

Degree heterogeneity in spatial networks with total cost constraint

WEIPING LIU, AN ZENG^(a) and YANBO ZHOU

Department of Physics, University of Fribourg - Chemin du Musée 3, CH-1700 Fribourg, Switzerland

Abstract – Recently, Li *et al.* (*Phys. Rev. Lett.*, **104** (2010) 018701) studied a spatial network which is constructed from a regular lattice by adding long-range edges (shortcuts) with probability $P_{ij} \sim r_{ij}^{-\alpha}$, where r_{ij} is the Manhattan length of the long-range edges. The total length of the additional edges is subject to a cost constraint ($\sum r = C$). This spatial network model displays an optimal exponent α for transportation (measured by the average shortest-path length). However, we observe that the degree in such spatial networks is homogeneously distributed, which is different from some real networks. In this letter, we propose a method to introduce degree heterogeneity in spatial networks with total cost constraint. Results show that with degree heterogeneity the optimal exponent shifts to a smaller value and the average shortest-path length can further decrease. Moreover, we find the optimal degree heterogeneity for transportation. We further consider the synchronization on the spatial networks and related results are discussed. Our new model may better explain the features of real transportation systems.

Introduction. – Spatial features play a significant role in the transportation networks [1], Internet [2], mobile phone networks [3], power grids [4], social networks [5] and neural networks [6]. In the past decade, many systems have been modeled by complex networks where nodes and links are embedded in space. In these models, nodes are located in the plane and the geometric distance between nodes is well defined. Then links are constructed according to rules based on spatial indices [7–12]. Growing spatial networks were also studied [13,14]. Generally, these models of spatial networks are able to reproduce some properties of the real systems such as community structure, scale-free connection length distribution and so on. Moreover, some methods have been proposed to design robust spatial networks for transportation [15–18]. For a detailed review of the field, see [19].

The most important feature of spatial network models is that there is a cost associated with the length of links, which has dramatic effects on the topology and function of these networks. A total cost constraint has been recently introduced to design the spatial networks [20–25]. The total cost C is defined as the total length of the links, $C = \sum r$, where r is the length of links. In [20], pairs of

sites ij in 2-dimensional lattices are randomly chosen to be linked with long-range connections with probability $P_{ij} \sim r_{ij}^{-\alpha}$, where r_{ij} is the Manhattan distance between sites i and j . New links are added until the total length of the links reaches the total cost C . The exponent α controls the trade-off between the link length and link number. A large value of α allows for the formation of many short-length links while the small α favors the creation of a few long-length links. The authors in [20] show that the optimal exponents for both the average shortest-path length and navigation steps are $\alpha = 3$ in 2-dimension and $\alpha = 2$ in 1-dimension (the 1-dimensional scenario has been proved analytically very recently in [23]). The authors claim that such model reveals the optimized aspect of airline networks under the conditions of geographical availability (for customer satisfaction) and cost limitations (for airline company profit). Similar works have been carried out to study dynamics such as traffic congestion and synchronization on spatial networks under total cost constraint, and the optimal link length exponents are found [22,24,25].

However, we observe that the degree distribution in these spatial network models with total cost constraint is homogenous, which is far different from real cases, including airline systems. Many previous works have revealed that many geographic-based networks have heterogeneous

^(a)E-mail: an.zeng@unifr.ch

degree distribution [1]. In this letter, we propose a method to introduce degree heterogeneity in spatial networks with total cost constraint. We find that degree heterogeneity affects the optimal exponent for the link length distribution. Specifically, the exponent shifts to a smaller value. Also, the average shortest-path length can further decrease due to degree heterogeneity and the optimal degree heterogeneity for transportation is found. Moreover, we study the effect of degree heterogeneity on the dynamics taking place on the spatial networks, and we observe the same shifting phenomenon of the optimal exponent.

