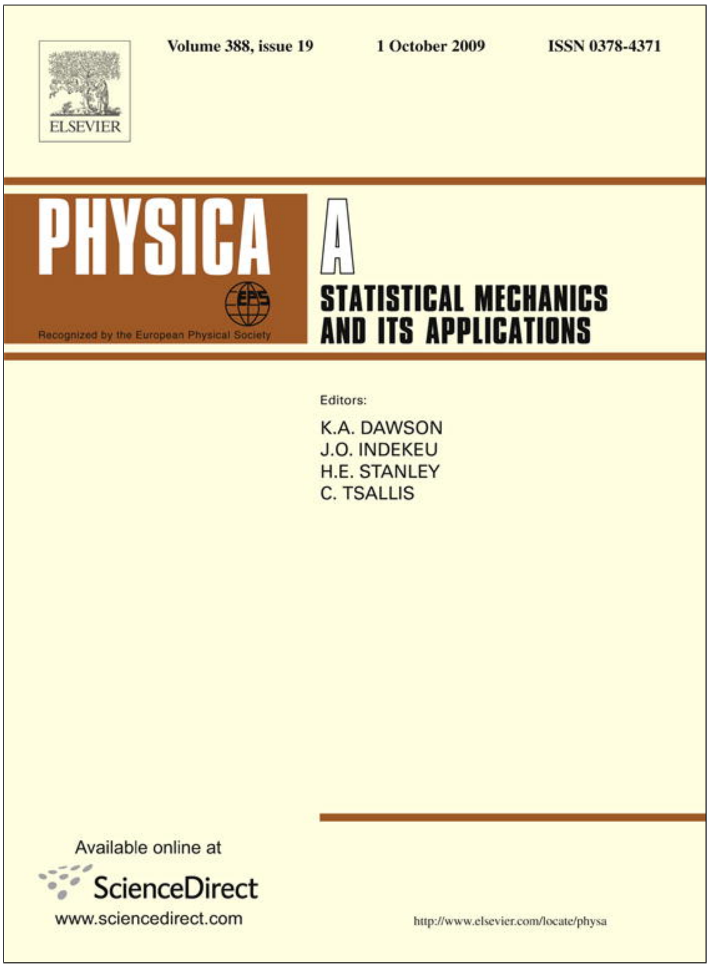
Volume 388, issue 19
1 October 2009
ISSN 0378-4371
ELSEVIER
PHYSICA
A
STATISTICAL MECHANICS
AND ITS APPLICATIONS
Recognized by the European Physical Society
Editors:
K.A. DAWSON
J.O. INDEKEU
H.E. STANLEY
C. TSALLIS
Available online at
ScienceDirect
www.sciencedirect.com
http://www.elsevier.com/locate/physa

Physica A 388 (2009) 4248–4258

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

A spatial weighted network model based on optimal expected traffic

Jiang-Hai Qian, Ding-Ding Han *

School of Information Science and Technology, East China Normal University, Shanghai 200241, China

ARTICLE INFO

Article history:
Received 9 December 2008
Received in revised form 17 April 2009
Available online 6 June 2009

PACS:
89.75.Hc
89.75.Da
89.40.Dd

Keywords:
Spatial weighted network model
Optimal expected traffic
Improved gravity equation
Fitness
Geographical constraints

ABSTRACT

We propose a spatial weighted network model based on the optimal expected traffic. The expected traffic represents the prediction of the flow created by two vertices and is calculated by the improved gravity equation. The model maximizes the total expected traffic of the network. By changing two parameters which control the fitness and the geographical constraints, the model can vary its topology and give rise to a variety of statistical properties observed in the real-world network. Notably, our study shows that a linear and a nonlinear strength–degree correlation can emerge when considering and neglecting the "transport effect", respectively.

1. Introduction

The empirical evidence coming from studies on all kinds of system have shown that most real-world networks have a power-law degree distribution. This nontrivial property has been captured by the BA model, in which the probability of a node to gain links is proportional to its degree, namely preferential attachment (PA) [1]. However, the rate at which nodes increase their connectivity depends not only on the degree but also on their fitness to compete for links [2,3]. Fitness represents the intrinsic weight of nodes and is interpreted as, for example, capital, social skills, activity levels and population of cities. Studies on fitness model have shown that scale-free degree distribution emerges even from a fitness distribution devoid of a power law [4]. In addition to the degree distribution, the empirical studies have demonstrated that the degree correlation and the clustering–degree correlation also display power-law behavior in some real systems [5,6].

The purely topological definition of networks, however, misses important attributes which are frequently encountered in real-world networks. In fact, networks are far from being Boolean structures and are better represented as weighted graphs with different intensities of links. Depending on the type of networks [5,7], the link weight can represent traffic, such as passengers in air flights or information flow in the Internet, as well as the collaboration frequency in a collaboration network or familiarity degree in a social network. The statistical properties of weights indicate non-trivial correlation with topological quantities. Notably, the strength–degree correlation follows a linear and a nonlinear form in a social network and a technical network, respectively [7,8]. The linear strength–degree correlation has been captured by some previous model such as the BBV model [9], while the nonlinear correlation has been catching more attention [10,11].

Another important element of many real networks is their embedding in real space. Examples of technical networks are communication networks [12,13], airline networks [7,14,15], railway networks [16] and the power grid [17]. The most

* Corresponding address: School of Information Science and Technology, East China Normal University, 3663 Zhong Shan Bei Ru, Shanghai 200241, China. Tel.: +86 21 28758701.
E-mail address: ddhan@ee.ecnu.edu.cn (D.-D. Han).

0378-4371/$ – see front matter
doi:10.1016/j.physa.2009.05.047

prominent feature of them is that their vertices have well-defined positions, and links occupy physical distance. This is not the case for other types of network, such as citation or biochemical networks, which exist only in an abstract "network space". Statistical measures of the lengths of connections have shown that the majority of edges in most spatial network are of short range [18]. Since the cost for long-range connections is high, vertices usually tend to link their geographical neighbors. In addition to this spatial dependence, the PA rule is another important mechanism responsible for the spatial network structure. This mechanism causes a small quantity of long-range links. Long-range links were reported both in the Internet and an airline network [18]. An interesting observation in these two networks is that the correlation of the connection number and the geographical length is not a monotonic function but a bimodal distribution [18]. Inspired by these observations, several spatial network models have been proposed. Some of them consider both the geographical constraints and the PA rule while others are based on the idea of effective length which is tunable from spatial length to the topological distance [18–22].

These studies provide some guidelines in modeling spatial networks. But most of them have focused on the cost of constructing the networks while the expected benefit was neglected. In a man-made system, however, the expected benefit could be a more important consideration. For example, when one airline company plans to open a new flight, the traffic of the flight will be estimated first. Only when the estimated traffic is large enough to benefit will the flight be opened. In a railway network, the link cost is much more important than that for an airline, but expected benefit is still a key concern. Motivated by this idea, we introduce the expected traffic and then propose a spatial weighted network model. The model can give rise to a variety of properties observed in a real-world network. Indeed, we find that a linear and a nonlinear strength–degree correlation can emerge when considering and neglecting the "transport effect", respectively.

