A LETTERS JOURNAL EXPLORING THE FRONTIERS OF PHYSICS

Scaling properties in spatial networks and their effects on topology and traffic dynamics

A LETTERS JOURNAL EXPLORING THE FRONTIERS OF PHYSICS

March 2010

EPL, **89** (2010) 58002 www.epljournal.org
doi: 10.1209/0295-5075/89/58002

Scaling properties in spatial networks and their effects on topology and traffic dynamics

H. YANG, Y. NIE, A. ZENG, Y. FAN, Y. HU[(a)] and Z. DI[(b)]

Department of Systems Science, School of Management, Beijing Normal University - Beijing 100875, PRC

received 18 August 2009; accepted in final form 16 February 2010
published online 16 March 2010

PACS 89.75.Hc – Networks and genealogical trees
PACS 89.75.Fb – Structures and organization in complex systems

Abstract – Empirical studies on the spatial structures in several real transport networks reveal that the distance distribution in these networks obeys a power law. To discuss the influence of the power law exponent on the network's structure and function, a spatial-network model is proposed. Based on a regular network and subject to a limited cost C, long-range connections are added with power law distance distribution $P(r) = ar^{-\delta}$. Some basic topological properties of the networks generated by the model with different δ are studied. It is found that the network has the smallest average shortest path when $\delta = 2$. Then a classic traffic model on our model networks is investigated. It is found that $\delta = 1.5$ is the optimization value for the traffic process in our model. All of these results give us some deep understanding about the relationship between spatial structure and network function.

Introduction. – In the last few years, the analysis and modeling of networked systems have received considerable attention within the physics community, including the World Wide Web, the Internet, and biological, social, and infrastructure networks [1–3]. Some of these networks exist only as abstract networks where the precise positions of the network nodes have no particular meaning, such as biochemical networks and social networks, while many others in which nodes have well-defined positions, are different. In all kinds of network, a particular class is the spatial network embedded in the real space. Many networks belong to this class like the neural network [4], communication networks [5], the electric-power grid [6], transportation systems including airport [7], street [8], railway and subway [9] networks. Most of the previous works on the studies of complex networks have focused on the characterization of the topological properties or other issues, while the spatial aspect has received more attention recently [10–13].

Actually, geography greatly matters. The geography information of the nodes and the distance between nodes would determine the characteristics of the network and play a more or less important role in the dynamics happening in the network. Ignoring it one would miss some of these systems' interesting features. Empirical studies have revealed some interesting phenomena about the spatial structure of networks. One is that the distance distribution of the edges obeys a power law. Examples include the Internet [14], social communication networks [15] and online social networks [16]. Recent research on circuit placement showed that the wire length of real circuits exhibits a power law distribution [17]. The anatomical distance distributions in human brain networks can also be well fitted by an exponentially truncated power law [18]. More evidence comes from the empirical research on transportation networks. For Japanese airline networks [19], even if there is an exponential decay in domestic flights, the distance distribution follows a power law when international flights are added. For the U.S. intercity passenger air transportation network, the distribution of the edge distance has a power law tail with exponent $\delta = 2.20 \pm 0.19$ [20].

Why does the geographical embedded network show this special spatial structure? How does the distance distribution affect the network's function? All these problems are interesting. Analyzing these problems will help us understand the real spatial network deeply and benefit us for the design of the transport system. In our opinion, the clue to answer these questions is the consideration of costs and efficiency when the geographical embedded networks were designed.

(a) E-mail: yanqing.hu.sc@gmail.com
(b) E-mail: zdi@bnu.edu.cn

In spatial embedded networks, especially transport networks, the connections between nodes are restricted by cost constraints, reflected through the distance distribution. The cost of establishing long-range connections between distant spots is usually higher than the cost of establishing short-range connections. For electric-power grids, the connection cost between farther spots is even higher, given that in long high-voltage lines a large amount of energy is lost during the transmission [21]. So we can easily understand that the number of short-range connections is much higher than the long-range connections in these networks.

To demonstrate how the distance distribution of the connections affect the structure, function and the traffic dynamics process of the networks, we proposed a spatial-network model in this paper. The model takes into account both the power law distribution of the distance and the total cost of the edges. We construct spatially constrained networks embedded in a geographical space, the distance distribution of the network obeys a power law distribution $P(r) \propto r^{-\delta}$ and the network has a limited total cost C to create edges. We analyse the spatial network in detail, it shows that the network has the smallest average shortest path when $\delta = 2$, which is not influenced by the value of the total cost C and the size of the networks. However, when we investigate a simple traffic model on the network, it is found that the network with $\delta = 1.5$ has the biggest transport capacity.

