



汪小帆 李 翔 陈关荣 编著

复杂网络 理论及其应用

清华大学出版社

汪小帆 李翔 陈关荣 编著

复杂网络 理论及其应用

清华大学出版社
北京

内 容 简 介

复杂网络研究正渗透到数理学科、生命学科和工程学科等众多不同的领域,对复杂网络的定量与定性特征的科学理解已成为网络时代科学研究中一个极其重要的挑战性课题。本书致力于系统地介绍复杂网络的基础知识和研究进展。由于复杂网络研究具有很强的跨学科特色,并且新的问题和研究成果不断涌现,因此本书着眼于复杂网络研究中已经取得的主要研究进展。主要包括:网络拓扑特性与模型,复杂网络上的传播行为、相继故障、搜索算法和社团结构,以及复杂网络的同步与控制。本书还详细介绍了作者们在复杂网络领域的有关工作。

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

版权所有,翻印必究。举报电话:010-62782989 13501256678 13801310933

图书在版编目(CIP)数据

复杂网络理论及其应用/汪小帆,李翔,陈关荣编著. —北京:清华大学出版社,2006.4

ISBN 7-302-12505-8

I. 复… II. ①汪… ②李… ③陈… III. 计算机网络—研究 IV. TP393

中国版本图书馆 CIP 数据核字(2006)第 013163 号

出 版 者:清华大学出版社 地 址:北京清华大学学研大厦
http://www.tup.com.cn 邮 编:100084
社 总 机:010-62770175 客 户 服 务:010-62776969

责任编辑:王听讲

版式设计:刘伟森

印 刷 者:北京嘉实印刷有限公司

装 订 者:北京国马印刷厂

发 行 者:新华书店总店北京发行所

开 本:185×260 印张:17 字数:403 千字

版 次:2006 年 4 月第 1 版 2006 年 4 月第 1 次印刷

书 号:ISBN 7-302-12505-8/TP·8017

印 数:1~3000

定 价:45.00 元

很多人可能都有这样的经历：偶尔碰到一个陌生人，同他聊了一会儿后发现你认识的某个人居然他也认识，然后你们会一起发出“这个世界真小”的感叹。那么对于地球上任意两个人来说，借助第三者、第四者这样的间接关系来建立起他们两人的联系，平均需要通过多少人呢？20世纪60年代美国哈佛大学心理学家 Milgram 的著名的小世界实验（社会调查）给出的推断是：地球上任意两个人之间的平均距离是6。也就是说，平均中间只要通过5个人的传递，一个人就能与地球上任何一个角落的另一个人建立联系。这就是著名的六度分离(six degrees of separation)推断。尽管这个平均数6不一定很准确，但它反映了人与人之间的平均距离与全球人口数量相比是一个非常小的数。

1998年6月，美国康奈尔(Cornell)大学理论和应用力学系的博士生 Watts 及其导师 Strogatz 在 *Nature* 杂志上发表了题为《“小世界”网络的群体动力行为》的文章，进一步揭示了复杂网络的小世界特性，并建立了一个小世界网络模型。1999年10月，美国圣母(Notre Dame)大学物理系的 Barabási 教授及其博士生 Albert 在 *Science* 杂志上发表了题为《随机网络中标度的涌现》一文，揭示了复杂网络的无标度性质，并建立了一个无标度网络模型。以这两篇文章的发表为标志，复杂网络研究进入了一个新时代。复杂网络理论研究再也不局限于数学的范畴，而是开始受到了从物理学到生物学，从工程技术到管理和社会科学的众多学科的研究人员越来越多的关注。

