

Multiple scaling law in networks with dynamic spatial constraint

Jiang-Hai Qian^{a,b,*}, Qi-Jia Liao^a, Jing Xu^a, Han-Yun Chang^a, Ding-Ding Han^{c,d,*}, Yu-Gang Ma^{e,*}

^a College of Mathematics and Physics, Shanghai University of Electric Power, Shanghai 200090, China

^b Engineering Research Center of Software/Hardware Co-design Technology and Application, Ministry of Education (East China Normal University), Shanghai 200062, China

^c School of Information Science and Technology, Fudan University, Shanghai 200433, China

^d Shanghai Artificial Intelligence Laboratory, Shanghai 200232, China

^e Institute of Modern Physics, Fudan University, Shanghai 200433, China

ARTICLE INFO

Keywords:

Scaling law
Double power law
Spatial networks
Spatial constraint
Airline network

ABSTRACT

We study a network model where the spatial constraint evolves with the network growth and demonstrate that the dynamic constraint can generally cause a non-stationary multiple scaling law in degree distribution. In specific, we show that the different power-law segments emerge successively during the network evolution, from a single power law at the beginning to a double one and to more scaling laws due to the particular model setting. Our findings indicate a plausible geometric origin of the double power-law distribution and predict the non-stationary nature accompanied by its emergence, both of which are supported by our empirical study on the Chinese airline network.

1. Introduction

Scaling law of degree distribution is known as a typical characteristic of many complex networks [1,2]. Numerous studies have devoted to the mechanisms of its emergence or distortion [3–14]. One of the factors that affect its performance is the spatial constraint, which usually originates from maintenance cost, transmission delay or energy dissipation and manifests itself as the preference of spatially short links in network formation [15]. As a consequence, a node is restricted to connect with others only nearby, which hampers the emergence of large degree and break the scaling law. Amounts of empirical studies on road networks [16,17], railway networks [18,19], airline networks [20–27], Internet [28], communication networks [29], power grids [30,31] and social interactions [32] have evidenced that the degree distribution displays a crossover from a skewed power law to a peaked Gaussian-like distribution with the increasing strength of the spatial constraint.

To model the spatial effect on the network structure, one can assume that each node has an influence zone. Nodes that locate within a particular zone can sense the influence of the corresponding node and the edges between them are more likely to occur. A typical example is the geometric graph [33–38], while others including the gravity models and the spatial growth models [39–44], where the spatial constraint is characterized as some distance decay function instead, can still form similar zone effect, as was analyzed in Ref. [45]. By tuning the strength of the spatial constraint, these models can produce almost all the observed degree distributions with the only exception—the double

power law [46], which is common in airline networks [20–27] and whose origin remains an open question.

The spatial constraint is considered to be related to a system's essential attribute. Thus it is usually assumed to be static in the sense that the zone area or the distance decay function is constant in time. Whereas, recent empirical studies on International trade network provide definitive evidence that the decay function in gravity law is time-dependent [47–49]. In other cases such as wireless sensor network, the communication radius decreases with energy dissipation [50]. Indeed, if the constraint is characterized as some kind of influence, the static assumption is inconsistent with the principle of locality. As action at a distance is prohibited, the influence of a new-born node cannot reach infinity instantaneously but instead might expand like field spreads in space. Motivated by these facts and considerations, we study the topology and the evolution resulting from the dynamic spatial constraint and focus on the scaling law of the degree distribution as a first step towards the issue. As will be shown, the dynamic constraint brings that final missing piece, i.e. the double power law, and goes further. It indicates the existence of the multiple scaling law and predicts the non-stationary nature in its evolution. All of them are validated by the empirical data.

The rest of the paper is organized as follows. In Section 2, we present the description of our model, including two different dynamic constraint rules in particular. In Section 3, we study the degree distribution both analytically and by simulation, showing their multiple scaling-law behavior and non-stationary property. In Section 4 we test

* Corresponding authors.

E-mail addresses: qianjianghai@shiep.edu.cn (J.-H. Qian), ddhan@fudan.edu.cn (D.-D. Han), mayugang@fudan.edu.cn (Y.-G. Ma).

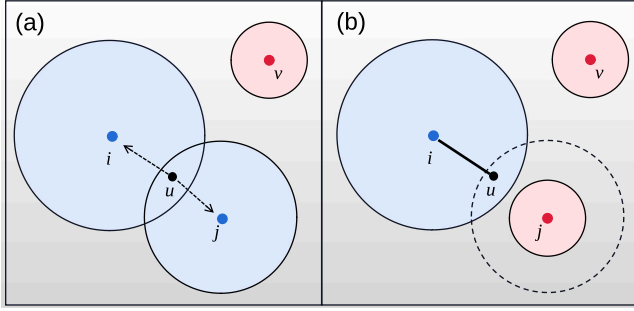


Fig. 1. An example of the key procedures in our model. (a). Creation of new links, corresponding to the evolution rule ii and iii. A node u (either a new-coming one or a chosen existing one) is located within the influence zone of node i and node j but outside of the zone of node v , which represents that we have $G_u = \{i, j\}$. It then chooses either i or j according to the preferential attachment to create a new link. (b). Change of zone area, corresponding to the evolution rule iv. A node changes its zone area according to a particular dynamic rule of the influence zone. In this case, node j is chosen. Its influence zone changes from the dashed circle to a small solid one, and correspondingly it is excluded from G_u .

our theory empirically and discuss the possible origin of the dynamic constraint observed in the airline network. In Section 5 we draw the conclusion.

2. Model description

Our model is constructed on a 10×10 square plane with periodic boundary. The network grows by continuously adding nodes and edges. For each node, it possesses an influence zone whose shape is assumed to be circular and whose area can evolve with time. The key principle of our model is that a node can only connect with others whose influence zones cover its position. The specific rules of our model are presented below and a concise example is illustrated in Fig. 1 to clarify some key procedures.

1. Evolution rule of the network

(i). At each time step, a new node u with the initial zone area S_0 is added to the network and is placed at a randomly chosen position.

(ii). The new-coming node u selects a node i within the set G_u according to the probability $\frac{k_i}{\sum_{i \in G_u} k_i}$ to create a new link, where k_i is the degree of node i that belongs to G_u and G_u is composed of all the existing nodes whose influence zones cover the position of u .

(iii). Select an existing node u' according to the probability $\frac{k_{u'}}{\sum_u k_u}$, and u' creates another new link by the same rule as that in (ii).

(iv). Refresh the zone area according to a particular dynamic rule of the influence zone (specified down below) and then repeat the whole process.

2. Dynamic rule of the influence zone

We introduce three different dynamic rules for the influence zone.

(a). Static rule: the zone area of each node keeps S_0 unchanged during the network evolution.

(b). Change with degree: Given an ascending degree sequence $\{k_1, k_2, \dots, k_N\}$ and a zone area sequence $\{S_1, S_2, \dots, S_N\}$, for each node, once its degree exceeds k_n , its zone area changes to S_n immediately, where $n \in \{1, 2, \dots, N\}$.

(c). Change by random selection: Given a zone area sequence $\{S_0, S_1, S_2, \dots, S_N\}$, at each time step, a randomly chosen node changes its zone area S_n to S_{n+1} , where $n \in \{0, 1, 2, \dots, N-1\}$. If the node has already reached the final area S_N , it remains unchanged.

It should be noted that the degree sequence here is not equivalent to that in the configuration model [3]. Its elements represent a series of predefined thresholds rather than the expected degree of nodes. Accordingly, the N here does not correspond to the number of nodes. To avoid possible confusion of the notations, we clarify that when we

use the subscript n , N or some specific numbers, S and k refer to the elements in the above-mentioned sequences, while for the subscript i and u , they mean in general the zone area and the degree of a particular node.

