Vol 17 No 10, October 2008
1674-1056/2008/17(10)/3580-08

Chinese Physics B

© 2008 Chin. Phys. Soc.
and IOP Publishing Ltd

Connectivity correlations in three topological spaces of urban bus-transport networks in China*

Chen Yong-Zhou(陈永洲)[a], Fu Chun-Hua(付春花)[b], Chang Hui(常 慧)[b], Li Nan(李 南)[a], and He Da-Ren(何大韧)[b]†

[a] *College of Economics and Management, Nanjing University of Aeronautics and Astronautics, Nanjing* 210016, *China*
[b] *College of Physics Science and Technology, Yangzhou University, Yangzhou* 225002, *China*

(Received 26 February 2008; revised manuscript received 17 March 2008)

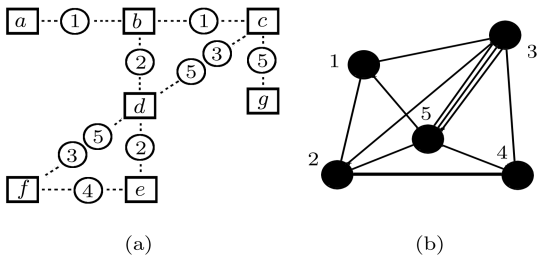
In this paper, an empirical investigation is presented, which focuses on unveiling the universality of connectivity correlations in three spaces (the route space, the stop geographical space and bus-transferring space) of urban bus-transport networks (BTNs) in four major cities of China. The underlying features of the connectivity correlations are shown in two statistical ways. One is the correlation between the (weighted) average degree of all the nearest neighbouring vertices with degree k, ($K^w_{nn}(k)$) $K_{nn}(k)$, and k, and the other is the correlations between the assortativity coefficient r and, respectively, the network size N, the network diameter D, the averaged clustering coefficient C, and the averaged distance $\langle l \rangle$. The obtained results show qualitatively the same connectivity correlations of all the considered cities under all the three spaces.

Keywords: connectivity correlation, bus-transport network, universality
PACC: 0520, 0250

1. Introduction

Metropolitan mass transportation networks can be mentioned as an important type of the complex networks.[1−12] Among them, the urban bus-transport system (BTS)[4−11] has attracted our interest, for the bus is still a main transport tool in most of Chinese cities. Hence in this paper we report on an empirical investigation on urban bus-transport networks (BTNs) in four major cities of China. Our attention is paid to a statistical measurement, which may be universal for the BTNs in many Chinese cities from some possible network viewpoints.

From different viewpoints, one can map a BTS into BTNs in three topological spaces. In some publications, BTNs are studied in two topological spaces: Space P and Space L.[3,7−11] In Space P, one vertex represents one stop, and one edge is connected between a pair of stops when at least one bus route provides direct service. In other words, an edge means that passengers can take at least one route for a direct travel between two stops. If passengers have to exchange routes then the pair of stops is connected by more than one edge. Therefore Space P may be called bus-transferring space. However in Space L an edge between two vertices exists if they are consecutive stops on at least one route although vertex definition is the same. In some sense, the network in Space L looks like a bus route map, thus Space L may be addressed as a stop geographical space. Some other publications propose an additional topological space: Space R[4,6] (or Space C in Ref.[5]). Different from Space P and Space L, routes in a BTN are defined as vertices and common stops (intersection stops, common ends or stops in overlapped intervals of routes, etc.) are defined as edges. Here an edge means that passengers can directly transfer between two routes without involving an additional route. Space R thus may be called route space. A schematic showing BTNs in these topological spaces is presented in Figs.1(b), 1(c) and 1(d) respectively.



*Project supported by the National Natural Science Foundation of China (Grant Nos 70671089 and 10635040) and the foundation of XM06-142.
†E-mail: darendo10@yahoo.com.cn

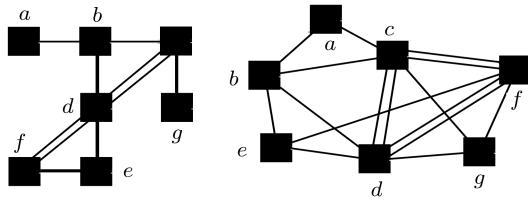


Fig.1. (a) Schematic diagram of a urban bus-transport system with five routes and seven stops. Numbers denote routes and letters denote stops. (b) The corresponding route space network (Space R), where edges represent the common stops. Note that the three edges between vertices 3 and 5 denote that there are three common stops in route 3 and route 5. However, in a traditional way, those three edges are regarded as one edge, whose 'edge weight' is the common stop number 3. Similar situations in (c) and (d) can be Seen in the same way. (c) The corresponding stop geographical space network (Space L). (d) The corresponding bus-transferring space network (Space P).

