

摘 要

在经典力学和经典电磁学中，研究的对象——运动物质，被想象为已经被分割为无限多个无限小，且在空间连续分布的基本单元的集合。这些基本单元被想象地放在规则空间中的各个规则格点位置上，因此使用许多年前数学家们创造的坐标体系就可以完善地描述所有基本单元的位置及其变化，即这个体系的运动。尽管各个基本单元的空间位置不同，而且一般来说在随时间变化，但是由于它们之间的相互作用遵从已经被认识的简单、普适基本法则，因此运用几百年前创立的微积分工具就可以非常简明地表示支配每一大类客观体系运动变化的普遍动力学规律。

然而，世界上的客观系统是规则、均匀分布的，全同的基本单元构成的，还是高度不规则、不均匀分布的，丰富多彩的基本单元构成的？基本单元的位置分布影响系统的动力学行为吗？基本单元之间的相互作用是遵从某种简单、普适法则，还是千变万化、错综复杂？这可能是物理学推广向复杂系统时要回答的首要问题。复杂网络应运而生，成为研究复杂系统的强有力工具。网络描述建立在简化描述模型的基础上，是对复杂世界的简化、抽象。这样的描述把基本单元看成是网中的一个“节点”，把基本单元之间的相互作用看作节点之间的“边”。在此基础上，提出了一系列的反映网络结构、特征的网络统计性质，如度分布、项目度分布、集群系数、度中心度、同类性等等。

复杂网络，尤其是社会网络的一个显著特征是，网络具有明显的群落、派系结构。本文的第二章主要研究了社会网络及类社会网络中的派系结构，提出了网络的新统计量—— k 方组项目度及其分布。并给出了简化的网络演化模型和解析，然后实证统计了十多个系统的二方组、三方组项目度及其分布，发现我们所研究的实际系统的二方组项目度分布和三方组项目度分布均为 SPL 分布，具有一定的普遍性，与模型得到的结论是一致的。

系统内部基本单元之间的相互作用是极其复杂多样的，若只考虑节点之间的相互合作关系，而忽略其他所有的相互作用（如竞争等关系），这样的网络称为合作网络，将主要反映网络中节点之间合作关系的统计性质，如项目大小、项目度

等，称为网络的合作性质。本文的第四章主要报道了四个实际系统的特殊合作性质，并给度分布与项目度分布的一致性提供了更多的实证。

然而，实际情况下，基本单元之间不仅存在合作关系，还存在着竞争关系，基本单元之间既合作又竞争是更为普遍的，像这样的网络称为合作竞争网络。既然合作竞争网络是普遍存在的，那么应该如何去描述基本单元之间的合作竞争关系？为此，我们提出了新的统计量——点权，它反映了合作竞争的结果。而且，通过我们的实证统计，发现实际网络的节点总权分布遵循 SPL 分布，也具有一定的普遍性。本文的第三章和第五章着重研究了合作竞争网络，及点权这一统计量在合作竞争网络中的统计规律。

[关键词]：合作网络，派系结构，特殊的合作性质，合作竞争网络，点权，SPL 分布，差异性

Abstract

In classical mechanics and classical electromagnetics, the investigation object, motorial matter, considered as been divided into infinitude parts, which tend to be infinitesimal and are consecutive distributed. Motorial matter is regarded as a collection of such parts. These basic parts supposed to be put on regular grid points in the regular space. Therefore, coordinate system, created by mathematician many years ago, can be used to describe the position and position change of all basic parts perfectly. Although every basic part's position in the space is different, generally changing with time, their mutual actions obey simple and universal recognized basic principles. So we can use calculus theory, founded hundreds of years ago, to show universal dynamic laws dominated these external systems.

However, are external systems regular or irregular, even distributed or uneven distributed? Are they composed of uniform basic cells or abundant disparate basic cells? Does the position distribution of basic cells affect system's dynamic behavior? This may be the first chief question to be answer when physics extend to complex systems. Complex network emerges as the times require and becomes a powerful tool to investigate complex systems. Network description based on simply described models. It's predigestion and abstract of complex world. Such kind of description regarded basic cells as "nodes" in the network and their mutual actions as "edges" between nodes. Then put forward a series of network properties, such as degree distribution, act degree distribution, clustering coefficient, assortativity and so on.

There is a very notable character in complex networks, especially social networks. That is this kind of networks have obvious community and clique structure. We mostly investigated clique structures of social and quasi-social networks in chapter three of this paper. A new statistical network property is proposed here. We called it k-clique act degree distribution. Then proposed a simplified network evolvement model and presented detailed analysis of this model. Finally, we did a lot of empirical statistical

work. Two-clique act degree distribution and three-clique act degree distribution of many real networks were presented in chapter two. Amazingly, nearly all the distributions we got followed shifted power law (i.e. SPL) distribution, which accorded with the model conclusion. We conferred this conclusion maybe universal in a certain extent.

As we know, mutual actions between basic cells in real systems are very complicated and various. If we just consider their collaboration and neglect all the other mutual actions, for example the competition relations, the networks are called collaboration networks in this way. And statistical network properties which mainly show collaboration relations between nodes in the network are called collaboration properties, such as act size, act degree and so forth. We have studied four real systems and their special collaboration properties in chapter four.

In fact, there is not only collaboration but also competition between basic cells in real systems. It should be more universal that collaboration and competition coexists. Such networks are called collaboration-competition networks in this paper. Since collaboration-competition networks exists universally, how should we describe this kind of collaboration and competition actions? Therefor, we proposed a new property called node weight, which reflects the competition result. Then we calculated total node weight distributions of more than ten real world networks. Surprisingly, all the gained distributions followed SPL distribution. This may be a universal law too. Chapter three and five mainly investigated collaboration-competition networks and the statistical law of node weight in ollaboration-competition networks.

[Keywords]: collaboration networks, clique, special collaboration properties, collaboration-competition networks, SPL distribution, discrepancy, node weight

扬州大学学位论文原创性声明和版权使用授权书

学位论文原创性声明

本人声明：所呈交的学位论文是在导师指导下独立进行研究工作所取得的研究成果。除文中已经标明引用的内容外，本论文不包含其他个人或集体已经发表的研究成果。对本文的研究做出贡献的个人和集体，均已在文中以明确方式标明。本声明的法律结果由本人承担。

学位论文作者签名：傅春花

签字日期：09年6月3日

学位论文版权使用授权书

本人完全了解学校有关保留、使用学位论文的规定，即：学校有权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文档，允许论文被查阅和借阅。本人授权扬州大学可以将学位论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存、汇编学位论文。同时授权中国科学技术信息研究所将本学位论文收录到《中国学位论文全文数据库》，并通过网络向社会公众提供信息服务。

学位论文作者签名：傅春花

导师签名：何大翔

签字日期：2009年6月3日

签字日期：09年6月4日

(本页为学位论文末页。如论文为密件可不授权，但论文原创必须声明。)

第一章 引言

经典物理学认为，研究对象（如：运动物质）可以被分割为无限多个无限小的基本单元（如：质点或电荷元）的集合，且这些基本单元在空间上是连续分布的。基本单元完全相同，它们之间的相互作用遵从已经被认识的简单、普适的基本法则，相互作用的迭加可以精确预言系统的运动，这就是传统的还原论的思想。但实际情况下，我们身边存在大量不适于或不能用还原论讨论、解释的系统。构成这些系统的基本单元并不是全同的，而是高度不规则、不均匀分布的，丰富多彩的。对于这些系统，还原论可能原则上适用，但影响因素太多、太敏感，也可能是原则上不适用的，因为此时基本单元的集合会“涌现”出分立个体不能展示的性质。总之，对于还原论不适用的系统，我们就不能用许多年前数学家们创造的坐标体系，来完善地描述所有基本单元的位置变化即该体系的运动。也无法用几百年前创立的微积分工具，简明地表示支配这类客观体系运动变化的普遍动力学规律，更无法准确地预言这些体系未来的行为。因此，对于这类复杂的系统我们需要寻找新的理论、新的工具进行研究和描述。复杂性科学应运而生。

复杂性科学是用以研究复杂系统和复杂性的一门方兴未艾的交叉学科。自上个世纪七八十年代开始，复杂性科学真正走入了国内外科学家的视野，成为一个新兴的研究领域。虽然它还处于萌芽时期，但已被有些科学家誉为是“21 世纪的科学”。

复杂性科学研究的复杂系统涉及的范围很广，包括自然、工程、生物、经济、管理、政治与社会等各个方面；它探索的复杂现象从一个细胞呈现出来的生命现象，到股票市场的涨落、城市交通的管理、自然灾害的预测，乃至社会的兴衰等等。目前，关于复杂性的研究受到了世界各国科学家们的广泛关注，许多杰出的科学家为复杂性科学的发展进行了不懈的努力。1999 年，美国《科学》杂志出版了一期以“复杂系统”为主题的专辑，这个专辑分别就化学、生物学、神经学、动物学、自然地理、气候学、经济学等学科领域中的复杂性研究进行了报道。概括起来，复杂系统都有一些共同的特点，就是在复杂多变的现象、活动背后，呈

现出某种捉摸不定的秩序,其中演化、涌现、自组织、自适应^[1]、自相似等被认为是复杂系统的共同特征。

近年来,复杂网络被认为是描述从生物到技术直至社会各类复杂系统的强有力工具^[2-3],复杂网络的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,引起了国际科学界的广泛重视,已成为统计物理学、随机图论、计算机网络、生态学、生物学以及社会学的研究热点,对复杂网络的定量与定性特征的科学研究已成为网络时代科学研究中一个极其重要的挑战性课题。

网络描述基于数学的一个分支—图论,传统图论中研究的许多问题,例如连绘图问题、最短路径问题、地图着色问题等,都可以归结为在一个二维平面或曲面上,把一些同一种类的点(即节点)用同一种类的线段(即边)连接起来的拓扑结构问题^[4-7]。在对复杂网络的研究中,随机图理论^[8]在近40年的时间里一直是研究复杂网络结构的基本理论,但绝大多数实际的复杂网络结构并不是完全随机的。例如,朋友关系,WWW上两个页面之间是否有超文本链接等都不会是完全靠抛硬币来决定的。在20世纪即将结束之际,对复杂网络的科学探索发生了重要的转变,复杂网络理论研究不再局限于数学领域。人们开始考虑节点数量众多、连接结构复杂的实际网络的整体特性,在从物理学到生物学的众多学科中掀起了研究复杂网络的热潮,甚至于被称为“网络的新科学^[9-10]”。

目前,复杂网络研究的内容主要包括:网络的形成机制,网络演化的统计规律,网络拓扑特性与模型,网络的演化动力学机制,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等问题^[2]。复杂网络的研究方法,是把复杂系统简化为节点以及连接节点的边的集合,其中节点代表系统中各个不同的个体,而边则用来表示个体之间的联系。有两篇开创性的文章可以看作是复杂网络研究新纪元开始的标志:一篇是美国 Cornell 大学理论和应用力学系的博士生 Watts 及其导师、非线性动力学专家 Strogatz 教授于 1998 年 6 月在 Nature 杂志上发表的题为《“小世界”网络的集体动力学》(Collective Dynamics of ‘small-world’ Networks)^[11];另一篇是美国 Notre Dame 大学物理系的 Barabási 教授及其博士生 Albert 于 1999 年 10 月在 Science 杂志上发表的题为《随机网络中标

度的涌现》(Emergence of Scaling in Random Networks)^[12]的文章。这两篇文章分别揭示了复杂网络的小世界特征和无标度性质，并建立了相应的模型以阐述这些特性的产生机理。

复杂网络可以分为技术网络、生物网络、社会网络等，其中社会网络是复杂网络中很重要的一类，它可以分为单模式网络（只有同一类节点，称为参与者(actor)）和双模式网络（有两类不同的参与者），甚至更多模式的网络。隶属网(affiliation network)是双模式网络中的最重要一种，其中一类节点是某种活动、事件或者组织的参与者（称为参与者节点(actor)），而另一类节点就是它们参与的活动、事件或者组织（称为项目(act)），通常用二分图^[13-14] (bipartite graph) 来描述这一类网络。

隶属网中的合作网络 (collaboration network)（即任意两个节点之间的边只表示节点之间的合作关系的网络，忽略节点之间的其他关系，例如竞争、支配等）引起了广泛的关注^[15-23]，如已被广泛研究过的科学家合作网和好莱坞演员合作网等。在合作网络的研究中，常常把二分图向一类节点（通常是参与者节点）投影，得到单模式网络。这时，参与每个项目的所有参与者节点之间都连有表示在此项目中合作关系的边，每个项目就表示为一个项目完全子图 (act complete subgraph)，整个单模式网络成为项目完全子图的集合。在此基础上提出了一系列的用来描述刻画网络的统计性质，更多的是拓扑统计性质，如度分布、项目度分布、平均最短道路长、最短道路长分布、平均集群系数、集群系数分布、项目大小分布、点强度分布和群落等。

然而，实际情况下节点之间的关系是有很多种的，除合作关系外可能还有竞争关系、支配关系等，甚至在一些复杂系统中，基本单元（节点）之间的合作关系是次要的、可以忽略的。合作网络的研究中完全忽略了这些关系，而仅仅关注节点之间的合作关系，这是过分简单的。节点之间只有纯粹的合作关系，或者纯粹的竞争关系等都是属于少数的极端情况。更多、更普遍的是节点之间的合作、竞争、支配等关系是同时存在的，且都不可忽略，这是显而易见的。我们把考虑节点之间的合作关系和竞争关系同时存在的网络称为合作竞争网络。接下来的问

题是，既然合作竞争网络是更普遍、更接近实际的，那我们又应该怎样去描述这样的合作竞争网络？以往包括度分布、集群系数分布等在内的传统的统计性质还能很好的刻画描述合作竞争网络吗？怎样去刻画？如果以往的统计性质都不适用了，那新的、更贴切的统计性质或新的统计量是什么？这都是我们亟待解决的问题。

本文的前半部分仍是致力于对合作网络的研究，包括集群系数对顶点度的依赖关系，以及合作网络中的派系结构，均给出了大量的实证结果及相应的网络演化模型，再好的模型也要有实证数据的支持，否则就是无源之水、无本之木。在对实证结果的总结及模型构造的基础上，对合作网络有了更进一步的认识。文章的后半部分着重阐述了我们与合作竞争网络的认识、理解和所做的简单研究，提出了几种可能的研究方法和途径，并提出了新的网络统计性质。希望对今后合作竞争网络的研究能有所启发。

参考文献：

- [1] Thilo Gross, Carlos J. Dommar D'Lima, and Bernd Blasius. Epidemic Dynamics on an Adaptive Network. *Phys. Rev. Lett.* 96, 2006, 208701
- [2] 汪小帆、李翔、陈关荣，《复杂网络理论及其应用》，清华大学出版社，北京，2006
- [3] 郭雷、许晓鸣等主编，《复杂网络》，上海，上海科技出版社，2006
- [4] 王朝瑞，《图论》，2004，第三版，北京：北京理工大学出版社
- [5] 卜月华，《图论及其应用》，2000，南京：东南大学出版社
- [6] 卢开澄、卢华明，《图论及其应用》，1995，第二版，北京：清华大学出版社
- [7] D. B. West (李建中、骆吉州译)，《图论导引》，2006，北京：机械工业出版社
- [8] Erdős P, Rényi A. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci.*, 1960, 5:17~60
- [9] Barabási A L. *Linked: The New Science of Networks*. Massachusetts: Persus Publishing, 2002

-
- [10] Watts D J. The 'new' science of network. *Annual Review of Sociology*, 2004, 30:243~270
- [11] Watts D J, Strogatz S H. Collective dynamics of 'small-world' networks. *Nature*, 1998, 393 (6684): 440~442
- [12] Barabási A L, Albert R. Emergence of scaling in random networks. *Science*, 1999, 286 (5439): 509~512
- [13] Petter Holme, Fredrik Liljeros, Christofer R. Edling, Beom Jun Kim. Network bipartivity. *Phys. Rev. E* 68, 2003, 056107
- [14] R. Lambiotte, M. Ausloos. N-body decomposition of bipartite author networks. *Phys. Rev. E* 72, 2005, 066117
- [15] Newman M. E. J., Scientific collaboration networks. I. Network construction and fundamental results, *Phys. Rev. E* 64, 2001, 016131; Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality, *Phys. Rev. E* 64, 2001, 016132
- [16] Barabási A. L., Jeong H., Neda Z. et al., Evolution of the social network of scientific collaboration, *Physica A* 311, 2002, 590
- [17] Bilke S. and Peterson C., Topological properties of citation and metabolic networks, *Phys. Rev. E* 64, 2001, 036106
- [18] Battiston S. and Catanzaro M., Statistical properties of corporate board and director networks, *Eur. Phys. J. B* 38, 2004, 345
- [19] Caldarella G and Catanzaro M., The corporate boards networks, *Physica A* 338, 2004, 98
- [20] Kou Z. and Zhang C., Reply networks on a bulletin board system, *Phys. Rev. E* 67, 2003, 036117
- [21] Newman M. E. J., Forrest S. and Balthrop J., Email networks and the spread of computer viruses, *Phys. Rev. E* 66, 2002, 035101(R)
- [22] Ebel H., Mielsch L-I. and Bornholdt S., Scale-free topology of e-mail networks,

Phys. Rev. E 66, 2002, 035103(R)

[23] Han D-D., Liu J-G, Ma Y-G, Cai X-Z. and Shen W-Q., Scale-free download network for publication, Chin. Phys. Lett. 21, 2004, 1855

第二章 合作网络中的派系结构

2.1 群落与派系

近年来, 复杂网络作为一门新兴的科学, 引起了人们广泛的关注^[1-2]。而在复杂网络的研究中, 社会网络越来越受到重视。社会网的一个显著的特征是基本单元(节点)具有群落结构(*community structure*), 群落中的节点间的联系比群落之间的节点间的联系密切得多^[3], 这表明一些活跃的、有个性的节点聚集在一起并构成了一个团簇。除了社会网络具有团簇结构, 本组其他成员(苏蓓蓓、常慧、张培培等)也对非社会网络中的团簇结构进行了一定的研究, 在这样的网络中, 网络中的基本单元(节点)并不活跃, 但是它们可能受到某些“控制因素”的操控, 并表现出间接的行为^[4-7], 这类网络我们称之为“类社会网络”, 如大家熟悉的交通网络中, 站点相当于“类参与者(*quasi-actors*)”, 世界语言网络中, 每种语言相当于“类参与者(*quasi-actors*)”, 等等一些人为的或人为控制的网络。

