

节点数加速增长的复杂网络生长模型^{*}

李 季^{1,2)} 汪秉宏^{1)†} 蒋品群^{1,3)} 周 涛¹⁾ 王文旭¹⁾

1) 中国科学技术大学近代物理系, 非线性科学中心, 合肥 230026)

2) 阜阳师范学院物理系, 阜阳 236041)

3) 广西师范大学物理与信息工程学院, 桂林 541004)

(2005 年 6 月 3 日收到, 2006 年 4 月 17 日收到修改稿)

受某些实际网络节点数按几何级数增长现象的启发, 构造了每个时间步中按当前网络规模成比例地同时加入多个节点的节点数加速增长的网络模型. 研究表明, 在增长率不是很大的情况下网络度分布仍然是幂律的, 但在不同的增长率 r 下幂律指数是不同的, 得到了幂律指数介于 2 到 3 之间可调的无标度网络模型, 并解析地给出了幂律指数随增长率变化的函数关系. 数值模拟还显示, 网络的平均最短距离随 r 减小而簇系数随 r 增大.

关键词: 复杂网络, 无标度网络, 生长网络模型, 节点数加速增长网络模型

PACC: 0590, 0175

1. 引 言

现实世界中个体之间复杂的相互作用关系可以用网络加以描述. 网络中的各个节点代表系统中的各个个体, 而个体之间的相互作用则可以抽象地用网络中的连边来表示. 自 20 世纪的 60 年代开始, 以 Erdős 为代表的一批数学家针对具有大量节点的网络系统提出并发展了随机图理论^[1]. 至今随机图论仍然是现代数学中的一个重要前沿领域. 然而从 20 世纪末开始, 计算机技术的发展使人们有可能对现实世界的一些复杂系统进行实证研究, 实证结果显示绝大多数实际系统的统计性质与随机图理论的结果是不同的. 这些新的统计性质的发现引起了物理学家极大的兴趣^[2-7]. 不论是以 Internet 为代表的技术网络^[8,9], 还是以科学家合作网为代表的社会网络^[10,11], 都显示出具有近似幂律特征的度分布, 而不是随机图理论所给出的泊松分布. 另外, 实际网络还表现出小世界效应, 即网络有大的簇系数和小的平均距离. 受现实网络特性的启发, Watts 和 Strogatz 以及 Barabási 和 Albert 分别提出了小世界网络模型 (WS 模型)^[12]和无标度网络模型 (BA 模型)^[13]. 前者

给出了较短的平均最短距离和较大的簇系数, 成功地描述了网络的小世界特性; 后一模型则构造出了具有稳定的幂律度分布的生长网络. 两种模型都具有非常简明的生成规则, 同时也忽略了影响实际网络生长的一些因素, 因而在某些统计性质上与实际网络相比有较大的偏离. 例如, BA 模型给出的簇系数太小, WS 模型计算出的度分布与实际网络不一致, 并且不生长. 近年来, 许多研究者考虑了更多影响网络生长的因素, 分别提出了许多不同的网络演化模型^[2-7,14-20]. 这些模型能够同时再现小世界效应和无标度特性, 而且各自从不同的方面更好地描述了实际网络的一些统计性质.

就我们所知, 在目前的绝大部分网络增长模型中, 每一时间步都是加入一个新的节点, 网络规模随时间线性增长. 然而在现实世界中, 网络规模一般随时间的变化是非线性的. 例如在互联网中, 在一个单位时间间隔内, 可能有多个新节点同时加入系统, 而且随着系统规模的增大, 同时加入的新节点数目也随之增多. 正如摩尔定律所指出的, 整个系统的规模按几何级数增长. 我们把这一现象称为节点数加速增长. 本文就是考虑到在实际网络演化过程中广泛

^{*} 国家自然科学基金 (批准号: 70271070, 10472116, 10532060, 10547004, 70471033) 和高等学校博士学科点专项科研基金 (批准号: 20020358009) 资助的课题.

[†] 通讯联系人. E-mail: bhwang@ustc.edu.cn

存在的节点数加速增长现象,提出了一种基于 BA 模型的节点数加速增长的网络模型.

