



Social Network Analysis for Startups

社会网络分析领域经典著作，
系统阐述社会网络分析的理论、方法与实践



YZLI0890190539

社会网络分析

方法与实践

O'REILLY®



机械工业出版社
China Machine Press

Maksim Tsvetovat & Alexander Kouznetsov 著

王薇 王成军 王颖 刘璟 译

阳志平 审校

社会网络分析：方法与实践

Maksim Tsvetovat

Alexander Kouznetsov 著

王薇 王成军 王颖 刘璟 译

阳志平 审校



YZLI0890190539

O'REILLY®

Beijing • Cambridge • Farnham • Köln • Sebastopol • Tokyo

O'Reilly Media, Inc. 授权机械工业出版社出版

机械工业出版社

图书在版编目(CIP)数据

社会网络分析:方法与实践/(美)茨韦特瓦(Tsvetovat, M.)等著;王微等译.
—北京:机械工业出版社,2013.5

(O'Reilly精品图书系列)

书名原文:Social Network Analysis for Startups

ISBN 978-7-111-42341-6

I. 社… II. ①茨… ②王… III. 社会关系—研究 IV. C912.3

中国版本图书馆CIP数据核字(2013)第090300号

北京市版权局著作权合同登记

图字:01-2012-5213号

Copyright ©2011 Maksim Tsvetovat and Alexander Kouznetsov.

Simplified Chinese Edition, jointly published by O'Reilly Media, Inc. and China Machine Press, 2013. Authorized translation of the English edition, 2011 O'Reilly Media, Inc., the owner of all rights to publish and sell the same.

All rights reserved including the rights of reproduction in whole or in part in any form.

英文原版由O'Reilly Media, Inc. 出版2011。

简体中文版由机械工业出版社出版2013。英文原版的翻译得到O'Reilly Media, Inc.的授权。此简体中文版的出版和销售得到出版权和销售权的所有者——O'Reilly Media, Inc.的许可。

版权所有,未得书面许可,本书的任何部分和全部不得以任何形式复制。

封底无防伪标均为盗版

本书法律顾问

北京市展达律师事务所

书 名/ 社会网络分析:方法与实践

书 号/ ISBN 978-7-111-42341-6

责任编辑/ 秦健

封面设计/ Karen Montgomery, 张健

出版发行/ 机械工业出版社

地 址/ 北京市西城区百万庄大街22号(邮政编码 100037)

印 刷/ 北京市荣盛彩色印刷有限公司印刷

开 本/ 178毫米×233毫米 16开本 12印张

版 次/ 2013年6月第1版 2013年6月第1次印刷

定 价/ 59.00元(册)

凡购本书,如有缺页、倒页、脱页,由本社发行部调换

客服热线:(010)88378991 88361066

购书热线:(010)68326294 88379649 68995259

投稿热线:(010)88379604

读者信箱:hzjsj@hzbook.com

O'Reilly Media, Inc.介绍

O'Reilly Media通过图书、杂志、在线服务、调查研究和会议等方式传播创新知识。自1978年开始，O'Reilly一直都是前沿发展的见证者和推动者。超级极客们正在开创着未来，而我们关注真正重要的技术趋势——通过放大那些“细微的信号”来刺激社会对新科技的应用。作为技术社区中活跃的参与者，O'Reilly的发展充满了对创新的倡导、创造和发扬光大。

O'Reilly为软件开发人员带来革命性的“动物书”；创建第一个商业网站（GNN）；组织了影响深远的开放源代码峰会，以至于开源软件运动以此命名；创立了Make杂志，从而成为DIY革命的主要先锋；公司一如既往地通过多种形式缔结信息与人的纽带。O'Reilly的会议和峰会集聚了众多超级极客和高瞻远瞩的商业领袖，共同描绘出开创新产业的革命性思想。作为技术人士获取信息的选择，O'Reilly现在还将先锋专家的知识传递给普通的计算机用户。无论是通过书籍出版，在线服务或者面授课程，每一项O'Reilly的产品都反映了公司不可动摇的理念——信息是激发创新的力量。

业界评论

“O'Reilly Radar博客有口皆碑。”

