

国家自然科学基金项目 (71173199)

教育部人文社会科学规划基金项目 (10YJA630001) 资助出版

河北省高校重点学科建设项目

复杂网络 在经济管理领域 的应用研究

FUZA WANGLUO ZAI JINGJI GUANLI LINGYU
DE YINGYONG YANJIU

安海岗 高湘昀 著

地质出版社

国家自然科学基金项目(71173199)

教育部人文社会科学规划基金项目(10YJA630001) 资助出版

河北省高校重点学科建设项目

复杂网络在经济管理 领域的应用研究

安海岗 高湘昀 著

地 质 出 版 社

· 北 京 ·

内 容 提 要

复杂网络在众多复杂理论中以能够形象地揭示系统中各要素之间的联系及其动态变化关系而受到越来越多的关注,如何运用复杂网络理论研究经济管理问题已成为学术界面临的新课题。本书正是针对这个问题,从复杂网络理论角度出发,系统地讨论了复杂网络在能源价格波动幅度、双能源价格联动波动、网络文本信息挖掘与舆情分析及电子商务领域科研合作情况四个方向的应用。本书共分为7章,第1~第3章主要对复杂网络的研究历史、基本模型与拓扑性质、传播与演化机理进行描述,第4~第7章运用复杂网络理论对4个不同的案例进行了实证分析。通过对能源价格波动及联动波动进行研究,可为我国能源价格的制定与预测提供参考;通过对网络文本信息进行挖掘与舆情分析,能够为我国舆情引导体制的建立提供决策参考;通过对电子商务领域科研合作情况统计分析,能够了解我国电子商务领域发展现状并对未来发展趋势做出预测。

本书可供从事经济管理研究的相关人员参阅。

图书在版编目(CIP)数据

复杂网络在经济管理领域的应用研究/安海岗,高
湘昀著. —北京:地质出版社,2014.8
ISBN 978-7-116-08904-4

I. ①复… II. ①安… ②高… III. ①信息产业—产
业经济—研究 IV. ①F49

中国版本图书馆CIP数据核字(2014)第174253号

责任编辑:柳 青

责任校对:李 玫

出版发行:地质出版社

社址邮编:北京市海淀区学院路31号,100083

电 话:(010)82324573(编辑室)

网 址:<http://www.gph.com.cn>

传 真:(010)82324579

经 销:北京中地金土图书发行有限公司

电 话:(010)82324508;(010)82324556

印 刷:北京地大天成印务有限公司

开 本:787 mm×1092 mm $\frac{1}{16}$

印 张:8.75

字 数:210千字

版 次:2014年8月北京第1版

印 次:2014年8月北京第1次印刷

定 价:35.00元

书 号:ISBN 978-7-116-08904-4

(如对本书有建议或意见,敬请致电本社;如本书有印装问题,本社负责调换)

前言

许多人可能会碰到过这样的事情：走在马路上，碰到了一个不熟悉的人，由于偶然原因你与他（她）聊起了天，聊天过程中，你竟然会发现，你的一个朋友同时也是他（她）的朋友。这时你们会深深地感叹：“真巧，这个世界真小啊！”那么地球上的任意两个人，中间平均需要经过多少个人能取得联系呢？哈佛大学教授 Stanley Milgram 通过调研得出的结果是：地球人任意两个人之间的平均距离是 6。在 Milgram 之后，其他学者又通过 Kevin Bacon 实验、互联网上的小世界实验进一步证明了“六度分离”理论的普遍适用性。

《Science》及《Nature》上的两篇论文被公认为是复杂网络领域最杰出的研究成果：一篇为《小世界网络的集体动力学 (Collective Dynamics of Small-World Networks)》，此论文由美国康奈尔大学的两个学者 Watts 与 Strogatz 于 1998 年在《Nature》杂志上联名发表；另一篇为《随机网络中标度的涌现 (Emergence of Scaling in Random Networks)》，此论文由 Barabási 与 Albert 在 1999 年的《Science》杂志上发表。以这两篇论文的发表为标志，从此复杂网络理论进入了新的研究时代。复杂网络理论从数学、物理学拓展到了管理学、经济学、生物学及社会科学等多个领域。

作者是从 2010 年开始酝酿撰写本书的，但是在撰写过程中发现撰写一本有关复杂网络理论的书籍非常困难，在撰写的过程中发现自己在理论及数理推理方面存在不足，撰写过程一度停滞不前。在复杂网络基本理论方面，上海交通大学汪小帆、李翔等专家学者编写的两本复杂网络著作《复杂网络理论及其应用》、《网络科学导论》已经对复杂网络理论做了系统详细的介绍。作者在撰写本书过程中也参考了这两本著作，在此向其作者表示感谢。

