

Urban economy as a scale-free network

Claes Andersson,^{*} Alexander Hellervik,[†] and Kristian Lindgren[‡]
Physical Resource Theory, Chalmers University of Technology, 412 96 Göteborg, Sweden

Anders Hagson[§] and Jonas Tornberg^{||}
City and Mobility, Chalmers University of Technology, 412 96 Göteborg, Sweden
 (Received 25 March 2003; published 23 September 2003)

We present empirical evidence that land values are scale free and introduce a network model that reproduces the observations. The network approach to urban modeling is based on the assumption that the market dynamics that generates land values can be represented as a growing scale-free network. Our results suggest that the network properties of trade between specialized activities cause land values, and likely also other observables such as population, to be power-law distributed. In addition to being an attractive avenue for further analytical inquiry, the network representation is also applicable to empirical data and is thereby attractive for predictive modeling.

DOI: 10.1103/PhysRevE.68.036124

PACS number(s): 89.75.Da, 61.43.Hv, 89.75.Hc, 89.65.Gh

I. INTRODUCTION

The Zipf rank-size law for city sizes is one of the most widely known power laws in science [1]. It is also but one out of many similar power laws from systems in biology, economy, and society. We continue this research by presenting empirical evidence that land values are scale free. The data we use are based on a database delivered by Sweden Statistics that covers estimations of the market value of all land in Sweden (2.9 million data points).

Although power laws are common, they are not easily reconstructed from realistic underlying dynamics. Their ubiquity suggests that they could be caused by some general systemic property common to a range of systems. Recent research suggests that many empirically observed power laws may be due to fundamental properties of these systems viewed as networks of interacting nodes [2–4]. We investigate the mechanisms causing land values to follow these statistics and present a network model that reproduces the empirical results. The model is based on basic definitions of city formation in urban economics theory [5].

According to urban economics theory, the formation of modern cities is primarily caused by the advantages of trade between specialized producers. The exchange of goods and services between localized and largely immobile activities in trade economies makes a network representation natural: the nodes are units of land and the edges represent the exchange of goods and services between them.

It has been shown that the node degrees of a certain class of growing networks are power law distributed [2,6]. This class of networks is important because their growth mechanisms can be mapped to the microscopic dynamics of several real-world systems. We demonstrate how trade in an urban

system can be represented as a scale-free network and that, as a consequence, land values can be expected to follow the same distribution. We also verify that the model retains these properties when spatial constraints are taken into account. We do this by using a spatial network model to reproduce the empirically observed distribution of land values.

The network approach solves a fundamental issue in the problem of modeling urban systems by representing the system at the level of the underlying market structure. This allows us to produce prices in units of currency rather than undefined and subjective fitness measures. It thereby opens up doors for several extensions of the scope of the model and provides a natural interface for integrating it with other models.

When we refer to an urban system we do not necessarily refer to individual cities but rather to systems of specialized trading activities. To clarify further, our use of the term *activity* refers to trade gains in units of *currency per unit area and unit time*. Activities can be resolved to any resolution down to individual transactions.

II. THE URBAN ECONOMY AS A NETWORK

In Sec. II A we define a nonspatial model of urban economic growth and in Sec. II B we extend it to a spatial model where growth is mapped to a two-dimensional surface. In Sec. II C we motivate the model ontology and the basic assumptions on which we have based the model. In particular we discuss the connection between node degree and land value and how the urban system meets the criteria for being a scale-free network.

A. Formulation of the nonspatial model

A geographic area on which an urban economic system can grow is represented by an enumerated set of nodes, $\{1, 2, \dots, N\}$ corresponding to nonoverlapping land areas. Trade of goods and services between activities in the nodes is represented by undirected edges. Since activities within the same node can trade with each other, an edge can connect

