Willhalm
  Technical Report, University of Konstanz (2002).
- Visone - Analysis and Visualization of Social Networks.
  Brandes U and Wagner D
  Lect. Comp. Sci. **2265** 463 (2002).
- Introduction to Complex Networks.
  Caldarelli G,
  Proceedings of the 7th Conference on Statistical and Computational Physics Granada (2002).
- Scale-Free Networks from Varying Vertex Intrinsic Fitness.
  Caldarelli G, A. Capocci, De Los Rios P, and M. A. Muñoz
  Phys. Rev. Lett. **89**, 258702 (2002).
- Visualization of the High Level Structure of the Internet with Hermes.
  A.Carmignani, G. Di Battista, W. Didimo, F. Matera and M. Pizzonia
  J. Graph Algor. and Appl. **6**,281 (2002).
- How Can Extremism Prevail? A study based on the relative agreement interaction model.

G. Deffuant, F. Amblard, Weisbuch G, and T. Faure
J. Artif. Soc. and Soc. Simul. **5**(2002).

- Exploration bias of Complex Networks.
  De Los Rios P,
  Proceedings of the 7th Conference on Statistical and Computational Physics Granada (2002).
- Drawing database schemas.
  G. Di Battista, W. Didimo, M. Patrignani and M. Pizzonia
  Software-Practice and Experience. **32**, 1065 (2002).
- Computing the types of the Relationships Between Autonomous Systems.
  G. Di Battista, M. Patrignani and M. Pizzonia
  Technnical Report University of Rome 3 **RT-DIA-73**(2002).
- Dynamical properties of model communication networks.
  R. Guimerà, Arenas A, A. Díaz-Guilera and F. Giralt
  Phys. Rev. E **66**, 026704 (2002).
- Optimal network topologies for local search with congestion.
  R. Guimerà, A. Díaz-Guilera, F. Vega-Redondo, A. Cabrales, and Arenas A,
  Phys. Rev. Lett. **89**, 248701 (2002).
- A Multi-layer model for the Web Graph
  L. Laura, S. Leonardi, Caldarelli G, and De Los Rios P,
  Contribution to 2nd Web Dynamics Workshop (2002).
- Critical load and congestion instabilities in scale-free networks.
  Y. Moreno, Pastor-Satorras R, A. Vázquez and Vespignani A,
  Europhys. Lett. **62**, 292-298 (2002).
- Scale-free behavior of the Internet global performance.
  Percacci R and Vespignani A,
  Eur. Phys. J. B **32** 411-414 (2003).
- Epidemics and immunization in scale-free networks.
  Pastor-Satorras R, and Vespignani A,
  Handbook of Graphs and Networks: From the Genome to the Internet, eds. S. Bornholdt and H. G. Schuster, Wiley-VCH, Berlin, pp. 113-132 (2002).
- Topology and correlations in structured scale-free networks.
  A. Vázquez, M. Boguñá, Y. Moreno, Pastor-Satorras R, and Vespignani A,
  Phys. Rev. E **67** 461111 (2003).
- Large-scale topological and dynamical properties of the Internet.
  A. Vázquez, Pastor-Satorras R, and Vespignani A,
  Phys. Rev. E **65**, 066130 (2002).
- Large-scale topological and dynamical properties of the Internet.
  A. Vázquez, Pastor-Satorras R, and Vespignani A,
  Phys. Rev. E **65** 1-12 (2002).

6. POTENTIAL IMPACT OF PROJECT RESULTS

To have an idea of the kind of impact achieved by COSIN project we can use the analysis of Mark Newman (related only to the first years of activity of the project) the network of authors in the field can be represented as follows (Physical Review E **69** 026113 (2004))



The network of production of papers on Networks. A first visualization through names, a second through communities and the larger is a mix of the two.

Using the word of the author describing this collaboration network:

"**The network is centered around the middle group shown in cyan (which consists of researchers primarily in southern Europe)**".