我们推测,在增长率较小时网络应是无标度的,特别是当增长率趋于零时,网络应等价于 BA 网络.但当增长率较大时,网络的性质应与 BA 网络有区别.因此,我们把研究重点放在了加速增长因素对网络的基本统计性质的影响上,研究了网络的度分布及其与网络的增长率之间的关系及不同的增长率下生成的网络的平均最短距离和簇系数.

2. 节点数加速增长的网络模型及其统计性质

在现实世界中,许多系统都呈现出一种加速增长趋势.作为例证,图 1 给出了 Internet 的节点数随时间的变化(数据取自网上公布的统计结果)以及美国人口的增长情况^[21],其数据都可以用指数函数很好地拟合,说明这类系统的规模是加速增长的.

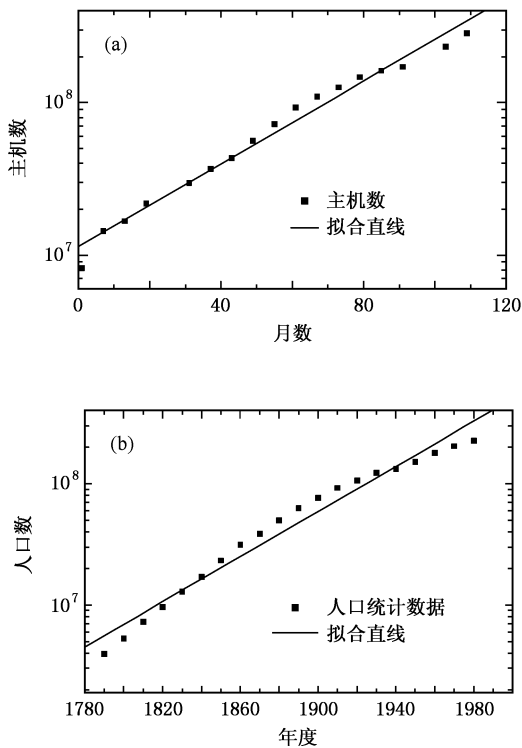


图 1 两个加速增长的系统实例 (a) Internet 上的主机数从 1995 年 7 月到 2004 年 7 月每 6 个月一次的统计结果 (b) 美国人口从 1780 年到 2000 年每 10 年一次的统计结果. 图中取的是单对数坐标, 因此拟合直线对应的实际是指数函数

在这些系统中,单位时间内加入的节点数(新增主机或新增人口)与系统的规模成比例地增长,而这些节点一般没有相互作用,不会直接相连.整个系统的规模是随时间按几何级数增长的.与之不同的是,在以往的增长网络模型中,任一时间步中新节点都只有一个.新节点加入网络后,在下一时间步中马上就变成了老节点.整个系统的规模随时间线性增长.

虽然我们可以通过对时间步 τ 与真实时间 t 之间作一非线性标度变换 $\tau = t^\alpha$,使得网络规模随时间呈几何级数增长,但这一做法不但不自然,而且并不能消除因同一批新节点之间互不相连而引起的对网络拓扑结构的影响.因此,受实际系统启发,我们对 BA 模型加以推广,提出了节点数加速增长的网络模型,试图通过此模型来研究加速增长因素对网络统计性质的影响.

2.1. 节点数加速增长网络模型

在我们的模型中,网络从初始的 n_0 个完全互连的节点上开始生长,每一时间步中,有 $r \times n(t)$ 个新节点同时加入,其中 $n(t)$ 为 t 时刻网络的节点数, r 为一正的比例系数,称为增长率.新加入的这一批节点,每一个都各自独立地与 m 个老节点相连,连接概率与老节点的度成正比,新节点之间则互不相连.这样,与 Internet 等的实证结果一致,网络的规模将以几何级数递增,是一种加速增长模型.我们注意到 Dorogovtsev 和 Mendes^[22]曾提出过加速增长的网络模型,他们的模型中考虑的是每个新节点与老节点连边数的持续增加,而每一时间步中仍是添加一个新节点.因此我们的模型考虑的生长机制与他们的模型是完全不同的.

2.2. 节点数加速增长网络的度分布

我们在不同网络规模下研究了增长率的不同取值对所生成网络的度分布的影响.数值模拟计算结果显示,增长率在相当宽的取值范围内($0 < r < 1$),网络的度分布都呈幂律.图 2 为增长率 r 取不同值运行到网络规模尽量接近 $N = 100000$ 时所得到的度分布.结果为重复运行 10 次的平均值. r 值的选取兼顾了使网络规模尽量接近 100000 的需要.图 2(a)(b)(c)都显示了很好的幂律度分布,图 2(d)在总体幂律趋势的基础上有周期性的波动.

