

doi:10.3969/j.issn.1007-7162.2015.02.011

微博公共事件演化分析研究综述

张钰莎¹, 蒋盛益²

(1. 广东外语外贸大学 南国商学院, 广东 广州 510545; 2. 广东外语外贸大学 思科信息学院, 广东 广州 510420)

摘要: 微博公共事件的演化分析研究有助于预警公共事件的形成和社会管理, 近年来已成为一个重要的研究方向和热点. 微博公共事件演化分析包括公共事件自动发现、演化模型及语料库构建, 通过梳理这 3 个方面的研究成果, 对相关研究理论和方法的现状进行评析. 最后, 对微博公共事件语料库的构建、微博公共事件自动发现及演化分析的发展趋势进行展望, 为进一步研究提供参考.

关键词: 微博; 公共事件发现; 演化模型; 公共事件语料库

中图分类号: TP311 **文献标志码:** A **文章编号:** 1007-7162(2015)02-0058-06

Research Review on Public Event Evolution in Microblogging

Zhang Yu-sha¹, Jiang Sheng-yi²

(1. South China Business College, Guangdong University of Foreign Studies, Guangzhou 510545, China;

2. School of Cisco Informatics, Guangdong University of Foreign Studies, Guangzhou 510420, China)

Abstract: In recent years, researchers have had much insight on public event evolution in microblogging, which can help the early warning of public events and social management. This paper attempts to summarize the existing research achievements on public events detection, evolution model and construction of public event corpus, and provide comments and analysis on current research and related theories and methods. In addition, it makes references for further research on the construction of public events corpus, public events automatic detection and evolution analysis in microblogging, which can provide future work with some guidelines.

Key words: microblogging; public event detection; evolution model; public event corpus

微博作为一种新兴传播媒介, 由于发布和传播信息即时、多样、快速以及互动性良好, 已成为网民信息获取和情感沟通的重要渠道. 微博的影响力已经渗透到社会政治、经济、文化等各个领域, 其公信力超过广播、杂志、手机和户外广告, 位列电视、网络、报纸等传统媒体之后, 居于第 4 位. 微博在引发公共议题、激发公众参与公共事务热情方面具有独特的优势, 体现出强大的舆论引导作用和传播效果. 微博就像一把双刃剑, 不仅为公民表达观点提供了极大的便捷, 充当了社会公共事件中的危机化解助推器, 也成为了危机言论的诱发器和不良言论的放大器, 给社会管理带来巨大冲击和严峻挑战.

由于社会转型期间的结构性原因、相关职能部

门的缺位和越位、社会民众的文化心理及综合原因等造成了近年来国内公共事件的频发. 所谓公共事件是指能引起社会普遍关心并引发热议或可能出现社会波动的事件. 而微博公共事件是指围绕现实中的某个社会事件产生, 而在微博平台引起网民广泛讨论并进而形成强大网络舆论, 影响党和政府决策或造成重大现实影响的公共事件. 突发公共事件的快速发酵性需要快速的反应机制, 这与微博发布、传播迅速的特点相一致, 利用微博的这一特性将更有利于快速处理公共事件. 微博传播方式的“蝴蝶效应”极易形成公共事件舆情, 微博内容的碎片化容易导致信息的不完整和误解, 微博言论的泛化容易造成独立事件的放大, 这些特点使得微博在重大突

发公共事件中发挥着重要作用。例如,微博打拐“彰显民众智慧理性”,而“余姚水灾、中南铁路爆炸事故、兰考火灾、厦门公交车着火、首都机场发生爆炸、房屋强拆致死”等系列公共事件通过微博迅速传播,膨胀发酵,造成社会不良情绪急剧蔓延,严重影响社会的安定。如何应对基于微博公共事件舆情的发展,实现对公共事件的有效监管,已成为一个重要课题。本文从微博公共事件语料库的构建、公共事件自动发现和演化模型研究等方面进行梳理,对研究现状进行评述并对未来研究工作进行展望。

