

文章编号:2095-3046(2012)01-0073-05

具有局部无标度特性的小世界网络模型

郭 印, 刘维清

(江西理工大学理学院, 江西 赣州 341000)

摘 要:提出了一种构建复杂网络模型的方法,利用该方法构建的复杂网络模型整体具有小世界效应,而局部具有无标度特性.详细分析了模型参数对复杂网络的主要特征如聚类系数,平均最短路径的影响;通过与实际网络的主要特征进行对比分析,结果表明利用该模型构建的复杂网络与实际网络具有相近的特征.

关键词:复杂网络;局部无标度特性;优化的小世界模型

中图分类号:TP393.02 **文献标志码:**A

Model of small world network with local scale free structure

GUO Yin, LIU Wei-qing

(Faculty of Science, Jiangxi University of Science and Technology, Ganzhou 341000, China)

Abstract: A kind of complex network model which has small world character with local structure being scale free structure is proposed. The effects of the parameter in the model on the cluster coefficient and the average shortest path are analyzed. The results indicate that this model of complex network has similar character with the real networks.

Key words: complex network; local scale free structure; optimal small world model

0 引 言

人类社会中存在形形色色的复杂系统,可以通过复杂网络来进行描述.通常将系统中的个体定义为(抽象的)节点,而将个体之间存在的关联途径定义为节点之间的边,通过这种方式来构建网络.

用网络的观点描述客观世界起源于1736年德国数学家 Euler 解决哥尼斯堡七桥问题.研究人员做了大量的工作,尤其是从1998年 Watts 和 Strogatz 在 Nature 上发表文章^[1]指出许多现实世界中的复杂网络不是规则网络和随机网络,而是具有 small-world 性质,即网络具有小的最短路径和高的聚类系数.这篇文章揭开了对复杂网络研究的序幕.1999

年,Barabasi 和 Albert 在 Science 上发表文章^[2]指出,许多现实世界的复杂网络还具有 scale-free 的特点,进一步引起了人们对复杂网络研究的热情,这两种网络模型的提出开启了复杂网络研究的新纪元,标志着复杂网络的研究进入了崭新的时代^[1-2].其中,利用节点增长与优先连接这两个机制构造的 BA 模型具有简明的网络演化规则,可以用来描述许多真实的网络,但仍然有一定的局限性.例如 BA 模型生成的网络的度分布指数 $\gamma=3$,而现实中的各种复杂网络中的度指数 γ 则在 1~3 之间^[3-4],甚至有些复杂网络还具有一定的非幂率特性(指数截止,小 k 饱和等^[5-7]).为了让生成的复杂网络更接近于现实网络,人们提出了许多改进的复杂网络生长模型如增长机制、各种类型的优先连接机制、内部演

收稿日期:2011-09-05

基金项目:江西理工大学研究生创新专项资金项目(YC05);江西省科技厅资助项目(2010GQW0021)

作者简介:郭 印(1983-),男,硕士研究生,主要从事复杂网络与病毒传播等方面的研究,E-mail: guoyin126@126.com.

化、局部事件、竞争等^[8-10]。刘宗华和来颖成^[11]指出经典的 BA 模型展示的幂律特性和 ER 随机网络展示指数度分布是两种极端情况,而自然界中的大部分网络的度分布介于这两者之间,从而提出采取具有部分优先,部分随机的生长机制的刘—来模型。这些机制在某些方面很好地符合实际情况,但又各具局限性。

自然界和人类社会中广泛存在各种类型的复杂网络,例如社会领域的社会关系网络、科学家合作者网络;信息领域的 WWW 网络、论文应用网络;技术领域的自治层 Internet、软件类网络;生物领域的食物链网络、新陈代谢网络、蛋白质相互作用网络等^[12-13]。这些现实世界中的网络整体上具有较短的平均最短路径和较大的聚类系数,即具有小世界效应,而在局部(某一范围内的小集团)具有少数节点与大量其它节点相连,多数点具有较少的连接,也就是说在小集团内部存在有少量核心节点,以这些核心节点为核心形成的子网络具有无标度的特性,这种网络的拓扑结构如图 1 所示。现实网络中的这些特性无法用以上提及的这些网络模型来实现。因此,提出一种优化的小世界模型(OSM 模型),以实现整体上具有小世界效应而局部具有无标度特性的复杂网络。并分析该模型的参数对所生成的复杂网络的平均最短路径和簇类系数的影响。