^{*}Electronic address: claesand@fy.chalmers.se

[†]Electronic address: f98alhe@dd.chalmers.se

[‡]Electronic address: frtkl@fy.chalmers.se

[§]Electronic address: hagson@arch.chalmers.se

^{||}Electronic address: jonas@arch.chalmers.se

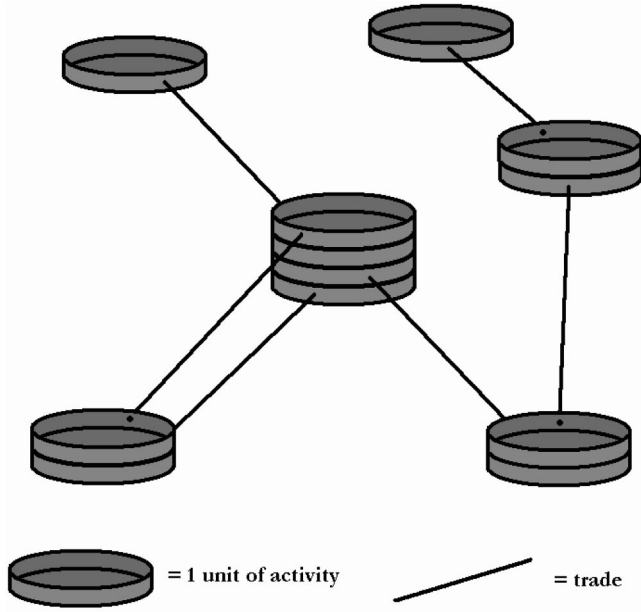


FIG. 1. The nodes are nonoverlapping areas of land. Areas with no activity are not shown in the figure.

a node with itself. The amount of activity of a site x_i is defined as the degree of the corresponding node. See Fig. 1.

The network is initialized by connecting n_0 nodes so that each has a degree of x_0 . The $N - n_0$ undeveloped nodes have no trade interactions and thus no activity.

At each time step we update the network as follows:

(1) With probability q_1 we add m edges between sites that are already developed: the first edge end point is selected uniformly among developed nodes. The probability of a node i to be selected is

$$\Pi_i^u = \delta_i^{(D)} \frac{1}{n_t^{(D)}}, \quad (1)$$

where $n_t^{(D)}$ is the number of developed nodes at time t and $\delta_i^{(D)} = 1$ if node i is developed and $\delta_i^{(D)} = 0$ otherwise.

The second end point is selected preferentially among developed nodes. Preferential selection corresponds to the uniform selection of an activity in the system and the subsequent location of its node. It was defined by Barabási *et al.* [2] as

$$\Pi_i^p = \frac{x_i}{\sum_j x_j}, \quad (2)$$

where Π_i^p is the probability of node i to be attached to a new edge and x_i is the degree of node i .

(2) With probability q_2 we add edges between m pairs of nodes that are both selected preferentially according to Eq. (2).

(3) With probability q_3 we add m units of initial activity on land that is previously undeveloped: the first end point is selected without bias similar to Eq. (1) but among undevel-

oped nodes. However, since nodes have no properties beside their degree in the nonspatial model, any undeveloped node can be added. The second end point is selected preferentially according to Eq. (2).

We will refer to the growth classes as type-1, type-2, and type-3 growth and to their relative rates as q_1 , q_2 , and q_3 with $q_1 + q_2 + q_3 = 1$ throughout the paper.

A continuum formulation

Instead of assuming any particular set of activities and interactions, we can use a continuum formulation where we consider each edge end point to be an average of a large set of urban activities. It follows that an average activity must be assumed to interact equally much with all other average activities. Rather than counting explicit interactions to determine the activity level, we study the evolution of the expected node degrees. The time evolution of activity on a developed site i follows the equation

$$x_i(t+1) = x_i(t) + q_1 \frac{m}{n_t^{(D)}} + m(q_1 + 2q_2 + q_3) \frac{x_i(t)}{\sum_j x_j(t)}, \quad (3)$$

which is solved by the continuous-time method introduced by Barabási *et al.* [7]. After sufficiently long time the degree distribution approaches the form

$$P[x_i = x] \sim (x + A)^{-\gamma}, \quad (4)$$

with

$$A = \frac{2mq_1}{q_3(1+q_2)} \quad (5)$$

and

$$\gamma = 1 + \frac{2}{1+q_2}. \quad (6)$$

According to Eq. (4) the node degree distribution will be power-law distributed for the nonspatial model.

B. Formulation of the spatial model

The nonspatial model does not include any information about inhomogeneous relations in the network. An important departure from such a simple model that can be readily observed in real systems is of course the fact that different pairs of sites have different distances between them. Because of transportation cost optimization this has the potential to affect the edge distribution of the network model.

We will now extend the basic model to incorporate spatial interactions. This allows us to verify that the power-law distribution of activities that is predicted by the nonspatial model is retained when space is introduced. Furthermore, it allows us to better map the model output to empirical measurements. Scale-free spatial network models that have been studied recently are not directly applicable to urban econom-

ics since they require an *a priori* distribution of node locations [8–10].