1 微博公共事件演化分析相关研究进展

有效应对突发公共事件是当前我国政府和社会面临的重要挑战,微博在突发公共事件传播中的作用日益凸显。国内外学者针对微博突发公共事件演化分析相关研究主要从宏观思辨和微观实证两个角度展开。

1.1 公共事件发现

虚拟网络中所构成的公共事件一旦进入到爆发期,将直接引起实体社会中社会情绪的爆发,从而增加社会的不安定因素。为此,提前发现和了解可能、潜在的微博公共事件,判断普通事件转变为公共事件的特征,预警公共事件的形成,有助于防止事态扩大、社会矛盾的激化。为政府、企事业单位和微博平台实现对微博信息过滤和传播引导,进而为实现对公共事件有效可行的监控提供有效的决策支持,有助于社会安定、社会和谐和社会管理。事件发现(或称检测、识别)在信息检索领域已有较长的研究历史,常规的方法是用短时间内突发的一系列关键词来描述事件。早期大部分网络事件关联检测的研究都集中于文本描述以及特征选择^[1-2]。

微博公共事件在发展初期应该是一个热点事件,如何提前自动发现和了解潜在的微博公共事件,判断其转变为公共事件的特征,预警公共事件的形成十分重要。现有微博公共事件自动发现方法分为基于统计特征和基于文档聚类两种类型。

1.1.1 基于统计特征的方法

基于统计特征的方法是通过分析词汇的分布并将词汇分组来发现事件。在微博平台上,一个公共事件通常在短时间内引起人们的广泛关注,并出现大量评论和转发。通过监控单位时间窗口内关键词出现的频率是否剧增来发现公共事件。主要包括两种策略:

(1) 基于统计词频的方法。通过实时挖掘相关

事件的关键词、词频和内容来监控事件的出现,结合语义分析和实时特征建立概率模型发现公共事件为较早采用的方法^[3]。利用短文本统计词频,把经过预处理的微博按等数量窗口划分^[4],构造短文本词频和增长速度的复合权值^[5],进而实现基于速度增长的微博热点话题发现方法。基于速度增长方式,以量化词语及上下文相关度模型进行增量式聚类算法,能更好地实现从微博中挖掘新闻话题。

(2) 基于加权改进的 TF-IDF 方法。基于统计单独词语在文本集及语料中的重要程度,一种命名实体加权改进的 TF-IDF 方法,通过调整特征权重来度量信息相似度方法以发现微博事件^[6]。文献[7-8]也从不同角度对 TF-IDF 方法进行改进来发现微博公共事件。

由于微博的时效特点和主题内容特点,许多学者在定义和发现微博公共事件时采纳了更丰富的因素,比如时间标签、用户、转发用户、关注用户、评论、内容标签、地理位置标签、引用链接等,这些因素为检测微博公共事件提供了更为丰富的信息内容,也面临有别于传统事件检测研究的新挑战。

1.1.2 基于文档聚类的方法

文档聚类方法通过基于语义相似性来发现公共事件,这类研究方法从多个视角入手。

(1) 语料库嵌入方法。通过具体的结合命名实体的话题模型描述、以时间为参数的权重与阈值估计方法融合嵌入新闻语料实现话题的识别与追踪^[9]。贾自艳等^[10]较早将命名实体融入 TDT 系统的研究,这些方法能够在一定程度上提高 TDT 系统性能,但并没有形成独立于话题检测与跟踪领域特有的研究框架与模型。在基于中心化的发现方法中通过开放 API 记录的结构化元数据信息,设计微博的元数据模型,将微博热点发现看作是原始语料到热点语料簇的生产加工增值过程,进一步发现微博热点话题^[11]。

(2) 基于动态滑动窗口机制的方法^[12]。微博突发话题检测方法中动态滑动窗口机制将处理问题的粒度定格在短文本上,利用词语语义相似性对 Single-Pass 算法加以改进,也成为基于文档聚类的有效方法之一。

(3) 基于聚类算法与模型的方法。一种基于“绝对聚类”^[13]的微博突发词文本聚类算法(ACFD)通过分析微博自身内容的简短性和突发事件在微博中的传播特性,使用突发词构建文本的特征向量,能够比较准确地实现微博突发事件的检测。通过消除孤

立点,采用 CURE 算法对有聚类价值的数据进行聚类,提高微博热点话题发现的准确性^[14].一种基于情感分布语言模型 ELM,在微博用户情感产生波动时所发微博中情感词的数量增多的情况下识别热点事件^[15].

