

文章编号 1000-5269(2007)06-0578-05

无标度网络 BA 模型的物理内涵 及其改进模型的比较

贾秀丽^{1,3}, 蔡绍洪^{2,3}, 刘浩广^{1,3}, 汪红^{1,3}, 张芙蓉^{1,3}

(1. 贵州省光电子技术与应用重点实验室, 贵州 贵阳 550025; 2. 贵州财经学院,
贵州 贵阳 550004; 3. 贵州大学 物理系, 贵州 贵阳 550025)

摘要: 建立合适的网络模型有助于理解和研究网络结构和网络的动力学行为。本文考察了复杂网络的 BA 模型以及它的改进模型, 分析比较每个模型的物理内涵及度分布, 讨论了网络演化中择优机制的差异可能对其度分布特征产生重要的影响。并在各模型的基础上提出了一个新的更符合实际的改进模型, 为复杂网络模型的更深一步探索提供了一个新思路。

关键词: 复杂网络; BA 模型; 改进模型; 无标度特征

中图分类号: O414.2 **文献标识码:** A

0 引言

复杂网络理论把真实的复杂系统抽象成一些节点和一些线条, 节点代表研究的对象个体, 线条则代表了这些个体之间的相互作用关系。目前, 复杂网络的研究受到了世界各国的科学家们的广泛关注, 主要原因是它对人们理解真实网络的复杂行为有重要的启示作用, 自然界中存在的大量复杂系统都可以通过各种各样的网络模型来加以描述, 例如, 神经系统可以看成是大量神经细胞通过神经纤维相互连接形成的网络^[1], 计算机网络可以看作是自主工作的计算机通过通信介质如光线、双绞线、同轴电缆等相互连接形成的网络^[2], 此外还有电力网络^[1]、社会关系网络^[1,3,4]、生态网络^[7,8]等等。

目前, 对复杂网络的研究主要集中在三个领域, 一个是网络生成机制及演化模型, 即通过生成机制建立模型, 模仿真实网络行为。二是复杂网络的稳定性, 研究限制条件对网络几何特征的影响, 如复杂网络承受意外故障和恶意攻击的能力等。三是复杂网络上的动力学, 这是人们研究复杂网络的最终目标, 也就是超越网络的拓扑结构, 掌握建立在这些网络上的系统工作方式和机理, 认识复杂网络内部深奥得难以理解的动力学。如复杂网络上的同步、病毒、信息、知识、舆论如何在网络上传播等。

学术界十分关心网络结构的复杂性研究主要原因之一是网络的结构在很大程度上决定了它的功能。网络模型就成了关注的焦点, 1999 年 Barabási 和 Albert 等探索了 WWW 几个大型网络的数据库, 这些实例首次提供了某些大型网络能够自组织成无标度网络的论据, 人们把节点的度服从幂律分布的网络叫做无标度网络。实证结果与理论分析表明这些网络的度分布以幂律衰减, 即 $P(k) \propto k^{-\gamma}$ 。该网络经过长时间的演化后将导致网络中绝大多数点只有少数连接边, 而存在少数点具有大量连接边。为了寻找无标度网络的形成机理, Barabási 和 Albert 发现增长和择优连接在网络演化中起重要作用。由此他们提出了著名的网络演化模型即 BA 模型。

BA 模型把网络的无标度性归结为两个简单的机制。人们在解决实际问题用 BA 模型时发现它的度分布跟实际不完全相符, 人们就想到是否是 BA 模型的形成机理过于简单, 因此, 为了使网络模型的度分布更接近实际网络, 人们在 BA 模型的形成机理的基础上提出了一些更加复杂的演化机理, 这就促使 BA 模型的改进模型的研究得到了迅速发展。本文主要分析 BA 模型的一些改进模型的形成机理, 比较每个模型的度分布, 讨论了网络演化中的一些差异对网络的度分布产生重要的影响。分析每个模型的物理内涵, 比