作者发现自己在理论及数理推导方面不足后，转换了一下思路。本书在内容上除了对复杂网络的基本概念和理论方法进行详细介绍外，着重于运用复杂网络理论对复杂网络领域新涌现的几个实例进行了实证研究。作者希望读者在阅读本书后能够具备较强的网络分析能力，同时不会陷入物理学和数学的理论推导，而是更关注网络科学的创新思维方式。

本书分为理论篇和实证篇两大部分，理论篇包括第 1~第 3 章，实证篇包括第 4~第 7 章。其中第 4~第 6 章为复杂网络领域最新涌现的实证研究；第 7 章主

要对电子商务领域的科研合作情况进行统计研究，此类实例国内外学者研究相对较多。本书虽然对复杂网络领域最新的一些实证研究进行了详细论述，但是难免存在不足，对于在某些方面不能满足读者期望之处，深表歉意。

本书在撰写过程中得到了石家庄经济学院管科学院电子商务教研室老师的大力支持，得到了中国地质大学（北京）的国家自然科学基金项目“基于复杂网络的虚拟社区关系挖掘与网络舆情演化分析”（71173199）、教育部人文社会科学规划基金项目（10YJA630001）、河北省高校重点学科建设项目、石家庄经济学院校内博士基金的资助，特此表示感谢。

安海岗

2014年夏

目 录

前 言

1 引 论	(1)
1.1 引 言	(1)
1.2 复杂网络发展历程	(2)
1.2.1 欧拉七桥	(2)
1.2.2 随机图理论	(3)
1.2.3 小世界现象	(3)
1.2.4 无标度网络	(5)
1.2.5 不同类型网络统计性质	(6)
1.3 本书主要内容	(7)
参考文献	(8)
2 复杂网络基本拓扑性质	(10)
2.1 网络的图的表示	(10)
2.2 基本拓扑性质	(11)
2.3 本章小结	(15)
参考文献	(16)
3 复杂网络的基本模型	(17)
3.1 随机网络模型	(17)
3.2 小世界网络模型	(18)
3.3 无标度网络模型	(20)
3.4 局域世界演化网络模型	(22)
3.5 本章小结	(23)
参考文献	(23)
4 时间序列单变量波动幅度研究	(24)
4.1 引 言	(24)
4.2 数据和方法	(25)
4.3 分析与结果	(29)
4.4 本章小结	(33)
参考文献	(33)

5	时间序列双变量联动波动复杂网络拓扑性质	(35)
5.1	引 言	(35)
5.2	价格联动性复杂网络的构建	(36)
5.3	联动性复杂网络拓扑性质	(39)
5.4	本章小结	(47)
	参考文献	(48)
6	网络新闻文本集关联复杂网络信息挖掘与可视化	(50)
6.1	引 言	(50)
6.2	相关理论概述	(52)
6.2.1	新闻的概念和要素	(52)
6.2.2	用户信息需求研究现状	(53)
6.2.3	复杂网络研究现状	(55)
6.2.4	文本挖掘研究现状	(57)
6.2.5	信息可视化研究现状	(57)
6.2.6	基于复杂网络的网络新闻文本集信息可视化研究	(60)
6.3	网络新闻文本集信息可视化理论模型的构建	(61)
6.3.1	网络新闻文本集信息可视化理论模型	(61)
6.3.2	网络新闻文本集信息可视化的要素及其形式化表示	(62)
6.3.3	网络新闻文本集信息可视化的过程	(66)
6.4	网络新闻文本集关键词的提取及其关联的建立	(67)
6.4.1	新闻文本关键词粒度的界定	(67)
6.4.2	网络新闻文本集关键词的提取	(70)
6.4.3	网络新闻文本集关键词关联的建立	(72)
6.5	实证研究	(74)
6.5.1	样本集的收集与用户需求分析	(74)
6.5.2	灾后重建网络新闻文本集关联可视化	(78)
6.5.3	汶川地震灾后重建网络新闻文本集演化可视化	(89)
6.6	本章小结	(99)
	参考文献	(100)
7	我国电子商务领域作者与机构合作网络分析	(104)
7.1	引 言	(104)
7.2	数据来源与研究方法	(104)
7.2.1	数据来源	(104)
7.2.2	研究方法	(104)
7.3	作者合作整体情况分析	(105)

