

加速增长的 HK 网络演化模型

崔爱香 傅彦

(电子科技大学互联网科学中心 成都 611731)

摘 要 近年来,随着复杂网络演化模型的深入研究,研究的焦点已从全局结构转移到了局部结构,其中,网络的集团度及其分布受到了广泛关注。实证研究表明,大量真实网络具有幂律的低阶集团度分布,而且分布指数随集团阶数的上升呈下降趋势。这一普遍规律无法由 Holme 和 Kim(HK)模型的熟人推荐机制再现。在 HK 模型中考虑网络演化的加速增长这一重要因素,提出一种改进的 HK 模型。数值模拟显示,改进的 HK 模型生成网络的簇系数较大,平均最短路径较小,不仅具有小世界效应和无标度特性,而且再现了真实网络的低阶幂律集团度分布特性。此研究有利于更好地认识真实网络中的各阶基元。

关键词 复杂网络,网络演化,演化模型,加速增长,集团度分布

中图法分类号 TP301.6 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2015.4.005

Accelerated-growth HK Network Evolution Model

CUI Ai-xiang FU Yan

(Web Science Center, University of Electronic Science and Technology of China, Chengdu 611731, China)

Abstract In recent years, with the further study of evolution model of complex networks, the research focus has shifted from the global structure to local structure. Empirical results show that many real networks exhibit power-law clique-degree distribution, and the distribution exponents decrease with the increase of the order of clique. This general regularity can't be produced by the acquaintance recommended mechanism of HK model proposed by Holme and Kim. This work considered the property of accelerated growth of networks and proposed an improved HK model. Numerical simulations indicate that accelerated growth HK model has large cluster coefficient and small average shortest path. It not only has the small-world effect and scale-free property, but also reproduces the observed power-law clique-degree distribution. This study is better to understand the motifs in the network.

Keywords Complex networks, Network evolution, Evolution model, Accelerated growth, Clique-degree distribution

1 引言

近年来,复杂网络的研究受到了广泛关注^[1-7]。WS 小世界网络模型^[8]和 BA 无标度网络模型^[9]开创了复杂网络研究的新纪元,前者揭示了复杂网络的小世界特性,后者阐述了无标度特性的产生机理,标志着人们对复杂网络的研究进入了一个新时代。真实网络的统计特性促使人们不断地从理论上构造出与真实网络更接近的演化模型,解释网络的形成机理,探索隐藏在网络背后的内在演化机制。寻找可能表征部分真实网络的演化规律并建立相应的模型,一直是推动复杂网络演化模型研究的根本动力^[10]。

在复杂网络研究初期,人们主要关注网络的宏观特性,如小世界特性和无标度特性等。随着对复杂网络演化模型研究的深入,研究的焦点开始转移到网络的局部结构^[11],如网络中的模块、紧密子图等。周涛等^[12,13]提出集团度的概念来刻画网络的局部结构,用来作为量化基元密度的一种尝试。在无向网络中, s 阶集团是指由 s 个节点构成的完全子图($s > 1$),记做 K_s ,其中 K_2 是一条边, K_3 是一个三角形, K_s 包括 s

个节点和 $s(s-1)/2$ 条边。图 1 给出了 K_2 到 K_5 的示意图。一个节点 i 的 s 阶集团度 k_i^s 是指包含该节点的不同的 s 阶集团的数目。显然,节点 i 的 2 阶集团度 k_i^2 等于该节点的度,3 阶集团度 k_i^3 等于包含节点 i 的三角形的个数等。从这个意义上讲,集团度可以看作是网络节点度概念的推广,2 阶集团度分布就是传统的度分布。为了更清楚地说明集团度的定义,图 2 给出了一个 9 个点的网络,节点 i 的 2 到 5 阶集团度分别为 $k_i^2=8, k_i^3=5, k_i^4=1, k_i^5=0$ 。周涛等^[12,13]对社会网络、技术网络、生物网络和信息网络 4 类复杂网络中典型真实网络的集团度分布进行了实证分析,如自主系统层次的计算机互联网、路由器层次的计算机互联网、万维网、科学家合作网等。对于每个真实网络,分别给出了其 2 阶到 5 阶的集团度分布。实证研究发现,尽管这些从不同领域抽象出来的网络的背景相差很大,但它们的集团度分布都表现出相似的性质,近似服从幂函数形式。对真实网络集团度分布的实证研究揭示了网络具有近似服从幂律的低阶集团度分布,而且分布指数随着所统计的集团阶数的上升呈下降的趋势。这一新统计特性为复杂网络演化模型的研究提供了新的比较平台。杨涵

