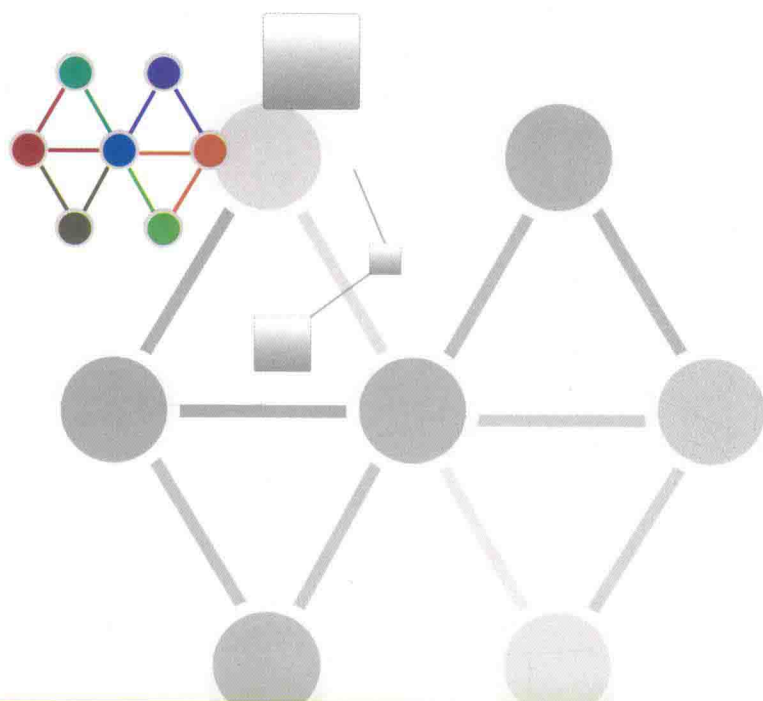


基于本体建模的微博信息管理机理研究

崔金栋 孙遥遥 王欣媛 著



科学出版社

基于本体建模的微博信息 管理机理研究

崔金栋 孙遥遥 王欣媛 著

教育部人文社会科学研究规划基金项目（基于微本体架构的微博信息管理机理研究（14YJA870001））、吉林省社会科学基金重点项目（吉林省微博信息生态分析与管理机制研究（2016A20））研究成果之一

科学出版社

北 京

内 容 简 介

微博凭借实时性和便捷性成为重要的网络应用,其聚集的大量用户和相对自由的言论信息使之成为把握社会脉搏的重要工具。在引入篇,本书从分析微博研究的理论价值和现实意义入手,通过分析关于微博信息管理的研究现状,发现微博现有研究的不足和缺陷,即语义性差、难以管控等。随着语义网技术的进步,本体技术的应用逐渐成熟,语义本体的研究和应用被提上日程。在高级篇,本书把复杂的本体架构简化后,利用自动构建技术在语义网中予以实现,并在此基础上提出语义网中本体匹配的相关算法和模型,为微博在语义网中的应用埋下技术伏笔。在应用篇,本书从微型本体架构入手,提出通过微型本体提升语义性,从而在微博信息组织和传播阶段提升微博信息管理的可控性和有效性。

本书为微博信息管理方面的专业书籍,专业性较强,可供信息管理、情报学和计算机相关领域的研究者阅读和参考。

图书在版编目(CIP)数据

基于本体建模的微博信息管理机理研究/崔金栋,孙遥遥,王欣媛著. —北京:科学出版社,2017.10

ISBN 978-7-03-053522-1

I. ①基… II. ①崔… ②孙… ③王… III. ①博客—信息管理—研究 IV. ①G255.76

中国版本图书馆 CIP 数据核字(2017)第 139213 号

责任编辑:王 哲 董素芹 / 责任校对:桂伟利

责任印制:张 倩 / 封面设计:迷底书装

科学出版社出版

北京东黄城根北街 16 号

邮政编码:100717

<http://www.sciencep.com>

新科印刷有限公司 印刷

科学出版社发行 各地新华书店经销

*

2017 年 10 月第 一 版 开本:720×1 000 1/16

2017 年 10 月第一次印刷 印张:9 1/2

字数:210 000

定价:58.00 元

(如有印装质量问题,我社负责调换)