目 录

7.3.1 合作网络数据统计分析	(105)
7.3.2 作者合作组群分析	(108)
7.4 机构合作网络构建与分析	(109)
7.4.1 机构信息预处理	(109)
7.4.2 选取权威机构	(110)
7.4.3 权威机构合作网络构建	(111)
7.4.4 机构合作网络整体分析	(112)
7.5 本章小结	(114)
参考文献	(114)
附录 A 伦敦金价格波动幅度及模态化表格	(116)
附录 B 各事件关键词	(125)
附录 C 各关键词总度数	(126)

1 引 论

1.1 引 言

复杂网络是物理学、数学、信息科学、运筹学等多学科交叉融合的一个学科。它诞生于 20 世纪中叶,进入 21 世纪后,掀起了一个复杂网络研究的新浪潮,即通过对人类社会中的各种网络结构和拓扑性质的研究来理解各种复杂社会现象。那么到底什么样的网络属于复杂网络呢?让我们从一些问题着手踏入复杂网络的学习旅程。

➤ 许多人可能会碰到过这样的事情:走在马路上,碰到了一个不熟悉的人,由于偶然原因你与他(她)聊起了天,聊天过程中,你竟然会发现,你的一个朋友同时也是他(她)的朋友。这时你们会深深的感叹:“真巧,这个世界真小啊!”

➤ 在互联网上畅游时,经常会思考从一个网页到另一个网页需要点击多少次鼠标?

➤ 当某一个计算机病毒在全球互联网上疯狂传播时,我们会好奇它是如何传播的,为什么传播速度如此迅速?

➤ 当在大都市乘车上下班时,为什么堵车现象如此严重?

➤ 2003 年,“非典”病毒为什么在全球造成了巨大伤亡,它是如何流行和传播的?

➤ 为什么一个国家或地区的局部金融动荡能够在较短时间内引发全球性的金融危机?

➤ 为什么人的大脑具有思维功能,它是如何运转的?

以上这些问题从表面上看,各不相同,但是它们之间却存在一个共同的特征——它们之间都存在着一个相互关联的复杂网络。越来越多的科学研究表明,这些互不相关的复杂网络存在着许多惊人的相似之处。

不同类型的复杂网络具有不同的特征,一般而言,其复杂性具有以下几个方面的共同特征^[1]:

(1) 节点复杂多样性

现实世界中包括许多不同类型的复杂网络,它们的节点一般都代表不同的含义,例如 WWW 复杂网络的节点可由不同网页构成,科研合作网络中节点由期刊论文中不同的专家学者构成,国际贸易复杂网络节点可由不同国家来表示等。

(2) 结构复杂性

不同类型的复杂网络,结构是不同的。从直观上来看,一般的复杂网络结构都是极其复杂、节点层层交叠的。此外,复杂网络的结构会随时间的推移发生一些难以预料的演化,例如互联网上每天都会新产生一些网页,同时一些过期的网页也会消亡,整个互联网的结构是复杂多变的。再者,有向复杂网络边还可能具有不同的权重和方向,例如神经网络、时间序列变量演变复杂网络等。

(3) 网络进化与演化

网络的进化与演化内涵比较具体,一般是指网络中某些节点(或边)的诞生或消亡,例如 WWW 网络中某些网站或网页发生改变后会导致整个复杂网络的结构发生变化。网络结构的进化与演化包括节点集的变化、边集合(或链接集合)的变化。

(4) 不同类型复杂因素的相互影响

现实社会中不同类型的复杂网络经常会受到各种因素的制约。例如,国际金融贸易网络会受到地区暴乱动荡影响,当某一地区发生动乱,可能会影响全球经济,进而对国际金融贸易复杂网络产生影响。复杂网络除了受到自身因素影响外,其他不同类型的复杂网络可能也会对其发生联系,对其产生影响。例如某一年,美国西部发生了大面积的停电事故,此事件不但造成了局部的交通瘫痪,甚至使得整个美国的互联网速度都变得异常缓慢。

当前社会属于一个网络社会,从我们人类的大脑神经网络到食物链复杂网络,从 2003 “非典” 在我国的迅速蔓延到互联网上各种计算机病毒的肆意传播,从国际金融贸易网络到全球性的金融危机,从大型电力网络到全球交通网络等,可以说,当前社会存在着各种各样的复杂网络,复杂网络已经把全人类包围了起来。被各种复杂网络包围起来的人类,面对网络化是喜忧参半,网络化可以提高整个人类社会生产和管理的效率,可以方便人们的生活,为人类社会的衣食住行提供各种方便,但是,它同时也给我们造成了许多不良影响,例如小范围停电导致大面积的停电事故、计算机网络病毒的快速传播等。因此,为了减少复杂网络对人类社会造成的危害,我们应当对其网络特征和传播机理进行深入的研究。长期以来,不同类型的网络属于不同学科的研究范畴,而复杂网络理论应对这些不同类型网络的共性特征及研究它们的普适方法进行研究。从复杂网络诞生到现在,它的研究已经渗透到管理科学、信息科学、生物科学、传播学等不同学科,复杂网络的研究被称为“网络的新科学”,已成为当前社会科学研究的一个极其重要的创新性课题^[2,3]。