Introducing degree heterogeneity to spatial networks with total cost constraint. – To begin our analysis, we first briefly describe the original spatial network model introduced in [20]. Nodes are located in a d -dimensional regular square lattice, where each site i is connected with its $2d$ nearest neighbors. Pairs of sites ij are randomly chosen to receive long-range connections with probability proportional to $r_{ij}^{-\alpha}$, where r_{ij} is the Manhattan distance between sites i and j (*i.e.*, the number of connections separating the nodes in the underlying regular lattice). Long-range connections are added to the system until their total length (cost) $\sum r_{ij}$ reaches a given value C . The exponent α controls the average length of the long-range connections, a smaller α is corresponding to fewer but longer connections, and vice versa. In this model, the most interesting phenomenon is that the average shortest-path length and the navigation steps of the spatial networks can be minimized under a certain value of the parameter α . Specifically, optimal values are found to be $\alpha^* = 3$ in 2-dimensional spatial networks and $\alpha^* = 2$ in 1-dimensional ones. Accordingly, it is concluded that the best transportation condition is obtained with an exponent $\alpha^* = d + 1$, where d is the dimension of the underlying lattice.

The probability P_{ij} for two nodes (i and j) to receive a long-range connection can be described by the probability density function (PDF) $p(r)$ of long-range connections with Manhattan length r . Actually, $p(r)$ is the summation of P_{ij} with $r_{ij} = r$. In d -dimensional space, the number of nodes which have Manhattan distance r from a given node is proportional to r^{d-1} . Therefore, $p(r)$ can be expressed as

$$p(r) \sim r^{d-1} r^{-\alpha} = r^{d-\alpha-1}, \quad (1)$$

which means $p(r) \sim r^{-\delta}$ and $\delta = \alpha + 1 - d$. Since the optimal exponent is $\alpha^* = d + 1$, $\delta^* = (d + 1) + 1 - d = 2$ independently of the spaces' dimension.

Based on this observation, the spatial network in [20] can be equivalently generated by the following procedure:

- 1) N nodes are arranged in a d -dimensional lattice with periodic boundary condition. 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 Manhattan distance.

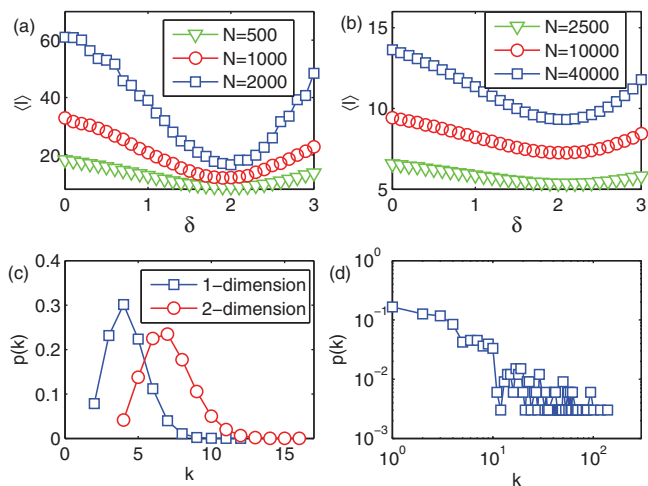


Fig. 1: (Color online) The average shortest-path length $\langle l \rangle$ as a function of δ in spatial networks [20] embedded in 1-dimensional lattices (a) and 2-dimensional lattices (b). Panels (c) and (d) are, respectively, the degree distribution of these network models ($N = 1000$ for 1-dimension and $N = 10000$ for 2-dimension, $\delta = 2$) and the degree distribution of the airline network of USA. Results for the 1-dimensional and 2-dimensional models are averaged over 100 and 10 independent realizations, respectively.

- 2) A certain distance r ($2 \leq r \leq r_{\max}$, where r_{\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 $\sum_{r=2}^{r_{\max}} P(r) = 1$.
- 3) A node i is chosen randomly, and one of the N_r nodes who are at distance r from node i is picked randomly, and an edge between them is created if there is no edge between them yet.
- 4) Update the total cost $\sum r$. Repeat step 2) and 3) until the total cost reaches $C = cN$ (in the following, we fix $c = 10$ for convenience).