到稿日期:2014-05-12 返修日期:2014-07-14 本文受国家自然科学基金(61300018,61103109),中央高校基本科研业务费(ZYGX2012 YB027,ZYGX2012J071,ZYGX2013J079,ZYGX2011J057)资助。

崔爱香(1985—),女,博士生,CCF 会员,主要研究方向为复杂网络、信息物理等,E-mail: cax2006@126.com;傅彦(1962—),女,教授,博士生导师,主要研究方向为数据挖掘、信息推荐、复杂系统等。

新等提出了一个基于随机行走的网络演化模型,能够生成具有幂律集团度分布的网络,但是集团度分布指数随着阶数呈上升趋势,与真实情况恰恰相反。

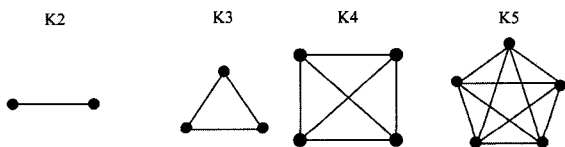


图1 集团结构示意图

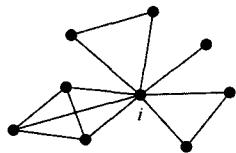


图2 简单网络示意图

Barabasi 和 Albert 实证研究了万维网的动态演化过程,发现万维网每天都有新网页增加,并且新网页总是喜欢与知名的网页相连接,据此提出了 BA 模型^[10],再现了真实网络所观察到的幂律度分布特性。BA 模型考虑了真实网络演化的两个重要特性:增长特性和优先连接(Preferential Attachment, PA)特性。前者说明随着新节点的加入,网络的规模不断增大;后者说明新加入的节点更倾向于与那些度大的节点相连接,又称为“富者愈富”。为了再现真实网络的高集聚特性(即较大的簇系数),Holme 和 Kim 在 BA 模型的基础上,引入了三角形形成(Triad Formation, TF)机制,又称“熟人推荐”机制,提出了 HK 模型^[14],该模型生成的网络不仅具有幂律度分布,而且具有较大的簇系数,实现了小世界特性和无标度特性的统一。但是, HK 模型只考虑了网络演化的线性增长,在网络演化过程中每个新引入的节点带来固定数目的连边,使得网络中的总边数随节点数呈线性增长。HK 模型生成的网络即使达到 10 万个节点的规模也不存在 5 阶及 5 阶以上的集团,不具有幂律的低阶集团度分布特性。加速增长是指网络演化的过程中连边数比节点数增长得更快,网络的平均度随时间增大,表现出一种超线性增长方式, Dorogovtsev 和 Mendes 研究发现加速增长在很大程度上影响网络的拓扑结构^[15]。网络的加速增长是比线性增长更一般的情况,这种加速增长现象广泛存在于现实演化网络中,是网络演化中的重要因素。加速增长是网络演化过程遵循的一个规则,而不是单单只存在于某一特例网络中。在很多情况下,如果不考虑演化网络的加速增长,则很难理解网络的特性。本文在经典的 HK 模型中引入加速增长这一因素,数值模拟显示加速增长的 HK 模型可同时再现真实网络的小世界特性、无标度特性和低阶幂律集团度分布特性,而且集团度分布指数的变化规律与实证研究一致。

