Superlinear and sublinear urban scaling in geographical networks modeling cities

K. Yakubo* and Y. Saijo
Department of Applied Physics, Hokkaido University, Sapporo 060-8628, Japan

D. Korošak†
University of Maribor, Slomškov trg 15, Maribor SI-2000, Slovenia

(Received 16 January 2014; revised manuscript received 28 May 2014; published 4 August 2014)

Using a geographical scale-free network to describe relations between people in a city, we explain both superlinear and sublinear allometric scaling of urban indicators that quantify activities or performances of the city. The urban indicator $Y(N)$ of a city with the population size N is analytically calculated by summing up all individual activities produced by person-to-person relationships. Our results show that the urban indicator scales superlinearly with the population, namely, $Y(N) \propto N^{\beta}$ with $\beta > 1$, if $Y(N)$ represents a creative productivity and the indicator scales sublinearly ($\beta < 1$) if $Y(N)$ is related to the degree of infrastructure development. These results coincide with allometric scaling observed in real-world urban indicators. We also show how the scaling exponent β depends on the strength of the geographical constraint in the network formation.

DOI: 10.1103/PhysRevE.90.022803 PACS number(s): 89.75.Da, 89.65.Lm, 89.75.Hc

I. INTRODUCTION

Cities are often compared to living organisms with a hierarchical organization consisting of cells, tissues, and organs. Likewise, people in a city form groups, groups form organizations serving certain functions, and interdependent complex relationships between functional organizations sustain the whole urban activities. Such similarities are found not only in the correspondence between constituent elements of cities and living organisms, but also in allometric scaling [1]. Similar to living organisms with the basal metabolic rate of an animal proportional to the 3/4 power of its body mass M [2,3] or the breathing rate (or heart rate) proportional to $M^{-1/4}$ [4], there are various quantities related to activities or performances of cities such as urban road systems [5], night illuminations [6], and size distribution of buildings [7] that are described by power-law relations [8]. Power-law scaling has also been found in the morphology and evolution of cities and argued from a viewpoint of fractal cities [9–13]. In particular, it has been elucidated that urban indicators quantifying city activities on average scale with the population size in a power-law manner. These include many creative productivities and infrastructure volumes [14–17] such as the gross domestic product (GDP), the number of patents, human online activities [18], prosocial behaviors [19], quantities related to highway systems [20], the number of crimes [21], the number of supply stations [22], and emissions of CO_2 [23] and NO_2 [24]. There exist several models explaining such allometric scaling behavior [25–28] and fluctuations around average power-law relations [29–31]. These findings allow us to write an urban indicator Y as a function of the population size N as

$$Y(N) \propto N^{\beta}, \tag{1}$$

where β is a scaling exponent. Bettencourt *et al.* [14–16] found that an urban indicator representing a creative productivity, such as the number of new patents, the GDP, the number of crimes, etc., obeys a superlinear scaling law ($\beta > 1$), while an indicator related to the degree of infrastructure development, such as the total length of electrical cables, the number of gas stations, the total road surface, etc., scales sublinearly with the population size ($\beta < 1$). Due to the nonlinear scaling (1), a meaningful comparison between characteristics of individual cities requires evaluation of deviations from this average scaling behavior, instead of considering per capita quantity $Y(N)/N$ [28–31].

It is crucial to understand the reason why urban indicators representing creative productivities scale superlinearly and those corresponding to material infrastructures scale sublinearly. Arbesman, Kleinberg, and Strogatz [25] proposed a network model to explain superlinear scaling found in creative productivities. They introduced hierarchical social distances between nodes representing people in a city. A network is formed by connecting nodes with the edge probability decaying exponentially with the social distance. Assuming that the individual productivity yielded by an edge increases exponentially with the social distance, the Arbesman-Kleinberg-Strogatz (AKS) model gives superlinear scaling of creative productivity $Y(N)$ if the total contribution from connected node pairs separated by the social distance d is an increasing function of d; otherwise the scaling is linear.

In order to explain both superlinear and sublinear scaling of urban indicators, Bettencourt [26] has worked with four simple assumptions. (i) Citizens explore the city fully to benefit from it and the city develops in a way to make this possible. (ii) The infrastructure network volume A_n grows in a decentralized way in order to connect each addition of a new inhabitant, namely, $A_n \propto Nr$, where N is the number of people in the city and r is the average distance between individuals. (iii) The product of average social output and the volume spanned by an individual's movement is a constant of city size N, which means that human effort is bounded. (iv) The urban indicator $Y(N)$ related to a creative productivity is proportional to the number of local social interactions. According to Bettencourt's model, the scaling exponent is given by $\beta = 1 + \delta$ for the superlinearly scaled creative productivity and $\beta = 1 - \delta$ for the sublinearly scaled infrastructure volume, where δ is a

*yakubo@eng.hokudai.ac.jp
†dean.korosak@um.si

positive exponent that depends on the fractal dimension of human travel paths.

Despite these two pioneering and suggestive theories, the mechanism of urban scaling is not yet completely understood. Although the AKS model [25] gives a possible explanation of superlinear (or linear) scaling of creative productivities and describes how the social structure (i.e., human relations) affects the scaling exponent β, sublinear scaling for urban indicators reflecting infrastructures has not been argued. On the other hand, Bettencourt's model [26] demonstrates both superlinear and sublinear scaling for a creative productivity and an infrastructure volume, respectively. However, it is not clear how the social structure influences urban scaling because this theory is based on a continuum model. Furthermore, the scaling exponent β always appears symmetrically as $\beta = 1 \pm \delta$ for superlinear and sublinear scaling, hence a variety of real-world nonlinear urban scaling cannot be described by this model. It is therefore important to explain consistently both superlinear and sublinear scaling in the context of the relation between the scaling behavior and the social structure in the city. This enables us to create alternative ways to understand the structure and function of cities.

In this paper we propose a model to account for urban scaling by representing human relations in a city by a geographical network in which nodes close to each other are more likely to be connected. It is assumed that an urban indicator $Y(N)$ is given by the sum of the activities produced by individual connected node pairs and that the individual activity y_{ij} depends on the Euclidean distance l_{ij} between connected nodes i and j. We show that the urban indicator scales superlinearly or linearly with the population size N when the activity y_{ij} represents a creative productivity that is an increasing function of the Euclidean distance l_{ij} and scales sublinearly or linearly if y_{ij} decreases with l_{ij} as the strength of the demand for infrastructure does. This result is consistent with observed urban scaling phenomena. We also predict that urban indicators representing either creative productivities or infrastructures are proportional to the population size (i.e., linear scaling) if the geographical constraint in the network formation is strong enough.

The paper is organized as follows. After presenting our model in Sec. II, the urban scaling exponent β is analytically calculated in Sec. III. Numerical confirmations for the analytical results are given in Sec. IV. We also show how the exponent β depends on parameters characterizing our model. Finally, we summarize our work in Sec. V.