——Wired

“O'Reilly凭借一系列（真希望当初我也想到了）非凡想法建立了数百万美元的业务。”

——Business 2.0

“O'Reilly Conference是聚集关键思想领袖的绝对典范。”

——CRN

“一本O'Reilly的书就代表一个有用、有前途、需要学习的主题。”

——Irish Times

“Tim是位特立独行的商人，他不光放眼于最长远、最广阔的视野并且切实地按照Yogi Berra的建议去做了：‘如果你在路上遇到岔路口，走小路（岔路）。’回顾过去Tim似乎每一次都选择了小路，而且有几次都是一闪即逝的机会，尽管大路也不错。”

——Linux Journal

译者序

——社会网络分析：探索关系背后的科学与技术

在一个社交网络中，谁是明星人物？谁是边缘人物？谁是八卦传播者？谁是沟通两个不同网络的搭桥者？谁是真正的实权人物，在社交网络中沉默却拥有极大权力？

为什么当基地恐怖分子发动攻击时，每组人数不能超过六人？为什么网上病毒传播关键临界节点是7%，当一条微博渗透率突破它，才从线性增长变为病毒传播式增长？

为什么在脸谱一步一步起家时，都先要在一个小范围内吸引社群关注达到6%的饱和度，才考虑从一个小的社群（如哈佛大学），跃迁到一个更大的社群（如常青藤高校）中去呢？

本书阐释了帮助我们理解这些现象背后的科学与技术。对各类关系谜题的研究，构成了人类学术史波澜壮阔的一面。始于20世纪30年代，由社会心理学家莫雷诺创立，社会学、经济学研究者发扬光大的学派——社会网络分析（Social Network Analysis, SNA）开启了一条将关系图论化的新道路。

社会网络分析将社会关系看做节点（node）和连带（tie）组成的网络（network）。节点表示网络内的独立行动者（actor），网络则用来表示行动者之间的关系。行动者主体可以从个体、组织到国家；同样，关系也可以是人际关系、组织关系与国家关系。如果说传统回归分析等统计技术关心的是个体层面各类属性的统计技术，那么社会网络分析致力于分析关系数据。

推特、脸谱等社交媒体已成为人们生活的一部分，每时每刻，天南海北，都在产生大量实时关系数据。人们第一次在历史上提供了对人类关系图谱在时间与空间上大尺度的取样。社会网络研究者可以轻易地下载分布在全世界数千万甚至上亿红男绿女长达数年的喜怒哀乐。其中蕴涵的价值令每一位社会网络爱好者激动不已。

然而，社会网络分析庞大的学术积累已制造了一个较高的入门门槛。社会网络分析新增的三块砖头：小世界网络、无标度网络与随机网络，更使得它渐渐与物理科学的复杂性研究、计算机科学的网络科学交错在一起，走向“可计算的社会科学”（Computational Social Science）。

初学者如何更好地迈过这个门槛？本书是一本帮助你快速掌握社会网络分析技术要点、核心概念与典型算法示例的优秀图书。本书与市面上既有图书不同，它具备一个鲜明特点：基于开源软件（NetworkX）讲授。

截至今天，还有不少社会网络分析研究者使用UCINET等基于Windows窗口的软件。我们可以将UCINET理解为社会网络分析中的类Excel软件。海量增长的数据、日新月异的计算模型，使得这类商业软件的弊端越来越明显。使用开源软件与商业软件更是在社区支持、知识再生产上差异极大。基于R或Python的开源社会网络分析包，如R语言的SNA或IGRAPH、Python的NetworkX，在设计新模型、开发新算法、分析大数据与进行可重复分析等方面目前已经完胜。

本书以基于Python的网络分析包NetworkX作为社会网络分析工具，但不是一本NetworkX使用手册。作者将重点放在如何从庞大的社会网络分析学术积累中，挑选最精要与最实用的知识点，以帮助读者形成关于社会网络分析的知识谱系图。全书可以分为四部分。第1章和第2章是基础知识，主要介绍社会网络分析的背景信息与图论基础知识。第3~5章主要介绍如何分析社会网络，分别从个体与群体两个层面，介绍社会网络的主要测量指标与分析方法。其中第3章重点介绍社会网络节点层面的四个核心指标：

