

# 基于实际拓扑结构的交通网络客运流量的理论研究

唐芙蓉, 杨先清

(中国矿业大学理学院, 江苏 徐州 221008)

[摘要] 针对 deMoura 等人提出因特网信息流的费米-狄拉克统计模型, 我们提出用玻色-爱因斯坦统计模型来研究一般交通客运网络中的客流分布问题. 将交通客运网络抽象成一个与外界有粒子和能量交换的热力学系统, 其中粒子代表乘客, 适当地定义能级, 分析得出交通网络中的每个站点平均客流的解析表述. 然后以中国铁路客运网络实际拓扑为基础, 用随机行走模型模拟分析交通网络中的客流分布情况, 模拟结果显示, 无论初态如何, 经过足够长时间后, 各站点的客流分布趋于平衡, 平衡时各站点的客流量与其点权正相关, 模拟客流分布曲线与理论结果基本一致.

[关键词] 复杂网络, 中国铁路网络, 交通客流

[中图分类号] U491.14 [文献标识码] A [文章编号] 1001-4616(2010)03-0045-06

## Theoretical Research of Passenger Flow Based on Actual Transport Network

Tang Furong, Yang Xianqing

(School of Sciences, University of Mining & Technology of China, Xuzhou 221008, China)

**Abstract** Comparing with deMoura's model of Fermi-Dirac statistics for traffic in a network, we studied the statistical properties of the general passenger transport network with Bose-Einstein distribution. A transport of passenger network can be modeled as a thermodynamics system exchanging particles and energy with its environment where passengers are considered as particles. With appropriately defined energy-level structure, the analytical expressions of the passenger flow of the transport network can be obtained. Based on the empirical properties of Chinese railway network, we propose an idealized model in which large numbers of passengers move randomly from one station to another in this railway network. Numerical simulations show that whatever the early state is, in the  $t \rightarrow \infty$  limit the network reaches a stationary state with the passenger flow in proportion with the node weight. The numerical simulations results are shown to be nearly in agreement with the analytical expressions.

**Key words** complex networks, Chinese railway network, passenger flow of transport network

自然界中存在大量的复杂系统都可以用复杂网络<sup>[1-6]</sup>描述, 如互联网<sup>[1,7,8]</sup>、食物链网<sup>[9]</sup>、新陈代谢网<sup>[10,11]</sup>、蛋白质相互作用网<sup>[12]</sup>、交通网络等等. 近几年, 由于网络的小世界效应和无标度特性引起物理学界广泛关注, 特别是网络拓扑结构对复杂网络上的动力学行为的影响成为研究焦点之一.

最近, 人们提出很多动力学模型<sup>[13-17]</sup>研究复杂网络的流量问题. 其中, deMoura<sup>[18]</sup>规定因特网中所有的节点最多被一个粒子(数据包)占有, 提出了复杂网络的费米-狄拉克统计模型研究因特网的信息流问题. 在此基础上, Renato Gernano 和 deMoura<sup>[19]</sup>规定每个节点上都对应有其所能支持的最大数据包数  $m_i$  ( $\geq 1$ ), 给出了  $m_i = 1, 2, 3$  的 3 种情况的解析解.

实际交通客运网络如航空网络<sup>[20-24]</sup>、铁路网络<sup>[25-28]</sup>和公交系统网络<sup>[29]</sup>等也是复杂网络, 具有固定节点和结构, 网络中客站是节点, 乘客看成粒子. 与因特网数据传输过程不同, 实际客运网中每个站点不可能

收稿日期: 2010-01-15

基金项目: 中国矿业大学青年科研基金(0K061024).

通讯联系人: 杨先清, 教授, 博士生导师, 研究方向: 统计物理学, 凝聚态理论. E-mail: xianqiang@163.com

均只有一个或几个乘客,而往往数以千、万计. deMoura等人提出费米-狄拉克统计模型尽管能给出严格的理论解析解,但要求每个节点上仅能容纳几个粒子,因此该模型不能用于实际交通客运网络客流分析.

