



# 一类变幂率的无标度网络模型构建和分析

那日萨\*, 张书超, 穆青

(大连理工大学 管理学院, 辽宁 大连 116024)

**摘要:** 多数现实网络的度分布指数(幂指数)介于(2,3),而BA网络的度分布指数恒等于3. 基于BA模型,引入老节点之间的择优连接机制,建立了一类变幂率的无标度网络模型,给出了这类复杂网络演化的解析结果,证明了在不同的参数下,其幂指数介于[2,3]. 同时指出,BA网络只是该模型的一个特例,通过一些实际网络数据的分析,说明了该网络模型的有效性和合理性.

**关键词:** 无标度网络;幂率;BA模型;择优连接

**中图分类号:** N94      **文献标志码:** A

## 0 引言

现实世界中许多系统都可以用复杂网络来描述,20世纪末,Albert等在对互联网的研究中发现了无标度网络(scale-free network),即节点度分布服从幂律分布,从而开辟了人类对于复杂网络系统认识的新天地.

BA模型<sup>[1]</sup>是第一个无标度网络演化模型,它捕捉到了无标度网络形成的增长和择优连接这两个必不可少的机制,说明了大规模复杂网络自组织成为无标度状态的原因.BA模型虽然比较准确地把握了现实世界中网络最基本的特点,较好地解释了无标度网络的形成机制,但其与现实网络相比,仍有一定的差距,如多数现实网络具有很大的集聚系数,它们的度分布指数 $\gamma$ 一般介于(2,3)<sup>[2~4]</sup>,相反BA网络的集聚系数随着网络规模增大趋于0且其度分布指数 $\gamma=3$ .

为了对现实的复杂网络进行更深入的研究,还需要对BA模型进行扩充,使它更加符合实际.刘慧等<sup>[5]</sup>兼顾局域演化、增长,以及局域与局域外存在较弱连接等3方面因素的加权网络,发现它服从幂律指数为(2,3)的幂律分布.Krapivsky等<sup>[6]</sup>研究了非线性的择优连接机制,并证明线性择优连接时的幂率分布指数 $\gamma \in (2,$

$\infty)$ .Dorogovtsev等<sup>[7]</sup>考虑了老节点之间的连接或者断开的情况,用节点连通的平均度的变化率来建立模型,并计算出该模型的网络幂指数 $\gamma = 2 + 1/(1+2c)$ ,其中 $c \in \mathbf{R}$ .Albert等<sup>[8]</sup>建立的模型则加入了一个老节点之间连接和已有边的重新连接这两种情况.但如何用连续域的解析方法建立和论证具有一般意义的复杂网络演化模型仍是需要解决的问题,这对于细致地分析复杂网络的演化过程具有重要意义.本文基于BA模型(新增节点的度为 $m$ ),引入老节点之间的线性择优连接机制(每次新增 $p$ 条边),建立一类变幂率的无标度网络模型,用连续域的方法<sup>[9]</sup>给出这类复杂网络演化的解析结果,并给出该网络的集聚系数;同时通过一些实际网络数据的分析,说明该网络模型的有效性和合理性.

## 1 模型的演化规则

BA模型只考虑了新加入节点与系统已存在节点的连接情况,忽略了系统已存的老节点间的连接情况,为此基于BA模型,本文引入老节点之间的连接机制,构建一类新的网络演化模型,其演化规则描述如下:

(1) 在初始时刻 $t=0$ ,假定系统中已有少量

( $m_0$  个) 节点.

(2)  $t = 1$  时刻, 新增一个度为  $m$  ( $m \leq m_0$ ) 的节点(与系统中原有的  $m_0$  个初始节点进行连接, 并且  $m_0$  个孤立点之间产生  $p$  条边 ( $p \leq m_0(m_0 - 1)/2$ )).

(3) 在以后每一个时间间隔内, 向系统增加一个度为  $m$  的新节点, 其  $m$  条边择优连接到网络中已经存在的不同节点上, 并且现有系统老节点

间择优产生  $p$  条边, 直至网络达到所需要的大小为止. 这样在经过  $t$  个时间间隔后, 便形成一个具有  $N = m_0 + t$  个节点、 $(m + p)t$  条边的网络.