序

应金栋之约写序，欣喜之余甚慰。小友近年来发表 CSSIC、EI 论文几十篇，承担基金项目十余项，获得国家自然科学和国家社会科学的奖励也很多，但是唯独缺部专著，现已补上，甚慰。

该书主要解决现有微博信息管理中的技术问题。从本体建模的思想出发，利用本体建模思想对微博信息的组织和传播进行规范，有效地提高了微博信息描述问题的规范性，同时利用 Folksonomy 和本体技术构造的微本体来组织微博信息，实现微博信息传播中的可检测和可控性，为微博信息管理在技术层面上提供了具体的实现方法，同时结合用户层面的管理提出基于本体的微博信息用户推荐机制。

该书的主要研究内容和创新点包括以下三个方面。首先，以形式概念分析和本体来规范微博信息组织。在形式概念分析的背景下构建一种基于 Folksonomy 的类似本体的结构——微本体，这种结构化的信息形式既可以结合 Folksonomy 的优点，让所有用户参与到本体构建中，降低微博信息本体的构建成本和复杂度；又可以拥有本体的优点，形成结构化的、无歧义的、易控制的信息组织结构，便于信息的管理和传播。其次，以系统学和本体来规范微博信息传播；以系统的视角，即以系统稳定性、系统临界行为、系统可控性来研究微博信息的传播；从信息建模的角度，通过对微博信息抽取而成的微博信息传播本体的监测从而达到微博信息传播预警的目的。最后，以本体技术来约束微博信息的用户推荐。在构建用户推荐本体的基础上，利用改进的 LDA 模型进行主题推荐，有效地避免了现有推荐机制缺乏语义性的缺点，提高了推荐精度。

该书的理论价值主要在于对微博的信息组织架构、传播模式和影响力形成机理进行研究，利用微本体技术改善微博信息资源的组织方式，寻找利于微博管理的最合适的信息架构，分析微博产生裂变式传播能量和海量信息的动因，提出基于微本体建模的微博信息管理体系，所有这些将有效地丰富现有微博研究的理论体系，丰富信息资源管理相关理论。

张海涛

吉林大学教授

2017 年 7 月

前 言

微博是互联网的新产物，已经具备了信息快速、广泛传播的基础条件，随着用户规模的迅速扩张，以及其内部运行机制的不断完善，它在社会信息传播中所能发挥作用的空間將越来越大。微博用户可以不受时间、空间的限制，持续不断地对某条信息作出评论和转发，使得该信息可能在极短时间内以“核裂变”的方式被成百上千的用户转发，从而演化成引起舆论关注的社会新闻事件；面对内容庞杂、数量巨大的信息，用户很容易眼花缭乱，为垃圾信息所困扰。因此，庞大而琐碎的信息、频繁变化的话题、名人效应等，都在一定程度上影响了微博的长期传播效果。

微博核心价值是建立在独特的信息发布机制、信息获取机制和信息传播机制上的，这些机制满足了大众及时、简单沟通的需求，迎合了“碎片化”的生活方式。微博的出现和迅猛发展大大拓展了网络信息传播渠道，同时对规范网络传播秩序、确保网络信息安全提出了新的要求。总体来看，微博具有的个体性、即时性、分享性、参与性，往往使其兼具“天使”和“魔鬼”的双重面孔，也给微博领域的信息评论引导带来了巨大的机遇和挑战。

