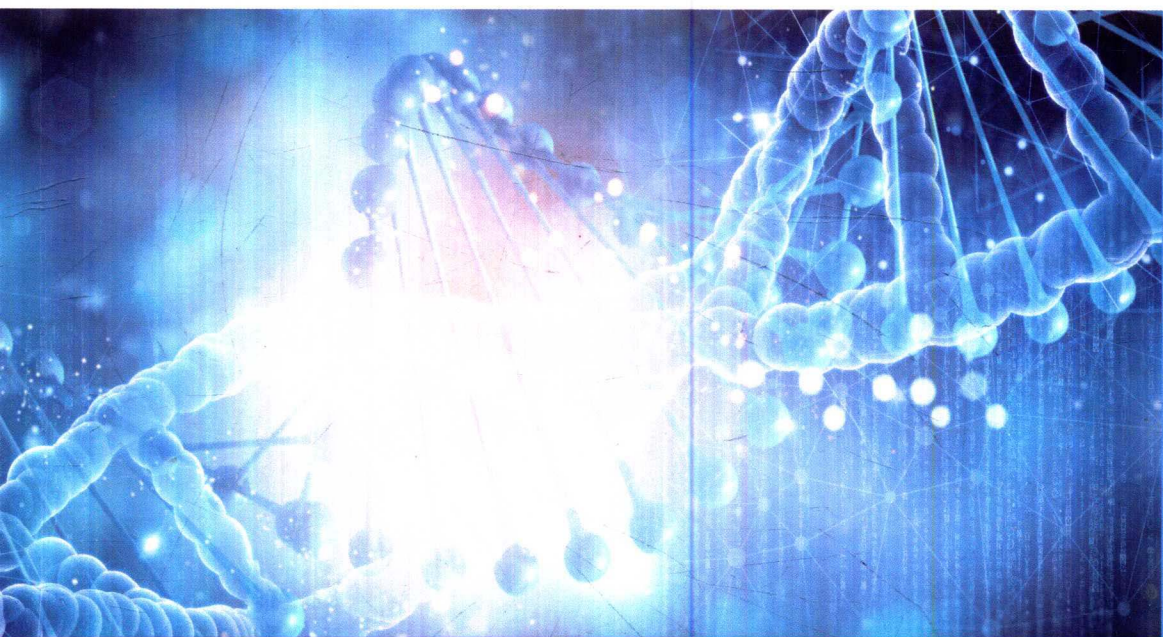


生物复杂网络中功能模块的挖掘

焦清局 ◎ 著



科学技术文献出版社
SCIENTIFIC AND TECHNICAL DOCUMENTATION PRESS

生物复杂网络中功能模块的挖掘

焦清局 著



科学技术文献出版社
SCIENTIFIC AND TECHNICAL DOCUMENTATION PRESS

· 北京 ·

图书在版编目 (CIP) 数据

生物复杂网络中功能模块的挖掘 / 焦清局著. —北京: 科学技术文献出版社, 2017. 11

ISBN 978-7-5189-3685-4

I. ①生… II. ①焦… III. ①智能机器人—智能模拟—研究 IV. ①TP242.6

中国版本图书馆 CIP 数据核字 (2017) 第 289207 号

生物复杂网络中功能模块的挖掘

策划编辑: 张 丹 责任编辑: 王瑞瑞 责任校对: 张咧噪 责任出版: 张志平

出 版 者 科学技术文献出版社
地 址 北京市复兴路15号 邮编 100038
编 务 部 (010) 58882938, 58882087 (传真)
发 行 部 (010) 58882868, 58882874 (传真)
邮 购 部 (010) 58882873
官 方 网 址 www.stdp.com.cn
发 行 者 科学技术文献出版社发行 全国各地新华书店经销
印 刷 者 虎彩印艺股份有限公司
版 次 2017 年 11 月第 1 版 2017 年 11 月第 1 次印刷
开 本 710×1000 1/16
字 数 160千
印 张 9
书 号 ISBN 978-7-5189-3685-4
定 价 39.00元