交通客运网络客流研究具有重要意义,然而目前交通客运网络的研究主要集中对网络拓扑结构分析上,很少讨论实际交通客运网络的客流问题. 本文根据交通客运网络的特点,要求每个站点所对应的最大乘客数不受限制,提出了交通客运网络的玻色-爱因斯坦统计模型,得到每个站点的平均客流量的解析表述式. 以中国铁路网实际拓扑和客流特征为基础,用随机行走模拟分析该网络的客流分布,得到结果与理论结果符合较好.

## 1 交通客运网络客流模型

在通常情况下,乘客若能从  $A$  站点通过某一车次到达  $B$  站点,则也能够从  $B$  站点沿同一或其他车次到达  $A$  站点,因此交通客运网络往往是无向的加权复杂网络<sup>[4 31]</sup>. 本文这样来构造和赋权交通客运网络:停靠站点为网络节点,若两站点之间有同一车次同向通过则看作这两个对应节点间有一条边连接,并且定义每条边的权重为在单位时间内该两站点所经过相同车次的数目;定义每个节点的点权为该节点所有相关边的边权之和,也可称之为节点度的加权值. 如图 1所示,设单位时间内有  $\omega_{ij}$  次车先后通过站点  $i$ 和站点  $j$ ,则节点  $i$ 和  $j$ 的边的权重为  $\omega_{ij}$ ,节点  $i$ 的点权为  $n_i = \sum_{j=1}^{k_i} \omega_{ij}$  其中  $k_i$  为节点  $i$ 的度,其值为该站点的直通车站个数.

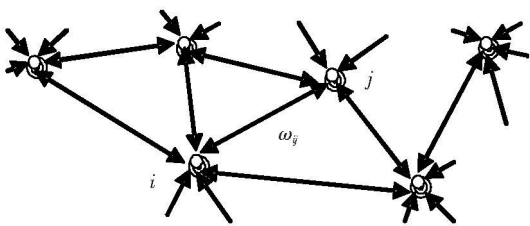


图 1 交通客运复杂网络示意图  
Fig.1 The sketch of the general passenger transport network

假设在一段时间内,某交通客运系统车次安排相对合理;网络中客流情况稳定,每个站点的进出站乘客数动态平衡;乘客在乘车方式选择上具有全同性;同一车次上的乘客数不受限制;乘客在任何时候都处于网络中的某种状态,将任何两个乘客互换,不改变整个系统的微观状态;忽略乘客之间的相互影响,认为乘客出行方式相互独立.

## 2 客流模型的统计物理分析

设想交通客运网络系统是一个热力学系统. 将当前每个车站进(出)站乘客的全体看成一个系统,发往不同的目标站点对应该系统的不同能级,也就是说每个车站发往不同目标站点的乘客对应系统内所处各自的能级的粒子,在单位时间内发往不同目标站点的车次数目为各自能级的简并度,即交通网络中两站点之间边的权值.

关于复杂网络能级的问题,文献[2]认为可用网络节点  $i$ 的某种固有属性  $\eta_i$ 来定义网络节点能级,并形式上定义网络节点能级为  $\varepsilon_i = -\frac{1}{\beta} \ln \eta_i$ ,文献[21]认为  $\beta \varepsilon_i = \ln(C k_i^{-1})$ ,其中  $k_i$ 是节点  $i$ 的度,  $C$ 是常量.

对于交通客运网络,能级反映了乘客总体的出行欲望. 根据能量最低原理,粒子总是尽量占据低能级状态. 站点点权越大,其所经过车次越多,能运送的乘客就越多,即粒子数就大,反映了该节点的能级越低. 能级与网络中节点的点权有关. 从统计物理观点看,一个粒子占据处于能级为  $\varepsilon_i$ 的几率为  $\frac{1}{e^{\frac{\alpha + \beta \varepsilon_i}{T}} - 1}$ . 另一方面,从交通客运网络角度看,客流量随着节点的点权数增加而加大,但是考虑到实际客运站点的容量,那么点权很大时客流量增加缓慢,因此可假设客流与  $\ln n_i$ 成正比. 若把交通客运网络系统看成统计物理的巨正则系统,满足:

$$\frac{1}{e^{\frac{\alpha + \beta \varepsilon_i}{T}} - 1} = c \ln n_i \tag{1}$$

交通客运网络以外的环境如政治、经济、文化及人口密度等等可以抽象为与该系统交换能量和粒子的源. 若出行人口占总人口的比例较小,可以认为外界源很大,且源与系统的作用很弱,交通客运网络相当于