The paper is organized as follows. In Section 2, we introduce the expected traffic and argue that it can be calculated by the improved gravity equation; Section 3 is devoted to the presentation of the gravity model; in Section 4 we study in detailed the statistical properties of our model, including the pure topology, weight and spatial aspects, and provide some explanation for the nontrivial strength–degree correlation.

2. Expected traffic and gravity equation

The traditional gravity equation is used to measure the force between two objects with certain mass. Furthermore, for years its improved version has been used in other areas such as geography, economics and marketing to estimate the traffic between cities or study trade flows [23,24]. The improved version of the gravity equation is

$$T_{ij} = K\frac{M_i^{\alpha} M_j^{\alpha}}{D_{ij}^{\gamma}} \tag{1}$$

where K is a constant coefficient and the fitnesses of i and j are represented by M_i and M_j. D_{ij} is usually defined as the Euclidean length; however, it can also have other meanings such as time. The two exponents, α and γ, represent the dependence of the system on fitness and spatial constraints. When describing the interaction between cities, it was theoretically proved that $\alpha = 1$, while γ was demonstrated to range from 0.2 to 2.7 by a variety of real data [24–26]. The very recent investigation on the Korea highway system has also confirmed that the traffic in the network follows the gravity law with $\alpha = 1$ and $\gamma = 2$ [27]. On the other hand, the statistical properties observed in real-world networks inspire us to derive Eq. (1) in another plausible way. The weight–degree correlation observed in several networks such as airlines follows $w_{ij} \sim x_{ij}(k_i k_j)^{\theta}$ [7], where x_{ij} is a random number and θ represents the strength of the correlation. k_i, k_j are the degrees of nodes i and j and w_{ij} is the traffic between them. If we assume that the fitness M increases with the degree k according to the relation $k \sim M^{\beta}$, the traffic w_{ij} is given by $w_{ij} \sim (M_i M_j)^{\alpha}$, with $\alpha = \beta\theta$. In addition, the traffic usually decreases with the spatial length since long-distance travel consumes much time and expense. Assuming $w_{ij} \sim \frac{1}{D_{ij}^{\gamma}}$ and combining it with $w_{ij} \sim (M_i M_j)^{\alpha}$, we found $w_{ij} = K\frac{M_i^{\alpha} M_j^{\alpha}}{D_{ij}^{\gamma}}$, which is exactly the form of Eq. (1).

The gravity model is considered to be suitable for describing the traffic in many real system. This result provides us with some ways to predict the flow in the network. Since Eq. (1) is independent of the network topology, it can be calculated before the vertices are connected and its value provides the information of the potential flow between two nodes. For this reason, we call T_{ij} the expected traffic. However, we emphasize that T_{ij} only describes the flow created by vertices i and j, that is, the traffic which originates from i (or j) and ends at the other. T_{ij} does not include traffic created by other pairs of nodes which travels through link (i, j). Thus T_{ij} could be different from the real weight in some networks. The relationship of the expected traffic and the real weight will be mentioned in Section 4.

3. Gravity model

Based on the idea of expected traffic, our model is described as follows. Suppose there are n nodes distributed on a two-dimensional plane and their fitness and geographical information are known, T_{ij} of all the node pairs can be calculated by Eq. (1). Then we connect preferentially those node pairs with larger expected traffic, which indicates that busier links are constructed earlier. This mechanism satisfies the main demand as well as increases the expected benefit of the network. To

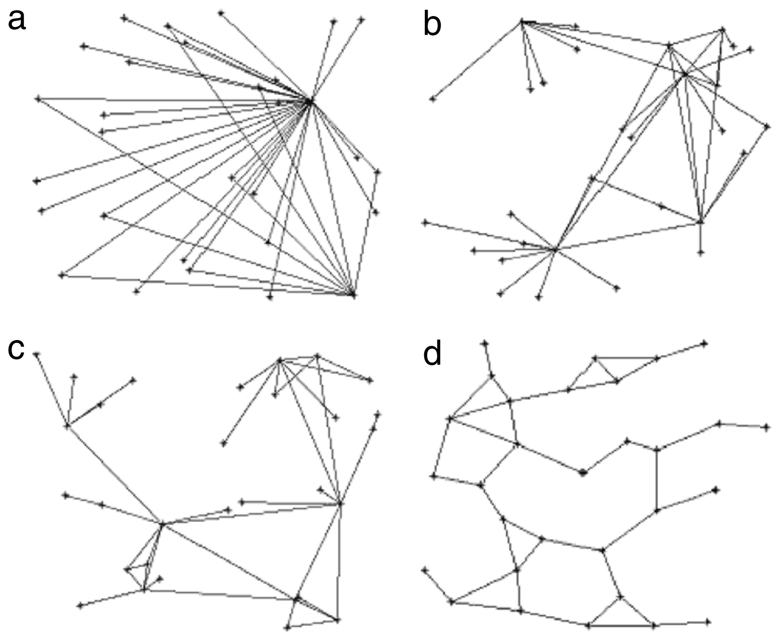

Fig. 1. Four networks with different topology, which is controlled by the parameters α and γ. (a) α takes any value, $\gamma = 0$. The network is dominated by two large hubs when $\gamma = 0$. (b) $\alpha = 1$, $\gamma = 1$. With γ increasing, the hubs become much smaller than in (a). (c) $\alpha = 1$, $\gamma = 2$; (d) $\alpha = 0$, γ takes any value. The topologies are strongly reminiscent of airlines and roads, respectively.

be specific, we arrange all the T_{ij} in descending order, denoted by T, and then start connecting these node pairs in sequence of T until the edge number comes to a value that we preset. This process can be described as an optimal model:

$$\text{Max} \quad W_{exp} = \sum_{i<j} T_{ij}\eta_{ij} = \sum_{i<j} K \frac{M_i^{\alpha} M_j^{\alpha}}{D_{ij}^{\gamma}} \eta_{ij} \tag{2}$$

$$s.t. \quad \sum_{i<j} \eta_{ij} = \epsilon \tag{3}$$

where η_{ij} is the adjacency matrix element of the network, ϵ is the number of edges that we preset, and W_{exp} represents the whole expected traffic of the network. Since this process may cause some isolated nodes, two more restrictions are introduced to ensure that each node is connected:

$$\sum_i \eta_{ij} \geq 1 \quad (j = 1, 2, 3, \ldots, n) \tag{4}$$

$$\sum_j \eta_{ij} \geq 1 \quad (i = 1, 2, 3, \ldots, n). \tag{5}$$

Using our model, a network with 30 nodes was simulated. (The detailed statistical properties of the model will be reported in Section 4. Here we only intuitively show how the topology changes with the parameters α and γ.) Set $\epsilon = 39$ and the coefficient $K = 1$ (actually K makes no difference to the model). By changing the values of α and γ, four networks with different topology were obtained, as is shown in Fig. 1.

