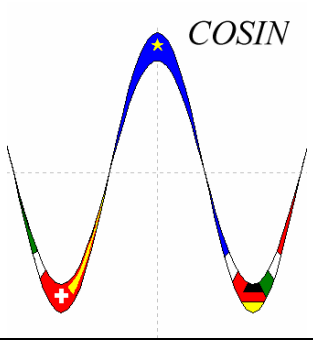


COSIN

IST–2001-33555

Coevolution and Self-Organization in Dynamical Networks

http://www.cosin.org

Periodic Progress Report N°:1

Covering period 1.3.2002-28.2.2003

Report Version: 2

Report Preparation Date: 14.03.2003

Classification:

Contract Start Date:1.3.2002 Duration: 36 months

Project Co-ordinator: Guido Caldarelli (gcalda@pil.phys.uniroma1.it)

Administrative Partners:

CO1 (INFM) INFM *Italy* (Ms. M. Arata),

CR2 (UDRLS) Università "La Sapienza" *Italy* (Prof. G. DI PILLO),

CR3 (UB) Universitat de Barcelona *Spain* (Dr. A. JUÁREZ),

CR4 (UNIL) Université de Lausanne *Switzerland* (Prof. W. WAHLI),

CR5 (ENS) Ecole Normal Superieure de Paris *France* (Dr. P. BRAND),

CR6 (UKON) Universitaet Konstanz *Germany* (Mr. H. HENGSTLER)

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

Table of Contents

1. Executive Summary

- **MAIN OBJECTIVES REACHED IN THE FIRST YEAR**

The main objective reached in this period is the onset of a interdisciplinary community to approach the challenges of modern Information Technology. This interdisciplinary approach is the best way to tackle the present problems in this area. Indeed, two of the most common communication networks, namely the Internet and the WWW are rapidly growing to unprecedented size for a human system. A statistical description of their properties represents therefore the only reliable choice to describe them. For their size, their self-similar properties and their irreversible dynamics, they attracted the interest of disciplines different from computer science. Due to such complex features, they are now a paradigmatic example for self-organized complex systems. COSIN projects was then set up in order to unify ideas, approaches technical analysis mainly from Computer Science, Statistical Physics and Graph Theory (see **Appendix A** for a list of milestones and expected achievements). We believe that this effort would not only produce the desired instruments to control the developing of the large communication networks but also constitute the core of the new science of complex systems. A crucial feature to obtain such result is to reach a critical mass of experts in different fields. Therefore not only a continuous feedback and exchange of information is needed within the project, but possibly the project itself should be conceived as the seed of this aggregation. The most immediate way to provide to everyone inside and outside the project the status the research has been realized through the web site of the project at **http://www.cosin.org**.

The importance of a reliable and efficient web site is twofold. For the people in the project it must represent the source of information, and the natural repository for the result of the research. For people interested in Information Technology it has present the current state of art as well as a common view to approach different problems. Even if the onset of such archive was a later deliverable, we rapidly realize the importance of such virtual home for the COSIN researchers. At the moment after 12 months we have been able to collect and make available the data in **Appendix B**

- **EXPECTED IMPACT**

The formation of a common framework of ideas, tools and scientists should help in realizing reliable models for Internet and WWW as well as helping in approaching problems of peer to peer connections and *ad hoc* networks. Due to the critical mass reached by the project we believe we can now trigger interest and activity in this field both on European and National basis.

- **PUBLIC VISIBILITY**

The project site has not been advertised yet, and in the present state of construction and shaping has an average visit rate of 15 persons per working day. The sections more visited are those related to the data and to the meeting announcements. We received requests from other scientists outside the project to have a link to them from the project page. To accomplish such requests we started the realization of a *who's who in complex network* html page

As for scientific production, more than 15 papers have been published or submitted with specific acknowledgement to the official COSIN number IST-2001-33555. The project is also acknowledged in the contribution presented by members in the proceedings of the Sitges Conference on Statistical Physics and VII conference on Statistical and Computational Physics (Granada). COSIN also organised a school in Udine devoted to large information networks. The intermediate results recovered in the project will be presented in a large midterm conference, whose poster is reported in **Appendix C**

Finally we intend to link the project activity to dissemination through an archive http://www.complexityreview.org

2. Work Progress overview

2.1 SPECIFIC OBJECTIVES MET IN THE FIRST YEAR

- Formation of a Common Background:

We tried to follow carefully the timetable indicated in the proposal of the project during our first year of activity. In particular we planned to start the project by collecting and merging different ideas and approaches from various field in order to approach the problem of analysis and possible control of large network structures. This process is now on its way, increasing collaborations between partners are presently approaching the same questions with a complete set of different frameworks from statistical physics to graph theory. In this period we realized the importance of a reliable web site. Therefore an extra effort has been paid in order to have this "deliverable" working well ahead of the due date.

- Formation of a Critical Mass:

Since the activity of the consortium we have been able to make COSIN as one of the most important organization working in the field. The activity of the consortium focused mainly in the realization of scientific papers and their presentation to conferences. This attracted interest also in people outside of the consortium. That is a crucial point, since increases dramatically the possibility of dissemination of results. Reaching this critical mass we now started directly in the dissemination activity. We are now monitoring the traffic on the website and most of the accesses outside consortium explore the site in order to find information on the midterm meeting and to download the data.

- Setting the standard

The COSIN consortium produced a large list of papers reported in the next section. In almost all of them the numerical analysis of the data goes beyond the simple analysis of the degree probability distribution. Network in their complexity are analyzed by means of several quantities like correlation functions between clustering and degree (from Statistical Physics), betweenness (Social Networks) and optimization of transport (Biology). Also the modelization explored other active ingredient rather than the Barabási's preferential attachment. Finally a standard for visualization large networks has been produced with the software visone (http://www.visone.de)

2.2 PROGRESS DURING THE FIRST YEAR

As reported in the project, the *main result we expected form this FET is the production of basic research* in the form of scientific publications. One of the goals of the project was indeed to form an interdisciplinary community that merging experience from computer science, graph theory and statistical physics could produce a unified framework to approach the problem of complexity in Information Technology. The result of this interaction should bring to common activity. A possible way to witness this activity and report the results is represented by joint publication between different sites members. It is worth noticing that the kind of publication we are referring here and in the following are not in the form of technical reports. We instead intend to present the COSIN progress on referred journals with preference to those of largest impact factor (i.e. Nature, Science, Proceedings of the National Academy of Sciences, Physical Review). A quantitative measure of the success of the consortium will then be related to the number of joint publications on these largest impact journals. In a next stage we want also to carefully monitor the number of citations per paper in order to measure the impact of COSIN activity on the various communities.

The main and most visible outcome of the scientific activity of the consortium is there presented in a series of publications (reported in section **5 "*Information Dissemination and exploitation of results*"**).

We will try to present the most striking results obtained in the many publication of the consortium. In particular, we select one publication per deliverable that we suggest to consider for report. For more information the whole set of papers are downloadable at http://www.cosin.org/Publications.html

In order to describe to what extent the success has been reached in the deliverables we give a brief description of the various deliverables for the first year.

- **D01, D02, D03**

  They are mainly related to the management of the project and are respectively a Powerpoint presentation of the project, a "dissemination and use plan" together with the set up of the advisory board. These points have already been contracted with a positive feedback by the commission.

  We nevertheless give a brief description for all of them.

  - D01 is a 3 pages Powerpoint presentation with the figures of the project and the participant to the consortium. A hardcopy of this presentation has been sent together with this report to the reviewers.
  - D02 is a dissemination and use plan where it is reported the actions to take in order to disseminate the results. A hardcopy of that has been sent together with this report to the reviewers.
  - D03 is simply the setup of an advisory board. We appointed for such role Prof. A.L. Barabasi and Dr. M. Buchanan

- **D04 Universality in Networks (WP 1).**

  This deliverable is mainly intended in finding quantities that remain constant in various experimental network as well as new topological quantities to describe graphs. This deliverable is particularly important for the project. Indeed it addresses the point of the common nature of different phenomena in various fields of research.

  As a necessary step in any scientific research we wanted to start by carefully analyzing the data sets available for Internet WWW and the other systems. Therefore a continuous collaboration and feedback has been necessary between this Work package and the WP 2 specifically devoted to the data collection. We extended the current state of art in literature where universality is claimed by a qualitative analysis of the degree distribution. Thanks to the concepts developed in WP 3 about the representation of social networks we classified the different systems analyzed through their clustering, centrality, connectivity and other topological measures. We are able to conclude that the qualitative universal behaviour observed in the degree distribution derives from a series of rules more general than the growth and preferential attachment that have been invoked by A.L. Barabasi et al.

The research of the minimum set of topological variables in order to classify different systems is a very promising field of research. We are going to devote to this task some of the future activity in this Work Package.

Amongst the various publications on the subject it has to be noticed that two of them independently arrived to the same conclusion. A possible measure of clustering can be related to the number of quadrilaterals in the structure. We have studied the properties of metrics aimed at the characterization of grid-like ordering in complex networks. These metrics are based on the global and local behaviour of cycles of order four which are the minimal structures able to identify rectangular clustering. The analysis of data from real networks reveals the ubiquitous presence of a high level of grid-like ordering that is non-trivially correlated with the local degree properties. These observations provide new insights on the hierarchical structure of complex networks.

Also, the problem of the reliability of present data, discussed against the possible influence of the collection method on the data themselves, has been discussed ("Exploration bias of Complex Networks", P. De Los Rios, *Proceedings of the 7th Conference on Statistical and Computational Physics Granada* (2002)).

These results have been presented in public conferences at the *7th Conference on Statistical and Computational Physics Granada* (September 2002) and at the Stadybis INFM Meeting in Florence (February 2003).

Another important step in the understanding of the onset of scale-free behaviour is represented by the idea that the fractal behaviour could not be necessarily related to growth and preferential attachment (see above). We want to point out that all the above publications already cited received a fairly large number of citations. In a series of papers (some of them published, some others in publication or submitted) this collaboration has proposed various models to explain most statistical properties of the Internet and of other scale-free networks such as protein networks.

We selected this paper (where three different nodes are present), as representative of the activity of the Work package in this deliverable.

- *Scale-Free Networks from Varying Vertex Intrinsic Fitness*
  G. Caldarelli, A. Capocci, P. De Los Rios and M. A. Muñoz ,
  Physical Review Letters **89**, 258702 (2002),

- **D05 Preliminary analysis of collected data (WP 2).**

THE DATA

In order to have the largest possible impact in the field we realized that we needed to collect very good sets of data to start with. We also realized that this issue was closely connected to the issue to make available all the data in real time to the member of consortium. We therefore started by realizing a friendly interface to download the data available from the project site at www.cosin.org.

This topic in our idea is the most important, since the complex behaviour of Information Networks has been revealed by unexpected statistical features of Internet and WWW.

The Home Page is a simple introduction to the network's structure, and important meeting will be advertised there. On the left side, a button panel offers the possibility to explore the details of the structure of COSIN (Objectives, Participants, Data and Codes, Publications, both cumulative and classed by deliverable, Visualization and Press about COSIN members). So the COSIN website already works as a "community home" for COSIN members, where they can find most relevant material for their research and presentations. In particular, with Thomas Erlebach (ETH, Zurich), the Data sub-page has been structured in such a way that downloading data is as easy as possible. At present, the contained data are

- Internet at AS scale (14 samples from 3000 to 10500 sites)
- Traceroute analysis generated from the COSIN server
- Protein map interactions for 5 different organisms
- Food webs (source from scientific publications)
- Board of Directors for Italian Stock Exchange, NYSE, Nasdaq
- Topholders for Italian Stock Exchange, NYSE, Nasdaq
- Actor movie dataset

- Taxonomic trees from different ecosystems

As the structure is now in place, more data will be added over time (the exploration of the Internet from a Swiss node being one of the first in the to-do list: contacts with SWITCH, the official manager of the Internet in Switzerland, have been made).