收稿日期: 2007-10-15

基金项目: 国家自然科学基金 (10647005); 贵州省科学技术基金 (20062006); 贵州省高层次人才特助基金 (2005012); 贵州省教育厅自然科学基金 (2005115)。

作者简介: 贾秀丽, 贵州大学理学院硕士研究生; 研究方向: 非平衡物理复杂性研究。联系方式: Email: lmjxllt@126.com

较每个模型的优缺点,在此基础上,提出了一种更符合实际网络的模型的构建原则,为复杂网络模型的更深一步探索提供了一个新的理论思路。

1 BA 模型

Barabási 等人在研究万维网的度分布时意外地发现它并不服从泊松度分布,进一步研究结果表明许多实际网络有两个重要特征:网络通过增添节点在不断增长和新节点总是择优连接到高连通的节点上。Barabási 和 Albert 提出了 BA 模型[6]。BA 模型第一次把幂律度分布引入网络,它描述的是一个生长的开放系统,从一小组核心节点开始,网络生长的整个过程中,外界会不断地有节点加入这个系统。BA 无标度模型的两个基本假设:(1) 网络增长:即从 m_0 个节点开始,每一个时间步长增加一个新的节点,在 m_0 个节点中选择 m ($m \leq m_0$) 个节点与新节点相连;(2) 择优连接:新节点与一个已经存在的节点 i 相连的概率 Π_i 与节点的度、节点的度之间满足如下关系:

$$\Pi_i = k_i / \sum k_j \quad (1)$$

在经过时间 t 步后,这种算法产生一个有 $N = t + m_0$ 个节点、 mt 条边的网络。BA 网络模型说明网络的节点是不断增加的,但是新节点的加入不是随机的与原网络中的节点相连,而是与原网络中的节点的度成正比,经过长时间的演化,最终演化成幂律指数为 3 的不变网络,即节点的度为 k 的概率与节点的度的 -3 次方成正比,这样就会出现少量的节点获得大量的连接边,而大量的节点只有少数连接边的网络。

BA 网络模型的度具有幂律型分布,其度分布的幂律规律也被证明广泛存在于现实网络。BA 模型重要的意义在于它把实际复杂网络的无标度特性,归结为增长和优先连接这两个非常简单明了的机制。BA 模型首次成功地提供了可以解释实际网络无标度的形成机理。而其中增长和择优连接是许多复杂网络自组织成无标度网络的带有普遍意义的形成机理。当然,这也不可避免地使得 BA 无标度网络模型和真实网络相比存在一些明显的限制。如它是生长网络模型,它只有节点的加入却没有节点的删除和节点间的重新连接,其实一般自然的或者人造的网络更多的是要与外界有节点交换的,而且节点间的连接也是不断变化的,所以说 BA 模型是不能反映现实网络的真实情况的。BA 模型中只能生成度分布的幂律指数固定为 3 的无标度网络,而各种实际复杂网络的幂律指数则不甚相同,且大都是 2 至 3 的范围内。因此 BA 模型的一些改进模型就成为科研工作者广泛关注的研究内容。

2 改进的 BA 模型

实际网络常常具有一些非幂律特征,如指数截断、小变量饱和等,这些都说明真实网络还可能受其它因素的影响,现实网络的演化过程中,各种精微的变化就可能影响网络的拓扑结构,而且不同的现实网络受到诸如成本、老化、竞争等因素的影响,网络演化的差异很大。BA 模型没有考虑到这些因素,为了弥补这些不足,许多研究人员从不同角度提出了对 BA 模型进行改进和推广,使模型能够对真实系统进行更好的描述,下面就对改进做一些简单介绍。