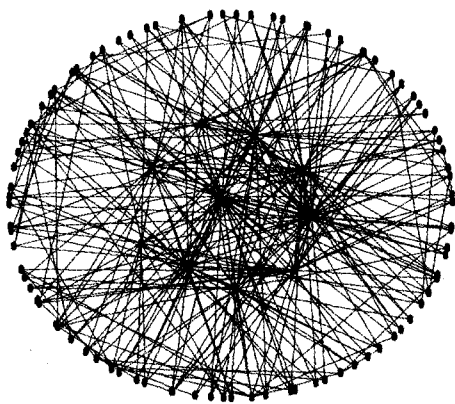


图 1 局部具有无标度特性的小世界网络示意图

1 OSM 模型

现实网络中,往往存在一些具有很好的连通性的核心节点。以这些核心节点为主体构成网络中的小集团。不同的小集团子网络之间具有随机连接性。例如:人们在查阅文献资料时,通常是先查找相关性高或者是被引用次数较高的文献,然后通

过该文献的参考文献,寻找相关的文献,或者进一步通过该文献的参考文献来寻找文献,然后再在这些文献中选择自己感兴趣的文献。

因此,本文拟采用以下规则来实现 OSM 模型,具体生成 N 个节点的 OSM 网络的演化过程主要分为三步:

(1)按 NW^[14]模型,以随机化重连概率 P 生成一个大小为 N_0 ,平均度为 k 的小世界网络。

(2)在所生成的小世界网络基础上,单位时间新加入 1 个节点。新增加的节点与已存在的节点间的连接规则如下:首先随机选择已存在的节点 i 进行连接,然后在节点 i 的邻居和次邻居节点中,按度优先原则选择 M 个节点与新加入的节点进行连接(若所选节点 i 的邻居和次邻居数小于 M 则按实际数量与新加入的节点连接)。

(3)在步骤(2)中新增加的 $(N-N_0)$ 个节点中,随机地选择 N_1 个节点,选中的每个节点再随机地与网络中的 M_2 个的节点连接。

通过 OSM 模型构建的复杂网络能够实现整体上具有小世界效应而局部具有无标度特性。为了描述复杂网络的拓扑结构特性,人们通常采用以下三个最基本统计特征量来刻画复杂网络:平均度(Average Degree, $\langle k \rangle$)、平均最短路径(Average Path Length, APL)、聚类系数(Clustering Coefficient, C)。其中,网络的(节点)平均度是指网络中所有节点的度的平均值,记为 $\langle k \rangle$,即:

$$\langle k \rangle = \frac{1}{N} \sum_{i=1}^N k_i \quad (1)$$

其中, k_i 为节点 i 的度(网络中与节点 i 相连的边数)。 k_i 的大小能直观的反映网络中节点 i 的影响力的大小。

平均路径长度 APL 是指复杂网络中所有可能的任意两个节点之间的最短路径 d_{ij} 的平均值,即:

$$APL = \frac{1}{N(N-1)} \sum_{i \neq j \in V} d_{ij} \quad (2)$$

其中, d_{ij} 是指节点 i 和 j 的最短路径(从节点 i 到节点 j 所需要经过最少的边数)。通常用于反映网络的传输性能。节点的聚类系数是指节点 i 的邻居之间相连的边数 E_i 与节点的邻居之间所有可能的边数之比,记为 C_i 。

$$C_i = 2E_i / [k_i(k_i-1)] \quad (3)$$

其中, k_i 是指节点 i 的度(邻居数)。显然,在这 k_i 个节点之间最多可能有 $k_i(k_i-1)/2$ 条边。整个网络的聚类系数 C 是指网络中的所有节点 i 的聚类系数