Strogatz 是从事非线性动力学研究的专家。他特别是在耦合振子同步方面做了许多重要的工作。夏日窗外蟋蟀叫声的共鸣就是耦合振子同步的一个典型例子。Watts 和 Strogatz 当初是从研究大量耦合振子构成的网络系统的同步问题中受到启发，而考虑社会网络中的小世界现象的。2001年春季开始，本书作者们以研究具有小世界和无标度拓扑结构的动态网络的同步化问题为切入点，开始进入复杂网络研究领域。当时，复杂网络研究在国内还未受到关注。然而，从2002年起，在短短几年中，国内不同学科的研究人员和青年学者对复杂网络研究的兴趣越来越浓，至今国内已召开过多次以复杂网络为主题的学术会议和论坛。我们近几年来多次应邀在国内做过有关复杂网络的学术报告，并作为主要组织者于2004年4月在无锡组织了有40余人参加的首届全国复杂动态网络学术论坛。2005年10月在北京召开的由中国高等学术研究中心组织的第二届全国复杂网络学术论坛的参加人数已经超过了150人。此外，中国系统工程学会于2004年10月在浙江大学组织了全国复杂网络研讨班；中国数学会于2005年1月在上海交通大学组织了复杂网络学习班；武汉大学在国内率先成立了校级复杂网络研究中心并于2005年春季组织了全国复杂网络学术会议；上海理工大学还成立了上海系统科学研究院并编辑出版

了复杂网络文集。在广泛的学术交流过程中,我们深切感受到了国内青年研究人员和学生渴望尽快了解复杂网络基础知识和研究进展的心情,也充分认识到了尽早出版一本系统介绍复杂网络理论与应用的中文著作的必要性。

一些国际著名大学(如 MIT,哥伦比亚大学和密歇根大学等)已相继开设了有关复杂网络的课程,本书作者汪小帆教授也在上海交通大学为研究生开设了复杂网络课程。读者可在上海交通大学复杂网络与控制研究室的主页^①上查到相关链接。我们相信今后国内会有更多的高校开设有关复杂网络的课程和学习班。

我们是于 2004 年春季开始考虑撰写本书的,但在真正开始写作之后,我们发现撰写这样一本著作的难度远远超出了我们最初的想像。近几年来,关于复杂网络的研究文献呈现出极快的增长趋势,每个月都有很多新文章出现。如何从数以千计的且仍在不断增长的文献中筛选出合适的内容,以及如何组织这些内容是我们面临的一个很大的挑战。更为困难的是,作为一门新兴学科,许多术语和概念没有完全统一,甚至在一些理论和方法的研究上还存在不同的观点。尽管国际上已经出版了几本关于复杂网络的专业著作,但是这些专著的侧重点各有不同。可以说,到目前为止,国际上还没有一本系统介绍复杂网络理论与应用的著作面世。因此,尽管困难重重,我们仍然决定尽自己最大的能力撰写这样一本入门性的著作。

本书致力于系统地介绍复杂网络的基础知识和研究进展。复杂网络研究具有极强的交叉学科特色,而从事复杂网络研究或对此感兴趣的人员来自众多不同的学科,并且还存在着一些不同的看法和观点。为了使本书能为更多的读者所接受,我们撰写的一个基本原则就是使得具有大学高等数学和普通物理基本知识的读者能够较为顺利地阅读其基本内容。此外,本书在材料的取舍和组织上也不可避免地反映了作者们自己的偏好和信息学科的背景。因此,数学家也许会觉得书中的一些叙述没有给出数学上的严格论证,而物理学家则可能会觉得本书对于大量从统计力学角度研究复杂网络的重要内容没有给予充分的介绍。对于不能满足读者期望的许多地方,我们深表歉意。

作者们衷心感谢方锦清、刘曾荣、汪秉宏、吕金虎、李春光、范正平等教授和博士。他们对本书的写作给予了热情的关注和协助,并以不同的形式和在不同程度上对本书的形成和出版作出了贡献。上海交通大学复杂网络与控制研究室的范谨、史明江、许丹、宋丽雅、解伟等研究生做了大量的辅助工作。

作者们十分感谢他们的家人对他们长期、连续和忙碌的科研工作的充分理解和大力支持。

作者汪小帆和李翔感谢国家杰出青年科学基金、国家自然科学基金重点项目和面上项目、教育部优秀青年教师资助计划、教育部高等学校博士点基金和教育部留学回国人员科研启动基金等的资助。作者陈关荣感谢香港研究资助局和香港城市大学的资助。作者们还要感谢清华大学出版社王听讲先生等对本书出版的大力支持,并感谢 IEEE、Springer-Verlag、World Scientific、Elsevier 等出版社和 American Physical Society 允许书中引用少量图形资料。本书封面图片是反映澳大利亚堪培拉的一群居民之间的社会关系网络的

