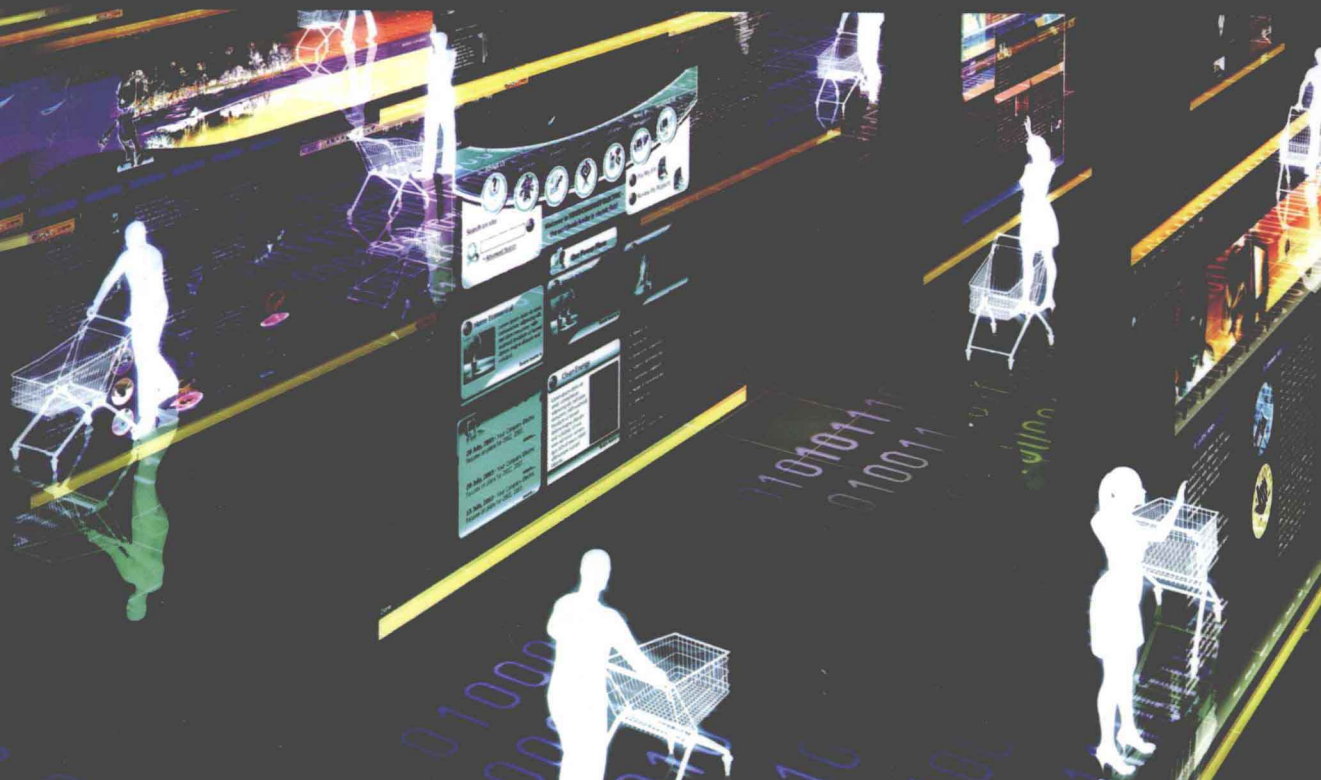


TURING

图灵程序设计丛书

CAMBRIDGE



Recommender Systems
An Introduction

推荐系统

[奥地利] Dietmar Jannach Markus Zanker 著
Alexander Felfernig Gerhard Friedrich 蒋凡 译

**推荐系统
必读经典**

百度技术委员会主席 廖若雪 新浪微博数据挖掘技术专家 张俊林
人民搜索商务搜索部总监 常兴龙 百分点信息科技有限公司首席运营官 张韶峰

**联袂
推荐**

人民邮电出版社
POSTS & TELECOM PRESS

TURING

图灵程序设计丛书



Recommender Systems
An Introduction

推荐系统

著 Markus Zanker
著 Gerhard Friedrich
译 蒋凡



人民邮电出版社
北京

图书在版编目 (C I P) 数据

推荐系统 / (奥地利) 詹尼士 (Jannach, D.) 等著 ;
蒋凡译. — 北京 : 人民邮电出版社, 2013. 7

(图灵程序设计丛书)