2.1 AB 模型

在不同的网络中,节点及节点间的相互连接情况是不一样的,比如说文献索引网和科学家协作网,在这些网络中,一旦两个节点组成一个连接,那么这个连接就不会被删除,相应的两个节点也不会离开这个网络。而对于 world wide web 中的节点,可以设想它一旦产生就总是保留在网络中,但是已经存在的连接却很可能会被轻易地删除,新的连接也可能被轻易地产生,因为网络节点本身或者节点之间的连接既是不稳定的也是不可靠的。Albert 和 Barabási 考虑到上述情况提出了 AB 模型^[11]。Albert 和 Barabási 提出了三个假设,每一个时间步随机操作下面的三个假设中的一个:(1) 初始有 m_0 个孤立节点,以 p 的概率增加 m ($m \leq m_0$) 条新的内部连接,即在已存在的节点间添加新的边;随机选取一个节点作为新的边的起始点,边的另一个端点由以下概率决定:

$$\Pi_i = (k_i + 1) / \sum (k_j + 1) \quad (2)$$

重复此过程 m 次。(2) 以 q 的概率随机断去 m 条边。随机选择一个点 i 与连接 i 的边 l_{ij} ,然后移走此边,以连接节点 i 和节点 j' 新边 $l_{ij'}$ 取代,每次根据(1)式中的概率选取 j' 来配置一条边,并重复此过程 m 次。

(3) 以 $1 - p - q$ 的概率增加一个节点, 根据(1) 式中的概率分别与网络中已存在的 m 个节点相连接。

显然, 在这改进模型中不仅仅包含有 BA 模型的生长机制和优化机制, 同时也引入了节点连接的变化机制。网络中节点的数目随时间是增长的, 新节点的加入与原网络中的节点的度加 1 成正比, 即与 $(k_j + 1)$ 成正比, 这就保证了那些孤立节点有新的连接。理论分析与模拟结果表明幂律分布与指数分布都有可以在网络中出现, 取决于 p, q 的值, p, q 的取值范围分别为, $0 \leq p < 1, 0 \leq q \leq 1 - p$, 当 $q < q_{\max}^{[10]}$ 时网络的度分布为幂律形式, 当 $q > q_{\max}$ 时为指数分布。当 $q \rightarrow 1$ 时, $1 - p - q \rightarrow 0$, 网络几乎不生长, 此时加边占统治地位, 必然使网络趋向完全网络(完全网络就是网络中任意一个节点都可以与其它的所有节点相连)。

2.2 适应度模型

在对 BA 模型进行改进时, 考虑竞争因素而提出的一种模型。在 BA 无标度网络的增长过程中, 节点的度也发生变化且满足如下幂律关系: $k_i(t) = (t/t_i)^{1/2}$ 其中 $k_i(t)$ 是第 i 个节点在 t 时刻的度, t_i 是第 i 个节点加入到网络中的时刻。上式表明无标度网络中, 越老的节点具有越高的度。然而在实际网络系统中, 节点的度及其增长速度并非只与该节点的年龄有关, 如一些高质量的科研论文在较短的时间内获得大量的引用, 像 Watts 和 Strogatz 关于小世界网络的论文, 以及 Barabási 和 Albert 关于无标度网络的论文发表之后很快就受到关注。显然, 这些例子都是与节点的内在性质相关的。Bianconi 和 Barabási 把这一性质称为节点的适应度, 并由此提出了适应度模型^[9], 适应度模型的两个假设: (1) 网络增长: 从一个具有 m_0 个节点的网络开始, 每次引入一个新的节点, 并且连到 m 个已存在的节点上, 这里 $m \leq m_0$ 。每个节点的适应度按概率分布 $\rho(\eta)$ 选取。(2) 优先连接: 一个新节点与一个已经存在的节点 i 相连的概率 Π_i 与节点 i 的度 k_i 、节点 j 的度 k_j 和适应度 η_i 之间满足如下关系:

$$\Pi_i = \eta_i k_i / \sum \eta_j k_j \quad (3)$$

可以看出, 适应度模型与 BA 无标度模型的区别在于, 在适应度模型中的优先连接概率与节点的度和适应度之积成正比, 而不是仅与节点的度成正比。这样, 在适应度模型中, 如果一个年轻的节点具有较高的适应度, 那么该节点就有可能在随后的网络演化过程中获取更多的边。取决于 $\rho(\eta)$ 的形式, 如果 $\rho(\eta) = \delta(\eta - 1)$, 则每个节点的适应度都一样, 这就与 BA 模型一样。另一种情况就是适应度相对高的节点, 在网络演化中将有绝对的优势获取新的连接, 将会有大量的连接边, 还有一种极限状态, 类似于爱因斯坦凝聚现象。

2.3 陈庆华、史定华模型^[12]

在 BA 模型中, 没考虑到许多实际网络中的节点之间的关系是变化的, 如供销网络中, 消费者随时都可以改变原来的供应商而连接到另一个供应商上; 如 WWW 中, 可能有部分旧的连线在演化中被删除, 陈庆华、史定华认为在删除连线时应与择优连接相反而应采用反择优概率可能更为合理, 他们考虑到这些原因提出三个假设的模型, 每个时间步操作三个假设: (1) 网络增长: 开始时系统有 m_0 个孤立节点, 新加入一个点与系统中原有的 m ($m \leq m_0$) 个不同点相连接, 产生 m 条新边, 且新点与旧点 i 相连接的择优概率由 $\Pi_i = k_i / \sum k_j$ 确定。(2) 旧节点之间择优产生 n 条新边: 新边的两个端点均以择优概率被除选取, 旧节点 i 成为新边一端的概率仍由 $\Pi_i = k_i / \sum k_j$ 确定。(3) 反择优删除 c 条旧连接: 删除的连线的两个端点均以反择优概率被选取, 点 i 成为被删除连线的端点的概率为:

$$\Pi_i = (1 - \Pi_i) / (N(t) - 1) \quad (4)$$

这个网络模型也体现了网络的生长性, 节点的加入到原网络中也不是任意的, 而是要进行择优的, 网络中节点连接边的产生也不是随机的, 而是与节点本身的度成正比, 节点的度越高, 相应产生新连接的概率也越高; 删除边的连接也不是随机的, 而是进行反择优的, 如某个节点增加新边的概率为 80%, 则该节点连接的边被删除的概率为 20%。这个网络模型通过节点的择优连接、加边时的择优连接、删除连接边的反择优连接最终演化成无标度网络, 度分布服从幂律形式, 而且幂律指数是可调的。

