Agent preferences and the topology of networks

Daniel O. Cajueiro
Department of Economics, Catholic University of Brasilia, 70790-160, Brasilia, DF, Brazil
(Received 28 April 2005; published 26 October 2005)

In this paper, a different framework to study weighed networks is introduced. The idea behind this methodology is to consider that each node of the network is an agent that desires to satisfy his/her preferences in an economic sense. Moreover, the formation of a link between two agents depends on the benefits and costs associated with this link. Therefore, an edge between two given nodes will arise only if the tradeoff between satisfaction and cost for building it is jointly positive. Using a computational framework, I intend to show that depending on the agents' combination of benefits and costs, some very well known networks can naturally arise.

DOI: 10.1103/PhysRevE.72.047104 PACS number(s): 89.75.Hc, 02.50.Le, 87.23.Ge, 89.65.−s

INTRODUCTION

During recent years, one of the main issues of the statistical physics literature has been the study of dynamic systems such as airports, wireless links, financial institutions, web pages, and other communication networks and social networks that may be described by complex weblike structures.[1]

On one hand, several models such as small world networks [2,3] and free scale networks [4] have been introduced to specially accommodate the particularities of these structures that could not be modeled by the seminal well known random graphs [5]. One should notice that although most attempts have been devoted to the study of unweighted undirected networks like the ones presented in [2,4], recently some researchers have also introduced models to deal with undirected weighted networks [6] and also directed digraphs [7].

On the other hand, several measures have been presented aiming at characterizing the properties of these networked systems, for instance, characteristic path length [8], clustering coefficient [2], efficiency [9,10], cost [10], node degree [4], degree correlation [11], weighted connectivity strength [6], and disparity [12]. The main advantage of using these measures to analyze these complex structures is the ability to compare different systems with each other and also to develop a unified theory to approach these systems.

This paper focuses particularly on undirected weighted graphs. It proposes another way based on economic and decision theory to cope with these systems. I suppose that each node of the network is an agent[2] that has his/her own preferences and is striving to maximize them. Since all agents in the network will interact in order to maximize their preferences, an edge between two given nodes will arise only if the tradeoff between satisfaction and cost for building it is jointly positive. It is assumed that this happens when the benefit brought to an agent is greater than his own cost and the cost left by the other agent (which sometimes is zero). Therefore, if the benefits brought to the agents by the edge are positive enough to compensate the cost of construction, then the edge will exist. This makes sense if one considers that a connection between agents always brings some kind of benefits, but the connection sometimes does not exist in a given network because of the high costs involved.

This tradeoff just presented above is very related to the formalism developed by [9,10] since the authors also seek a tradeoff between satisfaction (measured in a very specific way as efficiency of communication between the nodes) and cost (also measured in a very specific way).[3]

Preferences here are modeled as in economic or decision theory as utility functions. Specifically, I consider that each agent has utility function given by

$$u_i(G) = \sum_{\forall\ j \in \mathcal{N}(G)/i} a_{ij}(w_{ij} - c_{ij}) \quad \forall\ i \in G \tag{1}$$

where $\mathcal{N}(G)$ is the set of nodes in a graph (network) G, $A=[a_{ij}]$ is the adjacency matrix, $W=[w_{ij}]$ is the matrix of weights, and $C=[c_{ij}]$ is the matrix of costs.

In this context, I am particularly interested in the networks that are the solution to the problem

$$\max_{A} \sum_{i \in \mathcal{N}(G)} u_i(G). \tag{2}$$

[1]A comprehensive review of this literature may be found in [1].

[2]Throughout this paper nodes and agents are synonymous.

[3]Actually, these ideas were borrowed from engineering and operations research where researchers have been studying optimal paths in networks for a long time in order to maximize some measure of efficiency and/or minimize some measure of cost. These attempts were responsible for the arising of the seminal problems such as the minimum spanning tree problem, shortest path problem, maximum flow problem, etc. A review of these seminal problems may be found in [13]. However, although in [9,10] there is a similar flavor, the motivation here is totally different. I am not directly interested in characterizing the network topology by measuring its properties and the center of attention here is not necessarily small world networks. Moreover, the reference of the "best" network here is not necessarily the complete network, because it simply may not be the network that maximizes agent preferences.