为了更为细致地比较网络的度分布 $P(k)$ 随增长率 r 的变化情况,我们对数值模拟结果的数据进行了处理,给出了累积度分布.

图 3 给出的是增长率 r 取不同值时所生成网络

的累积度分布.与图 2 相比,图 3 可以以较高的精度来求取分布曲线线性部分的斜率,斜率加 1 即得度分布幂律指数 γ .

图 4 则直接给出了度分布的幂律指数 γ 与 r 的

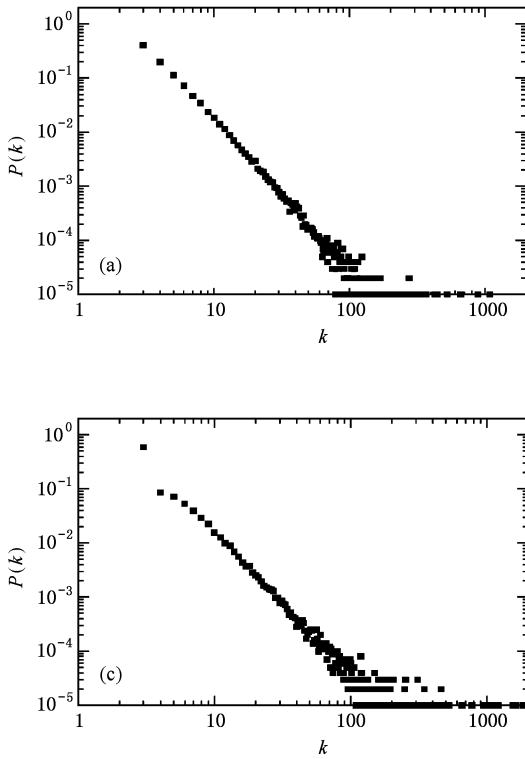


图 2 节点数加速增长模型在不同增长率下的度分布 $N = 100000$. (a) $r = 0.01$ (b) $r = 0.098$ (c) $r = 1.038$ (d) $r = 3.236$

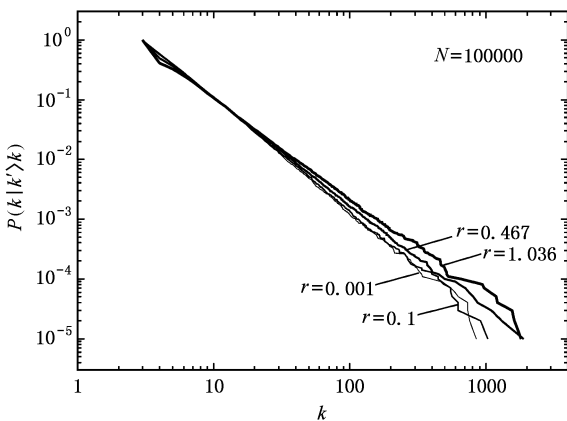
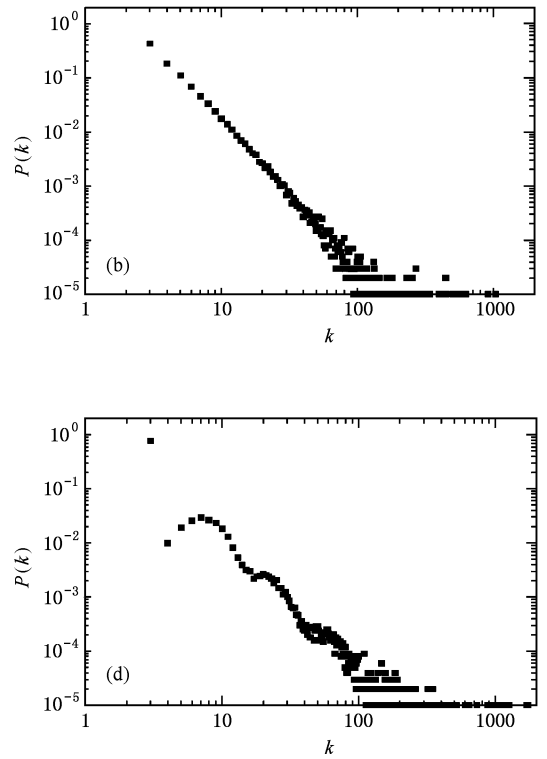


