

基于数字仿真模型的网络 舆论引导理论与应用

吕德生 等 著



科学出版社

国家社会科学基金项目成果

基于数字仿真模型的网络舆论 引导理论与应用

吕德生 等 著

科学出版社

北 京

内 容 简 介

中国互联网已经进入了繁荣发展时期,作为世界上网民规模最大的国家,网络对民众的影响已远远超过了其他传媒方式,互联网已成为思想文化信息的集散地和社会舆论的放大器,网络舆论引导的重要性日益凸显。

本书从学科交叉的角度展开,借助数字仿真模型来研究网络舆论引导规律,介绍并探讨网络舆论的自身规律和网络舆论引导相关技术,阐述并建立网络舆论领袖发现与筛选模型,设计并实现网络舆情倾向性分析系统与网络社区舆论引导系统。

本书可供网络传播领域的科研人员参考,也可作为高等院校相关专业的教材。

图书在版编目(CIP)数据

基于数字仿真模型的网络舆论引导理论与应用/吕德生等著. —北京:科学出版社, 2015

ISBN 978-7-03-043247-6

I. ①基… II. ①吕… III. ①互联网络—舆论—系统仿真—研究 IV. ①G219

中国版本图书馆 CIP 数据核字 (2015) 第 022842 号

责任编辑:王 哲 董素芹 / 责任校对:钟 洋

责任印制:赵 博 / 封面设计:陈 敬

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>

三河市骏杰印刷有限公司印刷

科学出版社发行 各地新华书店经销

*

2015 年 3 月第 一 版 开本:720×1 000 1/16

2015 年 3 月第一次印刷 印张:13 1/2 插页:2

字数:267 000

定价:65.00 元

(如有印装质量问题,我社负责调换)

编写委员会

主 编：吕德生

编 委：何 慧 梁 冰 张伟哲

前言

做好网上舆论工作是一项长期任务，要创新改进网上宣传，运用网络传播规律，弘扬主旋律，激发正能量，大力培育和践行社会主义核心价值观，把握好网上舆论引导的时、度、效，使网络空间清朗起来¹。

——习近平

中国互联网的规模价值正在日益放大，根据中国互联网络信息中心（China Internet Network Information Center, CNNIC）发布的《中国互联网络发展状况统计报告》，自2008年6月我国开始成为世界上网民规模最大的国家，并且网民规模和普及率持续增加，截至2014年6月，中国网民规模达6.32亿。随着网络的迅速发展，网络舆论引导的重要性日益凸显，前国家主席胡锦涛指出：“舆论引导正确，利党利国利民；舆论引导错误，误党误国误民”“互联网已成为思想文化信息的集散地和社会舆论的放大器，要充分认识以互联网为代表的新兴媒体的社会影响力，加强主流媒体建设和新兴媒体建设，形成舆论引导新格局”。

通过数字仿真的方法模拟，掌握复杂系统的运行规律，在很多科学领域被证明是行之有效的重要科技手段，因而借助数字仿真模型来研究网络舆论引导规律是一条颇具吸引力的道路。本书的研究工作从学科交叉的角度展开，主要内容包括：初探网络舆论、网络舆论引导相关技术、网络舆论领袖与影响力计算、网络舆论领袖的发现与筛选模型、网络舆情倾向性分析系统设计与实现、网络社区舆论引导系统设计与实现。

中国互联网起始于20世纪90年代，由于Web 2.0技术的出现，2003年年初互联网进入了繁荣发展时期，人们开始在网络上讨论公共话题。民众受到网络的影响已远超过其他传媒方式，网络舆论的主要特点包括即时监督性、相对匿名性、开放互动性、偶然突发性、“领袖”权威性、大众盲从性、群体“极化”性、预测可能性、预警必要性、引导可行性等。网络舆论本身具有一定的规律性，有效地对网络舆论的规律加以把握和利用是实现网络舆论正确引导的前提条件。

网络论坛是重要的民众意见传播渠道，其中“意见领袖”的网民影响力最大，容易引起其他网民的共鸣和追捧。本书从两个角度构建网络论坛（Bulletin Board System, BBS）舆论领袖发现模型：一是从语义的角度出发，通过计算词语在回帖传播链上的影响力，设计并实现基于高权重词集的论坛成员影响力计算模型；二是继承PageRank算法思想，设计并实现基于回复关系随机游走的BBS舆论领袖识别模型，并提出了用