1. Network evolution

In the continuum nonspatial model we were able to treat network evolution as a purely local phenomenon since interactions in the network are homogenous, see Eq. (3). In the spatial model we have to again explicitly separate the dynamics of the edge endpoints. This is because the spatial context of the first end point of a new edge modifies the probabilities with which other sites will become the second end point (transportation costs).

The same types of growth, which are used in the nonspatial model, are also used in the spatial model. But, as stated, the two end points here need to be dealt with separately. The first site (primary increase) is selected either uniformly or preferentially depending on growth type. This will be discussed shortly. The second end points of the trade relation edges (secondary increase) are always preferential.

2. Secondary increase

The spatial model tells us how a given primary increase p_j in activity at site j causes secondary increases s_i in activity at all other sites i where

$$s_i = p_j \frac{D_{ij}x_i}{\sum_k D_{kj}x_k}. \quad (7)$$

This is analogous to Eq. (2) but in the spatial model, transportation costs will bias the choice of trade partners. D_{ij} is a matrix representing the interaction strength, which is assumed to decay with the transportation costs. For the results in this paper we have used $D_{ij} = [1 + cd(i,j)]^{-\alpha}$, where $d(i,j)$ is the Euclidean distance between sites i and j . The non-negative parameters c and α control the impact of spatiality. Another function that can apply is $D_{ij} = \max(0, 1 - d(i,j))$ if the transportation characteristics of the activity is known, which can be the case in a model where activity types are modeled separately rather than as an average. Exponential $D_{ij} = e^{-d(i,j)}$ can also be interesting to the extent that shielding is important, i.e., an activity tends to trade exclusively with the nearest supplier.

3. Primary increase

When trade takes place there is a mutual benefit that is often used for further increasing the activity in the city. This feedback process makes it possible for the amount of activity in the urban system to increase considerably faster than the population. It is useful to think of Eq. (7) as a black box system to which a driving force, primary increase, is applied. It should also be noted that the primary effects used for the model we present here are, by no means, neither exhaustive nor final: most earlier models of urban growth could be introduced as primary effects in our framework.

4. Primary effects in type-1, type-2, and type-3 growth

Network evolution in the spatial model is similar to that in the nonspatial model. Secondary increases are always preferential following Eq. (7) and primary increases now reflect the spatial distribution:

Type-1 growth. The primary uniform increase is identical to the nonspatial case following Eq. (1).

Type-2 growth. The primary preferential increase is identical to the nonspatial case following Eq. (2).

Type-3 growth. This is split into two related processes where one corresponds to growth in the perimeter of clusters and the other corresponds to growth in connection to infrastructure in the rural areas between clusters. Such an infrastructure is not explicitly represented in our model and instead we use a parameter ϵ to tune the amount of ambient infrastructure and thus the rate with which seemingly isolated clusters will appear.

(1) Type-3a growth: With a probability of $q_3(1 - \epsilon)$ we set the activity of a perimeter node to m . Perimeter nodes are nodes that are not developed but border to at least one developed cell. The site of the new node is selected randomly and with uniform probability among the perimeter sites on the grid

$$\pi_i^a = \delta_i^{(P)} \frac{1}{n_t^{(P)}}, \quad (8)$$

where π_i^a is the probability with which node i is selected to undergo type-3a growth, $\delta_i^{(P)} = 1$ if the node i is on the perimeter, $\delta_i^{(P)} = 0$ if the node is not on the perimeter, and $n_t^{(P)}$ is the number of perimeter nodes at time t .

(2) Type-3b growth: With a probability of $q_3\epsilon$ we set the activity of an external node to m . An external node is a node that is undeveloped and that has no developed neighbors. The site of the new node is selected randomly and with uniform probability among external sites on the grid

$$\pi_i^b = \delta_i^{(E)} \frac{1}{n_t^{(E)}}, \quad (9)$$

where π_i^b is the probability with which node i is selected to undergo type-3b growth, $\delta_i^{(E)} = 1$ if the node i is external, $\delta_i^{(E)} = 0$ if the node is not external, and $n_t^{(E)}$ is the number of external nodes at time t .