图 3 节点数加速增长模型在不同增长率下度的累积分布

关系,其中网络规模 $N = 2000, 5000, 10000$ 的 γ 值都是从运行 100 次所得度分布的数据处理得到.而 $N = 100000$ 的 γ 值则是 10 次平均的结果.从图 4 可以看出,随着网络增长率的增加,幂律指数随之单调减小.这与下面解析得出的关系 $\gamma = 1 + \ln(1 +$

$r) / \ln(1 + r/2)$ 是基本一致的.

2.3. 节点数加速增长网络的平均最短距离

平均最短距离是复杂网络的一个重要统计特性,小的平均最短距离意味着网络具有小世界特性^[12].对一个连通的无向网络,任意一对节点之间都存在一条最短的路径,路径长度 d 以其中含有的连边数来计量.整个网络中所有节点对之间的最短路径长度的平均值 \bar{d} 称为该网络的平均最短距离.通过对不同增长率下网络的平均最短距离的计算,我们给出了图 5.

从图 5 可看出,网络是具有小世界特性的.图 5 内插图中 N 取对数坐标, \bar{d} 取线性坐标,对应的增长率为 $r = 0.1$,有 $\bar{d} \propto \ln N$.可以对此作进一步的解析分析.若整个网络由 T 个时间步的过程生成,则任意两节点间的距离不超过 $2T$,故可把 $2T$ 作为估计网络的平均最短距离的一个上界.容易证明,在节点数加速增长的情况下,网络规模 N 与演化时间

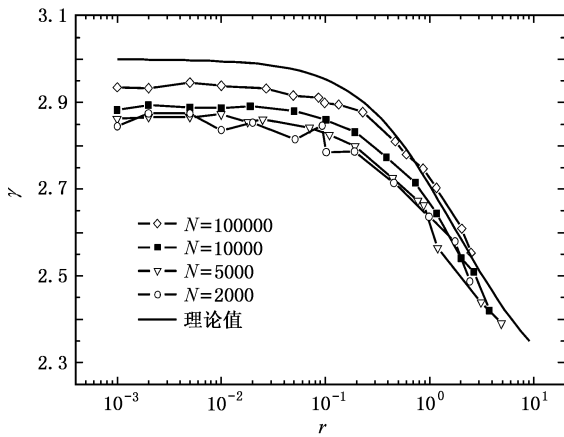


图 4 度分布幂律指数 γ 随增长率 r 变化关系

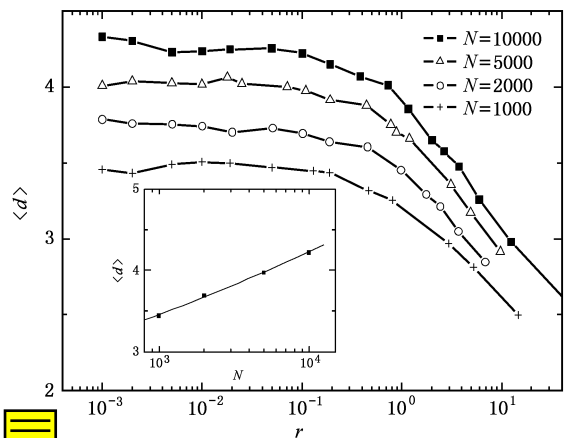


图 5 平均最短距离 d 随增长率 r 变化关系 内插图显示了 $r=0.1$ 时, 不同网络规模下 d 随 N 的变化趋势, 直线为拟合结果

步 T 之间存在关系 $N = n_0(1+r)^T$, 因而当 $N \rightarrow \infty$ 时有 $T = \ln N / \ln(1+r)$, 这样就有 $d < 2T \propto \ln N$. 因此, 我们认为在节点数加速增长的情况下网络仍然是具有小世界特性的. 与我们在考虑增长率对度分布指数的影响时看到的情况类似, 在 $0 < r < 0.1$ 的范围内增长率的改变对平均最短距离的影响并不显著, 而当 $r > 0.1$ 时, 随着 r 的增大网络的平均最短距离明显下降. 我们对此的解释是, 随着 r 的增大, 网络的拓扑结构逐渐趋向于一种“类星型”结构^[23]. 星型的中心由一些连通度较高的节点构成, 外层节点伸出多条边连向星型的中心, 整个网络比较“致密”.