具体的演化示意图见图 1, 初始时刻  $m_0 = 5$ , 若  $m = 3, p = 2$ .  $t = 1$  时, 新增节点 5 与节点 0、1、4 相连, 老节点 2 与 3、3 与 4 之间产生  $p = 2$  条边.  $t = 2$  时, 新增节点 6 与节点 3、4、5 相连, 已有节点 0 与 2、3 与 5 产生  $p = 2$  条边.

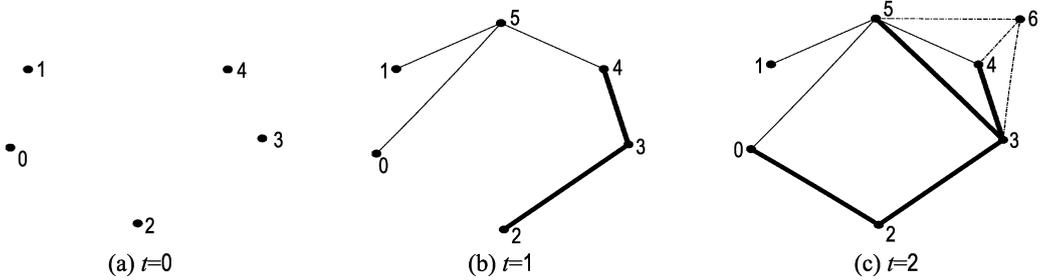


图 1 网络演化示意图

Fig. 1 Sketch diagram of network evolution

下面给出该网络演化的解析方程.

## 2 网络模型的解析解及计算分析

在  $t$  时刻系统中有  $m_0 + t$  个节点, 设模型中  $i$  节点的连通度为  $k_i(t)$ . 根据前述演化规则, 一方面考虑依据 BA 规则新增节点及其连接, 一方面考虑老节点间的连接.

根据 BA 模型演化的择优连接规则,  $i$  节点与新增节点连接的概率为  $P_i(t)$ , 其中  $P_i(t)$  取决于该节点的连通度,

$$P_i(t) = k_i(t) / \sum_{j=1}^{m_0+t-1} k_j(t) \quad (1)$$

假设  $k_i(t)$  连续, 由于新增节点与老节点间的连接, 以及老节点之间的连接均服从择优连接规则, 对已有节点  $i$ , 有

$$\frac{\partial k_i(t)}{\partial t} = \Delta k(t) \cdot P_i(t); \quad i = 1, 2, \dots, m_0 + t - 1 \quad (2)$$

其中  $\Delta k(t)$  表示  $t$  时刻网络系统连通度的增量. 在一个时间间隔网络系统连通度的变化  $\Delta k(t)$  由两部分构成, 一是由新增节点增加的  $m$  条边导致的新增连通度  $m$ , 二是由老节点产生  $p$  条边导致的新增连通度  $2p$ . 因此有  $\Delta k(t) = m + 2p$ , 又考

虑到  $\sum_{j=1}^{m_0+t-1} k_j(t) = 2(m + p)t$ , 则有

$$\frac{\partial k_i(t)}{\partial t} = \frac{m + 2p}{2(m + p)} \frac{k_i(t)}{t} \quad (3)$$

设节点  $i$  是  $t_i$  时刻添加到系统的, 则连通度  $k_i(t_i) = m$ , 此即为式(3)中微分方程的初始条件. 通过计算, 可得方程(3)的解

$$k_i(t) = m \left( \frac{t}{t_i} \right)^{\frac{m+2p}{2(m+p)}} \quad (4)$$

考虑节点连通度  $k_i(t)$  小于  $k$  的概率  $P(k_i(t) < k)$  可表示为

$$P(k_i(t) < k) = P\left(t_i > \left(\frac{m}{k}\right)^{\frac{2m+2p}{m+2p}} t\right) \quad (5)$$

假设时间  $t$  服从均匀分布, 即

$$P(t_i < T) = T / (m_0 + t) \quad (6)$$

则

$$P\left(t_i > \left(\frac{m}{k}\right)^{\frac{2m+2p}{m+2p}} t\right) = 1 - P\left(t_i < \left(\frac{m}{k}\right)^{\frac{2m+2p}{m+2p}} t\right) = 1 - \frac{1}{m_0 + t} \left(\frac{m}{k}\right)^{\frac{2m+2p}{m+2p}} t$$