II. MODEL

A. Geographical network model

It has been demonstrated that urban structure possesses a self-similar property [9–13,32–35]. It is therefore meaningful to explain urban scaling in a fractal-city framework, as allometric scaling in living organisms has been modeled with an idea of the fractal structure of living systems [3]. In order to incorporate the fractal population density into the theory, we adopt the geographical network model studied by Ref. [36] to describe the social structure in a city. In this model, N nodes embedded in an underlying Euclidean space, which represent people in a city, are inhomogeneously distributed such that their spatial distribution exhibits fractality with the fractal dimension D. The Euclidean distance is defined for any pair of nodes. The underlying Euclidean space has the linear size L and is assumed to be large enough and isotropic from any point such as a spherical surface [37]. Thus, the number of nodes or the population size of the city is presented by

$$n(l) \propto l^{D-1}, \tag{2}$$

where $n(l)dl$ is the number of nodes within the distance range of $[l,l+dl]$ from a given node. Each node (person) has its own ability or charm to attract others, such as wealth, communication skills, or leadership. Many of the quantities representing these personal characteristics follow fat-tail distributions [38]. For example, the distribution functions of measures of influence in cyberspace [39–41] and individual professional performance [42] obey power-law forms, in addition to the well-known Pareto distribution of income [43,44]. In order to quantify such personal attractiveness, a real continuous quantity x (referred to as attractivity hereafter) is randomly assigned for each node according to the power-law probability distribution function $s(x)$ expressed by

$$s(x) \propto x^{-\alpha} \quad (x \geqslant x_{\min}), \tag{3}$$

where $x_{\min}$ must be positive. Since our theory does not require the existence of the average $\langle x \rangle$ whereas $s(x)$ must be normalized, the exponent α can take any value in the range of $\alpha > 1$. It is natural to consider that two nodes spatially close to each other and having large attractivity values are more likely to be connected, thus two nodes i and j are connected if the following condition is satisfied [36]:

$$\frac{x_i x_j}{l_{ij}^m} > \Theta, \tag{4}$$

where l_{ij} denotes the Euclidean distance between the nodes i and j, m ($\geqslant 0$) is a parameter controlling the strength of the geographical constraint in the network formation, x_i is the attractivity of the node i, and Θ is a threshold value. For $m = 0$, the network connectivity is determined only by the attractivity regardless of node locations, whereas in the case of an infinitely large m each node is connected to n_Θ closest nodes independently of the attractivity, where n_Θ is a constant determined by Θ. Similar network models have already been studied [45–51] and the present model [36] is a straightforward extension of these models, by introducing the parameter m and the fractal dimension D.

Statistical properties of networks formed by the above procedures have been studied previously in the mean-field framework in which all nodes are statistically equivalent with respect to their placements [36,50]. We briefly summarize the results of these works here. First, the network exhibits the scale-free property [52,53], that is, the number of edges (degree) from a node is distributed in a power-law manner. It has been shown that the distribution function $P(k)$ of the degree k of the present network model obeys

$$P(k) \propto k^{-\gamma} \tag{5}$$

for large k [36,50]. The exponent γ is related to the model parameters D, α, and m through [36]

$$\gamma = \begin{cases} 2 & \text{for } D \geqslant d_c \\ 1 + \dfrac{d_c}{D} & \text{for } D < d_c, \end{cases} \tag{6}$$

where

$$d_c = m(\alpha - 1). \tag{7}$$

This result shows that the degree distribution becomes more homogeneous when the geographical constraint is enhanced by increasing m [54]. This is because the network formed by a large m value has a latticelike structure.

Second, the probability distribution function $R(l)$ of the edge length l is proportional to the average number of nodes $k(l)dl$ within the distance range of $[l,l+dl]$ from a given node that are connected to the node of the origin, averaged over all the possible attractivity values and node arrangements. The quantity $k(l)dl$ is nothing but the average number of edges of the length in the range of $[l,l+dl]$ from a given node. These are given by [36]

$$R(l) \propto k(l) \propto \begin{cases} l^{D-1} & \text{for } l \leqslant \xi \\ l^{D-1}\left(\dfrac{l}{\xi}\right)^{-d_c} & \text{for } l > \xi, \end{cases} \tag{8}$$

where

$$\xi = \left(\frac{x_{\min}^2}{\Theta}\right)^{1/m}. \tag{9}$$

The quantity ξ is the distance below which any two nodes are connected regardless of the attractivity x. Here we neglected a logarithmic correction term proportional to $l^{D-1}(l/\xi)^{-d_c}\ln(l/\xi)$ in Eq. (8) for $l > \xi$ because it does not influence the asymptotic power-law behaviors of $R(l)$ and $k(l)$ for large l. The probability of two nodes separated by the Euclidean distance l to be connected by an edge is directly obtained from Eq. (8). This probability $g(l)$ is presented by the ratio of $k(l)dl$ to the number of nodes $n(l)dl$ located at a distance within the range of $[l,l+dl]$ from a given node. Since $n(l) \propto l^{D-1}$, the relation $g(l) = k(l)/n(l)$ immediately leads to

$$g(l) = \begin{cases} 1 & \text{for } l \leqslant \xi \\ \left(\dfrac{l}{\xi}\right)^{-d_c} & \text{for } l > \xi. \end{cases} \tag{10}$$

The power-law decay of $g(l)$ for $l > \xi$ is consistent with the fact that the probability of two people separated by l to be socially connected decreases with l in a power-law manner [55–58]. The relation $g(l) = 1$ for $l \leqslant \xi$ is obvious from the meaning of the distance ξ.

Third, the average degree $\langle k\rangle$ can be controlled by tuning the threshold Θ. The quantity $\langle k\rangle$ is averaged under fixed values of D, α, and m. More precisely, the average degree in the present model is defined as

$$\langle k\rangle = \frac{1}{N}\sum_{i=1}^{N}\iint k_i(\boldsymbol{r},\boldsymbol{x})p(\boldsymbol{r},\boldsymbol{x})d\boldsymbol{r}d\boldsymbol{x}, \tag{11}$$

where $k_i(\boldsymbol{r},\boldsymbol{x})$ is the the degree of the node i in a specific network where nodes have their coordinates $\boldsymbol{r} = (r_1,r_2,\ldots,r_N)$ and attractivities $\boldsymbol{x} = (x_1,x_2,\ldots,x_N)$ and $p(\boldsymbol{r},\boldsymbol{x})$ represents the joint probability of N nodes to have the coordinates $\boldsymbol{r}$ and the attractivities $\boldsymbol{x}$. Although the Θ dependence of $\langle k\rangle$ has been already studied [36], here we clarify not only the Θ dependence but the N dependence of $\langle k\rangle$. The average degree is obviously given by

$$\langle k\rangle = \int_0^L k(l)dl, \tag{12}$$

where the linear size L of the city is related to the population size N through $N = \int_0^L n(l)dl$, with $n(l)$ given by Eq. (2). Substituting Eq. (8) into Eq. (12), $\langle k\rangle$ can be calculated as

$$\begin{aligned}\langle k\rangle &= c_1\int_0^{\xi} l^{D-1}dl + c_2\int_{\xi}^{L}\left(\frac{l}{\xi}\right)^{-d_c} l^{D-1}dl \\ &= \left(\frac{c_1}{D} - \frac{c_2}{D-d_c}\right)\xi^D + \frac{c_2}{D-d_c}\xi^{d_c}L^{D-d_c}, \end{aligned} \tag{13}$$

where c_1 and c_2 are irrelevant numerical coefficients. Here we define a relation symbol $\propto$: to represent the relation $A = cx + c'y$ by $A \propto: x + y$ if c and c' are nonzero constants independent of x and y. Using this notation, Eq. (13) can be written as $\langle k\rangle \propto: \xi^D + \xi^{d_c}L^{D-d_c}$. Thus, the relation $L \propto N^{1/D}$ from Eq. (2) and Eq. (9) lead to

$$\langle k\rangle \propto: \Theta^{-D/m} + \Theta^{-d_c/m}N^{1-d_c/D}. \tag{14}$$

Therefore, we obtain

$$\langle k\rangle \propto \begin{cases} \Theta^{-D/m} & \text{for } D \leqslant d_c \\ \Theta^{-d_c/m}N^{1-d_c/D} & \text{for } D > d_c \end{cases} \tag{15}$$

for a large enough value of N. We should note that the average degree $\langle k\rangle$ for $D > d_c$ depends on N under a fixed Θ. Consequently, if we keep $\langle k\rangle$ for $D > d_c$ constant for any N, Θ must be changed as $N^{m(1/d_c-1/D)}$. Equations (14) and (15) give the relation between different parameters of the model. In this context, the degree k of a node is also related to the attractivity x of the node. It is obvious from the connection condition (4) that a node with a large attractivity must have a large number of connections and vice versa. In fact, it has been shown by Ref. [36] that the degree k is given by

$$k \propto: \Theta^{-D/m}x^{D/m} + \Theta^{-d_c/m}N^{1-d_c/D}x^{d_c/m}, \tag{16}$$

which leads, for large N, to $k \propto x^{D/m}$ for $D \leqslant d_c$ and $k \propto x^{d_c/m}$ for $D > d_c$. These analytical results have been numerically confirmed for uniform node sets in which nodes are uniformly distributed in a two-dimensional space and for fractal node sets in which nodes are placed in a fractal manner [36].