① <http://automation.sjtu.edu.cn:81/cnc/>

示意图。该图版权属于澳大利亚国立大学(Australian National University)的 Alden S. Klov Dahl 博士。我们非常感谢 Alden S. Klov Dahl 博士同意本书使用该图作为封面图片。

读者可通过上海交通大学复杂网络与控制研究室的主页进入与本书配套的网站(见上页脚注)。该网站会根据复杂网络的研究进展不断补充新的内容。作为从事复杂系统反馈控制的研究人员,我们深知反馈对改进系统性能的重要性。因此,热忱欢迎读者们把对本书的任何意见及时反馈给我们,以便及时在网上更正和将来有重印机会时进一步提高本书的质量。

汪小帆、李翔,上海交通大学复杂网络与控制研究室
陈关荣,香港城市大学混沌控制与同步学术研究中心

2005 年 冬

第 1 章 引论	1
1.1 引言	1
1.2 复杂网络研究简史	3
1.2.1 从七桥问题谈起	3
1.2.2 随机图理论	4
1.2.3 小世界实验	5
1.2.4 弱连接的强度	7
1.2.5 复杂网络研究的新纪元	8
1.3 基本概念	9
1.3.1 网络的图表示	9
1.3.2 平均路径长度	10
1.3.3 聚类系数	10
1.3.4 度与度分布	11
1.3.5 实际网络的统计性质	14
1.4 本书内容简介	15
参考文献	16
第 2 章 网络拓扑基本模型及其性质	18
2.1 引言	18
2.2 规则网络	18
2.3 随机图	19
2.4 小世界网络模型	21
2.4.1 小世界模型	21
2.4.2 小世界网络的小波分析	23
2.5 无标度网络模型	27
2.5.1 BA 无标度网络	27
2.5.2 鲁棒性与脆弱性	29
2.5.3 适应度模型	33
2.6 局域世界演化网络模型	34
2.7 模块性与等级网络	37

2.7.1	模块与模体	37
2.7.2	等级网络	39
2.7.3	超家族	40
2.8	复杂网络的自相似性	42
	参考文献	46
第3章	Internet 拓扑特性及建模	49
3.1	引言	49
3.2	Internet 的拓扑特性	49
3.2.1	幂律分布	50
3.2.2	层次性	53
3.2.3	富人俱乐部特性	55
3.2.4	异配性	55
3.2.5	核数和介数	56
3.3	随机图产生器	57
3.4	结构产生器	58
3.4.1	Tiers 产生器	58
3.4.2	Transit-Stub 产生器	60
3.5	基于连接度的产生器	61
3.5.1	Inet	61
3.5.2	AB 模型	62
3.5.3	BRITE	62
3.5.4	GLP 模型	63
3.5.5	PFP 模型	63
3.5.6	DP 模型	64
3.5.7	T _{ANG} 模型	64
3.6	多局域世界模型	65
3.6.1	模型构造	65
3.6.2	度分布分析	66
3.7	各类模型的定性比较	68
	参考文献	70
第4章	复杂网络上的传播机理与动力学分析	72
4.1	引言	72
4.2	复杂网络的传播临界值理论	73
4.2.1	均匀网络的传播临界值	73
4.2.2	实际数据分析	74
4.2.3	无标度网络的传播临界值	75

4.2.4	BA 无标度网络的传播临界值	75
4.2.5	有限规模无标度网络的传播临界值	76
4.2.6	关联网络的传播临界值	77
4.2.7	更广义的复杂网络传播临界值研究	78
4.3	复杂网络的免疫策略	78
4.3.1	随机免疫	79
4.3.2	目标免疫	79
4.3.3	熟人免疫	80
4.4	复杂网络的传播动力学	81
4.4.1	d 维 NW 小世界网络的线性传播方程	81
4.4.2	小世界网络传播动力方程的分形、混沌与分岔	82
4.4.3	小世界网络的广义传播动力方程及其分岔	83
4.4.4	复杂网络传染动力方程的分岔与震荡	85
4.5	计算机病毒在 Internet 上的传播	86
4.5.1	红色代码蠕虫的随机常数传播模型	86
4.5.2	蠕虫的间隔模型	88
4.5.3	电子邮件病毒的传播模型	89
4.5.4	电子邮件病毒传播的仿真模拟	90
4.5.5	计算机病毒传播对网络拓扑结构的影响	94
4.6	复杂网络中的其他传播现象	95
4.6.1	谣言在复杂网络中的传播	95
4.6.2	传播蔓延现象的推广模型与普遍行为	97
	参考文献	98
第 5 章	复杂网络上的相继故障	101
5.1	引言	101
5.2	复杂网络相继故障的动态模型分析	101
5.2.1	负荷—容量模型	101
5.2.2	二值影响模型	110
5.2.3	沙堆模型	112
5.2.4	OPA 模型	113
5.2.5	CASCADE 模型	117
5.2.6	其他模型	120
5.3	基于耦合映象格子的相继故障模型	120
5.3.1	基于 CML 的相继故障模型	120
5.3.2	典型拓扑结构 CML 中的相继故障	121
5.3.3	一类指数可调的无标度耦合映象格子中的相继故障	125
	参考文献	129