5. A continuum formulation of the spatial model

A continuum formulation for the evolution of developed nodes in the spatial model can, as in Sec. II A 1, be constructed by studying the time evolution of expected node degrees:

$$x_i(t+1) = x_i(t) + E[p_i(t)] + E[s_i(t)], \quad (10)$$

where $E[s_i]$ is the expected secondary increase which can be calculated by a weighted summation over the expected primary increments,

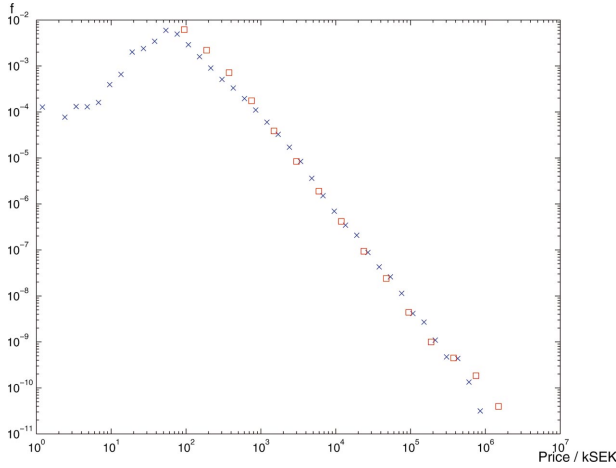


FIG. 2. (Color) Double logarithmic histogram of simulation results and empirical data. Simulation results are denoted with squares in the figure and they are a mean of the results of three runs of the spatial model. The exponent of the model output has been tuned to match the empirical data. As indicated in Eq. (6) this is done by setting the relative proportions of the growth types, in this case $q_1=0.1$, $q_2=0.6$, $\epsilon=0.01$, $c=0.2$, $\alpha=1$, $m=100$ kSEK, and $t=170\,000$. The exponent is roughly 2.1, which is close to the value of $\gamma=2.25$ that is predicted by Eq. (6) for these parameter values in the nonspatial model. A slightly different value for the spatial model must be expected because of spatial bias in the growth dynamics. The sharp transition that occurs around a land price of 60 kSEK (1 USD \approx 10 SEK) per 100×100 m² marks the difference in dynamics between trade based urban activities and rural activities whose values are not described by the presented model.

$$E[s_i] = \sum_j E[p_j] \frac{D_{ij} x_i}{\sum_k D_{kj} x_k}. \quad (11)$$

For the special case of our simple model for the primary effects we have

$$E[p_j] = \delta_j^{(D)} (p_j^{(1)} + p_j^{(2)}) + \delta_j^{(P)} p_j^{(3a)} + \delta_j^{(E)} p_j^{(3b)}, \quad (12)$$

with

$$p_j^{(1)} = q_1 \frac{m}{n_t^{(D)}}, \quad (13)$$

$$p_j^{(2)} = q_2 \frac{m x_j}{\sum_k x_k}, \quad (14)$$

$$p_j^{(3a)} = q_3 (1 - \epsilon) \frac{m}{n_t^{(P)}}, \quad (15)$$

$$p_j^{(3b)} = q_3 \epsilon \frac{m}{n_t^{(E)}}. \quad (16)$$

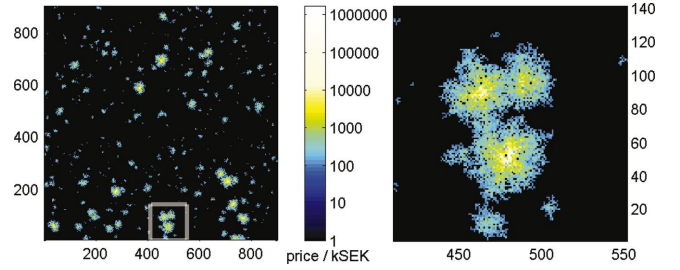


FIG. 3. (Color) Simulated land price configuration from one of the runs of the spatial model accounted for in Fig. 2. The left panel shows the full lattice and the right one shows a detailed picture of the largest cluster.

The total growth for a developed node can in this case be separated into a uniform and a preferential part as

$$x_i(t+1) = x_i(t) + \zeta_i(t) + \eta_i(t)x_i(t), \quad (17)$$

with

$$\zeta_i = q_1 \frac{m}{n_t^{(D)}} \quad (18)$$

and

$$\eta_i = \frac{q_2 m}{\sum_k x_k} + \sum_j E[p_j] \frac{D_{ij}}{\sum_k D_{kj} x_k}. \quad (19)$$

Noting that $\sum_i \eta_i x_i = m(q_1 + 2q_2 + q_3)$, Eq. (17) can be rewritten as

$$x_i(t+1) = x_i(t) + \zeta_i(t) + m(q_1 + 2q_2 + q_3) \frac{\eta_i(t)x_i(t)}{\sum_j \eta_j(t)x_j(t)}, \quad (20)$$

which, in a comparison with Eq. (3), reveals that the only difference between the spatial and the nonspatial models is the site and configuration-dependent parameter η_i . This is similar to the concept of node fitness, as presented in Refs. [11,12], which can affect the node degree distribution. However, our simulation results indicate that η_i falls within a sufficiently narrow interval for the power law to be essentially preserved (Fig. 2). This is also supported by calculations of η_i for both simulated and empirical data.

