New perspectives on organizational design based on the analysis of complex communication networks

NEW PERSPECTIVES ON ORGANIZATIONAL DESIGN BASED ON THE ANALYSIS OF COMPLEX COMMUNICATION NETWORKS

Memòria presentada per optar al títol de
Doctor per la Universitat Rovira i Virgili

ROGER GUIMERÀ MANRIQUE
Departament d'Enginyeria Química
Escola Tècnica Superior d'Enginyeria Química
Universitat Rovira i Virgili
Av. Països Catalans 26
43007 Tarragona

Universitat Rovira i Virgili
Tarragona, 2002

Els signants

FEM CONSTAR

Que el present treball que porta per títol

NEW PERSPECTIVES ON ORGANIZATIONAL DESIGN BASED ON THE ANALYSIS OF COMPLEX COMMUNICATION NETWORKS

i que presenta en Roger Guimerà Manrique per optar al grau de Doctor per la Universitat Rovira i Virgili, ha estat realitzat sota la nostra direcció i que tots els resultats presentats i la seva anàlisi són fruit de la investigació realitzada per l'esmentat doctorand.

I perquè se'n prengui coneixement i tingui els efectes que correspongui, signem aquest certificat.

Alex Arenas Moreno
Professor Titular
Dep. d'Enginyeria Informàtica i Matemàtiques
Universitat Rovira i Virgili

Albert Díaz Guilera
Professor Titular
Dep. de Física Fonamental
Universitat de Barcelona

Francesc Giralt i Prat
Catedràtic d'Universitat
Dep. d'Enginyeria Química
Universitat Rovira i Virgili

Tarragona, 16 de setembre del 2002

Als meus pares, l'Anna i el Domènec, i a l'Oriol.

A la Marta.

Als Iaios, la Carme i l'Emilio,
que haurien estat les persones més felices del món.

Agraïments

Això dels agraïments sempre és una cosa injusta. Un pensa en unes persones i se n'oblida d'altres de tal manera que els agraïts acostumen a ser les persones que han estat properes en els temps més recents i els sacrificats aquells que, potser havent ajudat molt al principi, s'han allunyat de mica en mica. Per tant, ben mirat, potser valdria més no fer-ne d'agraïments, i deixar que cadascú es donés per agraït si ho considerava just. Però bé, no agrair és també una mena d'injustícia. Ho faré, doncs, el millor que pugui, ho garanteixo, però no puc assegurar que no em deixaré ningú. Per minimitzar això dels oblits, però, intentaré seguir una mica d'ordre cronològic.

Per començar, és clar, he de donar les gràcies a l'ideòleg del projecte, el Joan Ramon Alabart. Encara que aquesta tesi està molt lluny del que havíem parlat les primeres vegades que ens vem trobar i que segurament el Joan Ramon hauria preferit una cosa més *soft* i més últil, per què no dir-ho, penso que en algun lloc s'ha de veure l'esforç que he fet per reconvertir-me al *management*. També a la Lourdes Vega, perquè sense el seu ajut probablement no hagués aconseguit la beca de què he gaudit tots aquests anys.

El relleu del Joan Ramon el van prendre l'Alex i l'Albert. Suposo que no es poden demanar millors directors de tesi. Perquè m'han ensenyat a fer recerca, perquè m'han dirigit als problemes més interessants, perquè s'han assegut al meu costat a programar, a calcular, a llegir i a escriure, i perquè sempre m'han fet sentir més com un company que no pas com un esudiant. I tot això, un dia darrera de l'altre durant quatre anys. També, més recentment, al Francesc Giralt, que ens va dirigir a la part que, ara com ara, em sembla més interessant de tota la tesi: l'estudi empíric de la xarxa de correu de la URV. I també a tots els altres professors que, d'una manera o altra, són corresponsables d'aquest treball: el Fernando Vega, l'Antonio Cabrales, el Juan Camacho i, especialment, el Luís Amaral.

A part del projecte de tesi, ara fa quatre anys hi havia els cursos de doctorat. He de donar les gràcies a tots els professors que realment van fer un esforç

perquè aprenguéssim alguna cosa, i en particular al Josep Bonet amb qui, més endavant, he mantingut algunes discussions científiques i no científiques, sempre edificants per l'esperit. Als que no van fer aquest esforç, res.

Seguint amb els cursos, un agraïment també se'l mereixen els meus companys de fornada de doctorat, amb qui hem compartit moltes més coses que no pas classes i problemes. Menció especial per la Cata, l'Albert Manyes, l'Àngel (linuxero convençut ara, però no pas quan ens vem conèixer), el Thanasis i el Mohammad. També a la Susana, amb qui vam compartir inacabables hores de docència al Laboratori d'Enginyeria Química I.

Simultàniament a tot això, entre curs i curs, jo feia les primeres passetes en la vida social de Tarragona. Molt a poc a poc, com em sol passar. Tot va començar a La Penya, a invitació de la Clara i del Piotr, els meus primers companys de despatx. Allà, a La Penya, vaig conèixer a alguns dels qui més de cor estic agraït: el Maxi, l'Albert Manyes (altre cop), l'Alvarito i el Robert.

També vaig acabar coneixent el Frank, no recordo ben bé com, però recordo que en algun punt va començar a venir a La Penya i de mica en mica ens vem fer molt amics. El Frank és de les persones que, fent sempre conyeta, això sí, està a prop quan cal. També va començar a venir a casa a veure futbol i a sopar i, de fet, a aquests sopars s'hi va anar afegint gent de mica en mica. Primer el Maxi, el meu germà bessó, sempre disposat a preguntar per l'estat dels meus articles, per què estic trist i en què pot ajudar, i a donar-me sopar quan arribava tard a Tarragona els diumenges. Després el Gökhan que amb la seva rialla permanent aconsegueix animar als morts. Després l'Alvarito i l'Albert, la parella feliç recentment separada per forces majors. I després l'Anton, el Michael i el Robert. Al final, els sopars eren tot un esdeveniment de reconegut interès xafardero-gastronòmic. Gràcies a tots per les bones estones que m'heu fet passar i perquè m'heu fet sentir estimat.

També menció obligada a una colla de persones que escapen a l'ordre cronològic perquè no sabria dir en quin punt van començar a entrar en la meva vida. Suposo que va ser de mica en mica. La Gabriela i el Carlitos perquè m'han fet riure molt i perquè hi he pogut compartir preocupacions científiques vàries: des de xarxes a funcions de correlació. Molt especialment estic agraït a la Rosa i el Josep Maria. Tenim tantes coses en comú que era inevitable que al final acabéssim sent amics: ens agrada discutir acaloradament del que sigui, ens agrada cuinar i menjar bé, som descreguts i "progres" i, bé, més coses que segur que em deixo. M'han fet passar molt bones hores i m'han estimulat el sentit crític, que sempre va bé.

I, és clar, també es mereix l'agraïment més de cor la Montse. Perquè mai fa cas dels meus consells però, tot i així, els segueix escoltant i perquè sempre està de part meva. Si mai algú busca una definició d'amistat, hauria de preguntar a la Montse.

No vull deixar-me, tampoc, a tres dels nous “precaris”: l’Orlando, la Paula i el Leon. Amb ells, tot i el poc temps que hem coincidit, he compartit un nombre sorprenentment gran de coses. I tampoc em vull deixar el Xavi Guardiola. Jo diria que amb el Xavi ens vem conèixer tard, perquè ell ja acabava la tesi i només vem poder treballar junts amb un parell d’idees boges de les seves (una de les quals, per cert, encara espera una introducció per esdevenir article!). Tot i això, amb el Xavi ens han unit, al final, moltes coses, i espero que ens segueixin unint per molt de temps. Valor, equilibri, força i seny.

I finalment, agraïments per la gent que ha estat sempre al meu costat: abans, durant i espero que després de la tesi. La gent de La Llar del Vent: el Xavi Oca, el Narcís, la Isabel, l’Òscar, el Santi, l’Isma, el Xavi Noguera i especialment el Jordi. Les seves visites a Tarragona per anar a sopar al Serrallo sempre han estat un motiu d’alegria. Gràcies a la Marta Terrín i a l’Orlando, que són la mena d’amics que mai deixaran de ser-ho, i també agraïments especials pel Capo: amb ell també he compartit tantes coses i tan diverses que no sabria per on començar.

Agraïments “ordinaris” es queden curts amb els meus pares, l’Anna i el Domènec. Als meus pares els ho dec tot. Aquesta tesi és el seu fruit més que no pas el meu perquè tot ho tinc gràcies a ells.

A l’Uri només hi ha una manera d’agrair-li que m’estimi tant com m’estima: estimant-lo tant com l’estimo.

També ho dec tot a la Marta. Ella, més que ningú, és qui m’aguanta les tristors i qui comparteix les meves alegries i les meves il.lusions. És l’empenta que em fa tirar endavant.

Contents

List of Figures

List of Tables

Chapter 1

INTRODUCTION

The typical chemical company is large, usually with thousands of employees. According to data from the European Union, in 1990 almost 70% of the total turnover generated by the chemical industry corresponded to companies with more than 250 employees (European Comission, 2001). The remaining 30% corresponded, in similar amounts, to small companies with less than 50 employees and to medium sized companies with 50 to 250 employees. Indeed, although some products have regional markets, chemical industry is essentially global and is dominated by large multinationals like Bayer, with 117,000 employees,[1] BASF, with 93.000 employees,[2] DuPont, with 79.000 employees,[3] or Dow Chemical, with 50.000 employees.[4] Specially for such large companies, organizational design and human capital management play a key role, as important, at least, as technology or management of material resources. A substantial part of the human workforce of such a company is devoted to information processing rather than to "make" or "sell" products in the narrow sense. However, most formal analysis of organizations have downplayed communication and information processing and focused on issues related to individual incentives. Only in the last decade, the importance of communication processes in organizations has started to be understood, mainly in the economics literature.

Parallel to these efforts to understand the role of communication in organizations, the appearance and fast development of huge technology-based communication networks such as the Internet, as well as their inherent complex structure and dynamics, has contributed to awaken the interest of the scien-

[1] http://www.bayer.com/en/bayer/bayerwelt.php#top
[2] http://www.basf.de/en/corporate/overview/mitarbeiter/
[3] http://www.dupont.com/corp/overview/glance/index.html
[4] http://www.dow.com/about/aboutdow/about.htm

tific community in the so-called "complex networks". Actually, the study of networks was already a topic by itself in social sciences and in mathematics. However, recent studies on these technology-based communication networks as well as the discovery of surprising properties in big and complex networks in fields as diverse as biology, physics, computer science, engineering or economics, has generated a great interest. In particular, statistical physics has played a particularly important role in understanding some of the properties of such networks. The reason is that some of the tools derived to understand complex collective behavior in physical systems (that differ from the addition of the individual behaviors of the parts of the system) are applicable in the field of complex networks.

The present work uses ideas from both the economics literature and complex networks literature to understand the role of communication processes in organizations. This chapter presents the most important developments in the two lines, communication and information processing in organizations on the one hand and complex networks on the other, as well as the scope and main objectives of the work. Section 1 discusses the main ideas, developed mostly during the last ten years in the economics literature, about communication and decentralized information processing in organizations. Section 2 explains how, from the study of social networks, such as friendship networks, and technological networks, such as the Internet and the Worldwide Web, a "science of networks" has born and developed rapidly finding applications in physics, biology, etc. Finally, section 3 presents the scope and main objectives of this work.

1. Communication and information processing in organizations

To a large extend, organizations exist to solve coordination problems, since each individual or group within an organization is unable to acquire unbounded knowledge and process an unlimited amount of information (Van Zandt, 1998). Specialization thus arises as a consequence of this limitation. In turn, the need for coordination and communication between individuals and groups arises as a consequence of specialization. The internal organization of firms is designed to minimize the costs of communication and to optimize information sharing and processing. In this sense, it has been established that the multi-divisional corporation, which is the most frequent form of organization in modern companies, is suited to handle a continuously increasing flow of information (Chandler, 1966, Chandler, 1990). It has also been shown that, as firms grow, more managers are hired and information processing is decentralized (Van Zandt, 1998). Moreover, according to Radner (Radner, 1993):

> The typical U.S. company is so large that a substantial part of its work force is devoted to information-processing, rather than to "making" or "selling" things in the narrow sense. Although precise definitions and data are not available, a reasonable estimate is that

more than one-half of U.S. workers (including managers) do information-processing as their primary activity.

From this perspective, it seems natural to develop a theory of decentralized information processing in organizations, in which decentralization arises due to information processing constrains or costs. However, most studies of organizations in the economic literature have focused on incentive problems and not on communication problems, in part because there exists a well established theoretical apparatus to analyze these issues. In those studies, the point is to allocate incentives in such a way that the global outcome of the organizations is maximized. The incentive approach cannot explain, among other things, the interaction between communication technology and organizational structure. For example, there seems to be some empirical evidence that computerization of companies tends to reduce the number of organizational layers (Brynjolfsson et al., 1989, Hangstrom, 1991).

Considering organizations in a broad sense as "collections of individuals or groups whose actions are coordinated and for which there is some criterion for evaluating the collective outcome" (Van Zandt, 1998), one can regard markets, for example, as organizations. Then, it is possible to retrace the study of communication, information processing, and centralization/decentralization to early debates about economic planning under socialist systems. However, studies about information processing in organizations considered in a narrower sense as firms or companies have appeared only in the last 30 years, and the most significant developments started in the 1990's with the work by Radner and van Zandt (Radner and van Zandt, 1992, Radner, 1993). These developments are presented in the following sections. First, we show how the idea of communication and information processing was introduced in the theory of organizations. Next, we turn to more recent developments that have stressed the key role of specialization and knowledge organization in relation to decentralized information processing.

1.1 Decentralized information processing

Probably, the most influential work in the area of information processing in organizations is the one by Radner (Radner, 1993). In the paper, the author presents a stylized model of an organization that performs a decentralized associative task, involving communication and information processing. The model can describe a number of real situations including accounting and control (linear operation), project selection, and pattern recognition.

The model is as follows. Consider a set of P processors (employees, for example) and a set of N data items, usually called a *cohort*, that need to be summed by these processors. Each processor has an *in-box* and a *register*, and is connected to a number of other processors by directed communication lines, such that information can flow only one way. At each *time step*, any processor

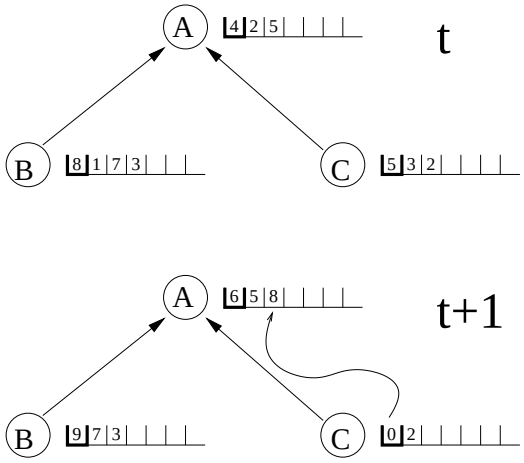


Figure 1.1. A sample time step in Radner's model for a network with three nodes. At time t, processor A has a value 4 stored in its register (bold box) and two more items waiting to be processed in its in-box. Similarly, nodes B and C have values 8 and 5 in their respective registers and items waiting in their in-boxes. In the next time step, all the nodes have added the first value of the in-box to the register. In addition, node C has sent the content of its register to its superior A, and has reset its register to 0.

can take one item from its in-box and add it to its register. Therefore, the in-box is a sort of *queue* where data items wait until they are processed, and the register contains the result of the processing. Additionally, at each time step processors can send the contents of their register to any of their neighbors and reset their register to zero, with no extra time cost. Such a time step is illustrated in figure 1.1. The basic design problem consists in finding the network of processors that can perform the sum of the N items with minimum delay. In addition, the *designer* has to specify the times of communication and the way the N items are distributed in the in-boxes of the processors at the beginning of the process: a network with this extra information is called a *programmed network*.

Next, Radner defined *efficient* programmed networks and turned to the problem of finding them. A programmed network is efficient for a given number of items, N, if the number of processors cannot be decreased without increasing the delay. In other words, a network is inefficient if it is possible to find a different network with the same number of processors that performs the summation of the N items with a smaller delay.

Finding exact efficient networks was out of the scope of the paper. However, the author was able to find lower bounds for the delay and to show that a certain type of hierarchies are close to this lower bounds. An example of such a *quasi-efficient* hierarchy is shown in figure 1.2. Moreover, the paper discussed that

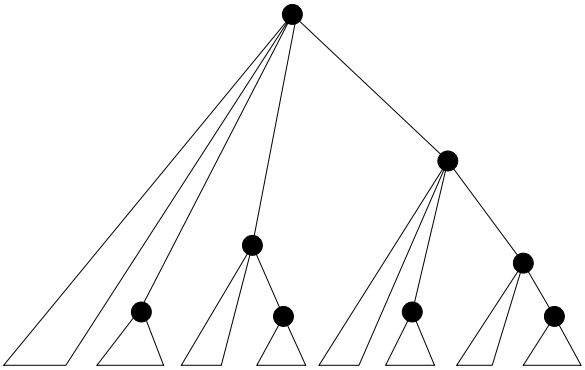


Figure 1.2. Hierarchical network that processes one cohort of items quasi-efficiently in Radner's model.

in presence of a continuous flow of items to be added, balanced hierarchies are also quasi-efficient.

Although the model can seem very abstract, it can account for a number of real situations. Consider, for example, the task of selecting the best project in a collection of N proposals. At each time step an employee can evaluate one project and compare it to her register. After the comparison, the employee will keep in the register the best of them and will discard the other. At certain times, the employee will send to another employee (probably a superior) the best project she has found so far. Similarly, the superior will compare this project, which has already been filtered by the first employee and therefore is in general one of the *good* projects, to the best one she has been able to find up to that moment.

In spite of the importance of the work by Radner, there are several criticisms that can be done. First, the model does not include specialization which, as discussed, is one of the key ingredients to understand the need for communication in companies. Second, quasi-efficient networks (as the one shown in figure 1.2) have a number of unrealistic features. For example, "skip-level reporting" is a common practice in real organizations, but it is hard to believe that employees at the lowest level can communicate directly with employees at the top level. Moreover, hierarchical structures are postulated *ad hoc* and there is no hint about the behavior of other arbitrary networks. Third, although the model can account for some real situations, it is not clear that it is a good metaphor for the global functioning of an organization. Some of these points were addressed latter in the literature.

1.2 Returns to specialization

The work by Bolton and Dewatripont addressed the issue of specialization (Bolton and Dewatripont, 1994). The authors recognized the contribution of Radner and van Zandt but stressed the trade-off between specialization and communication.

Although the formal expression of the model is slightly different from that of Radner, the main ideas are the same. The organization is immerse in an environment that makes information available at each time t.[5] Information arrives, again, in form of cohorts, that is groups of N data items. Each cohort has the same informational content, and all the N items must be processed so that the organization obtains some benefit.

Processing of items is costly: it takes a time τ to process each item. Communication is also costly although, as in the case of Radner, what takes time is to *read* the information that has been received and not to send it. Indeed, as described in the previous section, in Radner's work any processor could send the information in its register to another processor to whom it was connected without any additional cost (Radner, 1993). However, once the information was sent, it was stored in the in-box of the receiver and took always one time step to process it. In the work by Bolton and Dewatripont, there is also a cost C for reading a *report* containing aggregated information sent by a processor i, but the difficulty of reading such information depends on the amount of raw items that the report contains, n_i:

$$C(n_i) = \tau(\lambda + a\, n_i)\ , \tag{1.1}$$

where λ and a are parameters. In Radner's paper, $\lambda = 1$ since the cost of processing a report received from another processor is the same as processing a single raw item, and $a = 0$ since the cost is independent of the amount of processed items contained.

Except for this little generalization, Bolton and Dewatripont's model is equivalent, so far, to Radner's. Also as in reference (Radner, 1993), the design problem consists in finding the structure that minimizes the delay in processing cohorts. With these conditions, the authors show a number of important properties of efficient networks. First, assuming that the communication network is organized in layers and that reports (information) flow from the bottom to the top, they demonstrate that delegation of processing tasks from the top to the bottom only occurs when processors in the top are overloaded. In other words, agents do not delegate unless they cannot process the information themselves, in such a way that the number of agents through which a given item transits

[5]Note that in Radner's work, data items arrived only at certain time steps separated of one another, while here information is always available and the organization decides when to use it.

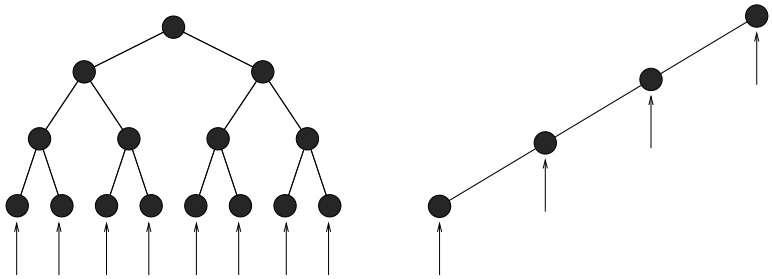


Figure 1.3. Efficient networks in presence of specialization considerations in Bolton and Dewatripont's model. The arrows indicate the introduction of raw data items. (a) Regular pyramidal network, efficient when agents are specialized in either processing or aggregating. (b) Conveyor belt network, efficient when it is better to be involved in both processing and aggregation.

is minimized. Even more significantly, they show that efficient networks are essentially pyramidal because each agent only reports to one superior.

But the main contribution of Bolton and Dewatripont is probably that they introduce the concept of specialization. Consider that cohorts are processed with a frequency x. They assume that the processing time τ is a decreasing function of x: $\tau = \tau(x)$ and

$$\frac{d\tau}{dx} < 0 \,, \tag{1.2}$$

so that the more an agent processes a certain type of information, the faster she is able to do it and the higher the payoff of the organization. Considering specialization, the paper shows that regular pyramidal networks (figure 1.3a) are efficient when agents are specialized in either processing or aggregating, and "conveyor belt" networks (that resemble assembly lines, as depicted in figure 1.3b) are efficient when it is better to be involved both in processing and aggregation.

1.3 Problem solving and organization of knowledge

Although the work by Bolton and Dewatripont represents an important step toward the understanding of the trade-off between specialization and communication, they simply equate specialization to a higher network throughput and do not consider explicitly task heterogeneity, that is different types of problem to solve. Garicano discusses that "if communication is available, workers do not need to acquire all the knowledge necessary to produce", and therefore proposes a model where distribution and organization of knowledge plays a fundamental role (Garicano, 2000).

Consider an organization formed by employees that are partitioned into L groups of different sizes $\beta_i, i = 1,\ 2, \ldots,\ L$. The organization is immerse in an

environment in which different problems arise. Each problem z appears with a certain probability, given by a distribution $F(z)$. Beyond its size, each group i is characterized by the set of problems it is able to solve A_i, an ordered list, l_i, that specifies the groups to which i can ask when facing an unknown problem, and the fraction of time that the employees in the group dedicate to produce, t_i^p (the rest of the time, $t_i^h = 1 - t_i^p$, is dedicated to solve problems arising in other groups). The output per capita is then:

$$Y = \sum_{i=1}^{L} \left[\beta_i t_i^p F \left(\bigcup_{k \in l_i} A_k \right) - c\beta_i \mu(A_i) \right] \tag{1.3}$$

where $F(\cdot)$ is the probability that a problem will find its solution in the list of i and therefore the probability that a problem will be solved, $\mu(A_i)$ is a measure of the size of A_i and c is the cost of knowledge. One can imagine that the organization obtains a benefit every time that a problem arises and is solved, but has to pay a cost to provide knowledge to employees (large sets A_i are more expensive than small ones).

Moreover, as in the case of Radner and Bolton and Dewatripont, to process information is costly in terms of time. In Garicano's model, this is included by assigning a time to solve problems:

$$t_i^h = \left(\sum_{k:i \in l_k} h\, \beta_k t_k^p \left[1 - F \left(\bigcup_{m <_k i} A_m \right) \right] \right) / \beta_i \tag{1.4}$$

where h is the helping cost and $F(\cdot)$ is now the probability that the problem is solved before i is asked by k (thus, $1 - F(\cdot)$ is the probability that i has to ask k). In words, the members of a group spend time when they are consulted by another agent, no matter if they have the solution or not.

With these ingredients, Garicano draw important conclusions about efficient networks, defined now as those that yield a maximum output. First, it is shown that workers specialize either in production or in solving problems and that only one class of workers specializes in production, while other classes specialize in solving different types of problems. Second, Garicano demonstrates that problem solvers learn how to solve the most exceptional problems. Finally, he shows that efficient organizations have a pyramidal structure with a lot of producers and successive smaller layers of problem solvers that know how to solve more and more rare problems.

2. Complex self-organized networks

Graphs, or in a less precise language *networks*, are mathematical objects that consist of a group of vertices or nodes, the agents, and a set of links that join vertices and represent the relations between them. Just as happens in social

networks (Wasserman and Faust, 1994), graphs are convenient representations of organizations because they are particularly suited to represent inter-individual relationships and to quantify important structural properties.

Social network analysis has a relatively short history. At the end of the World War II, Alex Bavelas founded the Group Networks Laboratory at M.I.T., although it was not until the 1970s—when modern discrete combinatorics, particularly graph theory, experienced rapid development and relatively powerful computers became readily available—that the study of social networks really began to take off as an interdisciplinary specialty (Wasserman and Faust, 1994).

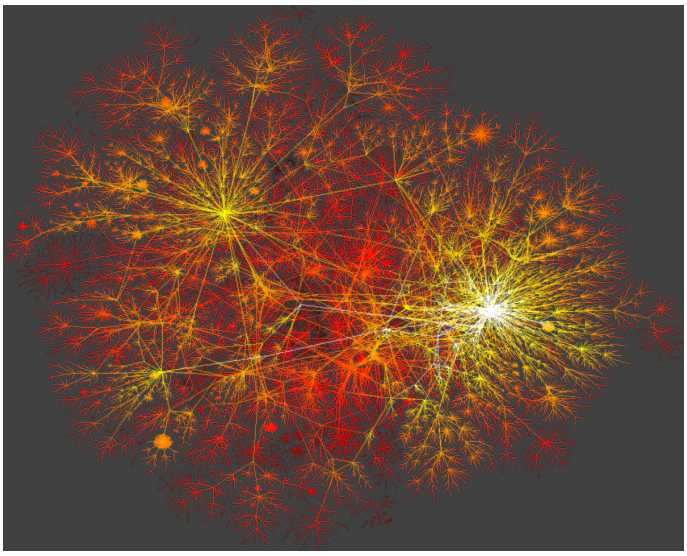
In 1998, a work by Watts and Strogatz, originally motivated by a problem on social networks, gave rise to a revolution in the scientific community. The idea was simple. The interaction between elements in many systems (not only social systems) gives rise to large and complex networks. Consider, for instance, computers and routers connected by means of physical or wireless connections in the Internet or chemicals in a cell connected by chemical reactions. For many years, the structure of such networks had been considered secondary in front of the interaction between the elements. As a consequence, simplistic approaches were used to model the topology of the interactions, although sophisticated models were used for the interactions themselves. Ordinarily, the connection topology was assumed to be either completely regular or completely random. The main contribution of the work by Watts and Strogatz was to show that real networks are neither regular nor random, as shown in figure 1.4.

Since 1998, the scientific community has realized that, in many situations, it can be completely wrong to assume such extreme cases of interaction topologies because an important part of the behavior of the system is due to the complex network of interactions and its non-trivial properties. Moreover, it has been shown that some non-trivial properties hold for networks that seem, in principle, completely unrelated: social, biological and technological networks share some similarities that are not reproduced by regular lattices or random graphs.

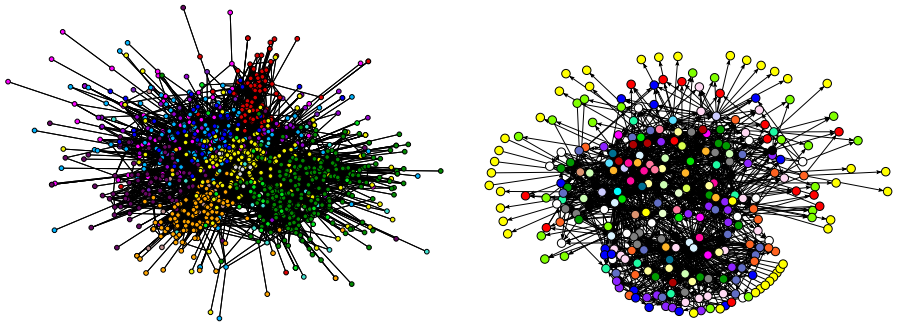
Although it might be pretentious to claim that nowadays a "science of networks" has been established, it is fair to recognize the relevance of this new "network approach" to problems in many different fields (Barabasi, 2002, Buchanan, 2002). It is true that there is not a solid and coherent theory of network structure, but there is a lot of knowledge that can be classified in the following way:

- Empirical generalities that hold for networks arising in sociology, biology, physics, chemistry and engineering. For example high clustering coefficient and low average path length (Watts and Strogatz, 1998), scale free degree distribution (Barabasi and Albert, 1999), assortativity and disassortativity (Newman, 2002), etc.

- *Evocative models* (Willinger et al., 2002) that reproduce some of those properties, and provide understanding on the structure of the networks. For



(a)



(b) (c)

Figure 1.4. Three examples of complex self-organized networks: (a) the Internet; (b) the e-mail network of the Universitat Rovira i Virgili, in Tarragona; and (c) the neural network of the worm *C. elegans*.

example the *small world* model of Watts and Strogatz (Watts and Strogatz, 1998), the scale free model of Barabasi and Albert (Barabasi and Albert, 1999), etc.

- Mathematical and computational tools that allow analytical analysis of both the models and the data. For example the renormalization group (Newman and Watts, 1999), the rate equation formalism (Dorogovtsev et al., 2000, Krapivsky et al., 2000), the generating function formalism (Newman et al., 2001), etc.

In the following, the main developments in this field are presented.

2.1 Random graphs

The probabilistic treatment of random graphs was introduced by Erdos and Renyi (Erdos and Renyi, 1959), as a counterpart of the enumeration and deterministic approach taken by other authors (Gilbert, 1956, Ford and Uhlenbeck, 1957, Austin et al., 1959). Within the probabilistic approach, the interest is in approximating a variety of exact values using probabilistic ideas rather than obtaining exact formulas, which are usually very complicated.

To use probabilistic ideas, it is usual to consider a probability space consisting of graphs of a particular type and a *typical* graph in this space. The simplest such probability space consists of all graphs with a given set of N nodes and M links, and each such graph is assigned the same probability. Another simple family of graphs is defined as follows: take a set of N nodes and for each of the $N(N-1)/2$ possible pairs of nodes establish a link with probability p. For the purpose of the present work, both families are equivalent when $N \to \infty$. Graphs constructed according to this procedures (in other words, graphs randomly selected in these probability spaces) will be called Erdos-Renyi (ER) random graphs or simply random graphs. Next, we discuss some of the properties of ER random graphs.

First, we focus on the so called "degree distribution". The degree of a node is simply the number of links of that node, i.e. its connectivity. Consider, again, the creation procedure: for each pair of nodes, a link is established with constant probability p. Therefore, the probability $f(k)$ of a node having degree k is given by a binomial distribution of parameters p and $N-1$. When the size of the network is large, the binomial can be approximated by a Poisson distribution:

$$f(k) = \frac{e^{-\lambda}\lambda^k}{k!} \tag{1.5}$$

with mean $\lambda = \langle k \rangle = pN$. For large values of k, $f(k)$ decays as a Gaussian function, meaning that degrees that significantly deviate from the average, λ, are extremely rare.

Next, we focus on the average distance between nodes. It has been shown that the diameter of a graph, δ_{ran}, (i.e. the maximum distance between pairs of nodes) is given by (Bollobas, 2001)

$$\delta_{ran} = \frac{\ln N}{\ln\langle k \rangle}. \tag{1.6}$$

Similarly, one expects that the average distance between nodes (or average path length), d, is given by the same scaling

$$d_{ran} \propto \frac{\ln N}{\ln\langle k \rangle}, \tag{1.7}$$

that is, that although the precise values of the average distance and the diameter are not the same, their dependence on the main parameters of the network will indeed coincide. It is worth noting that, fixed $\langle k \rangle$, the increase of the average distance (or diameter) with the size of the system is very slow. For example, for a network with $N = 10^6$ nodes, the average distance is only double than in a network of size $N = 10^3$.

Finally, we consider the so-called "clustering coefficient", that measures the transitivity of the links: if A is connected to S and to T, the clustering coefficient, C, is the probability that S is also connected to T (Wasserman and Faust, 1994, Watts and Strogatz, 1998). Consider a node i having degree k_i. The maximum number of links between its neighbors is $k_i(k_i - 1)/2$ and the clustering coefficient of that node is

$$C_i = \frac{2E_i}{k_i(k_i - 1)} \tag{1.8}$$

where E_i is the actual number of links between neighbors. The clustering coefficient of the network C is the average of all individual C_i's. In a random graph, since links are established independently with probability p, $C_i = p$ and

$$C_{ran} = \frac{\langle k \rangle}{N}. \tag{1.9}$$

Therefore, for relatively large networks the clustering coefficient becomes very small.

2.2 Regular lattices

Regular lattices are much simpler than random graphs, and one can easily compute the quantities that we have discussed in the previous section: degree distribution, average path length and clustering coefficient.

Consider the 1-dimensional (1D) and 2-dimensional (2D) lattices depicted in figure 1.5, where nodes have been linked to first and second nearest neighbors. Clearly, all the nodes have the same connectivity (at least when periodic boundary conditions are considered or when the size of the system is large enough so that boundary effects can be disregarded). Therefore the degree distribution is 1 for a given value of k and 0 otherwise.

Regarding the average path length, it is known that

$$d_{reg} \propto N^{1/D} \tag{1.10}$$

where N is, again, the number of nodes and D is the dimension of the embedding space. Compared to the case of the random graph, this potential growth is very fast. Recalling the example of the previous section, for a network with $N = 10^6$ nodes, the average distance would be 1000 times larger than in a network of size $N = 10^3$ in the 1D case, and more than 30 times larger in 2D.

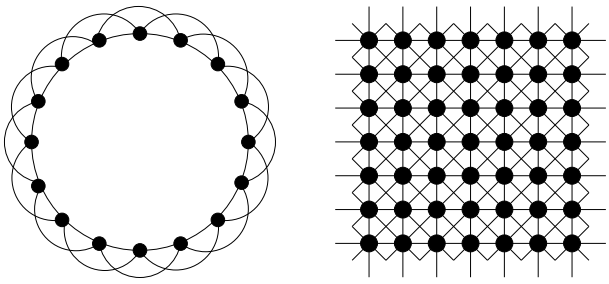


Figure 1.5. Low dimensional regular lattices with nodes connected to first and second nearest neighbors: (a) n-dimensional lattice and (b) two-dimensional lattice

Finally we consider the clustering coefficient. Since a high fraction of neighbors of a particular node are connected to each other, C will be large, that is close to 1. Moreover, for large enough systems C will not depend on N and therefore we have

$$C_{reg} \sim 1. \tag{1.11}$$

Again, this behavior is opposite to what happens in random graphs.

