

王 林 戴冠中 著

复杂网络的Scale-free性、 Scale-free现象及其控制



科学出版社

www.sciencep.com

复杂网络的 Scale-free 性、 Scale-free 现象及其控制

王 林 戴冠中 著

科 学 出 版 社

北 京

内 容 简 介

复杂网络是复杂系统的高度抽象,在现实世界中存在大量的复杂网络,如社会网络(朋友关系网、科研合作网络等)、技术网络(Internet、万维网及电力网等)、生物网(神经网络、食物链网络及新陈代谢网络)、经济网络等。近年来,在世界范围内掀起了一股复杂网络的研究热潮,且取得了一系列重要成果。

Scale-free (无标度)是复杂网络的一个重要特性。本书致力于系统地介绍复杂网络的 Scale-free 性质的基础知识以及在 Scale-free 性质上所取得的一系列重要成果。主要内容包括:复杂网络的拓扑参数及异质性指标, Scale-free 网络的拓扑结构及鲁棒性, BBS 网络的 Scale-free 性及其社区发现, Scale-free 网络的传播动力学——SIS 模型和 SIR 模型,并讨论了 Scale-free 网络的同步动力学和控制问题。本书还介绍了作者在这些领域的相关工作。

本书适合理工科大学研究生、博士生和教师阅读,也可供自然科学和工程技术领域的研究人员参考。

图书在版编目(CIP)数据

复杂网络的 Scale-free 性、Scale-free 现象及其控制/王林,戴冠中著.
—北京:科学出版社,2009

ISBN 978-7-03-025600-3

I. 复… II. ①王… ②戴… III. 计算机网络—研究 IV. TP393
中国版本图书馆 CIP 数据核字 (2009) 第 167371 号

责任编辑:姚庆爽/责任校对:陈玉凤
责任印制:赵 博/封面设计:耕者设计工作室

科学出版社出版

北京东黄城根北街 16 号
邮政编码:100717

<http://www.sciencep.com>

陈海印刷有限责任公司印刷

科学出版社发行 各地新华书店经销

*

2009 年 9 月第 一 版 开本: B5(720×1000)
2009 年 9 月第一次印刷 印张: 15
印数: 1—2 500 字数: 292 000

定价: 50.00 元

(如有印装质量问题,我社负责调换〈路通〉)

前 言

自然界、生物界和人类的工程界、经济界、社会界中的大量复杂系统是由在微观层次上的海量个体所组成,个体之间存在着作用,如物理作用、化学作用、实际(real)连接、虚拟(virtual)联系、资金链、朋友关系等等。如果把个体抽象为网络节点,而个体之间的作用抽象为节点之间的边,则复杂系统就可用一个复杂网络来描述。复杂网络描述了复杂系统的拓扑结构特性,研究复杂网络往往是研究复杂系统的一个切入点。

复杂网络(如地震网络、食物链网络、蛋白质网络、Internet、万维网、电力网、航空网、铁路网、演员合作网、科学引文网、朋友关系网、电子邮件网以及金融网络等等)主要研究其个体之间相互作用所产生的系统的整体性质与行为。深入研究复杂网络,可以揭示隐藏在复杂系统(其中一些复杂系统关乎国计民生)中的共同规律,这种一般性规律的揭示对于把握复杂系统的宏观特征,对于调节复杂系统上的动力学行为都将具有重要意义。

人们对复杂网络的研究历来已久。匈牙利数学家 Erdős 和 Rényi 最先将网络作为一个整体并利用严格的数学方法进行研究。他们将大型网络看成一个随机图,并建立起一套优美的随机图理论。然而,1998年和1999年分别在 *Nature* 和 *Science* 刊物上发表的两篇文章(Scale-free 网络、小世界网络)打破了随机图理论长达40年的统治地位,并宣告了现代复杂网络这一新的学科分支的诞生。

复杂网络中节点之间的连接数、网络访问数、社会关系的联系数等等称为节点的度。复杂网络中存在着不同的度分布,其中有两类分布比较特殊。一类是比较均匀的分布,它可以用泊松(Poisson)分布或高斯(Gauss)分布来描述,此时,复杂网络中大部分节点的度值分布在其均值附近,度值远大于或远小于均值的节点数都很少,Erdős 和 Rényi 研究的网络即属于这类。另一类是严重不均匀的分布,它可以用幂律分布(Power-law)来描述,此时,复杂网络中大部分节点的度值很小,少量节点(称为 Hub 点)的度值甚大。1998年,Barabási 等把度分布为幂律分布的复杂网络称为 Scale-free 网络,亦即这类网络具有 Scale-free(无标度)性。之所以冠以“无标度”这一如今非常流行的术语,是因为幂律分布函数具有标度不变性。

近年来,科学界兴起研究 Scale-free 网络的热潮,其主要原因有三:

(1) Scale-free 性质的普遍性。众多科学家用实证研究方法,相继发现无论在自然形成的系统(自然界、生物界)中,或者人造的或有人参与的系统(工程界、经济界、社会界)中,大量复杂网络的度分布均具有幂律分布特性。