Actually the large majoirity of this central cluster is made by COSIN people (Bianconi (CR4-CR9), Boguna, Munoz, Pastor-Satorras, (CR3), Caldarelli (C01), Capocci (C01-CR4), Vespignani and Barrat(CR8)) and their co-workers. COSIN scientists are also majority of the second pink community (Camacho, Arenas, through which many people downloaded data and publications

Complexity in Networks

Home Page of the European Project COSIN
COevolution and Self-Organization In dynamical Networks

About COSIN
People
Publications
Software
Visualization
Network data

Complexity
What is Complexity?

Complex networks
What are Complex Networks?

Node 1 INFM	Node 2 Roma "La Sapienza"	Node 3 Universidad Barcelona	Node 4 EPFL Lausanne	Node 5 ENS Paris	Node 6 Universitat Karlsruhe	Node 7 Université de Paris Sud

"COMPLEX" NEWS

COSIN

Postdoctoral position in complex networks: We are looking to hire a person with expertise in the analysis and modeling of complex networks to work in the group of Prof. Hernan Makse at City College of New York. The candidate should have a significant publication record and a solid experience in programming in C C++ Fortran. For those interested, please send a CV to Hernan Makse (makse@mailaps.org).

The COSIN project is a FET OPEN Project
It is funded by EU commission in the priority area of:
Information Society Technologies

Information Society Technologies

The Cosin Project

The Cosin Project is a research project aiming to develop statistical models to describe Networks growths and evolution. These models will be based on agents interactions and inspired by the theory of self-organisation and fractal growth.

At the same time, we are thinking to collect data mainly for the Internet and the World Wide Web. These data will be collected in order to validate our models.
We also want to device visualization tools in order to analyze large data sets both from numerical simulations and real world data.
Applications to economic networks will be also considered.

The European Physical Journal B

Special Issue on COSIN Conference

Evolution and structure of the Internet

The most beautiful book on technological networks

Home | Software | People | Data | Related Links | Publications

Contact webmaster

Given all this credit it is naturally to think that research done by COSIN participants draws a rather high attention in the field. Nevertheless it is always difficult to predict the potential impact of basic research so that in this section we only list the most promising (in our view) applications that could be pursued by using the project results.

- **EPIDEMICS** (CR3 UB – CR8 UPSUD R. Pastor-Satorras, A. Vespignani) The work that has been done by people working in these two nodes is related to the effect of topology on the dynamics of the spread of infections. In particular the knowledge of scale-free nature of the underlying network of connection (as it is the case in social networks) is able to change completely the dynamics of infections. In particular no critical mass is required to have a pandemic (world-infection) behaviour. Those phenomena have a small but finte probability to happen regardless the number of infected people. An application that can be envised is that of a new policy against computer virus diffusion (given a finite amount of resources it is better to use them to immunize the hubs of the net).

- **UNIVERSALITY** (C01 INFM – CR3 UB - CR4-CR9 EPFL CR8 UPSUD) The quest for a common principle behind the formation of different scale-free networks it is made more complicated by the experimental evidence that many of these systems happen to have very different statistical properties when considering their internal correlation or they cycle structure. The research done by COSIN in this field on systems as different as the WWW, the Internet, financial systems and the Food Webs stands as the basic step in this studies. We have been able to classify these networks showing their similarities and difference. From that data we will proceed to define a unified model in which presumably different active ingredients as ``growth and preferential attachment'' or fitnesses or load minimisation interact with various strength in order to produce the results observed.
- **NON TECHNOLOGICAL NETWORK** (C01 INFM – CR5 ENS) While the acitivity in the field of Information technology is already at a good stage of development, the study of network in Biology (zoology and botany) as well as in economics and financial systems is at their beginning. Papers by COSIN project have been the first ones in these areas and they attracted a lot of interest in these communities, helping the development of this research in those fields.
- **DATA REPOSITORY** (Consortium) The necessity of having a reliable set of data for testing algorithms, models and visualization tools is a world-wide necessity. Different projects are currently running in Europe and elsewhere to create and maintain these resources. We are already witnessing papers by scientists outside the consortium using the data collected by COSIN. For that reason we expect to maintain and increase for a while this repository after the end of the project.
- **VISUALIZATION** (C01 INFM – CR7 UNIKARL - CR8 UPSUD) Some of the ideas born during the COSIN activity are starting just now as the site for Large Network visualisation Lanetvi hosted by the University of Paris Sud and The Indiana University. We expect those tools to have application and diffusion in running European project (DELIS).