2 模型算法

Dorogovtsev 和 Mendes 对多个真实网络进行了实证分析^[16],发现万维网、因特网、论文引用网、科学家合作网等均存在加速增长现象。张国清等^[17]进一步研究了大规模因特网在自治系统级别的演化,发现网络中的连边数随网络规模呈指数形式增长,服从著名的摩尔定律,在统计上表现为 $E \sim N^\gamma$, 其中 $\gamma > 1$ 。对于自治系统的互联网^[17], $\gamma \approx 1.11$; 而对于万维网^[18], $\gamma \approx 1.29$ 。加速增长现象广泛存在于现实网络中,是网络演化过程应遵循的一个规则,然而这一重要因素

在之前的网络演化模型中却常常被忽略了。

本文提出一种改进的 HK 模型,其不仅考虑到网络的优先连接机制和熟人推荐机制,而且在经典的 HK 模型中引入加速增长这一重要因素。这里,考虑到网络中的连边数随网络规模呈指数形式增长,这种加速增长主要是由于新节点的加入与已存节点之间建立了越来越多的连接。具体的算法如下:

(1) 初始网络:包含 m_0 个完全连通的节点;

(2) 网络演化:在每个时间步引入一个新节点 v 和 Δm 条边,其中 $\Delta m = \lfloor E_{N+1} - E_N \rfloor$, 其中 $\lfloor \cdot \rfloor$ 表示向下取整, N 是引入新节点前网络的节点数目, $E \sim N^\gamma$ 是加速增长网络应该有的连边数目。不再像 BA 模型和 HK 模型,每加入一个新节点,引入固定的变数。具体的连边建立规则如下:

1) 每个时间步新引入的节点 v 首先采用优先连接机制^[10]与网络中已存节点 i 建立一条连边。新节点 v 与已存节点 i 建立连边的概率 Π_i 正比于节点 i 的度,如式(1)所示:

$$\Pi_i = \frac{k_i}{\sum_i k_i} \quad (1)$$

其中, k_i 表示节点 i 的度。

2) 以概率 p 采用熟人推荐机制^[14]或以概率 $1-p$ 采用优先连接机制添加其余的 $\Delta m - 1$ 条边。熟人推荐机制是指新节点 v 与已存节点 i 建立连接之后,进一步与节点 i 的邻居节点建立新连接,也就是通过朋友的朋友来建立新连接。

由以上算法可知,在网络演化的过程中,随着网络规模的不断增大,每个新引入的节点添加的边数 Δm 不断增大。与线性增长的 HK 模型相比,加速增长的 HK 模型生成网络的节点平均度不再是一个常数,而是不断增长的。概率 p 的取值影响生成网络的簇系数, p 的值越大,生成网络所包含的三角形个数越多,网络的簇系数也就越大。

3 数值仿真结果

当初始网络规模 m_0 远小于网络最终规模时,初始网络对最终的结果没有影响,这里数值仿真取 $m_0 = 3$ 。加速增长的指数设定为互联网和万维网相应指数的平均值,即 $\gamma = 1.2$ 。网络的簇系数大小与算法中概率 p 的取值紧密相关, p 的取值越大,生成网络的簇系数越大。特别地,当 $p = 0$ 时,仅考虑了网络的优先连接机制和加速增长特性,生成网络不具有高集聚特性;当 $p = 1$ 时,生成网络的簇系数最大,集聚特性最明显。在数值模拟的过程中,随机生成一个 $[0, 1]$ 之间的随机数 r ,如果随机数 r 的值小于等于 p ,即 $r \leq p$,则采用熟人推荐机制建立连接。这里,我们考虑概率 p 为 0.3、0.6、0.9 3 个具有代表性的取值。利用现有计算机对加速增长的 HK 模型进行建模得到相关仿真数据,所有绘图数据均采用 10 次仿真结果的平均值。

