

摘 要

现实世界中,存在大量的复杂系统,这些系统都可以用复杂网络来表示,例如,Internet 网络,演员合作网、新陈代谢网、航空网等。这些不同类型系统的拓扑结构有着纷繁复杂的表现,但普通网络又表现出三大共性:连接度呈幂率分布,较高的集聚系数以及较短的平均路径。这些共性引起了科学工作者的极大兴趣,并引起了复杂网络的研究浪潮。本文将数学解析方法、分形理论应用到复杂网络的研究当中,通过建立一些改进的复杂网络模型来对复杂网络的宏观拓扑结构进行研究,用来模拟真实网络的拓扑结构和相关性质。

本文首先介绍了复杂网络研究背景和意义以及复杂网络研究的现状,并详细地解释了复杂网络和分形所涉及的基本概念。然后针对复杂网络模型研究的现状,提出了以下三个模型:

1、提出了谢尔宾斯基支架的具有分形和小世界特性的网络。利用数学归纳的方法计算出了网络图的集聚系数,平均最短路径和网络图的直径,证明了网络图的小世界特性。用盒维数和相似维数来衡量网络图的分形性,并得到该网络的集聚系数为 0.53、网络的直径为 4、平均路径长度为 2.5,平均度分布为 7.5,证明了该网络的小世界特性;然后测定该网络的盒维数和相似维数都约为 2。

2、研究了谢尔宾斯基地毯的分形和小世界特性。利用数学归纳的方法计算出了网络图的集聚系数,平均最短路径和网络图的直径,证明了网络图的小世界特性。用盒维数和相似维数来衡量网络图的分形性,并得到该网络的集聚系数为 0.054,网络图的直径为 6,平均路径长度小于 6,测定盒维数和相似维数都约为 1.8928。

3、提出了一个具有分形和小世界特性的网络图。本文利用数学归纳的方法计算出了网络图的集聚系数,平均最短路径和网络图的直径,证明了网络图的小世界特性。用盒维数和相似维数来衡量网络图的分形性,得到该网络的集聚系数为 0.44、网络的直径为 4、平均路径长度为 2.67,平均度分布为 5.3,证明了该网络的小世界特性;然后测定该网络的盒维数和相似维数都约为 1.585。最后对网络图的构造方法作了进一步地拓展,并给出了拓展的网络图的相关拓扑特性的表达式,并将其和原来的网络图可归结为一类具有分形和小世界特性的网络图。

关键词: 谢尔宾斯基图形; 复杂网络; 分形; 小世界网络; 无标度网络

The Study on the Evolving Model of three Sierpinski Graphs And Their Fractality Abstract

In the real world, there are many systems, which can be represented by complex networks, such as Internet, movie actor collaboration network, metabolic networks, aviation network, and so on. Many empirical studies indicate that these various kinds of networks have some common characteristics: power-law distribution of degree, high clustering coefficient, and small average path length. These characteristics receive the researchers' great interest and considerable attention. This thesis takes mathematic analytical method, fractal, these methods into the complex network, and builds some improved models on the complex network to study on the character of the complex network.

This thesis first introduces the background and the significance of the complex networks and the relevant conceptions are explained, too. Then according to the situation of complex studying, three new better models are proposed:

1. On the basis of the Sierpinski tetra, we propose a kind of complex network in high dimensions with fractality and small-world effect. Using mathematical induction, we obtain that the clustering coefficient is 0.53, the diameter is 4, the average path length is 2.5, and the average degree is 7.5. The values of these properties testify that the network is small-world. Then we calculate the box-counting dimension and Similarity dimension as a measure of their fractality and the value of the dimension is about 2.

2. On the basis of the Sierpinski carpet, we construct a kind of Sierpinski network. Using mathematical induction, we obtain that the clustering coefficient is 0.054, the diameter is 6, the average path length is less than 6, and the average degree is 4.61. The values of these properties testify that the network is small-world. Then we calculate the box-counting dimension and Similarity dimension as a measure of their fractality and the value of the dimension is about 1.8928. The similarity and elaborate structure proves the fractality of the network.

3. A class of graphs with fractality and small-world effect is proposed. The clustering coefficient, the average length of the graphs, and the diameter are calculated analytically. Also we calculate the box-counting dimension and Hausdorff dimension as a measure of their fractality and the value of the dimension is 1.585. At last, the method of the constructing graphs is expanded, and the expressions of the relevant characteristics of the expanded graphs

are given. The expanded graphs and the graphs proposed before are considered as a class of graphs with fractality and small-world effect.

Key Words: Sierpinski Graphs; Complex Networks; Fracality; Small-world Networks; Scale-free Networks;

独创性说明

作者郑重声明：本硕士学位论文是我个人在导师指导下进行的研究工作及取得研究成果。尽我所知，除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得大连理工大学或者其他单位的学位或证书所使用过的材料。与我一同工作的同志对本研究所做的贡献均已在论文中做了明确的说明并表示了谢意。

作者签名： 日期：2008-6-11

大连理工大学学位论文版权使用授权书

本学位论文作者及指导教师完全了解“大连理工大学硕士、博士学位论文版权使用规定”，同意大连理工大学保留并向国家有关部门或机构送交学位论文的复印件和电子版，允许论文被查阅和借阅。本人授权大连理工大学可以将本学位论文的全部或部分内容编入有关数据库进行检索，也可采用影印、缩印或扫描等复制手段保存和汇编学位论文。

作者签名：

穆青

导师签名：

郝日祥

2008年6月10日

1 绪论

1.1 复杂网络的研究背景

随着科学技术的飞速发展和社会的快速进步,越来越多的复杂事物和现象呈现在人们的面前,例如生态、环境、计算机网络、统计方法、工程技术与人文社会相结合等诸多社会经济问题有待解决。正是在这样的背景下,复杂性科学开始孕育、萌芽,并受到越来越多的学者的关注,如今复杂性科学的研究方兴未艾,正如著名的物理学家霍金更是认为“二十一世纪是复杂性的世纪”。中国科学院院士戴汝为认为,“对整个科学的发展来讲,对于复杂系统的精确、更完整的描述,是科学最大的挑战”,并把复杂性科学提到了这样一个高度。而复杂网络理论是借助于图论和物理的一些方法,来研究各种看上去互不相同的复杂系统、复杂网络之间的共性和处理它们的普适方法。中国科学院院士许雷则认为,复杂网络可以用来描述从技术到生物直至社会各类开放复杂系统的骨架,是研究它们拓扑结构和动力学性质的有力工具^[4,68]。正因如此,从新世纪以来,复杂网络开始不断蓬勃地发展。复杂网络的研究最早可以追溯到欧拉的七桥问题^[1]。欧拉将一个小镇每周六所进行的娱乐活动——一次性走过所有七座桥的散步,画在了纸上(将陆地抽象为节点,而桥则成为了点与点之间的连线),并通过推理得出“不可能一次性不重复地走完这七座桥”,从而由此开创了数学的一个分支——图论。在欧拉解决七桥问题之后的相当长的一段时间里,图论并没有得到足够的发展。直到20世纪60年代,由两位匈牙利数学家Erdős和Rényi建立了随机图理论^[2],并被公认为是在数学上开创了复杂网络理论的系统性研究。他们认为两个节点是否相连不再是确定的事情,而是用由一个概率 P 决定,这种网络又称随机网络,网络中的节点的连接数目大致相同,其分布遵循泊松分布。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,被认为是描述真实最适宜的网络。

随着计算机技术的高速发展,存储各种复杂系统的拓扑性数据成为了可能,人们开始利用计算机高速计算的能力来分析这些大量的数据,来探查这些复杂系统的整体特性。在这些大量的统计分析中发现,人们发现万维网^[3]、好莱坞的演员合作网^{[7][8]}、科学引文网^[5]、细胞代谢系统^[6]等这些各种不同类型的复杂系统、复杂网络并不是完全随机的,但又表现出相同的统计属性。美国康奈尔大学的Watts和Strogatz在Nature杂志上发表的《Collective Dynamics of ‘Small-World’ Networks》^[7]和美国Notre dame大学的Barabási和Albert在Science杂志上发表的《Emergence of Scaling in Random Networks》^[8]这两篇文章开创了复杂网络的研究新纪元,它们分别揭示了复杂网络的小世界特性和无

标度性质，并建立了相应的模型以阐述这些特性的产生机理。这两种网络的出现和大量真实网络的实证研究成果，促使人们去探索现实网络的内在演化机制，从此掀起了研究复杂网络的热潮。

复杂网络模型各式各样，按照不同的方式，可以分成不同的类别^[21]，见下图：

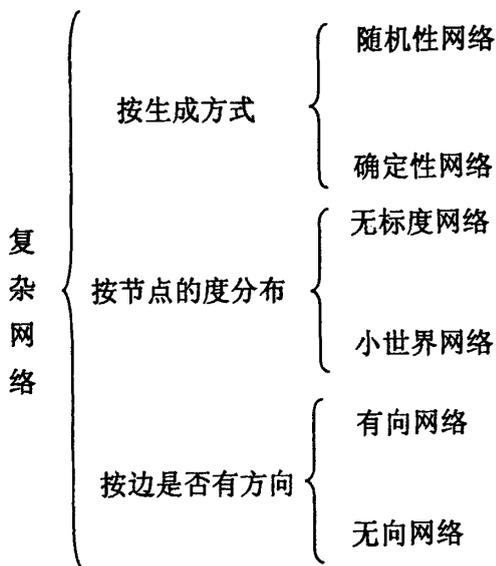


图1.1 复杂网络分类图

Fig1.1 The classification Chart of complex network

根据上面的分类，可将本文提出的具有分形和小世界的网络图是确定性的、小世界、无向网络模型。

1.2 复杂网络的研究意义

复杂网络的研究大多采用数学和统计物理学的方法，而这为研究者提供了一种复杂性研究的新视角、新方法。通过复杂网络的研究可以对各种复杂系统网络进行比较、研究和综合概括。正因为如此，复杂网络刚一提出，就在各个领域都呈现出广泛的应用前景。

世界上大部分的事物都可以看成一个有各个元素组并相互联系的系统，而系统就可以用网络图来表示，所以对网络的研究极具意义。科学家通过大量的统计分析发现大多数实际的系统都是复杂网络，从 Internet 上的路由器之间连成的网络^[9]，电力网^[10]，科学家之间的合作网络，演员的合作网，细胞和蛋白质系统，到人类性关系^[11]，论文之间的引用关系^[12]等，它们都构成某种复杂网络系统。在这些复杂网络系统中，如果发现一种能够提出它们共性的观点和方法，则能够对这类网络的形成深入的认识。而复杂网络

的实证研究发现这些不同的复杂系统之间,却有着共同的特征:节点的连接度呈幂率分布,较高的集聚系数、以及较短的平均路径^{[13][14]}。

网络的节点的连接度呈幂率分布,决定了网络由少数几个具有众多连接的节点所支配,这一性质决定了网络的鲁棒性和脆弱性。所谓鲁棒性,即指网络移走少量节点后,网络中的绝大部分节点仍是连通的,而脆弱性则是恰好相反。无标度性决定了现实网络的这双重特性,对随机节点故障具有极高的鲁棒性,而对蓄意攻击却具有高度的脆弱性。这给现实系统带来很大的启发意义。例如,在网络预防黑客攻击方面,为了更有效地保护网络,可以主要针对保护那些大的门户网站^[15];在疫苗注射上,可以采取直接或间接地针对那些与很多人具有连结关系的人接种疫苗,这样可以达到很好的效果^[16]。

网络的较高的集聚系数反映了事物在小世界的情况下自发走向有序的态势,较短的平均路径特征反映了演化速度快的特征。在疾病传播方面,科学研究者通过建立复杂网络模型^[17]发现,对于从规则网络到随机网络的所有传染概率 $p \in (0,1)$ 网络,其传染病的扩散时间刚好与最短路径一致,这很好地说明了最短路径这一几何量的作用。这使得我们明白在控制疾病传播时,不仅仅是要提高治疗的医学问题,而且考虑如何切断网络。

目前复杂网络研究已经扩展到复杂系统的各个方面,研究网络上信息的传播,网络上的相继故障,网络中的搜索,以及同步、控制等等,研究者都积极用复杂网络的理论解决自己专业领域的问题,已经在经济、社会、管理各个领域显示出强大的生命力^{[18]-[20]}。

1.3 复杂网络的研究现状及存在的问题

复杂网络演化模型作为复杂网络的重要研究方向,经历了一个从简单到复杂的过程,由于复杂系统之间所表现的共性,使得人们积极地去建立一种通用的模型来模拟现实网络。越来越多的复杂网络模型正在被提出来,并一步一步地接近现实网络。下面简要讨论复杂网络模型发展中的几个关键模型。

一、规则网络和随机网络

我们把一维链、二维正方晶格等称为规则网络^[23]。规则网络是指平移对称性晶格,任何一个晶格点的近邻数目都相同。比较典型的规则网络有:全局耦合网络、最近邻耦合网络和星形网络。全局耦合网络就是指任意两个点之间都有边直接相连,所以其平均最短路径和集聚系数的值均为 1(见图 1.2(b))。最近邻耦合网络是指网络中的每个节点只能和它周围的节点相连(见图 1.2(a))。而星形耦合网络是指网络中有一个中心点,其余的所有点都只与这个中心点相连,而它们彼此之间却不连接(见图 1.2(c))。具体下图 1.2:

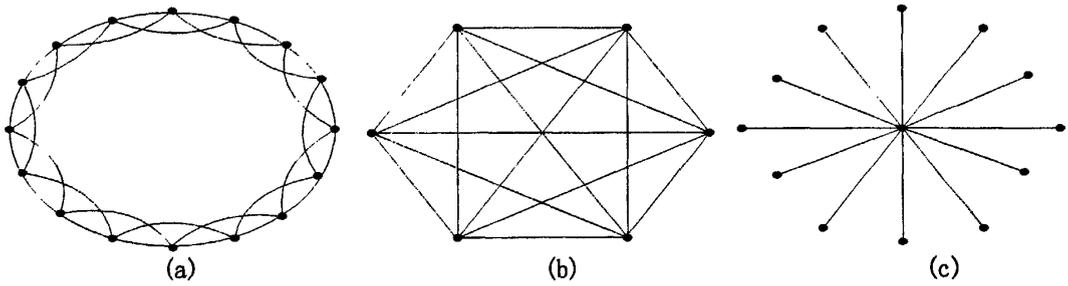


图 1.2 规则网络图

Fig 1.2 The Graph of rule network

20 世纪 50 年代末期, 匈牙利数学家 Erdős 和 Rényi 首次将随机性引入网络的研究, 提出了著名的随机网络模型, 简称 ER 模型。该模型是给定 N 个节点, 每一对节点以相同概率 p 进行连接边。如果 $p = 1$ 则表示 N 个节点完全连接, 总边数为 $N(N - 1)/2$; 如果 $p = 0$, 则 N 个节点相互孤立; $p \in (0, 1)$ 时, 网络中的总边数为 $pN(N - 1)/2$ 。具体见下图。

随机网络的主要特征是: 1) 聚集系数较小; 2) 网络平均距离小; 3) 节点度服从 Poisson 分布。

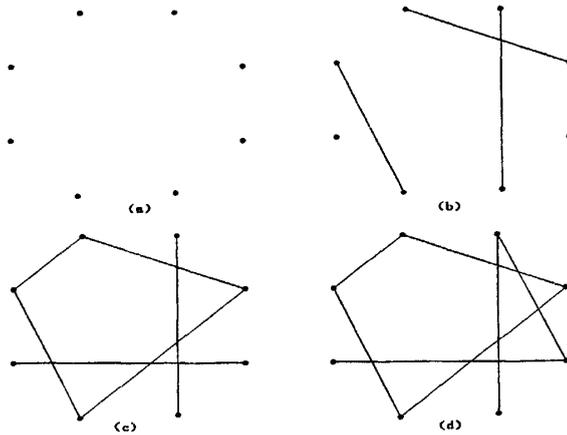


图 1.3 随机网络图

Fig 1.3 The Graph of random network

注 1) (a) $p=0$, 给定的 8 个孤立点; (b) ~ (d) 分别以概率 $p=0.1$ 、 $p=0.2$ 、 $p=0.25$ 生成的随机图。

二、小世界网络

规则网络模型和 ER 随机图模型都不能体现真实网络的一些重要特征, 毕竟大部分实际网络既不是完全规则的, 也不是完全随机的。作为从完全规则网络向完全随机图的过渡, Watts 和 Strogatz 在 1998 年建立了著名的小世界网络模型^[7], 称为 WS 小世界模

型。WS 模型是从一个具有 N 个节点的规则网络开始，每个节点与其 k 个最近邻连通；然后，以概率 P 重新选取随机边，也就是对每个节点的边，我们以概率 P 重连，保持边的总数目不变。具体见图 1.4。

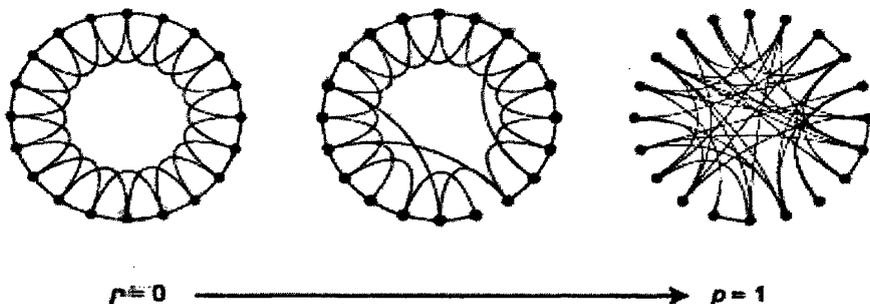


图 1.4 WS 模型图^[7]