2. Statistical measurements

Some averaged statistical measurements, such as averaged distance $\langle l\rangle$, averaged clustering coefficient C, averaged degree $\langle k\rangle$, and/or some maximal statistical measurement (e.g. the network diameter D) have been used for characterizing the topology of networks. The definitions of the above measurements can be found in Ref.[13]. The empirical data of these statistical network properties, together with the data describing the city and the city BTN basic features, are presented in Table 1.

Obviously the values of the averaged or maximal statistical measurements are dependent on network size, therefore they are not universal for BTSs in many cities. One has to consider some statistical distribution measurements to see whether the distribution function forms are universal or not. The well-known degree distribution might be a good candidate for the universal measurement. Degree of vertex i, k_i, is defined as the number of its adjacent vertices and can be expressed as

$$k_i = \sum\nolimits_{j\in\Gamma(i)} a_{ij}, \tag{1}$$

where $\Gamma(i)$ represents a set of the adjacent vertices of vertex i, and a_{ij} denotes the element of the adjacency matrix. However, some empirical investigation results strongly suggest that although degree distribution function form may be universal for different city BTNs, it is not universal for different network definitions (different 'Spaces'). For example, with the same network definition ('Space P'), Sienkiewicz and Holyst have shown that the degree distributions of eight Polish BTNs follow the same exponential decay function.[3] Xu *et al*[6] and Chen *et al*[7] obtained the same conclusion in an empirical investigation on BTNs of three or four main cities in China.[6,7] However, if the network definition is changed ("Space L"), the degree distributions are recognized as the approximate power laws as reported in Refs.[3, 6].

Table 1. Empirical data.

		Hangzhou	Nanjing	Beijing	Shanghai
	A	683.00	975.82	16807.80	6340.50
	I	1.29	2.66	14.23	10.47
	M	150	252	572	968
	N	827	1764	4199	4374
Space R	D	4	6	4	5
	C	0.65	0.59	0.57	0.54
	$\langle l\rangle$	1.76	1.99	1.79	2.20
	$\langle k\rangle$	47.04	53.68	136.41	89.95
Space L	D	38	47	66	28
	C	0.12	0.10	0.15	0.17
	$\langle l\rangle$	9.20	13.70	11.12	7.52
	$\langle k\rangle$	3.21	2.93	3.25	4.08
Space P	D	5	7	5	6
	C	0.78	0.76	0.78	0.74
	$\langle l\rangle$	2.60	2.74	2.53	2.89
	$\langle k\rangle$	33.03	50.65	91.45	57.29

A: city area (km^2); I: population (million); M: total number of bus routes; N: total number of bus stops; D: diameter of the network; C: average clustering coefficient; $\langle l\rangle$: average distance; $\langle k\rangle$: average degree. The data were collected in Nov. 2006.[14]

In spite of many possible choices, we decide to try the statistical measurements describing the degree connectivity correlations. We believe that for all the Chinese BTNs in all the three spaces the smaller vertices have a larger probability to connect other smaller vertices, while larger vertices have a bigger probability to connect other larger vertices than the smaller vertices. Therefore the degree connectivity correlations may show similar function forms in many BTNs in all the spaces. To the best of our knowledge, there are two ways of describing the degree connectivity correlations. Pastor-Satorras *et al*[15] proposed the first statistical measurement. They defined the average degree of the nearest neighbouring vertex of vertex i, $k_{nn,i}$, as

$$k_{nn,i} = \frac{1}{k_i} \sum_{j \in \Gamma(i)} k_j. \tag{2}$$

The average degree of all the nearest neighbouring vertices with degree k, $K_{nn}(k)$, is defined as[15]

$$K_{nn}(k) = \frac{1}{N_k} \sum_{k_i = k} k_{nn,i}, \tag{3}$$

where N_k is the number of vertices with degree k.