1.2 复杂网络发展历程

1.2.1 欧拉七桥

复杂网络诞生后,国内学者对其进行了大量的研究,主要是研究各种复杂网络的统计特征。对其进行研究时,它们一般会选择一种合适的方法进行研究,我们一般将这种方法叫作“图论”。每个复杂网络都可以看作是由一些节点及与节点相连的边组成的复杂系统,即用抽象的点表示具体网络中的节点,并用点之间的连线表示具体网络节点之间的联系。

1736 年,还未成名的欧拉向学校提交了一篇学术论文,关于哥尼斯堡七座桥的连接问题。在对这个问题解决过程中,一门新的学科——图论诞生了。哥尼斯堡七座桥的问题,是指在哥尼斯堡中央,有一条河穿过,在这条河的边上有 2 个岛屿,在 2 个岛屿之间有 7 座桥,能够把它们连接在一起,一个普通的民众曾经提出一个问题,引起了人们的深深思考,即一个人能不能一次就顺利地通过所有的桥,但是每个桥不能重复走 2 次,最后还回到出发的原始地点。这个问题提出以后,许多人尝试了不同方法来解决它。最后,欧拉把它用数学几何方法进行了完美的诠释(图 1.1)。

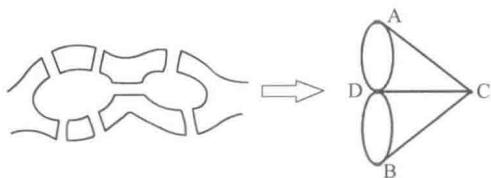


图 1.1 “欧拉七桥”问题

在欧拉的这篇论文中，它把每个陆地想象成一个顶点，把 7 座桥想象成 7 条边，这 7 条边用来连接不同的顶点。并由此得到了如图 1.1 中右图一样的几何图形。在论文中，如果分别用 A、B、C、D 等 4 个节点代表 4 块陆地，经过这样处理后，7 桥问题就可以看作是能不能用 7 条线连接 4 个节点的问题了。如果能够实现，那么在该网络图中，没有重复连线，并且连线的起点和最终点重合。在论文中，欧拉巧妙地解决了这个问题。欧拉对七桥问题的抽象论证开启了一个分支——图论。因此，他被称为“图论之父”^[4]。

1.2.2 随机图理论

哥尼斯堡七桥问题被数学家欧拉成功解决以后，由此问题产生的数学小分支——图论发展一直停滞不前。1930 年，图论的第一部专著问世，从此，图论进入了快速发展时期。20 世纪 60 年代，数学家 Erdős 和 Rényi 提出了一个对后人影响巨大的著名理论——随机图理论，该理论属于数学的一个分支，它从全新的角度开始了对复杂网络理论的研究^[5]。在随机图理论中，学者们把复杂网络中边的诞生描述为一定的概率，随机图理论被提出以后，许多学者对其进行了研究，提出了大量的模型，其中影响最大的是 ER 图模型。学者们提出的大量随机图模型中，一些影响较大且经典的模型，点的度数及其分布服从一定规律，例如泊松分布。以上我们所提的随机图一般称之为“狭义随机图”，在该理论诞生后的半个世纪后，“广义随机图”的概念也被提了出来。在“广义随机图”中，它的度数及分布可以拓展到任意情况，它的出现使得人类对复杂网络的理论研究达到了更深的层次^[6]。

自从随机图理论被提出以后，国内外的学者就开始运用该理论对复杂网络进行研究，随后取得了许多研究成果，但是其中有两个成果对复杂网络理论的发展影响最大。

1) 《小世界网络的集体动力学 (Collective Dynamics of Small-World Networks)》。该论文为《Nature》期刊上一篇评价极高的论文，是美国康奈尔大学的两个学者 Watts 与 Strogatz 联名发表的^[7]。