书名原文: Recommender systems: An introduction
ISBN 978-7-115-31069-9

I. ①推… II. ①詹… ②蒋… III. ①计算机网络
IV. ①TP393

中国版本图书馆CIP数据核字 (2013) 第049966号

内 容 提 要

本书全面阐述了开发最先进推荐系统的方法, 其中呈现了许多经典算法, 并讨论了如何衡量推荐系统的有效性。书中内容分为基本概念和最新进展两部分: 前者涉及协同推荐、基于内容的推荐、基于知识的推荐、混合推荐方法, 推荐系统的解释、评估推荐系统和实例分析; 后者包括针对推荐系统的攻击、在线消费决策、推荐系统和下一代互联网以及普适环境中的推荐。此外, 本书还包含大量的图、表和示例, 有助于读者理解和把握相关知识。

本书适用于从事搜索引擎、推荐算法、数据挖掘等研发工作的专业人员以及对推荐系统感兴趣的读者。

-
- ◆ 著 [奥地利] Dietmar Jannach Markus Zanker
Alexander Felfernig Gerhard Friedrich
译 蒋 凡
责任编辑 李松峰
执行编辑 李 洁
责任印制 焦志炜
- ◆ 人民邮电出版社出版发行 北京市崇文区夕照寺街14号
邮编 100061 电子邮件 315@ptpress.com.cn
网址 <http://www.ptpress.com.cn>
北京天宇星印刷厂印刷
- ◆ 开本: 800×1000 1/16
印张: 15.25
字数: 361千字 2013年7月第1版
印数: 1-4 000册 2013年7月北京第1次印刷
- 著作权合同登记号 图字: 01-2012-8417号
-

定价: 59.00元

读者服务热线: (010)51095186转604 印装质量热线: (010)67129223

反盗版热线: (010)67171154

广告经营许可证: 京崇工商广字第 0021 号

版权声明

Recommender Systems: An Introduction, First Edition (ISBN 978-0-521-49336-9) by Dietmar Jannach, Markus Zanker, Alexander Felfernig, Gerhard Friedrich, first published by Cambridge University Press 2011 All rights reserved.

This simplified Chinese edition for the People's Republic of China is published by arrangement with the Press Syndicate of the University of Cambridge, Cambridge, United Kingdom.

© Dietmar Jannach, Markus Zanker, Alexander Felfernig, Gerhard Friedrich 2011

This book is in copyright. No reproduction of any part may take place without the written permission of Cambridge University Press and Posts & Telecom Press.

This edition is for sale in the People's Republic of China (excluding Hong Kong SAR, Macao SAR and Taiwan Province) only.

此版本仅限在中华人民共和国境内（不包括香港、澳门特别行政区及台湾地区）销售。

推 荐 语

推荐引擎是近几年比较火爆的领域，很多人对它寄予了很高的期望。作为这个领域的实际从业者，我又深知这个领域在实际应用中的难度。本书不但介绍了比较成熟的经典算法，还介绍了最近几年的一些新进展，并辅之以实际应用的案例介绍。希望看到越来越多的朋友加入到推荐引擎的研究和应用中来！

北京百分点信息科技有限公司首席运营官兼技术副总裁 张韶峰

市场化社会的最大价值在于能互补有无，如果存在一种渠道把卖家商品信息准确无误地展示在真正需要它的买家面前，这个渠道照理应该是人们最喜欢的渠道，因此，在网民覆盖率日益攀高的今天，对应的互联网广告应该是最受人们推崇的产品形式之一。

然而互联网广告总是在用户体验的对立面被提及。当前的互联网广告大多就是做得不好的推荐，大多个性化相关性很差，内容真实性无法有效地被鉴别。糟糕的推荐不仅无法达成买卖双方交易的目标，反而干扰了正常的引导型消费，成为用户避之唯恐不及的洪水猛兽。与之相对，良好的推荐应该是雪中送炭，让用户眼前一亮，有觉得“找的就是它”的认同感。

道理虽然简单，做起来却并不容易，最关键的是如何利用互联网了解每一个人，并根据他们的兴趣在不同的地点、不同的时间及时送上恰当的信息，并在这些信息传递的同时，有效地针对这些信息单元建立不同的衡量标准及评估方法，去伪存真，促成信息的有效传递。这些内容是推荐引擎研究的范畴。