In the early stage of the evolution, there are only a few nodes distributed sparsely on the plane. It is very likely that the set G_u of some nodes is empty. If these nodes are chosen, no new links will be created and we just continue the procedure. This will result in some isolated nodes and the network is disconnected at this stage. With the increase of the node density, the number of the empty sets vanishes. Those isolated nodes get chances to form links and the network becomes connected with few exceptional nodes which will not affect the whole structure. In this paper, we are only interested in the state of the network with spatially high node density, so that the initial effects can be ignored.

3. Scaling law of degree distribution

Let A be the area of the whole plane. The probability of a node i to be connected by another node u is

$$p_i = \frac{S_i}{A} \frac{k_i}{\sum_{i \in G_u} k_i}. \quad (1)$$

From the view of the mean-field theory, each node i in the network contributes its degree weighted by $\frac{S_i}{A}$ to $\sum_{i \in G_u} k_i$, giving $\frac{1}{A} \sum_i S_i k_i$ in all. Then Eq. (1) is rewritten as

$$p_i = \frac{S_i k_i}{\sum_i S_i k_i}. \quad (2)$$

At each time step, the degree increment of node i comes from three contributions, including the contribution of the new link from a new-coming node, which gives increment p_i , the contribution of the preferential selection of an existing node, which gives increment $\frac{k_i}{\sum_i k_i}$ and the contribution of the new link from an existing node, which gives increment p_i . The increase rate of degree k_i is the sum of the three parts, resulting in the dynamical equation

$$\frac{dk_i}{dt} = \frac{2S_i k_i}{\sum_i S_i k_i} + \frac{k_i}{\sum_i k_i}. \quad (3)$$

3.1. Static zone area

If all the zone area remains S_0 , Eq. (2) degenerates into the classical preferential attachment

$$p_i = \frac{k_i}{\sum_i k_i}, \quad (4)$$

and Eq. (3) is reduced to

$$\frac{dk_i}{dt} = \frac{3k_i}{\sum_i k_i}, \quad (5)$$

where $\sum_i k_i \sim 4t$. Using the initial condition $k_i(t) = 1$, we get $k_i(t) = (\frac{t}{t_0})^{\frac{3}{4}}$. The standard treatment in Ref. [5] helps to derive the complementary cumulative degree distribution. It follows a single scaling law

$$P(k) = k^{-\frac{4}{3}}. \quad (6)$$

The simulations are in agreement with Eq. (6), as shown in Fig. 2.

3.2. Zone area changing with degree

We start with the simplest case that the degree sequence and the zone area sequence are $\{k_1\}$ and $\{S_1\}$ respectively. Before any node reaches k_1 , there exists only the zone area S_0 and the dynamical equation is exactly the same as Eq. (5). A single power law is therefore to be expected at the early stage of the evolution.

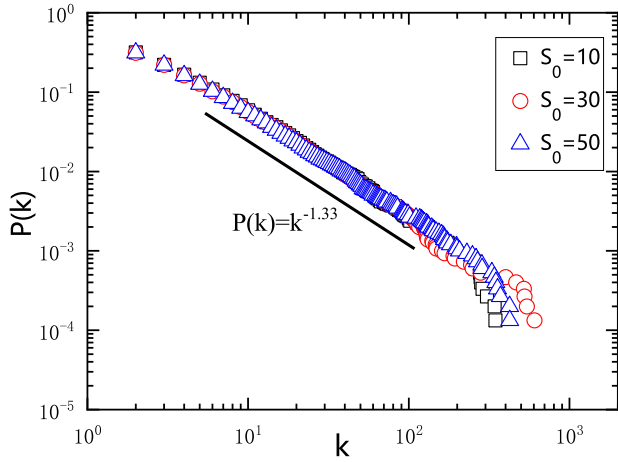


Fig. 2. The complementary cumulative degree distribution under the static rule. The simulations are executed for three different S_0 and reach 1.5×10^4 nodes in network size. The distribution displays a single power law with exponent $-\frac{4}{3}$ (represented by the solid line) regardless of S_0 , which is consistent with Eq. (6).

Once there is a node exceeding k_1 , Eq. (3) is split into two parts,

$$\begin{aligned} \frac{dk_i}{dt} &= \frac{2S_0k_i}{S_0 \sum_{i| \leq k_1} k_i + S_1 \sum_{i| > k_1} k_i} + \frac{k_i}{\sum_i k_i}, & k_i \leq k_1, \\ \frac{dk_i}{dt} &= \frac{2S_1k_i}{S_0 \sum_{i| \leq k_1} k_i + S_1 \sum_{i| > k_1} k_i} + \frac{k_i}{\sum_i k_i}, & k_i > k_1, \end{aligned} \quad (7)$$

where the notations $\sum_{i| \leq k_1}$ and $\sum_{i| > k_1}$ represent the sum over all nodes i with $k_i \leq k_1$ and $k_i > k_1$ respectively. Let $f_0(t) = \sum_{i| \leq k_1} k_i$ and $f_1(t) = \sum_{i| > k_1} k_i$. The addition of a new link between a new-coming node and an existing node always increases $f_0(t)$ by 1. In addition, it generates unit increment for $f_0(t)$ if it is attached to a node with $k < k_1$, but decreases $f_0(t)$ by k_1 if it is attached to a node with $k = k_1$ since this node's zone area changes to S_1 . The derivation of the two contributions requires the probability of this link to be attached to a node with k_1 , which is given by $\frac{S_0k_1p_1t}{S_0f_0(t)+S_1f_1(t)}$, where p_1 is the ratio of the nodes with degree k_1 . The contributions of a new link created between two existing nodes can be analyzed in the similar way. The increase rate $\frac{df_0(t)}{dt}$ is the sum of all these contributions and $\frac{df_1(t)}{dt}$ can be obtained analogously, resulting in

$$\begin{aligned} \frac{df_0(t)}{dt} &= 1 + 2 \frac{S_0f_0(t) - S_0p_1k_1^2t}{S_0f_0(t) + S_1f_1(t)} + \frac{f_0(t) - p_1k_1^2t}{4t}, \\ \frac{df_1(t)}{dt} &= 2 \frac{S_1f_1(t) + S_0p_1k_1^2t}{S_0f_0(t) + S_1f_1(t)} + \frac{f_1(t) + p_1k_1^2t}{4t}, \end{aligned} \quad (8)$$

where we use $k_1 + 1 \simeq k_1$.

Eq. (8) has a set of solutions in the form of $f_0(t) = C_0t$ and $f_1(t) = C_1t$, with C_0 and C_1 following the constraint $C_0 + C_1 = 4$ and the self-consistent relation

$$\begin{aligned} C_0 &= 1 + 2 \frac{S_0C_0 - S_0p_1k_1^2}{C} + \frac{C_0 - p_1k_1^2}{4}, \\ C_1 &= 2 \frac{S_1C_1 + S_0p_1k_1^2}{C} + \frac{C_1 + p_1k_1^2}{4}, \end{aligned} \quad (9)$$

where $C = S_0C_0 + S_1C_1$. We can solve $k_i(t)$ by substituting the solutions back into Eq. (7), deriving

$$\begin{aligned} k_i(t) &= \left(\frac{t}{t_1}\right)^{\frac{1}{\gamma_0}}, & k_i \leq k_1, \\ k_i(t) &= k_1 \left(\frac{t}{t_1}\right)^{\frac{1}{\gamma_1}}, & k_i > k_1, \end{aligned} \quad (10)$$

where $\gamma_0 = \frac{4C}{8S_0+C}$, $\gamma_1 = \frac{4C}{8S_1+C}$ and $t_1 = ik_1^{\gamma_0}$. Here t_1 is the time when the zone area of node i changes to S_1 , which is obtained from the first