版权所有 违法必究

购买本社图书, 凡字迹不清、缺页、倒页、脱页者, 本社发行部负责调换

前 言

网络科学作为一门交叉学科，它的基本理论正渗透到数理科学、生命科学、工程科学甚至社会科学等众多学科中去。复杂网络的研究引起了世界不同领域科学家的广泛关注。对复杂网络的定性与定量特征的认识和理解是网络时代中一个重要而又具有挑战性的课题。作为复杂网络中一个重要特性，模块结构（或称社团结构）是一个重要而又普遍存在的结构特性。准确挖掘和分析模块结构对理解复杂网络的演化、结构和动态性都有着理论和实践的意义。

模块结构作为生物复杂网络中的功能模块组织形式，在生命科学领域中有着重要的意义。人们虽然提出很多有效的算法来分析功能模块，如基于图论的方法，基于随机游走模型和谱聚类方法，但是这些方法在算法层面和生物网络局限性上都存在一定的缺陷。面对这些问题，需要我们有针对性地研究并提出新的功能模块挖掘方法。

在本书中，我们主要研究了如何挖掘生物复杂网络中的功能模块，并探索了多样性的功能模块组织形式。首先，针对现有模块分析算法存在的缺陷，提出一种衡量网络中任意两个结点的新相似性 ISIM，依据这种新的相似性和层次聚类思想对生物网络中的功能模块进行挖掘，并利用新结点相似性进一步揭示了生物网络中蛋白质复合物的层次性和功能模块的多尺度性。为了避免生物网络不完备性带来的影响，我们通过融合多条件下基因共表达谱数据构建完备的基因共表达网络，进而分析功能模块。接下来，我们突破模块结构是生物复杂网络中功能单元的唯一组织形式这一概念，在生物网络上发现一种新的功能模块组织形式，Bi-sparse 功能模块。并相应地提出一种能同时挖掘高聚合和稀疏功能模块的二叉树搜索方法。最后，通过大规模网络的统计分析，深入研究了 Bi-sparse 模块的一些特性。因此，本书的研究内容主要包括以下几个方面。



①为了有效克服现有功能模块挖掘方法的缺点，我们使用受限的随机游走模型，提出一种新的转移概率矩阵，进而定义一种新的结点相似性 ISIM 来衡量网络中任意两个结点之间的距离。新的结点相似性有 3 个良好的特性：一是它能成功地融合网络的全局和局部拓扑信息；二是新结点相似性不仅能有效地衡量两个结点之间的距离，而且能捕捉到两个结点在网络中的拓扑结构；三是它是在一个收敛的空间定义结点相似性。因此，在一系列不完备和含有噪声的生物网络中，具有良好的稳定性和鲁棒性。

使用新结点相似性和层次聚类思想，可以有效地分析生物复杂网络中的功能模块。首先，我们使用新结点相似性产生网络的相似性矩阵。其次，使用层次聚类思想建立网络模块的树状结构。最后，选择合理的目标函数自动地挖掘网络中的功能模块。在此基础上，通过改变新结点相似性中的调节因子，本书又提出一种新的方法（ISIMB 方法）揭示生物网络中蛋白质复合物的层次结构和功能模块的多尺度特性。

与现有的模块挖掘方法相比，基于新结点相似性的方法是一个无参数的方法，它能自动地确定网络中模块的个数。使用它挖掘到的模块不仅与真实的功能模块结构获得更好的匹配，而且能有效克服生物网络不完备性的缺陷。与单尺度的方法相比，本书把模块多尺度概念引入到生物网络中蛋白质复合物和功能模块的挖掘，这种新的理念不仅能成功地预测蛋白质复合物及其层次特性，而且能从具体到一般的视角揭示功能模块的动态过程。

②针对生物网络的不完备特性和基因共表达的不传递性，本书提出一种新的方法检测基因共表达网络中的功能模块。这种方法首先融合不同条件下基因共表达谱数据构建完备的基因共表达网络，随后使用最大团算法挖掘网络中的功能模块。这种新的方法与其他方法相比，预测的结果有较强的生物功能相似性。通过转录和调控关系分析，预测功能模块中的基因有较高的概率被同一个转录因子所调控，从而为构建基因调控网络提供丰富的结果。