注：1. 2014年2月27日中央网络安全和信息化领导小组第一次会议。

节点入度初始 PageRank 向量的方法来减少迭代次数,经验证该模型一般通过 50 次左右的迭代就可以收敛。本书提出了一种基于网民的论坛语料和网民网络行为统计属性特征的互补性网络舆论领袖筛选模型,所运用的“阶梯式”评价算法保持了基于对话链关系的影响力的持续性,通过多次机器筛选和人工筛选的对比实验,该模型筛选出的舆论领袖的成功率在 85%以上。

同时,本书研究并实现了一个网络舆情倾向性分析系统,该系统可以在从多个其他领域迁移到一个目标域的场景时,进行知识的迁移,提高迁移的精确度。通过跨领域的分析方法——基于 Markov 逻辑网的多任务迁移学习,充分利用其他领域已有的知识,比只从本领域学习能得到更多的知识,通过实验验证了模型的优越性,解决了跨领域舆情倾向性分析问题。

针对网络舆情引导自动化程度低、手段单一、资源匮乏的缺陷,实现了在网络社区进行舆情引导的自动化引导系统。针对不同的信息来源,实现了一种具有多通道信息有效采集功能的网络爬虫,具体包括网站通道、博客通道、BBS 通道、微博通道。传统的引导方式都是由人工登录社区发布信息的,本系统实现了自动化登录网络社区的功能:对于 BBS 类的社区,采用 WebBrowser 控件的方式编写程序,对于微博类的社区采用微博开放平台提供应用接口编写程序,两种登录发布信息的方式都有很好的稳定性。

本书由哈尔滨工业大学吕德生主笔,哈尔滨工业大学何慧、中国科学技术信息研究所梁冰、哈尔滨工业大学张伟哲参与了相关章节的研究工作。

感谢哈尔滨工业大学研究生郭琛、李颖、王丹、陈琳、李文娟、邸文晨、张东杰等的努力工作。感谢朴松昊、左旺孟、高超等的宝贵意见。

本书是国家社会科学基金项目“基于数字仿真模型的网络舆论引导规律研究”(项目编号:09CXW018)的研究成果。

由于作者学识有限,本书难免存在不足之处,欢迎读者朋友批评指正。

吕德生

2014 年 11 月

目 录

前言

第 1 章 初探网络舆论	1
1.1 舆论学相关研究	1
1.1.1 舆论学基础	1
1.1.2 网络舆论概念的形成	3
1.1.3 网络舆论引导	3
1.1.4 方兴未艾的网络舆论引导研究	4
1.2 网络舆论的形成	9
1.2.1 网络舆论客体的公共性	9
1.2.2 网络舆论主体的复杂性	10
1.2.3 网络舆论的传播模式	11
1.2.4 网络舆论的形成条件	13
1.2.5 网民态度的表现分类	17
1.3 网络舆论的信息失真	19
1.3.1 舆论环境的虚拟性	19
1.3.2 始作俑者的假意	19
1.3.3 公众成见的存在	20
1.3.4 大众媒体的介入	21
1.3.5 暴力语言引发的失控	21
1.3.6 社会环境的作用	22
1.4 网络舆论的特性分析	23
1.4.1 即时监督性	23
1.4.2 相对匿名性	24
1.4.3 开放互动性	25
1.4.4 偶然突发性	26
1.4.5 “领袖”权威性	27
1.4.6 大众盲从性	28
1.4.7 群体“极化”性	29
1.4.8 预测可能性	30
1.4.9 预警必要性	30