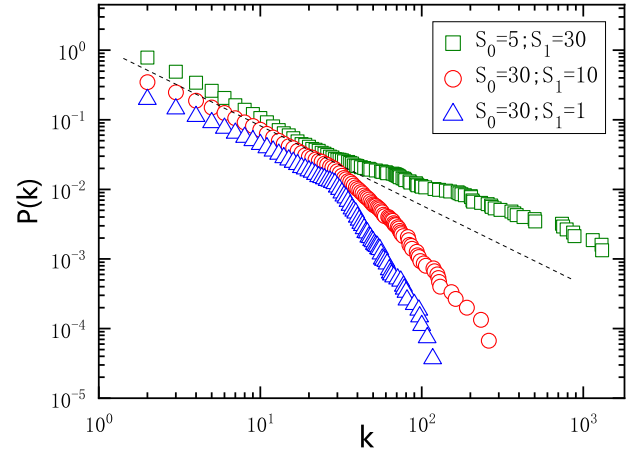


Fig. 3. The complementary cumulative degree distribution under the dynamic rule(b) with the sequence length $N = 1$. The simulations are executed for three different combinations of S_0 and S_1 . The degree sequence is set to $\{k_i\} = \{30\}$ and the network size reaches 1.5×10^4 nodes. The dashed line in between, drawn according to the exponent of the single power law (Eq. (6)), is a guide for eyes. The data points of different colors are shifted upwards or downwards for better visualization. When $S_0 > S_1$, $P(k)$ is a normal double power law with $\gamma_0 < \gamma_1$. When $S_0 < S_1$, it is an inverted one with $\gamma_0 > \gamma_1$. The result is in agreement with Eq. (11). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

equation in Eq. (10) by letting $k_i(t) = k_1$. Then the complementary cumulative degree distribution can be derived and we find

$$P(k) = \begin{cases} k^{-\gamma_0}, & k \leq k_1 \\ \lambda k^{-\gamma_1}, & k > k_1 \end{cases} \quad (11)$$

where $\lambda = k_1^{\frac{2\gamma_0\gamma_1}{C} \frac{S_0-S_1}{C}}$. From Eq. (11) we can calculate the ratio p_1 by $p_1 = -\frac{\partial P(k)}{\partial k} \Big|_{k=k_1}$. Substituting it into Eq. (9) and combining the constraint $C_0 + C_1 = 4$, we arrive at the self-consistent relation of C as given by Eq. (12) and complete the whole derivation,

$$4C(S_0 - S_1)k_1^{\frac{8S_0-3C}{8S_0+C}} = (4S_0 - C)(3C - 8S_1). \quad (12)$$

Eq. (11) shows that the second scaling-law segment of $P(k)$ is born at this stage of the evolution and the degree distribution becomes a double power law with a transition degree at k_1 , which is defined as the point where two different scaling-law segments meet [46]. If $S_0 > S_1$, we have $\gamma_0 < \gamma_1$. The degree distribution is a normal double power law which has been observed in a variety of complex systems [20–27,51–57]. If $S_0 < S_1$, $P(k)$ is an inverted double power law with $\gamma_0 > \gamma_1$, and for $S_0 = S_1$ a single scaling law is recovered. Figs. 3 and 4 present the simulation results of the degree distribution and the measurement of the power-law exponents respectively. They show good agreement with the analytical result. For details of our data analysis of the multiple scaling law, refer to the description in Appendix A.

We now turn to some more complicated cases. If the two sequences are $\{k_1, k_2\}$ and $\{S_1, S_2\}$, before any node reaches k_2 , the network evolution is the same with the simplest case as discussed above. Once there is a node exceeding k_2 , a new dynamical equation is to be added to Eq. (7) to involve the evolution of the nodes with zone area S_2 . One can use the same analytical technique to show $P(k)$ will evolve to be a triple power law with two transition degrees at k_1 and k_2 . The similar results can be further generalized to the situation of $\{k_1, k_2, \dots, k_N\}$ and $\{S_1, S_2, \dots, S_N\}$ with arbitrary N . Generally, the different scaling-law segments emerge successively, causing a non-stationary $P(k)$ during the evolution, and finally forms $N+1$ power-law parts with the $(n+1)$ th exponent $\gamma_n = \frac{4C}{8S_n+C}$ and N transition points at k_1, k_2, \dots, k_N . Technically, a clear presentation of the multiple scaling law depends on the parameters $\Delta S (= S_n - S_{n-1})$ and $\Delta \ln (= \ln k_n -$

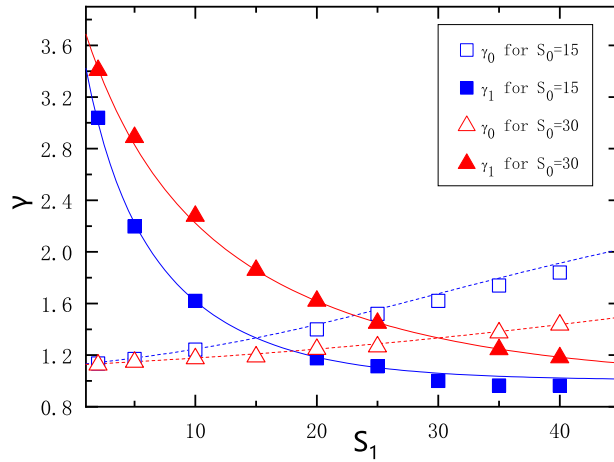


Fig. 4. The power-law exponents of the two tails of the degree distribution, measured under the dynamic rule(b) with the sequence length $N = 1$ and plotted versus S_1 for two different S_0 . The degree sequence is set to $\{k_i\} = \{30\}$ and the network size reaches 1.5×10^4 nodes. Each data point is an average over 20 independent runs. The dashed lines and the solid lines are the theoretical predictions of the exponents of the upper tail and the lower tail respectively. They present nice agreement with the simulation results.

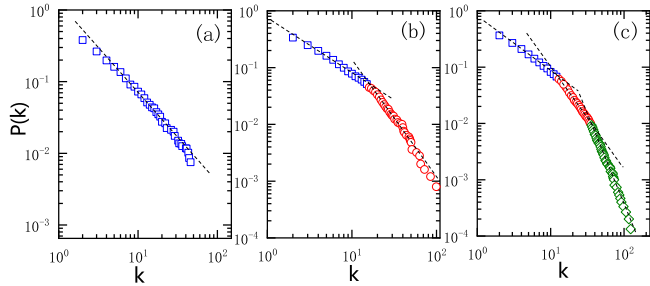


Fig. 5. The evolution of the complementary cumulative degree distribution under the dynamic rule(b) with the sequence length $N = 2$. The degree sequence and the zone area sequence are set to $\{k_1, k_2\} = \{15, 35\}$ and $\{S_0, S_1, S_2\} = \{50, 20, 5\}$ respectively. The three figures are the snapshots of a run at $t = 800$ (a), 2500 (b), 15000 (c). The different power-law segments are marked by different colors for better discrimination. The dashed lines, drawn according to the theoretical exponents, are served as guides for eyes. The degree distribution displays a crossover from a single power law to a double one and finally to a triple one, presenting a non-stationary evolution as predicted by the theoretical analysis. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

$\ln k_{n-1}$), which control the difference between two consecutive power-law exponents and the spacing between adjacent transition degree on log-log plot, respectively. Large ΔS and $\Delta \kappa$ are therefore essential for the recognizable power-law segments, otherwise the segments get either close in slope or short in length, blurring the multiple scaling-law performance. In Fig. 5 we show the evolution of a triple power-law degree distribution. The non-stationary dynamics and the scaling-law performance are consistent with the theoretical analysis.