Spatial-network model with limited total cost. – Generally, the cost and efficiency are equally important in transport networks. The network structure is the result caused by the tradeoff between cost and efficiency.

The model network is embedded in a k-dimensional regular network. The long-range connections is generated from a power law distance distribution by the approach suggested in [12]. Different from the previous model [12], we introduce a total cost C to this network model. Every edge has a cost c which is linearly proportional to its distance r. For simplification, the edge connecting node i and j would cost c_{ij}, and is represented by its length r_{ij} in the model. Naturally, the cost is limited in the network, which display as a certain limited total cost C. The network is constructed as follows.

1. N nodes are arranged in a k-dimensional lattice. Every node is connected with its nearest neighbors which can keep every node reachable. In addition, between any pair of nodes there is a well-defined Euclidean distance.
2. A node i is chosen randomly, and a certain distance r ($2 \leqslant r \leqslant N_{max}$, N_{max} is the largest distance between any nodes in the initial network) is generated with probability $P(r) = ar^{-\delta}$, where a is determined from the normalization condition $\sum_{r=2}^{N_{max}} P(r) = 1$.
3. One of the N_r nodes that are at a distance r from node i is picked randomly, for example node j. An edge between nodes i and j is created if there exist no edge between them yet. The denominator of the weight between i and j is plus 1 if there has already been an edge between them.
4. After step 3, a certain cost $c_{ij} = r_{ij}$ is generated. Repeat step 2 and 3 until the total cost reaches C.

After these steps, we can obtain a weighted network in which the weight reflects the closeness of nodes. For example, if nodes i and j are connected three times, the weight of the edge between them will be 1/3. Finally, we also transfer the weighted network into a binary one by imposing all the weight of the edges to 1. Obviously, there are two significant features of the spatial network generated by our model: the power law distribution of the long-range connections in the network and the restriction on total energy. In this model, the distance distribution and the total cost play important roles in the formation of the network. Consequently, we first focus on how the topological properties are affected by the two factors. We are interested in how the power law exponent δ influences the topological properties in our model, including the average degree (node intensity in weighted networks), and the average shortest path of the network. We have simulated the model both in a 1-dimensional chain and 2-dimensional lattice, with periodic boundary condition, respectively. They all give the same qualitative results. So in the following, we only report the results of the 1-dimensional chain with periodic boundary condition. The network size is typically $N = 500$ and the results are the average of 1000 realizations. The total cost is set as $C = Nc$ ($C = 500c$), where c is the average cost per node which is set as 10, 20, 30, 40, 50 successively in the simulation. The results are shown in fig. 1. Interestingly, it shows that in weighted networks, the network enjoys the minimum average shortest path when $\delta = 2$. In the binary networks, the average degree increases with δ. The average shortest path reaches its minimum when $\delta = 2$, which is consistent with the result in [22]. Moreover, these two kinds of networks can have the maximum average node intensity and average degree, respectively. So when the total cost is limited, the network with $\delta = 2$ has the lowest average shortest path, which may explain why the power law exponent of the distance distribution is close to 2 in many airline networks. As we know, in the public-transport networks, travelers prefer less transfers when traveling. The transport network with $\delta = 2$ has the lowest average shortest path in topological structure, which can make travelers have the least transfers and more convenience.

What is more, we also investigate how the total cost affects the networks properties, especially in the binary network which we are about to use to investigate how the spatial property affects traffic dynamics. As we know, because of the probability distribution $P(r) = ar^{-\delta}$, most of the connections are short while a few connections are relatively long when the connections distance shows a