The topology of networks formed by the above model [36] depends on the model parameters m, α, and D. Since the network topology is reflected in the distribution functions $P(k)$ and $R(l)$ and in the average degree $\langle k\rangle$, it is meaningful to summarize briefly how these statistical quantities are affected by such parameters. From Eqs. (5), (6), and (8), $P(k)$ and $R(l)$ depend on m and α only through d_c. Therefore, the parameters m and α influence both distribution functions in a similar way. We see from these equations that $P(k)$ and $R(l)$ become narrower with increasing m or α. On the other hand, $\langle k\rangle$ depends not only on d_c but also on m as presented by

Eq. (15). The average degree $\langle k \rangle$ is a decreasing function of m for $m < D/(\alpha - 1)$ and turns to an increasing function for $m \geqslant D/(\alpha - 1)$, while $\langle k \rangle$ decreases with α for $\alpha < 1 + D/m$ and becomes constant for $\alpha \geqslant 1 + D/m$. The increase of the fractal dimension D makes both distribution functions broader and enlarges $\langle k \rangle$.

B. Urban indicator

In order to clarify the scaling property of an urban indicator $Y(N)$ quantifying activities in a city, we must relate $Y(N)$ to relations between people in the city modeled by a geographical network described above. Although actual urban performances are sometimes produced by a cooperation between many people in a group or an organization, we consider here that the total urban performance stems from one-to-one relationships, namely, from individual connected node pairs of people in the network. Furthermore, we neglect nonlinear effects such as interactions between individual node-pair activities creating additional activities. These simplifications allows us to write the urban indicator as

$$Y(N) = \frac{1}{2} \sum_{i,j}^{N} a_{ij} y_{ij}, \tag{17}$$

where a_{ij} is the (i,j) element of the adjacency matrix of the network and y_{ij} is the individual activity between nodes i and j.

As in the case of the AKS model [25] in which the individual productivity is assumed to increase with the *social* distance d, it is natural to consider that the individual activity y_{ij} depends on the *Euclidean* distance l_{ij} between nodes i and j. Instead of the exponential d dependence in the AKS model, we assume a power-law dependence of y_{ij} on l_{ij}, i.e.,

$$y_{ij} \propto l_{ij}^{\eta}, \tag{18}$$

where the exponent η can take either positive or negative values. If η is positive, longer-distance connections give rise to higher activities. In this case, we can regard y_{ij} as a creative productivity because distant individuals usually have different experiences and values and the fusion of heterogeneous ideas often leads to greater creativity compared to combinations of homogeneous ideas. This interpretation is consistent with the geographical network model presented in the previous section. In the network model, a long-distance connection is established only when two nodes have large attractivity, namely, they are highly capable. Outputs by collaboration between such talented individuals must be innovative. Since we consider that y_{ij} between individuals close to each other is always small even if x_i and x_j are very large, the individual activity is not defined as a product of x_i and x_j in this work, though a node pair with large $x_i x_j$ is likely to be connected and gives large y_{ij}. Note that y_{ij} given by Eq. (18) is influenced by the values of x_i and x_j through l_{ij} despite the apparent independence from the attractivity and by defining y_{ij} separately from the attractivity, we can explain both superlinear and sublinear urban scaling in a unified way, as we will show. Due to this definition of y_{ij}, networks must be embedded in the Euclidean space in our model even for $m = 0$.

On the other hand, if η is negative and y_{ij} decreases with l_{ij}, short-distance connections contribute more significantly to the total urban indicator $Y(N)$ than long-distance ones. In this case, the following consideration suggests that $Y(N)$ represents an infrastructure volume. The degree of infrastructure development depends on how strong the demand for the infrastructure is. Since infrastructure facility, such as electrical power cables, railway stations, or green open urban spaces, provides services for inhabitants near the facility, the social need for the infrastructure arises from local consensus among neighboring residents in areas having no access to the infrastructure. Thus, the consensus between residents close to each other must be stronger than that between distant ones. If we regard y_{ij} given by Eq. (18) with negative η as the strength of the consensus between nodes i and j, $Y(N)$ provided by Eq. (17) quantifies the whole social need in the city. Considering that infrastructure facilities are realized in proportion to the social need, $Y(N)$ is proportional to the infrastructure volume.

III. URBAN SCALING

Here we concentrate on the urban indicator averaged over all possible cities with the same population size N but different spatial arrangements of people. We consider the quantity

$$Y(N) = \frac{1}{2} \left\langle \sum_{i,j}^{N} a_{ij} y_{ij} \right\rangle_{\mathrm{c}}, \tag{19}$$

where $\langle \cdots \rangle_{\mathrm{c}}$ denotes the average over network configurations with different attractivity values and positions of nodes under fixed values of D, α, m, and η. In the mean-field picture, we can write $Y(N) = N\bar{y}/2$, where $\bar{y}$ is the node activity $\sum_j a_{ij} y_{ij}$ averaged over all nodes in possible network configurations. Using the probability $g(l)$ of a node to be connected by an edge of the length l, the average urban indicator is then presented by

$$Y(N) \propto N \int_0^L g(l) y(l) n(l) dl, \tag{20}$$

where $n(l)dl$ is, as defined below Eq. (2), the number of nodes within the range of $[l, l + dl]$ from a given node and $y(l)$ is the individual activity between nodes separated from each other by the distance l. In this section we examine the scaling behavior of $Y(N)$ by evaluating Eq. (20).

Substituting the relations $y(l) \propto l^{\eta}$ from Eq. (18), Eq. (2) for $n(l)$, and Eq. (10) into Eq. (20), we have

$$\begin{aligned} \frac{Y(N)}{N} &\propto: \int_0^{\xi} l^{\eta} l^{D-1} dl + \int_{\xi}^{L} \left(\frac{l}{\xi}\right)^{-d_{\mathrm{c}}} l^{\eta} l^{D-1} dl \\ &\propto: \xi^{D+\eta} + L^{D+\eta} \left(\frac{L}{\xi}\right)^{-d_{\mathrm{c}}}, \end{aligned} \tag{21}$$

where the symbol $\propto:$ has been defined below Eq. (13). Here we assumed

$$\eta > -D \tag{22}$$

for the convergence of the integral at $l = 0$. This condition, however, is not important due to the existence of the minimum node-pair distance in actual spatial arrangements of people. Since the linear size L is related to N through the relation

$N = \int_0^L n(l)dl$, $Y(N)$ is written as

$$\frac{Y(N)}{N} \propto: \Theta^{-(D+\eta)/m} + \Theta^{-d_c/m} N^{1-(d_c-\eta)/D}, \tag{23}$$

where the characteristic length ξ in Eq. (21) was replaced with the threshold Θ by using Eq. (9). Equation (23) tells us how the urban indicator scales with the population size N under a fixed value of the threshold Θ.

