



论文

随机派系生长网络模型及其传输能力研究[†]丁益民^{①②}, 樊京芳^②, 周斌^①, 陈晓松^{②*}

① 湖北大学物理学与电子技术学院, 武汉 430062;

② 中国科学院理论物理研究所, 理论物理国家重点实验室, 北京 100190

*联系人, E-mail: chenxs@itp.ac.ac

†陈晓松编委供稿

收稿日期: 2013-10-15; 接受日期: 2013-11-29

国家自然科学基金(批准号: 11274102)和理论物理国家重点实验室开放式课题(编号: Y3KF321CJ1)资助项目

摘要 本文提出一种基于随机选择的派系生长网络模型, 该网络从一个 a -派系模体开始, 每个时间步 t , 在网络中随机选择 m 个节点构建一个新的 a -派系, 由此网络生长演化. 模拟研究表明: 该网络具有高的聚类系数和短的平均路径长度是一个小世界网络, 并且比值 $\rho=m/a$ 越小, 聚类系数越大. 而度分布则呈指数分布, 这些特征与许多交通系统的复杂网络的实证研究结果相符. 该网络的聚类系数与节点度呈幂律变化, 显示网络具有模块化层次结构的特征, 这一特征与近年来人们研究的大多数复杂网络具有模块化层次结构特征的实证研究结果相符. 我们还研究了该网络的传输能力, 研究表明该网络的传输能力随着比值 $\rho=m/a$ 的减小而增大. 这些研究结果对城市公共交通网络的构建具有一定的指导意义.

关键词 复杂网络, 派系, 层次结构, 传输能力

PACS: 89.75-Hc, 89.75.-k, 02.50.-r, 05.10.-a

doi: 10.1360/SSPMA2013-00083

1 引言

自 Watts 和 Barabási 分别提出小世界(Small-World)网络模型^[1]和无标度(Scale-Free)网络模型^[2]以来, 复杂网络的相关研究得到了迅速发展, 其研究涉及信息网络、技术网络、生物网络以及社会网络等诸多领域^[3-5]. 在复杂网络的研究中, 度分布、聚类系数和平均路径长度是描述复杂网络的重要物理量. 大量的实证研究表明, 许多的复杂网络同时具有高的聚类系数和短的平均路径长度, 呈现小世界特征^[1]. 还有一些复杂网络的度分布满足幂律分布, 表现出

无标度网络的特征^[2]. 然而现实世界中复杂网络的度分布规律比较复杂, 幂律分布是一种高度不均匀分布的特性, 它只能反映一部分复杂网络的结构特征, 事实上在现实世界中, 大量的复杂网络呈相对均匀的结构, 其度分布往往呈指数分布(或泊松分布)^[6]. 近年来人们的研究发现有些网络(如蛋白质网络)具有模块化的结构特征^[7,8]. 人们还发现对于许多复杂网络还存在着类似于分数维的自相似指数, 从而也揭示了复杂网络还具有某种内在的自相似性^[9]. 最近, 复杂网络的社团结构的研究越来越受到重视. 匈牙

引用格式: 丁益民, 樊京芳, 周斌, 等. 随机派系生长网络模型及其传输能力研究. 中国科学: 物理学 力学 天文学, 2014, 44: 299-304

Ding Y M, Fan J F, Zhou B, et al. Random clique evolving network model and their communicability (in Chinese). Sci Sin-Phys Mech Astron, 2014, 44: 299-304, doi: 10.1360/SSPMA2013-00083

利学者 Palla 等人^[10]首次提出派系(Clique)的概念, 并指出自然界和人类社会中的许多复杂网络都可描述为具有重叠连接的社团结构. 我国学者汪秉宏等人引入“Clique Degree”概念描述网络中节点所具有 Clique 的数目并研究了一些实际网络中的 Clique Degree 分布的规律^[11]. 最近, 美国学者 Clauset 等人^[12]在 Nature 上发表论文指出层次结构可能是建构复杂网络最重要的组织原则之一, 进一步揭示了复杂网络具有层次结构和社团结构的本质特征.