社会网络研究中, 一个大家都比较感兴趣的问题是群落结构是否能划分为一些基本的子图。这些基本子图包含了几个节点(两个或两个以上), 这几个节点在体现网络功能时总在一起, 如一部电影的男、女主角总是作为搭档出演一些著名电影中的夫妻、情侣等, 观众也习惯将他们看做一个固定的搭配, 如果这对男女主角中的其中一个和第三个演员饰演搭档, 观众可能会感到强烈的不满。在社会网络中, 一对无序的节点可以定义为一个二分体(*dyad*), 本章将讨论相互的二分体, 即两个节点之间都有指向对方的有向边, 我们最终用一条无向边来替代这两条有向边, 这条无向边仅表示这两个节点之间的联系。三分体的定义是类似的, 即三个节点相互之间都有连边。在社会网络中, 派系(*clique*)定义为含有三个或更多节点的最大完全子图。如果我们将派系的定义修改为“含有两个或更多节点的最大完全子图”, 那么, 我们就可以将二分体(*dyad*)称为2-派系, 将共存的三元组(*mutual triad*)称为3-派系, 类似的可以定义*k*-派系。

社会合作网络研究的一些有趣的研究成果已经发表, 如好莱坞演员合作网和科学家合作网^[8-11]。这样的合作网可以用二分图来描述, 在二分图中, 节点被划分

为两类，一类是参与者节点（参与了活动、事件、组织等），另一类即活动、事件、组织等，我们称之为“项目”。当我们只关注节点间的合作关系时，可将二分图向参与者节点投影，得到单粒子图。在投影图中，每个项目代表一个完全子图，项目中的任意两个节点之间都连接了边，每个项目完全图还可再分为更小的项目完全子图或派系，不同的项目完全子图可能共有一些相同的派系。一个派系参加的项目越多，从某种程度上说明它越重要。项目完全图中的节点总数成为项目大小（act size），用 T 表示。而项目度（act degree）则定义为节点参加了多少个项目，用 $h^{[4-7]}$ 表示。张培培、常慧、苏蓓蓓及她们的合作者研究了一些类社会合作网。

本章将着重关注社会网络及类社会网络中的一些重要的派系，参加了最大的项目并具有最大的 k -派系项目度（即 k -派系所参加的项目数），这样的项目可能是最重要的。因此， k -派系项目度（也称为 k 方组项目度）的研究将显得极为重要。第三节中，在非常理想和简化的情况下，我们构造了一个描述社会网络或类社会网络演化的模型。通过对模型的分析，我们给出了简化情况下 k 方组项目度分布的一般函数形式。通过数值模拟，我们发现 k 方组项目度分布的一般函数形式在更近实际情况下也是定性正确的。第四节给出了实际的类社会网络的实证统计结果，与模型得出的结论能很好的符合。同时，我们也找出了这些实际网络中最重要的二方组项目度和三方组项目度。

2.2 k -方组项目度

在二分图中，二方组项目度（即一对节点共同参加了多少个项目）用 D 表示，定义为： $D_{i,j} = \sum_m a_{i,m} a_{j,m}$ ，其中 i,j 是两个不同节点的标号， m 表示项目， $a_{i,m}$ 、 $a_{j,m}$ 均为二分图邻接矩阵元。 $a_{i,m}$ 、 $a_{j,m}$ 的定义为，若 $a_{i,m}=1$ ，表示节点 i 参加了项目 m ；反之，若 $a_{i,m}=0$ ，表示节点 i 没有参加项目 m 。类似地，三方组项目度（即由三个节点组成的共存三元组共参加了多少个项目），用 Tr 表示，定义为： $Tr_{i,j,k} = \sum_m a_{i,m} a_{j,m} a_{k,m}$ 。同样地， i, j, k 表示节点的标号， m 表示项目， $a_{i,m}$ 、 $a_{j,m}$ 、 $a_{k,m}$ 为邻接矩阵的矩阵元。以此类推，我们可以定义 k -方组项目度：

$q_{i_1, i_2, \dots, i_k} = \sum_m a_{i_1, m} a_{i_2, m} \dots a_{i_k, m}$ 。k-方组项目度分布 $P(q)$ 定义为网络中 k-方组项目度的值为 q 的出现几率^[12]。

2.3 模型

2.3.1 简化模型及解析

我们考虑最理想的简单情况，首先，每个项目包含的节点数是相同的，即项目大小 T 是常数，同时 $T = k \times n$ ，其中 n 是整数、 k 是一个常量，因此每个项目只能包含 n 个“合法的 k-派系”，所有参与者节点联合起来形成合法的 k-派系，网络中仅有 $k \times N$ 个参与者节点，因此，每个节点都参加且仅参加了一个合法的 k-派系，网络中没有孤立的节点。在网络的演化中，“合法的 k-派系”不会解散，而是作为一个固定的单元在演化。在每个时间步，加入一个新的“合法的 k-派系”，并按照一定的法则选择 $(n-1)$ 个旧的“合法的 k-派系”形成一个新项目。形成新项目之后，由于项目中的任意两个节点之间都有连边，所以就会出现一些与“合法的 k-派系”共用节点的“非法的 k-派系”。这里，我们只考虑模型的简单情况，计算时仅计算“合法 k-派系”的项目度。

首先，我们考虑在选择 $(n-1)$ 个旧的“合法的 k-派系”形成新项目时，与每个旧的“合法 k-派系”的 k-方组项目度值成正比，即 k-方组项目度值线性优选的法则。我们能得出与 BA^[1,9] 类似的结论，即“合法 k-派系”的 k-方组项目度分布 $P(q)$ 是幂函数形式。

其次，我们考虑完全随机选择 $(n-1)$ 个旧的“合法的 k-派系”形成新项目，并可以写出合法 k-方组项目度的演化方程，得出的结论是“合法 k-派系”的 k-方组项目度分布 $P(q)$ 是精确的指数函数形式。

为了表示出介于上述两种极端情况的中间情况^[4,5,13,14]，我们考虑在选择 $(n-1)$ 个旧的“合法的 k-派系”时，以一定的概率 p 随机选择，以 $(1-p)$ 的概率线性优选（即与 k-方组项目度值成正比）。类似的，当时间 t 趋于无穷大时，我们可以得到如下方程：

$$\frac{\partial q_i}{\partial t} = p \frac{n-1}{t} + (1-p) \frac{(n-1)q_i}{nt},$$

上述方程可以改写为:

$$\frac{\partial q_i}{\partial \ln t} = p \frac{T-k}{k} + (1-p) \frac{T-k}{T} q_i$$

方程的解为:

$$q_i = C_i t^{(T-k)(1-p)/T} - \frac{Tp}{k(1-p)}$$

其中, C_i 是积分常数, 通过条件 $q_i(t=t_i)=1$ 可以确定出 C_i 。令 $\alpha = Tp/[k(1-p)]$,

$\eta = T/[(T-k)(1-p)]$, 得出:

$$P(q_i < q) = P(t_i > t \left(\frac{q + \alpha}{1 + \alpha} \right)^{-\eta})$$

所以, 合法的 k -方组项目度分布 $P(q)$ 为:

$$P(q) = \frac{dP(q_i < q)}{dq} = \frac{\eta}{1 + \alpha} \left(\frac{q + \alpha}{1 + \alpha} \right)^{-\eta-1}$$

上述合法的 k -方组项目度分布函数称为“漂移幂律函数” (SPL) [5]。我们可以通过

上述方程检验极端情况, 当 $p=0$ 时, $\alpha=0$, $\eta=T/(T-k)$, 此时 $P(q) \propto q^{-\frac{kT-k}{T-k}}$; 当

$p \rightarrow 1$ 时, $\alpha \rightarrow \infty$, $\eta \rightarrow k\alpha/(T-k)$, 此时 $P(q) \propto e^{k(1-q)/(T-k)}$ 。所以, 我们得到 $0 < p < 1$

时介于指数分布和幂律分布之间的情况。当可调参数 p 从 0 连续的变化到 1 时,

k -方组项目度分布 $P(q)$ 从幂律分布连续的变化到指数函数分布。

2.3.2 更近实际情况的网络演化模型

在实际网络的演化过程中, 我们是无法区分“非法 k -派系”和“合法 k -派系”的, 所以忽略“非法 k -派系”的 k -方组项目度是不合理的。而且, 通常情况下项目大小 T

(项目中所包含的节点个数) 的值不会刚好是 $k \times n$, 因此有些项目可能只包含孤立的节点。在这一部分, 我们将考虑这两种实际情况。由于本组以前的研究结果表明, 项目大小 T 看成一个常数的简化是适用于许多实际网络的研究的 [4-5], 因此,

此处我们仍取 T 为常量。要想解析得出模型较复杂的情况（考虑实际演化情况在内时）是比较困难的，所以我们通过数值解析来说明。数值结果表明，考虑实际演化情况在内的结论去前面所得出的结论是定性一致的， $k=2$ 的结果已发表，见文献[15]，我们也做了 $k=3$ 的情况，结果是类似的。我们可以得出结论，在合作网络中，二方组项目度分布、三方组项目度分布均遵循 SPL 函数，意味着网络中的 k -方组项目度分布是很不均匀的，只有少数的 k -派系参加了较多的项目，因此可以将它们看成网络中的固定单元和重要的子图。而绝大多数 k -派系只参加了少数几个项目，由这些节点构成的派系并不是网络中的固定单元，在不同的项目中它们可能与不同的节点连边以表示合作关系，因此也不是网络中的重要子图。

2.4 实证统计

2.3.1 部分模型的解析与模拟仅仅是对一种固定的 k 值的情况，而模型的结论是适用于所有可能的 k 值的。然而，在实际的合作网络中，许多不同 k 值的 k 派系是共存的。若我们得到的实证结果，也显示不同 k 值的 k 方组项目度（如二方组项目度、三方组项目度）均遵循 SPL 函数，我们就能在一定程度上相信模型的结论是正确的，至少对于相当一部分比例的合作网络是正确的。下面是我们的统计得到的实证结果。

2.4.1 世界语言分布网

世界语言的分布情况引起了人们浓厚的兴趣^[16]。本组从复杂网络的角度，对世界语言及其分布进行了一定的研究。我们将各种语言所在的国家或地区定义为项目集，将所有的语言定义为节点集，在同一国家或地区使用的语言之间定义连边，这样构建起一个世界语言网络。数据来源于 2005 年出版的人类文化学^[17]第十五版，共列出了 6142 种语言、228 个国家和 8 个地区。图 2.4.1 是世界语言网的累计二方组项目度分布图，结果显示很好的幂律分布，从图中可以得到二方组项目度分布函数近似为 $P(D' \geq D) \propto D^{-5.9}$ 。若近似地认为数据量足够大且准连续，则通过积分与求导的关系可以推导出不累计时（原始的二方组项目度分布）遵循 $P(D) \propto D^{-6.9}$ ，这是 SPL 分布的一种极端情况（ $\alpha = 0$ ）。图 2.4.2 给出了世界语言

网的累计三方组项目度分布, 也显示较好的幂律分布 $P(T_r' \geq T_r) \propto T_r^{-9.6}$, 类似地, 在假设数据量足够大且准连续的情况下, 可得出累计三方组项目度分布遵循 $P(T_r) \propto T_r^{-10.6}$ 。

从我们的实证统计得到, 世界语言网的最重要的二方组如下: (1)Izora 和 Mangas (共同出现在 9 个国家或地区); (2)Mbulungish 和 Pular (共同出现在 8 个国家或地区); (3)Mann 和 Mbulungish(共同出现在 7 个国家或地区); (4)Mbulungish 和 Bainouk-Gunyu Adobe (共同出现在 7 个国家或地区); (5)Mann 和 Pular (共同出现在 7 个国家或地区)。最重要的三方组如下: (1) Mann 、 Mbulungish 和 Pular (共同参加了 7 个项目); (2) Mbulungish 、 Pular 和 Bainouk-Gunyu Adobe (共同参加了 7 个项目); (3) Mann 、 Mbulungish 和 Bainouk-Gunyu Adobe (共同参加了 6 个项目); (4) Mann 、 Pular 和 Bainouk-Gunyu Adobe (共同参加了 6 个项目); (5) Mann、 Nalu 和 Bainouk-Gunyu Adobe (共同参加了 5 个项目)。

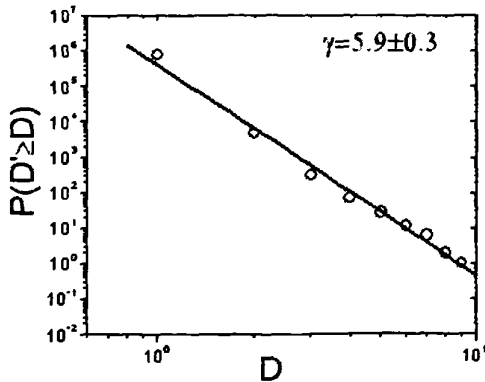


图 2.4.1 语言网二方组项目度分布

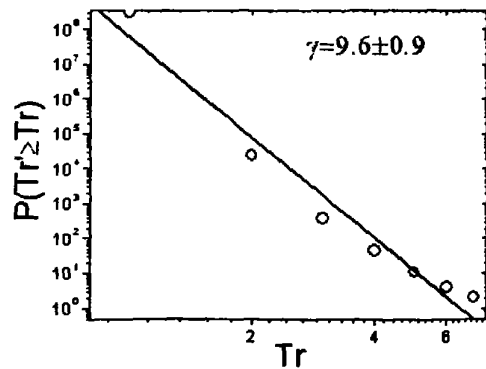


图 2.4.2 语言网三方组项目度分布

2.4.2 混合饮料网

在混合饮料网中, 每种混合饮料定义为项目, 饮料中的各种成分定义为节点, 出现在同一项目中的节点之间定义连边。混合饮料网的原始数据来源于 <http://www.drinknation.com>, 共统计得到 1501 种饮料成分和 7804 种混合饮料。

图 2.4.3 为混合饮料网的累计二方组项目度分布, 由图中可以得出 $P(D' \geq D) \propto (D + 20)^{-2.81}$, 遵循 SPL 分布。图 2.4.4 为混合饮料网的累计三方组项

目度分布也是 SPL 分布, 即 $P(T_r' \geq T_r) \propto (T_r + 3)^{-3.05}$ 。

混合饮料网的最重要的二方组如下: (1) Vodka 和 OrangeJuice (共同出现在 273 种饮料中); (2) OrangeJuice 和 PineappleJuice (共同出现在 201 种饮料中); (3) OrangeJuice 和 Grenadine (共同出现在 195 种饮料中); (4) Vodka 和 TripleSec (共同出现在 191 种饮料中); (5) Vodka 和 GranberryJuice (共同出现在 177 种饮料中)。

最重要的三方组如下: (1) OrangeJuice、PineappleJuice 和 Grenadine (共同出现在 66 种饮料中); (2) Vodka、OrangeJuice 和 Grenadine (共同出现在 59 种饮料中); (3) Vodka、OrangeJuice 和 PineappleJuice (共同出现在 57 种饮料中); (4) Vodka、OrangeJuice 和 CranberryJuice (共同出现在 56 种饮料中); (5) Vodka、Gin 和 TripleSec (共同出现在 53 种饮料中)。

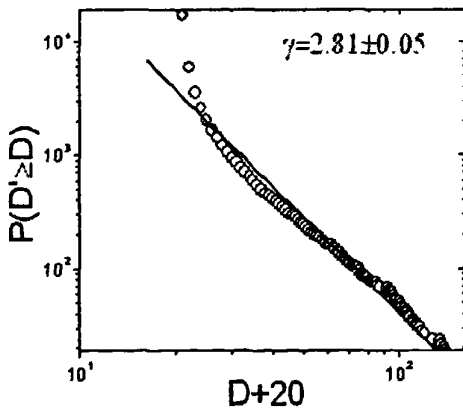


图 2.4.3 混合饮料网二方组项目度分布

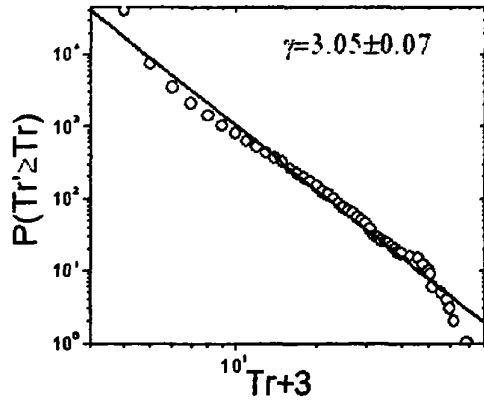


图 2.4.4 混合饮料网三方组项目度分布

2.4.3 IT 产品网

原始数据来源于 www.pcpop.com 和 www.it168.com 两个网站。在 IT 产品网中, 将各种 IT 产品定义为项目, 将产品的生产厂家定义为项目中的节点, 若两个厂家生产同一种 IT 产品, 则定义连边, 说明这两个厂家至少在一种 IT 产品市场上存在竞争。这样建立的 IT 产品网共包含 265 种 IT 产品和 2121 家 IT 产品生产厂家。图 2.4.5 是 IT 产品网的累计二方组项目度分布, 显示出很好的幂律分布, 此时 $P(D' \geq D) \propto D^{-3.7}$, 右边的图 2.4.6 是该网的累计三方组项

目度分布, 遵循 SPL 函数分布, 即 $P(T_i' \geq T_i) \propto (T_i + 3)^{-10.2}$ 。类似地, 可以分别近似推导出原始的二方组项目度分布和三方组项目度分布。