第 6 章 复杂网络中的搜索	131
6.1 引言	131
6.2 社会网络搜索	131
6.2.1 Kleinberg 网络模型	131
6.2.2 层次树结构网络模型	134
6.2.3 Email 网络验证	137
6.2.4 实现快速搜索的网络	140
6.3 几种复杂网络搜索策略分析	143
6.3.1 广度优先搜索策略	143
6.3.2 随机游走搜索策略	144
6.3.3 最大度搜索策略	150
6.4 P2P 网络中的搜索	152
6.4.1 P2P 网络简介	152
6.4.2 Gnutella 网络中的广播搜索	154
6.4.3 Gnutella 网络中的改进的搜索策略	155
6.5 复杂网络中的搜索和拥塞	159
参考文献	160
第 7 章 复杂网络中的社团结构	162
7.1 引言	162
7.2 Kernighan-Lin 算法	164
7.3 谱平分法	165
7.3.1 谱平分法的基本思想	165
7.3.2 一种线性时间的物理方法	166
7.3.3 基于 Normal 矩阵的谱平分法	169
7.4 分裂方法	171
7.4.1 GN 算法	171
7.4.2 采用节点集的 GN 算法	175
7.4.3 自包含 GN 算法	176
7.4.4 快速分裂算法	179
7.4.5 基于相异性的算法	180
7.4.6 基于信息中心度的算法	181
7.4.7 极值优化算法	183
7.5 凝聚算法	184
7.5.1 Newman 快速算法	184
7.5.2 利用堆结构的贪婪算法	185
7.5.3 结合谱分析的凝聚算法	186

7.6 派系过滤算法	188
参考文献	191
第8章 复杂网络中的同步	194
8.1 引言	194
8.2 复杂网络的完全同步判据	195
8.2.1 一般连续时间耦合网络完全同步判据	195
8.2.2 连续时间线性耗散耦合网络完全同步判据	198
8.3 复杂动力网络的完全同步	200
8.3.1 规则网络的完全同步	200
8.3.2 小世界网络的完全同步	201
8.3.3 无标度网络的完全同步	202
8.3.4 局域世界演化网络模型的完全同步	205
8.4 连续时间时变耦合网络完全同步	207
8.4.1 基于 Lyapunov 稳定性的判据	207
8.4.2 基于连接图的判据	209
8.4.3 基于连接图的判据的应用	211
8.5 其他网络完全同步判据	215
8.5.1 具有耦合时滞的连续时间网络完全同步判据	215
8.5.2 离散时间耦合网络完全同步判据	217
8.6 复杂网络中各个因子与完全同步的关系	218
8.7 改进复杂网络同步的方法	222
8.7.1 无序扰动改进同步特性	222
8.7.2 通过时滞提高网络同步特性	223
8.7.3 加权耦合提高网络同步特性	227
8.8 复杂网络的相位同步	229
8.8.1 Kuramoto 相位同步模型	229
8.8.2 小世界网络的相位同步	231
8.8.3 无标度网络模型的相位同步	232
8.8.4 网络模体的同步	235
8.8.5 非对称耦合网络的相位同步	236
8.8.6 离散时间耦合网络的相位同步	238
参考文献	240
第9章 复杂动态网络的控制	243
9.1 引言	243
9.2 规则网络时空混沌的牵制控制	244
9.3 无标度动态网络的牵制控制: 鲁棒性与脆弱性	246