2.4. 节点数加速增长网络的簇系数

簇系数是描述网络集聚性的一个统计量. 目前

对簇系数有不只一种定义, 本文采用 Watts 等^[12]原始的定义.

对一个节点 i , 度为 k_i , 其 k_i 个邻接点之间有 E_i 条连边, 则该节点的簇系数定义为

$$C_i = \frac{2E_i}{k_i(k_i - 1)} \quad (1)$$

整个网络的簇系数 C 为 C_i 的平均值, 即

$$C = \frac{1}{N} \sum_{i=1}^N C_i \quad (2)$$

在本模型中, 针对不同的网络规模计算了不同增长率下网络的簇系数, 结果如图 6 所示.

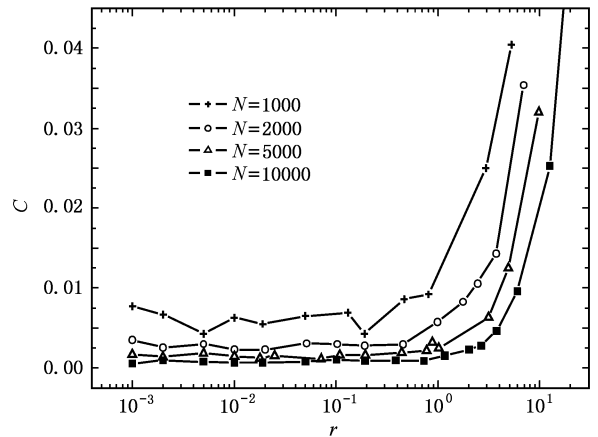


图 6 簇系数 C 随增长率 r 的变化关系

从图 6 可以看出, 在 $r < 1$ 时, 簇系数随 r 的变化不大, C 值一直较小. 而且网络规模越大, C 值越小. 这一点与 BA 模型一致. 而在 $r > 1$ 时, 随着 r 的增大 C 值明显增大. 与 2.3 节的结果一样, 也提示网络拓扑结构发生了明显的改变. 我们对此的解释依然是网络趋向于一种“类星型”结构. 这种结构的中心是由一些节点构成的高连通子图, 而连向这个中心的节点每个都是伸出多条边连到结构中心的不同节点上. 在 r 比较大时, 中心外的节点占有较大的比例, 其簇系数对整个网络的簇系数的贡献也较大, 因而结构中心的高连通度使整个网络的簇系数变大.

3. 增长率对网络度分布指数影响的解析分析

设整个网络从具有 n_0 个节点的完全图上开始生长, 则 n_0 个初始节点的度均为 $m = n_0 - 1$. 在节点数加速增长的情况下, 每次新加入的节点数

$\Delta n(t)$ 与系统当前的总节点数 $n(t)$ 成正比,比例系数记为 r .每个时间步中我们加入一批新节点,每个新节点伸向老节点的连边数为 m .这样,第一批加入的节点数为 $n_0 r$,加入后系统的总节点数变为 $n(1) = n_0(1+r)$;第二批加入的节点数则为 $n_0(1+r)r$,加入后系统的总节点数变为 $n_0(1+r)^2$;……;第 i 批加入的节点数为 $n_0(1+r)^{i-1}r$,加入后系统的总节点数变为 $n_0(1+r)^i$.

我们记 $K(t, i)$ 为加入第 t 批节点后,第 i 批节点的度之和.由于每批新加入的节点各自独立地只与老节点相连,连接概率与老节点的度成正比,因此有

$$K(t, i) = K(t-1, i) + P(t, i)K(t, t) \quad (t = 1, 2, 3, \dots; i = 0, 1, 2, \dots, t-1). \quad (3)$$

这里,

$$K(t, t) = mn_0 r(1+r)^{t-1}$$

既表示 t 时刻加入的新节点的度之和,又是这批新节点要连到老节点上的连边数;

$$P(t, i) = \frac{K(t-1, i)}{\sum_{j=0}^{t-1} K(t-1, j)} \quad (4)$$

表示新节点中的任一条边连到第 i 批节点中的概率.