1.2 公共事件演化

微博公共事件是动态变化的,具有生成演化的特性,对事件的演化过程通常采用以时间为维度的方法,将微博事件的演化分为潜伏、扩散、成熟和消退4个阶段,涵盖了事件的曝光、舆论合力、政府介入和问题解决几个影响事件发展的阶段^[16].在潜伏期,微博的作用主要体现在危机预警方面;在扩散期,主要表现为信息发布、现场直播和组织全民参与方面;在成熟期,主要起到对事件连续性报道和动态追踪的作用;在消退期,提供情感交流平台和舆论监督的作用.微博舆论生成演变过程会因焦点事件的性质差异和事态演变而各有不同.舆论事件能否进入大众视野,成为公共事件,与这些事件和社会公众的普遍社会心态是否契合息息相关.

国外在公共事件演化研究方面, Twitter 已被用于许多事件的实时提醒,比如大规模火灾、网络内容服务的故障以及交通状况实时更新等.研究中多采用建立模型的方式,比如将追踪事件的演化模型分为两个任务,即对事件受欢迎程度的追踪,以及对事件主题的追踪;这两项任务同时考虑了文档的演化以及 Twitter 所代表的社区结构的演化^[17].通过对 Twitter 关于某些主题的发布频率建立模型,文献[18]成功地预测了电影票房收入.用 Twitter 数据建立了外源性输入的自动回归模型,可预测流感病情的变化趋势^[19].

国内在公共事件演化研究方面主要有以下两类方法:

(1) 基于动态分析方法.把突发公共事件看作一个复杂自组织系统,运用自组织理论和演化博弈理论来对突发公共事件演化过程的实质进行动态研究^[20-21];在突发公共事件演化分析的基础上运用案例推理理论和情景分析方法提出突发公共事件应对预案构建体系.在后期的演化影响指标中针对突发性群体事件演化系统的高阶次、多重反馈回路、高度非线性的复杂性,提出了影响突发性群体事件演化的主要因素为静态结构和动态流程的共性特征.以“三亚宰客门”事件为例,应用共振等社会物理学原理^[22],剖析微博传播事件形成的作用机理,描绘微博传播作用下舆情事件的“发生→发展→消解”演

进模式;通过定量统计与微博传播影响力分析,剖析该事件舆情演进的动因.

(2) 基于舆情演变过程的特征分析方法.从舆情的五要素及舆情发展过程的角度^[23],以微博网络结构特性以及舆情话题在微博网络中传播特性,基于传染病理理论,分析具有免疫速率的舆情话题的传播模型^[24].通过定性与定量相结合、要素分析与过程分析相结合,描述影响网络舆情演变的内在机制来发现舆情演变规律,并分析信源特征^[25]对微博舆情传播效果的影响.在比较分析网络数据的基础上,基于用户社会属性和行为特征吸引度的微博粉丝网络演化模型^[26],及混合模型的平均场方程^[27],从偏移和拉伸后的幂律分布对类社交网络的形成和演化特征进行描述分析粉丝网络演化规律.利用新浪微博分析权威信息的发布对微博内容发展趋势的影响,发现随着权威信息的发布,行动类和疑问类微博的比例减少,观点类和情绪类微博趋于平稳^[28];这说明政府在处理突发公共事件时,及时发布权威信息并建立畅通的信息沟通渠道,有利于正确控制和引导舆论.

