第一图书网, tushu007.com

<<链接>>

图书基本信息

前言

链接，大数据之钥　　程学旗　　中科院计算所所长助理　　中科院网络数据科学与技术重点实验室主任　　著于世纪之初的《链接》一书，展现了世界万物之间泛在的关联关系，揭示了社会、信息、生物、技术、经济等各种复杂系统的网络化存在。
书中，巴拉巴西教授着重以网络思维来刻画和解析了复杂系统的链接特性和相关的普适性规律。
如今，十余年过去了，人类可获取和待处理的数据规模、数据增长速度和数据复杂性经历了大幅度提升，社会进入了“大数据”时代。
一方面，大数据为我们认识自然界和人类社会自身提供了前所未有的机遇；另一方面，技术上所面临的挑战需要我们从本质上发现和总结大数据计算与大数据科学的基本规则与理论体系。
在这样的背景下，一起重温《链接》一书，领略科学家们在网络科学伊始对链接泛在性、数据复杂性、规律普适性的认识和思考，对我们在大数据时代抓住机遇、迎接挑战将大有裨益。
　　复杂性引领《链接》。
从随机宇宙到无尺度小世界，记录着人类对复杂性认识的进步。
复杂性蕴含于万物之间的关联关系中，体现为无序和有序的共存，反映在无处不在的无尺度和幂律上。
链接既是刻画复杂性的方式，又是认识复杂性的途径。
大数据的复杂性，同样体现为数据之间的复杂关联，正是这种复杂关联使得数据复杂性随着数据规模非线性增长。
因此，认识数据复杂性最终需要回到理解和掌握数据之间的复杂关联关系上。
正如《链接》书中所言：在沿着还原论这条路飞奔时，我们撞上了复杂性这堵墙。
网络让世界变得不同，链接是人类认识复杂性的脚手架。
　　普适性贯穿《链接》。
规律的普适性是科学发展孜孜追求的目标，从亚里士多德到伽利略，从哥白尼、开普勒到牛顿，从普朗克到爱因斯坦，正是普适性引领科学一点点拨开复杂性的面纱，一步步揭示复杂性背后简单而深刻的规律。
各个领域形形色色的网络，呈现出共有的无尺度、小世界和高聚团性；富者愈富和适者生存的简单法则，支配着各个领域网络的生长和演化；从微观尺度的基因网络到宏观尺度的经济网络，无不蕴含着枢纽节点和层级结构。
探求普适性规律，小处着手、大处普适，对于我们应对和处理规模不断增长的大数据世界具有重要的借鉴和指导意义。
　　计算蕴含于《链接》。
自然界和人类社会不仅仅是一个静态的客观存在，更是一个具有自我计算能力的动态系统。
蛋白质能够快速地完成转录、鸟群能够高速有序地飞行和有效地觅食、通过互联网的通信而自动完成的寄生计算，这些无不在启示我们自然界固有的计算模式。
从计算的视角来看，计算是对复杂性所蕴含规律的一种实现，高效的计算需要找到有效而简单的规律。
每一种计算都体现着一种规律，简单而深刻的规律对应着高效而简洁的计算。
寻找数据复杂性背后的固有规律，是我们有效处理和利用大数据的钥匙，大数据时代的计算亟需探索大数据复杂性的普适规律。
　　温故而知新，《链接》一书将带我们回顾人类对复杂性的探索历程，领会网络思维的启迪，探寻数据复杂性的普适规律，为大数据时代的计算找回那把迷失的钥匙。