THE ANALYSIS

In the site we then collected form various sources 13 maps of INTERNET at the AS level starting from 3000 to 10.500 nodes. We also analyzed the largest WWW crawl available at the moment and we also collected a series of other social and biological networks (US patents, Protein Interaction Maps, Food Webs). An extended analysis of the Internet data sets is presented in the publications of A. Vespignani, T. Erlebach and R. Pastor-Satorras. Pastor-Satorras and Vespignani will be shortly publishing a book with some of the activity made in the COSIN project. As regards the analysis on the biological networks one paper on Food webs has been accepted on Nature and cannot be present on the site (at the moment) for copyright reasons, another one on Protein Interaction Maps is in preparation.

Since the universality of statistical behaviour with respect to different systems, we did not restrict ourselves to the analysis of Internet and WWW. Nevertheless a special emphasis has been paid in order to describe these two networks.

We selected this paper (where three different nodes are present), as representative of the activity of the Work package in this deliverable.

- *Large-scale Topologicaland Dynamical Properties of the Internet*
  A. Vazquez, R. Pastor Satorras and A. Vespignani,
  Physical Review E **65**, 066130 (2002),

- **D06 Algorithms for Network Centrality (WP 3).**

A first deliverable is a case study of visualization tools for large graphs. Under http://www.inf.uni-konstanz.de/algo/cosin/ mirrored by the project site we provide a collection and description of tools for visualization of graphs and networks and an evaluation of some tools for visualization of very large graph. Starting from this list of available software we devoted most of the activity in this field on the realization of the new visualization software **visone.**

Visone is a software for visual social network analysis. A social networks consists most of the times of persons linked by some kind of relations (sometime even the presence of link could be matter of debate between the two vertices). Visone consists of a graph editor tailored to social networks with specialized analysis and visualization components.

The software visone is a tool that facilitates the visual exploration of social networks. Social network analysis is a methodological approach in the social sciences using graph-theoretic concepts to describe, understand and explain social structure. The visone software is an attempt to integrate analysis and visualization of social networks and is intended to be used in research and teaching. While we are primarily focussing on users in the social sciences, several features provided in the tool will be useful in other fields as well.

In contrast to more conventional mathematical software in the social sciences that aim at providing a comprehensive suite of analytical options, our emphasis is on complementing every option we provide with tailored means of graphical interaction. We attempt to make complicated types of analysis and data handling transparent, intuitive, and more readily accessible. User feedback indicates that many who usually regard data exploration and analysis complicated and unnerving enjoy the playful nature of visual interaction. Consequently, much of the tool is about graph drawing methods specifically adapted to facilitate visual data exploration.

The origins of visone lie in an interdisciplinary cooperation with researchers from political science which resulted in innovative uses of graph drawing methods for social network visualization, and prototypical implementations thereof. In a nutshell, visone is a

- tool for interactive analysis and visualization of networks, in which
- originality is preferred over comprehensiveness, and that
- caters especially to social scientists.

We developed an algorithm for radial graph layouts to support centrality visualizations. This method is available in visone. The actors are assigned to circuits according to their centrality value. In a three phase method, formal and informal communication is incorporated. Using colors, different shapes and sizes for the graphical variables, semantic attributes are included. Algorithmic challenges are, e.g. avoidance of overlaps and minimization of crossings. The visone software is provided as a standalone executable for systems running Linux, Solaris, or Windows, and is free for academic purposes. See http://www.visone.de. So far visone does not provide algorithms for visualizing large and very large networks. Therefore we pursue as a long term goal the design of algorithms that support visualization of very large graph based on clustering. As a first step towards our long term goal we are developing an efficient clustering algorithm based on the natural notion of graph clustering as separation of sparsely connected dense sub-graphs from each other. Several formalizations have been proposed. However, the understanding of current algorithms and indices is still rather intuitive. As a first step towards understanding the consequences of particular conceptions, we concentrate on indices that focus on the relation between the number of intra-cluster and inter-cluster edges. Investigations to use our algorithm for clustering autonomous system graphs are promising.

Since this deliverable comes out essentially in the form of a software, we enclose the necessary specification and standards in the following publication:

- *Visone – Analysis and Visualization f Social Networks*
  U. Brandes, and D. Wagner ,
  To be published on Graph Drawing Software Springer Series "Mathematics and Visualization" M. Junger, P. Mutzel (eds).

- **D07 Centrality and Groups in Social Networks (WP 4).**

We study some particular cases of social networks that can be related to the economical and financial systems. Namely the system investigated is the stock exchange with respect to the various Board of directors and with respect to the portfolio composition of the different companies.

We propose a measure of the impact of the interlock on the decision making, which could help identifying boards susceptible to be influenced by a minority. We apply our models and methodology to the data of the boards of the largest US corporations in 1999. We also collected (and we are going to publish on the site) information related to the Board of Directors and Top-holders of the companies quoted in US and European Stock Exchanges. The report on the statistical properties of such graph is in preparation.

For the shareholders network we have studied the topological properties of the graph in relationship to the ability of firms to control other firms. A financial system forms an intricate network of ownership relationships between owners and controlled companies. On a large scale, = such a structure can occasionally experience dramatic cooperative phenomena which follow some empirically well-established scaling laws. It is therefore very important to understand if any underlying structure exists in order to predict when such organized phenomena take place. =
Despite multiple efforts in describing criteria for the robustness of a financial systems and the best portfolio strategy, poor attention have been devoted to the topology of such a network.

The set of companies quoted on a stock market togheter with their respective topholders form the Shareholding Network (SN). If we consider only those topholders that are as well companies quoted on the same market, we obtained a subnetwork, the Stock Shareholding Network (SSN).

We have performed a systematic study of the topological properties of the SN and the SSN for two US stockmarket (NYSE and NASDAQ) and one european stockmarket (Milan). Besides the fact that these kind of networks have never, to our knowledge, been studied before, we pay here particular attention to the weights of the links in the graph, when measuring the topological properties related to centrality and connectivity degree. While some network properties seems to be common to different market, others are dramatically different and may be used to =
classify financial systems. These financial agent networks exihibit statistical properties typical of some self-organized ecological networks shaped by natural evolution based on robustness criteria.

We study the distribution of the basin of companies controlled by a given company and we propose a measure of the goodness of a portfolio strategy based on the topological relationships between stocks in the portfolio.

These results are reported in an article that is about to be submitted for publication and will be presented at the conference WEHIA 2003 in Kiel, Germany.

For the Board Network, in the framework of decision making dynamics, we have investigated the role of special subgroups of directors in driving global decisions. Members of boards of directors of large corporations are connected in networks through their multiple appointments on different boards. Members of a board who also serve together on an outside board, form a special subgraph, that we call interlock graph of the board. We investigate the extent to which a minority of well connected directors can drive the decision of the whole board. We present a first model of decision making dynamics inspired to herd behavior and a second one more realistic in the context of boards. We study how the size and the topology of the interlock graph affect the probability that the board approves a strategy proposed by the Chief Executive Officer.

The work we selected to represent the deliverable is the following

- *Decision making dynamics in corporate boards*
  S.Battiston, E.Bonabeau, G.Weisbuch
  to be published on *Physica A*.

- **D08 Modelling WWW (WP 5).**

As regards the WWW we analyzed the largest available Web crawl made of $3 \cdot 10^8$ html documents. The results we find are rather different from those available in the literature that refer to a very small and bias set of $3 \cdot 10^5$ documents. We are currently preparing the final version of this report and we are anonymizing this data set in order to make it available on the web site of the project. Analysis of this data set (made in collaboration with WP 2) represent a non trivial task since the large size of the graph. Nevertheless thanks to this analysis we have been able to define a suitable model to reproduce the system growth and behaviour.

Amongst the main results of this analysis we want to present the following:

- Kumar et. al. and Barabasi and Albert, suggested that both the indegree and outdegree distribution of the nodes of the WebGraph follow a power-law distribution. Experiments on larger scalemade by Broder et. al. [9] confirmed it as a basic web property. The probability that the indegree of a vertex is distributed by a power-law, whose exponent is –2.1. The outdegree of a vertex is also distributed with a power law with exponent roughly equal to -2.7. The average number of edges per vertex is about 7.
  We find instead the distribution of out-degree is not a power-law.

- The recent study of Dill et al. [12] gives a picture of the web explaining its fractal structure as produced by the presence in the web of multiple regions generated by independent stochastic processes. The different regions being different in size and aggregation criteria, for instance topic, geography or Internet domain. All these regions are connected together by a "connectivity backbone" formed by pages that are part of multiple regions. In fact, all previous models present the Web as a flat organism, every page may potentially connect to every other page of the Web. This is indeed far from reality.
  We propose a "Multi-Layer" model in which every new page that enters the graph is assigned with a constant number of regions it belongs to and it is allowed to link only to vertices in the same region.

- We compare the results of our model, with respect to other models, by considering the following quantities
  - In- and out- degree distribution
  - Distribution of disjoint bipartite cliques
  - Number of vertices within a certain distance d from a sample vertex
  - Clustering coefficient per vertex

We find that with respect to other WWW models only the multilayer reproduces correctly all the above quantities.

This deliverable is composed by this joint publication

- *A Multi-layer model for the Web Graph*
  L. Laura, S. Leopardi, G. Caldarelli and P. De Los Rios,
  Contribution to 2nd Web Dynamics Workshop (2002).