近年来人们对空间网络(Spatial Network)的研究产生了浓厚的兴趣^[13]. 空间网络是一种考虑网络拓扑结构受到空间限制的复杂网络^[5]. 大量的实证研究表明, 铁路网络、城市公共汽车交通网络和城市地铁网络等空间网络都显示出小世界网络的特征, 即具有高的聚类系数和短的平均路径长度. 印度铁路网络的实证研究结果表明: 印度铁路网络具有高的聚类系数和短的平均路径长度($C=0.69, L=2.16$), 是一种小世界网络^[14]. 波士顿地铁网络的实证研究表明该网络也显示出小世界网络的特征($C>0.9, L<2.0$), 不过与一般的小世界网络所不同的是这些网络具有更高的聚类系数^[15]. 关于空间网络度分布的研究, 人们却得出了不同的结果. 印度铁路网络的度分布服从指数规律^[14]; 波兰 22 个城市的公共汽车交通网络的度分布既有服从幂律分布, 也有服从指数分布, 但服从指数分布居多^[16]; 我国许多大城市如北京、上海、南京和杭州等大城市的公共汽车交通网络的度分布服从指数分布^[17]. 对于以上空间网络的特征, 我们可以这样理解: 在空间网络中, 网络节点的度要受到空间的限制, 即节点的度不能太大. 例如城市公共交通网络中, 每个站点的线路数受平面空间的限制, 不能太多, 因此其度分布往往呈相对均匀(与幂律分

布相比)的指数分布, 另一方面, 在空间网络中, 长距离的连接也受到几何距离的限制, 距离太长连接成本太高, 因此网络往往呈模块化的层次结构^[18,19].

综上所述, 实证研究发现: 许多空间网络, 尤其是交通网络大多具有高的聚类系数和短的平均路径长度、模块化的层次结构和指数型度分布特征. 这些网络特征不能用 WS 小世界网络、BA 无标度网络等现有的网络模型来解释. 为此, 本文在文献[19,20]的基础上提出了随机派系生长网络模型并用数值方法对其网络特性进行了模拟研究. 其研究结果与许多空间网络具有小世界、模块化和指数型度分布的实证研究结果相符. 我们还通过计算该网络的 Estrada 指数研究了该网络的传输能力, 并研究了网络参量对网络特性和传输能力的影响. 进一步分析了提高交通网络传输能力的措施.

2 网络模型

(1) 开始. 初始时刻 $t=0$, 从一个具有 a 个节点, 每个节点具有 $a-1$ 条边的 a -派系(a -Clique)完全子图开始.

(2) 选择. 每个时间步, 随机地在网络中选择 $m(m<a)$ 个节点作为新加入的 a -派系的起点.

(3) 生长. 以选择的 m 个节点为起点构建一个新的 a -派系, 以此类推, 网络生长演化为一个具有重叠的多派系结构的网络(如图 1 所示).

该网络模型与 Takemoto 等人^[20]提出 Evolving Networks by Merging Cliques 的网络模型所不同的是我们采用了随机选择的生长方式. 用这种方式构建的网络呈相对均匀的指数型度分布, 可以描述交通网络等相对均匀分布的空间网络的演化过程. 下面

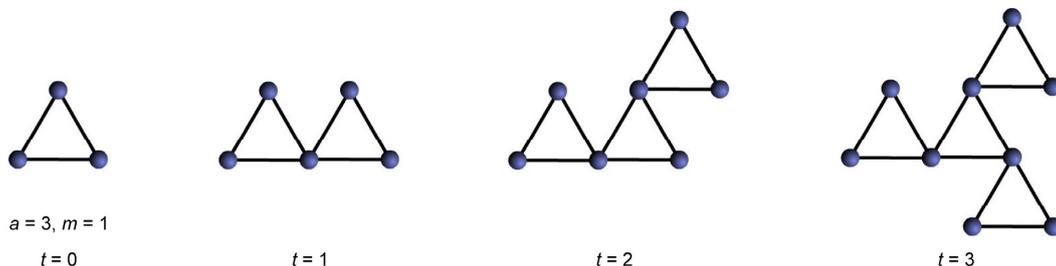


图 1 (网络版彩图)网络模型演化图

$a=3, m=1, a$ 是派系的节点数, m 是每次选择的节点数, t 是时间步

Figure 1 (Color online) Schematic illustration of the evolutionary procedure of our model network with $a=3$ and $m=1$. Each clique is attached through common nodes without adding extra edges.