众所周知，互联网信息经历了早期黄页导航时代，相关检索的搜索引擎时代，目前处于阶段性的成熟突破期，推荐引擎在这种情况下越来越多地作为一个独立的知识主题被提及，无疑预示着下一步互联网的发展趋势。由蒋凡执笔翻译的这本《推荐系统》是一本从基础介绍推荐引擎的难得的好书，给人启迪良多。愿越来越多的互联网爱好者认真阅读本书，走在互联网发展大潮的前沿，成为下一代互联网产品真正需要的人才。

人民搜索商务搜索部总监 常兴龙

最近几年推荐系统很火爆，几乎成为大型网站的标准配置和功能，这反映了在信息过载环境下对信息消费需求的可能解决方案，即如何根据用户的品味爱好自适应地发现满足用户消费习惯的信息。

本书比较全面地介绍了推荐系统涉及的相关知识点，虽然细节不是非常深入，但是在内容广度和新颖性方面做得较好，尤其突出的一点是能够辟出多个章节讲述推荐系统研究较新的进展，比如如何防范针对推荐系统的作弊和攻击，社交环境及上下文感知环境下的推荐系统设计方案等。译文也清晰流畅，有很好的可读性。

本书很适合对于推荐系统感兴趣的相关人员作为入门教程，目前能够系统全面介绍相关技术的中文书籍还显得匮乏，相信这本译著对于缓解这种情况大有裨益。

新浪微博数据挖掘技术专家 张俊林

推 荐 序

我写这篇推荐序时正值周末，对于一个美好的周末而言，没有什么比愉悦的音乐和阅读更重要。打开百度音乐随心听的私人频道，房间里回荡起陌生而又对味的旋律；打开百度首页，猜你喜欢频道已经为我准备好新鲜又正中下怀的新闻资讯；浏览完新闻，打开亚马逊网站，跃入眼帘的是为我推荐的 Kindle 电子书，不少都是我想看的……

还记得 20 年前的我们，小心翼翼地为一本书包上封皮，支起收音机搜索每一个波段，在只有三个台的电视前欢天喜地地度过一整个假期。20 年后的我们，互联网上成千上万部电影即点即看，却不知道要看哪一部；几百万首歌可以实时播放，却找不到想听的。信息技术带来资讯极度丰富的同时，也带来了用户的选择困难症。而最好的选择，就是不需要选择。这就是推荐系统，一个经过十多年技术发展，而今逐渐浮出水面的前沿技术，所要解决的问题。

自上世纪 90 年代诞生以来，推荐系统一直在互联网的众多领域发挥着核心作用，例如亚马逊的电子商务系统、Hulu 的视频推荐系统、Pandora 的音乐推荐系统等。近年来移动互联网的高速发展，用户的移动个人终端设备越来越丰富。移动设备往往体积较小，移动端页面承载的信息量是有限的，这也就意味着展现的信息需要更精准地符合每个用户的个体需求。而媒体的碎片化、文化的多样性，也带来人们对信息的需求越来越个性化。这些因素推动着个性化及推荐系统应用的普及化和深入化。我也相信在不远的未来，推荐系统会像搜索引擎一样，成为互联网产品必不可少的基础性应用。

正因为推荐系统的重要性，百度在 2011 年成立了独立的部门，专注于个性化和推荐系统的技术研发。我们正在把诸多该领域的最新技术，融入百度丰富的产品线中，加快各类产品的快速进化。经过两年的努力，百度推荐系统已崭露头角。除了上文提到的百度音乐随心听、百度首页个性化推荐，百度视频、地图、新闻、知道等产品线都植入了推荐系统，并体现出很高的价值。

在负责百度推荐与个性化部门的过程中，我也发现，和推荐系统应用的快速发展相比，推荐技术人才近年来却处于非常匮乏的状态。一方面，机器学习、数据挖掘等推荐系统的相关基础技术，也正处于人才的培养期；另一方面，推荐技术相关书籍的相对空白，使得这一领域无法进入

更多人的视野，同时对这一领域感兴趣的人也没有系统性的材料可供参考学习。

在此书翻译为中文之前，我有幸阅读过此书的原版。该书以浅显易懂的方式，系统介绍了推荐系统的相关技术，并提供了一个非常好的理论结合实践的学习方式。通过对本书的学习，读者不仅可以全面系统地了解该领域的基础原理，还能试验如何搭建一套真正的推荐系统。