③传统的生物网络中功能模块的挖掘都是基于高聚合的模块结构是功能模块组织的唯一形式。然而，这个结论在生物网络中，特别是在蛋白质相互作用网络中存在可疑性。因此，我们发现一种与高聚合模块不同的 Bi-sparse 模块，然后结合二叉树理论和矩阵论提出一种新的方法（BTS 方法）来挖掘两种类型的功能模块。BTS 方法在蛋白质相互作用网络中挖掘的高聚合模块和 Bi-sparse 模块都组成功能单元。与其他方法相比，BTS 方法具有良好的性能：一是不需要预先设置模块的个数；二是挖掘的高聚合模块和 Bi-

sparse 模块都具有显著性的生物功能相似性。

④我们把高聚合模块和 Bi-sparse 模块作为功能单元的组织形式共存于同一网络中这一概念进行泛化。我们整理了 4 种类型共 25 个网络，用 BTS 方法对 25 个网络中的模块进行分析，结果发现：(a) Bi-sparse 模块具有普遍性。(b) 在社会网络中，Bi-sparse 模块中的人们充当着经纪人的角色，负责协调不同群体之间的矛盾，促进信息、技术和知识的交流等作用；在计算机软件网络中，Bi-sparse 模块中的结点具有相似的软件包属性；在生物复杂网络中，Bi-sparse 模块中的蛋白质或基因具有显著性的功能相似性。(c) 复杂网络中的 Bi-sparse 模块拥有一些特性：一是 Bi-sparse 模块和高聚合模块相比，Bi-sparse 模块含有的结点较少；二是 Bi-sparse 模块在不同类型的网络中，具有一定的偏好性；三是存在复杂网络中的 Bi-sparse 模块有两种明显的拓扑结构。

本书主要内容来自笔者的博士毕业论文，它详细介绍了笔者攻读博士期间在生物网络中功能模块领域所做的工作。本书重点描述了笔者所做工作的创新之处：提出了一种衡量网络中结点之间相似的新距离，并在此基础上，揭示了生物复杂网络中多尺度模块特性能有效揭示蛋白质的从具体到一般的生物功能。不仅如此，本书还揭示了单一的模块结构组织形式不能很好地分割网络结点的相同属性。虽然本书尽可能地介绍生物复杂网络中功能模块挖掘的各个方面的内容，但由于笔者水平有限，书中难免存在疏漏和不足之处，欢迎各位专家和读者批评指正。

本书的相关工作得到了河南省高等学校重点科研项目（17B520001、16B413001）的大力支持，在此表示衷心的感谢。本书最后列举了主要的参考文献，在此对所引参考文献中的作者和出版机构表示感谢。

目 录

第1章 绪 论	1
1.1 复杂网络	1
1.1.1 复杂网络的概念	1
1.1.2 复杂网络中的主要参数	4
1.1.3 复杂网络研究概况	5
1.2 复杂网络的模块结构	9
1.2.1 社团的定义及其研究意义	9
1.2.2 社团结构研究现状	10
1.2.2.1 谱聚类方法	10
1.2.2.2 分裂方法	11
1.2.2.3 合并方法	13
1.2.2.4 基于模块度的优化和扩展方法	15
1.2.2.5 基于随机游走模型算法	16
1.2.2.6 多目标优化方法	17
1.3 生物网络中模块结构研究进展	19
1.3.1 基于图论的方法	19
1.3.2 流体力学的方法	21
1.3.3 核连接方法	22
1.3.4 群体智能的方法	23
1.4 模块的多尺度性	24