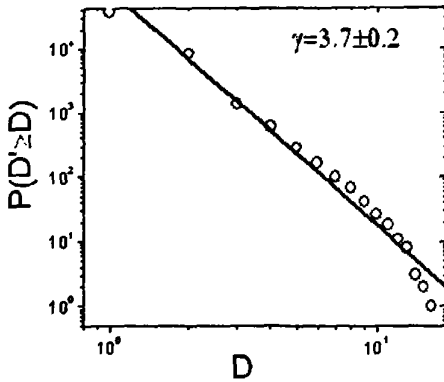


图 2.4.5 IT 产品网二方组项目度分布

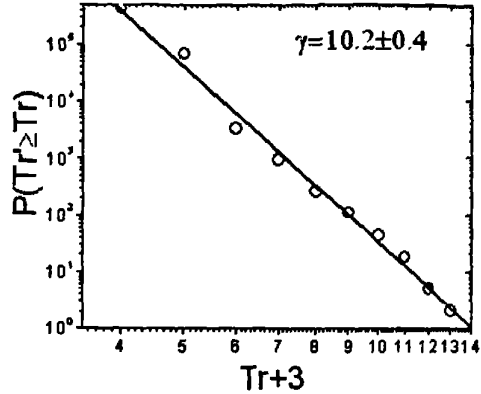


图 2.4.6 IT 产品网三方组项目度分布

IT 产品网最重要的二方组如下: (1) 大水牛和 MASTER (共同生产了 16 种 IT 产品); (2) 康冠和双敏 (共同生产了 15 种 IT 产品); (3) 三星和新中大 (共同生产了 14 种 IT 产品); (4) 三星和 Adobe (共同生产了 13 种 IT 产品); (5) MASTER 和三星 (共同生产了 13 种 IT 产品)。

最重要的三方组如下: (1) 大水牛、MASTER 和 TESSM (共同生产了 10 种 IT 产品); (2) 大水牛、MASTER 和三星 (共同生产了 10 种 IT 产品); (3) 大水牛、MASTER 和三星 (共同生产了 9 种 IT 产品); (4) 康冠、双敏和 Oneforall (共同生产了 9 种 IT 产品); (5) MASTER、三星和新中大 (共同生产了 9 种 IT 产品)。

2.4.4 其他实际合作网络的统计结果

除了上述三个网络外, 我们还实证研究了另外 9 个实际合作网络, 统计结果显示其中有 8 个网络的二方组项目度分布和三方组项目度分布都是 SPL 分布, 只有一个是例外, 以下是对这 8 个实际网络的简要介绍。

(1) 扬大公选课网^[7]: 节点: 121 门公共选修课; 边: 两门课同属于一个科学领域; 项目: 78 个科学领域。原始数据来源于扬州大学教务处。

(2) 中国专业培训机构网: 节点: 2674 门培训课程; 边: 两门课在同一个培训机构开设; 项目: 398 个培训机构。原始数据来源于 <http://www.ot51.com>。

<http://www.00100.cc>、<http://philosophy.cass.cn> 和 <http://www.people.com.cn> 等。

(3) 中国大陆电影网^[18]: 节点: 3085 部电影; 边: 两部电影具有相同的演员; 项目: 920 个著名演员。原始数据来源于 <http://soft6.com> 和 <http://www.mtime.com/movie> 两个网站。

(4) 2004 年奥运会网: 节点: 4496 名运动员; 边: 两个运动员参加同一个体育比赛项目; 项目: 229 个体育比赛项目。原始数据来源于 <http://2004.sina.com.cn/results/summary>。

(5) 中药网: 节点: 681 种中草药; 边: 两种中草药出现在同一副药方中; 项目: 1536 副药方。原始数据来源见参考文献[19]、[20]。

(6) 北京公交网: 节点: 4199 个公交站点; 边: 两个公交站点在出现同一条公交线路中; 项目: 1572 条公交线路; 原始数据来源于 <http://www.bjbus.com>。

(7) 旅游网: 节点: 171 个旅游景点; 边: 两个旅游景点出现在同一条旅游线路中; 项目: 240 条旅游线路。原始数据来源于 <http://www.cnta.com/8-ssls/lyqd.asp>。

(8) 淮扬菜网: 节点: 242 种食材; 边: 两种食材出现在同一食谱中; 项目: 242 种食谱。原始数据来源参见文献[21]。

在我们的实证研究中, 二方组项目度分布、三方组项目度分布不是 SPL 分布的唯一的例外是水果营养网^[22]。水果营养网的节点定义为一种营养成分(节点总数为 45), 项目定义为一种水果(项目总数为 151), 若两种营养成分包含在同一种水果中则定义为边, 水果营养网的原始数据来源于 <http://www.fumuqin.com>。通过我们的实证统计, 得到水果营养网的二方组项目度分布和三方组项目度分布均为正态分布。因此, 我们的实证统计结果强烈的支持模型得出的结论。

2.5 结论与讨论

我们通过一个非常简化的网络演化模型, 解析地给出合作网络的 k-方组项目度分布很可能都遵循 SPL 分布。并给出了我们的实证统计结果, 几乎所有被研究的系统的二方组项目度分布和三方组项目度分布都遵循 SPL 分布,

从很大程度上验证了我们模型的结论。实证结果表明, 在合作网络中(可能包含一些非社会网络)确实存在一些小完全子图。然而, 它们中仅有少数一些参加了许多项目, 所以可被看成网络中重要的基本单元。同时, 我们也发现有些实际网络的 k -方组项目度分布并不遵循 SPL 分布, 这些网络的动力学演化机制是值得进一步深入研究的。

参考文献:

- [1] Albert, R., Barabasi, A-L. Statistical mechanics of complex network. *Rev. Mod.Phys.* 74, 47–97 (2002)
- [2] Newman, M.E.J. The structure and function of complex networks. *SIAM Review* 45, 167–225 (2003)
- [3] Wasserman, S., Faust, K. *Social Network Analysis: Methods and Applications*. Cambridge Univ. Press, Cambridge (1994)
- [4] Zhang, P.P., Chen, K., He, Y., et al. Model and empirical study on some collabo-ration networks. *Physica A* 360, 599–616 (2006)
- [5] Chang, H., Su, B-B., Zhou, Y-P., He, D-R. Assortativity and act degree distribution of some collaboration networks. *Physica A* 383, 687–702 (2007)
- [6] Su, B-B., Chang, H., Chen, Y-Z., He, D-R. A game theory model of urban public traffic networks. *Physica A* 379, 291–297 (2007)
- [7] Fu, C-H., Zhang, Z-P., Chang, H., et al. A kind of collaboration-competition net-works. *Physica A* 387, 1411–1420 (2008)
- [8] Watts, D.J., Strogatz, S.H. Collective dynamics of 'small-world' networks. *Nature*, 393, 440–442 (1998)
- [9] Barabasi, A-L., Albert, R. Emergence of scaling in random networks. *Science* 286, 509–512 (1999)
- [10] Newman, M.E.J. Scientific collaboration networks. I. Network construction and fundamental results. *Phys. Rev. E* 64, 016131 (2001); Newman, M.E.J.

- Scientific collaboration networks. II. Shortest paths, weighted networks and Centrality. *Phys.Rev. E* 64, 016132 (2001)
- [11] Ramasco, J.J., Dorogovtsev, S.N., Pastot-Satorras, R. Self-organization of collaboration networks. *Phys. Rev. E* 70, 036106 (2004)
- [12] Krapivsky, P.L., Redner, S. Rate equation approach for growing networks. In: P-Satorras, R., Rubi, M., D-Guilera, A. (eds.) *Statistical Mechanics of Complex networks*. pp. 4. Springer-Verlag, Berlin, Heidelberg, 2003
- [13] Liu, Z., Lai, Y-C., et al. Connectivity distribution and attack tolerance of general networks with both preferential and random attachments. *Phys. Lett. A* 303, 337-344 (2002)
- [14] Li, X., Chen, G. A local world evolving network model. *Physica A* 328, 274-286(2003)
- [15] 常慧, 何大韧. 广义合作网络, 《复杂网络》第八章, 上海科技教育出版社, 2006. 11
- [16] Gomes, M.A.F., Vasconcelos, G.L., Tsang, I.J., Tsang, I.R. Scaling relations for diversity of languages. *Physica A* 271, 489-495 (1999)
- [17] <http://www.ethnologue.com/>
- [18] 刘爱芬, 付春花, 张增平, 常慧, 何大韧. 中国大陆电影网络的实证统计研究, *复杂系统与复杂性科学*, 第4卷第3期, 10-16
- [19] 朱玉祥, 《中医处方指南》, 金盾出版社(第二版), 北京, 1996
- [20] 刘东亮等, 《中医常用方剂手册》, 人民军医出版社, 北京, 1996
- [21] Compiling group of Beijing Nationality Restaurant. *Huai-Yang Bill of Fare*. Chinese Travel Publisher, Beijing, (1993) (in Chinese)
- [22] Qu, Y.Q., Jiang, Y.M., He, D-R. Fruit Nutritive Factor Network. *Jrl. Syst. Sci. and Complexity* 21, 1-9 (2008)

第三章 合作竞争网络的初步研究

3.1 引言

在复杂网络^[1-7]的研究中,尤其是在社会网络中^[8],合作网一直被认为是一个重要的研究课题。这类网络一般包含两类节点,项目(可以是某种活动、事件、组织等)和参与者(是指参与活动、事件、组织等的个体),像这样的网络可以用二分图很好的描述。在二分图中,关注的是不同类节点之间的连边,每条边都表示了一个隶属关系,即节点参与了项目。在合作网络的研究中,常常把二分图向一类节点(一般是参与者节点)投影,得到单模式网络。这时参与每个项目的所有参与者节点之间都连有表示在此项目中合作关系的边,而忽略了边所表示的所有其他关系(如竞争、监督、不同角色间的合作等)。例如,在被广泛研究过的好莱坞演员网^[1,2,6,7]和科学家合作网^[5,6,7]中,每两个演员或者每两个科研论文合作者之间的合作关系都看成是均等无差别的。然而,由于电影演员的角色和科研论文的作者实际情况下是有很大差别的,因此,认为演员之间的合作及科研作者之间的合作都是无差别的假设是过分简单的。

在投影后的图中,所有参加同一个项目的参与者之间都连接了相同的无权的边,因此每个项目都表示一个项目完全子图(ACSG)。在投影后的单模式网络中,每个 h 完全图都可以再划分为 m 个完全子图($m < h$),所以在投影后的网中,一个完全子图可能不是一个项目完全子图。在社会网络中,派系是指包含三个或更多节点的一个最大的完全子图^[8]。虽然,大多数项目完全子图都是派系,但是,通过实证研究我们发现,一些项目完全子图可能仅仅包含两个甚至一个参与者节点。例如,一篇科研论文的作者可能只有一个。一个项目完全子图也可能包含了更小的项目完全子图,而这个被包含的项目完全子图显然不是最大项目完全子图。如一篇科研论文的部分作者也可能合作写了另一篇论文。因此,项目完全子图不等同于派系。

在二分图中,一个项目完全子图包含了所有参加同一个项目的参与者节点,这是毋庸置疑的。实证数据已经包含了项目完全子图的信息,即每个项目完全子

图各自包含了哪些参与者节点。如图 3.1，是用来解释二分图及其投影图的一个示意图，其中 T1、T2、T3 表示项目，用数字 1 至 9 标记的是节点。在二分图中，只考虑不同类节点（即项目节点和参与者节点）之间的连边，投影时，这些边将被投影并形成了一系列的项目完全子图，如图中的节点[1-4]、[2-6]、[6-9]所构成的项目完全子图。

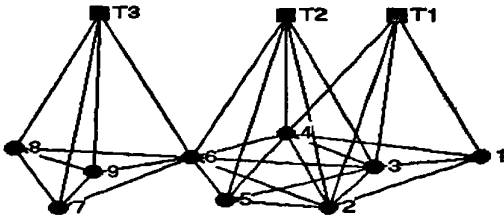


图 3.1 二分图及其投影图

大多数实际的合作网，如已被研究过的好莱坞演员合作网、中国旅游网、一些城市公交网、传统的中药网^[6,7]和中国的食谱网都不是社会网络。如果我们只关注节点之间的合作关系，如不同的药材构成了一副药方、许多景点组成一条旅游线路、多个站点组成一条公交线路等等，我们可以重新定义这些网络的节点（演员、公交站点、中药药材、旅游景点、食物素材）和项目（电影、公交线路、药方、旅游线路、菜肴），并让这些网络按图 3.1 投影，投影后的图就是一系列项目完全子图的集合。因此，我们仍可以称这些非社会网络为合作网络^[6,7]。

实际情况下，在复杂系统中，完全合作或者完全竞争是属于少数的极端情况。大多数情况下，基本单元之间同时存在合作和竞争。如何去描述竞争，这引起了大家很大的兴趣^[9,10]。在这儿，将简要介绍一下，我们所做的一些粗浅的工作。我们认为仍然可以用二分图及其投影图来描述这类合作竞争网络，在投影图中，一个项目完全子图仍指一个合作的项目，在项目完全子图中，所有节点之间都连了边。节点之间的连边则表示了它们之间的合作关系，为了描述节点间合作的地位等的差别，我们引入“点权”的概念。对于各种不同的实际网络，点权的定义可能而且应该是不相同的。例如，在电影演员合作网中，每个演员所得到的不同的报酬就可以定义为节点权，反映了他们在同一部电影中的不同地位或重要性。当然，每个电影演员都希望得到更多的报酬，所以，节点权就意味着一种竞争，我们称之为竞争因素。然而，所有的演员都明白只有大家都努力合作才能将电影演

好, 所以, 合作仍是基本的。以前所提出的没有考虑点权在内的网络的统计性质, 仍然能描述节点间合作的态势, 而新提出的包含点权在内的统计性质, 则可能反映了竞争的结果或其他一些竞争的性质。我们的兴趣在于这类网络的共同的统计性质, 尤其是包含权重的统计性质。

3.2 网络的统计性质

3.2.1 不含权统计性质

在网络描述中, 度分布^[2](degree distribution)一直深受重视, 因为它反映了网络的重要的拓扑性质, 并提供了网络演化的信息。例如, 无标度网的度分布是幂律分布, 隐含着网络的演化时是偏好连接的。某个节点 i 的度 k_i 定义为这个节点的邻边数: $k_i = \sum_{j \in \Gamma(i)} a_{ij}$, 其中 $\Gamma(i)$ 表示节点 i 的邻点集, a_{ij} 表示邻接矩阵元。若 $a_{ij}=1$ 表示节点 i 和节点 j 是连边的, 若 $a_{ij}=0$ 表示节点 i 和节点 j 是不连边的。而在合作网络中, 用来描述节点是怎样参加项目的统计量更为重要, 这些统计量中其中之一就是有多少节点参加了同一个项目完全子图, 我们称之为项目大小(act size), 用 T 表示: $T_j = \sum_i b_{ij}$ 。其中 b_{ij} 是二分图的邻接矩阵元, 若节点 i 参加了项目 j , 则 $b_{ij}=1$; 反之亦然。另一个重要的统计量则是节点的项目度(act degree), 反映了一个节点共参加了多少个项目, 用 h 表示: $h_i = \sum_j b_{ij}$ 。这儿, 我们用 $p(k)$ ^[11] 表示网络中 degree 为 k 的节点的个数, 以此来表示一个节点的 degree 为 k 的概率, $p(h)$ 、 $p(T)$ 也是如此。

在几乎所有被研究过的合作网络中, 节点之间的边都是不含权重的等同的边, 仅仅表示了它们之间的合作关系^[1-6], 不能反映出不同节点对之间的合作的差别。而实际上, 节点对之间的合作有强有弱, 是不等同的。一对节点共同参加了多少个项目, 可以作为边权来反映它们之间合作的强弱程度^[12-18]。例如, 若两个电影演员共同出演了 20 部电影, 则 20 可以作为它们之间连边的边权 w_{ij} 。所以, 在含边权的网络中, 节点 i 的度应该用点强度来表示 s_i 来表示, 定义为: $s_i = \sum_{j \in \Gamma(i)} w_{ij}$ 。

其中 w_{ij} 是含边权网络的邻接矩阵元，点强度分布则是考虑了合作强度在内的拓扑性质。

3.2.2 含权统计性质

我们引入节点点权来表示一类网络的竞争结果，用 w_n 表示， $p(w_n)$ 则表示节点点权分布，与前面说的 $p(k)$ 、 $p(h)$ 等定义类似。节点点权分布可能反映了网络中节点间竞争的重要信息。例如，若节点点权分布 $p(w_n)$ 是幂律分布，如已报道的很多国家的财富分布^[19,20]，这很可能表示节点间的竞争是很激烈的，而且竞争的结果依赖于个人的努力等。因此，只有少数的节点占有了竞争的大部分结果，而大量的节点所分得的竞争结果是很小的。若节点点权分布 $p(w_n)$ 是正态分布，如某大学学生的身高分布，这很可能反映了节点之间的竞争是很弱的，并且竞争的结果不是由个人的努力所决定的，而是依赖于某种概率，所以大多数节点的竞争结果都在平均值附近。

那么，节点点权与节点的项目度是否有内在的联系？是否节点的项目度越大（即节点参加的项目越多）节点的点权也越大？这也是一个值得探究的问题。例如，如果在含权的电影演员网中，节点点权与节点的项目度是正相关的，这意味着平均来看，一个演员演的电影越多所得的报酬也越多。由于点强度分布规律与项目度分布规律通常都相同（而度分布与项目度分布可能形式不同），因此我们相信，点权与点强度的统计规律一定和点权与项目度的统计规律是一致的^[7]。

3.2.3 漂移幂律分布

在合作网络中，我们提出度分布、项目度分布和点强度分布通常都是漂移幂律分布^[7](Shifted Power Law)，简称为 SPL 分布。我们以度分布为例，则度分布的 SPL 函数表达式应写为： $p(k) \propto (k + \alpha)^{-\eta}$ ，其中 α 和 η 都是常量。当 $\alpha = 0$ 时， $p(k) \propto k^{-\eta}$ ，这就是幂函数分布；而当 $\alpha \rightarrow \infty$ 时， $p(k) \propto (k + \alpha)^{-\eta}$ 近似为指数函数分布，所以 SPL 分布是介于指数分布和幂律分布之间的一种分布，幂律分布和指