3.3. Zone area changing by random selection

Consider a node i with zone area $S_i = S_n$, where $n \in \{0, 1, 2, \dots, N\}$, Eq. (3) in this case can be rewritten as

$$\frac{dk_i}{dt} = \frac{2S_n k_i}{\sum_n \sum_{i|S_i=S_n} S_n k_i} + \frac{k_i}{\sum_i k_i}, \quad (13)$$

where the notation $\sum_{i|S_i=S_n}$ represents the sum over all nodes i with $S_i = S_n$. Let $f_n(t) = \sum_{i|S_i=S_n} k_i$, the increase rate of $f_n(t)$ follows

$$\frac{df_n(t)}{dt} = \delta_{0n} + \frac{2S_n f_n(t)}{\sum_n S_n f_n(t)} + \frac{f_n(t)}{\sum_i k_i} + \frac{(1 - \delta_{0n})f_{n-1}(t) - (1 - \delta_{Nn})f_n(t)}{t}, \quad (14)$$

where δ_{mn} is the Kronecker Delta. The last term on the right hand side of Eq. (14) is the contribution from the change of the zone area. Since a node changes its zone area from S_n to S_{n+1} with probability $\frac{1}{t}$, it decreases $f_n(t)$ by $\frac{k_i}{t}$ and increases $f_{n+1}(t)$ by the same quantity. The total relevant contributions to $f_n(t)$ is the sum of all the increments from the nodes with S_{n-1} and all the decrements transferred to $f_{n+1}(t)$ from the nodes with S_n , giving the last term in Eq. (14).

Eq. (14) has a set of solutions in the form of $f_n(t) = C_n t$, with C_n following the self-consistent relation

$$C_n = \delta_{0n} + \frac{2S_n C_n}{C} + \frac{C_n}{4} + (1 - \delta_{0n})C_{n-1} - (1 - \delta_{Nn})C_n, \quad (15)$$

where $C = \sum_n S_n C_n$. This recursive formula allows us to derive C_n as an expression with C . Combining this with the constraint $\sum_n C_n = 4$, we can determine C by

$$\frac{(4C)^{N+1}}{3C - 8S_N} \prod_{j=0}^{N-1} \frac{1}{7C - 8S_j} + \sum_{n=0}^{N-1} (4C)^{n+1} \prod_{j=0}^n \frac{1}{7C - 8S_j} = 4, \quad (16)$$

and complete the derivation of $f_n(t)$. Substituting the solution of $f_n(t)$ back into Eq. (13), we can solve the dynamical equation and find

$$k_i(t) = k_n \left(\frac{t}{t_n}\right)^{\frac{1}{\gamma_n}}, \quad k_n \leq k_i < k_{n+1} \quad (17)$$

where $\gamma_n = \frac{4C}{8S_n + C}$, t_n is the time when S_i changes to S_n and k_n is the corresponding degree. The scale of t_n follows $\int_0^{t_n} \frac{1}{i+\tau} d\tau = n$, which gives

$$t_n \sim i e^n, \quad 0 \leq n \leq N \quad (18)$$

Taking $t = t_{n+1}$ in Eq. (17), we get the recursive formula $\ln k_{n+1} - \ln k_n = \frac{1}{\gamma_n}$ and derive the transition degree

$$k_n \sim \exp\left(\sum_{i=0}^{n-1} \frac{1}{\gamma_i}\right), \quad (19)$$

where we stipulate the summation takes zero if $n = 0$ and infinity if $n = N + 1$, which is consistent with the initial degree $k_0 = 1$ and the boundless k_{N+1} respectively. Finally we arrive at

$$P(k) = e^{-n} \left(\frac{k}{k_n}\right)^{-\gamma_n}, \quad k_n \leq k < k_{n+1} \quad (20)$$

Eq. (20) demonstrates that the final steady complementary cumulative degree distribution is a multiple scaling law with $N + 1$ power-law segments. Note that large ΔS is still essential for a clear presentation of $P(k)$. However, noticing $\Delta \kappa \sim S_n$, the scaling-law segment for small k could be practically invisible. In Fig. 6 we plot the simulation

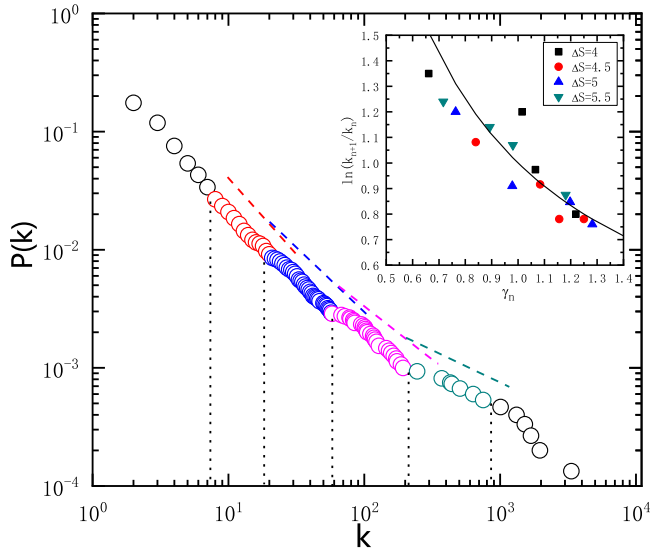


Fig. 6. The complementary cumulative degree distribution under the dynamic rule(c). The zone area sequence $\{S_n\}_{n=0}^N$ is set to $S_n = S_0 + n\Delta S$ with $S_0 = 0.8$, $\Delta S = 4$ and $N = 10$. The network size reaches 1.5×10^4 nodes. The different power-law segments in the body of the distribution are marked by four different colors, while those unrecognizable ones at both tails, caused by small S_n or finite size effect, are marked by black cycles. The dashed lines and the vertical dotted lines, drawn according to the theoretical predictions of the scaling-law exponents and the transition degrees respectively, are served as guides for eyes. Inset: The relation between the transition degree and the power-law exponent. We measure the k_n and the γ_n for those identifiable scaling-law segments and plot the logarithmic difference $\ln k_{n+1} - \ln k_n$ versus γ_n . To produce more data points, we run the simulation for four different ΔS while keeping the other parameters unchanged. The solid curve represents the recursive formula of our analytical result, showing nice agreement with the measurement results. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

result of the degree distribution and the measurement result of the transition degree versus the power-law exponent (inset). In Fig. 7, we plot the relation between the zone area and the corresponding power-law exponent. All of them are consistent with the analytical results.

The scaling-law segments do not emerge simultaneously. Indeed, before S_n appears, the network evolves in the same way with the case $\{S_0, S_1, S_2, \dots, S_{n-1}\}$ and only n power-law segments are formed. Therefore the $(n + 1)$ th segment will not develop until the first time that a node's zone area reaches S_n , presenting a non-stationary picture of the successive emergence of the different scaling laws. This non-stationary property can be characterized by the long-term dynamics of C before it reaches its final value derived from Eq. (16). When a new zone area S_n appears in the network, C starts a new evolution and gradually stabilizes at the final value corresponding to the case $\{S_0, S_1, S_2, \dots, S_n\}$. As a consequence, the increment ΔC between two successive middle steady states can be evaluated by the difference between the two corresponding final value, say C and C' , which by definition is written as $C = \sum_{j=0}^n S_j C_j$ and $C' = \sum_{j=0}^{n+1} S_j C'_j$. Since the emergence of a new zone area needs exponentially long time, as predicted by Eq. (18), we expect $C(t)$ to be a slowly changing function. To show this, we consider an example of constant ΔS . The increment ΔC is given by

$$\begin{aligned} \Delta C &= C' - C \\ &= S_{n+1} C'_{n+1} + \sum_{j=0}^n S_j \Delta C_j, \end{aligned} \tag{21}$$

where $\Delta C_j = C'_j - C_j$ and $S_j = S_0 + j\Delta S$. The constraint $\sum_{j=0}^n C_j = \sum_{j=0}^{n+1} C'_j = 4$ implies $C_j \sim \frac{4}{n+1}$ and $C'_j \sim \frac{4}{n+2}$, leading to $\Delta C_j \sim$

$-\frac{4}{(n+1)(n+2)}$. Substituting these results into Eq. (21), we get $\Delta C \sim 2\Delta S$ for arbitrary n . Combining with Eq. (18), we find

$$C(t) \sim 2\Delta S \ln t. \tag{22}$$

In Figs. 8 and 9, we plot the evolution of the degree distribution and the evolution of the parameter C . Both of them are consistent with the theoretical analysis.