DELIVERABLES TABLE

Project Number:	**IST-2001-33555**
Project Acronym:	**COSIN**
Title:	**Coevolution and Self-Organization in Dynamical Networks**

Del. No.	Revision	Title	Type[1]	Classification[2]	Due Date	Issue Date
1		Project Presentation	R	Pub	3	3
2		Dissemination and Use Plan	R	Pub	6	9
3		Setup of the Advisory Board	R	Pub	6	3
4		Universality in Networks	R	Pub	12	12
5		Preliminary analysis of the collected data	R	Pub	12	12
6		Algorithms for Network Centrality	R	Pub	12	12
7		Centrality and Groups in Social Networks	R	Pub	12	12
8		Modelling WWW	R	Pub	12	12
9		First Progress Report	R	Pub	12	12

[1] *R: Report; D: Demonstrator; S: Software; W: Workshop; O: Other – Specify in footnote*

[2] *Int.: Internal circulation within project (and Commission Project Officer + reviewers if requested)*
Rest.: Restricted circulation list (specify in footnote) and Commission SO + reviewers only
IST: Circulation within IST Programme participants
FP5: Circulation within Framework Programme participants
Pub.: Public docu

DELIVERABLE SUMMARY SHEET

Project Number:	IST-2001-33555
Project Acronym:	COSIN
Title:	Coevolution and Self-Organization in Dynamical Networks

Deliverable N°: 1 **Project Presentation**

Due date: 3 Month

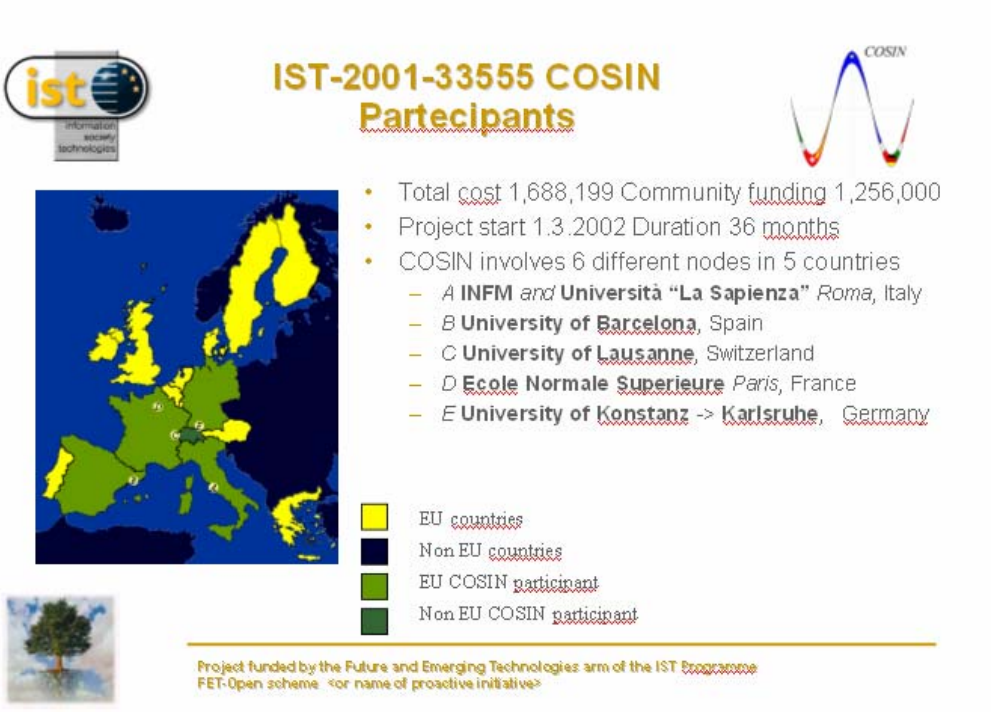
Delivery Date: 3 Month

Short Description:

Power Point presentation of Project. It is realized with the template enclosed in FET guidelines. It presents the consortium and the objectives of the projects as well as the relevant information about funding and time.

Downloadable at

http://www.cosin.org/deliverables/D01.ppt



Partners owning: **CO1 (INFM)**

Partners contributed:

Made available to: public

DELIVERABLE SUMMARY SHEET

Project Number:	IST-2001-33555
Project Acronym:	COSIN
Title:	Coevolution and Self-Organization in Dynamical Networks

Deliverable N°: 2 and 3 **Dissemination and Use Plan (it contains Advisory Board)**

Due date: 6 Month

Delivery Date: 9 Month

Short Description:

As suggested in the FET guidelines it is a document to present the expected impact of project objectives as well as the description of the dissemination initiatives.

The final form of this deliverable was presented in slight delay after the submission of a preliminary version. Substantial corrections were needed in order to include the actions promoted by the VI Framework Programme. We devoted one section to present the advisory board that is composed by

A.-L. Barabasi and

M. Buchanan

Downloadable at

http://www.cosin.org/deliverables/D02.doc

Partners owning: **CO1 (INFM)**

Partners contributed:

Made available to: public

DELIVERABLE SUMMARY SHEET

Project Number:	IST-2001-33555
Project Acronym:	COSIN
Title:	Coevolution and Self-Organization in Dynamical Networks

Deliverable N°: 4 **Universality in Networks**

Due date: 12 Month

Delivery Date: 12 Month

Short Description:

In works done by three of the groups in the consortium we have analyzed the hierarchical structure of complex networks by looking at the formation of cycles of different orders.

In publications below we have studied the properties of metrics aimed at the characterization of grid-like ordering in complex networks. These metrics are based on the global and local behaviour of cycles of order four which are the minimal structures able to identify rectangular clustering. The analysis of data from real networks reveals the ubiquitous presence of a high level of grid-like ordering that is non-trivially correlated with the local degree properties. These observations provide new insights on the hierarchical structure of complex networks. In a separate publication by the Swiss group the number of scaling of the number of loops as a function of the system size is investigated in particular analytical results are obtained for the Barabási-Albert scale-free model.

Finally in another publication we relate the formation of social capital with the number of squares formed in the network. We study in detail the community of users of the PGP encryption system and show that this community creates social capital very effectively.

The deliverable is formed by the publications

2-3-5-7-11-15-16-18-20-21-25-26-27-38 of the list in Section 5

All documents are downloadable at

http://www.cosin.org/deliverables/D04

Partners owning: **CR3 (UB)**

Partners contributed: CO1, C04

Made available to: public

DELIVERABLE SUMMARY SHEET

Project Number:	IST-2001-33555
Project Acronym:	COSIN
Title:	Coevolution and Self-Organization in Dynamical Networks

Deliverable N°: 5 **Preliminary analysis of the collected data**

Due date: 12 Month

Delivery Date: 12 Month

Short Description:

We tried to collect and make publicly available on the web site the most reliable datasets for the Information Networks. At the moment we have 13 maps of Internet at the AS level and the largest Web crawl ever realized. On top of that we also collected and presented some examples of social networks like the actor movie dataset and the dataset of 13 years of citations between US patents. From the point of view of the biological networks we present the Protein Interaction Maps and the Food webs available in the literature. The last data set has been the subject of a paper accepted on Nature on the statistical properties of Food Webs. For the protein network we also found very interesting results. The status of this latter piece of research is a preprint since the delayed starting date of the Swiss node.

An extensive analysis on these data sets is present in a series of publications on the most important international journals. Pastor-Satorras and Vespignani will present most of their analysis on the Internet case in their next book to be published by Cambridge University Press.

This deliverable consists of the following publications

8-14-19-23-37-41 of the enclosed list

Downloadable at

http://www.cosin.org/deliverables/D05

Partners owning: **CR4 (UNIL)**

Partners contributed:CO1, CR3

Made available to: public

DELIVERABLE SUMMARY SHEET

Project Number:	IST-2001-33555
Project Acronym:	COSIN
Title:	Coevolution and Self-Organization in Dynamical Networks

Deliverable N°: 6 **Algorithms for Network Centrality**

Due date: 12 Month

Delivery Date: 12 Month

Short Description:

Firstly it has been developed a new algorithm for representing centrality in social networks.
U. Brandes, P. Kenis, and D. Wagner: **Communicating Centrality in Policy Network Drawings**.
IEEE Transactions on Visualization and Computer Graphics. *IEEE Press*, to appear.

Then a method for representing ranking of link structures was designed and evaluated.
U. Brandes and S. Cornelsen: **Visual Ranking of Link Structures.**
Journal of Graph Algorithms and Applications, to appear.

Furthermore the software package "visone" has been developed (see http://www.visone.de/).
Visone is particularly indicated to represent the relevant network indices in social networks
U. Brandes and D. Wagner **visone - Analysis and Visualization of Social Networks**
to appear in special issue on *Graph Drawing Software, Springer series "Mathematics and Visualization*", Michael Juenger and Petra Mutzel (eds.)

Finally we developed an efficient clustering algorithm.
This is in order to the design of algorithms and software to represent groups in very large graphs.
In an experimental study it is compared to other graph clustering methods.
U. Brandes, M. Gaertler and D. Wagner **On Clustering large graphs**
in preparation

A series of other papers on visualization tools applied to the Internet case has been produced by G. Di Battista and coworkers.

This deliverable is formed by the software tool visone (http://www.visone.de) and by the following publications:

9-12-13-14-29-30-31-32-34-35-36

Downloadable at

http://www.cosin.org/deliverables/D06

Partners owning: **CO6 (UKON)**

Partners contributed:CR2

Made available to: public

DELIVERABLE SUMMARY SHEET

Project Number:	IST-2001-33555
Project Acronym:	COSIN
Title:	Coevolution and Self-Organization in Dynamical Networks

Deliverable N°: 7 **Centrality and Groups in Social Networks**

Due date: 12 Month

Delivery Date: 12 Month

Short Description:

To address the issue of this deliverable, we have focussed on two exempla of social networks: the network of directors of large corporations' boards (Board Network) and the network of shareholders in a stock market (Shareholder Network). In both cases we deal with real data collected by the project. Results for the first work have been published in one article. Results for the second work are about to be submitted for publication. In both cases articles are downloadable from COSIN website for prompt dissemination.

For the Board Network, in the framework of decision making dynamics, we investigated the role of special subgroups of directors in driving global decisions. Members of boards of directors of large corporations are connected in networks through their multiple appointments on different boards. Members of a board who also serve together on an outside board, form a special subgraph, that we call interlock graph of the board.. We present a first model of decision making dynamics inspired to herd behavior and a second one more realistic in the context of boards. We study how the size and the topology of the interlock graph affect the probability that the board approves a strategy proposed by the Chief Executive Officer. We propose a measure of the impact of the interlock on the decision making, which could help identifying boards susceptible to be influenced by a minority. We apply our models and methodology to the data of the boards of the largest US corporations in 1999. This work is reported in **Decision making dynamics in corporate boards** by S.Battiston, E.Bonabeau, G.Weisbuch to be published on *Physica A*.

For the shareholders network we have studied the topological properties of the graph in relationship to the ability of firms to control other firms. A financial system forms an intricate network of ownership relationships between owners and controlled companies. On a large scale, such a structure can occasionally experience dramatic cooperative phenomena which follow some empirically well-established scaling laws. It is therefore very important to understand if any underlying structure exists in order to predict when such organized phenomena take place. Despite multiple efforts in describing criteria for the robustness of a financial systems and the best portfolio strategy, poor attention have been devoted to the topology of such a network.

This deliverable is formed by the following publications

1-4-10-28

Downloadable at

http://www.cosin.org/deliverables/D07

Partners owning: **CR5 (ENS)**

Partners contributed:CO1

Made available to: public

DELIVERABLE SUMMARY SHEET

Project Number:	IST-2001-33555
Project Acronym:	COSIN
Title:	Coevolution and Self-Organization in Dynamical Networks

Deliverable N°: 8 **Modelling WWW**

Due date: 12 Month

Delivery Date: 12 Month

Short Description:

In the first publication we proposed a "Multi-Layer" model of the Webgraph in which every new page that enters the graph is assigned with a constant number of regions it belongs to and it is allowed to link only to vertices in the same region. When deciding the end-points of the edges we adopt a combination of Copying and Evolving Network in the sub graph of the specific region. In particular, if an edge is not copied from the prototype vertex, its end-point is chosen with probability proportional to the in-degree in the existing graph. The final outcome of the stochastic process is the graph obtained by merging the edges inserted between vertices of all layers. This model shows surprising stability properties with respect to the variability of the parameters.

In the second of the publications reported below we presented the first extensive comparison study of the statistical properties of large scale simulations of the most important models for the Webgraph presented so far in literature. These models are compared on several relevant observables such as in-degree, out-degree, PageRank, correlation between different measures, e.g. in-degree and Pagerank, number of small bipartite cliques. These models are also compared with a large sample of 300Ml documents and about 1 Billion edges crawled by Alexa in 2001. The analysis of this new sample has allowed discovering some new properties of the Webgraph. For instance, the out-degree does not follow a power law distribution and we can observe a much larger number of cyber-communities with respect to a 3 years older sample always collected by Alexa. This comparison study has required the development of external and semi-external algorithms for computing several graph properties and large scale simulation of stochastic graph models.