- 程度中心性：哪些是明星人物？哪些是边缘者？程度中心性回答类似问题。这是最为人们理解的社会网络测量指标。以微博为例，程度中心性就是粉丝的数量，那些程度中心性高的人就是微博中的明星。
- 亲近中心性：亲近中心性通过点与其他点的距离来测量。那些在社交网络中经常与人互动、人际关系颇好的人，比如公司中的八卦传播者，往往亲近中心性得分较高。
- 居间中心性：节点的居间程度，表示一个网络中经过该点最短路径的数量。在网络中，节点的居间程度越大，那么它在节点相互之间的信息传播起到的作用也就越大。在两个社会网络之间的人，比如跨界者，往往拥有较高的居间中心性。
- 特征向量中心性：那些在社交网络中沉默却拥有极大权力的人物，如《教父》中的主人翁柯里昂。社会网络研究者将他们称为“灰衣主教”。特征向量中心性就是找出他们的办法。基本原理是，一个有着高特征向量中心性的行动者，与他建立连接

的很多行动者往往也被其他很多行动者所连接。在社交网络中，有这样一种人，很多明星与其做朋友，即使他沉默不语，也可能是一位重要的人物。

社会网络分析不仅仅在节点层面测量。第4章、第5章介绍如何分析群体。其中，第4章主要介绍社群划分的基础知识：如何将庞大的社会网络划分为小的组块？如何利用社会网络中的结构洞牟利？如何进行三元组普查与分析？例如，如何通过岛屿方法逐步找出推特上埃及革命的成千上万条转发的核心人物？又如，如何评估埃及革命中一个人的信息传播能力？显然，如果你的朋友们相互信任，将比那种一个明星发言，粉丝们单纯收听的星形网络传播能力更强。第5章主要介绍二模网络与多模网络的基础知识。关系还会存在于不同类型的主体之间，比如公司雇佣员工、投资者购买公司股票、人们占有信息与资源等。这些关系称为二模关系。现实生活中的关系往往是二模或多模。比如在微博上，可以通过你的兴趣、地域、使用的标签来为你推荐新的朋友，或者基于你对一些公共事件的看法，将你划分到特定政治群体中，这些都是基于二模或多模网络的分析得出的。

第6章是全书最精彩的部分，关注信息如何传播，初步展示分析动态社会网络发展的建模技巧。一条微博如何从一两个人关注突然成为流行用语？作者在实验中发现，当网络密度接近7%的时候，将从线性增长（每次增加一条连接）转化为病毒式扩散，也就是说，如果转发一条微博、加入一个网络社群等的人数比例达到7%，其他人将会在关键阶段马上跟进。这是一个推动脸谱走出哈佛大学的神奇数字。脸谱一步一步地跃迁，总是遵循一个规则——在一个社群里到达饱和点之后才移入一个更大的社群。作者通过手写Python算法，为读者打开动态社会网络与网络仿真的大门——我们如何用算法来模拟人类社会各类关系的变迁？有了自己亲手实践算法的经验，读者未来使用netlogo等网络仿真软件，将更加得心应手。

对于初学者来说，第4~6章这三章有一定难度，需要同时理解社会科学与编程技巧两方面知识。第7章则简单明了，主要介绍获取网络数据的入门知识。如果希望深入了解，可以阅读作者推荐的相关资源。附录A介绍收集社会网络分析所需数据的传统方法、伦理准则与相关API。附录B介绍如何安装本书涉及的相关软件，如NetworkX、matplotlib等。

总而言之，作为一本技术非常新颖的入门读物，本书通俗易懂，基于Python进行分析使得其灵活性变得更高。可以说，本书令学习者从一开始就具有上手实践的能力，除介绍网络数据获取技巧、网络抽样方法、网络在个体与群体两个层面的基本属性之外，还涉及目前日益热门的网络模拟方法，融合基础理论与算法于一身。简约却不简单，上升空间非常大！无论你是对社会网络感兴趣的大众读者，还是社会网络的专业研究者、开发者，相信本书都会在社会网络的理论与实践两方面给予启发！