微博使用环境属于典型的语义网，提升微博信息管理的语义性能有效地解决现有管控难题，而本体又是语义网信息建模的主要工具，本书力图使用本体技术提升语义性来解决微博信息管理存在的问题。基于上述认识，在本书的编写中，在对现有语义网信息检索模型进行充分研究的基础上，作者结合自身提出的语义网本体自动构建理论和改进的匹配模式，提出基于本体自构的语义网信息检索模型，为后面本体在微博中的应用打下基础。后半部分主要利用前述模型解决现有微博信息管理中的技术问题。其理论价值主要在于对微博的信息组织架构、传播模式和影响力形成机理进行研究，利用微本体技术改善微博信息资源的组织方式，寻找利于微博管理的最合适的信息架构，分析微博产生裂变式传播能量和海量信息的动因，提出基于微本体建模的微博信息管理体系，所有这些将有效地丰富现有微博研究的理论体系，丰富信息资源管理相关理论。应用价值就是从本体建模的思想出发，利用本体建模思想对微博信息的组织和传播进行规范，有效地提高微博信息描述问题的规范性，同时利用自由分类法和本体技术构造的微本体来组织微博信息，实现微博信息传播中的可检测和可控性，为微博信息管理在技术层面上提供具体的实现方法，同时结合用户层面的管理提出基于本体的微博信息用户推荐机制。最后，通过构建微博用户推荐本体，并用改良后的 LDA 模型与传统 VSM 模型加 LDA 模型的比较说

明在用户推荐环节, 微博用户推荐本体的建立更有利于推荐精度的提升。利用这种方式能够更好地控制社会舆论并加以正确引导, 保障社会和谐安定。

本书是在借鉴诸多学者辛勤劳动成果的基础上编写而成的, 这些成果已列于参考文献中, 在此表示深深的感谢。感谢为本书提供参考资料的同学、朋友, 感谢科学出版社对本书的出版工作的支持。

作 者

2017 年 7 月

目 录

序 前言

引入篇：微博及其研究现状

第 1 章	绪论	2
1.1	背景及意义	2
1.1.1	研究背景	2
1.1.2	研究意义	3
1.2	国内外研究综述	4
1.2.1	微博信息组织研究	5
1.2.2	微博信息传播研究	6
1.2.3	微博用户推荐研究	7
1.3	本书的主要内容	10
1.4	本书的创新点	12
1.5	本章小结	13

基础篇：微博与网络信息管理技术

第 2 章	微博与信息管理技术	16
2.1	微博与信息管理概述	16
2.1.1	微博信息管理的现状	16
2.1.2	微博信息管理的手段	17
2.2	本体	18
2.2.1	概念及内涵	18
2.2.2	分类与功能	19
2.2.3	构建方法与技术	20
2.3	Folksonomy	21
2.3.1	概念及内涵	21
2.3.2	特征与功能	23

2.3.3	应用	25
2.4	网络信息传播	26
2.4.1	概念及内涵	26
2.4.2	特征与功能	27
2.4.3	构建方法与技术	29
2.5	本章小结	29

高级篇：本体在网络（语义网）信息检索中的应用

第 3 章	语义网信息检索中本体自动构建的研究	32
3.1	语义网信息检索中引入本体构建技术的目的	32
3.2	语义网本体建模中用到的其他技术	33
3.3	语义网环境下本体的自动构建理论	34
3.3.1	语义网本体自动构建的基本思想	34
3.3.2	语义网本体自动构建的原则	34
3.3.3	语义网本体自动构建方法的选择	35
3.4	语义网领域本体库的自动构建	35
3.4.1	语义网领域本体自动构建的信息需求	36
3.4.2	语义网环境下的 LDAP 目录服务	37
3.4.3	语义网环境下 LDAP 信息的存储	38
3.4.4	基于 LDAP 的语义网领域本体的自动构建	39
3.5	语义网服务本体库的自动构建	40
3.5.1	语义网服务本体	40
3.5.2	语义网服务本体的形式化描述分析	41
3.5.3	语义网服务本体自动构建过程	43
3.6	本章小结	45
第 4 章	语义网信息检索中本体匹配问题的研究	46
4.1	语义网信息检索中引入本体匹配技术的作用	46
4.2	本体匹配理论与相关技术分析	47
4.2.1	本体匹配理论	47
4.2.2	本体匹配技术	52
4.3	语义网本体自动构建理论下的本体匹配	57
4.3.1	语义网自动构建理论下本体匹配的要求	57
4.3.2	语义网本体匹配技术的选择	58