将(3)式写成

$$K(t, i) - K(t-1, i) = \frac{K(t-1, i)}{\sum_{j=0}^{t-1} K(t-1, j)} K(t, t), \quad (5)$$

并对两边同时进行求和运算,可得

$$\sum_{i=0}^{t-1} K(t, i) - \sum_{i=0}^{t-1} K(t-1, i) = K(t, t). \quad (6)$$

(6)式有明确的物理意义:等号左端表示系统中的全部老节点的度之和从 $t-1$ 时刻到 t 时刻的增量,等号右端则表示新节点带来的连边数.显然,老节点的度之和的增加正是来源于新节点带来的连边.

利用(6)式,并反复使用其所代表的递推关系,可求得

$$\begin{aligned} & \sum_{i=0}^{t-1} K(t-1, i) \\ &= K(t-1, t-1) + \sum_{i=0}^{t-2} K(t-1, i) \\ &= K(t-1, t-1) + K(t-1, t-1) \\ & \quad + \sum_{i=0}^{t-2} K(t-2, i) \end{aligned}$$

$$= 2 \sum_{i=0}^{t-1} K(i, i). \quad (7)$$

将

$$K(i, i) = mn_0 r(1+r)^{i-1} \quad (i \neq 0), \\ K(0, 0) = mn_0$$

代入(7)式,即得

$$\sum_{i=0}^{t-1} K(t-1, i) = 2mn_0(1+r)^{t-1}. \quad (8)$$

将(8)式及

$$K(t, t) = mn_0 r(1+r)^{t-1}$$

代入(5)式,可求得

$$K(t, i) = \left(1 + \frac{r}{2}\right) K(t-1, i). \quad (9)$$

由此递推关系并考虑到

$$K(i, i) = mn_0 r(1+r)^{i-1},$$

可求得第 i 批节点度之和随时间的演化关系

$$K(t, i) = mn_0 r(1+r)^{i-1} \left(1 + \frac{r}{2}\right)^{t-i}. \quad (10)$$

由于第 i 批加入的节点数为 $n_0 r(1+r)^{i-1}$ 个,故这批节点的平均度为

$$\bar{k}(t, i) = m \left(1 + \frac{r}{2}\right)^{t-i}. \quad (11)$$

考虑 $\bar{k}(t, i)$ 的分布 P_{k_i} ,有

$$\begin{aligned} P_{k_i} &= \frac{n_0 r(1+r)^{i-1}}{n_0 + n_0 r + \dots + n_0 r(1+r)^{t-1}} \\ &= \frac{r}{1+r} (1+r)^{-(t-i)}. \end{aligned} \quad (12)$$

分别对(11)(12)式两边取对数,然后联立之,可得

$$\begin{aligned} \ln P_{k_i} &= \ln \frac{r}{1+r} + \frac{\ln m \ln(1+r)}{\ln\left(1 + \frac{r}{2}\right)} \\ & \quad - \frac{\ln(1+r)}{\ln\left(1 + \frac{r}{2}\right)} \ln \bar{k}_i. \end{aligned} \quad (13)$$

这样,我们就通过求具有累积分布性质的平均度的分布,得到了节点数加速增长的网络度分布幂律指数 γ 与增长率 r 之间的关系

$$\gamma = 1 + \frac{\ln(1+r)}{\ln\left(1 + \frac{r}{2}\right)}. \quad (14)$$

依此关系,当 $r \ll 1$ 时, $\gamma \approx 3$; $r \gg 1$ 时, $\gamma \approx 2$;一般情况下 $2 < \gamma < 3$.从图4可以看出,解析结果与数值模拟结果是基本一致的,而且当 N 越大时,模拟值与解析值越接近.

4. 结 论

考虑到节点数加速增长因素,本文得到了一个网络规模呈几何级数增长且 γ 可调的无标度网络. 根据对节点数加速增长网络的统计性质研究可知,

在增长率 $r \ll 1$ 时,加速增长因素不论是对网络的度分布、平均最短距离还是簇系数的影响都不明显. 但当 r 接近或大于 0.1 时,加速增长因素不论是对度分布、平均最短距离还是对网络簇系数的影响都是显著的.