我们用数值模拟方法对该模型的网络特性进行研究.

3 网络特性

我们用 Monte Carlo 方法模拟了该网络的度分布、聚类谱、聚类系数和平均路径长度.

3.1 度分布(Degree Distribution)

度是描述节点在网络中重要性的物理量. 节点 i 的度 k_i 定义为与此节点连接的边的数目, 网络中的所有节点的度的平均值称为网络平均度, 定义为 $\langle k \rangle$. 网络中节点的度分布一般用分布函数 $p(k)$ 表示, 它描述的是在网络中任意选择一个节点的度恰好为 k 的概率. 为了减小统计起伏和涨落, 我们通常采用累积度分布 $P(k) = \sum_{k'=k}^{\infty} p(k')$ 来描述网络中节点的度分布状况, 本文将研究网络的累积度分布.

该网络的累积度分布情况如图 2 所示. 由此可见, 该网络的累积度分布在对数坐标中近似为一条直线, 服从指数分布. 我们知道, 指数函数是差别程度比较小的不均匀分布^[6]. 如果是指数分布, 极大测量值出现的概率基本是零, 而在幂律分布中很大的值都有相对高的多的出现概率. 本网络模型的累积度分布呈指数型分布是由于该网络是按照随机选择的方式生长演化的. 我们的模拟研究结果与许多交通网络的度分布满足指数分布实证研究结果相一致.

3.2 聚类谱(Clustering Spectrum)

聚类谱是一种描述网络群体的层次结构特征的物理量^[6]. 聚类谱可以定义为

$$C(k) = \frac{1}{NP(k)} \sum_{i=1}^N c_i \times \delta(k - k_i), \quad (1)$$

这里 $\delta(x)$ 为 Kronecker 函数, 若 $k=k_i$, $\delta(k-k_i)=1$, 若 $k \neq k_i$, $\delta(k-k_i)=0$, c_i 为节点 i 的聚类系数定义为

$$c_i = \frac{2e_i}{k_i(k_i - 1)}. \quad (2)$$

k_i 是节点 i 的度, e_i 为节点 i 的所有邻居节点之间实际存在的边数. 聚类谱实际上就是度为 k 的所有节点聚类系数的平均值. 一般认为, 如果 $C(k) \propto k^{-\alpha}$, 则网络存在层次结构, 即网络可以划分为一个一个明显的层次. 我们对本网络模型的聚类谱进行了模拟计算, 计算结果如图 3 所示.