数分布是 SPL 分布的两个极端。由前可知 $0 < \alpha < \infty$ ，当 α 的取值由 0 逐渐趋向于无穷大时，SPL 分布由幂律分布逐渐变化为指数分布。实际情况下，当 α 的取值大于 100 时，SPL 分布就接近于指数分布；当 α 的取值小于 1 时，SPL 分布就接近于幂律分布。所以，典型的 SPL 分布是当 α 的取值介于 1 至 100 之间时情况。

3.3 实证研究

这里，我们以两个实际网络——扬州大学公选课网和扬州大学物理类图书借阅网作为简单的例子，来讲述这种含点权的描述合作竞争网的方法及这两个实际网络相应的实证统计结果。

3.3.1 扬州大学公选课网

扬州大学是 1992 年由七所学校合并的，虽然它合并办学的历史不长，但是近几年的发展很快。与其他高等院校相同，扬州大学实行学分制，在校学生除了学习一些院内开设的必修课程，还需选修学校统一开设的一些公共选修课。在 2002-2006 年之间，学校共开设了 121 门公共选修课，这些选修课涵盖了自然科学和人文社科两个方面，如数学、物理、文学、艺术、体育等，每个学生必须从中选修至少四门课程。我们的数据来源于扬州大学教务处，共包含了 65536 条记录，每条记录表示一个学生的选课情况。我们将每门公共选修课程定义为网络的节点，将系科定义为网络中的项目，若两门同属于一个系科门类，则定义连边。一个节点可能参加不止一个项目，因此将各个项目完全子图连接起来，这样构建成一个扬州大学公选课网。

图 3.3.1 是扬州大学公选课网的项目大小累计分布图，呈一较平滑的半支泊松分布。这表明，隶属于各个系科门类的课程数虽然是在变化的，但都在平均值附近。图 3.3.2 是扬大公选课网的度累计分布图，从图中可以看出度累计分布呈线性规律，这与图 3.3.4 所显示的项目度累计分布规律是截然不同的。扬大公选课网的项目度累计分布在单对数坐标上呈一条直线，表明项目度累计分布遵循指数函数规律。

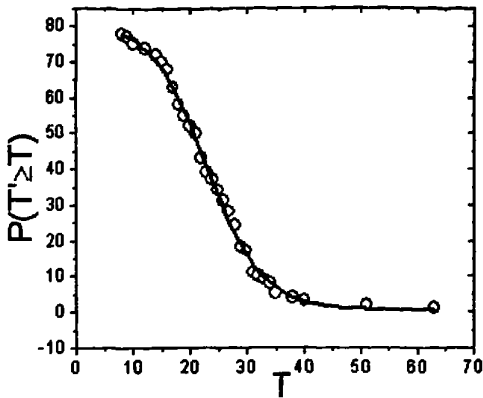


图 3.3.1 扬大公选课网项目大小累计分布图

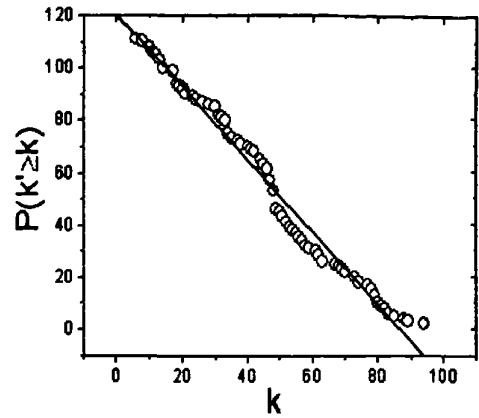


图 3.3.2 扬大公选课网度累计分布图

图 3.3.3 是扬大公选课网的点强度累计分布图, 与项目度累计分布规律类似, 也是在单对数坐标上呈一直线即遵循指数函数规律。与文献^[7]中所得出的, 同一网络的项目度分布规律与点强度分布规律是一致的结论相符。

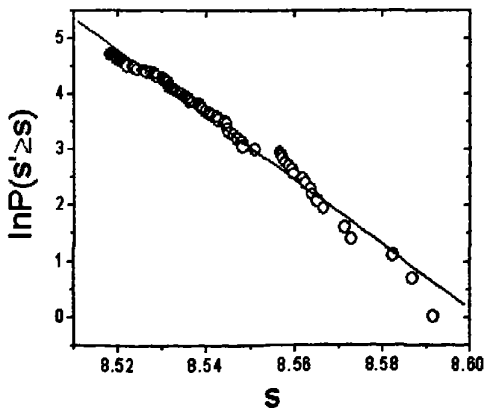


图 3.3.3 扬大公选课网点强度累计分布图

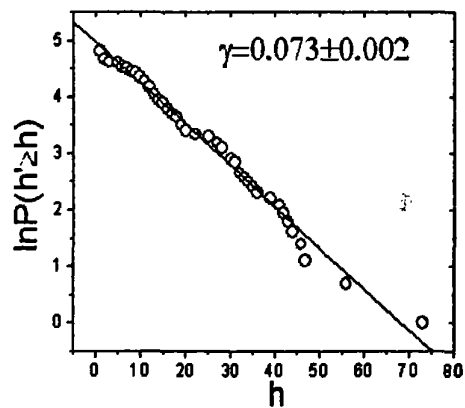


图 3.3.4 扬大公选课网项目度累计分布图

图 3.3.5 和图 3.3.6 分别是扬大公选课网的二方组项目度累计分布图和三方组项目度累计分布图。从图中可以看出, 二方组项目度和三方组项目度均满足幂律分布, 即 SPL 分布的一种极端情况, 这与前面第三章中得出的结论是一致的。

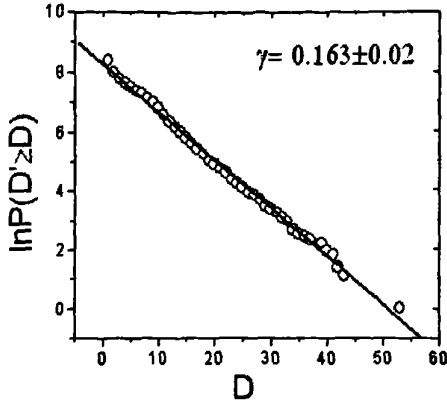


图 3.3.5 扬大公选课网

二方组项目度累计分布图

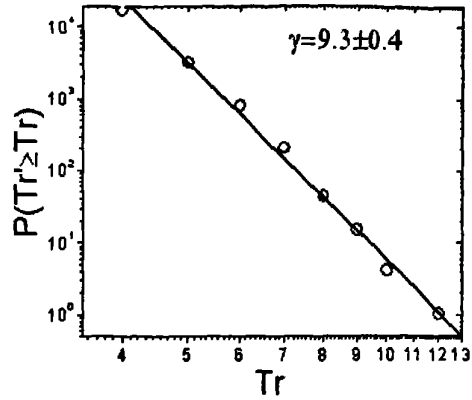


图 3.3.6 扬大公选课网

三方组项目度累计分布图

图 3.3.7 至图 3.3.9 分别为扬大公选课网的集群系数对顶点度、点强度和顶点项目度的依赖关系。由图可知，集群系数对顶点度的依赖关系满足较好的泊松函数关系，而集群系数对点强度和项目度的依赖关系，在单对数坐标上呈现一条很好的直线，即均满足很好的幂函数关系。

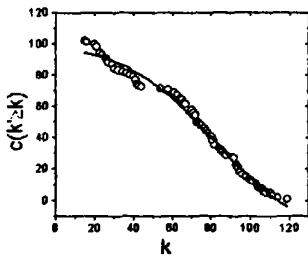


图 3.3.7 集群系数对顶点度的依赖关系图

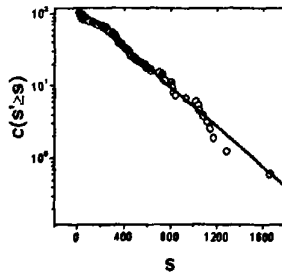


图 3.3.8 集群系数对点强度的依赖关系图

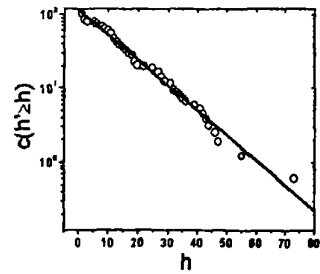


图 3.3.9 集群系数对项目度的依赖关系图

前面已经提及，扬大公选课网的节点是每门课，项目是系科，点权定义为选择每门课程的人数。选择某门课的人数越多，意味着这门课越受欢迎，从某种程度上说明，这门课在竞争中更胜一筹。所以，将选课人数定义为课程的点权具有一定的合理性，新的含权统计性质也是我们所关注的。图 3.3.10 是扬大选课网的点权累计分布图，从图中我们可以得出结论，点权分布是很好的幂律分布规律。这意味着，各门公共选修类课程之间的竞争是很激烈的，任课教师也在不断的努

力提高所教课程的质量以吸引更多的学生。然而, 由于学术能力和各自的努力有很大的差别, 所以, 竞争的结果是仅有少数课程吸引了大量的学生, 而大多数课程的选课人数是较少的, 导致了最终的幂律分布。图 3.3.11 显示的是扬大公选课网点权和项目度的相关性, 我们发现在选课网中点权和项目度是呈正相关的。这说明, 平均来看, 一门课程所涉及的系科门类越多, 它所吸引的学生也将越多。

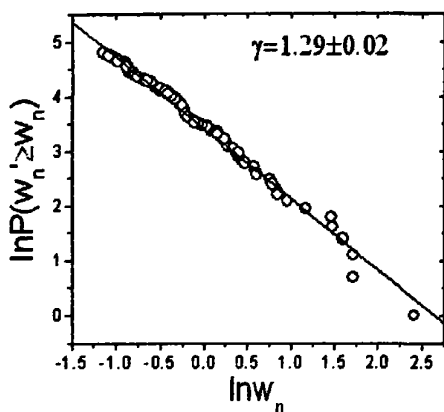


图 3.3.10 扬大公选课网

点权累计分布图

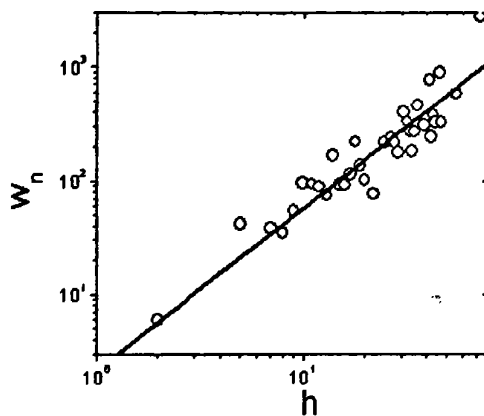


图 3.3.11 扬大公选课网

点权和项目度的相关性

3.3.2 扬州大学物理类图书借阅网

我们对扬州大学图书馆物理类图书的借阅情况(直至2006年)进行了实证研究。根据我们的统计结果, 共有 15204 册物理类图书, 其中有些书的书名是相同的(例如, 对于一些重要课程, 通常有 30 甚至更多本相同书名的参考书), 还有一些书几乎从来没有被借过(如一些刚进馆不久的新书、很偏僻的研究领域的书籍等)。考虑到这些情况, 我们共统计了 3207 册均有借阅记录的图书, 共涉及到 227 个科学领域(如电磁学、力学等)。一本书可能涉及多个研究领域, 在 88981 条借阅记录的基础上, 我们建立了一个物理类图书借阅网, 将每本书定义为节点, 研究领域定义为项目, 若两本书涉及同一个研究领域(意味着这两本书都能给学生提供这个领域的一些必要的知识)则定义连边。正如前面所说, 一本书可能涉及多个研究领域, 因此将孤立的项目完全子图连接起来, 从而构成整个物理类图书借阅网。如果一本书被频繁借阅了很多次, 说明这本书是很受欢迎的, 对学生的学习起着较重要

的作用，因此我们可以将每本书的借阅次数定义为它的点权。根据我们的统计，所有这些书中，点权最大的有 1541，最小的是 1。下面是该网络的一些实证统计结果。

图 3.3.12 和图 3.3.13 分别显示了物理图书借阅网的度累计分布和项目度累计分布情况，均呈指数函数分布，具有一致性。由累计与不累计的关系，我们可以近似推导出不累计情况下，物理图书借阅网的度分布和项目度分布仍满足指数函数关系。

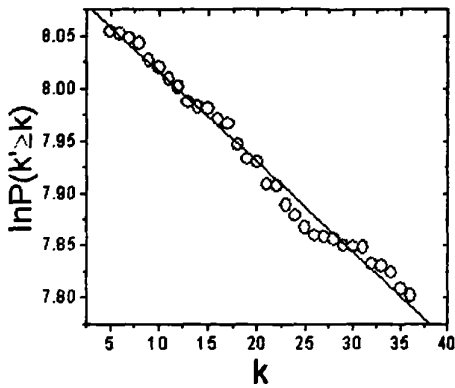


图 3.3.12 物理图书借阅网度累计分布图

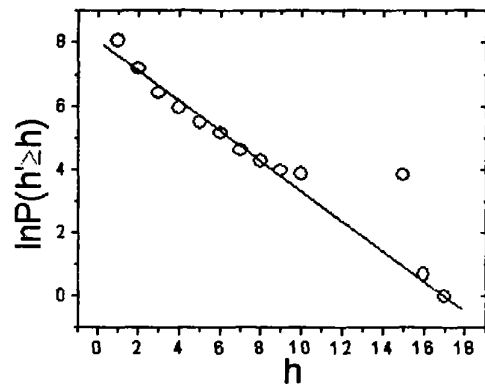


图 3.3.13 物理图书借阅网项目度累计分布图

图 3.3.14 是物理图书借阅网的点强度累计分布图，我们发现点强度分布与项目度分布的规律仍是定性一致的，仍是指数函数分布。图 3.3.15 是物理图书借阅网的项目大小分布情况，与以往报道的实际网络的实证统计结果有很大不同的是，物理图书借阅网的项目大小是呈指数衰减分布的，而以往报道的实际网络的项目大小分布大多是单峰（正态）分布。这表明，就物理学来说，少数领域的图书很多，是大量的，而绝大多数领域的图书是少量的，分布很不均匀，因此呈指数衰减，这或许能反映出科学领域之间和图书之间存在着竞争。

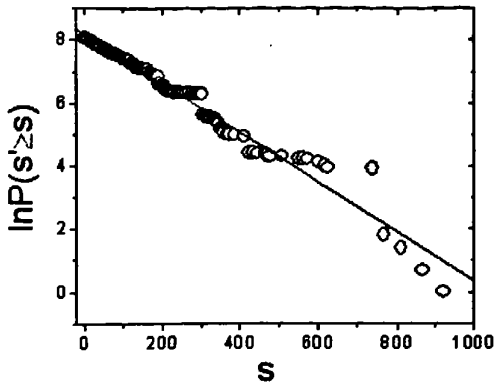


图 3.3.14 物理图书借阅网点强度累计分布图

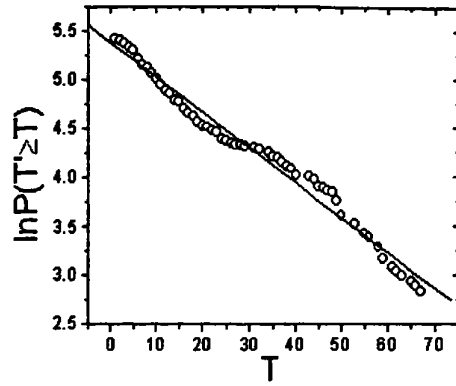
图 3.3.15 物理图书借阅网
项目大小累计分布图

图 3.3.16 是物理图书借阅网的二方组项目度累计分布图，呈现出较好的幂律分布，即 $P(D' \geq D) \propto D^{-2.3}$ 。图 3.3.17 是该网的三方组项目度累计分布图，从图中可以看出，三方组项目度的值只有 1、2、3、4 四个，可以说物理图书借阅网的派系结构是不明显的。

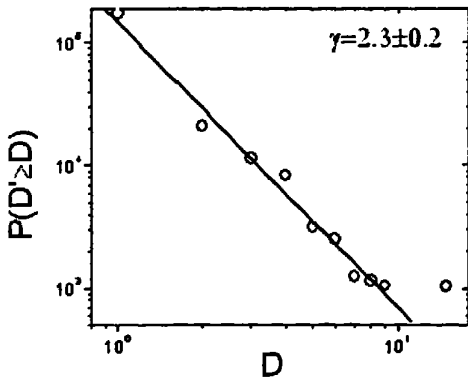
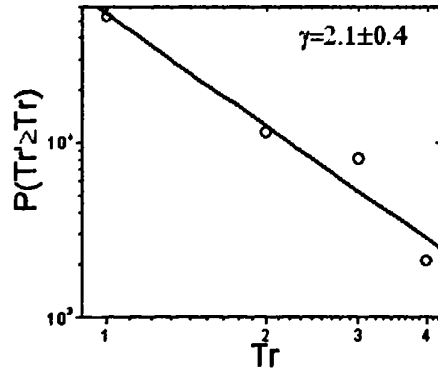
图 3.3.16 物理图书借阅网
二方组项目度累计分布图图 3.3.17 物理图书借阅网
三方组项目度累计分布图

图 3.3.18 至图 3.3.20 为物理图书借阅网的集群系数对顶点度、点强度和项目度的依赖关系，由图可知，集群系数对顶点度的依赖关系仍满足泊松函数关系，而集群系数对点强度和项目度的依赖关系仍近似地满足幂函数关系。