Newman[16] has proposed a more accurate and convenient parameter, assortativity coefficient, for describing degree-degree correlation. Assortativity coefficient is denoted by r, which lies in a range: $-1 \leq r \leq 1$. If $r > 0$, the network shows a positive nearest-neighbouring degree-degree correlation, also named 'Assortative Mixing'. If $r < 0$, the network indicates a negative nearest-neighbouring degree–degree correlation, also named 'Disassortative Mixing'. The formula, from which Newman has proposed to calculate the assortativity coefficient in an undirected network,[16] is

$$r = \frac{M^{-1} \sum_{\phi} \left(\prod_{i \in F(\phi)} k_i \right) - \left(\frac{M^{-1}}{2} \sum_{\phi} \left(\sum_{i \in F(\phi)} k_i \right) \right)^2}{\frac{M^{-1}}{2} \sum_{\phi} \left(\sum_{i \in F(\phi)} k_i^2 \right) - \left(\frac{M^{-1}}{2} \sum_{\phi} \left(\sum_{i \in F(\phi)} k_i \right) \right)^2}, \tag{4}$$

where $F(\phi)$ denotes a set of two vertices connected by the ϕth link and M is the total number of links in the network.

To the best of our knowledge, there is no such a systematic empirical investigation on the two statistical measurements describing degree-degree correlation in all the three spaces of BTNs. Reference [3] is the only one that is devoted to the empirical investigation on assortativity in Space P and Space L of some public transport networks (PTSs). One conclusion in Ref.[3] is that the assortativity coefficient r basically shows a positive dependence on network size N in Space P. The small public transport systems containing less than 500 vertices basically show a negative assortativity correlation, while the larger ones with the vertex number larger than 500 basically show a positive assortativity correlation. On the contrary, values of the assortativity coefficient in Space L are independent of network size, and they are always positive.[3] It will be interesting to perform the above-mentioned systematic investigation in some Chinese BTNs to see the measurement behaviours in all the three spaces and to find whether the properties in Chinese BTNs are exactly the same as those in Polish transport networks.

In Refs.[3, 5, 7], the elements of adjacency matrixes of the BTNs take a value of either 1 (two vertices are connected) or 0 (otherwise). This only shows that, between two vertices in BTNs, an edge is either present or not. Reference [6] might be the only one that is for investigating BTNs, where the elements of adjacency matrix could take integers greater than 1 in Space R of BTNs, but in Space L and Space P only binary matrices were discussed too. Actually, in transportation networks the edges are different, characterized by different capacities or amounts of traffic. The information should be attached to an edge as its weight.[17,18] In such weighted BTNs the elements of adjacency matrixes, w_{ij}, can take an integer, which is equal to the edge weight, and may be greater than 1.

Corresponding to $k_{nn,i}$, in weighted BTNs, the weighted average degree of the nearest neighbouring vertices of vertex i is defined as

$$k^w_{nn,i} = \frac{1}{s_i} \sum\nolimits_{j \in \Gamma(i)} w_{ij} k_j, \tag{5}$$

where s_i, denoted by 'the vertex i's strength', is defined as

$$s_i = \sum\nolimits_{j\in\Gamma(i)} w_{ij}. \tag{6}$$

The weighted average degree of all the nearest neighbouring vertices with degree k, $K^w_{nn}(k)$, can be defined as[18,19]:

$$K^w_{nn}(k) = \frac{1}{N_k}\sum\nolimits_{k_i=k} k^w_{nn,i}. \tag{7}$$

Reference [20] may be the only reference where a definition of weighted assortativity coefficient was suggested to be

$$r^w = \frac{H^{-1}\sum_\phi \left(w_\phi \prod_{i\in F(\phi)} k_i\right) - \left(\frac{H^{-1}}{2}\sum_\phi \left(w_\phi \sum_{i\in F(\phi)} k_i\right)\right)^2}{\frac{H^{-1}}{2}\sum_\phi (w_\phi \sum_{i\in F(\phi)} k_i^2) - \left(\frac{H^{-1}}{2}\sum_\phi \left(w_\phi \sum_{i\in F(\phi)} k_i\right)\right)^2}, \tag{8}$$

where w_ϕ denotes the weight of the ϕth link and H is the total weight of all links in the network.[20] All the three weighted quantities, s_i, $K^w_{nn}(k)$, and r^w, are equal to their un-weighted ones if all the nonzero w_{ij} are set to be 1.[18−20]

