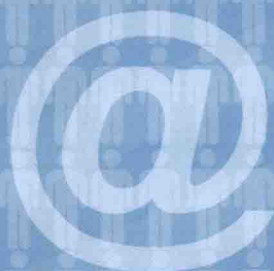


复杂网络中的人类 双向选择行为研究

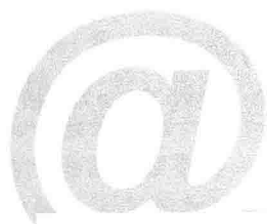
周斌 著



大学出版社
UNIVERSITY PRESS

复杂网络中的人类 双向选择行为研究

周斌 著



 江苏大学出版社
JIANGSU UNIVERSITY PRESS

镇江

图书在版编目(CIP)数据

复杂网络中的人类双向选择行为研究 / 周斌著. —
镇江: 江苏大学出版社, 2016. 5
ISBN 978-7-5684-0223-1

I. ①复… II. ①周… III. ①双向选择—计算模型—
研究 IV. ①F241.32

中国版本图书馆 CIP 数据核字(2016)第 127702 号

复杂网络中的人类双向选择行为研究

Fuza Wangluo zhong de Renlei Shuangxiang Xuanze Xingwei Yanjiu

著 者/周 斌

责任编辑/柳 艳

出版发行/江苏大学出版社

地 址/江苏省镇江市梦溪园巷 30 号(邮编: 212003)

电 话/0511-84446464(传真)

网 址/http: //press. ujs. edu. cn

排 版/镇江文苑制版印刷有限责任公司

印 刷/虎彩印艺股份有限公司

经 销/江苏省新华书店

开 本/890 mm×1 240 mm 1/32

印 张/6. 625

字 数/200 千字

版 次/2016 年 5 月第 1 版 2016 年 5 月第 1 次印刷

书 号/ISBN 978-7-5684-0223-1

定 价/36. 00 元

如有印装质量问题请与本社营销部联系(电话: 0511-84440882)

前 言

在日常生活中,人类的双向选择行为是非常普遍的。为了研究隐藏在人类双向选择行为背后的机制和规律,本书建立了一个关于任意双向选择系统的最小模型,并且得到了这个模型的解析解。模型包括三个影响双向选择的主要因素,分别是双向选择双方各自的个体总数 k_1, k_2 和双向选择双方个体所共有的可供选择的总特性数 n 。在三种不同的条件下,我们对模型的精确解释式做了近似分析,得到了更简洁的函数关系:配对成功率 p 和相关变量成简单的反比关系,并和模型的仿真结果一致。现实中男女之间的爱情选择是典型的双向选择问题,为了验证模型在现实中的有效性,笔者搜集了现实中 82 组关于相亲大会的实证数据。由于相亲大会提供的数据信息不完整,只有参加的总人数和配对成功总对数,没有提供详细的男女双方各自的人数,所以笔者无法用模型和实证数据做精确的对比,只能对模型和实证数据做一个粗糙的验证,结果发现大多数实证数都落在了模型可以预测的范围内,但是当参加相亲大会的总人数相对较少时,会有少量实证数据落在模型所能预测的区域范围之外,对此做出了合理的解释。因此,总体上讲,实证数据理论上验证了模型的合理性,这个双向选择模型虽然简单,但对了解现实中双向选择系统的内在机制有一定的价值和现实意义。

上面的研究属于双向选择在复杂系统和网络科学领域的一个基础性初步研究,其双向选择机制是建立在一个全联通网络上的。为了进一步加深对双向选择课题的研究,本书将双向选择机制拓展到了任意一个网络上,包括主流的网络模型及现实中的网络,并且给出了模型在任意一个网络上的解析解,模型仿真结果和模型的数学解析结果吻合得非常好,证明了解析结果的精确性和正确性。笔

者主要选做了双向选择机制在三个主流网络上的研究：ER 随机网络、WS 小世界网络、BA 无标度网络。经过研究发现，一个个体的邻居的数量和状态直接决定了个体的匹配成功率，而且网络的平均度越高，个体的匹配概率也越高。在三个主流网络模型中，我们发现 WS 小世界网络的拓扑结构最有利于匹配，而 BA 无标度网络的拓扑结构最不利于匹配。WS 小世界网络是以现实中社交网络为基础研究建立起来的，那么双向选择机制研究可以暗示现实中社交网络的演化形成很可能遵守着有利于异性个体之间的匹配这一基本规则。这项研究工作将现实中的双向选择研究更推进了一步，让对双向选择的基础研究从一开始的特殊性网络（全联通网络）拓展到任意一个网络上，这对后续关于现实中双向选择现象的理论和应用研究有一定的意义和作用。