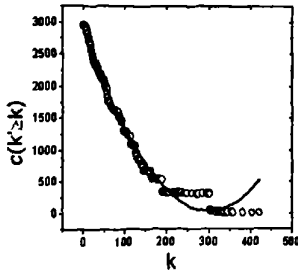


图 3.3.18 集群系数对顶
点度的依赖关系图

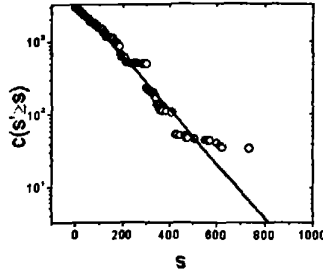


图 3.3.19 集群系数对点
强度的依赖关系图

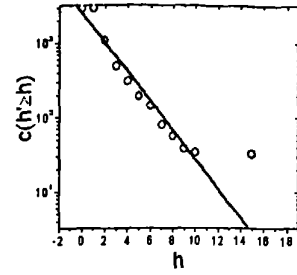


图 3.3.20 集群系数对项目
度的依赖关系图

前面已经提到，物理图书借阅网的节点是每本书，节点的点权是每本书的借阅次数，反映了这本书的受欢迎程度。由于每本书的借阅次数各不相同，从某种程度上说明所有的书本相互之间存在着竞争关系。一本书的点权越大，说明该书越受读者欢迎，或者说该书在竞争中获得的竞争结果越大。图 3.3.15 显示的是物理图书借阅网的点权累计分布情况，由统计结果我们可以看出，图书借阅网的点权分布满足 SPL 分布，其 SPL 参数 α 值为 5，由 3.2.3 的内容可知，该 SPL 分布是更接近幂律的。这说明物理图书借阅网中，每个节点的点权分布是极不均匀的，少数图书的借阅次数很多，拥有很大的节点点权，这些书很可能是一些考试用的参考书、习题册等。而相比之下，大量图书的借阅次数都远不及那些少数图书，即大多数节点的节点点权都相对很小。这从一定程度上反映了图书借阅网中节点之间的竞争是很激烈的，导致竞争结果分布的极不均匀性。

图 3.3.16 研究的是图书借阅网的点权与项目度的相关性，统计结果显示在该网中点权与项目度是负相关的，更精确地说，节点点权与项目度是指数衰减的关系。这意味着，一本书涉及的科学领域越多（即这本书的专业性越低），它的受欢迎程度也越低，反之，一本书的专业性越强，它的受欢迎程度也就越高，这很可能是对于物理类图书借阅网才成立的。

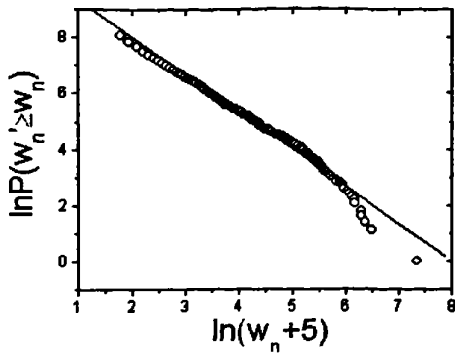
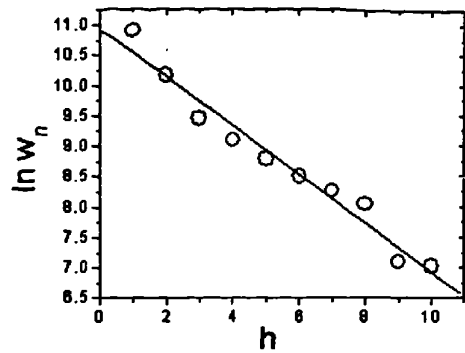


图 3.3.15 物理图书借阅网点权累计分布图

图 3.3.16 物理图书借阅网
点权和项目度的相关性

另外, 我们统计得到的度同类性为 0.83, 点强度同类性为 0.73, 结果表示此物理图书借阅网无论是度还是点强度都是同类匹配的, 即度大的节点倾向于与度大的节点相连。

3.4 小结与讨论

在这一章中, 我们提出了对一类合作竞争网的一种新的描述方法。网络中的基本单元可以分为两大类, 一类是事件、组织、活动等, 称为项目; 另一类是事件、组织、活动等的参与者, 称为节点, 可以用二分图及其投影图来很好的描述这一类网络, 所有以前提出的网络的拓扑结构的统计性质仍可以用来描述节点之间合作的情况与合作的态势。由于传统的边权可以看成是对节点对之间重复边的统计描述, 因此在本文中, 我们仍把含边权的统计性质归类为无权统计性质。本文所提出的重要的统计性质是点权分布及点权与项目度的相关性, 点权用来描述竞争的结果或节点的竞争因素, 点权分布则显示了节点间竞争的重要信息, 点权与项目度的相关性则可以反映出节点的竞争因素与其参加的项目多少之间的关系。我们以两个简单的实际网络为例简要说明了我们的研究方法, 并较详细的给出了这两个实际网络的实证统计结果。从两个网络的实证统计结果中, 我们得出的一个结论是项目度分布与点强度分布规律是一致的, 具有相同的函数形式。这个结论在文献^[6,7]中有详细的报道, 并通过一个网络演化模型作了进一步的阐述。另一个

统计性的结论是，这两个网的点权分布都是近似的幂律分布，然而，这个结论是否适用于所有这一类网络，我们现在还不能下此定论，值得进一步研究。在本文的第六章将详细地讨论点权与合作竞争网络。

参考文献：

- [1] S. Wasserman and K. Faust. *Social Network Analysis: Methods and Applications*, Cambridge Univ. Press, Cambridge (1994).
- [2] D. J. Watts and S. H. Strogatz. Collective dynamics of 'small-world' networks . *Nature*, 393: 440-442 (1998)
- [3] Albert-László Barabási, Réka Albert. Emergence of Scaling in Random Networks. *Science* Vol. 286. No. 5439, 509 – 512 (1999)
- [4] Newman, M.E.J. Scientific collaboration networks. I. Network construction and fundamental results. *Phys. Rev. E* 64, 016131 (2001); Newman, M.E.J. Scientific collaboration networks. II. Shortest paths. weighted networks and Centrality. *Phys.Rev. E* 64, 016132 (2001)
- [5] Barabasi A-L, Jeong H, Neda Z et al. Evolution of the social network of scientific collaboration. *Physica A* 311: 590-614 (2002)
- [6] Y. Fan, M.Li, J.Chen, etal.. Network of econophysicists: a weighted network to investidate the development of econophyscs. *Int.J.Mod.Phys.B* 17–19 (2004) 2505
- [7] P.Zhang, K.Chen, Y.He, T.Zhou, B.-B.Su, Y.Jin, H.Chang, Y.-P.Zhou, L.-C.Sun, B.-H.Wang, D.-R.He, *PhysicaA* 360 (2006) 599
- [8] H. Chang, B.-B. Su, Y.-P. Zhou, D.-R. He, *Physica A* 383 (2007) 687. [8]S. Wasserman, K.Faust, *Social Network Analysis: Methods and Applications*, Cambridge Univ. Press,Cambridge,1994
- [9] Maria Bengtsson and Sören Kock. "Coopetition" in Business Networks—to Cooperate and Compete Simultaneously. *Ind. Mark. Manag.* 29 (2000) 411
- [10] J.Yang, L.Liu, W. Xie, G.Chen, D.Zhuang. On competitive relationship networks: A

- new method for industrial competition analysis. *Physica A* 382 (2007) 704-714
- [11] P. L. Krapivsky, S. Redner, in: R.P. Satorras, M. Rubi, A.D. Guilerá (Eds.), *Statistical Mechanics of Complex Networks*, Springer-Verlag, Berlin, Heidelberg, 2003, p.4
- [12] M.E.J.Newman. Analysis of weighted networks. *Phys.Rev.E* 70 (2004) 056131
- [13] Z.Pan, X.Li, X.Wang. Generalized local-world models for weighted networks. *Phys.Rev.E* 73 (2006) 056109
- [14] M.Barthélelem, A.Barrat, R.Pastor-Satorras, A.Vespignani. Characterization and modeling of weighted networks. *Physica A* 346 (2005) 34
- [15] M. Li, Y.Fan, J.Chen, L.Gao, Z.Di, J.Wu. Weighted networks of scientific communication: the measurement and topological role of weight . *Physica A* 350 (2005) 643
- [16] W.-X.Wang, B.-H.Wang, B.Hu, G.Yan, Q.Ou. General Dynamics of Topology and Traffic on Weighted Technological Networks. *Phys.Rev.Lett.* 94 (2005) 188702
- [17] A.De Montis, M Barthelemy, A.Chessa,A.Vespignani.arXiv:physics/0507106
- [18] C.C.Leung, H.F.Chau.ar Xiv:physics/0607134
- [19] S.Sinha. Evidence for power-law tail of the wealth distribution in India. *Physica A* 359 (2006) 555–562
- [20] S.Battiston. Inner structure of capital control networks. *Physica A* 338 (2004) 107

第四章 四个合作竞争系统的特殊合作性质

4.1 实证研究的系统

在最典型的复杂系统，例如生命系统或社会系统中，基本单元常常具有自主的行为能力。在其它许多系统（例如人造的或者由人挑选构成的系统）中，似乎基本单元不具有自主行为能力，然而它们常常会反映具有自主能力的“操纵者”的影响，从而间接地表现出一些自主行为特征。从一个角度可以发问：这些直接或间接地具有自主行为的基本单元之间是在合作还是在竞争？合作与竞争的具体性质如何？是否存在某些共性？这是一直受到关注的问题。

近十年来复杂网络的研究风靡整个科学界。描述基本单元之间合作关系的“合作网络”也受到了重视，电影演员或者科研人员之间的合作关系网络得到了特别多的研究^[1-4]。本课题组曾经建议把合作网络的描述范围扩大到上述基本单元可以间接地表现出自主行为特征的系统，例如中药方剂网、淮扬菜系网、中国大陆的旅游线路网、北京、上海、南京、杭州、扬州五个城市的公交线路网。加上好莱坞演员合作网，本课题组一共实证研究过九个社会的或者非社会的合作网络^[5-9]。

文献^[1-9]建议用二分图描述的合作网是“隶属网”的一种。它们的节点可以分为互补的两类。其中一类节点是某种活动、事件或者组织中的“参与者”(actor)，而另一类节点就是它们参与的活动、事件或者组织（称为“项目(act)”)。合作网络强调节点在参加项目时期的合作关系，忽略节点之间的其它关系，例如竞争、支配等，认为在一个项目中所有参与者节点都在无差别地合作，因此两两之间都连接一条等同的边。在合作网络的研究中，常常把二分图向一类节点（常常是参与者节点）投影，得到单模式网络。这种投影图中每个项目表示为一个项目完全子图。

认为所有节点都在无差别地合作显然不符事实，差别就反映了竞争。然而，如果认为所有节点都在绝对地竞争，那也不符事实。即使交战的两国，随着人类文明的发展，结果也不会是一方完全消灭另一方。因此，在自然界中，既有合作、又有竞争才是普遍情况。本课题组已经实证研究了二十五个合作竞争网络，其中少数成果已经发表^[10-12]。我们建议仍旧使用二分图来描述这些合作竞争网络，仅

仅加上标志竞争结果的“点权”^[10-12]。表 1 列出了这 9 个合作网络和 25 个合作竞争网络的名称和节点、项目、节点权定义，以及相关论文和数据来源。其中节点权项下标志“无”的是理想化的合作网。我们强调：这 34 个实际网络的数据（包括点权）都来源于公开发表的书籍或者对公众开放的网址（其中扬州大学的两个需有关部门授权），读者们可以根据表 1 提供的信息继续对它们的研究。同时，由于要获得标志竞争的点权数据（尽管点权不是本文主要报道的内容），所以不能使用国外同行普遍应用的几个合作网范例（例如好莱坞演员、科研论文合作、企业董事会等），我们必须自己搜索合适的数据库。

在复杂网络研究中，常常最重视“度分布”的定性规律。所谓“度”就是一个节点连接的邻边数。度的分布 $p(k)$ 是网络中随机选择一个节点，它的度为确定值 k 的概率。有时，为了减小涨落，还可以统计累计度分布 $p(k \geq k)$ ，它表示随机选择一个节点，它的度值不小于 k 的概率。二分图中有两类节点，而传统图论中只考虑不同类节点之间的边，因此有两种度及其度分布。一个参与者节点的二分图度表示它参与的项目数，称为“节点项目度”，用 h 表示；一个项目的二分图度表示它包含的参与者数，称为“项目大小”，用 T 表示。不管在理想化的合作网中，还是在合作竞争网络中，这三个性质都表示系统的合作特性，竞争特性的表述必须借助点权^[10-12]。本文集中讨论这 34 个实际网络中的 14 个所表现出的这三种性质的新特征。

表 1：9 个合作网络和 25 个合作竞争网络简况

序号	系统	节点	项目	点权
1	高校自主招生	高中	高校	人数
2	高考录取	高校	录取批次	人数
3	世界语言	语言	国家和地区	人数
4	中国餐饮	餐饮企业	行业	占有率
5	中国商业银行	商业银行	行业	占有率
6	全国连锁企业	连锁企业	行业	占有率
7	中国电子	电子企业	行业	占有率

8	上网栏目访问	上网栏目	时间段	占有率
9	央视电视节目收视率	电视节目	时间段	占有率
10	NBA 冠军联赛	球队	赛季	得分
11	台湾汽车销售	汽车厂家	行业	占有率
12	物理图书借阅	物理图书	分类	人数
13	扬州大学选课	课程	专业系科	人数
14	水果营养	营养成分	水果	占有率
15	混合饮料	成分	饮料	占有率
16	经络穴位	穴位	疾病	占有率
17	大陆铁轨线	火车站	铁轨线	占有率
18	2004 年奥运	运动员	比赛项目	得分
19	大陆电影	电影	演员	点评分
20	淘宝网计算机销售	淘宝商家	笔记本型号	销售价格
21	软件下载	软件	网站	占有率
22	北京餐馆	餐馆	菜肴	点评分
23	培训机构	培训课程	培训机构	人数
24	综合商场	商场	商品	点评分
25	IT 产品	生产厂家	IT 产品	占有率
26	好莱坞演员	演员	影片	无
27	中药方剂	药材	方剂	无
28	淮扬菜肴	食品	菜肴	无
29	大陆旅游	旅游点	旅游线路	无
30	北京公交	车站	线路	无
31	上海公交	车站	线路	无
32	南京公交	车站	线路	无
33	杭州公交	车站	线路	无

34	扬州公交	车站	线路	无
----	------	----	----	---

这三十四个实际系统中，全部九个实际合作网络和六个实际合作竞争网络的上述三项性质都显示比较容易理解的规律，即项目大小遵循正态分布，度分布与项目度分布为幂函数和指数函数之间的所谓“漂移幂律”分布（函数形式为 $P(h) \propto (h + \alpha)^{-\gamma}$ ，其中 α 和 γ 为常量，随 α 从 0 到 ∞ 的变化，函数从幂律连续地过渡到指数），而且每个系统的这两种分布定性相同^[5,9]。张培培、常慧等人曾建议了一个合作网的演化模型来解释 $p(h)$ 和 $p(k)$ 采取漂移幂律的机制。模型的主要假设是网络演化中每次增加的新节点以概率 p 随机选择，以 $1-p$ 概率按照节点项目度 h 线性优选地选择 $T-1$ 个旧节点构成一个新项目。在 T 为常量的假设下，这个模型得到的演化方程可以解析求解，得到 $p(h)$ 和 $p(k)$ 的漂移幂律函数^[5,6]。这个主要假设告诉我们，即使在每个项目中假设参与者都在无差别地合作，各个参与者在网络演化中起的作用也非常不同。 h 大（即已经参加项目多）的参与者参加新项目的概率也高，从而参加项目就更多。然而，另一部分概率决定的随机选择机制又反映了合作，即不管什么样的节点都还有一部分概率平等地参加新项目。所以部分随机选择、部分 h 线性优选选择就反映了网络演化中的合作与竞争。我们以前一直认为只有在第二个假设也成立，也就是说 T 为常量，或者为正态分布（可以看作在平均值常量附近存在扰动）时，这核心假设才会导致 $p(h)$ 和 $p(k)$ 的漂移幂律形式。只有在完成了本文报道的研究后，这种看法才得到了修正。

在上述 25 个合作—竞争网中，剔除 5 个只有一个项目和 6 个遵从 T 正态分布的系统之后，其余 14 个实际合作竞争网络却显示不是单峰分布，而是漂移幂律分布的项目大小分布。从占据系统总数的比例来看，它们也代表一种不同的普遍现象，值得研究。本文比较详细地报道代表这 14 个“另类”系统的 4 个系统（大陆电影网、北京餐馆网、2004 年奥运会网、淘宝笔记本电脑销售网）作为例子，并且仍旧运用张培培、常慧等人的演化模型来进行解释。下两节将介绍对 4 个项目大小、度以及项目度的统计结果。由于统计这三项性质的程序比较简单，我们自己编写了程序。结果的可靠性经过验证。

4.2 项目大小分布

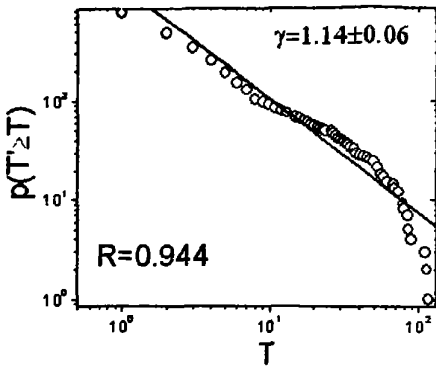


图 4.2.1 大陆电影网的项目大小 T 累计分布

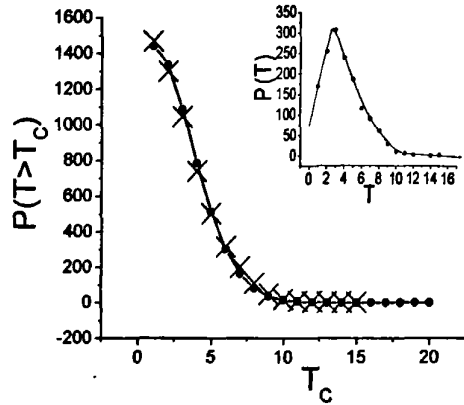


图 4.2.2 中药方剂网的项目大小 T 累计分布（主图）、不累计分布（插入图）

为了强调影片在同一类观众中的竞争，大陆电影网的节点定义为中国大陆拍摄的电影，影片中的主要演员定义为项目，节点点权定义为网民对该电影的打分。如果两部影片聘请同一个演员，则这两部电影之间连接一条边。我们共收集了 3085 部影片和 920 个主要演员的数据^[10]。图 4.2.1 显示了大陆电影网的项目大小 T 的累计分布（含义是随机选择一个项目，其项目大小不小于 T 的几率），数据在双对数坐标下直线拟合的相关系数 $R=0.944$ ，说明 T 的累计分布满足幂律分布 $P(T' \geq T) \propto T^{-\gamma}$ ，其中累计分布幂指数 $\gamma=1.14 \pm 0.06$ （可以解析地证明如果一个随机变量的累计分布遵循幂律分布、指数函数分布、或居于其间的漂移幂律分布，则它的不累计原始分布也是同样的分布^[48]）。为了对比，图 4.2.2 显示了作为间接自主行为理想合作网代表的中药方剂网的项目大小 T 分布（取自文献^[6]）。图中圆点表示实证数据，叉子和曲线表示正态分布函数的拟合。可以看出图 1 和图 2 表示的确实是两种类型。