Fig 1.4 The Graph of WS model

在 WS 模型中， $P=0$ 对应于随机网络， $P=1$ 则对应于完全随机网络，通过调节 P 的值，就可以控制从完全规则网络到完全随机网络的过渡。WS 模型结合了规则网络较大的簇系数和随机网络较小的平均最短距离的特征，很好地描述了真实网络的小世界特性。

但 WS 网络模型的随机化构成有可能破坏网络的连通性。另一个研究比较多的是 Newman 和 Watts 提出的小世界模型^[24]，它是通过“随机化加边”的方式来取代 WS 模型中的“随机化重连”，并且对充分小的概率 P 和充分大的 N ，该模型等价于 WS 模型，同时也避免了 WS 模型中由于重连而可能产生孤立点的发生。

三、BA 无标度网络模型

Barabási 和 Albert 在对许多真实世界的网络进行研究时发现许多网络的度的分布函数呈现出幂律分布而非钟形的泊松分布的特点^[8]，如图 1.5 和图 1.6。他们分析万维网的产生机理，提出了无标度网络产生的两个基本机制：增长和择优连接。增长是指网络不是静态不变的，而是不断演化不断增长的过程，是区别于小世界网络和随机网络的固定节点个数的；而择优连接意味着网络中节点之间的连接不是均等的，而是有偏好的。他们认为网络是在这两种规则下不断地生长演化的，并建立了著名的无标度网络演化模型，简称 BA 模型^[25]。

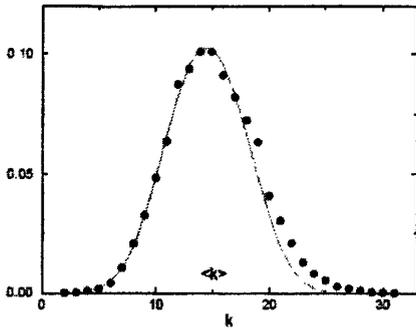


图 1.5 泊松度分布
Fig 1.5 The Poisson distribution

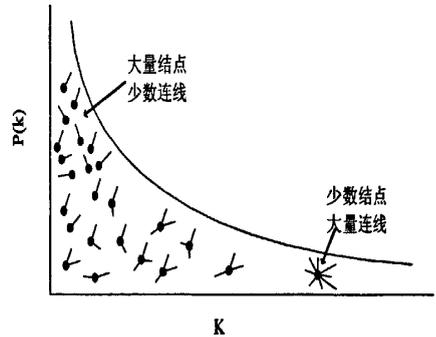


图 1.6 幂律度分布
Fig 1.6 The power law distribution

小世界网络模型和BA无标度网络模型的提出，掀起了复杂网络的研究浪潮。在这两个模型的基础上，许多研究者又提出了许多扩展模型和新的演化机制。Krapivsky等人^[26]研究了非线性的择优连接机制，并证明线性择优连接时的幂率分布指数 $\gamma \in (2, \infty)$ 。

Dorogovtsev和Mendes^[27]考虑了老节点之间的连接或断开的情况，用节点的连通的平均度的变化率来建立模型，并计算出该模型的网络幂指数 $\gamma = 2 + \frac{1}{1+2c}$ ，其中 $c \in R$ 。他

们还将节点的老化机制加入网络的演化当中，让新节点与老节点连接的概率和老节点的年龄相关^[28]。Klemm K, Eguíluz V M也考虑了节点的老化情况^[29]，提出了基于节点有限记忆量的网络增长模型，证明了现实网络中的节点的年龄和它被连接的概率呈负相关关系。

自从小世界网络的经典模型WS模型^[7]的提出，人们对具体网络的小世界特性进行了深入的研究，并提出了大量的小世界网络模型^[31-32]。在2005年的nature上的“Self-similarity of complex networks”中^[46]，Song等人用分形理论中的分维来描述网络中的自相似性，给人提供了研究复杂网络的新角度。他们采用盒维数的方法来计算网络的维数证明其自相似性，并进一步通过重整化过程，证明自相似性和无标度的分布在网络的所有粗粒化过程阶段都成立。他们在随后的研究中，又将复杂网络出现的这种分形性归因于连接度高的节点之间的互斥性^[33]。Gim和Goh等人^[34-36]则从一种称为骨架（skeleton）的树结构出发，来研究分形网络的成因，认为最初的分形网络可以看作是这种分形骨架和一些点与点之间的捷径组成。目前关于分形网络形成原因的研究还不是太多。

从Dorogovtsev等人提出的伪分形网络模型^[47]开始，研究者开始建立各种具有分形特征的网络模型。其中以阿波罗网络^[37]最为著名，章忠志等人先后建立了高维的阿波罗网

络^[38]，高维的随机阿波罗网络^[62]，以及阿波罗网络的一般模型^[63]，周涛等则建立了二维随机阿波罗网络模型^[64]。古志鸣^[65]则用单纯形剖分生成了一个类似阿波罗网络的模型。在这些模型演化的网络中虽然都具有一定的分形特征，但并没有用分形的相关理论方法证明其分形性。

在复杂网络模型以几乎每天产生一个新模型的速度产生时^{[39]-[45]}，2005年的nature上一篇名为“Self-similarity of complex networks”文章，给人提供了研究复杂网络的新角度^[46]。文章作者Song等创造性地用分形理论中的分维来描述网络中的自相似性，采用盒维数的方法来计算网络的维数证明其自相似性，并进一步通过重整化过程，证明自相似性和无标度的分布在网络的所有粗粒化过程阶段都成立。其实从2002年Dorogovtsev等人提出的伪分形网络模型开始^[47]，研究者开始建立各种具有分形特征的网络模型。阿波罗网络模型就是具有明显的自相似特征的^[38]。古志鸣用单纯形剖分生成了一个类似阿波罗网络的模型^[43]。在song这篇文章之后，研究者逐步开始建立具有分形性的网络演化模型，并用这种盒维数方法来证明其分形性^{[48]-[54]}。Gim和Goh等人则从一种称为骨架(skeleton)的树结构出发来建立网络模型，并分析研究分形网络的成因，认为最初的分形网络可以看作是这种分形骨架和一些点与点之间的捷径组成^{[48]-[50]}。Hernán等人则从伪分形模型进行扩展，提出迭代的无标度网络，并证明其分形性^[52]。Didier Sornette等人则提出了用边覆盖来计算复杂网络分维数^[53]。D.J.B.Soaes建立了一个空间填充的复杂网络模型，并用盒维数的方法证明其分形性^[54]。Song在随后的研究中，又将复杂网络出现的这种分形性归因于连接度高的节点之间的互斥性^[55]，并提出了计算盒维数的算法^[56]。可以认为复杂网络的自相似性正在成为一个重要的研究方向。

复杂网络演化模型研究发展地很快，模型考虑的因素也越来越多，也越来越接近于实际现实网络，但仍存在着以下问题：

- 1、虽然复杂网络的演化模型从复杂网络的兴起至今，已具有相当的数目，并且研究了各个方面的复杂网络特性，我们发现绝大多数文章，都是在计算机的帮助下，通过随机模拟等算法，对产生出的统计数据进行分析，这一方面有利于我们认识错综复杂的外部世界，但另一方面的要认识复杂世界，还需要对我们定量的认识复杂网络，多角度观察会有助于我们对复杂网络的认识。

- 2、复杂网络的自相似的形成原因研究还不是太多，对网络复杂性的测度没有一个行之有效的方法，模型更多的只是用盒维数的方法去证明其分形性，而没有将分形理论更多的内容和复杂网络结合起来考虑。

3、模型与现实网络结合的不够紧密。尽管复杂网络有大量的实证研究，但是实证与模型演化脱节情况尚未解决，网络模型设定的一些生成演化机制，还不能反映真实网络的真正特性。根据实际网络来定具体机制的模型还比较少。

1.4 本文的研究内容

根据 1.3 所述的复杂网络演化模型及自相似研究的现状和所存在的问题，研究了由谢尔宾斯基经典图形构造的一系列网络拓扑性质，并得到这些网络的平均路径长度随着网络规模的无限扩大趋于一个固定的常数。本文主要做了以下方面的工作：

(1) 提出了谢尔宾斯基支架的具有分形和小世界特性的网络。利用数学归纳的方法计算出了网络图的集聚系数，平均最短路径和网络图的直径，证明了网络图的小世界特性。用盒维数和相似维数来衡量网络图的分形性，并得到该网络的集聚系数为 0.53、网络的直径为 4、平均路径长度为 2.5，平均度分布为 7.5，证明了该网络的小世界特性；然后测定该网络的盒维数和相似维数都约为 2。

(2) 研究了谢尔宾斯基地毯的分形和小世界特性。利用数学归纳的方法计算出了网络图的集聚系数，平均最短路径和网络图的直径，证明了网络图的小世界特性。用盒维数和相似维数来衡量网络图的分形性，并得到该网络的集聚系数为 0.054，网络图的直径为 6，平均路径长度小于 6，测定盒维数和相似维数都约为 1.8928。

(3) 提出了一个具有分形和小世界特性的网络图。本文利用数学归纳的方法计算出了网络图的集聚系数，平均最短路径和网络图的直径，证明了网络图的小世界特性。用盒维数和相似维数来衡量网络图的分形性，得到该网络的集聚系数为 0.44、网络的直径为 4、平均路径长度为 2.67，平均度分布为 5.3，证明了该网络的小世界特性；然后测定该网络的盒维数和相似维数都约为 1.585。最后对网络图的构造方法作了进一步地拓展，并给出了拓展的网络图的相关拓扑特性的表达式，并将其和原来的网络图可归结为一类具有分形和小世界特性的网络图。

1.5 本文的结构

文章剩余部分组织如下：

第 2 章着重介绍了复杂网络的基本理论概念。首先对图论的一些基本概念作了介绍，描述了复杂网络了一些基本属性，然后对复杂网络的重要的特征度量参数作了介绍；然后介绍了分形的有关背景及相关知识，最后解释了复杂网络的自相似性。

第 3 章分析了基于谢尔宾斯基支架的具有分形和小世界特性的网络。本章首先介绍了基于谢尔宾斯基支架的构造过程。然后提出网络模型的演化机制，并解析计算出其网

络的集聚系数、平均路径长度、及网络的平均度，并发现当网络的规模无限增大时网络的平均路径长度趋于一个常数。

第 4 章建立并分析了基于谢尔宾斯基地毯的具有分形和小世界特性的网络。本章首先介绍了基于谢尔宾斯基地毯的构造过程。然后提出网络模型的演化机制，并解析计算出其网络的集聚系数、平均路径长度，并发现当网络的规模无限增大时网络的平均路径长度在一定的区间内部。

第 5 章建立了具有分形和小世界特性的网络图。首先介绍谢尔宾斯基垫片的相关内容和分形维数的基本概念。然后提出本网络图的构造法则，并解析计算出网络图的相关拓扑特性证明其小世界和分形性。最后对网络图的进一步的拓展深化，定义了一个一般性的分形和小世界特性的网络图。

最后，对研究结果进行总结，给出了本文的结论和不足之处，并对下一步的研究工作进行了展望。

2 复杂网络和分形的基本概念简述

2.1 复杂网络的定义及表示方式

自随机图理论提出至今，在复杂网络领域提出了许多概念和术语。这里介绍一些常用的定义。

网络 (Network) 在数学上以图 (Graph) 来表示，图的研究最早起源于18 世纪瑞士著名数学家Euler 的哥尼斯堡七桥问题。复杂网络可以用图论的语言和符号精确简洁地加以描述。图论不仅为数学家和物理学家提供了描述网络的语言和研究的平台，而且其结论和技巧已经被广泛地移植到复杂网络的研究中。为了能够方便准确地表达复杂网络的统计性质，有必要先介绍一些网络和图论的基本知识。

网络 (图) 是指二元组 (V, E) ，其中 V 为节点集， E 为边集， V 中元素称为节点或顶点 (Vertex 或 Node)， E 中元素称为边 (Edge 或 Link)，且 E 中的每条边 l_i 有 V 的一对节点 (u, v) 与之对应，如果 E 中任意的节点对 (u, v) 和 (v, u) 对应同一条边，则该网络称为无向网络，否则为有向网络；如果 E 中所有边的长度均为 1，即 $l_i = 1$ ，则称网络为无权网络，否则为加权网络。 V 中元素个数和 E 中元素个数分别称为网络的阶 (Order) 和边数 (Size)。阶和边数都有限的图为有限图 (Finite Graph)。边所连接的节点称为端点 (End-vertices)，两端点相同的边称为环 (Loop)。有公共起点并且有公共终点的两条边称为平行边 (Parallel Edges) 或重边 (Multi-edge)。无环且无平行边的图称为简单图 (Simple Graph)。任何不同的两节点之间都有边相连的简单无向图称为完全图 (Complete Graph)。本文提到的图，若没有特别的说明，均指没有重边的无向简单图。

2.2 复杂网络的基本属性

复杂网络之所以被广泛地研究，在于许多现实网络表现出相同或相似的特性，有一些共同的属性。下面就简单介绍一下这些属性。

小世界效应。20 世纪 60 年代，美国哈佛大学的心理学家 Milgram 做了一个社会调查，发现地球上任意两个人之间的平均距离是 6，即是平均只要通过 5 个人的传递，一个人就能同地球上任何一个角度的另一个人建立联系。这就是著名的小世界实验，又称为“六度分离”。它证明了现实许多网络大多数结点都可以通过一条较短的路径相连。一个网络如果称为是小世界效应的，是指如果对于固定的网络的节点平均度，平均路径长度的增加速度至多与网络规模 N 的对数成正比 (度和平均路径的概念将在 2.3 节详述)。小世界效应同时也表明，在现实网络中，信息的传播是十分迅速的过程。

无标度特性。所谓无标度特性，是指网络中存在着极少数连接度特别大的点，连接着网络中的绝大多数的连接度小的点，网络中的节点连接度分布符合幂率分布，即“无标度”。绝大多数的现实网络的这种特性直接决定了它同时具有鲁棒性和脆弱性的双重特性。

网络弹性。所谓网络弹性，主要是指网络节点的删除对网络连通性的影响。对于有些节点删除，网络依然能保持良好，而有些节点删除，则对网络是致命的。

模块性。所谓模块，是指一组物理上或功能上连接在一起的、共同完成一个相对独立功能的节点。现实网络中许多都包含模块，万维网上相似主题的网站、社会网络中一群朋友或不同阶层，以及相对固定的蛋白质-蛋白质联合体^[58]等等。这就让我们除了研究小世界和无标度这种共同的全局结构特征之外，也要重视网络的不同的局部结构特征。

自相似性。一个复杂系统的自相似性是指某种结构或过程的特征从不同的空间角度或时间尺度来看都是相似的，或者某系统或结构的局域性质或局域结构与整体类似。现实网络中，好多呈现出这种自相似性，如万维网、社会网络、蛋白质网等等^[46]。关于复杂网络的自相似性将在 2.4 节详述。

当然，复杂网络的特性还有很多，随着复杂网络的逐渐深入，必然会有更多的属性被人发现并应用起来。

2.3 复杂网络的特征度量

随着对复杂网络研究的深入，研究者提出了许多特征度量参数，来描述各种不同类型的复杂网络的共同特征，同时也用来衡量各种复杂网络演化模型的准确性和有效性。下面只介绍几个常用到的基本概念。

2.3.1 节点的度分布

度是单独节点的属性中简单而又重要的概念。节点 i 的度 k_i 指的是与该节点连接的其他节点的数目，也等价于该节点所有的邻居的数目。而度分布 $P(k)$ 定义为随机地选择一个节点，度为 k 的概率，或者等价地描述为网络中度为 k 的节点数占网络节点总数的比例。对于有向网络，由于一个节点的度又分为出度（从该节点指向其他节点的边的数目）和入度（从其他节点指向该节点的边的数目），其度分布又有入度分布和出度分布之分。

节点的度分布是网络的一个重要的统计特征。目前，主要有两种度分布：一是指数分布（又称 Poisson 分布）；另一种分布是幂律分布，即 $P(k) \propto k^{-\gamma}$ ， γ 是度指数。近几

年的大量实证研究表明,许多网络的度分布可以用幂律分布来更好地描述,并且度指数 γ 一般介于 2 到 3 之间。幂率分布曲线要比 Poisson 指数分布曲线下降要缓慢得多。当然,并不是所有的网络都服从幂率分布,如美国西部的电力网络服从指数分布;电影演员合作网络、蛋白质相互作用网络的度分布则是一个幂律加指数截断的度分布的网络 [59]。

以下是度分布的两个实例:

①www 子网,在双对数坐标系下度分布基本为直线形式,符合幂律分布,见图 2.1(a) [59];

