Information flow in generalized hierarchical networks

Juan A. Almendral *, Luis López and Miguel A. F. Sanjuán

Grupo de Dinámica no Lineal y Teoría del Caos
E.S.C.E.T., Universidad Rey Juan Carlos
Tulipán s/n, 28933 Móstoles, Madrid, Spain.

Abstract

The theory of complex networks is used to study different aspects of a topology that we propose to describe the relationships between members of a social group. This model is a generalized hierarchical model since relations between members of the same group are also considered. We derive the existence of a natural limit in the size of a group, and besides, an insight into hierarchical networks is given, which explains why they are so spread despite its global inefficiency.

Key words: Hierarchical social networks, information flow, diminishing returns.
PACS: 89.65.-s, 89.75.-k

1 Introduction

Network analysis is a tool which has been successfully used in different scientific fields, such as neurobiology [1], Internet [2], the financial markets [3], the social interactions, etc. A network is just a set of entities, which interact among each other following certain topology. These elements and the topology can be represented by a graph. The elements are represented by a set of points, called nodes or vertices, and the interactions are represented by a set of lines between them, called edges or links. Then, in the definition of a network, the first step is to determine the vertices and the property determining if there exists a connection between any couple of them (the edges).

* Corresponding author.
Email address: j.a.almendral@escet.urjc.es (Juan A. Almendral).

It is indeed possible to use a network model to describe a company. Thus, the nodes might be the employees and the links could be the social relations or the information flow. Actually, in this paper we are particularly interested in the study and characterization of the information flow between the members of a company. The nodes are the employees and the links represent the social or professional interactions among them.

2 Information flow in social networks

Traditionally, the research in graph theory has assumed that information in a graph travels through edges without degradation. This approach is useful to model some particular types of phenomena, like disease spread [4] or virus infection in a group and error propagation in computer networks [5].

Nevertheless, this is not appropriate when trying to model processes that take place in collaborative social networks. In order to create a model for this particular situation, we define a quantity that we call the *coordination degree*, which measures the ability of the vertices in a graph to interchange information. There are several manners to model this magnitude, but one of the easiest ways is to consider the coordination degree to be exponentially related to the distance between the vertices [6]. In this way, we define the coordination degree γ_{ij} between two vertices i and j as $\gamma_{ij} = e^{-\xi d_{ij}}$, where d_{ij} is the distance between the two vertices and ξ is a real positive constant, measuring the strength of the relationship which we call the *coordination strength*.

Quantities similar to the coordination degree have been already discussed in the literature. The most remarkable work in this field is the one by Katz [7], where the author considers the sum of $e^{-\xi d_{ij}}$ over all paths to a particular vertex. However, our model postulates that only the shortest paths are appropriate for this purpose. We think that our model is more appropriate than the one proposed by Katz for several reasons. First, the Katz measure can only be expressed in terms of the adjacency matrix of the graph, making the analysis and computations much more complex. Second, the fact that all the paths have the same priority for the spread of information produces some inconsistencies in the interpretation of the results, mainly when considering closed loops, where the information can be somehow amplified using this approach. Opposite to this, the coordination degree may be easily evaluated and can be considered as a very good approximation in sparse graphs, just by considering that the information travelling through secondary routes is negligible.

Accepting these assumptions, we can define the *total coordination degree of a vertex i* in a graph as the sum of all the coordination degrees between that particular vertex and the rest. Namely, $\Gamma_i = \sum_{j=1}^{N} \gamma_{ij}$, where N is the order of the graph (the total number of vertices in that particular graph). The total coordination degree of a vertex is a measure of the amount of information that the vertex is able to receive belonging to that particular network.

In the same way, we define the *average coordination degree of the graph* as $\overline{\Gamma} = \frac{1}{N}\sum_{i=1}^{N}\Gamma_i$. This average can be interpreted as a measure of the efficiency of a particular community or organization, since it suggests how much an individual contributes to the community.

3 The law of diminishing marginal returns

When analyzing the efficiency of social networks in terms of the average coordination degree, an interesting phenomenon appears (see Fig. 1). The efficiency of networks does not vary linearly with the order, but it tends to saturate to a value which depends on the topology of the network.