C_i 的平均值,即

$$C = \frac{1}{N} \sum_{i=1}^N C_i \quad (4)$$

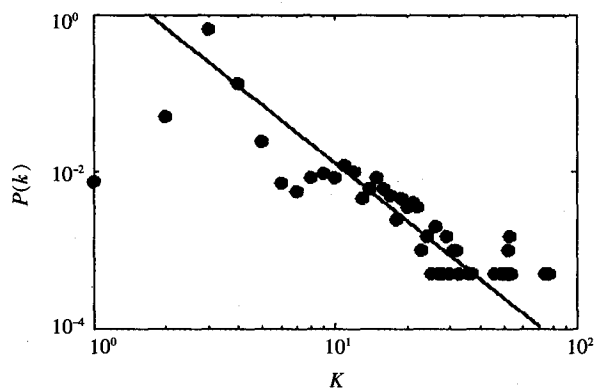
节点的聚类系数反映该节点的一级近邻之间的集团性质,近邻之间联系越紧密,该节点的聚类系数越高。

为了分析依据 OSM 模型所生成的复杂网络的主要特征,我们分别计算了依据 OSM 模型生成的网络的聚类系数、平均最短路径以及度分布等特征量。其中,模型的参数 $N, N_0, K, P, M, N_1, M_2$ 分别取 2000、50、2、0.05、2、200、1。通过与几个真实网络的主要网络参数进行对比分析,我们发现该模型生成网络参数和真实网络比较的相近,该模型生成的网络具有较小的平均最短路径和较大的聚类系数。即该网络整体具有小世界特性。这几个真实网络为:①线虫新陈代谢网络,该线虫新陈代谢网络(*C. elegans*)中的节点表示为代谢物,边表示化学反应,即如果代谢物 A 通过某种生物化学反应得到代谢物 B,则在 A 与 B 之间建立一条边。②美国政治书网络,该网络(Polbooks)是由 V. Krebs 根据 Amazon

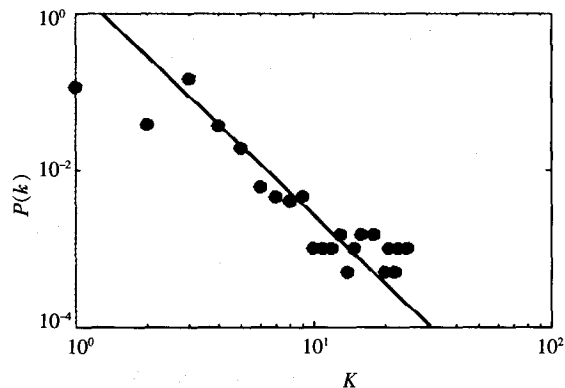
上销售的美国政治相关书籍页面而建立起来的网络。其节点代表在线书店 Amazon 上销售的美国政治相关图书,边代表一定数量的读者同时购买了某两本图书(由抽取网页上的“购买了某本书的读者同时也购买了其他某些图书”而得到)。③空手道俱乐部成员关系网络,该网络(Karate)是由 Wayne Zachary 用三年时间观察了美国一所大学空手道俱乐部成员间的社会关系,并构造出了社会关系网(Zachary's karate club network)。网络的主要特征详见表 1, OSM 模型的整体度分布与现实网络的度分布情况(现实网络中所用的数据均来源于 <http://www.public.asu.edu/~ryang8/resourse.html>)如图 2。