②电力网络,在半对数坐标下度分布为直线形式,服从指数分布,见图 2.1(b) [7]。

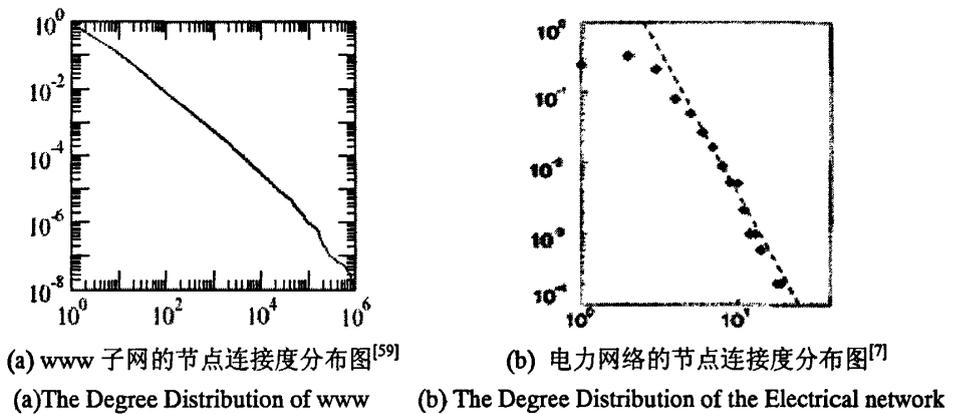


图 2.1 真实网络的节点连接度分布图

Fig 2.1 The Degree Distribution of real networks

2.3.2 集聚系数

集聚系数是用来描述网络中节点的聚集情况,是衡量网络集团化程度的重要参数。比如在社会网络中,你朋友的朋友可能也是你的朋友或者你的两个朋友可能彼此也是朋友。节点 i 集聚系数 C_i 定义为,它的 k_i 个直接邻居之间实际存在的边数 E_i 占所有可能存在的边数 $k_i(k_i - 1)/2$ 的比例,即 $C_i = 2E_i / [k_i(k_i - 1)]$ 。整个网络的集聚系数 C 指的是,所有节点集聚系数的算术平均值。很显然, $C \in [0,1]$, 当 $C = 0$ 时,所有的节点均为孤立点,当 $C = 1$ 时,网络中任意两个节点都是直接相连的。如图 2.2, 节点 1 的集聚系数为 $1/3$, 节点 2 的集聚系数为 $1/3$, 节点 4 的集聚系数为 1, 节点 3, 5 只有一个邻居节点,集聚系数可认为是 0 (也有些研究文献认为是 1), 则整个网络的集聚系数是:

$$C = \frac{1/3 + 1/3 + 1 + 0 + 0}{5} = \frac{1}{3}.$$

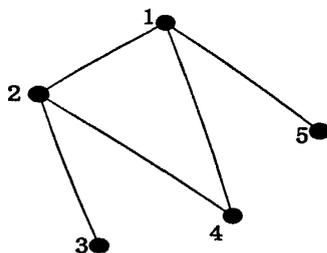


图2.2 集聚系数示意图

Fig 2.2 The Figure of Cluster Coefficient

许多现实网络都具有明显点的集聚效应，它们的集聚系数尽管远小于1但却比 $o(N^{-1})$ 要大得多。并且发现，你的朋友之间也是互为朋友的概率将随着网络的不断扩大，而趋向于一个常数，而这恰恰体现了“物以类聚，人以群分”这句话的道理。

2.3.3 平均路径长度

两个节点之间的路径长度是指从某一节点沿网络到达另一节点所经过的距离，又指从一个节点到达另一个节点所经过的边的数目。最短路径长度即是指两点之间的最短的路径长度，也就是经过的最少的边的数目，又称两点之间的距离。所有节点间的距离的最大值称为网络的直径，记为 $D = \text{Max}(d_{ij})$ 。网络的平均路径长度 L 定义为网络中任意一对节点之间的距离的平均值，数学表达式为：

$$L = \frac{1}{N(N-1)} \sum_{i,j \in G} d_{ij} \quad (2.1)$$

由上面的定义可知，图 2.2 的网络直径是 $D = d_{35} = 3$ ，平均路径长度 $L=1.6$ 。近期研究发现，大多数现实网络虽然节点数巨大，但网络的平均路径长度却小得惊人。如具有 449913 个节点的电影演员网络的平均路径长度为 3.48，具有 153727 个节点的万维网的平均路径长度为 3.1。

上述三个特征度量参数是描述复杂网络的研究基础，是网络的三个基本属性，研究者提出的大量的网络拓扑演化模型，也大多通过数学解析或者计算机模拟仿真的方法来获取这三个参数的值或表达式，从而进一步验证模型的有效性。随着复杂网络研究的不断深入，人们也不断地提出了其它一些特征度量参数，如网络的介数，最大连同分支，模块性，网络的度相关性，网络的簇度相关性等等。

2.4 分形与复杂网络的自相似性

2.4.1 分形的起源

自然界是宇宙万物的总称，是各种物质系统相互作用相互联系的总体，它包括大至宇宙天体的形成演化，小至微观世界中基本粒子的运动，呈现在人们面前是如此的千变万化、瑰丽多彩，又是广阔无垠、奥秘无穷。人类在认识自然改造自然的过程中，正在一层层的揭去其面纱，来探索其“庐山真面目”。应该说，物理学家们在解析宇宙和基本粒子方面花了极大的精力。随着牛顿经典力学的创立，爱因斯坦相对论，以及量子力学的发展，人类在自然科学方面已经取得了辉煌的成就；随着天体物理学以及其它相关学科的迅速发展，人类已经登上月球，进入太空；人类对微观世界由质点组成的简单系统的运动规律也有了全面而正确的认知。

尽管如此，只要人们稍微留心一下周围环境中发生的大量非线性不可逆现象，就会发现，人们对这些现象所知甚少，有许多问题甚至束手无策。就以天空中发生的大家习以为常的现象为例，当你仰望蔚蓝的天空，往往可以看到一团团白云漂浮其间，一派诗情画意，但如果用不同倍数的望远镜来观察云团时，就会发现，白云的形态似乎和望远镜的放大倍数无关，不管放大倍数多大，它的形态几乎总是保持不变。

再看一个与天气有关的例子，那就是气象预报。事实证明，长期的气象预报是不可能很准确的，因为随机性总是存在的，而它是无法事先预见的。另外对一个特定的地点而言，完全相同的天气(指气温、湿度、风速、风向、阳光、雨雪等)也是绝对不会重现的。

以上几个例子都是一些与天气相关的自然现象，其主要特点是不可逆性和随机性。除了气象之外，还有许许多多的非线性不可逆现象在科学研究和日常生活中存在，如流体力学中的湍流、对流、电子线路的电噪声、某些化学反应等，远离平衡的宏观体系中自发产生时空有序状态是十分普遍的自然现象和社会现象。自然界的各种变化过程中都包含着偶然性和必然性的统一。

1975年，曼德勃罗出版了他的法文专著《分形对象：形、机遇与维数》(Les objets fractals: forme, hazard et dimension)，标志着分形理论正式诞生。1977年他又出版了该书的英译本。1982年曼德勃罗的另一部历史性著作《大自然的分形几何学》与读者见面，该书虽然是前书的增补本，但在曼德勃罗看来却是分形理论的“宣言书”，而在分形迷的眼中，它无疑是一部“圣经”。该书旁征博引、图文并茂，从分形的角度考察了自然界中的诸多现象，引起了学术界的广泛注意，曼德勃罗也因此一举成名。

分形在英文中为 fractal, 是由美籍法国数学家曼德勃罗 (Benoit Mandelbrot) 创造出来的。此词源于拉丁文形容词 fractus, 对应的拉丁文动词是 frangere (破碎、产生无规则碎片)。此外, 它与英文的 fraction (碎片、分数) 及 fragment (碎片) 具有相同的词根。在 20 世纪 70 年代中期以前, 曼德勃罗一直使用英文 fractional 一词来表示它的分形思想。因此, 去拉丁词之头, 撮英文之尾所合成的 fractal, 本意是不规则的、破碎的、分数的。曼德勃罗是想用此词来描述自然界中传统欧氏几何学所不能描述的一大类复杂无规的几何对象, 例如, 蜿蜒曲折的海岸线、起伏不定的山脉, 粗糙不堪的断面, 变幻无常的浮云, 九曲回肠的河流, 纵横交错的血管, 令人眼花缭乱的满天繁星等。它们的特点是, 极不规则或极不光滑。直观而粗略地说, 这些对象都是分形。

目前, 分形是非线性科学中的一个前沿课题, 在不同的文献中, 分形被赋予不同的名称, 如“分数维集合”、“豪斯道夫测度集合”、“S 集合”、“非规整集合”、以及“具有精细结构集合”等等。一般地可把分形看作大小碎片聚集的状态, 是没有特征长度的图形和构造以及现象的总称。由于在许多学科中的迅速发展, 分形已成为一门描述自然界中许多不规则事物的规律性的学科。

长期以来, 自然科学工作者, 尤其是物理学家和数学家, 由于受欧几里得几何学及纯数学方法的影响, 习惯于对复杂的研究对象进行简化和抽象, 建立起各种理想模型(绝大多数是线性模型), 把问题纳入可以解决的范畴。对这种逻辑思维方法, 大家都是很熟悉的, 因为从中学到大学, 每个学生在课堂学习中, 已经多次反复地被灌输、熏陶, 已经习以为常。应该指出的是, 这种线性的近似处理方法也很有效, 在许多学科中得到了广泛的应用, 解决了许多理论问题和实际问题, 取得了丰硕的成果, 推动了各门学科的发展。但是在复杂的动力学系统中, 简单的线性近似方法不可能认识与非线性有关的特征, 如流体中的湍流、对流等等。虽然从数学上, 这种近似方法也可以对一些非线性系统列出微分方程(组)来加以定量描述, 但是除了极个别的例子可以在某一特定条件下, 求出其特解以外, 大多至今都解不出来。对于复杂一些的非线性系统和过程, 则连微分方程(组)也列不出来。而分形则是直接从非线性复杂系统的本身入手, 从未经简化和抽象的研究对象本身去认识其内在的规律性, 这一点就是分形理论与线性近似处理方法本质上的区别。

需要指出的是, 应用分形理论来研究非线性科学中的各种课题, 丝毫不贬低线性近似处理方法的重要性, 因为在一定的范围之内, 应用线性近似处理方法可以迅速地得到有效的结果。但是对远离平衡的非线性复杂系统(过程)来说, 就只能用分形理论来进行研究, 正如对低速运动的物体, 用牛顿三大定律来处理完全正确; 而对微观世界中粒子的高速运动, 就只能用量子力学和相对论来加以描述。

古希腊人创立的经典几何学一直是人们认识自然物体的有力工具。宇宙理论，宏伟的建筑都是建立在这个基础上的。以致伟大的科学家伽利略曾经断言：大自然的语言是数学，“它的标志是三角形、圆和其他几何图形”。然而自然界存在着大量无法用经典几何方法准确描述的复杂图形，如自然界中长着分岔的树枝、变幻的云彩、高低不平的山脉、弯曲的河流，生活中股市上的股票价格曲线、水文测量中的水位变化曲线、一个地区的气候变化曲线等等。从整体上看，它们的几何图形是处处不规则的，但在不同的尺度上图形的规则性又是相同的，也就是从整体到局部的各个层次上都有自相似的结构，在一个花样的内部还有更小的同样的花样^{[60][61]}。

曼德勃罗曾经为分形下过两个定义：

(1) 满足条件 $Dim(A) = \dim(A)$ 的集合 A ，称为分形集。其中， $Dim(A)$ 为集合 A 的分维数， $\dim(A)$ 为其拓扑维数。一般说来， $Dim(A)$ 不是整数，而是分数。

(2) 部分与整体以某种形式相似的形，称为分形。

然而，经过理论和应用的检验，人们发现这两个定义很难包括分形如此丰富的内容。实际上，对于什么是分形，到目前为止还不能给出一个确切的定义，正如生物学中对“生命”也没有严格明确的定义一样，人们通常是列出生命体的一系列特性来加以说明。分形具有五个基本特性：

①分形都具有任意小尺度下的比例细节，或者说它具有精细的结构；

②分形不能用传统的几何语言来描述，它既不是满足某些条件的点的轨迹，也不是某些简单方程的解集；

③分形具有某种自相似形式，可能是近似的自相似或者统计的自相似；

④一般，分形的“分形维数”，严格大于它相应的拓扑维数；

⑤在大多数令人感兴趣的情形下，分形由非常简单的方法定义，可能以变换的迭代产生。

2.4.2 分形的几何特征

分形作为几何对象，首先是破碎的、不规则的，但不是所有破碎的、不规则的形状都是分形。曼德勃罗（1986年）曾经给分形写过这样一个定义：组成部分与整体以某种方式是相似的形。也就是说，分形一般具有自相似性。但分形理论发展到今天，已经不仅限于研究对象的自相似性质了，如果一个对象的部分与整体具有自仿射变换关系，我们也可以称它为分形。今后，条件可能还会进一步拓宽，只要是部分与整体以某种规则

联系起来，通过某种变换使之对应，我们都可以将其看成是分形，因为分形的本质就是标度变换下的不变性，而这层意思是可以拓宽的。

1. 自相似性

自相似便是局部与整体的相似，或者说，局部是整体的缩影等。下面举几个典型的例子来使读者更好的理解对象的自相似性。

(1) Cantor 三分集

集合论创始人，德国数学家康托 (G. Cantor, 1845-1918 年) 在 1883 年曾构造了一种三分集，其几何表示如下：

取一条欧氏长度为 L_0 的直线段， L_0 叫做初始操作长度。将这条直线段三等分之后，保留两端的线段，将中间的一段扔掉，如图 2.3 所示 $N=1$ 的操作；再将剩下的两条直线段分别三等分，然后将其中间部分扔掉，如图 2.3 所示 $N=2$ 的操作，以此类推，直至无穷，便形成了无数个尘埃似的点，这便是 Cantor 三分集。它们的数目无穷多，但长度为零。这种构造的自相矛盾性质曾使 19 世纪的数学家感到困惑。但从几何关系来看，最终生成点的分布是局部相似的，甚至，这个过程中每一步图形之间也是局部相似的，这便是自相似。

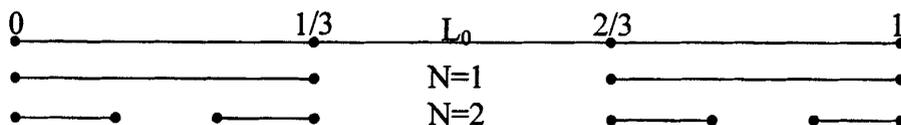


图 2.3 康托三分集

Fig 2.3 Cantor curve

(2) Koch 曲线

1904 年，瑞典数学家科赫 (H. von Koch, 1870-1924 年) 构造了一种“妖魔曲线”，被称为 Koch 曲线。其构造过程如下：

取一条欧氏长度为 L_0 的直线段，将其三等分，保留两端的线段，将中间的一段改换成夹角为 60° 的两个等长的直线，如图 2.4 中 $N=1$ 的操作。将长度为 $L_0/3$ 的 4 个直线段分别进行三等分，并将它们中间的一段均改换成夹角为 60° 的两段长为 $L_0/9$ 的直线段，得到图 2.4 中 $N=2$ 的操作。重复上述操作直至无穷，便得到一条具有自相似结构的折线，如右图便是 Koch 雪花曲线。

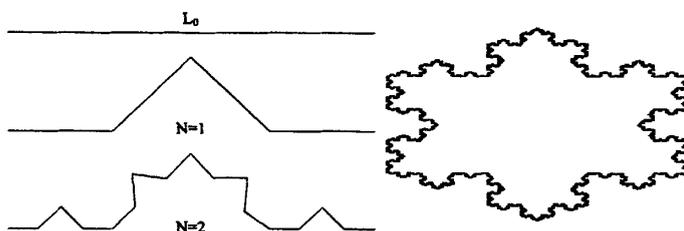


图 2.4 科赫曲线

Fig 2.4 The Koch curve

(3) Sierpinski 垫片

以上两个自相似图形都是基于一条欧氏直线段生成的, Cantor 三分集是将线段删去一部分, 最终得到的是一个离散点集, 而 Koch 曲线是将线段增加一部分, 最终得到的是一个处处不光滑的折线集。波兰数学家谢尔宾斯基 (W. Sierpinski, 1882-1969 年) 于 1915 年给出了一个从平面上的二维图形出发做曲线的有趣例子。所构造的 Sierpinski 垫片相当于将上述构造方法推广到平面上, 起初始图形是一个等边三角形面, 构造过程如下:

首先, 我们将这个等边三角形面四等分, 得到 4 个小等边三角形面, 去掉中间一个。将剩下的 3 个小等边三角形面分别进行四等分, 再分别去掉中间的一个。重复以上操作直至无穷, 可以得到如图 2.5 所示的图形, 可以看出它的每一小部分在结构上都与整体相同, 这也是一个典型的自相似图形。

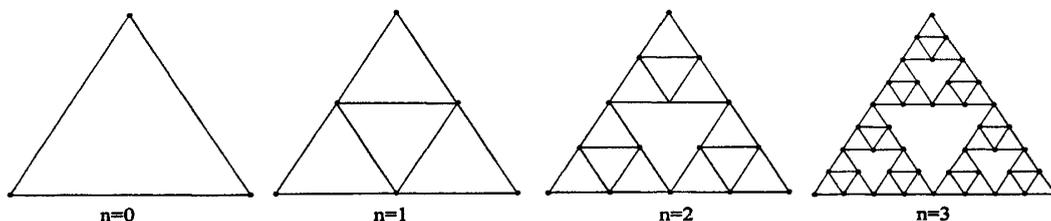


图 2.5 谢尔宾斯基垫片

Fig 2.5 Sierpinski gasket

2. 自仿射性

自仿射性是自相似性的一种拓展。如果将自相似性看成是局部到整体在各个方向上的等比例变换的结果的话。那么, 自仿射性就是局部到整体在不同方向上的不等比例变换的结果。前者称为自相似变换, 后者称为自仿射变换。

3. 精细结构

分形还有一个更重要的特征, 即精细结构。在理论上, Koch 曲线是按一定规则无限变换的结果, 所以, 假如用一个数学放大镜来看 Koch 曲线的话, 无论放大多少倍, 都能看到里面还有与整体相似的结构。这一点非常接近于自然界中的对象, 也符合中国古代的哲学思想: “一尺之棰, 日取其半, 万世不竭” (庄子《天下篇》)。在这里, 我们打算讨论物质是否无限可分, 我们只是注意到分形和自然对象都具有极多层次的结构, 这是分形体最基本的特征。但是, 自然界中的对象与数学中的分形还是有所不同的。