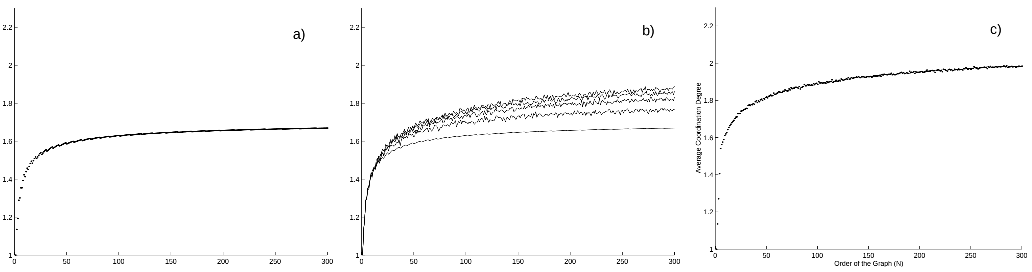

Fig. 1. Average coordination degree for 3 different graphs (all of them with $k = 4$ and $\xi = 2$). Case (a) is a regular 2D lattice. Case (b) is made of different small world networks. Case (c) is a random graph.

This result can be seen from the point of view of the well known *law of diminishing marginal returns*. This law states that when the amount of a variable resource is increased, while other resources are kept fixed, the resulting change in the output will eventually diminish. This is precisely what occurs in the models, more members in the organization does not produce an increase in the average coordination degree. This means that the increase in information of each individual diminishes as the number of members grows. As a consequence, it is reasonable to think that there exist a maximum group size, since values greater than a certain N imply marginal returns close to zero.

Actually, some scientists propose the existence of this limit in the maximum number of members of a social group by other means. Probably the most important work in this direction is the one carried out by the British anthropologist R. Dunbar [8], who related the size of the neocortex (a part of the brain related to social and language capabilities) and the maximum group size for primates. When applying this relation for the Homo sapiens, the group estimate maximum size is 147.8, or roughly 150.

Nevertheless, the analysis performed in this paper shows that the size of an organization cannot be only understood in terms of the intrinsic psychological properties of its members. The relational structure and the properties of the information transfer on the network may also play a definitive role.

4 Information in hierarchical networks

In this subsection, we focus on the analysis of social networks having hierarchical topologies [9]. Examples of graphs having this structure are regular trees. A regular tree is a regular graph (all vertices have the same degree c) that is connected (there is a path joining any two of its vertices) and that contain no circuits (there is no path going from one actor to itself that does not visit the same vertex twice). Every regular tree has a particular vertex, called root node or top of the tree, that is the most central vertex in the graph.

In order to generalize hierarchical topologies based on regular trees, we work with a regular tree that each vertex has $c-1$ order 1 lower neighbors and $c-2$ order 1 neighbors in the same level (see Fig. 2). The edges that link vertices in different levels and the edges that link vertices in the same level, have different coordination strength, and hence, there are two different coordination degrees.

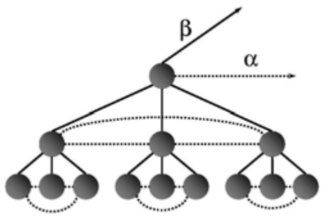

Fig. 2. Representation of a hierarchical topology with links between member of the same group. Notice that there are two different coordination degrees α and β.

Let ξ and ζ be the coordination strength which measures the strength of the relationship between vertices in different levels and vertices in the same level, respectively. Then, the coordination degree between two vertices order 1 neighbors in the same level is $\alpha = e^{-\zeta}$, and the coordination degree between two vertices order 1 neighbors in different levels is $\beta = e^{-\xi}$. Our objective is to obtain a formula giving the information flow, for the former topology representing social networks, in terms of the coordination degrees α and β.

As it is mentioned in the introduction of this section, we assume that the information travels through the shortest path. This implies that α has really an effect on the model only when $\alpha > \beta^2$. In that case, the following formula for the coordination degree is obtained

$$\widetilde{\Gamma}_i(\alpha,\beta) = \begin{cases} \left.\begin{array}{l} \frac{(c-2)\alpha}{1-(c-1)\beta}\left[\frac{\beta-\beta^{N-i}}{1-\beta} - \frac{1-[(c-1)\beta^2]^{N-i-1}}{1-(c-1)\beta^2}(c-1)^{i+1}\beta^{i+2}\right] \\ +\frac{\beta-\beta^{N-i+1}}{1-\beta} + [1+(c-2)\alpha]\frac{1-[(c-1)\beta]^i}{1-(c-1)\beta} \end{array}\right\} & i \leq N-1 \\ \frac{1-[(c-1)\beta]^i}{1-(c-1)\beta} & i = N \end{cases} \quad (1)$$