(2) Scale-free 现象的非寻常性。研究表明 Scale-free 网络 (主要是 Hub 点) 具有非同寻常的动力学行为。

(3) Scale-free 问题的重要性。人类工程界、经济界、社会界中的一些 Scale-free 系统对国计民生有重要意义, 我们必须对其加强研究。

在前人研究的基础上, 本书从复杂网络的 Scale-free 性、复杂网络所呈现的 Scale-free 现象及其控制等多个方面对复杂网络进行了系统深入的研究, 主要包括如下三个方面的内容:

(1) 复杂网络的 Scale-free 性质。主要从 Scale-free 网络幂律分布的度分布指数的变化范围、Scale-free 网络的 Hub 点的数量及其连接数、复杂网络的连接率和吸引率以及复杂网络的社区发现等方面对复杂网络的 Scale-free 性质进行了探讨。

(2) 复杂网络上的 Scale-free 现象。主要研究 Scale-free 网络在内、外部作用下的动力学行为, 包括复杂网络上的两个传染病模型 (SIS 模型和 SIR 模型) 的稳定性问题、传染病模型趋于稳态后系统遭到破坏的程度与复杂网络的 Scale-free 性质的关系等。

(3) 复杂网络的控制。主要研究复杂网络动力学行为所呈现的 Scale-free 现象对于复杂网络 Scale-free 性质的依赖性。复杂网络上的 Scale-free 现象表现出正、负两面性, 我们研究了影响 Scale-free 现象的两面性的相关因素及其控制问题, 并提出了关于复杂网络的一般性控制框架。

期望本书能起到抛砖引玉的作用, 促进不同领域的科学工作者密切注视复杂网络的 Scale-free 问题, 深入对其进行研究。

本书的出版得到国家 863 计划项目和陕西省自然科学基金的资助, 在此深表谢意; 感谢西北工业大学和西安理工大学的覃森、李海林、张婧婧、付宁、赵焕成等研究生, 他们做了大量的辅助工作。

作者

于西北工业大学

目 录

| | |
|--------------------------------|----|
| 前言 | |
| 第 1 章 绪论 | 1 |
| 1.1 研究背景及意义 | 1 |
| 1.2 国内外研究现状 | 3 |
| 1.3 本书主要研究内容及框架 | 8 |
| 参考文献 | 10 |
| 第 2 章 复杂网络的两个拓扑参数 | 15 |
| 2.1 连接率 | 15 |
| 2.2 吸引率 | 17 |
| 2.3 复杂网络的一个静态概率模型 | 19 |
| 2.4 定理 2.1 的证明 | 25 |
| 2.5 小结 | 36 |
| 参考文献 | 37 |
| 第 3 章 复杂网络的异质性指标 | 38 |
| 3.1 度分布指数 | 38 |
| 3.2 网络结构熵 | 38 |
| 3.3 度分布熵 | 40 |
| 3.4 洛伦兹曲线与基尼系数 | 43 |
| 3.5 异质性指标的比较 | 51 |
| 3.6 小结 | 52 |
| 参考文献 | 52 |
| 第 4 章 Scale-free 网络的拓扑结构及其形成机制 | 54 |
| 4.1 度分布指数的取值范围 | 54 |
| 4.2 Scale-free 网络中的 Hub 节点 | 58 |
| 4.3 Scale-free 性质的形成机制 | 64 |
| 4.4 小结 | 65 |
| 参考文献 | 65 |
| 第 5 章 复杂网络的社区发现理论与应用 | 67 |
| 5.1 复杂网络中的社区结构 | 67 |
| 5.2 社区结构发现算法 | 68 |