在自然界中，对象即使存在自相似性，也是在有限层次区间中，所以在对自然界的分形研究中，一般我们只关心有限层次的自相似性或自仿射性结构。

4. 分形与欧氏几何图形的区别

可以看出，分形与普通的欧氏几何图形有明显的区别：

①欧氏图形是规则的，而分形是不规则的，也就是说，欧氏图形一般是逐段光滑的，而分形往往在任何区间内都不具有光滑性。

②欧氏图形层次是有限的，而分形从数学角度上讲，层次是无限的。

③欧氏图形一般不会从局部的到整体的信息，因为它们不强调局部与整体的关系，而分形强调这种关系，所以，分形往往可以从局部“看出”整体。

④欧氏图形越复杂，其背后的规则也必定越复杂，而对于分形图形，虽然看上去十分复杂，但其背后的规则却是相当简单的。

因此，我们要构造、绘制的图形是与中学学到的欧氏图形完全不同的，首先必须找出分形对象“不规则”的规则，然后在计算机的帮助下才能将其生成出来。

分形理论真正发展起来才一、二十年，并且方兴未艾，目前仍处于不断发展之中。分形，为探讨自然界的复杂事物的客观规律及其内在联系提供了新的概念和方法，在其发展过程中，许多传统的科学难题，由于分形的引入而取得显著进展，分形理论的应用发展远远超过了理论的发展。分形作为一种新的概念和方法，正在许多领域开展应用探索。美国著名物理学家惠勒说过：今后谁不熟悉分形，谁就不能被称为科学上的文化人。

2.4.3 复杂网络的自相似性

现实中很多网络的部分与整体有着很明显的相似性；而局部在某种意义上与整体相似，即自相似性。

衡量复杂网络的小世界特征一个重要特征就是网络平均路径长度 L 可用网络模型 N 的对数函数来表示，即 $L \propto \ln N$ ，或表示为

$$N \propto e^{\frac{L}{L_0}} \quad (2.2)$$

其中常数 L_0 表示特征长度。式(2.2)意味着网络规模是平均路径长度的指数函数，而自相似性则要求满足某种幂律关系。

Song 等人^[46]创造性地将分形理论中的盒维数的方法推广用于复杂网络，计算出复杂网络中的分数维，从而指出，许多实际网络，包括 WWW、社会网络、蛋白质交互作用网络(PIN)和细胞网络等，在某种长度-标度下确实是自相似的。盒维数方法用于复杂网络的主要困难是：对于大多数显示网络来说，并不存在包含这些网络的自然的欧氏空

间，而且复杂网络上两个节点之间的距离并不是指两个节点的欧氏距离，而是连接两个节点的最短路径所包含的边数。他们将用于覆盖复杂网络的尺寸为 l_B 的盒子规定为：盒子中任意两个节点之间的距离都小于 l_B 。用这些盒子覆盖网络，将网络中的节点不重复地放在了一个个盒子里。Song 等人通过计算许多现实网络发现，覆盖网络所需的最少盒子数 N_B 与盒子尺寸 l_B 之间的关系，是服从分形所遵循的幂律标度公式，即：

$$N_B(l_B) \approx l_B^{-d_B} \tag{2.3}$$

其中 d_B 是分形维数，从而可以证明了复杂网络的自相似性。

Song 等人在此基础上又做了进一步的研究。他们把所有节点都分配到盒子中之后，再把每个盒子用单个节点来表示，并称为重整化节点(renormalized node)，重整化节点是否相连取决于之前两个盒子之间是否有边相连。这样就得到了一个新的重整化网络，这种重整化过程可以一直进行下去，直到整个网络被归约为单个节点。他们通过这种重整化过程，发现自相似和无标度的度分布在网络的所有粗粒化阶段都成立。图 2.6 表示出这一重整化的过程。并且进一步发现，在一次一次的重整化中，网络的无标度分布具有标度不变性，这也似乎证明了分形理论中局部反应整体，局部与整体相似的特征。

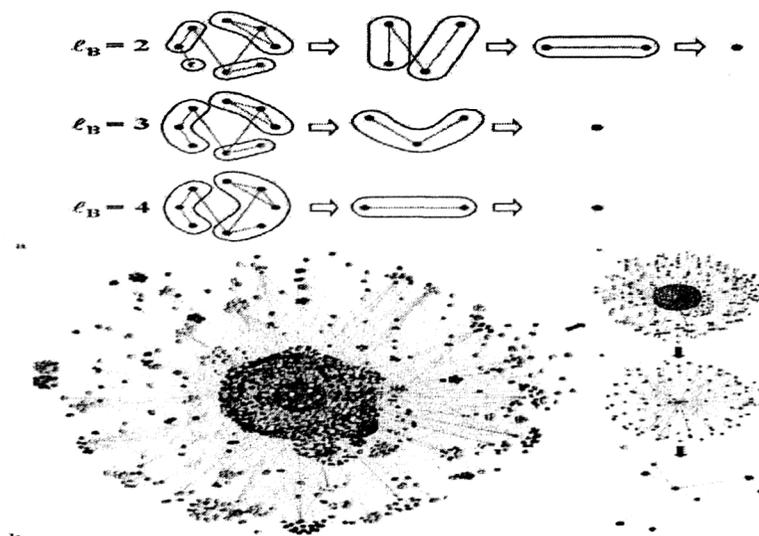


图 2.6 复杂网络的重整化过程

Fig 2.6 The renormalization of the complexwork

图 2.6(a)显示了一个包含 8 个节点的网络在不同 l_B 情形下的重整化。图 2.6(b)显示了 WWW 的重整化过程的各个阶段，固定盒子尺寸 $l_B = 3$ 。Song 等人这一开创举的工作，将复杂网络的研究又提供了新的视角和方法。许多研究者开始沿着这条路开始研究复杂

网络这种自相似或者说分形特征的原因是什么。Song 等人在随后的研究中，将这种自相似特性归因于连接度高的节点（又称集散节点）的互斥性造成的。Gim 和 Goh 等人则从一种称为骨架（skeleton）的树结构出发，认为最初的分形网络可以看作是这种分形骨架和一些点与点之间的捷径组成。而盒维数的方法成为复杂网络演化模型又一工具，用来验证了模型演化网络的自相似性或者分形特征。目前，关于复杂网络的分形性（自相似性）的形成原因的研究还不是太多，更多是还是在提出模型来验证这种分形性（自相似性）。

3 基于谢尔宾斯基支架的具有分形和小世界特性的网络

本章针对谢尔宾斯基支架的演化机制及网络拓扑性质，构造出高维的具有分形和小世界特性的确定性网络，利用数学归纳的方法计算出了该网络的集聚系数为 0.53、网络的直径为 4、平均路径长度为 2.5，平均度为 7.5，证明了该网络的小世界特性；然后测定该网络的盒维数和相似维数都约为 2，由网络的自相似性及其具有的精细结构证明了该网络的分形特性。

3.1 谢尔宾斯基支架

3.1.1 谢尔宾斯基支架的构造

谢尔宾斯基支架的构造方法如下^[66]：

- a. 初始一个四面体 ST_0 ；
- b. 用四个四面体 ST_0 组成一个边长是 ST_0 边长两倍四面体（不包括中间的小八面体）；
- c. 按照上面的方法进行下去，可知 ST_n 是用四个四面体 ST_{n-1} 组成的一个边长是 ST_{n-1} 边长两倍的四面体（不包括中间的小八面体），具体的过程见下面的图 3.1：

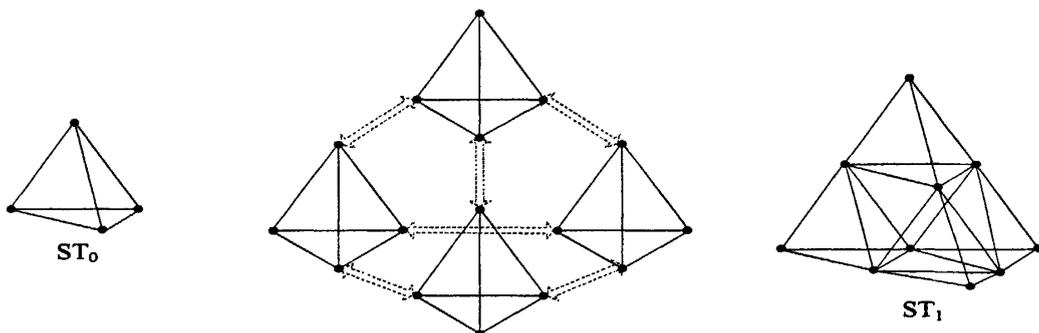


图 3.1 谢尔宾斯基支架的迭代过程示意图

Fig 3.1 The recursive constructions of ST_n

我们之所以称之为谢尔宾斯基支架，是因为 tetra 在英文中是指一种叫锯脂鲤的鱼，这种鱼生活在亚马逊河、奥利诺科河流域，以凶猛而出名。平时以吃小鱼为生，饥饿时，陆上动物如牛、马和人等若坠落水中，会群集袭击，只数分钟便吃得只剩骨头。它那有劲的颌肌肉，能撕裂猎物的皮并咬断肉。而我们这里介绍的谢尔宾斯基支架就像锯脂鲤的颌骨那样结实、稳健。将四面体的顶点看作是网络中的点，四面体的边可看作网络中

的结点之间存在着的连接，所以谢尔宾斯基支架也可以被看作是一个较复杂的网络图。为方便讨论，我们简记第 n 代谢尔宾斯基支架(Sierpinski tetra)为 ST_n 。

3.1.2 谢尔宾斯基支架的性质

引理1 第 n 代谢尔宾斯基支架 ST_n ($n \geq 0$)，具有以下特性^[66]：

- a. ST_n 具有的结点数为 $V_n = 2(4^n + 1)$ ，边数为 $E_n = 6 \cdot 4^n$ 。
- b. ST_n 的直径为 $D_n = 2^n$ 。
- c. ST_n 的集聚系数均为 $C_n = \frac{5 \cdot 4^n + 16}{10(4^n + 1)}$ 。

证明：根据 ST_n 的构造过程，当 $n \geq 1$ 时，我们有

$$\begin{aligned} V_n &= 4V_{n-1} - 6, & V_0 &= 4; \\ E_n &= 4E_{n-1}, & E_0 &= 6. \end{aligned}$$

解得： $V_n = 2(4^n + 1)$ 及 $E_n = 6 \cdot 4^n$ ；

同理，由 $D_n = 2D_{n-1}$ ，且 $D_0 = 1$ ，可得： $D_n = 2^n$ ；

ST_n 中结点度数为3的只有 ST_n 的四个顶点，且它们具有集聚系数为1；当 $n \geq 1$ 时，除了上述四个结点外，剩余结点的集聚系数或是 $8/15$ 或是 $2/5$ 。设其中集聚系数为 $8/15$ 的结点的个数为 x_n ，集聚系数为 $2/5$ 的结点的个数为 y_n ，则有下面递推公式成立：

$$\begin{aligned} x_n &= 4x_{n-1}, & x_1 &= 6; \\ y_n &= 4y_{n-1} + 6, & y_1 &= 0. \end{aligned}$$

解得： $x_n = 6 \cdot 4^{n-1}$ ， $y_n = 2(4^{n-1} - 1)$ ，所以 ST_n 的集聚系数为

$$C_n = \frac{4 \times 1 + 2(4^{n-1} - 1) \times \frac{2}{5} + 6 \cdot 4^{n-1} \times \frac{8}{15}}{2(4^n + 1)} = \frac{5 \cdot 4^n + 16}{10(4^n + 1)}.$$

3.2 基于谢尔宾斯基支架构建的网络

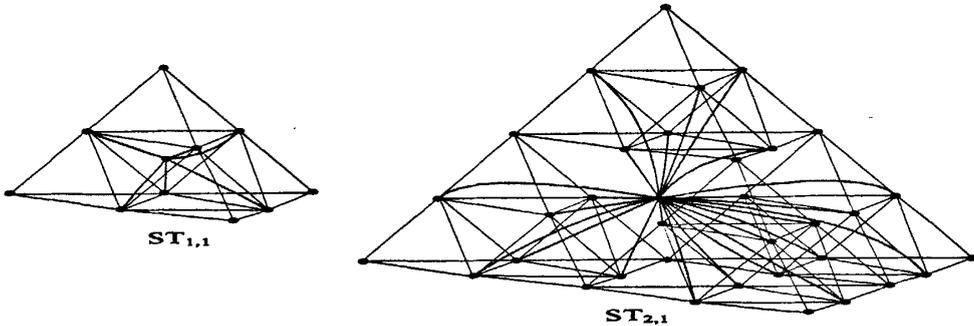


图 3.2 定义在谢尔宾斯基支架构建上的复杂网络

Fig 3.2 The networks based on Sierpinski tetra

定义在谢尔宾斯基支架构建上的复杂网络如下：定义 $ST_{n,1}$ 为在第 n 代谢尔宾斯基支架 ST_n 中加入一中心结点，然后将此中心结点与所有第一代谢尔宾斯基支架的六条边的中点相连，形成过程如图3.2所示。

引理2 由第 n 代谢尔宾斯基支架构成的复杂网络 $ST_{n,1}$ ($n \geq 1$)，具有以下特性：

- a. $ST_{n,1}$ 具有的结点数为 $V_{n,1} = 2 \cdot 4^n + 3$ ，边数为 $E_{n,1} = 6 \cdot 4^n + 6 \cdot 4^{n-1}$ 。
- b. $ST_{n,1}$ 的直径为 $D_{n,1} = 4$ 。

3.3 基于谢尔宾斯基支架构建的网络的小世界特性

3.3.1 网络的集聚系数

集聚系数是用来描述网络中结点的聚集情况，即网络有多紧密，比如在社会网络中，你朋友的朋友可能也是你的朋友或者你的两个朋友可能彼此也是朋友。结点 i 集聚系数 C_i 定义为：它的 k_i 个直接邻居之间实际存在的边数 E_i 占所有可能存在的边数 $k_i(k_i - 1)/2$ 的比例，即 $C_i = 2E_i / [k_i(k_i - 1)]$ 。整个网络的集聚系数指的是，所有结点集聚系数的算术平均值。

引理3 由第 n 代谢尔宾斯基支架构成的网络 $ST_{n,1}$ ($n \geq 1$)，具有集聚系数为

$$C_{n,1} = \frac{\frac{37}{35} \cdot 4^n + \frac{16}{5} + \frac{4}{6 \cdot 4^{n-1}}}{2 \cdot 4^n + 3}。$$

证明：根据引理2及网络的构造我们知道， $ST_{n,1}$ 具有的结点数为 $V_{n,1} = 2 \cdot 4^n + 3$ ，边数为 $E_{n,1} = 6 \cdot 4^n + 6 \cdot 4^{n-1}$ ；对于 $ST_{n,1}$ 而言，有下列四类结点：

- (1) $ST_{n,1}$ 的四个顶点，它们的集聚系数为1；
- (2) 新加入的中心结点，其连接度为 $6 \cdot 4^{n-1}$ ，所连结的结点的之间存在的边数也为 $12 \cdot 4^{n-1}$ ，所以该结点的集聚系数是 $4/(6 \cdot 4^{n-1} - 1)$ ；
- (3) 和中心结点相连的结点，共有 $6 \cdot 4^{n-1}$ 个结点，它们的连接度为7，每个结点的邻居间存在的边数为12，所以此类结点的集聚系数是 $4/7$ ；
- (4) 剩下的结点还有 $2(4^{n-1} - 1)$ 个，其连接度为6，此类结点邻居之间有6条边，所以该类结点的集聚系数是 $2/5$ 。

所以对于 $ST_{n,1}$ ，其集聚系数为

$$C_{n,1} = \frac{4 \times 1 + 1 \times \frac{4}{6 \cdot 4^{n-1} - 1} + 6 \cdot 4^{n-1} \times \frac{4}{7} + 2 \cdot (4^{n-1} - 1) \times \frac{2}{5}}{2 \cdot 4^n + 3}$$

$$= \frac{\frac{37}{35} \cdot 4^n + \frac{16}{5} + \frac{4}{6 \cdot 4^{n-1}}}{2 \cdot 4^n + 3}。$$

由上式可得，当 $n \rightarrow \infty$ 时， $\lim_{n \rightarrow \infty} C_{n,1} = 37/70 \approx 0.53$ 。

3.3.2 网络的平均路径长度

所谓网络的平均路径长度是指网络中所有结点对之间的平均最短距离，它描述了网络结点之间的分离程度。这里结点间的距离指的是从一结点到另一结点所要经历的边的最小数目。网络的直径(Diameter)为任意两点间的最大距离。平均路径长度的计算公式为

$$APL = \frac{1}{N(N-1)} \sum_{i,j} d_{ij}$$

其中 d_{ij} 为结点 i 和 j 之间的最短距离， N 为网络的规模。

由谢尔宾斯基支架构造的网络的定义我们得到 $ST_{n,1}$ 的直径为 $D_{n,1} = 4$ 。可知该网络中结点之间的距离有四种情况，分别为1, 2, 3, 4。所以对于引理4中所论述的四类结点，这里我们将 $ST_{n,1}$ ($n \geq 2$)网络中结点对之间的距离列于下表3.1。