4. Empirical study

We test our theory in the Chinese airline network (CAN), a typical system with both the double power law structure and the spatial relevance. We focus on two key signatures indicated by our model, the non-stationary evolution of the degree distribution and the inverse change of the zone area with the power-law exponent. If the double power law does arise from the dynamic spatial constraint, a crossover of the degree distribution from a single scaling law to a double one, and a change of the link length before and after the transition degree are supposed to be observed.

Our network data come from the *China Traffic Yearbook (1989–2014)*. They provide the information about the airports(nodes) and the flight routes (links) for 25 years, which allow us to trace the evolution of CAN from the very beginning. The link length can be obtained from the Great Circle Mapper (<http://www.gcmap.com/>), where the distance between any pair of the cities is available. For the information about the number of nodes and the mean degree of CAN, refer to Fig. 13 in Appendix B.

Most studies on CAN focused on the years after 2002 [20–25], when the double power law had already formed. To seek the dynamics of the successive emergence of the scaling laws, we begin our observation from 1989, which is much earlier than all the previous studies. As is shown in Fig. 10(a) and (b), we find the complementary cumulative degree distribution of CAN was actually a single power law in the first place, while it displayed a double power law in the later years (Fig. 10(c) and (d)). The careful examination shows that it began to turn at the year 1999 and reached the double power law in 2002, which presents a positive result for the non-stationary property.

To capture the connection between the zone area and the power-law exponent, we calculate for each node the average length of the links created before and after it reaches the transition degree k_c (≈ 22), denoted as R_1 and R_2 respectively. The ratio R_1/R_2 , as the assessment of the relative change of the zone area, is expected to be inverse to the relative change of the power-law exponent γ_1/γ_2 according to our model, resulting in $R_1/R_2 > 1$ for CAN that displays $\gamma_1 < \gamma_2$. It is noted that due to the scale-free nature, the majority of the nodes never reach k_c and cannot be included in the analysis. After filtering out these nodes, the results of the remaining 29 (13.3% of the total) nodes are plotted in Fig. 11. We find that 22 nodes out of a total of 29 present $R_1/R_2 > 1$. When removing those data points closest to $R_1/R_2 = 1$ (shaded region in Fig. 11), which might be questioned about their reliability, one can verify that the number of the nodes with $R_1/R_2 > 1$ goes up to almost four times higher than those with $R_1/R_2 < 1$. If the zone area is static, we expect the data points to fluctuate around the dashed line in Fig. 11. Instead, the empirical result shows a remarkable trend of shrinking area that is in agreement with the prediction of our theory, which illustrates the reality of the dynamic spatial constraint and indicates its connection with the scaling laws.

An interesting question is why the zone area in the real system tends to shrink instead of the contrast. Indeed, the inverted double power law seems barely reported. The similar phenomenon, which is known as the globalization puzzle, has also been found in the international trade system [47–49]. For transportation network, a possible clue to the puzzle relates to the budget and the cost. We assume the links in the spatial network can be categorized into two types according to their length scale: the long-range links associated with the cost that is proportional to the spatial distance and the local costless links

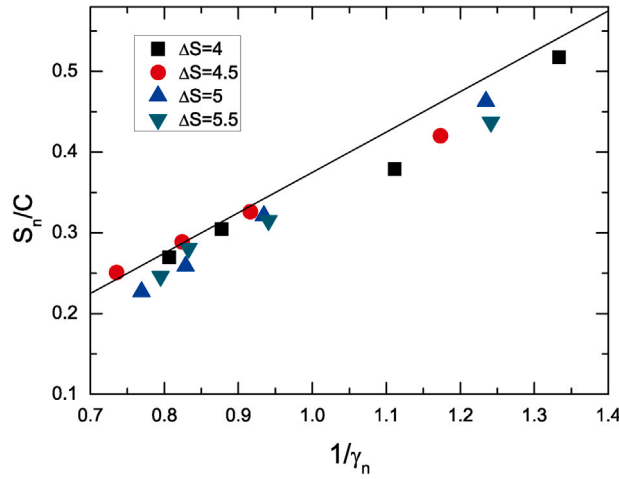


Fig. 7. The relation between the zone area and the corresponding power-law exponent, measured under the dynamic rule(c). The zone area sequence $\{S_n\}_{n=0}^N$ is set to $S_n = S_0 + n\Delta S$ with $S_0 = 0.8$, $N = 10$ and ΔS of four different values as shown in the figure. The network size reaches 1.5×10^4 nodes. For each ΔS , we measure the γ_n and determine the corresponding S_n for each identifiable scaling-law segment, and estimate the parameter C by $\sum_i S_i k_i / t$ at $t = 15000$. Each data point is an average over 20 independent runs. The measurement results are consistent with the theoretical prediction as represented by the solid line.

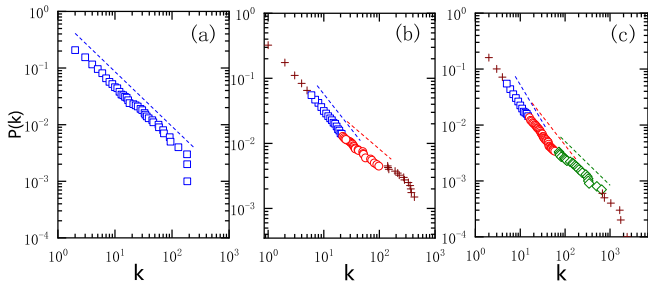


Fig. 8. The evolution of the complementary cumulative degree distribution under the dynamic rule(c). The zone area sequence $\{S_n\}_{n=0}^N$ is set to $S_n = S_0 + n\Delta S$ with $S_0 = 0.8$, $\Delta S = 4$ and $N = 10$. The three figures are the snapshots of a run at $t = 1500$ (a), 4000 (b), 10000 (c). The different power-law segments are marked by different colors for better discrimination. The dashed lines are served as guides for eyes. The unrecognizable parts at both tails are marked by brown +. The degree distribution displays a crossover from a single power law to a multiple one, showing a non-stationary evolution as predicted by the theoretical analysis. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

between geographically neighboring nodes. The same assumption has already been applied in the generalization of the Kleinberg model [58]. Consider for each node there is a budget $B \sim k_c \sqrt{A}$, at the beginning the sufficient budget allows its connection to any place, resulting in a long-range link with high probability. The corresponding length scales as $O(\sqrt{A})$ and B is reduced by the quantity. The procedure continues until $B \sim \sqrt{A}$. Then just one more connection can exhaust the budget and only local connections are available in the subsequent evolution. This causes just the effect of the discontinues reduction of the zone area as described by our model. Generally, this procedure could occur in any type of the transportation network. However if the spatial constraint is too strong, the preferential attachment will be suppressed. Only those with relatively weak constraint, such as airline networks, are closed to the situation of our model, which helps to understand why double power-law degree distribution is so common in airline networks.

There are some additional evidences that support our theory. In Ref. [59], the author observed a triple power-law degree distribution in US airlines, justifying the existence of the scaling laws with more than two segments. We also find another plausible case in the Brazilian airline network, as shown in Appendix C. Also in airline network Gastner and Newman reported a bimodal distribution in link length [60], which is consistent with the dual scales of the link length as discussed above.