2) 《随机网络中标度的涌现 (Emergence of Scaling in Random Networks)》。该论文是《Science》期刊上一篇同样评价极高的论文，是美国圣母诺特丹大学的两位学者 Barabási 与 Albert 联名发表的^[8]。

这两篇论文的诞生，对复杂网络的研究产生了巨大影响，它们从宏观层面上对复杂网络的两个重要特征进行了系统科学的论述，这两个重要特征即小世界特征与无标度特征。从此以后，复杂网络的研究进入健康有序的发展轨道。

1.2.3 小世界现象

许多人可能会碰到过这样的事情：走在马路上，碰到了一个不熟悉的人，由于偶然原

因你与他（她）聊起了天，聊天过程中，你竟然会发现，你的一个朋友同时也是他（她）的朋友。这时你们会深深的感叹：“真巧，这个世界真小啊！”那么对于地球上的任何两个人，中间平均需要通过多少人能够建立联系？匈牙利著名作家 F. Karinthy 于 1929 年提出了“小世界”现象的假设。他大胆地假设：在地球上，任意两个互不认识的人，通过他们中间的亲人或朋友可以发生联系，而发生联系，中间平均通过 6 位朋友或亲人即可。例如，你要跟美国总统奥巴马发生联系，中间只要经过 6 个人即可。

1.2.3.1 Milgram 实验^[9]

小世界现象假设被提出后，20 世纪 60 年代，一个由哈佛大学教授 Stanley Milgram 设计的著名实验引起了美国人的关注。该实验过程如下：Stanley Milgram 给奥马哈市的任意 300 个人发送了一封信。在这封信中，他恳求收信者，通过自己认为最可能简短的方式，把这封信邮寄给一个在波士顿居住的确定的目标人，该实验得到了大部分收信人的配合。Milgram 告诉了每个发信人有关目标人的信息，包括姓名、所在地、职业，如果发信人不认识这个目标人，他们把这封信寄给他们认为有可能认识目标人的熟人。每个发信人根据 Milgram 的要求，通过他们认为最简短的方式，把信邮寄给了他们的亲人、朋友、同事。通过这个实验，Milgram 发现，最终有 60 个路径，信件能够最后交付到目标人手中，每条路径上的平均距离是 6，也就是说，地球上的任意两个陌生人，中间通过他们熟悉的 6 个人能够彼此发生联系。这就是“六度分离理论”的产生经过。

在复杂网络的发展历程中，Milgram 的小世界实验影响深远。该实验虽然取得了成功，但是也存在一些问题。比如实际过程中，只有 20% 的信件最终交付到了目标人手中，实验成功率较低；此外，由于实验样本数据较少，所以此实验结果的可信性有待推敲。虽然此实验存在以上不足，但是它毕竟从实证上验证了小世界现象的可信性，在复杂网络的发展历程上影响较大。

1.2.3.2 Kevin Bacon 实验

在 Milgram 的实验之后，为进一步验证“六度分离”理论的真实性和可信性，人们又进行了一些其他实验。另一个影响深远的例子是“凯文·贝肯游戏（game of Kevin Bacon）”。

游戏的方法是通过不停地寻找共同出演同一电影的演员，最终“找到”另一个“目标”演员。在贝肯游戏中，每一个演员都有一个统计指标：贝肯数。如果某一个演员与 Kevin Bacon 共同参演过一部电影，那么他（她）的贝肯数即为 1；如果一个演员没有与 Kevin Bacon 合作过，但是同与 Kevin Bacon 合作过的其他演员合作过，那么他（她）的贝肯数为 2，以此类推即可得到数据库中所有演员的贝肯数^[10]。例如，李连杰在《敢死队》中与史泰龙合作过，而史泰龙又跟贝肯合作过，那样李连杰的贝肯数即为 2。该实验中，收集了 130 多万名世界各个国家的演员，通过对他们的贝肯数进行统计，并求平均值，最后发现整个网络的平均贝肯数是 2.98，网络中最小的贝肯数为 0，是贝肯与他自己，最大的是 8。

另一个证明小世界现象的实验由数学家 Paul Erdős 提出的。Paul Erdős 是一位著名的数学系统科学家。该实验通过统计与 Paul Erdős 合作过论文的学者的 Paul Erdős 数来证明小世界现象，即与 Paul Erdős 合作过的专家的 Paul Erdős 数是 1，同 Paul Erdős 数为 1 的专家合作发表过文章的学者的 Paul Erdős 数是 2，以此类推可得到其他 Paul Erdős

数高于 2 的专家学者。该实验中,对美国的 40 多万名专家学者的 Paul Erdős 数进行了统计,然后求平均,最终得到平均值为 4.65,而最大的 Paul Erdős 数为 13^[11]。该实验,又一次证明了小世界现象的普遍性。