1.4.10	引导可行性	31
第 2 章	网络舆论引导相关技术	33
2.1	聚焦网络舆论领袖的网络爬虫技术	33
2.1.1	聚焦网络爬虫技术	33
2.1.2	网络爬虫的爬行策略	34
2.1.3	网页分析算法	35
2.2	面向热点话题发现的文本聚类技术	36
2.2.1	文本聚类的数学描述	36
2.2.2	文本的特征抽取模型	37
2.2.3	相似度计算方法	37
2.2.4	基于文本聚类分析的话题提取方法	38
2.3	网络舆情倾向性分析的相关理论与技术	40
2.3.1	基于语义的网络舆情倾向性分析方法	40
2.3.2	基于机器学习的网络舆情倾向性分析方法	40
2.3.3	基于支持向量机的网络舆情倾向性分析	41
2.3.4	基于 K 最近邻方法的网络舆情倾向性分析	42
2.3.5	基于 N -gram 方法的网络舆情倾向性分析	43
2.4	网络舆论传播可视化技术	44
2.4.1	关于信息可视化的研究	45
2.4.2	层级化传播路径	45
2.4.3	面向网络舆论领袖的可视化模型	46
2.4.4	可视化模型的实现流程	48
第 3 章	网络舆论领袖与影响力计算	49
3.1	舆论领袖研究概述	49
3.1.1	舆论领袖的相关概念关系界定	50
3.1.2	舆论领袖的分类	50
3.1.3	网络舆论领袖的研究现状	51
3.1.4	舆论领袖的应用过程	53
3.2	舆论领袖的基本特征及其测量方法	55
3.2.1	传统舆论领袖的基本特征	55
3.2.2	现代网络舆论领袖的基本特征	56
3.2.3	传统舆论领袖测量方法	57
3.2.4	网络舆论领袖测量方法	59
3.3	网络舆论领袖的影响力计算模型	59
3.3.1	基于内容的影响力计算模型	59

3.3.2	基于链接分析的影响力计算模型	61
3.3.3	评价指标与皮尔逊 (Pearson) 相关系数	63
3.3.4	基于 NodeXL 的节点重要性评价指标	64
3.3.5	简单影响力扩散模型	65
第 4 章	网络舆论领袖的发现与筛选模型	69
4.1	BBS 舆论领袖发现模型	69
4.1.1	数据获取	70
4.1.2	BBS 网页基本信息抽取	70
4.1.3	BBS 网页聚类	72
4.1.4	基于高权重词集的 BBS 舆论领袖发现技术	73
4.1.5	基于随机游走的 BBS 舆论领袖发现技术	75
4.2	BBS 舆论领袖发现模型合理性测试和分析	77
4.2.1	单帖实验验证模型合理性	78
4.2.2	版块实验验证模型合理性	83
4.2.3	计算结果评估	84
4.2.4	舆论领袖地位稳定性分析	88
4.2.5	帖子生命周期分析	89
4.3	网络舆论领袖筛选模型及其评价	91
4.3.1	论坛舆论领袖筛选模型	92
4.3.2	基于简单语料的影响力阶梯式评价方法	95
4.3.3	基于舆论领袖网络统计属性特征的评价技术	98
4.3.4	舆论领袖综合评价	100
4.4	网络舆论领袖筛选模型实现及实验结果分析	101
4.4.1	网络舆论领袖筛选的模块化实验过程及测试环境	101
4.4.2	舆论领袖筛选的模块化实现	102
4.4.3	基于语料的评价算法实现	104
4.4.4	基于网络统计属性特征的评价算法实现	106
4.4.5	舆论领袖筛选实验结果分析	110
第 5 章	网络舆情倾向性分析系统设计与实现	117
5.1	研究现状分析	117
5.1.1	网络舆情研究现状	118
5.1.2	网络舆情倾向性研究现状	119
5.1.3	存在的问题	120
5.2	跨领域舆情倾向性分析技术	121
5.2.1	多任务迁移学习	122

5.2.2	Markov 逻辑网	124
5.2.3	基于 Markov 逻辑网的迁移学习	129
5.2.4	跨领域舆情倾向性分析的关键理论研究	130
5.2.5	不确定超图的概念和相关定义	131
5.2.6	一个大型的不确定超图频繁子图挖掘相关定义	132
5.2.7	不确定超图复杂性证明	133
5.2.8	近似算法设计	137
5.3	网络舆情倾向性分析系统设计	139
5.3.1	Alchemy 系统软件包	139
5.3.2	系统设计思想	140
5.3.3	模块描述	140
5.4	网络舆情倾向性分析系统实现	143
5.4.1	系统环境	143
5.4.2	数据预处理	143
5.4.3	系统的实现	145
5.4.4	舆情倾向性分析结果	149
第 6 章	网络社区舆论引导系统设计与实现	152
6.1	面向网络社区舆论引导系统的爬虫设计	152
6.1.1	爬虫主要模块设计	152
6.1.2	多通道爬虫设计	154
6.2	多通道信息采集技术运用	155
6.2.1	网页正文提取	156
6.2.2	网站通道内容采集	156
6.2.3	博客通道内容采集	158
6.2.4	BBS 通道内容采集	160
6.2.5	微博通道内容采集	163
6.2.6	信息采集单机吞吐量实验	165
6.3	图形验证码识别技术运用	169
6.3.1	图形验证码识别的研究现状	169
6.3.2	图形验证码识别步骤	170
6.3.3	图形验证码识别测试实验	179
6.4	话语复述技术运用	180
6.4.1	话语复述研究现状	180
6.4.2	逆向最大匹配算法	181
6.4.3	同义词替换	183