In actual cities, the average number of acquaintances corresponding to $\langle k \rangle$ is almost independent of N, though $\langle k \rangle$ depends on N if Θ is constant, as we remarked below Eq. (15). Therefore, we must reveal the scaling behavior of $Y(N)$ under a fixed value of $\langle k \rangle$ instead of a fixed Θ. In order to express $Y(N)$ as a function of N and $\langle k \rangle$, taking into account the argument below Eq. (15), we rewrite Eq. (15) as

$$\Theta \propto \begin{cases} \langle k \rangle^{-m/D} & \text{for } D \leqslant d_c \quad \text{(24a)} \\ \langle k \rangle^{-m/d_c} N^{m(D-d_c)/Dd_c} & \text{for } D > d_c. \quad \text{(24b)} \end{cases}$$

In the case of $D \leqslant d_c$, substitution of Eq. (24a) into Eq. (23) yields

$$Y(N) \propto: \langle k \rangle^{1+\eta/D} N + \langle k \rangle^{d_c/D} N^{2+(\eta-d_c)/D}. \tag{25}$$

This relation is valid for a large enough population size because Eq. (24) derived from Eq. (15) holds for a large N. In this case, one of two terms in Eq. (25) dominates $Y(N)$ depending on the value of the exponent of N. If $2 + (\eta - d_c)/D \leqslant 1$, namely, $D \leqslant d_c - \eta$, the first term grows with N faster than the second term and we have linear scaling of $Y(N)$, i.e.,

$$Y(N) \propto N \quad \text{for } D \leqslant d_c,\ D \leqslant d_c - \eta. \tag{26}$$

For $D > d_c - \eta$, however, the second term of Eq. (25) dominates $Y(N)$. Thus, $Y(N)$ scales as

$$Y(N) \propto N^{2+(\eta-d_c)/D} \quad \text{for } d_c - \eta < D \leqslant d_c. \tag{27}$$

On the other hand, for $D > d_c$, substitution of Eq. (24b) into Eq. (23) leads to

$$Y(N) \propto: \langle k \rangle^{(D+\eta)/d_c} N^{[d_c(2D+\eta)-D(D+\eta)]/Dd_c} + \langle k \rangle N^{1+\eta/D}. \tag{28}$$

Similarly to the case of Eq. (25), the comparison between the exponents $[d_c(2D+\eta) - D(D+\eta)]/Dd_c$ and $1 + \eta/D$ gives

$$Y(N) \propto N^{[d_c(2D+\eta)-D(D+\eta)]/Dd_c} \quad \text{for } d_c < D \leqslant d_c - \eta \tag{29}$$

and

$$Y(N) \propto N^{1+\eta/D} \quad \text{for } D > d_c,\ D > d_c - \eta. \tag{30}$$

These relations provide nonlinear scaling of the urban indicator $Y(N)$.

Summarizing the above results, the scaling exponent β in Eq. (1) is given by

$$\beta = \begin{cases} 1 & \text{for } D \leqslant d_c,\ D \leqslant d_c - \eta \quad \text{(31a)} \\ 2 + \dfrac{\eta - d_c}{D} & \text{for } d_c - \eta < D \leqslant d_c \quad \text{(31b)} \\ 2 + \dfrac{\eta}{D} - \dfrac{D+\eta}{d_c} & \text{for } d_c < D \leqslant d_c - \eta \quad \text{(31c)} \\ 1 + \dfrac{\eta}{D} & \text{for } D > d_c,\ D > d_c - \eta. \quad \text{(31d)} \end{cases}$$

TABLE I. Urban scaling exponent β at the extreme values of the model parameters.

Parameter	Extreme value	β
m	0	$1 + \eta/D$
	∞	1
α	1	$1 + \eta/D$
	∞	1
D	0	1 [for $\eta \leqslant m(\alpha - 1)$]
		∞ [for $\eta > m(\alpha - 1)$]
	∞	1
η	$-D$	1
	∞	∞

The exponent β can take any positive value by controlling the four parameters D, α, m, and η. This implies that the urban indicator in our model scales superlinearly ($\beta > 1$), linearly ($\beta = 1$), or sublinearly ($\beta < 1$) with the population size N. It is meaningful to know the exponent β in extreme situations. Table I lists values (or forms) of β at the limiting values of the four parameters calculated by Eq. (31). From the value of β at $m = 0$, we see that superlinear or sublinear urban scaling is realized depending on the sign of η even for networks formed by a nongeographical mechanism. However, since $Y(N)$ is defined through Euclidean distances between nodes, this does not mean that networks do not need to be embedded in Euclidean space. The urban indicator $Y(N)$ always obeys linear scaling for large enough m, when the geographical constraint in the network formation is very strong. Since the network formed by a large m value has a latticelike structure as mentioned below Eq. (7), the lengths of edges in the network are almost constant. This is also confirmed by the fact that the edge-length distribution $R(l)$ given by Eq. (8) becomes narrower as m increases. If edge lengths are constant, individual node-pair activities given by $y_{ij} \propto l_{ij}^{\eta}$ are also constant. Denoting this constant by y_0, Eq. (19) gives $Y(N) = N\langle k \rangle y_0/2$, which leads to linear scaling. The exponent α close to unity implies a very inhomogeneous distribution of the attractivity. In this case, the network formation is governed by $x_i x_j$ and the geographical effect becomes relatively weak. Thus, we have the same result of β as that for $m = 0$. In contrast, the infinite α makes the effect of the attractivity weak and enhances relatively the geographical effect. Therefore, we have the same result as that for the infinite m. Although the situations of $D = 0$ and $D \to \infty$ are unrealistic from the viewpoint of real-world population densities, the exponent β in these limiting cases is explained as follows. The divergent β for $\eta > m(\alpha - 1)$ in the limit of $D = 0$ is reasonable because the system must have very long edges to keep $\langle k \rangle$ constant and $Y(N)$ increases rapidly with N. In addition, an infinitely large value of D means that all nodes are the nearest neighbors of a node and the consequent constant edge length leads the same result as β for $m \to \infty$. Furthermore, for the smallest value of η equal to $-D$, the average node activity $\bar{y}$ defined below Eq. (19) is dominated by the shortest edge length $l_{\min}$. The quantity $l_{\min}$ is almost constant for any nodes and for any N under the condition of the constant node density. Thus, Y is

simply proportional to N. Finally, in the limit of $\eta \to \infty$, y_{ij} given by Eq. (18) becomes infinite and Y diverges even for a finite N. Thus, β must be infinite.

Let us consider more precisely the range of values β can take. The exponent β presented by Eq. (31a) obviously leads to linear scaling of $Y(N)$. In this case, the exponent η can be positive or negative. If $\eta \geqslant 0$, the condition for Eq. (31a) is read as $D \leqslant d_c - \eta$, namely, $D + \eta \leqslant m(\alpha - 1)$, while it becomes $D \leqslant d_c$ [i.e., $D \leqslant m(\alpha - 1)$] for $\eta < 0$. Next, β by Eq. (31b) is always larger than 1 because $(\eta - d_c)/D$ is larger than -1 from the condition $d_c - \eta < D$. We should note that the condition for Eq. (31b) requires $\eta > 0$. On the contrary, Eq. (31c) is the case only when $\eta < 0$. Taking into account Eq. (22), η in Eq. (31c) must satisfy $-D < \eta < 0$ actually. Since $(D + \eta)/d_c \leqslant 1$ for Eq. (31c), we have $\beta \geqslant 1 + \eta/D$. In addition, the condition $\eta > -D$ gives $\beta > 0$. Furthermore, β given by Eq. (31c) is expressed as $\beta = 1 + (D + \eta)(1/D - 1/d_c)$. Since $D + \eta > 0$ because of $\eta > -D$ and $(1/D - 1/d_c) < 0$ from $d_c < D$, the value of β is less than 1. Therefore, the exponent β presented by Eq. (31c) can take a value in the interval $0 < \beta < 1$. Finally, for Eq. (31d), η can be positive or negative. If $\eta \geqslant 0$, obviously $\beta \geqslant 1$, whereas $0 < \beta < 1$ for $-D < \eta < 0$.