Therefore this paper does not approach the mechanisms of network formation but it seeks the best topology for a given set of parameters.

The concept of "efficiency" provided by Eq. (2), which focuses on the total "productivity" of the network and how this allocation is made among individual agents,[4] is the same one used in [13–16] to approach—in a game theoretical framework—the dynamics of network formation and the relation between the concepts of efficiency (introduced above) and stability.[5]

The focus of this paper, differently from [13–16], is to provide a computational framework to relate agent preferences to network topologies. Thus, one has to maximize Eq. (2) to reach the desired solution.[6] One should notice that since Eq. (2) has been specified as a linear function, this can be solved as a linear binary programming problem.

BINARY LINEAR PROGRAMMING

Binary linear programming is a problem very well studied in the field of operations research and there are several methods to solve it. Unfortunately, however, due to its combinatorial nature, this problem is not trivially solved. Sometimes due to its computational cost, the size of the problem is constrained or a heuristic method that can provide only a suboptimal solution instead of an optimal one is used.

In this paper, since there are no constraints and, in Eq. (1), the choices of edges are independent of each other, the solution of (2) is trivial.[7]

LATTICES WITH *K* NEIGHBORS

The arising of a regular network where each node has K neighbors as a solution of problem (2) is in general possible only if all the agents have homogeneous preferences with constant benefits over all agents and a cost that depends only on some measure of the distance between them (not necessarily physical distance). In spite of the latter hypothesis being reasonable in the real world, the former is very hard, since agents in general have different interests. If the agents are labeled with ordinal indices from 1 to n, where n is the number of nodes, without loss of generality, one may suppose in this case that

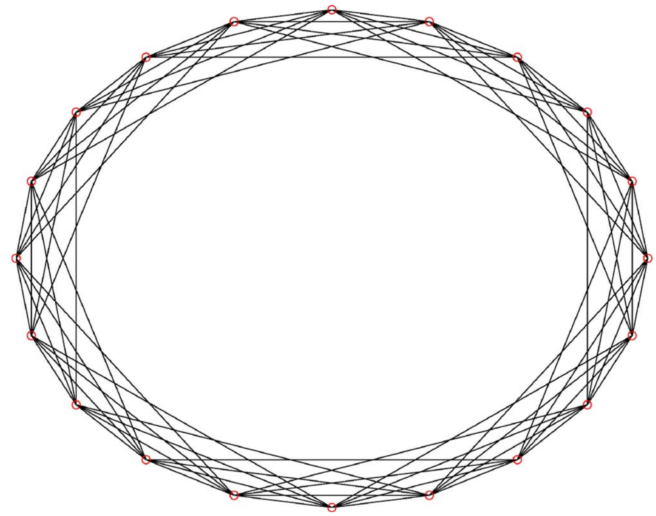

FIG. 1. (Color online) A typical regular lattice that arises with $n=20$ and $K=8$.

$$w_{ij} = \frac{K}{2\ \mathrm{floor}(n/2)} \tag{3}$$

and

$$c_{ij} = \frac{\min(|i-j|, n-|i-j|)}{\mathrm{floor}(n/2)} \tag{4}$$

where floor(x) is a function that evaluates the biggest integer less than x and $|x|$ is the absolute value of x. A typical lattice that arises in this case when $n=20$ and $K=8$ is shown in Fig. 1.

RANDOM GRAPHS

Random graphs are the opposite of regular lattices with k neighbors. The agents take random preferences into account. This specially works if the benefits brought by the connections between two nodes are random with magnitude given by a variable p and the cost of building this connection is constant as, for instance,

$$w_{ij} = p + \epsilon_{ij} \tag{5}$$

and

$$c_{ij} = 1 \tag{6}$$

where p is the probability of an edge connecting nodes $i,j \in \mathcal{N}(g)$ and ϵ_{ij} is a random variable with uniform distribution in the set [0, 1]. A typical network that arises in this case when one solves (2) with $n=20$ and $p=0.2$ is shown in Fig. 2.