北京餐馆网中定义餐馆为节点，每道菜为项目，销售同一道菜的两个餐馆之间连接一条表示竞争的边，点权定义为网民的评分。我们统计了 3337 个餐馆和 688 道菜。图 4.2.3 显示了北京餐馆网的幂律项目大小分布。累计分布幂指数 $\gamma=$

1.26 ± 0.05 , 直线拟合相关系数 $R=0.969$ 。

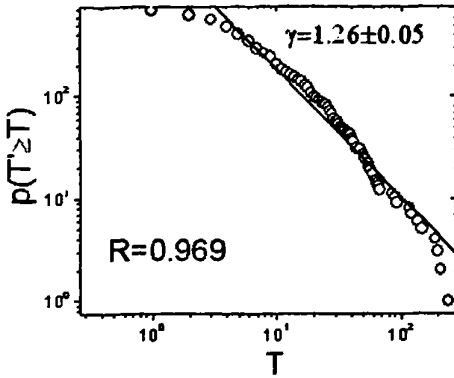


图 4.2.3 北京餐馆网的项目大小 T 累计分布

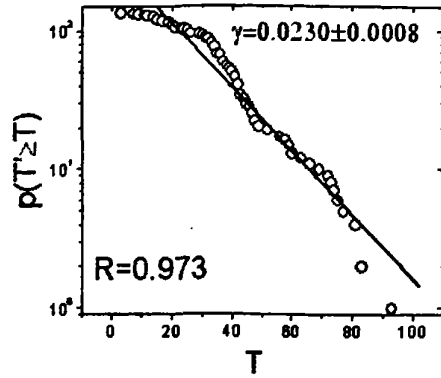


图 4.2.4 2004 年奥运会网的项目大小 T 累计分布

2004 年奥运会网中定义运动员为节点, 比赛项目为项目, 参加同一个比赛项目的两个运动员之间连一条表示竞争的边, 点权定义为运动员的比赛成绩分。我们统计了 229 个比赛项目和 4496 个运动员。图 4.2.4 显示了项目大小累计分布图。结果显示单对数下的直线, 拟合系数 $R=0.973$, 即累计分布为指数函数分布, $P(T' \geq T) \propto e^{-T}$ 。

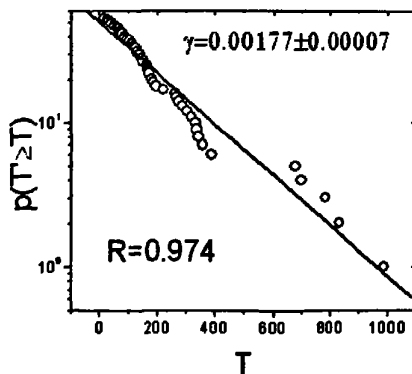


图 4.2.5 淘宝笔记本电脑销售网的项目大小 T 累计分布

在淘宝网上销售笔记本电脑的网络中, 把商家定义为节点, 笔记本的型号定义为项目, 销售同一型号笔记本电脑的商家之间连接一条表示竞争的边, 点权定

义为销售价格。我们共收集到 53 个型号的笔记本电脑, 4723 个淘宝网商家的数据。图 4.2.5 显示了淘宝笔记本电脑销售网络的项目大小分布累计图, 单对数下为直线, 拟合系数 $R=0.974$, 即服从指数函数规律。

以上四幅图显示这四个网络的项目大小不遵循单峰分布。它们遵循的幂函数或指数函数分布 (是漂移幂律的两个极端情况) 表示高度不均匀的分布, 绝大多数项目只包含很少参与者, 而极少数项目包含大量的参与者。

4.3 度分布与项目度分布

如上所述, 项目度分布是二分图中另一项重要统计性质, 表示节点参加项目个数的分布。大陆电影网的项目度表示每部影片在网上列出的主演 (常常是名角) 的人数。图 4.3.1 显示大陆电影网的项目度累计分布 $p(h \geq h)$ (随机选择一个节点其项目度不小于某一特定值 h 的概率), 显示很好的指数函数分布。在北京餐饮业合作竞争网络中, 项目度指每一个餐馆做多少道菜。由图 4.3.2 可以看出它的项目度也满足指数分布。2004 年奥运会网的项目度则是一个运动员共参加了多少个比赛项目。图 4.3.3 显示它的项目度累计分布, 也是一个指数分布。淘宝笔记本电脑销售网的项目度指每一个商家销售的笔记本电脑型号数目。图 4.3.4 显示它的项目度累计分布图, 所显示的还是指数分布。它们的直线拟合相关系数分别为 0.996、0.991、0.979、0.997。

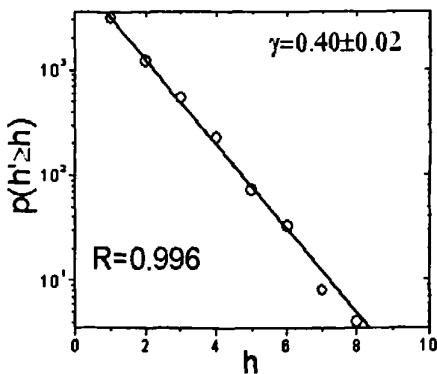


图 4.3.1 大陆电影网项目度 h 累计分布

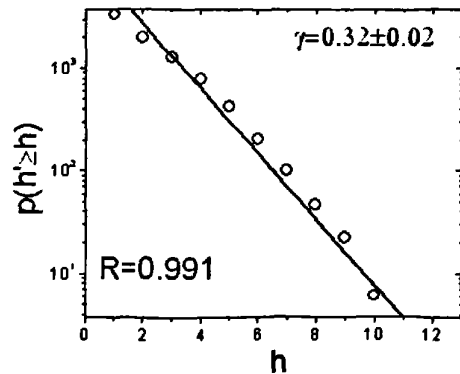


图 4.3.2 北京餐馆网项目度 h 累计分布

合作竞争二分图向参与者的投影图中, 节点的度 k 表示它和多少个其它节点一

起共同参加过至少一个项目。显然, 一个节点的度完全取决于它的项目度(参加几个项目)以及它参加项目的大小(其中一共有多少其它节点)。如果网络的项目大小遵循单峰分布, 则可能近似地认为项目大小等于其平均值, 则任一个节点的度与它的项目度成正比, 度分布自然也与项目度分布遵循相同的规律。如果网络的项目大小遵循很不均匀的幂函数或指数函数分布, 就不能做这样的推断。就我们所知, 在此之前还没有这样的实证或模型研究。

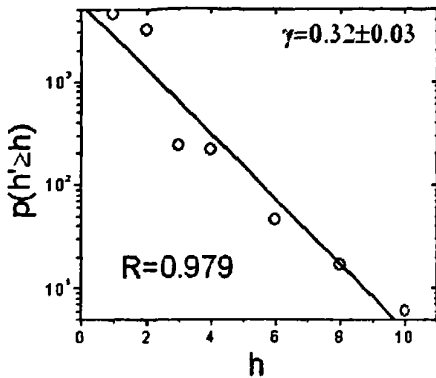


图 4.3.3 2004 年奥运会网

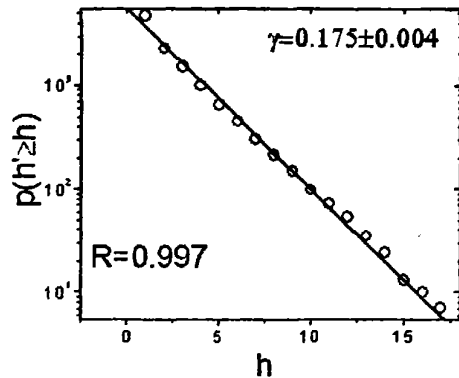
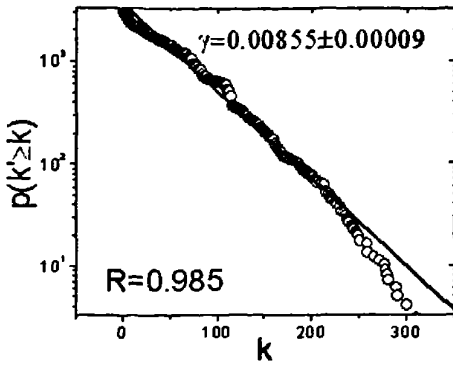
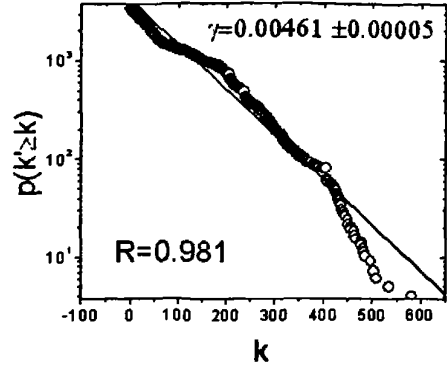
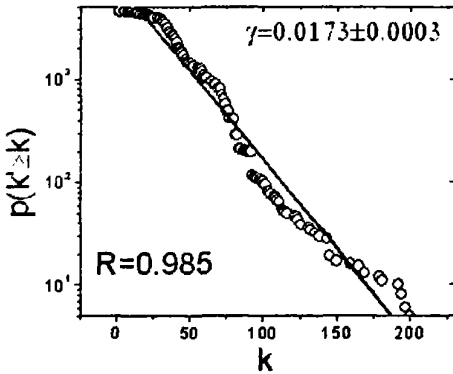
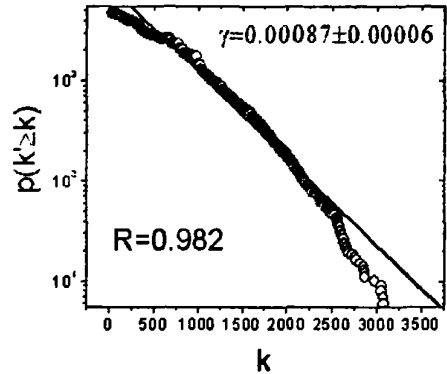
项目度 h 累计分布

图 4.3.4 淘宝笔记本电脑

销售网项目度 h 累计分布

大陆电影网的度表示与一部影片和其他多少部影片至少共用过一个主要演员。图 4.3.5 显示大陆电影网的累计度分布图 $P(k \geq k)$, 呈指数函数分布。北京餐饮网的度表示它和多少餐馆做过至少一道相同的菜肴。图 4.3.6 显示北京餐饮网的节点度累计分布, 同样呈指数函数分布。2004 年奥运会网中节点的度表示与该运动员和多少运动员一起参加过至少一项相同的比赛项目。图 4.3.7 显示 2004 年奥运会网的节点度累计分布图, 也是一个指数分布。在淘宝笔记本电脑销售网中, 节点的度表示与某一商家与多少个其他商家一起销售过至少一个相同型号笔记本电脑。图 4.3.8 显示淘宝笔记本电脑销售网的节点度累计分布图, 再一次显示很好的指数分布。直线拟合相关系数分别为 0.998、0.981、0.985、0.982。

这样, 四个系统研究的结论是: 尽管网络的项目大小遵循很不均匀的幂函数或指数函数分布, 但是它们的度分布仍旧与项目度分布遵循相同的规律。

图 4.3.5 大陆电影网度 k 累计分布图 4.3.6 北京餐馆网度 k 累计分布图 4.3.7 2004 年奥运会网度 k 累计分布图 4.3.8 淘宝笔记本电脑销售网度 k 累计分布

4.4 讨论

本文的实证研究结果使我们怀疑以前的观点，即只有在 T 为常量，或者为正态分布时，前面介绍的张培培、常慧等人建议的合作网的演化模型的核心假设才会导致 $p(h)$ 和 $p(k)$ 的漂移幂律形式。为此我们在项目大小 T 为幂函数或者指数函数分布的情况下分别计算了模型分别取不同的随机选择概率 p 时得到的度分布与项目度分布。图 4.4.1—4.4.4 显示了部分结果。

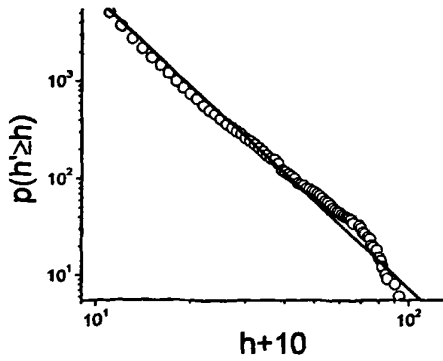


图 4.4.1 张培培、常慧模型
 T 幂律, $p=0.5$ 时的 h 累计分布

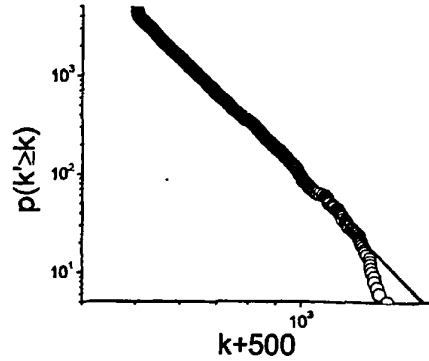


图 4.4.2 张培培、常慧模型
 T 幂律, $p=0.5$ 时的 k 累计分布

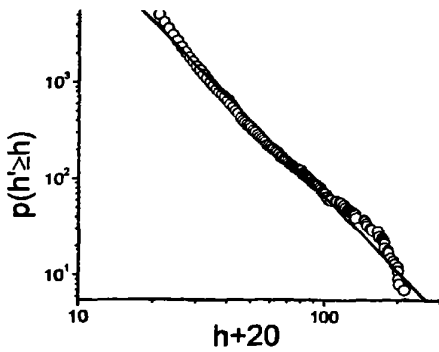


图 4.4.3 张培培、常慧模型
 T 指数, $p=0.5$ 时的 h 累计分布

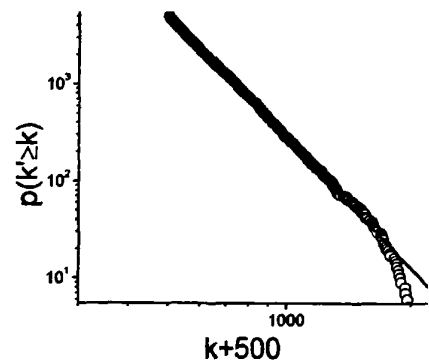


图 4.4.4 张培培、常慧模型
 T 指数, $p=0.5$ 时的 k 累计分布

这些新计算结果说明, 不管 T 分布采取幂律或指数, 模型的项目度分布和度分布在各种 p 参数下都遵从漂移幂律。因此, 张培培、常慧模型中的核心假设, 即“网络演化中每次增加的新节点以概率 p 随机选择, 以 $1-p$ 概率按照节点项目度 h 线性优选地选择 $T-1$ 个旧节点构成一个新项目”, 完全决定了项目度分布和度分布的函数, 与项目大小分布的形式无关。如果网络演化时完全采取反映竞争的线性优选法则, 项目度分布和度分布都遵循幂律, 差异最大; 如果网络演化时完全采取反映合作的随机法则, 项目度分布和度分布都遵循指数函数, 差异最小。这既说明了本文报道现象的机制, 也把张培培、常慧模型的适用范围作了推广。

我们在本文中报道的实证研究涉及三十四个系统，分属文化、教育、传媒、体育、经济、旅游、交通、人体科学等大领域，导致了可能跨越这些大领域的普遍规律。研究历时五年，除本文作者外，曾有姜玉梅博士，何阅、许田、张培培、常慧、周月平、苏蓓蓓、沈丹等硕士，陈洁、王彬彬、韩雪芳、邱荣、唐继英、李娟、张培方等五届、近四十位本科毕业生贡献过辛勤的劳动。本文报道的结论是：所有合作网络的一部分合作竞争网络的项目大小分布遵循单峰分布，另一部分的项目大小分布遵循漂移幂律分布。张培培、常慧等人提出的网络演化模型可以解释这两种情况的度分布与项目度分布为何遵循漂移幂律分布。

参考文献：

- [1] Watts D J and Strogatz S H. Collective dynamics of small-world networks[J]. *Nature*, 1998, 393: 440-442
- [2] Barabasi A.-L and Albert R. Emergence of Scaling in Random Networks[J]. *Science*, 1999, 286: 509-512.
- [3] Newman M E J. Scientific collaboration networks. I. Network construction and fundamental results[J]. *Phys. Rev. E*, 2001, 64: 016131; Newman M E J. Scientific collaboration networks. II. Shortest paths, weighted networks and Centrality[J]. *Phys. Rev. E*, 2001, 64: 016132.
- [4] Barabasi A-L, Jeong H, Neda Z, Ravasz E, Schubert A and Vicsek T. Evolution of the social network of scientific collaborations[J]. *Physica A*, 2002, 311: 590-614
- [5] Zhang P P, Chen K, He Y, Zhou T, Su B B, Jin Y, Chang H, Zhou Y-P, Sun L-C, Wang B-H, He D-R. Model and empirical study on some collaboration networks[J]. *Physica A*, 360 (2006) 599-616
- [6] Chang H, Su B B, Zhou Y-P, He D-R. Assortativity and act degree distribution of some collaboration networks[J]. *Physica A*, 383 (2007) 687-702
- [7] 常慧, 何阅, 张义勇, 苏蓓蓓, 何大初. 中国旅游线路的合作网络描述[J]. *科技导报*, 2006, 24 (9): 84-87

