ELSEVIER

Physica A 299 (2001) 247–252

PHYSICA A

www.elsevier.com/locate/physsa

Communication and optimal hierarchical networks ☆

R. Guimerà[a], A. Arenas[b], A. Díaz-Guilera[c,*]

[a] *Departament d'Enginyeria Química, Universitat Rovira i Virgili, Carretera Salou s/n, E-43006 Tarragona, Spain*

[b] *Departament d'Enginyeria Informàtica i Matemàtica, Universitat Rovira i Virgili, Carretera Salou s/n, E-43006 Tarragona, Spain*

[c] *Departament de Física Fonamental, Facultat de Fisica, Universitat de Barcelona, Diagonal 647, E-08028 Barcelona, Spain*

Abstract

We study a general and simple model for communication processes. In the model, agents in a network (in particular, an organization) interchange information packets following simple rules that take into account the limited capability of the agents to deal with packets and the cost associated with the existence of open communication channels. Due to the limitation in the capability, the network collapses under certain conditions. We focus on when the collapse occurs for hierarchical networks and also on the influence of the flatness or steepness of the structure. We find that the need for hierarchy is related to the existence of costly connections.

Keywords: Complexity; Communication flow; Organizations; Hierarchy

---

Nowadays, a lot of attention is paid to the dynamics of complex social and economic systems [1–3]. In particular, the influence of the topology of the underlying interactions on the behavior of such systems [4–6] deserves special interest. We focus on the behavior of hierarchical structures formed by agents (or elements, in general) that interact with each other via communication processes. This framework is especially adequate to study for instance packet flow in computer networks like the Internet [7,8], traffic networks [9], river networks [10] and particularly information flows in

---

☆ Partially supported by DGES of the Spanish Government, Grants PB96-0168 and PB97-0131, and EU TMR grant ERBFMRXCT980183.

* Corresponding author.

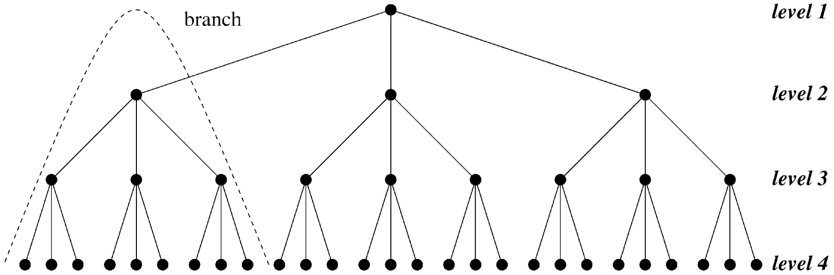

Fig. 1. Typical hierarchical tree structure used for simulations and calculations. Dashed line: definition of *branch*.

organizations [11–13]. Using Radner's words [11]:

> The typical U.S. company is so large that a substantial part of its workforce 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.

Thus, it is worth considering an organization as a system of information processors.

In this work, we extend a general and simple model for communication processes that has been recently proposed [14]. The original model considers agents that deliver information packets through well established channels. These agents have an infinite capacity to store packets and the only limitation comes from the fact that they do not have an infinite capacity to deliver. Despite its simplicity, the model reproduces the main characteristics of the flow of information packets in a network and a continuous phase transition from a free to a congested regime is observed and properly characterized by means of an order parameter [14].

In the present extension of the model, we introduce a cost associated with the establishment of links so that the agents in the communication network cannot be linked to an arbitrary number of neighbors or, at least, it has a negative influence on their performance. In both cases, with and without cost associated with the links, the optimal organizational structures are studied.

The organization is mapped into a hierarchical (Bethe) lattice (see Fig. 1), where nodes represent the communicating agents (employees) and the links between them represent communication lines. These tree like structures are characterized by two quantities: the branching factor, z, and the number of levels, m.