有趣的是，全书通过基于GitHub网站协作，使用Markdown语法完成。全书第1章、第2章由中科院心理所硕士、北京语言大学心理学教师王颖翻译；第3章由中科院心理所硕士、安人公司同事王薇翻译；第4~6章由香港城市大学媒体与传播系博士候选人、互联网挖掘实验室研究员王成军翻译；第7章由香港城市大学媒体与传播系博士候选人刘璟翻译；附录A由北京大学教育学硕士、安人公司同事曲元周翻译；附录B由吉林大学心理学学士、安人公司同事牟百会翻译。全书由阳志平审校成书、统一定稿。

感谢各位参与翻译本书的友人，更感谢协助校阅的豆瓣网友孔明、BOOMER与哈哈。最后特别感谢在组织翻译的过程中，机械工业出版社华章公司编辑吴怡老师与秦健老师。在去年杂事甚多的情况之下，他们一再包容与理解我拖延交稿。他们的细致工作，更使得本书增色不少。希望这本书的面市能够帮助更多人迈入一个广阔迷人的社会网络分析世界。

始生之物，其形必丑。水平有限，不当之处，敬请读者指出。全书如有任何翻译疑问或者关于社会网络分析软件的使用问题，欢迎与我联系。未来有关勘误，也敬请留意华章公司相关网站。我的联系方式是：

- Email: im@yangzhiping.com
- 个人网站: yangzhiping.com

如果读者希望深度了解社会网络分析，可以阅读我们整理的图书资源：

- 阳志平：《社会网络分析入门书目导读》
网址: <http://www.yangzhiping.com/tech/sna-book.html>
- 王成军整理的豆列：《社会网络和复杂网络研究》
网址: <http://book.douban.com/doulist/212560/>
- 阳志平整理的豆列：《社会网络分析技术：探索复杂性与自组织性》
网址: <http://book.douban.com/doulist/1380669/>

阳志平

2013年5月6日于北京

目录

前言	1
第1章 导论	5
分析关系，理解人与群体	6
从关系到网络——超乎所见	7
社会网络与连接分析	8
非正式网络的力量	10
恐怖分子与革命者：社会网络的力量	14
推特上的革命	17
第2章 图论速览	22
什么是图	22
图的遍历与距离	28
图的距离	38
为什么重要	39
六度理论神话	40
小世界网络	40
第3章 中心性、权力与瓶颈	42
样本数据	42
中心性	48
中心性测量不能告诉我们什么	64

第4章 派系、聚类 and 组元	65
组元和子图	65
子图——自我中心网	69
三元组	71
派系	83
分层聚类	85
三元组、网络密度和冲突	92
第5章 二模网络	97
竞选资金是否影响选举	97
二模网络的理论	100
扩展多模网络	109
第6章 信息扩散：像病毒一样传播开来	112
病毒视频剖析	112
信息如何影响网络	119
Python中的一个简单动态模型	123
网络和信息的共同演化	129
第7章 在现实世界中绘图	138
中等规模数据：传统SQL关系数据库	139
大数据：未来，从今天开始	139
小数据——平面文件表达	139
中等规模数据：数据库表达	143
使用二模数据工作	150
社会网络和大数据	152
运行大数据	157
附录A 收集数据	163
附录B 安装软件	173

前言

2011年，很多创业公司都在他们的商业计划中提到“social”这个词——虽然事实上并没有人知道如何分析和理解这能够决定公司成败的社交过程。如果你从事如下工作：社交媒体、社交CRM、社交营销、管理咨询等，你应该读一读这本书，它会告诉你社交系统是如何发展、演变以及运转的。

这本书的内容不仅仅适用于创业公司。实际上，整本书就是一门系统的课程，它囊括了几乎一个学期的理论知识和实际操作材料——阅读以后，你就会对于社会网络分析是“危险的”有足够的理解。如果你是这个研究领域的学生，我们强烈鼓励你去寻找并阅读脚注里提到的每一篇论文或每一本书。这样做会让你非常了解这个领域的经典文献，也可以让你自信地开展研究课题。