1.5	生物网络中功能单元组织形式的研究	26
1.6	本书的研究内容	26
1.7	本书的结构和组织	29
第2章	基于收敛空间内结点相似性度量的社团分析	31
2.1	基于新结点相似性的社团分析	31
2.1.1	收敛空间内结点相似性的定义	31
2.1.2	新结点相似性的收敛性证明	35
2.1.3	新结点相似性的鲁棒性和稳定性	36
2.1.4	调节因子 α 的选取	39
2.1.5	层次聚类	40
2.1.6	算法时间复杂度分析	41
2.2	新结点相似性在挖掘网络社团中的应用	42
2.2.1	ISIM 算法在合成网络上的应用	42
2.2.2	ISIM 算法在真实网络上的应用	44
2.3	ISIM 算法捕捉局部和全局的网络拓扑结构	46
2.4	小结	47
第3章	生物网络中多尺度功能模块挖掘	48
3.1	新结点相似性揭示生物网络中模块的多尺度性	49
3.1.1	ISIMB 算法在标准数据集上的性能验证	50
3.1.2	蛋白质相互作用网络中复合物的模块多尺度分析	53
3.1.3	蛋白质和基因相互作用网络中功能模块的多尺度 分析	59
3.2	基于网络分割状态的重要模块结构筛选	62
3.3	单尺度模块挖掘方法和多尺度模块挖掘方法之间的关系	65
3.4	小结	67

第 4 章 多条件下基因共表达网络中功能模块的挖掘	68
4.1 完备基因共表达网络中功能模块的挖掘	68
4.1.1 多条件下基因共表达网络的构建	68
4.1.2 基于最大团的功能模块预测	69
4.2 预测模块的生物功能分析	72
4.2.1 预测模块的基因本体论分析	72
4.2.2 预测模块的示例分析	74
4.3 小结	76
第 5 章 蛋白质相互作用网络中稀疏和高聚合功能模块的共挖掘	77
5.1 生物网络中功能单元组织形式的研究背景	77
5.2 基于二叉树搜索的功能模块挖掘算法	79
5.2.1 二叉树搜索算法	79
5.2.2 三阈值的选取	83
5.3 BTS 方法在不同网络中的应用	84
5.3.1 BTS 方法在合成网络中的应用	84
5.3.2 BTS 方法在蛋白质相互作用网络中的应用	85
5.3.3 蛋白质相互作用网络中 Bi-sparse 模块功能分析	91
5.3.4 蛋白质相互作用网络中高聚合模块功能分析	93
5.4 小结	93
第 6 章 复杂网络的多样性功能模块组织形式探讨	95
6.1 BTS 算法性能测试	95
6.2 复杂网络中高聚合模块和 Bi-sparse 模块的意义	97
6.2.1 社会网络中 Bi-sparse 模块的意义	97
6.2.2 计算机软件网络中高聚合模块和 Bi-sparse 模块的 意义	98
6.2.3 基因共表达网络中高聚合模块和 Bi-sparse 模块的 意义	100



6.3	复杂网络中高聚合模块和 Bi-sparse 模块的共存性	101
6.3.1	高聚合模块和 Bi-sparse 模块的共存性	103
6.3.2	Bi-sparse 模块存在的合理性分析	104
6.3.3	Bi-sparse 模块的偏好性	106
6.3.4	Bi-sparse 模块的大小	107
6.3.5	Bi-sparse 模块的组织形式	108
6.4	BTS 方法中三阈值的选取	109
6.5	小结	110
第7章	总结和展望	111
7.1	总结	111
7.2	展望	113
参考文献		117

绪 论

当今社会，高新技术的创新和发展超过了以往任何一个时期。处在包括信息技术在内的高新技术发展的年代，自然世界和人类社会都得到了蓬勃发展。它们的发展需要人们拥有更多、更强的认识自然和改造社会的工具。面对纷繁复杂的自然界和人类社会，复杂网络的出现为人们研究现实社会提供了有力的武器。复杂网络是复杂系统最为常用的建模形式。复杂网络作为一个新兴的研究领域，它的基本理论与计算机科学、生命科学、数理科学甚至社会科学等众多领域紧密相关。在本章中，首先介绍了什么是复杂网络及其国内外的研究现状；然后主要介绍复杂网络中模块结构性质及其研究现状，以及模块结构在生物网络中的研究意义和研究现状；最后给出本书的主要研究内容和组织结构。