1.3 公共事件语料库构建

公共事件语料库是公共事件发现和演化分析的基础.公共事件语料库动态记录已有公共事件从发起到消亡的过程,将被用于辅助发现和识别潜在公共事件;依据语料库中公共事件数据在不同发展等级的属性特征,判断微博事件转变为公共事件的形成契机和数据特征,利用采集到的微博议题以及相关信息和语言特征数据,制定关注度指标、传播指标和议题特征指标,确定议题的敏感度、影响程度、类别、性质、转化特征等,借助语料库定性和量化分析哪类事件将容易转化为公共事件、哪些语言和事件处置方式将容易引发公共事件等,有助于及时预警公共事件的形成.

由于突发事件语料库的构建中语料信息均来自于新闻网页,充分利用新闻网页结构与其内容分布特征,通过分析新闻特定格式与用词特点出发提出了基于网页分割的噪声排除与主题内容提取方法;根据新闻事件的易碎性,提出按发布日期分群,并在此基础上进行网页去重;通过改进的 TF-IDF 权重计算方法对突发事件新闻网页重复的特点和实例用词进行分析.在语料库构建过程中包括几个阶段:大规模新闻网页语料收集获取(原始语料)、语料的人工分类、网页去重、文本格式处理、语料编码以及分词与词性标注等加工过程^[29-30].低成本、短周期构建

大规模语料库是目前研究工作的难点. 超大规模分类语料库构建中分别采用网页内容抽取语料并按照信息字段和网页分类模块确定所需语料字段进行XML信息标注和构建语料库^[31], 以及利用网格技术的大规模计算能力与Wiki的开放编辑环境去收集和整理语料^[32], 根据可信度模型挑选出不可信的语料并由人工进行校对, 计算校对后结果的可信度, 选择出最可信的结果作为正确语料存储到语料库中.

2 研究展望

目前对于微博突发公共事件, 如何根据公共事件的不同类型以及公共事件的动态发展状态, 进行行之有效的舆情引导和监管, 还存在很多挑战. 需要结合新闻传播理论、信息检索技术和自然语言处理等技术, 研究微博公共事件的采集、标注和存储方法, 并在此基础上构建微博公共事件语料库, 进而研究潜在在微博公共事件的自动发现方法及微博公共事件演化机制等, 以便及早发现潜在的微博突发公共事件, 进而根据事件的不同类型和演化阶段, 确定公共事件的引导策略.

2.1 微博公共事件语料库构建

通过对文献的梳理, 可以看出微博公共事件的自动发现, 对公共事件信息进行挖掘, 预测发展趋势并评估危机态势已成为国内外学者的研究热点. 基于微博公共事件发现有着不同于传统事件的特殊难点和挑战. 对于微博公共事件发现机制, 目前尚缺乏大规模的公共事件语料库支撑, 现有微博公共事件自动识别方法没有充分考虑到事件的动态传播特征; 公共事件语料库数据不仅要包含博主或用户的各种信息、微博内容、微博元数据、用户关系、热门话题等常见数据属性和事件元数据, 还将具有公共事件的类别属性, 用于验证潜在公共事件的识别模型和演化模型. 针对微博公共事件的语义和内容特点, 在制定规则和标准的基础上, 利用信息提取技术, 研究从新浪微博、腾讯微博等自媒体网站中自动采集和标注微博信息的方法, 螺旋式地扩展语料库. 公共事件语料库构建需要包括两个方面: 其一是针对微博中已有的典型公共事件, 采集其微博语料数据. 针对典型公共事件数据, 研究其数据特征、类型特征、级别演化特征, 及其转化为公共事件的标志性特点, 为识别公共事件提供高质量的标注机制. 其二是结合机器学习方法和人工标注, 研究半监督的自动标注方法, 即通过少量人工标注语料, 使用机器学习方法从已标注语料中学习其规律, 利用这些规律实现

对未标注的语料进行自动标注, 从而螺旋式地快速扩展语料库, 以高效获得大规模高质量的公共事件语料.