6.5	网络社区发布技术运用	184
6.6	系统总体结构及模块设计	186
6.6.1	精品语料采集模块	187
6.6.2	小号语料采集模块	187
6.6.3	虚拟人物培养模块	188
6.6.4	话语复述模块	189
6.6.5	虚拟人物舆论引导模块	189
6.6.6	账号管理模块	190
6.6.7	配置文件管理模块	191
6.7	系统功能测试	191
6.7.1	网络社区中帖子人工置顶测试	191
6.7.2	人物影响力比较	192
	参考文献	194

第 1 章 初探网络舆论

网络迅速普及让越来越多的声音在网络上出现,不管声音发出者的身份和地位如何,他们在网络上享有同等程度的自由。数量巨大的“草根网民”形成了强大的网络舆论主体,上传、互动和共享使得网络舆论正以一种前所未有的阵势主导社会总体舆论走向,同时作为一种新生活话语力量改变着现实社会。

广义的网络舆论包括一切以网络作为传播渠道形成的有声和无声的舆论的总和;狭义的网络舆论指论坛和新闻跟帖所组成的网络文本集合。网络媒体、传统媒体和公众对公共话题事件综合交错的传播,既作用于现实,又作用于网络,才能形成持续而广泛的网络舆论。

1.1 舆论学相关研究

“舆”的本义是指车子,《周礼·考工记·舆人》中说“舆人为车”,意思是舆人制造车子。“舆人”泛指下层民众,舆论的本义是公众意见。《汉语大词典》中解释【舆论】条为:公众的言论。舆论一词最早出现在《三国志·魏书·王朗传》中“设其傲狠,殊无入志,惧彼舆论之未畅也,并怀伊邑”。南京大学王雄^[1]认为,舆论是人们对于一切感兴趣的社会生活领域尤其是社会公共事务所产生的一切意见、意愿、态度、情绪、期待、评价和想象的公开表达,是持续性与变化性、公共兴趣和个人兴趣的统一;舆论包含着相互独立、又相互关联的自发的公众舆论和自为的新闻舆论,是多种主导性社会意见的集合体。舆论往往是自发形成的,早期的传播形式主要有人际传播、组织传播和群体传播等。大众媒体的加入使得舆论有了新的内容和形式,首先,大众媒体扩大了社会事件的传播范围,提高了传播速度,迅速吸引人群的注意,覆盖更广泛的群体;其次大众媒体对新闻的报道是有选择的,对大众舆论的方向有着明显的引导作用。有关舆论的定义有很多,本书采用陈力丹^[2]在《舆论学——舆论导向研究》中的定义:舆论是公众关于现实社会和社会中各种现象、问题所表达的信念、态度、意见和情绪表现的总和,具有相对的一致性、强烈程度和持续性,对社会发展和有关事态的进程产生影响,其中混杂着理智和非理智的成分。

1.1.1 舆论学基础

西方舆论学的研究开始较早并且较为成熟。卢梭早在 18 世纪的著作《社会契约论》中首次将拉丁文体系中的“公众”与“意见”两个词汇联系起来,用来表达人们

对于社会性的或者公共事务方面的意见,即“舆论”。1922年美国记者沃尔特·李普曼(Walter Lippman)的《公众舆论》(*Public Opinion*)标志着舆论学的诞生。他对公众舆论是民主的原动力提出质疑。他认为公众舆论并不能代表真正的民主,而是由政治家运用模糊、有多种解读方式的具有煽动性的话语来操纵的^[3]。其书中的很多观点对当今的舆论研究仍然有着一定的启示作用。