This deliverable consists of the following publications

33-40

Downloadable at

http://www.cosin.org/deliverables/D08

Partners owning: **CR2 (UDRLS)**

Partners contributed:CO1,CR3

Made available to: public

DELIVERABLE SUMMARY SHEET

Project Number:	IST-2001-33555
Project Acronym:	COSIN
Title:	Coevolution and Self-Organization in Dynamical Networks

Deliverable N°: 9 **First year report**

Due date: 12 Month

Delivery Date: 12 Month

Short Description:

This document

Downloadable at

http://www.cosin.org/deliverables/D09.doc

Partners owning: **CO1 (INFM)**

Partners contributed:CR2,CR3,CR4,CR5,CR6

Made available to: public

2.3 COMPARISON WITH PLANNED ACTIVITIES

We planed for the first year to establish a solid collaboration between the different members of the various nodes in the consortium. At the moment we have a fairly large number of publications and scientific activity amongst the various groups. As planned we expect to increase the joint research in the incoming year. As regards the training of young researchers, many nodes experienced a slight delay (from 1 month to about 3-4 months) in finding good candidates for post-doc positions. We are not worried by this deviation from the plan. Indeed it is quite problematic to start directly on the starting date of the project. In one case (node CO1) one postdoc had to retire for illness, and that also comported a delay in the use of resource planned. In all these cases we supplied temporarily with resources not financed by the project in order to avoid delay in the deliverables.

2.4 ACTIVITY IN THE FIRST YEAR

Progress Overview Sheet (Partner CO1 INFM)

PROGRESS OVERVIEW SHEET[1]

Organisation: Istituto Nazionale Fisica Per la Materia

Workpackage/ Task	Planned effort[2]	Planned Date[3]		Actual Date[4]		Resources employed[2]	Cumulative Resources[2]
	Whole Project	Start	End	Start	End	This Period	Since start
WP 1	18	1	36	1	36	2	2
WP 2	6	1	36	1	36	2	2
WP 3	0	1	36	1	36		
WP 4	20	1	36	1	36	4	4
WP 5	12	1	36	1	36	1	1
WP 6	10	1	36	1	36	3	3
WP 7	14	1	36	1	36	2	2
WP 8	3	1	36	1	36	1	1
Total	83					19	19
One person month is equal to		170 [5]		Person hours			

Main contribution during this period	
Workpackage/Task	**Action**
WP 1	***Mathematical Tools for Complex Systems***
Task 1.2	In collaboration with CR4 and the University of Fribourg and the University of Granada, we have proposed a new mechanism in order to form complex networks. This is particularly suitable for the Internet Network. We applied such mechanism in order to analyze data we collected on Protein Interactions Maps and WWW. This resulted in two papers that enter in the deliverable D05 one devoted to WWW that has been submitted to Physical Review E, and another on Protein Interaction that is in the form of a draft.
WP 2	***Data Collection and Analysis***
Task 2.1	With the Collaboration of CR4 and Thomas Erlebach (ETH Zurich) we set up a database of data and software tools at **www.cosin.org .** The data collected until now are about Internet, Social and Biological Systems and Protein Interactions
Task 2.3	See Task 2.1
WP 4	***Dynamics of Social Networks***
Task 4.1	Study of statistical properties of shareholders network and study of correlation in financial data (Pub. 4,42)
WP 5	***Model for Communication Networks***
Task 5.1	See Task 1.2
WP 6	***Dissemination of the Results***
Task 6.1-6.2	See following list of publication and talks
WP 7-8	***Management-Assessment and Evaluation***
	See section 5 for conference and workshops. See attached advisors comments.

[1] Each partner should fill in its own Progress Overview Sheet for a period in question. The Project Co-ordinator will check and approve the forms and attach them to the corresponding PPR.
[2] In person months (or in person hours)
[3] Project month when the activity was planned to be started or to be completed
[4] Project month when the activity was actually started or completed
[5] Give a figure used for converting person hours to a person month

Deliverables due this period		
Deliverable number	**Title of Deliverable**	**Status (Draft Final, Pending)**
D01-D02-D03-D09	Presentation-Dissemination and Use Plan –Setup of advisory board-first year report	Final

Dissemination actions (articles, workshops, conferences etc.)

Publications:

5 G. Caldarelli *Proceeding of the VII conference in Granada (to appear)*
6 G. Caldarelli, A. Capocci, P. De Los Rios and M.A.Muñoz *Physical Review Letters* **89** 258702 (2002)
7 G. Caldarelli, R. Pastor-Satorras and A. Vespignani. preprint submitted to *Physical Review Letters*
8 A. Capocci, G. Caldarelli and P. De Los Rios preprint submitted to *Physical Review E*
37 D. Garlaschelli**,** G. Caldarelli, and L. Pietronero to be published on Nature.
40 G. Caldarelli, P. De Los Rios S. Leonardi, S. Millozzi and A. Vespignani preprint
41 G. Caldarelli, P. De Los Rios and F. Squartini preprint
42 S. Battiston, G. Caldarelli, M. Castri, D. Garlaschelli and L. Pietronero preprint

Talks:

XVIII Sitges Conference on Statistical Mechanics June 2002
VII Conference on Statistical and Computational Physics Granada 2-7/09/2002

Deviations from the planned work schedule/reasons/corrective actions/special attention required

The postdoc rented for the first year had to retire due to illness problems. We used resources not financed by Commission to end related projects.

Planned actions for the next period

A Postdoc has been hired in July 2002. Starting Fall 2003, we are planning to hire two other postdoc

Progress Overview Sheet (Partner CR2 UDRLS)

PROGRESS OVERVIEW SHEET[1]

Organisation: Università di Roma "La Sapienza"

Workpackage/ Task	Planned effort[2]	Planned Date[3]		Actual Date[4]		Resources employed[2]	Cumulative Resources[2]
	Whole Project	Start	End	Start	End	This Period	Since start
WP 1	2	1	36	1	36		
WP 2	6	1	36	1	36		
WP 3	13	1	36	1	36		
WP 4	0	1	36	1	36		
WP 5	36	1	36	1	36		
WP 6	2	1	36	1	36		
WP 7	1	1	36	1	36		
WP 8	1	1	36	1	36		
Total	61						
One person month is equal to		170 [5]		Person hours			

Main contribution during this period	
Workpackage/Task	**Action**
WP 3	***Large Networks Visualization Tools***
Task 3.1	Realization of algorithms
WP 5	***Models for Communication Networks***
Task 5.1	Modelization of WWW
WP 6	See Dissemination Actions Below.

Deliverables due this period		
Deliverable number	**Title of Deliverable**	**Status (Draft Final, Pending)**
D08	Modelling and Visualizing WWW	Draft

Dissemination actions (articles, workshops, conferences etc.)

Publications:

9 A. Carmignani, G Di Battista, W. Didimo, F. Matera and M. Pizzonia Journal of Graph Algorithms(2002).
12 G. Di Battista, W. Didimo, M. Patrignani and M. Pizzonia Software Practice and Experience
13 G. Di Battista, F. Mariani, M.Patrignani and M. Pizzonia preprint
14 G. Di Battista, M. Patrignani and M. Pizzonia preprint
17 L. Laura, S. Leonardi, G. Caldarelli and P. De Los Rios *Contribution to 2nd Web Dynamics Workshop* (2002).
40 G. Caldarelli, P. De Los Rios, S. Leonardi, S. Millozzi and A. Vespignani preprint

Talks:

2nd Web Dynamics Workshop

Deviations from the planned work schedule/reasons/corrective actions/special attention required

Planned actions for the next period

[1] Each partner should fill in its own Progress Overview Sheet for a period in question. The Project Co-ordinator will check and approve the forms and attach them to the corresponding PPR.
[2] In person months (or in person hours)
[3] Project month when the activity was planned to be started or to be completed
[4] Project month when the activity was actually started or completed
[5] Give a figure used for converting person hours to a person month

Progress Overview Sheet (Partner CR3 UB)

PROGRESS OVERVIEW SHEET[1]

Organisation: Universidad de Barcelona

Workpackage/ Task	Planned effort[2]	Planned Date[3]		Actual Date[4]		Resources employed[2]	Cumulative Resources[2]
	Whole Project	Start	End	Start	End	This Period	Since start
WP 1	6	1	36	1	36		
WP 2	36	1	36	1	36	8	8
WP 3	4	1	36	1	36		
WP 4	4	1	36	1	36		
WP 5	4	1	36	1	36		
WP 6	2	1	36	1	36		
WP 7	1	1	36	1	36		
WP 8	1	1	36	1	36		
Total	58						
One person month is equal to		170 [5]		Person hours			

Main contribution during this period	
Workpackage/Task	**Action**
WP 1	***Mathematical Tools for Complex networks***
Task 1.1	In works done by three of the groups in the consortium we have analysed the hierarchical structure of complex networks by looking at the formation of cycles of different orders.
Task 1.2	There is some effort devoted to the study of self-organization in networks. In particular we want to analyze the effect of local rules on the global behavior of the system and under which circumstances the evolution of the system is critical in the sense that there are no characteristic time or length scales. In particular we are currently applying rules from a model of diffusion of innovations where agents can change their neighborhood of connections in order to achieve a better local performance. The resulting global performance is then under discussion.
Task 1.3	We have studied in detail the role of network structure in the optimization of search algorithms. We find analytical expressions for the load of the network in terms of the topological properties of the graph and of the search algorithm. This procedure enables to find optimal structures based on purely static properties instead of the dynamic properties. Currently we are investigating the effect of the different protocols on the optimization of the network. We are assuming that the networks structure is given but the emphasis now is on the protocols to deliver the packets from node to node.

Deliverables due this period		
Deliverable number	**Title of Deliverable**	**Status (Draft Final, Pending)**
D04	Universality in Networks	Final

Dissemination actions (articles, workshops, conferences etc.)
Publications: **3** M. Boguñá and R. Pastor-Satorras

[1] Each partner should fill in its own Progress Overview Sheet for a period in question. The Project Co-ordinator will check and approve the forms and attach them to the corresponding PPR.
[2] In person months (or in person hours)
[3] Project month when the activity was planned to be started or to be completed
[4] Project month when the activity was actually started or completed
[5] Give a figure used for converting person hours to a person month

Epidemic spreading in correlated complex networks.
Phys. Rev. E. 66, 047104 (2002) [cond-mat/0205621].
15 R. Guimerà, A. Arenas, A. Diaz-Guilera, F. Giralt
Dynamical properties of model communication networks
Phys. Rev. E 66, 026704 (2002). [cond-mat/0206077]
16 R. Guimerà, A. Diaz-Guilera, F. Vega-Redondo, A. Cabrales, A. Arenas
Optimal network topologies for local search with congestion
Phys. Rev. Lett. 89, 248701 (2002)
37 R. Guimera, L. Danon, A. Diaz-Guilera, A. Arenas
Self-similar community structure in organizations
38 R. Guimerà, X. Guardiola, A. Arenas, A. Diaz-Guilera, D. Streib, L.A.N. Amaral
Quantifying the creation of social capital in a digital community Preprint
43 R. Pastor-Satorras, E. Smith, and R. V. Solé
Evolving protein interaction networks through gene duplication.
J. Theor. Biol. (2003) [in press] [SF Working Paper 02-02-008].