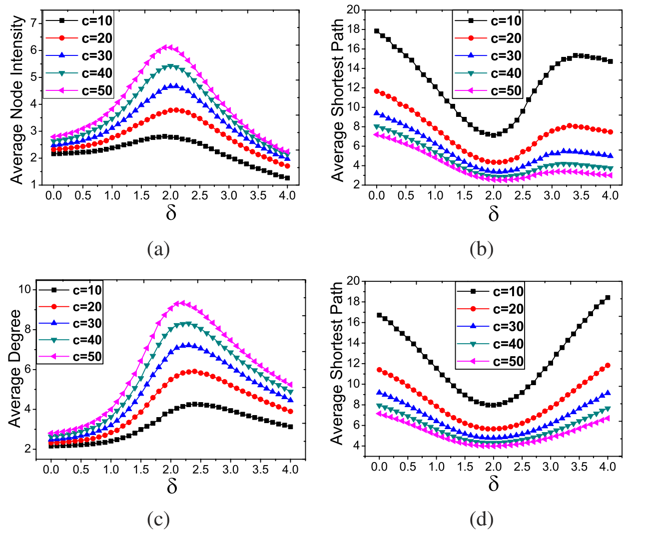


Fig. 1: (Colour on-line) Basic topological properties in the model network. (a), (b) represent the node intensity and average shortest path of different limited total cost $C = 500c$ ($c = 10, 20, 30, 40, 50$) respectively in weighted spatial networks. (c), (d) show the average degree and average shortest path in corresponding binary networks. Each point is averaged 1000 times.

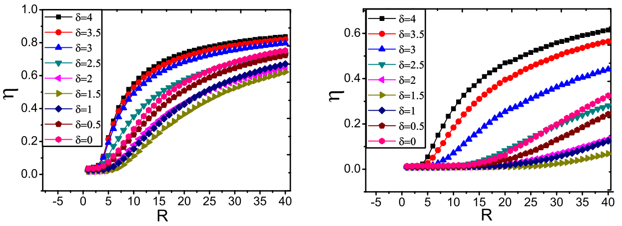
power law distribution. However, when the total cost reaches a certain value in the binary network, the network with a given size would become nearly full connected and the power law distance distribution will be destroyed. Specifically, when the power law exponent δ is close to 4, the underlying network cannot provide enough short long-range connections when $\delta = 2$, 3, or a higher-dimensional space, although the total cost is not too high, it may result that the distance distribution shifts from standard power law to truncated power law, which is prevalent in the real world. In this model, most of the long-range connections show truncated power law with only a small number of short long-range connections shifted. With the increase of total cost, more long-range edges are created. The average degree increases, while the average shortest path decreases. One interesting result is that for a certain range of total cost, the relationship between topological properties and the parameter δ keeps the same qualitatively as shown in fig. 1. Then, how this structure is related with the traffic dynamics? Next section will show us some results.

Traffic process on the model spatial network. – From the analysis of the spatial network above, the average shortest path in the model network may explain why some public networks's exponent is close to 2 to a certain extent. What will it happen when we consider the traffic dynamics on the above spatial embedded networks? To investigate the traffic dynamics it may be helpful to understand real systems such as the express-delivery network. The express-delivery network is constructed based on all kinds of public-transport networks, especially the airline networks. But it has its own features, which are different from the public-transport networks. First, the express-delivery network is constructed from the perspective of overall optimization, while the public-transport networks are constructed by self-organization, based on a local optimization process [23]. Second, the express-delivery network really cares about the traffic process on the network. Both efficiency and cost are important factors to shape the network structure. Third, as the traffic model on networks, the bottleneck of the express-delivery network lies in the node, the capability of the node determines the whole network's efficiency. So to investigate the traffic flow on such kind of network may help us to find how the exponent δ influences the traffic flow on the model network.

We employ a typical traffic dynamics [24] on this spatial network. Firstly, generate the underlying network infrastructure with the method we propose in the previous section. We also take the network as a binary one. Then a traffic dynamics is modeled on the network. All the nodes embedded in the spatial network are treated as both hosts and routers. We assume that every node can deliver at most D packets one step toward their destinations. At each time step, there are R packets generated homogeneously on the nodes in the system. The packets are delivered from their own origin nodes to destination nodes by a special routing strategy. There are many kinds of different routing strategies, such as the shortest pathways routing strategy, the local dynamical strategy [25] and so on. Here the shortest-pathways routing strategy is used. A packet, upon reaching its destination, is removed from the system. The order parameter

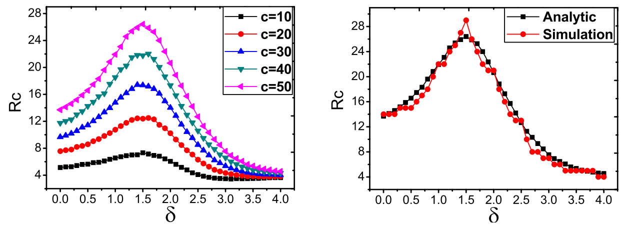
$$\eta(R) = \lim_{t\to\infty} \frac{1}{R} \frac{\langle \Delta N_p \rangle}{\Delta t} \tag{1}$$