表 1 复杂网络主要特征参数表

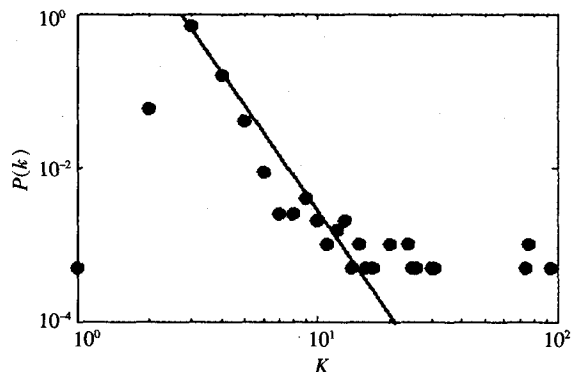
网络类型	平均最短路径	聚类系数	平均度	幂率值
<i>C. elegans</i>	2.4553	0.2924	14.465	2.652
Polbooks	3.0788	0.4875	8.4000	4.706
Karate	2.4082	0.5706	4.5882	4.179
OSM 模型	3.3639	0.2079	5.9940	4.071



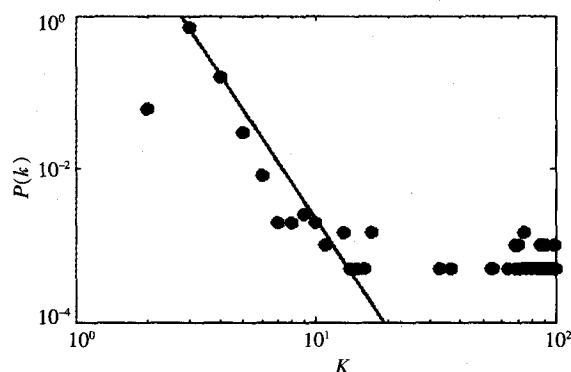
(a) *C. elegans* 中各节点的度分布



(b) Polbooks 中各节点的度分布



(c) Karate 中各节点的度分布



(d) OSM 中各节点的度分布

图 2 OSM 模型与现实网络的度分布情况

2 OSM 网络模型特性

为了考查该模型中 N_0 、 P 、 N_1 、 M_2 这 4 个主要参

数对最终生成网络的聚类系数和平均最短路径的影响,分别给出了这 4 个参数与所生成的网络的聚类系数和平均最短路径的关系图如图 3 所示.