3. Empirical investigations on degree-degree correlations

As shown in Fig.2, large fluctuations in $K_{nn}(k)$-k and $K^w_{nn}(k)$-k curves bring on the difficulty in recognizing the accurate function forms. In most of the cases $K_{nn}(k)$ or $K^w_{nn}(k)$ basically increases with k and it indicates positive nearest-neighbouring degree-degree correlations. However, in some intervals of parameter k, it is hard to determine whether the unweighted or weighted correlations are positive, negative or uncorrelated. To solve the problem, we propose computing the corresponding cumulative functions, which should greatly suppress the statistical fluctuations.

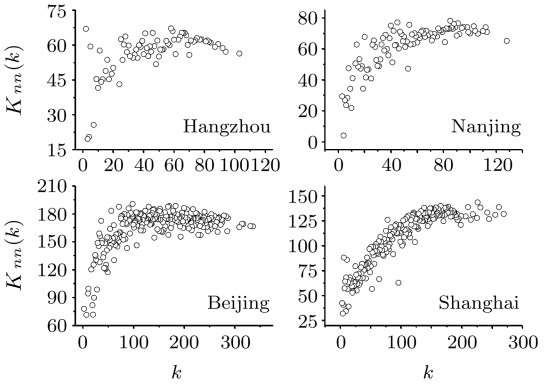
In Fig.2 one obtains the conclusion that all the cumulative relationships can be well fitted by Gaussian distribution functions. The function form can be written as

$$y \propto \mathrm{e}^{-(x-x_\mathrm{c})^2}, \tag{9}$$

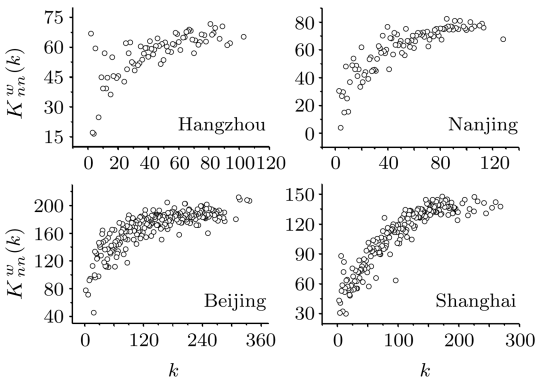
where x_c is a constant. By calculating the derivative (the quasi-continuous approximation), one knows the original non-cumulated $K_{nn}(k)$-k, or $K^w_{nn}(k)$-k function form, which is embedded in fluctuations, as

$$y \propto (x - x_c)\mathrm{e}^{-(x-x_\mathrm{c})^2}. \tag{10}$$

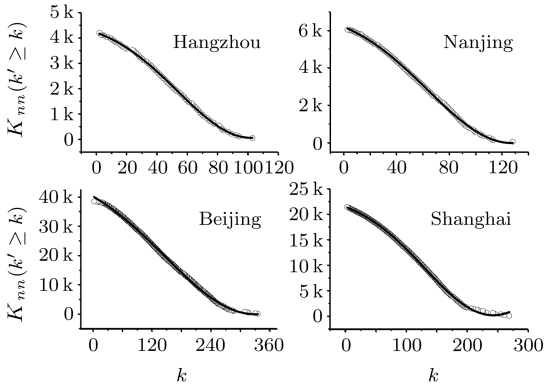
The results indicate that all the weighted $K^w_{nn}(k) - k$ and un-weighted $K_{nn}(k) - k$, or their corresponding cumulated and un-cumulated curves of the four major cities' BTNs in Space R, Space L, and Space P can be described by the same function form. The findings of the BTNs' underlying features are new.