4.3.3	语义网本体自动构建理论下本体匹配的指导思想	59
4.3.4	语义网本体自动构建理论下本体匹配的预处理	59
4.3.5	语义网本体自动构建理论下领域本体的匹配	62
4.3.6	语义网本体自动构建理论下服务本体的匹配	66
4.4	本章小结	75
第 5 章	语义网信息检索中本体组合及其匹配的研究	76
5.1	语义网环境下本体组合研究的必要性	76
5.2	语义网环境下本体组合的技术基础	77
5.3	语义网环境下本体组合问题的实质	81
5.4	语义网服务本体的组合及匹配	82
5.4.1	语义网用户服务请求判定	82
5.4.2	语义网服务本体组合问题	84
5.4.3	语义网服务本体组合的匹配	89
5.5	本章小结	91

应用篇：基于微本体的微博信息管理机理

第 6 章	基于本体技术的微博信息管理模型	94
6.1	微博的信息资源特点	94
6.2	微博的信息组织方式	95
6.3	微博信息质量分析	97
6.4	用户层面微博平台的信息管理	98
6.4.1	网络提供给微博用户的信息组织界面	98
6.4.2	用户的自我信息组织方式	98
6.4.3	用户之间的互动方式	99
6.4.4	用户兴趣建模方法	99
6.4.5	用户兴趣模型表示	100
6.4.6	用户兴趣信息获取	101
6.5	本章小结	102
第 7 章	基于本体的微博的信息组织研究	103
7.1	微博现有的信息组织方式——Folksonomy	103
7.1.1	Folksonomy 的特点	103
7.1.2	Folksonomy 的不足	104
7.2	微博信息组织方式的改进	105

7.2.1	Folksonomy 的优化角度的选择	105
7.2.2	网络信息资源 Folksonomy 的优化途径	105
7.2.3	具体优化建议	105
7.3	基于本体技术的微博信息组织机理	106
7.3.1	原则	106
7.3.2	基于本体规范的微博信息组织模型	108
7.3.3	Folksonomy 与本体的融合及微本体	110
7.3.4	微博信息组织中的微本体构建具体过程	111
7.3.5	微博信息组织中微本体的更新	113
7.4	本章小结	113
第 8 章	基于本体的微博信息传播	114
8.1	微博信息传播	114
8.1.1	信息传播的特征	114
8.1.2	微博信息传播的影响因素	115
8.1.3	微博信息传播过程	117
8.2	微博信息传播建模	118
8.2.1	建模要素	118
8.2.2	建模原理	119
8.2.3	模型构建	120
8.2.4	微博传播与微博热度	121
8.2.5	特征提取	123
8.3	本体规范下的微博信息传播	124
8.3.1	微博信息传播本体模型构建	124
8.3.2	基于本体的微博信息传播管理机理	125
8.4	本章小结	126
第 9 章	基于 Folksonomy 和本体融合的微博信息推荐机制研究	127
9.1	微博信息推荐中的 Folksonomy 和本体融合——微本体	127
9.2	基于微本体架构的微博信息推荐方法	128
9.2.1	微博主题微本体的构建	128
9.2.2	基于改良 LDA 模型的微博信息推荐	130
9.3	实证检验与分析对比	131
9.4	本章小结	132
结语		133
参考文献		134

引入篇：微博及其研究现状

第 1 章 绪 论

1.1 背景及意义

1.1.1 研究背景

微博从产生至今,凭借自身快速、便捷和实时的优势受到人们的推崇与喜爱,相对来说,微博具有更自由的言论信息,这使得大量的用户开始使用微博,微博迅速成为把握社会脉搏的有力工具。2013 年,微博使用人数急剧增加,以 208.9% 的增长速度,令用户数突破 3.95 亿^[1,2]。此外,微博的使用领域也十分宽泛,例如,灾难事故中的微博寻亲、热点话题的微访谈等,这些均展示出微博巨大的社会影响力和人文气息^[3]。虽然微博极具发展优势,但是微博裂变式传播的特点同样极具社会危害隐患,这种传播方式很容易引起舆论狂潮,若不积极引导将产生很大的负面效应,其中伦理问题日益凸显出来,尤其是道德空间的紊乱和道德行为失范问题。

微博是互联网不断发展的必然产物,拥有信息广泛、传播快速的基本条件,随着微博使用人数的不断增加和内部运行机制的进一步改善,微博将在社会信息传播中占据越来越大的比例和作用^[4]。与此同时,我们仍应清楚地认识到我国正处在微博发展的初期阶段,微博内在运行机制(信息生产、筛选、传播)还不健全,由此造成信息冗杂、侵权等一系列问题的产生^[5],日后也将存在很多不确定性因素,需要不断加以完善。