-
- [1] Erdős P, Renyi A 1960 *Publications of the Mathematica Institute of the Hungarian Academy of Sciences* **5** 17
- [2] Albert R, Barabási A L 2002 *Rev. Mod. Phys.* **74** 47
- [3] Newman M E J 2003 *SIAM Rev.* **45** 167
- [4] Wu J S, Di Z R 2004 *Prog. Phys.* **24** 18 (in Chinese) [吴金闪、狄增如 2004 物理学进展 **24** 18]
- [5] Fang J Q, Wang X F, Liu Z R 2004 *Sci. Techn. Rev.* **2** 9 (in Chinese) [方锦清、汪小帆、刘曾荣 2004 科技导报 **2** 9]
- [6] Wang X F 2002 *Int. J. Bifurc. Chaos* **12** 885
- [7] Zhou T, Bai W J, Wang B H *et al* 2005 *Physics* **34** 1 (in Chinese) [周 涛、柏文洁、汪秉宏等 2005 物理 **34** 1]
- [8] Faloutsos M, Faloutsos P, Faloutsos C 1999 *Comput. Commun. Rev.* **29** 251
- [9] Pastor-Satorras R, Vázquez A, Vespignani A 2001 *Phys. Rev. Lett.* **87** 258701
- [10] Newman M E J 2001 *PNAS* **98** 404
- [11] Zhang P P, Chen K, He Y *et al* 2006 *Physica A* **360** 599
- [12] Watts D J, Strogatz S H 1998 *Nature* **393** 440
- [13] Barabási A L, Albert R, Jeong H *et al* 2000 *Science* **287** 2115
- [14] Holme P, Kim B J 2002 *Phys. Rev. E* **65** 026107
- [15] Klenm K, Eguíluz V E 2002 *Phys. Rev. E* **65** 036123
- [16] Andrade J S Jr, Herrmann H J, Andrade R F S *et al* 2005 *Phys. Rev. Lett.* **94** 018702
- [17] Wang W X, Wang B H, Hu B 2005 *Phys. Rev. Lett.* **94** 188702
- [18] Zhou T, Yan G, Wang B H 2005 *Phys. Rev. E* **71** 046141
- [19] He Y, Zhang P P, Xu T *et al* 2004 *Acta Phys. Sin.* **53** 1710 (in Chinese) [何 阅、张培培、许 田等 2004 物理学报 **53** 1710]
- [20] Wang W X, Hu B, Zhou T *et al* 2005 *Phys. Rev. E* **72** 046140
- [21] Peter J B, Richard A D 1991 *Time Series: Theory and Methods* (New York: Springer-Verlag)
- [22] Dorogovtsev S N, Mendes J F F 2001 *Phys. Rev. E* **63** 025101
- [23] Bollobás B 1998 *Modern Graph Theory* (New York: Springer-Verlag)

Growing complex network model with acceleratingly increasing number of nodes^{*}

Li Ji^{1,2)} Wang Bing-Hong^{1)†} Jiang Pin-Qun^{1,3)} Zhou Tao¹⁾ Wang Wen-Xu¹⁾

¹ *Nonlinear Science Center, Department of Modern Physics, University of Science and Technology of China, Hefei 230026, China*

² *Department of Physics, Fuyang Normal College, Fuyang 236041, China*

³ *College of Physics and Information Engineering, Guangxi Normal University, Guilin 541004, China*

(Received 3 June 2005 ; revised manuscript received 17 April 2006)

Abstract

Inspired by the observation that some real-life networks' sizes grow as a geometric series, a growing complex network model with acceleratingly increasing number of nodes is proposed. At each time step, the number of newly added nodes is proportional to the size of the network. This network shows scale-free property when the growing rate r is not large, and its power-law exponent is tunable from 2 to 3 through r . The average path length decreases and clustering coefficient increases with r respectively. In addition, we also give an analytical solution about power-law exponent versus r that agrees well with the simulation result.

Keywords : complex networks, scale-free networks, growth network model, network model with accelerating increasing number of nodes

PACC : 0590, 0175

^{*} Project supported by the National Natural Science Foundation of China (Grant Nos. 70271070, 10472116, 10532060, 10547004, 70471033) and the Doctoral Program Foundation of Institution of Higher Education of China (Grant No. 20020358009).

[†] Corresponding author. E-mail: bhwang@ustc.edu.cn