结果表明:OSM 模型构建的网络的聚类系数

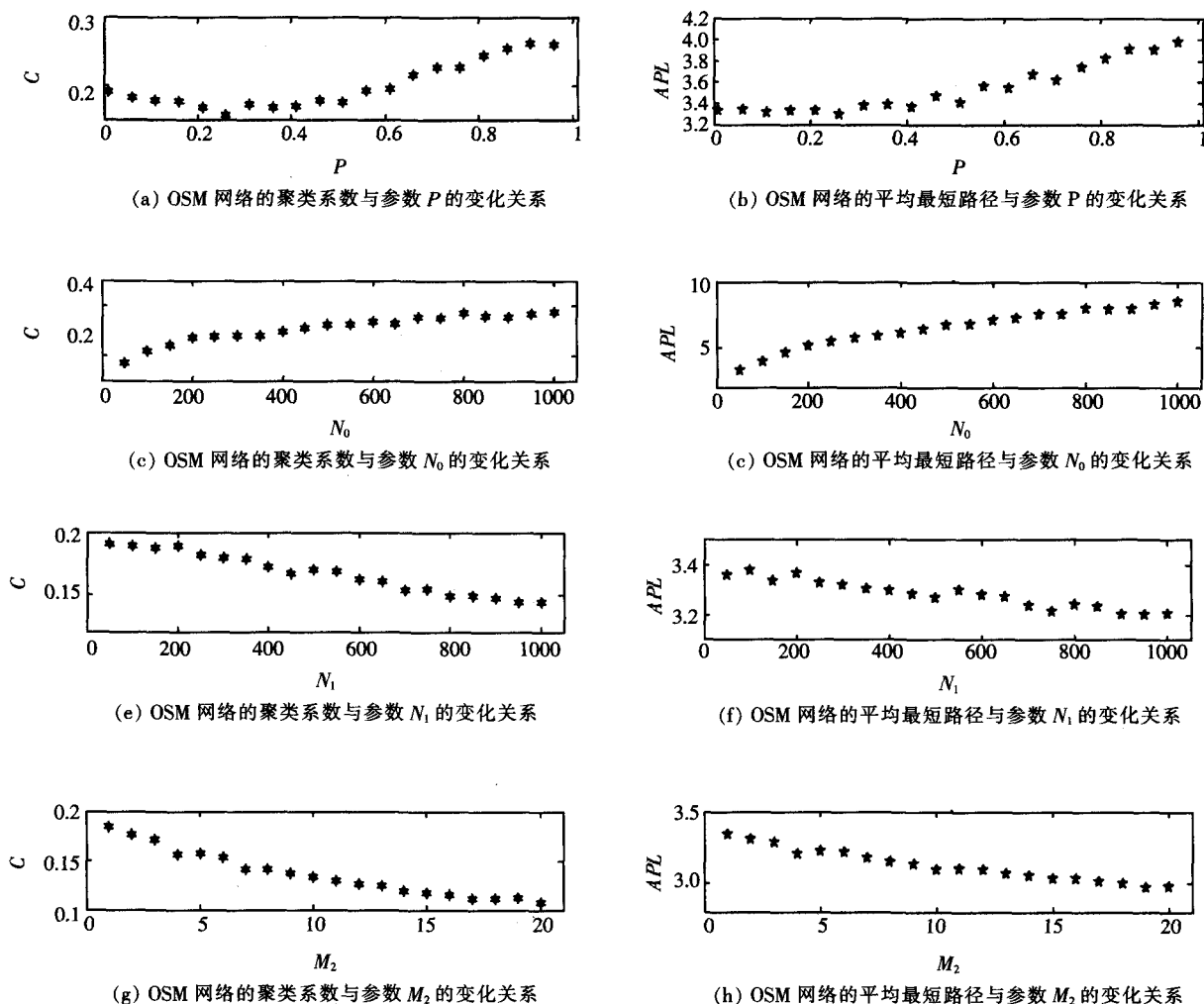
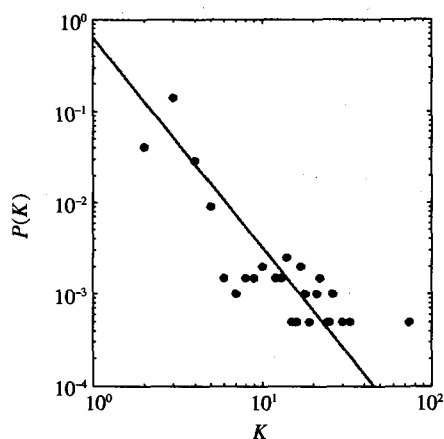


图 3 OSM 模型网络的聚类系数和平均最短路径与主要参数的关系图

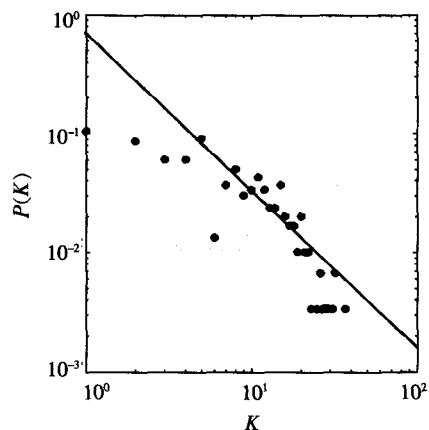
基本参数取 $N=2000$ 、 $N_0=50$ 、 $K=2$ 、 $M=2$ 、 $N_1=200$ 、 $M_2=1$ 、 $P=0.05$

会随着参数 P (初始生成的小世界网络时的断边重连概率) 的增加而先减小再增大. 在 P 约为 0.3 时, 聚类系数具有最小值. 而 OSM 模型网络的聚类系数会随着初始生成小世界网络的规模 (N_0) 增加而增加, 随着随机选择网络节点数 (N_1) 和每个节点连边数 (M_2) 的增加而减少. OSM 模型网络的平均最短路径在 $P < 0.3$ 基本不随 P 变化, 而当 $P > 0.3$ 时会随着断边重连概率 P 值的增大而增加. 而 OSM 模型网络的平均最短路径随着初始生成小世界网络的规模 (N_0) 增加而增加, 随着随机选择网络节点数 (N_1) 和每个节点连边数 (M_2) 的增加而减少.