We first study the average shortest-path length $\langle l \rangle$ in spatial networks generated by the above procedure (see fig. 1(a) and (b)). As expected, the optimal $\langle l \rangle$ appears under the exponent $\delta = 2$ in both 1-dimensional and 2-dimensional networks, and this feature is independent of the network size. In addition, we observe that the degree in such spatial networks is homogeneously distributed, as shown in fig. 1(c). However, the degree distribution in real systems is far more heterogeneous [19]. For example, the Chinese airline [26], the Italian airline [27] and the worldwide airline [28] networks are found to have truncated power-law degree distribution. We show in fig. 1(d) that for the airline network of USA (USAir) [29] the degree distribution approximately obeys a power law. To capture this feature of real networks, it is necessary to introduce degree heterogeneity in spatial network models and investigate how the topology and function of the networks are

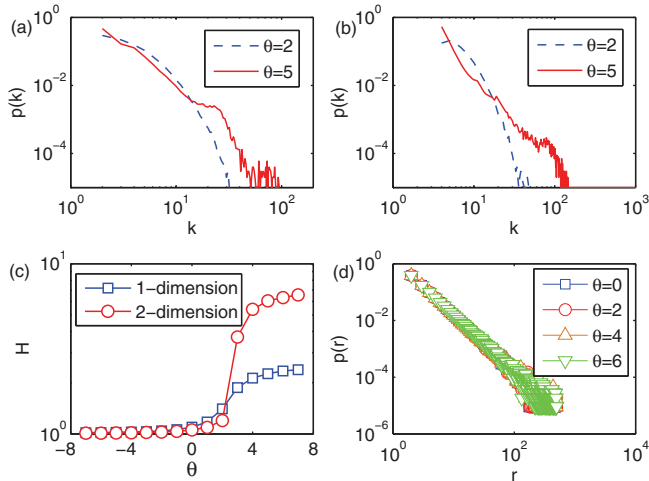


Fig. 2: (Color online) The degree distribution of the (a) 1-dimensional and (b) 2-dimensional modified spatial network model under different θ . Panel (c) shows the index H as a function of the parameter θ and panel (d) shows the link length distributions for different θ in 1-dimensional spatial networks. The parameters are set as $\delta=2$, $N=1000$ for 1-dimension and $N=10000$ for 2-dimension. Results for the 1-dimensional and 2-dimensional models are averaged over 100 and 10 independent realizations, respectively.

affected. Accordingly, we propose a revised spatial network generating procedure. Instead of randomly adding long-range connections to the underlying lattice in step 3), we first calculate a score S_{ij} for each pair of node ij with distance r (generated in step 2)) and add a link between the pair ij with probability proportional to its score. S_{ij} is obtained by

$$S_{ij} = (k_i k_j)^\theta, \quad (2)$$

where θ is a tuneable parameter. Actually, similar weighting method is widely used in many researches [30,31].

In the new step 3), we no longer randomly choose a node i first. Instead, we take all the pairs with distance r into consideration and calculate the S score for each pair. Since there are many long-range links to be added, the calculation of S is a dynamic process. For each new long-range link, S score should be recalculated according to the change of node degree caused by the adding of previous long-range link. Note that only S of the pairs without link will be calculated. If there is already a link between a pair, its S score will be directly set as 0. Finally, S_{ij} is normalized by dividing the sum of all the score S of the pairs with distance r to obtain the probability for adding long-range link.

When $\theta=0$, the generated network reduces to the original model in [20]. When $\theta < 0$, new links will connect the nodes with lower degree, which leads to an even more homogeneous degree distribution. When $\theta > 0$, new links will preferentially attach to the nodes with higher degree, and a heterogenous degree distribution emerges. Two examples are shown in fig. 2(a) and (b). Obviously,

by adjusting the parameter θ , the degree distribution is no longer Poissonian as in the original model. When θ is small (e.g., $\theta=2$), the spatial networks exhibit a power-law-like degree distribution. In this case, the network has several sub-hubs which are connected to each other by some long-distance links. When θ is relatively large (e.g., $\theta=5$), a few super-hubs will emerge and most of the links are connected to such nodes. Therefore, the tail of the degree distribution will become longer and longer as θ increases. The degree heterogeneity can be measured by the index $H = \langle k^2 \rangle / \langle k \rangle^2$. The higher H , the larger the degree heterogeneity. As shown in fig. 2(c), the index H increases with θ , which confirms that the parameter θ can indeed control the degree heterogeneity. In the USAir network, $H=3.46$. It suggests that the degree heterogeneity of real system is far from that of the original network model where $\theta=0$ and $H \approx 1$. Furthermore, the modified spatial model can perfectly preserve the link length distribution as the original model since it does not affect the link length generating process but only rearranges the location of the links. We can see from fig. 2(d) that the link length distributions for different θ are identical as for the case $\theta=0$.