2.3 Small-world behavior

As already explained, before the seminal paper of Watts and Strogatz in 1998 (Watts and Strogatz, 1998), most of the systems were modeled either as completely regular or completely random. One of the main findings of this paper was to show that many real networks have properties of random graphs and properties of regular low dimensional lattices. In particular, real networks usually show average path lengths similar to those of a random graph, but clustering coefficients that are much larger than expected for this sort of networks. This behavior is known since then as "small world" behavior. Table 1.1 shows the average path length and clustering for some networks reported in the literature and the corresponding comparison with the values expected for a completely random graph.

Beyond realizing that, even for simple properties such as the average path length or the clustering coefficient, simple network models fail to reproduce real world networks, Watts and Strogatz hypothesized a mechanism leading to the small world phenomenon. The idea of their model is very simple. To fix ideas, consider a social network. In a social system, it is plausible that individuals lay in a low dimensional *social space* where nodes are connected to nearest neighbors. For instance, people living in the same town are more likely to be connected to one another than to other individuals living in a different town; people having a degree in philology are more likely to be connected to one another than to people with a degree in biology; etc. However, there is a certain

Net	Size	$\langle k \rangle$	d	d_{ran}	C	C_{ran}	Reference
1	225 226	61	3.65	2.99	0.79	0.00027	(Watts and Strogatz, 1998)
2	52 909	9.7	5.9	4.79	0.43	0.00018	(Newman, 2001b)
3	1 520 251	18.1	4.6	4.91	0.066	0.000011	(Newman, 2001b)
4	315	28.3	2.62	1.98	0.59	0.09	(Wagner and Fell, 2000)
5	460 902	70.13	2.67	3.03	0.437	0.0001	(Ferrer and Sole, 2001)
6	4941	2.67	18.7	12.4	0.08	0.005	(Watts and Strogatz, 1998)
7	282	14	2.65	2.25	0.28	0.05	(Watts and Strogatz, 1998)
8	153 127	35.21	3.1	3.35	0.1078	0.00023	(Adamic, 1999)

Table 1.1. Values of the size, average connectivity, average path length and clustering coefficient for various real world networks, and comparison with the corresponding values of a random graph. 1. Movie actors; 2. LANL co-authorship; 3. MEDLINE co-authorship; 4. *E. coli* reaction graph; 5. Words coocurrence; 6. Power grid of the USA; 7. Neural network of the *C. elegans*; 8. World-Wide Web.

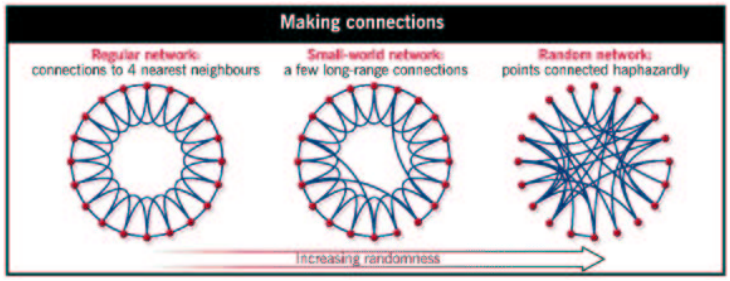


Figure 1.6. The small world model of Watts and Strogatz. Starting from a low dimensional regular lattice (left), some links are randomly rewired. When the fraction of rewired links is small (center) the network still has the low dimensional structure and a high clustering coefficient, but the rewired links act as shortcuts reducing the average distance between nodes . When the fraction of rewired links is high (right) the graph is completely random.

probability that a person knows another person living at 10000km from his or her hometown, or of a biologist knowing a philologist, for example. This links can be seen as long range links in the social space. Therefore, roughly speaking, a social network would look like a combination of a low dimensional (more or less) regular social space plus some long range links. Watts and Strogatz put these two ingredients in a very stylized model (figure 1.6). Nodes (individuals, chemicals, computers, etc.) are assumed to be located initially in a regular low dimensional space, say 1D. Then, with a certain probability p, each link is rewired at random, in such a way that for $p = 0$ we recover the regular low dimensional lattice, and for $p = 1$ one obtains a completely random network. For small values of p such that $pN \sim 1$ (in 1D), the clustering coefficient is

still high because most of the links are still established with neighbors in the low dimensional lattice (figure 1.6). However, some random long range links acting as shortcuts start to appear and the average path length between nodes decreases dramatically, reaching a regime where d scales as $\ln N$ as in a random graph.

2.4 Degree distribution: scale free networks and growth constrains

The paper by Watts and Strogatz (WS) showed the inability of traditional models to capture the complexity of real world networks and proposed a conceptual framework to understand the small world behavior observed in fields as diverse as biology, sociology and engineering. However, the model is too simple to capture all the complexity observed in real networks.

The first important shortcoming of the model was found in the degree distribution. As in the case of ER random graphs, WS graphs show a degree distribution that decays very fast (as a Gaussian function) for values k of the degree larger than $\langle k \rangle$. Surprisingly, many real world networks show a highly skewed degree distribution, usually with power law tails

$$p(k) \propto k^{-\alpha}, \tag{1.12}$$

as shown in figure A.3. This fact indicates that high degree nodes, with k being orders of magnitude larger than $\langle k \rangle$, are present in the network. In other words, real networks are much less homogeneous than expected from all models presented so far (regular, ER and WS).

The power law dependence of the degree distribution is certainly surprising, and networks with such a degree distribution have been called scale-free networks. In physical sciences, the appearance of power laws is usually related to the proximity of a critical point with the corresponding emerging scale invariance (Stanley, 1987). However, there are many other mechanisms that have nothing to do with criticality and that can lead to power law distributions. Significantly, multiplicative processes have been proposed as the underlying mechanism generating power law distributions in many systems in biology, sociology and economy (Ijiri and Simon, 1977). In a multiplicative process, the *size* of the variables is incremented by an amount which is proportional to this size. Barabasi and Albert proposed a model to explain the ubiquity of scale-free networks (Barabasi and Albert, 1999) that is, indeed, a multiplicative model. They identified two main ingredients in the formation of real complex networks: growth and preferential linking. First, they realized that many networks (and particularly the Internet) are subject to a continuous growth process by the introduction of new nodes and links. Second, they postulated that when new links are established, they are often directed to those nodes that are already highly connected: in communication networks, establishing links with highly

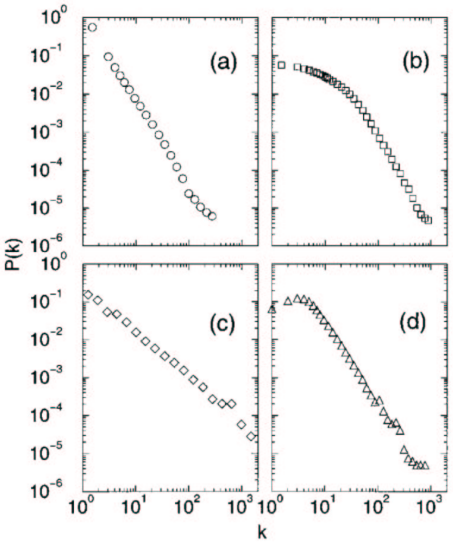


Figure 1.7. Scale free degree distribution of several real networks. Note that in a log-log scale, a power law becomes a straight line whose slope is the exponent of the power law. (a) Internet. (b) Movie actor collaboration network. (c) Co-authorship network of high-energy physicists. (d) Co-authorship network of neuroscientists. The figure has been taken from (Barabasi, 2002), publicly available at http://xxx.arxiv.org/abs/cond-mat/0106096

connected nodes can be good because they act as hubs; in social networks, highly connected individuals can be though as having an important social role and it can be useful to connect to them; etc. The Barabasi-Albert (BA) model combines these two ingredients in the following way. Consider an initial small set of nodes connected to each other. Then, at each time step, add a new node to the network. When a new node is added, it establishes m links with already existing nodes, in such a way that nodes with high degree have a high probability of being linked. In particular, a node i with degree k_i has a probability of being linked, Π_i, given by

$$\Pi_i = \frac{k_i}{\sum_j k_j} \, . \tag{1.13}$$

It has been shown (Dorogovtsev et al., 2000, Krapivsky et al., 2000) that this process generates a network whose degrees are power law distributed

$$P(k) = \frac{2m(m+1)}{k(k+1)(k+2)} \sim k^{-3} \tag{1.14}$$

Moreover, the exponent $\alpha = 3$ is similar to the exponent observed in most scale-free networks. Many works have shown that slight modifications of the BA model yield a continuum of exponents from $\alpha = 2$ to $\alpha = \infty$.

After the discovery of scale-free networks, many systems were reported to show scale-free topology. However, not all real complex networks are scale-free. In some situations, preferential attachment simply does not work, but even when preferential attachment is the mechanism driving the formation and growth of the network there are usually restrictions that prevent the network from being scale-free. Amaral and coworkers (Amaral et al., 2000) showed that small-world networks can be classified in three groups according to their degree distribution: scale-free networks (with power law degree distributions), single scale networks (with exponential or Gaussian degree distributions) and broad-scale networks (with degree distributions that show a power law region followed by an exponential or Gaussian cutoff). Moreover, they discussed that the existence of costs or restrictions in the establishment of links is the responsible for the existence of a well defined scale in single-scale networks and for the truncation of the scale-free behavior in broad-scale networks. Consider, for example, the network of world airports, where every node corresponds to a city and two cities are connected by a link if there is at least one direct flight connecting them. Although for reasons of efficiency it is good to have a small number of hubs connecting all flights, because of space and time constrains it is impossible for an airport to become the hub of all companies in the world. These sort of limitations prevent many real networks from being scale-free.

2.5 Percolation theory: fragility and robustness of complex networks

The topological properties discussed so far have important consequences in the behavior of real complex networks. One of the most relevant consequences of the topology of the network is its fragility or robustness against the removal of some of its nodes or, in other words, its vulnerability.

The effect of node or link removal on regular lattices has been studied within percolation theory (Stauffer and Aharony, 1992). Consider, for example, a large square lattice where every node is connected to its four nearest neighbors. Then, start to remove nodes (an the corresponding links) randomly. When the fraction of removed nodes, f, is small, the network will still essentially be formed by one big cluster containing a finite fraction of the nodes in the network, the main component, although some small groups of nodes will become disconnected form the rest. Conversely, when f is very large all the nodes will belong to small clusters. The transition from the former state, in which there is a main component, to the last state, in which all nodes belong to small clusters, is called the "percolation transition", and occurs suddenly for a critical fraction f_c.[6] The

[6]Strictly, the transition is only sudden in the so-called thermodynamic limit, that is, when the size of the systems tends to infinity.

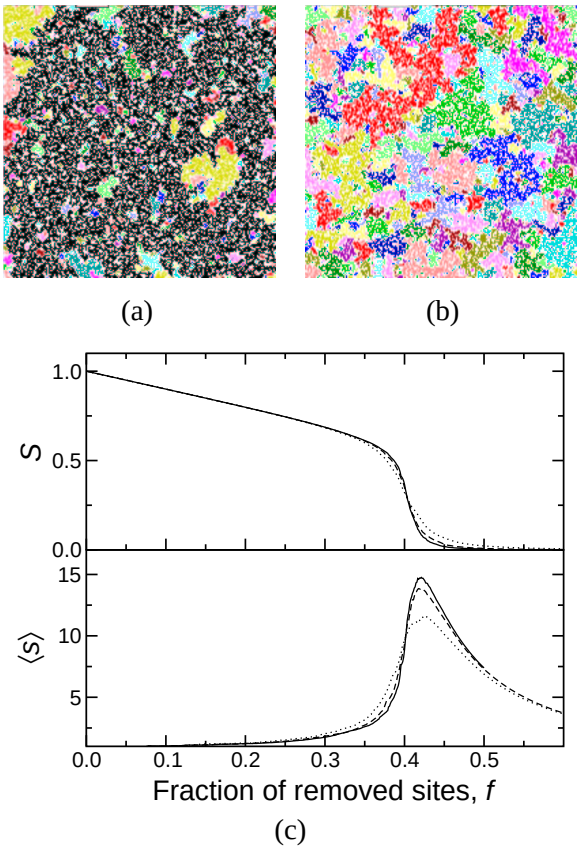


Figure 1.8. Percolation transition in a 2D lattice. Lattice below (a) and above (b) the percolation point f_c. Each cluster is plotted in a different color. Below the percolation point the main component comprises almost all nodes in the lattice, while above f_c only small clusters are left. (c) The top panel shows the behavior of the fraction of nodes belonging to the main component, S, as a function of the fraction of removed nodes, f. The bottom panel shows the average size of the remaining clusters, $\langle s \rangle$, as a function of f. Different lines correspond to different system sizes: 100×100 (dotted line), 200×200 (dashed line), and 300×300 (full line). As the system size grows, the order parameter, S, shows a sharp decay around $f_c = 0.4073$ and the susceptibility, $\langle s \rangle$, develops a peak around the same value.

transition is properly characterized defining two quantities: the fraction of nodes that belong to the largest cluster in the network, S, and the average size of all the other clusters, $\langle s \rangle$. The first, S, is close to 1 for small values of f, since most of the nodes belong to the main component. Conversely, for $f > f_c$, S is very close to 0, since all the clusters left are very small. Regarding $\langle s \rangle$, for $f \to 0$ the few nodes that do not belong to the main component are essentially isolated and $\langle s \rangle \approx 1$. Similarly, for $f \to 1$ all nodes are isolated and $\langle s \rangle \approx 1$. Near the critical point f_c, the main component is broken, some large clusters

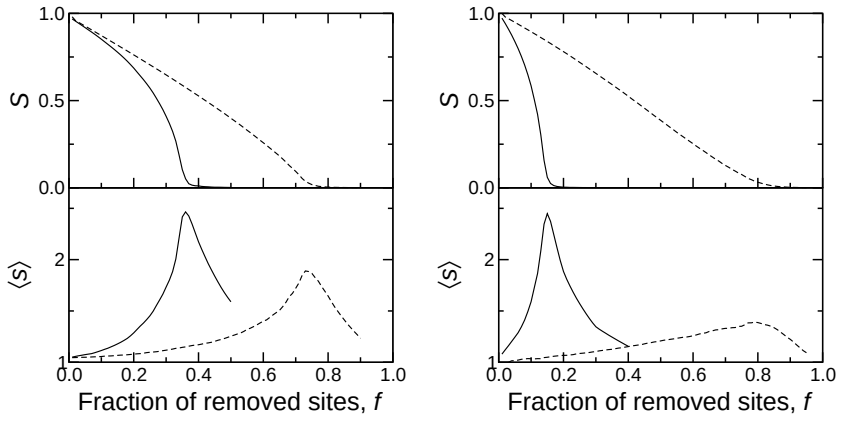


Figure 1.9. Effect of random removal (dashed line) and directed attack (full line) of nodes in: (a) ER random graphs and (b) BA scale-free networks. In both cases, the size of the network is $N = 10000$ and $\langle k \rangle = 4$. While the BA network is slightly more robust against random removal of nodes, it is significantly more sensitive to directed attacks of the most connected nodes.

are left, and $\langle s \rangle \gg 1$. Therefore, at f_c, S becomes 0 and $\langle s \rangle$ has a peak, as shown in figure 1.8. The percolation transition turns out to be a phase transition in which S and $\langle s \rangle$ play the role of the order parameter and the susceptibility respectively.[7]

Similarly, one can study percolation-related properties of complex networks, which will give an idea about their robustness against failure or intentional removal of nodes. Albert and coworkers (Albert et al., 2000) found important results. First, they showed that when nodes are removed randomly, BA scale-free networks are slightly more robust than ER random graphs, that is, for BA networks it is more difficult to break the main component (figure 1.9). Analytical results have confirmed the resilience of scale-free networks in front of random failure of the nodes (Cohen et al., 2000): when the exponent of the degree distribution is $\alpha \leq 3$ the percolation transition never occurs, meaning that for any value of f there is always a main component.

Conversely, the effect of "directed attacks" can be very destructive in scale-free networks. Consider now that, instead of removing nodes at random, one intentionally attacks those nodes with highest degree (Albert et al., 2000). As indicated by the highly skewed degree distribution, the most connected nodes in a scale-free network concentrate an important fraction of the total number of links. Therefore, the removal of these nodes will destroy the main component

[7] A classical book on phase transitions is (Stanley, 1987).

very fast (figure 1.9). This effect is less important in ER random graphs. There are also analytical results supporting these findings (Cohen et al., 2001).

3. Scope of the work

The aim of the present work is to obtain a deeper understanding about communication processes and apply it to the study, design and redesign of organizations. The problem of getting insights into communication processes is tackled from a double perspective: theoretical and empirical.

From the theoretical perspective, a model for communication and information processing is presented. In the model, task heterogeneity and specialization are present and communication arises due to the need for solving problems that are unknown. Although in this sense the model is closer to Garicano's, the precise formulation of the model is similar to Radner's. The organization (or, in general, the communication network) is immerse in an environment that generates "problems" that need to be solved by the organization. Therefore, individuals send and receive problems that travel through the network in form of packets or information items. Unlike all the models discussed in this introduction, the arrival of packets and the communication process itself are subject to randomness. Actually, Radner already stressed the potential importance of considering the effects of randomness (Radner, 1993):

> The decentralization of information typically implies that the amount of data in a cohort is subject to stochastic variation, a circumstance that gives rise to stochastic queuing in the network of processors.

The presence of stochasticity in the model makes the tools of statistical mechanics particularly useful. The model has been studied from this perspective (Arenas et al., 2001, Guimera et al., 2001a, Guimera et al., 2002a) and the results are presented in chapter 2.

Next, we move to the question of finding optimal network structures. Consider a situation in which agents do not have a complete knowledge of the structure of the network and therefore do not know exactly to whom they should address a packet once they receive it. This scenario is adequate in situations where employees are constantly subject to the arrival of new problems (new in the sense that they have not been faced before) or simply when being aware of the knowledge sets of each agent (using the Garicano's language) is a scarce result. In such a situation, the problem on network congestion as described before is superposed to the problem of local search and, in general, there is a trade-off between the first—that benefits from network decentralization—and the second—that requires centralization so that information is available through central nodes. Like the works in the economics literature, we focus on finding structures that minimize the average delivery time of the packets, but we proceed by exhaustive numerical calculation. First the search is restricted to particular families of networks (Guimera et al., 2001b) and finally we look for

global optimal networks without any restriction (Guimera et al., 2002c). This results are discussed in chapter 3.

In chapter 4 we take a complementary approach to unravel the role of communication and information processing in organizations. We present a large scale empirical analysis of the real communication network of an organization with about 1,700 employees (Guimera et al., 2002b). Using tools from complex networks theory it is possible extract information which is valuable from managerial and fundamental points of view. These results, as well as the techniques used to obtain them, are explained in chapter 4. Another real communication network is studied in appendix A (Guardiola et al., 2002).

Concluding remarks and perspectives for future work are given in the last chapter of the work.

Chapter 2

MODELING OF COMMUNICATION PROCESSES

As explained in the first chapter, it seems clear that the modeling of communication processes should provide valuable hints about which organizational structures are more adequate in a given environment. Although this has been understood by economists and there is a complete corpus of work in the economics literature devoted to address the relation between communication efficiency and organizational design (Radner, 1993, Bolton and Dewatripont, 1994, Van Zandt, 1998, Garicano, 2000), a stochastic *microscopic* model for general communication processes, describing how communication between individuals actually occurs, is still lacking. By *microscopic* we mean a model that is defined through a set of rules that prescribe the behavior of the *atoms*, in this case the communication agents and packets.

In the field of computer science and, in particular, of artificial intelligence, there has been an important effort to model organizations formed by complex intelligent agents from a microscopic perspective (Prietula et al., 1998). Agents act according to quite exhaustive sets of rules that determine how decisions are taken based on information inputs received from the other agents and also from the environment. Within this approach, however, it is difficult to deal with large numbers of agents due to the complexity of each one of these agents: the behavior of each agent depends on a large collection of variables. On the other hand, the great level of detail used in the definition of the agents makes the models and the conclusions drawn from them very specific and problem dependent.

The approach taken in this work lies somewhere between the abstract economics approach and the very detailed artificial intelligence approach. Although a *microscopic* model is proposed as in the artificial intelligence approach, the rules that the agents apply are intended to be as simple as possible to allow settings with large number of agents. Therefore, the interesting out-

put of the model is not the particular behavior of one node but the aggregated emergent behavior of the collective of agents. This emergent behavior can be understood using tools taken from statistical mechanics. Similar models exist for computer based communication networks such as the Internet (Tretyakov et al., 1998, Ohira and Sawatari, 1998, Sole and Valverde, 2001), but the models are restricted to the very particular behaviors of simple computer queues where packets are delivered according to very simple rules.

Section 1 presents the basic communication model in hierarchical networks and discusses its main properties. Section 2 is devoted to discuss various possible extension of the model.

1. Model for communication processes in hierarchical networks

The model assumes that agents are connected by communication channels. Agents and channels constitute the physical support for the communication process, which can be mapped onto a graph where nodes represent agents and links between nodes represent communication channels. This physical support remains fixed in time, although it is possible to compare different communication network structures in different simulations of the model. As a first step, this section considers only hierarchical networks which seem to be the basic structure underlying complex organizational systems, even when real organizations tend to more decentralized charts (Warnecke, 1993). Hierarchical networks have also been used in the economics literature to model organizations (Radner, 1993, Bolton and Dewatripont, 1994). Actually, as described latter in chapter 4, the structure of real complex organizations can be mapped onto graphs which are, in general, far from being strictly hierarchical. However, hierarchical networks provide a *zeroth order* approximation to real structures. To check the generality and limitations of the conclusions drawn for hierarchical networks, the model is extended to other topologies in the last section of the present chapter.

Beyond the physical support for the communication process, the information itself, the *object* of the communication process, is needed. The model considers that the information is formed by discrete packets that are sent from an origin node to a destination node. Each node can store as many information packets as needed. However, the capacity of nodes to deliver information cannot be infinite. In other words, any realistic model of communication must consider that delivering, for instance, two information packets takes more time than delivering just one packet. A particular example of this would be to assume that nodes are able to deliver one (or any constant number) information packet per time step independently of their load, as happens in the communication model by Radner (Radner, 1993) and in simple models of computer queues (Tretyakov et al., 1998, Ohira and Sawatari, 1998, Sole and Valverde, 2001),

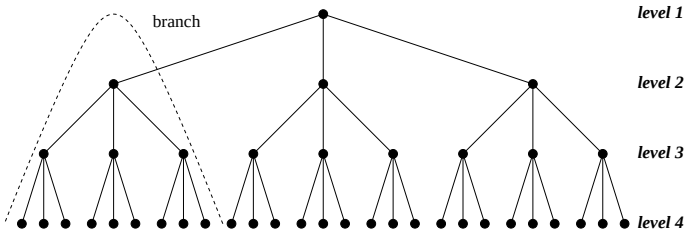


Figure 2.1. Typical hierarchical tree structure used for simulations and calculations: in particular, it is a tree $(3, 4)$. Dashed line: definition of branch, as used for some of the calculations.

but note that many alternative situations are possible. In the present model, each node has a certain *capability* that defines how long does it take on average to deliver a single packet and decreases as the number of accumulated packets at the node—or load—increases. This limitation in the capability of agents to deliver information can result in congestion of the network. Indeed, when the amount of information is too large, agents are not able to handle all the packets and some of them remain undelivered for extremely long periods of time. The maximum amount of information that a network can manage gives a measure of the adequacy of its organizational structure. In the study of the model, the interest is focused in both *when* the congestion occurs and *how* it occurs.

1.1 Description of the model

As explained in the previous section, the model has three basic components: (i) the physical support for the communication process—agents and communication channels—, (ii) the discrete information packets that are interchanged, and (iii) the limited capability of the agents to handle such packets.

The communication network is mapped onto a hierarchical Cayley tree as depicted in figure 2.1 where nodes mimic the communicating agents (employees in the company or, in other scenarios, routers and servers in a computer network, etc.) and the links between them represent communication lines. Cayley trees are identified by their branching z and their number of generations or levels m, and will be hereafter denoted by (z, m).

The dynamics of the model is as follows. At each time step t, an information packet is created at every node with probability ρ. Therefore ρ is the control parameter: small values of ρ correspond to low density of packets and high values of ρ correspond to high density of packets. When a new packet is created, a destination node, different from the origin, is chosen randomly in the network. Thus, during the following time steps $t + 1$, $t + 2, \ldots$, $t + T$, the packet *travels* toward its destination. Once the packet reaches the destination node, it is delivered and disappears from the network. Another interpretation is possible for this information transfer scenario. Packets can be regarded as

problems that arise at a certain ratio anywhere in the company. When one of such problems arises, it must be solved by an arbitrary agent of the network. Thus, in subsequent time steps the problem flows toward its *solution* until it is actually solved. This problem solving scenario can be considered a particular illustrative case of the more general information transfer scenario. The problem solving interpretation suggests a model similar to Garicano's (Garicano, 2000) in that there is task diversity and agents are specialized in solving only certain types of tasks.

The organization will be regarded as hierarchical not only from a bureaucratic point of view but also from a knowledge point of view. It is assumed in the model that agents have complete knowledge of the structure of the network in the subbranch below them. Therefore, when an agent receives a packet, he or she can evaluate whether the destination is to be found somewhere below. If so, the packet is sent in the right direction; otherwise, the agent sends the packet to his or her supervisor. Using this simple routing algorithm, the packets travel always following the shortest path between their origin and their destination.

The time that a packet remains in the network is related not only to the distance between the source and the target nodes, but also to the amount of packets in its path. Indeed, nodes with high loads—i.e. high quantities of accumulated packets—will need long times to deliver the packets or, in other words, it will take long times for packets to cross regions of the network that are highly congested. In particular, at each time step, all the packets move from their current position, i, to the next node in their path, j, with a probability q_{ij}. This probability q_{ij} is called the *quality of the channel* between i and j, and is defined in as

$$q_{ij} = \sqrt{k_i k_j}\,, \tag{2.1}$$

where k_i represents the *capability* of agent i and, in general, changes in time. Note that the capability of a node gives information about how each one of the individual packets accumulated at the node will be delivered. The quality of a channel is, thus, the geometric average of the capabilities of the two nodes involved, so that when one of the agents has capability 0, the channel is disabled. High qualities ($q_{ij} \approx 1$) imply that packets move easily while low qualities ($q_{ij} \approx 0$) imply that it takes a long time for a packet to jump from one node to the next. The algorithmic representation of the model is as follows:

1 Start with a network with no packets at time $t = 0$.

2 For each node in the network:

- Create a packet with probability ρ. When the packet is created, its destination is fixed at random.

3 For all the packets in the network:

- Determine the next node in the path and the quality of the corresponding channel, q_{ij}.
- Generate a random number $r \in [0,1]$. If $r < q_{ij}$, move the packet to the next node.
- If the packet has arrived to its destination, remove it from the network.

4 Increase time $t \leftarrow t+1$ and repeat from 2.

It is assumed that k_i depends only on the number of packets at node i, ν_i, through:

$$k_i = f(\nu_i) \tag{2.2}$$

The function $f(n)$ determines how the capability evolves when the number of packets at a given node changes. We propose a general form

$$f(\nu) = \begin{cases} 1 & \text{for} \quad \nu = 0 \\ \nu^{-\xi} & \text{for} \quad \nu = 1,2,3,\ldots \end{cases} \tag{2.3}$$

with $\xi \geq 0$. Equation (2.3) defines a complete collection of models with agents that behave qualitatively different depending on the exponent ξ.

The election of this functional forms for the quality of the channels and the capability of the nodes is arbitrary. Regarding the first, equation (2.1) is plausible for situations in which an effort is needed from both agents involved in the communication process. For instance, this is true if information is to be transmitted during a face to face meeting. If, on the contrary, information can be transmitted without the collaboration of the receiver, an equation of the form

$$q_{ij} = k_i \,, \tag{2.4}$$

would be more adequate. This would be the case, for instance, in an e-mail communication, where the receiver does not play an active role, and an arbitrary amount of e-mails can be received, but not sent, without any time cost. Equation (2.4) will be used for analytical understanding of the problem in chapter 3, but for the moment the interest is focused in (2.1). Most of the relevant features of the model, however, do not depend on this election.

Regarding the election of (2.3) for the capability of the nodes, it will be shown to give rise to a general enough phenomenology. Indeed, the average number of packets delivered during one time step by a node i to another node j, ν_{ij}, is

$$\nu_{ij} = \alpha \nu_i q_{ij} = \alpha \nu_i / (\nu_i^{\xi/2} \nu_j^{\xi/2}) \,, \tag{2.5}$$

where α represents the fraction of nodes at i that are trying to jump to j. Assuming that $\nu_i \propto \nu_j$ the former expression is proportional to $\nu_i^{1-\xi}$. The proportionality is exact, as shown later, not only in the hierarchical lattice

but also in other topologies. Therefore, for $\xi < 1$ the number of delivered packets increases with the number of accumulated packets. For $\xi > 1$ the number of delivered packets decreases as the number of accumulated packets increases. Finally, for the particular case $\xi = 1$, the number of delivered packets is independent of the number of accumulated packets. These three behaviors correspond to agents that react qualitatively different against load. The first case $\xi < 1$ correspond to agents that increase their efficiency as their load increase. Although it might be difficult to find such a situation in real communication environments, it is possible to justify it from a psychological point of view, considering that agents react to an increased pressure with a higher performance. On the other extreme, $\xi > 1$ represents agents that get *stressed* when the pressure increases, a situation which is also understandable from a psychological point of view. Finally, $\xi = 1$ correspond to agents that do not react against pressure and therefore their performance is unaltered. Note that this particular case is consistent with simple models of computer queues (Ohira and Sawatari, 1998), although the precise definition of the models may differ from ours.

1.2 Communication dynamics

It has been shown that the cases $\xi < 1$, $\xi = 1$ and $\xi > 1$ correspond to qualitatively different behaviors of the agents. As a consequence, each of them results in a completely different phenomenology.

For $\xi > 1$, the number of transmitted packets decreases as ν_i grows. For very small values of the probability of packet generation per node and time step, ρ, packets are generated only rarely and they can travel freely to their destination without encountering other packets in their path. Therefore all the packets are delivered and the average total number of packets in the network, $N = \sum_i \nu_i$, remains constant. For slightly higher values of ρ, all the packets are still delivered to their destination and, after a transient, the system reaches a steady state in which N fluctuates around a constant value. However, if we continue to increase ρ, at some point the total number of packets will be so large that the network will not be able to handle them, N will increase continuously and, at the end, no packets at all will be delivered to their destination. This state in which some packets are accumulated in the network at each time step is referred to as *congested*.

On the contrary, for $\xi < 1$, the number of transmitted packets grows with ν_i. Thus, the number of delivered packets increases with N until an equilibrium between generated and delivered packets is reached: at this point, N remains constant (except fluctuations).

In the case $\xi = 1$, the number of delivered packets is constant irrespective of the number of stored packets. This particular behavior is less obvious and will be treated accurately from the viewpoint of critical systems. The remaining of

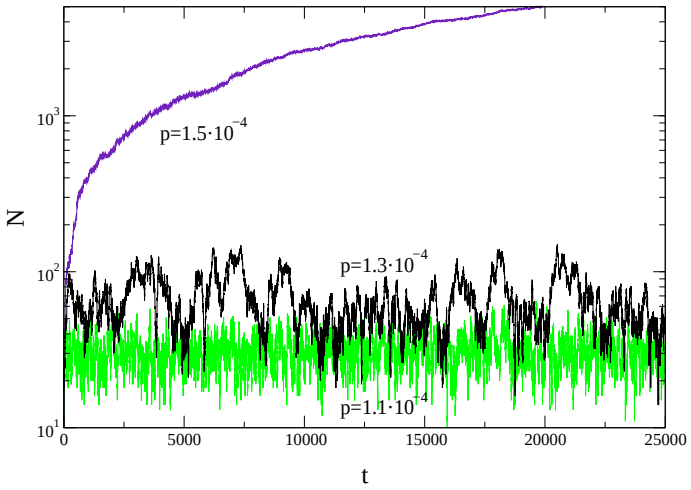


Figure 2.2. Evolution of the total number of packets, N, as a function of time for a (5,7) Cayley tree and different values of ρ, below the critical congestion point ($\rho = 1.1 \cdot 10^{-4} < \rho_c$), above the critical congestion point ($\rho = 1.5 \cdot 10^{-4} > \rho_c$), and close to the critical congestion point ($\rho = 1.3 \cdot 10^{-4} \approx \rho_c$). Note the logarithmic scale in the Y axis.

section 1 and most of section 2 assume $\xi = 1$, and the discussion of the other cases in some more detail is left for section 2.3.2.

1.2.1 Congestion and network capacity

Consider, thus, the case $\xi = 1$. Depending on the ratio of generation of packets ρ, two different behaviors are observed. When ρ is small, or in other words when the amount of flowing packets is small, the network is able to deliver, all the packets that are generated and, after a transient, the total load N of the network achieves a stationary state and fluctuates around a constant value. These fluctuations are indeed quite small. Conversely, when ρ is large enough the number of generated packets is larger than the number of packets that the network can manage to solve and the network enters a state of congestion. As already noted, in the case $\xi = 1$ the number of packets that a node can deliver at any time step is independent of its load. Therefore, when the number of packets that arrive to a node at each time step is, on average, larger than the number of packets it is able to deliver some packets are accumulated. Therefore, N does never reach the stationary state but grows indefinitely in time. The transition from the *free regime*, ρ small, to the *congested regime*, ρ large, occurs for a well defined value of ρ, that will be denoted by ρ_c. For values smaller than but close to ρ_c, the steady state is reached but large fluctuations arise. Moreover, the correlation times of these correlations become also enormous.

The three behaviors (free, congested and close to the transition) are depicted in figure 2.2. For $\rho < \rho_c$, the width of the fluctuations is small, indicating short characteristic times. This means, among other thinks, that the average time required to deliver a packet to the destination is small. It also means that time correlations are short, that is, the state of the network at one time step has little influence on the state of the network only a few time steps latter. As ρ approaches ρ_c, the fluctuations are wider and one can conclude that correlations become important. In other words, as one approaches ρ_c the time needed to deliver a packet grows and the state of the network at one instant is determinant for its state many time steps latter. In the congested regime, the amount of delivered packets is independent of the load and thus remains constant along time, while the number of generated packets is also constant, but larger that the amount of delivered packets. Thus, at each time step the number of accumulated packets is increased by a constant amount, and $N(t)$ grows linearly in time.