译者蒋凡是我的同事，他在搜索技术领域有着多年的相关工作经验，同时推荐技术也是他的兴趣所在，投入了大量时间进行钻研和实践。相信他带来的该书中文译本，可以在很大程度上弥补目前中文推荐系统学习材料的不足，为更多希望进入推荐领域的读者提供帮助。

百度主任架构师、百度技术委员会主席 廖若雪

译者序

2011年8月，我去 San Diego 参加 KDD 大会，剑桥大学出版社的高级编辑 Lauren Cowles 女士向我介绍了这本书。她正好也是本书的编辑，像《信息检索导论》、《网络、群体与市场》和《程序设计语言概念》的英文版也都出自她的手笔。书打完折是 50 美元，是我买过最贵的一本了，但是现在看起来很值。

首先，这是一本全面介绍推荐系统概念、源流、算法的专业著作，体系完整，贴近实用，结合很多实例告诉你如何尽快地将推荐系统领域里的成熟算法运用起来构建应用。著名的推荐系统 MovieLens 的创始人 Joseph A. Konstan 教授还特地为本书撰写了序。因此这本书非常适合刚刚进入推荐系统领域，希望了解一些成熟算法，解决实际问题的工程师学习使用。

其次，本书更有价值的第二部分介绍了推荐系统在防攻击、消费决策、社交网络和移动四个前沿领域的研究进展，收录了很多已经在实际商业产品中经过反复验证的套路和经验。本书取之有据，言之成理，非常适合希望能在国内各类推荐应用中少走弯路，尽快发挥推荐系统作用的产品负责人参考。

此外，本书在每章之后都给出了参考文献，但与我们常见的按照论文格式给出引用的方式不同，而是单辟一节注释参考书目，介绍你该去看哪些相关内容，这些资料的来龙去脉，当时是为了解决什么问题，解决得怎样等。非常适合不满足于已有结论，愿意梳爬文献资料，探究背后学理的科研人士将此作为推动推荐系统领域进步的新起点。

推荐系统已经发展了十几年，但仍然是一个需要持续创新的领域。我和我的同事、同行们也正在努力将这些研究成果与搜索引擎技术结合起来。本书总结出推荐系统的一些经验和规律已经开始融入到了我们的产品中，正在改变着广大互联网用户的体验。因此可以说，本书提供了一个窗口，非常适合希望洞悉互联网产品的未来发展脉络，把握技术和商业相结合的下一波浪潮的有志同学。

正是对本书这些价值的感悟，使得我能坚持 5 个月，利用工作之余的深夜和周末坚持翻译，经常是加班回家之后又打开电脑翻译几页。在此过程中，要感谢傅志红老师和李洁编辑在译稿、

编辑阶段的帮助，感谢李松峰老师在回稿过程中细致严谨的工作，感谢谢工老师邀请我参加图灵公司的活动，感谢李盼向我介绍图灵的翻译计划，同样感谢陈黎教授、胡蓉博士和张谧老师不辞辛苦地审稿。最后要感谢的是我的妻子和父母，他们帮我承担了两个女儿的抚育责任，我挪用了本该属于他们的时间。

非常抱歉的是，由于能挤压出的时间有限以及本人在专业和翻译水平上的局限，大家最终看到的译稿肯定还会有很多疏漏之处。希望读者能持续关爱这本著作，通过新浪微博@蒋凡 Baidu 和邮箱 jf@ustc.edu，不吝提出宝贵的意见和建议，一同改进。

蒋凡

2012年12月

序

利用网络上数百万人的意见帮助我们发现更有用和有趣的内容，这个简单的想法在 20 世纪 90 年代早期极具诱惑力。事实上，这个简单想法的确在不同的领域以不同的形式被证明是有效的。帕洛阿尔托研究中心的 Tapestry 系统（Goldberg et al. 1992）引入了协同过滤的思想和概念，展示了如何将显式标注数据和隐式行为数据存入可查询的数据库中，以及用户如何利用这些数据进行个性化过滤。不到两年，GroupLens 系统（Resnick et al. 1994）展示了协同过滤方法既能跨网计算又能自动完成。GroupLens 是针对 Usenet 新闻消息进行自动协同过滤，而麻省理工学院（MIT）的 Ringo 系统是针对音乐唱片和艺术家进行过滤，贝尔通信研究中心的视频推荐系统（Hill et al. 1995）针对影片进行过滤。每个系统都是使用相似的自动化技术——通过算法识别出其他有相似爱好的用户，然后综合他们的评分给出个性化的加权平均值。简单的“k 最近邻”算法非常有效，在所有协同过滤算法中脱颖而出，很快就成为了黄金标准。

