

节点数固定的 ER 网络演化模型拓扑性质的研究

李发旭

(青海师范大学计算机科学系, 西宁 810008)

摘要: 复杂网络模型的研究主要集中在 ER 网络模型、小世界网络模型和无标度网络模型, 其中 ER 网络模型是最经典的复杂网络模型之一。文章研究了节点数固定情况下的 ER 网络模型, 根据网络内部边的动态演化特点, 将节点数固定的网络模型进行了细分, 给出了模型的生成算法, 并模拟了算法生成的随机网络的平均路径长度和聚类系数。仿真结果表明, 节点数固定的 ER 网络具有小世界特性, 出现这种小世界特性的主要原因是由于网络内部边的异质性, 而与网络的节点数变化没有必然联系。

关键词: 复杂网络; ER 网络模型; 动态演化; 小世界网络; 节点数固定

中图分类号: TP393 N94

文献标识码: A

近年来, 复杂网络已成为学术界研究的一个热点。现实社会中的很多真实网络, 如 Internet 网、WWW 网、社会网、生物链网、神经网络等均可用复杂网络作为研究工具。研究者大多用随机图理论、统计物理方法和微分动力系统理论等来研究复杂网络^[1-4]。有关复杂网络模型的研究主要集中在 3 种重要的复杂网络模型——随机网络模型 (ER 模型)^[1, 5]、小世界网络模型 (WS 模型)^[2, 4, 6-7] 和无标度网络模型 (BA 模型)^[8-11] 上。其中, ER 模型是最经典的复杂网络模型之一, 该模型包含有 N 个节点, 每个节点之间连接边的概率为 p , 约 $pN(N-1)/2$ 条边的 ER 随机网络。WS 模型的特点是具有小的平均最短路径长度和较高的聚类系数。Barabási 和 Albert 在其 BA 模型中最早观察到一些复杂网络模型的度分布服从幂律分布^[8], 并分析了产生幂律分布的两个原因: 网络增长性 (growth) 和择优连接性。但是, BA 模型中的线性择优连接假设并不符合所有的现实网络^[2, 4]。

尽管 ER 随机网络作为实际复杂网络的模型存在明显的缺陷, 在 20 世纪的后 40 年中, ER 随机网络理论一直是研究复杂网络拓扑的基本理论, 其中的一些基本思想在目前的复杂网络理论研究中仍然重要。由于节点与边是复杂网络拓扑的两个基本要素, 在节点数固定的 ER 网络演化时, 改变其结构和性质的唯一原因是网络连接边的变化。本文以 ER 随机网络理论为基础, 主要

对节点数固定的 ER 网络模型进行了研究, 依据网络内部边的动态演化特点, 得到了两种不同的 ER 网络生长模型, 给出了模型的生成算法, 并通过数据仿真分析了两类模型生成的 ER 随机网络的拓扑性质。

1 定义

文中的网络拓扑图用 $G = \{V, E\}$ 表示, 其中 $V = \{v_1, v_2, \dots, v_n\}$ 表示网络 G 中的节点集合, $E = \{e_1, e_2, \dots, e_n\}$ 表示网络 G 中边的集合, 节点 v_i 的度 k_i 是与该节点相连的边的数目, 节点 v_i 的邻居节点之间存在的边数用 E_i 表示。

定义 1^[12] 网络的特征路径长度 L 是所有节点对之间的最短路径的平均值, 表示为:

$$L = \frac{1}{n(n-1)} \sum_{i \neq j} d_{ij}$$

其中 d_{ij} 表示节点 v_i 与 v_j 之间的最短路径值。

定义 2^[21] 节点 v_i 的聚类度 c_i 为 v_i 的所有邻居节点之间实际的连接数与理论存在的最大连接数之比, 表示为:

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

定义 3^[21] 平均聚类系数 C 定义为网络中所有节点的聚类系数的平均值, 表示为:

$$C = \frac{1}{n} \sum_{i=1}^n C_i$$

收稿日期: 2010-03-11

基金项目: 国家自然科学基金项目资助 (60863006)

作者简介: 李发旭 (1976-), 女, 青海海东人, 讲师, 硕士, 主要从事复杂网络、理论计算机方面的研究。http://www.cnki.net

2 节点数固定的 ER 网络模型生成算法

本文依据节点数固定的 ER 网络模型,先设定具体的网络规模 N (网络节点数),通过分别设定网络的边数 M 和网络中每个节点的最大关联度 K ,生成相应的 ER 网络模型。通过利用该算法生成的 ER 网络模型,可以获得相关数据进行仿真分析。

算法描述如下:

输入:图 G 的节点数 N , 边数 M , 每个节点度的最大关联度 K 。

算法步骤:

Step1: 用 $\text{init}()$ 函数初始化 N 个节点,使得 $\text{visit}(N) = 0$ 。其中, $\text{visit}()$ 数组表示 v_i 节点是否被访问过, 0-表示未被访问, 1-表示已被访问;