如果你有计算机技术背景，可以从这本书学到主要的社会学概念，并从中提取出可以用来编程的信息和分析数据；如果你有社会学或市场营销背景，你会发现一些熟悉的材料，与此同时也会学到用定量和定性的方法去了解社交背景下的人们。

阅读本书的前提

由于这本书的受众非常广泛，因此我们尽可能少地使用专业术语，同时为书中的专业术语做出解释。但是，本书会有大量的技术内容（这是O'Reilly出版图书的一项要求）。

我们希望至少稍微熟悉Python，即，能自己写脚本、了解语言的基本控制结构和数据结构。如果你不了解技术内容，建议你使用Python的在线教程或者参考Paul Barry（O'Reilly出版）写的《Head First Python》开始学习。

本书不会涉及从Twitter、Facebook以及其他数据来源收集数据的详细过程，O'Reilly出版的“Animal Guide”系列的其他书籍已经提供了充分的资料，如Kevin Makice写的《Twitter API: Up and Running》以及Matthew Russell写的《Mining the Social Web》。

开源工具

本书使用开源的Python库，包括：NetworkX、NumPy和MatPlotLib。

- NetworkX工具包和相关文档可以在<http://networkx.lanl.gov>找到。
参考文献：Aric A. Hagberg, Daniel A. Schult and Pieter J. Swart, “Exploring network structure, dynamics, and function using NetworkX”, in *Proceedings of the 7th Python in Science Conference (SciPy2008)*, eds. Gael Varoquaux, Travis Vaught, and Jarrod Millman, pp. 11–15, Aug 2008.
- NumPy工具包和相关文档可以在<http://numpy.scipy.org>找到。
参考文献：Oliphant, Travis E. “Python for Scientific Computing”. *Computing in Science & Engineering* 9(2007), 10-20 Ascher, D. et al. *Numerical Python*, tech. report UCRLMA-128569, Lawrence Livermore National Laboratory, 2001.
- Matplotlib 可以在 <http://matplotlib.sourceforge.net>找到。
参考文献：Hunter, J.D. “Matplotlib: A 2D Graphics Environment”. *Computing in Science & Engineering* 9(2007), 90-95.

本书的约定

本书使用了以下排版约定：

斜体 (*Italic*)

用于新术语、URL、电子邮件地址、文件名与文件扩展名。

等宽字体 (**Constant width**)

用于表明程序清单，以及在段落中引用的程序中的元素，如变量、函数名、数据库、数据类型、环境变量、语句、关键字。

等宽粗体 (**Constant width bold**)

用于表明命令，或者需要读者逐字输入的文本内容。

等宽斜体 (*Constant width italic*)

用于表示需要使用用户提供的值或者由上下文决定的值来替代的文本内容。

注意：这个图标代表一个技巧、建议或一般性说明。

警告：这个图标代表一个警告或注意事项。

示例代码的使用

本书提供代码的目的是帮你快速完成工作。一般情况下，你可以在你的程序或文档中使用本书中的代码，而不必取得我们的许可，除非你想复制书中很大一部分代码。例如，你在编写程序时，用到了本书中的几个代码片段，这不必取得我们的许可。但若将O'Reilly图书中的代码制作成光盘并进行出售或传播，则需获得我们的许可。引用示例代码或书中内容来解答问题无需许可。将书中很大一部分的示例代码用于你个人的产品文档，这需要我们的许可。

注意：所有的代码、数据和示例都可以从GitHub下载：<https://github.com/maksim2042/SNABook>。

如果你引用了本书的内容并标明版权归属声明，我们对此表示感谢，但这不是必需的。版权归属声明通常包括：标题、作者、出版社和ISBN号，例如：“Social Network Analysis for Startups by Maksim Tsvetovat and Alexander Kouznetsov (O'Reilly). 2011 Maksim Tsvetovat and Alexander Kouznetsov, 978-1-449-30646-5”。