面向系统的探索。现在看来，推荐系统发展过程中有四个相互交叠的阶段，这些早期的协同过滤系统很显然是其第一阶段中的重要实例。面向系统的探索阶段（不仅是通过协同过滤，还有基于知识的系统，比如 FindMe 系统（Burke et al. 1996））证明了推荐系统的可行性和效果，极大地激发了人们推动该领域在科研及商业实践方面不断向前发展。（我并不是说这些早期研究成果没有研究算法和设计替代方案，而是想说，在很大程度上让人兴奋的是歌声已经响起，我们并不担心它是不是天籁之音。）

这一阶段的关键事件是 1996 年 3 月在伯克利举办的协同过滤专题研讨会。这次集会将工作在个性化与非个性化系统、不同算法（从统计汇总、k 最近邻到贝叶斯聚类）、不同领域的人们聚集在了一起，最后达成了一项共识：这些研究内容都围绕一个主题，也就是后来人们所说的推荐系统展开，而这在某种程度上要感谢那一期专题学术研讨会的《ACM 通信》专刊（Resnick and Varian 1997）。

快速商业化——规模和价值上的挑战。推荐系统出现在快速膨胀的互联网商业氛围中，商业化几乎刻不容缓。MIT 的 Pattie Maes 研究组于 1995 年创立了 Agents 公司（后来更名为萤火虫网络，Firefly Networks）。我们在明尼苏达州的 GroupLens 研究组于 1996 年创立了 Net Perceptions。还有其他很多公司奋起直追。很快，我们就开始面对实验室里未曾遇到的真实挑战。为了成功，公司不得不在如何精确地预测结果上更进一步。我们必须证明能够提供有价值的推荐（一般情况下会选择一些特定产品进行推荐，从而产生额外的购买量），而且要在不降低现有 Web 站点速度的情况下做到这些。这些系统必须能够在大大超越实验室规模的情况下运行（处理上百万的用户

和物品以及每秒成百上千的交易)。所以 John Riedl 和我合著的关于推荐系统的第一本书 *Word of Mouse*，其目标读者不是研究人员而是营销专家，或许就不那么令人惊讶了。

这一时期研究工作的目标是要解决这些技术挑战。人们开发新的算法以降低在线计算时间，包括沿用至今的基于物品的关联算法和降维方法。研究人员更感兴趣的是根据对于 top- n 推荐列表的各种性能的度量来评估各个推荐系统。研究领域非常广泛，包括隐式评分、新用户和新物品的冷启动问题以及可信度、可解释性和透明度等用户体验相关问题。

研究大爆发——推荐成为主流。2000~2005 年之间，随着互联网泡沫的破灭，或者由于无法抗衡将推荐整合到更全面的商业产品线的主流公司，许多推荐系统公司逐渐消亡了。然而，推荐系统作为一门技术仍然存在，并广泛应用在电子商务、大规模零售业和各种知识管理应用中。

与此同时，随着各个学科研究人员的参与及方法的引入，推荐系统研究得到迅猛发展。来自人工智能、信息检索、数据挖掘、安全与隐私以及商业与营销等各个领域的研究，都为推荐系统提供了新的分析和方法。由于可以获取到海量数据，算法研究方面取得了很大进步，在 2006 年更是被悬赏 100 万美元将预测准确度提高 10% 的 Netflix 大奖推上高峰。

前进——基于上下文的推荐。Netflix 大奖激发了许多研究人员共同不懈努力来提高推荐准确度。但是，即便这些研究人员成功地完成了任务，还是有一批研究人员和专业人士主张要回归到探索和价值的准则上。2006 年，MyStrands 组织了 Recommenders06，这是一个介绍推荐系统现状和未来的暑期班。2007 年，我组织了第一届 ACM 推荐系统大会。这个大会最初有 120 人参加，到 2009 年已有 300 多人。回顾这些事件，可以看出人们对基于上下文的推荐越来越感兴趣，乐于改进研究方向使其立足于理解人们如何与机构或企业互动，推荐如何使得这些互动更为便利。虽然 Netflix 成功引入一些新的想法，使推荐领域得到了极大的促动，但实际上我们大多数人都知道，一个仅能提高“用户讨厌某些差电影”预测精准度的复杂算法，并不会对 Netflix 或用户有所帮助。这就解释了为什么 2009 年最佳论文颁发了一篇证明推荐领域评估算法中传统的“隐藏某些数据”方法存在缺陷的论文 (Marlin and Zemel 2009)，而近年来被引用次数最多的推荐领域论文是一篇设计如何让评估适合用户需求的论文 (Herlocker et al. 2004)。