We can draw the phase diagram of our model from the above arguments. Figure 1(a) shows the regions of three distinct scaling behaviors in the parameter space of η and D under fixed values of m and α and Fig. 1(b) demonstrates those in the parameter space of η and m under fixed values of D and α. The phase boundaries in Fig. 1(b) are translated from Fig. 1(a) by using Eq. (7). These results clearly show that superlinear scaling appears if η is positive and sublinear scaling for negative η. Since the urban indicators $Y(N)$ constructed by positive and negative η correspond to a creative productivity and infrastructure, respectively, these analytical results are consistent with urban scaling observed in the real world [14]. Note that we have linear scaling ($\beta = 1$) on the phase boundaries. Thus, the condition $\eta = 0$ always gives linear scaling regardless of the values of other parameters. This is reasonable because $Y(N)$ for $\eta = 0$ is nothing but the number of edges M in the network and M is proportional to N when $\langle k \rangle$ is independent of N. The urban indicator that scales linearly corresponds to individual human needs such as the total number of houses.

Typical profiles of β given by Eq. (31) are presented in Figs. 2 and 3. Figure 2 shows the η dependence of β for various values of m under $D = 2.0$ and $\alpha = 2.0$. It is verified that superlinear scaling of $Y(N)$ requires $\eta > 0$ and sublinear scaling is allowed only for $\eta < 0$. For any combination of m, α, and D, the exponent β linearly increases with η if η is large enough. The m dependence of β is depicted in Fig. 3(a) for various values of η. This figure clearly demonstrates that

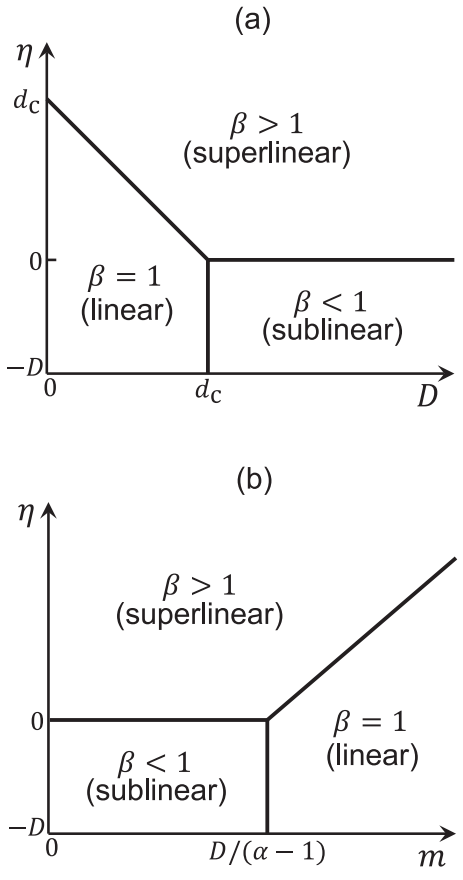

FIG. 1. Phase diagrams of our model in (a) D-η space with fixed values of m and α and (b) m-η space with fixed values of D and α. On the phase boundaries represented by thick lines, β is equal to 1 (linear scaling).

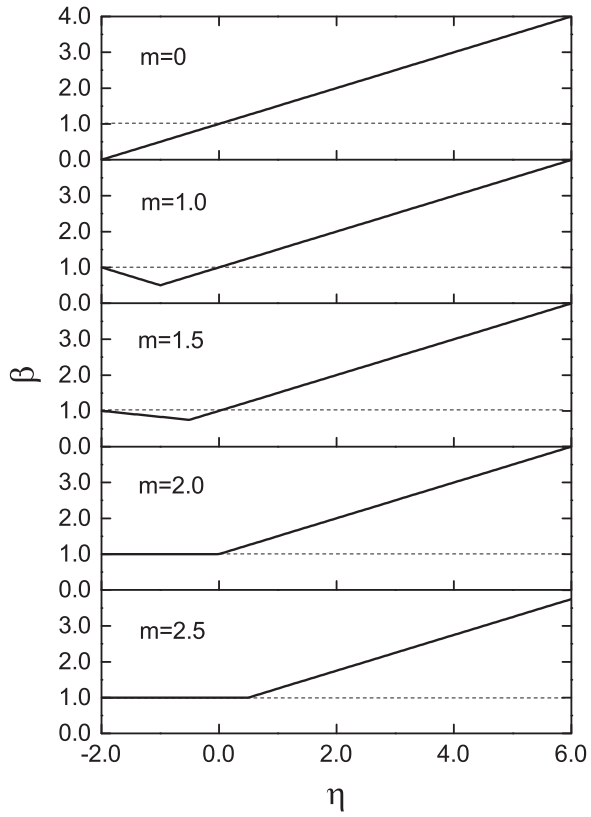

FIG. 2. Urban scaling exponent β as a function of η for several values of m. The exponent α and the fractal dimension D are fixed at $\alpha = 2.0$ and $D = 2.0$. Horizontal dashed lines at $\beta = 1$ are guides to the eye, which separate the superlinear scaling region from the sublinear one.

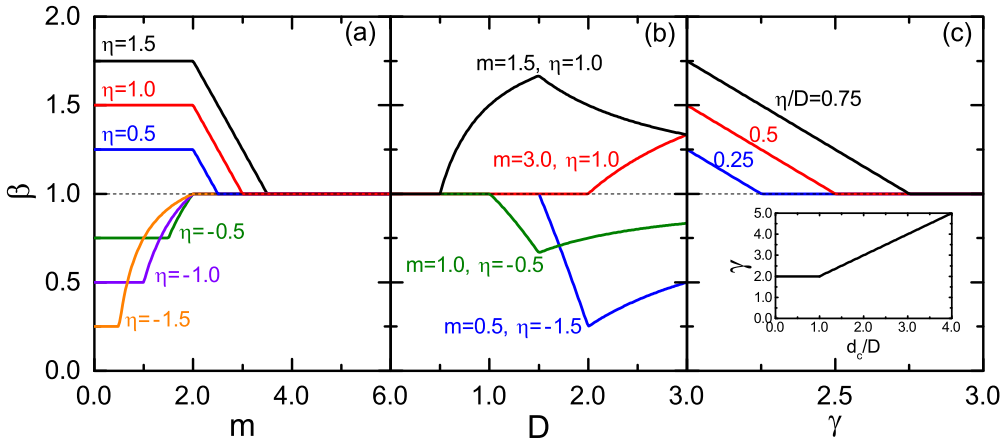

FIG. 3. (Color online) Profiles of the urban scaling exponent β as a function of m, D, and γ. (a) Plot of β versus m for various values of η. The exponents α and D are fixed as $\alpha = 2.0$ and $D = 2.0$. (b) Plot of β versus D for various combinations of m and η. The exponent α is fixed at $\alpha = 2.0$. (c) Plot of β versus γ for various positive values of η/D. The dashed line at $\beta = 1$ in each panel separates the superlinear scaling region from the sublinear one. The inset of (c) shows the d_c/D dependence of the exponent γ given by Eq. (6).

the urban indicator scales linearly with the population size if m is large enough, as pointed out in the previous section. The fact that $\beta \neq 1$ at $m = 0$ shows that the geographical constraint *in the network formation* is not a necessary for nonlinear urban scaling, which does not mean, however, that networks are not required to be embedded in the Euclidean space to obtain nonlinear scaling of $Y(N)$. The exponent β changes with the fractal dimension D as shown in Fig. 3(b). In contrast to the m dependence, β depends nonmonotonically on D. Although only results for $\eta < d_c$ are shown here, β for $\eta > d_c$ monotonically decreases with D and diverges at $D = 0$.