The influence on the network topology and function. – Since the average shortest-path length $\langle l \rangle$ reflects the general transportation ability for a network, we conduct simulations for different values of θ and see how the degree heterogeneity affects the optimizing process of $\langle l \rangle$ based on different δ . The simulations are carried out in both 1-dimensional and 2-dimensional spatial models and the results are reported in fig. 3. When $\theta \leq 0$, the minimal $\langle l \rangle$ is always achieved at $\delta^*=2$. However, when the degree distribution becomes heterogenous (i.e., $\theta > 0$), the optimal exponent δ^* will shift to a smaller value. The results are similar in both 1-dimensional and 2-dimensional spatial networks (see fig. 3(a) and (b), respectively), and imply that a few more longer connections are needed to minimize $\langle l \rangle$ when the network degree distribution is heterogenous.

As shown in fig. 3(c) and (d), for each θ there is an optimal δ^* and a minimal $\langle l^* \rangle$, another interesting aspect is hence studying the relation between θ and $\langle l^* \rangle$. Figures 3(e) and (f) show that introducing the degree heterogeneity can indeed further decrease $\langle l^* \rangle$. However, if one keeps enlarging θ , $\langle l^* \rangle$ will start to increase. We can see from fig. 3(e) and (f) that optimal θ^* for $\langle l^* \rangle$ exist in both 1-dimensional and 2-dimensional spatial networks. Interestingly, these optimal θ^* are independent of the network size but dependent on the network dimension. Specifically, $\theta^*=2$ in 1-dimensional spatial networks and $\theta^*=3$ in 2-dimensional spatial networks, which leads us to the conjecture that the optimal value is obtained at $\theta^*=1+d$. We also consider the greedy navigation algorithm used in ref. [20] to model the case when only local information of the network is known during the transport process. A similar optimal exponent shifting phenomenon

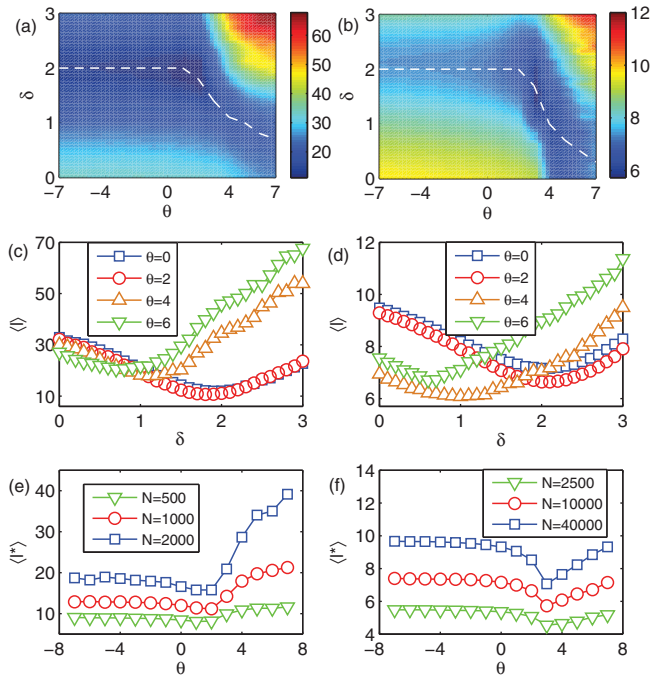


Fig. 3: (Color online) The dependence of $\langle l \rangle$ on δ under different θ in spatial networks generated from 1-dimensional lattices with $N = 1000$ (a) and 2-dimensional lattices with $N = 10^4$ (b). The white lines mark the optimal δ^* under different θ . The detail curves of $\langle l \rangle$ against δ for different values of θ in 1-dimensional and 2-dimensional spatial networks are shown in (c) and (d), respectively. Panels (e) and (f) show the relation between θ and the minimal $\langle l^* \rangle$ in 1-dimensional and 2-dimensional spatial networks. The results for the 1-dimensional and 2-dimensional models are averaged over 100 and 10 independent realizations, respectively.

is observed and the minimum navigation steps can further decrease due to the degree heterogeneity.