一个巨正则系综. 这是一个玻色系统, 巨配分函数和平均总粒子数为:

$$\Xi = \prod_l \Xi_l = \prod_l \frac{1}{1 - e^{-\alpha - \beta \epsilon_l}}, \tag{2}$$

$$N = - \frac{\partial}{\partial \alpha} \ln \Xi = \sum_l \frac{1}{e^{\alpha + \beta \epsilon_l} - 1} = \sum_l a_l \tag{3}$$

其中  $\epsilon_i$  是与第  $i$  个节点有关的能量, 每个能级  $\epsilon_l$  上的粒子数  $a_l$ , 参数  $\beta = 1/kT$  和  $\alpha = - \mu/kT$  分别对应交通网络系统的温度和化学势  $\mu$ , 是比例常数. 化学势  $\mu$  取决于网络中的平均客流.

考虑某段时间内出发乘客总数为  $a_i$ , 站点系统  $i$  中有  $a_{ij}$  乘客发往站点  $j$ , 则这  $a_{ij}$  个乘客所处能级为  $\epsilon_j$ , 其对应简并度为  $\omega_j$ . 当乘客在各车站能级的分布为  $\{a_i\}$  时,  $a_{ij}$  可以独立地取各种可能值, 则站点  $i$  发往站点  $j$  的平均客流  $a_{ij}$ :

$$a_{ij} = \frac{\omega_{ij}}{e^{\alpha + \beta \epsilon_j} - 1}. \tag{4}$$

从上式可以看出站点  $i$  乘客选择站点  $j$  为目标站的概率为:

$$p_{ij} = \frac{\frac{\omega_{ij}}{e^{\alpha + \beta \epsilon_j} - 1}}{\sum_{j=1}^{k_i} \frac{\omega_{ij}}{e^{\alpha + \beta \epsilon_j} - 1}}. \tag{5}$$

对于实际交通客运网络, 在客流稳定状态下, 各站点的平均客流量为:

$$a_i = \sum_j a_{ij} = \sum_{l=1}^{k_i} (\omega_{il} \cdot c \ln(n_j)), \tag{6}$$

其中参数  $c$  可由客流总量确定:

$$a = \sum_i a_i = c \sum_i \sum_{l=1}^{k_i} (\omega_{il} \cdot \ln(n_j)). \tag{7}$$

由于这是一个巨正则系综, 实际的客流都在  $a$  周围涨落, 但对  $a$  很大的情况, 这种涨落是很小的. 用上述模型研究中国铁路交通客运网络的客流.

中国铁路交通网络现由 3 182 个火车客运站点、2 350 列车次组成. 统计目前中国大陆境内的所有火车客运站及 24 h 内所有列车次, 取 24 h 为一个时间步长, 这些相关数据全部来源于互联网. 需要强调的是, 本文是以实际火车客运站作为研究的节点, 若一个城市有多个站点也是分开独立计算的. 例如, 北京有北京站和北京西站, 沈阳有沈阳站和沈阳北站, 哈尔滨有哈尔滨站和哈尔滨东站. 各站点点权值见图 2 其中横轴为站点序号 (按点权值从高到底排序).

假设在一段时间内, 中国铁路网络中流动的乘客数恒定, 每个站点的进出站乘客数动态平衡, 鉴于非铁运高峰期, 我国铁路客流日发送量一般在 250 万人左右. 根据式 (6)、式 (7) 可理论计算客流分布, 结果 (见图 3) 表明各站点客流量与其对应点权呈正相关关系, 且站点对应点权值越高 (能级就越低), 就能分布更多的乘客. 如此一来, 解决出行高峰的拥堵问题最简单的方法就是通过客流总量来计算各节点点权, 用调节点权的方法来改变网络中的拓扑结构, 从而非均匀地分配有限资源.