第一图书网, tushu007.com

内容概要

第一图书网, tushu007.com

作者简介

书籍目录

推荐序一链接，大数据之钥　程学旗　中科院计算所所长助理　中科院网络数据科学与技术重点实验室主任　推荐序二链接：泽万物以生机　周涛　电子科技大学教授　互联网科学中心主任　第1链 网络让世界变得不同　让雅虎网站瘫痪的少年黑客　谁在支配网络的结构与演化　当还原论撞上复杂性　探寻下一个大变革　第一部分 复杂网络的起源　第2链 随机宇宙　欧拉的图论与哥尼斯堡七桥问题　网络构造，理解复杂世界的关键　只需30分钟，一个无形社会网络的形成　世界是随机的吗　寻找复杂网络背后的秩序　第3链 六度分隔　《链》与六度分隔的最早表述　任意两个网页之间平均相隔19次点击　对数让大网络缩小了六度，社会间隔的上限　“小世界”，网络的普遍性质　第4链小世界　格兰诺维特与“弱关系的优势”　趋同与聚团　高度聚团的代价，消失的小世界　抛弃随机世界观　第二部分 复杂网络的本质　第5链枢纽节点和连接者　枢纽节点，颠覆“平等网络空间”　贝肯数与埃尔德什数　平均没有意义，多少不是关键　重新思考网络　第6链 幂律　帕累托与80/20定律　幂律，复杂网络背后的规律　有序如何从无序中涌现　从“随机”灌木丛到“自组织”舞台　幂律无处不在　第7链 富者愈富　幂律为什么会出现　生长机制，先发先至　偏好连接，让强者愈强　生长机制和偏好连接，支配真实网络的两大定律　不断完善的无尺度网络理论　第8链 爱因斯坦的馈赠　为什么雅虎选择了谷歌　新星效应打破先发先至　适应度主导一切　“适者愈富”与“胜者通吃”　节点永远在为链接而竞争　第9链 阿喀琉斯之踵　美国西部大停电与互联导致的脆弱性　有效的攻击：攻击枢纽节点　丢失枢纽节点，网络变成碎片　健壮性与脆弱性并存　将对网络的认识转化为实践　第三部分 复杂网络的影响　第10链 病毒和时尚　互联网，让一夜成名的梦想变为现实　意见领袖的力量　无尺度拓扑，病毒得以传播和存活的基础　优先治疗枢纽节点，优先对付“毒王”　社会网络的变化影响传播与扩散规律　第11链 觉醒中的互联网　保罗·巴兰与最优的抗击打系统　将互不兼容的机器连起来——互联网的诞生　人类创造的互联网有了自己的生命　互联网中的“权力制衡”　寄生计算，让所有的计算机都为你工作　第12链 分裂的万维网　万维网的结构影响一切最多不一定最好万维网上的四块“大陆”代码与架构不断扩大的互联网黑洞　第13链 生命的地图　寻找“躁郁症”基因的竞赛　破译人类基因组，打造生命之书　细胞网络的无尺度拓扑，少数分子参与多数反应　个性化药物瞄准问题细胞　网络思维引发生物学大变革　第14链 网络新经济　AOL吞并时代华纳　公司网络，从树形结构到网状结构　复杂董事网络中的完美“内部人士”　市场，带权有向网络　商业模式的转变，互联网带来的真正财富　第15链 一张没有蜘蛛的网　网络研究的冒险之旅网络理论，描述互联互通世界的新语言开启复杂性科学的世纪　后记 复杂网络的未来 299　一场范式变革应对多任务模块化，补上无尺度网络缺失的一链何时才能驯服复杂性　注释　译者后记