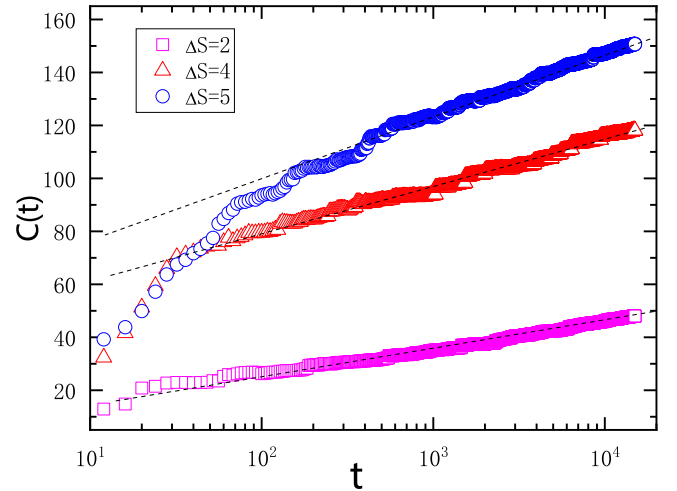


Fig. 9. The evolution of the parameter C under the dynamic rule(c). The zone area sequence $\{S_n\}_{n=0}^N$ is set to $S_n = S_0 + n\Delta S$ with $S_0 = 0.5$, $N = 15$, and $\Delta S = 2, 4, 5$. The network size reaches 1.5×10^4 nodes. The large N ensures C will not reach its final state under the present network size so that we can study its non-stationary behavior in the whole evolution time. For each ΔS , we determine C at time t by $\sum_i S_i k_i / t$. The resulting $C(t)$ displays an asymptotic logarithmic increase with the measured slope 4.5, 8.2, 10.4 for $\Delta S = 2, 4, 5$ respectively, as indicated by the fitting lines(dashed). The results are in agreement with Eq. (22).

To the best of our knowledge, there have been no other effective models that can account for all these observations.

5. Conclusion

We demonstrate that a dynamic spatial constraint can generate a multiple scaling law in network structure, irrespective of whether it changes with some topological property or by random selection. In particular, the multiple scaling law evolves in a non-stationary manner that its power-law segments emerge successively in the sequence of their corresponding zone area. The empirical study on the Chinese airline system verifies the non-stationary property and obtains a direct observation of the change of the average link length before and after the transition degree, which is consistent with the theoretical prediction. Other predictions including the existence of the triple power law and the dual scales of the link length have also been witnessed. These results

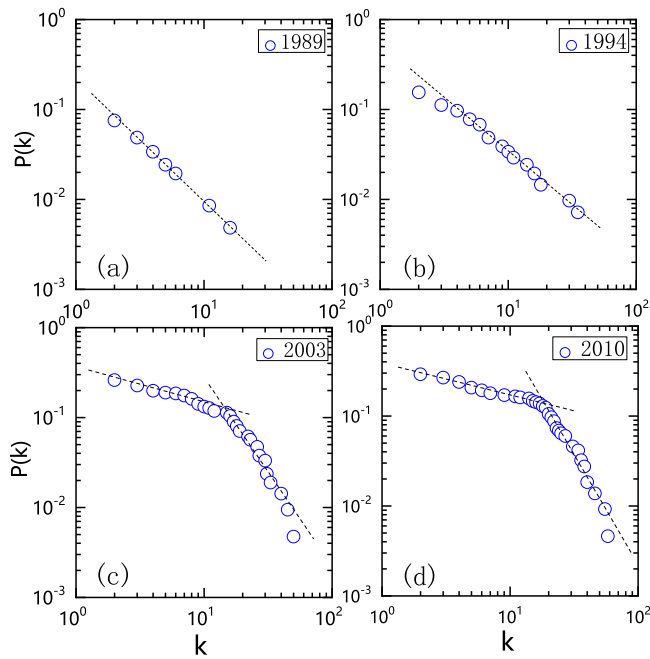


Fig. 10. The complementary cumulative degree distribution of CAN for four different years. The distribution in 1989(a) and 1994(b) was a single power law while it displayed a double power law in year 2003(c) and 2010(d). The fitting lines (dashed) are served as guides for eyes. The empirical observation indicates that the two power-law segments do not emerge simultaneously but in a consecutive manner, presenting a non-stationary property in agreement with our theory.

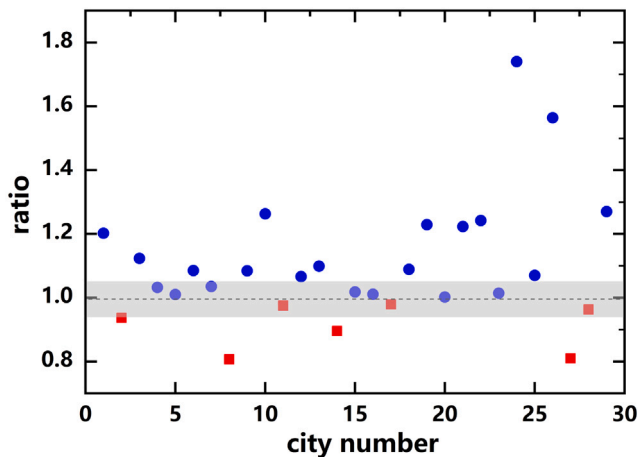


Fig. 11. The ratio R_1/R_2 of each node(city) in CAN, served as an assessment of the relative change of the zone area. R_1 is obtained by averaging the link length of a node at its last year before it reaches k_c and R_2 is the average over those links created after the node reaches k_c . The dashed horizontal line represents the ratio equaling to one and the x-axis labeled as city number represents the identity assigned to each city. The blue data points above the dashed line correspond to the shrinking zone area while the red ones represent the opposite case. The shaded ribbon is a visualization of the region closest to $R_1/R_2 = 1$. The figure shows a remarkable trend of shrinking area when node degree exceeds k_c . (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

validate the dependency of the multiple scaling law on the dynamic spatial constraint and convince us of the high reliability of our theory in explaining the origin of the double power-law distribution.

There are some other models yielding double power law. Reed proposed one based on the geometric Brownian motion [54], which was further expounded by Mitzenmacher [46] and spawned several

variants [57,61–63]. Dorogovtsev and Mendes proposed another one by assuming a linear acceleration in network growth [55]. However neither of them can yield the non-stationary dynamics or the scaling laws with more than two segments. By contrast, our model presents a wide range of consistency with the airline networks, which indicates its high validity and particular significance to these systems.

An interesting insight brought by the dynamic spatial constraint is the order of the creation of different types of links. Despite the fact that the spatial network prefer short links, our empirical study indicates a counter-intuitive result that long links have a higher priority to access the airline network. We have shown its impact on the network structure, while it could be even more significant for the network function such as navigability or robustness [32,64–69], which is closely related to the link length distribution. Understanding how systems evolves to reach these functions requires the knowledge of how and why the spatial constraint changes. We have tried such an explanation for airline networks from the perspective of budget and cost, but are still in dark for many others.

CRedit authorship contribution statement

Jiang-Hai Qian: Conceptualization, Methodology, Formal analysis, Writing – original draft, Writing – review & editing, Funding acquisition. **Qi-Jia Liao:** Formal analysis, Visualization, Investigation, Software. **Jing Xu:** Software, Formal analysis, Investigation. **Han-Yun Chang:** Visualization, Software. **Ding-Ding Han:** Funding acquisition. **Yu-Gang Ma:** Funding acquisition, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant No. 11875133, 12147101, and 11075057), the Science and Technology Commission of Shanghai Municipality (Grant No. 22JC1402500), and the Open Research Fund of Engineering Research Center of Software/Hardware Co-design Technology and Application, Ministry of Education (East China Normal University) (Grant no. OP202102).

Appendix A. Methodology of the data analysis

1. Identification of the Transition Degree

In most cases, the transition degree in our data can be visually identified from the doubly logarithmic plot of the distribution. Nonetheless, to avoid subjectivity in the data analysis, we develop a principled approach that makes the results more reliable.