When $\alpha < \beta^2$, the shortest path is through the 1 order upper neighbor, as in a traditional hierarchical tree. Consequently, the former equation cannot be used to compute the coordination degree. However, from Eq. (1) it is possible to derive the coordination degree in a traditional hierarchical tree, by introducing the following change $\alpha \to \beta^2$. Hence, the coordination degree in our model can be written in the following terms

$$\Gamma_i(\alpha,\beta) = \begin{cases} \widetilde{\Gamma}_i(\alpha,\beta) & \alpha > \beta^2 \\ \widetilde{\Gamma}_i(\beta^2,\beta) & \alpha \leq \beta^2 \end{cases} \quad (2)$$

As a basic ingredient of our model, it is important to remark the common perception that the number of close relationships a person may have within a community is necessary limited to a quite small number, independently on the type of organization. This may be the consequence of the fact that establishing close relationships with people is normally very time consuming, and time is a limited resource for every individual.

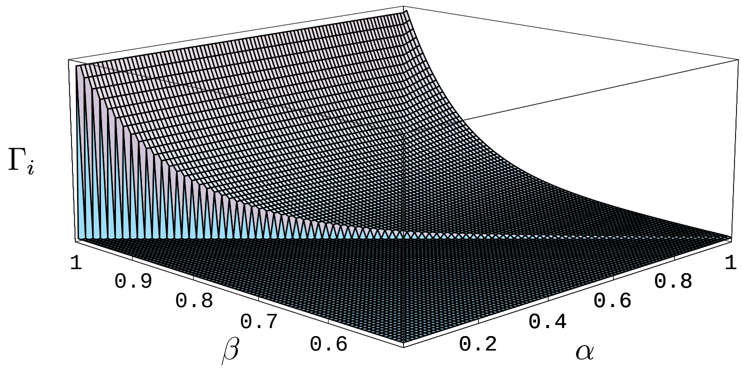

Fig. 3. When the constraint $(c-2)\alpha + c\beta = cons.$ is included in the model, the coordination degree is a curve depending on α and β whose maximum is at $\alpha = 0$.

Therefore, it is reasonable to think that each member devotes time to his neighbors proportionally to the information obtained. That is, each actor shares his time between neighbors in the same level and neighbors in different levels, proportional to α and β respectively. Thus, there is a constraint on α and β given by $(c-2)\alpha + c\beta = const$, which is a plane in the space $\{\alpha, \beta, \Gamma_i\}$. Hence, the coordination degree is a curve, the intersection of that plane and the surface defined by Γ_i (see fig. 3). And the result is that the maximum information is received when each actor devotes all his time to neighbors in upper levels.

5 Conclusions

The traditional hierarchical tree represents a topology globally inefficient as compared with others. However, it is rather spread because this structure arises when each actor only looks for maximizing his information. The result is a structure which mainly benefits the higher levels, by providing them a higher information centrality and improving their dominance of information.

When edges, between vertices with the same upper neighbor, are added to a hierarchical tree, we show that the information each actor manages decreases. This means that a hierarchical tree is a stable network against relationships between members of the same group. This stability can be seen as another reason which explains why hierarchical trees are so spread in companies all over the world. A hierarchical tree backs the leader's superiority of information despite the strength of the relationship which links the members of a group.

Nevertheless, it should be noticed that in our model edges between vertices in the same level with different upper neighbor are not included, or between vertices in different levels. This study may yield a different result.

References

[1] D. Golomb and D. Hansel, Neural Comput. **12**, 1095 (2000)

[2] M. Faloutsos, P. Faloutsos and C. Faloutsos, Proc. ACM SIGCOMM, Comput. Comm. Rev. **29**, 251 (1999)

[3] A. Kirman, C. Oddou and S. Weber, Econometrica **54**, 129 (1986)

[4] F. Ball, J. Mollison and G. Scalia-Tomba, Ann, Appl. Prob. **7**, 46 (1997)

[5] R. Albert, H. Jeong and A. L. Barabasi, Nature **406**, 378 (2000)

[6] L. López and M. A. F. Sanjuán, Phy. Rev. E **65**, 036107 (2002)

[7] L. Katz, Psychometrika **18**, 39 (1953)

[8] R. I. M. Dunbar, Behav. Brain Sci. **16**, 681 (1993)

[9] L. López, J. F. F. Mendes and M. A. F. Sanjuán, Physica A **316**, (2002) 591