1.2.3.3 互联网上的小世界实验

尽管科学家们对电影演员合作网络和数学家合作网络的小世界现象进行了验证,但是它们的规模尺度太小,并不能科学地说明“小世界现象”的准确性。也就是说,近年来证明小世界现象的 Milgram 实验、Kevin Bacon 实验及 Paul Erdős 数实验,并不能直接应用到我们整个地球上 60 多亿人存在的复杂世界,即整个物理世界不一定符合小世界现象。

2001 年秋天,为了验证小世界现象的普遍适应性,美国哥伦比亚大学教授 Watss 带领他的几个学生与同事,开始了一个新的实验。他们首先建立了一个网站,然后通过该网站平台,在互联网上验证六度分离理论的正确性。该实验得到了不同国家、民族、年龄和社会阶层各类人的帮助。作为该实验的目标对象,他们积极参与了该实验。该小组招募了一些志愿者,志愿者在网站注册后会被告知关于目标对象的一些信息。这些志愿者主要任务是通过电子邮件,把消息发送到固定的目标对象手中,类似 Milgram 的小世界实验。如果志愿者不认识指定的目标对象的话,就给网站提供他觉得比较合适的一个朋友的电子邮件地址。然后,该网站会把这个实验的目的、意义及实验过程告诉这个对象,如果该对象愿意主动参加这个实验的话,那么这个实验就可以继续下去,以此类推。Watts 等把研究成果发表在了 2003 年 6 月的《Science》杂志上,该论文得到了广泛关注。在一年多时间里,总共有 13 个国家的 18 名目标对象和 166 个国家(地区)的 6 万多名志愿者参与实验,最后有 381 个志愿者的电子邮件抵达了目的地。通过对实验结果进行统计计算,发现每封邮件要交付到目标对象手中,平均需要转发 5~7 次。该实验虽然取得了成功,但是也存在一些侥幸。比如,某些对象可能对实验本身就没有兴趣或者自身很忙,没有时间参与,并且互联网上垃圾邮件的肆虐也在一定程度上对实验对象参与实验的热情有所冲击,从而使得该实验的顺利完成变得异常困难。

1.2.4 无标度网络

无标度(scale-free)是复杂网络的另一个重要特征。

在现实世界中存在的各种复杂网络,大部分都属于随机网络。许多网络都具有一些共同特征,即少量的节点往往包含大量的连接,而大量的节点却连接较少,这一般符合马太定律(即二八定律)。

不同类型的复杂网络度及度分布具有一些共同特征,度分布符合幂律分布的网络一般称为无标度网络。无标度网络具有严重的异质性,也就是说复杂网络中,不同节点之间的度数分布严重不均匀,度数分布服从马太定律。在复杂网络中,往往少数节点在网络中具有重要地位,对网络的运行起着关键作用,但是网络度数分布却具有严重的不均匀性。因此,从这个意义上来说,其无标度性也是复杂网络系统整体上所具有的一种内在属性^[12]。

1998 年 6 月,当《小世界网络的集体动力学》一文在《Nature》上发表的时候,Barabási 小组也开始以 WWW 为例研究复杂网络的结构。该小组通过大量的统计研究,发现 WWW 复杂有向网络的度(出度与入度)分布具有一定的规律性,即有幂律分布特征。这一点,与小世界网络模型及 ER 随机图网络模型的度分布明显不同,后者的度及度

分布具有明显的泊松分布特征。该小组将这一惊人发现发布在了 1999 年 9 月的《Nature》上。为了进一步验证幂律分布的普遍性，Barabási 教授向 Watts 发送邮件索要了电力网络和线粒虫神经网络的数据，并且发现这两类网络的度分布也服从幂律分布。于是，复杂网络领域另一个创新性的、具有重要影响的成果：“基于幂律分布产生的增长和优先连接机制建立的 BA 无标度网络模型”诞生了，Barabási 与 Albert 将研究成果发表在了 1999 年的《Science》杂志上，即《随机网络中标度的涌论》一文^[8]。

该成果发布后，Watts 等都非常懊悔，因为这本应该是属于他们的研究成果。他在后来回忆这件事情的时候，曾经这样说过：“当我在 1999 年 4 月的一个周末收到 Barabási 索要网络数据的邮件时，我还不知道他们想要什么。”“我们没有检查！我们深信非正态的度分布是不相关的，因此我们从来没有想到要看一看到底哪些网络服从正态分布，哪些网络则不服从正态分布。数据在我们手中几乎有两年的时间，而我们只需半个小时就可以做完检查。但是，我们从来没有想过去做。”