As to Fig. 1(a), the value of $\gamma = 0$ makes the network only rely on the fitness and causes a star-like network which is dominated by two hubs. With γ increasing, stronger geographical constraints make the node tend to connect to the closer ones and the network gradually becomes homogeneous. When $\alpha = 1$, $\gamma = 2$ (Fig. 1(c)), the network exhibits some features similar to an airline network. When $\alpha = 0$ (Fig. 1(d)), the network is entirely constrained by the geography and forms a topology which is strongly reminiscent of roads.

4. Properties of the gravity model

Extensive numerical simulations confirm that our gravity model is able to reproduce many properties of spatial weighted networks. In the following simulations, the nodes are distributed randomly in a 10×10 square. Distance D_{ij} is defined as the Euclidean length between vertex i and vertex j. The fitness is chosen from a given distribution $\rho(M)$. Some studies have shown that in some transportation networks such as airlines and highways, the vertices (cities) usually have a power-law distribution fitness (population). Thus we focus on $\rho(M) \sim M^{\phi}$ and select $\phi = -3$ (Cases with other values of ϕ are also discussed. See Appendix A.) The value of parameters α and γ, however, are not arbitrary. Real data and theoretical proof suggest that $\alpha = 1$ and $\gamma \in [0.2, 2.7]$ [24–26]. To simplify our study, we set $\alpha = 1$ and vary γ by taking the values $\{0.5, 1, 2, 3\}$. We

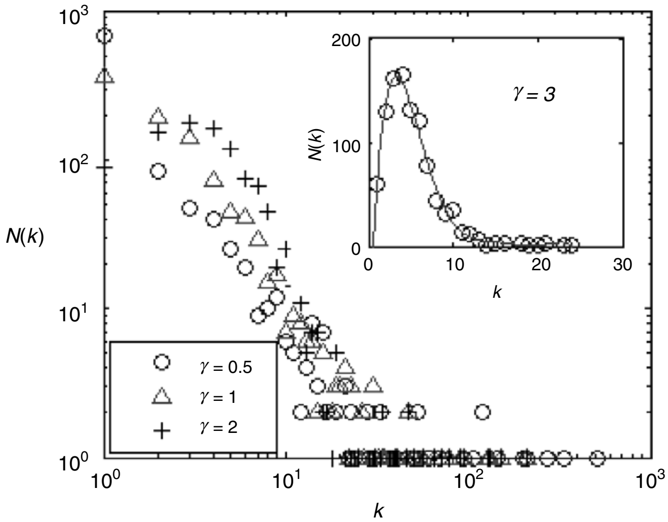

Fig. 2. The degree distribution for the four cases (The case of $\gamma = 3$ is drawn in the insert.). $N(k)$ is the number of vertices with degree k and the degree distribution $p(k)$ is written as $p(k) = \frac{N(k)}{N}$; when $\gamma = 0.5$ and $\gamma = 1$, the degree distribution is a power law with the exponents -1.75 and -2.01, respectively. With the spatial constraints becoming stronger, the distribution deviates from the power-law form, as the case of $\gamma = 2$ shows. When $\gamma = 3$, a deflective Poisson distribution emerges. The simulation is averaged over 20 networks.

fix the graph size $N = 1000$ and average degree $\langle k \rangle = 5$. (Although it is a small graph, numerical simulations indicate that the structure does not change significantly when $N > 500$. To prove that the properties of our model are stabilized when $N = 1000$, the degree distribution and the average nearest-neighbor degree function of larger graph size are illustrated in Appendix B.) We will first study the purely topological properties and then discuss in detail the strength–degree correlation as well as some other weight properties. Finally the spatial features of the present model will be reported.

4.1. Pure topology

In Fig. 2, we plot the degree distribution with four different values of γ. The degree distribution changes evidently as the geographical constraints are gradually strengthened. Both the cases of $\gamma = 0.5$ and $\gamma = 1$ exhibit a power-law distribution: $p(k) \sim k^{-1.75}$ ($\gamma = 0.5$) and $p(k) \sim k^{-2.01}$ ($\gamma = 1$). Note that the two values are close to the real ones observed in airline networks [7]. One may notice that the PA rule, which is believed to be responsible for the power-law degree distribution, is absent in the present model. However, we argue that the mechanism of optimal expected traffic can naturally result in the preferential attachment. According to Eq. (1), vertices with larger fitness are able to extend their interaction further away and create larger expected traffic with more nodes. Since our model favors large expected traffic, these vertices will be preferentially connected, indicating the emergence of the PA rule. When $\gamma = 2$, strong spatial constraints make the network homogeneous and its degree distribution apparently deviates from the power-law form. In the case $\gamma = 3$ (see the inset picture), a deflective Poisson distribution, which usually occurs in railway and road networks [28,29], emerges.

The average nearest-neighbor degree function $k_{nn}(k)$ is shown in Fig. 3. Similar to the situation in degree distribution, when $\gamma = 0.5$ and $\gamma = 1$, $k_{nn}(k)$ decays with k in a power-law form. However, when we look more closely, $k_{nn}(k)$ does not decay with k at the beginning but has a flat area when the degree is small. This result has been observed in some networks in which disassortative mixing [30] emerges only when the degree is larger than a certain value k_c while $k_{nn}(k)$ has no correlation when $k < k_c$ [31]. When the spatial constraints are stronger, the topology becomes homogeneous and the disassortativity disappears, as we see in the case $\gamma = 2$, where a nearly constant $k_{nn}(k)$ indicates an uncorrelated network topology. When $\gamma = 3$, $k_{nn}(k)$ seems to be slightly increasing as k grows. This slightly assortative mixing could be caused by the spatial community effect. Since strong spatial constraints make the nodes easily connect to their geographical neighbors, the nodes which are distributed closely will connect densely with each other and form a clustering in a certain local area.

The clustering–degree correlation $C(k)$ also exhibits some behavior close to the fact in a real network (see Fig. 4) [6]. When $\gamma = 0.5$ and $\gamma = 1$, the clustering correlation follows $C(k) \sim k^{\sigma}$ with exponent $\sigma_1 = 1.5$ and $\sigma_2 = 1$, indicating a hierarchical organization in the network formation [6]. When the spatial constraints are strong enough, as is shown in the cases of $\gamma = 2$ and $\gamma = 3$, the correlations are not so pronounced.

4.2. Vertex strength and link weight

In addition to the pure topology, the weight properties in our model also show many nontrivial correlations. In this section, we will mainly discuss the strength–degree correlation and then briefly report other weight features.