3 分析改进模型, 提出新模型的构造假设

BA 模型能够很好解释许多实际网络如 Internet、WWW 以及新陈代谢等网络中连接度分布函数具有

幂律特征的现象,解释这些网络的无标度特征。但是这一模型存在许多不足,一般自然的或人造的网络更多的是要与外界交换物质或能量,而且节点之间的连接也是不断变化的。BA模型只能生成度分布的幂律指数固定为3的无标度网络,而各种实际复杂网络的幂律指数则不甚相同,且大多数属于2至3的范围内,这些说明真实网络在演化过程中还可能受到许多其它因素的影响。AB模型、适应度模型、陈庆华、史定华模型这三种模型都是在BA模型的基础上提出来的,这些模型都是等时间间隔增加新节点来反应网络的增长性,这些模型在一定条件下可退化成BA模型。但是每一种模型都从不同角度对BA模型加以修正,各具侧重点。AB模型在BA模型的两个假设上又增添了连接边的假设,网络中的节点的连接边是变化的,不像BA网络中的节点的连接边一旦形成,它就不会发生变化,而且对BA模型的择优连接假设也进行了修改,把BA模型中的线性择优改为非线性择优,这样这保证了孤立节点有连接边的机会,这样形成的无标度网络的幂律指数也发生了变化,幂律指数由不变的变为可调的。但AB模型连接度分布呈现幂律分布时,可以解释Internet的度分布,幂律指数接近Internet的统计值。但是,AB模型中加边、重连可能会使整个图不连通,且出现自环和重边;在最大连接度、次最大连接度方面与Internet存在较大的偏差。而且加入到网络中的节点就一直在网络中,不发生变化,这一点AB模型未考虑到。适应度模型就是在BA模型的两个假设中的第二个假设加以修正,它不仅考虑了节点的度对以后连接边的影响,而且还考虑了节点本身的适应度,这两方面都考虑就出现了比BA模型更加垄断现象,即适应度高而且度也高的节点获得绝对多的边的连接。但适应度模型只考虑了节点的适应度,节点的生长和择优,没考虑该节点在以后的网络演化中会离开网络,也没考虑连接边的断开、连接的变更,所以它不能刻画网络中节点的删除、断边等。陈庆华、史定华模型在BA模型的基础上增加了删除边和加边的假设,在删除边的假设上面也与其他人的假设有些不同,他们采用了反择优的删除,这样更符合实际网络中的要求,网络中的节点连接边的越少的节点,在以后的演化中,这些连接边就越容易被删除,但没考虑节点离开网络。因此非线性择优、节点的适应度、节点之间的连接变更、去边、加边、这些原因都会影响无标度网络幂律指数的大小。这些模型都能解释实际网络中的一部分网络。

鉴于上述分析,每个模型都它的优点,同时也有它的局限性。现我们考虑这样的一种网络,如world wide web中的节点,节点容易加入到入到网络中,也容易离开网络,同时已经存在的连接却很可能被轻易地删除,新的连接也可能被轻易地产生,当然,当一个节点所有的连接都被删除时就可以认为这个节点本身被删除了,若再次被连接则可视作新节点的加入。再如,在在市场经济中,企业之间的供销关系并非一成不变的,而总是处在不断变化之中,不断的有新企业加入到网络中来,也有旧企业由于各种原因离开了原网络,同时在老企业之间很可能会产生新的连接、老的连接也可能轻易的删除。显然上面的模型都是无法描述这一种类型的网络特征的,因为网络节点本身或者节点之间的连接既是不稳定的也是不可靠的。因此,在BA模型基础上,考虑到原有网络中的节点随时都有可能离开网络,老节点之间的连接既是不稳定的也是不可靠的,即既可以产生新连接,也随时会断开一些老连接。为了描述这种网络的特征,结合上面模型的优点,建立节点和连边都变化的网络模型,具体的构造法如下:

开始有少量 m_0 个孤立节点,在每一个时间步长内操作四个过程:

- (1) 生长择优:增加一个新节点,新节点与网络中原有的 $m(m \leq m_0)$ 个不同点相连接,产生 m 条新边,且新节点与旧节点 i 相连接的择优概率由(1)式确定;
- (2) 在原网络中删除 c 个旧节点:随机地删除原网络中 c 个节点,以及与 c 个节点相连接的连边都被删除;
- (3) 在原网络中增加 r 条新连边:首先随机地选取一个节点,另一个节点根据(1)式的概率确定;
- (4) 在原网络中删除 n 条旧连边:首先随机地选取一个节点,另一个节点由如下概率确定:

$$\prod_i = (1 - \prod_i) / (N(t) - 1) \quad (5)$$

上面的模型来解释电影演员的合作网络,在这个模型中,每一个演员代表一个节点,两个演员在同一部电影中扮演角色就产生一个连接,这个网络通过新演员的不断加入网络是不断增长的(步骤1),在网络中的演员由于各方面的原因会不再出演,那之前与他(她)相连的边就一起跟着断开(步骤2),一个老演员出演一部新电影,会在网络与许多演员产生合作,建立许多新的连接(步骤3),一个老演员由于形象、演