1.2.5 不同类型网络统计性质

当复杂网络领域两篇标志性的论文发表以后，复杂网络领域进入健康发展轨道。国内外学者对不同类型的复杂网络的拓扑结构进行了大量的统计研究，取得了大量的研究成果。表 1.1 列出了一些研究成果，包括复杂网络的各种统计参数，例如：网络类型（有向或无向）、节点总数 N 、平均路径长度 L 、平均度数 $\langle k \rangle$ 、边的总数 M 、聚类系数 C 。在该表中，如果该网络服从幂律分布，则给出幂指数 γ ，对于有向网络，则分别给出入度指数和出度指数；如果两者都不符合，则为“—”，表格中的空格代表该类网络没有切实可行的数据^[13]。

表 1.1 不同类型网络实际统计性质^[4]

网络	类型	N	M	$\langle k \rangle$	L	γ	C
电影演员	无向	449913	25516482	113.00	3.48	2.3	0.78
公司董事	无向	7673	55392	14.40	4.60	—	0.88
数学家合作	无向	253339	496459	3.92	7.57	—	0.34
合作物理学家	无向	52909	245300	9.27	6.19	—	0.56
合作生物学家	无向	1520251	11803064	15.50	4.92	—	0.60
电话呼叫图	无向	4.7E + 07	8E + 07	3.16			
电子邮件	有向	59912	86300	1.44	4.95	1.5/2.0	0.16
电子邮件地址	有向	16881	57029	3.38	5.22	—	0.13
学生关系	无向	373	477	1.66	16.00	—	0
性关系	无向	2810				3.2	
WWW (nd. edu)	有向	269504	1497135	5.55	11.30	2.1/2.4	0.29
WWW (Altavista)	有向	203549046	2.13E+0.9	10.50	16.20	2.1/2.7	
电力网	无向	4941	6954	2.67	19.00	—	0.08
铁路网	无向	587	19603	66.80	2.16	—	0.69
神经网络	有向	307	2359	7.68	3.97	—	0.28

1.3 本书主要内容

仍然健在的最伟大的物理学家,被称为宇宙之王的霍金曾经说过:“我认为下个世纪将是复杂性的世纪(I think the next century will be the century of complexity)”。网络与复杂网络已成为 21 世纪的新科学领域。网络科学具有很强的跨学科特色,既有从纯物理学、数学角度进行的完全抽象的纯理论性分析,也有针对具体网络的实证和应用研究,并且新的研究成果不断涌现。本书在材料选择上着重于复杂网络的基本概念和方法,并运用复杂网络理论对几个新涌现的不同实例进行实证研究,希望读者在阅读本书后能够具备较强的网络分析能力,同时又不会陷入物理学和数学的理论推导,而是更关注网络科学的创新思维方式。

本书分为理论篇和实证篇两大部分。理论篇包括第 1~第 3 章,实证篇包括第 4~第 7 章。其中第 4~第 6 章为复杂网络领域新涌现的研究实证;第 7 章主要对电子商务领域的科研合作情况进行统计研究,此类实例国内外学者研究相对较多。

“1 引论”,主要系统介绍了复杂网络理论诞生与发展的历程,现实社会中存在的复杂网络类型及各种统计属性。

“2 复杂网络基本拓扑性质”,首先对网络的图的表示进行了详细论述,然后对复杂网络各种拓朴性质的概念及计算方法进行了详细的论述,例如,度与度的分布、平均路径长度、聚类系数、中介中心性及凝集子群等。

“3 复杂网络的基本模型”,主要对自复杂网络理论诞生以后,国内外学者提出的影响较大的网络模型进行了详细介绍,包括随机网络模型、小世界网络模型、无标度网络模型及局域世界网络演化模型,对各种网络模型的内涵、特征及构造算法进行了详细论述。

“4 时间序列单变量波动幅度研究”,选择伦敦金下午收盘价格作为样本数据,借鉴统计物理学的方法进行研究。利用粗粒化方法建立了价格波动幅度变化模态,运用复杂网络理论对时间序列单变量波动幅度模态的统计、变化规律和演化规律进行了分析。统计结果表明,时间序列单变量波动幅度具有一定的规律,例如群簇性、度分布的幂律性、网络演化的周期性。研究结果还表明,网络演化波动过程中,波动幅度模态的演化与转化重点通过几种主要模态实现。研究成果不仅可以作为不同类型时间序列单变量波动幅度研究的方法,同时可为多变量波动幅度及其联动波动规律研究提供思路。