The basic idea of our method is simple. Consider a power-law cumulative degree distribution on a log–log plot, the goodness of fit for any of its part by a linear regression function is expected to be almost identically good because they all fall on a single line. But if the part we fit contains a transition degree, the fit will be significantly worse. Therefore, the goodness of fit can tell us about the information of the scaling property and help to locate the transition degree and equivalently to identify the corresponding power-law segment. The detailed procedure includes two steps.

Step 1: Find out the possible transition degree. We start from the first or first several data points of the cumulative distribution and renew

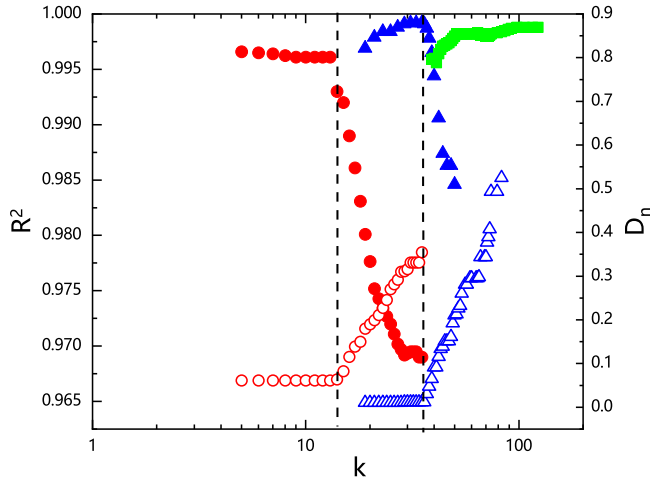


Fig. 12. The illustration of the identification of the transition degree in Fig. 5(c). The solid points are the result of step 1, corresponding to the left axis that represents R^2 . The green points show no significant decrease, indicating no further transition degree. The hollow ones are the result of step 2, corresponding to the right axis that represents D_n . The dashed vertical lines represent the locations of the two identified transition degree. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the data continually by adding in the remainders one at a time in sequence. Meanwhile we fit the renewed data each time by performing a least-squares linear regression on the doubly logarithmic plot and monitor the change of the corresponding goodness of fit, quantified by the coefficient of determination R^2 . The point at which the R^2 starts to continuously decrease is our estimation of the first transition degree, say k_1^* . Then we remove the data points whose degree is less than k_1^* and do all these iteratively to find out $k_2^*, k_3^*, k_4^*, \dots$, until there is no significant decrease in the R^2 , which indicates no further transition degree.

To ensure the correctness of the results, we introduce another measure of the goodness of fit, the distance D_n , to test the candidate transition degree k_n^* found in step 1, which is defined as

$$D_n(k) = \max_{k_{n-1}^* < x < k} |\ln S_n(x) - \ln P_n(x)|, \quad (A.1)$$

where $S_n(x)$ is the cumulative distribution function of the data with value at least k_{n-1}^* (we stipulate $k_0^* = 1$), and $P_n(x)$ is the cumulative distribution function for the power-law model that best fits the data in the region $k_{n-1}^* < x < k_n^*$. Step 2 is described as follow.

Step 2: Test the identified transition degrees. We calculate $D_n(k)$ for each $k > k_{n-1}^*$ and monitor its change with k . If the critical point where $D_n(k)$ starts to continuously increase is consistent with k_n^* , k_n^* passes the test and is considered as an effective measurement of the n th transition degree k_n . Otherwise we abandon it and let $k_n^* = k_{n+1}^*$ for all the subsequent n . We start from $n = 1$ and do this iteratively until we complete all the tests.

As an example, Fig. 12 illustrates the identification result of Fig. 5(c). The transition degree is identified as $\{k_1^*, k_2^*\} = \{15, 35\}$, in agreement with the predefined $\{k_1, k_2\}$. When we identify the transition degree, we equivalently identify the power-law segment. Each segment is simply established by the corresponding transition degree at its two ends, as we have illustrated in Figs. 5, 6 and 8. One may argue that the method is based on the prior assumption that each segment is power law, which needs to be verified in the first place. However we can do the hypothesis test for these identified segments(as we will do next) and if they pass the test, the whole method is self-consistent.

2. Hypothesis Test

After having identified the transition degree, we set about checking whether the distribution is a multiple power law. We can do the

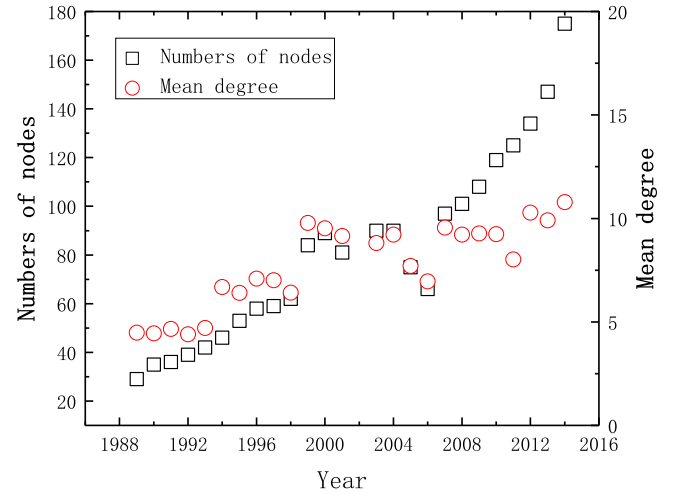


Fig. 13. The evolution of the number of nodes and the mean degree of CAN. The size of CAN grows from 29 cities in 1989 up to 175 cities in 2014. The mean degree increases from 4.9 and has been stabilizing at about 9.5 since 1999.

hypothesis test for each identified segment and each two adjacent segments as a whole respectively by applying the method in Ref. [2]. If all the segments pass the test while all the two adjacent ones fail, it can be concluded that each single segment follows power law but neither of the any adjacent two has the same scaling property, which suggests a multiple scaling law in terms of the present segment identification.

We choose the commonly used threshold of p value, $p_c = 0.05$, to rule out the hypothesis. If the p value is significantly higher than the threshold (say $p > 0.1$ as suggested by Ref. [2]), the power-law distribution is acceptable. If p is lower than p_c , it is ruled out. We do the test for both the simulation data and the empirical data of CAN, and find that all the single segments of the distributions pass the test (namely $p > 0.1$). In contrast, p values of the adjacent ones as a whole are significantly small. Most of them drop directly to zero and the others also fall below the threshold. The test results confirm the multiple scaling-law property and can also be served as a double check for the validity of the identification of the power-law segments.

3. Exponent Estimation

Based on the above results, we can fit each segment with power-law model confidently and estimate the exponents by the method of maximum likelihood estimation [2], which gives the results illustrated in Fig. 4, inset of Figs. 6 and 7 in this paper.

Appendix B. Evolution of the number of nodes and the mean degree of CAN

See Fig. 13.

Appendix C. Degree distribution of Brazilian airline network

The data come from the Brazilian National Agency of Civil Aviation (<http://www.anac.gov.br>) and are extracted from the electronic files of the Air Transportation Annual Report which contains specific information about flights and statistics about the airline companies operating in Brazil. The complementary cumulative degree distribution of the network for year 2018 is plotted in Fig. 14. It displays a triple power-law-like distribution.

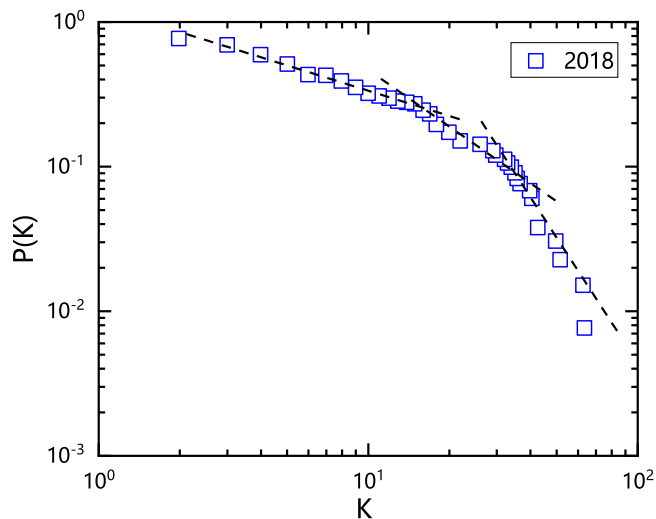


Fig. 14. The complementary cumulative degree distribution of Brazilian airline network for year 2018. It displays a triple power-law-like distribution. The fitting lines (dashed) are served as guides for eyes.