技等问题,原来与他(她)同时演出的演员不再愿意与他(她)合作,他们之间的连接随之断开(步骤4)。这个网络模型除了解释电影演员合作网,还能解释 WWW 网络、企业之间关系的供销网络等。这种模型符合一些实际网络,也能解释很多网络中出现的现象。

4 结束语

复杂网络的研究起步比较晚,但它是以迅猛的速度发展,网络的拓扑结构决定着网络所拥有的特性,许多实际网络都有无标度性质。本文主要分析和讨论了 BA 模型以及它的几个改进模型,BA 模型概括了无标度网络的基本特征,改进模型只是在它的基础上增加一些因素,使网络的性质更接近真实网络,尽管如此,改进模型也不能反映网络的全部性质,只能使它更接近真实网络,网络内节点之间的边断开或重联、不同节点之间的竞争、节点本身的性质、网络中的退化机制、使网络中节点的度分布发生变化。

本文在分析了以上模型不足的基础上,综合了各模型的优点,提出了一个新的改进模型。我们下一步的工作是采用理论分析的方法对该改进模型的度分布进行研究,如采用连续介质理论、主方程、速率方程等方法进行研究,用计算机进行数值模拟。并用以解释现实中存在的一些实际网络。

参考文献:

- [1] WATTS DJ, STROGATZ SH. Collective dynamics of small-world network[J]. Nature(London), 1998;393:440-442.
- [2] FALOUTSOS M, FALOUTSOS P, FALOUTSOS C. On Power-law Relationships of the Internet Topology[J]. ACM SIGCOMM Computer communications review, 1999, 29(4):251-262.
- [3] LILJEROS F, et al. The Web of Human Sexual Contacts[J]. Nature, 2001, 411:907-908.
- [4] EBEL H, MIELSCH LI, BORBHOLDT S. Scale-free topology of e-mail networks[J]. Phys. Rev. E. 2002, 66:035103.
- [5] SEN P, et al. Small-world properties of the Indian Railway network Phys[J]. Rev. E, 2003, 67:036106.
- [6] BARABÁSE A L, ALBERT R. Emergence of scaling in random networks[J]. Science, 1999, 286(5439):509-512.
- [7] MONTROYA, J M, RICARD V. SOLE. Small world patterns in Food webs. SFI. Working papers 00-10-059.
- [8] WILLIAMS, RJ. et al. Two degree of separation in complex in Food webs. SFI. Working papers 01-07-036.
- [9] BIANCONI G, BARABÁSE A L. Bose-Einstein condensation in complex networks[J]. Phys. Rev. Let., 2001, 86:287-296.
- [10] 吴金闪, 狄增加. 从统计物理学看复杂网络研究[J]. 物理学进展, 2004, 24:18-46.
- [11] ALBERT R, BARABÁSE A L. Topology of evolving networks: Local events and universality[J]. Phys. Rev. Let., 2000, 85(24):5234-5237.
- [12] CHEN QINGHUA, SHI DINGHUA. The modeling of scale-free networks[J]. Phys A, 2004, 335:240-248.

Physical Connotation of BA Scale-free Network Model and Comparative Research on Mending Models

JIA Xiu-li^{1,3}, CAI Shao-hong^{2,3*}, LIU Hao-guang^{1,3}, WANG Hong^{1,3}, ZHANG Furong^{1,3}

(1. Laboratory for Photoelectric Technology and Application, Guizhou Guiyang 550025, China;

2. The Institute of finance and economics, Guiyang 550004, China;

3. Department of Physics, Guizhou University, Guiyang 550025, China)

Abstract: Designing a suitable model is owing to understanding network's structure and dynamic behavior. In this paper, Researching on the BA model and its mending models, analyzing and comparing physical connotation and degree distribution of every model, Authors discuss the difference of preference which is important to affect network's degree distribution. Based on these models, a generic model which is according with practical network is designed. It gives a more study for complex network model.

Key words: complex networks BA model mending model scale-free property