The strength s of a node, given by $s_i = \sum_{j \in V(i)} w_{ij}$, has been found to have two different behaviors with its connectivity k_i: $s_i \sim k_i$ for class I networks and $s_i \sim k_i^{\beta}$ for class II networks [32]. In real systems an example of class I networks is the scientific coauthorship network (SCN) [7], while examples of class II networks are the airline networks and the Internet

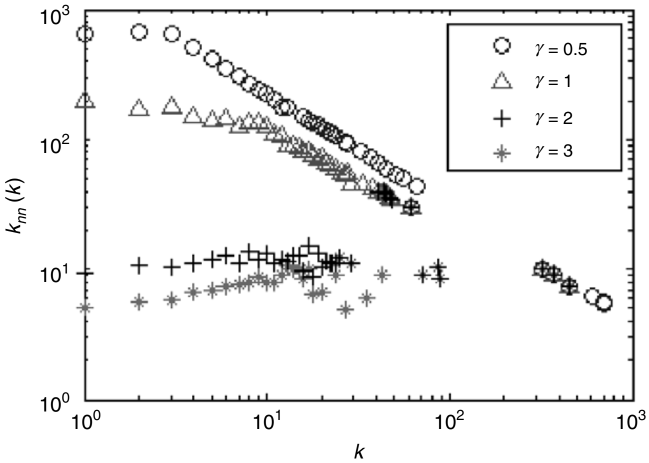

Fig. 3. The average nearest-neighbor degree function $k_{nn}(k)$ for the four cases. When $\gamma = 0.5$ and $\gamma = 1$, $k_{nn}(k)$ decays with k in a power-law form. Noted that there is a flat area at the very beginning for both cases and the turning point k_c is about 4 and 10 for the case of $\gamma = 0.5$ and $\gamma = 1$, respectively. When $\gamma = 2$, the network seems to be uncorrelated, and a slightly assortative mixing emerges when $\gamma = 3$. The simulation is averaged over 20 networks.

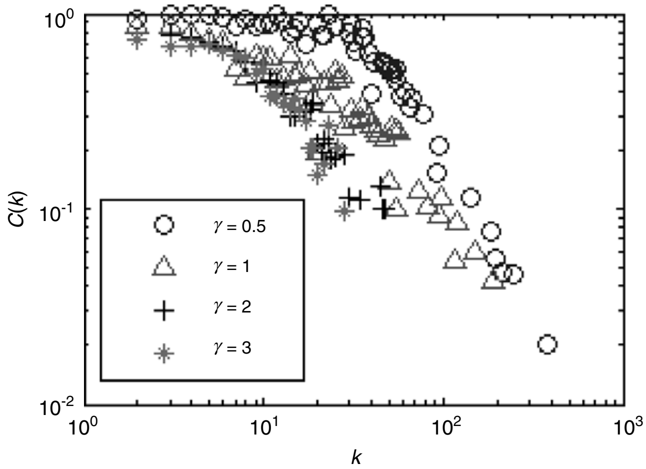

Fig. 4. The clustering degree correlation $C(k)$ for the four cases. When $\gamma = 0.5$ and $\gamma = 1$, $C(k)$ follows $C(k) \sim k^{\sigma}$ with exponents $\sigma_1 = 1.5$ and $\sigma_2 = 1$. In the cases of $\gamma = 2$ and $\gamma = 3$, the power-law correlation is not pronounced. The simulation is averaged over 20 networks.

[7,8]. In the SCN, the weight represents the collaborative frequency of two scientists, which is created directly by the two people. But this is not the case for an airline network in which the link weight, represented by the traffic, includes not only the passengers of direct flights but also others of indirect flights. Similarly and more typically, in the Internet the routers transfer a large amount information flow created by other pairs of nodes. Here we call this kind of mechanism the "transport effect". Since the expected traffic represents the flow which originates from node i (or j) and ends at the other (as we have emphasized in Section 2), it is only suitable to describe the weight in the networks without the "transport effect", namely the class I networks. However, by using the assumption proposed in Ref. [33], the relationship of the expected traffic T_{ij} and the real strength s in class II networks can be easily derived. Assuming that the traffic travels only along the shortest path between the pair and that when it encounters one or more branching points during the transport, it is supposed to be divided evenly by the number of branches, then the real node strength is expressed as

$$s_i = \sum_j T_{ij} + \frac{1}{2} \sum_{(w,w')} T_{ww'} \frac{\sigma'_{ww'}(i)}{\sigma_{ww'}} \tag{6}$$

where the second summation is over all pairs of vertices such that $w \neq w'$ and $w, w' \neq i$, $\sigma_{ww'}$ is the number of the shortest paths between vertex w and w' and $\sigma'_{ww'}(i)$ is the number of shortest paths between w and w' that passes vertex i. $\sum_j T_{ij}$ represents the total traffic that ends at node i while $\frac{1}{2} \sum_{(w,w')} T_{ww'} \frac{\sigma'_{ww'}(i)}{\sigma_{ww'}}$ represents the traffic created by other node pairs but flowing through the vertex i. One can derive the real link weight in the same way.

The strength–degree correlation for both classes of network was simulated. First, let the expected traffic T_{ij} be the link weight, namely neglecting the "transport effect". The correlation is shown in Fig. 5. The power-law correlation is not pronounced when $\gamma = 2$ and $\gamma = 3$ while it forms a linear correlation for both the cases $\gamma = 0.5$ and $\gamma = 1$. Numerical simulations suggest the when $\gamma \leq 1$, whatever its value, the model can only give rise to a linear strength–degree correlation. Now

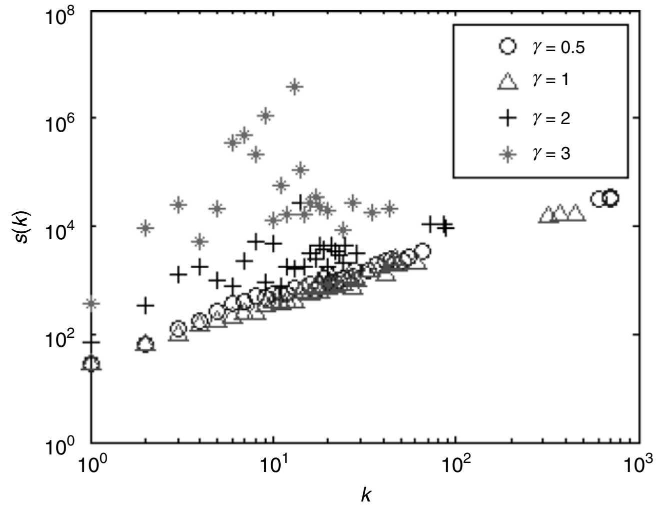

Fig. 5. The strength–degree correlation for a class I network. Here the transport effect is not taken into account. Both the cases of $\gamma = 0.5$ and $\gamma = 1$ exhibit a linear correlation, namely $s(k) \sim k$. A nonlinear correlation cannot emerge whatever value γ takes. When $\gamma = 2$ and $\gamma = 3$, the strength–degree correlations are not pronounced. The simulation is averaged over 20 networks.