如果你认为你对示例代码的使用已经超出上述范围，或者你对是否需要获得示例代码的授权还不清楚，请随时联系我们：permissions@oreilly.com。

联系我们

有关本书的任何建议和疑问，可以通过下列方式与我们取得联系：

美国：

O'Reilly Media, Inc.
1005 Gravenstein Highway North
Sebastopol, CA 95472

中国：

北京市西城区西直门南大街2号成铭大厦C座807室（100035）
奥莱利技术咨询（北京）有限公司

我们会在本书的网页中列出勘误表、示例和其他信息。可以通过<http://www.oreilly.com/catalog/0636920020424>访问该页面。

要评论或询问本书的技术问题，请发送电子邮件到：

bookquestions@oreilly.com

想了解关于O'Reilly图书、课程、会议和新闻的更多信息，请访问以下网站：

<http://www.oreilly.com.cn>

<http://www.oreilly.com>

还可以通过以下网站关注我们：

我们在Facebook上的主页：<http://facebook.com/oreilly>

我们在Twitter上的主页：<http://twitter.com/oreillymedia>

我们在Youtube上的主页：<http://www.youtube.com/oreillymedia>

感谢

Max Tsvetovat 感谢他的老师：Sergei Fyodorovich Ivanov(High School #367, St. Petersburg, Russia)、Maria Gini（明尼苏达大学）、Kathleen Carley和Katya Sycara（卡内基·梅隆大学），教了一个不守规则的学生许多年，最重要的是，教会了他独立思考、主动学习的能力。他感谢家人的鼓励和耐心，在他深夜写作的时候（尤其是小David，整晚都在睡觉，这样可以让父亲安心地写作）。感谢Miles Davis的音乐、NorthSide Social的咖啡、Wardrobe……我像是在开玩笑吗？

Alex Kouznetsov 要感谢很多朋友、老师和大学同学的帮助。他感谢家人的支持，尤其是他的妻子Natalia，感谢她的悉心照顾。还要感谢Nikolai Afanasievich Kolobov(莫斯科电子和数学学院)，在哪里我体验到非正统思想的乐趣、科学的严谨，并使我走上科学研究的道路。感谢老友Javier Sanchez,让我学会了辩证地思考，并主动地为了兴趣和利益去学习。

“额……你是做什么的？”

“我在研究社会网络分析。”

“哦，那就是说你在玩脸谱（Facebook），而且还能靠它赚钱？”

在社会网络分析研究者的生活中，经常会听到这样的对话。但是如果只是说：“没错，但并不仅仅是脸谱。”还是没有回答“做什么”的问题。事实上，社会网络分析是一种特别通用的方法学，它的诞生至少比推特（Twitter）和脸谱早三十年。

简单来说，社会网络分析（Social Network Analysis, SNA）就是“通过图论研究人类关系的一门学问”。当然，这句解释还远远不够。

在某种程度上，社会网络分析与许多统计方法相似。经济学家广泛使用回归分析的事实并不意味着这项技术只局限于经济学研究。研究社交媒体是应用社会网络分析技术的一个很好方式，同样的道理，社会网络分析技术的应用也不局限于此——数据容易获得，研究机会众多且有价值。事实上，我敢打赌，在十位我尊敬的读者当中，就有九位是基于这个原因而翻开这本书。

十年前，社会网络分析领域还是科学的一潭死水。我们是同时被主流社会学和主流计算机科学拒之门外的异数，我们用奇怪的数学技术处理奇怪的数据类型，生成漂亮的但十分难懂的图片，以及只有在我们的小圈子里才说得通的结论。那时收集和获取社会网络数据很困难（参见附录A）；这个领域的学生都是通过数据集的一个小的标准集来工作，很少冒险离开这个标准集去采集自己的数据。社交媒体的出现改变了所有这一切。每一天，推特产生的社会网络数据比十年前我们整个领域的数据还要多；每个社交媒体

网站提供一个API以方便数据检索；世界上很多政府机构也开放数据由社会网络分析技术来处理。

本书将分析社交媒体数据，我们将收集来自推特（附录A）、脸谱以及LiveJournal的数据，学习识别在线社群，研究如何解析一段病毒视频和一次快闪（flashmob）。