国外(以美国为主)舆论学研究经历了三次研究方向的大转向后,发展到现代阶段^[4]:19世纪中期从哲学本位阶段转向到社会学研究阶段;20世纪初从社会学研究阶段转向社会心理学阶段;20世纪中期从社会心理学阶段转向舆论研究的本体阶段,开辟出了集体行为和社会心理研究、态度和意见研究、政治行为研究和大众传播研究并举的局面。国外舆论研究专著倾向于讨论基础性和较为宏观的舆论学问题,已经进入较为成熟的阶段,舆论学家的研究重点正由舆论基础学转向舆论与现代科技的结合。

中国和西方的学者因处于不同的社会语境,对舆论研究各自有其不同的价值取向和实际关注。国内对舆论学的研究目前尚未形成完备体系,甚至对“舆论”本身的定义仍然存在分歧。争论的焦点在于,舆论是否同时包含社会公众的一致性意见和少数群体特别是统治阶级的集体意志^[5]。纵观历史,19世纪以前中国关于舆论的研究几乎空白,虽然舆论在现实生活中已经深刻地影响了社会的发展变化,然而关于舆论概念的明确研究却几乎为零。

虽然舆论在我国长期的封建统治中早有运用,但真正作为专业化名词在我国出现大约开始于20世纪初叶的晚清政府动荡时期。最早将西方“公共”概念与舆论研究相结合大约始于辛亥革命时期,早在1909年胡汉民就在《近年中国革命报之发达》中提到了舆论的公共性质。几乎与胡汉民同时研究社会舆论问题而影响更大的是梁启超,他在《舆论之母与舆论之仆》《〈国风报〉序例》等文章中对舆论的性质、作用和制造舆论的方法等问题进行了专门的研究,就其论述所涉及的舆论问题范围之广、见解之深、影响之大而言,堪称中国研究舆论问题的第一位大师。而他的观点与胡汉民在很多地方是不谋而合的。1914年汪馥炎在《甲寅》杂志发表的《舆论与社会》已经可以看出舆论学萌芽的趋势,五四时期,罗家伦、徐宝璜等新式知识分子在西方学术精英的启蒙推动下,开创了中国专业化的现代舆论学研究传统。徐宝璜的《舆论之研究》是中国现代舆论学研究的开山之作,将舆论研究引入现代学术园地,运用现代学术观点和方法比较系统、全面、清晰地论述了现代舆论学的一些基本问题^[6]。

继徐宝璜之后,早期留学日本的新闻学者如邵飘萍、任白涛等在自己的著作中也将“公共性”的概念引入新闻学的研究之中,使之成为中国现代新闻舆论研究的重要内容。1941年叶明发表的《舆论的形成》是目前所知的我国最早出版的舆论学专著。1981年第4期《陕西新闻研究》刊登的《舆论学初探》一文,是当代中国舆论研究的开端,此后,李祖兴、刘允州、张晓虎等也发表了舆论学论文^[7]。1988年起,《基础舆论学》《当代舆论学》《现代舆论学》《民意学》《舆论学——舆论导向研究》等一批著作的诞生,标志着我国舆论学繁荣时期的到来。

1.1.2 网络舆论概念的形成

中国互联网起始于20世纪90年代,由于Web 2.0技术的出现,2003年年初互联网进入了繁荣发展时期,人们开始在网络上讨论公共话题。至此,“网络舆论”才作为一个概念被正式提出,属于舆论学的研究范畴,从此网络舆论受到越来越多学者的关注。由于研究历史较短,2008年前针对网络舆论进行研究的专著不多,仅《网络传播理论》《中国网络新闻事业管理》《网络传播学:一种形而上的透视》三本著作中的部分章节提到网络舆论,着墨不多,论述也不够系统深刻^[8]。

长期关注网络的新闻传播学者彭兰说“就我所知,中国网络舆论发端的标志性事件是1998年5月印尼排华事件后全世界华人(也包括国内)在网上的抗议活动”^[9]。最终网络舆论的巨大压力迫使印尼政府成立一个独立委员会,负责调查这次事件。这是中国网络舆论的发端。而其后人民网为抗议北大西洋公约组织轰炸我国驻南斯拉夫联盟共和国大使馆而开设抗议论坛,“刘涌案”“木子美事件”等网民在网络上大量发表意见,中国网络舆论在不断扩大。网络舆论不断引起人们的重视并提上日程。此后,针对网络上的各种案例,学者进行了大量研究,大多是关于网络舆论的作用和网络舆论控制引导方面的研究。