9.4 一般复杂动态网络的牵制控制	248
9.4.1 一般复杂动态网络牵制控制稳定性分析	248
9.4.2 一般复杂动态网络牵制控制的虚拟控制原理	250
9.4.3 无标度网络与随机网络的牵制控制与虚拟控制	251
9.4.4 牵制控制陈氏无标度动态网络到不稳定平衡点	254
9.5 随机驱动下动态网络的有序性与动力学	256
参考文献	257
附录 名词对照	259

引 论

1.1 引言

地球上任意两个人之间要通过多少个朋友才能互相认识？万维网(WWW)上从一个页面到另一个页面平均需要点击多少次鼠标？层出不穷的计算机病毒是如何在互联网(internet)上传播的？各种传染病(艾滋病、非典型性肺炎和禽流感等)是如何在人类和动物中流行的？为什么流言蜚语会散布得很快？全球或地区性金融危机是如何发生的？局部故障是如何触发大面积停电事故的？大城市的交通堵塞问题是如何引起的？应该如何建立合理的公共卫生与安全网络？为什么大脑能够具有思维的功能？这些问题尽管看上去各不相同，但每一个问题中都涉及很复杂的网络，包括 WWW、Internet、社会关系网络、经济网络、电力网络、交通网络、神经网络等等。更为重要的是，越来越多的研究表明，这些看上去各不相同的网络之间有着许多惊人的相似之处。

20 世纪 90 年代以来，以 Internet 为代表的信息技术的迅猛发展使人类社会大步迈入了网络时代。从 Internet 到 WWW，从大型电力网络到全球交通网络，从生物体中的大脑到各种新陈代谢网络，从科研合作网络到各种经济、政治、社会关系网络等，可以说，人们已经生活在一个充满着各种各样的复杂网络的世界中。人类社会的网络化是一把“双刃剑”：它既给人类社会生产与生活带来了极大的便利，提高了人类生产效率和生活质量，但也给人类社会生活带来了一定的负面冲击，如传染病和计算机病毒的快速传播以及大面积的停电事故等。因此，人类社会的日益网络化需要人类对各种人工和自然的复杂网络的行为有更好的认识。长期以来，通信网络、电力网络、生物网络和社会网络等分别是通信科学、电力科学、生命科学和社会学等不同学科的研究对象，而复杂网络理论所要研究的则是各种看上去互不相同的复杂网络之间的共性和处理它们的普适方法。从 20 世纪末开始，复杂网络研究正渗透到数理学科、生命学科和工程学科等众多不同的领域，对复杂网络的定量与定性特征的科学理解，已成为网络时代科学研究中一个极其重要的挑战性课题，甚至被称为“网络的新科学(new science of networks)”(见文献[1~2])。

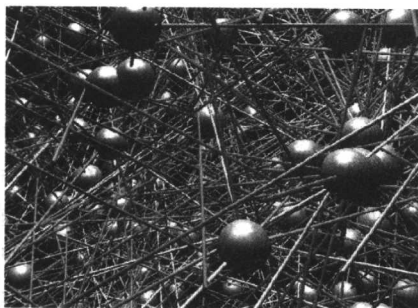
以生命科学为例，20 世纪的生命科学研究主流是建立在还原论基础上的分子生物学。还原论的基本前提是，在由不同层次组成的系统内，高层次的行为是由低层次的行为所决定的。具有还原论观点的生物学家通常认为，只要认识了构成生命的分子基础(如基因和蛋白质)就可以理解细胞或个体的活动规律，而组分之间的相互作用常常被忽略不计。尽管基于还原论的分子生物学极大地促进了人类对单个分子功能的认识，然而绝大多数生物特征都来自于细胞的大量不同组分，如蛋白质、DNA、RNA 和小分子之间的交