is used to characterize the phase transition. Here $\Delta N_p = N_p(t + \Delta t) - N_p(t)$, $\langle \dots \rangle$ denotes taking the average over a time window of width Δt. $N_p(t)$ is the number of packets in the system at time t. Hence, the order parameter η is actually corresponding to the average packets number rate congested in the system. We are most interested in the critical value R_c (as measured by the number of packets created within the network per unit time), where a phase transition takes place from free flow to congested traffic. This critical value can best reflect the maximum capability of a system handling its traffic. In particular, for $R < R_c$, the numbers of created and delivered packets are balanced, leading to a steady free traffic flow. At this time, the order parameter η has a small value close to 0 because there is no congestion in the system. For $R > R_c$, traffic congestion occurs as the number of accumulated packets increases with time, simply because the capacities of the nodes for delivering packets are limited. Under this circumstance, of course the order parameter η will become larger than the small value. Therefore, in our paper, we determine R_c in this way: once the order parameter is larger than the small value mentioned before, R_c equals the corresponding R. Here, the small value is chosen as 0.05.



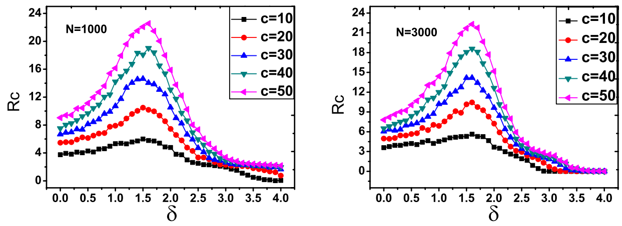
(a) (b)

Fig. 2: (Colour on-line) The simulation for the critical value $\eta(R)$ as a function of R, different colours mean different δ. The network size here is $N = 500$ while every node has the same delivery ability $D = 1$. (a) The total cost C is 500×10. (b) The total cost C is 500×50. This figure shows that the network with $\delta = 1.5$ is the best for the traffic process, which has no relationship with the value of the total cost C. Each point is averaged 50 times.



(a) (b)

Fig. 3: (Colour on-line) The network size here is $N = 500$ while every node has the same delivery ability $D = 1$. (a) The analytical results of R_c (using the method in [24]) with different values of the parameter δ. The total cost C increases from bottom to top. (b) The critical R_c *vs.* δ with the total cost $C = 500 \times 50$. Both simulation and analysis indicate that the maximum R_c corresponds to $\delta = 1.5$. Each point is averaged 100 times.



(a) (b)

Fig. 4: (Colour on-line) The analytical results of R_c with different values of the parameter δ when the network size is (a) $N = 1000$ and (b) $N = 3000$, while every node has the same delivery ability $D = 1$. The results indicate that the $\delta = 1.5$ has no relationship with the network size. Each point is averaged 50 times.

For simplicity, here we construct a network with 500 nodes under the total cost $C = 500 \times 10$ and $C = 500 \times 50$, and set that every node has the same delivery ability $D = 1$. We will adjust the network parameters δ to generate different networks, and then investigate how δ affects the critical value R_c and the order parameter $\eta(R)$.

The simulation results for the critical value $\eta(R)$ as a function of R on the model networks are reported in fig. 2. It shows that the network with $\delta = 1.5$ has the biggest R_c and the smallest $\eta(R)$.

In previous works, the traffic handling capacity of a particular network has been estimated by a simple analysis method [24]. The betweenness coefficient of node v can be calculated as $g(v) = \Sigma_{s \neq t} \frac{\sigma_{st}(v)}{\sigma_{st}}$. Here σ_{st} is the number of shortest paths going from s to t and $\sigma_{st}(v)$ is the number of shortest paths going from s to t and passing through v. Note that with the increasing of parameter R (number of packets generated every step), the system undergoes a continuous phase transition to a congested phase. Below the critical value R_c, there is no accumulation at any node in the network and the number of packets that arrive at node u is $Rg_u/N(N-1)$ on average. Therefore, a particular node will collapse when $Rg_u/N(N-1) > D_u$, where g_u is the betweenness coefficient and D_u is the transferring capacity of node u. So, congestion occurs at the node with the largest betweenness. Thus R_c can be estimated as $R_c = D_u N(N-1)/g_{max}$, where g_{max} is the largest-betweenness coefficient of the network.

In fig. 3(a), the analysis results of R_c with network parameters δ are shown. It indicates that this kind of network has the biggest R_c when δ is close to 1.5, which is in good agreement with the simulation results as shown in fig. 3(b). In addition, we also calculated the situations for 1000 nodes and 3000 nodes, the R_c also maximizes when $\delta = 1.5$. Details are shown in fig. 4.