- [8] 何阅, 张培培, 唐继英, 韩雪芳, 邱荣, 陈其娟, 周月平, 常慧, 何大韧. 中药方剂的合作网络描述[J]. 科技导报, 2005, 23(11): 36-39
- [9] 张培培, 侯威, 何阅, 何大韧. 复杂系统与复杂性科学. 淮扬菜系的网络描述[J]. 2005, 2 (2): 9-53
- [10] 刘爱芬, 付春花, 张增平, 常慧, 何大韧. 中国大陆电影网络的实证统计研究[J]. 复杂系统与复杂性科学, 2007, 4(3), 10-16.
- [11] Fu C-H, Zhang Z-P, Chang H, Tao J-R, Chen Z-H, Dai Y-L, Zhang W, He D-R. A kind of collaboration-competition networks[J]. Physica A, 2008, 387: 1411-1420
- [12] Qu Y, Jiang Y-M and He D-R. Fruit nutritive factor network[J]. J. of Systems Science and Complexity, 2008, 21: 1-9.
- [13] <http://www.chsi.com.cn/>
- [14] <http://www.hneeb.cn/>, <http://www.jszs.net/>, <http://www.lnzsks.com/>, <http://www1.nm.zsks.cn/> et al.
- [15] <http://www.ethnologue.com/>
- [16] <http://www.dianping.com/beijing>
- [17] <http://www.enorth.com.cn>
- [18] <http://www.jrj.com>
- [19] <http://www.enorth.com.cn>
- [20] <http://www.cctv.com/service/special/range/01/index.shtml>
- [21] <http://www.cctv.com>
- [22] 邱荣. NBA 比赛的动力学模型. 扬州大学毕业论文, 2005
- [23] <http://www.china.nba.com>
- [24] <http://club.autohome.com.cn/bbs/>, <http://ido.thethirdmedia.com/article/frame>
- [25] <http://lib.yzu.edu.cn/>
- [26] <http://jiaowu.yzu.edu.cn/>.
- [27] <http://www.fumuqin.com/>.
- [28] <http://www.drinknation.com/>

- [29] <http://www.acutimes.com/index.asp>
- [30] <http://www.chinamor.cn.net/>
- [31] <http://2004.sina.com.cn/>
- [32] <http://www.movdb.com>, <http://soft6.com/>, <http://www.mtime.com/movie/>
- [33] <http://list.taobao.com/browse/cat-1101.htm>
- [34] 吴琴. 软件下载竞争网. 扬州大学毕业论文 (数据从 180 个网址下载), 2007
- [35] <http://www.dianping.com/>
- [36] <http://www.ot51.com/>, <http://www.00100.cc/>, <http://philosophy.cass.cn/>
<http://www.people.com.cn/et al.>
- [37] <http://www.dianping.com/shanghai/shopping>
- [38] <http://www.pcpop.com/>; <http://www.it168.com/>
- [39] <http://us.imdb.com/>
- [40] 朱玉祥. 中医处方指南[M]. 北京: 金盾出版社(第二版), 1996; 刘东亮等. 中医常用方剂手册[M]. 人民军医出版社, 北京: 1996
- [41] 北京民族饭店菜谱编写组. 淮扬菜谱[M]. 中国旅游出版社, 1993.
- [42] 孙文昌. 现代旅游开发学[M]. 青岛: 青岛出版社, 1999; 实用交通旅游图典编辑组. 实用交通旅游图典[M]. 福州: 福建省地图出版社, 2002.
- [43] <http://www.bjbus.com/>
- [44] <http://www.jt.sh.cn/>
- [45] <http://www.njkgc.cn>
- [46] <http://www.hzbus.com.cn>
- [47] <http://www.yzren.com/life-busstation.aspi>.
- [48] 何大韧, 官山, 朱陈平. 复杂网络实证研究中的一些问题. 《复杂网络理论与应用研究》论文集[M]. 上海: 上海系统科学出版社, 20

第五章 点权与合作竞争网

5.1 基本概念

5.1.1 合作网与二分图

社会网络是复杂网络中重要的一类，可以分为“单模式网络”（只有同一类节点，称为参与者）和“双模式网络”（有两类不同的参与者），甚至更多模式的网络。双模式网络中的最重要一种称为“隶属网”（affiliation network），其中一类节点是某种活动、事件或者组织的参与者，而另一类节点就是它们参与的活动、事件或者组织，称为“项目(act)”。这种网络可以用一个二分图^[1,3-9](bi-partite graph)来比较好地描述。合作网^[1-2](collaboration network)是隶属网中的一种，合作网的研究只考虑基本单元之间的合作关系，即任意两个节点之间的边只表示节点之间的合作关系，忽略节点之间的其它关系，如竞争、支配等。在合作网络的研究中，常常把二分图向一类节点（通常是参与者节点）投影，得到单模式网络。这时参与每个项目的所有参与者节点之间都连有表示在此项目中合作关系的边，每个项目就表示为一个项目完全图(complete act graph)，整个单模式网络成为项目完全图的集合，如图 5.1 所示。

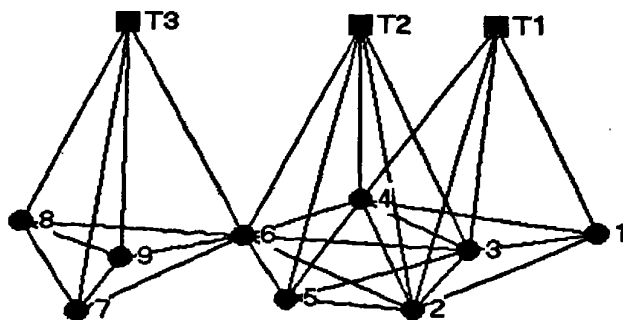


图 5.1 二分图及其投影图

5.1.2 合作网的重要统计性质

合作网的统计性质已提出很多，下面简要介绍几种常用的统计性质。度

(degree), 表示网络中每个节点的邻点个数, 用 k 表示。从一定程度上可以说明, 度越大节点的重要性也越大。 $p(k)$ 表示网络中度为 k 的节点出现的概率, $p(k)-k$ 即度分布。项目度(act degree), 表示节点参与的项目的个数, 用 h 表示。 $p(h)$ 表示网络中项目度为 h 的节点出现的概率, $p(h)-h$ 即项目度分布。研究表明, 合作网络中度分布与项目度分布是趋于一致的^[7]。

项目大小, T , 则用来表示每个项目中所包含的节点个数^[7-9]。同理, $p(T)-T$ 表示项目大小分布。集群系数, C , 表示节点间联系的紧密程度。节点 i 的集群系数定义为 C_i 定义为:

$C_i = \frac{2\{e_{jk}\}}{k_i(k_i-1)}$, 其中 $v_j, v_k \in N_i$, N_i 为节点 i 的邻点集合,

若 $e_{jk} = 1$, 表示 i 节点的邻点 j 和 k 之间有连边, 反之, 若 $e_{jk} = 0$, 表示 i 节点的

邻点 j 和 k 之间没有连边。 $\frac{k_i(k_i-1)}{2}$ 表示节点 i 的所有邻点之间最大可能连边数。

整个网络的平均集群系数为 $\bar{C} = \frac{1}{n} \sum_{i=1}^n C_i$, n 为网络的节点总数。

除了上面所讲的度分布、项目度分布、项目大小分布和集群系数之外, 合作网的统计性质还有最短道路长、同类性、交联度和群落等等。

5.1.3 合作竞争网

合作网络的研究忽略了节点之间的其它关系, 如竞争、支配等, 只考虑合作关系。而实际情况下, 基本单元之间只有纯粹合作关系的情况是很少见的, 更多更普遍的是基本单元之间同时存在合作与竞争, 完全合作与完全竞争只是两种极端情况。如演员合作网中, 每个演员在电影中的地位是不同的(如主演和配角), 每个演员所得到的报酬通常也是不同的(主演的报酬肯定多于配角的报酬), 每个演员都希望自己成为主演, 都希望获得更多的报酬。因此, 合演同一部电影的演员之间不仅存在合作关系, 他们之间还存在着竞争关系。再如已被广泛研究过的科学家合作网^[3-6], 我们都知道一篇科研论文的作者往往不只一个, 尤其是实验成果类的论文作者是更多的。所以, 毫无疑问, 同一篇科研论文的作者之间存在合

作关系。合作的同时，每个人都希望能成为论文的主要贡献者，在论文中起主要的作用，这就意味着他们之间也存在着竞争关系。我们将这类网络统称为合作竞争网络，仍可以用二分图来描述。

5.1.4 节点项目权与节点总权

二分图及其投影图仍可以用来描述合作竞争网，但是传统的合作网络统计性质已不能很好的刻画出节点之间的合作竞争关系。

通过前面的叙述，我们知道节点之间的竞争主要来源于它们在合作中的地位等的差别。如演员合作网中，有主演和配角的差别，因此有报酬多少的差别，这正是由于竞争产生的结果，我们说报酬的多少就是竞争的结果。从一定程度上来说，获得的报酬越多说明了节点的竞争力越强，在竞争中越占优势。因此，我们用节点项目权来表示节点在项目中的竞争结果^[10]，用 w_{il} 表示， i 是节点的标号， l 是项目的标号。有的演员在这部电影中获得的报酬多，在另一部电影中获得的报酬少。所以，节点项目权通常是不同的。节点 i 的节点总权，用 w_i 表示，是指把

节点 i 的所有节点项目权累加，即 $w_i = \sum_{l=1}^{N_i} w_{il}$ ，其中 N_i 为项目总数。对于同一个实

际网络，节点项目权的定义可以是不同的。这样，二分图投影后的网络就是含权网络，原有的合作网的统计性质描述的是合作竞争的态势，而我们此处定义的节点总权描述的是合作竞争的结果。

与度分布等类似，我们可以统计网络的节点总权分布。 $p(w)$ 表示点权为 w 时的出现概率， $p(w)-w$ 表示点权分布。

5.2 实证结果

我们共统计了 14 个实际网络的节点总权分布。表 5.2 是 14 个实际网络的简介。

表 5.2 14 个实际合作竞争网的简介

实际网络	项目	节点	点权	数据来源
水果网	水果	营养成分	成分的含量	www.fumuqin.com
IT产品网	产品	企业	排名	www.pcpop.com; www.it168.com

选课网 ^[10]	系科	课程	选课人数	jiaowu.yzu.edu.cn
高校自主招生网	高校	中学	录取人数	www.chsi.com.cn
穴位网	疾病	穴位	针灸次数	www.acutimes.com/index.asp
电子百强	产品	企业	市场占有率	www.enorth.com.cn
混合饮料网	饮料	饮料成分	成分的含量	www.drinknation.com
物理图书借阅网 ^[10]	科学领域	书	借阅次数	lib.yzu.edu.cn
软件下载网	网站	软件	下载次数	www.baidu.com
中国培训机构网	培训机构	培训课程	课程收费	www.ot51.com; www.00100.cc et al.
北京餐馆网	菜肴	餐馆	菜的价格	www.dianping.com
餐饮百强	产品	餐馆	资产	www.dianping.com/beijing
大陆电影网 ^[11]	电影演员	电影	电影的得票数	www.movdb.com; soft6.com et al.
世界语言网	国家或地区	语言	说某语言的人数	www.ethnologue.com

图5.2.1至5.2.14分别为14个实际合作竞争网的累计节点总权分布，每个系统的节点总权都是归一的，通过我们的实证统计，惊讶地发现这14个系统的节点总权分布均为SPL^[9](Shifted Power Law)分布，SPL函数形式为： $P(w) = (w + \alpha)^{-\gamma}$ 。 α 和 γ 是SPL分布的两个重要参数，当 $\alpha=0$ 时， $P(w) = w^{-\gamma}$ ，此时SPL函数就是幂函数形式^[9]，此时是分布最不均匀的情况；当 $\alpha \rightarrow \infty$ 时，SPL函数近似为指数函数，分布由最不均匀逐渐向均匀转化。幂函数和指数函数是SPL函数的两种极端情况，当 α 由0逐渐向无穷大增大时，SPL函数逐渐由幂函数转化为指数函数，SPL分布也越来越均匀。