在“网络舆论”概念出现之前,有“在线舆论”“网上舆论”“互联网舆论”等提法。2003年之后,“网络舆论”作为一个固定的概念被使用;2004—2006年是网络舆论的发展时期,网络舆论的相关研究呈逐年上涨趋势,研究的主要问题涉及网络舆论主体性、网络舆论的引导与监控、网络舆论的法律伦理问题、网络舆论的功能及力量等内容;2007年至今是网络舆论的繁荣期,出现了大量的相关研究文献,网络舆论的广度和深度也有了很大提高。很多文献是从热点事件进行透析来窥探网络舆论的特点的,如以“虐猫事件”“华南虎事件”等为案例的网络舆论研究。构建和谐社会的议题开始出现,一些文献进一步探索网络舆论中涉及的道德和法律问题,此外还出现了从社会心理学的角度对网络舆论进行研究的文献,网络舆论已逐渐发展为一个跨领域、跨学科的重要研究方向。

1.1.3 网络舆论引导

我国对舆论引导问题始终十分关注,明确提出了舆论引导的思想和要求。1989年11月,江泽民在《关于党的新闻工作的几个问题》的讲话中指出要“按照党和人民的意志、利益进行舆论导向”。党的十六届四中全会通过的《中共中央关于加强党的执政能力建设的决定》指出要“牢牢把握舆论导向,正确引导社会舆论”。

20世纪90年代以来,如何运用传播、沟通、控制等手段对舆论进行疏导、引领、控制逐渐成为热点问题,许多学者开始了相关的研究工作,“舆论导向”“舆论调控”“舆论引导”都是该问题的不同提法。2008年,国家主席胡锦涛提出“舆论引导正确,利党利国利民;舆论引导错误,误党误国误民”,之后该问题基本都使用“舆论引导”一词。

民众受到网络的影响远远超过了任何一种传媒方式。网络传播的速度之快,也是其他任何传播方式无法比拟的。大量的信息不断地充斥着人们的眼球和思想,其中包括积

极的方面,同时也包括消极的方面。网络让民众能够公开地表达自己对现有的社会政策、社会管理、社会制度的看法、意见和态度。网络舆论引导的主体,主要是政府;网络舆论的执行主体,主要是网络各个媒体;网络舆论的参与者,主要是网民。

坚持正确的舆论引导,是要求各个媒体在发布信息的同时,要坚守国家相关法律法规,做到真实可靠,避免具有消极影响的煽动意识。网络新闻也要具有正确的引导,坚持网络新闻的真实性原则,才能使网络舆情朝着健康的方向发展。网络是民众心声表达的最重要的方式,通过网络,可以了解到民生、民意、民情,然后认真解决存在的问题。

舆论引导可以对社会发展产生积极的作用,例如,新浪网 2012 年全国两会报道,据广州日报官方微博和大洋网官方微博发起的调查来看,有 76%的网友支持将公共供暖延伸到南方,称南方湿冷,应安装暖气,但前提是要以环保的方式,而有 23%的网友却反对,反对的理由是“恐怕会造成资源浪费”“会加剧环境污染”等。为了使两会达到公开透明的原则,网络媒体时刻对其进行跟踪报道。主要包括民生热点、政策聚焦、提案议案、分析评论、深度报道等,其中民生热点放在了篇幅的首尾,能够切实地维护民众的利益的网络舆论体现了其积极的作用。

网络舆论在彰显积极一面的同时,也存在着民意的偏激和随意,给社会带来了消极的反作用,如果不能进行正确的舆论引导,就会影响经济社会的发展和稳定。

舆论引导的作用,主要包括三方面。

(1) 舆论引导与政治。社会的舆论是民众心理的反映,常常影响着各个国家的局势的发展和人们的行为。互联网的及时和公开,为民众和政府提供了一个良好的沟通平台。政府注重网络舆论的反应,有利于建设和谐社会。