目 录

第 1 章 绪论 / 001

- 1.1 研究背景 / 001
- 1.2 网络的描述 / 003

第 2 章 复杂网络基础知识 / 005

- 2.1 网络的概念 / 005
- 2.2 网络的分类 / 007
- 2.3 网络的表示 / 011
- 2.4 网络拓扑性质 / 013
- 2.5 网络模型 / 019

附录 A 常用基本网络模型程序示例 / 027

第 3 章 人类双向选择行为模型 / 046

- 3.1 引言 / 046
- 3.2 模型和解析 / 046
- 3.3 实证数据验证 / 052
- 3.4 本章小结 / 063

第 4 章 双向选择机制在复杂网络中的研究 / 065

- 4.1 引言 / 065
- 4.2 模型和仿真方法 / 066
- 4.3 仿真结果 / 068
- 4.4 本章小结 / 075

附录 B 双向选择机制在复杂网络中的模型程序示例 / 077

第 5 章 描述网络拓扑的一种新方法 / 099

5.1 研究背景/ 099

5.2 描述网络结构特征的新方法/ 100

5.3 新方法在复杂网络中的应用/ 103

5.4 本章小结/ 107

附录 C 网络模型一层度和二层度相似性系数程序示例/ 109

第 6 章 多人多任务时间优化策略研究 / 154

6.1 研究背景/ 154

6.2 模型介绍/ 155

6.3 仿真结果/ 155

6.4 本章小结/ 161

附录 D 多人多任务时间优化策略研究程序示例/ 163

第 7 章 研究结论与展望 / 178

7.1 研究结论/ 178

7.2 工作展望/ 181

附录 E 有关的数学知识/ 185

参考文献 / 189

后 记 / 201

第1章 绪 论

1.1 研究背景

由几十亿人组成的社会系统非常庞大而且复杂,人们越来越深刻地认识到认知人类社会自身规律的重要性。随着科学技术和计算机网络技术的不断发展,人们开始有能力记录和统计关于人类社会行为的海量数据。挖掘这些大数据,对其进行统计特征分析,探索可能存在的动力学机制和内在机理,建立各种模型,可以帮助人们更加深入地理解和认知各种相关的人类社会行为。这方面,国内外众多科研工作者已经做了大量的研究工作,这些工作主要体现在时间统计特性方面^[1-17]和空间统计特性方面^[18-25]。其中很多新发现颠覆了人们的传统认知^[26,27],从而引发了探索人类社会各种行为规律的研究热潮,并建立了很多动力学模型来解释人类社会各种行为现象,阐述隐藏于行为背后的内在机制^[28-43]。当前对于人类社会行为的复杂性研究如火如荼,大量的研究和发现具有重要的理论意义和应用价值,也必将对人类的生活和发展产生重大影响。

到目前为止,对于人类行为的传统研究主要集中在时间和空间方面。人类社会极其复杂,很多隐藏在人类行为背后的内在规律和机制仍然不为人所知。因此,要想对人类社会行为有更加深入、客观、全面的认识和了解,就必须从多方面、多角度,突破常规去研究。人类社会双向选择行为在社会中是非常普遍的,同时也是非常复杂的。例如每年大学毕业生的就业和公司的招聘,人类社会日常生活中男女之间的情感和婚姻匹配,以及商业活动中的买方和卖方之间的需求互换等。图 1-1 和图 1-2 分别是招聘大会和相亲大会场景。

这些非常重要而又庞杂的社会现象归根结底都是双向选择问题。迄今为止,我们对人类双向选择行为规律的认知还不够深入和重视,研究成果比较单一,缺乏全面性、系统性。如果我们能对现实中双向选择系统的大数据进行统计特征分析,发现双向选择系统内在的机理和动力学机制,那么就可以对现实中的各种双向选择系统进行建模预测。通过评估预测模型的反馈信息,改进模型规则,最终找到提高模型预测结果的手段和方法,建立高精度预测模型。这将加深和拓展人们对人类行为的认知和理论研究,也可以在大学生就业、家庭婚姻、商业贸易活动等领域产生重要的指导作用 and 实际应用价值。