1.1 复杂网络

1.1.1 复杂网络的概念

网络 (network) 一般可以表示为一个由结点 (node) 集 V 和边 (edge) 集 E 构成的图 $G = (V, E)$ ^[1]，如图 1-1 所示。结点是某种具体事物的抽象，边对应着现实世界中事物与事物之间存在的某种关系。网络的研究是图论中研究的重点内容，而最早的图论研究可以追溯到 18 世纪著名的数学家欧拉 (Euler) 对大家熟知的“Konigsberg 七桥问题”的解决^[2]。欧拉利用数学抽象法把 4 块陆地抽象为 4 个结点，而 7 座桥抽象为连接陆地的 7 条线。进而“Konigsberg 七桥问题”的研究就转化为图论的研究。复杂网络的



研究和图论的发展紧密相连、一脉相承。

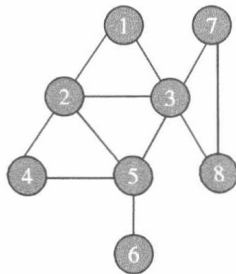


图 1-1 一个包含 8 个结点和 11 条边的网络

在现实世界中，处处有复杂网络的存在。例如，我们生活在一个庞大的社会系统中，人的社会属性决定我们在社会中和他人会发生各种联系和关系。在这个人类社会中，如果把社会中每个个体看作一个结点，而个体和个体之间的关系当作边，那么，我们生活的社会系统可以用各种属性的社会网络得到很好的刻画。例如，人们之间亲戚社会网络、朋友社会网络、商人之间的商业合作网络，研究人员之间技术交流及文献引用网络^[3]（图 1-2）。

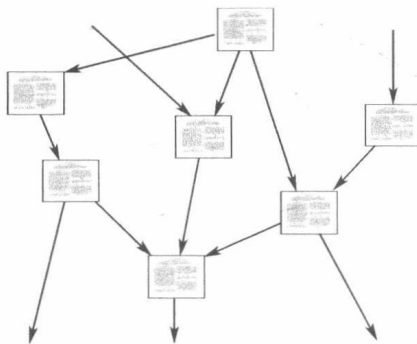


图 1-2 文献引用网络

以计算机和万维网（World Wide Web）为代表的高科技领域同样存在复杂网络。因特网（Internet）是计算机领域最为典型的一种网络^[4]，这种网络的结点是主机（host），边是主机和主机之间的物理连接。因特网的出现代表着一场革命的到来。它的出现使整个世界变得越来越小，地球一端可以随时联系到另一端的人们；也可以使分散在世界不同地区的人们迅速地共享资源。因特网的出现也促使了另外一种新的通信网络（communication network）：万维网。它由不同网页（web pages）之间的超链接构成，是迄今

为止人们构建的最大网络^[5]，并且还在逐渐增大的一个网络。与因特网不同，万维网是一个有向网络，网络中的结点代表的是网页，边是不同网页之间的超链接。因特网和万维网彻底改变了人们的思维方式：保守的思维从此被打破，资源和知识的共享达到了空前绝后的程度。不仅如此，万维网的大规模性为人们研究网络的拓扑结构提供了可靠的统计分析数据。