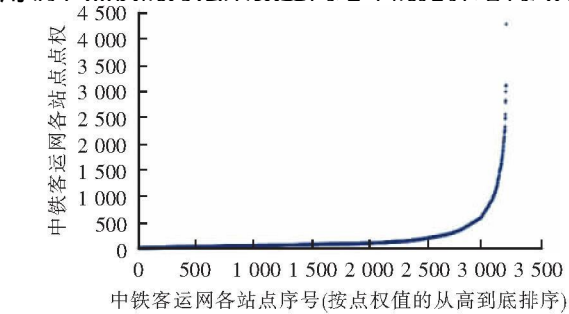


图 2 中铁客运网各站点点权值

Fig.2 The node weight of the station of chinese railway network

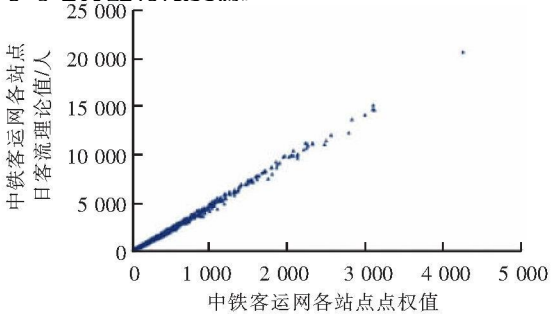


图 3 各站点客流理论结果

Fig.3 The theoretical results of the passenger flow of the various stations

### 3 随机行走模型分析客流

也可用随机行走模型来研究交通客运网络中的客流问题.

假设某交通客运系统某段时间内客流情况稳定, 每个站点的进出站乘客数动态平衡, 乘客可在网络中随机行走, 在乘车方式选择上具有全同性. 也就是说, 某时刻相对某出发站点的乘客而言, 那些进出车次越多的目标站点(点权越大) 被乘客选择的可能性就越大; 对于同样点权的站点, 通往它的车次越多(即两站之间边权越大), 其被乘客选择的可能性也越大. 为便于与理论结果相比较, 随机行走概率我们仍然选用(5)式, 将(1)式代入得:

$$p_{\bar{j}} = \frac{\omega_{\bar{j}} \ln n_j}{\sum_{j=1}^{k_i} (\omega_{\bar{j}} \ln n_j)}.$$

(8)

该过程的模拟算法如下: 我们设定初始  $t = 0$ 时刻, 大量乘客随机地分布在网络中不同站点, 每一个时间步长里, 网络中的所有乘客都将在其出发站点以式(8)的概率选择目标站点. 在  $t$ 时刻,  $i$ 站点上有  $a_{\bar{j} \ t} = a_{i \ t} p_{\bar{j}}$  乘客跑往站点  $j$ . 在  $t + 1$ 时刻, 从各个站点来站点  $j$ 上的乘客数为  $a_{j \ t+1} = \sum_{i=1}^{k_j} a_{ij \ t}$ . 设在局域时间内网络中乘客总数守恒, 每个站点进出站客流动态平衡, 这个随机过程被很多次重复以后, 该系统将达到一个稳定客流状态.

特别要说明的是, 在某个站点上车的乘客一定会在另外的某个站点下车, 讨论上车的乘客数和讨论下车的乘客数具有等效性, 因此我们只讨论下车的乘客数, 当然也可以只考虑上车的乘客数.

仍以中国铁路网络实际拓扑特征为基础进行客流分析. 设定初始  $t = 0$ 时刻, 铁路客流日发送量达 250 万人, 随机分配初始人数分布在全国 3 182个站点上. 考虑到一列车大约 2 500人左右座位数及可能超载情况, 故初始分配每个站点的人数上限是  $3\ 500n_i$  人,  $n_i$ 是该站点的点权.