微博使用者对某一条信息可以重复不断地评论或转发,完全不受时间和空间的限制,这可能导致该条信息以“核裂变”的方式在极短的时间内被数以千计的用户评论或转发,进而演化发展成为社会新闻事件,广受舆论关注;使用者在面对如此庞杂的信息时,十分容易被其中的垃圾信息迷惑,产生混乱和困扰^[6]。因此,杂乱庞大的信息、纷繁多变的热点话题以及名人效应等均会对微博信息长期传播效果产生某种程度的影响;另外,微博传播可以在无形之中影响用户的是非观、价值观等长期思想道德的形成,拥有很强的渗透力。这都说明无论从社会还是用户的角度,对微博信息的有效管理势在必行。

学术界也逐渐开始重视对微博进行有效管理的研究,取得了一些成果,如制定微博服务自律公约;指导微博运营商采取措施,做好微博的内容管理工作;倡议微

博使用者提高自身约束能力、思想道德意识和社会责任感,在使用微博的过程中积极监督其所发现的不良信息,不断提升自己的道德修养等。但是,这些研究只是从宏观的角度上规范微博的使用,寄希望于社会的力量去规范微博用户的行为,少量对微博信息组织和传播管理的技术性研究也只是提出大体思路,并没有给出具体可行的方案。本书的研究试图用本体建模理论规范微博信息的组织和传播,通过构建微博信息组织模型和传播模型,在技术上提出相应的解决办法,实现微博信息的有效管理。

1.1.2 研究意义

通过对信息在微博中的组织、传播及其用户推荐的性质和规律的研究,继而调查研究如何更好地控制社会舆论并加以正确引导,具有十分重要的理论价值和现实意义。本书在理论和实际应用方面都有较高的价值。

微博可以很好地满足用户对简单、实时交流的这一需求,迎合大众琐碎的生活状态。这要归功于微博网络独特的信息获取、发布和传播机制,这也是微博的核心价值所在。微博网络的产生和发展,拓宽了网络信息传播的渠道,微博可以实现信息共享、实时获取和大众参与,同时兼顾个体差异,这些特性既是微博网络发展的优势,也是微博网络的潜在隐患,如何规范其信息传播的秩序,确保网络安全等是一大难题。因此,需要对微博的迅猛发展不断完善并加强引导,抓住其中的发展机遇,更要迎接其带来的挑战。微博的获取和使用十分简单,安装软件注册即可对其操作,加上微博的即时性特点,在对于突发事件的报道和传播上,通常具有信息获取、发布和传播的领先优势。然而,一般而言,微博用户对所浏览的信息的认知是片面的、不对称的,外加一些虚拟互联网信息的掺杂,极易产生信息失真和信息造假等问题^[7,8]。微博信息多短小精湛,且微博网络不受时间和空间的限制,评论和转发十分自由、便捷,加上信息的不对称和用户认知的片面性,很多时候使用者会忽视或遗漏事实的真相,导致信息失真进一步扩大。面对这一问题,政府如果对其进行深入的调查和数据搜集,对此分析后再进行信息发布和辟谣,虽然会使信息的内容更加真实可靠兼具权威性,但是这一过程必然耗费较长时间,公众关注度和舆论热度已经下降,其恶劣影响也已产生,这样做往往事倍功半;政府如果简单地采取强制手段,勒令用户删除微博信息,禁止用户对信息进行评论和转发,虽然可以在一定程度上减少谣言的传播,但是效果不佳,还可能会激发群众的不满,降低群众对政府整体舆论的评价和政府公信力,激发社会矛盾^[9-12]。因此,实现微博信息的有效管理成为大势所趋。