Again, as in the case of the regular lattices, this kind of network is not likely to arise in real life due to the constant cost.

[4]Considering the simple formulation of Eq. (1), this notion is also a Paretian one.

[5]The definition of a stable network comes from the thought that agents have the discretion to form or reject links. The formation of a link requires the consent of both parties involved, but severance can be done unilaterally. This concept is not considered here.

[6]This is not the first time that a kind of maximization principle is used to understand the topology of complex networks. In [17], coping with natural drainage networks, it is showed that fractal and multifractal properties evolve from arbitrary initial conditions by minimizing the local and global rates of energy expenditure in the system.

[7]However, in the general case, the branch and bound technique [18–21] is usually considered. The basic concept underlying this technique is to divide and conquer. Since the original "large" problem is so difficult to solve directly, it is divided into smaller subproblems until these problems can be conquered—this is the branch step. The conquering step is done partially by bounding how good the best solution in the subset can be and then discarding the subset if its bound indicates that the optimal solution is not in it. A detailed review of the methods may be found in [22].

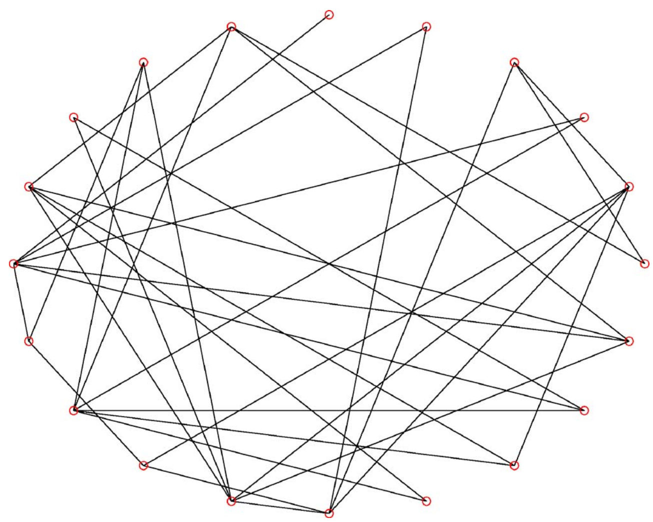

FIG. 2. (Color online) A typical random graph that arises with $n=20$ and $p=0.2$.

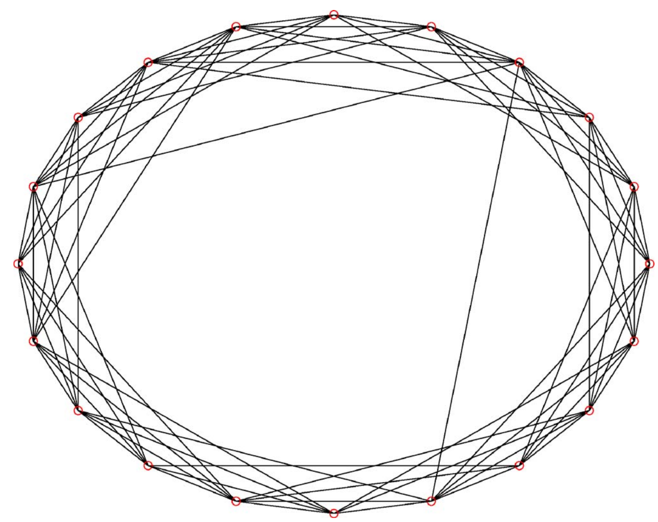

FIG. 3. (Color online) A typical small world that arises with $n=20$, $K=8$, and $p=0.2$.

SMALL WORLDS

If one leaves the two extremes presented above, as in [2,3], one may arrive at small world networks. Therefore, one should now consider a set of agents where with probability p the connection with another agent in the network brings a benefit modeled by a random variable ϵ_{ij}, for $i,j \in \mathcal{N}(g)$, with uniform distribution in the set [0, 1], and where with probability $(1-p)$ the benefit is given by a constant. The first mechanism described above models the unusual phenomenon of receiving a large benefit from a distant agent or not receiving a good benefit from a close agent. The latter mechanism models the usual phenomenon of receiving a good benefit from a close agent. Additionally, as in real life the cost of establishing a connection depends on some measure of distance.