| | | |
|--------------|-------------------------------------|------------|
| 5.3 | N-G 模块度 | 75 |
| 5.4 | 关于 N-G 模块度的进一步讨论 | 78 |
| 5.5 | 复杂网络的自相似特性与社区分解 | 80 |
| 5.6 | BBS 网络的 Scale-free 性质 | 93 |
| 5.7 | 复杂网络中的社区发现在 BBS 热点话题发现中的应用 | 98 |
| 5.8 | 小结 | 101 |
| | 参考文献 | 102 |
| 第 6 章 | 复杂网络的中心化理论与应用 | 104 |
| 6.1 | 复杂网络中心化问题的发展综述 | 104 |
| 6.2 | 复杂网络的重要中心化指标 | 107 |
| 6.3 | 建立在现实系统中的网络中心化研究 | 114 |
| 6.4 | 建立在拓扑层次上的网络中心化研究 | 122 |
| 6.5 | 网络中心化程度的量化与中心化指标的评定 | 134 |
| 6.6 | 网络中心化与网络脆弱攻击关系的讨论 | 139 |
| 6.7 | 小结 | 143 |
| | 参考文献 | 144 |
| 第 7 章 | Scale-free 网络的鲁棒性 | 147 |
| 7.1 | 静态鲁棒性 | 147 |
| 7.2 | 雪崩动力学 | 148 |
| 7.3 | 小结 | 149 |
| | 参考文献 | 150 |
| 第 8 章 | 复杂网络中病毒传播的稳态和瞬态特征 | 151 |
| 8.1 | 引言 | 151 |
| 8.2 | 网络病毒传播模型 | 156 |
| 8.3 | 病毒传播的稳态特性 | 167 |
| 8.4 | 病毒传播的瞬态特性 | 176 |
| 8.5 | 小结 | 185 |
| | 参考文献 | 186 |
| 第 9 章 | Scale-free 网络上的传播动力学——SIS 模型 | 188 |
| 9.1 | 引言 | 188 |
| 9.2 | 复杂网络中 SIS 模型的平衡点的若干性质 | 189 |
| 9.3 | 复杂网络中 SIS 模型的解的若干性质 | 193 |
| 9.4 | 复杂网络中 SIS 模型的稳定性分析 | 198 |
| 9.5 | 复杂网络中 SIS 模型的收敛速度分析 | 199 |
| 9.6 | Scale-free 网络中 SIS 模型的感染规模分析 | 203 |

| | |
|--|------------|
| 9.7 小结 | 205 |
| 参考文献 | 206 |
| 第 10 章 Scale-free 网络上的传播动力学——SIR 模型 | 207 |
| 10.1 引言 | 207 |
| 10.2 任意连接边指向移除节点的概率 | 209 |
| 10.3 若干引理 | 211 |
| 10.4 复杂网络中 SIR 模型的病毒爆发的阈值 | 216 |
| 10.5 复杂网络中 SIR 模型的病毒爆发规模分析 | 216 |
| 10.6 小结 | 220 |
| 参考文献 | 220 |
| 第 11 章 Scale-free 网络上的同步动力学 | 221 |
| 11.1 复杂网络中的同步 | 221 |
| 11.2 主稳定函数法 | 222 |
| 11.3 网络的异质性对于同步的影响 | 223 |
| 11.4 小结 | 225 |
| 参考文献 | 225 |
| 第 12 章 Scale-free 网络的控制 | 226 |
| 12.1 一般性框架 | 226 |
| 12.2 Internet 拓扑结构的演化控制模型 | 228 |
| 12.3 小结 | 231 |
| 参考文献 | 232 |

第1章 绪 论

1.1 研究背景及意义

复杂系统由相互作用的众多子系统组成,当把子系统抽象成节点,把子系统之间的相互作用抽象成节点之间的边,则复杂系统就可以抽象成一个复杂网络。复杂网络的复杂性通常体现在两个方面:其一,复杂网络常常包含海量的节点,即节点数量非常庞大;其二,复杂网络中节点之间的连接关系通常较为复杂,既不像规则网络那样具有完全确定的连接关系,也不像随机网络那样具有完全不确定的连接关系,而是介于两者之间。

复杂网络广泛存在于自然界、生物界、工程界和人类社会界^[1~11],如地震网络、食物链网络、蛋白质网络 (protein interaction network)、新陈代谢网络 (metabolic network)、基因调控网络 (genetic regulatory network)、Internet、万维网、电力网、航空网、铁路网、高速公路网、演员合作网、科学家合作网、科学引文网、朋友关系网、电子邮件网以及金融网络等。深入研究复杂网络,可以揭示隐藏在自然界、生物界、工程界和人类社会界中大量复杂系统(其中一些复杂系统是关乎国计民生的)中的共同规律。这种一般性规律的揭示对于把握复杂系统的宏观特征,对于调节复杂网络上的动力学行为都将具有重要意义。

实证研究表明,绝大多数实际网络在节点之间的连接关系方面表现出惊人的相似。我们用 $P(k)$ 表示网络中节点的度分布,即 $P(k)$ 表示从网络中任选一个节点其度值(与该节点直接相连接的边数)为 k 的概率。对于一个实际网络来说, $P(k)$ 实际上代表的是网络中度值为 k 的节点数所占的比例。在 1998 年以后的几年中,科学界惊奇地发现,绝大多数实际网络的度分布服从幂律分布^[11],而不是像著名数学家 Erdős 和 Rényi 所断言的那样服从 Poisson 分布^[12]。由于幂律分布具有标度不变性,因而将度分布服从幂律分布的网络统称为 Scale-free (无标度)网络。因此,可以断言,绝大多数实际的复杂网络表现出 Scale-free 性。下面我们将列举部分典型实例来验证这一结论。

Abe 和 Suzuki 等研究了地震网络^[1]。他们将所考察的地理区域划分成若干小方格,地震网络的节点指的是其中发生过地震的小方格。在时间上相继发生的两次地震如果分属不同的小方格区域,则在这两个小方格区域所对应的节点之间连上一条边;如果属于同一小方格区域,则在对应的节点上加上一个自环。该地震网络中的节点数及边数显然将随时间的进展而不断增加。Abe 和 Suzuki 等的研究表明