Step2 在 N 个节点中任选一个节点 v_i , 编号 i ;

Step3 通过 $\text{rand}()$ 函数随机确定该节点的度 K ;

Step4 通过 $\text{rand}()$ 函数随机得到与节点 v_i 相连的 K 个节点编号,并根据节点数 N 限定节点编号的上限,剔除不符合要求的随机节点编号(即与节点 v_i 相同编号的节点,或者已经与节点 v_i 确定无关联的节点编号等),并保证每个节点的度不大于节点的最大度 K ;

Step5 将产生的符合要求的随机节点编号逐一入队列保存,并修改对应节点对的邻接矩阵 $a_{ij} = 1$,然后在判断队列不空的前提下,从队列中取出节点编号,重复 Step3 直到队列空。如果产生的随机边数与给定边数数相同,跳转到 Step7

Step6 随机产生一个不同于已创建的节点编号,跳转到 Step3

Step7 判断构建的随机网络图的边数是否符合要求,如果符合,保存该随机图,跳转到 Step8 否则,跳转到 Step6

Step8 输出构建的随机网络的邻接矩阵,算法结束。

3 节点数固定的两类 ER 模型的仿真分析

ER 随机网络的平均度 $\langle k \rangle = p(N-1) \approx pN$,为了数据仿真分析需要,在这里将网络的平均度定义为 $\langle k \rangle = 2M/N$,因此根据网络边和平均度之间的关系,将网络内部边的演化模型分成两类:

模型 A 网络的平均度 $\langle k \rangle$ 保持不变,将每个节点的最大关联度 K 逐步增加,即边数固定情况下边的随机化重连。

模型 B 网络中每个节点的最大关联度 K 保持不变,将网络平均度 $\langle k \rangle$ 逐步增加,即边数增加情况下的随机化加边模型。

在模型 A 的演化过程中,网络的平均度 $\langle k \rangle$ 不变,网络中每个节点的最大关联度逐步增加,也就是说,随机选择一条边复制后重新连接它的两个节点,以此来

模拟网络的内部边的动态演化。在模型 B 的演化过程中,网络中每个节点的最大关联度 K 不变,网络平均度 $\langle k \rangle$ 逐步增加,也就是说在随机选择的一对节点之间加上一条边,则其度分布也会不断改变,以此来模拟随机网络生长的情况。

在模型 A 中,网络的平均度 $\langle k \rangle$ 不变,随机连接一条边的两个节点发生了改变,在模 B 中,网络的平均度 $\langle k \rangle$ 变化,随机选择的一对节点之间加上一条边,因此模型 B 在演化过程中其随机性比模型 A 增加得更快。对于模型 A 和 B 的平均最短路径和聚类系数,数值模拟结果如图 1 和图 2 所示。数值是当网络的节点数 $N = 20$ 节点的最大关联度分别是 $K = 10\ 15\ 19$ 网络的平均度分别是 $\langle k \rangle = 4\ 5\ 6\ 7\ 8\ 9\ 10$ 时,取 10 次模型的平均值。

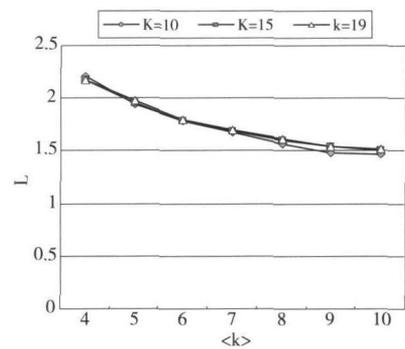


图 1 两类模型的平均路径长度

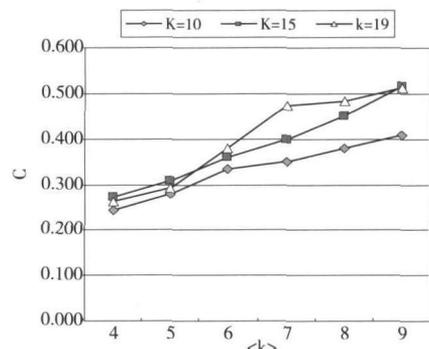


图 2 两类模型的平均聚集系数

直观上,对于 ER 随机网络中选取的一个点,网络中大约有 $[k]^L$ 个其他的点与该点之间的距离等于或非常接近于 L , L 是 ER 随机网络的平均路径长度,即 $L \propto \ln N / \ln \langle k \rangle$ 。这种平均路径长度为网络规模的对数增长函数的特性就是典型的小世界特性。由图 1 可知,随着 $\langle k \rangle$ 的增加,整个网络中的总边数不断增加,其平均最短路径 L 按 $\ln k$ 不断减小,当网络的平均度 $\langle k \rangle$ 不变,节点的最大关联度 K 不断增加时,网络的平均最短路径 L 几乎不变,这说明了在节点数固定条件下的 ER 模型的平均最短路径 L 只与网络的平均度 $\langle k \rangle$ 有关与节点的最大关联度无关。在两种网络模型下,生成的网络都具有较小的平均最短路径 L ,说明