表 3.1 结点之间的距离表

Table 3.1 The distance between vertexes

结 点 对 个 数	1	2	3	4
-----------------------	---	---	---	---

连结度为 3 (共 4 个结点)	3	7	$6 \cdot 4^{n-1} - 6$	$2 \cdot 4^{n-1} - 2$
连结度为 $6 \cdot 4^{n-1}$ (即中心结点)	$6 \cdot 4^{n-1}$	$2 \cdot 4^{n-1} + 2$	0	0
连结度为 7 (共 $6 \cdot 4^{n-1}$ 个结点)	7	$6 \cdot 4^{n-1} - 3$	$2 \cdot 4^{n-1} - 2$	0
连结度为 6 (共 $2 \cdot 4^{n-1} - 2$ 个结点)	6	13	$6 \cdot 4^{n-1} - 12$	$2 \cdot 4^{n-1} - 5$

从表 3.1 可得：结点连结度为 3 的距离总和是：

$$4 \times [1 \times 3 + 2 \times 7 + 3 \times (6 \cdot 4^{n-1} - 6) + 4 \times (2 \cdot 4^{n-1} - 2)] = 26 \cdot 4^n - 36。$$

同理可得到，结点连结度为 $6 \cdot 4^{n-1}$ 的距离总和是 $10 \cdot 4^{n-1} + 4$ ；结点连结度为 7 的距离总和是 $27 \cdot 4^{2n-1} - 30 \cdot 4^{n-1}$ ；结点连结度为 6 的距离总和是 $13 \cdot 4^{2n-1} - 25 \cdot 4^n + 48$ 。由 $N = 2 \cdot 4^n + 3$ ，根据网络的平均路径长度的计算公式可得， $ST_{n,1}$ 的平均路径长度为：

$$APL = \frac{1}{N(N-1)} \sum_{i,j} d_{ij} = \frac{2.5N^2 - 17N + 44.5}{N^2 - N}，$$

由上式易得当 N 较大时有， $\lim_{N \rightarrow \infty} APL = \lim_{N \rightarrow \infty} \frac{2.5N^2 - 17N + 44.5}{N^2 - N} = 2.5$ 。

3.3.3 网络的平均度

网络中结点 i 的度 k_i 为结点 i 连接的边的总数目，所有结点 i 的度 k_i 的平均值称为网络的平均度，记为 $\langle k \rangle$ 。在我们构造的复杂网络中结点的度分布如下：度为 $6 \cdot 4^{n-1}$ 的结点，即中心结点，只有一个；度为 3 的结点，只有四个顶点；度为 6 的结点，共有 $2 \cdot 4^{n-1} - 2$ 个；度为 7 的结点，共有 $6 \cdot 4^{n-1}$ 个。

所以网络的平均度为

$$\langle k_n \rangle = \frac{1 \times 6 \cdot 4^{n-1} + 3 \times 4 + 6 \times (2 \cdot 4^{n-1} - 2) + 7 \times (6 \cdot 4^{n-1})}{2 \cdot 4^n + 3} = \frac{15 \cdot 4^n}{2 \cdot 4^n + 3}，$$

即，当 $n \rightarrow \infty$ 时， $\lim_{n \rightarrow \infty} \langle k_n \rangle = 15/2 = 7.5$ 。

3.4 基于谢尔宾斯基支架构建的网络的分形性

引理4 第 n 代谢尔宾斯基支架 ST_n 的盒维数和相似维数相等，都为 2，即 $d_B = D_S = 2$ 。

证明：根据文[46]用盒维数的方法来计算网络图的维数，用长度为 l_B 的盒子来覆盖网络图，节点对之间的距离不大于 $l_B - 1$ ，将有可能在一个盒子里，得到所需的盒子数为 N_B ，对于 ST_n ，示意图见图 3.3，

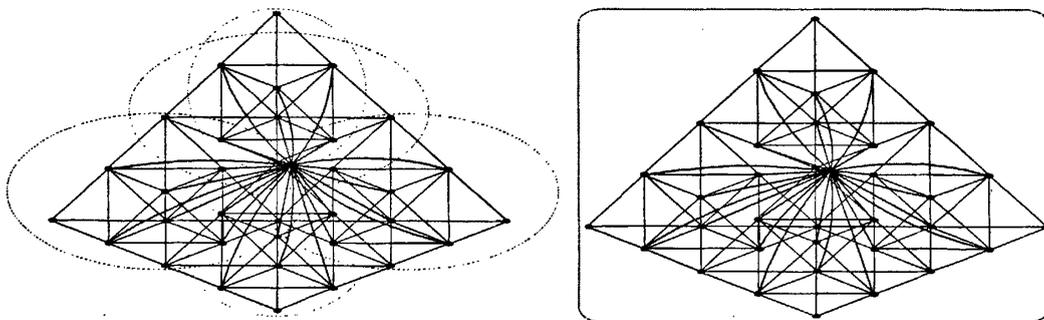


图 3.3 对 $n=2$ 时，即 $ST_{2,1}$ 时，分别用 $l_B=2$ 和 $l_B=3$ 来覆盖网络图的示意图

Fig 3.3 Boxes in $ST_{2,1}$ with $l_B=2$ and $l_B=3$

具体结果如下表 3.2。

表 3.2 盒维数计算结果

Table 3.2 The box-counting results

l_B	2	3	5	$2^n + 1$
N_B	4^n	4^{n-1}	4^{n-2}	1

所以可得该网络图的盒维数 $d_B = \frac{\ln 4}{\ln 2} = 2$ 。

现在我们采用相似维数^[67]的方法来计算网络图的维数，显然新加入的结点不影响网络的维数。如图 3.1 把四面体 ST_n 的边长放大 $L = 2$ 倍，则得到与 ST_{n+1} 相似的四面体，由于 ST_{n+1} 的面积是 ST_n 的 $K = 4$ 倍。所以，利用相似维数的计算公式：

$$D_s = \frac{\ln K}{\ln L} = \frac{\ln 4}{\ln 2} = 2,$$

所以 $d_B = D_s = 2$ 。

3.5 网络图的拓展讨论

因为 $ST_{n,1}$ 是由 ST_n 加入一个中心节点与所有第一代谢尔宾斯基支架的六条边的中点相连而得到的，我们自然联想到，考虑加入的这个中心节点与第 i 代谢尔宾斯基支架的六条边的中点相连而形成的网络图的特性，我们不妨记为 $ST_{n,i}$ ，其中 $n \geq i \geq 2$ 。具体可见图 3.4。

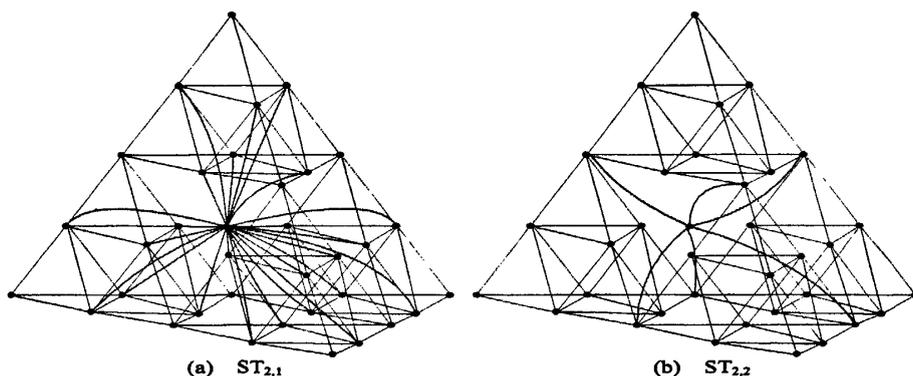


图 3.4 这是 $n = 2$ 时的连接示意图。(a)是连接第 1 代谢尔宾斯基支架的六条边的中点, 形成 $ST_{2,1}$;
(b)是连接第 2 代谢尔宾斯基支架的六条边的中点, 形成 $ST_{2,2}$

Fig 3.4 The figure of the network when $n = 2$. (a) $ST_{2,1}$; (b) $ST_{2,2}$

分析发现, $ST_{n,i}$ 的网络拓扑特性存在一定的规律性, 并且和 $ST_{n,1}$ 存在一定的区别。这里从集聚系数、网络的直径和分形维数三方面进行考察。

1、集聚系数。对于 $ST_{n,i}$ ($i \geq 2$), (1) $ST_{n,i}$ 的四个顶点, 它们的集聚系数为 1;
(2) 新加入的中心结点, 其连接度为 $6 \cdot 4^{n-i}$, 所连结的结点的之间存在的边数也为 $12 \cdot 4^{n-i}$, 所以该结点的集聚系数是 $4/(6 \cdot 4^{n-i} - 1)$; (3) 和中心结点相连的结点, 共有 $6 \cdot 4^{n-i}$ 个结点, 它们的连接度为 7, 每个结点的邻居间存在的边数为 12, 所以此类结点的集聚系数是 $4/7$; (4) 剩下的结点还有 $2 \cdot 4^n - 6 \cdot 4^{n-i} - 2$ 个, 其连接度为 6, 此类结点邻居之间有 6 条边, 所以该类结点的集聚系数是 $2/5$ 。

所以对于 $ST_{n,1}$, 其集聚系数为

$$C_{n,i} = \frac{4 \times 1 + 1 \times \frac{4}{6 \cdot 4^{n-i} - 1} + 6 \cdot 4^{n-i} \times \frac{4}{7} + (2 \cdot 4^n - 6 \cdot 4^{n-i} - 2) \times \frac{2}{5}}{2 \cdot 4^n + 3}$$

$$= \frac{\frac{1}{5} \cdot 4^{n+1} + \frac{9}{35} \cdot 4^{n-i+1} + \frac{16}{5} + \frac{4}{6 \cdot 4^{n-i} - 1}}{2 \cdot 4^n + 3}。$$

由上式可得, 当 $n \rightarrow \infty$ 时, $\lim_{n \rightarrow \infty} C_{n,i} = \frac{2}{5} + \frac{9}{70 \cdot 4^{i-1}}$ 。

2、网络的直径。很显然，随着 i 的增大，网络的直径必将随着增大。用数学归纳的方法不难求解，对于 $ST_{n,i}$ ，其直径为 $D_{n,i} = 2^i + 2$ 。可以得出当 $i \leq \log_2^n$ 时， $D_{n,i} \leq n + 2 = o(\log_2^n)$ ，其中 N 为网络的规模，即点的个数为 $2 \cdot 4^n + 3$ 。

3、网络的维数。由于只是增加了一个节点，所以其维数和 $ST_{n,i}$ 一样，仍为 $D_s \approx 2$ 。

从上面3个属性可以看出，拓展的网络图也是同时具有小世界和分形特性的。从而 $ST_{n,i}$ 和 $ST_{n,i}$ 是一类具有分形和小世界特性的网络图，并且 $ST_{n,i}$ 的分形和小世界特性随着代数 i 的增加而逐渐减弱。

3.6 小结

经典的谢尔宾斯基支架网络并不具备小世界特性，其集聚系数较小、网络直径指数级增长。而本章以谢尔宾斯基支架为基础，构造出了高维的具有分形和小世界特性的网络，通过数学归纳方法推导出该网络具有较大的集聚系数 0.53 和较短的平均路径长度 2.5，证明了其小世界特性，并用分形理论中相似维数和盒维数两种方法计算出网络的分维数约为 2，由网络的自相似性及其具有的精细结构证明了该网络的分形特性。

本章构造的具有小世界特性的网络中，新加入的中心结点就像是网络中的路由器，在不同的网路中，起到大量缩短网络距离的作用，使着无限扩大的网络有着相互联系；就像 Milgram 的小世界实验，虽然茫茫人海中人与人之间的距离表面上看很远，但是往往因为个别亲朋好友的存在使得表面上很远的关系并不存在。该网络图对于同时存在分形性和小世界效应的实际网络的研究具有一定的参考价值。另外，该网络支架的演化还为新衍架结构的设计及其系统稳定性、可靠性提供了新的启示。

4 谢尔宾斯基地毯的分形和小世界特性

本章将引入谢尔宾斯基地毯的网络拓扑性质，并在此基础上构造具有分形和小世界特性的网络图，利用数学归纳的方法得到了该网络图的集聚系数为 0.054，网络图的直径为 6，平均路径长度小于 6，平均度分布为 4.61，证明了该网络的小世界特性；盒维数和相似维数都约为 1.8928，由网络的自相似性及其具有的精细结构得到该网络的分形特性，由此证明了其分形和小世界特性。

4.1 谢尔宾斯基地毯

4.1.1 谢尔宾斯基地毯的构造

谢尔宾斯基地毯的构造方法如下^[66]：

- a. 初始一个正方形 SC_0 ；
- b. 用八个正方形 SC_0 组成一个边长是 SC_0 边长三倍的正方形（不包括中间的小正方形）；
- c. 按照上面的方法进行下去，可知 SC_n 是用八个正方形 SC_{n-1} 组成的一个边长是 SC_{n-1} 边长三倍的正方形（不包括中间的小正方形），具体的过程见下面的图4.1：

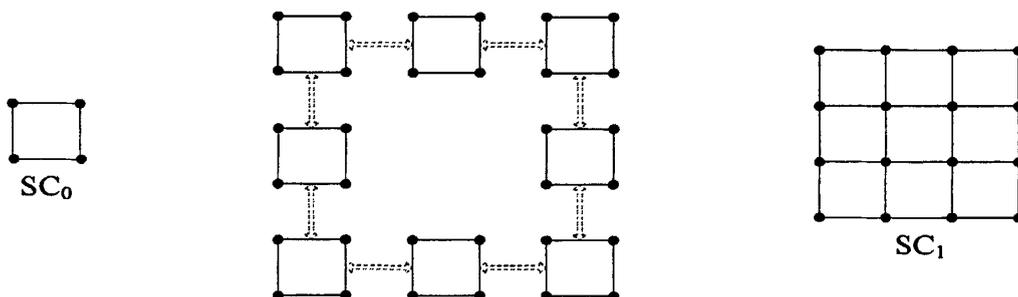


图 4.1 谢尔宾斯基地毯的迭代过程示意图

Fig 4.1 The recursive constructions of SC_n

将正方形的顶点看作是网络中的点，正方形的边可看作网络中的结点之间存在着的连接，所以谢尔宾斯基地毯也可以被看作是一个较复杂的网络图。为方便讨论，我们简记第 n 代谢尔宾斯基地毯(Sierpinski carpet)为 SC_n 。

4.1.2 谢尔宾斯基地毯的性质

引理5 第 n 代谢尔宾斯基地毯 SC_n ($n \geq 0$)，具有以下特性^[66]：

- a. SC_n 具有的结点数为 $V_n = \frac{11}{70} \cdot 8^{n+1} + \frac{8}{5} \cdot 3^n + \frac{8}{7}$, 边数为 $E_n = \frac{3}{10} \cdot 8^{n+1} + \frac{8}{5} \cdot 3^n$ 。
- b. SC_n 的直径为 $D_n = 2 \cdot 3^n$ 。
- c. SC_n 的集聚系数均为零。

证明:

设 SC_n 的边长为 L_n , V_n 、 E_n 分别代表 SC_n 的结点数和边数, 则由 SC_n 的构造过程我们得到如下递推公式:

$$L_n = 3 \cdot L_{n-1}, \quad L_0 = 1;$$

$$V_n = 8(V_{n-1} - L_{n-1} - 1) \quad V_0 = 4;$$

$$E_n = 8(E_{n-1} - L_{n-1}) \quad E_0 = 4。$$

解上面三个递推公式我们得到 $L_n = 3^n$, $V_n = \frac{11}{70} \cdot 8^{n+1} + \frac{8}{5} \cdot 3^n + \frac{8}{7}$ 和 $E_n = \frac{3}{10} \cdot 8^{n+1} + \frac{8}{5} \cdot 3^n$;

显然, SC_n 的直径 D_n 为边长 L_n 的两倍, 故有 $D_n = 2 \cdot 3^n$;

最后, 由于在 SC_n 中没有任何三角形, 所以 SC_n 中任何点的集聚系数均为零, 故 SC_n 的集聚系数均为零。

4.2 基于谢尔宾斯基地毯构建的网络

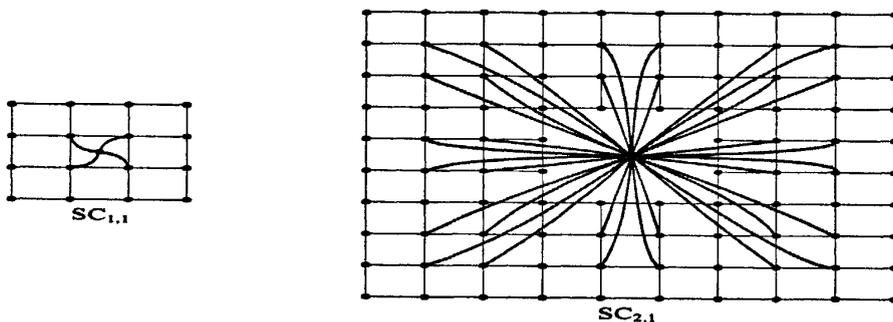


图4.2 定义在谢尔宾斯基地毯基础上的复杂网络

Fig 4.2 The networks based on Sierpinski carpet

定义在谢尔宾斯基地毯基础上的复杂网络如下: 定义 $SC_{n,1}$ 为在第 n 代谢尔宾斯基地毯 SC_n 中加入一中心结点, 然后将此中心结点与所有第一代谢尔宾斯基地毯的四个内点相连, 形成过程如图 4.2 所示。

引理6 由第 n 代谢尔宾斯基地毯构成的复杂网络 $SC_{n,1}$ ($n \geq 1$), 具有以下特性:

- a. $SC_{n,1}$ 具有的结点数为 $V_{n,1} = \frac{11}{70} \cdot 8^{n+1} + \frac{8}{5} \cdot 3^n + \frac{15}{7}$, 边数为 $E_{n,1} = \frac{3}{10} \cdot 8^{n+1} + \frac{8}{5} \cdot 3^n + 4 \cdot 8^{n-1}$ 。

b. $SC_{n,1}$ 的直径为 $D_{n,1} = 6$

4.3 基于谢尔宾斯基地毯构建的网络的小世界特性

4.3.1 网络的集聚系数

集聚系数是用来描述网络中结点的聚集情况，即网络有多紧密，比如在社会网络中，你朋友的朋友可能也是你的朋友或者你的两个朋友可能彼此也是朋友。结点 i 集聚系数 C_i 定义为：它的 k_i 个直接邻居之间实际存在的边数 E_i 占所有可能存在的边数 $k_i(k_i - 1)/2$ 的比例，即 $C_i = 2E_i / [k_i(k_i - 1)]$ 。整个网络的集聚系数指的是，所有结点集聚系数的算术平均值。

引理7 由第 n 代谢尔宾斯基地毯构成的复杂网络 $SC_{n,1}$ ($n \geq 1$)，具有的集聚系数为

$$C_{n,1} = \frac{\frac{2}{4 \cdot 8^{n-1} - 1} + \frac{4}{5} \cdot 8^{n-1}}{\frac{11}{70} \cdot 8^{n+1} + \frac{8}{5} \cdot 3^n + \frac{15}{7}}。$$

证明：根据引理6及网络的构造我们知道， $SC_{n,1}$ 具有的结点数为

$V_{n,1} = \frac{11}{70} \cdot 8^{n+1} + \frac{8}{5} \cdot 3^n + \frac{15}{7}$ ，边数为 $E_{n,1} = \frac{3}{10} \cdot 8^{n+1} + \frac{8}{5} \cdot 3^n + 4 \cdot 8^{n-1}$ ；对于 $SC_{n,1}$ 而言，有下列两类结点：

(1) 新加入的中心结点，其连接度为 $4 \cdot 8^{n-1}$ ，所连结的结点的之间存在的边数也为 $4 \cdot 8^{n-1}$ ，所以该结点的集聚系数是 $2/(4 \cdot 8^{n-1} - 1)$ ；

(2) 和中心结点相连的结点，共有 $4 \cdot 8^{n-1}$ 个结点，它们的连接度为5，每个结点的邻居实际存在的边数为2，所以此类结点的集聚系数是 $1/5$ 。

所以对于 $SC_{n,1}$ ，其集聚系数为

$$C_{n,1} = \frac{\frac{2}{4 \cdot 8^{n-1} - 1} + \frac{4}{5} \cdot 8^{n-1}}{\frac{11}{70} \cdot 8^{n+1} + \frac{8}{5} \cdot 3^n + \frac{15}{7}}。$$

由上式可得，当 $n \rightarrow \infty$ 时， $\lim_{n \rightarrow \infty} C_{n,1} = 7/130 \approx 0.054$ 。

4.3.2 网络的平均路径长度

所谓网络的平均路径长度是指网络中所有结点对之间的平均最短距离，它描述了网络结点之间的分离程度。这里结点间的距离指的是从一结点到另一结点所要经历的边的最小数目。网络的直径(Diameter)为任意两点间的最大距离。平均路径长度的计算公式为 $APL = \frac{1}{N(N-1)} \sum_{i,j} d_{ij}$ ，其中 d_{ij} 为结点 i 和 j 之间的最短距离， N 为网络的规模。由谢尔宾斯基地毯构造的复杂网络的定义我们得到 $SC_{n,1}$ 的直径为 $D_{n,1} = 6$ 。

由于网络的复杂性，我们没能给出一个关于网络的平均路径长度的表达式，但是我们已知网络的直径为 $D_{n,1} = 6$ ，所以网络的平均路径长度 $APL < 6$ 一定成立。这一小世界效应正与 Milgram 的小世界实验^[19]所得到的平均路径长度一样。

4.3.3 网络的平均度

网络中结点 i 的度 k_i 为结点 i 连接的边的总数目，所有结点 i 的度 k_i 的平均值称为网络的平均度，记为 $\langle k \rangle$ 。在我们构造的复杂网络中结点的度分布如下：

1. 度为 $4 \cdot 8^{n-1}$ 的结点，即中心结点，只有一个；
2. 度为 2 的结点，只有四个顶点；
3. 度为 3 的结点，共有 $\frac{8^{n+1}}{35} + \frac{16}{5} \cdot 3^n - \frac{24}{7}$ 个；
4. 度为 4 的结点，共有 $\frac{148}{35} \cdot 8^{n-1} - \frac{8}{5} \cdot 3^n + \frac{4}{7}$ 个；
5. 度为 5 的结点，共有 $4 \cdot 8^{n-1}$ 个。

所以网络的平均度为

$$\begin{aligned} \langle k_n \rangle &= \frac{4 \cdot 8^{n-1} + 2 \times 4 + 3 \times \left(\frac{8^{n+1}}{35} + \frac{16}{5} \cdot 3^n - \frac{24}{7} \right) + 4 \times \left(\frac{148}{35} \cdot 8^{n-1} - \frac{8}{5} \cdot 3^n + \frac{4}{7} \right) + 5 \times 4 \cdot 8^{n-1}}{\frac{11}{70} \cdot 8^{n+1} + \frac{8}{5} \cdot 3^n + \frac{15}{7}} \\ &= \frac{\frac{203}{35} \cdot 8^n + \frac{16}{5} \cdot 3^n}{\frac{11}{70} \cdot 8^{n+1} + \frac{8}{5} \cdot 3^n + \frac{15}{7}} \end{aligned}$$

由此可得，当 $n \rightarrow \infty$ 时，

$$\lim_{n \rightarrow \infty} \langle k_n \rangle = \lim_{n \rightarrow \infty} \frac{\frac{203}{35} \cdot \frac{1}{8} + \frac{16}{5} \cdot \left(\frac{3}{8}\right)^n}{\frac{11}{70} + \frac{1}{5} \cdot \left(\frac{3}{8}\right)^n + \frac{15}{7} \left(\frac{1}{8}\right)^{n+1}} = \frac{203 \cdot \frac{1}{8}}{\frac{11}{70}} = 203/44 \approx 4.61.$$

4.4 基于谢尔宾斯基地毯构建的网络的分形性

引理 8 由第 n 代谢尔宾斯基地毯构成的复杂网络 $SC_{n,1}$ ($n \geq 1$) 的盒维数和相似维数与 SC_n 的盒维数和相似维数相同, 即 $d_B = D_S \approx 1.8928$ 。

证明: 根据文[46]用盒维数的方法来计算网络图的维数, 用长度为 l_B 的盒子来覆盖网络图, 结点对之间的距离不大于 $l_B - 1$, 将有可能在一个盒子里, 得到所需的盒子数为 N_B , 对于 $SC_{n,1}$, 具体示意图见图 4.3,

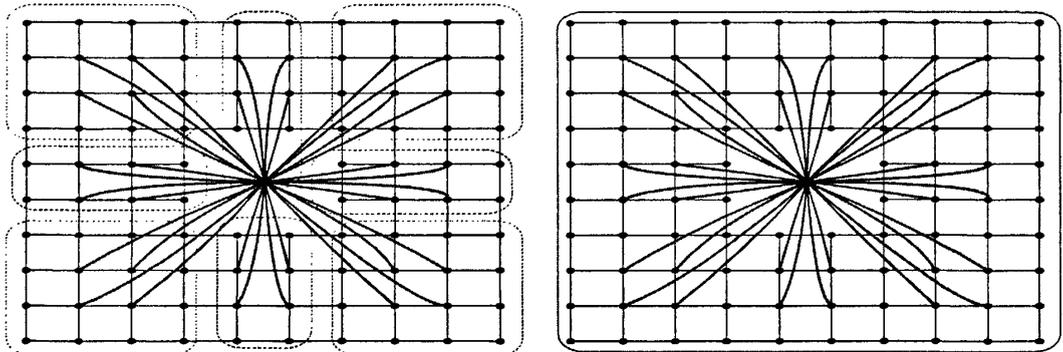


图 4.3 对 $n=2$ 时, 即 $SC_{2,1}$ 时, 分别用 $l_B=7$ 和 $l_B=19$ 来覆盖网络图的示意图

Fig 4.3 Boxes in $SC_{2,1}$ with $l_B=7$ and $l_B=19$

结果如下表 4.1。

表 4.1 盒维数计算结果

Table 4.1 The box-counting results

l_B	3	7	19	$2 \cdot 3^n + 1$
N_B	8^n	8^{n-1}	8^{n-2}	1

所以可得该网络图的盒维数 $d_B = \frac{\ln 8}{\ln 3} \approx 1.8928$ 。

现在我们采用相似维数^[67]的方法来计算网络图的维数，显然新加入的结点不影响网络的维数。如图 4.1 把正方形 SC_n 的边长放大 $L = 3$ 倍，则得到与 SC_{n+1} 相似的正方形，由于 SC_{n+1} 的面积是 SC_n 的 $K = 8$ 倍。所以，利用相似维数的计算公式：

$$D_s = \frac{\ln K}{\ln L} = \frac{\ln 8}{\ln 3} \approx 1.8928。$$

所以 $d_B = D_s \approx 1.8928。$

4.5 网络图的拓展讨论

因为 SC_{n+1} 是由 SC_n 加入一个中心节点与所有第一代谢尔宾斯基地毯的四个内顶点相连而得到的，我们自然联想到，考虑加入的这个中心节点与第 i 代谢尔宾斯基地毯四个内顶点相连而形成的网络图的特性，我们不妨记为 $SC_{n,i}$ ，其中 $n \geq i \geq 2$ 。具体可见图 4.4。

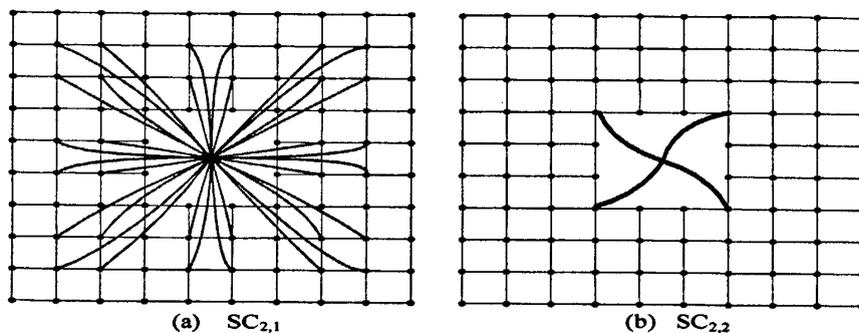


图 4.4 这是 $n = 2$ 时的连接示意图。(a)是连接第 1 代谢尔宾斯基地毯的四个内顶点，形成 $SC_{2,1}$ ；
(b)是连接第 2 代谢尔宾斯基地毯的四个内顶点，形成 $SC_{2,2}$

Fig 4.4 The figure of the network when $n = 2$. (a) $SC_{2,1}$; (b) $SC_{2,2}$

分析发现， $SC_{n,i}$ 的网络拓扑特性存在一定的规律性，并且和 $SC_{n,1}$ 存在一定的区别。这里从集聚系数、网络的直径和分形维数三方面进行考察。

(1) 集聚系数。对于 $SC_{n,i}$ ($i \geq 2$)，只存在一类集聚系数非零结点，即新加入的中心结点，其连接度为 $4 \cdot 8^{n-i}$ ，所连结的结点的之间存在的边数也为 $4 \cdot 8^{n-i}$ ，所以该结点的集

聚系数是 $2/(4 \cdot 8^{n-i} - 1)$ 。所以对于 $SC_{n,i}$ ($i \geq 2$)，其集聚系数 $C_{n,i} = \frac{2}{4 \cdot 8^{n-i} - 1}$ ，
 $\frac{11}{70} \cdot 8^{n+1} + \frac{8}{5} \cdot 3^n + \frac{15}{7}$

且 $\lim_{n \rightarrow \infty} C_{n,i} = 0。$

(2) 网络的直径。很显然, 随着 i 的增大, 网络的直径必将随着增大。用数学归纳的方法不难求解, 对于 $SC_{n,i}$, 其直径为 $D_{n,i} = 4 \cdot 3^{i-1} + 2$ 。可以得出当 $i \leq \log_3^n$ 时,

$$D_{n,i} \leq n + 2 = o(\log_3^n), \text{ 其中 } N \text{ 为网络的规模, 即点的个数为 } \frac{11}{70} \cdot 8^{n+1} + \frac{8}{5} \cdot 3^n + \frac{15}{7}。$$

(3) 网络的维数。由于只是增加了一个节点, 所以其维数和 $SC_{n,1}$ 一样, 仍为 $D_s \approx 1.8928$ 。

从上面 3 个属性可以看出, 拓展的网络图也是同时具有小世界和分形特性的。从而 $SC_{n,i}$ 和 $SC_{n,1}$ 是一类具有分形和小世界特性的网络图, 并且 $SC_{n,i}$ 的分形和小世界特性随着代数 i 的增加而逐渐减弱。

4.6 小结

本章给出了谢尔宾斯基地毯的一些网络特性, 包括网络直径、集聚系数, 发现其集聚系数为零, 网络直径呈指数极增长, 不具有小世界效益。于是我们在这谢尔宾斯基地毯的基础上, 构造出具有小世界特性的复杂网络, 研究了它们的网络拓扑特性, 得到由谢尔宾斯基地毯生成的网络具有的网络维数为 1.8928、集聚系数 0.054、平均路径长度小于 6 和平均度 4.61 具有小世界网络的特性; 通过计算也得到该网络的盒维数和相似维数均约为 1.8928, 由其自身的自相似性和具有的精细结构, 其分形特征也得到以体现。

5 一类具有分形和小世界特性的网络图

本章提出了一类具有分形和小世界特性的网络图。利用数学归纳的方法计算出了网络图的集聚系数，平均最短路径和网络图的直径，证明了网络图的小世界特性。用盒维数和豪斯道夫维数来衡量网络图的分形性，得到其维数均为 1.585。最后对网络图的构造方法作了进一步地拓展，并给出了拓展的网络图的相关拓扑特性的表达式，并认为其和原来的网络图可归结为一类具有分形和小世界特性的网络图。

5.1 谢尔宾斯基垫片

5.1.1 谢尔宾斯基垫片的构造

谢尔宾斯基垫片是俄国数学家谢尔宾斯基早在 1915 年~1916 年构造出的一些“病态”图形之一（同期还提出了谢尔宾斯基地毯、谢尔宾斯基支架,将在下一小节中介绍），其构造方法如下^[66]：

- a. 初始一个正三角形 SG_0 ；
- b. 用三个正三角形 SG_0 组成一个边长是 SG_0 边长两倍的正三角形（不包括中间的倒三角形）；
- c. 按照上面的方法进行下去，可知 SG_n 是用三个正三角形 SG_{n-1} 组成的一个边长是 SG_{n-1} 边长两倍的正三角形（不包括中间的倒三角形），具体的过程见下面的图 5.1：

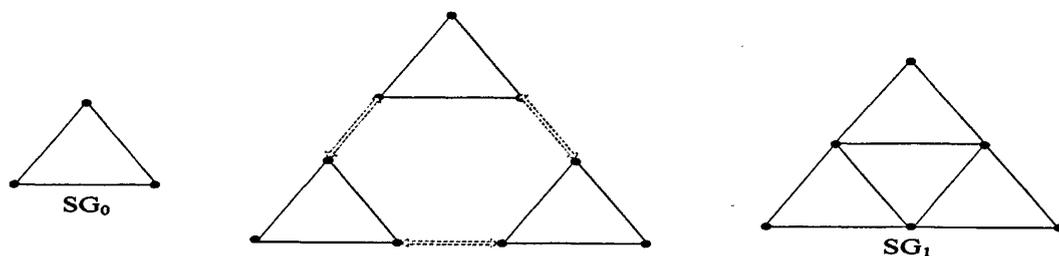


图 5.1 谢尔宾斯基垫片的迭代过程示意图

Fig5.1 iterative process of Sierpinski gasket

将三角形的顶点看作是网络中的点，三角形的边可看作网络中的结点之间存在着的连接，所以谢尔宾斯基垫片也可以被看作是一个较复杂的网络图。为方便讨论，我们简记第 n 代谢尔宾斯基垫片(Sierpinski gasket)为 SG_n 。

很显然，谢尔宾斯基垫片图形具有很明显的特点：

(1) 图形具有严格的自相似性。由一个正三角形，按上面的方法经过了无穷多次分割舍弃后所形成的“空三角”，图形中每一个“空三角”其局部与整体是相似的。由图

5.1 中可以观察将虚线圈出的局部与整体对比观察可知，由于它们都按相同的方式进行无限细分、舍弃这一操作，所以内部结构是完全相同的，只是它们之间的大小不同而已。整体与局部是严格相似的，是很典型的自相似性，即是一个规则分形。

(2) 图形的维数介于 1 与 2 之间。从图 5.1 四个图的变化趋势可以看出，图形的总体轮廓可以想象为一个中间有空隙的三角形，其内部仍按上述规则将三角形进行分割、舍弃，从而形成了众多小三角形。当分割、舍弃操作无限延续时，其中一些小三角形的数目越来越多，空隙也越来越多。于是这些大大小小的三角形的边，其线段数目越来越多，长度越来越短，而线段的总长度越来越长，图形的线段总长度趋于无穷大，而总面积则趋于 0。换句话说，当用一维尺寸去量度图形时，测量得到的结果是无穷大，而当用二维的尺寸去量度图形时，测量得到的结果是 0，由此可推断出图形的维数介于 1 与 2 之间。