地震网络是一个 Scale-free 网络, 这一结论与传统的 Gutenberg-Richter 定律相吻合, 因为根据 Gutenberg-Richter 定律^[13], 地震强度服从幂律分布。

Montoya 和 Sole 对三种不同类型的食物链网络 (Ythan estuary web、Silwood park web 和 Little Rock lake web) 进行了研究^[2], 发现其中节点的度值服从幂律分布, 该结论说明生态系统也具有 Scale-free 性质。

蛋白质网络探讨的是生物体中蛋白质的相互作用问题, 其中网络的节点是蛋白质, 若两个蛋白质之间存在相互作用, 则在相应的节点之间连上边。Jeong 和 Tombor 等分析了蛋白质网络的结构^[3], 发现除具有其他性质外, 蛋白质网络也是 Scale-free 网络。这说明 Scale-free 性质也存在于生物界。

Internet 可以在两种意义上形成复杂网络: 路由器层和自治系统 (autonomous system) 层。路由器层次上的节点是路由器, 边是连接路由器的物理链路; 自治系统表示 Internet 中遵循共同的路由策略并且受同一个实体控制的 IP 网络及路由器的集合, 自治系统层次上的节点是自治系统, 边是自治系统间的物理链路。Faloutsos 三兄弟对上述两种层次上的 Internet 的结构进行了开创性的研究工作^[4], 除发现了 Scale-free 性质之外, 还发现了许多其他的幂律关系。

万维网是一个有向图, 其中的节点是位于网站服务器中的超文本, 边是从一个超文本指向另一个超文本的超连接 (URLs)。Albert、Jeong 以及 Barabási 构造了一个网络机器人对万维网进行了递归探测^[7,8], 并且发现无论是节点的出度 (从该节点指向其他节点的连接边数) 还是节点的入度 (从其他节点指向该节点的连接边数) 均服从幂律分布。

科学引文网中的节点是科学工作者所发表的文章, 如果一篇文章引用了另一篇文章, 则在相应的两个节点之间连上一条边。Redner 发现科学引文网具有 Scale-free 性质^[5]。

个人收入以及个人财富均服从幂律分布, 这说明幂律关系在经济生活中也是存在的。澳大利亚学者 Matteo 和 Aste 等通过构造财富交换网络说明个人收入以及个人财富可以看成网络中节点的度值^[6], 并且从理论和实证研究两方面论证了节点的度值服从幂律分布, 即财富交换网络具有 Scale-free 性质。

上述例子说明, Scale-free 性质广泛出现在自然界、生物界、工程界以及人类社会界的复杂系统中, 无论是在人造系统中, 还是在自然形成的系统中, Scale-free 性质均会显现出来。有理由认为, Scale-free 性质是一种客观存在。

由于复杂网络上的动力学行为通常依赖于复杂网络的拓扑结构, 因此研究复杂网络的 Scale-free 性质对于深入理解复杂网络上的动力学行为具有重要意义。复杂网络上的动力学行为可以理解成复杂网络的内在性质的外部表现, 复杂网络所具有的 Scale-free 性质一定导致复杂网络上的动力学行为从本质上有别于其他网络 (如规则网络、随机网络等) 上的动力学行为, 因此我们将复杂网络上的动力学行为

所表现出的独特性质统称为复杂网络上的 Scale-free 现象。这样命名主要是为了突出 Scale-free 性质在复杂网络的所有性质中的中心地位。复杂网络上的 Scale-free 现象体现在以下几个方面：

第一，复杂网络的 Scale-free 性质导致其中大多数节点仅拥有很少的连接边数，对于万维网、Internet、科学引文网、新陈代谢网络以及电话呼叫网这样的系统而言，80% 的节点只拥有 30% 的连接，因而这些节点发生故障或者遭受攻击将不会对系统产生根本性的影响，即复杂网络的 Scale-free 性质使得复杂网络对于随机攻击或故障具有很强的鲁棒性^[14~17]。

第二，复杂网络中存在极少数的拥有大量连接边数的 Hub 节点，对于万维网、Internet、科学引文网、新陈代谢网络以及电话呼叫网这样的系统而言，10% 的节点控制了 60% 的连接，而 5% 的节点控制了 50% 的连接。因此，网络攻击者只要花费较少的代价针对 Hub 点进行攻击，就可能造成系统的崩溃，即复杂网络的 Scale-free 性质使得复杂网络对于定向攻击具有极大的脆弱性^[14~17]。

第三，极少数 Hub 节点构成了复杂网络的通信枢纽，因而复杂网络信息传播的效率很高，维护成本较低。

第四，但是，复杂网络对于信息传播的有效性的负面效果是对于病毒传播的不可抗拒性，只要存在病毒节点，这些节点就可以迅速接近 Hub 节点，从而造成大面积的传播。事实上，理论上已经证明，度分布指数位于 2~3 的复杂网络对于病毒的传播不存在阈值^[18]。因而，为防止病毒传播，Hub 节点应该是重点防范对象。