网络的簇系数是一个局部特征量,衡量的是网络的集团化程度。网络的簇系数 C 是网络中所有节点的簇系数的平均值。图 3 给出了加速增长的 HK 模型在概率 p 取值分别为 0.3、0.6、0.9 时网络的簇系数 C 随网络规模 N 的变化。概率 p 的取值越大,生成网络的簇系数越大。当概率 p 的取值一定,网络簇系数 C 随着网络规模 N 的增大逐渐趋于一个常数,最终分别稳定在 0.09、0.21、0.37 左右。

网络的平均最短路径长度是网络的一个重要的全局几何量,描述网络的传输性能与效率,可以对网络的总体特征进行评价。网络的平均最短路径长度定义为网络中所有节点对之

间的最短路径长度的平均值。图 4 分别给出了加速增长的 HK 模型在概率 p 取值分别为 0.3、0.6、0.9 时平均最短路径 $\langle L \rangle$ 随网络规模 N 的变化规律。平均最短路径随网络规模呈对数形式增长。由图 3 和图 4 可知,加速增长的 HK 模型不仅再现了真实网络的集聚特性,而且具有小的平均最短长度,生成的网络具有明显的小世界效应。

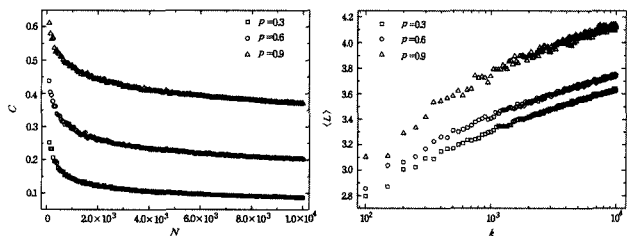


图 3 簇系数随网络规模的变化图 图 4 平均最短路径随网络规模的变化图

图 5 给出了本文所提模型在网络规模 $N=10^4$ 时的 2 到 5 阶集团度分布,其中正方形、圆圈和三角形分别代表 p 值为 0.3、0.6 和 0.9 3 种情况。可以看出,这些分布都近似服从幂律分布,再现了真实网络的幂律集团度分布特性。更进一步,我们将改进的 HK 模型和经典的 HK 模型进行了比较,图 6 给出了 Holme 和 Kim 提出的经典 HK 模型在网络规模 $N=10^4$ 时生成网络的 2 到 5 阶的集团度分布。由图 6 可知,即使在概率 p 的取值为 0.9 时也只存在少量的 4 阶集团,不存在 5 阶集团。可见,本文提出的模型更接近真实网络,加速增长在网络演化过程中的重要因素对网络中集团的形成起着重要作用。