The dynamics of the model is the following. At each time step t, an information packet is created by every agent with probability p. When a new packet is created, a destination agent, different from the origin, is chosen at random in the network. Thus, during the following time steps $t, t+1, \ldots, t+T$, the packet is traveling toward its destination: once the packet reaches this destination agent, it is delivered and disappears from the network. One can think in a problem solving scenario [13] and say that, from time to time, problems arise in the organization; these problems need to be solved somewhere in the network. When an agent receives a packet (problem), she knows whether the destination (solution) is to be found somewhere below her. If so, she directs the packet downwards in the right direction. Otherwise, she transmits it upwards to the agent overseeing her. Thus, the information packets move toward their destination following the shortest path. The time a packet remains in the network is related not only to the distance between the origin and the destination agents, but also to the amount of packets in the network. In particular, at each time step, all the agents try to send each one of the packets they are handling. For each packet, there is a probability q_{ij} to go from the present agent i to the next one j. We call q_{ij} the *quality of communication* between agents i and j and it is defined as

$$q_{ij} = \sqrt{k_i k_j}\,, \tag{1}$$

where k_α represents the capability of agent α to communicate at each time step. For k_α we propose

$$k_\alpha = Q_L(c_\alpha) f(n_\alpha)\,, \tag{2}$$

where c_α is the number of links of agent α, $0 < Q_L(c) \leqslant 1$ 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), L is the *linking capability* that tunes the magnitude of this cost (higher values of L correspond to low linking cost and viceversa), n_α is the total number of packets currently at agent α, and $0 < f(n) \leqslant 1$ is the function that determines how the capability of a particular agent decreases when the number of information packets to handle grows (again, $f(n)$ is a decreasing function of the argument).

A suitable election for $f(n)$ seems to be

$$f(n) = \begin{cases} 1 & \text{for } n = 0\,, \\ 1/n & \text{for } n = 1, 2, 3, \ldots\,, \end{cases} \tag{3}$$

although other functional forms can be considered and one observes different interesting behaviors [14,15].

As the first step, let us focus in the simpler case $L \to \infty$, i.e., cost-less connections. The probability of generating a packet per agent and time unit, p, is an exogenous parameter that controls the behavior of the system. For small values of p, all the packets are delivered and so, after a transient, the system reaches a steady state in which the total number of packets, N, fluctuates around a constant value, i.e., the number of delivered packets is equal, on average, to the number of generated packets. However, for large values of p, not all the packets can be delivered, and N grows in time without limit. The transition between one regime and the other is a continuous phase transition and occurs for a well defined critical value of p, p_c [14].

It is possible to give an analytical estimation of p_c. Within a mean field approach, it is if we do not consider fluctuations and we assume that the behavior of all the agents in the same level is statistically identical, one gets the following expression for p_c [14]

$$p_c = \frac{\sqrt{z}}{(z(z^{m-1}-1)^2)/(z^m-1)+1} . \tag{4}$$

For values of z and m such that $z^{m-1} \gg 1$ this expression can be approximated by

$$p_c \approx z^{3/2-m} . \tag{5}$$

Although strictly speaking (4) (and its approximation (5)) provides an upper bound to p_c, it is an excellent estimation for $z \geqslant 4$, as can be seen in Fig. 2 of Ref. [14].

More interesting to us is the maximum number of information packets that can be generated in a time step without collapsing the organization, $N_c = p_c S$, with S standing for the size of the organization. It is given by

$$N_c = \frac{\sqrt{z}}{(z(z^{m-1}-1)^2)/(z^m-1)+1} \, \frac{z^m-1}{z-1} \approx \frac{z^{3/2}}{z-1} \tag{6}$$

again with the same approximation as in (5). Thus the total number of packets a network can deal with does not depend on the number of hierarchical levels. Furthermore 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 $z = S - 1$.