在生命科学领域，特别是在系统生物学（System Biology）领域存在多种多样的生物复杂网络。在复杂的生命系统中，蛋白质执行其生物功能时，它往往和其他蛋白质共同合作，这种合作通过蛋白质与蛋白质之间的物理相互作用来实现。蛋白质相互作用网络^[6-7]（protein-protein interaction network, PPI network）是一个包含了一系列蛋白质相互作用的集合。在这个网络中，结点代表蛋白质个体，边代表的是两个蛋白质之间的相互作用关系。图1-3^[7]给出一个蛋白质相互作用网络的例子。另外一种典型的生物网络是代谢网络^[8]（metabolic network）。代谢网络中结点代表的是代谢底物（metabolic substrates）或者产物（produces），而边代表的是底物或产物之间的代谢反映。代谢网络处于生命活动的末端，是生命体存在最为重要的化学引擎。代谢网络的重构对研究人们的重大疾病有着开拓性的意义。在基因调控网络中^[9-10]（gene regulatory network），结点可以是转录因子（transcription factor）或基因，而边表示转录因子和基因之间的调控关系。基因调控网络的研究，对药物研究有着非常重要的意义。一个引起疾病的突变基因受一个转录因子的调控，如果我们能够研制出割断转录因子调控基因的链条，那么就能对疾病有很好的控制。同样，基因和基因之间也存在共表达（co-expressed）现象，进而构成了基因共表达网络（gene co-expressed network）^[11-12]。与蛋白质共同合作执行生物功能一样，基因和基因之间通过共同表达来控制生命体的多样性。

另外，关系着国计民生的基础设施方面也存在很多网络，如电力网络^[13]、通信网络和交通网络^[2]。经济领域，存在经济网络和金融网络。总之，网络与人们生活、学习和工作密切相关。以上我们给出的只是复杂网络在现实世界中的冰山一角。其实，复杂网络无处不在，无时不有。复杂网络的研究有着重要的意义：第一，复杂网络的出现改变了人们的研究方式。对于一个复杂系统，以前的研究人员总是把复杂系统分解成单独的部分，进而对单独的部分进行详细的研究。人们认为只要对部分有很好的研究，那么就能对系统有全面的把握。然而，事实证明，这种研究方式并不能系统地了解

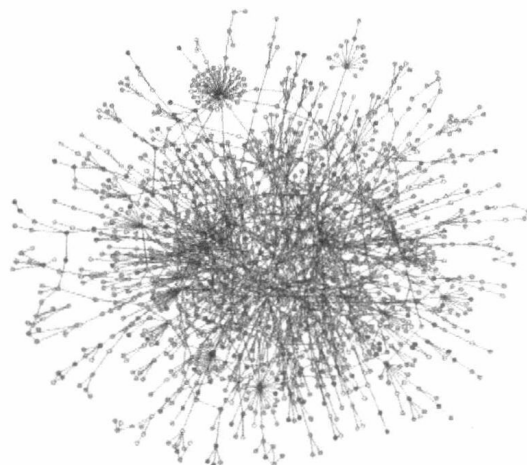


图 1-3 一个蛋白质相互作用网络

整体。甚至，部分研究得越详细，有可能对系统的研究带来相反的作用^[14]。而复杂网络可以很好地抽象复杂系统，进而可以从整体上展开研究。第二，当今的社会是一个信息大爆炸的时代，数据已经覆盖到社会的方方面面。大数据时代的到来，要求我们拥有强有力的工具处理大数据，进而认识自然界和人类社会。在这种形式下，复杂网络就成了我们研究大数据的先进设备。第三，复杂网络的基本理论渗透了计算机科学、数学科学和生命科学等多个领域，因此，它的研究也可以带动交叉学科的蓬勃发展。复杂网络理论的日益完善和其良好的特性推动着网络科学以一个前所未有的速度向前发展。

1.1.2 复杂网络中的主要参数

度 (degree) 是复杂网络中一个非常基本的参数，一个结点的度可以描述为这个结点的邻接结点数或者是这个结点连接边的个数。如果一个结点 i 的度表示为 k_i ，那么 $k_i = \sum_{j=1}^N w_{ij}$ ，如果结点 i 与结点 j 有边相连， $w_{ij} = 1$ ；如果两个结点没有相连， $w_{ij} = 0$ ， N 是网络中结点数。网络结点的平均度 k 可以部分地反映网络的稀疏程度。与度相关的另外一个概念是度分布 (degree distribution)^[3]。如果我们把结点度为 k 的数目占网络结点总数目的比例记为 P_k ，那么网络中结点不同度的统计分布即为度分布。不同的网络