图 1-1 大学生就业招聘大会



图 1-2 男女择偶相亲大会

目前很少有用复杂性科学的统计研究方法和手段对现实中双向选择系统进行全面和深入的研究。Resnick、周涛和吕林媛等对推荐系统进行了研究^[44-46],认为推荐系统可以被看作是双向选择中的一种特殊形式。在男女婚姻方面,Michael Dzierzawa, Oméro M J, Caldarelli G 和张翼成等做了一些相关的研究工作^[47-53],对现实中男女之间稳定婚姻匹配的内在机制进行了研究和探索,提出权重性优先排序互选模型机制。该方法的优点是比较直观,容易理解,符合日常人们的感性认识,对于理解男女之间在婚姻方面的双向选择行为有一定的现实意义,丰富了人们在人类双向选择行为方面的认知和理论,但是定量描述能力有所欠缺,没有从本质上给出男女两个个体之间在婚姻双向选择上的机制和规律,也不能

对男女个体之间在婚姻双向选择方面的多种现象(例如关于结婚就有三种现象:男女双方都愿意结婚;只有一方愿意;双方都不愿意)给出明确合理的解释。因此,需要对实证数据进行更加深入的统计特征分析,从全新的视角探索双向选择行为背后的内在机理,建立更加符合现实、具有应用价值的预测模型。

1.2 网络的描述

笔者在2014年初发表的文章^[54]中提出了一种双向选择机制来解释人类社会中的双向选择行为,在全联通网络结构上,建立了一个最小的双向选择模型。此种机制可以定量描述男女之间在婚姻方面的双向选择行为,也可以对现实中存在的多种现象进行明确合理的解释,例如:男女双方都愿意结婚(机制中的两个条件都满足);只有一方愿意(只满足机制中的一个条件);双方都不愿意(机制中的两个条件都不满足)。笔者搜集了现实中不同时间、不同地点82组相亲大会的数据,对模型进行了验证和评估,发现最小双向选择模型能够在一定程度上预测相亲大会的男女配对成功率。若要得到更高精度的预测模型,需要进一步对人类双向选择行为大数据进行深入探索和分析。

笔者在2014年发表的文章^[55]中,将双向选择机制拓展到任意复杂网络上,包括各种网络模型(ER随机网络、WS小世界网络、BA无标度网络)及现实世界中的各种真实网络,使得工作^[54]变成了文章中的一个特例。研究表明,网络结构对双向选择个体匹配率有重要影响。可以发现,WS小世界网络的拓扑结构最有利于匹配,而BA无标度网络则最不利于匹配。这似乎表明,WS小世界网络的形成对应于现实生活中社交网络的演化形成是朝着更有利于异性个体之间匹配趋势进行的。笔者通过断边重连改变WS小世界的网络结构,发现双向选择系统的匹配结果随断边重连概率之变化而呈现丰富多样的变化。其内在原因及与现实中社交网络结构之间的关系,有待在后续工作中进一步深入研究。

综上所述,目前用复杂网络科学的方法和手段对社会中普遍存在的人类双向选择行为的研究还处于刚刚起步的阶段,对双向选择系统的研究是一项非常具有创新性的课题,通过对现实中双向选择系统大数据进行深入的统计特征分析,发现其内在规律和动力学机制,具有重要的理论价值,在人类社会生活中也具有潜在的应用价值。因此,未来对人类双向选择行为的研究很可能会开启一个新的研究热潮。

第 2 章 复杂网络基础知识

2.1 网络的概念

在人类社会和自然界中有着各种数不清的大大大小小的系统,大到天文领域的星系系统(见图 2-1),小到一个小的神经元细胞(见图 2-2)。对于每一个系统来说,包含了数量众多的个体。从某一个角度出发研究这些系统中个体之间的相互作用及整个系统的特征,每个个体可以被抽象成一个点,用边来表示个体与个体之间存在的相互作用,那么整个系统就被抽象成一个网络。例如图 2-3(a)是一个局域网计算机设备之间的连接图,把每一个设备看成是一个点,设备与设备之间的连接关系用一条边来代替,那么就得到一个局域网的网络结构,如图 2-3(b)所示。



图 2-1 后发座 NGC 4414 螺旋星系(直径 55 000 光年,距离地球 6 520 万光年)