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 to maintain such a enormous number of communication lines. Thus, it is necessary to introduce the cost for establishing links in order to get a more realistic picture of the problem. In this case, following arguments analogous to that used in the case of cost-less connections, we can arrive at the following expression for p_c:

$$p_c = \frac{\sqrt{z Q_L(z) Q_L(z+1)}}{(z(z^{m-1}-1)^2)/(z^m-1)+1} . \tag{7}$$

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

$$N_c \approx \frac{z^{3/2}(Q_L(z) Q_L(z-1))^{1/2}}{z-1} . \tag{8}$$

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

$$Q_L(c) = 1 - \tanh\frac{c}{L} . \tag{9}$$

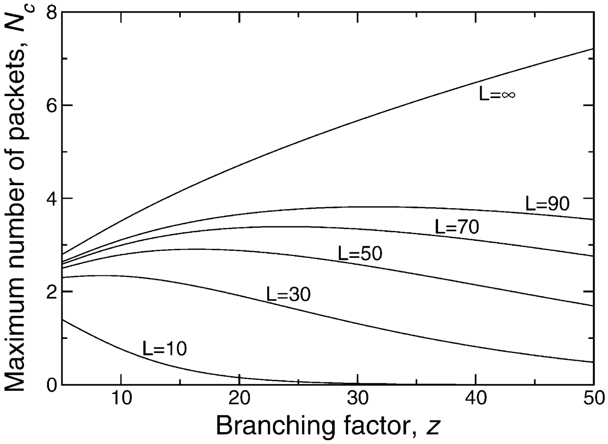

Fig. 2. Maximum number of packets that can be generated in an organization per time unit without collapsing it, plotted as a function of z. Different curves correspond to different values of the linking capability, L.

Although the election of Q_L is completely arbitrary, (9) 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 viceversa.

As can be seen from Fig. 2, 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 costless 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 $z=S-1$.

Summarizing, we have studied a model of communication that includes cost for establishing communication channels. While in the absence of such cost the flat structure is the most efficient, the need for hierarchy arises because of the existence of even small costs. In both cases with and without costs, the capacity of the network to handle information packets does not depend on the number of levels of the hierarchy but in the branching factor of the structure.

References

[1] W. Arthur, S. Durlauf, D. Lane, The Economy as an Evolving Complex System, Addison-Wesley, Reading, MA, 1997.
[2] R. Axelrod, The complexity of cooperation, Princeton University Press, Princeton, NJ, 1997.
[3] R. Axelrod, M.D. Cohen, Harnessing Complexity, Free Press, New York, 1999.
[4] D. Watts, S. Strogatz, Collective dynamics of 'small-world' networks, Nature 393 (1998) 440.
[5] A.-L. Barabási, R. Albert, Emergenge of scaling in random networks, Science 286 (1999) 509–512.
[6] L.A.N. Amaral, A. Scala, M. Barthelemy, H.E. Stanley, Classes of small-world networks, Proc. Nat. Acad. Sci. 97 (2000) 11149–11152.

[7] M. Takayasu, H. Takayasu, T. Sato, Critical behaviors and $1/f$ noise in information traffic, Physica A 233 (1996) 824.
[8] T. Ohira, R. Sawatari, Phase transition in a computer network traffic model, Physical Review E 58 (1998) 193.
[9] D. Chowdhury, L. Sanken, A. Schadschneider, Statistical physics of vehicular traffic and some related systems, Physics Reports 329 (2000) 199.
[10] J. Banavar, A. Maritan, A. Rinaldo, Size and form in efficient transportation networks, Nature 399 (1999) 130.
[11] R. Radner, The organization of decentralized information processing, Econometrica 61 (1993) 1109–1146.
[12] T. van Zandt, Contemporary Economic Issues, Vol. 4: Economic design and behavior, MacMillan, New York, 1999.
[13] L. Garicano, Hierarchies and the organization of knowledge in production, Journal of Political Economy 108 (5) (2000) 874–904.
[14] A. Arenas, A. Diaz-Guilera, R. Guimera, Communication in networks with hierarchical branching, Physical Review Letters 86 (2001) 3196.
[15] R. Guimera, A. Arenas, A. Diaz-Guilera, unpublished.