COLLABORATIONS WITH OTHER TEAMS:
6 G. Caldarelli, A. Capocci, P. de los Rios, M.A. Muñoz
Scale-free networks from varying vertex intrinsic fitness.
Phys. Rev. Lett. 89, 258702 (2002).
7 G. Caldarelli, R. Pastor-Satorras, and A. Vespignani
Cycles structure and local ordering in complex network.[cond-mat/0212026].
18 Y. Moreno, R. Pastor-Satorras, A. Vázquez, and A. Vespignani
Critical load and congestion instabilities in scale-free networks.
Europhys. Lett. (2003) [in press] [cond-mat/0209474].
20 R. Pastor-Satorras and A. Vespignani
Epidemics and immunization in scale-free networks.
In "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)
[cond-mat/0205260].
21 A. Vázquez, M. Boguñá, Y. Moreno, R. Pastor-Satorras, and A. Vespignani
Topology and correlations in structured scale-free networks. PRE (in press).[cond-mat/0209183].
22 A. Vázquez, R. Pastor-Satorras, and A. Vespignani
Large-scale topological and dynamical properties of the Internet.
Phys. Rev. E 65, 066130 (2002) [cond-mat/0112400].
23 A. Vázquez, R. Pastor-Satorras, and A. Vespignani
Internet topology at the router and autonomous system level.
[cond-mat/0206084].
26 M. Boguñá, R. Pastor-Satorras, and A. Vespignani
Absence of epidemic threshold in scale-free networks with degree correlations.
Phys. Rev. Lett. 90, 028701 (2003)
[cond-mat/0208163].
27 M. Boguñá, R. Pastor-Satorras, and A. Vespignani
Epidemic spreading in complex networks with degree correlations.
[cond-mat/0301149].

Deviations from the planned work schedule/reasons/corrective actions/special attention required

Due to the delay in the hiring of the post-doc the work on the role of self-organisation in networks has been also delayed. Nevertheless, some work on optimisation that had been scheduled to start on the second year has been advanced because of the work of some members of the team.

Planned actions for the next period

Progress Overview Sheet (Partner CR4 UNIL)

PROGRESS OVERVIEW SHEET[1]

Organisation: Université de Lausanne

Workpackage/ Task	Planned effort[2]	Planned Date[3]		Actual Date[4]		Resources employed[2]	Cumulative Resources[2]
	Whole Project	Start	End	Start	End	This Period	Since start
WP 1	6	1	36	4	39		
WP 2	36	1	36	4	39	8	8
WP 3	4	1	36	4	39		
WP 4	4	1	36	4	39		
WP 5	4	1	36	4	39		
WP 6	2	1	36	4	39		
WP 7	1	1	36	4	39		
WP 8	1	1	36	4	39		
Total	58						
One person month is equal to		170 [5]		Person hours			

Main contribution during this period	
Workpackage/Task	**Action**
WP 1	***Mathematical Tools For Complex Systems***
Task 1.2	In collaboration with CO1 and the University of Fribourg and the University of Granada, we have proposed a new mechanism in order to form complex networks. This is particularly suitable for the Internet Network. We applied such mechanism in order to analyze data we collected on Protein Interactions Maps and WWW. This resulted in two papers that enter in the deliverable D05 one devoted to WWW that has been submitted to Physical Review E, and another on Protein Interaction that is in the form of a draft.
WP 2	***Data Collection***
Task 2.1	With the Collaboration of CO1 and Thomas Erlebach (ETH Zurich) we set up a database of data and software tools at **www.cosin.org .** The data collected until now are about Internet, Social and Biological Systems and Protein Interactions
Task 2.2	We realized a C++ implementation of the Brandes algorithm to compute the betweenness measure of centrality in the networks. This has been done by Julien Pitteloud (EPF Lausanne). The algorithm is currently used by other members of COSIN (CO1, CR5)
Task 2.3	See Task 2.1
WP 5	***Model for Communication Networks***
Task 5.1	See Task 1.2
WP 6	***Dissemination of the Results***

Deliverables due this period		
Deliverable number	**Title of Deliverable**	**Status (Draft Final, Pending)**
D05	Preliminary Analysis of Collected data	(two papers) 1 Pending + 1 Draft

[1] Each partner should fill in its own Progress Overview Sheet for a period in question. The Project Co-ordinator will check and approve the forms and attach them to the corresponding PPR.
[2] In person months (or in person hours)
[3] Project month when the activity was planned to be started or to be completed
[4] Project month when the activity was actually started or completed
[5] Give a figure used for converting person hours to a person month

Dissemination actions (articles, workshops, conferences etc.)

Publications:

6 G. Caldarelli, A. Capocci, P. De Los Rios and M.A.Muñoz *Physical Review Letters* **89** 258702 (2002)
8 A. Capocci, G. Caldarelli and P. De Los Rios preprint submitted to *Physical Review E*
11 P. De Los Rios *Proceeding of the VII conference in Granada (to appear)*
40 G. Caldarelli, P. De Los Rios, S. Leonardi, S. Millozzi and A. Vespignani preprint
41 G. Caldarelli, P. De Los Rios and F. Squartini preprint

Talks:

VII Conference on Statistical and Computational Physics Granada 2-7/09/2002
School of Biology Firenze 3-5/02/2003
Stadybis INFM meeting Firenze 7/02/2003

Algorithms:

C++ implementation of Brandes algorithm.

Deviations from the planned work schedule/reasons/corrective actions/special attention required

Node 4 is not directly paid by European Union, but rather by the Swiss Confederation through the OFES-Bern department. They began to pay at June 2002 and can continue until May 2005. CR4 is then formally still in the first year of work.

Planned actions for the next period

A PhD student has been hired in June 2002. Starting Fall 2003, after a physiological first year of training, he is expected to produce scientific activity and actively contribute to meet the deliverables of the project.

Progress Overview Sheet (Partner CR5 ENS)

PROGRESS OVERVIEW SHEET[1]

Organisation: Ecole Normale Superiéure

Workpackage/ Task	Planned effort[2]	Planned Date[3]		Actual Date[4]		Resources employed[2]	Cumulative Resources[2]
	Whole Project	Start	End	Start	End	This Period	Since start
WP 1	4	1	36	1	36		
WP 2	7	1	36	1	36	2	2
WP 3	5.5	1	36	1	36	2	2
WP 4	36	1	36	1	36	11	11
WP 5	2	1	36	1	36		
WP 6	2	1	36	1	36		
WP 7	1	1	36	1	36		
WP 8	1	1	36	1	36		
Total	58.5					15	15
One person month is equal to		170 [5]		Person hours			

Main contribution during this period	
Workpackage/Task	**Action**
WP 2	***Data Collection***
Task 2.1	• Collection of data set of shareholding networks for the following stock markets: Milan, Nasdaq, Nyse
WP 3	***Large Network Visualization Tools***
Task 3.1	• Realization of prototype visualization software tools allowing for visualization of subnetworks of nodes with specific features or nodes at a chosen degree of separation from a starting node.
WP 4	***Dynamics of Social Networks***
Task 4.1	• Characterization of the role of groups in decision networks • Study of statistical properties of shareholding networks
WP 6	***Dissemination of the Results***
	See attached list

Deliverables due this period		
Deliverable number	**Title of Deliverable**	**Status (Draft Final, Pending)**
D07	Centrality and groups in social networks	Final

Dissemination actions (articles, workshops, conferences etc.)

Publications:
S.Battiston, E.Bonabeau, G.Weisbuch, Opinion Dynamics in the corporate board network, to appear on Physica A.
article: S.Battiston, G.Caldarelli, M.Castri, D.Garlaschelli, L.Pietronero, Statistical Properties of Shareholding Networks, in preparation

Talks:
oral presentation at WEHIA 2002

Deviations from the planned work schedule/reasons/corrective actions/special attention required

11 person months have been devoted to task 4.1 instead of the planned 12. The deliverable has been accomplished. No need for any action.

Planned actions for the next period

Modelling of visions spread across the board network, special attention to external triggering events. Investigate relationships between firm network topology and stock values correlation

[1] Each partner should fill in its own Progress Overview Sheet for a period in question. The Project Co-ordinator will check and approve the forms and attach them to the corresponding PPR.
[2] In person months (or in person hours)
[3] Project month when the activity was planned to be started or to be completed
[4] Project month when the activity was actually started or completed
[5] Give a figure used for converting person hours to a person month

Progress Overview Sheet (Partner CR6 UKON)

PROGRESS OVERVIEW SHEET[1]

Organisation: Universitaet Konstanz

Workpackage/ Task	Planned effort[2]	Planned Date[3]		Actual Date[4]		Resources employed[2]	Cumulative Resources[2]
	Whole Project	Start	End	Start	End	This Period	Since start
WP 1	1	1	36	1	36		
WP 2	1	1	36	1	36	8	8
WP 3	36	1	36	1	36		
WP 4	2	1	36	1	36		
WP 5	1	1	36	1	36		
WP 6	2	1	36	1	36		
WP 7	1	1	36	1	36		
WP 8	1	1	36	1	36		
Total	45						
One person month is equal to		170 [5]		Person hours			

Main contribution during this period	
Workpackage/Task	**Action**
WP 3	***Large Networks Visualization Tools***
Task 3.1	Realization of the software visone as well as publication of algorithms and results
WP 6	***Dissemination of the results***

Deliverables due this period		
Deliverable number	**Title of Deliverable**	**Status (Draft Final, Pending)**
D06	Algorithms for representing Network Centrality	Final

Dissemination actions (articles, workshops, conferences etc.)

Publications:

29 Ulrik Brandes, Markus Eiglsperger, Ivan Herman, Michael Himsolt, and M. Scott Marshall
Proc. 9th Intl. Symp. Graph Drawing (GD '01), LNCS 2265, pp. 501-512. © Springer-Verlag, 2002.
30 Volker Maag and Thomas Willhalm
preprint
31 U. Brandes and D. Wagner *accepted on Graph Drawing Software Springer Series "Mathematics and Visualization" M. Junger P. Mutzel (eds).*
32 Dorothea Wagner *to appear in Proceedings of WEA 2003, LNCS*
34 U. Brandes, P. Kenis and D. Wagner accepted on IEEE transactions on Visualization and Computer Graphics.
35 U. Brandes and S. Cornelsen accepted on Journals of Graph Algorithms and Applications.
36 U. Brandes, M. Gaertler and D. Wagner preprint.

Deviations from the planned work schedule/reasons/corrective actions/special attention required

.