Since the exponent γ characterizing the scale-free property of the network depends on d_c and D as presented by Eq. (6), it seems interesting to elucidate how the urban scaling exponent β varies with γ. The model parameter d_c giving $\gamma = 2$ for a fixed D, however, is not uniquely determined if $D \geqslant d_c$ [see the inset of Fig. 3(c)]. Thus, β for sublinear scaling that requires $D \geqslant d_c$ cannot be related to γ. On the other hand, there is a one-to-one correspondence between γ and d_c for a fixed D if $D < d_c$ that leads to superlinear or linear scaling. In this case, from Eqs. (31a) and (31b), the exponent β is expressed as

$$\beta = \begin{cases} 3 + \dfrac{\eta}{D} - \gamma & \text{for } 2 < \gamma < 2 + \dfrac{\eta}{D} \\ 1 & \text{for } \gamma \geqslant 2 + \dfrac{\eta}{D}, \end{cases} \tag{32}$$

where η must be positive. The γ dependence of β for $\eta > 0$ and $\gamma > 2$ is illustrated in Fig. 3(c). From this argument, we can conclude that sublinear scaling is realized in a network with $\gamma = 2$ and superlinear scaling appears for $2 < \gamma < 2 + \eta/D$ in our model.

In this section we showed how the scaling exponent β depends on the model parameters. The parameter dependence of β points to important implications for real life scenarios. The m (or α) dependence of β shows that urban scaling depends on how strongly human relationships are restricted by a geographical constraint. If we desire strong superlinear scaling to realize a highly innovative city or strong sublinear scaling for saving infrastructure development costs, for example, our result suggests that will happen in cities with people having heterogeneous personalities (small α) and communicating widely at a distance with other people (small m). Urban scaling is also influenced by the fractal dimension D that characterizes where people live in a city. The population density distribution is usually affected by the topographic factors in or surrounding the city, such as rivers, mountains, or coast lines. The D dependence of β offers some insights into the idea of a suitable place (topographic condition) to build an efficient new city. Our result shows that the maximum β (i.e., the most remarkable superlinear scaling) is attained at $D = d_c$ for $d_c > \eta > 0$ and the minimum β (i.e., the most remarkable sublinear scaling) is presented at $D = d_c - \eta$ for $\eta < 0$. Even in the case that the topographic condition of the city has been already fixed, we can enhance the superlinear or sublinear scaling behavior of the city by controlling the fractal dimension D of the residential area in the city. Regarding the η dependence of β, η must be enlarged as much as possible to obtain remarkable superlinear scaling, while there is a suitable value of η $(= d_c - D)$ for realizing the most remarkable sublinear scaling.

IV. SIMULATION RESULTS

Let us confirm the above analytical results by numerical simulations. First, we calculate the urban indicator Y as a function of N for geographical networks with two different kinds of node sets. One is the uniform node set with an integer dimension and the other is the percolation node set with a noninteger fractal dimension. In a network with the uniform node set, N nodes are uniformly distributed at random in a two-dimensional square space, which leads $D = 2$. The linear size L of the square space is adjusted to keep the node density constant independently of N. On the other hand, in a network with the percolation node set, nodes are placed on sites of the largest cluster of a site-percolation system formed in a two-dimensional square lattice with the periodic boundary conditions at the critical percolation probability $p_c = 0.593$ [59]. The fractal dimension of the largest percolation cluster at criticality is known to be $D = 91/48 = 1.896$ [59]. To keep

the node density constant for various N (the number of sites in the largest cluster), the lattice constant of the square lattice is set to be unity. In both cases (i.e., uniform and percolation node sets), the attractivity x_i is assigned to each node according to the distribution function (3) with $\alpha = 2.0$ and $x_{\min} = 1.0$. The Euclidean distance l_{ij} between nodes i and j is measured under periodic boundary conditions and the threshold value Θ in Eq. (4) is chosen so that the average degree becomes $\langle k \rangle = 10.0$. The urban indicator $Y(N)$ is calculated directly from the definition (19). Figures 4(a) and 4(b) show the N dependence of $Y(N)$ for various combinations of η and m on a double logarithmic scale for networks with uniform and percolation node sets, respectively. The longitudinal axis

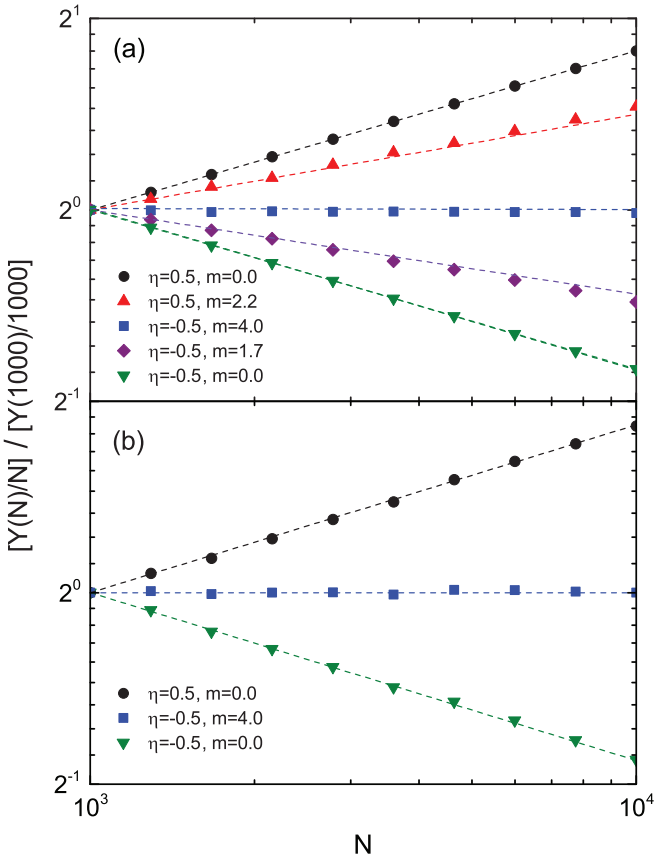

FIG. 4. (Color online) Numerically calculated urban indicators as a function of the population size (number of nodes) N. The exponent α and the parameter $x_{\min}$ characterizing the attractivity distribution given by Eq. (3) are chosen as $\alpha = 2.0$ and $x_{\min} = 1.0$. The longitudinal axis indicates $Y(N)/N$ rescaled by its value at $N = 1000$. (a) Results for geographical networks with uniform node sets ($D = 2.0$). Circles, triangles, squares, diamonds, and inverted triangles are the results for ($\eta = 0.5$, $m = 0$), ($\eta = 0.5$, $m = 2.2$), ($\eta = -0.5, m = 4.0$), ($\eta = -0.5, m = 1.7$), and ($\eta = -0.5, m = 0$), respectively. Dashed lines through symbols from the top to the bottom give the theoretically predicted slopes of $\beta - 1 = 0.25, 0.155$, 0.0, -0.132, and -0.25, respectively. (b) Results for geographical networks with percolation node sets ($D = 1.896$). Circles, squares, and inverted triangles are the results for ($\eta = 0.5, m = 0$), ($\eta = -0.5$, $m = 4.0$), and ($\eta = -0.5, m = 0$), respectively. Dashed lines through symbols from the top to the bottom give the theoretically predicted slopes of $\beta - 1 = 0.264$, 0.0, and -0.264, respectively.

in each panel represents $Y(N)/N$ rescaled by its value at the minimum N ($= 1000$) to improve the legibility of the results. Thus, an increasing, decreasing, or constant straight line indicates superlinear, sublinear, or linear scaling of $Y(N)$, respectively. Each symbol represents the result averaged over 1000 realizations with different node layouts and attractivity assignments. In contrast to the case of the uniform node set, it is difficult for the percolation node set to specify in advance the number of nodes N in the largest cluster and to average over different node configurations under the same network size. In order to perform the averaging procedure in the case of the percolation node set, we gathered 1000 samples of the largest critical clusters in which the numbers of nodes lie within the range of $\pm 1\%$ centered on a desired number of nodes among many clusters. Standard errors are smaller than the size of symbols. Our numerical results clearly show that $Y(N)$ obeys a power law with respect to N and the slopes representing $\beta - 1$ agree, for both node sets, with the theoretical predictions indicated by dashed lines. Triangles ($\eta = 0.5$ and $m = 2.2$) and diamonds ($\eta = -0.5$ and $m = 1.7$) in Fig. 4(a) slightly deviate from the corresponding theoretical lines. These deviations are caused by the finite-size effect as discussed below.