一般情况下都具有不同的度分布。如随机网络和小世界网络服从 Poisson 分布,许多实际网络具有幂率分布。

最短路径 (shortest path) 和平均最短路径长度 (average shortest path length): 最短路径表示的是经过网络中两个结点之间最少的边数。最短路径在网络中的通信和运输方面都发挥着重要的作用。网络中任意两个结点之间最长的最短路径为网络的直径 (diameter)。一个网络的平均最短路径:

$$L = \frac{1}{N(N-1)} \sum_{i,j \in V, i \neq j} SP_{ij}$$
。结点之间的最短路径在网络中模块结构 (modular structure) 的形成过程中也起着重要的作用。

聚类系数 (clustering coefficient)^[1,15] 是用来描述一个结点的直接相邻结点之间的边的连接情况。一个结点 i 有 U_i 个邻接结点, 这 U_i 个结点之间存在 l_i 条边, 那么这结点的聚类系数为: $CC_i = \frac{l_i}{U_i \times (U_i - 1) / 2}$ 。整个网络的聚类系数是所有网络中结点的聚类系数的平均值。结点的聚类系数定义了网络的局部特性。与聚类系数相关的另外一个网络特性是聚集性 (clustering), 又可以称为传递性 (transitivity)。传递性描述的是一个结点的邻接结点的属性: $T = \frac{\text{包含结点 } i \text{ 的三角形的数量}}{\text{以结点 } i \text{ 为中心的连通三元组的数量}}$ 。

结点介数^[16] (betweenness) 和边介数^[17]: 给定一个结点 i , 它的结点介数是网络中任意两个结点之间通过该结点的最短路径的数目。我们可以把结点 i 的介数表示为: $b_i = \sum_{j,k \in V, j \neq k} \frac{n_{jk}(i)}{n_{jk}}$, n_{jk} 是连接结点 j 和 k 的最短路径的数目, $n_{jk}(i)$ 是连接结点 j 和 k 的最短路径, 并经过结点 i 的数目。结点介数来源于社会网络中评估个体重要性的指标, 它是衡量结点中心性 (node centrality) 的标准指标之一。在实际的网络中, 结点介数也可以用来衡量一个结点通信能力的好坏。边介数是结点介数的一种扩展形式。它的定义是通过某一条边的最短路径的数目。

1.1.3 复杂网络研究概况

虽然复杂网络的研究最早可以追溯到欧拉研究 “Konigsberg 七桥问题”, 但是在欧拉研究上述问题的很长时间后, 人们并没有对图论问题进行深入的



研究。到 1936 年，著名的匈牙利数学家 Erdos 和 Renyi 建立了随机图理论 (random graph theory)。他们研究的随机图模型 (ER 随机图) 是从数学的角度研究复杂网络的开始。在随机图模型中，图中任意两个结点都以相同的概率 p 相连，那么一个含有 N 个结点的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。由此可以得到，产生一个含有 N 个结点和 M 条边的 ER 随机图的概率为 $p^M (1-p)^{N(N-1)/2-M}$ 。掀起人们对复杂网络研究狂潮的是来自《Nature》和《Science》杂志上的两篇创新性的论文。一篇是 D.J. Watts 和 S.H. Strogatz 发表在《Nature》杂志上的有关网络小世界 (small-world) 的论文 “Collective dynamics of ‘small-world’ networks”，这篇论文深入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^[13]。另外一篇是 A.L. Barabási 和 R. Albert 发表在《Science》杂志上题为 “Emergence of Scaling in Random Networks” 的论文^[18]。它揭示了复杂网络的无标度 (scale-free) 性质。随后，世界各地的研究人员对复杂网络的研究工作如火如荼地展开了，学术成果层出不穷。小世界模型在生物网络和技术网络 (technological network) 中都有所表现，这些网络中都具有小的平均最短路径和聚类系数。最为明显的小世界模型是网络中的社团或者模块结构 (community or module structure)。除了 Watts 等人研究的小世界特性基本模型外，人们也研究了很多网络改进的模型，如 Newman 和 Watts 提出性能较好的 NW 小世界模型^[19]，这个模型更有利于人们对网络的分析。随着网络的无标度性质出现后，研究人员分析了很多不同类型网络的无标度属性的参数，网络的综述文章^[5]给出了这些统计的结果。随后，人们进一步研究了一些网络的模型^[5]，如随机图 (random graphs) 模型、静态无标度 (static scale-free) 网络和演化无标度 (evolving scale-free) 网络。最近几年，随着 Barabási 实验室发表在《Nature》杂志上的一篇文章 “Controllability of complex networks”^[20] 出现后，可控性也成为复杂网络的研究热点^[21]。与复杂网络属性相关的鲁棒性也吸引了人们的研究^[22-23]，并取得丰硕的成果。从 2002 年起，网络中的社团 (或者模块) 结构也成为网络属性的一个研究热点。本书也是从复杂网络中社团结构出发，进一步研究生物网络中的模块结构。