除上述 Scale-free 现象外，复杂网络还存在许多其他能够反映 Scale-free 现象的物理过程和动力学行为，如传播、同步、自组织临界等^[19~30]，这些物理过程和动力学行为均依赖于网络的拓扑结构，即依赖于复杂网络的 Scale-free 性质。

复杂网络上的 Scale-free 现象常常表现出正、负两面性，因而应针对不同情况采取不同措施，即应该对于导致 Scale-free 现象的 Scale-free 性质进行控制，使得复杂网络上的动力学行为朝着有利的方向发展。例如，对于国民经济系统的宏观调控，实际上就是试图调整经济网络的拓扑结构使得国家的经济能够持续、有效并协调发展。

总之，研究复杂网络的 Scale-free 性、Scale-free 现象及其控制问题既具有重要的理论意义，也具有现实的指导意义。

1.2 国内外研究现状

复杂系统作为一门独立的学科出现于 20 世纪 90 年代，而复杂网络作为研究复杂系统的一个行之有效的方法则产生于 90 年代末。一般将具有自组织、自相似、小世界以及 Scale-free 中的一个或几个特性的大型网络称为复杂网络。由 1.1 节的

讨论可以看到, 复杂网络广泛存在于现实世界中。

最早将复杂系统作为网络进行研究的是社会学家, 社会学家研究的是人群关系网络, 其研究的侧重点是考察一个个体对于整个系统的贡献, 强调的是网络中个别节点的作用, 而不是研究网络的整体特性。为此社会学家们定义了许多参数, 如介数 (betweenness)、紧密度 (closeness) 以及度 (degree) 等指标, 这些指标都是用来度量网络中的一个节点在网络中的某种意义之下的重要性^[31]。

首先将网络作为一个整体并利用严格的数学方法进行研究的是匈牙利的数学家 Erdős 和 Rényi^[12]。他们将大型网络看成一个随机图, 并建立起一套优美的随机图理论。构造随机图的过程比较直观: 给定一个固定的概率 $0 \leq p \leq 1$ 以及 $N(N \geq 2)$ 个相互之间不存在任何连接的节点, 以概率 p 在任意一对节点之间连上边。随机图不是一个图, 而是若干图的一个集合 (样本空间), 记为 $G(N, p)$ 。随机图理论所研究的主要问题是当概率 p 由 0 向 1 变化时, 随机图 $G(N, p)$ 的几何性质的变化, 尤其是要确定一个临界概率 p_c , 使得当 $p < p_c$ 时, 随机图不具备某种几何性质; 而当 $p > p_c$ 时, 随机图突然拥有这种几何特性。例如, 可以研究随机图何时连通、随机图何时拥有极大连通分支等性质, 这些性质都是网络的整体拓扑特性。根据随机图理论, 网络中的任意一对节点之间存在连接边的概率相等, 因而节点的度分布 $P(k)$ (节点的度是指节点的连接边数, $P(k)$ 是指任意一个节点的度等于 k 的概率) 服从二项式分布, 当 N 充分大时, $P(k)$ 近似服从 Poisson 分布。因此, 网络中绝大多数节点的度值分布在均值附近, 只有极少数节点具有很大或很小的度值。

从 20 世纪 50 年代到 90 年代末, 随机图理论统治了 40 余年^[11], 在此期间, 人们习惯于将巨型网络理解成随机图, 并利用随机图理论的思想和方法来研究巨型网络。1998 年和 1999 年分别在 *Nature* 和 *Science* 这两个世界级顶尖刊物上发表的两篇文章打破了随机图理论长达 40 年的统治地位^[8,32], 并宣告了复杂网络这一新的学科分支的诞生。

Watts 和 Strogatz 通过对规则网络中少量的边进行重连, 得到了小世界模型。在小世界模型中, 节点之间的平均路径长度较短 (其数量级相当于对节点数取对数), 且平均聚集系数较高。我们称满足上述两条性质的网络具有小世界性质。而在随机网络中, 虽然平均路径长度较短, 但平均聚集系数很低。现实世界中很多网络都具备小世界性质, 如演员合作网、电力网以及线虫 (*Caenorhabditis elegans*) 网络等, 可见小世界性质广泛出现在社会网络、技术网络以及生态网络之中^[32]。

小世界网络中的两条特性在某种意义上体现了网络的自组织功能^[33,34]: 规则网络具有高聚集特性但其网络直径很大, 而随机网络具有较小的网络直径但其聚集性很低。小世界网络正好居于这两种极端情形之间, 即小世界网络处于完全确定与完全不确定之间, 这正是网络根据一定的效用函数自我优化与自我组织的结

果。现实世界中的网络一定是某种组成原则 (organizing principle) 在某种随机外界因素作用下的产物, 在这个意义上, 小世界模型揭示了现实世界中大型网络的一种内在规律。Watts 和 Strogatz 等工作因此受到了广泛的重视。