Another important aspect to consider in spatial networks is how the network function (dynamics) can be enhanced. Specifically, refs. [24,25] have revealed that there is an optimal link length distribution for the network synchronizability. Under the framework of master stability analysis, the synchronizability of an undirected network can be quantified by the eigenvalue ratio of the corresponding Laplacian matrix of this network, namely $R = \lambda_N / \lambda_2$, where λ_N and λ_2 are, respectively, the largest and the smallest non-zero eigenvalues of the Laplacian matrix [32]. Generally, the smaller the value of R , the stronger the network synchronizability. In [24,25], starting from 1-dimensional lattice, the authors find the optimal length distribution of links (power-law distribution with exponent $\delta = 1.5$) to add to the network for minimizing R . However, these previous network models also yields homogeneous degree distribution. Here we are interested in how the degree heterogeneity affects the synchronizability of spatial networks. The results are reported in fig. 4. Obviously, the optimal δ^* for R shifts from $\delta^* = 1.5$ to a smaller value as θ increases. However, the optimal

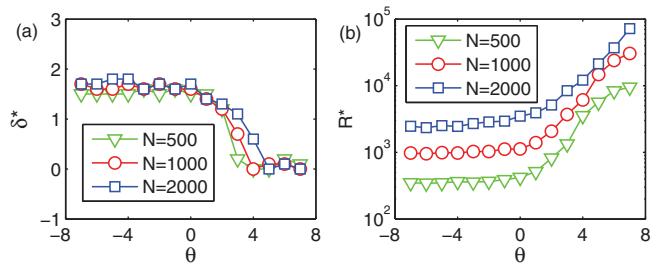


Fig. 4: (Color online) (a) The dependence of the optimal δ^* for synchronizability R on θ . (b) The optimal synchronizability R^* as a function of θ . The spatial networks are built on 1-dimensional lattices. The results are averaged over 100 independent realizations.

synchronizability R^* cannot be further enhanced by introducing the degree heterogeneity as shown in fig. 4(b). It is well known that degree heterogeneity will weaken the synchronizability while a small average shortest-path length is generally favorable for synchronizability [32]. As we discussed above, introducing the degree heterogeneity can further decrease the average shortest path. However, the decreased average shortest-path length in the current case is not sufficient to overcome the effect of degree heterogeneity, and hence R^* keeps increasing with θ .

Conclusion. – To summarize, we have proposed a method to introduce degree heterogeneity in spatial networks without destroying the power-law link length distribution of the original model [20]. We have investigated its effect on network topology and function. When optimizing the average shortest-path length and the network synchronizability, we find that the best power-law exponent for the link length distribution shifts to a smaller value than that of the original model. Additionally, the average shortest-path length can further decrease due to the degree heterogeneity while the network synchronizability is weakened if the degree distribution is heterogenous.

From the practical point of view, the original spatial network model [20] can explain the power-law connection length distribution with a given exponent in some real systems. By building the relation between the optimal exponent and the degree heterogeneity, our model can extend such explanation of real transportation networks to a wider range of connection length distribution which is found to be existing by many empirical studies [19]. Our model can not only better understand and model real transportation systems, but also help to design efficient ones.

We would like to thank YI-CHENG ZHANG, GIULIO CIMINI, MATUS MEDO and CHI HO YEUNG for helpful suggestions. This work is supported by the Swiss National Science Foundation (No. 200020-132253).