Project's Achievements Fiche

Questions about project's outcomes	Number	Comments
1. Scientific and technological achievements of the project (and why are they so ?)		
Question 1.1. Which is the 'Breakthrough' or 'real' innovation achieved in the considered period	6	Brief description: • We defined some new active ingredients (the presence of disorder) able to explain the onset of scale-free behaviour. This means that scale-free graph can arise also without the rules of growth and preferential attachment • We are currently realizing algorithms based on the ideas of renormalization procedure in order to visualize large graphs as those of the Internet and the WWW. • We also started the characterization of communities presence in graphs with a variety of different approaches ranging from divisive methods to spectral analysis. • We have started the study and modeling of weighted networks; this is an important step forward since many networks display a strong heterogeneity of interactions. • We have a new analysis of the WWW displaying some unexpected statistical properties • We made a first step towards an agent based description of the Internet evolution
2. Impact on Science and Technology: Scientific Publications in scientific magazines		
Question 2.1. Scientific or technical publications on reviewed journals and conferences	149	Title and journals/conference and partners involved[1] See attached list of publications
Question 2.2. Scientific or technical publications on non-reviewed journals and conferences	6	Title and journals/conference and partners involved[2] See in the attached list of publications the preprint submitted to archive (i.e. cond-mat, physics ecc.)
Question 2.3.		

[1] Please submit these information in an 'excel' sheet with title of publication/authors/journal or conference/date etc.
[2] Please submit these information in an 'excel' sheet with title of publication/authors/journal or conference/date etc.

Invited papers published in scientific or technical journal or conference.	Yes	Title and journals/conference and partners involved[3] European Journal of Physics B special issue
3. Impact on Innovation and Micro-economy		
A – Patents		
Question 3.1. Patents filed and pending	0	When and in which country(ies): Brief explanation of the field covered by the patent:
Question 3.2. Patents awarded	0	When and in which country(ies): Brief explanation of the field covered by the patent* (if different from above):
Question 3.3. Patents sold	0	When and in which country(ies): Brief explanation of the field covered by the patent* (if different from above):
Questions about project's outcomes	**Number**	**Comments or suggestions for further investigation**
B - Start-ups		
Question 3.4. Creation of start-up	No	If YES, details: - date of creation: - company name - subject of activity: - location: - headcount: - turnover: - profitable : yes / no / when expected

[3] Please submit these information in an 'excel' sheet with title of publication/authors/journal or conference/date etc.

Question 3.5. Creation of new department of research (ie: organisational change)	No	Name of department:
C – Technology transfer of project's results		
Question 3.6. Collaboration/ partnership with a company ?	No	Which partner : Which company : What kind of collaboration ?
4. Other effects		
A - Participation to Conferences/Symposium/Workshops or other dissemination events		
Question 4.1. Active participation[4] to Conferences in EU Member states, Candidate countries and NAS. (specify if one partner or "collaborative" between partners)	Yes	Names/ Dates/ Subject area / Country: Conference on Complex Systems Torino Villa Gualino 5-7 December 2004 CISM School in Udine Italy June 2002 Participation in Sitges Conference June 2002. Midterm Conference of COSIN Rome Italy September 1-5 2003 Final conference of COSIN Salou Spain March 14-18 2005
Question 4.2. Active participation to Conferences outside the above countries (specify if one partner or "collaborative" between partners)	yes	Names/ Dates/ Subject area / Country: American Physical Society March Meeting Austin March 23 2003 USA American Physical Society March Meeting Los Angeles March 21-25 USA NIKKEI Econophysics Conference Tokyo November 2004
B – Training effect		
Question 4.3. Number of PhD students hired for project's completion	6	In what field : Theoretical Physics Computer Science

[4] 'Active Participation' in the means of organising a workshop / session / stand / exhibition directly related to the project (apart from events presented in section 2).