Mathematically, with probability p

$$w_{ij} = \epsilon_{ij} \tag{7}$$

where ϵ_{ij}, for $i,j \in \mathcal{N}(g)$, is a random variable with uniform distribution in the set [0, 1] and with probability $(1-p)$

$$w_{ij} = \frac{K}{2\ \mathrm{floor}(n/2)}. \tag{8}$$

On the other hand,

$$c_{ij} = \frac{\min(|i-j|, n-|i-j|)}{\mathrm{floor}(n/2)}. \tag{9}$$

Therefore, the solution of Eq. (2) provides a network with small world behavior.

As we know, several examples of real networks follow this kind of behavior. If one analyzes the preferences of the agents, it makes sense. An agent, for example, receives constant benefits (in average) from being connected to other agents, but there are some agents who receive lower or bigger benefits than the average. In Fig. 3, a typicall small world that arises in this case when one solves (2) with $n=20$, $K=8$, and $p=0.2$ is shown.

FREE SCALE NETWORKS

Differently from the other situations considered in this paper, the phenomenon behind the generation of free scale networks seems to be a kind of cost hierarchy between the nodes, i.e., there are some nodes that are less costly than others. Therefore, some agents will preferentially attach to these nodes. More specifically, without loss of generality, let w_{ij} and c_{ij} be defined as

$$w_{ij} = \epsilon_{ij}. \tag{10}$$

ϵ_{ij}, for $i,j \in \mathcal{N}(g)$, is a random variable with uniform distribution in the set [0, 1] and

$$c_{ij} = \frac{i}{n}. \tag{11}$$

In Eq. (11) it was supposed that the nodes with minor indices are less costly than the others. Hence, these nodes will likely present the highest degrees in this case. These networks, like the small world networks, are very likely to be found in real life. One could think, for instance, of a network of airports. There are some airports that due to their geographic locations are less costly than others. In Fig. 4, there is a typical free scale network that arises when one solves (2) with $n=20$. In fact, one may clearly notice the preferential attachment presented in the network of this figure.

Moreover, simulations with bigger sets like $n=1000$ yielded networks with $\gamma=2.4\pm0.2$ where γ is the exponent of equation $P(k)\sim k^{-\gamma}$ and k is the degree of a node in the network.

FINAL REMARKS

In this Brief Report, I have presented a computational framework to characterize complex networks, i.e., one that may characterize the networks by the preferences of their agents (nodes). Actually, although only the four most common classes of networks have been considered, this frame-

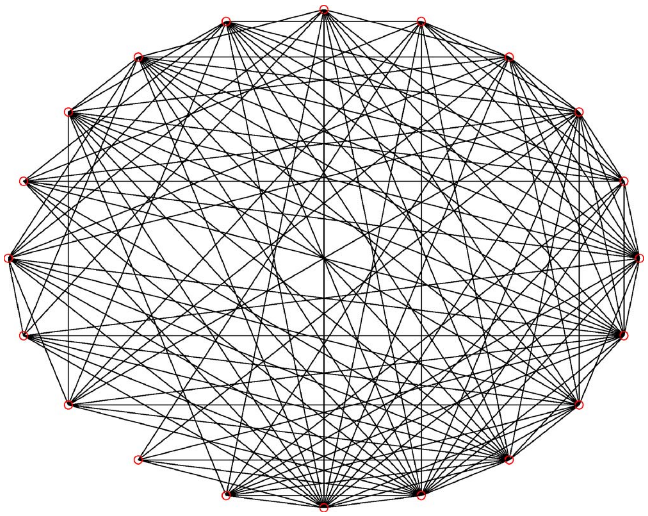

FIG. 4. (Color online) A typical free scale network that arises when $n=20$.

work can be used for many classes. In particular, by mixing the preferences of the agents presented in Eqs. (8)–(11), one may find networks with small world behavior and also attach preferences. Moreover, this methodology also works for weighted digraphs.