REFERENCES

- [1] GUIMERA R., MOSSA S., TURTSCHI A. and AMARAL L. A. N., *Proc. Natl. Acad. Sci. U.S.A.*, **102** (2005) 7794.
- [2] YOON S.-H., JEONG H. and BARABASI A.-L., *Proc. Natl. Acad. Sci. U.S.A.*, **99** (2002) 13382.
- [3] WANG P., GONZALEZ M., HIDALGO C. A. and BARABASI A.-L., *Science*, **324** (2009) 1071.
- [4] ALBERT R., ALBERT I. and NAKARADO G. L., *Phys. Rev. E*, **69** (2004) 025103.
- [5] LIBEN-NOWELL D., NOVAK J., KUMAR R., RAGHAVAN P. and TOMKINS A., *Proc. Natl. Acad. Sci. U.S.A.*, **102** (2005) 11623.
- [6] WATTS D. J. and STROGATZ S. H., *Nature*, **393** (1998) 440.
- [7] SEN P., BANERJEE K. and BISWAS T., *Phys. Rev. E*, **66** (2002) 037102.
- [8] CARMÍ S., CARTER S., SUN J. and BEN-AVRAHAM D., *Phys. Rev. Lett.*, **102** (2009) 238702.
- [9] CARTOZO C. C. and LOS RIOS P. D., *Phys. Rev. Lett.*, **102** (2009) 238703.
- [10] LI D., KOSMIDIS K., BUNDE A. and HAVLIN S., *Nat. Phys.*, **7** (2011) 481.
- [11] MEDO M. and SMREK J., *Eur. Phys. J. B*, **63** (2008) 273.
- [12] BOGUNA M., PASTOR-SATORRAS R., DIAZ-GUILERA A. and ARENAS A., *Phys. Rev. E*, **70** (2004) 056122.
- [13] XULVI-BRUNET R. and SOKOLOV I. M., *Phys. Rev. E*, **66** (2002) 026118.
- [14] XIE Y.-B., ZHOU T., BAI W.-J., CHEN G.-R., XIAO W.-K. and WANG B.-H., *Phys. Rev. E*, **75** (2007) 036106.
- [15] SCHNEIDER C. M., MOREIRA A. A., ANDRADE J. S. jr., HAVLIN S. and HERRMANN H. J., *Proc. Natl. Acad. Sci. U.S.A.*, **108** (2011) 3838.
- [16] HERRMANN H. J., SCHNEIDER C. M., MOREIRA A. A., ANDRADE J. S. jr. and HAVLIN S., *J. Stat. Mech.* (2011) P01027.
- [17] MORAIS P. A., ANDRADE J. S. jr., NASCIMENTO E. M. and LYRA M. L., *Phys. Rev. E*, **84** (2011) 041110.
- [18] WU Z.-X. and HOLME P., *Phys. Rev. E*, **84** (2011) 026106.
- [19] BARTHELEMY M., *Phys. Rep.*, **499** (2011) 1.
- [20] 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.
- [21] HU Y., WANG Y., LI D., HAVLIN S. and DI Z., *Phys. Rev. Lett.*, **106** (2011) 108701.
- [22] YANG H., NIE Y., ZENG A., FAN Y., HU Y. and DI Z., *EPL*, **89** (2010) 58002.
- [23] LI Y., ZHOU D., HU Y., ZHANG J. and DI Z., *EPL*, **92** (2010) 58002.
- [24] ZENG A., ZHOU D., HU Y., FAN Y. and DI Z., *Physica A*, **390** (2011) 3962.
- [25] BREDE M., *Phys. Rev. E*, **81** (2010) 025202(R).
- [26] LI W. and CAI X., *Phys. Rev. E*, **69** (2004) 046106.
- [27] GUIDA M. and MARIA F., *Chaos, Solitons Fractals*, **31** (2007) 527.
- [28] GUIMERA R., MOSSA S., TURTSCHI A. and AMARAL L. A. N., *Proc. Natl. Acad. Sci. U.S.A.*, **102** (2005) 7794.
- [29] Pajek Datasets, available at <http://vlado.fmf.uni-lj.si/pub/networks/pajek/data/gphs.htm>.
- [30] MOREIRA A. A., ANDRADE J. S. jr., HERRMANN H. J. and INDEKEU J. O., *Phys. Rev. Lett.*, **102** (2009) 018701.
- [31] HOYBERGHS H., SCHAEYBROECK B. V., MOREIRA A. A., ANDRADE J. S. jr., HERRMANN H. J. and INDEKEU J. O., *Phys. Rev. E*, **81** (2010) 011102.
- [32] ARENAS A., DIAZ-GUILERA A., KURTHS J., MORENO Y. and ZHOU C.-S., *Phys. Rep.*, **469** (2008) 93.