Planned actions for the next period

[1] Each partner should fill in its own Progress Overview Sheet for a period in question. The Project Co-ordinator will check and approve the forms and attach them to the corresponding PPR.
[2] In person months (or in person hours)
[3] Project month when the activity was planned to be started or to be completed
[4] Project month when the activity was actually started or completed
[5] Give a figure used for converting person hours to a person month

Comparative Information on Resources (Person months)

Effort in person months reporting period 01/03/2002-28/02/2003

	CO1 INFM				CR2 UDRLS				CR3 UB				CR4 UNIL			
	Period		Total		Period		Total		Period		Total		Period		Total	
WP/Task	Est.	Act.	Est.	Act.	Est.	Act.	Est.	Act.	Est.	Act.	Est.	Act.	Est.	Act.	Est.	Act.
WP1																
Task 1.1									12	12	12	12				
Task 1.2	4	2	4	2					12	0	12	0				
Task 1.3									0	6	0	6				
WP2																
Task 2.1	2	1	2	1									8	8	8	8
Task 2.2																
Task 2.3	1	1	1	1												
WP3																
Task 3.1					12	8	12	8								
WP4																
Task 4.1	4	4	4	4												
WP5																
Task 5.1	1	1	1	1	8	7	8	7								
WP6	3	3	3	3												
WP7 - WP8	3	3	3	3												
Total	**18**	**15**	**18**	**15**	**20**	**15**	**20**	**15**	**24**	**18**	**24**	**18**	**8**	**8**	**8**	**8**

Effort in person months period 01/03/2002-28/02/2003

	CR5 ENS				CR6 UKON				Total			
	Period		Total		Period		Total		Period		Total	
WP/Task	Est.	Act.	Est.	Act.	Est.	Act.	Est.	Act.	Est.	Act.	Est.	Act.
WP1												
Task 1.1									12	12	12	12
Task 1.2									16	2	16	2
Task 1.3									0	6	0	6
WP2												
Task 2.1	2	2	2	2					12	11	12	11
WP3												
Task 3.1	2	2	2	2	8	8	8	8	22	18	22	18
WP4												
Task 4.1	12	11	12	11					16	15	16	15
WP5												
Task 5.1	1	1	1	1					10	9	10	9
WP6	3	3	3	3					3	3	3	3
WP7 - WP8	3	3	3	3					3	3	3	3
Total	**16**	**15**	**16**	**15**	**8**	**8**	**8**	**8**	**94**	**79**	**94**	**79**

Period: Est.: estimated effort in contract for period Act.: effort actually spent in period
Total: Est.: estimated cumulative effort to date in contract Act.: cumulative effort to date actually spent

2.5 WORLD WIDE STATE OF THE ART

The present complexity of the information networks has reached a level that puts them beyond our ability to manage them, and keep them functioning correctly through traditional techniques. Therefore in the last years we witnessed the formation of a new discipline that took inspiration from different fields as statistical physics, social and computer science as well as biology.

Indeed when approaching communication networks (i.e. INTERNET and WWW) we note that a static representation in the form of an Erdős-Renyí Random Graph is no more accurate. They display a scale-free behaviour in the degree distribution rather than a Poissonian distribution. They show also large deviations from average clustering as a signature of presence of communities[5]. As for the distance this is usually peaked on a small value[6]. More generally, a complete set of quantities ranging from the betweenness (i.e. how much a vertex/edge is central) to the connectivity (i.e. robustness against random failure) display unexpected statistical properties. Such behaviour is not only limited to communication networks. Actually, the network structure may describe many different real-world units such as internet providers, electricity providers, economical agents, and ecological species. In all these situations dramatic deviations from the random graph paradigm are present. They emerge as a complex collective phenomenon from the dynamics of the system (such as the internet traffic, electricity supply service, market trend, environmental resources depletion etc.). New models and paradigms are now needed in order to predict the behaviour of such systems.

The present state of art of research in graph models is able to reproduce only partially the rich phenomenology of the data observed. Then at the moment only a list of the most successful ideas can be produced. We are therefore still working on the general understanding of these phenomena. The goals mentioned below can be presently found and we are trying to work in our project along these directions.

- A possible optimisation of networks in order to comply with desired cost function.
- A dynamic representation of graphs. That is to study not only the topology, but rather the *load* of graphs, that is to assign a *weight* on the links in the graphs. Such weight can represent the traffic between Autonomous Systems in the Internet network or the number of clicks between html documents in the WWW.
- Systems with adding/removing components, (as peer to peer networks, mobile ad hoc networks etc.) arising from the highly dynamic environment in which they are deployed.
- Collaboration as well as competition as in real market systems or in a biological evolution context.

2.6 SUGGESTION FROM ADVISORY BOARD

Albert-Laszlo Barabasi
Emil T. Hofman Professor of Physics

The COSIN project, in a bit more than a year, has become a quite visible force in network modeling. This visibility resulted from a series of new projects, that materialized in important papers, attracting the attention of the research community. While the nature of most research is such that typically large project start producing visible results only from the second or the third year of their existence, COSIN was different: important results started to emerge from the first day.

Let me comment on a few areas that might be important for the future development of the project:

1. **Research**: so far that has been first rate. The results range from virus spreading on the internet, to modeling the Internet, or the truly nice paper on the possibility of generating scale-free networks using quenched fitness, which was highly unexpected. The focus on economic systems is also unique: economists are typically way behind applying network thinking to economic systems. The members of the COSIN project represent a notable exception. I feel, however, that the list of papers on the website is far from complete: papers that I know that have come out from the group members are missing. Thus a more up-to-date paper database would serve well the project.

2. **Deliverables**: many deliverables, in the form of the various models that the group has proposed to deliver, appear to be on track based on the available publications. It is hard for me to judge, based on the available information, the most recent results, but what I have seen so far was quite encouraging.

3. **Outreach**: The creation of the Complexity Review is an important step. It is of future importance, however, to add content to it. Right now the infrastructure for that appears to be present, which is a major milestone.

4. **Conferences and Workshops**: COSIN has played a leading role in maintaining a presence at all major network related conferences, as well as in initiating new workshops. Their presence was quite important for the evolution of the field. The upcoming conference in Rome does promise to be a landmark event as well. The future plans for a satellite meeting to the Statphys, as well as a closing conference, are all excellent ideas that would serve the project's needs.

5. **Recommendations**: If the group continues its research and activity along the planned path that should lead to important results. Several challenges lie ahead, however. The first one, as I mentioned, is to convince the members to add content to Complexity Review, as well as to find sources to maintain its future content. Second, some of the datasets the group members have collected offer information not only about network topology, but dynamics as well. A closer focus on network dynamics would continue keeping the collaboration at the cutting edge. Third, for new people to join the network community one needs to lower the entry barriers. The COSIN project has a unique opportunity to disseminate programs. Making available not only new programs, but simpler programs as well, that can generate random or scale-free models, would be valuable for many from other fields, that are less familiar with

programming. Just a collection of programs for the few much used basic models would be a great start. Finally, the project is too important to see it fade away in two years. It is the time for the members to start charting the path for new funding that would continue this important research going.

Mark Buchanan
Science Writer

This is a brief report offering my assessment of progress on the COSIN project through February 2003. The past five years have witnessed a small explosion of scientific activity aimed toward building a deeper understanding of complex networks as they arise both in the natural world and in the context of human affairs. Examples range from food webs and the networks of protein-protein interactions that underpin cellular biochemistry to the Internet and the World Wide Web. In the context of this rapidly advancing body of work, the COSIN group envisioned, at the project's inception, to aim for the achievement of three general goals:

- first to develop effective techniques for analysing, characterising and visualising complex networks having an extremely large number of elements (the Internet and WWW representing two especially important examples);
- second, to use these techniques in conjunction with theoretical and analytical concepts from statistical physics and complex systems theory to probe similarities and differences in the growth dynamics of distinct networks, to elucidate the relationship between growth dynamics and resulting topology of a network, and to explore and understand how important properties such as network stability and functional efficiency develop upon this topology;
- third, to use the understanding so developed to disseminate practical knowledge regarding potential improvements in the design and management of both the Internet and WWW, as well as of other large, complex networks, especially social networks.

To date, and in keeping with the schedule of proposed work and deliverable items, the COSIN team has made what I believe to be excellent progress in each of these areas.

1. Notable Achievements:

a. COSIN researchers have contributed significantly in the past year to improving our understanding of the Internet's topology. The Internet is a scale-free network, with a degree distribution that follows a power law $P(k) \sim k^{-\gamma}$. This property can be understood as following from a simple growth dynamics in which nodes gain new links at a rate that grows in linear proportion to the number they already have ("preferential attachment"). However, the COSIN group has demonstrated empirically that the Internet possesses a number of subtle and potentially very important topological features (correlations and clustering) that cannot be explained by this rudimentary model. Moreover, they have shown that these features are naturally reproduced in models that include geographical information about the nodes (reflecting the extra wiring costs involved in establishing links over longer distances, for example) as well as other influences such as the aging of nodes. This work offers a significantly improved empirical picture of the statistical topology of the Internet, and should pay practical dividends in the future. Improved simulation and understanding of Internet dynamics – from the operation of routing algorithms to the spreading of viruses – can only follow from an improved model of the network's topology.

b. The COSIN group has also contributed a number of important results to the basic theory of network topology. One of these is the development of novel theoretical measures for exploring the clustering properties of nodes within a network. For several years researchers have explored the clustering of nodes within networks primarily through the use of a simple "clustering coefficient". The new measures offer a more powerful probe of cyclic structures. In particular, they reveal both the Internet and WWW to have a surprising "grid-like" ordering, and a strongly hierarchical organization. This work not only offers an improved picture of several real-world networks, but a powerful tool that can be widely applied. Members of the COSIN group have also shown that scale-free networks may arise through a process that need not involve growth at all. Although the implications of this discovery remain to be assessed, it illustrates that the development of basic theory in this area remains subject to major surprises. The COSIN group has also explored and achieved important results regarding network topologies that allow for the optimal performance of search algorithms. These results have clear relevance to the design and improvement of communication networks, yet also offers potential insights into the design of organizational structures (within business firms, for example).

c. Another central aim in the study of complex networks is to understand dynamical processes that take place within them, and especially to appreciate in detail how the architecture of a network affects such dynamics. In this context, the COSIN group has already contributed several seminal studies that offer deep insights into both theoretical and practical issues. The group has shown, for example, that there is no "epidemic threshold" for the spreading of a virus (or other influence) within a scale-free network. This work offers a dramatic illustration of the potential for findings of the COSIN project to translate into practical knowledge relevant to diverse social needs. Although epidemiological researchers have understood for many years that the pattern of linkages within a social network may strongly affect how a disease may spread, the COSIN results suggest specific strategies for inhibiting the spread of either computer viruses or sexually transmitted diseases such as AIDS. A number of COSIN researchers have also achieved important initial results in understanding the propagation of cascading failures in scale-free networks. They have shown, for example, that a scale-free network undergoes a natural transition from a free phase to a congested phase as the level of traffic increases. This represents a significant beginning step in the modelling and hopefully management of traffic instabilities in real communication networks.

d. The COSIN group has made early progress toward one of it more ambitious goals – contributing to a deeper understanding of social networks and their function. In particular, COSIN researchers have developed algorithms for identifying and visualizing the community structure (significant groups and cliques) of real social networks. These algorithms are of interest to theorists in organization and management as they offer tools for exploring the patterns of interpersonal interactions that underlie an organization's social architecture. With these methods, for example, one might identify the lack of an important "bridging link" between distinct sub-communities, and so identify a clear target for organizational improvement. COSIN researchers have also taken seminal steps in modeling the spread of social innovations, such as technological innovations. Work in the area of social networks and their dynamics might obviously evolve in innumerable distinct directions. So far, I believe the COSIN group has focused appropriately on taking specific knowledge and techniques relevant to networks *per se* and extending these in a sensible way to understanding aspects of the social world.

The specific achievements listed above are a few of those that seem notable to me. Overall, the COSIN group appears to be moving steadily toward the achievement of its stated goals.

Suggestions on Dissemination and Use/Future Activities: Although the COSIN work is relevant to a wide variety of scientific studies, the results will ultimately, I believe, be of particular value to social scientists – especially to scientists concerned with organization and management in the context of business firms, communities, markets, etc., as well as to those attempting to understand social and economic organization and dynamics from a "bottom up" or "agent based" perspective. Advances in these areas have been hampered by the lack of powerful quantitative methods for exploring complex, disorderly social networks, and for discerning the influence of topology (the social "connectedness" of a community, or the lack of it, for example) on social processes. Hence, I suggest that the COSIN group might beneficially attempt to spread its results to researchers in these areas, with the expectation that useful information/insights/ideas/data will flow in both directions.