互作用。对这些极其复杂的交互作用网络的结构和动力学的理解已成为 21 世纪生命科学的关键性研究课题和挑战之一。

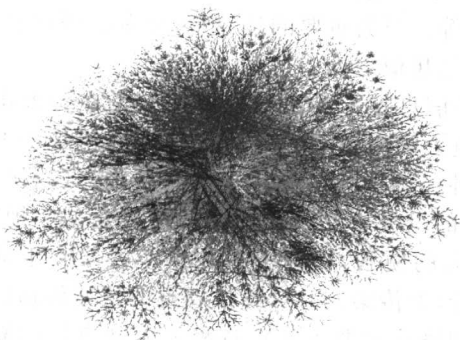
一般而言,网络系统的复杂性体现在以下几个方面(见文献[3])。

(1) 结构复杂性

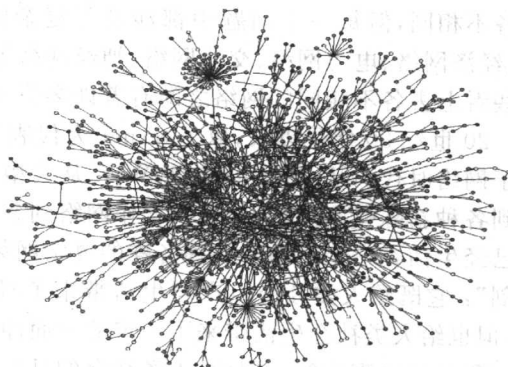
网络连接结构看上去错综复杂、极其混乱(图 1-1)。而且网络连接结构可能是随时间变化的,例如,WWW 上每天都不停地有页面和链接的产生和删除。此外,节点之间的连接可能具有不同的权重或方向。例如,神经系统中的突触有强有弱,可以是抑制的也可以是兴奋的。



(a)



(b)



(c)

图 1-1 三个具有复杂连接结构的实际网络示意图

(a) 澳大利亚堪培拉的一个社会关系网络结构示意图(A. S. Klov Dahl);

(b) Internet 上部分 IP 地址的连接结构示意图(W. R. Cheswick);

(c) 蛋白质相互作用网络结构示意图(H. Jeong)

(2) 节点复杂性

网络中的节点可能是具有分岔和混沌等复杂非线性行为的动力系统。例如,基因网络和 Josephson 结阵列中每个节点都具有复杂的时间演化行为。而且,一个网络中可能存在多种不同类型的节点。例如,控制哺乳动物中细胞分裂的生化网络就包含各种各样的基质和酶。

(3) 各种复杂性因素的相互影响

实际的复杂网络会受到各种各样因素的影响和作用。例如,耦合神经元重复地被同时

激活,那么它们之间的连接就会加强,这被认为是记忆和学习的基础。此外,各种网络之间也存在密切的联系,这使得对复杂网络的分析变得更为困难。例如,电力网络的故障可能会导致 Internet 流量变慢、金融机构关闭、运输系统失去控制等一系列不同网络之间的连锁反应。

20 世纪中叶以来,非线性动力学已逐步发展成为分析复杂动力系统的基本工具之一。其中,网络系统的动力学分析受到了很多人的关注。但是在 20 世纪的非线性动力学研究中,绝大多数网络动态系统模型均假设网络具有规则而且固定的拓扑结构,而把重点放在由节点的非线性动力学行为所产生的复杂性,如斑图(pattern)的涌现和时空混沌(spatio-temporal chaos)的产生等。典型的例子包括耦合映象格子(coupled map lattice, CML)(见文献[4])和细胞神经/非线性网络(cellular neural/nonlinear network, CNN)(见文献[5])。采用简单的规则结构的主要好处是使人们可以集中精力研究节点的复杂动力学行为对整个网络复杂性的影响,并且规则的网络结构也便于用集成电路来实现。

1.2 复杂网络研究简史

1.2.1 从七桥问题谈起

近年来复杂网络研究的兴起,使得人们开始广泛关注网络结构复杂性及其与网络行为之间的关系。要研究各种不同的复杂网络在结构上的共性,首先需要有一种描述网络的统一工具。这种工具在数学上称为图(graph)。任何一个网络都可以看作是由一些节点按某种方式连接在一起而构成的一个系统。具体网络的抽象图表示,就是用抽象的点表示具体网络中的节点,并用节点之间的连线来表示具体网络中节点之间的连接关系。

实际网络的图表示方法可以追溯到 18 世纪伟大的数学家欧拉(Euler)对著名的“Königsberg 七桥问题”的研究。Königsberg 是东普鲁士(现俄罗斯)的一个城镇,城中有一条横贯城区的河流,河中有两个岛,两岸和两岛之间共架有七座桥,如图 1-2 所示。传说当地居民常常议论这样一个有趣的问题:

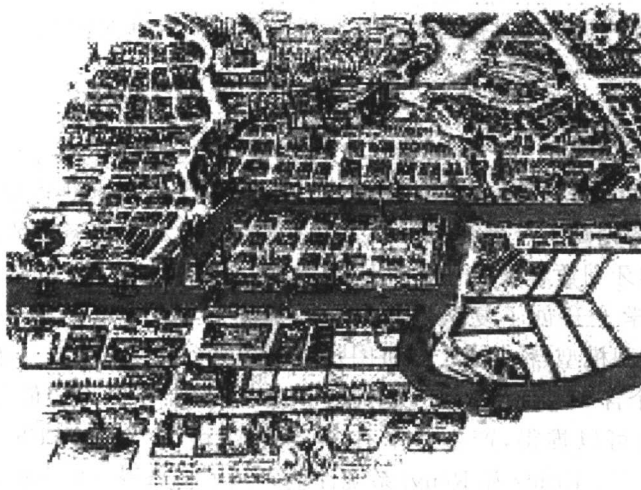


图 1-2 1736 年的 Königsberg 镇

一个人能否在一次散步中走过所有的七座桥,而且每座桥只经过一次,最后返回原地?这个问题看起来似乎相当简单,但长期以来小镇上没有一个人能走出这样一条路径。

1736年,欧拉仔细地研究了这个问题。他利用数学抽象法,将被河流分隔开的四块陆地抽象为四个点,分别用A、B、C和D表示,而将连接这四块陆地之间的七座桥抽象为连接四个点的七条线,分别用a、b、c、d、e、f、g表示。这样就得到了由四个点和七条线构成的一个图,如图1-3所示。于是七桥问题就转化为如下的数学问题:

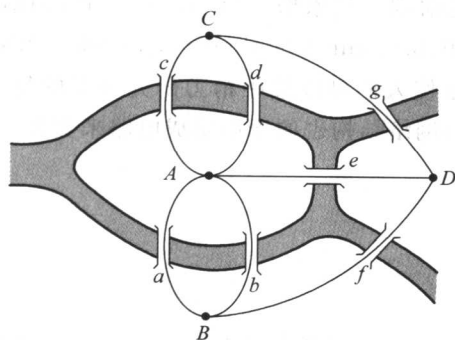


图 1-3 Königsberg 七桥问题的图示

从图1-3中任一点出发,经过每条边一次而后返回原点的回路是否存在?

欧拉给出了存在这样一条回路的充要条件,并由此推得上述七桥问题是没有解的。

那么欧拉当时是怎么解决这个问题的呢?他考虑图1-3如果能够一笔画成,则七桥问题迎刃而解。于是,欧拉转而研究一笔画成的图形应该具备什么性质?可以想像,能一笔画出的图形,一定只有一个起点和一个终点(这里要求起点和终点重合),中间经过的每一点总是包含进去的一条线和出来的一条线,这样除起点和终点外,每一点都只能有偶数条线与之相连。因此,如果要求起点和终点重合的话,那么能够一笔画出的图形中所有的点都必然有偶数条线与之相连。欧拉进一步阐明了这一条件也是充分的。而从图1-3中四个点来看,每个点都是有三条或五条线通过,所以不能一笔画出这个图形。这就是说不重复地一次走遍图1-2中的七座桥是绝对不可能的。

欧拉对七桥问题的抽象和论证思想,开创了数学中的一个分支——图论(graph theory)的研究。因此,欧拉被公认为图论之父,而图1-3也被称为欧拉图。事实上,今天人们关于复杂网络的研究与欧拉当年关于七桥问题的研究在某种程度上是一脉相承的,即网络结构与网络性质密切相关。

1.2.2 随机图理论

在欧拉解决七桥问题之后的相当长一段时间里,图论并未获得足够的发展。直到1936年才出版了图论的第一部专著,此后图论开始进入发展与突破的快车道。20世纪60年代,由两位匈牙利数学家 Erdős 和 Rényi 建立的随机图理论(random graph theory)被公认是在数学上开创了复杂网络理论的系统性研究(见文献[6])。在 Erdős 和 Rényi 研究的随机图模型(称为 ER 随机图)中,任意两个节点之间有一条边相连接的概率都为 p 。因此,一个含 N 个节点的 ER 随机图中边的总数是一个期望值为 $p[N(N-1)/2]$ 的随机变量。由此可以推得,产生一个有 N 个节点和 M 条边的 ER 随机图的概率为 $p^M (1-p)^{N(N-1)/2-M}$ 。Erdős 和 Rényi 系统性地研究了当 $N \rightarrow \infty$ 时 ER 随机图的性质(如连通性等)与概率 p 之间的关系。他们采用了如下定义:几乎每一个 ER 随机图都具有