各站点客流模拟值见图 4 其中坐标横轴为中铁客运网各站点点权值, 纵轴为稳定时各站点的日客流模拟值. 模拟结果显示, 无论初态如何将 250 万乘客随机地分布在全国 3 182个站点上, 经过足够长时间后, 各站点的客流分布趋于稳定, 稳定时各站点客流量与其点权正相关.

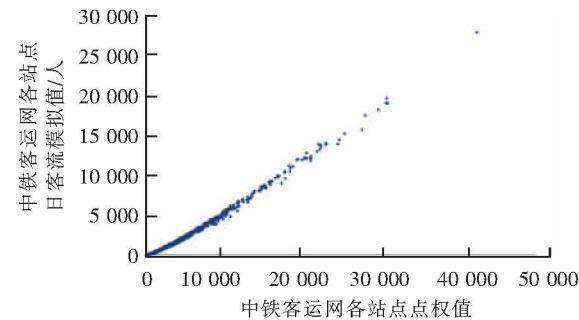


图 4 各站点客流模拟结果

Fig.4 The numerical simulation results of the passenger flow of the various stations

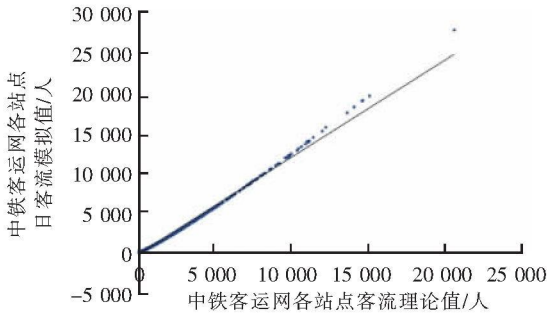


图 5 各站点理论值与模拟值的关系

Fig.5 The comparison between the results of simulation and the theoretical values of the various stations

### 4 理论和模拟结果比较

模拟与理论客流分布结果(见图 3)趋势相同, 尽管从模拟与理论的站点客流数据上看有个别站点模拟结果与理论结果的排名顺序有一些差异, 但是差异细微, 模拟和理论结果均表明各站点客流量与其点权呈正相关. 这一结果说明考虑交通客运网络系统的流量问题, 可以不用追随乘客个体状态的复杂变化, 用热力学方法求解微观量的统计平均值是可行的. 这也说明一个交通客运网络系统可以看成玻色系统, 其客流分布可以用玻色-爱因斯坦分布来描述.

为了更好地比较模拟结果与理论结果, 我们不仅在图 3和图 4中给出了各站点日客流量值模拟和理论值, 还在图 5中给出了各站点理论值与模拟值的对应关系, 其中散点对应中国铁路网络所有站点对应的

理论值与各自的模拟值; 实线为  $y = x$  直线.

从图 5 可以看出, 绝大部分站点客流量的模拟与理论的绝对数值差并不大. 尽管如此, 客流最大的前 200 个站点所对应的模拟与理论结果的绝对数值相差却很大. 究其原因, 一方面是因为这些站点客流基数很大, 模拟时每步均有小数进位的误差, 最后导致累积效应; 也可能是能级定义在点权值很大时不相符合的结果. 为了明晰这种误差, 我们在表 1 中列出了中国铁路客运网络计算误差最大的排名前 30 个站点的客流模拟值、理论值及站点名称, 由于上述结果是基于实际网络拓扑特征而得到的, 所以本表也一起列出点权值和度值.

表 1 中国铁路客运网络中客流、点权值及度值排名前 30 的站点

Table 1 The passenger flow, the node weight and the degree of the top 30 stations among Chinese railway network