让我们回到这本书。本书涵盖了推荐系统领域的全部知识，并为应对未来新的挑战提供了前瞻性建议。书中全面解释了一系列生成推荐的经典算法和方法，概述了源自社交计算和语义网的新手段对推荐系统的作用。希望这本书能够点燃你的激情，释放你的创造力和进取精神，把推荐系统的研究与应用推向新的高度。

Joseph A. Konstan
杰出 McKnight 教授
明尼苏达大学计算机科学与工程系

前 言

“我该买哪款数码相机？我们全家要在哪儿度过最完美的假期？对孩子教育的最佳投资是什么？我该租哪部电影？我会对哪些网站感兴趣？我下次休假时该买哪本书看？哪个学位和大学对我的前途最有利？”

人们在决定该如何花钱，或者更宽泛地说，如何对未来做出决策时都会提出这样的问题，类似的例子还有很多。

传统上，人们用过各种各样的方法来解决这些决策问题：找朋友聊聊、从可信的第三方获取信息、雇用专家团队、在互联网上咨询、使用决策论的各种方法（如果他想更理性些）、凭直觉或是索性随大流。

然而，几乎每个人都有过这样的经历：推销员大献殷勤的建议并不那么有用；凭感觉跟着富人邻居投资，却没有真正给我们带来收益；无休止地花费时间在互联网上会导致困惑，而不能做出迅速而正确的决定。总而言之，好的建议难得一遇。大多数情况下，需要花费大量时间或金钱，即便如此还总是让人半信半疑。

如果有个能付得起的私人顾问帮助我们高效地做出正确的决策该有多好啊！

构建支持用户在线决策的系统正是推荐系统领域的主要目标。这个目标强调要为大规模用户提供便捷访问的高质量推荐。

强调数据规模和易于访问使得这项技术非常强大。尽管推荐系统的目标是用户的个人决策，但大量的应用使得该系统在更广泛的意义上产生了重要影响，比如 Amazon.com 的推荐引擎。由于互联网市场的深入渗透，这个问题尤其不能被忽视，因为掌握推荐系统就可以在更广泛的意义上控制市场。想想，比如百货公司所有销售员只能根据订单推销某种商品。

有人会争辩说，推荐系统是为那些负担不起或不愿为专家的高质量建议付费的人群服务的。从某种程度上，在一些领域这是正确的，比如金融服务或医疗；然而，做出好决策的目标还包括要超越该领域的专家。这显然不太可能，也不是在所有的领域都有必要，但还是可以从很多实例中发现群体智慧能够用于改进决策。因此，考虑到互联网上有着大量可以获取的信息，我们能否开发出一个系统，提供比人为推荐更好的推荐？

努力推荐用户支付得起、个性化、匹配度高的产品是推荐领域的核心问题，这也对技术和心理学提出了很多挑战。尽管在技术层面上，我们关心的是发现尽可能有效利用可用信息和知识的方法，但在设计最终用户交互过程时，必须考虑到心理层面的因素。这些交流过程的设计会极大影响随后推荐的信任度，最终会影响到决策本身。用户没法像理性经济人那样行事，因为后者完

全清楚自己想要什么。甚至在推荐过程中询问用户偏好的方式，或者提供哪些决策选项都会影响到用户的选择。因此，推荐系统不能被简化为简单的决策理论概念。

现在被称为“推荐系统”的软件最早出现于十五年前。从那时起，研究人员不断地开发实现推荐系统的新方法。今天，我们中的大多数人已经习惯于推荐系统的服务，比如 Amazon.com 使用的推荐引擎。历史上，推荐系统由于应用了人工智能和信息过滤领域的方法而广受关注，采用这些方法可以推荐 Web 站点或对新闻进行排序、过滤。事实上，基于实例或规则技术的推荐方法，正是源于 20 世纪 80 年代的专家系统。然而，推荐系统的应用范围远远超过了纯粹的信息过滤方法，现在的推荐技术正在不同领域提供解决方案，比如金融产品、房地产、电子消费产品、电影、书籍、音乐、新闻和 Web 站点等等。