The transition from the free regime to the congested regime is therefore captured by the slope of $N(t)$ in the stationary state. When all the packets are delivered and there is no accumulation, the average slope is 0 while it is larger than 0 for $\rho > \rho_c$. We use this property to introduce an *order parameter*, η, that is able to characterize the transition from one regime to the other:

$$\eta(p) = \lim_{t\to\infty} \frac{1}{\rho S} \frac{\langle \Delta N \rangle}{\Delta t}, \tag{2.6}$$

In this equation $\Delta N = N(t+\Delta t) - N(t)$, $\langle \ldots \rangle$ indicates average over time windows of width Δt and S is the number of nodes in the system. These averages can be over one or many realizations, yielding the same result. Essentially, the order parameter represents the ratio between undelivered and generated packets calculated at long enough times such that $\Delta N \propto \Delta t$. Thus, η is only a function of the probability of packet generation per node and time step, ρ. For $\rho > \rho_c$, the system collapses, $\langle \Delta N \rangle$ grows linearly with Δt and thus η is a function of ρ only. For $\rho < \rho_c$, $\langle \Delta N \rangle = 0$ and $\eta = 0$. The behavior of the order parameter is shown in figure 2.3. Since the order parameter is continuous at ρ_c, the transition to congestion is a critical phenomena and ρ_c is a critical point as usually defined in statistical physics (Stanley, 1987).

Once the transition is characterized, the first issue that deserves attention is the location of the transition point ρ_c as a function of the parameters of the hierarchical lattice. This transition point gives information about the capacity of a given network. Indeed, the maximum number of packets that a network can handle per time step will be $N_c = S\rho_c$. Therefore, ρ_c is a measure of the amount of information an organization is able to handle and thus of the adequacy of a given organizational structure. One reasonable problem to put is, therefore, which is the network that maximizes ρ_c fixed a certain set of available resources (agents and links).

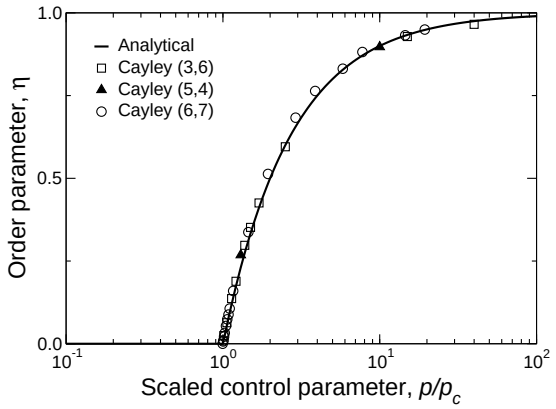


Figure 2.3. Behavior of the order parameter. The solid line corresponds to the analytical calculation for two nodes exchanging information packets (equation (2.11)). Symbols correspond to simulations performed on different Cayley trees.

It is possible to estimate numerically the value of ρ_c. One possibility would be to observe the plot $N(t)$ and determine when the slope in the steady state is different from 0. However, due to the existence of large fluctuations arising near the critical point, it is difficult to establish precisely the location of ρ_c. For a better estimation of the transition point, it is possible to take advantage of these fluctuations, the idea being that the fluctuations diverge at the critical point ρ_c. A susceptibility-like function $\chi(p)$ can be defined by analogy with equilibrium thermal critical phenomena (Stanley, 1987, Binder, 1987), and used to estimate more accurately the value of the critical probability of packet generation, ρ_c. The susceptibility χ is related to the fluctuations of the order parameter by

$$\chi(p) = \lim_{T\to\infty} T\sigma_\eta(T) \tag{2.7}$$

where T is the width of a time window, and $\sigma_\eta(T)$ is the standard deviation of the order parameter estimated from the analysis of many different time windows of width T. Thus a calculation implies a long realization of $N(t)$, its division into windows of width T, calculation of the average value of the order parameter in each window and finally the determination of the standard deviation of these values. As shown in figure 2.4, the susceptibility has a peak at ρ_c. This peak becomes sharper as T grows, as expected, and allows a quite precise determination of the transition point.

1.2.2 Analytical estimation of the transition point

As happens in other problems in statistical physics (Stauffer and Aharony, 1992), the particular symmetry of the hierarchical tree allows an analytical estimation of the critical point ρ_c. In particular, the approach taken here is

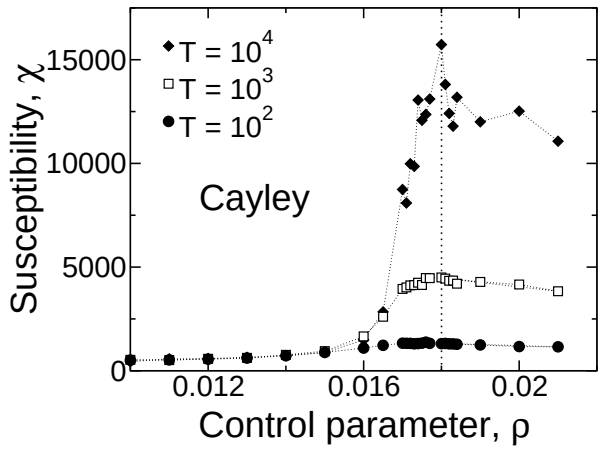


Figure 2.4. Susceptibility for a (5,4) Cayley tree, for different time windows T. The vertical dotted line corresponds to the mean field calculation of the critical point in equation (2.9).

mean field in the sense that fluctuations are disregarded and only average values are considered. Since in the steady state regime there is no accumulation of packets, the number of packets arriving to the top of the hierarchical structure (level 1) per time unit, ν_1^a, is, on average, equal to the number of packets that are created in one branch of the network and have their destination in a different branch (see figure 2.1). Since the origin and the destination of the packets are chosen at random, from purely geometric considerations it is straightforward to estimate this number of packets per unit time as:

$$\nu_1^a = \rho \left(\frac{z \left(z^{m-1} - 1 \right)^2}{z^m - 1} + 1 \right). \tag{2.8}$$

Within this mean field approach, it can be easily shown that it is indeed the top node which is the most congested.

On the other hand, in our mean field calculation q_{12} is the average probability that a given packet moves from one of the nodes in the second level to the top node and vice versa[1], and is given, as a first approximation, by $q_{12} = 1/\sqrt{\nu_1 \nu_2}$, where ν_1 is the average number of packets at level one and ν_2 is the average number of packets at each of the z nodes in the second level. Thus the average number of packets leaving the top at each time step will be $\nu_1^l = \nu_1 q_{12}$, and the average number of packets going from the z nodes in the second level to the top will be $\nu_1^a = z\alpha\nu_2 q_{12}$, where α stands for the fraction of packets in the second level that are trying to go up (some of the packets in level 2 are, of course, trying to go down to level 3).

[1]Note that, within the mean field approach, all the nodes at the second level are equivalent

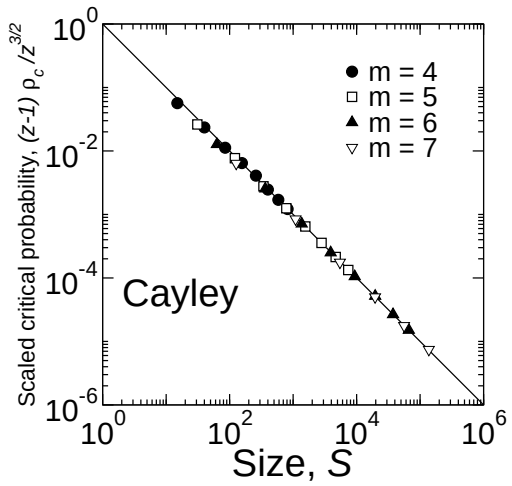


Figure 2.5. Comparison between analytical (lines) and numerical (symbols) values of ρ_c obtained for hierarchical trees. The error bars of the numerical points are smaller than the size of the symbols.

At the critical point the top agent becomes collapsed and the communications between the first and the second level are much more congested than the communications between levels 2 and 3 so one can assume that $\alpha \approx 1$. At this point, by imposing the steady state condition $\nu_1^a = \nu_1^l$ one arrives to the relations $\nu_1 = z\nu_2$ and $\nu_1^a = \sqrt{z}$. Using equation (2.8), the final expression for ρ_c is obtained:

$$\rho_c = \frac{\sqrt{z}}{\frac{z(z^{m-1}-1)^2}{z^m-1} + 1} \tag{2.9}$$

Although strictly speaking the condition $\alpha = 1$ provides an upper bound to ρ_c, equation (2.9) is an excellent approximation for $z \geq 3$, as depicted in figure 2.5.

The critical total number of generated packets, $N_c = \rho_c S$, with S standing for the size of the system, can be approximated, for large enough values of z and m such that $z^{m-1} \gg 1$, by

$$N_c = \frac{z^{3/2}}{z-1}, \tag{2.10}$$

which is independent of the number of levels in the tree. It suggests that the behavior of the top node is only affected by the total number of packets arriving from each node of the second level, which is consistent with the mean field hypothesis.

According to equation (2.10), the total number of packets a network can deal with, N_c, is a monotonically increasing function of z, suggesting that, fixed the

number of agents in the organization, S, the optimal organizational structure, understood as the structure with higher capacity to handle information, is the flattest one, with $m = 2$ and $z = S - 1$.

To understand this result it is necessary to take into account the following considerations:

- We are restricting our comparison only to different hierarchical networks and in any hierarchical network, the top node will receive most of the packets. Since origins and destinations are generated with uniform independent probabilities, roughly $(z - 1)/z$ of the packets will pass through the top node. Therefore, although it might seem that in a flat hierarchy the top node receives much higher amount of packets, this turns out to be false. For example, for a quite small value of the branching $z = 5$ the top node must already process 80% of packets.

- Still, it could seem that having small z is *slightly* better according to the previous consideration. However, it is important to note that, in the present model (in particular due to equation (2.1)), the loads of both the sender and the receiver are important to have a good communication quality. In a network with small z, the nodes in the second level have also a high load, while in a network with a high z the nodes in the second level are much less loaded. This effect is responsible of the observed behavior and the situation would change if the communication would only depend on the sender of the packet as considered in the next chapter.

- We have implicitly assumed that there is not any cost for an agent to have a large amount of communication channels active. Thus, for the top agent it is exactly the same to communicate with two agents every time step than to communicate with 20 agents every time step. In section 2.2 we extend the model to consider costly communication channels.

1.2.3 Analytical estimation of the order parameter

The behavior of the order parameter, which measures the ratio of accumulation of packets, is studied next. It is possible to derive an analytical expression for the simplest case where there are only two nodes that exchange packets. Since from symmetry considerations $\nu_1 = \nu_2$, the average number of packets eliminated in one time step is 2, while the number of generated packets is 2ρ. Thus $\rho_c = 1$ and with the present formulation of the model it is not possible to reach the super-critical congested regime. However, ρ can be extended to be the average number of generated packets per node at each step (instead of a probability) and in this case it can actually be as large as needed. As a result, the order parameter for the super-critical phase is $\eta = (\rho - 1)/\rho$. As observed

in figure 2.3, the general form

$$\eta(\rho/\rho_c) = \frac{\rho/\rho_c - 1}{\rho/\rho_c} \tag{2.11}$$

fits very accurately the behavior of the order parameter for any Cayley tree.

1.2.4 Power spectrum and characteristic time

Next we consider the behavior of the characteristic times of the system as a function of the probability of packet generation per node and time step ρ. By characteristic time we mean any time that is relevant in the behavior of the system, including the average time to deliver a packet, the correlation time, etc. From the theory of critical phenomena we know that any of these characteristic times will behave in a similar way (Stanley, 1987).

Beyond the inherent interest of the study on the mean delay to deliver a packet, for example, the understanding of the behavior of characteristic times is interesting as it is related to other key quantities like the total load of the network. Indeed, if τ is the average time needed to deliver a packet and $\overline{N}$ is the average load of the network,

$$\frac{\overline{N}}{\tau} = \rho S \tag{2.12}$$

from Little's law of queuing theory (Allen, 1990). This law states that, in steady state, the number of delivered packets and the number of generated packets are equal. The number of delivered packets is simply the total amount of packets $\overline{N}$ multiplied by the probability that a packet arrives to its destination in one time step, that is $1/\tau$.

Instead of studying directly the characteristic time, we consider its inverse the characteristic frequency, $f \sim 1/\tau$, by means of the power spectrum (i.e. the square of the modulus of the Fourier transform) of the temporal series $N(t)$. The analysis of the power spectrum shows that in the sub-critical regime, i.e. in the free phase, the spectrum is well fitted by a Lorentzian characterized by a frequency, f_c. This means that the spectrum is flat—there are no correlations—for frequencies smaller than f_c, that is, for long enough times. It is known that Lorentzian power spectra correspond to exponentially decaying correlations with a characteristic time τ. Figure 2.6 shows that, as ρ approaches ρ_c, $f_c \to 0$ and the power spectrum becomes $1/f^2$ for the whole range of frequencies. Alternatively the characteristic time diverges: $\tau \to \infty$, meaning that packets are essentially never delivered.

It is interesting to study how the characteristic frequency drops to 0 for each network topology. Near the critical point, one expects the scaling behavior (Stanley, 1987)

$$f_c \propto (\rho_c - \rho)^{\gamma}. \tag{2.13}$$

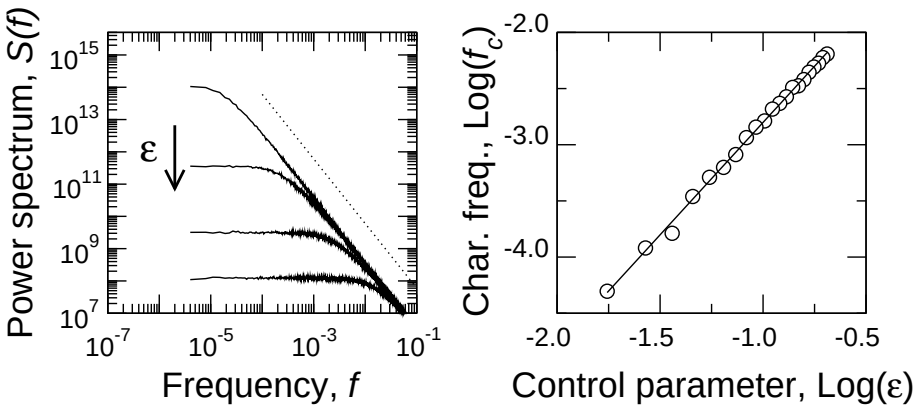


Figure 2.6. Left: Power spectrum of $N(t)$ for different values of the control parameter ϵ and a $(7, 5)$ Cayley tree. Power spectra have been obtained averaging over 100 realizations of $N(t)$. Dotted lines represent a power law with exponent -2. Right: Characteristic frequency as a function of the control parameter $\epsilon = (\rho_c - \rho)/\rho_c$. As observed, the characteristic frequency tends to 0 as $\rho \to \rho_c$ following a power law. The straight lines correspond to fittings of equation (2.13).

The value of the critical exponent γ can be estimated by fitting equation (2.13) to values of $f_c(p)$ close enough to the critical point, as shown in figure 2.6. Note that we fit ρ_c and γ simultaneously. This procedure yields very accurate values of ρ_c but the values of γ are subject to large errors. Figure 2.6 yields $\gamma \approx 2$ for a (7,5) Cayley tree.

The determination of γ is interesting not only from an academic point of view, but also from an engineering perspective. As explained, this exponent is related to divergences of other relevant quantities near the critical point. Any characteristic time τ—the average time to deliver a packet, for instance—will diverge as

$$\tau \propto (\rho_c - \rho)^{-\gamma} \tag{2.14}$$

and similarly the total number of packets

$$\overline{N} \propto (\rho_c - \rho)^{-\gamma} \tag{2.15}$$

The estimation of γ is particularly interesting in electronic communication protocols. Indeed, equation (2.14) is used to determine the waiting time before a packet is considered lost in the network and therefore sent again (Jacobson, 1988). In practice, the exponent $\gamma = 1$ predicted by classical queue theory (Allen, 1990) is assumed, while our current results suggest that more complex settings can lead to significantly different exponents.

2. Generalizations of the model

In the previous section we have introduced the simplest communication model and have studied its main dynamical properties. The model can be extended in many directions to account for a number of more realistic situations. In the following, we concentrate in some of these possible extensions and their consequences regarding organization design. Note that these extensions are independent of one another and they are only considered separately.

2.1 Agent heterogeneity

So far, it has been assumed that all agents in the network have the same capability. However, in real situations some agents are more efficient than others and, moreover, one individual work efficiently with some of his or her colleagues and inefficiently with others due to personal reasons. This effect can be taken into account in the model modifying equation (2.2) for the capability of one agent. Now we consider that the capability of node i to communicate with node j is

$$k_{ij} = \zeta_{ij} f(\nu_i), \tag{2.16}$$

where $0 \leq \zeta_{ij} \leq 1$ is a number that characterizes the communication line between i and j.

For simplicity, let us consider the case in which the variables ζ_{ij} take random values uniformly distributed in $[0, 1]$. A particular realization of these random variables will result in a network configuration that will behave as the case without *disorder*. For small values of ρ, the network will be in the free phase and beyond a certain critical point ρ_c, the network will transit to the congested phase. The main difference is that, now, due to the particularities of the pair communications, the network will not collapse at the top node in general. However, the top node will still support the heaviest traffic and, therefore, it will be optimum to place the node with highest capabilities in this position.

Now, we consider the *average behavior* of the system. Even for very small values of ρ, a particular realization of the disorder can provoke a very weak communication line and the congestion of the whole network. Thus there is no transition controlled by ρ. However, it is still possible to define the order parameter as in (2.6), just considering that the average $\langle \ldots \rangle$ has to be taken over time and over disorder realizations (figure 2.7).

Again, it is possible to obtain an analytical expression of the order parameter in the case of two nodes. As in the ordered case, the number of packets generated in a time step will be 2ρ. Now, however, for a particular realization, ζ_{12} and ζ_{21}, the maximum number of delivered packets will be $2\sqrt{\zeta_{12}\zeta_{21}}$. Thus, if $\zeta_{12}\zeta_{21} > \rho^2$ the system will reach the steady state and the configuration will not contribute to the order parameter, while if $\zeta_{12}\zeta_{21} < \rho^2$ the system will collapse and the contribution will be $\eta_{\zeta_{12}\zeta_{21}} = 1 - \sqrt{\zeta_{12}\zeta_{21}}/\rho$.

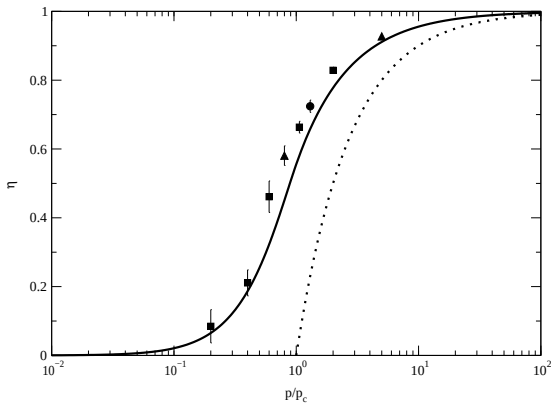


Figure 2.7. Order parameter in the case of agent heterogeneity. Symbols represent the same structures than in figure 2.3. The bold line corresponds to the analytic prediction of equation (2.19). The dotted line represents the critical behavior observed in the case without agent heterogeneity.

Thus we can define:

$$\eta(\rho, \zeta_{12}, \zeta_{21}) = \begin{cases} 0 & \text{for} \quad \zeta_{12}\zeta_{21} > \rho^2 \\ 1 - \sqrt{\zeta_{12}\zeta_{21}}/\rho & \text{for} \quad \zeta_{12}\zeta_{21} < \rho^2 \end{cases} \tag{2.17}$$

and the order parameter will be given by the average over the random variables:

$$\eta(\rho) = \int_0^1 d\zeta_{12} \int_0^1 d\zeta_{21} \eta(\rho, \zeta_{12}, \zeta_{21}). \tag{2.18}$$

It is straightforward to obtain the result:

$$\eta(\rho) = \begin{cases} 1 - 4/(9\rho) & \text{for} \quad \rho > 1 \\ \left(5\rho^2 - 3\rho^2 \ln \rho^2\right)/9 & \text{for} \quad \rho < 1 \end{cases} \tag{2.19}$$

As depicted in figure 2.7, there is reasonable agreement between this analytical expression and the points obtained by simulation, always keeping in mind the simplicity of our approach.

2.2 Costly communication channels

As it has been already discussed, in the basic model it turns out that the hierarchical network that is able to cope with a largest amount of packets is the flattest one with one top node and all the others directly connected to it. However, from a practical point of view this structure is not possible: an organization with 10,000 employees, for instance, cannot be organized in only two hierarchical levels, since it is impossible for the central node to maintain such a

enormous number of communication lines. Thus, it is necessary to introduce a cost for establishing links in order to get a more realistic picture of the problem.

Consider now a situation in which keeping communication channels open has a certain cost even when there is no information transfer. This can be introduced in the model modifying, again, equation (2.2) for the capability of one agent. Consider, in particular, that

$$k_i = Q_L(c_i) f(\nu_i), \tag{2.20}$$

where c_i is the number of links or degree of agent i, $0 < Q_L(c) \leq 1$ is a cost factor related to these links (note that, the higher the number of links, the smaller Q_L, so Q_L is a monotonically decreasing function of its argument), and L is the *linking capability* that tunes the magnitude of this cost (higher values of L correspond to low linking cost and vice versa).

In this case, following arguments analogous to those used in the case of costless connections, we can arrive to the following expression for ρ_c:

$$\rho_c = \frac{\sqrt{z Q_L(z) Q_L(z+1)}}{\frac{z(z^{m-1}-1)^2}{z^m-1} + 1}. \tag{2.21}$$

Again, for z and m such that $z^{m-1} \gg 1$, the maximum number of packets that can be generated per time step without collapsing the system is independent of m, and is given by

$$N_c \approx \frac{z^{3/2}(Q_L(z)Q_L(z-1))^{1/2}}{z-1}. \tag{2.22}$$

Note that N_c is the maximum number of packets that one can generate at each time step without collapsing the network and not the total load of the network at p_c.

To check the effect of the cost factor, we propose the following form for $Q_L(c)$

$$Q_L(c) = 1 - \tanh \frac{c}{L}. \tag{2.23}$$

Although the election of Q_L is completely arbitrary, (2.23) has two desirable properties: (i) it is a monotonically decreasing strictly positive function and (ii) Q_L decreases linearly for small values of c (compared to L). Also, Q_L decreases faster for small values of L and vice versa.

As can be seen from figure 2.8, the scenario that arises with the introduction of the cost factor is much more interesting. Now, the cost term compete with the behavior we have found for the critical number of generated packets, N_c, in the case of cost less connections. Thus, there is a maximum in N_c related to an optimum value of z, z^*, which defines an optimal organizational structure different from the trivial $m = 2$ and $z = S - 1$.

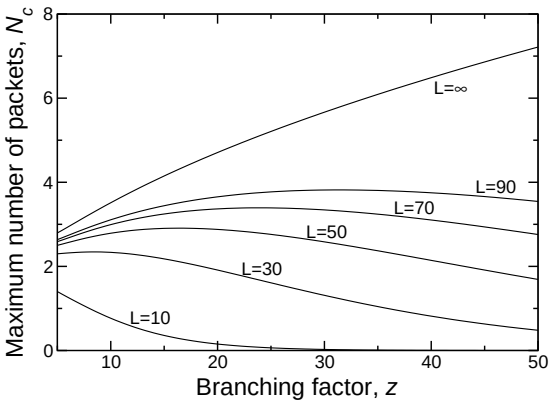


Figure 2.8. Maximum number of packets that can be generated in an organization per time unit without collapsing it, plotted as a function of z. Different curves correspond to different values of the linking capability, L.

From an organizational point of view, L is related to the availability of communication technology. When efficient communication technology is available, then the cost for communication is low and L is large. Conversely, when the communication technology is outdated and not efficient, L is small meaning that it is expensive in terms of time to keep channels open. With this interpretation, our results suggest that when communication technologies are improved, organizations should tend to more flattened charts or, in other words, to an increase of the span of control of managers. This reduction of layers related to the improvement of communication technologies has been observed in real organizations (Batt, 1996).

2.3 Non hierarchical networks

As we have discussed, the hierarchical network is probably the best *zeroth order* approximation to the communication network of an organization. However, this chapter is devoted to understanding the communication model and, at least from this perspective, it is interesting to consider different network topologies. In particular, we will focus in one-dimensional (1D) and two-dimensional (2D) regular lattices. In the first part of this section, we will compare the critical dynamics of the communication process in the different networks when $\xi = 1$, that is, in the communication model considered so far. In the second part, we will study the non-critical cases $\xi < 1$ and $\xi > 1$ that have not been considered for the hierarchical lattice.

2.3.1 Critical congestion behavior in 1D and 2D lattices

Similarly to what we have done for the hierarchical lattice, in this section we will study the position of the congestion critical point and its dependence on the size of the network, and the behavior of the order parameter and of the characteristic time of the system. As before, packets are forced to follow minimum paths from their origins to their destinations. In 1D lattices the minimum path between origin and destination is unique as happened in the hierarchical networks. However, in 2D lattices there are many such minimum paths and one of them is chosen at random. As will be shown, the coexistence of many possible paths is important for the dynamic behavior of the system.

As in the hierarchical network, it is possible to derive a mean field expression of ρ_c for the 1D lattice. Since the most congested node is, from symmetry arguments, the central one—the node at $\ell = S/2$—, the network will collapse when the amount of packets received by this central node is higher than the maximum number of packets that it is able to deliver. Since in a large enough network it is safe to assume that the central node will be congested similarly to its neighbors, $\nu_{\ell-1} = \nu_\ell = \nu_{\ell+1}$, the maximum number of delivered packets should be 1. On the other hand, the number of packets arriving to the central node at each time step is the number of packets that are generated at each time step at the left half of the network and have their destination at the right half and conversely, this is $\rho S/2$. Then the critical condition is given by

$$1 = \frac{\rho_c^{1D} S}{2} \Rightarrow \rho_c^{1D} = \frac{2}{S}. \tag{2.24}$$

Figure 2.9.b shows the excellent agreement between this equation and simulation results.

For the 2D lattice it is more difficult to obtain even a mean field expression for ρ_c. However, since for 1D lattices and hierarchical trees the scaling relation $\rho_c \propto S^{-1}$ holds, one may expect the same behavior for the 2D lattice. Using the susceptibility to numerically determine ρ_c from simulations, one finds that this turns out to be incorrect. Although it is difficult to obtain a precise value of the exponent, figure 2.9.d shows that it is close to 0.6 instead of 1.0:

$$\rho_c^{2D} \propto S^{-0.6}. \tag{2.25}$$

This result suggests that the existence of multiple paths to get from the origin to the destination has important consequences, not only in shifting the value of ρ_c, but actually changing its critical scaling behavior.

The behavior of the order parameter is studied next. As observed in figure 2.10, the general form

$$\eta(\rho/\rho_c) = \frac{\rho/\rho_c - 1}{\rho/\rho_c} \tag{2.26}$$

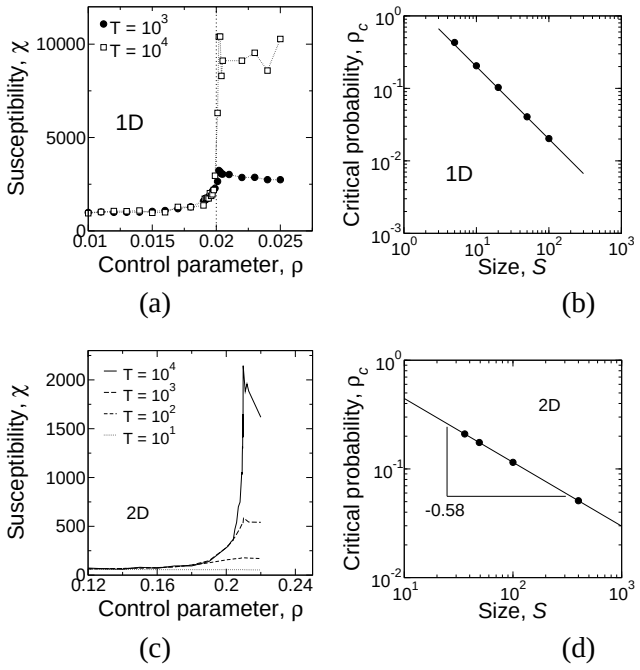


Figure 2.9. (a) and (c) Susceptibility for different time windows: (a) 1D and (c) 2D. The dotted vertical line in (a) represents the mean field estimation of the congestion point. (b) and (d) Dependence of the critical congestion point with the size of the network: (b) 1D and (d) 2D. The line corresponds in (b) to the mean field estimation and in (d) is simply a power law fitting of the points, that yields an exponent of -0.58.

obtained before fits very accurately the behavior of the order parameter not only for trees but also for any 1D lattice. Two-dimensional lattices again deviate from this behavior, although the deviation is small.

Finally we consider the behavior of the characteristic times. The analysis of the power spectrum shows that in the sub-critical regime, i.e. in the free phase, the spectrum is well fitted by a Lorentzian characterized by a frequency, f_c, as happened in the hierarchical case (figure 2.11). Near the critical point, we have seen that scaling behavior holds (equation (2.13)). As before, the value of the critical exponent γ is estimated by fitting equation (2.13) to values of $f_c(p)$ close enough to the critical point, as shown in figure 2.11. This procedure yields $\gamma \approx 0.9$ for a 1D network with $S = 100$, and $\gamma \approx 2.5$ for a 2D network with $S = 7 \times 7$, which are different from the values obtained for the hierarchical lattice.

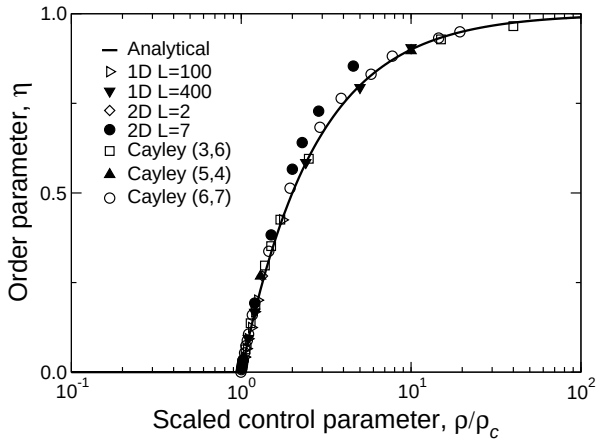


Figure 2.10. Behavior of the order parameter in the critical case for different network topologies. The solid line corresponds to the analytical calculation for two nodes exchanging information packets. Symbols correspond to simulations performed in 1D, 2D and hierarchical lattices.

2.3.2 Non critical cases $\xi < 1$ and $\xi > 1$

We have shown in section 1.1 that the number of packets delivered by node i is $\nu_i^{1-\xi}$ and thus, when $\xi < 1$, it increases with the number of packets that this node accumulates. It is difficult to imagine a real scenario with this characteristic. However, this case has been included to understand the critical behavior when $\xi = 1$, i.e., to show the relationship between criticality and the amount of packets that can be delivered when load increases. As a consequence of the increase of the deliver capability with the load, the transition to collapse will never occur because, at some point in time, the number of accumulated packets will be large enough and the number of delivered and created packets will balance each other. Thus, the order parameter will be zero for any value of the control parameter ρ, and the correlations will decay exponentially. As shown in figure 2.12, the characteristic frequency tends asymptotically to f_c^* as ρ increases. This asymptotic value depends on the size of the system.

For a 1D lattice with a high density of packets ($\rho \to 1$), the number of packets that are delivered by a node is $\nu_i^{1-\xi}$ while the number of packets that are being delivered to this node is proportional to S (for instance, for the central node, this number is simply $\rho S/2$). Therefore, $\nu_i \propto S^{1/(1-\xi)}$. The total number of packets is $N = \sum_i \nu_i \sim S^{1+1/(1-\xi)}$ and according to Little's Law

$$f_c^* \propto \frac{pS}{N} \propto S^{\frac{-1}{1-\xi}} \tag{2.27}$$

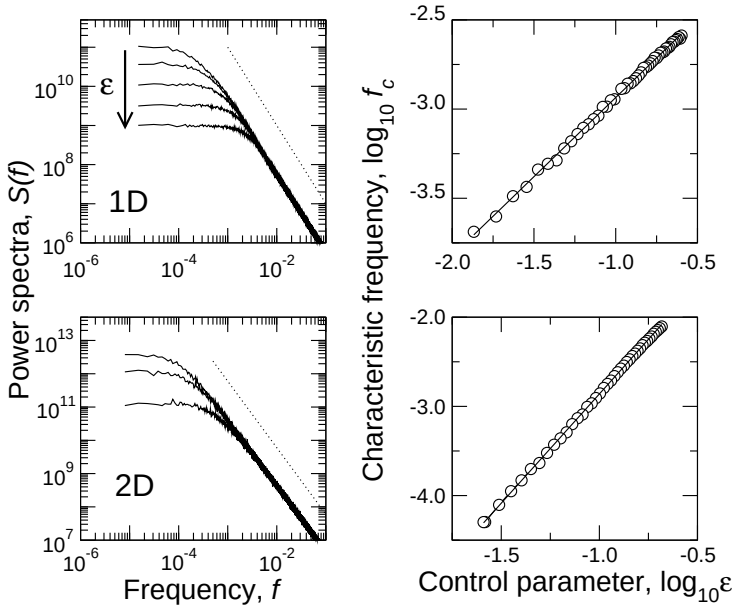


Figure 2.11. Left: Log-Log plot of the power spectrum of $N(t)$ for different values of the control parameter $\epsilon = (\rho_c - \rho)/\rho_c$ and for different topologies: the 1D case ($S = 100$) and the 2D case ($S = 7 \times 7$). Power spectra have been obtained averaging over 100 realizations of $N(t)$. Dotted lines represent a power law with exponent -2. Right: Characteristic frequency as a function of the control parameter ϵ for the different topologies. As observed, the characteristic frequency tends to 0 as $\rho \to p_c$ following a power law. The straight lines correspond to fittings of equation (2.13).

On the other hand, for $\rho \to 0$ the scaling of the characteristic frequency is given by

$$f_c^0 \propto S^{-1} \tag{2.28}$$

since the packets success to jump from one node to the next at all time steps, and therefore the characteristic time is directly the average path length between nodes. Therefore, although there is no phase transition in this case $\xi < 1$, there is a cross-over from a low density behavior to a high density behavior, as shown in figure 2.12. This crossover is also observed in 2D lattices and Cayley trees.

The phase transition observed for $\xi = 1$ is recovered when $\xi > 1$. Above a certain threshold, some packets are accumulated in the network and the order parameter differs from 0. However, the number of packets delivered by a node i in this case $\xi > 1$ decreases with the number of packets accumulated at that node. Therefore, when some packets are accumulated, ν_i grows and finally no packets are delivered at all. Thus suddenly above the transition, which is discontinuous, the order parameter becomes 1.