5.1.2 谢尔宾斯基垫片的性质

引理9 第 n 代谢尔宾斯基垫片 SG_n ($n \geq 0$)，具有以下特性：

a. SG_n 具有的结点数为 $V_n = \frac{3^{n+1} + 3}{2}$ ，边数为 $E_n = 3^{n+1}$ 。

b. SG_n 的直径为 $D_n = 2^n$ 。

c. SG_n 的集聚系数为 $C_n = \frac{4 \cdot 3^{n-1} + 5}{3^{n+1} + 3}$ 。

证明：

a) 设在 SG_n 中的结点数和边数分别为 V_n 和 E_n ，则由 SG_n 的构造过程我们得到下式成立：

$$\begin{aligned} V_n &= 3V_{n-1} - 3 \\ E_n &= 3E_{n-1} \end{aligned}, \quad (n \geq 1)$$

又由于 $V_0 = 3$ ， $E_0 = 3$ ，根据上边的递推公式及初始条件我们得到(a)成立；

b) 同理由于 SG_n 的直径 D_n 满足 $D_n = 2D_{n-1}$ ，且 $D_0 = 1$ ，所以 $D_n = 2^n$ ；

c) 关于 SG_n 集聚系数的证明我们首先将结点分为如下三类：一类是外围大三角形的三个顶点，它们的连接度为 2，所连接的两个结点也相互连接，所以结点的集聚系数为 1；显然剩余所有结点的度数都为 4，其中 ($n \geq 1$)，在这些结点中又有两类，一类结点集聚系数为 1/2，设其个数为 a_n ；另一类结点集聚系数为 1/3，设其个数为 b_n ；由 SG_n 的构造过程我们得下式成立：

$$a_n = 3a_{n-1}, \quad a_1 = 3,$$

$$b_n = 3b_{n-1} + 3, \quad b_1 = 0.$$

解这两个递归方程得 $a_n = 3^n$, $b_n = \frac{3^n - 3}{2}$, 所以 SG_n 的集聚系数为

$$C_n = \frac{3 \times 1 + 3^n \times \frac{1}{2} + \frac{3^n - 3}{2} \times \frac{1}{3}}{\frac{3^{n+1} + 3}{2}} = \frac{4 \cdot 3^{n-1} + 5}{3^{n+1} + 3}.$$

5.2 基于谢尔宾斯基垫片构建的网络

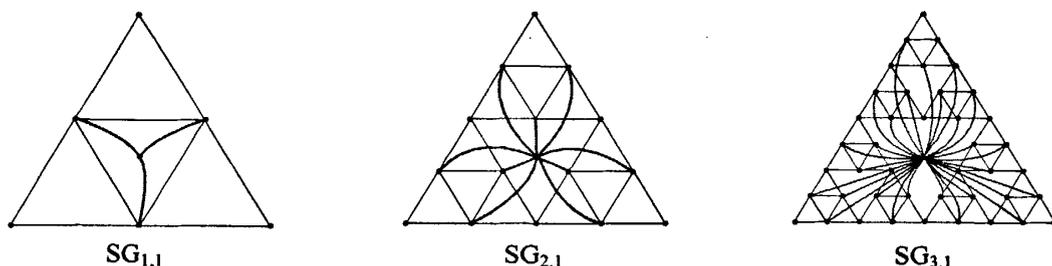


图 5.2 定义在谢尔宾斯基垫片基础上的复杂网络

Fig 5.2 The networks based on Sierpinski gasket

定义在谢尔宾斯基垫片基础上的复杂网络如下：定义 $SG_{n,1}$ 为在第 n 代谢尔宾斯基垫片 SG_n 中加入一中心结点，然后将此中心结点与所有第一代谢尔宾斯基垫片的三条边的中点相连，形成过程如图 5.2 所示。

引理 10 由第 n 代谢尔宾斯基垫片构成的复杂网络 $SG_{n,1}$ ($n \geq 1$)，具有以下特性：

a. $SG_{n,1}$ 具有的结点数为 $V_{n,1} = \frac{3^{n+1} + 5}{2}$ ，边数为 $E_{n,1} = 3^{n+1} + 3^n$ 。

b. $SG_{n,1}$ 的直径为 $D_{n,1} = 4$ 。

5.3 基于谢尔宾斯基垫片构建的网络的小世界特性

5.3.1 网络的集聚系数

集聚系数是用来描述网络中节点的聚集情况，即网络有多紧密，比如在社会网络中，你朋友的朋友可能也是你的朋友或者你的两个朋友可能彼此也是朋友。节点 i 集聚系数 C_i 定义为：它的 k_i 个直接邻居之间实际存在的边数 E_i 占所有可能存在的边数 $k_i(k_i - 1)/2$

的比例, 即 $C_i = 2E_i / [k_i(k_i - 1)]$ 。整个网络的集聚系数指的是, 所有节点集聚系数的算术平均值。

引理11 由第 n 代谢尔宾斯基垫片构成的复杂网络 $SG_{n,1}$ ($n \geq 1$), 具有的集聚系数为

$$C_{n,1} = \frac{4 \cdot 3^{2n-1} + 11 \cdot 3^{n-1} - 1}{3^{2n+1} + 2 \cdot 3^n - 5}。$$

证明: 根据引理 10 及网络的构造我们知道, $SG_{n,1}$ 具有的结点数为 $V_{n,1} = \frac{3^{n+1} + 5}{2}$, 边数为

$E_{n,1} = 3^{n+1} + 3^n$; 对于 $SG_{n,1}$ 而言, 有下列四类结点:

(1) $SG_{n,1}$ 的三个顶点, 它们的集聚系数为 1;

(2) 和中心结点相连的结点, 共有 3^n 个结点, 它们的连接度为 5, 每个结点的邻居间存在的边数为 5, 所以此类结点的集聚系数是 $1/2$;

(3) 新加入的中心结点, 其连接度为 3^n , 所连结的结点的之间存在的边数也为 3^n , 所以该结点的集聚系数是 $2/(3^n - 1)$;

(4) 剩下的结点还有 $(3^n - 3)/2$ 个, 其连接度为 4, 此类结点邻居之间有 2 条边, 所以该类结点的集聚系数是 $1/3$ 。

所以对于 $SG_{n,1}$, 其集聚系数为

$$C_{n,1} = \frac{3 \times 1 + 3^n \times \frac{1}{2} + 1 \times \frac{2}{3^n - 1} + \frac{3^n - 3}{2} \times \frac{1}{3}}{\frac{3^{n+1} + 5}{2}} = \frac{4 \cdot 3^{2n-1} + 11 \cdot 3^{n-1} - 1}{3^{2n+1} + 2 \cdot 3^n - 5}。$$

由上式可得, 当 $n \rightarrow \infty$ 时,

$$\lim_{n \rightarrow \infty} C_{n,1} = 4/9 \approx 0.44。$$

5.3.2 网络的平均路径长度

所谓网络的平均路径长度是指网络中所有节点对之间的平均最短距离, 它描述了网络节点之间的分离程度。这里节点间的距离指的是从一节点到另一节点所要经历的边的最小数目。网络的直径(Diameter)为任意两点间的最大距离。平均路径长度的计算公式为

$APL = \frac{1}{N(N-1)} \sum_{i,j} d_{ij}$, 其中 d_{ij} 为节点 i 和 j 之间的最短距离, N 为网络的规模。

显然该网络中结点之间的距离有四种情况，分别为 1, 2, 3, 4。所以对于引理 3 中所论述的四类结点，这里我们将 $SG_{n,1}$ ($n \geq 2$) 网络中结点对之间的距离列于下表 5.1。

表 5.1 节点之间的距离表
Table 5.1 The distance between vertexes

节点对 点的种类	距离 个数	1	2	3	4
节点的连结度为 2 (共有 3 个节点)		2	4	$3^n - 3$	$(3^n - 3)/2$
节点的连结度为 4 (共有 $(3^n - 3)/2$ 个节点)		4	7	$3^n - 6$	$(3^n - 7)/2$
节点的连结度为 5 (共有 3^n 个节点)		5	$3^n - 2$	$(3^n - 3)/2$	0
节点的连结度为 3^n (即中心节点)		3^n	$(3^n + 3)/2$	0	0

从表 5.1 可得：

节点连结度为 2 的距离总和是：

$$3 \times [1 \times 2 + 2 \times 4 + 3 \cdot (3^n - 3) + 4 \cdot (3^n - 3)/2] = 15 \cdot 3^n - 15$$

节点连结度为 4 的距离总和是：

$$\frac{3^n - 3}{2} \times [4 + 2 \times 7 + 3(3^n - 6) + 4(3^n - 7)/2] = \frac{5 \cdot 3^{2n} - 29 \cdot 3^n + 42}{2}$$

节点连结度为 5 的距离总和是：

$$3^n \times [5 + 2(3^n - 2) + 3(3^n - 3)/2] = \frac{7 \cdot 3^{2n} - 7 \cdot 3^n}{2}$$

节点连结度为 3^n 的距离总和是：

$$1 \times [3^n + 2 \cdot (3^n + 3)/2] = 2 \cdot 3^n + 3$$

由网络的平均距离的计算公式可得， $SG_{n,1}$ 的平均距离为：

$$APL = \frac{1}{N(N-1)} \sum_{i \neq j} d_{ij} = \frac{24 \cdot 3^{2n} - 4 \cdot 3^n + 36}{9 \cdot 3^{2n} + 24 \cdot 3^n + 15}$$

由 $N = (3^{n+1} + 5)/2$ ，可得 $3^n = (2N - 5)/3$ ，代入 (3) 式，可得：

$$APL = \frac{8N^2 - 42N + 28}{3N^2 - 3N}$$

由上式易得当 N 较大时, $APL < \log(N)$, 且

$$\lim_{N \rightarrow \infty} APL = \lim_{N \rightarrow \infty} \frac{8N^2 - 42N + 28}{3N^2 - 3N} = \frac{8}{3} \approx 2.67.$$

5.3.3 网络的平均度

网络中结点 i 的度 k_i 为结点 i 连接的边的总数目, 所有结点 i 的度 k_i 的平均值称为网络的平均度, 记为 $\langle k \rangle$ 。在我们构造的复杂网络中结点的度分布如下: 度为 3^n 的结点, 即中心结点, 只有一个; 度为 2 的结点, 只有 3 个顶点; 度为 4 的结点, 共有 $(3^n - 3)/2$ 个; 度为 5 的结点, 共有 3^n 个。

所以网络的平均度为

$$\langle k_n \rangle = \frac{1 \times 3^n + 2 \times 3 + 4 \times (3^n - 3)/2 + 5 \times 3^n}{(3^{n+1} + 5)/2} = \frac{16 \cdot 3^n}{3 \cdot 3^n + 5},$$

即, 当 $n \rightarrow \infty$ 时, $\lim_{n \rightarrow \infty} \langle k_n \rangle = 16/3 = 5.3$ 。

5.4 基于谢尔宾斯基垫片构建的网络的分形性

引理12 由第 n 代谢尔宾斯基垫片(Sierpinski gasket)构成的复杂网络 $SG_{n,1}$ ($n \geq 1$) 的盒维数和相似维数与 SG_n 的盒维数和相似维数相同, 即 $d_B = D_S \approx 1.585$ 。

证明: 根据文[46]用盒维数的方法来计算网络图的维数, 用长度为 l_B 的盒子来覆盖网络图, 节点对之间的距离不大于 $l_B - 1$, 将有可能在一个盒子里, 得到所需的盒子数为 N_B , 对于 SG_n , 具体结果如下表 5.2, 示意图见图 5.3。

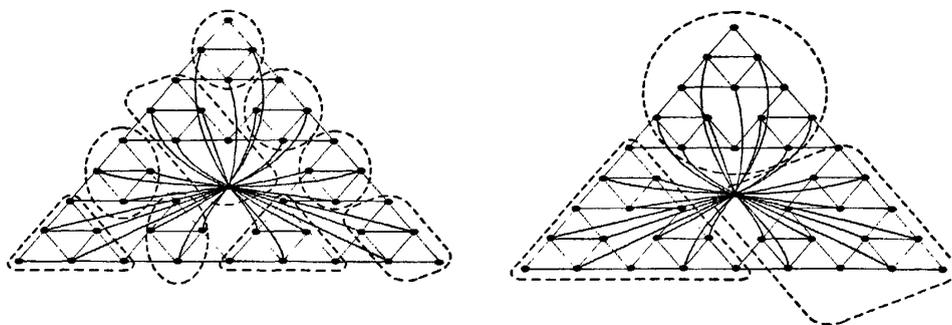


图 5.3 对 $n = 3$ 时, 即 G_3' 时, 分别用 $l_B = 3$ 和 $l_B = 5$ 来覆盖网络图的示意图

Fig 5.3 The Graph covered with different l_B

表 5.2 盒维数计算结果

Table 5.2 The calculation result of the box-counting dimension

l_B	2	3	5	$2^n + 1$
N_B	3^n	3^{n-1}	3^{n-2}	1

所以可得该网络图的盒维数:

$$d_B = \frac{\ln 3}{\ln 2} \approx 1.585$$

现在我们采用相似维数^[67]的方法来计算网络图的维数, 显然新加入的结点不影响网络的维数。如图 5.1 把三角形 SG_n 的边长放大 $L = 2$ 倍, 则得到与 SG_{n+1} 相似三角形, 由于 SG_{n+1} 的面积是 SG_n 的 $K = 3$ 倍。所以, 利用相似维数的计算公式:

$$D_s = \frac{\ln K}{\ln L} = \frac{\ln 3}{\ln 2} \approx 1.585。$$

所以 $d_B = D_s \approx 1.585$ 。

5.5 网络图的拓展讨论

因为 SG_{n+1} 是由 SG_n 加入一个中心节点与所有最小的倒三角形的顶点相连而得到的, 我们自然联想到, 加入的这个中心节点与次小的倒三角形或者是更大的倒三角形的顶点相连形成的网络图的特性, 我们设中心节点选择连接的倒三角形边长是最小倒三角形的 $i (i > 1)$ 倍, 则形成的网络图, 不妨设为 $SG_{n,i}$ 。具体可见下图 5.4。

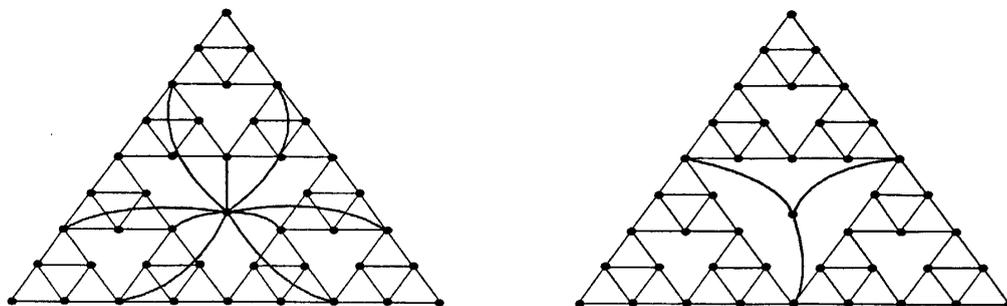


图 5.4 $n = 3$ 时的连接示意图

Fig 5.4 The figure of the network when $n = 3$

注 1) 左图是中心节点连接边长是最小倒三角形的 2 倍的倒三角形的顶点, 形成 $SG_{3,2}$; 右图是中心节点连接边长是最小倒三角形的 3 倍的倒三角形的顶点, 形成 $SG_{3,3}$ 。

分析发现, $SG_{n,i}$ 的网络拓扑特性存在一定的规律性, 并且和 $SG_{n,1}$ 存在一定的区别。这里从集聚系数、网络的直径和分形维数三方面进行考察。

(1) 集聚系数。对于 $SG_{n,i}$ ，也存在四类点，其中连接度为 2 的节点有 3 个，集聚系数为 1；连接度为 5 的节点有 3^{n-i+1} 个，集聚系数为 $1/5$ ；连接度为 4 的节点集聚系数有 $1/2$ 和 $1/3$ 两种情况， $1/2$ 的有 3^n 个， $1/3$ 的有 $\frac{3^n-3}{2} - 3^{n-i+1}$ 个；中心节点的连接度为 3^{n-i+1} 的集聚系数为 0。

所以对于 $SG_{n,i}$ ，其集聚系数 $C_{n,i} = \frac{20 \cdot 3^n - 4 \cdot 3^{n-i+1} + 75}{45 \cdot 3^n + 75}$ ，且 $\lim_{n \rightarrow \infty} C_{n,i} = \frac{4}{9} \approx 0.44$ 。

(2) 网络的直径。很显然，随着 i 的增大，网络的直径必将随着增大。用数学归纳的方法不难求解，对于 $SG_{n,i}$ ，当 $i < n$ 时， $D_{n,i} = 2^i + 2$ ；当 $i = n$ ， $D_{n,i} = 2^i$ 。可以得出当 $i \leq \log_2 n$ 时， $D_{n,i} \leq n + 2 = o(\log_2 N)$ ，其中 N 为网络的规模，即点的个数，值为 $(3^{n+1} + 5)/2$ 。

(3) 网络的维数。由于也是加了一个节点，所以其维数和 $SG_{n,1}$ 一样，仍为 $\frac{\ln 3}{\ln 2} \approx 1.585$ 。

从上面 3 个属性可以看出，拓展的网络图也是同时具有小世界和分形特征的。从而 $SG_{n,i}$ 和 $SG_{n,1}$ 可以看作是一类具有分形和小世界特性的网络图。

5.6 小结

本章从谢尔宾斯基垫片的构造方法出发，提出了一个同时具有分形特性和小世界特性的网络图，计算了网络图的拓扑特性，得出其集聚系数约为 0.51，平均最短路径约为 2.67，证明了它的小世界效应。同时，我们又同时用盒维数和豪斯道夫维数计算网络图的分维，并得到共同结果，分维约为 1.585，从而验证了网络图是一个分形网络。最后，我们对模型又进行了进一步拓展，并分析了其相关拓扑属性。在网络图中加入的中心节点相当于有些实际网络中的核心节点，所以本文构造的网络图对于实际网络同时存在的分形性和小世界效应现象具有一定的参考价值，希望本模型能够为这方面的研究提供一些借鉴。