(1) 本书的理论价值主要在于对微博的信息组织架构、传播模式和影响力形成机理进行研究,利用微本体技术改善微博信息资源的组织方式,寻找利于微博管理

的最合适的信息架构,分析微博产生裂变式传播能量和海量信息的动因,提出基于微本体建模的微博信息管理体系。

(2) 本书的应用价值在于通过微博理论研究、总结、归纳和发现微博发展的模式、趋势与规律,增强微博用户对微博性质和功能的全方位、深层次理解,增强用户对其传播特征进行准确判断和认知的能力,对微博媒介中的网络信息进行有效的排除和选择,从而减少或避免由于微博的不正确使用而对个人、社会造成的危害。并在此基础上提出规范我国微博发展的有效方案,改变现有用户存在认识程度有限、政府微博信息管理工作机制不健全、运营商功能定位不清、盈利性差等问题的现象,进一步加强、推广和完善微博建设,提升微博的应用水平和能力。

1.2 国内外研究综述

国内外学者对于微博信息管理这一主题的研究内容具有较为明显的差异。国外在这一方面的研究主要以“应用”为主,作者在中国高等教育文献保障系统(China Academic Library&Information System, CALIS)外文期刊网等英文数据库中,对 micro-blog 等主题词进行交叉搜索,发现在 2006~2013 年的时间范围内, CALIS 外文期刊网中相关文章的题目中含有“微博信息组织”或者“微博信息传播”字样或词语的文章总计为 763 篇,其中提及 Twitter 的文章题目共计 647 篇。由此可大胆推出,在国外 Twitter 就是微博的代名词。Hall 等^[13]选用了一种超文本引导主题搜索(Hypertext-Induced Topic Search, HITS)算法来计算 Twitter 上用户的中心度和权威度,并将用户之间的关系分成三类:信息共享关系、信息搜集关系和朋友关系;Kipp^[14]则将用户分为九大类,其中信息共享(Information Share, IS)、意见/抱怨(Opinion/Complain, OC)、随机想法(Random Think, RT)和关于我的一切(My Everything, ME)这四类占据主体地位,其理论依据在于对 Twitter 中用户发布的状态信息的内容进行详细分析。Kaplan 等^[15]发现网络信息的传播与微博媒介的网络结构以及用户个体的特性两者息息相关,因此,提出一个多层次的网络(Multi Layer Network, MLN)结构,用来描述微博用户个体所特有的多项属性,并且在线性阈值(Linear Threshold, LT)模型的基础上,在 MLN 的结构上构建了一个多阈值模型。Welch 等^[16]将独立级联(Independent Cascade, IC)模型中两个用户之间的激活概率设置为分布,收集汇总 Twitter 的微博转发数据,通过数据分析,利用马尔可夫蒙特卡罗方法,评估模型的传播概率,并且可以较好地预测 Twitter 中标签和 URL(Uniform Resource Locator)的传播路径。Kwak 等^[17]提出以下几方面相关信息传播的内容:影响因素、话题语义、网络结构和用户随时间推移其活跃态的变化,利用贝叶斯逻辑回归方法,在属性集成架构(Attributes Integrate Construction, AsIC)模型的基础上,推断 Twitter 网络各个节点之间随时间变化的传播概率。Ma 等^[18]进行了大量关

于用户行为的实证研究,提出了LI(Linear Influence)模型,该模型对各节点之间的影响力进行有效评估,在此基础上预测其信息传播的流向和趋势,并通过实际证明了该模型的精准性。国外的微博领域中, Twitter所占比例大,占有绝对优势, Twitter也因此成为微博的代名词,学术界也将 Twitter的信息管理问题作为直接研究对象,这些研究有很强的个案性,很多研究无法在其他微博系统上推广使用。

国内关于微博信息管理研究的起步比较晚,真正意义上的研究开始于2010年,而关注的焦点在于微博信息传播与微博用户个性化推荐这两方面。在少数理论文献研究中,介绍性的文章又占了多数,对于微博中信息组织、传播方式、影响力形成要素等方面的研究尤为匮乏,而这些却恰恰是微博发展的基础。

1.2.1 微博信息组织研究

Xu等^[19]研究微博允许用户之间交换短篇内容,如句子、图像和视频链接等信息组织形式的机理; Velikovich等^[20]主要研究了微博三个方面的特点:微博热点话题、消息内容和消息活跃时间; Lu等^[21]比较了国会主题词表中的主题词与 Library-Thing 标签,发现用户标签是提高图书馆资源的可获取程度的有力工具; Lee等^[22]认为目前关于用户标签和关键词的比较研究还只停留在学术领域,且研究目标多为专业性较强的调查研究,一般互联网用户的研究少之又少,尤其是微博信息的研究尤为欠缺。