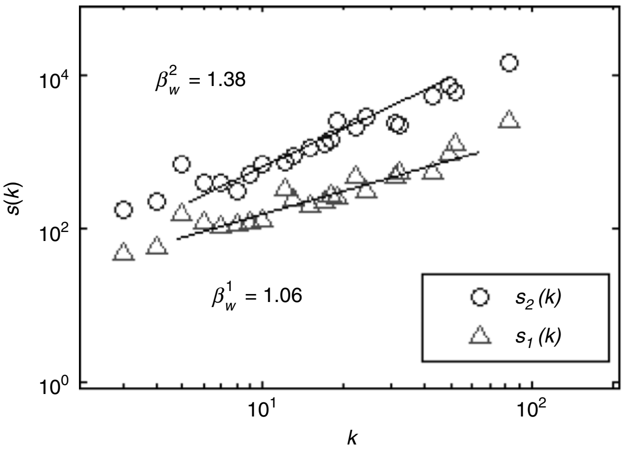

Fig. 6. The comparison of the strength–degree correlation when considering and neglecting the transport effect. Here only the case of $\gamma = 1$ is plotted. The result is qualitatively the same as the case of $\gamma = 0.5$. When considering the transport effect, the strength–degree correlation $s_2(k)$ behaves in a nonlinear form with exponent $\beta_w^2 \simeq 1.38$. When neglecting the transport effect, the strength–degree correlation $s_1(k)$ follows $s_1(k) \sim k^{1.06}$, indicating a nearly linear correlation. The simulations suggest that the nonlinear strength–degree correlation can only emerge when the transport effect is taken into account. The simulation is averaged over 20 networks.

we take the "transport effect" into account. According to Eq. (5), the strength–degree correlation of the class II networks $s_2(k)$ for the case of $\gamma = 1$ was calculated, denoting its correlation exponent β_w^2. (The result is qualitatively the same as the case of $\gamma = 0.5$.) As a comparison, the strength–degree correlation of the class I networks $s_1(k)$ was also drawn, with β_w^1 denoting its correlation exponent. As is shown in Fig. 6, both $s_1(k)$ and $s_2(k)$ follow the power-law form. But compared to the linear correlation for $s_1(k)$ ($\beta_w^1 = 1.06$), $s_2(k)$ behaves as a nonlinear correlation with exponent $\beta_w^2 = 1.38$. A large amount of numerical simulations show that as long as the "transport effect" is taken into account, a nonlinear strength–degree correlation emerges immediately. Therefore, the "transport effect" could be responsible for the nonlinear strength–degree correlation.

To prove this hypothesis, we approximately analyzed Eq. (6) and showed that the emergence of class I and class II networks was possible to be explained under a common framework. Eq. (6) can be written as

$$s_i = \sum_{j\in V(i)} T_{ij} + \left(\sum_{j\notin V(i)} T_{ij} + \frac{1}{2} \sum_{(w,w')} T_{ww'} \frac{\sigma'_{ww'}(i)}{\sigma_{ww'}} \right) \qquad (7)$$

where $V(i)$ represents the neighboring vertices of vertex i. $\sum_{j\in V(i)} T_{ij}$ is the sum of the expected traffic created by node i and its neighbors. Thus this item is not associated with the "transport effect" while the item in the brackets, namely $\sum_{j\notin V(i)} T_{ij} + \frac{1}{2}\sum_{(w,w')} T_{ww'} \frac{\sigma'_{ww'}(i)}{\sigma_{ww'}}$, completely relies on the transport mechanism. Let $T_{ij} = 1$ (this indicates that every T_{ij} is considered to be equal and we assume this approximation will not affect the result qualitatively); then Eq. (7) becomes

$$s_i = k_i + \left(N - 1 - k_i + \frac{1}{2} \sum_{(w,w')} \frac{\sigma'_{ww'}(i)}{\sigma_{ww'}} \right) \qquad (8)$$

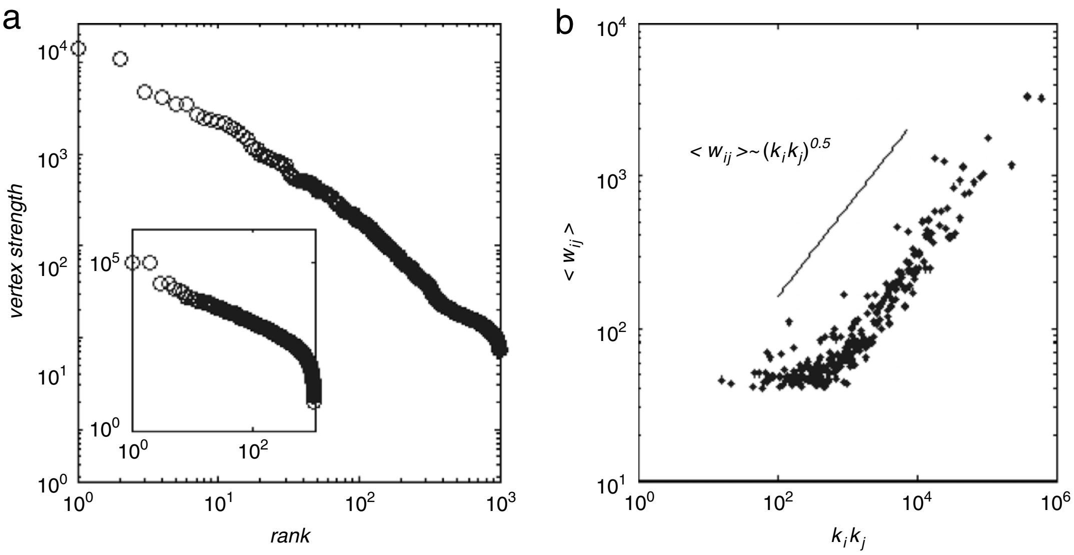

Fig. 7. (a) The distribution of the node strength. Since the strength is a continuous variable, we draw the Zipf plot. Here only the cases of $\gamma = 0.5$ and $\gamma = 3$ are drawn. (The latter is shown in the insert.) The cases of $\gamma = 1$ and $\gamma = 2$ exhibit similar properties to the cases of $\gamma = 0.5$ and $\gamma = 3$, respectively. When $\gamma = 0.5$, the node strength follows a power-law distribution. When $\gamma = 3$ (see the insert), the distribution is a little deviated from the typical power-law form when k is large, but it is still a heavy-tail distribution. (b) The weight–degree correlation $w_{ij}(k_ik_j)$ for the case of $\gamma = 0.5$. It was found to follow $\langle w_{ij}\rangle \sim (k_ik_j)^{0.5}$. Simulations indicate that a correlation can emerge when $\gamma < 0.7$. The simulation is averaged over 20 networks.