(2) 舆论引导与社会风气。较之政治和法律,网络舆论更加具有感情性和自由性。公众的好恶,可以更好地体现时下的社会风气和社会制度的形成。通过正确的舆论引导,可以起到社会的调节功能,促进人与社会的和谐发展。

(3) 舆论引导与教育。民众是形成网络舆论的主要参与者,民众的教育和道德是不容忽视的。政府的舆论监督如果疏忽,就可能会产生错误的舆论导向,因此需要民众的监督和意见的反馈。从另一方面讲,也可通过网络提升民众的道德品质水平。民众整体道德水平的提高有利于促进社会文明的进步、社会政治的安定。

网上的信息量十分巨大,真伪度造成的舆论影响大相径庭,仅依靠人工的方法难以解决网上巨大的信息含量的收集和处理,需要加强筛选信息技术的研究,及时应对网络舆情,起到积极的引导作用。但当前网络社区舆论引导技术的主要问题在于手段缺乏、功能单一、自动化程度低、应对问题不及时。

1.1.4 方兴未艾的网络舆论引导研究

2000 年魏永刚^[10]从意识形态的角度提出了网络舆论传播中的导向问题,但并未引起学界的重视。直到 2003 年,SARS 爆发、宝马撞人案等网络舆论意见的真正形成,舆论引导问题才被予以关注。以此为滥觞,网络舆论的相关研究日益活跃。

1. 时间分布

自 2003 年 1 月 1 日，截至 2010 年 12 月 31 日，以“网络舆论”并含“引导”进行全文检索，检得文献 814 篇，2005 年以来共发表 788 篇，占总数的 96.80%。以“网络舆情”并含“引导”进行全文检索，检得文献 234 篇，2005 年以来共发表 234 篇，占总数的 100%。经处理，剔除非规范文献 15 篇，重合文献 12 篇，有效文献共有 1021 篇。文献基本情况如表 1-1 所示。同时，绘出了此阶段文献总量变化趋势图如图 1-1 所示。根据文献增长变化曲线图可以发现，此阶段科学文献是严格以指数方式快速增长的。按照普赖斯（Price）的逻辑曲线增长理论^[11]推测，国内对此领域的研究还处在学科的发展壮大时期，相对不成熟，体系待完善。

表 1-1 网络舆论引导类文献逐年变化一览表（2003—2010 年）

发表时间/年	文献数/篇	核心期刊/篇	优秀硕士学位论文/篇	博士学位论文/篇	重要报纸全文/篇	重要会议论文/篇
2003	8	4	2	0	0	0
2004	16	3	3	0	3	0
2005	26	6	8	0	1	1
2006	39	3	8	1	1	2
2007	64	10	23	0	5	0
2008	150	15	41	3	10	0
2009	249	17	41	7	15	2
2010	469	48	87	3	10	4

注：表中核心期刊来源于北京大学图书馆中文核心期刊目录第 5 版；核心期刊是衡量学术水平的重要指标，所以单列于此

由表 1-1 可见，优秀硕士学位论文、博士学位论文、重要报纸全文和重要会议论文的文献数并不稳定，呈现跳跃式的变化。博士学位论文和重要会议论文都有两年 0 篇数的出现。科学计量学中的大量数据表明^[12]，成熟学科的文献累积量呈线性增长趋势，文献数量增长减缓，仅维持相对固定的文献增长量。而一个新兴的研究领域在其发展初期，专业理论迅速发展，文献数量急剧增加，文献增长模型往往表现为严格的指数增长型，且后劲十足，发展迅速。这也就是现阶段舆论引导类研究文献的明显特征。2005—2007 年文献数增长率相对较小，从 2007 年开始增长速度逐年加快。

核心期刊数也在不断上升，仅 2010 年一年发表的核心期刊数几乎是前三年的总和。进一步研究发现，2007 年是网络舆论引导研究的转型之年。之前大部分文献的视野都集中在社会科学领域，呈现多定性少定量、多经验少批判、方法单一且随意的特点。但随着信息科学和网络技术的飞速发展，依托舆论学和信息学，网络舆情开辟了舆论研究的新走向。网络舆论研究在人文社会科学和信息科学两个领域中得到了飞速发展。