国内学者的研究成果同样有很大的局限性。刘鲁等^[23]对基于机器学习的中文微博情感分类进行了实证研究;王晶等^[24]对当前基于信息数据分析的微博研究进行系统梳理,提出微博信息传播三大构件的概念,归纳了此类研究的主要研究内容及方法,总结了国内外围绕微博信息传播三大构件所取得的主要研究成果;杜伟夫^[25]、章成志等^[26]以腾讯微博作为研究和调查对象,研究了不同领域用户对标签主题表达能力的差异,并对发现的差异进行了详细研究;吴丹等^[27]发现用户在使用微博的过程中给予信息一个个性化易分辨的标签,并通过标签的相关程度对信息进行组织,这一行为和过程极大地简化了信息检索过程,带来了极大便利;潘婵等^[28]当前的研究还十分局限,范围较小,通过对标签和关键词的用户行为进行分析,发现娱乐和学术两个领域的标签相似程度存在很大的差别;李林红等^[29]认为微博社会网络是一个自组织系统,依据自组织理论从整体网络、个体网络、小团体、小世界效应四个方面构建模型,并以“可持续发展”话题为例,采用“滚雪球抽样方法”,考虑网络用户间的“发布、转发、评论、@、回复”关系,进行实证研究;刘风光^[30]、朱爱菊^[31]主要介绍了网络博客、微博的信息组织方式,探讨了网络博客、微博信息质量的分析原则,强调微博是在对人的关注和浏览中获取信息;胡媛^[32]利用自然语言处理和数据挖掘技术,通过对微博中的信息内容、用户的社会网络关系和交互情况以及某条信息传播过程进行分析与研究来发现微博中非正式信息交流的规律及其信息

流机制；夏雨禾^[33]、廖福生等^[34]研究发现微博信息传播过程有信息碎片、聚合受众和即时通信等特点；王晓光^[11]通过随机抽取新浪微博中的 3000 条微博，并对其博主、发博方式、微博内容、博文转发数量、评论数量、个人关注数及粉丝数等七项数据进行统计分析，研究信息传播的一般流程和微博的基本架构，观察微博用户基本行为和联系特征；夏雨禾^[33]以传统社会互动论为基础，统计分析了 438 个样本，并且探究了新浪微博信息产生互动的结构性要素和发生机制。

1.2.2 微博信息传播研究

平亮等^[35]以社会网络理论为理论依据，结合微博用户间“关注”与“被关注”的网络拓扑关系，以多种中心度和中心势测度指数为工具，分别从三个中心性（点度、中间和接近）着手对微博社会网络进行了具体分析，最后提出了具有深远意义的启示。

社会网络理论被微博研究者充分挖掘，研究者以“关注”与“被关注”为关联纽带，以用户为节点，通过分析用户节点和整个社会网络的中心性等若干指标，解析微博社会网络，通过定量方式发现微博传播特征。于洪等^[36]以统计方法为基础，找到了微博信息传播所具有的集中规律和信息传播途径所具有的三种典型传播模式：一触即发传播模式、多点触发传播模式和多级传播模式；董海军等^[37]、田占伟等^[38]利用复杂网络理论方法，对构建的微博信息传播网络，进行基于度、路径统计指标的分析，发现该网络具有集群性、小世界、高度中心化等特征；吴凯等^[39]基于行为预测的微博网络信息传播建模，提取了影响转发行为的四类特征，利用机器学习中的逻辑回归模型分析预测个体转发行为，并在此基础上融入用户个体差异，建立了一种基于行为预测的信息传播模型；田占伟等^[40]分析了微博信息传播模式、分享预测理论方法，基于被动攻击（Passive Aggressive, PA）算法提出了信息分享预测模型，以新浪微博数据为例验证了预测模型；邢立双^[41]研究突发事件中微博信息的传播与管理，发现微博信息具有时效性强、碎片化、病毒式传播等特征，在社会公众事件、突发性事件的传播方面，已经对传统媒体造成了冲击；樊鹏翼等^[42]认为信息传播过程中，微博的重要特征是用户之间的“关注”关系为弱联系，这使得微博信息传播网络的平均路径小，信息传播效率高；袁毅^[43]以新浪微博为研究平台，采集事件传播路径中的用户属性数据及行为数据，利用社会网络分析软件绘制信息传播网络图，并对传播网络的结构、路径及其影响因素进行分析，发现传播网络的形态与用户的影响力、节点的合理布局和外部干扰因素有关；方杲等^[44]探讨了哲学生活方式对微博信息传播的三点启示：微博信息传播的自我认识，微博信息传播者应有自己的理性原则，认识、承认和接受微博信息传播的限度；刘燕锦^[45]以当前应用最热门的社交网站和微博为分析对象，运用社会网络分析法，对两者的传播模式进行分析对比，分析同属于关系型的媒介技术平台是否因为功能设置的差异而造成信息在使用者之