where k_i is the degree of node i and N is the number of the nodes. $\sum_{(w,w')}\frac{\sigma'_{ww'}(i)}{\sigma_{ww'}}$ is exactly the vertex betweenness centrality [34]. In a network with a transport mechanism, $s_i = N-1+\frac{1}{2}\sum_{(w,w')}\frac{\sigma'_{ww'}(i)}{\sigma_{ww'}}$. Therefore, the strength–degree correlation varies proportionately with the betweenness–degree correlation which has been demonstrated to form a nonlinear power-law distribution in some real-world networks [35]. Then $s_i \sim \sum_{(w,w')}\frac{\sigma'_{ww'}(i)}{\sigma_{ww'}} \sim k_i^{\beta}$. However, when the network lacks the "transport effect", such as SCN, the items $N-1-k_i+\frac{1}{2}\sum_{(w,w')}\frac{\sigma'_{ww'}(i)}{\sigma_{ww'}}$ are removed from Eq. (8). Then $s_i = k_i$, indicating a linear correlation. The above analysis has taken the approximation of $T_{ij} = 1$ but numerical simulations have confirmed that this approximation does not affect the result qualitatively. In summary, we consider that the "transport effect" could be responsible for the nonlinear strength–degree correlation and that a linear correlation can emerge when such effect is lacking.

Another two weighted properties are the strength distribution $p(s)$ and the weight–degree correlation $w_{ij}(k_ik_j)$. It was found that the strength distribution follows that of a heavy tail distribution in both heterogeneous and homogeneous networks while the weight–degree correlation obeys $w_{ij} \sim (k_ik_j)^{0.5}$ in airline networks [7,27]. Both of these properties are captured by our model (see Fig. 7). In Fig. 7(a), the strength distributions are drawn for the cases of $\gamma = 0.5$ and $\gamma = 3$. When $\gamma = 0.5$, $p(s)$ is a power-law distribution and when $\gamma = 3$ (see the insert), although the topology is homogeneous, $p(s)$ still follows a heavy-tail form. Fig. 7(b) shows the weight–degree correlation for the case $\gamma = 0.5$. The correlation is found to be $\langle w_{ij}\rangle \sim (k_ik_j)^{0.5}$, which is exactly the result found in airline networks. Numerical studies suggest that the correlation emerges when $\gamma < 0.7$. When $0.7 < \gamma < 1$, the correlation is not obvious, and when $\gamma > 1$, this property is absent. Both of these weight properties are under the consideration of the "transport effect".

4.3. Spatial analysis

A complete characterization of the spatial network must take into account the geographical effect. In Fig. 8, the edge length distributions for the case $\gamma = 0.5$, $\gamma = 1$ and $\gamma = 2$ are compared. (The case $\gamma = 3$ is not plotted here since it is similar to the case $\gamma = 2$.) In all three cases, the number of edges decays with the edge length, which reflects that short edges are usually preferred in a network with spatial constraints. When $\gamma = 2$, a strong geographical effect drives most links to concentrate within the area of [0, 2]. When the spatial constraints are gradually loosened, the distribution decays more and more slowly as we see in the case of $\gamma = 1$ and $\gamma = 0.5$. Furthermore, if we look more closely, it is found that when $\gamma = 1$ and $\gamma = 0.5$, the distributions are not monotonic but have a small peak around 7. This phenomenon was first observed in Ref. [18], where both the Internet and the airline network exhibit the property. Since the present model connects preferentially two nodes with larger expected traffic, according to Eq. (1), long-range links can only occur among those node pairs with large fitness. Thus the links corresponding to the small peak are typically the connections between large-fitness nodes. On the other hand, to increase the total expected traffic, the small-fitness nodes will connect locally to a certain vertex with relatively larger fitness. This behavior corresponds to the fact that in the airline networks small airports usually connect to a hub nearby while the large airports are able to connect to each other despite the long geographical distance.

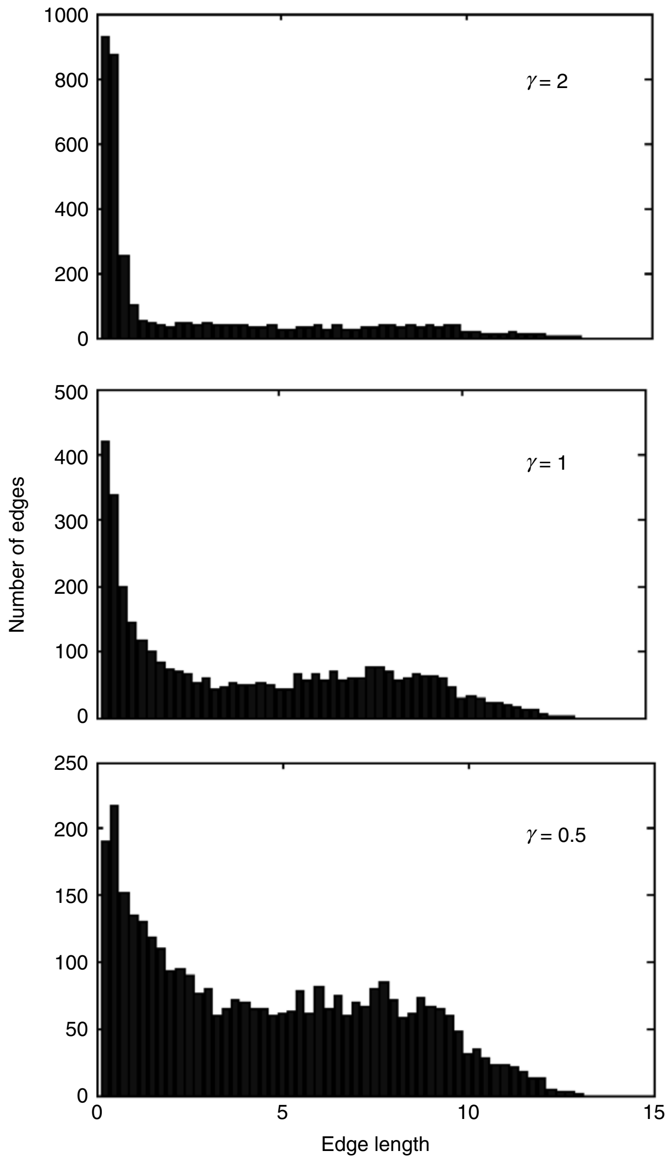

Fig. 8. The edge length distributions for the cases of $\gamma = 0.5$, $\gamma = 1$ and $\gamma = 2$. In all three cases, the number of edges decays with the edge length. But when the spatial constraints are gradually loosened, the distribution decays more and more slowly. In the case of $\gamma = 1$ and $\gamma = 0.5$, the distribution are not monotonic but have a small peak around 7. The simulation is averaged over 20 networks.

In Fig. 9, the distance strength–degree correlation is drawn. The distance strength is defined as $s_d(k_i) = \sum_{j\in V(i)} D_{ij}$ [36]. This quantity gives the cumulated distances of all connections from the neighboring nodes. Similar to the strength–degree correlation, the distance strength–degree correlation follows $s_d(k) \sim k^{\beta_d}$. When $\gamma = 0.5$, $\beta_d \simeq 0.97$, indicating a linear correlation. When $\gamma = 1$, a nonlinear correlation emerges with exponent $\beta_d \simeq 1.23$. In the case $\gamma = 2$ and $\gamma = 3$, the correlations are not so pronounced as for the first two cases.