In this regard, it may be worthwhile to foster the exchange of ideas by inviting key social scientists or management/organization theorists to future COSIN workshops or meetings. In the area of agent-based modeling, I have in mind names such as Robert Axtell (Brooking Institute) or Robert Axelrod (University of Michigan), and in organization and management, Kathleen Carley (Carnegie Mellon University). As another name, Mark Granovetter (Stanford University) has long experience studying community organization and its influence on social dynamics from both theoretical and experimental perspectives. [I am not aware of specific names, but I am sure there are European researchers with similar expertise.] Also, perhaps it might be possible to organize a small session on complex networks at one of the many meetings on agent-based model in economics or social science.

2.7 ACTIONS TAKEN WITH RESPECT TO SUGGESTIONS AND PLANNED WORK

We are particularly grateful to the board for the work done. We report hereafter a point to point answer to both advisors suggestions

- As regards the updating of the web site, we decided in a first approach to present only the papers already published. We realized that in a field fast changing as the growing network field it is very important to help ideas circulation also through dissemination of preprints. In the present version of the site we have all the papers produced by teams as soon as they are ready for publication (whenever this is possible according to journals rules).
- As regards the archive of Complexity Review we are currently dealing with FET NoE "Exystence", with the Santa Fe Institute and Los Alamos Laboratory in order to form a critical mass to ensure success to archive. We expect to find an agreement for late May when most of the project members will be in Santa Fe for a conference on Network. In any case launch of the archive is scheduled for Autumn 2003.

Project Management and Coordination

The management of the consortium has been facilitated by the great scientific activity in the field. Since the large number of conference and the activities related to the presentation of the VI Framework Programme by European Community WP leaders met on a regularly basis for all the first year. The creation of a virtual meeting point represented by the project site also contributed in the exchange of ideas and feedback for the activities of the partners.

The principal deviation from the scheduled plan in coordination of activities regards the first year workshop, in order to have a more timely presence in the field COSIN helped in organizing the Granada seminar on Computational Physics and it is currently organizing a very large midterm meeting. We hired a part-time administrative assistant for the first year of the project and she is currently taking care of this first year report and of the midterm meeting.

Since the continuous feedback and consultation of WP leaders we believe we obtained rather good results in the various tasks of the project. For such a reason no particular coordination actions has been necessary until now. We are planning after the midterm meeting to take into consideration new ideas and developments that should arise in the field. It is possible that the activity of the project could be slightly rearranged in order to take into account novel lines of research of great interest. Possible candidates could be biology inspired algorithms and peer to peer networks.

Cost Breakdown

Appendix 4 (b) - Comparative Information on Resources (Costs)

Costs in euro for reporting period 1/3/2002 -28/2/2003

	CO1 INFM				CR2 UDRLS				CR3 UB				CR4 UNIL			
	Period		Total		Period		Total		Period		Total		Period		Total	
Cost category	Est.	Act.	Est.	Act.	Est.	Act.	Est.	Act.	Est.	Act.	Est.	Act.	Est.	Act.	Est.	Act.
Direct costs																
1. Personnel	48594.00	24267.33	176567.00	24267.33	74730.00	81392.06	224190.00	81392.06	32585.00	7699.69	133000.00	7699.69	51428.00	23304.78	162722.00	23304.78
2. Durable equipment	4000.00	685.98	4000.00	685.98	12000.00	1599.51	12000.00	1599.51	13000.00	4290.00	13000.00	4290.00				
3. Subcontracting						7300		7300								
4. Travel and subsistence	3000.00	5358.45	9434.00	5358.45	5620.00	2778.83	16859.00	2778.83	6000.00	5484.74	18000.00	5484.74	4360.00	1742.87	13642.00	1742.87
5. Consumables																
6. Computing																
7. Protection of knowledge																
8. Other specific costs							8000.00									
Subtotal	55594.00	30311.76	190001.00	30311.76	92350.00	93070.4	253049.00	93070.4	51585.00	17474.43	164000.00	17474.43	55788.00	25047.65	176364.00	25047.65
Indirect costs																
9. Overheads	11119.00	6062.35	38001.00	6062.35	59784.00	65113.65	179352.00	65113.65	10317.00	3494.89	32800.00	3494.89	10317.00	1753..33	17636.00	1753.33
Total	**66713.00**	**36374.12**	**228002.00**	**36374.12**	**152134.00**	**158184.05**	**440401.00**	**158184.05**	**61902.00**	**20969.32**	**196800.00**	**20969.32**	**61902.00**	**26800.98**	**194000.00**	**26800.98**

Period: Est.: estimated costs in contract for period Act.: actual costs in period

Total: Est.: estimated cumulative costs to date in contract Act.: cumulative actual costs to date

Appendix 4 (b) - Comparative Information on Resources (Costs)

Costs in euro for reporting period 1/3/2002 -28/2/2003

	CR5 ENS				CR6 UKON				Total			
	Period		Total		Period		Total		Period		Total	
Cost category	Est.	Act.	Est.	Act.	Est.	Act.	Est.	Act.	Est.	Act.	Est.	Act.
Direct costs												
1. Personnel	75892.00	52662.75	231108.00	52662.75	36960	32418.00	153199	32418.00	320189.00	221744.61	1080786.00	221744.61
2. Durable equipment									29000.00	6575.49	29000.00	6575.49
3. Subcontracting										7300		7300
4. Travel and subsistence					6000	2053.22	17634	2053.22	24980.00	17418.11	75569.00	17418.11
5. Consumables												
6. Computing												
7. Protection of knowledge												
8. Other specific costs			8000.00						16000.00		16000.00	
Subtotal	75892.00	52662.75	231108.00	52662.75	42960	34471.22	170883	34471.22	374169.00	253038.21	1185355.00	253038.21
Indirect costs												
9. Overheads	60714.00	42130.00	184888.00	42130.00	8592	6894.24	34167	6894.24	156105.00	125448.46	486844.00	125448.46
Total	**136606.00**	**94792.75**	**423996.00**	**94792.75**	**51552**	**41356.46**	**205000**	**41356.46**	**530274.00**	**378486.67**	**1688199.00**	**378486.67**

Period: Est.: estimated costs in contract for period Act.: actual costs in period
Total: Est.: estimated cumulative costs to date in contract Act.: cumulative actual costs to date

Information dissemination and exploitation of results

The following list of activities of the consortium is divided according to publications, conferences and schools. Most of the following publications are downloadable from the project site at http://www.cosin.org/Publications.html

PUBLICATIONS

1. Decision making dynamics in corporate boards.
   S. Battiston, E. Bonabeau and G. Weisbuch
   ArXiv:cond-mat/0209590 (2002).
2. Number of Loops of size *h* in growing scale-free networks.
   G. Bianconi and A.Capocci
   ArXiv:cond-mat/0212028 (2002).
3. Epidemic spreading in correlated complex networks.
   M. Boguñá and R. Pastor-Satorras
   Physical Review E **66** 047104 (2002).
4. Topology of correlation based minimal spanning trees in real and model markets.
   G. Bonanno, G. Caldarelli, F. Lillo and R. Mantegna
   ArXiv:cond-mat/0211546 (2002).
5. Introduction to Complex Networks.
   G. Caldarelli
   Proceedings of the 7th Conference on Statistical and Computational Physics Granada (2002).
6. Scale-Free Networks from Varying Vertex Intrinsic Fitness.
   G. Caldarelli, A. Capocci, P. De Los Rios and M. A. Muñoz
   Physical Review Letters **89**, 258702 (2002).
7. Cycles structure and local ordering in complex networks.
   G. Caldarelli, R. Pastor-Satorras and A. Vespignani
   ArXiv:cond-mat/0212026 (2002).
8. Quantitative description and modelling of real networks.
   A.Capocci, G. Caldarelli, and P. De Los Rios
   ArXiv:cond-mat/0206336 (2002).
9. Visualization of the High Level Structure of the Internet with Hermes.
   A.Carmignani, G. Di Battista, W. Didimo, F. Matera and M. Pizzonia
   Journal of Graph Algorithms and Applications **6**,281 (2002).
10. How Can Extremism Prevail? A study based on the relative agreement interaction model.
    G. Deffuant, F. Amblard, G. Weisbuch and T. Faure
    Journal of Artificial Societies and Social Simulation **5** (2002).
11. Exploration bias of Complex Networks.
    P. De Los Rios
    Proceedings of the 7th Conference on Statistical and Computational Physics Granada (2002).
12. Drawing database schemas.
    G. Di Battista, W. Didimo, M. Patrignani and M. Pizzonia
    Software-Practice and Experience. **32**, 1065 (2002).
13. Archive of BGP Updates: Integration and Visualization.
    G. Di Battista, F. Mariani, M. Patrignani and M. Pizzonia
    Preprint (2002).
14. 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).
15. Dynamical properties of model communication networks.
    R. Guimerà, A. Arenas, A. Díaz-Guilera and F. Giralt
    Physical Review E **66**, 026704 (2002).
16. Optimal network topologies for local search with congestion.
    R. Guimerà, A. Díaz-Guilera, F. Vega-Redondo, A. Cabrales, and A. Arenas
    Physical Review Letters **89**, 248701 (2002).
17. A Multi-layer model for the Web Graph
    L. Laura, S. Leonardi, G. Caldarelli and P. De Los Rios
    Contribution to 2nd Web Dynamics Workshop (2002).
18. Critical load and congestion instabilities in scale-free networks.
    Y. Moreno, R. Pastor-Satorras A. Vázquez and A. Vespignani
    ArXiv:cond-mat/0209474 (2002).

19. Scale-free behavior of the Internet global performance.
R. Percacci and A. Vespignani
ArXiv:cond-mat/0209619 to be published in European Journal of Physics B (2002).
20. Epidemics and immunization in scale-free networks.
R. Pastor-Satorras and A. Vespignani
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).
21. Topology and correlations in structured scale-free networks.
A. Vázquez, M. Boguñá, Y. Moreno, R. Pastor-Satorras and A. Vespignani
ArXiv:cond-mat/0209183 (2002).
22. Large-scale topological and dynamical properties of the Internet.
A. Vázquez, R. Pastor-Satorras, and A. Vespignani
Physical Review E **65**, 066130 (2002).
23. Internet topology at the router and autonomous system level.
A. Vázquez, R. Pastor-Satorras, and A. Vespignani
ArXiv:cond-mat/0206084 (2002).
24. Adjustment and social choice.
G. Weisbuch and D. Stauffer
submitted to *Physica A* (2002).