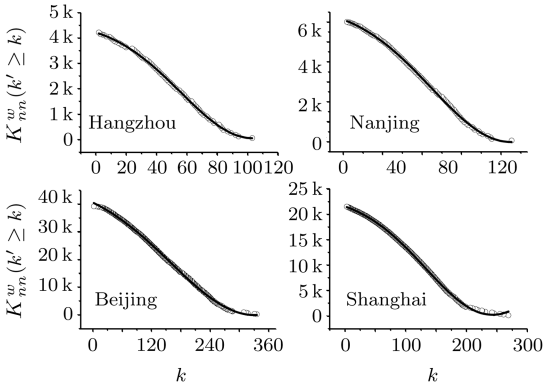
(a) The $K_{nn}(k) - k$ curves in Space R



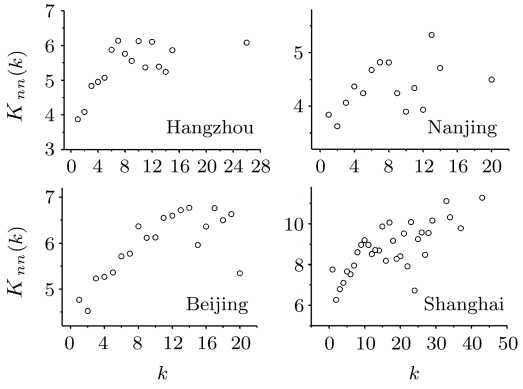
(b) The $K^w_{nn}(k) - k$ curves in Space R



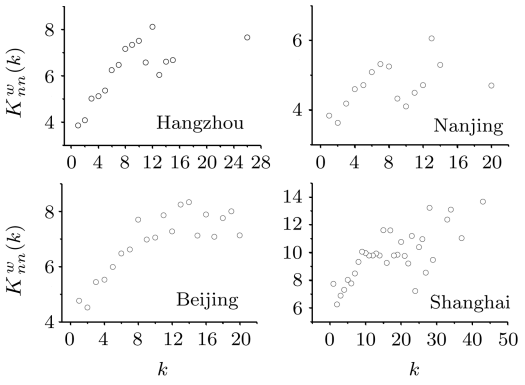
(c) The cumulated $K_{nn}(k) - k$ curves in Space R



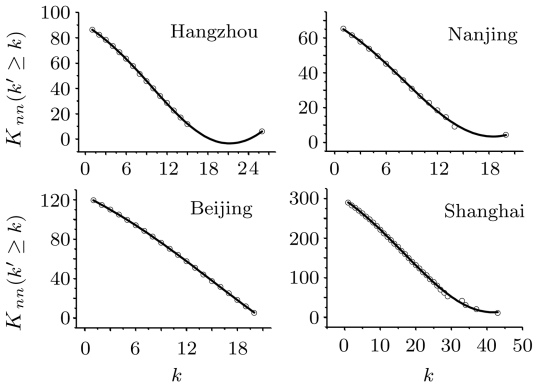
(d) The cumulated $K^w_{nn}(k) - k$ curves in Space R



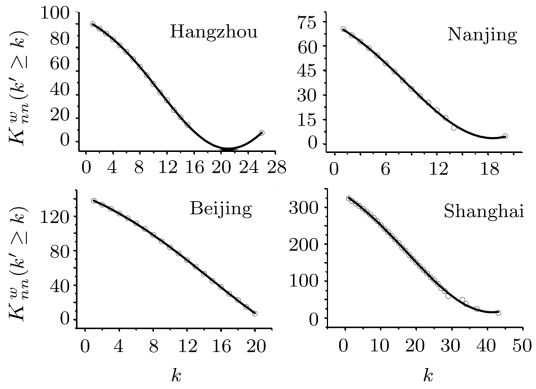
(e) The $K_{nn}(k) - k$ curves in Space L



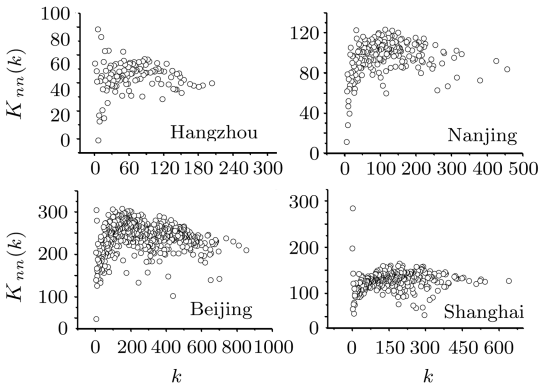
(f) The $K^w_{nn}(k) - k$ curves in Space L