References

- [1] Newman MEJ. *Contemp Phys* 2005;46:323.
- [2] Clauset A, Shalizi CR, Newman MEJ. *SIAM Rev* 2009;51:661.
- [3] Newman MEJ. *SIAM Rev* 2003;45:167.
- [4] Boccaletti S, Latora V, Moreno Y, Chavez M, Hwang DU. *Phys Rep* 2006;424:175.
- [5] Barabási AL, Albert R. *Science* 1999;286:509.
- [6] Krapivsky PL, Redner S, Leyvraz F. *Phys Rev Lett* 2000;85:4629.
- [7] Bianconi G, Barabási AL. *Europhys Lett* 2001;54:436.
- [8] Blasius B, Tönjes R. *Phys Rev Lett* 2009;103:218701.
- [9] Goh KI, Kahng B, Kim D. *Phys Rev Lett* 2002;88:108701.
- [10] Caldarelli G, Capocci A, De Los Rios P, Munoz MA. *Phys Rev Lett* 2002;89:258702.
- [11] Fujii K, Berengut JC. *Phys Rev Lett* 2021;126:102502.
- [12] Amaral LAN, Scala A, Barthélemy M, Stanley HE. *Proc Natl Acad Sci USA* 2000;97:11149.
- [13] Mossa S, Barthélemy M, Stanley HE, Amaral LAN. *Phys Rev Lett* 2002;88:138701.
- [14] Barthélemy M. *Europhys Lett* 2003;63:915.
- [15] Barthélemy M. *Phys Rep* 2013;499:1.
- [16] Barthélemy M, Flammini A. *Phys Rev Lett* 2008;100:138702.
- [17] Masucci AP, Smith D, Crooks A, Batty M. *Eur Phys J B* 2009;71:259.
- [18] Latora V, Marchiori M. *Physica A* 2002;314:109.
- [19] Kurant M, Thiran P. *Phys Rev E* 2006;74:036114.
- [20] Li W, Cai X. *Phys Rev E* 2004;69:046106.
- [21] Liu HK, Zhou T. *Acta Phys Sin* 2007;56:0106.
- [22] Qian JH, Han DD, Ma YG. *Acta Phys Sin* 2011;60:098901.
- [23] Zhang J, Cao XB, Du WB, Cai KQ. *Physica A* 2010;389:3922.
- [24] Paleari S, Redondi R, Malighetti P. *Transp Res* 2010;46:198.
- [25] Li W, Wang QA, Nivanen L, Méhauté AL. *Physica A* 2006;368:262.
- [26] Lin JY, Ban YF. *Physica A* 2014;410:302.
- [27] Chi LP, Wang R, Su H, Xu XP, Zhao JS, Li W, et al. *Chin Phys Lett* 2003;20:1393.
- [28] Yook SH, Jeong H, Barabási AL. *Proc Natl Acad Sci USA* 2002;99:13382.
- [29] Lambiotte R, Blondel VD, Kerchove C, Huens E, Prieur C, Smoreda Z, et al. *Physica A* 2008;387:5317.
- [30] Albert R, Albert I, Nakarado GL. *Phys Rev E* 2004;69:025103.
- [31] Solé RV, Rosas-Casals M, Corominas-Murtra B, Valverde S. *Phys Rev E* 2008;77:026102.
- [32] Liben-Nowell D, Novak J, Kumar R, Raghavan P, Tomkins A. *Proc Natl Acad Sci USA* 2005;102:11623.
- [33] Dall J, Christensen M. *Phys Rev E* 2002;66:016121.
- [34] Herrmann C, Barthélemy M, Provero P. *Phys Rev E* 2003;68:026128.
- [35] González MC, Lind PG, Herrmann HJ. *Phys Rev Lett* 2006;96:088702.
- [36] serrano MA, Krioukov D, Boguná M. *Phys Rev Lett* 2008;100:078701.
- [37] Krioukov D, Papadopoulos F, Kitsak M, Vahdat A, Boguná M. *Phys Rev E* 2010;82:036106.
- [38] Dettmann CP, Georgiou O, Knight G. *Europhys Lett* 2017;118:18003.
- [39] Qian JH, Han DD. *Physica A* 2009;388:4248.
- [40] Louf R, Jensen P, Barthélemy M. *Proc Natl Acad Sci USA* 2013;110:8824.
- [41] Xulvi-Brunet R, Sokolov IM. *Phys Rev E* 2002;66:026118.
- [42] Manna SS, Sen P. *Phys Rev E* 2002;66:066114.
- [43] Jost J, Joy MP. *Phys Rev E* 2002;66:036126.
- [44] Ferretti L, Corteleszi M. *Phys Rev E* 2011;84:016103.
- [45] Xie YB, Zhou T, Bai WJ, Chen GR, Xiao WK, Wang BH. *Phys Rev E* 2007;75:036106.
- [46] Mitzenmacher M. *Internet Math* 2004;1:226.
- [47] Brun JF, Carrère C, Guillaumont P, De Melo J. *World Bank Econ Rev* 2005;19:1.
- [48] Disdier AC, Head K. *Rev Econ Stat* 2008;90:37.
- [49] Karpiarz M, Fronczak P, Fronczak A. *Phys Rev Lett* 2014;113:248701.
- [50] Guan ZH, Ding L, Kong ZM. *Physica A* 2010;389:198.
- [51] Pinto Carla MA, Mendes Lopes A, Tenreiro Machado JA. *Appl Math Model* 2014;38:4019.
- [52] Kwon O, Son WS, Jung WS. *Physica A* 2016;461:85.
- [53] Hajargasht G, Griffiths WE. *Econ Model* 2013;33:593.
- [54] Reed WJ. *Physica A* 2003;319:469.
- [55] Dorogovtsev SN, Mendes JFF. *Proc R Soc London Ser B* 2001;268:2603.
- [56] Toda AA. *J Econ Behav Organ* 2012;84:364.
- [57] Yuan WG, Liu Y. *Physica A* 2015;432:167.
- [58] Li G, Reis SDS, Moreira AA, Havlin S, Stanley HE, Andrade Jr. JS. *Phys Rev Lett* 2010;104:018701.
- [59] Bounova GA. *Topological evolution of networks: case studies in the US airlines and language wikipeidias*. (Phd thesis), Cambridge: Massachusetts Institute of Technology; 2009.
- [60] Gastner MT, Newman MEJ. *Eur Phys J B* 2006;49:247.
- [61] Mitzenmacher M. *Internet Math* 2004;1:305.
- [62] Han DD, Qian JH, Ma YG. *Europhys Lett* 2011;94:28006.
- [63] Toda AA. *Phys Rev E* 2011;83:046122.
- [64] Kleinberg JM. *Nature* 2000;406:845.
- [65] Simsek Ö, Jensen D. *Proc Natl Acad Sci USA* 2008;105:12758.
- [66] Oliveira CLN, Morais PA, Moreira AA, Andrade Jr. JS. *Phys Rev Lett* 2014;112:148701.
- [67] Chen Q, Qian JH, Zhu L, Han DD. *Phys Rev E* 2016;93:032321.
- [68] Li D, Kosmidis K, Bunde A, Havlin S. *Nat Phys* 2011;7:481.
- [69] Li D, Li G, Kosmidis K, Stanley HE, Bunde A, Havlin S. *Europhys Lett* 2011;93:68004.