继实际网络中的小世界特性被揭示之后, Barabási 和 Albert 对万维网进行了深入研究^[7~9], 万维网中 HTML 文档是网络节点, 而 HTML 文档之间的超链接则构成了节点之间的连接边。通过网络机器人, 他们获得了万维网的节点连接关系。Barabási 和 Albert 非常惊奇地发现, 万维网中节点的度分布并不像 Erdős 和 Rényi 的随机图理论所预料的那样服从 Poisson 分布, 而是服从幂律分布, 即

$$P(k) \propto k^{-\gamma}, \quad (1.1)$$

其中, γ 位于 2~3。式 (1.1) 的分布具有标度不变性, 故幂律分布也可称为 Scale-free 分布。类似地, 科学引文网中的节点的度分布也满足上述关系。各种大型网络经过自我演化最终达到这种与标度无关的状态。那么导致这种 Scale-free 状态的内部机制是什么呢?

Barabási 和 Albert 创造性地提出了大型网络自我蜕变成 Scale-free 网络的两条根本原则^[8]。

(1) 增长性: 随着时间的进展, 网络的节点数不断增加。

(2) 择优连接性: 新增节点会优先选择连接到已经拥有较高连接数的节点。

理论分析和仿真结果均说明, 按照上述两条原则所演化出的大型网络具有 Scale-free 性质, 其节点的度分布服从的幂律分布。上述演化模型在文献中通常称为 BA 模型, 幂指数通常称为度分布指数, 也称为标度指数 (scaling exponent)。

Scale-free 性质的发现掀起了一场大规模的针对实际网络的实证研究。科学界既期望又不无惊诧地发现, 在自然界、生物界、工程界以及社会界中几乎所有的实际网络都具有 Scale-free 特性^[1~7, 35~41], 并且绝大多数实际网络的度分布指数均位于 2~3 (Barabási 和 Albert 称无法解释其原因^[9])。BA 模型所演化出的网络虽然具有 Scale-free 性质, 但其度分布指数等于 3, 这一点与实际不符。因此, 一些学者对 BA 模型进行了多方面的推广^[42~61], 如考虑节点的老化 (aging)、节点连接的成本限制、新增节点选择连接时的地理因素等, 所得到的模型一方面与实际情况更加符合, 另一方面也使得度分布指数能够在 2~3 进行调节。但是需要指出的是, 尽管如此, Barabási 和 Albert 的工作仍然是本质性的、原创性的贡献。

小世界性质和 Scale-free 性质是复杂网络最为本质的特性, 除此之外, 研究人员还发现了某些实际网络的其他性质, 如度相关性 (degree correlation)、存在网络基序 (network motif) 等^[62~76]。

复杂网络小世界性质和 Scale-free 性质的发现推动了复杂网络上动力学的研究, 主要体现在三个方面: 传播动力学 (spreading process)、同步和集团动力学 (syn-

chronization and collective dynamics) 以及静态和动态鲁棒性 (static and dynamic robustness) 等。

复杂网络上传播动力学的代表性成果是成功地将经典的建立在均质网络上的传染病模型移植到复杂网络上, 并据此解释了病毒传播的若干现象^[18,77~95]。

最典型的传染病模型是 SIS 模型^[96]。在 SIS 模型中, 网络中的节点处于两种状态之一: 健康状态 (S) 和染病状态 (I)。在每一个离散时间步, 与 I 节点接触的 S 节点将以概率 μ 转变成 I 节点, 每个 I 节点将以概率 δ 转变成 S 节点。由于均质网络中绝大多数节点的连接数集中在平均连接数 $\langle k \rangle$ 附近, 可以用一个统一的基于平均场的反应方程 (mean-field reaction equation) 描述 SIS 模型如下^[96]:

$$\dot{\rho}(t) = -\delta\rho(t) + \mu\langle k \rangle\rho(t)(1 - \rho(t)), \quad (1.2)$$

其中, $\rho(t)$ 表示 t 时刻 I 节点所占的比例。再定义有效传播速率 (effective spreading rate) 为 $\lambda = \mu/\delta$ 。SIS 模型的一个最重要的研究成果是: 存在 λ 的一个阈值 $\lambda_c > 0$, 使得当 $\lambda > \lambda_c$ 时, 病毒能够在网络中传播并最终稳定在一定的比例上, 而当 $\lambda < \lambda_c$ 时, 病毒将以指数速度消亡^[96]。

复杂网络具有 Scale-free 特性, 其中大多数节点的度值很小, 少数节点 (称为 Hub 节点) 度值很大, 因而绝大多数节点的度值偏离平均度值 $\langle k \rangle$ 。由于复杂网络的节点的连接数的这种异质性, Pastor-Satorras 和 Vespignani 将网络中的节点按度值进行分组, 用 $\rho_k(t)$ 表示 t 时刻度值为 k 的节点中染病节点所占的比例, 建立了如下基于平均场的反应方程组^[18]:

$$\frac{d\rho_k(t)}{dt} = -\rho_k(t) + \lambda k[1 - \rho_k(t)]\theta(t), \quad (1.3)$$