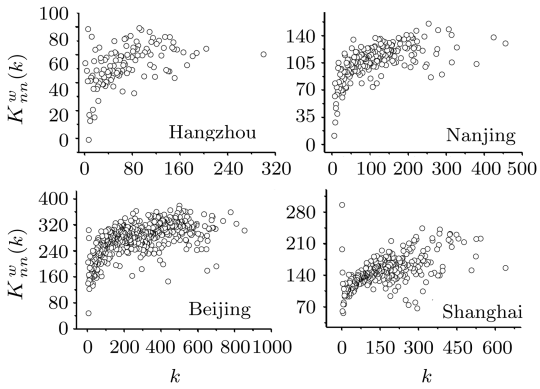
(g) The cumulated $K_{nn}(k) - k$ curves in Space L



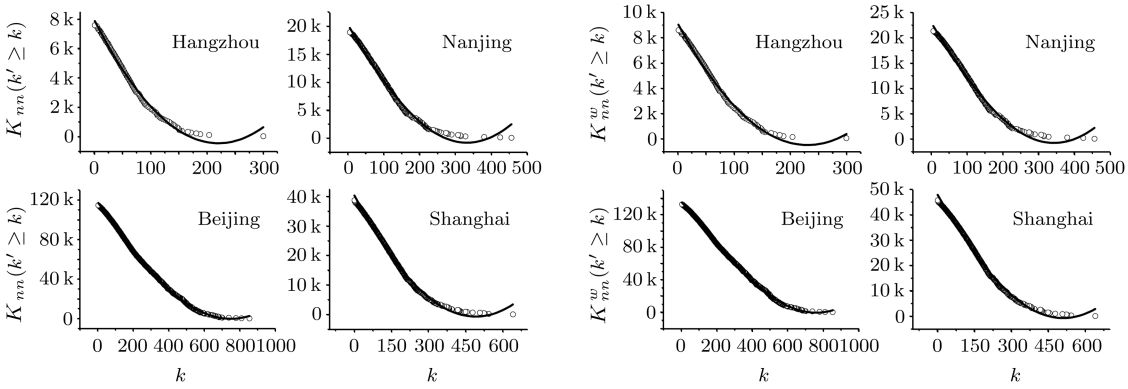
(h) The cumulated $K^w_{nn}(k) - k$ curves in Space L



(i) The $K_{nn}(k) - k$ curves in Space P



(j) The $K^w_{nn}(k) - k$ curves in Space P



(k) The cumulated $K_{nn}(k) - k$ curves in Space P (l) The cumulated $K^w_{nn}(k) - k$ curves in Space P

Fig.2. The cumulated and the un-cumulated un-weighted $K_{nn}(k) - k$ and weighted $K^w_{nn}(k) - k$ curves of the four major cities' BTNs in Space R, Space L, and Space P in China. In subfigures (c), (d), (g), (h), (k), and (l) the solid lines represent their Gaussian-like function fitting results.

The results of the un-weighted and weighted assortativity, i.e. r and r^w, are listed in Table 2, showing $r^w > r > 0$ for the four major cities in China. This is in agreement with conclusions reported in Refs.[3, 20].

In order to verify the conclusion that the small transport systems containing less than 500 vertices basically show a negative assortativity, while the larger ones (vertex number larger than 500) basically show a positive assortativity[3], we report on an empirical investigation of r^w and r of the BTNs in 2006 in a small city Yangzhou in China[21] (there were 506 stops and 52 bus routes in that period). The results are also presented in Table 2. As can be seen, the small city BTN shows a negative un-weighted assortativity only in Space P and Space R, but it shows a positive weighted assortativity in all the three spaces.

Table 2. Un-weighted and weighted assortativity values for BTNs in four major cities in China, i.e. Hangzhou, Nanjing, Beijing and Shanghai in 2006, and for the BTN in a small city, Yangzhou, in 2006 for comparison, in Space R, Space L, and Space P respectively.