Questions about project's outcomes	**Number**	**Comments or suggestions for further investigation**
C - Public Visibility		
Question 4.4. Media appearances and general publications (articles, press releases, etc.)	Yes	References: Various publication mainly in Italian and French newspapers (Please attach relevant information)
Question 4.5. Web-pages created or other web-site links related to the project	1	References: http://www.cosin.org (Please attach relevant links)
Question 4.6. Video produced or other dissemination material	1	References: Book in press on World Scientific Press Some press release and videos
Question 4.7. Key pictures of results	No	References: (Please attach relevant material .jpeg or .gif)
D - Spill-over effects		
Question 4.8. Any spill-over to national programs	No	If YES, which national programme(s):
Question 4.9. Any spill-over to another part of EU IST Programme	No	If YES, which IST programme(s):
Question 4.10. Are other team(s) involved in the same type of research as the one in your project ?	Yes	If YES, which organisation(s): Notre Dame University, North Western University, Los Alamos National Laboratory, University of Cambridge, Indiana University, University of Porto

7. FUTURE OUTLOOK

As the project arrives to its natural end, there are some considerations that can be done, regarding the multidisciplinary community that has been shaped in the three years of activity. Collaboration started within the project will certainly remain at work, some nodes will keep collaboration under the support of the Commission both in running projects (i.e. DELIS) and in ones yet to come. Still it is a pity that the expertise of the consortium as a whole is fragmented after three years of activity (most of which has been devoted in the creation of a common language and a basic framework of theoretical tools). Under this respect COSIN is absolutely similar to other FET projects who probably face similar problems. Given the particular scientific situation at the moment (that is witnessing the first step of a creation of the new discipline of complexity), maybe one can think of a specific action that collects and maintains the basic results of the various projects in order to form a repository of the most interesting tools, papers, lectures etc.
One possibility could be to have a specific website where this material is organised and classified and links to the various project are preserved. Another way could be the organisation of regular meetings amongst the various participants to the projects. In any case some action to consolidate the results obtained is really necessary in order to understand the present situation of the field and answer to the question *"Where do we go from here?"* as required by the Commission.

As regards the scientific outlook we can have after the COSIN activity, we continue the analysis done in the methodology section regarding the data collection, the modelling and the formation of a general theory. It is very likely that in a few years the work on data collection will be over and the core of the activity in Technological Networks will split in two different directions. From one side there will probably be some specific modelling whose aim is to reproduce completely the behaviour of the systems. This very-refined models will be intended as instruments to monitor and check the activity of the various networks as for example Internet. On the other side there will be probably an activity of modelling that will try to put together different networks reproducing both their common features and their differences. This activity will be aimed at the investigation of the basic mechanisms that are behind network shaping in their evolution. As regards the more general issue of the formation of network structures in Nature, it is easy to predict that the application to social science and biology will be the future of this field. It is important to note that these studies have an immediate application in the field of Information Technology. From one side structures like WWW or blogs have an evident social structure that needs to be analysed in order to understand the systems. On the other side biology can help in devising more efficient ``bio-inspired'' algorithms and give inspiration to the apparent self-organised features of most of the technological networks under consideration.

Acknowledgements
At the end of this project there are some things that we still would like to say. Is the opinion of people involved in the project that European Commission has been a fantastic support, without which the work done would have been impossible, not only for the financial support, but also because the various proactive initiatives helped in the formation of a consortium such that we could reach the critical mass necessary to boost our activity. A special acknowledgement must be given to the different reviewers and advisors (Laszlo Barabasi and Mark Buchanan who work for the project without any support if not this acknowledgement). Through their suggestions we certainly improve not only the quality of the presentation (we admit it was rather poor at the beginning of the project) but also the scientific value of our research. As regards the partners, I would only say that has been a privilege to work with them and I am certainly looking forward to having again this possibility.