其中, $\theta(t)$ 表示任一条边连接到染病节点的概率, $\theta(t)$ 满足

$$\theta(t) = \langle k \rangle^{-1} \sum_{i=1}^n iP(i)\rho_i(t). \quad (1.4)$$

通过对上述方程组进行理论分析, Pastor-Satorras 和 Vespignani 发现对于度分布满足式 (1.1) 的复杂网络, 如果网络具有无限多个节点且度分布指数位于 $2 \sim 3$, 则病毒传播不存在阈值, 即任何病毒都可以在这种网络中蔓延并且染病节点始终占据一定比例。这一发现使得病毒防治工作者需要重新审查复杂网络中的病毒防治策略。

研究发现, 复杂网络中的 SIR 模型也具有上述相同结论, 即满足 SIR 模型的病毒在复杂网络中传播时也不存在阈值^[80,93,94]。另外, 对于度值相关的复杂网络^[81,82,84,95], 上述结论仍然成立。因此, 复杂网络的 Scale-free 特性对于病毒的传播起到了推波助澜的作用。复杂网络中另一种受到关注的动力学是同步动力

学^[44]。在网络中的所有节点处各放置一个振荡器，这些振荡器之间通过节点之间的边进行耦合作用。同步动力学所研究的问题是这些振荡器的相位最终是否可以趋于一致（同步），达到同步需要多少时间，网络拓扑结构如何影响同步的实现。可以用下述方程组描述每个节点处所放置的振荡器^[44]：

$$\dot{\mathbf{x}}_i = \mathbf{F}_i(\mathbf{x}_i) - \sigma \sum_{j=1}^N b_{ij} \mathbf{H}(\mathbf{x}_j), \quad i = 1, \dots, N, \quad (1.5)$$

其中， N 表示网络中的节点数； \mathbf{x}_i 代表定义在节点 i 处的一个 m 维向量场； $\mathbf{H}(\mathbf{x}) : \mathbf{R}^m \rightarrow \mathbf{R}^m$ 是一个映射， b_{ij} 满足

$$\sum_{j=1}^N b_{ij} = 0, \quad i = 1, \dots, N. \quad (1.6)$$

研究人员引进了多种同步概念，如完全同步 (complete synchronization)、位相同步 (phase synchronization)、延迟同步 (lag synchronization)、广义同步 (generalized synchronization)、间歇延迟同步 (intermittent lag synchronization)、非完全位相同步 (imperfect phase synchronization) 以及几乎同步 (almost synchronization) 等，并进行了深入研究。研究表明，复杂网络的 Scale-free 性质影响到同步的实现及达到同步的时间^[44]。

复杂网络的鲁棒性研究集中在静态鲁棒性和动态鲁棒性两个方面。静态鲁棒性研究的是取出一定比例的节点或连接边之后对于网络拓扑结构的影响。研究表明，Scale-free 网络对于随机攻击或故障具有极强的鲁棒性，而对于选择性攻击却表现出明显的脆弱性。动态鲁棒性研究的是复杂网络上的雪崩效应。雪崩效应是指当每个节点的负载能力有限时，一个节点超出其负载可能导致整个网络出现连锁反应，从而使得全网瘫痪。研究表明，具有 Scale-free 特性的网络似乎更容易产生雪崩效应^[14~17,44]。

根据上述简要综述，可以看出，关于复杂网络的研究尚存在一些需要进一步探讨的问题：

(1) 复杂网络具有 Scale-free 性质，实证研究表明绝大多数实际网络的度分布指数位于 2~3，这个事实的理论根由尚需证明。

(2) Scale-free 网络的特点是大多数节点的度值很小，而少量节点 (Hub 点) 有很大的度值。但是关于 Hub 点的数量以及 Hub 点的度值与度分布指数之间的关系需要进一步探讨。

(3) 复杂网络具有 Scale-free 性质说明复杂网络是一个异质网络，如何对异质性程度进行度量？

(4) 实证研究表明, Internet 存在 AS 核 (AS core), 如何从数学上给出 AS 核的明确定义?