该模型可以作为随机网络的模型。

ER 随机网络中两个节点之间不论是否具有共同的邻居节点, 其连接概率均为 p 。因此 ER 随机网络的聚类系数是 $C = p = \langle k \rangle / N \ll 1$ 这就意味着大规模的稀疏的 ER 随机网络没有聚类特性。但由图 2 可知, 随着网络的平均度 $\langle k \rangle$ 不断增加, 由于整个网络中边数是不不断增加的, 其聚类系数不断增加, 当网络的平均度 $\langle k \rangle$ 不变时, 网络中的总边数不变, 网络的聚集系数在不同的 K 值下没有明显的变化规律, 这说明了在 ER 网络模型中, ER 随机网络中两个节点之间不论是否具有共同的邻居节点, 其连接概率均为 p 时, 由于每条边的出现与否都是独立的, 因此在相同的网络平均度下, 聚集系数没有明显的变化规律。总之, 通过对两类模型生成的随机网络的数据仿真分析, 发现这两类模型生成的随机网络都具有较高的聚集系数 C 。

4 结 论

通过对两种模型生成的 ER 随机网络的数值模拟结果, 可以得到以下两个结论:

(1) 由模型 A 和模型 B 的平均聚类系数和平均最短路径长度的数值模拟结果可知, 它们均具有小的平均最短路径长度和高的聚类系数, 即表现出小世界特性。

(2) 对模型 A 和模型 B 的数值模拟可知, 复杂网络出现小世界特性主要是由于网络节边的异质性, 而与网络的节点数变化没有必然联系。当然, 虽然两类模型均表现出小世界特性, 但它们的聚类系数与平均最短路径长度具有各自不同的特性。

5 结束语

本文研究了节点数固定情况下的 ER 网络模型, 根据网络内部边的动态演化特点, 得到了节点数固定情况下的 ER 随机网络模型的生产算法, 通过该算法生成了两类随机网络, 并对其进行了数据仿真分析, 模拟分析

了这两类模型生成的网络的平均路径长度和聚类系数。仿真分析结果表明, 节点数固定的 ER 网络具有小世界特性, 小世界特性产生源于节点与边的异质性, 与节点数是否改变无关。

参 考 文 献:

- [1] Albert R, Barabási A L. Statistical mechanics of complex networks[J]. RevMod Phys 2002, 74(1): 47-97.
- [2] Watts D J Strogatz S H. Collective dynamics of small-world networks[J]. Nature 1998, 393(6): 440-442.
- [3] Strogatz S H. Exploring complex networks[J]. Nature 2001(8), 410 268-276
- [4] Newman M E J Moore C, Watts D J Mean-field solution of the smallworld network model[J]. Phys Rev Lett 2000, 84(14): 3 201-3 204
- [5] Newman M E J The structure and function of complex networks[J]. SIAM Review, 2003, 45(2): 167-256
- [6] Masuda N, Konno N. Transmission of severe acute respiratory syndrome in dynamical smallworld networks [J]. Phys Rev E, 2004, 69(3): 031917.
- [7] Zanette D H. Dynamics of rumor propagation on small-world networks[DB/OL]. arXiv 0110324 2001.
- [8] Barabasi A L, Albert R. Emergence of scaling in random networks[J]. Science 1999 286 509-512
- [9] Barabasi A L, Albert R, Jeong H. Mean-field theory for scale-free random networks [J]. PhysA, 1999, 272 173-187.
- [10] Dorogovtsev S N, Mendes J F F, Sanukhin A N. Structure of growing networks with preferential linking[J]. Phys Rev Lett 2000 85 4 633-4 636
- [11] Barabasi A L, Bonabeau E. Scale-free networks[J]. Scientific American, 2003(5), 52 60-69
- [12] 汪小帆, 李翔, 陈关荣. 复杂网络理论及其应用[M]. 北京: 清华大学出版社, 2006

Research on Topological Properties of Fixed Nodes ER Networks

LI Faxu

(Computer Department of Qinghai Normal University, Xining 810008, China)

Abstract The complex network have three kinds of major model: ER model, smallworld network model and the scale-free network model, and the ER model is the most classical one. In this paper, we mainly study the ER network model based on fixed nodes. According to the dynamic evolution characteristic of the ER network, sorting the ER networks into two different kinds, and the model production algorithm is given. According to two different modes of connective edge, average shortest path lengths and clustering coefficients are compared. It has been indicated that ER networks with fixed nodes have small-world characteristics, and the primary reason of phenomena existing is different methods of connective edge, but not node's number.

Key words complex network, ER random graph model, smallworld model, fixed nodes