国内学者也较早地开始对复杂网络进行研究。2002 年上海交通大学的汪小帆在国外杂志发表了一篇有关复杂网络研究的文章^[24]，文中主要回顾了国外复杂网络研究的进展，其中包括网络的平均路径长度、聚类系数、度

分布等网络特性,并介绍了网络的一些模型,如规则网络、随机网络等网络模型,以及复杂网络的同步性问题。国内杂志上最早发表有关复杂网络的论文是朱涵2003年在《物理》杂志上发表的一篇有关网络建筑学的文章^[25],论文介绍了网络的产生、抽象、参数及网络的未来前景等问题。2005年,周涛等人撰写综述文章,着重阐述了网络中传播动力学的问题^[26]。李翔等认为择优连接机制不会在整个网络上都起作用,而只会在某个局域世界里发挥作用,通过引入局域世界的概念对BA网络模型进行推广,提出局域世界演化网络模型^[27]。研究表明,该模型保持了无标度网络的鲁棒性,同时改善了无标度网络固有的面对恶意攻击的脆弱性。陈庆华等为了进一步研究无标度网络的增长性和择优连接性,通过对网络结点进行重新连接建立BA网络模型的两个拓展模型。章忠志和荣莉莉提出BA网络的一个等价的演化模型^[28],在这个演化模型上,利用平均场近似方法,深入探索并得到了模型上网络的结点度分布和聚集系数,并对网络的平均路径长度进行数值模拟。结果表明,其演化成的结构特性与BA网络相同。2013年席裕庚发表文章阐述了大系统控制论和复杂网络之间的关系^[29]。2014年有众多的国内网络研究专家阐述了复杂网络研究的机遇和挑战^[30]。在中文书籍方面,2006年,汪小帆等人出版网络方面书籍^[2],该书籍首先介绍了网络基本理论知识和网络的基本模型及其性质。然后逐步介绍了网络的动力学、搜索和社团结构、同步和控制等相关问题。2011年,上海大学史定华编写网络专著教材^[31],该教材先从网络的建模和统计出发,紧接着重点介绍了网络的稳定性、结点的度及其相关的问题。2012年,汪小帆、李翔和陈关荣再一次出版了新的网络书籍《网络科学导论》^[1],该书籍详细介绍了复杂网络方面最新的研究进展和基础理论知识。

复杂网络可以详细地分为多种类型的网络,如社会网络、技术网络、经济与金融网络、生物网络等不同类型的网络。生物网络是复杂网络中研究比较深入的网路,因为生物网络的研究和生命科学紧密相关。用计算机、数学和物理等角度研究生物网络是生物信息学的研究范畴。生物信息学是一门交叉学科^[32-33],与复杂网络的研究一样,它的研究涉及了计算机科学、数理科学和物理科学等众多学科的内容。从模式识别的角度看待生物信息学,研究内容主要包括蛋白质组学、基因组学、代谢组学和系统生物学。而生物网络几乎渗透了生物信息学研究的各个领域,但是最为瞩目的成就体现在系统生物学(system biology)上。系统生物学是以整体的视角研究生物学,打破