5. Conclusion

In this paper, we have proposed a spatial weighted network model based on the expected traffic. The expected traffic is the prediction of the flow created by two vertices and is calculated by the improved gravity equation. In contrast to the previous studied spatial graph, we consider that such expected benefit plays a key role in network design. The model is constructed by optimizing the total expected traffic of the network. This mechanism suggests that busier links are usually established earlier.

The topology of the model depends on the two parameters α and γ, which control the fitness and the geographical constraints, respectively. By changing them, we found structures ranging from star-like networks to decentralized networks. Moreover, we have studied in detail the statistical properties of the model, including pure topology, weight and spatial analysis. Extensive numerical simulations have confirmed that our gravity model is able to reproduce many properties observed in real-world networks. Notably, we found that linear and nonlinear strength–degree correlation could emerge when considering and neglecting the "transport effect", respectively. This result could offer an explanation for the different strength–degree correlation observed in social networks and technical networks.

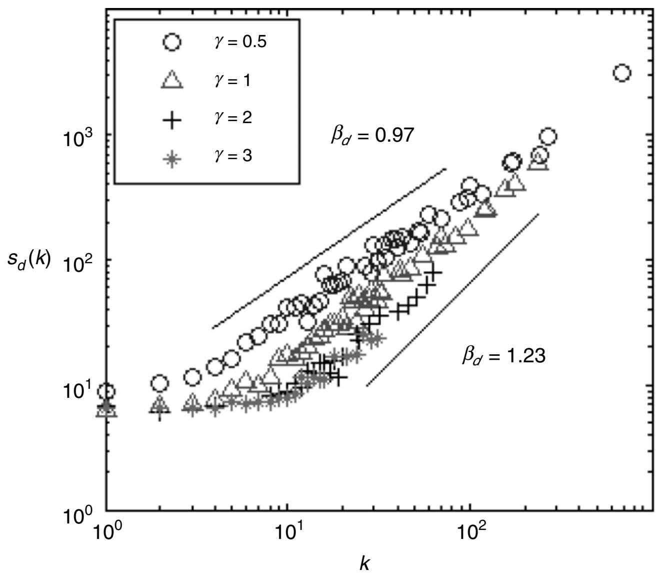

Fig. 9. The distance strength–degree correlation $s_d(k)$ for the four cases. $s_d(k)$ is found to follow $s_d(k) \sim k^{0.97}$ and $s_d(k) \sim k^{1.23}$ for the cases of $\gamma = 0.5$ and $\gamma = 1$, respectively. When $\gamma = 2$ and $\gamma = 3$, the correlations are not so pronounced as for the first two cases. The simulation is averaged over 20 networks.

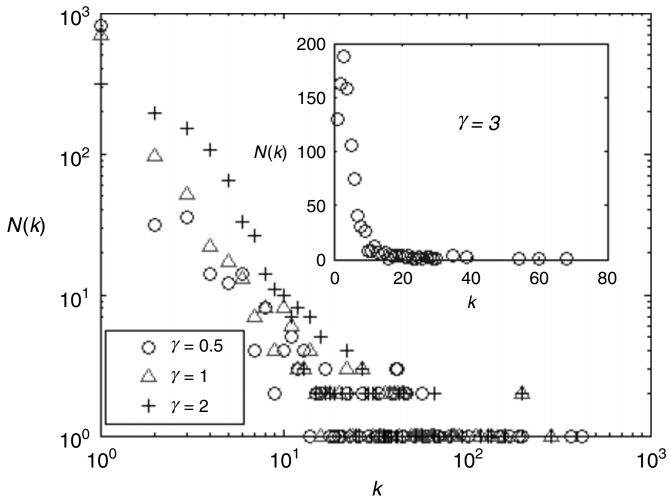

Fig. 10. Degree distribution for the case of $\phi = -2$.

Our model can not only help us understand the formation and dynamics of the network, but also provides a way to optimize our resources distribution, such as the bandwidth or the capacity of the system. In our model, topology and weight properties are determined as long as the fitness and the geographical information are known. Thus one can predict the topology and dynamics before the network is constructed and then reasonably distribute the resources to the necessary places.

Acknowledgements

This work was partially supported by Shanghai Development Foundation for Science and Technology under Grant Numbers 06JC14082 and 05XD14021.

Appendix A

The power-law exponent of the fitness distribution in our study is $\phi = -3$. However, most power-law distributions occurring in nature have their exponents $\phi \in [-2, -3]$ [37]. To examine if the structure of our model changes with ϕ significantly, we simply test the degree distribution for $\phi = -2$ and $\phi = -2.5$. The result is illustrated in Figs. 10 and 11. Compared with Fig. 2 ($\phi = -3$), it is found that the properties are not sensitive to the change of ϕ. Thus despite the different power-law exponents for all kinds of real systems, they can still share common properties. If ϕ is far from $[-2, -3]$, the properties can be recovered by extending the value of γ.

Appendix B

To prove that the properties of our model are stabilized when $N = 1000$, the degree distribution and the average nearest-neighbor degree function for size $N = 8000$ are illustrated in Figs. 12 and 13. The properties do not change significantly.

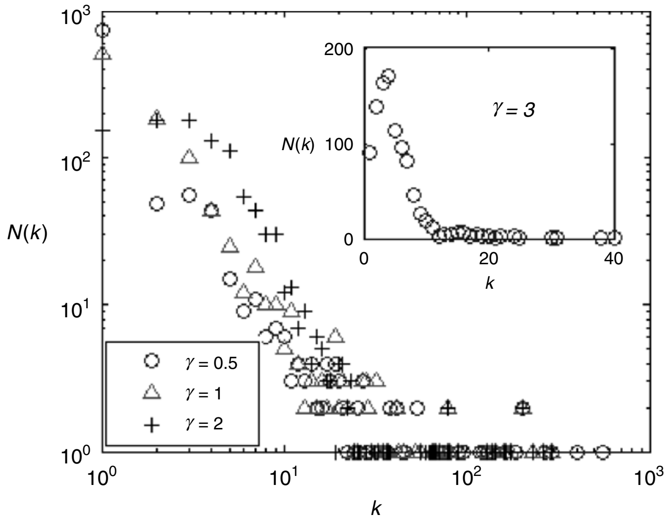

Fig. 11. Degree distribution for the case of $\phi = -2.5$.