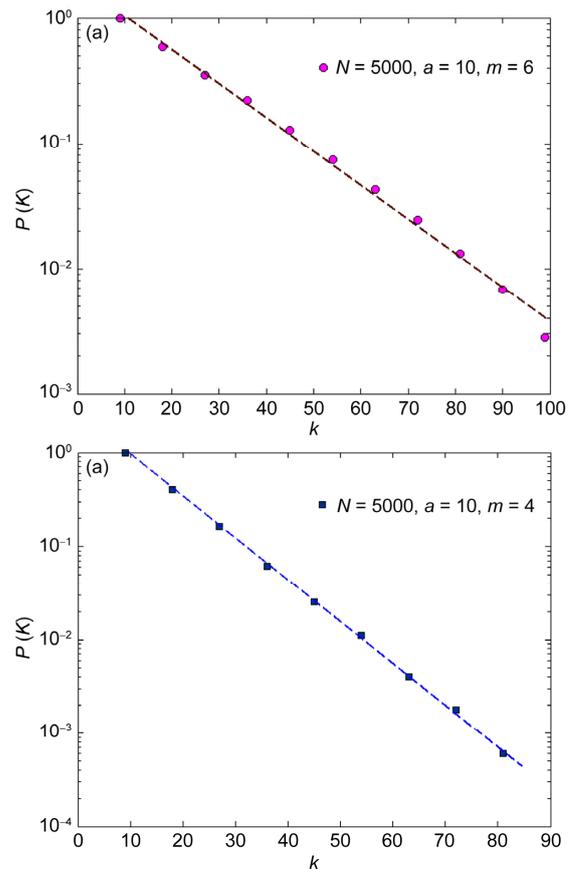


图 2 (网络版彩图)网络模型累积度分布模拟图

(a) 节点数 $N=5000, a=10, m=4$. 矩形表示模拟数据, 直线表示拟合直线. (b) 节点数 $N=5000, a=10, m=6$. 圆形表示模拟数据, 直线表示拟合直线

Figure 2 (Color online) The simulation of the cumulative degree distribution of our model for: (a) $N=5000, a=10, m=4$; (b) $N=5000, a=10, m=6$.

从图 3 可以看出, $C(k) \sim k$ 在双对数图中近似为一条直线, 即满足 $C(k) \propto k^{-\alpha}$, 可见本网络模型生成具有模块化层次结构的网络. 这一点我们可以这样理解, 由于构成网络的基本单元是派系 (a -Clique), 它是一种完全子图 (Complete Subgraphs), 也可以看成是一种小模块. 我们的网络是在这种小模块的基础上通过随机选择节点生长演化的, 因此网络总体上呈现模块化层次结构特性. 模块化层次结构是复杂网络的重要特征之一^[14], 本网络模型很好地解释了这一网络特征.

3.3 聚类系数与平均路径长度

聚类系数是表征复杂网络的局部特征的物理量.

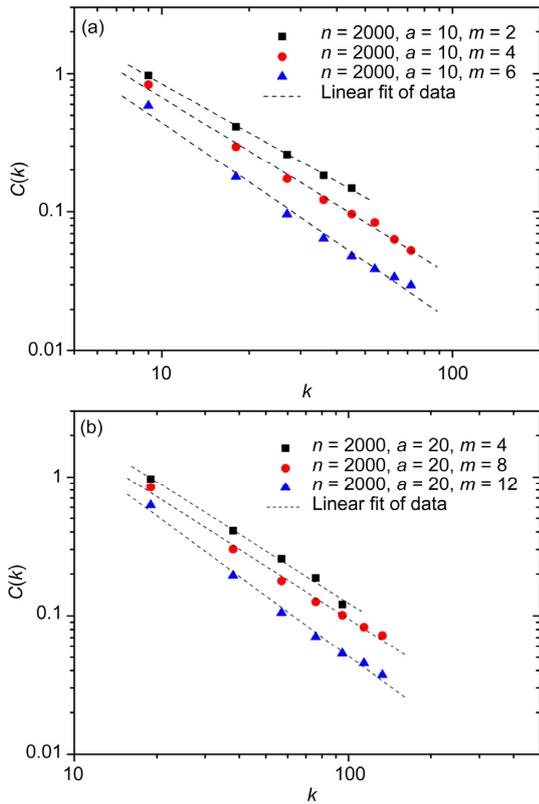


图3 (网络版彩图) 聚类谱 $C(k)$ 模拟图

(a) 节点数 $N=2000$, $a=10$, $m=2, 4, 6$. (b) 节点数 $N=2000$, $a=20$, $m=4, 8, 12$

Figure 3 (Color online) The plot of the $C(k)$ as a function of k for: (a) the clique sizes $a=10$; (b) the clique sizes $a=20$. Where, the node number $N=2000$, and the merging nodes $m=4, 8, 12$.