第一图书网, tushu007.com

章节摘录

连接者，拥有大量链接的节点　　《纽约客》特约撰稿人马尔科姆 · 格拉德威尔（Malcolm Gladwell）在他的著作《引爆点》（The Tipping Point）中，介绍了一种测试人们社交程度的简单方法。
　　他给出了一个包含248个姓氏的列表——这些姓氏选自曼哈顿电话黄页，然后请被试对照该姓氏列表给自己打分，每认识一个姓氏出现在该列表中的人得一分，可以重复计分。
假如列表中有一个姓氏是琼斯，而被试认识三个姓琼斯的人，就可以得三分。
格拉德威尔对曼哈顿城市学院的大学生们进行了测试。
这些学生大多20岁出头，都是近几年才搬到曼哈顿的，他们的平均得分为21分。
换句话说，他们一般认识21个姓氏出现在格拉德威尔姓氏列表中的人。
格拉德威尔在以白种人为主且受过高等教育的人群中随机选出一组人，又进行了一次测试，他们的平均得分为39分，几乎是大学生得分的两倍。
这个结果并不令人意外，反倒是被试得分的分布范围之广引起了格拉德威尔的注意。
在大学生测试中，得分分布在2到95之间，而在另一个测试中，最低分是9分，而最高分为118分。
即使是有着相似年龄、相似教育程度、相似收入的高度同质化人群，得分的变化也很大：最低分为16分，最高分则达到108分。
格拉德威尔共计测试了大约400人，在每个测试组中都观测到了少数高得分者。
最终，他得出了这样一个结论：“在各行各业中都有一小群特别善于交际的人，他们是社会中的连接者。
”　　连接者是社会网络中极其重要的元素，他们引领潮流趋势、促成重要交易、传播流行时尚并帮助推广餐馆。
他们是社会的连线，于轻描淡写间把不同种族、不同教育水平、不同社会背景的人联系在一起。
当格拉德威尔发现连接者时，他认为自己找到了一种人类独有的特质。
但事实上，他在无意间发现了一个重要现象。
在《引爆点》一书出版之前，这一现象就一直困扰着我的研究小组。
　　连接者是拥有大量链接的节点，他们广泛存在于包括经济系统、细胞等在内的多种复杂系统中，是大多数网络的根本特质之一。
连接者激发了许多领域的科学家的兴趣，包括生物学、计算机科学和生态学。
连接者现象的发现颠覆了我们对网络的所有认识，让我们重新回到了起点。
聚团现象暴露了埃尔德什—莱利随机世界观的第一个裂隙。
前一章探讨的瓦茨—斯托加茨模型挽救了随机世界观，使朋友圈现象和六度分隔现象在网络中和谐共处。
现在，连接者现象发起了对这两个模型的最后一击。
要解释这些高度连接的节点，我们必须彻底抛弃随机世界观。
枢纽节点，颠覆“平等网络空间”　　在网络空间，人们享有极大的言论自由。
有些人受其困扰，有些人热衷于它，而网页内容确实难以审查。
网页内容一旦发布，便有数以亿计的人可以看到。
这种无与伦比的言论发表方式，配合低廉的发布成本，使万维网成为了民主论坛的终极形式。
在这里，每个人的声音被听到的机会都是均等的，至少政客律师和商业杂志是这样认为的。
如果万维网是随机网络，那这些观点都是对的，可惜它不是。
我们的万维网项目得出了一个有趣的发现：万维网并非民主、公平和平等的，万维网的拓扑使我们仅能看到十亿计文档中很少的一部分。
　　在万维网上，你的观点是否可以发布不再是关键问题。
实际上，所有观点都可以发布，而且一旦发布，就立刻能被全世界任何一个有条件上网的人看到。
现在的问题是，当你把信息发布在万维网上后，它能否在数十亿文档中引起别人的关注。
　　要想被读到，先得被看到。
无论是对小说作者还是科学家而言，这都是至理名言。