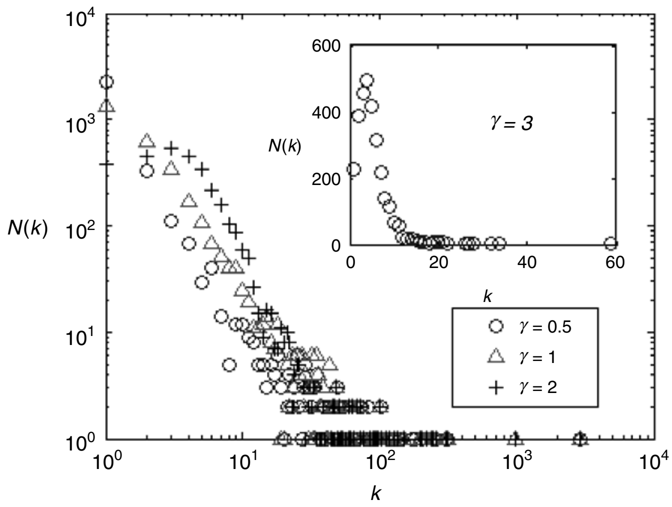

Fig. 12. Degree distribution for $N = 8000$.

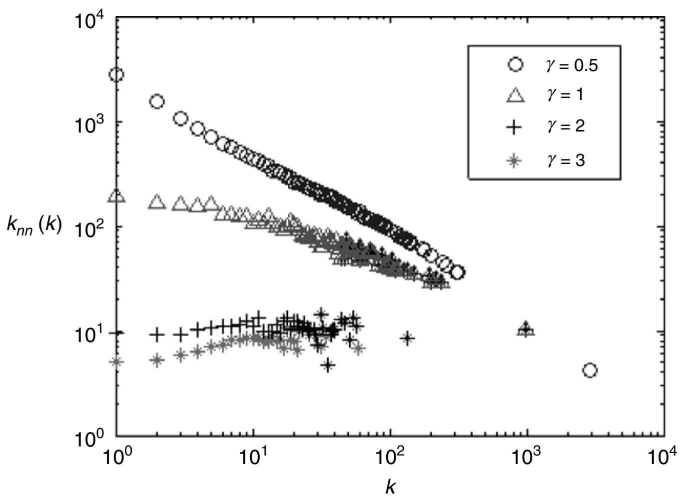

Fig. 13. The average nearest-neighbor degree function $k_{nn}(k)$ for $N = 8000$.

References

[1] A.-L. Barabási, R. Albert, Science 286 (1999) 509;
A.-L. Barabási, R. Albert, H. Jeong, Physica A 272 (1999) 173.
[2] G. Bianconi, A.-L. Barabási, Europhys. Lett. 54 (2001) 436.
[3] G. Bianconi, A.-L. Barabási, Phys. Rev. Lett. 86 (2001) 5632.
[4] M. Boguná, R. Pastor-Satorras, Phys. Rev. E 68 (2003) 036112.
[5] A. Vázquez, R. Pastor-Satorras, A. Vespignani, Phys. Rev. E 65 (2002) 066130.
[6] Erzsébet Ravasz, Albert-László Barabási, Phys. Rev. E 67 (2003) 026112.
[7] A. Barrat, M. Barthélemy, R. Pastor-Satorras, A. Vespignani, Proc. Natl. Acad. Sci. USA 101 (2004) 3747.
[8] M.A. Serrano, M. Boguná, A. Díaz-Guilera, Phys. Rev. Lett. 94 (2005) 038701.
[9] A. Barrat, M. Barthélemy, A. Vespignani, Phys. Rev. Lett. 92 (2004) 228701.
[10] Y.-H. Eom, C. Jeon, H. Heong, B. Kahng, Phys. Rev. E 77 (2008) 056105.
[11] G. Bianconi, cond-mat/0412399 (2004).
[12] R. Pastor-Satorras, A. Vespignani, Evolution and Structure of the Internet: A Statistical Physics Approach, Cambridge University Press, Cambridge, 2004.
[13] V. Latora, M. Marchiori, Phys. Rev. E 71 (2005) 015103.
[14] Ding-Ding Han, Jiang-Hai Qian, Jin-Gao Liu, Physica A 388 (2009) 71–81;
Ding-Ding Han, Jin-Gao Liu, Yu-Gang Ma, Xiang-Zhou Cai, Wen-Qing Shen, Chinese Phy. Lett 21 (2004) 1855;
Ding-Ding Han, Jin-Gao Liu, Yu-Gang Ma, Chinese Phy. Lett 25 (2008) 765;
Ding-Ding Han, Complex network property and reliability analysis of wireless receiver circuit, Electronic Measurement Technology 31 (9) (2008) (in Chinese).
[15] R. Guimerà, S. Mossa, A. Turtschi, L.A.N. Amaral, Proc. Natl. Acad. Sci. USA 102 (2005) 7794;
R. Guimerà, L.A.N. Amaral, Eur. Phys. J. B 38 (2004) 381.
[16] V. Latora, M. Marchiori, Physica A 314 (2002) 109.
[17] R. Albert, I. Albert, G.L. Nakarado, Phys. Rev. E 69 (2004) 025103.
[18] M.T. Gastner, M.E.J. Newman, Eur. Phys. J. B 49 (2006) 247–252.
[19] M.T. Gastner, M.E.J. Newman, Phys. Rev. E 74 (2006) 016117.
[20] B. Waxman, Routing of multipoint connections, IEEE J. Selec. Areas Commun. 6 (1988) 1617.
[21] S.-H. Yook, H. Jeong, A.-L. Barabási, Proc. Natl. Acad. Sci. USA 99 (2002) 13382.
[22] R. Xulvi-Brunet, I.M. Sokolov, Phys. Rev. E 75 (2007) 046117.
[23] F.J. Jin, Geographical Res. 20 (2001) 31.
[24] B. Latané, Personality Social Psych. Bull 21 (1995) 795.
[25] J.S. Liu, Y.G. Chen, Scintia Geographica Sinica 20 (2000) 0528.
[26] A.S. Fotheringham, Ann. Assoc. Am. Geographers 71 (1981) 425.
[27] Woo-Sung Jung, Fengzhong Wang, H. Eugene Stanley, Europhys. Lett. 81 (2008) 48005.
[28] V. Latora, M. Marchiori, Physica A 314 (2002) 109–113.
[29] K.A. Seaton, L.M. Hackett, Physica A 339 (2004) 635–644.
[30] M.E.J. Newman, Phys. Rev. E 64 (2003) 026126.
[31] Hong-Kun Liu, Tao Zhou, Acta Physica Sinica 56 (2007) 0106.
[32] G. Bianconi, cond-mat/0412399 (2004).
[33] K.-I. Goh, B. Kahng, D. Kim, Phys. Rev. Lett. 87 (2001) 278701.
[34] Petter Holme, Beom Jun Kim, Phys. Rev. E 65 (2002) 066109.
[35] K.I. Goh, E. Oh, B. Kahng, D. Kim, Phys. Rev. E 67 (2003) 017101.
[36] A. Barrat, M. Barthélemy, A. Vespignani, arXiv:physics/0504029v1.
[37] M.E.J. Newman, arXiv:cond-mat/0412004v3.