客流模拟值		客流理论值		点权值		度值	
郑州	27 873	郑州	20 629	郑州	4 250	哈尔滨	485
南京	19 601	南京	15 082	沈阳北	3 104	沈阳北	449
株洲	19 028	株洲	14 646	南京	3 090	沈阳	437
沈阳北	18 985	沈阳北	14 600	株洲	3 082	北京西	433
西安	18 208	西安	14 072	西安	2 982	郑州	430
徐州	17 485	徐州	13 611	徐州	2 816	南京	400
哈尔滨	15 670	哈尔滨	12 219	哈尔滨	2 777	西安	393
鹰潭	15 196	鹰潭	11 980	鹰潭	2 548	齐齐哈尔	376
北京西	14 411	北京西	11 398	北京西	2 484	成都	372
洛阳	13 956	洛阳	11 142	沈阳	2 465	徐州	364
沈阳	13 896	上海	11 115	洛阳	2 310	石家庄	349
上海	13 849	蚌埠	11 059	金华西	2 310	长春	344
金华西	13 848	金华西	11 055	山海关	2 250	山海关	342
蚌埠	13 798	沈阳	11 002	蚌埠	2 240	北京	340
山海关	13 583	山海关	10 881	石家庄	2 233	四平	336
石家庄	13 249	石家庄	10 623	上海	2 217	株洲	329
衡阳	12 994	衡阳	10 442	衡阳	2 195	洛阳	324
无锡	12 775	无锡	10 349	四平	2 119	唐山	319
四平	12 346	四平	9 967	南昌	2 118	蚌埠	317
济南	12 207	济南	9 906	长春	2 112	广州	315
长春	12 189	长春	9 845	无锡	2 063	汉口	311
义乌	12 098	义乌	9 816	宝鸡	2 063	太原	309
宝鸡	12 093	苏州	9 798	义乌	2 040	锦州	309
苏州	12 007	宝鸡	9 797	济南	2 037	上海	305
唐山	11 919	常州	9 710	唐山	2 005	济南	305
常州	11 883	唐山	9 692	苏州	1 954	鹰潭	302
南昌	11 844	南昌	9 563	常州	1 934	天津	301
镇江	11 159	镇江	9 179	长沙	1 887	辽阳	301
长沙	11 083	长沙	9 086	商丘	1 856	秦皇岛	299
商丘	10 863	商丘	8 924	广州	1 848	南昌	297

5 讨论

用统计物理的方法把铁路、航空、汽运及公交等交通客运网络系统抽象成为一种与外界有能量和粒子交换的热平衡系统, 其中客流是一个重要参量, 为了研究这些具有固定节点和结构的交通网络中的客流状况, 本文以中国铁路网实际拓扑和客流特征为基础, 提出一个简单的动力学模型, 模拟分析了这种情况下的客流分布情况. 模拟结果显示, 无论初态如何, 经过足够长时间后, 各站点的客流分布趋于稳定. 然后利用巨正则系综理论分析交通客运网络系统对应的客流分布, 其理论结果与模拟结果总体上符合很好.

这两种分析对于其它网络流量分析也成立. 遗憾的是, 我们拿不到实际的全部铁路客流数据, 无法将理论结果与模拟结果直接与实际客流做对比分析. 这个模型也有点理想化, 从统计物理的角度来看, 交通客运网络实际是非平衡耗散系统, 在客流高峰或低谷时, 不能认为每个站点进出站客流动态平衡. 另外, 模

拟是基于中国铁路网络的实证数据进行的,为了简单而用式(1)来定义每个站点的能级,这可能是点权很大的站点上客流绝对数值相差很大的原因,不过也可能是误差的累积效应.对于这些问题,我们要做进一步的研究和改进.