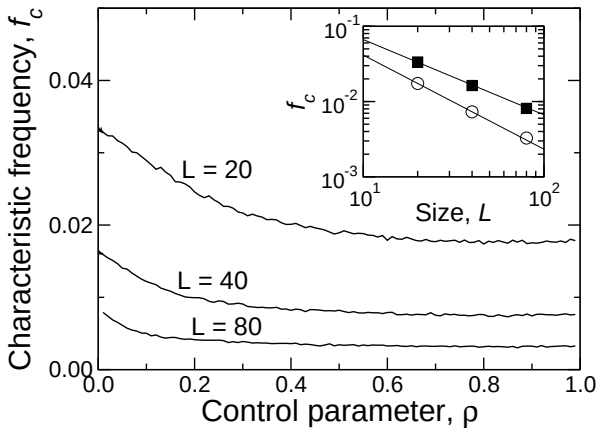
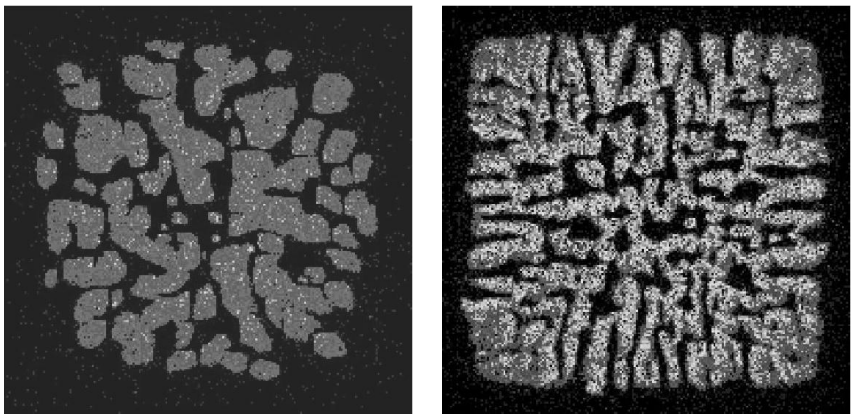


Figure 2.12. Characteristic frequency f_c as a function of the probability of packet generation ρ, for $\xi = 0.2$ and different sizes of a 1D lattice. As observed, f_c never becomes 0 as happens in the critical $\xi = 1$ case. Inset: Characteristic frequency at $\rho \to 0$, f_c^0 (squares), and characteristic frequency at large ρ, f_c^* (circles). The lines represent the fittings provided by equation (2.28) $f_c^0 \propto S^{-1}$, and equation (2.27) $f_c^* \propto S^{-1/(1-\xi)}$, respectively.

The change in the order of the phase transition (from continuous to discontinuous) affects the spreading of the collapse over the network. In the critical case $\xi = 1$, the collapse starts at the most *central* node and then spreads from this point to the rest of the network. In this case $\xi > 1$, the reinforcement effect—the fact that the more collapsed a node is, the more collapsed will get in the future—leads to the formation of many congestion nuclei generated by fluctuations, that spread over the whole network. Figure 2.13 illustrates the formation of these congestion nuclei for 2D lattices with $\xi = 5$ and $\rho = 0.001$, and $\xi = 2$ and $\rho = 0.01$

3. Summary

After stressing the importance of communication and information processing in the theoretical analysis of organizations, in this chapter we have proposed and studied a simple and general collection of stochastic models for communication processes. The models include only the essential elements present in a communication process between two elements: (i) information packets to be transmitted (delivered), (ii) communication channels to transmit the packets, and (iii) limited capability of agents to handle information packets. Despite its simplicity, the model reproduces the main characteristics of the flow of information packets in a real environment. In particular, we focus in a scenario where nodes are able to deliver on average a fixed number of packets per time step independently of their load, and observe the appearance of long queues



(a) (b)

Figure 2.13. Congestion nuclei formation for large 2D lattices with 200×200 nodes, in the non critical case $\xi > 1$. Dark regions represent regions with small congestion levels while bright regions correspond to highly congested regions. (a) $\xi = 5$ and $\rho = 0.001$. (b) $\xi = 2$ and $\rho = 0.01$.

that give raise to delays in the delivery of the packets, and the emergence of scale-free fluctuations in the total amount of information packets in the network, as reported in empirical analysis of real communication networks.

The behavior of the model is tuned by a parameter, ρ, that represents the probability of packet generation per node and time step. When ρ is small (close to 0) there are only a few packets traveling through the network and there is no congestion. However, as ρ grows congestion starts to play an important role and the delay to deliver a packet also increases. At a given critical point ρ_c the network collapses and the average delivery time diverges. For hierarchical networks, we have been able to characterize the phase transition and to estimate the position of the critical point, that gives a measure of the ability of the network to handle information. We have also shown that, when it is costless for an agent to keep a communication line open, the optimal hierarchical network is the flattest one, with a central node and all the others connected to it.

Moreover, the basic model has been extended in a number of independent directions. First, we have introduced the fact that agents are heterogeneous. Second, we have considered that keeping communication channels open can have a cost for agents. In this situation, we have shown that the optimal hierarchical structure is not the flattest one in general. Rather, there is an optimal *span of control* which is larger as the communication technology improves, as happens in real organizations. Third, we have considered networks that are not hierarchical and have concluded that when there is more than one path from the origin to the destination, the critical behavior of the system changes quan-

titatively. Finally, we have studied, in both hierarchical and non-hierarchical networks, the behavior of the system when the number of packets that a node is able to deliver is not independent of its load.

Chapter 3

OPTIMAL COMMUNICATION NETWORKS

In the previous chapter we have discussed the dynamical properties of a simple and general model of communication processes. At the end of the chapter, some considerations about the optimal design problem have been included when discussing the effect of costly communication channels. However, this study has been restricted to the comparison between different hierarchical networks. The purpose of the present chapter is to tackle the problem of optimal design from a more general perspective, in the line expressed in the Introduction. The question we try to answer is the following: given a set of agents and a limited amount of links to connect them, which is the network setup that optimizes the flow of information? As in previous works in the economics literature (Radner, 1993, Bolton and Dewatripont, 1994), optimality will be defined as the minimization of the average delay to process a certain information item.

The model presented in the preceding chapter, or a slight modification of it, will be used to measure the performance of a certain network configuration. However, as long as arbitrary networks will be considered, it will be necessary to extend the model. We have discussed before that, in the hierarchical network, the different levels did not represent bureaucratic hierarchy but knowledge hierarchy. Indeed, for every agent in the hierarchical lattice we have assumed perfect knowledge of the corresponding subtree below that agent, in such a way that packets always travel following the shortest path from their origin to their destination. Similarly, when one-dimensional and two-dimensional lattices have been introduced, we have also assumed that packets are still able to find optimal paths. When networks with arbitrary topology are considered, this assumption needs to be revised and the issue of *search* for the destination without precise information about the network will become important. In general, nodes will have at least local knowledge of the network, meaning that they will be able to identify whether the destination of a given packet is one of their

neighbors. Therefore, there will be a trade-off between centralization, which is positive in absence of global knowledge, and decentralization, which avoids network congestion.

The chapter is organized as follows. In the first section, we present some recent developments on the issue of search in complex networks. Next, we move to the study of optimal networks. In section 2 we consider model complex networks that are combination of regular lattices, random ER graphs and preferential BA networks, in which agents have local knowledge of the structure of the network plus *diffuse* global knowledge of the location of the nodes. Finally, in section 3 we consider a general analytical framework that allows to tackle the general problem of optimality in arbitrary networks with different levels of knowledge about the network, from purely local to complete knowledge.

1. Search in complex networks

After the *discovery* of complex networks, one of the issues that has attracted a lot of attention is "search". Real complex communication networks such as the Internet or the Worldwide Web are continuously changing and then it is not possible to draw a *map* that allows to navigate in them. Rather, it is necessary to develop algorithms that efficiently search for the desired computers or the desired contents.

The origin of the study of this problem is again in sociology since the seminal experiment of Travers and Milgram (Travers and Milgram, 1969). In the experiment, randomly selected individuals from Boston, Massachusetts, and Omaha, Nebraska, where asked to direct letters to a target individual in Boston, each forwarding her letter to a single acquaintance whom she judged to be closer than herself to the target. Surprisingly, Travers and Milgram found that the average length of the resulting acquaintance chains was about six. This means not only that short chains exist in social networks as reported, for example, in the "small world paper" by Watts and Strogatz (Watts and Strogatz, 1998), but even more striking that these short chains can be found using local strategies, that is without knowing exactly the structure of the whole social network.

The first attempt to understand theoretically the problem of *searchability* in complex networks was provided by Kleinberg (Kleinberg, 1999, Kleinberg, 2000). In his work, Kleinberg proposes a scenario where the network is modeled as a combination of a two-dimensional regular lattice plus a number of long-range links.[1] The distance Δ_{ij} between two nodes i and j is defined as the number of "lattice-steps" separating them in the regular lattice, that is disregarding long-range links (see figure 3.1). Long range links are not established at random. Instead, when a node i establishes one of such links, it connects

[1]Note that the long range links are *added* to the lattice without removing short range ones.

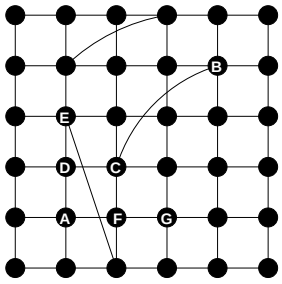


Figure 3.1. Network topology and search in Kleinberg's scenario. Consider nodes A and B. The distance between them is $\Delta_{AB} = 6$ although the shortest path is only 3. A search process to get from A to B would proceed as follows. From A, we would jump with equal probability to D or F, since $\Delta_{DB} = \Delta_{FB} = 5$: suppose we choose F. The next jump would then be to G or C with equal probability since $\Delta_{CB} = \Delta_{GB} = 4$, although from C it is possible to jump directly to B. This is a consequence of the local knowledge of the network assumed by Kleinberg.

with higher probability with those nodes that are closer in terms of the distance Δ. In particular, the probability that the link is established with node j is

$$\Pi_{ij} \propto (\Delta_{ij})^{-r} \tag{3.1}$$

where r is a parameter.

The search algorithm proposed by Kleinberg is the following. A packet standing at one node will be sent to the neighbor of the node that is closer to the destination in terms of the distance Δ. The algorithm is local because, as shown in figure 3.1, the heuristics of minimizing Δ does not warrant that the packet will follow the shortest path between its current position and its destination. Therefore, the underlying two-dimensional lattice has an imprecise global informational content. Going back to the Travers and Milgram's experiment, one could imagine that each participant in Ohama started sending the letter to an acquaintance in Boston or at least in Massachusetts, expecting that from there it would be easier to get to the destination. Alternatively, in the Worldwide Web environment, a user looking for a web page on a particular species of birds will probably begin looking for a web page devoted to animals and from there will probably move to a page devoted to birds and so on. Therefore, the underlying space does not need to have a geographical meaning but is just a way to model the fact that nodes are organized according to some criteria.

Kleinberg showed that with this essentially local scenario (with imprecise global information), short paths cannot be found in general, unless the parameter r is fixed to $r = 2$.[2] This raised the question of why real networks are

[2]In general, if one considers a n-dimensional embedding lattice instead of a two-dimensional lattice, the condition is $r = n$.

then searchable, that is, how is it possible that in real networks local strategies are able to find paths that scale as $\log N$, where N is the size of the network. Recently, Watts and coworkers have shown that with a similar idea to Kleinberg's, one can easily obtain searchable networks (Watts et al., 2002). Their contribution consists in substituting the underlying low-dimensional lattice by an *ultra-metric* space where individuals are organized in a hierarchical fashion according to their preferences, similitudes, etc. In this case, a broad collection of networks turn out to be searchable.

Parallel to these efforts, there have been some attempts to exploit the scale free nature of some networks to design algorithms that, being local in nature, are still quite efficient (Adamic et al., 2001, Tadic, 2001, Adamic et al., 2002). The idea in all these works is to take profit of the scale-free nature of networks such as the Internet and bias the search towards those nodes that have a high connectivity and therefore act as hubs.

Our approach is complementary to these efforts. The question we pose is the following: given a search algorithm that uses essentially local information and a fixed set of resources—i.e. a fixed number of nodes and links—, which is the topology that optimizes the search process? Moreover, we give an answer to this question in a general situation where the network has to tackle several simultaneous (or parallel) search problems, which in turn rises the important issue of congestion (Jacobson, 1988, Arenas et al., 2001, Ohira and Sawatari, 1998, Sole and Valverde, 2001) at overburdened nodes, a question that has been disregarded in the literature so far. Indeed, for a single search problem the optimal network is clearly a highly polarized star-like structure, with one or various nodes in the center and all the rest connected to them. This structure is *cheap* to assemble in terms of number of links and efficient in terms of searchability, since the average cost (number of steps) to find a given node is always bounded (2 steps), independently of the size of the system. However, the polarized star-like structure will become inefficient when many search processes coexist in parallel in the network, due to the limitation of the central node to process all information.

The discovery of optimal structures will be a useful guide to design, redesign and drive the evolution of communication networks. Although it can be argued that such a redesign process is not possible in networks like the Internet, it is worth noting that other systems like peer-to-peer networks, distributed databases, and most significantly in the present context, organizations can actually be designed and redesigned.

2. Communication and search in model networks

In this section we extend previous studies about local search in model networks in two directions. First, we consider networks that, as in Kleinberg's work, are embedded in a two-dimensional space, but study the effect not only

of long range random links but also of long range preferential links, directed to nodes that are already highly connected as in the BA model. Second and more significantly, we consider the effect of congestion when multiple searches are carried out simultaneously. As we will show, this effect has drastic consequences regarding optimal network design.

2.1 Network topology

The small world model by Watts and Strogatz (Watts and Strogatz, 1998) considered two main components: local linking with neighbors and random long range links giving rise to short average distance between nodes. The idea of Kleinberg is that local linking provides information about the social structure and can be exploited to heuristically direct the search process. Latter, Barabasi and Albert showed that growth and preferential attachment play a fundamental role in the formation of many real networks (Barabasi and Albert, 1999). Even though this model captures the correct mechanism for the emergence of highly-connected nodes, it is not likely that it captures all mechanisms responsible for the evolution of "real-world" scale-free networks. In particular, it seems plausible that in many of the networks that show scale-free behavior there is also an underlying structure as in the WS model. To illustrate this idea, consider web-pages in the World Wide Web. It is plausible to assume that a page devoted to physics is more likely to be connected to another page devoted to physics than to a page devoted to sociology. That is, a set of pages devoted to physics is likely more inter-connected than a set including pages devoted to physics and sociology.

Therefore we consider networks with four basic components: growth, preferential attachment, local attachment and random attachment. To create the network the following algorithm is used:

1 Nodes are located in a two-dimensional square grid with no links interconnecting them.

2 A node i is chosen at random.

3 We create m links starting at the selected node. With probability ϕ, the destination node is selected preferentially. With probability $1-\phi$ the destination node is one of the neighbors of the selected node. When the destination node is selected preferentially, we apply the following rule: the probability of a given destination node j being chosen is a function of its connectivity

$$\Pi_j \propto k_j^{\gamma}, \tag{3.2}$$

where k_j is the number of links of node j and γ is a parameter that allows to tune the network from maximum preferentiallity to no preferentiallity. Indeed, for $\gamma = 0$ the links are random and for $\gamma = 1$ we recover the BA

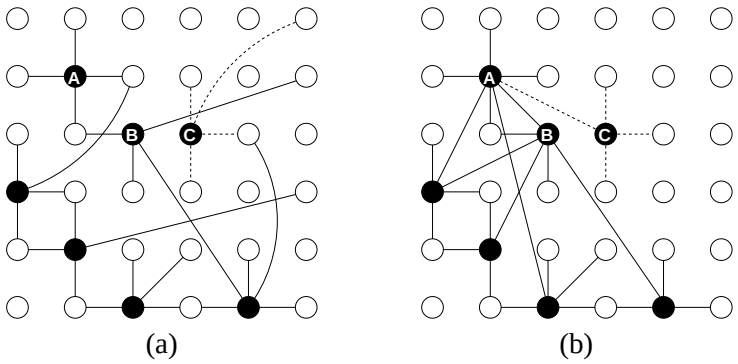


Figure 3.2. Construction of networks with multiple linking mechanisms. In both cases $\phi = 0.25$ in such a way that approximately one fourth of the links are long range. A random node is selected at each time step and $m = 4$ new links starting from that node are created. Black nodes represent nodes that have already been selected. Dotted lines represent the links created during the last time step in which node C was selected. In (a), the destination of long range links is created at random ($\gamma = 0$), while in (b) they are created preferentially ($\gamma > 0$) and nodes A and B are attracting most of them.

model, that generates scale free networks in the case $\phi = 1$. For $\gamma > 1$, one node tends to accumulate all the links.

4 A new node is chosen and the process is repeated from step 3, until all the nodes have been chosen once.

Figure 3.2 shows two examples of networks in the process of being created according to this algorithm.

Note that in this case, the number of links is fixed and the existence of long range links implies that some local links are not present and therefore that the information contained in the two-dimensional lattice is less precise.

2.2 Communication model and search algorithm

After the definition of the network creation algorithm, we move to the specification of the communication model and the search algorithm. For the communication model, we will use the general model presented and discussed in chapter 2. As already stated, this model is general enough and considers the effect of congestion due to node limitation to handle information.

In comparison with hierarchical networks, there is only one ingredient of the communication model that needs to be reformulated. In the hierarchical version of the model, when a node receives a packet, it decides to send it downwards in the right direction if the solution is there, or upward to the agent overseeing her otherwise. This simple *routing algorithm* arises from the fact that we implicitly assume that the hierarchy is not only a communicational hierarchy, but also a knowledge hierarchy, where nodes know perfectly the structure of the network

below them. In a complex network, this informational content of the hierarchy is lost. Here we will use Kleinberg's approach (Kleinberg, 1999, Kleinberg, 2000). When an agent receives a packet, she knows the coordinates in the underlying two-dimensional space of its destination. Therefore, she forwards the packet to the neighbor that is *closer* to the destination according to the lattice distance Δ defined in section 1, provided that the packet has not visited that node previously.[3] Note, however, that distance is refereed to the two-dimensional space, but not necessarily to the topology of the complex network and, as in Kleinberg's work, the algorithm will ignore some short paths just because it is necessary to increase Δ before getting to the right shortcut. Moreover, here long range links *replace* short range links and are not simply added to short range links. Therefore it is possible that following the direction of minimization of Δ the packet arrives to a dead end and has to go back.

Considering this algorithm, it is worth noting that the three mechanisms to establish links (local, random and preferential) are somehow complementary. A completely regular lattice (all links are local) contains a lot of information since all the agents efficiently send their packets in the best possible direction. However, the average path length is extremely high in this networks and therefore the number of packets that are flowing in the network at a given time is also very high. The addition of random links can reduce dramatically the average path length as happens in small worlds. However, if the number of random links is very high, then the number of local links is small and thus sending the packet to the node closer to the destination is probably quite inefficient (since it may happen that, even if it is very close in the underlying two-dimensional space, there is not a short path in the actual topology of the network). Finally, preferential links seem to solve both problems. They obviously solve the long average path length problem but, in addition, there is not a big lost of information because there are highly connected nodes that actually concentrate this information. The star configuration is an extreme example of this: although there are not local links, the central node is capable of sending all the packets in the right directions. However, when the amount of information to handle is big, preferential links are especially inadequate because highly connected nodes act as centers of congestion. Therefore, optimal structures should be networks where all the mechanisms coexist: complex networks.

2.3 Results

We simulate the behavior of the communication model in networks built according to the algorithm presented in section 2.1. First, a value of the probability of packet generation per node and time step, ρ, is fixed. For that particular

[3]Packets are sent to previously visited nodes only if it is strictly necessary. This *memory* restriction avoids packets getting trapped in loops

value, we compare the performance of different networks: networks with different preferentiallity, from random ($\gamma = 0$) to maximum centralization ($\gamma \gg 1$), and with different fraction of long range links, from pure regular lattices with no long range links ($\phi = 0$) to networks with no local component ($\phi = 1$). For each collection of the parameters ρ, γ, and ϕ, the network load, $\overline{N}$, is calculated and averaged over a certain time window and over 100 realizations of the network, so that fluctuations due to particular simulations of the packet generation and of the network creation are minimized. As in the economics literature, the objective is to minimize the average delay τ to arrive from the origin to the destination.

According to Little's Law of queuing theory (Allen, 1990), already presented in the previous chapter, the characteristic time is proportional to the average total load, $\overline{N}$, of the network:

$$\frac{\overline{N}}{\tau} = \rho S \Rightarrow \tau = \frac{\overline{N}}{\rho S} \tag{3.3}$$

where ρ is the probability of packet generation per node and time step. Thus, minimizing the average cost of a search is equivalent to minimizing the total load $\overline{N}$ of the network. The main results are shown in figure 3.3.

Consider first the behavior of the networks at low values of ρ. Figure 3.3.a shows the load of the network for $\rho = 0.01$ as a function of the fraction of long range links, ϕ, both when they are random $\gamma = 0$ and when they are extremely preferential $\gamma = 6$. In the last case, long range links are established only with the most connected node. In this case of small ρ, centralization is not a big problem because congestion effects are still not important. Therefore, preferential links are, in general, better than random long range links. In the case of preferential links, it is interesting to understand the behavior of the curve $\overline{N}(\phi)$. For $\phi = 0$ the network is a two-dimensional regular lattice and then the average distance between nodes is large. As some long range links are introduced, the average path length decreases as happens in the WS model (Watts and Strogatz, 1998), and therefore the load of the network is smaller because packets reach their destination faster. However, the addition of long range links implies the lack of local links and when ϕ is further increased, the heuristic of minimizing the lattice distance Δ becomes worse and worse. This fact explains that for $\phi \approx 0.15$ (the network is similar to the one depicted in figure 3.3.d) the load has a local minimum that arises due to the trade-off between the two effects of introducing long range preferential links: shortening of the distances that tends to decrease $\overline{N}$ and destruction of the lattice structure that tends to decrease the utility of the search heuristic and then to increase $\overline{N}$. If ϕ is further increased, one node tends to concentrate all the links and for $\phi = 1$ (figure 3.3.e) the network is strictly a star with one central node and the rest connected to it. In this completely centralized situation, the lack

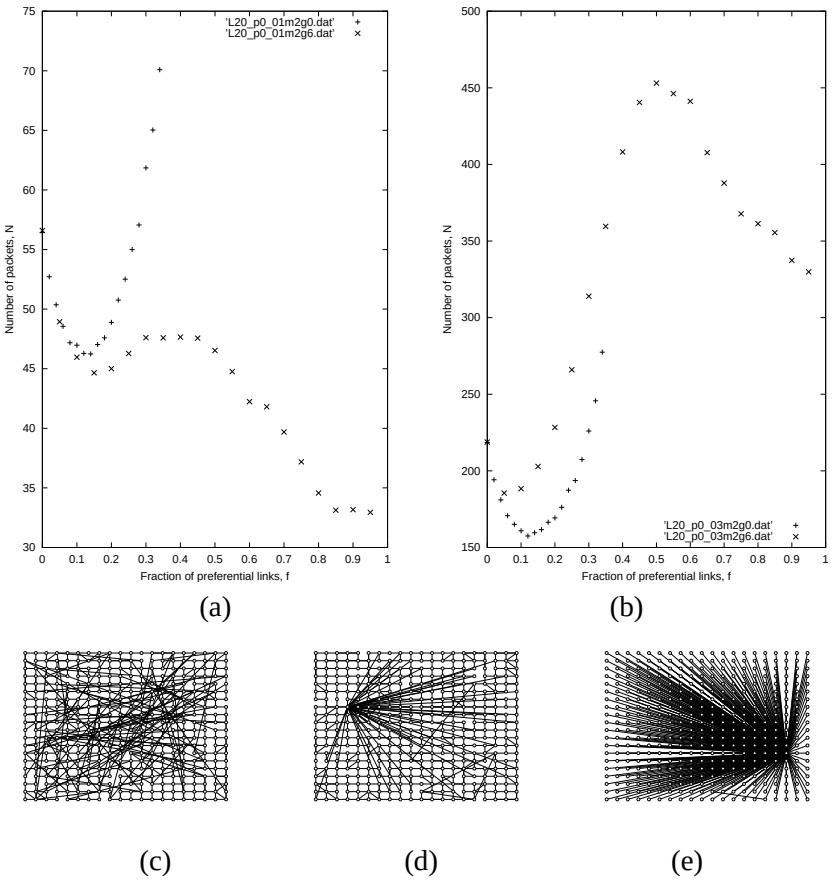


Figure 3.3. (a) and (b) Average number of packets flowing in the network as a function of the fraction of preferential links: (a) $\rho = 0.01$ and (b) $\rho = 0.03$. Symbol $(+)$ corresponds to $\gamma = 0$ (random links) and symbol $(\times)$ corresponds to $\gamma = 6$ (extremely focused links). Figures (c),(d) and (e) show the typical shape of complex networks with particularly efficient configurations: (c) $\gamma = 0$ and $\phi = 0.12$; (d) $\gamma = 6$ and $\phi = 0.07$; and (e) $\gamma = 6$ and $\phi = 1.0$;

of two-dimensional lattice is not important because the packets will be sent to the central node and from there directly to the destination. Since for small ρ congestion is not an issue, this structure turns out to be even better than the locally optimal structure with $\phi \approx 0.15$.

The situation is different when considering higher values of the probability of packet generation (figure 3.3.b displays the the results for $\rho = 0.03$). Regarding preferential linking, the two locally optimal structures with $\phi = 0.7$ and $\phi = 1$ (figures 3.3.d and 3.3.e respectively) persist. However, in this situation and due to congestion considerations the first is better than the second. Thus, at some intermediate value of $0.01 < \rho < 0.03$, there is a transition such that

the optimal structure changes from being the star configuration to being the *mixed* configuration with local as well as preferential links. Significantly, this transition is sharp, meaning that there is not a continuous pass from the star to the mixed.

Beyond the behavior of networks build with preferential long range links, it is worth noting that when the effect of the congestion is important (figure 3.3.b), the structure depicted in figure 3.3.c, where the long range links are actually thrown at random, becomes better than the structure in 3.3.d. In other words, the optimal network is, in this case, a completely decentralized small world network *a la* Watts-Strogatz.

3. Search, congestion, and optimal networks

So far, we have been able to compare the behavior of different networks build a priory following different rules (nearest neighbors linking, preferential attachment, etc.). The main reason for focusing on a particular set of networks is that it is very costly to compare the performance of two networks: it is necessary to run a simulation, wait for the stationary state and calculate the average load of the network. Specially close to the critical congestion point, the time needed to reach the stationary state become prohibitively long. Here, we present a formalism that is able to cope with search and congestion simultaneously, allowing the determination of optimal topologies. This formalism avoids the problem of simulating the dynamics of the search-communication process which turns out to be impracticable. We do not focus on detailed models of any of the above mentioned communication networks (organizations, computer networks, etc). Rather, we study a general scenario applicable to *any* communication process. First we calculate the average number of steps (search cost) needed to find a certain node in the network given the search algorithm and the topology of the network. The calculation is exact if the search algorithm is Markovian. Next, congestion is introduced assuming that the network is formed by nodes that behave like queues, meaning that are able to deliver a finite number of packets at each time step (Allen, 1990, Ohira and Sawatari, 1998, Arenas et al., 2001). In this context, we are able (i) to calculate explicitly the point at which the arrival rate of packets leads to network collapse, in the sense that the average time needed to perform a search becomes unbounded, and (ii) to determine, below the point of collapse, how the average search time depends on the rate at which search processes are started. In both cases, the relevant quantities are expressed in terms of the topology of the network and the search algorithm. Finally we obtain optimal structures by performing exhaustive generalized simulated annealing (Tsallis and Stariolo, 1994, Penna, 1995) in the space of the networks with fixed size and mean connectivity.

3.1 Search cost in absence of congestion

Before we take into account the effect of congestion, we consider the average cost to find a given node in an arbitrary communication network when there is no congestion. In other words, we focus on a single *information packet* at node i whose destination is node k, that is, a packet searching for k. The probability for the packet to go from i to a new node j in its next movement is p_{ij}^k. In particular, $p_{kj}^k = 0 \ \forall j$ so that the packet is *removed* as soon as it arrives to its destination. This formulation is completely general, and the precise form of p_{ij}^k will depend on the search algorithm. In particular, when the search is Markovian, p_{ij}^k does not depend on previous steps of the packet. In this case, the probability of going from i to j in n steps is given by

$$P_{ij}^k(n) = \sum_{l_1,l_2,\ldots,l_{n-1}} p_{il_1}^k p_{l_1 l_2}^k \cdots p_{l_{n-1}j}^k. \tag{3.4}$$

Thus defining the matrices p^k and $P^k(n)$, whose elements are p_{ij}^k and $P_{ij}^k(n)$, we have

$$P^k(n) = \left(p^k\right)^n. \tag{3.5}$$

The objective is to obtain the search cost or, in other words the *effective distance*, that a packet has to travel before it gets to its destination. For a packet starting at node i which destination is k the effective distance is

$$d_{ik} = P_{ik}^k(1) + 2\,P_{ik}^k(2) + 3\,P_{ik}^k(3) + \ldots \tag{3.6}$$

Using matrix notation we can define the matrices D^k whose elements D_{ij}^k are such that $d_{ik} = D_{ik}^k$. These matrices are given by

$$D^k = \sum_{n=0}^{\infty} nP^k(n) = \sum_{n=0}^{\infty} n\left(p^k\right)^n, \tag{3.7}$$

and using standard matrix algebra one obtains for this summation

$$D^k = \left[(I - p^k)^{-1}\right]^2 p^k, \tag{3.8}$$

where I is the identity matrix. The elements D_{ij}^k are the average number of steps needed to go from i to j for a packet traveling towards k.[4] In particular, as stated, the element D_{ik}^k is the average number of steps needed to get from i to k when using the search algorithm given by the set of matrices p^k. When the search

[4]It is assumed that the eigenvalues of the p^k matrix are smaller than 1, which must be true if the number of times that a packet goes through a certain node is finite. This condition will not hold, for instance, if the network is formed by more than one connected component.

algorithm has global knowledge of the structure of the network and the packets follow minimum paths between nodes, the effective distance will coincide with the topological minimum distance; otherwise, the effective distance between nodes will be, in general, larger than the topological minimum distance.

Finally, the average search cost in the network when there is not congestion is

$$\overline{d} = \frac{\sum_{i,k} d_{ik}}{S(S-1)} = \frac{\sum_{i,k} D^k_{ik}}{S(S-1)}, \tag{3.9}$$

where S is the number of nodes in the network. This expression allows to calculate exactly the average search cost performing simple matrix algebra. Note that simulation based calculation of this quantity would require, in principle, to generate an infinite amount of packets and let them travel from all possible origins to all possible destinations following all possible paths, which are in general arbitrarily long. This is why the analytical result is extremely useful.

3.2 Search cost in presence of congestion

Next, we consider a situation in which multiple search processes are performed in parallel. As discussed in chapter 2, this will give rise to accumulation of packets in the nodes due to the limitation of agents to handle information. In this case, the effective distance is not a good measure of performance since, even when the distance is small, accumulation of packets can generate long delays. Rather, the characteristic time, τ, needed to get from the origin to the destination is the right measure. As discussed before, minimizing τ is equivalent to minimizing $\overline{N}$. In the following, we show how to calculate the load of a network using only the p^k matrices as has been done for the case of no congestion.

Let us consider a measure of the *centrality* of each of the nodes in the communication network. First, we calculate the average number of times, b^k_{ij}, that a packet generated at i and with destination k passes through j. According to the previous definitions

$$b^k = \sum_{n=1}^{\infty} P^k(n) = \sum_{n=1}^{\infty} \left(p^k\right)^n = (I - p^k)^{-1} p^k. \tag{3.10}$$

Note that the elements b^k_{ij} are sums of probabilities but are not probabilities themselves. For example, imagine a packet starting at i and going to k, and assume that the probability for the packet to be in node j is one in steps 3 and 13 and zero otherwise—according to the search algorithm represented by the p^k matrices. Then, b^k_{ij} will be 2, since that packet will pass two times through j.

The *effective* betweenness of node j, B_j, is defined as the sum over all possible origins and destinations of the packets, and represents the total number

of packets that would pass through j if one packet would be generated at each node at each time step with destination to any other node:

$$B_j = \sum_{i,k} b_{ij}^k. \tag{3.11}$$

Again, as in the case of the effective distance, when the search algorithm is able to find the minimum paths between nodes, the effective betweenness will coincide with the *topological* betweenness, β_j, as usually defined (Freeman, 1977, Newman, 2001a). The topological betweenness, β_j, is the number of minimum paths connecting pairs of nodes in the network that go through node j. The effective betweenness of the nodes in a network contains valuable information about its behavior when multiple searches are performed simultaneously and congestion considerations become relevant.

Now consider the following general scenario. In the communication network, each node generates one packet at each time step with probability ρ independently of the rest of the nodes. The destination of each of these packets is randomly fixed at the moment of its creation. On the other hand, the nodes are queues that can store as many packets as needed but can deliver, on average, only a finite number of them at each time step—without lost of generality, we fix this number to 1. The model in chapter 2 is a particular example of this general scenario and it has been shown that for low values of ρ the system reaches a steady state in which the total number of *floating* packets in the network $N(t)$ fluctuates around a finite value. As ρ increases, the system undergoes a continuous phase transition from this *free phase* to a *congested phase* in which $N(t) \propto t$ (Arenas et al., 2001). Right at the critical point, ρ_c, quantities such as $N(t)$ and the characteristic time diverge (Guimera et al., 2001a). In the free phase, there is no accumulation at any node in the network and the number of packets that arrive to node j is, on average, $\rho B_j/(S-1)$. Therefore, a particular node will collapse when $\rho B_j/(S-1) > 1$ and the critical congestion point of the network will be

$$\rho_c = \frac{S-1}{B^*} \tag{3.12}$$

where B^* is the maximum effective betweenness in the network, that corresponds to the most central node.