某种性质 Q , 如果当 $N \rightarrow \infty$ 时产生具有这种性质 Q 的 ER 随机图的概率为 1。Erdős 和 Rényi 的最重要的发现是: ER 随机图的许多重要的性质都是突然涌现的。也就是说, 对于任一给定的概率 p , 要么几乎每一个图都具有某个性质 Q (比如说, 连通性), 要么几乎每一个图都不具有该性质。

Erdős 是 20 世纪最杰出的数学家之一, 他喜爱数学达到了如痴如醉的程度。Erdős 在素数理论、不定方程、组合论、概率论、数学分析和集合论等多个数学分支都做了大量创新性的工作, 他一生写作了大约 1600 篇文章, 其中大部分文章是分别与大约 500 位研究人员合作完成的! 曾经流传这样的说法: Erdős 跨进一位数学家的大门, 宣布“我的头脑敞开着 (My brain is open)”, 然后就和对方展开数学研究。临走时, 则说“另一个屋顶, 另一个证明 (another roof, another proof)”。如果读者希望进一步了解 Erdős 传奇式的科学人生, 可以阅读国内已经翻译出版的两本介绍 Erdős 的书: 《我的大脑敞开了》(见文献[7]) 和《数字情种》(见文献[8])。

在 20 世纪的后 40 年中, 随机图理论一直是研究复杂网络的基本理论。在此期间, 人们也做了试图揭示社会网络特征的一些实验, 例如, 下面介绍的著名的小世界实验 (small world experiment)。

1.2.3 小世界实验

1. Milgram 的小世界实验

一个社会网络就是一群人或团体按某种关系连接在一起而构成的一个系统。这里的关系可以多种多样, 如个人之间的朋友关系、同事之间的合作关系、家庭之间的联姻关系和公司之间的商业关系等等。图 1-1(a) 反映的就是澳大利亚堪培拉的一个社会关系网络。以朋友关系为例, 很多人可能都有这样的经历: 偶尔碰到一个陌生人, 同他聊了一会儿后发现你认识的某个人居然他也认识, 然后一起发出“这个世界真小”的感叹。那么对于地球上任意两个人来说, 借助第三者、第四者这样的间接关系来建立起他们两人的联系, 平均需要通过多少人呢? 20 世纪 60 年代美国哈佛大学的社会心理学家 Stanley Milgram 通过一些社会调查后给出的推断是: 地球上任意两个人之间的平均距离是 6。也就是说, 平均中间只要通过 5 个人, 你就能与地球上任何一个角落的任何一个人发生联系。这就是著名的六度分离 (six degrees of separation) 推断。1990 年在美国上演的一部名为《六度分离》的戏剧, 使 Milgram 的推断得到进一步的传播。

下面来看看 Milgram 的社会实验是怎么做的。首先, 他选定了两个目标对象: 一个是美国马萨诸塞州 (Massachusetts) 沙朗 (Sharon) 的一位神学院研究生的妻子, 另一位是波士顿 (Boston) 的一个证券经纪人。然后他在遥远的堪萨斯州 (Kansas) 和内布拉斯加州 (Nebraska) 招募到了一批志愿者。Milgram 要求这些志愿者通过自己所认识的人, 用自己认为尽可能少的传递次数, 设法把一封信最终转交到一个给定的目标对象手中。Milgram 在发表于 1967 年 5 月美国出版的《今日心理学》杂志上的一篇论文中 (见文献[9]), 描述了一份信件是如何仅用 3 步就从堪萨斯州的一位农场主手中转交到马萨诸塞州的那位神学院学生的妻子手中的: 农场主将信件寄给一个圣公会教父, 教父将其转寄给住在沙朗的一位同事, 然后信件就到了神学院学生妻子的手中。尽管并不是每一个实验对象都如此