为了考查网络的局部结构, 即子网络的拓扑结构, 随机选择 OSM 网络中度较大的一个节点, 以该点为核心, 考查其与邻居和次邻居所组成的网络. 从该子网络的度分布如图 4(a) 可知, 该子网络的度分布为幂率分布, 幂率约为 3.062. 这说明以 OSM 模型生成的网络中, 以度较大的节点为核心形成的子网络也具有“马太效应”, 即该网络中的子网络具有无标度的特性. 该结论与实际网络中的子网络度分布具有相同的规律, 我们用同样的方法提取了线虫新陈代谢网络 (*C.elegans*) 中度较大的一个节点形成的子网络, 如图 4(b) 所示, 该子网络的度分布为幂率分布, 其幂率约为 2.619.



(a) OSM 模型中的子网络



(b) 线虫新陈代谢网络的子网络

图4 子网络的度分布图

3 结 论

复杂网络的构建方法很多,虽然这些模型中有很多人工雕琢的痕迹,但是也体现了某些复杂事物发生、发展的若干重要规律.文中提出了一种新的网络构建方法,利用该方法可以构建一种整体具有小世界效应,而以度大的节点为核心形成的局部子网络则具有无标度特性的网络.并分析了模型中的参数对所生成的网络的聚类系数和最短平均路

径的影响.通过与实际网络对比分析表明,OSM网络与现实网络的特征相近.OSM模型为构建与现实网络接近复杂网络提供了一种新方法.

参考文献:

- [1] Duncan J Watts, Steven H Strogatz. Collective dynamics of 'small-world' networks[J]. Nature, 1998(393):440-442.
- [2] Albert-László Barabási and Réka Albert. Emergence of scaling in random networks[J]. Science, 1999, 286(5439):509-512.
- [3] Jon Kleinberg, Steve Lawrence. Network analysis, The structure of the web[J]. Science, 2001, 294(5548):1849-1850.
- [4] Wang X F, Chen G R. Complex networks:small-world,scale-free and beyond[J]. IEEE circuits and systems magazine, 2003, 3(1):6-20.
- [5] Naoki Masuda, Norio Konno. Transmission of severe acute respiratory syndrome in dynamical small-world Networks[J]. Phys Rev E, 2004(69):31917-31923.
- [6] L A N Amaral, A Scala, M Barthe'le'my and H E Stanley. Classes of small-world networks[C]. Proc Natl Acad Sci, 2000(97):11149-11152.
- [7] M E J Newman. Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality[J]. Phys Rev E, 2001(64):16132-16139.
- [8] S N Dorogovtsev, J F F Mendes. Effect of the Accelerating Growth of Communications Networks on their Structure[J]. Phys Rev E, 2001(63):25101(R)-25104(R).
- [9] Réka Albert and Albert-László Barabási. Topology of evolving networks: local events and universality[J]. Phys Rev Lett, 2000(85):5234-5237.
- [10] P L Krapivsky, S Redner, F Leyvraz. Connectivity of growing random networks [J]. Phys Rev Lett, 2000(85):4629 - 4632.
- [11] Z Liu Y-C Lai, N Ye and P Dasgupta. Connectivity distribution and attack tolerance of general networks with both preferential and random attachments[J]. Phys Lett A, 2002(303):337-344.
- [12] 方锦清,汪小帆,郑志刚,等. 一门崭新的交叉科学:网络科学(上)[J]. 物理学进展, 2007, 27(3):239-342.
- [13] 方锦清,汪小帆,郑志刚,等. 一门崭新的交叉科学:网络科学(下)[J]. 物理学进展, 2007, 27(4):361-448.
- [14] M E J Newan, D J Watts. Renormalization group analysis of the small-world network model[J]. Phys Lett A, 1999(263):341-346.