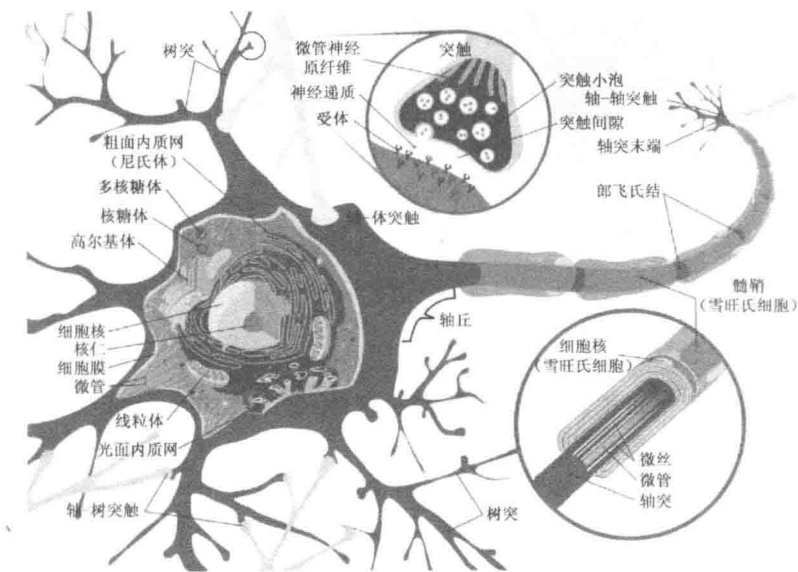


图 2-2 神经元细胞结构示意图

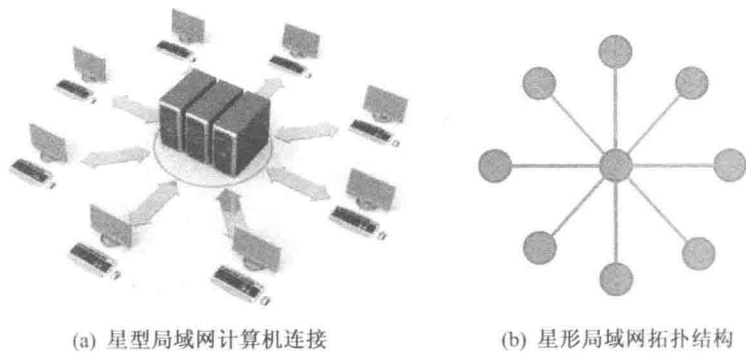


图 2-3 局域网的拓扑结构示意图

按照研究领域分类,现实中的网络主要可以分为 Internet 网络、万维网络、生物网络、神经网络、宇宙天体网络、社会网络、经济与金融网络、电力与交通网络、科研和教育网络等几大领域。这些实际系统与网络的对应关系,见表 2-1。通过把现实系统抽象网络化,可

以研究系统的结构复杂性、节点复杂性、结构与节点之间的相互影响及网络之间的相互影响等重要特征,从而可以让我们对自然及社会有更加深入和全面的理解。

表 2-1 实际系统与网络的对应关系^[56]

系统	节点	边
互联网	计算机、服务器、路由器等	电缆、光纤、无线连接等
万维网	网页	超链接
引文系统	论文、书籍等	引用关系
电力系统	发电厂、电站、中转站等	电缆、电线等
交通系统	汽车站、火车站、飞机场等	公路、铁路、航空路线等
社交系统	人	人际关系
经济系统	部门、企业、个人等	贸易、资金关系等
代谢系统	代谢物	代谢反应
神经系统	神经元	突触
食物链	生物	捕食关系
晶体	原子	化学键

2.2 网络的分类

根据本书研究内容的需要,在此主要给出有向网络与无向网络、加权网络与无权网络、加权有向网络与无权无向网络、稀疏网络与稠密网络,以及二分图网络的定义。

2.2.1 有向网络和无向网络

在有些现实系统中,个体与个体之间的相互作用对双方而言并不是同时存在的,例如在人与人之间的社交网络中,很多人知道某些知名人士,但是这些知名人士却不可能认识所有人,那么这种人与人之间的相互作用关系就具有单向性。因此按照网络结构中边是否有向,可以将网络分为有向网络和无向网络,如图 2-4 所示。

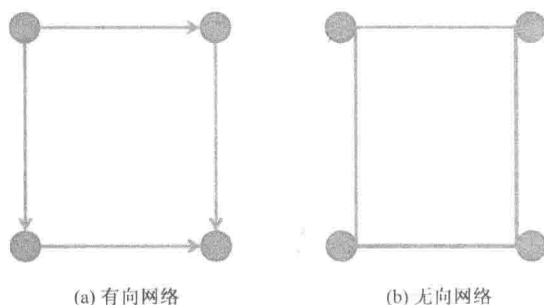


图 2-4 有相网络和无向网络示意图

2.2.2 加权网络和无权网络

在有些现实系统中,不同个体之间的相互作用强弱不同,也可以理解为不同个体间的相互作用在整个网络中的强弱或者重要性不同。例如一个城市的交通道路网络中,有些道路属于主干道,承担着巨大的车流量,在城市整个交通网络中占有重要地位;有些道路属于支路,只有少量车通行,那么在整个交通网络中的地位就很弱。因此按照网络结构中不同边的强弱及重要性的不同,可以将网络分为加权网络和无权网络,如果网络中所有边的相对重要性都一样,或者地位一样,那么每条边的权值可以定为 1,也可以默认不标,即为无权网络。如果网络中各边的相对重要性不一样,地位有高有低,即为加权网络,如图 2-5 所示。