科技文献和科研项目之间具有密切的相关性，以 2007 年为时间节点，作为导向性、示范性、权威性和代表性的国家社会科学基金和国家自然科学基金，相关网络舆论/舆情的研究课题立项多达 8 个，而在此之前只有 2006 年华中科技大学新闻与信息

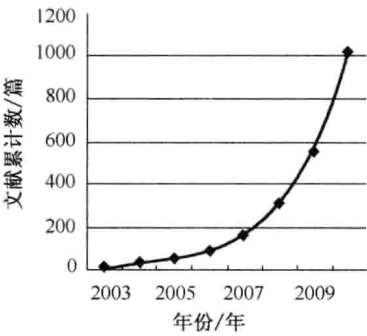


图 1-1 2003—2010 年网络舆论引导规律研究文献总量变化趋势

传播学院钟瑛的“互联网管理中的平衡与协调”课题的立项。表 1-2 和表 1-3 分别统计了 2007—2010 年的国家社会科学基金和国家自然科学基金中相关网络舆论的获准立项的研究信息，结合国家级课题的项目类别、名称、主持人、关键词等几个特征，分别对国家社会科学基金和国家自然科学基金的资助情况进行统计。

表 1-2 国家自然科学基金资助项目一览表（2007—2010 年）

年 份	项目序号	课 题 名 称	主 持 人	关 键 词
2007 年 (3 项)	1	面向公共安全的网络舆情预警方法研究	清华大学疏学明	网络舆情；突发事件；公共安全；热点信息
	2	网络舆论发展趋势分析核心技术研究	上海交通大学刘公申	文本倾向性分类；停用词表；特征选择；特征加权
	3	网络化数据挖掘方法研究	中国人民解放军总参谋部第六十一研究所李德毅	网络舆情；拓扑势；节点重要性
2008 年 (3 项)	4	基于特征概念网的网上离散文本信息舆情分析研究	上海交通大学李生红	概念网；本征特征；义元；概念消歧
	5	网络突现智能的突现计算模型研究	重庆邮电大学吴渝	复杂网络；突现行为；突现模型
	6	面向互联网舆情分析的文档自动摘要关键技术研究	北京大学万小军	文本分类；特征向量；子包含
2009 年 (3 项)	7	非常规突发事件应对的集群决策行为与法则研究	武汉理工大学谢科范	突发事件；生命周期；集群决策
	8	非常规突发事件网络舆情分析方法和预警机制的研究	哈尔滨工业大学陈晓东	非常规突发事件；网络舆情；生成过程
	9	非常规突发事件中网络舆情作用机制与相关技术研究	中国科学院计算技术研究所方滨兴	非常规突发事件；系统动力学建模
2010 年 (4 项)	10	支持网络舆情分析的观点挖掘及观点社群发现关键技术研究	东北大学王大玲	观点挖掘；情感词典；倾向性分析
	11	搜索与挖掘的新理论和新方法——支持舆情监控的 Web 搜索与挖掘的理论与方法研究	中国科学院计算技术研究所方滨兴	舆情监控；Web 搜索；信息挖掘
	12	基于跨媒体信息挖掘的网络舆情分析研究	中国科学院自动化研究所刘静	网络舆情；跨媒体；信息分析
	13	互联网信息的聚合与观点演化机制研究	北京交通大学沈波	信息抽取；模板；文档对象模型；XPath

表 1-3 国家社会科学基金资助项目一览表（2007—2010 年）

年 份	项目序号	课 题 名 称	主 持 人	关 键 词
2007 年 (5 项)	1	新形势下舆论引导新格局、新机制研究	中国传媒大学雷跃捷	舆论引导；工作机制
	2	WEB2.0 对社会舆论的影响研究	复旦大学新闻学院周葆华	突发事件；新媒体事件
	3	新形势下的舆论引导新格局研究	南京师范大学骆正林	舆论环境；舆论引导；新格局
	4	网络谣言的传播规律及应对策略研究	深圳大学传播学院周裕琼	传播过程；控制实验；谣言
	5	网络信息资源内容安全主动监控体系研究与实现	解放军南京政治学院上海分院李留英	分级分类；安全模型
2008 年 (4 项)	6	舆论引导力与社会舆情预警系统研究	清华大学李希光	突发事件；危机传播管理
	7	新形势下提升舆论引导力对策研究	四川大学蒋晓丽、侯雄飞	舆论引导；突发事件