C. The network model in an urban economics context

1. The connection between node degree and land value

An approximate linear relationship between node degree in the model and land value in the real system is crucial for the interpretation of our results. The motivation follows from (i) market pricing of goods and services and (ii) the connection between trade benefits and land value.

(i) Market pricing of commodities provides an adaptive measure that allows us to compare the activities that generate

them. Hence, on an average, an edge contributes identically to the value of both nodes to which it connects. This contribution is exactly our definition of activity, which implies that the degree of a node is proportional to its benefits due to trade.

(ii) This connection consists of two proportionalities. For a node i we have

$$v_i \propto r_i \propto x_i, \quad (21)$$

where v_i is the value of the corresponding land area, r_i is the bid rent [13,14], and x_i is the total trade benefits as outlined above. Capitalizing periodic rent income from the site i gives land value $v_i = r_i / i$ where i is the interest rate [5]. The second proportionality is a weak form of the leftover principle from urban economics, which states that, in a competitive land market, rent equals the amount of money leftover after all expenses (except rent) are paid. This amount of money equals the sum of all trade benefits at the site. For our results it is sufficient that, on average, a certain proportion of each new unit of trade benefit goes to the landowner.

Together, (i) and (ii) suggest an approximately linear relationship between node degrees and land values.

2. Types of growth

Urban activity can increase in essentially two ways: either a new activity is related to, or it is unrelated to, an existing activity at the site. In the former case (preferential growth) this could be a new employee hired as a response to increased demand, in the latter case (uniform) it could be the establishment of a new firm. Preferential growth corresponds to a per-unit activity rate. Uniform growth corresponds to establishment among lots on a competitive land market where, for the average land use, we cannot expect any lot to be more profitable than any other.

3. Reasons for treating perimeter growth separately from internal growth

The jagged perimeter of urban areas exposes large areas of undeveloped land to urban infrastructure, thus making it attractive for urban land use. Because perimeter land is in ample supply and currently has a low revenue, even land uses with a very low trade gain can be competitive. Many low-activity land uses in the outskirts of the urban area can likely just barely out-bid agriculture and would not be viable in competition with other urban land uses. Among high-activity land uses some have very specific demands on land improvements and therefore cannot benefit from buying existing buildings inside the urban area. This creates a special case for perimeter land. Note that just like for type-1 growth, competition prevents prediction of where the next growth event will take place among the perimeter nodes, and no preferential growth is possible since there is no previous activity that can expand.

4. Node fitness and growth

If we only regard trading activities, the only difference between two sites with identical activity is the value of their

spatial context. Therefore, in the spatial model, secondary growth is not homogenous, see Eq. (7). Node fitness [Eq. (17)] can be viewed as a local interest rate that predicts the growth rate of activity investments made in that site. The result of this is nontrivial growth predictions since the expected amount of new local development does not become a simple fraction of current development.

Most notably, the model predicts the emergence of urban subcenters. This is realized by examining the expected secondary effects, $E[s_i]$, in the spatial model [Eq. (11)]. Apart from being proportional to the amount of present activity, $E[s_i]$ is also subject to site competition and will be relatively high for nodes that have high activity in relation to their own neighborhoods. For each possible primary effect in j , the node development x_i is weighted with the fraction between trade intensity between site j and i and the sum of all other trading options for the primary effect under consideration. Thus, nearby high-intensity nodes will not necessarily benefit a small neighbor.

III. RESULTS

A. Land values are power-law distributed

From empirical data, land values in Sweden are demonstrated to follow a power-law distribution for the higher range of land values (see Fig. 2). The sharp transition that can be observed around 60 kSEK suggests that two truly different mechanisms generate the prices below and above this point. This is in agreement with the observation that the pricing mechanism we suggest would apply only to trading activity.