Next, we numerically fit values of β as a function of m and compare the obtained results with the theoretical predictions. Here we concentrate on the case of the uniform node set. In order to examine the m dependence of β, values of the scaling exponent β are estimated by the least-squares fit for numerical data of $Y(N)$ within the range of $10^3 \leqslant N \leqslant 10^4$. Results for $\eta = 0.5$ and -0.5 are presented by closed circles and squares in Fig. 5, respectively. Parameters other than η and m and the computational conditions, such as the boundary conditions and the number of realizations for the sample average, are the

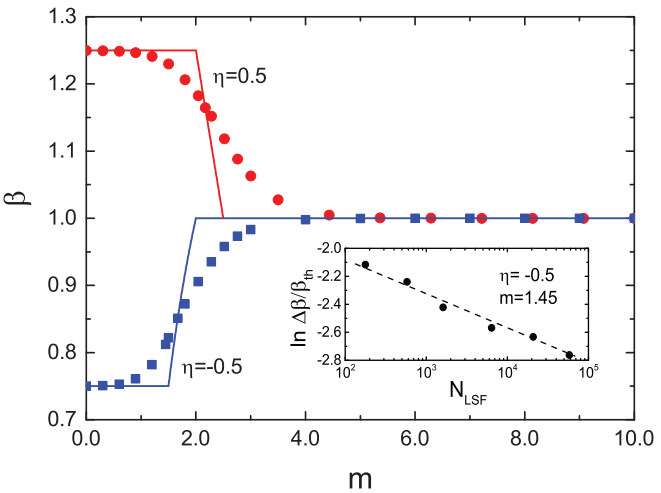

FIG. 5. (Color online) Numerically calculated m dependence of the exponent β. Circles and squares represent the results for $\eta = 0.5$ and -0.5, respectively. All the conditions other than m and η are the same as those in Fig. 4. Solid lines give the theoretical predictions by Eq. (31) for $\eta = 0.5$ and -0.5 ($\alpha = 2.0$ and $D = 2.0$ for both lines). The inset shows the relative error $\Delta\beta/\beta_{\text{th}}$ for $\eta = -0.5$ and $m = 1.45$ as a function of N_{LSF}, around which the least-squares fit is performed within a narrow window of N. The dashed line in the inset is a guide to the eye.

same as those for Fig. 4(a). Standard errors over samples are smaller than the symbol size. Solid lines in Fig. 5 represent the theoretical predictions given by Eq. (31) for $\eta = 0.5$ and -0.5. Numerical results roughly coincide with the theoretical curves. In particular, data for $m > 4$ and $m < 1$ agree quite well with the theoretical curves. However, simulation results near $m = D/(\alpha - 1)$ and $(D + \eta)/(\alpha - 1)$ that give the turnoff points of $\beta(m)$ (i.e., $m = 2.0$ and 2.5 for $\eta = 0.5$ and $m = 2.0$ and 1.5 for $\eta = -0.5$) deviate from the theoretical values. This is due to the finite-size effect. In the analytical calculation of the exponent β, we assume a large enough number of nodes to determine the dominant terms of Eqs. (14), (25), and (28). If two exponents of N in each of these equations becomes close to each other (i.e., approaching the turnoff point), both terms almost equally contribute to $Y(N)$ [or to $\langle k \rangle$ for Eq. (14)] and $Y(N)$ for numerically accessible N does not obey a power law anymore. In order to demonstrate that the deviation $\Delta\beta$ of numerically calculated β from its theoretical value β_{th} near the turnoff point is caused by the finite-size effect, we show the network-size dependence of the relative error $\Delta\beta/\beta_{\text{th}}$ in the inset of Fig. 5. To obtain this inset, we calculated numerically $Y(N)$ for $\eta = -0.5$ and $m = 1.45$ within the range of $10^3 \leqslant N \leqslant 10^5$ and estimated β by the least-squares fit for these data in relatively narrow windows of N around N_{LSF}. The result in the inset displays that the relative error $\Delta\beta/\beta_{\text{th}}$ decreases with increasing N_{LSF}, which suggests $\Delta\beta/\beta_{\text{th}} = 0$ in the large-network-size limit ($N \to \infty$).

V. CONCLUSION

The origin of superlinear and sublinear scaling observed in urban indicators has been analytically argued by modeling the interrelationship of people in a city by a geographical scale-free network. In this network model, nodes close to each other are more likely to be connected than long distant nodes. We assumed that the urban indicator Y of a city is given by the sum of individual node-pair activities $\{y_{ij}\}$ produced by personal, one-to-one human relationships in the city and y_{ij} is proportional to l_{ij}^{η}, where l_{ij} is the Euclidean distance between directly connected nodes i and j. For a positive or negative exponent η, the urban indicator represents a creative productivity or a degree of infrastructure development, respectively. We showed that the urban indicator obeys a power law $Y(N) \propto N^{\beta}$ for a large enough population size N. The exponent β is larger than or equal to one if $\eta > 0$, while $0 < \beta \leqslant 1$ for $\eta < 0$, which implies that $Y(N)$ corresponding to a creative productivity scales superlinearly or linearly with respect to the population size N and it scales sublinearly or linearly if $Y(N)$ is a quantity related to infrastructure. This result coincides with the scaling behavior of real-world urban indicators. It has been also found that $Y(N)$ is proportional to N if networks are formed under a strong geographical constraint. These results have been confirmed by numerical simulations.

In our argument, nodes are assumed to be placed on a D-dimensional Euclidean space and the geographical distance plays a crucial role in understanding urban scaling. To interpret $Y(N)$ under a negative η as a degree of infrastructure development, the nodes must be arranged in a physical (geographic) space. This condition, however, can be relaxed for superlinear scaling. We can derive the same result for superlinear scaling of $Y(N)$ even in the case that nodes are placed on a more general metric space in which Eqs. (4) and (18) with the abstract distance l_{ij} are a reasonable condition for the network formation and a plausible relation for the individual activity, respectively. For example, in a sociometric space with defined social distances between nodes we can consider that nodes socially close to each other are more likely to be connected and a socially more distant node pair yields a higher productivity. Therefore, the scaling exponent β is also presented by Eq. (31) for $\eta > 0$, if Eqs. (4) and (18) with the social distance l_{ij} do actually hold. We should note that in such a case D must be the (fractal) dimension of the sociometric space.

Actual urban indicators of individual cities deviate from the average values of $Y(N)$ expected from their population sizes. Statistical properties of such fluctuations of $Y(N)$ have been extensively studied in recent works [28–31]. In the present work, no analysis of fluctuations is performed, nor is any methodology to compute them given. Furthermore, the model requires empirical validation with the actual set of parameters used in order to be considered a useful tool to assess the fluctuations of actual observed values on real cities; hence, as it is presented here, although it represents an initial step towards the understanding of these fluctuations, it cannot be considered an actual benchmark.

ACKNOWLEDGMENTS

This work was supported by a Grant-in-Aid for Scientific Research (No. 25390113) from Japan Society for the Promotion of Science, by the project SEETechnology "Co-operation of SEE science parks for the promotion of transnational market uptake of R&D results and technologies by SMEs" cofunded by South East Europe Transnational Cooperation Programme, and by the operation entitled "Centre for Open Innovation and Research of the University of Maribor." The last operation is cofunded by the European Regional Development Fund and conducted within the framework of the Operational Programme for Strengthening Regional Development Potentials, development priority 1: "Competitiveness of companies and research excellence," priority axis 1.1: "Encouraging competitive potential of enterprises and research excellence." Numerical calculations in this work were performed in part on the facilities of the Supercomputer Center, Institute for Solid State Physics, University of Tokyo.