To calculate the time average of the load of the network, $\overline{N}$, it is necessary to establish the behavior of the queues. In the general scenario proposed above, the arrival of packets to a given node j is a Poisson process with mean $\mu_j = \rho B_j/(S-1)$. Regarding the delivery of the packets, consider the following simplification of the communication model proposed in chapter 2. For a node j with ν_j packets stored in its queue, each packet jumps to the next node (chosen according to the algorithm defined through the matrices p^k) with probability $1/\nu_j$. With respect to the previous communication model, the only change is

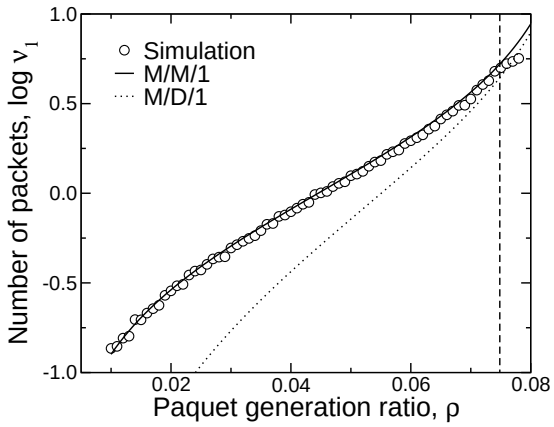


Figure 3.4. Comparison between simulated and analytical load of a node in the communication model described in section 3 of the present chapter. As observed, the behavior of the nodes is in excellent agreement, as expected, with a queue M/M/1. The behavior of an M/D/1 queue is shown for comparison. Note that there is not any adjustable parameter to fit, since the load is calculated according to equation (3.13). The vertical dashed line corresponds to the critical congestion point of the network, ρ_c at which the most central node starts to collapse. Then, some packets are accumulated at that node and the load of the considered node is less than predicted by equation (3.13). It does not represent a shortcoming of the calculation because, at this point, the total load of the network diverges.

that now communication depends only on the sender of the packet and not on the receiver. This allows analytical treatment and yield more interesting results as discussed later. In this new model, the delivery of packets is also a Poisson process. In this simple case in which both the arrival and the delivery are Poisson processes, queues are called M/M/1 in the computer science literature and the average size of the queues is given by (Allen, 1990)

$$\langle \nu_j \rangle = \frac{\mu_j}{1-\mu_j} = \frac{\frac{\rho B_j}{S-1}}{1-\frac{\rho B_j}{S-1}}. \tag{3.13}$$

Figure 3.4 shows the perfect agreement between simulation of the model and the values predicted by equation (3.13). The average load of the network $\overline{N}$ is

$$\overline{N} = \sum_{j=1}^{S} \langle \nu_j \rangle = \sum_{j=1}^{S} \frac{\frac{\rho B_j}{S-1}}{1-\frac{\rho B_j}{S-1}}. \tag{3.14}$$

It is straightforward to extend the calculations to other types of queues. For instance, the queues used in (Ohira and Sawatari, 1998) are such that one packet is delivered deterministically at each time step. These queues are called M/D/1 and the corresponding expression for the size of the queues is $\langle \nu_j \rangle =$

$\mu_j^2/(1-\mu_j)$. Moreover, it is worth noting that, although we started considering a modification of the communication model presented in the previous chapter, the fact that we are able to map the behavior of the nodes to that of M/M/1 queues implies that any conclusion that we are able to draw will be valid in general for any system of M/M/1 queues, and with small modifications for other types of queues.

There are two interesting limiting cases of equation (3.14). When ρ is very small, $\langle \nu_j \rangle \approx \mu_j$ and taking into account that $\sum_j B_j = \sum_{i,k} d_{ik}^k$, one obtains

$$\overline{N} \approx \rho S \overline{d} \qquad \rho \to 0. \tag{3.15}$$

On the other hand, when ρ approaches ρ_c most of the load of the network comes from the most congested node, and therefore

$$\overline{N} \approx \frac{1}{1 - \frac{\rho B^*}{S-1}} \qquad \rho \to \rho_c, \tag{3.16}$$

where B^* is, as before, the betweenness of the most central node. The last two expressions suggest the following interesting problem: to minimize the load of a network it is necessary to minimize the effective distance between nodes if the amount of packets is small, but it is necessary to minimize the largest effective betweenness of the network if the amount of packets is large. The first is accomplished by a *star-like* network, that is, a network with one central node and all the others connected to it. Rather, the second is accomplished by a very decentralized network in which all the nodes support a similar load. This behavior is common to any system of queues provided that the communication depends only on the sender. In queues M/D/1, for example, equation (3.15) reads $\overline{N} \approx (\rho S \overline{d})^2$ (thus, minimization of $\overline{N}$ still implies minimization of $\overline{d}$) and equation (3.16) is unchanged.

3.3 Limitations of the calculation and bounds to other models

It is worth noting that there are only two assumptions in the calculations above. The first one has already been mentioned: the trajectory of the packets needs to be Markovian to define the jump probability matrices p^k. Although this is not strictly true in real communication networks—where packets are not allowed usually to go through a given node more than once—it can be seen as a first approximation (Sole and Valverde, 2001, Arenas et al., 2001, Ohira and Sawatari, 1998). The second assumption is that the jump probabilities p_{ij}^k do not depend on the congestion state of the network, although communication protocols sometimes try to avoid congested regions, and then $B_j = B_j(\rho)$. However, all the derivations above will still be true in a number of general situations, including situations in which the paths that the packets follow are

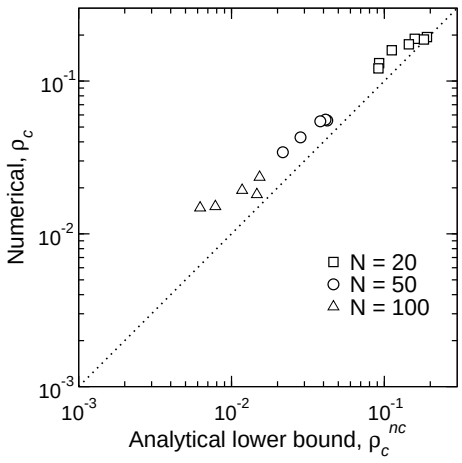


Figure 3.5. Comparison between the predictions of equation (3.12) for ρ_c and the results obtained for the communication model discussed in chapter 2. The analytical value is a lower bound to the actual value. To keep the figure simple, we do not show results corresponding to the model discussed in section 3, but the points would lay exactly on the diagonal line, since all the assumptions of the calculation are fulfilled.

unique, in which the routing tables are fixed, or situations in which the structure of the network is very homogeneous and thus the congestion of all the nodes is similar.

When these two assumptions are fulfilled the calculations are exact. For example, the calculation of ρ_c using equation (3.12) coincides exactly (within the simulation error) with simulations of the communication model introduced in this section where the communication only depends on the sender of the packet. Compared to situations in which packets avoid congested regions, equations (3.12)–(3.16) correspond to the worst case scenario and thus provide bounds to more realistic scenarios in which the search algorithm interactively avoids congestion. Consider, for example, ρ_c in the model presented in the previous chapter, where the communication depends not only on the sender but also on the availability of the receiver. As discussed in section 2.3.1 of that chapter, the fact that the packets are sent with higher probability to less congested nodes implies that the flow is better balanced among nodes. Although the assumptions of the present calculation do not apply, one would expect that the value of ρ_c estimated analytically will be a lower bound to the real situation in which load is more balanced. Figure 3.5 shows that this is indeed true and, moreover, that the analytical estimation provides a good approximation to the simulated value. This figure also confirms another expected and useful result. For a given size of the network and a given number of links, the most robust networks, that is those with higher ρ_c, are those with better balanced load. For

these networks, the effect of avoiding congestion is less important and therefore the analytical estimation turns out to be more accurate.

Of course, one can think of other interesting generalizations that can be included in the formalism. For example, one could imagine that the generation of packets with their corresponding destinations is not uniform, but that some origins and destinations are more common and even there are correlations between origin and destination. To include this effect, it would be enough to introduce a collection of weights in equation (3.11) such that

$$B_j = \sum_{i,k} \omega_{ik} b_{ij}^k. \tag{3.17}$$

3.4 Optimal network structures for local search

Equations (3.10), (3.11) and (3.14) enable us to tackle the problem of finding optimal structures for local search. An optimal structure is defined as the one that minimizes the average time needed to perform a search and therefore minimizes $\overline{N}$. In a purely local search scenario, nodes face the problem of forwarding a given packet: if the destination of the packet is one of the neighbors of the node, then the packet is sent to it; otherwise, the packet is just sent at random to one of the neighbors of the node. The corresponding p^k matrices are given by

$$p_{ij}^k = a_{ik}\delta_{jk} + (1 - a_{ik} - \delta_{ik})\frac{a_{ij}}{\sum_l a_{il}}. \tag{3.18}$$

where a_{ij} are the elements of the adjacency matrix of the network: $a_{ij} = 1$ if i and j are connected in the network and $a_{ij} = 0$ otherwise. The first term corresponds to i and k being neighbors: then the packet will go to j if and only if $j = k$, i.e. the packet will be sent directly to the destination. The second term corresponds to i and k not being neighbors: in this case, j is chosen at random among the neighbors of i. Therefore, in the absence of information, packets are distributed uniformly among neighbors. Finally, the delta symbol ensures that $p_{kj}^k = 0\ \forall j$ and the packet *disappears* from the network.

3.4.1 Optimization algorithm

The optimization process is carried out using generalized simulated annealing (GSA) as described in (Tsallis and Stariolo, 1994, Penna, 1995). Classical simulated annealing (CSA) is a stochastic optimization technique based on ideas from statistical mechanics that resembles the real annealing that experimental physicists and metallurgist do in the laboratory (Kirkpatrick et al., 1983). The idea is the following. Consider a crystalline solid whose atoms are organized mostly in a regular cubic lattice. Some atoms, however, lay outside this cubic lattice generating defects in the crystalline structure. Understanding why most

atoms lay in the cubic lattice and why some of them do not are the key questions to understand simulated annealing. Most of the atoms lay in the cubic lattice because this is the best way to minimize the interaction energy between them. However, some atoms are trapped in configurations that do not correspond to this minimum energy configuration and they cannot scape because, even when their configuration is not optimal, between their current configuration and the optimal one there is a big energy barrier which is difficult to cross. Annealing of solids is a technique to eliminate such defects, consisting in heating the material so that atoms can move around their positions and then cooling slowly so that atoms progressively tend to occupy the minimum energy configurations.

Simulated annealing proceeds similarly. Consider a system and a *cost function*, E, that depends on its configuration. The objective of an optimization technique is to minimize such cost function, that is to find the global optimum avoiding configurations that are only locally optimal. Then, a *computational temperature*, T is introduced in such a way that for low temperatures the system can only evolve towards a direction that minimizes E. Rather, at high temperatures the system can evolve in any direction. The process starts at high temperatures and then T is decreased slowly so that the system moves towards the optimal minimum avoiding local minima. This technique is thus adequate for systems with a complicated cost function with lots of local minima, where the system could get trapped with conventional optimization techniques.

In CSA this is accomplished in the following way. Consider the system in a given initial configuration whose energy (or cost) is E_i. Then a random modification is performed to the system, and the energy is changed to E_f. If $E_f < E_i$, the change is accepted. Otherwise, the change is accepted with a probability, P, that depends on $\Delta E = E_f - E_i$ and on T:

$$P_{CSA} = \exp -\frac{\Delta E}{T} \ . \tag{3.19}$$

This process is repeated and the temperature is decreased progressively until the system gets frozen in a given configuration.

In GSA as described in (Penna, 1995), the only difference is the form of the acceptance probability, which is given by:

$$P_{GSA} = \left(1 - (1-q)\frac{\Delta E}{T}\right)^{1/(1-q)} , \tag{3.20}$$

where q is a parameter that can be adjusted. In the limit $q = 1$, CSA is recovered. In general, GSA performs better than CSA provided that q is chosen properly. Figure 3.6 shows the result of several optimization processes with different values of q. The best performance is obtained for $q = -5$, and this is the value that will be used in the remaining of the chapter.

For the problem of network topology optimization, the cost function is the total load of the network $\overline{N}$ and the following procedure is used:

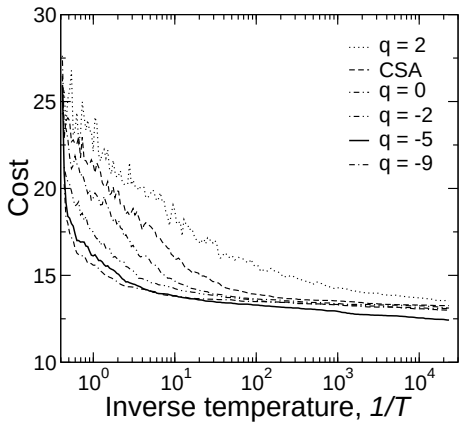


Figure 3.6. Performance of classical simulated annealing and generalized simulated annealing. Each line corresponds to a single run of the optimization process. As temperature is decreased, configurations with smaller and smaller cost (load) are obtained. Generalized simulated annealing with $q = -5$ (full line) yields the best results.

- Start with an initial network configuration with a fixed number of nodes, S, links, L, and ratio of packet generation, ρ, and with an initial temperature T.
- Repeat until $T < T_{fin}$ or the network has remained unchanged for more than $5 \times S \times S$ iterations:
  - Repeat S times:
    1. Choose one node at random and redirect one of its links to a new destination.
    2. Evaluate the cost $\overline{N}_{new}$ of the new configuration according to equation (3.14).
    3. If $\overline{N}_{new} < \overline{N}_{old}$ accept the change and continue; otherwise, accept according to the probability in equation (3.20).
  - Decrease the temperature according to $T_{new} = 0.99 \times T_{old}$.

Different sets of initial conditions are explored: for a given value of ρ, the optimization process is started from 100 different random initial configurations and also from networks that turned out to be optimal at similar values of ρ. Of all the realizations, only the network with a smallest cost is considered as optimal.

3.4.2 Results

The formalism introduced allows to perform an exhaustive search for optimal topologies in terms of parallel *searchability* avoiding the simulation of

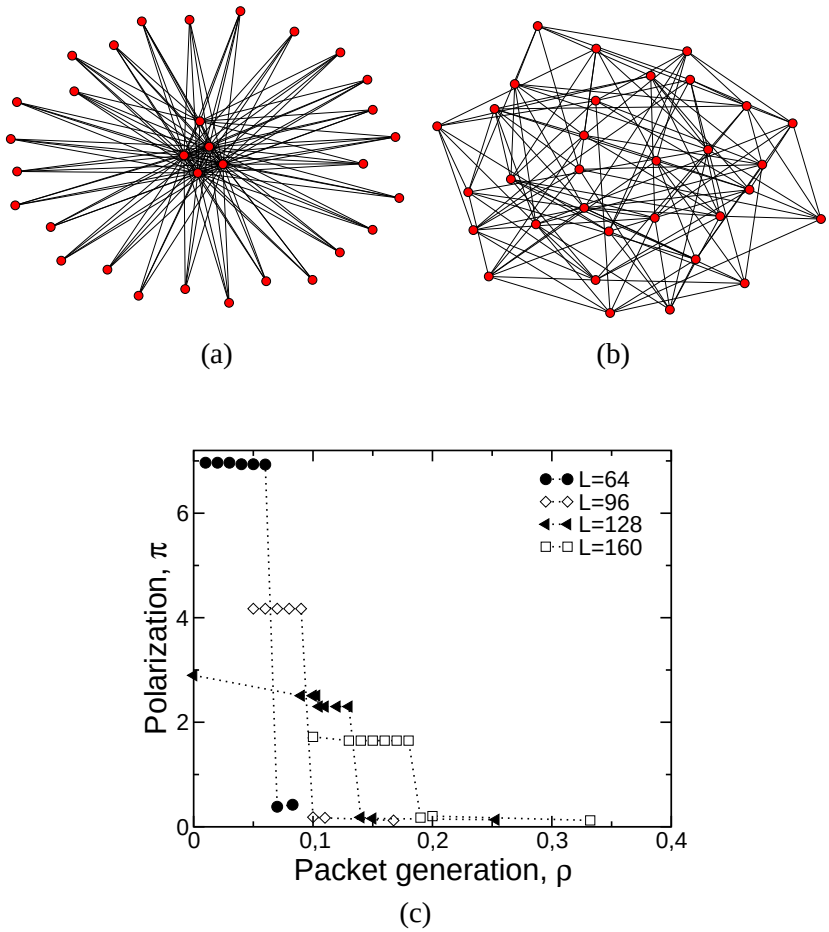


Figure 3.7. Optimal structures for local search with congestion. (a) Star-like configuration optimal for $\rho < \rho^*$. (b) Homogeneous-isotropic configuration optimal for $\rho > \rho^*$. (c) Polarization of the optimal structure as a function of ρ, for networks of size $S = 32$ and different number of links L.

the dynamics of the search-communication process. These simulations would result prohibitive in computational time, specially in situations in which one approaches the critical congestion point, ρ_c, and therefore the characteristic time diverges.

The main results of the optimization process are shown in figure 3.7. As predicted by equation (3.15), for $\rho \to 0$, the optimal network has a star-like centralized structure as expected, which corresponds to the minimization of the average effective distance between nodes. On the other extreme, for high values of ρ, the optimal structure has to minimize the maximum betweenness of the network, according to (3.16). This is accomplished by creating a homogeneous

network where all the nodes have essentially the same degree, betweenness, etc.

To clearly distinguish between the two opposite network topologies, star-like and homogeneous, we introduce a measure of the *polarization*, π, of the network:

$$\pi = \frac{\beta^* - \langle\beta\rangle}{\langle\beta\rangle} \tag{3.21}$$

where β is, as before, the topological betweenness of the nodes, and β^* is the largest betweenness in the network. For very homogeneous networks $\pi_h \approx 0$. Conversely, for star-like networks the central node belongs to all minimum paths and therefore, for $N \to \infty$, $\beta^* \propto N^2$ while $\langle\beta\rangle \propto N^5$ and thus $\pi_s \propto N$.

One could expect that the transition centralized-decentralized occurs progressively. Surprisingly, the results of the optimization process reveal a completely different scenario (figure 3.7.c). According to simulations, star-like configurations are optimal for $\rho < \rho^*$; at this point, the homogeneous networks that minimize B^* become optimal. Therefore there are only two type of structures that can be optimal for a local search process: star-like networks for $\rho < \rho^*$ and homogeneous networks for $\rho > \rho^*$. This result is similar to the one obtained in the previous section for model networks. In that case, it would be possible to argue that it is due to the restriction of the search space but not in the present situation.

3.5 Discussion

As already mentioned in the introductory chapter, the debate of centralization/decentralization in organizations has a long history (Van Zandt, 1998). However, none of the communication models in the economics literature can account for a transition from clearly centralized structures to clearly decentralized ones, both being optimal in different situations. Our developments in this section show that both centralized and decentralized network structures can be optimal in certain situations. In particular, centralization is good when the amount of information to handle is small and, conversely, decentralization is the best option in situations of information overload. Therefore, the need for decentralization arises as a consequence of the existence of limitations in agents' communication and information processing capabilities. This result is also related to the experimental evidence existing in the management literature in that simple tasks are better carried out by centralized groups while complex tasks are more properly accomplished by decentralized groups.

[5]Indeed, using again that $\sum_j \beta_j = \sum_{i,k} d_{ik}$ and that $d_{ik} = 2$ in most cases, one obtains $\langle\beta\rangle \approx 2N$ in the limit of large networks.

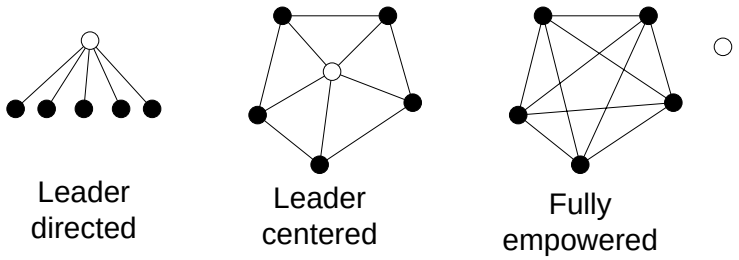


Figure 3.8. Pictorial representation of the empowerment process according to Dow Chemical's Strategic Blueprint. Here, the position of nodes and links should not be understood strictly as in the communication networks considered in the rest of the work. Rather, the drawing metaphorically represents a process by means of which leadership is decentralized and management tasks are assumed by the employees.

Our results also provide an explanation to the tendency existing in real multinational companies (Dow Chemical, in the chemical sector, is a good example) towards *empowerment*, or leadership decentralization, of their work teams. Figure 3.8 shows the concept of empowerment according to the Strategic Blueprint of Dow Chemical. It has been stated that empowered teams have a higher performance than traditional leader directed teams. For instance, according to Samuel L. Smolik (Smolik, 2001)

> Just to show you a relationship between level of Empowerment and safety performance; in 1999, we compared levels of empowerment in various plants to employee and contractor safety performance and found that the leader-directed organizations had a combined injury/illness rate of 4.47 per 200,000 man-hours. The plants that had achieved a Stage One level of Empowerment were performing at a 1.16 injury/illness rate; and the plants that had reached the top level Stage Two level of Empowerment were operating at a 0.62 injury rate. This is a significant demonstration of the multiple benefits gained from Empowerment of our employees.

It is worth emphasizing that, accordingly, our model predicts that decentralized networks can be more efficient that centralized ones in situations with information overload. None of the models in the economics literature can explain this fact in terms of communication and information processing capacities.

Beyond the existence of both centralized and decentralized optimal networks, it is remarkable that the transition from one sort of networks to the other is abrupt, meaning that there are not intermediate structures between total centralization and total decentralization. As already mentioned, this property is shared by the model networks in the previous section. The reason for the existence of such an abrupt transition is the following. Since we are considering (in both the present and the last sections) local knowledge of the network topology, centered star-like configurations are extremely efficient in searching destinations and minimizing, thus, the effective distance between nodes. This explains that stars are optimal for a wide range of values of ρ, until the central node (or

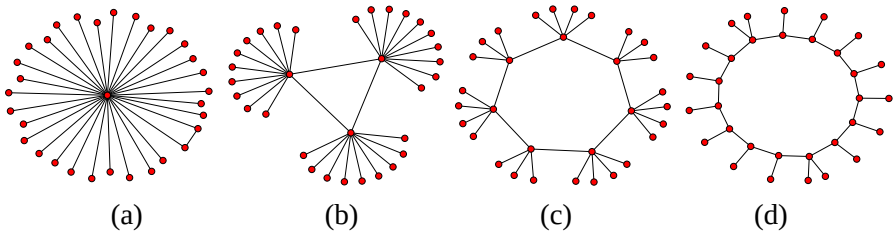


Figure 3.9. Optimal topologies for networks with $S = 32$ nodes, $L = 32$ links and global knowledge. (a) $\rho = 0.010$. (b) $\rho = 0.020$. (c) $\rho = 0.050$. (d) $\rho = 0.080$. In this case of global knowledge, the transition from centralization to decentralization seems smooth.

nodes) becomes congested. At this point, structures similar to stars will have the same problem and will we much worse regarding search; at this point, the only alternative is something completely decentralized, for which the absence of congestion can compensate the dramatic increase of the effective distance between nodes. In this situation, one should be able to obtain a smooth transition from centralization to decentralization by considering global knowledge of the network, in such a way that the average effective distance (that in this case coincides with the average path length) is not much larger in an arbitrary network than in the star. Our results for simple network parameters, figure 3.9, show that this is indeed the case.

This global knowledge situation is also interesting for one last question: the effect of the size of the network in centralization/decentralization. In this case, $\beta^* = B^*$ because effective and topological quantities coincide. For the star $B_s^* \propto N^2$, and for a completely decentralized homogeneous network, $B_h^* = \sum_{i,k} d_{ik}/N \propto N \log N$, where we have assumed that the average distance between nodes in this network scales as in a random graph, which seems very reasonable. This means that

$$\frac{\rho_c^s}{\rho_c^h} = \frac{B_h^*}{B_s^*} \propto \frac{\log N}{N}, \tag{3.22}$$

which is a decreasing function of the size of the system. In other words, it means that as the organization grows it should tend to decentralization. This is still another fact that has been observed in real organizations (Van Zandt, 1998).

4. Summary

With the knowledge acquired in the previous chapter about the dynamics of communication processes, here we have been able to address the main objective of the theoretical part of the present work: to design optimal communication networks. In particular, we have focused in a scenario in which agents do not have complete knowledge of the structure of the network but just local knowledge. In such a situation, there is a trade-off between search cost, that is

minimized in centralized structures, and congestion effects, that are minimized in decentralized structures. Therefore, in general it is not straightforward to find the optimal structure. Here we tackle the problem with two different approaches.

First, we have proposed an approach that is more intuitive but less general and reliable. The main idea is to generate networks with rules defined a priori and compare their efficiency. For a certain probability of packet generation per node and time step, ρ, we simulate a communication dynamics according to the model presented in the precious chapter, measure the total load of the network and define the optimal network as the one that has a lower load.[6]

Networks are built combining mechanisms that have been proposed in the literature of complex networks. First, there is a low-dimensional component resembling social structure that has an informational content as assumed in previous studies. Second, there is a preferential component, that can drift the establishment of links towards those nodes that have already a higher connectivity. Third, there is a random component. Each of these mechanisms has positive as well as negative aspects regarding communication. In a pure low-dimensional space, the informational content makes easy to find the destination of the packets. However, the average distance between nodes is large. When we substitute some of the links in the low-dimensional network by random long range links, we loose part of the informational content but we get much shorter distances between nodes. The average distance also decreases if we substitute links in the low-dimensional lattice by preferential ones and in this case the lost of informational content is less important than in the random case because the highly connected nodes will be connected to a lot of other nodes and therefore will have a lot of knowledge of the network. However, congestion effects will be very important in such centralized networks.

Simulations have confirmed these ideas and have yielded some surprising results. When the amount of packets is small, the optimal network is totally centralized. Rather, when there are a lot of packets, the optimal network is a combination of low-dimensional lattice and a few long range random links, that is, a small world a la Watts-Strogatz. One could expect that, when the amount of packets increases, the optimal structure was progressively less centralized, but this intuition turns out to be incorrect. Surprisingly, the optimal structure is totally centralized until for a certain value of ρ the optimal network becomes suddenly as Watts-Strogatz small world.

However, the objective of this chapter was more ambitious, and this has led us to the second approach. Without restricting our considerations to a particular family of a priory built networks, is it possible to find optimal network

[6]Actually, minimize the load is equivalent to minimize the average time needed to deliver a packet.

structures? With the elements presented so far, the most general procedure to find such optimal networks would be the following: generate a network, simulate the communication dynamics and calculate the load in the stationary state, modify the network and repeat the procedure until a network that cannot be further improved is found. However, such a process is prohibitively expensive in terms of time, specially when the amount of packets is such that the network is close to the collapse point and the time needed to get to the stationary state becomes arbitrarily long. To avoid this problem, we have introduced a formalism that is able to cope with search and congestion simultaneously and that allows to calculate exactly the load of the network. With this, the optimization procedure described above can be carried out. We have finally shown that, in the considered case in which agents have local knowledge of the network, there is, again, an abrupt transition between centralized and decentralized networks.

Although the debate centralization-decentralization in organizations is old, there is not any model that allows to explain, in communicational terms, why and in which conditions one sort of the structure is better than the other. Our results also allow to explain the empirical evidence that simple tasks are better carried out by centralized groups and, conversely, that complex tasks are better carried out by decentralized groups. Finally, these results also provide theoretical foundation to the tendency, existing in the management of large corporations, to *empower* (that is, decentralize) work groups at all levels.

Chapter 4

COMPLEX SELF-ORGANIZED COMMUNICATION NETWORKS AND ORGANIZATIONS

In chapter 2, communication networks have been modeled mainly as hierarchical networks, although some attempts of generalization have also been presented and optimal networks have been studied in chapter 3. Hierarchical networks are indeed a good approximation to model computer based communication networks such as the Internet (where there is a hierarchy of routers and servers) and also the formal chart of classical organizations, with the CEO at the top, her advisers at the second level and so on. For modern companies, however, the formal chart is not strictly hierarchical and less centralized settings have been shown to be much more efficient (Warnecke, 1993). This chapter is not devoted to the formal chart but to the informal communication network that naturally arises in a real organization. Krackhardt and Hanson (Krackhardt and Hanson, 1993) have established the following parallelism:

> If the formal organization is the skeleton of a company, the informal is the central nervous system driving the collective thought processes, actions, and reactions of its business units.

The formal chart of an organization is designed to handle routine and easily anticipated problems, but when unexpected problems arise, new ties are formed so that tasks can be accomplished properly. Ties in an organization also arise for personal, political and cultural reasons. The understanding of the informal networks underlying the formal chart is a key element for successful management (Mayo, 1949, Krackhardt and Hanson, 1993, Morgan, 1997), and therefore managers are interested in knowing how these networks work. The traditional way of investigating informal networks consists of two steps (Krackhardt and Hanson, 1993). First a network survey using employee questionnaires is conducted. However, employees answers often contain subjective elements such as "political" motives and the worry about offending colleagues. This effect can be minimized by the second step: cross-checking of the answers which is not

free of subjectiveness either. A more significant limitation of the questionnaire based analysis is that time and effort costs make it prohibitively expensive to map the entire network even for medium sized organizations.

The rapid development of electronic communications provides a powerful alternative for the analysis of informal networks. Indeed, the interchange of e-mails between individuals in organizations reveal a lot about how people interact and therefore should provide valuable hints about the real network structure behind the formal chart (Economist, 2001, Ebel et al., 2002, Adamic and Adar, 2002). This is interesting not only from a managerial point of view, but also from a theoretical and a fundamental points of view, if one wants to understand how organizations work and why do they work the way they do. However, obtaining information from communication networks is not straightforward. For instance, it is not possible to discriminate between different sorts of informal networks by analyzing an e-mail network. Krackhardt and Hanson (Krackhardt and Hanson, 1993) stressed the differences between different informal networks (advice network, trust network, etc.) and the importance of knowing them separately. In an e-mail network all the informal networks and even the formal chart contribute, interacting in a complex way. Nevertheless, the information obtained from communication network studies is still valuable. The second problem is methodological. The analysis of large and complex networks is not straightforward and the extraction of information requires the use of specific statistical techniques, developed recently in the field of statistical physics of complex networks (Watts and Strogatz, 1998, Barabasi and Albert, 1999, Amaral et al., 2000, Albert and Barabasi, 2002, Dorogovtsev and Mendes, 2002).

The purpose of this chapter is to study the properties and the community structure of the e-mail network of the University Rovira i Virgili (URV), at Tarragona, as an example of how these techniques developed for complex networks can be used in a real organization. In section 1 we describe how the network is built and study some of its statistical properties, such as the degree distribution and the clustering coefficient. Section 2 describes how can we obtain insights about the community structure of the network and present this information in a useful way. Next, we study the properties of the community structure. Strikingly, we find that it shows emergent self-similar properties as occurs in other natural systems like, for example, river networks. Finally, we show that the results obtained for the community structure can be used for management purposes. Appendix A, shows the study of a different communication network, the so-called “web of trust”.

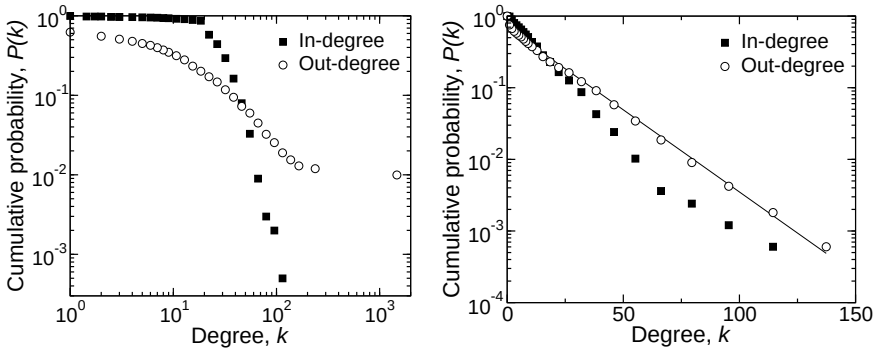


Figure 4.1. Degree distribution of the e-mail network of the Universitat Rovira i Virgili. (a) In- and out-degree distributions when all e-mails are considered. While the in-degree distribution decays exponentially, the out-degree distribution is highly skewed due to the presence of e-mail lists. (b) In- and out-degree distributions when e-mails sent to more than $\kappa = 50$ users are discarded. In this case, both distributions decay exponentially.

1. Characterization of the e-mail network of the Universitat Rovira i Virgili

Every time that an e-mail is sent, some information is registered in the corresponding server, including the address of the sender and the address of the receiver. Therefore, an *e-mail* network can be built regarding each address as a node and joining two nodes with a link if there is an e-mail communication between them. Considering that e-mails are directed and that one can easily distinguish between sender and receiver, the resulting graph will be, in principle, directed. We build such a network considering the e-mails sent within URV during the first three months of the year 2002. Significantly, only e-mails with sender and receiver belonging to the university were regarded, and *external* e-mails were not. At the URV, there are three different servers that manage the e-mail accounts of all the staff of the university (including academic and administrative staff, graduate students, managers, etc.). The total number of users is approximately 1700, which corresponds to the size of a medium sized company or of a site of a multinational company. Privacy is preserved by assigning a random code to each address, in such a way that the study is performed keeping the anonymity of the users. Moreover, it is worth noting that all the information used is routinely recorded by any e-mail server.

First, we study the degree distribution. In principle, the in- and out-degree distributions are measured considering the network as explained so far: nodes are addresses and directed links represent e-mails sent from one address to another. The results are shown in figure 4.1.a. The asymmetry between the distribution of incoming (received) and outgoing (sent) e-mails is observable

from the plot. While the maximum in-degree (that is, the maximum number of users that are sending e-mails to the same address) is about 100, the maximum out-degree (that is, the maximum number of addresses that a given user is sending e-mails to) is more than 1000. Actually, the in-degree distribution decays exponentially, while the out-degree distribution is highly skewed, due to a few nodes sending e-mails to more than 1000 addresses. The origin of this a priori surprising result is related to the existence of e-mail lists. Even when some addresses are removed explicitly because they represent lists of users instead of single users (all the academic staff in a department, for instance), some of them cannot be removed because when an e-mail is sent to one of these addresses it is *expanded* by the server, sent to all the addresses in the list individually and registered in the server as many different e-mails sent by the same user to different addresses. To overcome this problem, we fix a threshold κ: when a user sends an e-mail to more than κ different users, this e-mail is disregarded. The new in- and out-degree distributions obtained with $\kappa = 50$ are shown in figure 4.1.b. In this case, both follow similar distributions with exponentially decaying tails.

This result contrasts with the result obtained by Ebel and coworkers (Ebel et al., 2002), that showed that the degree distribution of a different e-mail network decays as a power law. There are different explanations to this apparent contradiction. First, they considered all the e-mails sent and/or received by users inside the university, while in the present study only e-mails with both origin and destination inside the university are considered. Second, we have shown that the skewness of the degree distribution could be due to the existence of lists of addresses and have decided to disregard these e-mails. There is nothing implicitly right or wrong in removing or not removing the lists. If one is interested in virus spreading, for instance, it is worth keeping all the e-mails because all of them could be infected. Rather, if one is interested in understanding how does relevant information flow (this is indeed the case of the present study) then it is fair to remove massive e-mails since they usually have low informational content.

Still another approach consists in considering that two nodes are connected if, and only if, there are e-mails flowing in both directions during the time period considered. Again, this is a way to identify *relevant* communication channels. In this case, in- and out- degree distributions coincide because all the links are bidirectional, and the total degree distribution is shown in figure 4.2. Significantly, with this restriction lists do not play an important role and the degree distribution depends only very slightly in the value of κ ($\kappa = \infty$ in figure 4.2.a and $\kappa = 50$ figure 4.2.b). The exponential decay is clear in this case.

Regarding the cluster structure, it is worth noting that, using $\kappa = 50$ and considering only bidirectional e-mails, the largest cluster contains 1133 nodes.