网络的聚类系数 C 为网络中所有节点(节点总数为 N) 的聚类系数的平均值

$$C = \langle c \rangle = \frac{1}{N} \sum_{i \in N} c_i. \quad (3)$$

平均路径长度是表征复杂网络全局特性的物理量. 网络平均路径长度 L 为网络中所有节点对之间距离的平均值

$$L = \frac{1}{N(N-1)} \sum_{i, j \in N, i \neq j} d_{ij}. \quad (4)$$

其中, N 为网络节点总数. d_{ij} 为网络中两个节点 i 和 j 之间的距离.

我们计算了该网络的聚类系数 C 与平均路径长度 L 的模拟值, 并与规模和平均度相当的 ER 随机网络模型的模拟值进行比较, 其数值如表 1 所示.

我们定义选择节点数 m 与派系节点数 a 的比值

表 1 随机派系生长网络模型统计参数表^{a)}

Table 1 The statistical parameters of the random clique evolving network

网络参数	1	2	3	4
N	1000	1000	1000	1000
a	10	10	10	10
m	2	4	6	8
$\langle k \rangle$	11.24	14.94	22.37	44.64
L	4.458	3.409	3.034	2.812
C	0.856	0.596	0.309	0.0683
L_{ER}	3.108	2.8218	2.5655	2.0811
C_{ER}	0.0111	0.0127	0.0222	0.0447

a) N : 网络节点数; $\langle k \rangle$: 平均度; C : 聚类系数; L : 平均路径长度; a : 派系节点数; m : 选择节点数

为 $\rho = m/a$, 则表 1 中的 4 组模拟数据的比值 ρ 依次为 0.2, 0.4, 0.6, 0.8. 由表 1 可以看出: 随机派系生长网络与规模和平均度大至相等 ER 网络相比, 平均路径长度比较接近即都具有较小的平均路径长度显示小世界特征. 而对于聚类系数而言, 随机派系生长网络的聚类系数明显大于 ER 网络的聚类系数, 并且 ρ 越小, 聚类系数越大. 这是因为 ρ 较小时, 派系节点数 a 与选择(合并的)节点数 m 相比较较大, 而每个 a 派系 (a -Clique) 为完全子图, 其每个节点(不考虑重叠的节点)的聚类系数为 1. 当 ρ 较小时则 m 较小, 网络中重叠的节点数也较小, 即网络中有更多聚类系数为 1 的节点, 从而使整个网络节点的平均聚类系数 C 较大.

4 网络的传输能力

网络传输能力的概念是 Estrada 和 Hatano^[21] 在 2008 年提出的描述网络传输特性的物理量. 过去人们考查网络中节点间的相互作用和通信通常只考虑两节点间的最短路径. Estrada 和 Hatano 认为, 网络中一对节点通信的产生不仅可以通过连接节点的最短路径, 而且可以通过连接节点的所有可能路径, 这种考虑所有可能路径情况的通信用一个新物理量 Communicability 来描述. 网络的 Estrada 指数一种表征网络整体的传输能力的物理量, 它等于所有节点子图中心度之和, 代表网络的一种全局指数. 中心度是指对每个节点处于网络中心地位程度的定量描述, 从而确定整个网络系统是否存在核心, 以及存在什么样的核心^[22].

设 A 为网络的邻接矩阵, $\lambda_j (j=1, 2, \dots)$ 为矩阵 A 的特征值, 则 Estrada 指数可表达为^[22]

$$EE(G) = \text{tr}(e^A) = \sum_{j=1}^n e^{\lambda_j}. \quad (5)$$

设 e^A 的特征值为 $\lambda_j (j=1, 2, \dots)$, 则可以证明

$$EE(G) = \text{tr}(e^A) = \sum_{j=1}^n e^{\lambda_j} = \sum_{j=1}^n \lambda_j'. \quad (6)$$

作为网络的一种全局指数, Estrada 指数可以表征网络整体的传输能力, 本文对该网络模型的 Estrada 指数进行模拟计算并将其结果与 ER 网络的模拟结果进行比较(如图 4 所示).