2.2 潜在微博公共事件发现

现有的事件发现方法在微博领域应用还存在一些不足, 如何针对微博内容的碎片特性和快速传播特性, 及早发现潜在的微博公共事件是一个挑战.

2.3 微博公共事件演化分析

已有突发公共事件的演化模型研究中, 主要从事事件的发生过程、已有公共事件的成因、特征着手分析, 针对事件发生的内在演化机理以及动态分析方法进行研究较少, 在现有演化模型中也没有考虑信息在微博传播中的方向性和裂变性^[27]以及影响微博用户传播行为等因素. 因此针对不同事件本身特性结合演化的动力学特征找到适应中国国情发展的公共事件演化传播特点势在必行.

微博平台上的事件信息在传播过程中会受到多种因素的影响, 而粉丝的转发或评论是促进事件传播的主要因素. 事件信息的转发或评论者的数量越大、相邻两次转发或评论的时间间隔越小, 则该事件越有可能成为公共事件. 如果把相邻两次转发的时间间隔和转发者粉丝数量及转发者粉丝评论数量的统计分布作为随机状态, 由于相邻两次转发的时间间隔是隐藏的, 则微博事件的传播、演化过程可以看成隐马尔科夫过程. 利用公共事件语料库和隐马尔科夫模型, 针对事件在不同发展等级阶段(即危险级、严重级和警示级)的特征, 建立从微博内容和传播角度识别事件在性质、严重程度、可控性和影响范围等方面的衡量指标, 自适应、动态的学习方法, 用于微博公共事件发展态势度量, 判断议题是否会对社会造成更大的冲击, 即转为爆发型公共事件或消沉型公共事件; 研究判断事件发展等级的标准、属性特点和智能判断模型, 确定事件演化等级的自动判断方式, 及时预测公共事件的发展趋势, 为公共事件应急引导奠定基础.

2.4 微博公共事件演化在舆情应急管理中的应用

微博日益成为舆论生成的策源地、传播的集散地与交锋的重要阵地, 目前一些地方政府、企事业单位对微博平台的使用及微博公共事件的认识还不到位, 不适应新媒体形势的变化, 导致应对公共事件的策略上还存在诸多问题. 目前, 突发公共事件中网络舆论引导方面存在两个方面的突出问题, 其一是政府与民众间的互动渠道不畅, 其二是被动防守、信息公布不及时. 为此, 需要完善舆情预警机制, 研究微

博公共事件的主要特征,及时发现潜在的公共事件,掌握其演变规律,提高舆情处置能力.依据微博公共事件演化分析的监测结果,以及事件在不同发展等级阶段的趋势变化分析,探讨应急管理操作机制,最大化地将公共事件的负面影响控制在警戒线以下.

通过科学的应急管理策略针对公共事件的不同情况建立相应的分类引导机制,采取一系列必要的措施,以引导公共事件向积极向上的方向发展,从而实现对公共事件进行有效可行的监管和控制.通过建立公共事件演化模型监测事件的发展动态,实现对微博公共事件的预警;预测特定议题对公共事件的影响力,以便针对公共事件的不同类型和公共事件的不同演化状态,制定不同的引导策略,为政府相关部门或网络运营商制定针对不同级别(警示级,严重级,危险级)公共事件的应急处理策略提供参考.

参考文献:

- [1] Allan J, Lavrenko V, Swan R.. Explorations within topic tracking and detection[M] // Topic Detection and Tracing: Event-based Information Organization. Kluwer Academic; Massachusetts, 2002,197-224.
- [2] Petrović S, Osborne M, Lavrenko V. Streaming first story detection with application to Twitter[C] // In Proceedings of the 11th Annual Conference of the North American Chapter of the Association for Computational Linguistics. [s. n.], 2010;181-189.
- [3] Sakaki T, Okazaki M, Matsuo Y. Earthquake shakes Twitter users: Real-time event detection by social sensors[C] // In Proceedings of the 19th International World Wide Web Conference. New York:ACM,2010,851-860.
- [4] 薛素芝,鲁燃,任圆圆. 基于速度增长的微博热点话题发现[J]. 计算机应用研究,2013,30(9): 2598-2601.
Xue S Z, Lu R, Ren Y Y. Hot topics found on micro-blog based on speed growth[J]. Application Research of Computers,2013,30(9): 2598-2601.
- [5] 郑斐然,苗夺谦,张志飞,等. 一种中文微博新闻话题检测的方法[J]. 计算机科学,2012,39(1): 138-141.
Zheng F R, Miao D Q, Zhang Z F, et al. News topic detection approach on chinese microblog[J]. Computer Science. 2012,39(1): 138-141.
- [6] Phuvipadawat S, Murata T. Breaking news detection and tracking in Twitter[C] // Web Intelligence and Intelligent Agent Technology In Proceedings of IEEE/WIC/ACM International Conference. Toronto, Canada:IEEE, 2010: 120-123.
- [7] 童薇,陈威,孟小峰. EDM: 高效的微博事件检测算法[J]. 计算机科学与探索,2012, 6(12):1076-1086.
Tong W, Chen W, Meng X F. EDM: an efficient algorithm for event detection in microblogs[J]. Journal of Frontiers of Computer Science and Technology, 2012, 6(12): 1076-1086.
- [8] 冯永,韩楠,贾东风. 云计算环境下基于代表点增量层次密度聚类的微博事件检测及跟踪[J]. 计算机应用, 2013,33(12):3559-3562.
Feng Y, Han N, Jia D F. Microblog events detection and tracking with incremental hierarchical DBSCAN based on representative posts using cloud framework[J]. Journal of Computer Applications, 2013,33(12):3559-3562.
- [9] 洪宇,张宇,刘挺,等. 话题检测与跟踪的评测及研究综述[J]. 中文信息学报,2007,11(6): 71-87.
Hong Y, Zhang Y, Liu T, et al. Topic detection and tracking review [J]. Journal of Chinese Processing, 2007, 11(6): 71-87.
- [10] 贾自艳,何清,张海俊,等. 一种基于动态进化模型的事件探测和追踪算法[J]. 计算机研究与发展, 2004,41(7):1273-1280.
Jia Z Y, He Q, Zhang H J, et al. A news event detection and tracking algorithm based on dynamic evolution model [J]. Journal of Computer Research and Development, 2004,41(7):1273-1280.
- [11] 蔡淑琴,张静,王旻,等. 基于中心化的微博热点发现方法[J]. 管理学报, 2012,9(6):874-879.
Cai S Q, Zhang J, Wang Y, et al. Micro-blogging hotspot discovery method based on centralization [J]. Chinese Journal of Management, 2012,9(6):874-879.
- [12] 邱云飞,程亮. 微博突发话题检测方法研究[J]. 计算机工程, 2012,38(9): 288-290.
Qiu Y F, Cheng L. Research on sudden topic detection method for microblog[J]. Computer Engineering, 2012,38(9): 288-290.
- [13] 王勇,肖诗斌,郭跬秀,等. 中文微博突发事件检测研究[J]. 现代图书情报技术, 2013,2: 57-62.
Wang Y, Xiao S B, Guo Y X, et al. Research on chinese micro-blog bursty topics detection[J]. New Technology of Library and Information Service, 2013,2: 57-62.
- [14] 赖锦辉,梁松. 一种消除孤立点的微博热点话题发现方法[J]. 计算机应用与软件,2014, 31(1): 105-107.
Lai J H, Liang S. A microblogging hot topics discovery method based on outliers elimination[J]. Computer Applications and Software,2014, 31(1): 105-107.
- [15] 杨亮,林原,林鸿飞. 基于情感分布的微博热点事件发现[J]. 中文信息学报,2012,26(1): 84-90.
Yang L, Lin Y, Lin H F. Micro-blog hot events detection based on emotion distribution[J]. Journal of Chinese Pro-