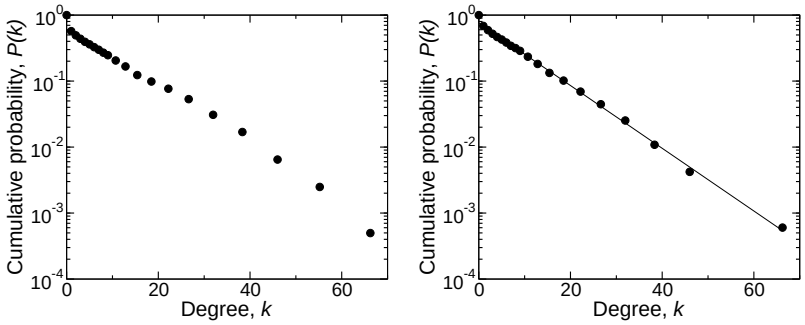


Figure 4.2. Degree distribution of the e-mail network of the Universitat Rovira i Virgili when only bidirectional e-mails are considered. (a) Lists are not eliminated. (b) Lists, that is e-mails sent to more than $\kappa = 50$ users, are disregarded. In this case, most of the useless e-mails are removed by the bi-directionality restriction and, therefore, the effect of removing lists is small. In other words, most of the e-mails that are sent to large amounts of people are not answered and thus are not considered.

The rest of the network is formed mostly by isolated nodes. For the largest component with 1133 nodes the clustering coefficient is $C = 0.254$, which is approximately 30 times larger than the expected value for a random graph with the same size and average degree. Such a high value of C suggests a scenario where the network is comprised of several highly connected communities—with a lot of redundancy in the linking—which are loosely connected to other highly connected communities. In fact it has been shown that there is a close relation between highly clustered regions of a graph and the existence of communities (Eckmann and Moses, 2002). In the next sections, we focus on the identification of such communities and on the characterization of their structure.

2. Community analysis methodology

In the previous section, we have studied some of the statistical properties of the e-mail network of the URV such as the degree distribution and the clustering coefficient. Although some interesting properties of the network can already be obtained from this statistical analysis (resilience of the network against removal of nodes or spreading of viruses, for example), there is a lot of information that is interesting from a managerial point of view that still remains unexploited.

As described in the introduction of this chapter, the intention of the analysis of the communication structure in an organization is to uncover the real (informal) chart behind the formal one, and to establish who collaborates with whom, which groups work together, etc. This analysis can be hardly performed by *visual inspection* as it is usually done when using employee questionnaires on small groups of 10–20 persons. Rather, it is necessary to design heuristic

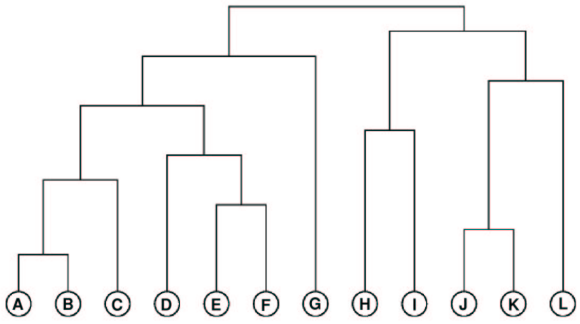


Figure 4.3. Example of a small dendogram. The circles at the bottom represent the nodes of the original network, and they are joined according to the hierarchical clustering. The vertical axis represents the order in which the clusters are joined together. In this case, A and B are joined first, J and K second and E and F third. Then the group formed by A and B is joined to C, and so on.

algorithms capable of identifying groups and communities from the topological properties of the complex communication network. In this section, such algorithms are described and some measures that help to characterize the community structure are presented.

2.1 Community identification using hierarchical clustering methods

The traditional method for identifying communities in networks is hierarchical clustering. The idea is the following. For each pair of nodes i and j in the network, define a weight W_{ij} that quantifies how closely connected they are. Then create an empty network with all the nodes but with no links between them, and start adding links between the nodes with highest W_{ij}. This procedure gives rise to a nested set of increasingly large components, that can be conveniently represented using the so-called *dendograms* (see figure 4.3).

A typical measure of the weights W_{ij} is the number of node (or edge) independent paths connecting the 2 nodes, that is the number of paths that connect the two nodes without sharing any node (or edge) in common. Actually, from the "Max Flow–Min Cut" theorem (Menger, 1927) it is known that the number of node (edge) independent paths between i and j equals the minimum number of nodes (edges) that need to be removed from the network to separate i and j from one another.

In a recent article, Flake and coworkers have used the Max Flow–Min Cut result with a different approach (Flake et al., 2002). The idea is to start from a set of *seed vertexes* and identify the community *around* these seed vertexes. This approach can be very useful for searching communities in the World Wide

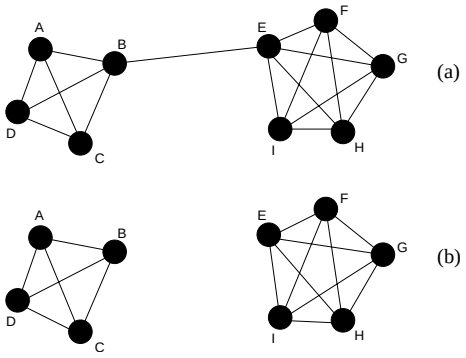


Figure 4.4. Identification of most central links in the GN algorithm. (a) The network in the drawing contains two clearly distinguished communities. The GN algorithm identifies the link that belongs to a higher number of minimum paths between all pairs of nodes: in this case the link $\overline{BE}$. (b) Removal of this link yields two separate networks that correspond to the original communities.

Web, for example, but does not provide an exhaustive map of communities of the whole network. Therefore, we will use a different algorithm.

2.2 Girvan-Newman algorithm for community identification and visualization of the community structure

The algorithm proposed by Girvan and Newman (GN) (Girvan and Newman, 2002) proceeds with a similar idea than the hierarchical clustering algorithms—that is identifying the most important links in the network—but removing connections from the initial network instead of adding them from an empty network as before. The main point is to identify the *most important links* in the network, that is those links that connect a maximum number of pairs of nodes, and remove them so that groups that are only slightly connected through these very important links become separated of each other. This is more easily understood considering figure 4.4. Imagine a network formed by 9 nodes, that will be denoted with letters A to I, connected as in the picture (figure 4.4.a). The most important link turns out to be the segment $\overline{BE}$ since it is necessary to use this link to go from any of the nodes in the left to any of the nodes in the right. The removal of the link $\overline{BE}$ actually separates the two communities correctly (figure 4.4.b). In general, however, it will be necessary to remove more than one link in order to split a given network.

The idea of *most important links* is formalized by means of the *betweenness* of the links, which is a measure of their centrality in the network. Consider all the minimum paths connecting pairs of nodes in the network. For instance, in figure 4.4.a, the minimum path connecting A and G consists of three steps:

$\overline{AB}$, $\overline{BE}$, and $\overline{EG}$. The betweenness of a link is defined as the number of such minimum paths that the link belongs to. It is straightforward to see that this definition of betweenness identifies the most important nodes and yield the desired results. Moreover, for a network with m links and n nodes, it is possible to calculate the betweenness of all the nodes in a time of the order $O(mn)$ (Newman, 2001a), and therefore the calculation can be easily performed even for relatively large networks (up to sizes of the order of 10^5 nodes).

The GN community identification algorithm proceeds as follows:

1 Calculate the betweenness for all the links in the network.

2 Remove the link with the highest betweenness.

3 Repeat from step 1 until no edges remain.

Girvan and Newman showed (Girvan and Newman, 2002) that this algorithm provides striking results even in networks (both real world and computer generated networks) in which traditional community identification algorithms systematically fail.

The output of the GN algorithm is a binary community tree that can, again, be represented as a dendogram. However, we choose not to plot it as a dendogram. Starting from the original connected network, links are removed until the network is split into two pieces. Then, each one of these two pieces is regarded as a new network and links are removed until they also split into two. The process is repeated until only isolated nodes are left. It is clear from figure 4.4 that when two communities are clearly separated, the GN split procedure will separate them. However, it is also interesting to understand what will happen when there is not a real community structure inside a group of nodes or, in other words, when all the nodes in a network belong to a well defined community. Consider the two examples given in figure 4.5. In the first case (a) all the nodes are equivalent and are connected to all the rest of the nodes. In this situation, all the links are also equivalent and, therefore, one of them will be selected arbitrarily by the GN algorithm. Imagine that the selected link is, for instance, $\overline{AB}$. After the removal of this link, the link between C and D will remain unchanged, but the other links, that are connected either to A or B, will have a higher betweenness than before because they have to *absorb* part of the minimum paths that previously went through $\overline{AB}$. Therefore, the next removal will affect one of these nodes, say $\overline{AD}$. Repeating the argument, it is easy to see that the next removed link will be $\overline{AC}$, and that the network will be finally split into two pieces, one formed by node A and another one formed by a new complete graph containing nodes B, C and D. This is represented in the right side of the figure: at the beginning of the process, there is a single community, 1, containing nodes A, B, C and D; after some removals the network is split into two pieces, one formed by node A and a new community, 2, formed by B,

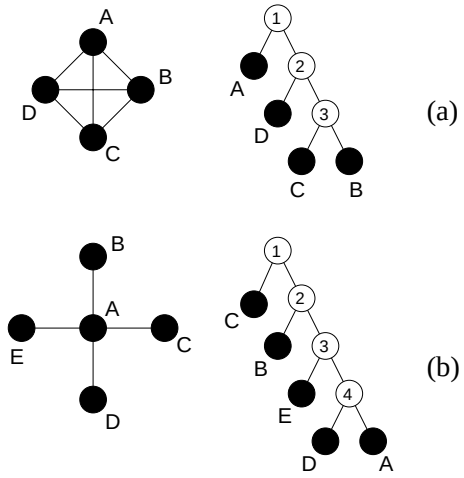


Figure 4.5. The GN algorithm on well defined communities. (a) When the network is completely uniform, the GN algorithm separates one node from the rest. Iterating this procedure, nodes are removed 1 by 1 and the resulting split binary tree is a linear branch. (b) When the network is star-like, nodes are also removed 1 by 1 but the central node will be the last one being separated.

C and D. The process would continue removing links and another isolate node would again be separated of the *main* component, for instance node D. Finally, the two remaining nodes would be separated of one another and the community identification process would stop. Similarly, in figure 4.5.b all the links are equivalent at the beginning and then would be removed arbitrarily. Again, individual nodes would be separated from the main component and finally one would be left with a pair of nodes formed by the central node A and one of its neighbors, for instance D.

With all this, it is easy to understand, figure 4.6, that the GN algorithm will separate different communities and that these communities will appear as relatively well defined branches in the binary tree. Moreover, according to figure 4.5, the final nodes of the branches can represent the most central nodes in the community. Actually, these ideas provide a powerful method to identify communities in large networks such as the e-mail of the Universitat Rovira i Virgili. Moreover, the study of the topology of the binary tree will provide quantitative information about the community structure of the network. For example, a poorly ramified structure will represent a network with diffuse network structure. Our main finding is that relevant information can be obtained by studying the topological properties of the resulting binary tree.

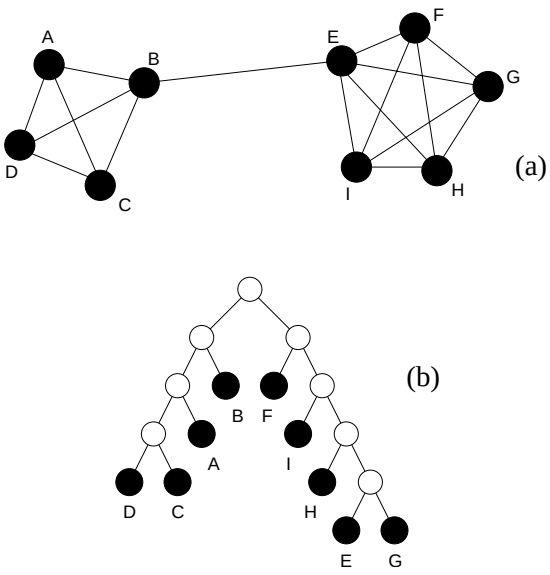


Figure 4.6. Communities and branches in the binary tree. When communities are identified as in (a), they appear in the binary tree as clearly differentiated branches (b).

2.3 Topological measures of the binary community tree

Binary trees have been extensively studied in many different areas, from discrete mathematics to computer science, geology or physics, and useful classification schemes and measures have been proposed. Some of them will be used in the study of the e-mail communication network and are introduced in this section.

2.3.1 Community size distribution

The first quantity that will be considered is the distribution of sizes of communities. Figure 4.7.a represents a hypothetical tree generated by the community identification algorithm (for clarity, the tree is represented *upside down*). Black nodes represent the actual nodes of the original graph while white nodes are just graphical representations of communities that arise as a product of the split procedure. Indeed, nodes A and B belong to a community of size 2, and together with E form a community of size 3. Similarly, C, D and F form another community of size 3. This two groups together form a higher lever community of size 6. Following up to higher and higher levels, the community structure can be regarded as a set of nested groups as depicted in figure 4.7.b.

A natural way of characterizing the community structure is to study the distribution of community sizes. In figure 4.7, for instance, there are 3 communities of size 2, 3 communities of size 3, 1 community of size 6, 1 community of size

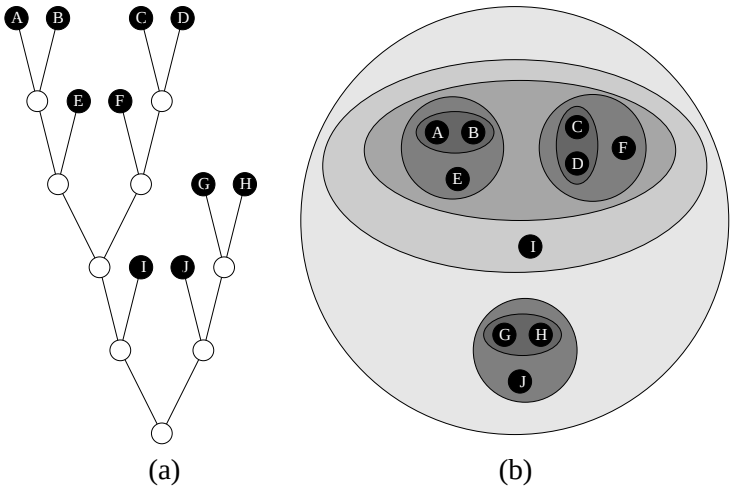


Figure 4.7. Community structure from the binary tree. The community structure represented by the binary community tree (a) can be regarded as a set of nested groups (b).

7, and 1 community of size 10. Note that a single node belongs, at different levels, to different communities.

The characterization of the community binary tree using the cumulative size distribution has its analogous in the river network literature (Rinaldo et al., 1993, Rodriguez-Iturbe and Rinaldo, 1996, Maritan et al., 1996). The equivalent measure is the distribution of drainage areas, that represents the amount of water that is generated upstream of a given point. Consider how the community size distribution is calculated. Assign, as shown in figure 4.8.a, a value 1 to all the leaves in the binary tree or, in other words, to all the nodes that represent single nodes in the original networks (black nodes). Then, the size of a community i, s_i, is simply the sum of the values s_{j_1} and s_{j_2} of the two offspring communities (or individual nodes), j_1 and j_2, in which i is split by the community identification algorithm. Figure 4.8.b shows how the drainage area of a given point in a river network is calculated. Consider that at any *node* of the river network there is a source of 1 unit of water (per time unit). The drainage area of a given point is the number of nodes upstream of it plus one. For a point i with offspring j_1 and j_2, $s_i = s_{j_1} + s_{j_2} + 1$. Therefore, the community size distribution would be equivalent to the drainage area distribution of a river where water is generated only at the leaves of the branched structure.

2.3.2 Horton-Strahler index and topological self-similarity

One of the most fundamental quantities developed to describe the topological properties of binary trees was introduced by Strahler (Strahler, 1952) as a refinement of the scheme proposed by Horton (Horton, 1945) to quantify the

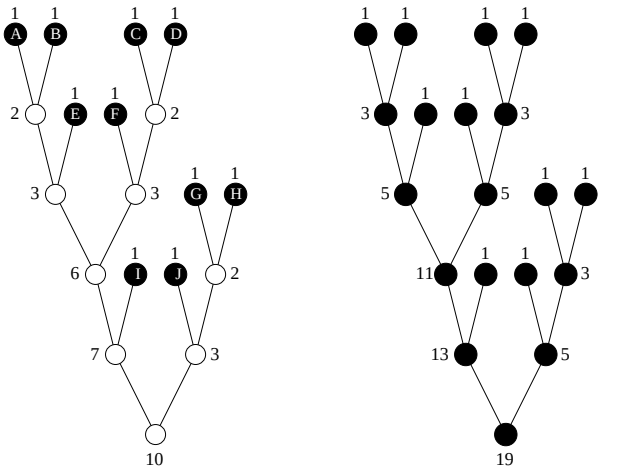


Figure 4.8. Calculation of the community size distribution and analogy with drainage area distribution in river networks. (a) Community sizes. A and B form a community of size 2. Together with E they form a community of size 3: this size is obtained by summing 1 from node E plus 2 from the community formed by nodes A and B. The procedure is repeated from the leaves downward, being the size of each community the sum of the sizes of the two offspring communities in the level immediately above. (b) Drainage area. The area drained by one node equals the number of nodes upstream from that node plus one. For a given node this area can be obtained summing up the areas of the two offspring nodes in the level immediately above plus one.

topological properties of river networks. Consider the binary tree depicted in figure 4.9.a. As shown in figure 4.9.b, the leaves of the tree are assigned Strahler index $k = 1$. For any other branch that ramifies into two branches with Strahler indexes k_1 and k_2, the Strahler index is calculated according to the following rule:

$$k = \begin{cases} k_1 + 1 & \text{if} \quad k_1 = k_2 \\ \max(k_1, k_2) & \text{if} \quad k_1 \neq k_2 \end{cases} \tag{4.1}$$

Therefore the index of a branch changes when it meets a branch with higher index or when it meets a branch with the same value and both of them join forming a branch with higher index (see figure 4.9.b).

The number of branches N_i with index i can be determined once the HS index of each branch is known . Note that, for this computation, a branch with many side branches of indexes smaller than its own index k, is counted as a single branch. Therefore, in figure 4.9.b, $N_1 = 10$, $N_2 = 3$ and $N_3 = 1$. The bifurcation ratios, B_k, are then defined as the ratio between the number of

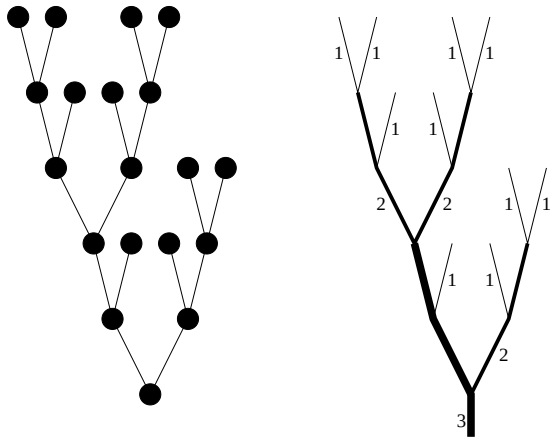


Figure 4.9. Horton-Strahler index. Arbitrary binary tree (a) and the corresponding values of the HS index of the branches. When two branches of size k meet, they give rise to a branch of index $k+1$. When two branches of sizes k_1 and k_2, with $k_1 > k_2$, meet the branch with index k_2 is absorbed by the branch with index k_1.

branches of size k and the number of branches of size $k+1$:

$$B_k = \frac{N_k}{N_{k+1}}. \tag{4.2}$$

To understand the meaning of this ratio, consider the branching structure of figure 4.9 and imagine that only branches of sizes 3 and 2 are present. The ratio B_3 indicates how many branches of size 2 appear from the branch of size 3. Similarly, if we consider only the branches of sizes 1 and 2, B_2 represents the number of branches that appear from each branch of size 2. When $B_k \approx B$ for all the indexes k, the structure is said to be topologically self-similar, because the overall tree can be viewed as being constructed of B trees, which in turn are constructed by B smaller trees with similar structures and so forth down to all scales.

Many systems in nature display topological self-similarity. Among others, some examples of topologically self-similar systems are river networks, with $3 < B < 5$ (Horton, 1945, Strahler, 1952), diffusion limited aggregates, with $B = 5.2$ in 2 dimensional spaces (Halsey, 2000), and random binary trees, with $B = 4$ (Halsey, 1997).

The meaning of the Horton-Strahler index in terms of communities and organization is less clear than the meaning of the community size distribution. Let us try to give an explanation. The index of a segment remains constant until another segment of the same magnitude is found. In other words, the index of a community changes when it joins a community of the same index. Consider, for

instance, the lowest levels: individuals (index $k = 1$) join in groups ($k = 2$); the index of a lowest level group ($k = 2$) will change when it joins another group to give a second level group, that is a group formed by different groups. Therefore, the index reflects the *level* of aggregation of communities. In the university, for example, one could expect to find the following levels: individuals ($k = 1$), groups ($k = 2$), departments ($k = 3$), faculties and schools ($k = 4$), and the whole university ($k = 5$).

3. Communities in informal communication networks: assessment of status and evolution of organizations

In the previous section we have shown how, from the original complex network, one can obtain the community binary tree using the GN algorithm. Moreover, we have shown that it is possible to quantify the community structure by means of measures carried out on this binary tree. In this section, we analyze in detail the e-mail network of the URV and show some interesting properties of its community structure.

3.1 Community analysis of the e-mail network of the Universitat Rovira i Virgili

First, the community structure of the Universitat Rovira i Virgili is analyzed. The network is built, as described in section 1, by considering all the e-mails that were sent within the university during January, February and March of 2002. For community identification purposes, the following considerations are taken into account:

- All e-mails sent to or received from addresses outside the university are discarded.
- Addresses corresponding to lists of users are discarded by identifying such addresses in the e-mail servers.
- E-mails sent to lists of users are discarded even when these lists are explicit lists of the users and do not correspond to any address in the e-mail servers. This is accomplished by disregarding e-mails sent to more than a certain number of users. In the present study, this maximum number of the receivers is set to 50.
- Only bidirectional e-mails are considered. Since the interest of the study is to identify *relevant* communication channels, a link between A and B is established if, and only if, A have sent an e-mail to B and B have sent an e-mail to A during the time period considered. Together with the elimination of user lists, this is a way to minimize the effect of massive and useless e-mail.

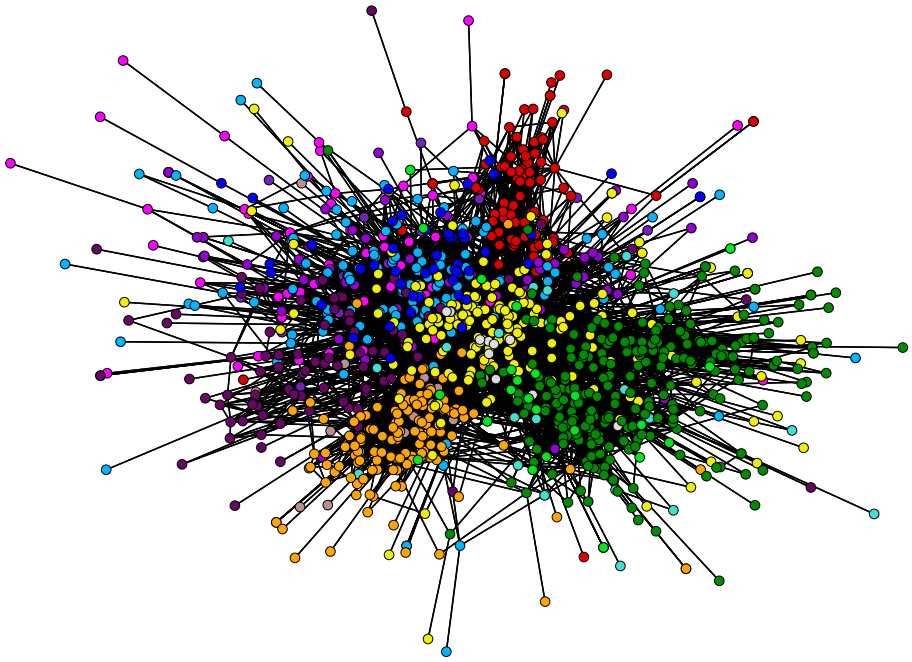


Figure 4.10. E-mail network of the Universitat Rovira i Virgili. The plot represents the main component of the e-mail network containing 1133 nodes and 5451 links. Only bidirectional e-mails are considered, and lists of size larger than 50 are also disregarded. Each color represents a center of the university.

With all these considerations, the main component of the e-mail network, depicted in figure 4.10, comprises 1133 nodes and 5451 links among them. As can be observed from the figure, the size and the complexity of the network prevent from any attempt to perform visual analysis of its properties, even when the nodes have been already plotted using the algorithm by Kamada and Kawai (Kamada and Kawai, 1989) to optimize the layout.

However, some initial hints can already be obtained from this plot. The university is divided into 18 different *centers*, including faculties or colleges, and management units such as the office of the Rector of the university. Nodes that belong to the same center are plotted in the same color. It is apparent from figure 4.10 that nodes in the same center tend to be close to one another in the e-mail communication graph as one would expect. Indeed, the high clustering that has been found in section 1 is related to the community structure as happens in other complex networks (Eckmann and Moses, 2002). It is also apparent that yellow nodes tend to occupy the central region of the graph, indicating that the corresponding center acts as a sort of hub for the other centers. This allows to

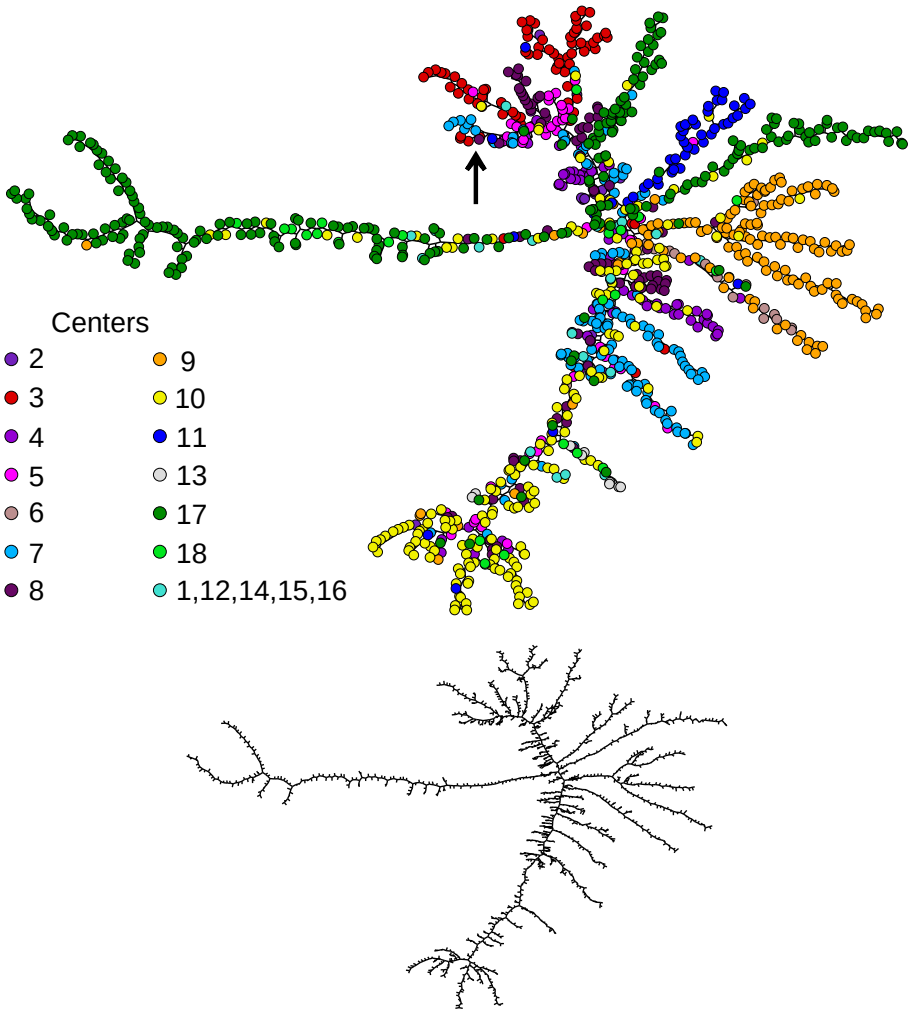


Figure 4.11. Community identification tree from the e-mail network. Each branch represents a community as identified by the GN algorithm. It is apparent that branches are mostly mono-color, indicating that the algorithm is actually successful in identifying the communities. The figure in the bottom shows more clearly the branching structure of the binary tree.

infer that yellow nodes represent individuals in the central management unit of the university.

It is hard to obtain more precise information from this preliminary plot of the e-mail network. The next step is therefore to apply the GN community identification algorithm and study the results. The binary community identification tree is depicted in a convenient way in figure 4.11. The root of the tree is indicated by the arrow in the upper left corner of the plot.

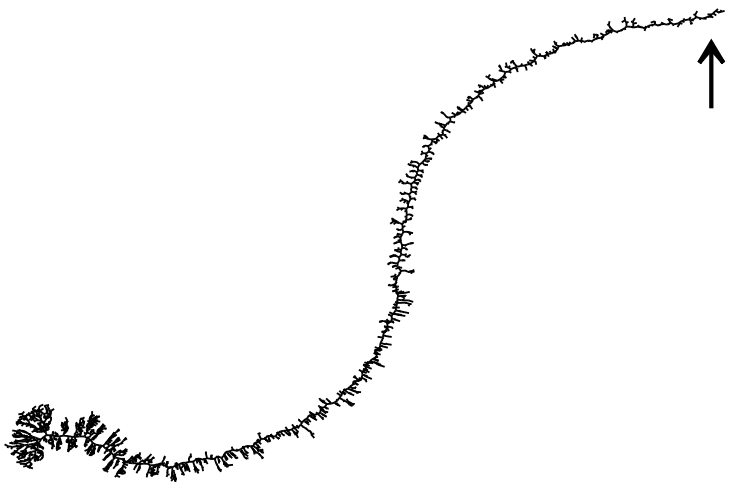


Figure 4.12. Community identification tree from a random network. The binary tree shows that there is not a community structure in the network, as one would expect.

With the discussion of the previous sections about the equivalence between branches and communities, the results of the community identification process plotted in figure 4.11 are quite convincing. The branches obtained by the GN procedure are essentially mono-color, indicating that we are correctly identifying communities. This is specially true if one focus in the ends of the branches since, as explained, this ends correspond to the most central nodes in the community. In regions close to the origin of the branches, the coexistence of colors correspond to the boundaries of the communities. It is worth insisting in the fact that the community identification has been carried out using *only* topological information from the e-mail communication network. Therefore, as speculated before, the structure of the communication network of an organization contains information about how groups and teams are formed and interact with each other, not only at the level of centers (in the case of the university) but also at smaller scales (subbranches inside branches). It is also worth noting that previous works on community identification using the GN algorithm dealt with much smaller networks (Girvan and Newman, 2002). For comparison, it is interesting to see what would happen if, instead of the real e-mail communication network, one would try to identify communities in a random network (Bollobas, 2001) with the same number of nodes and links than the original e-mail graph. The resulting binary tree is shown in figure 4.12. As expected, no communities are identified, and the contrast with the tree obtained for the e-mail communities is clear.

After the identification of communities, it is possible to study quantitatively the properties of both the original network and the binary tree, and therefore

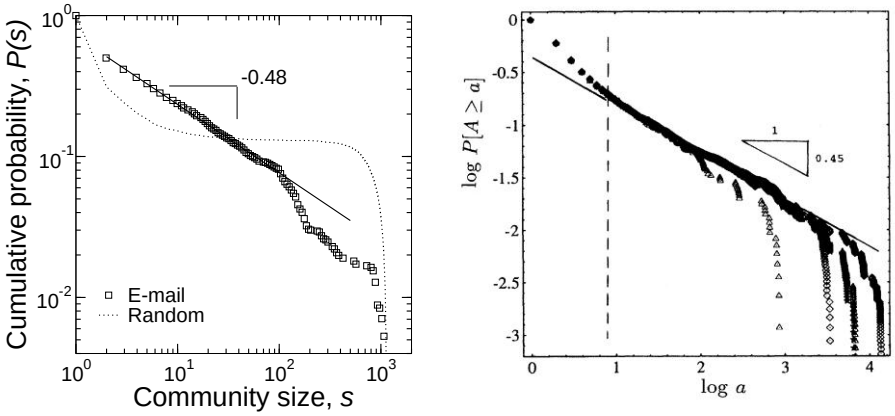


Figure 4.13. Community size and drainage area distributions. (a) Community size distribution for the e-mail network of the university. The distribution shows a power law region with exponent -0.48 between 2 and 100, followed by a sharp decay at 100 and a cutoff at 1000. The dotted line represents the community size distribution for the random graph. (b) Drainage area distribution for the river Fella, in Italy, and some of its affluents (sub-basins). Consider, for example, the triangles in the figure, that correspond to a sub-basin of approximately the size of the e-mail network. The distribution also shows a power law region with exponent -0.45, followed by a sharp decay at 100 and the cutoff at 1000 (figure taken from (Maritan et al., 1996)).

quantify concepts that have been traditionally used in a qualitative way in the management literature. Also, it will be possible to find properties of the network that will give hints about how organizations evolve. In the next two sections both problems are covered. First, some properties of the binary tree are studied to conclude that organizations show surprising emergent scaling and self-similarity properties. Second, quantitative measures on the network and the binary tree are related to concepts from organizational design and management.

3.2 Self-similarity properties in the community structure

In this section, some emerging properties of the community structure of the e-mail network are discussed. First we focus on the community size distribution as defined in section 2.3.1, and in the analogy existing with river networks. Figure 4.13.a shows the cumulative distribution of community sizes, that is, the probability $P(s)$ that the size of a community is larger than s. Between $s = 2$ and $s = 100$, the distribution is well fitted by a power law $P(s) \propto s^{-\alpha}$ with exponent $\alpha = 0.48$. At $s \approx 100$, $P(s)$ shows an abrupt decay, and at $s \approx 1000$ the distribution shows a cutoff that corresponds to the size of the system (the whole network contains 1133 nodes). The power law of the community size distribution suggests that there is not a characteristic community size in the

network (up to size 100). The community size distribution corresponding to the random graph (with the same size and connectivity than the e-mail network) shows a completely different behavior. In this case, there are no communities of sizes between 10 and 600, as observed in the plateau existing in the cumulative distribution between these to values.

The similarity between the community size distribution of the e-mail network and the area distribution of a river network is striking. Consider figure 4.13.b, that represents the drainage area distribution of the river Fella, in Italy, and in particular the triangles that correspond to a sub-basin on the river of approximately the same size (number of nodes) than the e-mail community tree. As observed, the area distribution shows a power law behavior at low values of the area (but larger than a certain lower threshold which is approximately 10). Then there is an abrupt decay at $a \approx 100$ and, at $a \approx 1000$, the cutoff corresponding to the size of the system. Moreover, the exponent of the power law region is $\alpha_{river} = 0.45$, very close to the value $\alpha = 0.48$ obtained for the community tree. Similar exponents have been obtained for many other rivers around the world (Rodriguez-Iturbe and Rinaldo, 1996, Maritan et al., 1996).

After discovering the functional analogy between the community size distribution and the drainage area distribution of river networks, one question arises: is it just chance or are there other properties shared by community trees and river networks? To answer this question the Horton-Strahler index is studied next.