“5 时间序列双变量联动波动复杂网络拓扑性质”,选择时间序列双变量(伦敦金与 Au9999 下午收盘价格)作为样本数据。依据粗粒化方法,将 Au9999 与伦敦标准金价格的联动波动转化为由 {P, O, M} 3 个字符构成的字符串;以周为单位,将每周价格波动转化为 5 个 {P, O, M} 字符组成的字符串,每一个字符串代表每 5 天的黄金现货价格联动波动模态,共产生 106 个模态。构建国内外黄金价格复杂网络时,以黄金价格联动波动模态为节点,不同模态之间的转化为边。构建复杂网络后,对此网络的各种统计参数进行统计研究,结果表明点强度较大的 32 个节点累积强度占比较大,为 92.05%,点强

度与度分布、点强度与其等级均呈幂律分布；加权集聚系数与点强度之间并没有表现出良好的相关性，网络中存在 14 个小群簇；黄金价格联动波动具有一定的周期，复杂网络平均路径长度为 7.67，即周期约等于 8 天；节点中介性差异不太明显，前 40 个节点对整个网络的中介中心性贡献率为 62.29%；8 个节点的凝集子群有 2 个，8 个节点以上的子群不存在。上述研究成果，从网络基本拓扑性质方面证实了此复杂网络的各种特征，这些特征对于掌握我国国内外黄金价格变化规律及国内外黄金价格联动波动的规律具有参考价值；能够为我国黄金价格制定、风险投资和规避经济风险提供决策参考。研究成果不仅可以作为不同类型时间序列双变量联动波动的研究方法，同时可为多变量联动波动规律研究提供思路。

“6 网络新闻文本集关联复杂网络信息挖掘与可视化”，首先对信息挖掘与可视化的相关理论、概念及模型进行综合分析之后，根据新文本要素自身的特点，结合网络新闻文本集用户的一些信息需求，对网络新闻文本集信息实体的概念进行了描述；提出网络新闻文本集信息可视化理论模型，并对模型的要素和过程进行详细分析。然后以从互联网上收集的有关汶川大地震灾后重建的网络新闻文本集为例做实证研究，在对用户进行信息需求分析之后，构建网络新闻文本集关联矩阵，然后对网络文本集复杂网络进行信息挖掘与分析，并对可视化的结果进行分析。

“7 我国电子商务领域作者与机构合作网络分析”，从社会网络研究角度出发，选取 CNKI 北大核心期刊中 2010~2012 年标题包含“电子商务”的文章为研究对象，对作者合作与机构合作网络进行研究。对 2010~2012 年 3 年的作者合作网络及其中几个颇有影响的团体合作网络参数进行了统计分析；引入核心机构和权威机构的概念，并绘制权威机构合作网络图，对其网络密度和中心性等进行统计分析，有助于了解现阶段我国电子商务领域合作的整体现状和特征。

参 考 文 献

- [1] Strogatz S H. Exploring complex networks [J]. Nature, 2001, 410 (6825): 268-276.
- [2] Barabási A-L. Linked: The New Science of Networks [M]. Massachusetts: Persus Publishing, 2002.
- [3] Watts D J. The “new” science of networks [J]. Annual Review of Sociology, 2004, 30: 243-270.
- [4] 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.
- [5] Erdős P, Rényi A. On the evolution of random graphs [J]. Publ. Math. Inst. Hung. Acad. Sci., 1960 (5): 17-60.
- [6] Newman M E J. The structure and function of complex networks [J]. SIAM Review, 2003, 45: 167-256.
- [7] Watts D J, Strogatz S H. Collective dynamics of small-world networks [J]. Nature, 1998, 393 (6684): 440-442.
- [8] Barabási A-L, Albert R. Emergence of scaling in random networks [J]. Science, 1999, 286 (5439): 509-512.

- [9] Milgram S. The small world problem [J] .Psychology Today, 1967 (2): 60-67.
- [10] 汪小帆, 李翔, 陈关荣. 网络科学导论 [M] . 北京: 高等教育出版社, 2012.
- [11] Borosh I, Chui C K, Erdős P. On changes of signs in infinite series [J] .Anal. Math. , 1978, 4 (1): 3-12.
- [12] Barabási A-L, Albert R, Jeong H. Mean-field theory for scale-free random networks [J] .Physica A, 1999, 272: 173-187.
- [13] Newman M E J. The structure function of complex networks [J] .SIAM Review, 2003, 45: 167-256.