即

$$P(k_i(t) < k) = 1 - \frac{1}{m_0 + t} \left(\frac{m}{k}\right)^{\frac{2m+2p}{m+2p}} \quad (7)$$

那么节点连通度的分布函数  $P(k)$  为

$$P(k) = \frac{\partial P[k_i(t) < k]}{\partial k} = \frac{t}{m_0 + t} \frac{2(m + p)}{m + 2p} m^{\frac{2m+2p}{m+2p}} k^{-(\frac{2m+2p}{m+2p} + 1)} \quad (8)$$

从中可得节点度分布呈幂率分布,其幂指数为

$$\gamma = \frac{2(m+p)}{m+2p} + 1 \quad (9)$$

式(9)可改写为

$$\gamma = 2 + \frac{1}{1+2p/m} \quad (10)$$

如果令  $c = p/m$ , 则式(10)与文献[7]的结果相同. 由式(10), 节点度分布的幂指数  $\gamma \in (2, 3]$ , 这与目前实际网络幂指数的值是符合的.  $p$  与  $m$  取值不同时  $\gamma$  的变化如下:

- (1) 当  $p = 0$  时,  $\gamma = 3$ , 即转化为 BA 模型;
- (2) 当  $p \ll m$  时, 即老节点连接的边远小于新节点连接的边时,  $\gamma$  接近 3;
- (3) 当  $p \gg m$  时, 即老节点之间的连接远大于新加入节点产生的边时,  $\gamma$  接近 2.

这说明, 本文模型能够更好地描述实际网络演化问题, 并且, BA 演化模型只是其一个特例.

图 2 给出了选取不同的  $p, m$  下利用理论结果式(8)计算得到的节点连通度分布函数图形.

图 2 是初始有 7 个点按照本文所述的演化规则演化到 10 000 个点的网络后节点的度分布情况.

通过图 2 可以看出, 在  $m = 3$  的情况下, 随着  $p$  的增大, 曲线斜率的绝对值减小, 即随着  $p/m$  的增大,  $\gamma$  减小, 结合式(10)可得  $\gamma_{m=3, p=1} = 2.6$ ,  $\gamma_{m=3, p=7} = 2.18$ . 图 3 给出了一个随机模拟节点连通度分布与理论分布函数比较的图.

通过图 3 可以看出, 随机模拟理论模型的节点连通度的分布结果一致, 此时理论节点度分布的幂率指数  $\gamma_{m=3, p=1} = 2.6$ .

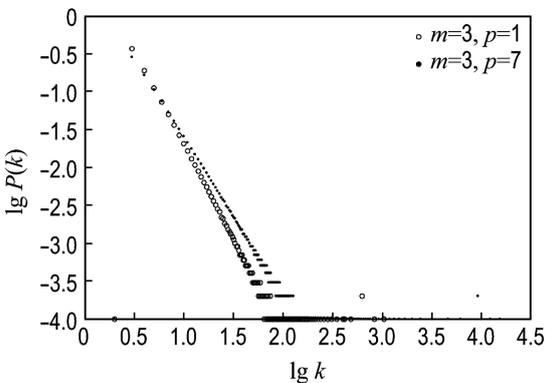


图 2 不同的  $p, m$  下节点连通度分布函数图形

Fig. 2 The diagram of distribution function of node connectivity in different  $p, m$

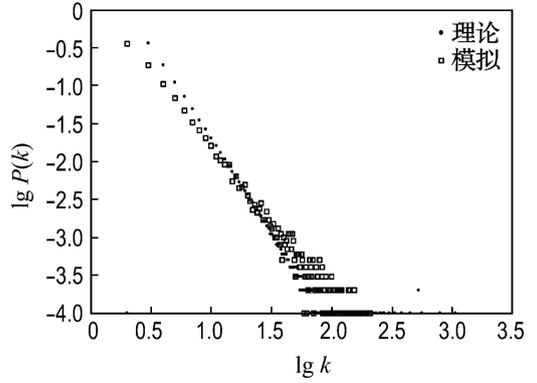


图 3 连通度分布与理论分布的比较 ( $m_0 = 5, m = 3, p = 1$ )