Conclusions. – In this paper, we tried to investigate how the spatial properties affect the topology as well as the function of networks. According to many former empirical researches, the distances of real systems such as airline systems and express-delivery systems are inclined to obey a power law distribution. Actually, our results intend to help us understand why they are like this and what their effects are. Specifically, a spatial-network model is proposed and the details of the model are studied. In the model, long-range connections are added with probability $P(r) = ar^{-\delta}$ in a regular lattice while the total cost is limited to C. Some basic topological properties of the network generated by the model are investigated. It is found that the network has the smallest average topological shortest path when the power law exponent $\delta = 2$. This may be the reason why the distance distribution in an airport network shows a power law with exponent close to 2, in fact people care more about the convenience and prefer less transfers when traveling by air. Then a traffic model is studied on the model network. We find that the network with $\delta = 1.5$ is the most conducive to the traffic process. Interestingly, it has not the smallest average topological shortest path. Our results indicate that spatial constraints have an important influence on the transport networks and should be taken into account when modeling real complex systems.

***

This work is partially supported by the 985 Project and NSFC under the grants No. 70771011 and No. 60974084. YH was supported by the BNU excellent PhD Thesis Project.

REFERENCES

[1] Albert R. and Barabasi A.-L., *Rev. Mod. Phys.*, **74** (2002) 47.
[2] Dorogovtsev S. N. and Mendes J. F. F., *Adv. Phys.*, **51** (2002) 1079.
[3] Newman M. E. J., *SIAM. Rev.*, **45** (2003) 167.
[4] Sporns O., *Complexity*, **8** (2002) 56.
[5] Latora V. and Marchiori M., *Phys. Rev. E*, **71** (2005) 015103.
[6] Albert R., Albert I. and Nakarado G. L., *Phys. Rev. E*, **69** (2004) 025103.
[7] Guimerá R., Mossa S., Turtschi A. and Amaral L. A. N., *Proc. Natl. Acad. Sci. U.S.A.*, **102** (2005) 7794.
[8] Crucitti P., Latora V. and Porta S., *Phys. Rev. E*, **73** (2006) 036125.
[9] Latora V. and Marchiori M., *Physica A*, **314** (2002) 109.
[10] Gastner M. T. and Newman M. E. J., *Eur. Phys. J. B*, **49** (2006) 247.
[11] Cardillo A., Scellato S., Latora V. and Porta S., *Phys. Rev. E*, **73** (2006) 066107.
[12] Kosmidis K., Havlin S. and Bunde A., *EPL*, **82** (2008) 48005.
[13] Hayashi Y. and Matsukubo J., *Phys. Rev. E*, **73** (2006) 066113.
[14] Yook S.-H., Jeong H. and Barabási A.-L., *Proc. Natl. Acad. Sci. U.S.A.*, **99** (2002) 13382.
[15] Lambiotte R., Blondel V. D., Kerchove C. D., Huens H., Prieur C., Smoreda Z. and Dooren P. V., *Physica A*, **387** (2008) 5317.
[16] Liben-Nowell D., Novak J., Kumar R., Raghavan P. and Tomkins A., *Proc. Natl. Acad. Sci. U.S.A.*, **102** (2005) 11623.
[17] Dambre J., PhD Thesis, Ghent University, Faculty of Engineering (2003).
[18] He Y., Chen Z. J. and Evans A. C., *Cereb. Cortex*, **17** (2007) 2407.
[19] Hayashi Y., *IPSJ Digit. Cour.*, **2** (2006) 155.
[20] Xu Z. and Harriss R., *GeoJournal*, **73** (2008) 102.
[21] Xulvi-Brunet R. and Sokolov I. M., *Phys. Rev. E*, **75** (2007) 046117.
[22] Li G., Reis S. D. S., Moreira A. A., Havlin S., Stanley H. E. and Andrade J. S. jr., *Phys. Rev. Lett.*, **104** (2010) 018701.
[23] Barthelem M. and Flammini A., *Phys. Rev. Lett.*, **100** (2008) 138702.
[24] Zhao L., Lai Y.-C., Park K. and Ye N., *Phys. Rev. E*, **71** (2005) 026125.
[25] Hu M.-B., Jiang R., Wu Y.-H., Wang W.-X. and Wu Q.-S., *Physica A*, **387** (2008) 4967.