	Space R		Space L		Space P	
	r	r^w	r	r^w	r	r^w
Hangzhou	0.1002	0.2328	0.1449	0.2256	0.0186	0.1256
Nanjing	0.2449	0.3604	0.1129	0.1217	0.0480	0.1586
Beijing	0.0403	0.1793	0.1560	0.2092	0.0328	0.1720
Shanghai	0.3342	0.3849	0.1441	0.2036	0.0533	0.1988
Yangzhou	–0.0642	0.0552	0.1079	0.2058	–0.0650	0.0630

4. Empirical investigations on dependence of r on some network topology parameters

Although the values of the averaged or maximal statistical measurements, such as the network size N, the network diameter D, the averaged clustering coefficient C, and the averaged distance $\langle l \rangle$, are not universal for BTNs, it is interesting to verify whether the assortativity, r, shows a universal correlation function form on them for the five cities' BTNs in all the three spaces.

In Fig.3 one again feels hard to recognize the exact function forms of the dependence of r on network topology parameters due to the large fluctuations. We hesitated whether to use the cumulative method, for now we have only 5 data points in each curve, which is far from the quasi-continuous assumption. However, we finally decide to do it only as an argument. It is shown that the cumulative $r - N$ curves can be well fitted by quadratic equations. This implies that the original un-cumulative $r - N$ function forms are linear. This conclusion is different from that reported in Ref.[3] where the authors summarized that the values of the assortativity coefficient in Space L were inde-

pendent of the network size. We think that the difference may indicate some differences in characteristic between Polish and Chinese cities. Another possibility is that the authors of Ref.[3] could not recognize the embedded rule because they did not use the cumulative method. The cumulative $r-D$, $r-C$ and $r-\langle l\rangle$ curves can also be well fitted by quadratic equations. The original un-cumulative $r-D$, $r-C$ and $r-\langle l\rangle$ function forms are also roughly linear.