本书介绍了很多推荐系统技术以及最新的进展。目标读者包括该领域的研究生或刚进入这一领域的博士，开始设计并实现实际推荐系统的专业人士和 IT 专家。更多的高级资料可以在《推荐系统手册》(*Recommender Systems Handbook*, Ricci et al.2010) 中找到，书中全面收录了这一领域先行者的研究成果。

本书由两部分组成。第一部分先是总结了实现推荐系统的基本方法，并讨论了它们各自的优点和缺点。除了描述如何构建这样的系统，我们还重点讨论了评估推荐精准度和通过在线用户行为检验推荐效果的方法。第二部分重点讨论最新进展，涉及了诸如推荐系统信任度和基于 Web 2.0 及语义网技术的新兴应用。配合本书主题内容的教学资料可以从 <http://www.recommenderbook.net/> 获取。

我们感谢所有对本书做出贡献的人，特别是剑桥大学出版社的 Heather Bergman 和 Lauren Cowles，他们在整个编辑过程中为我们提供了支持。也要特别感谢 Arthur Pitman、Kostyantyn Shchekotykhin、Carla Delgado-Battenfeld 和 Fatih Gedikli，是他们帮助校对了原稿。还要感谢几位学术同事帮助审核，并给了我们很多有益的反馈。

Dietmar Jannach

Markus Zanker

Alexander Felfernig

Gerhard Friedrich

2010 年分别于多特蒙德，克拉根福，格拉茨

目 录

第 1 章 引言	1	2.6 讨论和小结	30
1.1 第一部分：基本概念	2	2.7 书目注释	31
1.1.1 协同过滤推荐	2	第 3 章 基于内容的推荐	32
1.1.2 基于内容的推荐	2	3.1 内容表示和相似度	33
1.1.3 基于知识的推荐	3	3.1.1 向量空间模型和 TF-IDF	34
1.1.4 混合推荐方法	4	3.1.2 向量空间模型的改进及局限	35
1.1.5 推荐系统的解释	4	3.2 基于内容相似度检索	36
1.1.6 评估推荐系统	4	3.2.1 最近邻	36
1.1.7 案例研究	5	3.2.2 相关性反馈——Rocchio 方法	37
1.2 第二部分：最新进展	5	3.3 其他文本分类方法	40
第一部分 基本概念		3.3.1 基于概率模型的方法	40
第 2 章 协同过滤推荐	8	3.3.2 其他线性分类器和机器学习	43
2.1 基于用户的最近邻推荐	8	3.3.3 显式决策模型	44
2.1.1 第一个例子	8	3.3.4 特征选择	45
2.1.2 更好的相似度和赋权体系	10	3.4 讨论	47
2.1.3 选择近邻	11	3.4.1 对比评估	47
2.2 基于物品的最近邻推荐	11	3.4.2 局限	47
2.2.1 余弦相似度量	12	3.5 小结	48
2.2.2 基于物品过滤的数据预处理	13	3.6 书目注释	49
2.3 关于评分	14	第 4 章 基于知识的推荐	51
2.3.1 隐式和显式评分	14	4.1 介绍	51
2.3.2 数据稀疏和冷启动问题	15	4.2 知识表示法和推理	52
2.4 更多基于模型和预处理的方法	16	4.2.1 约束	52
2.4.1 矩阵因子分解	17	4.2.2 实例与相似度	54
2.4.2 关联规则挖掘	20	4.3 与基于约束推荐系统交互	55
2.4.3 基于概率分析的推荐方法	22	4.3.1 默认设置	55
2.5 近来实际的方法和系统	25	4.3.2 处理不满意的需求和空结果集	57
2.5.1 Slope One 预测器	26	4.3.3 提出对未满足需求的修改建议	61
2.5.2 Google 新闻个性化推荐引擎	28		