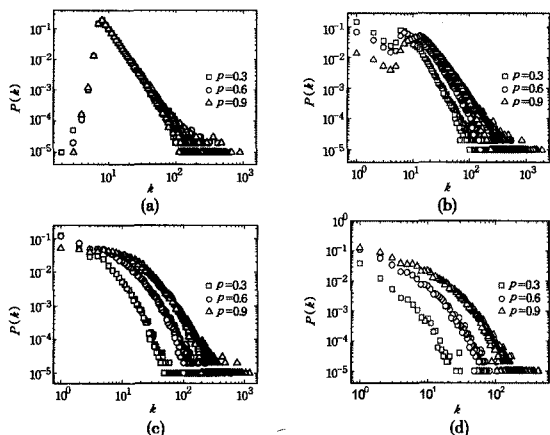


图 5 子图(a)–(d)分别对应为加速增长的 HK 网络的 2 到 5 阶集团度分布

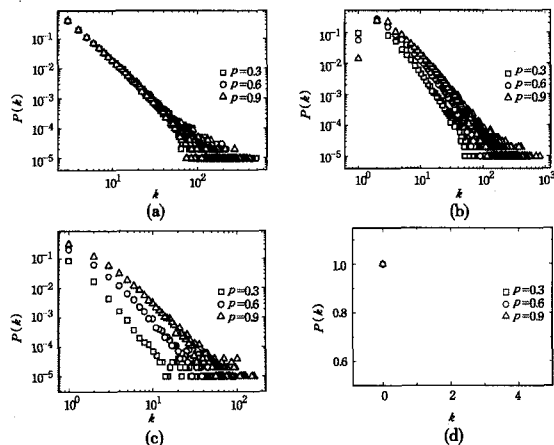


图 6 子图(a)–(d)分别对应为 Holme 和 Kim 提出的 HK 网络的 2 到 5 阶集团度分布

结束语 真实网络中广泛存在的低阶幂律集团度分布特性为网络演化模型提供了新的比较平台。本文对真实网络的集团度分布特性进行了研究,考虑了网络演化过程中的加速增长因素,提出一种加速增长的 HK 模型。每个时间步,增加一个新节点和 Δm 条连边,其中 Δm 不再是一个固定的常数,而是随时间不断增大。数值仿真结果表明,本文模型生成的网络不仅具有和经典 HK 模型生成网络一样的小世界效应和无标度特性,而且再现了真实网络中的低阶幂律集团度分布特性。未来的工作将进一步深入研究真实网络中的加速增长对其集团度分布的影响和基于真实网络研究加速增长的具体表现形式对集团度分布指数的影响。进一步地,通过在加速增长的 HK 模型上研究信息传播、数据包传递和网络同步等问题,可以观察网络的局部结构对网络动力学性质的影响。

参考文献

- [1] Strogatz S H. Exploring Complex Networks[J]. Nature, 2001, 410:268-276
- [2] Albert R, Barabasi A L. Statistical Mechanics of Complex Networks[J]. Reviews of Modern Physics, 2002, 74(1):47-97
- [3] Newman M E J. The Structure and Function of Complex Networks[J]. SIAM Review, 2003, 45(2):167-256
- [4] Boccaletti S, Latora V, Moreno Y, et al. Complex Networks: Structure and Dynamics[J]. Physics Reports, 2006, 424(4-5):175-308
- [5] 吴金冈, 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] 周涛, 柏文洁, 汪宏宏, 等. 复杂网络研究概述[J]. 物理, 2005, 34(1):31-36
- [7] 方锦清. 网络科学的理论模型探索及其进展[J]. 科技导报, 2006, 24(12):67-72
- [8] Watts D J, Strogatz S H. Collective Dynamics of small-world Networks[J]. Nature, 1998, 393(6684):440-442
- [9] Barabasi A L, Albert R. Emergence of Scaling in Random Networks[J]. Science, 1999, 286(5439):509-512
- [10] Barabasi A L. Scale-Free Networks: A Decade and Beyond[J]. Science, 2009, 325(5939):412-413
- [11] Alon U. Network motifs: theory and experimental approaches [J]. Nature Review Genetics, 2007, 8:450-461
- [12] Xiao W K, Ren J, Qi F, et al. Empirical study on clique-degree distribution of networks[J]. Physical Review E, 2007, 76(3):037102
- [13] 周涛, 肖伟科, 任捷, 等. 网络集团度的幂律分布[J]. 复杂系统与复杂性科学, 2007, 4(2):10-17
- [14] Holme P, Kim B J. Growing Scale-free Networks with Tunable Clustering[J]. Physical Review E, 2002, 65(2):026107
- [15] Dorogovtsev S N, Mendes J F F. Effect of the Accelerating Growth of Communications Networks on their Structure[J]. Physical Review E, 2001, 63(2):025101
- [16] Dorogovtsev S N, Mendes J F F. Accelerated Growth of Networks[J]. arXiv:cond-mat/0204102, 2002
- [17] Zhang G Q, Zhang G Q, Yang Q F, et al. Evolution of the Internet and its Cores[J]. New Journal of Physics, 2008, 10:123027
- [18] Border A, Kumar R, Maghoul F, et al. Graph Structure in the Web[J]. Computer Networks, 2000, 33(1/6):309-320