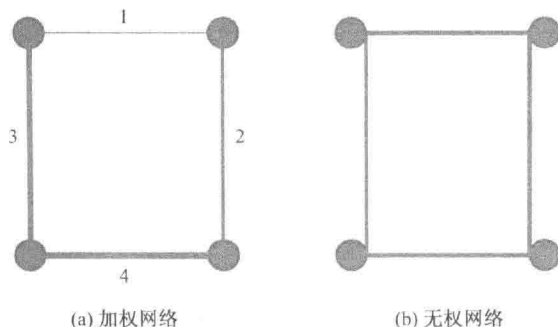


图 2-5 加权网络和无权网络示意图

2.2.3 加权有向网络和无权无向网络

根据上面两节的定义,加权有向网络是指网络中的边既有向又有权;无权无向网络是指网络中的边无权也无向,如图 2-6 所示。本书涉及的网络皆为无权无向网络,且网络内无同一节点的自环连接和两不同节点之间的重连。

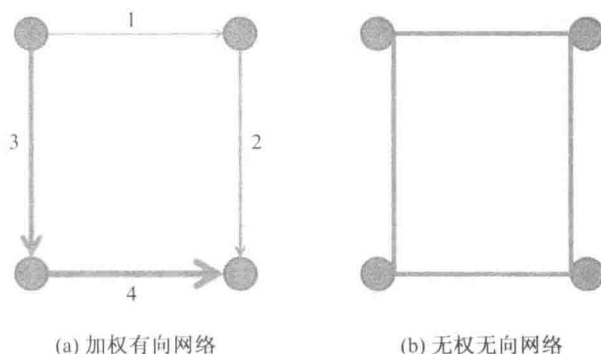


图 2-6 加权有向网络和无权无向网络

2.2.4 稀疏网络和稠密网络

通过定义网络密度的概念可以将网络分为稀疏网络与稠密网络。网络密度是网络中现有的边数与网络中可以存在的最大边数之比,用符号 ρ 标记,公式如下:

$$\rho = \frac{M}{\frac{1}{2}N(N-1)} \quad (2-1)$$

式(2-1)中, M 表示网络中现有的边数, N 表示整个网络的节点数, $N(N-1)/2$ 表示网络中可以存在的最大边数。 $\rho=0$,意味着网络中无边; $\rho=1$,意味着网络为全联通网络,即网络中任意两点之间都直接相连。

当网络规模 $N \rightarrow \infty$ 时,如果 ρ 满足:

$$\rho \rightarrow 0 \quad (N \rightarrow \infty) \quad (2-2)$$

意味着网络中现有的边数是比 N^2 低阶的,也就是说网络中现有的边数远小于网络中可以存在的最大边数,那么这种网络称之为稀疏网络,现实中大多数网络皆为稀疏网络^[57,58],故科学研究中的网络也以稀疏网络为主。

当网络规模 $N \rightarrow \infty$ 时,如果 ρ 满足:

$$\rho \rightarrow c (0 < c \leq 1) \quad (N \rightarrow \infty) \quad (2-3)$$

意味着网络中现有的边数跟 N^2 是同阶的,那么这种网络称之为稠密网络。

2.2.5 二分图网络

对于一个网络来说,如果网络中所有节点可以被分成没有交集的两个集合,网络中的每一条边两端所连接的两个点分别属于这两个不同的集合,那么这种网络结构可以被定义为二分图网络,如图 2-7 所示。

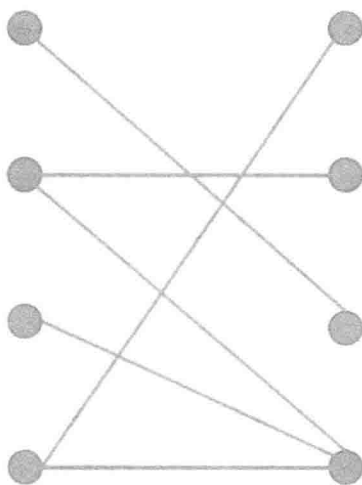


图 2-7 二分图网络