除此以外，本书也会演示社会网络分析很多不同的应用方法。本书的研究对象是社交媒体，但也并不仅限于社交媒体。我们可以通过公司的投资网络及其董事会的共同人员研究公司之间的关系。也可以深入机构内部，去发现饮水机和小吃店附近的社交网络如何影响一家公司的执行力，以及对此不以为然的公司将如何自食其果。我们通过研究竞选资金，可以发现某个特别利益集团是如何控制整个竞选结果的。还可以研究恐怖主义分子、革命者以及激进分子的活动——从1998年的霍巴塔爆炸到“9.11”袭击再到最近的埃及革命。我们将把这些方向和趋势一一解剖开来，而这些往往是通过推特和脸谱实现的，虽然这些事件本质上也是线下的。

本书将告诉你网络数据无处不在——你只需要学习识别它、分析它。一旦你开始这样做了，就会出现新的灵感和想法。

分析关系，理解人与群体

社会网络分析（SNA）科学的核心概念是人与人之间的关系，这些关系定义了你是谁，以及如何行事。我们的人格、教育、背景、种族、民族等所有的一切都与我们的关系模式发生互动，并在关系中留下不可磨灭的印记。因此，通过对这些模式的观察和研究，我们就能够回答许多关于社会的问题。

什么是关系？在人际关系中，它可以是友谊、影响、情感、信任，或者反过来，它可以是不喜欢、冲突或许多其他东西。

二元与赋值关系

关系可以是二元的，也可以被赋值：“张三在推特上关注了李四”是一种二元关系，而“李四转发了张三的四条推文”则是赋值关系。在推特的世界里，很容易对这些关系进行量化，但是在“更无形”的社会生活中，如果想对一段人际关系的状况进行界定并量化，则非常困难。

沟通频率在表示人际关系的程度时十分有用。除了用作客观测量，科学家还发现它能够对情感内容以及人际影响做出准确反应。当然，有时可能并非如此（亲爱的读者，也许您现在的人际关系就是一个反例），但是在很多情况下，由于没有更好的数据，使用沟通频率来衡量人际关系往往是有效的。

对称与非对称关系

有些关系本质上是**不对称的**，这很容易判断。例如老师与学生或老板与员工之间的关系，这样的角色设定了方向性，关系不是对称的。而推特和LiveJournal上的关注尽管从定义上来说有方向，但是还存在一种反关注的联系，因此形成了对称关系。

有些关系是对称的。脸谱上的朋友以及职场社交网站LinkedIn上的联系人都需要相互确认——即使在真实的人际关系中不对称，网站也是强制要求对称的。

在现实生活中，友谊与亲密关系是不对称的，虽然我们不希望这样。因此，我们才会在单相思、一厢情愿的友情以及受欢迎的妄想中挣扎。如果有很好的数据，我们可以使用社会网络分析研究这些现象——但是要想获得这些数据非常困难，而且也往往会受个体报告以及其他误差影响。

多模关系

最后，我们将提到，关系还会存在于不同类型的主体之间——公司雇佣员工、投资者购买公司的股票、人们占有信息与资源等。这些关系被称为双峰关系（bimodal）或二模关系（2-mode）——这些将在第5章进行讨论。

从关系到网络——超乎所见

如果一位传统的量化社会学家或计量经济学家拿到社会网络数据，他会按照以下方法处理：

- 我们可以从中了解什么？有人口统计学数据吗？年龄、种族、宗教、收入、教育程度、地点等任何有可能被测量的定性或定量变量。
- 可以从网络数据中得到什么量化指标？很有可能，这些指标会包括各种形式的中心性（参见第3章）。
- 哪些定性或定量的结果是可以测量的？也就是说，一个公民融入社会的能力、尝试非法药物的可能性等。

然后他将建立一个多变量回归模型，控制一些变量，并将其他一些变量与结果联系起来。这是一个非常有效的方法——事实上，这个方法仍然在社会网络分析会议上占有一席之地。

这种方法传统应用之一是研究同质性（homophily，希腊文，意思是对相似者的喜爱），或者，宽泛地说，是研究“物以类聚、人以群分”。例如，有人提出年龄相近的人比年龄段不同的人更容易成为朋友，还有人说相同种族的人往往会聚在一起。尽管有些猜想