[1] S. Nordbeck, Geogr. Ann. B **53**, 54 (1971).
[2] M. Kleiber, Hilgardia **6**, 315 (1932).
[3] G. B. West, J. H. Brown, and B. J. Enquist, Science **276**, 122 (1997).
[4] S. L. Lindstedt and W. A. Calder, Q. Rev. Biol. **56**, 1 (1981).
[5] H. Samaniego and M. E. Moses, J. Transp. Land Use **1**, 21 (2008).
[6] C. P. Lo, Ann. Assoc. Am. Geogr. **92**, 225 (2002).

[7] M. Batty, R. Carvalho, A. Hudson-Smith, R. Milton, D. Smith, and P. Steadman, Eur. Phys. J. B **63**, 303 (2008).

[8] D. Pumain, Santa Fe Institute Report No. 2004-02-002 (2004).

[9] M. Batty and P. A. Longley, *Fractal Cities: A Geometry of Form and Function* (Academic, London, 1994).

[10] M. Batty, Science **319**, 769 (2008).

[11] M. Batty, Cities **29**, S9 (2012).

[12] Y. Chen and S. Jiang, Chaos Solitons Fractals **39**, 49 (2009).

[13] Y. Chen, Discrete Dyn. Nat. Soc. **2010**, 194715 (2010).

[14] L. M. A. Bettencourt, J. Lobo, D. Helbing, C. Kühnert, and G. B. West, Proc. Natl. Acad. Sci. USA **104**, 7301 (2007).

[15] L. M. A. Bettencourt, J. Lobo, and D. Strumsky, Res. Policy **36**, 107 (2007).

[16] L. M. A. Bettencourt, J. Lobo, and G. B. West, Eur. Phys. J. B **63**, 285 (2008).

[17] L. M. A. Bettencourt, J. Lobo, and H. Youn, Santa Fe Institute Report No. 2013-01-004 (2013).

[18] L. Wu and J. Zhang, Phys. Rev. E **84**, 026113 (2011).

[19] S. Arbesman and N. A. Christakis, Physica A **390**, 2155 (2011).

[20] M. A. Changizi and M. Destefano, Complexity **15**, 11 (2010).

[21] E. L. Glaeser and B. Sacerdote, J. Polit. Econ. **107**, S225 (1999).

[22] C. Kühnert, D. Helbing, and G. West, Physica A **363**, 96 (2006).

[23] M. Fragkias, J. Lobo, D. Strumsky, and K. C. Seto, PLoS ONE **8**, e64727 (2013).

[24] L. N. Lamsal, R. V. Martin, D. D. Parrish, and N. A. Krotkov, Environ. Sci. Technol. **47**, 7855 (2013).

[25] S. Arbesman, J. M. Kleinberg, and S. H. Strogatz, Phys. Rev. E **79**, 016115 (2009).

[26] L. M. A. Bettencourt, Science **340**, 1438 (2013).

[27] W. Pan, G. Ghoshal, C. Krumme, M. Cebrian, and A. Pentland, Nat. Commun. **4**, 1961 (2013).

[28] J. Lobo, L. M. A. Bettencourt, D. Strumsky, and G. B. West, PLoS ONE **8**, e58407 (2013).

[29] L. M. A. Bettencourt, J. Lobo, D. Strumsky, and G. B. West, PLoS ONE **5**, e13541 (2010).

[30] A. Gomez-Lievano, H. Youn, and L. M. A. Bettencourt, PLoS ONE **7**, e40393 (2012).

[31] L. G. A. Alves, H. V. Ribeiro, E. K. Lenzi, and R. S. Mendes, PLoS ONE **8**, e69580 (2013).

[32] R. J. Smeed, J. Inst. Highw. Eng. **10**, 5 (1963).

[33] P. Frankhauser, Population (English Edition) **10**, 205 (1998).

[34] S.-H. Yook, H. Jeong, and A.-L. Barabási, Proc. Natl. Acad. Sci. USA **99**, 13382 (2002).

[35] Y. Chen, Int. J. Urban Sustain. Dev. **1**, 89 (2010).

[36] K. Yakubo and D. Korošak, Phys. Rev. E **83**, 066111 (2011).

[37] Strictly speaking, a compact curved space such as a spherical or torus surface does not satisfy Eq. (2) when l approaches the linear size of the compact space. Although a fractal node arrangement in an infinite space is isotropic from any point and satisfies Eq. (2), N must be finite for arguing urban scaling. An artificial realization is to regard a city as a region within a distance L from a node of focus in infinite nodes distributed in an infinite space.

[38] M. E. J. Newman, Contemp. Phys. **46**, 323 (2005).

[39] L. A. Adamic and B. A. Huberman, Q. J. Electron. Commer. **1**, 5 (2000).

[40] M. Cha, H. Haddadi, F. Benevenuto, and K. P. Gummadi, *Proceedings of the Fourth International Conference on Weblogs and Social Media* (AAAI, Menlo Park, 2010), p. 10.

[41] D. R. Bild, Y. Liu, R. P. Dick, Z. M. Mao, and D. S. Wallach, arXiv:1402.2671.

[42] E. O'Boyle, Jr. and H. Aguinis, Pers. Psychol. **65**, 79 (2012).

[43] V. Pareto, *Cours d'Économie Politique* (Rouge, Lausanne, 1897).

[44] B. C. Arnold, *Pareto and Generalized Pareto Distributions* (Springer, New York, 2008).

[45] M. A. Serrano, D. Krioukov, and M. Boguñá, Phys. Rev. Lett. **100**, 078701 (2008).

[46] M. Boguñá, D. Krioukov, and K. C. Claffy, Nat. Phys. **5**, 74 (2009).

[47] M. Boguñá and D. Krioukov, Phys. Rev. Lett. **102**, 058701 (2009).

[48] D. Krioukov, F. Papadopoulos, A. Vahdat, and M. Boguñá, Phys. Rev. E **80**, 035101 (2009).

[49] D. Krioukov, F. Papadopoulos, M. Kitsak, A. Vahdat, and M. Boguñá, Phys. Rev. E **82**, 036106 (2010).

[50] N. Masuda, H. Miwa, and N. Konno, Phys. Rev. E **71**, 036108 (2005).

[51] S. Morita, Phys. Rev. E **73**, 035104 (2006).

[52] A.-L. Barabasi and R. Albert, Science **286**, 509 (1999).

[53] A.-L. Barabasi, Network Science, http://barabasilab.neu.edu/networksciencebook/

[54] For many real-world networks, the exponent γ is usually bounded by $2 < \gamma < 3$, while γ of our model can be larger than 3. There surely exists a scale-free network with $\gamma > 3$, such as the mobile phone network ($\gamma = 8.4$) reported by J.-P. Onnela, J. Saramäki, J. Hyvönen, G. Szabó, D. Lazer, K. Kaski, J. Kertész, and A.-L. Barabási, Proc. Natl. Acad. Sci. USA **104**, 7332 (2007).

[55] D. Liben-Nowell, J. Nowak, R. Kumar, P. Raghavan, and A. Tomkins, Proc. Natl. Acad. Sci. USA **102**, 11623 (2005).

[56] L. Adamic and E. Adar, Soc. Netw. **27**, 187 (2005).

[57] J. Goldenberg and M. Levy, arXiv:0906.3202.

[58] R. Lambiotte, V. D. Blondel, C. de Kerchove, E. Huens, C. Prieur, Z. Smoreda, and P. Van Dooren, Physica A **387**, 5317 (2008).

[59] D. Stauffer and A. Aharony, *Introduction to Percolation Theory* (Taylor & Francis, London, 1994).