由图 4 可以看出, 在双对数图中, 随机派系生长网络与 ER 网络的 $EE(G)$ 指数均随网络节点数 N 线性增加, 满足幂次规律, 对于网络规模 N 和平均度 $\langle k \rangle$ 相当的网络, 随机派系生长网络一般比 ER 网络的 $EE(G)$ 指数要大, 并且当比值 $\rho = m/a$ 越小时, $EE(G)$ 指数越大. 这一模拟结果我们可以这样理解, 当比值 ρ 越小时, 每个 a -派系节点数 a 相对较大, 而当 a 较大时网络中节点的连接路径较多, 从而有更多可供选择传输的路径, 因此这时网络整体上具有较大的传输能力. 我们知道, 在交通网络中, 如何提高网络的传输能力是一个重要研究课题, 我们的研究结果对城市地铁网络、城市公共汽车网络的构建具有一定的参考价值.

5 结论

本文构造了一种基于随机选择的派系生长网络模型, 该模型充分考虑了现实世界中的复杂网络具有模块化的层次结构与交通网络具有相对均匀度分布的特征. 模拟研究表明: 该网络具有高的聚类系数

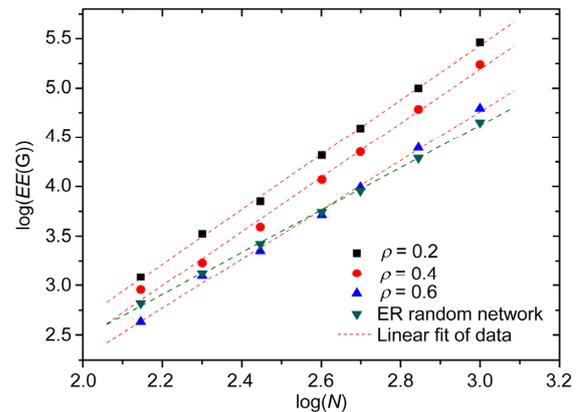


图 4 (网络版彩图) $EE(G)$ 指数模拟图

随机派系生长网络与 ER 网络的 $EE(G)$ 指数随网络节点数 N 的变化. 这里, 节点数 N 从 100 到 1000, $a=10$, $\rho=0.2, 0.4, 0.6$

Figure 4 (Color online) The plot the $EE(G)$ as a function of N for ER random network and the random clique evolving network. Where the clique sizes $a=10$, $\rho=0.2, 0.4, 0.6$, and the node number N from 100 to 1000.

和短的平均路径长度是一个小世界网络, 并且当比值 $\rho = m/a$ 越小时, 聚类系数越大. 而度分布则呈指数分布, 显示相对均匀网络的特征. 这些特征与多数交通系统的复杂网络的实证研究结果相符. 该网络的聚类系数与节点度呈幂律变化, 显示该网络具有模块化层次结构的特征, 这一特征与近年来人们研究的大多数复杂网络具有层次结构特征的实证研究结果相符. 我们还研究了该网络的传输能力, 研究表明该网络的 $EE(G)$ 指数一般高于 ER 随机网络, 并且比值 ρ 越小, $EE(G)$ 指数越大, 网络的传输能力越强. 这些研究结果对我国城市公共汽车网络、城市地铁网络的构建具有一定的指导意义.

参考文献

- 1 Watts D J, Strogatz S H. Collective dynamics of "small-world" networks. *Nature*, 1998, 393: 440–442
- 2 Barabási A L, Albert R. Emergence of scaling in random networks. *Science*, 1999, 286: 509–512
- 3 Albert R, Barabási A L. Statistical mechanics of complex networks. *Rev Mod Phys*, 2002, 74: 47–97
- 4 Newman M E J. The structure and function of complex networks. *SIAM Rev*, 2003, 45: 167–256
- 5 Boccaletti S, Latora V, Moreno Y, et al. Complex networks: Structure and dynamics. *Phys Rep*, 2006, 424: 175–308
- 6 何大初, 刘宗华, 汪秉宏. 复杂系统与复杂网络. 北京: 高等教育出版社, 2009. 95–100
- 7 Milo R, Shen-Orr S, Itzkovitz S, et al. Network motifs: Simple building blocks of complex networks. *Science*, 2002, 298: 824–827
- 8 Milo R, Itzkovitz S, Kashtan N, et al. Superfamilies organization of modularity in metabolic networks. *Science*, 2004, 303: 1538–1542
- 9 Song C, Havlim S, Makse H A. Self-similarity of complex networks. *Nature*, 2005, 433: 392–395
- 10 Palla G, Derenyi I, Farkas I, et al. Uncovering the overlapping community structure of complex networks in nature and society. *Nature*, 2005, 435: 814–818