As explained in section 2.3.2, for a self-similar tree the bifurcation ratio $B_k = N_k/N_{k+1}$ is independent of the index k: $B_k = B,\ \forall k$. In this case, it is straightforward to show that the number of segments of index k is given by

$$N_k = \frac{N_1}{B^{k-1}}. \tag{4.3}$$

The number of branches of index k, N_k, is measured for the binary community tree and the results are shown in figure 4.14. As observed, equation (4.3) fits perfectly the points obtained from the e-mail community tree with a bifurcation ratio $B = 5.76$. This value of the bifurcation ration is large compared to other systems displaying topological self-similarity. Indeed, for river networks $3 < B < 5$ and in diffusion limited aggregates in two dimensions $B = 5.2$. This means that the community tree is more bifurcated than the others. It is also remarkable that we find 5 levels in the tree, that could perfectly correspond to the levels outlined in section 2.3.2: individuals ($k = 1$), groups ($k = 2$), departments ($k = 3$), faculties and schools ($k = 4$), and the whole university ($k = 5$).

Once more, it is interesting to compare these results with those obtained for the community tree corresponding to the random network. As it is also shown in figure 4.14, topological self-similarity does not hold in this case, since the points

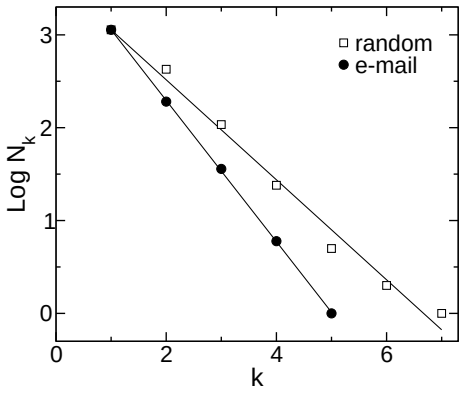


Figure 4.14. Topological self-similarity of the community binary tree. Filled circles represent the number of segments of Horton-Strahler index k, as a function of k, for the e-mail community tree. The fact that $\log N_k$ decreases linearly with k shows that the tree is topologically self-similar. This linearity does not hold for the tree obtained from a random network (void squares).

do not fall in a straight line, although N_k is still a monotonously decreasing function of k. In any case, the best fit of equation (4.3) yields a much smaller bifurcation ratio $B = 3.46$.

Summarizing: by means of the community size distribution and the Horton-Strahler index, it has been shown that the community tree displays non-trivial emerging properties. The comparison with the random case allows to conclude that these properties are not a consequence of the community identification algorithm. Rather, they must be related to the community structure of the organization considered. Then, it is natural to wonder why such properties emerge from the working relationships that the individuals have grown mostly *locally*—that is, without considering the whole organization. A similar question has been posed in river networks: while scaling and self-similarity emerge from local erosion and flow rules? In the latter case, it has been shown that the scaling relations and the self-similarity actually yield optimal networks, meaning that they optimize some global quantity such as the energy dissipation. Moreover, it has also been shown that local erosion rules can actually lead to these global optimum (Rinaldo et al., 1993, Sinclair and Ball, 1996). This fact suggests that the community organization could also be growing in such a way that some global quantity (maybe related to communication flow) is being optimized.

Beyond the fundamental interest of understanding how organizations are assembled and grow, the study of these properties can probably help managers. If the existence of emerging scaling and self-similarity properties are related to optimality as happens in other natural systems (Banavar et al., 1999), it should be possible to relate quantities such as the exponent α or the bifurcation ratio B to measures of performance and efficiency. Also, they could be used to measure

evolution of structures: highly evolved structures showing nice emerging properties in contrast to poorly evolved ones. These speculations would certainly require more investigations. Study of other organizations (other universities, firms, etc.) in different cultural environments and with different traditions are needed before the conclusions proposed here can be considered definitive. However, these results open an interesting line of research that deserves interest from both the theoretical/fundamental and the managerial points of view.

3.3 Communities and management

From a managerial point of view, there are many measures on the communication network and on the community binary tree that can provide valuable information. While the measures presented in the previous section would only be useful in an indirect way, here we present some examples of measures with direct applicability to management.

3.3.1 Levels of organizational complexity

The HS index also turns out to be an excellent measure to assess the levels of complexity in organizations. First, let us consider the interpretation of the index in terms of communities within an organization. The index of a branch remains constant until another segment of the same magnitude is found. In other words, the index of a community changes when it joins a community of the same index. Consider, for instance, the lowest levels: individuals ($i = 1$) join to form a group (or team, with $i = 2$), which in turn will join other groups to form a *second level* group (or department, $i = 3$). Therefore, the index reflects the *level* of aggregation of communities. For example, in URV one could expect to find the following levels: individuals ($i = 1$), research teams ($i = 2$), departments ($i = 3$), faculties and colleges ($i = 4$), and the whole university ($i = 5$). Strikingly, the maximum HS index of the community tree is indeed 5, as shown in figure 4.14.

Figure 4.15 shows the community tree of the e-mail network with different colors for different HS indices. This helps to distinguish the individual, team and department levels within a branch. Actually, the *university level* is the "backbone" of the network along which the separation of communities occurs (from the top to the bottom of the figure). From this backbone, colleges, departments and some research teams separate, although it is worth noting that colleges or, in general, centers which are small and have no internal structure will be classified with a HS index corresponding to a department or even a team. Therefore, the HS index does not represent administrative hierarchy but organizational complexity. For comparison figure 4.15c shows in color the HS index for the binary tree of a random graph.

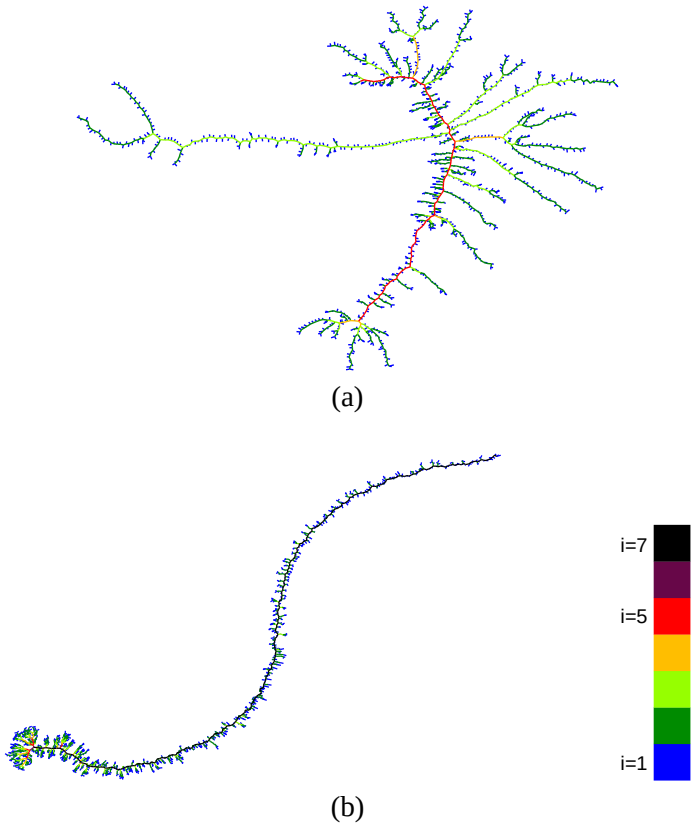


Figure 4.15. (a) Binary community tree for the e-mail network as in figure 4.11, but without showing the nodes so that the structure of the tree is clearly shown. Branches are colored according to their Horton-Strahler index (b) Binary tree for a random graph with the same size and connectivity than the e-mail network. Again, colors correspond to Horton-Strahler indices.

3.3.2 Measures of interaction within the organization

The original e-mail network also contains information about interactions within the organization. We propose two indicators to measure interactions in the original network: (i) the average distance between centers (formal university communities) and (ii) the probability of a node being connected to nodes in other centers.

First, we focus on the average distance between centers. We take each node in the network and measure the number of steps across the e-mail network needed to reach any other node. Then we average over all the nodes in the same center and obtain average distances between centers. To visualize this information, we proceed as follows. First, we calculate the distance from one center A to all other centers, d_{AB}, d_{AC}, etc. Then we compute the average

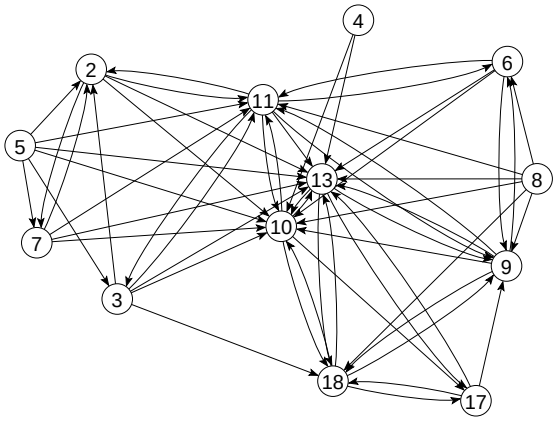


Figure 4.16. Inter-center relations from distances in the e-mail network. A directed link is established from A to B when the average distance between nodes in A and B is short (see text). Five small centers with less than 10 persons have been disregarded.

distance from A to the other centers $\langle d_A \rangle$. Finally, node A (that now represents a center, not an individual) is linked to another node B if $d_{AB} < \langle d_A \rangle$. In this case, the network is directed because, in general, $d_{AB} < \langle d_A \rangle$ does not imply $d_{BA} = d_{AB} < \langle d_B \rangle$. The result is shown in figure 4.16. Numbers in this figure correspond to the numbered communities (colors) in figure 4.15. According to the figure there are three central communities (10, 11 and 13) that would correspond to central offices and administrative centers of the university. These three interact on the left with a group of four centers and on the right with another one formed by five centers. There is one center that is only connected to two of the central nodes and somehow isolated from the rest of the university. No further comments can be made here due to confidentiality constrains.

The second important aspect is the probability of being connected to nodes in other centers. For each node in the network, we just regard its neighbors and, again, we average over all the nodes that belong to the same center. Two typical cases are shown in figure 4.17. Center 13 is one of the three central nodes in figure 4.16. As can be seen from figure 4.17, individuals that belong to center 13 are connected with a reasonably high probability not only to other individuals in the same center but also to individuals belonging to most of the other centers. Conversely, individuals in center 4 are mostly connected to others from the same center and also to individuals in center 10, that has been already identified as a central management unit. Extreme cases of these two patterns could be considered pathological: groups with lots of outside connections and very few internal connections are said to show anomalous communication patterns,

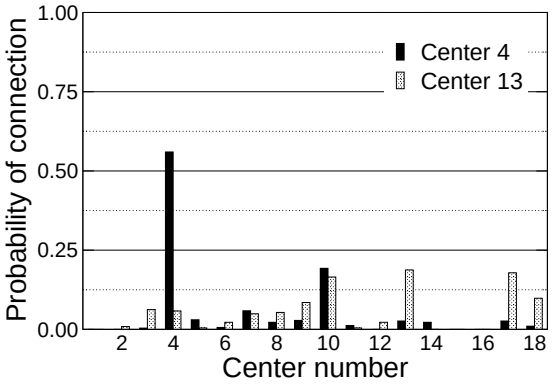


Figure 4.17. Probability of being connected to nodes that belong to other centers. Each bar represents the probability of a node in center 4 or 13, to be connected with a node in another center.

while groups with an extremely high fraction of internal connections but weakly connected to other groups are said to show imploded relationships (Krackhardt and Hanson, 1993).

4. Summary

In this chapter we have shown how to extract valuable information describing real complex communication networks behind the formal chart of an organization. We take advantage of the automatic registration of communication processes, in particular e-mails log files, to reconstruct the real network of interactions within the organization. The structure of this complex network has been unraveled by the identification of the whole hierarchy of communities using the Girvan-Newman algorithm. We have proposed a representation procedure that allows the identification of these communities by visual inspection. Moreover, we have suggested measures that allow to characterize quantitatively the community structure of the organization. As a real case study, we have studied the e-mail network of the University Rovira i Virgili. From this analysis, we have been able to identify the real organization of the individuals of the university into working teams, departments, faculties or colleges, and the whole university, as well as the interrelations between them.

From a theoretical point of view, the methodology identifies emerging scaling and self-similarity properties in the community structure as happens in some other self-organized systems such as river networks. This result that opens interesting questions about the mechanisms underlying the interactions between individuals within an organization and, thus, the formation of complex networks. Self-similarity is a fingerprint of the replication of the structure at different levels of organization, and could be the result of the trade-off between

the need for cooperation and the physical constrains to establish connections at any organizational level (in our case study: individuals, research teams, centers, etc.). At the same time, the similitude with river networks suggests that a common principle of optimization (of flow of information in our case or of flow of water in rivers) could be the underlying driving force in the formation and evolution of informal networks in organizations.

Finally, we have shown that the study of the informal communication network using tools recently developed in the field of complex networks theory could be useful for management purposes, for example, to assess formal charts or to measure the degree of attainment over time of proposed organizational changes.

Chapter 5

CONCLUSIONS AND PERSPECTIVES

1. Conclusions

In this theses we have studied, from both a theoretical and an empirical points of view, the role of communication and information processing in organizations and its implications for organizational design. The following conclusions can be drawn from the work.

- Any communication process with a stochastic component in which agents have limited capability to handle information packets, that is discrete pieces of information, gives rise to the formation of queues or, in other words, to the accumulation of packets waiting to be delivered. When the average number of packets that nodes can deliver during a period is bounded, the formation of queues results at some point in the transition to a collapsed state in which packets have to wait, on average, an infinitely long time to be delivered. The transition is tuned by the probability of packet generation, ρ: for small values of ρ, the traffic in the network is light and there is no collapse; for ρ above a certain point ρ_c, the network collapses. The critical value ρ_c is a measure of the amount of information the communication network (say the organization) is able to handle without collapsing.

  When the average number of packets delivered per time period by a node is fixed and independent of the load of the node, the transition to the collapsed state is a continuous phase transition and the total load of the network, the fluctuations of the total load, the average characteristic time, and other related quantities diverge. The transition is properly characterized considering that the ratio of accumulation of packets is the order parameter. When the average number of packets delivered per time period by a node is a decreasing function of the load of the node, the transition to the collapsed state is

discontinuous, and collapse arises in several nuclei that finally spread over the whole network.

- We focus on a situation where nodes deliver, on average, a fixed number of packets and where the roles of the sender and the receiver of the information are completely symmetric. Moreover, we consider hierarchical networks in which nodes have complete knowledge of the subtrees below them. In this situation, mean field estimations of the critical collapse point ρ_c are obtained. These analytical expressions are in excellent agreement with simulations of the communication model, and show that for hierarchical networks, the optimal design is the flattest possible one with only one node at the top and all the others connected to it. When one considers that keeping communication channels open has a cost for agents, the optimal hierarchical structure is not the flattest one in general. Rather, there is an optimal *span of control* which is larger as the communication technology improves, as observed in real organizations.

- Consider a situation in which the nodes do not have global knowledge of the structure of the network. In such a scenario, the effects of *congestion*, as discussed in the previous items, and *search* for the destination of the packets coexist. In general, from a search point of view centralization is positive, but it is negative from a congestion point of view. As far as optimal network designs are concerned, this trade-off between search and congestion results in a transition from centralization to decentralization.

  We have proposed a formalism that allows to cope with the problem of search in presence of congestion analytically. Moreover, we have found that the optimal network topologies for local search considering congestion are split in two categories: star-like networks, that are optimal for small number of parallel searches, and homogeneous-isotropic networks, that are optimal for large numbers of parallel searches. Strikingly, the transition between these categories is sharp, i.e. we are not able to find any optimal network topology different from these two classes.

- We have studied the communication network (in particular, the e-mail network) of a real organization with almost 1,700 employees: the University Rovira i Virgili. It has been shown that the application of community identification algorithms developed recently in the literature of complex networks is very successful in identifying the existence of centers, departments, and even research teams, provided that the data is treated conveniently to eliminate massive and *spam* e-mail. The new methodology proposed allows to identify all the communities mainly by visual inspection.

  Moreover, we have been able to characterize quantitatively the community structure. Our results reveal the emergence of self-similar properties that

suggest that some universal mechanism could be the underlying driving force in the formation and evolution of informal networks in organizations, as happens in other self-organized complex systems. It is worth noting that the quantitative analysis of the community structure is a useful tool for managers. In particular, we have shown that some concepts used frequently in the management literature can be quantified with the proposed methodology, and that large scale studies of organizations can be performed.

2. Perspectives

Although the objectives fixed at the beginning of the thesis have been mostly accomplished, the research developed has opened a number of interesting lines that are worth considering in the future. The most relevant of them are outlined next.

- With the study of the communication models in chapter 2, we have been able to get insights in the dynamics of communication processes. In particular, we have established that, for *regular* queue systems, the transition to collapse is a continuous phase transition. Our study has also shown that changes in the topology of the network result in changes of the critical properties of the congestion phenomena and, singularly, of the critical exponents that describe how quantities such as the load or the delivery time diverge near the collapse point. It is probably worth applying all the knowledge existing in the theory of critical phenomena to design better communication protocols (specially in computer based communication networks).

  For example, classical queuing theory assumes that the average delivery time diverges as $(1-\rho/\rho_c)^{-1}$, where ρ/ρ_c is the utilization ratio of the network. We have shown that, in some situations the divergence can be much more abrupt with an exponent of even -2.5 instead of 1. This has consequences regarding how packets are resent in a real communication protocol.

- We have faced the problem of optimal organization design in situations where agents have purely local knowledge of the communication network. However, the formalism developed allows to deal with other situations, including global knowledge and any intermediate situation between purely local search and global search.

  Different knowledge scenarios will result, probably, in different optimal networks and even in different scenarios for the transition from centralization to decentralization, as already pointed out in chapter 3. It would also be very interesting to study in which situations real networks (with scale-free and small-world topology) might be optimal or almost optimal.

- Still regarding optimal communication networks, the formalism we have proposed open new doors for a deeper analytical treatment of the complex

problem of finding optimal structures. In particular, it could be interesting to apply tools from physics of disordered media to obtain and characterize optimal networks.

- Finally, from an empirical perspective, it would be extremely interesting to test the results obtained from the theoretical optimization of communication networks in a real environment. For example, it can be useful to test in a controlled experiment whether centralized structures perform better than decentralized and in which conditions they do. Although similar experiments have been carried out and are in the literature, it would be worth to use the results of the theoretical optimization analysis to design and guide new experimental setups.

- The empirical analysis of the e-mail network has also opened a very interesting research line in the interface between complex networks theory and management. The work can be extended at least in two directions: one methodological and another one applied.

  Although it is true that the main components of the analysis methodology have been established, it is possible, for example, to refine the community identification algorithm or to propose new ways of characterizing the community structure. In particular, it is necessary to take into account that some of the techniques used so far (singularly the Girvan-Newman algorithm) were not specially devised to deal with community identification *in organizations* and therefore are susceptible of being improved using heuristics or components that can only be applied to this very particular problem, and not to other community identification problems in other areas.

  Also interesting, specially from a managerial perspective, is the applied line. The present study of the e-mail network of the University has shown some properties of the organization that might very well be universal, as happens in other complex systems that display self-organization. However, comparison of our results with similar studies in different environments are needed to be conclusive on this. Moreover, comparative studies would definitely help to understand the impact of different elements. For instance, it would be possible to determine quantitatively which is the influence of the cultural environment in which the study is carried out, the difference between public and private organizations, the evolution in time of a given organization, the level of attainment of a intended redesign process, etc. It would also be interesting to compare the results obtained for the e-mail network with those obtained for the phone-call network, or the internal mail network, for which it is also probably relatively easy to obtain massive data.

Appendix A
The "web of trust"

Electronic communication networks are examples of communication networks that have experienced an amazing growth in the last years, and that are easy to quantify and study. However, these networks have some special properties as compared to other types of communication networks. Since information travels, in general, through a publicly available space (the Internet, for instance), privacy can be forged relatively easy. The attempt to avoid the violation of privacy has let to the creation of encryption algorithms that ensure that messages can only be read by the desired user. In the Pretty-Good-Privacy (PGP) algorithm this is accomplished by a pair of keys for encryption: a public key, which encrypts data, and a corresponding private, or secret key for decryption (Garfinkel, 1994, Stallings, 1995). Users publish their public key to the worldwide keeping their private key secret. Anyone with a copy of the public key of user A can encrypt information that only A can read. It is computationally infeasible to deduce the private key from the public key and, therefore, only the person who has the corresponding private key can decrypt the information.

Used in the inverse way, public key cryptography also provides a method for employing *digital authentication*. Digital authentication enables the recipient of information to verify the authenticity of the information's origin, and also verify that the information is intact. The basic manner in which digital signatures are created is the following. Instead of encrypting information using someone else's public key, user A encrypts the information with his or her private key. If the information can be decrypted with his or her public key, then it must have originated with A. This authentication process requires that the recipient B is able to check, at least the first time that he or she receives information from A, that A is indeed who claims to be. This is accomplished by signing A's public key. When B signs the public key of user A, it means that user B trusts that A is indeed A, and in the future B will have the certainty that information allegedly sent by A has indeed been sent by A.

One interesting point is that signatures of public keys are stored in some servers and made publicly available. One may regard each key as a node, and each signature to define a directed link between nodes; the resulting directed graph is known as the "web of trust". Because of the way it is formed, this web is a communication network since signature of keys means potential interchange of secure information. Therefore, the study of the web of trust should help in the understanding of other communication networks.

First, let us focus on the growth of the web of trust. The whole network contains a lot of keys and many signatures among them, forming a large and complex web. However, the web

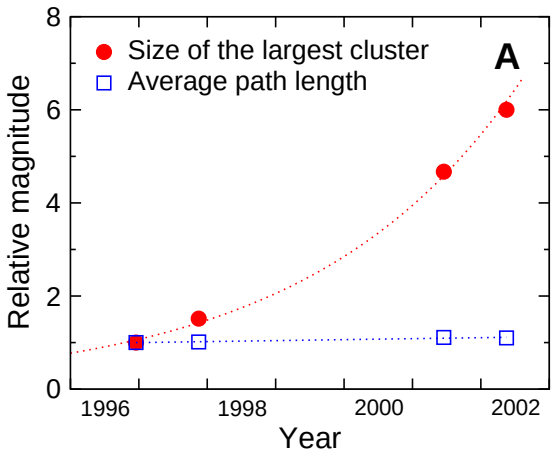


Figure A.1. Growth of the web of trust. The circles represent the growth in the number of keys that belong to the largest cluster. This growth is compatible with an exponential function (dotted line). The squares represent the growth of the average distance between nodes. As observed, this growth is much slower than that of the size, indicating that the web of trust has grown efficiently and nodes are still only a few steps away from each other. For both the size of the largest cluster and average path length, the representation shows the relative magnitude with respect to December 1996 as a reference.

is not connected or, in other words, is formed by many different clusters containing a variable number of keys. In a cluster, it is possible to jump from any node to any other node following the directed links. Using the precise language of graph theory, our clusters are indeed *strongly connected sets*.[1] In terms of trust, one can only be confident in the nodes that belong to the same cluster: if A trusts B, and B trusts C, A can, in principle, trust C. Therefore, a good measure of the efficiency of the web of trust is the size of the largest cluster and an ideal situation would be that millions of persons belong to the same cluster. The real situation is still far from this, but during the last 5 years, the largest cluster of the web of trust has grown by a factor of 6, as shown in figure A.1. Actually, the growth is compatible with an exponential function, although the number of data points available is probably too small to be conclusive.

The size of the largest cluster, however, is not the best measure of efficiency. Although in principle the trust relation is transitive (if A trusts B and B trusts C, then A can trust C), some experts state that one can only trust individuals directly, i.e. first neighbors in the network or, at most, individuals at a distance small enough. Quite surprisingly, the web of trust has also grown efficiently in this more restricted sense since, as shown also in figure A.1, the distance between nodes has grown only very slightly (a factor of 1.1) in the 5 years studied, even when the size of the largest cluster has been multiplied by 6.

By studying the sizes of the different clusters that form the network, it is possible to get more insights on how the network is growing. Unfortunately, information about the different clusters is not registered in the elder databases (those from 1996 and 1997). However, newer databases contain information that allow to retrace the evolution of the web of trust. In particular, we focus

[1]Weakly connected sets are formed by nodes that are mutually reachable but considering that all links are bidirectional.

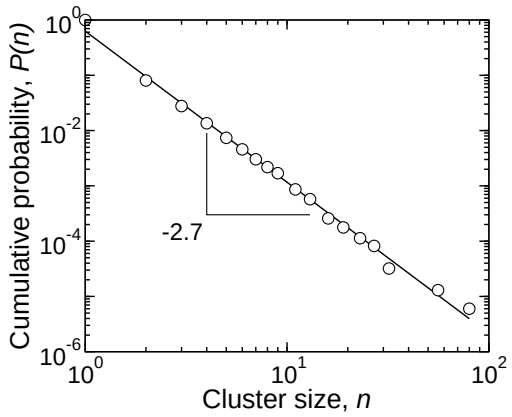


Figure A.2. Cluster size distribution for the web of trust. The web of trust is not a connected network but contains many disconnected clusters. The largest one contains 9652 nodes. The points represent the cumulative distribution of sizes of the rest of the clusters, and the straight line is a power law fit, with exponent -2.7.

in the network as recorded at http://dtype.org on July 2001, when it comprised 191,548 keys and 286,290 signatures. This 191,548 keys include only keys with at least one signature; many others without any signature are discarded. From this database, the distribution of cluster sizes, $P(n)$, is calculated as shown in figure A.2. The distribution follows a power law with exponent -2.7, $P(n) \propto n^{-2.7}$, between 2 and 100 approximately. Beyond these small clusters, the largest one contains (usually called the main component) 9562 keys (nodes). The power law distribution of cluster sizes suggests that the largest cluster has emerged as a result of a percolation transition (Stauffer and Aharony, 1992).

The next step is to study in detail the structure of the web of trust, and in particular its state as recorded, again, at http://dtype.org on July 2001. The average path length between nodes is small: in the largest cluster, that contains 9562 keys that have signed 5.80 different keys on average, the mean distance is 6.58. More surprisingly, as shown in figure A.3, it is found that the in- and out-degree distributions $P(k)$, that is the distribution of number of incoming or outgoing connections k_{in} and k_{out}, have scale-free power law decays, $P(k) \propto k^{-\gamma}$ with exponents $\gamma_{in} = 1.8$ and $\gamma_{out} = 1.7$ as happens in other complex networks (Barabasi and Albert, 1999, Amaral et al., 2000). In particular, at least two different computer based communication networks have been shown to display a similar scale free behavior: the e-mail network of the Kiel university in Germany (Ebel et al., 2002), and the instant messaging network (Smith, 2002). For the first one, only the total number of links was considered (that is, the study did not distinguish between in-coming and out-going links), and the exponent found was quite small, $\gamma = 0.81$, at least in a region including nodes with degrees between 5 and 100. In the last one, the exponents were $\gamma_{in} = 1.2$ and $\gamma_{out} = 1.4$.

The existence of a scale free degree distribution has been shown to have important implications. First, Albert and coworkers (Albert et al., 2000) showed that scale free networks are robust against random elimination of their nodes but very fragile against selective removal of the most connected nodes. Second, Pastor-Satorras and Vespignani (Pastor-Satorras and Vespignani, 2001) demonstrated both numerically and analytically that virus (either biological or electronic) spread very easily in such networks.

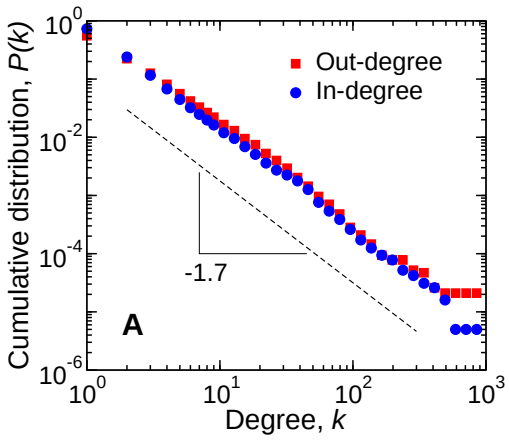


Figure A.3. Cumulative degree distribution for the web of trust. Both the in-degree and the out-degree distributions show a power law decay in nearly three decades. The exponent is approximately -1.7.

However, the web of trust has some properties that have not been reported in other scale-free networks, specially related to the clustering coefficient. As reported in the introduction, in a random graph, the clustering coefficient, C, is given by

$$C_{random} = \frac{\langle k \rangle}{S} \tag{A.1}$$

where $\langle k \rangle$ is the average connectivity and S is the number of nodes in the network. Similarly, in a scale free network as proposed by Barabasi and Albert (Barabasi and Albert, 1999), C decreases very fast with the system size (Klemm and Eguiluz, 2002)

$$C_{BA} \propto \frac{(\log S)^2}{S} \tag{A.2}$$

The fact that the network is not connected but formed by a collection of separate clusters allows a study of the clustering coefficient as a function of the system size. Figure A.4 shows the results. Independently of the number of nodes in the cluster, the clustering coefficient is approximately constant as happens both in the small world model of Watts and Strogatz and in low dimensional lattices (Watts and Strogatz, 1998).

Although it is usually implicitly assumed that the degree distribution contains all the information necessary to understand the network behavior, the correlations in the establishment of connections (quantified, in this case, by the clustering coefficient) have important implications. In the case of the web of trust, the response to intentional attacks (i.e. removal of the most connected nodes) is modified by the existence of highly interconnected groups of nodes. Figure A.5 shows a comparison between the response to attack of the web of trust and of a random graph with exactly the same degree distribution. The figure shows the behavior of the largest strongly-connected cluster of the web of trust and of a random graph with the same in- and out-degree distributions that the web of trust. Initially, both graphs have 9562 nodes with an average degree of 5.80. As the fraction f of nodes removed increases, the cluster is split into smaller components. Figure A.5.a shows the relative size S of the largest strongly-connected cluster, and A.5.b the average size $\langle s \rangle$ of the other clusters. Note that for the web of trust the largest strongly-connected cluster breaks down faster but that the other strongly-connected clusters have average

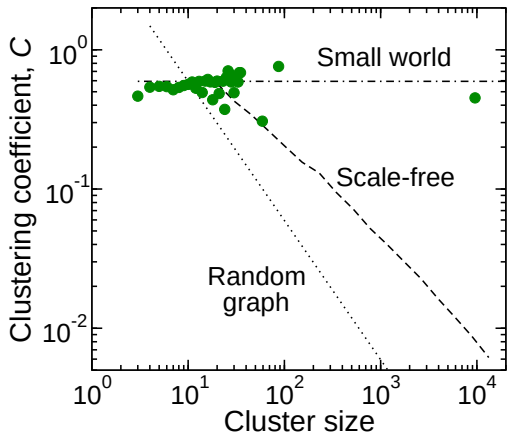


Figure A.4. Clustering coefficient for the different clusters in the web of trust. As expected for a low dimensional lattice and for a small world network a la Watts and Strogatz, the clustering coefficient is essentially independent of the size of the cluster. For a Barabasi-Albert scale free network and for an Erdos-Renyi random graph the clustering coefficient would decay very fast as the cluster size increases. In particular, for a model scale-free with the same connectivity and size than the largest cluster, the clustering coefficient would be approximately 100 of times smaller than its actual value.

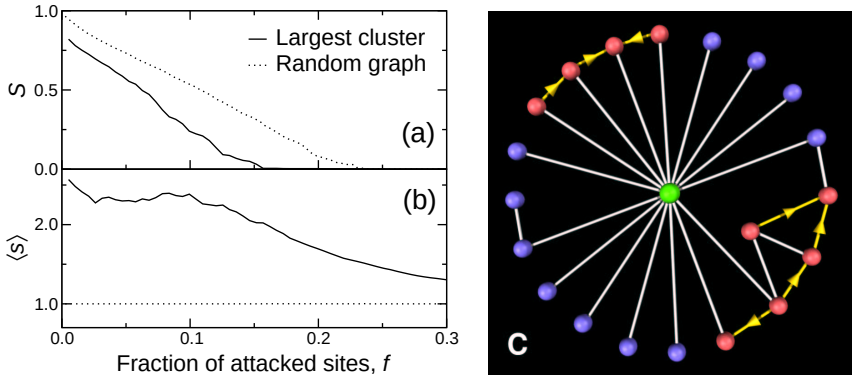


Figure A.5. Structure and resilience of the web of trust. (a) and (b) Intentional attack on the nodes with the highest in-degree of the largest cluster of the web of trust (full line) and a random graph with the same in- and out-degree distributions (dotted line). (a) Relative size S of the largest strongly-connected cluster. (b) Average size $\langle s \rangle$ of the other strongly-connected clusters. (c) A strongly-connected cluster comprising 21 nodes. White lines indicate bi-directional links while yellow arrows indicate unidirectional links. This cluster is strongly connected because every node is reachable from any other node. The red nodes indicate the groups that give rise to a large clustering coefficient.

sizes that remain unchanged up to the total destruction of the largest strongly-connected cluster. For the random graph, the small clusters formed by removing nodes are almost all isolated nodes,

which explains the slower decrease of S and also the constant value of $\langle s \rangle = 1$. Therefore the web of trust is disintegrated more easily but the low level structure remains essentially unaltered. The explanation to this is related to the structure of the web of trust (figure A.5.c), that contains highly interconnected groups that contribute to the clustering, and a few hubs organized in a hierarchical fashion that give rise to the scale free degree distribution.

Resum de la Tesi

1. Introducció

La companyia química típica és una companyia gran, sovint amb milers de treballadors. Segons dades de la Unió Europea, l'any 1990 quasi el 70% de la facturació en el sector químic corresponia a empreses amb més de 250 treballadors. La resta se la repartien a parts iguals empreses petites, de menys de 50 treballadors, i mitjanes, d'entre 50 i 250 treballadors. De fet, encara que alguns productes tenen mercats d'abast regional, la indústria química és essencialment global i la dominen multinacionals com ara Bayer, amb 117.000 treballadors, BASF, amb 93.000 treballadors, DuPont, amb 79.000 treballadors, o Dow Chemical, amb 50.000 treballadors. Especialment per aquestes grans multinacionals, el disseny organitzatiu i la gestió dels recursos humans juguen un paper clau, tan important, si més no, com la tecnologia o la gestió dels recursos materials.

L'estudi teòric de les organitzacions i el seu disseny han estat tradicionalment tractat dins l'àmbit de l'economia. La majoria de treballs en la literatura econòmica s'han concentrat en problemes relacionats amb incentius, en part perquè existeix un marc teòric ben establert que permet tractar aquesta mena de problemes. En els últims deu anys, però, s'ha proposat i desenvolupat un nou enfoc del problema de l'organització: el que podríem anomenar l'enfoc *comunicacional*. La idea essencial és que les organitzacions existeixen, en bona mesura, per resoldre el problema de coordinació que planteja la limitació de capacitat dels individus per processar informació. Efectivament, l'especialització sorgeix com a conseqüència d'aquesta limitació i, simultàniament, la necessitat de coordinació i comunicació sorgeix com a conseqüència de l'especialització. Per altra banda, és un fet establert empíricament que bona part del treball en una organització consisteix en processar informació i comunicar més que no pas en *fer* o *vendre* coses en el sentit més estricte. Des d'aquesta perspectiva, en l'enfoc comunicacional s'entén que l'organització és un processador d'informació i, per

tant, s'estableix que el problema de disseny és un problema de minimització dels costos de comunicació i d'optimització del flux d'informació.