# [参考文献]

- [1] Barabási A L, Albert R. Emergence of scaling in random networks[J]. Science, 1999, 286: 509-512.
- [2] Albert R, Barabási A L. Statistical mechanics of complex networks[J]. Rev Mod Phys, 2002, 74(1): 47-97.
- [3] Barabási A L, Bonabeau E. Scale-free networks[J]. Scientific American, 2003, 288(5): 50-59.
- [4] Newman M E J. The structure and function of complex networks[J]. SIAM Rev, 2003, 45(2): 167-256.
- [5] 吴金闪,狄增如.从统计物理学看复杂网络研究[J].物理学进展,2004,24(1):18-46.
- [6] Zhao L, Park K, Lai Y C. Attack vulnerability of scale-free networks due to cascading breakdown[J]. Phys Rev, 2004, 70(3): 035101(R).
- [7] Ravasz E, Barabási A L. Hierarchical organization in complex networks[J]. Phys Rev E, 2003, 67(2): 026112.
- [8] Albert R, Jeong H, Barabási A L. Diameter of the world wide web[J]. Nature, 1999, 401: 130-131.
- [9] Camacho J, Guinera R, Amaral L A N. Robust patterns in food web structure[J]. Phys Rev Lett, 2002, 88(22): 228102.
- [10] Newman M E J, Strogatz S H, Watts D J. Random graphs with arbitrary degree distributions and their applications[J]. Phys Rev E, 2001, 64(2): 026118.
- [11] Jeong H, Tombor B, Albert R, et al. The large-scale organization of metabolic networks[J]. Nature, 2000, 407: 651-654.
- [12] Jeong H, Mason S, Barabási A L, et al. Lethality and centrality in protein networks[J]. Nature, 2001, 411: 41-42.
- [13] Barthélemy M, Gondran B, Guichard E. Large scale cross-correlations in internet traffic[J]. Phys Rev E, 2002, 66(5): 056110.
- [14] Barthelemy M, Thumer S, Rodgers G J. Traffic on complex networks: towards understanding global statistical properties from microscopic density fluctuations[J]. Phys Rev E, 2004, 69(3): 036102.
- [15] Singh B K, Gupta N. Congestion and decongestion in a communication network[J]. Phys Rev E, 2005, 71(5): 055103.
- [16] Wang W X, Wang B H, Yin C Y, et al. Traffic dynamics based on local routing protocol on a scale-free network[J]. Phys Rev E, 2006, 73(2): 026111.
- [17] Zhao L, Lai Y C, Park K, et al. Onset of traffic congestion in complex networks[J]. Phys Rev E, 2005, 71(2): 026125.
- [18] de Moura A P S. Fermi-dirac statistics and traffic in complex networks[J]. Phys Rev E, 2005, 71(6): 066114.
- [19] Germano R, de Moura A P S. Traffic of particles in complex networks[J]. Phys Rev E, 2006, 74(3): 036117.
- [20] Guinera R, Amaral L A N. Modeling the world-wide airport network[J]. Eur Phys J B, 2004, 38: 381-387.
- [21] Guinera R, Mossa S, Tuftschi A, et al. The worldwide air transportation network: anomalous centrality, community structure, and cities' global roles[J]. Proc Natl Acad Sci USA, 2005, 102(22): 7794-7799.
- [22] Wu Z H, Braumstein L A, Colizza V, et al. Optimal paths in complex networks with correlated weights: the world wide airport network[J]. Phys Rev E, 2006, 74(5): 056104.
- [23] Barthélemy M, Barrat A, Pastor-Satorras R, et al. Characterization and modeling of weighted networks[J]. Physica A, 2005, 346(1/2): 34-43.
- [24] 刘宏鲲,周涛.中国城市航空网络的实证研究与分析[J].物理学报,2007,56(01):106-112.
- [25] Latora V, Marchiori M. Is the Boston subway a small world network[J]. Physica A, 2002, 314(1/4): 109-113.
- [26] Sen P, Dasgupta S, Chatterjee A, et al. Small world properties of the Indian railway network[J]. Phys Rev E, 2003, 67(3): 036106.
- [27] 谭江峡,王杜鹃,王鑫,等.与地理环境相关的中国铁路客运网拓扑结构[J].物理学报,2008,57(11):6771-6776.
- [28] 赵伟,何红生,林中才,等.中国铁路客运网网络性质的研究[J].物理学报,2006,55(08):3906-3911.
- [29] 赵金山,狄增如,王大辉.北京市公共汽车交通网络几何性质的实证研究[J].复杂系统与复杂性科学,2005,2(2):45-48.
- [30] Newman M E J. Analysis of weighted networks[J]. Phys Rev E, 2004, 70(5): 056131.
- [31] 汪秉宏,王文旭,周涛.交通流驱动的含权网络[J].物理,2006,35(4):304-310.

[责任编辑:顾晓天]