(5) 病毒传播的 SIS 模型已经推广到复杂网络, 随着时间的进展, 感染节点的比例是否能够稳定到一个固定比例尚需严格的数学证明, 复杂网络的 Scale-free 性如何影响这个最终比例的大小也需要进一步探讨。

(6) 病毒传播的 SIR 模型在复杂网络上的推广也存在上述类似问题。

(7) 研究复杂网络的根本原因之一是探究如何对复杂网络上的动力学行为进行控制, 关于如何对复杂网络的 Scale-free 性质以及复杂网络上的动力学行为所呈现的 Scale-free 现象进行控制的研究, 具有重大的理论意义和实际意义。

1.3 本书主要研究内容及框架

复杂网络是一个综合性、跨学科的研究分支, 它涉及统计物理、数据处理、图论、自动控制及计算机仿真等多个学科领域。至今为止, 复杂网络尚未形成专门的研究方法, 从事复杂网络研究的人员来自于众多学科, 包括生物、社会学、统计物理、数学、自动控制和计算机等领域。目前复杂网络的很多方面仍处于实证研究和数学建模阶段。在前人研究的基础上, 本书从复杂网络的 Scale-free 性、复杂网络所呈现的 Scale-free 现象及其控制等多个方面对复杂网络进行了系统深入的研究, 主要包括以下三个方面的内容:

(1) 复杂网络的 Scale-free 性质。它主要从 Scale-free 网络的度分布指数的变化范围、Scale-free 网络的 Hub 点的数量及其连接数、复杂网络的连接率和吸引率以及复杂网络的社区发现等方面对复杂网络的 Scale-free 性质进行了探讨。

(2) 复杂网络上的 Scale-free 现象。它主要研究复杂网络上的两个传染病模型 (SIS 模型和 SIR 模型) 的稳定性问题, 以及传染病模型趋于稳态后系统遭到破坏的程度与复杂网络的 Scale-free 性质的关系。

(3) 复杂网络的控制。它主要研究复杂网络上的动力学行为所呈现出的 Scale-free 现象对于复杂网络的 Scale-free 性质的依赖性。复杂网络上的 Scale-free 现象表现出正、负两面性, 我们研究了影响 Scale-free 现象的两面性的相关因素及其控制问题, 并提出了关于复杂网络的一般性控制框架。

本书内容安排如下:

第 1 章是绪论, 主要介绍了复杂网络研究的背景及意义、国内外的研究现状、本书主要研究内容及框架。首先通过列举大量实际网络说明复杂网络在现实世界中是广泛存在的, 并简要论述了研究复杂网络的 Scale-free 性、Scale-free 现象及其控制的意义; 然后简要介绍了复杂网络研究中与本书相关的主题; 最后概要地叙述了本书的主要研究内容。

第 2 章主要介绍了作者所提出的复杂网络的两个新的拓扑参数：连接率和吸引率，并提出了复杂网络的一个静态概率模型，从理论和实证两个方面阐述了连接率和吸引率服从幂律及其重要的应用。

第 3 章介绍了复杂网络的几个异质性指标，包括度分布指数、网络结构熵、度分布熵以及作者所引进的洛伦兹曲线与基尼系数，并对这些指标的异同进行了较为全面的比较。

第 4 章讨论 Scale-free 网络的拓扑结构及其形成机制，研究了 Scale-free 网络的一个重要参数——度分布指数的取值范围，揭示了现实世界中大多数网络的度分布指数位于 2~3 的理论根由。同时，研究了 Scale-free 网络中的 Hub 点的若干独特性质，如 Hub 点的数量、Hub 点的最高度值等问题。最后，对 Scale-free 性质的形成机制进行了探讨。

第 5 章我们针对 BBS 这样一个具体的复杂网络的 Scale-free 性质进行了实证研究，并且将复杂网络中的社区结构及其发现算法成功地应用于 BBS 的热点话题发现。

第 6 章讨论了复杂网络的几个重要的中心化指标，包括度指标、紧密度指标、特征向量指标、介数指标和流介数指标等，指出了其各自的应用场合，并通过解析现实系统中网络的中心化问题，总结了中心化研究的普遍思路和方法，说明了不同网络采用相应中心化测试指标的必要性。

第 7 章从静态鲁棒性和雪崩动力学两个方面对 Scale-free 网络的结构稳定性进行了研究，探讨了 Scale-free 网络的鲁棒性与 Scale-free 网络的异质性之间的关系。

第 8 章的主要研究工作可以分为两个部分：对复杂网络拓扑结构及病毒传播模型的理解，并对病毒传播的稳态特性和瞬态特性进行了分析，着重对病毒传播的瞬态特性进行仿真模拟。

第 9 章讨论了 Scale-free 网络上的传播动力学模型之一：SIS 模型。通过对模型的深入分析，解决了 SIS 模型的全局稳定性问题，并对其收敛速度进行了理论分析。最后给出了 SIS 模型的感染规模与网络拓扑结构之间的关系。

第 10 章讨论了另一个著名的 Scale-free 网络上的传播动力学模型：SIR 模型。深入分析了复杂网络中 SIR 模型的病毒爆发的阈值以及病毒爆发规模与网络拓扑结构之间的关系。

第 11 章介绍了 Scale-free 网络上的同步动力学，着重讨论了如何利用主稳定函数法判断网络上的动力学是否可以达到同步，并分析了网络的异质性对于同步的影响。

第 12 章研究 Scale-free 网络的控制，提出了 Scale-free 网络中控制的一般性框架和控制方法，在此基础上提出了 Interent 拓扑结构的一个新的演化控制模型。