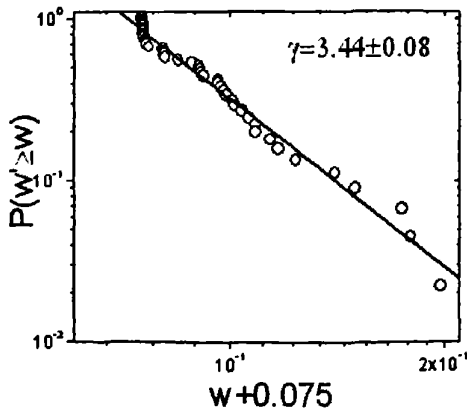


图 5.2.1 水果网累计点权分布

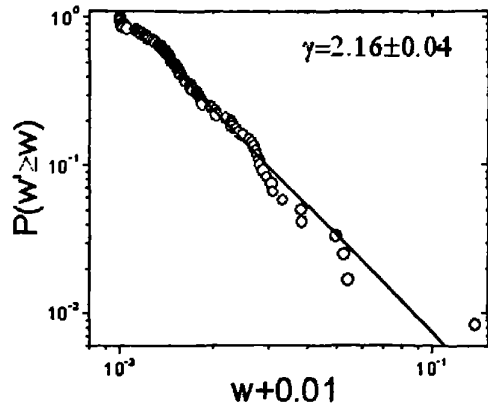


图 5.2.2 扬大选课网累计点权分布

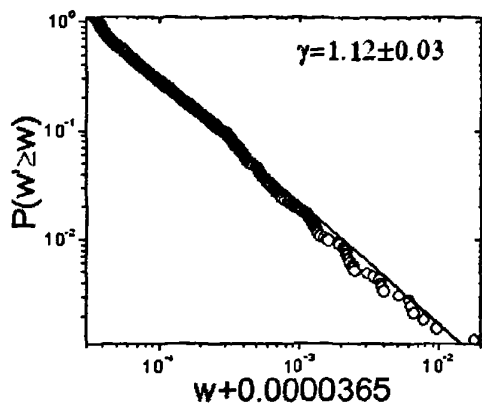


图 5.2.3 北京餐馆网累计点权分布

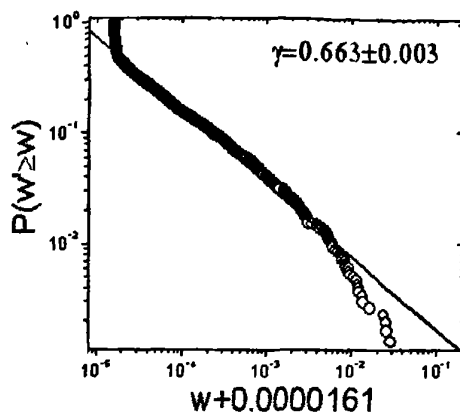


图 5.2.4 大陆电影网累计点权分布

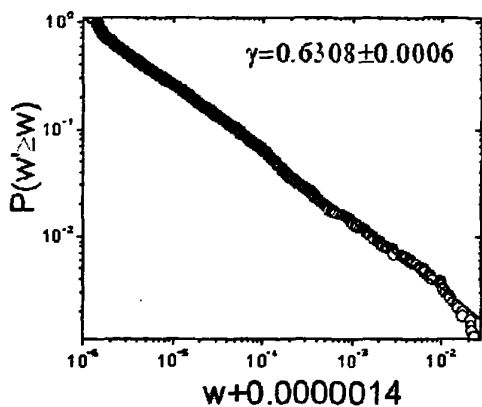


图 5.2.5 世界语言网累计点权分布

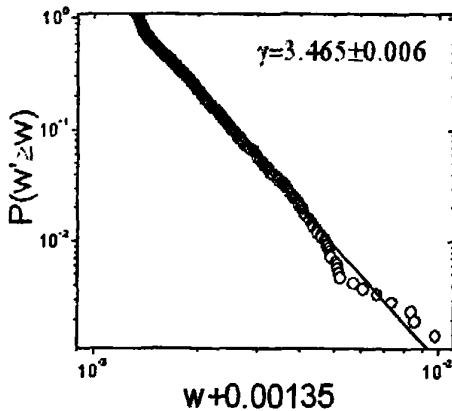


图 5.2.6 IT产品网累计点权分布

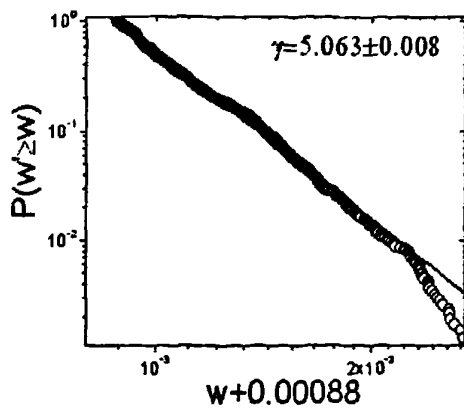


图 5.2.7 淘宝电脑销售网累计点权分布

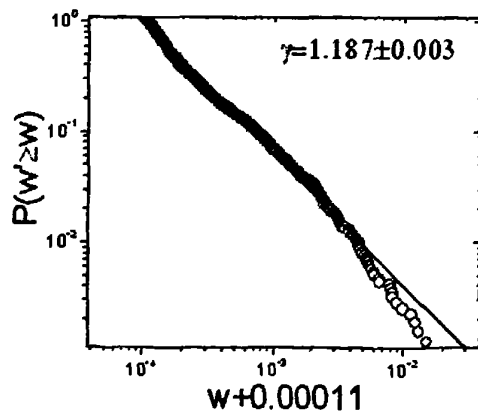


图 5.2.8 物理图书借阅网累计点权分布

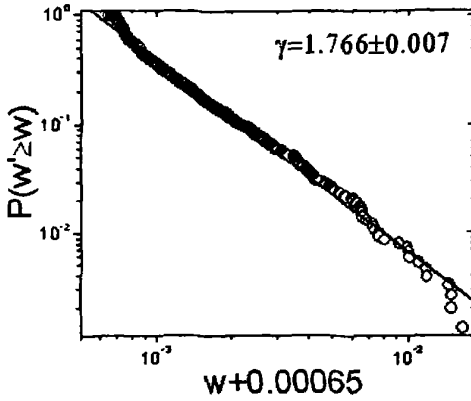


图 5.2.9 高校自主招生网累计点权分

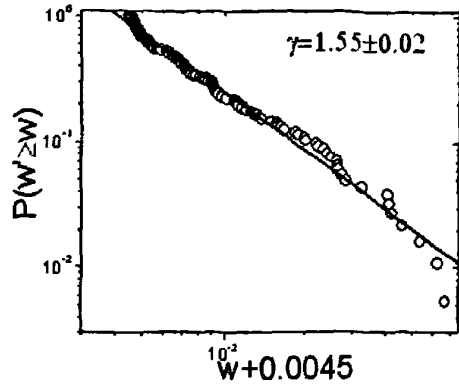


图 5.2.10 穴位网累计点权分布

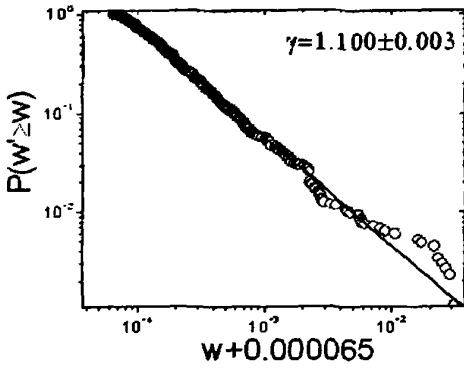


图 5.2.11 中国培训机构网累计点权分布

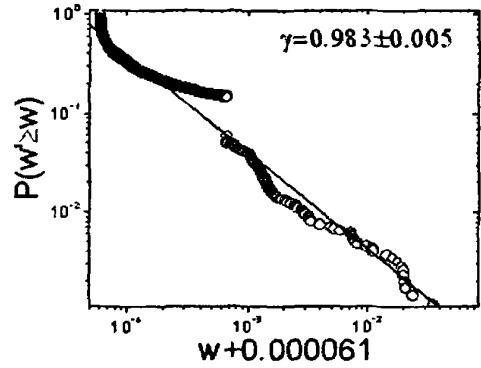


图 5.2.12 软件下载网累计点权分布

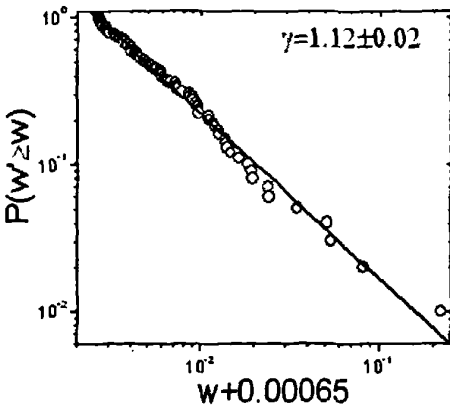


图 5.2.13 中国餐饮网累计点权分布

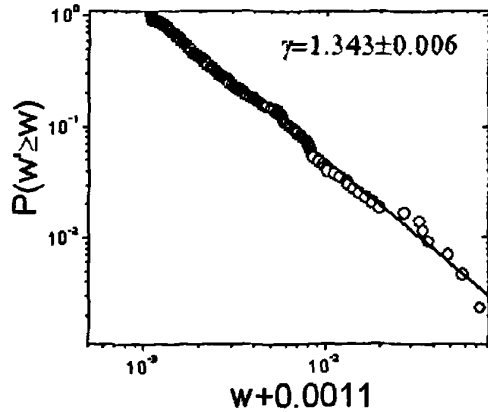


图 5.2.14 混合饮料网累计点权分布

从上述 12 幅图中我们可以看出，世界语言网的 α 参数值最小，最接近于 0，即世界语言网的点权分布最接近幂函数分布，是在这 14 个系统中分布最不均匀的系统。说明在世界语言网中，仅有少数节点的权重是很大的，而绝大多数节点的

点权都是很小的。事实上，对于不同的网络， $\frac{\alpha}{\langle w \rangle}$ 才是具有可比性的，由于整个网络的节点总权是归一的，所以 $\langle w \rangle = \frac{1}{N}$ ，在 6.3 节中，表格 6.3.1 列出了 14 个实际网络的 $\frac{\alpha}{\langle w \rangle}$ 的数值，从中我们也能得到世界语言网的 $\frac{\alpha}{\langle w \rangle}$ 数值是最小的，为 0.009。

5.3 差异性与 SPL 分布参数 α 、 γ

5.3.1 差异性的概念

点权分布是否均匀还可以通过差异性来反映。差异性用 Y 表示， Y_l 用来表示节点项目权的差异性，其定义式为：

$$Y_l = \frac{\sum_{i < j, i, j \in V_l} (w_{ii} - w_{jj})^2}{(T_l - 1) \sum_{m=1}^{T_l} w_{lm}^2}$$

其中， V_l 是项目 l 中的节点的集合， w_{ii} 、 w_{jj} 分别为节点 i 、节点 j 在项目 l 中的权重， T_l 为项目 l 的项目大小。则整个网络的点权差异性 Y 为：

$$Y = \frac{\sum_l Y_l}{N_l}$$

其中， N_l 为项目总数。差异性 Y 的取值范围为 $0 \leq Y \leq 1$ ， Y 的值越大表示点权分布越不均匀。若 $Y=0$ ，表示点权分布完全均匀，每个节点的点权都是相同的，此时差异性最小。若 $Y=1$ ，表示点权分布完全不均匀，只有一个节点获得了所有的点权，其他所有节点的点权均为 0，此时差异性最大。显然， $Y=0$ 或 $Y=1$ 是点权分布的两种极端情况，通常情况下， Y 的取值应介于 0 和 1 之间。

表 5.3 中列出了 14 个实际合作竞争网的各参数值，表格的最后一列为各网络的差异性值。从表中我们可以看出，世界语言网的差异性为 0.862，是这 14 个网络中最大的。这说明，说各种语言的人数的分布是极不均匀的，如英语、汉语等大语种，占压倒性优势，英语是世界通用的语言，因此说英语的人数是很多的。

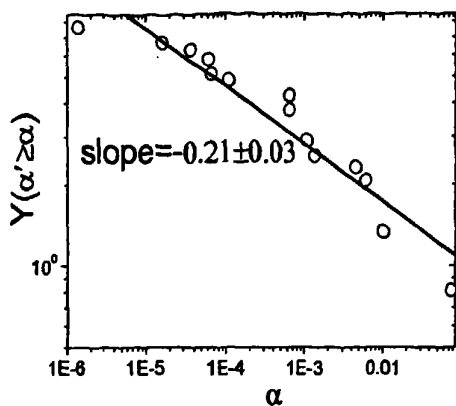
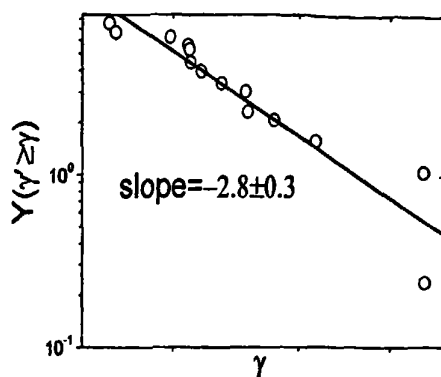
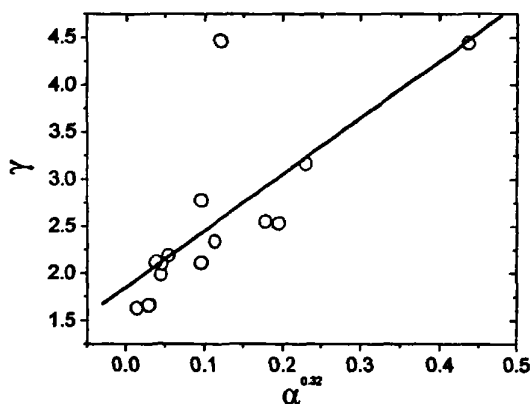
汉语是中国的官方语言，中国的人口是很多的，因此，说汉语的人也很多。但是，其他地区尤其是较偏僻的地区的人们说的语言，几乎只在当地使用，所以说这些语言的人数是很少的。IT 产品网的差异性是最小的，值为 0.238，说明 IT 产品网的点权分布相对来说是比较均匀的。

表 5.3 14 个实际合作竞争网的各参数值

实际网络	项目数	节点数	α	$\alpha / \langle w \rangle$	γ	Y
水果网	151	45	0.075	3.375	4.44	0.809
IT 产品网	265	2121	0.00135	2.863	4.465	0.238
选课网	78	121	0.01	1.21	3.16	0.520
高校自主招生网	52	1547	0.00065	1.006	2.773	0.492
穴位网	108	187	0.0045	0.842	2.55	0.242
电子百强	1	100	0.006	0.600	2.53	0.734
混合饮料网	7804	1501	0.0011	1.651	2.34	0.348
物理图书借阅网	227	3207	0.00011	0.333	2.19	0.582
软件下载网	128	3562	0.000061	0.217	1.983	0.662
中国培训机构网	398	2674	0.000065	0.174	2.1	0.308
北京餐馆网	688	3337	0.0000365	0.122	2.118	0.512
餐饮百强	1	100	0.00065	0.065	2.11	0.860
大陆电影网	920	3084	0.0000161	0.050	1.663	0.418
世界语言网	236	6142	0.0000014	0.009	1.631	0.862

5.3.2 Y 、 α 和 γ 之间的相互关系

图 5.3.1 显示的是差异性与 SPL 参数 α 之间的关系，图中每个点代表一个实际系统，由图可以看出 Y 与 α 之间是幂律关系，可以根据累计与不累计的关系近似推导出 $Y \propto \alpha^{-1.21}$ 。图 5.3.2 显示的是差异性与 SPL 参数 γ 之间的关系，类似地，图中每个点代表一个实际系统，由图可以看出 Y 与 γ 之间也是幂律关系，同理可以推导出 $Y \propto \gamma^{-3.8}$ 。因此，由 $Y \propto \alpha^{-1.21}$ 和 $Y \propto \gamma^{-3.8}$ 可以得出， $\gamma \propto \alpha^{0.32}$ 。图 5.3.3 是 14 个实际系统的实证统计结果拟合得到的曲线，除了一个系统（IT 产品网）偏离较远外，其他 13 个系统的拟合结果还是比较好的。

图 5.3.1 差异性与 α 的关系图 5.3.2 差异性与 γ 的关系图 5.3.3 SPL 分布参数 α 、 γ 间的关系

5.4 讨论

通过实证研究，我们发现世界语言网等 14 个实际网络的点权分布均为 SPL 分布，这很可能具有一定的普遍性。而且，我们发现 SPL 的两个关键参数 α 和 γ 是正相关的，即 γ 随着 α 的增大而增大。SPL 分布的参数 α 值的大小可以反映出点权分布的均匀程度， α 越小点权分布越不均匀，反之， α 越大点权分布相对来说越均匀。节点点权的差异性从另一个角度反映了点权分布的均匀程度，差异性的值越接近于 1，点权分布越不均匀，反之，差异性的值越接近于 0，点权分布越均匀。从两个角度去看这 14 个实际系统的点权分布均匀程度，得到的结论是一致的（即都得出世界语言网的点权分布最不均匀），从一定程度上验证了这两种方法的正确

性。

同时，我们也发现 $Y(\alpha)$ 和 $Y(\gamma)$ 都是幂函数关系，并由此能推导出 γ 和 α 这两个参数这间的关系。 Y 、 γ 、 α 这几个量对于我们所研究的 14 个实际合作竞争网都是逐渐单调变化的，其中是否存在一个动力学机制，这将是一个值得我们进一步探索的问题。

参考文献：

- [1] Wasserman S and Faust K. *Social Network Analysis: Methods and Applications* (Cambridge: Cambridge Univ. Press) (1994)
- [2] Xuan Qi, Li Yan-Jun and Wu Tie-Jun. Does the Compelled Cooperation Determine the Structure of a Complex Network? *Chin. Phys. Lett.* 25 (2) 363 (2008)
- [3] Watts D J and Strogatz S H. Collective dynamics of 'small-world' networks. *Nature* 393 440 (1998)
- [4] Albert-László Barabási, Réka Albert. Emergence of Scaling in Random Networks. *Science* Vol. 286. No. 5439, 509 – 512 (1999)
- [5] Newman, M.E.J. Scientific collaboration networks. I. Network construction and fundamental results. *Phys. Rev. E* 64, 016131 (2001); Newman, M.E.J. Scientific collaboration networks. II. Shortest paths. weighted networks and Centrality. *Phys.Rev. E* 64, 016132 (2001)
- [6] Barabasi A-L, Jeong H, Neda Z et al. Evolution of the social network of scientific collaboration. *Physica A* 311: 590-614 (2002)
- [7] Zhang, P.P., Chen, K., He, Y., et al. Model and empirical study on some collaboration networks. *Physica A* 360, 599–616 (2006)
- [8] Su, B-B., Chang, H., Chen, Y-Z., He, D-R. A game theory model of urban public traffic networks. *Physica A* 379, 291–297 (2007)
- [9] Chang, H., Su, B-B., Zhou, Y-P., He, D-R. Assortativity and act degree distribution of

- some collaboration networks. *Physica A* 383, 687–702 (2007)
- [10] Fu, C-H., Zhang, Z-P., Chang, H., et al. A kind of collaboration-competition net-works. *Physica A* 387, 1411–1420 (2008)
- [11] 刘爱芬, 付春花, 张增平, 等. 中国大陆电影网络的实证统计研究[J]. 复杂系统与复杂性科学, 2007, 4(3): 10-16.

致 谢

首先，我要特别地感谢我的导师何大韧教授，感谢他耐心地将我引入科研领域，感谢他三年来无论是生活上还是学习上给予了我很多的关爱和帮助。在此，向何老师表达由衷的尊敬和谢意。无声润物三春雨，有志育才终不悔，何老师专注的科研精神、高尚的品格和平易亲和的人生态度深深地感染了我，并将影响着我今后的人生道路。

感谢我的父母，是他们养育了我、栽培了我，感谢他们一直以来对我学业的支持和鼓励。

感谢物理科学与技术学院的各位领导、老师在这几年中对我诸多的教导和帮助。

感谢陪伴我一起度过这几年美好岁月的同学和朋友们，是他们让我感受到了友情和温暖。

感谢扬州大学复杂性科学研究中心的全体成员，感谢他们给予我方方面面的关照和帮助。本论文报告的研究内容包含了非常大的工作量，经过了几届研究生和本科生劳动成果的积累才得到今天的成果。感谢师姐常慧、师姐沈丹、师兄潘四军在论文研究工作中给予的支持和帮助，特别感谢师兄周月平、师兄顾长贵在编程方面给予的帮助。本论文的工作中，包含了徐秀莲老师、师姐张培培、师兄周月平、王永利、石建军、李凯军、瞿艳青、冯爱霞、师妹刘爱芬等人的辛勤劳动和帮助。在此，一并致以我深深的谢意！

特别感谢徐秀莲老师在研究工作和日常生活中给予了我很多的帮助和关照。

特别感谢张增平老师在程序编写中给予的大力支持和帮助。

最后，谨以这篇论文作为一份微薄的礼物，

敬献给：

我的父母——是他们用无私的爱养育者我；

我的各位老师——是他们用自己的生命之光，照亮了我！

攻读学位期间发表的学术论文

- 1、A kind of collaboration-competition networks, Chun-Hua Fu, Zeng-ping Zhang, Hui Chang, Jian-Ru Tao, Zhuo-Hui Chen, Yun-Long Dai, Wei Zhang, Da-Ren He, *Physica A* 387 (2008) 1411-1420
- 2、Node Weight Distribution and Disparity of Some Collaboration-Competition Networks, FU Chun-Hua (付春花), XU Xiu-Lian (徐秀莲), LIU Ai-Fen (刘爱芬), WU Yong-Ping (吴永萍), SHEN Dan (沈丹), LIU Shui-Jing (刘水晶), QIAN Xia (钱霞), FENG Ya-Chao (冯雅超), WEI Cheng-Liang (韦成亮), HE Da-Ren (何大韧), *Chinese Physics Letters*, Vol.25, No.11 (2008) 4181
- 3、合作-竞争网络的二分图描述建议, 付春花、刘爱芬、陈卓辉、何大韧, 中国数学力学物理学高新技术交叉研究会第12届学术年会, 科学出版社出版
- 4、Recognition of important subgraphs in collaboration networks, Chun-Hua Fu, Yue-Ping Zhou, Xiu-Lian Xu, Hui Chang, Ai-Xia Feng, Jian-Jun Shi, and Da-Ren He, *Proceedings of the FIRST INTERNATIONAL CONFERENCE ON COMPLEX SCIENCES: THEORY AND APPLICATIONS*, 23–25 February 2009, Shanghai, China, paper No. 5264. Published by: Springer in the series of ICST Lecture Notes, ISBN: 978-963-9799-46-2 (EI).
- 5、Connectivity correlations in three topological spaces of urban bus-transport networks in China, Chen Yong-Zhou (陈永洲), Fu Chun-Hua (付春花), Chang Hui (常慧), Li Nan (李南), He Da-Ren (何大韧), *Chinese Physics B*, Vol.17 No.10 (2008) 3580-08
- 6、中国大陆电影网络的实证统计研究, 刘爱芬, 付春花, 张增平, 常慧, 何大韧, *复杂系统与复杂性科学*, 第4卷第3期, 10-16
- 7、Empirical study on course selection network of Yangzhou University, Chun-Hua Fu,

Hui Chang, Da-Ren He, Bulletin of APS (2007)

- 8、四个合作竞争系统的特殊合作性质, 刘水晶, 钱霞, 冯雅超, 韦成亮, 付春花, 徐秀莲, 刘爱芬, 吴永萍, 何大韧, 科技导报, 2008, 26 (21)
- 9、城市公交的网络操纵者博弈模型, 常慧, 冯爱霞, 付春花, 中国高等科学技术中心系列文集 CCAST-06-12