- 11 Xiao W K, Ren J, Qi F, et al. Empirical study on clique-degree distribution of networks. *Phys Rev E*, 2007, 76: 037102
- 12 Clauset A, Moore C, Newman M E J. Hierarchical structure and the prediction of missing links in networks. *Nature*, 2008, 453: 98–101
- 13 Bathelemy M. Spatial networks. *Phys Rep*, 2011, 499: 1–101
- 14 Sen P, Dasgupta S, Chatterjee A. Small-world properties of the Indian railway network. *Phys Rev E*, 2003, 67: 036106
- 15 Seaton K A, Hackett L M. Stations, trains and small-world networks. *Phys A*, 2004, 339: 635–644
- 16 Sienkiewicz J, Holyst J A. Statistical analysis of 22 public transport networks in Poland. *Phy Rev E*, 2005, 72: 046127
- 17 Chen Y Z, Li N, He D R. A study on some urban bus transport networks. *Phys A*, 2007, 376: 747–754
- 18 Yang X H, Chen G, Sun B, et al. Bus transport network model with ideal n -depth clique network topology. *Phys A*, 2011, 390: 4660–4672
- 19 Ding Y M, Ding Z, Yang C P. The network model of urban subway networks with community structure (in Chinese). *Acta Phys Sin*, 2013, 62(09): 098901 [丁益民, 丁卓, 杨昌平. 基于社团结构的城市地铁网络模型研究. *物理学报*, 2013, 62(09): 098901]
- 20 Takemoto K, Oosawa C. Evolving networks by merging cliques. *Phy Rev E*, 2005, 72: 0461117
- 21 Estrada E, Hatano N. Communicability in complex networks. *Phys Rev E*, 2008, 77: 036111
- 22 Estrada E, Hatano N, Benzi M. The physics of communicability in complex networks. *Phys Rep*, 2012, 514: 89–119

Random clique evolving network model and their communicability[†]

DING YiMin^{1,2}, FAN JingFang², ZHOU Bin¹ & CHEN XiaoSong^{2*}

¹ Faculty of Physics and Electronic, Hubei University, Wuhan 430062, China;

² State Key Laboratory of Theoretical Physics, Institute of Theoretical Physics, Chinese Academy of Sciences, Beijing 100190, China

We introduce a notion of random clique evolving network, this network start from a complete subgraph of a -clique, where a is the size of the clique. In every time step T , m nodes are chosen from this network randomly, and forming a new complete subgraph of a -clique. In this way, this network grows in time steps. The numerical investigation shows that the cumulative degree distribution of this network takes an exponential function, which is the property of the homogeneous networks, and the clustering coefficient of this network is larger than that of the ER network. However, the characteristic path length of this network is the similar to that of the ER network, so this network shows the behaviors of small-world networks. Subsequent study shows that this network exhibits hierarchical modular structure for the clustering spectrum vs. k takes power-law. These results are in good agreement with the empirical results on many real-world complex networks, such as urban bus translation network or urban subway network, our model can explain the evolutionary procedure of these spatial networks. What's more, we present a numerical investigation on the communicability of our model by the Estrada index $EE(G)$, the Estrada index $EE(G)$ of this network increases with decreasing the rata m/a at the same size N and the same average degree $\langle k \rangle$. The communicability of the urban public translation networks is very important, our results have a certain guiding significance for the construction of urban bus translation network and urban subway network.

delay complex networks, clique, hierarchical structure, communicability

PACS: 89.75-Hc, 89.75-k, 02.50.-r, 05.10.-a

doi: 10.1360/SSPMA2013-00083