4.3.4 对基于物品/效用推荐结果的 排序.....	61	6.4 协同过滤推荐系统的解释.....	106
4.4 与基于实例的推荐系统交互.....	64	6.5 小结.....	108
4.4.1 评价.....	65	第7章 评估推荐系统	109
4.4.2 混合评价.....	67	7.1 介绍.....	109
4.4.3 动态评价.....	67	7.2 评估研究的一般特性.....	110
4.4.4 高级的物品推荐方法.....	70	7.2.1 总论.....	110
4.4.5 评价多样性.....	71	7.2.2 评估方案的实验对象.....	111
4.5 应用实例.....	72	7.2.3 研究方法.....	113
4.5.1 VITA——基于约束的推荐系 统.....	72	7.2.4 评估环境.....	115
4.5.2 Entree——基于实例的推荐 系统.....	77	7.3 主流推荐方案.....	115
4.6 书目注释.....	79	7.4 历史数据集评估.....	116
第5章 混合推荐方法	80	7.4.1 方法论.....	116
5.1 混合推荐的时机.....	81	7.4.2 衡量标准.....	117
5.1.1 推荐理论框架.....	81	7.4.3 结果的分析.....	121
5.1.2 混合设计.....	82	7.5 其他评估方案.....	121
5.2 整体式混合设计.....	83	7.5.1 实验性研究方案.....	122
5.2.1 特征组合的混合方案.....	84	7.5.2 准实验研究方案.....	122
5.2.2 特征补充的混合方案.....	85	7.5.3 非实验研究方案.....	123
5.3 并行式混合设计.....	87	7.6 小结.....	123
5.3.1 交叉式混合.....	87	7.7 书目注释.....	124
5.3.2 加权式混合.....	88	第8章 案例分析：移动互联网个性化 游戏推荐	125
5.3.3 切换式混合.....	89	8.1 应用与个性化概述.....	126
5.4 流水线混合设计.....	90	8.2 算法和评级.....	128
5.4.1 串联混合.....	90	8.3 评估.....	128
5.4.2 分级混合.....	91	8.3.1 测量1：我的推荐.....	129
5.5 讨论和小结.....	92	8.3.2 测量2：售后推荐.....	131
5.6 书目注释.....	92	8.3.3 测量3：起始页推荐.....	133
第6章 推荐系统的解释	94	8.3.4 测量4：演示版下载的整体 效果.....	135
6.1 介绍.....	94	8.3.5 测量5：整体效果.....	136
6.2 基于约束的推荐系统中的解释.....	96	8.4 小结与结论.....	138
6.2.1 实例.....	97	第二部分 最新进展	
6.2.2 通过推导生成解释.....	99	第9章 针对协同推荐系统的攻击	140
6.2.3 可靠解释的分析与概述.....	100	9.1 第一个例子.....	141
6.2.4 可靠解释.....	102	9.2 攻击维度.....	141
6.3 基于实例推荐系统的解释.....	103	9.3 攻击类型.....	142

9.3.1 随机攻击	142	11.1.1 利用显式的信任网络	169
9.3.2 均值攻击	143	11.1.2 信任度度量方法和效果	171
9.3.3 造势攻击	143	11.1.3 相关方法和近期进展	172
9.3.4 局部攻击	143	11.2 大众分类法及其他	174
9.3.5 针对性的打压攻击	144	11.2.1 基于大众分类法的推荐	174
9.3.6 点击流攻击和隐式反馈	144	11.2.2 推荐标签	181
9.4 效果评估和对策	145	11.2.3 在分享媒体中推荐内容	183
9.4.1 推举攻击	145	11.3 本体过滤	185
9.4.2 打压攻击	146	11.3.1 通过分类改进过滤	185
9.5 对策	146	11.3.2 通过属性改进过滤	188
9.6 隐私方面——分布式协同过滤	148	11.4 从网络抽取语义	189
9.6.1 集中方法：数据扰动	149	11.5 小结	191
9.6.2 分布式协同过滤	150	第 12 章 普适环境中的推荐	192
9.7 讨论	153	12.1 介绍	192
第 10 章 在线消费决策	155	12.2 上下文感知推荐	193
10.1 介绍	155	12.3 应用领域	195
10.2 环境效应	156	12.4 小结	197
10.3 首位/新近效应	159	第 13 章 总结和展望	198
10.4 其他效应	160	13.1 总结	198
10.5 个人和社会心理学	161	13.2 展望	198
10.6 书目注释	167	参考文献	201
第 11 章 推荐系统和下一代互联网	168	索引	223
11.1 基于信任网络的推荐系统	169		