Fig. 3 The comparison of connectivity distribution and theoretical distribution function ( $m_0 = 5, m = 3, p = 1$ )

### 3 网络的集聚系数和平均路径

根据定义, 节点  $i$  的集聚系数  $C_i$  等于它的  $k_i$  个直接邻居之间实际存在的边数  $E_i$  占所有可能存在的边数  $k_i(k_i - 1)/2$  的比例, 即  $C_i = 2E_i/[k_i(k_i - 1)]$ . 整个网络的集聚系数指的是所有节点集聚系数的算术平均值. 可见,  $C_i$  只与两个变量  $E_i$  和  $k_i$  有关, 只有当  $E_i$  和  $k_i$  变化时,  $C_i$  才随着改变. 章志忠等<sup>[9]</sup> 曾给出与 BA 网络等价的一个网络的集聚系数. 由于在本模型中, 只有新节点与老节点之间才建立连接, 当新节点连接到节点  $i$  且与  $i$  的邻居进行连接时, 或者节点  $i$  的邻居之间发生连接时, 节点的  $k_i$  和  $E_i$  会发生改变, 从而改变  $C_i$  的值.  $C_i$  满足下面的动态方程:

$$\frac{dC_i}{dt} = \sum_{n=0}^{m+p-1} P_n \Delta C_m \quad (11)$$

其中  $\Delta C_m$  表示新节点连接到节点  $i$  和它的  $n$  个邻居时集聚系数的变化值,  $P_n$  表示这一变化的概率.  $\Delta C_m$  满足下式:

$$\Delta C_m = \frac{2(E_i + n)}{k_i(k_i + 1)} - \frac{2E_i}{k_i(k_i - 1)} = -\frac{2C_i}{k_i + 1} + \frac{2n}{k_i(k_i + 1)} \quad (12)$$

$P_n$  由两部分的乘积构成: 第 1 部分是有一条新边连接到节点  $i$  的概率, 这一概率由式(3)给出; 第 2 部分指的是其余  $(m + p - 1)$  条新边连接到节点  $i$  的  $n$  个邻居的概率, 它等价于每次成功概率为  $n(i)/(2m + 2p)t$  的  $(m + p - 1)$  重贝努利试验中, 成功  $n$  次的概率.  $n(i)$  表示节点  $i$  的邻居数目, 其值可通过下面的积分获得.

$$n(i) = \int_1^i k_x P(i_x) dt_x = \frac{m+p}{2} k_i \ln t \quad (13)$$

综上可得  $P_m$  的关系式为

$$P_m = \frac{(m+2p)k_i}{2(m+p)t} C_{m+p}^n \times \left(\frac{k_i \ln t}{4t}\right)^n \left(1 - \frac{k_i \ln t}{4t}\right)^{m+p-1-n} \quad (14)$$

由式(14)知:  $n > 1$  时,  $P_m$  很小, 可以忽略. 将式(4)、(12)、(14)代入方程(11), 忽略低阶项, 得

$$\frac{dC_i}{dt} = \sum_{n=0}^{m-1} P_m \Delta C_m \cong P_{i0} \Delta C_{i0} + P_{i1} \Delta C_{i1} = -\frac{m+2p}{(m+p)t+t\sqrt{tt_i}} C_i + \frac{(m+2p)(m+p-1)}{4(mt+t\sqrt{tt_i})} C_i \quad (15)$$

由式(15), 可得到  $C_i$  随时间的演化公式:

$$C_i(t) = \frac{(m+2p)(m+p-1)}{8(m+p)(\sqrt{t} + \sqrt{t_i}/m+p)^2} \times \left( (\ln t)^2 + \frac{4(m+p)}{(m+p-1)^2} (\ln t_i)^2 \right) \quad (16)$$

对式(16)进行积分计算, 就可得到整个网络的平均集聚系数

$$C = \frac{(m+p)(m+2p)(m+p+1)^2}{4(m+p-1)} \times \left[ \ln \left( \frac{m+p+1}{m+p} \right) - \frac{1}{m+p+1} \right] \frac{(\ln t)^2}{t} \quad (17)$$

可见, 根据该演化模型, 可推定其平均集聚系数随着网络规模的增大而迅速下降. 当  $p=0$  时,

$$C = \frac{m^2(m+1)^2}{4(m-1)} \times \left[ \ln \left( \frac{m+1}{m} \right) - \frac{1}{m+1} \right] \frac{(\ln t)^2}{t} \quad (18)$$

由式(18)知该演化模型的平均集聚系数与 BA 网络完全相同, BA 模型只是该演化模型的特例.