On one hand, linear utility functions, which means that the agents are indifferent to the risk, were the only class of utility functions considered here. A question that arises is the following: What effect is expected in the topology of the networks if the agents are, for instance, averse to the risk with concave utility functions.[8] Furthermore, no constraint has been considered in the optimization problem provided by (2). What kind of constraints are the agents in the real world subjected to and what kind of effect will these constraints cause in the topology of networks?

On the other hand, the matrices W and C here were considered exogenous, i.e., they were formed prior to the solution of the problem. It is also possible to suppose that these matrices have elements that depend on the parameters of a given iteration of the problem. For instance, the benefit brought by node i to node j could depend on the number of nodes that i actually possesses.[9] This could be the root for the study of network formation using this kind of framework.

In summary, this proposed framework may be used to improve the understanding of these complex networks that are present everywhere.

[8]Clearly, if the utility functions of the agents are not linear, linear binary programming cannot be used to find the optimal solution of this problem, but another method may be applied. One of the most common choices in the general situation is the genetic algorithm [23].

[9]Again, this cannot be solved by binary linear programming, but another method could be applied using the framework of this paper.

[1] R. Albert and A. L. Barabasi, Rev. Mod. Phys. **74**, 47 (2002).
[2] D. J. Watts and S. H. Strogatz, Nature (London) **393**, 440 (1998).
[3] D. J. Watts, *Small Worlds: The Dynamics of Networks Between Order and Randomness* (Princeton University Press, Princeton, NJ, 1999).
[4] A. L. Barabasi and R. Albert, Science **286**, 509 (1999).
[5] P. Erdós and A. Rényi, Bull. Int. Stat. Inst. 38, 343 (1960).
[6] S. H. Yook, H. Jeong, A. L. Barabasi, and Y. Tu, Phys. Rev. Lett. **86**, 5835 (2001).
[7] P. L. Krapivsky, G. J. Rodgers, and S. Redner, Phys. Rev. Lett. **86**, 5401 (2001).
[8] H. Wiener, J. Am. Chem. Soc. **69**, 17 (1947).
[9] V. Latora and M. Marchiori, Phys. Rev. Lett. **87**, 198701 (2001).
[10] V. Latora and M. Marchiori, Eur. Phys. J. B **32**, 249 (2002).
[11] R. Pastor-Satorras, A. Vazquez, and A. Vespignani, Phys. Rev. Lett. **87**, 258701 (2001).
[12] M. Barthélemy, A. Barrat, R. Pastor-Satorras, and A. Vespignani, Physica A **346**, 34 (2005).
[13] M. O. Jackson and A. Wolinsky, J. Econ. Theory **71**, 44 (1996).
[14] B. Dutta and S. Mutuswami, J. Econ. Theory **76**, 322 (1997).
[15] V. Bala and S. Goyal, Econometrica **68**, 1181 (2000).
[16] M. O. Jackson, in *Models of the Formation of Networks and Groups*, edited by B. Dutta and M. O. Jackson (Springer-Verlag, Heidelber, 2001).
[17] I. Rodriguez-Iturbe, A. Rinaldo, R. Rigon, R. L. Bras, E. Ijjas-zvasquez, and A. Marani, Geophys. Res. Lett. **19**, 889 (1992).
[18] E. Balas, Oper. Res. **13**, 517 (1965).
[19] T. J. V. Roy and L. A. Wolsey, Oper. Res. **35**, 45 (1987).
[20] H. Crowder, E. L. Johnson, and M. Padberg, Oper. Res. **31**, 803 (1983).
[21] E. L. Johnson, M. M. Kostreva, and U. H. Suhl, Oper. Res. **33**, 803 (1985).
[22] F. S. Hillier and G. J. Lieberman, *Introduction to Operations Research* (McGraw-Hill, Boston, 2001).
[23] D. E. Golberg, *Genetic Algorithms in Search, Optimization and Machine Learning* (Addison-Wesley, Reading, MA, 1989).