2003

25. Size of Quantum Networks.
G. Bianconi
ArXiv:cond-mat/0301551.
26. Absence of Epidemic threshold in scale-free networks with degree correlations.
M. Boguñá R. Pastor-Satorras and A. Vespignani
Physical Review Letters **90**, 028701 (2003).
27. Epidemic spreading in complex networks with degree correlations.
M. Boguñá R. Pastor-Satorras and A. Vespignani
ArXiv:cond-mat/0301149 to appear in Lecture notes in Physics (2003).
28. Topology of the World Trade Web.
M. A. Serrano, and M. Boguñá
ArXiv:cond-mat/0301015.
29. GraphML Progress Report: Structural Layer Proposal.
Ulrik Brandes, Markus Eiglsperger, Ivan Herman, Michael Himsolt, and M. Scott Marshall
Proc. 9th Intl. Symp. Graph Drawing (GD '01), LNCS 2265, pp. 501-512. © Springer-Verlag, 2002.
30. Guide - Changes
Volker Maag and Thomas Willhalm
preprint
31. visone - Analysis and Visualization of Social Networks.
U. Brandes and D. Wagner
accepted on Graph Drawing Software Springer Series "Mathematics and Visualization" M. Junger P. Mutzel (eds).
32. Analysis and Visualization of Social Networks
Dorothea Wagner
to appear in Proceedings of WEA 2003, LNCS
33. A Multi-Layer Model for the Web Graph
L. Laura, S. Leonardi, G. Caldarelli and P. De Los Rios
Proceedings of the 2nd International Workshop on Web Dynamics
34. Communicating Centrality in Policy Network Drawings.
U. Brandes, P. Kenis and D. Wagner
accepted on IEEE transactions on Visualization and Computer Graphics.
35. Visual Ranking of Link Structures
U. Brandes and S. Cornelsen
accepted on Journals of Graph Algorithms and Applications.
36. on Clustering Large graphs.
U. Brandes, M. Gaertler and D. Wagner
preprint.
37. Universality in Food webs
D. Garlaschelli, G. Caldarelli, L. Pietronero
accepted on Nature.

38. Self-similar community structure in organizations
    R. Guimerà, L. Danon, A. Diaz-Guilera, A. Arenas
    preprint
39. Quantifying the creation of social capital in a digital community
    R. Guimerà, X. Guardiola, A. Arenas, A. Diaz-Guilera, D. Streib, L.A.N. Amaral
    Preprint
40. A study of stocastic model for the Web Graph
    G. Caldarelli, P. De Los Rios, S. Leonardi, S. Millozzi, A. Vespignani
    Preprint
41. Topology of Protein Protein Interaction Networks from Physical Principles
    G. Caldarelli, P. De Los Rios, F. Squartini
    Preprint
42. Statistical Properties of Shareholders Network.
    S. Battiston, G. Caldarelli, M. Castri, D. Garlaschelli, L. Pietronero
    preprint
43. Evolving Protein Interaction Network through gene duplication
    R. Pastor-Satorras, E. Smith, R. V: sole
    preprint

CONFERENCES AND SCHOOLS

1 XVIII Sitges Conference on Statistical Mechanics, 10-14 June 2002.

2 "Models and Algorithms for the WWW" School, Udine July 2002. (organised by COSIN)

3 VII Conference on Statistical and Computational Physics Granada 2-7 September 2002 (sponsored by COSIN).

4 School of Biology Firenze 3-5 February 2003

5 Stadybis INFM meeting Firenze 7 February 2003

6 APS Conference Austin (USA) 2-7 March 2003.

7 Midterm Conference (see Appendix C) 1-5 September 2003 (organised by COSIN).

Appendix A. Objectives and milestones from proposal

Project Abstract

In this project we want to develop statistical models to describe networks growth and evolution. These models will be based on agents interactions and inspired to the theory of Self-Organization and Fractal Growth. At the same time we are thinking to collect data mainly for Internet and World Wide Web structure in order to validate the models and we want to devise visualization tools in order to analyze large data sets both from numerical simulations and from real-world data. Application to social and economic networks will be also considered

Objectives

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.

Description of Work

The study of the dynamics of complex networks requires an interdisciplinary approach that will use methodologies developed in different areas. Complex Systems theory is the natural candidate framework for the study of such networks. It provides a unified language and a set of operative tools to address the fundamental issues involved in the study of such systems. In this framework we want to devise common methodologies for running empirical simulations, computing the relevant quantities on the simulated models and on real samples and visualizing the evolution of networks even on very large dimensions. The unifying feature of these networks is that their global structure and dynamical evolution are the result of locally interacting agents distributed in the system.

Shape of the network and agent requests coevolve to form the final structure. We plan to devise new stochastic models for Internet to be validated through an extensive comparison with a multiview observation from different locations. We also want to devise content sensitive stochastic models of the world Wide Web where documents link to other documents that are relevant for a common subject of interest. The study of such large networks require the development of ad-hoc visualization algorithms and data structures for drawing and browsing virtually infinite structures rapidly evolving in time. We shall also consider the analysis of social and economic networks as cyber communities, correlations between stocks and firm etc. We finally want to test the existing models for the Internet and the World Wide Web by comparing

the behaviour of suitable quantities identified through theoretical analysis and large computer simulations. In particular we refer to the various frequency distributions observed in real samples as the degree distribution, the number of hops required to reach boundary of the system etc.

Milestones

Realization of a stocastic model for the Internet. Realization of a stocastic model for the World Wide Web. Realization of a stocastic model for the social networks. Tools for visualizing large scale networks. We expect at the end of this project to have also a common framework of knowledge in order to proper describe growing networks and Complex Systems in general.

Appendix B. The project site

DATA

We have more than 30 datafiles publicly downloadable from site. They are commented and standardize with respect to the format. We present the source (if not COSIN) and when possible any scientific publication related to them.

1. Internet at AS scale (14 samples from 3000 to 10500 sites)
2. Traceroute analysis generated from the COSIN server
3. Protein map interactions for 5 different organisms
4. Food webs (source from scientific publications)
5. Board of Directors for Italian Stock Exchange, NYSE, Nasdaq
6. Topholders for Italian Stock Exchange, NYSE, Nasdaq
7. Actor movie dataset
8. Taxonomic trees from different ecosystems

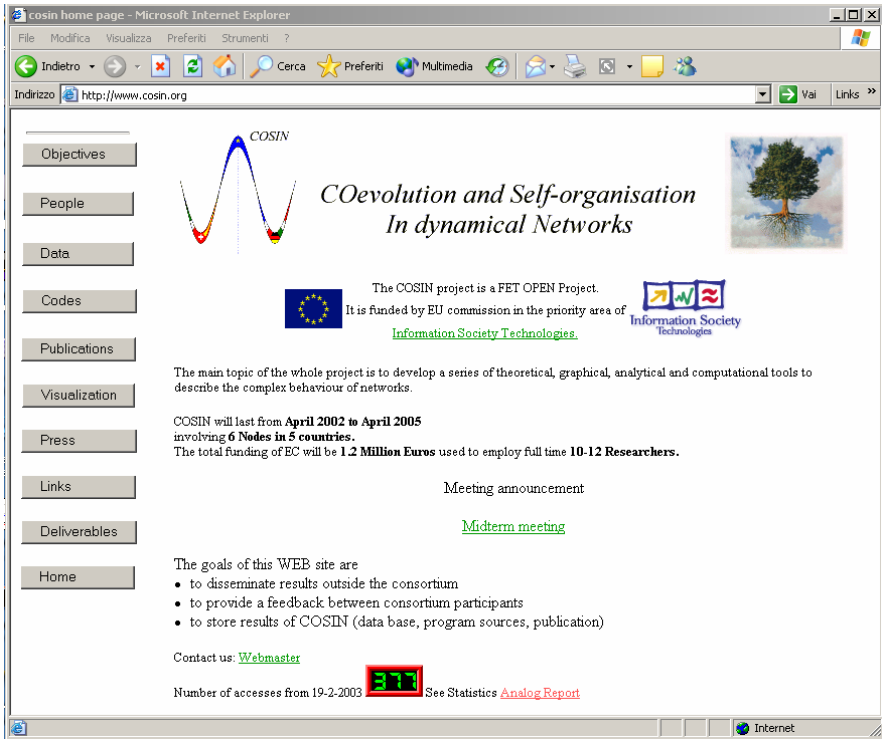
Some of these set of data **(part of 1, 2, 5, 6, 7)** have been specifically collected during the project. Links to other datasets are also present. At the moment we are putting in anonymous way the largest www data set available collected by Alexa corporation. Since the large size this data set is only available under request.

VISUALIZATION TOOLS

All the currently used tools in the state of art have been collected, document and presented in the COSIN site as well as in the mirror site http://www.cadada/COSIN

PUBLICATIONS

We collected all the publications related to the project dividing them according to the topic (deliverable) and in alphabetical list (reported in section 5). Following suggestions of advisors we extended the list of publications to preprints if they publicly available.



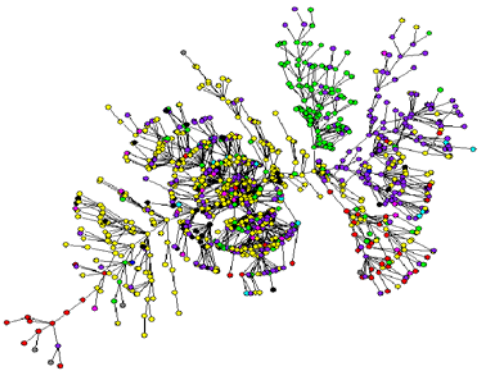
Appendix C. The midterm conference

MIDTERM CONFERENCE COSIN (FET Open Project IST-2001-33555)

Conference on Growing Networks and Graphs in Statistical Physics, Finance, Biology and Social Systems

Organizing Committee: **G. Caldarelli, F. Davide, S. Leonardi, R. Pastor-Satorras, L. Pietronero, A. Vespignani**

Sponsored by: ***Istituto Nazionale Fisica per la Materia (INFM), Telecom Italia, Università "La Sapienza"***



Roma 1-5 September 2003

Aule S. Pietro in Vincoli, Facoltà di Ingegneria, Università di Roma "La Sapienza" v. Eudossiana 18

http://www.cosin.org/midterm.html

Tentative List of Invited Speakers

L. ADAMIC *(HP Palo Alto)*	**A. ERZAN** *(Istanbul Univ.)*	**S. REDNER** *(Boston University)*
L. AMARAL *(NorthWestern University)*	**S. HAVLIN** *(Bar Ilan Univ.)*	**R. V. SOLÉ** *(UPF Barcelona)*
A. ARENAS *(Univ. Tarragona)*	**B. KAHNG** *(Seoul Univ.)*	**A. M. SPACCAMELA** *(Roma)*
A.-L. BARABÁSI *(Notre Dame)*	**J. KERTESZ** *(Eotvos Institute)*	**H. E. STANLEY** *(Boston Univ.)*
B. BOLLOBAS *(Memphis)*	**R. KUMAR*** *(IBM Almaden)*	**Z. TOROCZKAI** *(LANL)*
A. BRODER* *(Altavista)*	**R. MANTEGNA** *(Palermo)*	**D. WAGNER** *(Konstanz Univ.)*
W. R. CHESWICK *(Lumeta)*	**N. MARTINEZ*** *(S. Francisco Univ.)*	**D. WATTS*** *(Columbia)*
P. DE LOS RIOS *(IPT, Lausanne)*	**F. MENCZER*** *(Iowa Univ.)*	**G. WEISBUCH** *(ENS)*
S. DOROGOTSEV *(Porto)*	**J. J. MENDES** *(Porto)*	

- **REGISTRATION**
  Details about registration forms, contribute submissions, accommodation and tentative programme at web site http://www.cosin.org/midterm.html or e-mail Ms. Francesca Consales consales@pil.phys.uniroma1.it
  First Deadline 15 June 2003 Reduced Fee 270 €, Reduced Student Fee 150 € .
  Last Deadline 15 July 2003 Fee 350 €, Student Fee 250 €

* To be confirmed