在万维网上，网页的可见度可以用导入链接数刻画。
拥有的导入链接越多，你的网页越有可能被看到。
如果万维网中的每一个网页都有链接指向你的网页，那么，所有人都能在很短的时间内知道你在万维网上说了什么。
但是，一个网页平均只拥有大约5到7个链接，每个链接指向数十亿网页中的一个。
因此，一个网页将链接指向你的网页的可能性接近于零。
　　上述结论完全适用于我的主页：www.nd.edu/~alb。
根据远景公司（AltaVista）的调查结果，全世界大约有40个网页有链接指向我的主页。
坦率地讲，鉴于我主页涉及领域极其狭窄，40个链接已经很多了。
但是，全世界有大约十亿个网页可供选择，因此，你能发现我的网页的可能性仅有十亿分之四十。
这也就是说，如果你在万维网上随机浏览，浏览每个页面只花费短短10秒钟，你也需要不分昼夜地浏览八年，才能碰到一个指向我主页的链接。
　　每个人都有着不同的兴趣、价值观、信仰和品味，这种多样性反映在我们网页的链接上。
这些链接可以指向非洲部落艺术，也可以指向电子商务门户。
因为有十亿多个节点可供选择，所以网络链接模式看上去相当随机。
链接的随机性意味着，埃尔德什—莱利模型主宰着网络的链接模式。
由于埃尔德什—莱利理论保证了所有节点彼此相似，每一个节点拥有的导入链接数目大致相同，所以随机万维网将是平等主义的最终载体。
　　然而，我们的发现推翻了随机万维网这一预言。
我们的网页采集机器人发回的网络地图证实了万维网拓扑中的高度不均匀。
我们研究了圣母大学的325 000个网页，仅拥有3个导入链接的网页有270 000个，占总网页数的82%。
但是，有大约42个网页被超过1 000个其他网页指向，它们拥有的导入链接数超过了1 000！
随后，我们对由2.03亿个网页组成的样本进行观察，发现了更严重的不均匀：多达90%的网页只有不到10个导入链接，同时，有3个网页被近100万个其他网页引用！
　　万维网上的上述现象和人类社会类似。
人类社会中，少数连接者的朋友之多超乎寻常；万维网中，少数连接数非常高的节点主导着万维网的结构，它们被称为枢纽节点。
雅虎或亚马逊这样的枢纽节点的可见度非常之高，无论在哪里，我们都能看到指向这些节点的链接。
但在万维网中，还有很多不受欢迎或者很少被注意到的节点，它们通过少数枢纽节点连在一起。
　　枢纽节点的存在，颠覆了“平等网络空间”的乌托邦幻想。
我们确实有权利把任何东西放到万维网上，但是会有人注意到吗？
如果万维网是一个随机网络，我们被看到和听到的机会应该是相等的。
从群体的角度来看，我们以某种方式创建枢纽节点，每个人都与其建立链接。
不管你处于万维网中的什么位置，都能很容易地找到这些枢纽节点。
和枢纽节点相比，万维网的其他部分是不可见的。
出于实用目的，只被一个或两个网络链接的网页是不存在的，它们几乎不可能被找到。
在搜索整个万维网寻找热门站点时，搜索引擎对这些链接数非常少的节点同样视而不见。

第一图书网, tushu007.com

<<链接>>

编辑推荐

★复杂网络研究权威，无尺度网络创立者，H-指数高达96的论文狂人，诺贝尔奖大热人选，超越《黑天鹅》的惊世之作《爆发》的作者艾伯特-拉斯洛·巴拉巴西经典力作★中科院计算所所长助理、中国科学院网络数据科学与技术重点实验室主任程学旗，电子科技大学教授、互联网科学中心主任周涛专文推荐。
★巴拉巴西博士后，中科院计算所副研究员沈华伟打造唯一权威版本。
★湛庐文化出品

名人推荐

重温《链接》一书，领略科学家们在网络科学伊始对链接泛在性、数据复杂性、规律普适性的认识和思考，对我们在大数据时代抓住机遇、迎接挑战将大有裨益。
——程学旗 中科院计算所所长助理，中国科学院网络数据科学与技术重点实验室主任虽然《链接》写于十年前，但这本书的精神到现在丝毫没有褪色。
它带给了我们一种整体的、关联的、系统论的审视世界的方式，使我们不仅仅将视野局限于孤立的单元。
我相信，广泛存在的链接是从简单到复杂、从单一到多样、从平凡到璀璨的桥梁。
——周涛 电子科技大学教授，互联网科学中心主任《链接》自由地穿梭于多个科学领域之间，展现了复杂网络这一新领域快速发展的新景象。
巴拉巴西不仅创立了一门新理论，还能够用非技术语言清晰地描述深层的数学观点，这一点难能可贵。
阅读此书，犹如品味一场智慧的饕餮盛宴。
——马克·格兰诺维特 斯坦福大学社会学教授，弱连接理论奠基之作《找工作》作者在《链接》中，巴拉巴西向大家展示了复杂网络如何在空间展开，在《爆发》中，巴拉巴西又剖析了复杂网络如何在时间中展开。
——克莱·舍基 互联网革命最伟大的思考者，畅销书《人人时代》《认知盈余》作者　　巴拉巴西是这个世界上少数几个能够洞悉经验现实背后深意的人之一。
——纳西姆·塔勒布 纽约大学理工学院教授

第一图书网, tushu007.com

<<链接>>

版权说明

本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问:http://www.tushu007.com