The data we have used are based on the land value component of market value estimations of about 2.9 million units of real estate in Sweden. The data were originally compiled by the Swedish National Land Survey and coded by Sweden Statistics to geographical coordinates. The data points we use are aggregated land values into 100×100 m² squares.

Our results are supported by a recent study by Kaizoji that shows scale-free behavior of land prices in Japan, with an exponent γ ranging from 2.53 to 2.76 [15].

B. Power-law distributed prices are predicted by the nonspatial model

We have developed a simple network model of the urban economy based on the Barabási-Albert model by mapping fundamental assumptions from urban economics to the ontology of the network model (see Sec. II A). In our derivation, we have determined the exponent of the node degree distribution, and thus, per our definition, also the predicted land values, to follow Eq. (6). To the extent that our interpretation of the underlying dynamics is correct this demonstrates why urban land values can be expected to follow a power law and how the exponent may depend on parameters.

C. The spatial model retains power-law statistics

As discussed in Sec. II B 5, the impact of spatial constraints closely resembles that of node fitness [11,12]. This

could potentially result in the distribution for the spatial network model becoming a sum of many power laws with different exponents.

In Fig. 2 we plot results from simulations showing that node degrees in the spatial model follow a power-law distribution. The model parameters have been tuned (see Sec. II B 5) to reproduce the distribution of the observed land prices. To visually illustrate how the model works, Fig. 3 shows the spatial configuration of the land values from one of the simulations.

IV. CONCLUSION

We present empirical evidence that land prices follow a power-law distribution for urban land uses, and present a generic model, based on the underlying trade network, that reproduces this behavior. The model is a version of the Barabási-Albert scale-free network that is also extended to incorporate spatial constraints.

The applicability of the network paradigm to urban growth suggests that scale invariance in urban systems is caused by: (i) growth and (ii) preferentiality in how new trade connections are formed between areas of land. Growth in this context refers to the continual development of new land. Preferentiality is a consequence of point-to-point inter-

actions between activities occupying the land areas. Note that many other observables, such as population and urban land use intensity, might be highly correlated with land value.

The spatial model that we present can have more general applicability beyond urban economics. Other spatially growing networks are communication networks, transportation networks, electricity and utility networks. These can be expected to follow a similar type of growth since they are intimately connected to urban activity.

The network architecture is generic and allows for addition of any amount of detail. Also, being based on trade relations, the model produces output in units of currency. Because of this, such network models can provide a bridge between realistic microscopic dynamics and empirically observed emergent economic properties.

Further possible directions for research on urban economic networks include interpretation of other theoretical network results in terms of urban dynamics, finding empirical parameters for scenario predictions and model validation.

ACKNOWLEDGMENTS

We would like to thank Martin Nilsson, Steen Rasmussen, and Roger White for valuable input and discussions.

-
- [1] G.K. Zipf, *Human Behavior and the Principle of Least Effort* (Addison-Wesley, Reading, MA, 1949).
 - [2] Albert-László Barabási and Réka Albert, *Science* **286**, 509 (1999).
 - [3] A.-L. Barabási, H. Jeong, R. Ravasz, Z. Nédá, T. Vicsek, and A. Schubert, *Rev. Mod. Phys.* **74**, 47 (2002).
 - [4] S.N. Dorogovtsev and J.F.F. Mendes, *Adv. Phys.* **51**, 1079 (2002).
 - [5] Arthur O'Sullivan, *Urban Economics* (McGraw-Hill, New York, 2002).
 - [6] S.N. Dorogovtsev, J.F.F. Mendes, and A.N. Samukhin, *Phys. Rev. Lett.* **85**, 4633 (2000).
 - [7] Réka Albert, Albert-László Barabási, and Hawoong Jeong, *Physica A* **272**, 173 (1999).
 - [8] Marc Barthélemy, e-print cond-mat/0212086.
 - [9] R. Xulvi-Brunet and I.M. Sokolov, *Phys. Rev. E* **66**, 026118 (2002).
 - [10] S.S. Manna and Parongama Sen, *Phys. Rev. E* **66**, 066114 (2002).
 - [11] G. Bianconi and A.-L. Barabási, *Europhys. Lett.* **54**, 436 (2001).
 - [12] G. Ergün and G.J. Rodgers, e-print cond-mat/0103423.
 - [13] W. Alonso, *Readings Urban Econ.* **6**, 149 (1960).
 - [14] W. Alonso, *Location and Land Use* (Harvard University Press, Cambridge, 1964).
 - [15] Taisei Kaizoji, e-print cond-mat/0302470.