结 论

自 WS 模型和 BA 模型的提出以来, 复杂网络的研究浪潮已经发展了近九年了, 在这期间, 复杂网络的演化模型、实证研究以及网络的动力学方面都取得了巨大的进步。本文针对目前复杂网络演化及自相似性研究的现状, 并深入探求其存在的不足, 现将本文的研究工作做一下总结:

(1) 经典的谢尔宾斯基支架网络并不具备小世界特性, 其集聚系数较小、网络直径指数级增长。而本章以谢尔宾斯基支架为基础, 构造出了高维的具有分形和小世界特性的网络, 通过数学归纳方法推导出该网络具有较大的集聚系数 0.53 和较短的平均路径长度 2.5, 证明了其小世界特性, 并用分形理论中相似维数和盒维数两种方法计算出网络的分维数约为 2, 由网络的自相似性及其具有的精细结构证明了该网络的分形特性。

本章构造的具有小世界特性的网络中, 新加入的中心结点就像是网络中的路由器, 在不同的网路中, 起到大量缩短网络距离的作用, 使着无限扩大的网络有着相互联系; 就像 Milgram 的小世界实验, 虽然茫茫人海中人与人之间的距离表面上看很远, 但是往往因为个别亲朋好友的存在使得表面上很远的关系并不存在。该网络图对于同时存在分形性和小世界效应的实际网络的研究具有一定的参考价值。另外, 该网络支架的演化还为新衍架结构的设计及其系统稳定性、可靠性提供了新的启示。

(2) 本章给出了谢尔宾斯基地毯的一些网络特性, 包括网络直径、集聚系数, 发现其集聚系数为零, 网络直径呈指数极增长, 不具有小世界效益。于是我们在这谢尔宾斯基地毯的基础上, 构造出具有小世界特性的复杂网络, 研究了它们的网络拓扑特性, 得到由谢尔宾斯基地毯生成的网络具有的网络维数为 1.8928、集聚系数 0.054、平均路径长度小于 6 和平均度 4.61 具有小世界网络的特性; 通过计算也得到该网络的盒维数和相似维数均约为 1.8928, 由其自身的自相似性和具有的精细结构, 其分形特征也得到以体现。

(3) 本章从谢尔宾斯基垫片的构造方法出发, 提出了一个同时具有分形特性和小世界特性的网络图, 计算了网络图的拓扑特性, 得出其集聚系数约为 0.51, 平均最短路径约为 2.67, 证明了它的小世界效应。同时, 我们又同时用盒维数和豪斯道夫维数计算网络图的分维, 并得到共同结果, 分维约为 1.585, 从而验证了网络图是一个分形网络。最后, 我们对模型又进行了进一步拓展, 并分析了其相关拓归属性。在网络图中加入的中心节点相当于有些实际网络中的核心节点, 所以本文构造的网络图对于实际网络同时存在的分形性和小世界效应现象具有一定的参考价值, 希望本模型能够为这方面的研究提供一些借鉴。

本文的不足之处及进一步的工作：

复杂网络高速发展，新理论、新成果也是日新月异，本文所提出的三个模型虽然是对目前模型做了一些改进，并将某些理论的内容加入了复杂网络的研究，但还只是某一方面的研究，还有许多需要完善的工作，现将其总结如下：

(1) 将实证研究加入演化模型的研究中来。虽然每个现实网络有着许多共性特征，但对于某个特定网络来说，一定有其特殊性。建立一个具体现实网络的模型，并利用现有复杂网络的特征参数，并将是一个具有挑战性的工作。目前本人已有某地区电信行业的海量数据，着手准备研究一下电信网络的统计特性，并提出一个电信网络特有的网络演化模型。

(2) 复杂网络的自相似原因的进一步探究。nature 上那篇自相似性文章给人以研究复杂网络的新视角，研究者的复杂网络演化模型的自相似证明开始利用这篇文章说的盒维数方法来证明。但究其自相似的原因，目前研究工作还只是起步阶段。分形理论十分复杂和庞大，本人只是将一个计算分形维数的方法引入，还有许多东西值得探索。

(3) 复杂网络的动力学研究。动力学研究是复杂网络的一个重要研究方向，本文并没有涉及，对本文提出的模型的动力学特征的研究是一个有意义的工作。而动力学必然和自组织现象、混沌现象相关联，而这些又和分形理论息息相关，所以将其综合起来，必将产生新的火花。

参 考 文 献

- [1] Fred Buckley, Marty Lewinter. 图论简明教程. 北京:清华大学出版社, 2005.
- [2] Erdős P, Rényi A. On the evolution of random graphs. *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, 1960, 5: 17-61.
- [3] Albert R, Jeong H, Barabási A-L. Diameter of the World Wide Web. *Nature*, 1999, 401:130-131.
- [4] 汪小帆, 李翔, 陈关荣. 复杂网络理论及其应用. 北京: 清华大学出版社, 2006.
- [5] Newman M E J. The structure of scientific collaboration networks. *Proc. Natl. Acad. Sci. USA*, 2001, 98: 404-409.
- [6] Jeong H, Tombor B, Albert R. et al. The large scale organization of metabolic networks. *Nature*, 2000, 407: 651-654.
- [7] Watts D J, Strogatz S H. Collective dynamics of 'small-world' networks. *Nature*, 1998, 393: 440-442.
- [8] Barabási A-L, Albert R. Emergence of scaling in random networks. *Science*, 1999, 286: 509-512.
- [9] Faloutsos M, Faloutsos P, Faloutsos C, On power-law relationships of the internet topology. *Comput. Commun. Rev.* 1999, 29: 251-260.
- [10] Albert R, Albert I, Nakarado G L. Structural vulnerability of the North American power grid. *Physical Review E*, 2004, 69:025103(R).
- [11] Liljeros F, Edling C R, Amaral L A N et al. The web of human sexual contacts. *Nature*, 2001, 411:907-908.
- [12] Redner S. How popular is your paper? An empirical study of the citation distribution. *The European Physical Journal B*, 1998, 4: 131-134.
- [13] Albert R, Barabási A-L. Statistical mechanics of complex networks. *Reviews of Modern Physics*, 2002, 74(1): 47-97.
- [14] 车宏安, 顾基发. 无标度网络及其系统科学意义. *系统工程理论与实践*, 2004, 44(4):11-16.
- [15] Newman M E J, Forrest S, Balthrop J. Email networks and the spread of computer viruses. *Physical Review E*, 2002, 66:035101 (R).
- [16] Duan W-Q, Chen Z, Liu Z-R, et al. Efficient target strategies for contagion in scale-free networks. *Physical Review E*, 2005, 72:026133.
- [17] 吴彤. 复杂网络研究及其意义. *哲学研究*, 2004, 8: 58-63.
- [18] Huang Li, Park K, Lai Y-C. Information propagation on modular networks. *Physical Review E*, 2006, 73:035103(R).
- [19] Dodds P S, Watts D J, Sabel C F. Information exchange and the robustness of organizational networks. *Proc. Natl. Acad. Sci. USA*, 2003, 100(21): 12516-12521.

- [20] 张宁. 复杂网络实证研究——中国教育网. 系统工程学报, 2006, 21, (4):337-341.
- [21] 章忠志. 复杂网络演化模型研究: (博士论文)大连: 大连理工大学, 2006.
- [22] 王冰. 复杂网络的演化机制及若干动力学研究: (博士论文)大连: 大连理工大学, 2006.
- [23] 刘军. 复杂网络研究: 模型和统计性质: (硕士论文)兰州: 兰州大学, 2006
- [24] Newman M E J, Watts D J. Renormalization group analysis of the small-world network model. *Physics Letters. A*, 1999, 263:341-346.
- [25] Barabási A-L, Albert R, Jeong H. Mean-field theory for scale-free random networks. *Physica A* 1999, 272:173-187
- [26] Krapivsky P L, Redner S, Leyvraz F. Connectivity of growing random networks [J]. *Physical Review Letters*, 2000, 85(21): 4629 - 4632 (2000).
- [27] S.N. Dorogovtsev, J.F.F. Mendes, Scaling behaviour of developing and decaying networks, *Europhys. Lett.* 52 (2000) 33 - 39.
- [28] R. Albert, A.-L. Barabasi, Topology of evolving networks: local events and universality, *Phys. Rev. Lett.* 85 (2000) 5234 - 5237.
- [29] Klemm K, Eguíluz V M. Highly clustered scale-free networks. *Physical Review E*, 2002, 65: 036123
- [30] M. E. J. Newman, D. J. Watts. Scaling and percolation in the small-world network model [J]. *Physical Review Letters*, 1999, 60(6):7332-7342.
- [31] J. Kleinberg. Navigation in a small world [J]. *Nature*, 2000, 406:845.
- [32] Zhongzhi Zhang, Lili Rong, Chonghui, Guo. A deterministic small-world network created by edge iterations [J]. *Physica A*, 2006, 363:567-572
- [33] Song C, Havlin S and Makse H A. Origins of fractality in the growth of complex networks [J]. *Nature Physics*. 2006, 2(4):275 - 281
- [34] Dong-Hee Kim, Jae Dong Noh, Hawoong Jeong. Scale-free trees: The skeletons of complex networks [J]. *Physical Review E*, 2004, 70:046126
- [35] Goh K I, Salvi G, Kahng B and Kim D. Skeleton and fractal scaling in complex networks [J]. *Physical Review Letter*, 2006, 96: 018701
- [36] J. S. Kim, K-I. Goh, G Salvi, E. Oh, B. Kahng, D. Kim. Fractality in complex networks: critical and supercritical skeletons [J]. arXiv:cond-mat/0605324
- [37] J. P. K. Doye, C. P. Massen. Self-similar disk packings as model spatial scale-free networks [J]. *Physical Review E*, 2005, 71:016128.
- [38] Z. Z. Zhang, F. Comellas, G. Fertin, L. L. Rong, High dimensional Apollonian networks [J]. arXiv:cond-mat/0503316.
- [39] 陈禹, 宗晓, 郝杰, 许彦. BA模型的三种扩展. 系统工程学报, 2005, 20(2):120-127.
- [40] 李增扬, 韩秀萍, 陆君安, 何克清. 内部演化的BA无标度网络模型. 复杂系统与复杂性科学, 2005, 2(2):1-6.

- [41] 唐芙蓉, 蔡绍洪, 李朝辉. 无标度网络的嵌入-删除-补偿模型的建立及分析. 中国矿业大学学报, 2005, 34(3):390-393.
- [42] Chen qinghua, Shi Dinghua. The modeling of scale-free networks. *Physica A*, 2004, 335: 240-248.
- [43] ZM Gu, T Zhou, BH Wang, G Yan, CP Zhu, ZQ Fu. Simplex triangulation induced scale-free networks [J]. arXiv: cond-mat/0505175
- [44] ZhongZhi Zhang, Li-Li Rong and France Comellas. Evolving small-world networks With geographical attachment preference, *J. Phys. A:Math. Gen*, 2006, 39:3253-3261.
- [45] 李守伟, 钱省三. 均匀增长无标度网络的等价模型. 上海理工大学学报, 2006, 28(3):218-212.
- [46] Song C, Havlin S and Makse H A. Self-similarity of complex networks [J]. *Nature*, 2005, 433:392 - 395
- [47] S. N. Dorogovtsev, A. V. Goltsev, J. F. F. Mendes. Pseudofractal scale-free web [J]. *Physical Review E*, 2002, 65:066122
- [48] Dong-Hee Kim, Jae Dong Noh, Hawoong Jeong. Scale-free trees: The skeletons of complex networks [J]. *Physical Review E*, 2004, 70:046126
- [49] Goh K I, Salvi G, Kahng B and Kim D. Skeleton and fractal scaling in complex networks [J]. *Physical Review Letter*, 2006, 96: 018701
- [50] J. S. Kim, K-I. Goh, G Salvi, E. Oh, B. Kahng, D. Kim. Fractality in complex networks: critical and supercritical skeletons [J], arXiv:cond-mat/0605324
- [51] Soon-Hyung Yook, Filippo Radicchi, et al. Self-similar scale-free networks and disassortativity, *Physical Review E*, 2005, 72:045105.
- [52] Hermón D. Rozenfeld, Shlomo Havlin, and Daniel ben-Avranham. Fractal and Transfractal Recursive Scale-Free Nets, arXiv:cond-mat/0612330.
- [53] D. J. B. Soares, J. Ribeiro Filho, et al. A Fractal Space-filling Complex Network, arXiv:cond-mat/0508359
- [54] Wei-Xing Zhou, Zhi-Qiang Jiang, Didier Sornette. Exploring self-similarity of complex cellular network: the edge-covering method with simulated annealing and log-periodic sampling, arXiv:cond-mat/0605676
- [55] Song C, Havlin S and Makse H A. Origins of fractality in the growth of complex networks [J]. *Nature Physics*. 2006, 2(4):275 - 281
- [56] Chaoming Song, Lazaros K. Gallos, et al. How to calculate the fractal dimension of a complex network: the box covering algorithm, arXiv:cond-mat/0701216.
- [57] Yook S H, Jeong H, Barabási A-L. Weighted evolving networks. *Physical Review Letters*, 2001, 86:5835-5838.
- [58] Han J D, Bertin N, Hao T et al. Evidence for dynamically organized modularity in the yeast protein-protein interaction network. *Nature*, 2004, 430:88-93.

- [59] Newman M E J. The structure and function of complex networks. *SIAM Review*, 2003, 45(2):167-256.
- [60] 袁明刚. 分形与网络自相似性特征的哲学意义. *武汉科技学院学报*, 2005, 18(2):53-56.
- [61] 李兴宁, 张小超. 浅论自然界的自相似性. *泰州职业技术学院学报*, 2004, 4(1):1-3.
- [62] Zhongzhi Zhang, Lili Rong, Francesc Comellas. High-dimensional random Apollonian networks [J]. *Physica A*, 2006, 363:567-572
- [63] Zhongzhi Zhang, Lili Rong. General model for Apollonian networks [J]. arXiv:cond-mat/0512229.
- [64] T. Zhou, G. Yan, B.H. Wang. Maximal planar networks with large clustering coefficient and power-law degree distribution [J]. *Physical Review E*, 2005, 71:046141.
- [65] ZM Gu, T Zhou, BH Wang, G Yan, CP Zhu, ZQ Fu. Simplex triangulation induced scale-free networks [J]. arXiv: cond-mat/0505175
- [66] 黄润生, 黄浩. 混沌及其应用(第2版). 武汉: 武汉大学出版社, 2005. 192-195.
- [67] Lali Barrière, Francesc Comellas and Cristina Dalfó. Fractality and the small-world effect in Sierpinski graphs[J]. *Journal of Physical A: Mathematical and General*. 2006, 39:11739 - 11753
- [68] 张济忠. 分形. 清华大学出版社, 2001. 42-46
- [69] 许雷, 许晓鸣等. 复杂网络. 上海: 上海科技教育出版社, 2006.

攻读硕士学位期间发表学术论文情况

[1] 那日萨, 穆青, 张书超. 谢尔宾斯基地毯的复杂网络特征. 系统工程学报 (中国系统工程学会会刊 (双月刊)), 已录用, 论文第 4 章).

[2] 那日萨, 张书超, 穆青. 具有分形和小世界特性的网络图. 系统工程 (国家自然科学基金会管理科学重要期刊, 2007, 25(3), 115-119, 论文第 5 章).

[3] 那日萨, 穆青, 张书超. 基于 Sierpinski 支架的具有分形和小世界特性的网络. 系统工程 (国家自然科学基金会管理科学重要期刊, 评审中, 论文第 3 章).

致 谢

在导师那日萨副教授的悉心指导下,本论文终于出稿了。回顾这两年的研究生生活,在这段生命历程中,导师的教导和帮助给予了我宝贵的财富,他为我指引了研究的领域和方向,教导了我学术研究的方法。导师平时工作很忙,但仍抽出宝贵的时间为我细心指导。从论文的选题到论文的成稿,导师为我排解了无数的困惑,提出了许多关键性的意见。师生之间建立的融洽关系也使我在学业上和生活上受益匪浅。在此,谨对导师的辛勤培养和关怀表示衷心的感谢。

同时我要衷心感谢郭崇慧副教授。在管理学院 404 的学习的那段时间里,郭老师给予我的关心和指导,为我硕士的学习研究打下坚实的基础,郭老师学术上的严谨、一丝不苟,生活上的平易近人都深深激励和感动了我。

同时我也要感谢大连理工大学应用数学系的其他所有老师,在数学系这个温暖的大家庭里,他们不仅让我在课堂上学到了丰富的知识,引领我步入了学术的殿堂。各位老师的言传身教让我深刻认识到做人和做学问应有的态度。

感谢师兄张书超,师姐孟灵犀,以及张显峰、李力、许传磊平时的支持和帮助,我们共同度过了一段美好的时光。

感谢我的好友骆庆开、张小路和孔昭坤,黄伟、白庆朝、付朝晖、窦永超、焦力宾、卢中亮、高爱花、云水静、吉玉霞、路龙凯、张中文等人,我们互相帮助、共同进步,建立了深厚的友谊,和他们一起度过的快乐时光和彼此的真挚友情,都将是我人生旅途中的唯美珍藏。

感谢我的家人及亲戚朋友们、马俊杰及其家人对我的支持和鼓励,我的研究生学习生活因此才得以圆满的结束。