网络的平均路径长度是指网络中所有节点之间的平均最短距离, 它描述了网络节点之间的分离程度. 这里节点间的距离指的是从一节点到另一节点所要经历的边的最小数目. 平均路径长度的计算公式为  $L = \frac{1}{N(N-1)} \sum_{i \neq j} d_{ij}$ , 其中  $d_{ij}$

为节点  $i$  和  $j$  之间的最短距离,  $N$  为网络的规模. 关于随机网络平均路径长度的解析计算, 目前还没有普适的一般性方法. BA 模型平均路径的较精确解  $\lg n / \lg(\lg n)$  直到最近才被 Chen 等<sup>[10]</sup> 给

出, 可见复杂网络平均路径长度求解的困难. 在今后的工作中, 将对本文模型的平均路径长度的解析计算进行深入研究.

## 4 几个实际网络参数的估计

在现实生活中, 实际的许多网络如电影演员合作网等呈现出无标度现象, 这些网络相应规模下的幂率分布指数和节点平均连通度见表 1.

表 1 实际网络的无标度现象

Tab. 1 The scale-free phenomenon in the real networks

所研究的网络	网络规模	$\gamma$	$\langle k \rangle$
电影演员合作 <sup>[9,11]</sup>	212,250	2.30	28.78
数学家合作 <sup>[12]</sup>	70,975	2.50	3.90
引文 <sup>[2]</sup>	783,339	3.00	8.57
语言 <sup>[13]</sup>	460,902	2.70	70.13
互联网路由器 <sup>[2]</sup>	228,298	2.57	2.48

注:  $\gamma$  为实际网络的幂率分布指数,  $\langle k \rangle$  表示网络的平均度

这里假设实际的网络是由 BA 模型演化而来, 设网络初始有  $m_0$  个节点, 经过  $\omega$  时间, 则有  $\langle k \rangle = 2m\omega / (m_0 + \omega)$ , 当  $\omega \rightarrow \infty$  时,  $m$  则近似为实际网络的平均度  $\langle k \rangle$  的  $1/2$ , 将  $m, \gamma$  代入式(10)可以计算出  $p$ . 本文模型是同时考虑新增节点产生连接和老节点间的连接, 则有  $\langle k \rangle = 2(m+p)\omega / (m_0 + \omega)$ , 所以可近似地认为  $\langle k \rangle = (m+p)/2$ , 这样可以由算出的  $p$  和  $\langle k \rangle$  推出  $m$  的值, 结合式(10)可得本文模型的  $\gamma$ , 具体结果见表 2.

表 2 实际网络参数的估计

Tab. 2 The estimation of parameters on the real networks

所研究的网络	$\gamma$	$\gamma_{BA}$	$m_{BA}$	$p_{we}$	$m_{we}$	$\gamma_{we}$
电影演员合作	2.30	3	14	16.30	41.20	2.55
数学家合作	2.50	3	2	1.00	6.80	2.77
引文	3.00	3	4	0	17.14	3.00
语言	2.70	3	35	7.50	132.76	2.89
互联网路由器	2.57	3	1	0.37	4.50	2.85

表 2 中  $\gamma_{BA}$  表示 BA 模型推出的幂率分布指数,  $m_{BA}$  表示假设实际网络遵照 BA 模型演化时近似推得的 BA 模型中的  $m$  值,  $p_{we}$  表示由  $\gamma, m_{BA}$  结合式(10)计算的  $p$  值,  $m_{we}$  是根据  $p_{we}$  和  $\langle k \rangle$  近似得到的本模型中  $m$  值,  $\gamma_{we}$  则是将  $p_{we}, m_{we}$  代入式(10)计算得到的本模型的幂率分布指数值.