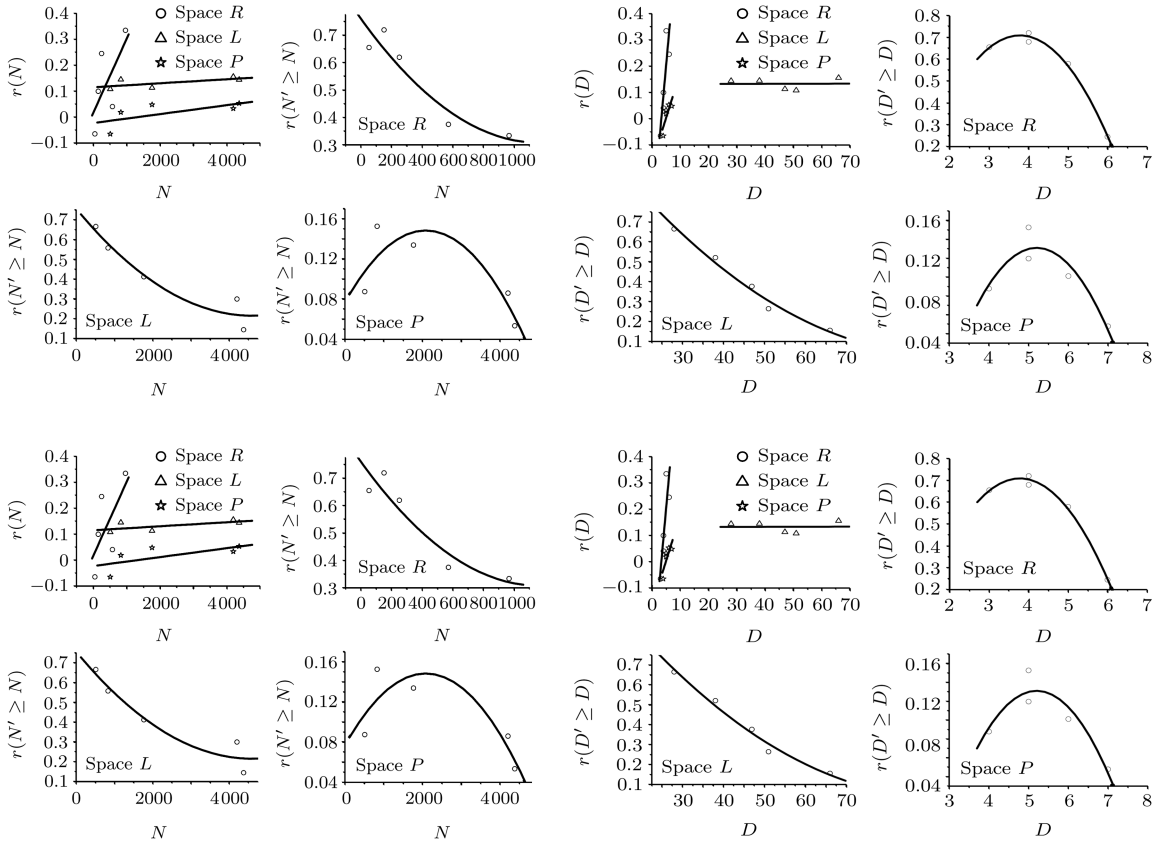


Fig.3. The cumulated and un-cumulated correlations between r and the network size N, the network diameter D, the averaged clustering coefficient C, and the averaged distance $< l >$ in the five city BTNs in all the three spaces. The solid lines show the fitting results.

5. Conclusion

Our empirical investigation on five Chinese BTNs strongly suggests that the degree-degree correlation functions uncover their universality in all the cities in all the discussed three spaces, i.e. route space, stop geographical space and bus-transferring space. All the weighted $K^w_{nn}(k)-k$ and un-weighted $K_{nn}(k)-k$ curves of the four major cities' BTNs in Spaces R, Space L, and Space P in China can be qualitatively described by the same Gaussian-like distribution function. The data of the five BTNs strongly suggest that all the $r-N$, $r-C$, $r-D$ and $r-\langle l\rangle$ function forms are linear. For the universality, we can have only a qualitative and very rough understanding, which has been introduced at the beginning of the text. The question about the quantitative understanding is still open.

References

[1] Latora V and Marchiori M 2002 *Physica* A **314** 109
[2] Angeloudis P and Fisk D 2006 *Physica* A **367** 553
[3] Sienkiewicz J and Holyst J A 2005 *Phys. Rev.* E **72** 046127
[4] Kurant M and Thiran P 2006 *Phys. Rev.* E **74** 036114
[5] Ferber C V, Holovatch T, Holovatch Y and Palchykov V 2007 *Physica* A **380** 585
[6] Xu X, Hu J, Liu F and Liu L 2007 *Physica* A **374** 441
[7] Chen Y Z, Li N and He D R 2007 *Physica* A **376** 747
[8] Zhang P P, Chen K, He Y, Zhou T, Su B B, Jin Y, Chang H, Zhou Y P, Sun L C, Wang L C, Wang B H and He D

R 2006 *Physica* A **360** 599

[9] Su B B, Chang H, Chen Y Z and He D R 2007 *Physica* A **379** 291

[10] Chang H, Su B B, Zhou Y P and He D R 2007 *Physica* A **383** 687

[11] Zhao J, Di Z and Wang D 2005 *Complex Systems and Complexity Science* **2** 45 (in Chinese)

[12] Zhao W and He H S 2006 *Acta Phys. Sin.* **55** 3906 (in Chinese)

[13] Albert R and Barabási A L 2002 *Rev. Mod. Phys.* **74** 47
Newman M E J 2003 *SIAM Review* **45** 167

[14] Hangzhou: http://www.hzbus.com.cn/; Nanjing: http://www.njkgc.cn/; Beijing: http://www.bjbus.com/; Shanghai: http://www.jt.sh.cn/.

[15] Pastor-Satorras R, Vazquez A and Vespignani A 2001 *Phys. Rev. Lett.* **87** 258701
Pastor-Satorras R and Rubi J M 2001 *Phys. Rev.* E **64** 016103

[16] Newman M E J 2002 *Phys. Rev. Lett.* **89** 208701
Newman M E J 2003 *Phys. Rev.* E **67** 0261126

[17] Serrano M A, Boguna M and Pastor-Satorras R 2006 *Phys. Rev.* E **74** 055101(R)

[18] Barthélemy M, Barrat A, Pastor-Satorras R and Vespignani A 2005 *Physica* A **346** 34

[19] De Montis A, Barthélemy M, Chessa A and Vespignani A 2007 *Env. and Planning* B **34** 905

[20] Leung C C and Chau H F 2007 *Physica* A **378** 591

[21] http://www.yzren.com/life-busstation.asp.