间流动的差异及其具体表现；刘继等^[46]发现网络舆情传播模式中单关键点型传播模式以强势节点为主，没有桥节点支持，传播链一般较短，但传播速度快，多关键点型传播中多个强势节点通过桥节点进行信息交互，传播影响大，链式型传播模式由于缺乏强势节点，信息传播扩散能力较弱，但传播层次较深；黄淑敏^[47]通过对网络社区危机事件传播模式进行分析，发现传统媒体、网络媒体和社区网站构成了层层传递、相互影响的非线性链条传播体系；刘颖等^[48]将网络媒体按信息来源进行区分，发现网络论坛信息传播随时间变化的相似性与论坛作者发表量上的不平等特征；陈波等^[49]将传染病模型推广到泛在媒体环境下的开放系统中，利用带免疫的舆情传播模型对舆情传播媒介进行控制；钱颖等^[50]发现舆情信息具有记忆性，传染病模型在有些场合和舆情信息传播并不是很一致；Lü 等^[51]发现网络信息传播不仅依赖于小世界网络中的捷径，还与网络行为的多次社会性强化有关；唐晓波等^[52]利用共词网络对微博舆情进行分析，使用图论对现实中的舆情问题进行建模；郭海霞^[53]通过对微博的传播方式、传播路径的研究，提出社交网络开放平台中信息传播的模型和特点；郑蕾等^[54]发现当信息敏感度大于某一临界值后，明星用户在信息传播中的意见领袖角色随着信息敏感度的增加而逐渐减弱；康伟^[55]基于社会网络分析（Social Network Analysis, SNA）的突发事件网络舆情关键节点识别，发现网络结构与节点位置决定成员的“影响力”程度，中心关键节点的资源控制能力与信息输入、输出效率呈显著的正相关；张赛等^[56]提出了一种三角和算法用于探测用户粉丝数阈值，该算法根据散点分布的统计规律来估计使微博热度达到某一值的粉丝数的临界值，发现为使微博热度大于 10，用户粉丝数应大于 150；郭浩等^[57]为了定量地研究用户影响力，提出基于用户消息传播范围的用户影响力量化定义，并给出用户影响力的计算方法。张旸等^[58]分析了影响微博转发的特征因素，利用机器学习中的分类方法建立了特征加权的预测模型。

Westman 等^[59]社会科学领域的学者对微博信息的传播多采用统计学方法，对大规模数据进行研究，发现规律；Narayanam 等^[60]介绍了一种线性阈值模型，其基本思想是一个节点转换为活跃状态的概率会随着周围活跃节点个数的增多而增大，如果一个节点邻居中活跃节点数量超过一定阈值，该节点将转换为活跃状态；Lahiri 等^[61]利用系统动力学模型对社会网络中的信息传播进行了分析；Efron^[62]针对微博信息传播的研究，集中在数据挖掘、电子商务应用等领域；Suh 等^[63]重点研究微博信息分享过程涉及的两个主体：用户和信息的个人行为对微博信息传播的影响；Yu 等^[64]对我国微博信息的特有现象进行了研究。

1.2.3 微博用户推荐研究

早期的微博用户推荐源于微博用户的个性化研究的需要。王广新^[65]研究了基于