Paral.lelament, l'aparició i el rapidíssim desenvolupament de xarxes de comunicacions tecnològiques com ara l'Internet, i també la complexitat de la seva estructura i la seva dinàmica, han despertat l'interès de la comunitat científica, en particular la de la física estadística. L'estudi de xarxes (o grafs) ja era una matèria amb entitat pròpia en l'àmbit de la sociologia o de la matemàtica, però el descobriment de propietats sorprenents en grans xarxes complexes reals ha portat a estudiar sistemes propis de la biologia o de l'enginyeria des d'una nova perspectiva.

En aquest treball fem servir idees i eines tant de la literatura econòmica com de l'emergent *física de les xarxes complexes* per entendre el paper de la comunicació i del processament d'informació en organitzacions. Ho fem des d'una doble perspectiva: teòrica i empírica. Des del vessant teòric, proposem i estudiem un model general i simple per processos de comunicació. Amb el coneixement adquirit del model, ataquem el problema de trobar estructures de comunicació òptimes. Des del vessant empíric, estudiem les xarxes de comunicació d'organitzacions reals i en traiem informació sobre l'estructura de comunitats, informació que es pot fer servir com a indicador quantitatiu de l'estat i evolució de l'organització.

2. Modelització de processos de comunicació

Un cop emfatitzada la importància que l'intercanvi i el processament d'informació tenen en l'anàlisi teòrica de les organitzacions, en aquest capítol es proposa i s'estudia una col.lecció de models simples i generals pels processos de comunicació. Els models inclouen només els ingredients bàsics que prenen part en qualsevol comunicació entre dos agents: (i) els paquets d'informació que hom vol transmetre, (ii) els canals de comunicació a través dels quals els paquets han de ser transmesos i (iii) la capacitat limitada que els agents tenen per tractar paquets d'informació. Tot i la seva simplicitat, els models reprodueixen les característiques principals del flux d'informació en un entorn real. De tota la col.lecció de models ens concentrem, en particular, en un escenari en què els nodes són capaços de lliurar en mitjana un nombre fixat de paquets independentment de la seva càrrega, és a dir, de la quantitat de paquets que en aquell precís moment tinguin acumulats. En aquesta situació, observem que apareixen cues llargues de paquets que esperen per ser transmesos i també fluctuacions sense una escala característica en la quantitat total de paquets que viatgen per la xarxa, tal com s'ha observat en estudis empírics de xarxes de comunicacions reals.

El comportament del model depèn d'un paràmetre extern, ρ, que determina la probabilitat que, en un node de la xarxa de comunicacions i en un instant de temps donat, es generi un paquet per ser lliurat. Quan ρ és petit (proper

a 0) hi ha pocs paquets circulant per la xarxa, no s'interfereixen els uns amb els altres i, per tant, no hi ha congestió. Quan ρ creix, però, aquest efecte de congestió comença a jugar un paper important. De fet, a un cert punt crític, ρ_c, la xarxa de comunicacions col.lapsa i el temps mig necessari per lliurar un paquet al seu destinatari divergeix. Per xarxes de tipus jeràrquic, som capaços de caracteritzar la transició de fase entre el règim col.lapsat i el règim lliure i d'estimar la posició del punt crític, ρ_c, que en certa manera mesura la capacitat de la xarxa per tractar informació. També hem demostrat que, quan mantenir canals de comunicació no té cap cost pels agents, l'estructura òptima, entesa com la que té més capacitat, és la més plana possible, amb només un node al primer nivell i tota la resta d'agents en el segon i connectats a ell.

A més a més, aquest model bàsic l'hem estès en diferents sentits i hem estudiat com les extensions modifiquen el comportament de la xarxa de comunicacions. Primer, hem introduït el fet que els agents són heterogenis pel que fa a les seves capacitats i les seves relacions amb altres agents. En segon lloc, s'ha tingut en compte que mantenir molts canals de comunicació oberts pot ser costós per als agents en termes de temps. En aquesta situació, s'ha demostrat que l'estructura jeràrquica òptima no serà, en general, la més plana possible. Per contra, hi ha una jerarquia òptima que és més plana a mesura que el cost dels canals de comunicació és més baix i a la inversa. Aquest és un fet contrastat empíricament en empreses reals. En tercer lloc, hem considerat xarxes que no tinguin estructura jeràrquica i hem conclòs que el fet que hi hagi diferents camins per arribar de l'origen a la destinació dels paquets té conseqüències importants pel que fa al comportament crític que es produeix en el punt de col.lapse. Finalment, s'ha estudiat, tan en xarxes jeràrquiques com en xarxes no jeràrquiques, el comportament del sistema quan la quantitat de paquets que un agent és capaç de lliurar no és independent de la seva càrrega.

3. Xarxes de comunicacions òptimes

El capítol anterior ens ha permès entendre la dinàmica dels processos de comunicació. Amb aquest coneixement, en el present capítol ens centrem en el que és l'objectiu principal de la part teòrica d'aquest treball de tesi: dissenyar estructures de comunicació òptimes. En particular, ens centrem en un escenari en què els agents no tenen un coneixement precís de l'estructura de la xarxa sinó només un coneixement local. En aquest cas, es produeix un conflicte d'interessos: per una banda, estructures centralitzades minimitzen el cost de cerca però, per l'altra, maximitzen els problemes de congestió. En general, doncs, no és trivial esbrinar quina serà la millor xarxa de comunicacions. En aquest capítol ataquem el problema amb dos enfocs diferents.

El nostre primer enfoc és intuïtiu però molt poc eficient. La idea és senzilla, generar xarxes amb regles definides a priori i estudiar-ne l'eficiència. Per una determinada probabilitat de generació de paquets per node i instant de temps, ρ,

generem una dinàmica de comunicació d'acord amb el model que s'ha presentat en el capítol anterior, mesurem quina és la càrrega total de la xarxa i definim la xarxa òptima com aquella que té una càrrega menor.[1]

Les xarxes les construïm combinant mecanismes que han estat proposats en la literatura de xarxes complexes. En primer lloc, hi ha un component bidimensional, que té un contingut informacional similar a estudis anteriors. En segon lloc, hi ha un component preferencial que fa que els nodes tinguin tendència a connectar-se amb aquells altres nodes de la xarxa que ja tenen una connectivitat més alta. En tercer lloc, hi ha un component aleatori. Resulta, a més, que cada component té avantatges i inconvenients pel què fa a la comunicació. En una xarxa bidimensional pura, el contingut informacional fa que sigui fàcil trobar camins que duen a la destinació dels paquets. Les distàncies mitges entre nodes, però, són grans. Quan substituïm part de les connexions en la xarxa bidireccional per connexions aleatòries de llarg abast, perdem part del contingut informacional de la xarxa bidimensional però, a canvi, fem decréixer dràsticament la distància mitja entre nodes. La distància mitja també decreix si les connexions de llarg abast són preferencials i, en aquest cas, la centralització de les connexions en uns quants (pocs) nodes fa que la pèrdua de contingut informacional no sigui tan greu, perquè els nodes importants tenen, de fet, tota la informació necessària. En canvi, aquesta centralització té efectes molt negatius pel que fa a la congestió.

Les simulacions confirmen aquestes idees i aboquen, a més, alguns resultats sorprenents. Quan la quantitat de paquets que es generen és petita, l'estructura òptima és totalment centralitzada. En canvi, quan hi ha molts paquets, l'estructura òptima és una combinació de xarxa bidimensional i xarxa aleatòria, és a dir, un *món petit* a la Watts-Strogatz (WS). De fet, es podria esperar que, a mesura que augmenta la quantitat de paquets, l'estructura òptima fos cada vegada més descentralitzada: aquesta intuïció resulta ser equivocada. El que s'observa, sorprenentment, és que l'estructura òptima és totalment centralitzada fins que, en un cert punt, esdevé una xarxa WS.

L'objectiu del present capítol, però, és més ambiciós. Sense cenyir-nos a una família concreta de xarxes construïdes apriorísticament, és possible trobar xarxes òptimes? Amb els elements que tenim fins aquest punt, el procediment més general seria el següent: generem una xarxa, simulem la dinàmica de comunicació i mesurem la càrrega de la xarxa, fem un petit canvi i tornem a mesurar la càrrega, acceptem o rebutgem el canvi d'acord amb alguna regla consistent i iterem el procediment. Això, però, és prohibitiu degut a que la simulació de la dinàmica pot ser molt lenta, especialment a prop del punt crític de col.lapse. Per tant, un segon enfoc consisteix, en primer lloc, en desenvolupar

[1]De fet, minimitzar la càrrega de la xarxa és equivalent a minimitzar el temps mig necessari per lliurar un paquet.

un formalisme que permeti calcular la càrrega d'una xarxa de manera més ràpida. Un cop fet això, ja podem obtenir estructures òptimes en general. En el cas que ens interessa en què els nodes tenen només informació local de l'estructura de la xarxa, els resultats mostren, novament, una transició sobtada centralització-descentralització.

Encara que el debat centralització-descentralització en organitzacions és molt antic, no hi ha cap model que permeti explicar, en termes comunicacionals, per què i en quines condicions són millors unes o altres estructures. El nostre resultat també permet explicar l'evidència empírica constatada en la literatura que problemes senzills són resolts millor per grups centralitzats i a la inversa, i la tendència de les grans multinacionals actuals a descentralitzar la presa de decisions a tots els nivells.

4. Xarxes de comunicacions complexes en organitzacions reals

Un cop assolit l'objectiu d'estudiar des d'un punt de vista teòric les estructures de comunicació òptimes, en aquest capítol adoptem una perspectiva totalment empírica dels processos de comunicació en organitzacions. Mostrem, en particular, com és possible treure informació de la xarxa de comunicacions complexa que hi ha darrera l'organigrama formal d'una organització. Per tal de reconstruir aquesta xarxa *informal*, aprofitem el registre automàtic de les comunicacions, i en particular del correu electrònic, que es fa en una organització. Un cop construïda la xarxa, n'estudiem l'estructura de comunitats, és a dir l'organització d'individus i grups i les seves interaccions, mitjançant l'algorisme de Girvan-Newman que s'ha proposat molt recentment. Més enllà del propi algorisme, proposem una tècnica de visualització que permet identificar aquesta estructura de comunitats per inspecció visual. A més a més, suggerim tota una colla de mesures que permeten caracteritzar quantitativament l'estructura de comunitats de l'organització. Com a cas d'estudi, ens concentrem en la xarxa de correu electrònic de la Universitat Rovira i Virgili, que representa les comunicacions electròniques entre les aproximadament 1700 persones que hi treballen. Amb aquesta anàlisi som capaços d'identificar, efectivament, l'organització dels individus en equips de recerca, departaments i escoles i facultats, a part de les diferents unitats administratives, i també les interaccions entre uns i altres.

Des d'un punt de vista teòric, aquesta metodologia ens permet identificar, també, certes propietats emergents d'invariància d'escala i d'autosimilitud, similars a les que s'observen en certs sistemes naturals com les xarxes de rius. Aquest resultat obre preguntes molt interessants sobre els mecanismes subjacents a les interaccions entre individus en una organització. L'autosimilitud és un senyal de replicació de l'estructura a diferents nivells organitzatius, i podria ser el resultat de la competència, a tots els nivells, de dos factors: la necessitat

de comunicació i les limitacions de capacitat. Alhora, el paral.lelisme amb les xarxes de rius suggereix que algun principi d'optimització global podria ser el responsable de l'emergència de propietats d'escala. Efectivament, en rius, aquestes propietats emergeixen quan s'optimitza la dissipació d'energia. És possible que en el cas de les xarxes de comunicacions s'optimitzi d'alguna manera el flux d'informació.

Finalment, mostrem que l'estudi de la xarxa informal de comunicacions fent servir eines desenvolupades en l'àmbit de la *física de xarxes complexes* pot ser molt útil per l'administració d'organitzacions. Per exemple, pot ser una manera objectiva de mesurar l'evolució d'una organització o el grau d'adequació a un cert objectiu.

5. Conclusions

En aquest treball de tesi hem estudiat, tant des d'una perspectiva teòrica com empírica, el paper de la comunicació i el processament d'informació en organitzacions i les seves implicacions en els problemes de disseny organitzacional. Podem treure'n les següents conclusions.

- Qualsevol procés de comunicació amb un component estocàstic en què els agents tenen una capacitat limitada per tractar paquets d'informació (els paquets són unitats discretes d'informació) dóna lloc a la formació de cues o, en altres paraules, a l'acumulació de paquets que esperen per ser lliurats. Quan el nombre mig de paquets que els nodes poden lliurar per unitat de temps és finit i afitat, la formació de cues pot resultar en un estat de col.lapse de la xarxa, en què el temps d'espera dels paquets esdevé infinitament llarg. La quantitat de paquets que una xarxa de comunicacions (per exemple, una organització) pot lliurar per unitat de temps sense col.lapsar es una mesura de la seva capacitat.

  La transició a l'estat de col.lapse és una transició de fase. Quan la quantitat de paquets lliurats per un agent és independent de la seva càrrega, la transició es contínua i podem aplicar l'instrumental desenvolupat en la teoria de fenòmens crítics. Per contra, quan el nombre de paquets lliurats disminueix amb la càrrega, la transició és discontínua i el col.lapse es produeix mitjançant la formació de nuclis congestionats.

- Fixem-nos en una situació en què els nodes de la xarxa de comunicació lliuren, en mitjana, un nombre fixat de paquets independentment de la seva càrrega i en què el paper que juguen l'emissor d'un paquet i el seu receptor són totalment simètrics. A més, considerem xarxes jeràrquiques en què els nodes tenen un coneixement total de la subxarxa que penja per sota seu. En aquesta situació som capaços d'obtenir, mitjançant una teoria de camp mig, el valor del punt de col.lapse. Les expressions obtingudes coincideixen perfectament amb els valors obtinguts en simulacions numèriques.

D'aquests resultats se'n pot concloure que, quan mantenir connexions amb altres agents no té cap mena de cost, l'estructura òptima (és a dir, aquella que pot tractar una quantitat de paquets més gran sense col.lapsar) és la jerarquia plana o estrella, en que un sol node supervisa tota la resta. Per contra, quan mantenir connexions actives té un cost en termes de temps, apareixen jerarquies òptimes que no són, en general, tan planes. A mesura que el cost de les connexions disminueix, la jerarquia òptima esdevé més i més plana, fet que s'ha observat empíricament en organitzacions reals.

- Considerem una situació en què els nodes no tenen un coneixement global exacte de la topologia de la xarxa. En aquest cas, els efectes de la *congestió*, com s'ha esmentat en els punts anteriors, i de la *cerca* de la destinació dels paquets coexisteixen. En general, des del punt de vista de la cerca la centralització de les comunicacions és positiva, però és negativa des del punt de vista de la congestió. Pel que fa a les estructures òptimes, la competició d'aquests dos efectes dóna lloc a una transició centralització-descentralització.

  Hem proposat un formalisme que permet tractar el problema de cerca amb congestió analíticament. A més a més, hem trobat que les estructures òptimes per cerca amb congestió es poden dividir en dues famílies: xarxes de tipus estrella, quan la quantitat d'informació és suficientment petita, i xarxes homogènies i isotròpiques, quan la quantitat d'informació creix. Sorprenentment, la transició entre xarxes d'una família i d'una altra és sobtada, és a dir, no hi ha estructures òptimes per cerca local amb congestió que no pertanyin a una d'aquestes dues classes.

- Hem estudiat la xarxa de comunicacions (en particular, la xarxa de correu electrònic) d'una organització amb gairebé 1.700 treballadors: la Universitat Rovira i Virgili. Hem demostrat que l'aplicació d'algorismes per identificar comunitats desenvolupats recentment en l'àrea de les xarxes complexes és molt eficient a l'hora d'identificar centres, departaments i fins i tot equips de recerca dins la Universitat. Per això, cal tractar la informació del correu electrònic convenientment, per tal d'eliminar correu massiu que no té un autèntic valor des del punt de vista informacional. La nova metodologia que hem proposat permet identificar les diferents comunitats per inspecció visual.

  A més a més, hem estat capaços de caracteritzar quantitativament l'estructura de comunitats. Els resultats revelen l'existència de propietats d'autisimilitud que suggereixen que el mecanisme subjacent a la formació de comunitats podria ser, en algun sentit, universal i no pas particular de la Universitat considerada, com passa en altres sistemes complexos. També és destacable que l'anàlisi quantitativa de l'estructura de comunitats és una eina útil des

d'un punt de vista organitzacional. En particular hem explicat que alguns conceptes que sovint es fan servir en la literatura de *management* es poden quantificar amb la metodologia que hem proposat i que és possible aplicar-la a estudis d'organitzacions a gran escala.

Publication list

Communication networks and organizations

- Arenas, A., Diaz-Guilera, A., and Guimera, R. (2001), Communication in networks with hierarchical branching, *Phys. Rev. Lett.* 86 (14), 3196–3199.

- Guimera, R., Arenas, A., Diaz-Guilera, A. (2001), Communication and optimal hierarchical networks, *Physica A* 299, 247–252.

- Guimera, R., Arenas, A., Diaz-Guilera, A., Vega-Redondo, F. (2001) Information processing and optimal organizational structures, *Proceedings of the Workshop on Economics with Heterogeneous Interacting Agents*, Maastricht, Netherlands.

- Guimera, R., Arenas, A., Diaz-Guilera, A., Giralt, F. (2002) Dynamical properties of model communication networks, *Phys. Rev. E* 66, 026704.

- Guimera, R., Diaz-Guilera, A., Vega-Redondo, F., Cabrales, A., Arenas, A. (2002) Optimal network topologies for local search with congestion, accepted for publication in *Phys. Rev. Lett*.

- Guimera, R., Danon, L., Diaz-Guilera, A., Giralt, F., Arenas, A. (2002) The real communication network behind the formal chart: community structure in organizations, submitted.

- Guardiola, X., Guimera, R., Arenas, A., Diaz-Guilera, A., Streib, D., Amaral, L.A.N. (2002) Micro- and macro-structure of trust networks, submitted.

Other publications

- Camacho, J., Guimera, R., Amaral, L.A.N., (2002) Analytical solution of a model for complex food webs, *Phys. Rev. E* 65, 030901(R).
- Camacho, J., Guimera, R., Amaral, L.A.N., (2002) Robust patterns in food web structure, *Phys. Rev. Lett.* 88, 228102.

References

[Adamic, 1999] Adamic, L. A. (1999). The small world web. In *Proceedings of the third european conference, ECDL'99*, pages 443–452, Paris. Springer-Verlag, Berlin.

[Adamic and Adar, 2002] Adamic, L. A. and Adar, E. (2002). Friends and neighbors on the web. *unpublished*. http://citeseer.nj.nec.com/380967.html.

[Adamic et al., 2002] Adamic, L. A., Lukose, R. M., and Huberman, B. A. (2002). Local search in unstructured networks. In Bornholdt, S. and Schuster, H. G., editors, *Handbook of graphs and networks: from genome to the internet*, Berlin, Germany. Wyley-VCH.

[Adamic et al., 2001] Adamic, L. A., Lukose, R. M., Puniyani, A. R., and Huberman, B. A. (2001). Search in power-law networks. *Phys. Rev. E*, 64:046135.

[Albert and Barabasi, 2002] Albert, R. and Barabasi, A.-L. (2002). Statistical mechanics of complex networks. *Reviews of Modern Physics*, 74:47–97.

[Albert et al., 2000] Albert, R., Jeong, H., and Barabasi, A.-L. (2000). Topology of evolving networks: local events and universality. *Phys. Rev. Lett.*, 85(24):5234–5237.

[Allen, 1990] Allen, O. (1990). *Probability, statistics and queueing theory with computer science application*. Academic Press, New York, 2nd edition.

[Amaral et al., 2000] Amaral, L. A. N., Scala, A., Barthelemy, M., and Stanley, H. E. (2000). Classes of small-world networks. *Proc. Nat. Acad. Sci. USA*, 97:11149–11152.

[Arenas et al., 2001] Arenas, A., Diaz-Guilera, A., and Guimera, R. (2001). Communication in networks with hierarchical branching. *Phys. Rev. Lett.*, 86(14):3196–3199.

[Austin et al., 1959] Austin, T. L., Fagen, R. E., Penney, W. F., and Riordan, J. (1959). The number of components in random linear graphs. *Annals Math. Statist.*, 30:747–754.

[Banavar et al., 1999] Banavar, J., Maritan, A., and Rinaldo, A. (1999). Size and form in efficient transportation networks. *Nature*, 399:130.

[Barabasi, 2002] Barabasi, A.-L. (2002). *Linked: The new science of networks*. Perseus Publishing, Cambridge, MA, USA.

[Barabasi and Albert, 1999] Barabasi, A.-L. and Albert, R. (1999). Emergenge of scaling in random networs. *Science*, 286:509–512.

[Batt, 1996] Batt, R. (1996). From bureaucracy to enterprise? In Osterman, P., editor, *Brokken Ladders: managerial careers in the new economy*, pages 55–80, New York. Oxford University Press.

[Binder, 1987] Binder, K., editor (1987). *Applications of the Monte Carlo method in statistical physics*. Number 36 in Topics in Current Physics. Springer-Verlag, Berlin, Germany, second edition.

[Bollobas, 2001] Bollobas, B. (2001). *Random graphs*. Cambridge University Press, 2nd edition.

[Bolton and Dewatripont, 1994] Bolton, P. and Dewatripont, M. (1994). The firm as a communication network. *Quarterly Journal of Economics*, 109:809.

[Brynjolfsson et al., 1989] Brynjolfsson, E., Malone, T. W., Gurvaxani, V., and Kambil, A. (1989). Does information thechnology lead to smaller firms? Technical report, Center for coordination science, MIT.

[Buchanan, 2002] Buchanan, M. (2002). *Nexus: Small worlds and the groudbraking science of networks*. W. W. North & Company, New York, USA.

[Chandler, 1966] Chandler, A. D. (1966). *Strategy and structure*. Doubleday, New York, USA.

[Chandler, 1990] Chandler, A. D. (1990). *Scale and scope: the dynamics of industrial capitalism*. Harvard University Press, Cambridge, MA, USA.

[Cohen et al., 2000] Cohen, R., Erez, K., ben Avraham, D., and Havlin, S. (2000). Resilience of the internet to random breakdowns. *Phys. Rev. Lett.*, 85(21):4626–4628.

[Cohen et al., 2001] Cohen, R., Erez, K., ben Avraham, D., and Havlin, S. (2001). Breakdown of the internet under intentional attack. *Phys. Rev. Lett.*, 86(16):3682–3685.

[Dorogovtsev and Mendes, 2002] Dorogovtsev, S. and Mendes, J. F. F. (2002). Evolution of networks. *Advances in Physics*, 51:1079.

[Dorogovtsev et al., 2000] Dorogovtsev, S. N., Mendes, J. F. F., and Samukhin, A. N. (2000). Structure of growing networks with preferential linking. *Phys. Rev. Lett.*, 85(21):4633–4636.

[Ebel et al., 2002] Ebel, H., Mielsch, L.-I., and Bornholdt, S. (2002). Scale-free topology of e-mail networks. *unpublished*. cond-mat/0201476.

[Eckmann and Moses, 2002] Eckmann, J.-P. and Moses, E. (2002). Curvature of co-links uncovers hidden thematic layers in the world wide web. *Proc. Nat. Acad. Sci. USA*, 99(9):5825–5829.

[Economist, 2001] Economist (2001). The big picture (network collaborations). *The Economist*, January 4th.

[Erdos and Renyi, 1959] Erdos, P. and Renyi, A. (1959). On random graphs i. *Publ. Math. Debrecen*, 6:290–297.

[European Comission, 2001] European Comission (2001). An industrial competitiveness policy for the european chemical industry: an example. Communication from the Comission to the Council, the European Parliament and the Economic and Social Commitee [ref: COM(96)187].

[Ferrer and Sole, 2001] Ferrer, R. and Sole, R. V. (2001). The small-world of human language. *Proc. Roy. Soc. London B*, 268:2261–2266.

[Flake et al., 2002] Flake, G. W., Lawrence, S., Giles, C. L., and Coetzee, F. M. (2002). Self-organization and identification of communities. *IEEE Computer*, 35:66–71.

[Ford and Uhlenbeck, 1957] Ford, G. W. and Uhlenbeck, G. E. (1957). Combinatorial problems in the theory of graphs. *Proc. Nat. Acad. Sci. USA*, 43:163–167.

[Freeman, 1977] Freeman, L. C. (1977). *Sociometry*, 40:35.

[Garfinkel, 1994] Garfinkel, S. (1994). *PGP: Pretty Good Privacy*. O'Reilly and Associates, Cambridge, MA.

[Garicano, 2000] Garicano, L. (2000). Hierarchies and the organization of knowledge in production. *Journal of Political Economy*, 108(5):874–904.

[Gilbert, 1956] Gilbert, E. N. (1956). Enumeration of labelled graphs. *Canad. J. Math.*, 8:405–411.

[Girvan and Newman, 2002] Girvan, M. and Newman, M. E. J. (2002). Community structure in social and biological networks. *Proc. Nat. Ac. Sci. USA*, 99:7821–7826.

[Guardiola et al., 2002] Guardiola, X., Guimera, R., Arenas, A., Diaz-Guilera, A., and Amaral, L. A. N. (2002). Micro- and macro-structure of trust networks. *unpublished*. cond-mat/0206240.

[Guimera et al., 2001a] Guimera, R., Arenas, A., and Diaz-Guilera, A. (2001a). Communication and optimal hierarchical networks. *Physica A*, 299:247–252.

[Guimera et al., 2002a] Guimera, R., Arenas, A., Diaz-Guilera, A., and Giralt, F. (2002a). Dynamical properties of model communication networks. *Phys. Rev. E*, 66:026704.

[Guimera et al., 2001b] Guimera, R., Arenas, A., Diaz-Guilera, A., and Vega-Redondo, F. (2001b). Information processing and optimal organizational structures. In *Proceedings of the WEHIA'01*.

[Guimera et al., 2002b] Guimera, R., Danon, L., Diaz-Guilera, A., Giralt, F., and Arenas, A. (2002b). Self-similar community structures in organizations. *Unpublished*.

[Guimera et al., 2002c] Guimera, R., Diaz-Guilera, A., Vega-Redondo, F., Cabrales, A., and Arenas, A. (2002c). Optimal networks for local search with congestion. *Unpublished*. cond-mat/206410.

[Halsey, 1997] Halsey, T. C. (1997). The branching structure of diffusion-limited aggregates. *Europhysics Letters*, 39(1):43–48.

[Halsey, 2000] Halsey, T. C. (2000). Diffusion-limited aggregation: A model for pattern formation. *Physics Today*, 53(11):36.

[Hangstrom, 1991] Hangstrom, P. (1991). *The "wired" multinational corporation: The role of information systems for structural change in complex organizations*. PhD thesis, Stockholm School of Economics.

[Horton, 1945] Horton, R. E. (1945). Erosional development of streams and their drainage basins: hydrophysical approach to quantitative morphology. *Bulletin of the Geological Society of America*, 56:275–370.

[Ijiri and Simon, 1977] Ijiri, Y. and Simon, H. A. (1977). *Skew distributions and the sizes of business firms*, volume 24 of *Studies in mathematical and managerial economics*. North-Holland, Amsterdam, Netherlands.

[Jacobson, 1988] Jacobson, V. (1988). Congestion avoidance and control. In *Proceedings of SIGCOMM '88*, Standford, CA. ACM.

[Kamada and Kawai, 1989] Kamada, T. and Kawai, S. (1989). An algorithm for drawing general undirected graphs. *Inform. Process. Lett.*, 31:7–15.

[Kirkpatrick et al., 1983] Kirkpatrick, S., Gelatt, C. D., and Vecchi, M. P. (1983). Optimization by simulated annealing. *Science*, 220:671–680.

[Kleinberg, 1999] Kleinberg, J. (1999). The small-world phenomenon: an algorithmic perspective. Technical Report 99-1776, Cornell Computer Science.

[Kleinberg, 2000] Kleinberg, J. (2000). Navigation in a small world. *Nature*, 406:845.

[Klemm and Eguiluz, 2002] Klemm, C. and Eguiluz, V. M. (2002). Growing scale-free networks with small-world behavior. *Phys. Rev. E*, 65:057102.

[Krackhardt and Hanson, 1993] Krackhardt, D. and Hanson, J. R. (1993). Informal networks: the company behind the chart. *Harvard Bussiness Review*, 71:104–113.

[Krapivsky et al., 2000] Krapivsky, P. L., Redner, S., and Leyvraz, F. (2000). Connectivity of growing random networks. *Phys. Rev. Lett.*, 85(21):4629–4632.

[Maritan et al., 1996] Maritan, A., Rinaldo, A., Rigon, R., Giacometti, A., and Rodriguez-Iturbe, I. (1996). Scaling laws for river networks. *Phys. Rev. E*, 53(2):1510–1515.

[Mayo, 1949] Mayo, E. (1949). *The social problems of an industrial civilization*. Routhledge.

[Menger, 1927] Menger, K. (1927). Zur allgemeinen kurventheorie. *Fundamenta Mathematicae*, 10:96–115.

[Morgan, 1997] Morgan, G. (1997). *Images of organization*. SAGE Publications, London, 2nd edition.

[Newman, 2002] Newman, M. (2002). Assortative mixing in networks. *unpublished*. cond-mat/0205405.

[Newman and Watts, 1999] Newman, M. and Watts, D. J. (1999). Renormalization group analysis of the small-world network model. *Phys. Lett. A*, 263(4–6):341–346.

[Newman, 2001a] Newman, M. E. J. (2001a). Scientific collaboration networks. ii. shortest paths, weighted networks, and centrality. *Phys. Rev. E*, 64:016132.

[Newman, 2001b] Newman, M. E. J. (2001b). The structure of scientific collaboration networks. *Proc. Nat. Acad. Sci. USA*, 98(2):404–409.

[Newman et al., 2001] Newman, M. E. J., Strogatz, S. H., and Watts, D. J. (2001). Random graphs with arbitrary degree distributions and their applications. *Phys. Rev. E*, 64:026118.

[Ohira and Sawatari, 1998] Ohira, T. and Sawatari, R. (1998). Phase transition in a computer network traffic model. *Phys. Rev. E*, 58:193.

[Pastor-Satorras and Vespignani, 2001] Pastor-Satorras, R. and Vespignani, A. (2001). Epidemic spreading in scale-free networks. *Phys. Rev. Lett.*, 86:3200–3203.

[Penna, 1995] Penna, T. J. P. (1995). Traveling salesman problem and Tsallis statistics. *Phys. Rev. E*, 51(1):R1–R3.

[Prietula et al., 1998] Prietula, M., Carley, K., and Gasser, L. (1998). *Simmulating organizations (computational models of institutions and groups)*. AAAI Press & MIT Press.

[Radner, 1993] Radner, R. (1993). The organization of decentralized information processing. *Econometrica*, 61:1109–1146.

[Radner and van Zandt, 1992] Radner, R. and van Zandt, T. (1992). Information processing in firms and returns to scale. *Annales d'Economie Statistique*, XXV–XXVI:240–254.

[Rinaldo et al., 1993] Rinaldo, A., Rodriguez-Iturbe, I., Rigon, R., Ijjasz-Vasquez, E., and Bras, R. L. (1993). Self-organized fractal river networks. *Phys. Rev. Lett.*, 70(6):822–825.

[Rodriguez-Iturbe and Rinaldo, 1996] Rodriguez-Iturbe, I. and Rinaldo, A. (1996). *Fractal river basins: chance and self-organization*. Cambridge University Press, Cambridge.

[Sinclair and Ball, 1996] Sinclair, K. and Ball, R. C. (1996). Mechanism for global optimization of river networks from local erosion rules. *Phys. Rev. Lett.*, 76(18):3360–3363.

[Smith, 2002] Smith, R. (2002). Instant messaging as a scale-free network. *unpublished*. cond-mat/0206378.

[Smolik, 2001] Smolik, S. L. (2001). Environment, health, and safety performance improvement at Dow. Technical report, Dow Chemical Company. Available on-line at http://www.dow.com/environment/reports/2001/20010801.html.

[Sole and Valverde, 2001] Sole, R. and Valverde, S. (2001). Information transfer and phase transitions in a model of internet traffic. *Physica A*, 289(3–4):595–605.

[Stallings, 1995] Stallings, W. (1995). The pgp web of trust. *Byte*, 20:161–162.

[Stanley, 1987] Stanley, H. E. (1987). *Introduction to phase transitions and critical phenomena*. Oxford University Press, Oxford, UK.

[Stauffer and Aharony, 1992] Stauffer, D. and Aharony, A. (1992). *Introduction to percolation theory*. Taylor & Francis, second edition.

[Strahler, 1952] Strahler, A. N. (1952). Dynamic basis of geomorphology. *Bulletin of the Geological Society of America*, 63:923–938.

[Tadic, 2001] Tadic, B. (2001). Adaptive random walks on the class of web graph. *Eur. Phys. J. B*, 23:221–228.

[Travers and Milgram, 1969] Travers, J. and Milgram, S. (1969). *Sociometry*, 32:425.

[Tretyakov et al., 1998] Tretyakov, A., Takayasu, H., and Takayasu, M. (1998). Phase transition in a computer network model. *Physica A*, 253:315.

[Tsallis and Stariolo, 1994] Tsallis, C. and Stariolo, D. A. (1994). *Annual Review of Computational Physics II*. World Scientific, Singapore. edited by D. Stauffer.

[Van Zandt, 1998] Van Zandt, T. (1998). Decentralized information processing in the theory of organizations. In Sertel, M., editor, *Contemporary Economic Development Reviewed, Volume 4: The Enterprise and its Environment*. Macmillan, London.

[Wagner and Fell, 2000] Wagner, A. and Fell, D. (2000). The small world inside large metabolic networks. Technical Report 00-07-041, Santa Fe Institue.

[Warnecke, 1993] Warnecke, H.-J. (1993). *The fractal company*. Springer-Verlag, Berlin.

[Wasserman and Faust, 1994] Wasserman, S. and Faust, K. (1994). *Social network analysis*. Cambridge University Press, Cambridge, U.K.

[Watts and Strogatz, 1998] Watts, D. and Strogatz, S. (1998). Collective dynamics of 'small-world' networks. *Nature*, 393:440.

[Watts et al., 2002] Watts, D. J., Dodds, P. S., and Newman, M. E. J. (2002). Identity and search in social networks. *Science*, 296:1302–1305.

[Willinger et al., 2002] Willinger, W., Govindan, R., Jamin, S., Paxson, V., and Shenker, S. (2002). Scaling phenomena in the internet: critically examining criticality. *Proc. Nat. Acad. Sci. USA*, 99:2573–2580.