通过表 2 可以看出本模型所得到的幂率分布指数比 BA 模型更接近于实际网络, 从而也证明了本模型的合理性和适用性.

## 5 结 语

用连续域的解析方法建立和论证具有一般意义的复杂网络演化模型,对于分析大规模复杂网络的演化规律、准确计算网络各项综合指标(如集聚系数)具有重要意义.本文提出了一类变幂率的无标度网络模型,模型在BA模型的基础上,加入系统已存节点之间的连接(其连接隐含择优连接机制),并用连续域方法给出了该模型节点连通度分布函数的解析解和网络的集聚系数.通过模拟计算分析,验证了本文所提出的变幂率的网络模型为无标度网络,并且其幂率为 $2(m+p)/(m+2p)+1$ ,通过调节 $m$ 和 $p$ 的值,可以让其幂率在 $(2,3)$ 变化.当 $p=0$ 时,该模型与BA模型是一致的.通过一些实际网络的数据,近似计算出本模型的幂率分布指数,验证了该模型较之BA模型更具合理性和有效性.当然这里仅考虑了老节点之间产生新连接的问题,现实网络涉及的其他问题并没有加入到模型中,在今后的研究中,将进一步完善.

## 参 考 文 献:

- [1] BARABÁSI A L, ALBERT R, JEONG H. Mean-field theory for scale-free random networks [J]. *Physica A*, 1999, **272**(68):173-187
- [2] ALBERT R, BARABÁSI A L. Statistical mechanics of complex networks [J]. *Reviews of Modern Physics*, 2002, **74**(1):47-97
- [3] DOROGOVTSSEV S N, MENDES J F F. Evolution

- of networks [J]. *Advances in Physics*, 2002, **51**(4):1079-1187
- [4] NEWMAN M E J. The structure and function of complex networks [J]. *SIAM Review*, 2003, **45**(2):167-256
- [5] 刘 慧,李增扬,陆君安. 局域演化的加权网络模型 [J]. *复杂系统与复杂性科学*, 2006, **3**(1):36-43
- [6] KRAPIVSKY P L, REDNER S, LEYVRAZ F. Connectivity of growing random networks [J]. *Physical Review Letters*, 2000, **85**(21):4629-4632
- [7] DOROGOVTSSEV S N, MENDES J F F. Scaling behaviour of developing and decaying networks [J]. *Europhysics Letters*, 2000, **52**(33):33-39
- [8] ALBERT R, BARABÁSI A L. Topology of evolving networks; local events and universality [J]. *Physics Review Letters*, 2000, **85**(24):5234-5237
- [9] 章忠志,荣莉莉. BA网络的一个等价演化模型 [J]. *系统工程*, 2005, **23**(2):1-5
- [10] CHEN Fei, CHEN Zeng-qiang, WANG Xiu-feng, *et al.* The average path length of scale free networks [J]. *Communications in Nonlinear Science and Numerical Simulation*, 2008, **13**(7):1405-1410
- [11] BARABÁSI A L, ALBERT R. Emergence of scaling in random networks [J]. *Science*, 1999, **286**(5439):509-519
- [12] NEWMAN M E J. Scientific collaboration networks. I. Network construction and fundamental results [J]. *Physics Review E*, 2001, **64**(1):016131
- [13] CANCHO R F I, SOLE R V. The small-world of human language [J]. *Proceedings of the Royal Society B*, 2001, **268**(1482):2261-2265

## Model construction and analyses of a class of scale-free network with variable power law exponent

ZHAO Narisa\*, ZHANG Shu-chao, MU Qing

( School of Management, Dalian University of Technology, Dalian 116024, China )

**Abstract:** Most of the scaling exponents in real network are scattered between 2 and 3, while the exponent of BA network is 3 constantly. Based on the BA model, the preferential attachment mechanism between the existing vertexes is imported and a scale-free network whose scaling exponent can change in some certain scope is proposed. The degree distribution is calculated analytically and the analytical results indicate that the scaling exponent varies from 2 to 3 under different values of the parameters. It is also pointed out that the BA network model is the particular case of the model. And through the analyses of some data of the real networks, the model is proved to be reasonable and effective.

**Key words:** scale-free network; power law exponent; BA model; preferential attachment