- cessing, 2012,26(1): 84-90.
- [16] 熊光清. 中国网络公共事件的演变逻辑——基于过程分析视角[J]. 社会科学, 2013,4:4-15.
- Xiong G Q. The Evolutional logic of internet public incidents in China: from the perspective of process analysis [J]. Social Science,2013,4:4-15.
- [17] Lin J, Snow R, Morgan W. Smoothing techniques for adaptive online language models: topic tracking in tweet streams[C] // In Proceedings of the 17th ACM SIGKDD international Conference on Knowledge Discovery and Data Mining. New York:ACM,2011: 422-429.
- [18] Asur S, Huberman B A. Predicting the Future With Social Media[C] // In Proceedings of the ACM International Conference on Web Intelligence. New Nork:ACM. 2010.
- [19] Achrekar H, Gandhe A, Lazarus R, et al. Predicting flu trends using Twitter data[C] // In Proceedings of Computer Communications Workshops (INFOCOM WKSHPS). Shanghai: IEEE,2011:702-707.
- [20] 罗成琳. 突发群体事件演化及其应对预案构建研究[D]. 哈尔滨: 哈尔滨工业大学管理学院,2009.
- [21] 罗成琳,李向阳. 突发性群体事件及其演化机理分析[J]. 中国软科学, 2009,6:163-171.
- Luo C L, Li X Y. Analysis on mass emergency and its evolutive mechanism[J]. China Soft Science, 2009,6:163-171.
- [22] 付宏,田丽. 基于微博传播的舆情演进案例研究[J]. 图书情报工作,2013,57(15): 34-38.
- Fu H,Tian L. The Case Study on the evolution of public opinion based on microblog-spread[J]. Library and Information Service, 2013, 57(15): 34-38.
- [23] 王平,谢耘耕. 突发公共事件网络舆情的形成及演变机制研究[J]. 现代传播,2013,3:63-69.
- [24] 李青,朱恒民,杨东超. 微博网络中舆情话题传播演化模型[J]. 现代图书情报技术,2013,12:74-80.
- Li Q, Zhu H M, Yang D C. The topic evolution model of the public opinion in micro-blogging network[J]. New Technology of Library and Information Service, 2013,12: 74-80.
- [25] 张玥,孙霄凌,浦正宁,等. 微博舆情传播影响因素研究[J]. 情报资料工作,2014,5: 59-64.
- Zhang Y, Sun X L,Pu Z N, et al. Influencing factors of microblog public opinion dissemination[J]. Information and Documentation Services,2014,5: 59-64.
- [26] 王晶,朱珂,汪斌强. 基于用户社会属性及行为特征吸引度的微博粉丝网络演化模型[J]. 计算机应用, 2013, 33(10):2753-2756.
- Wang J, Zhu K, Wang B Q. Microblog fans network evolving model based on user social characteristics and attractiveness of behavior properties [J]. Journal of Computer Application,2013,33(10):2753-2756.
- [27] 熊熙,曹伟,周欣,等. 社交网络形成和演化的特征模型研究[J]. 四川大学学报:工程科学版, 2012,44(4): 140-144.
- Xiong X, Cao W, Zhou X, et al. Research on the feature model of the formation and evolution of social networks [J]. Jouranl of Sichuan University:Engineering Science Edition, 2012,44(4):140-144.
- [28] 罗潇潇,何跃,熊涛. 突发公共事件中权威信息对微博内容的影响研究——以柳州镉污染事件为例[J]. 图书情报工作,2012,56(11): 123-127.
- Luo X X, He Y, Xiong T. The influence of authoritative information upon the content of micro-blogging in public emergency[J]. Library and Information Service, 2012, 56(11): 123-127.
- [29] 张永奎,张晴,王磊. 面向信息处理的突发事件新闻语料库建设与应用研究[J]. 山西大学学报:自然科学版, 2009,32(4):546-551.
- Zhang Y K, Zhang Q, Wang L. The research on establishment and applications of emergency news corpus for information processing [J]. Journal of Shanxi University: Natural Science Edition, 2009,32(4):546-551.
- [30] 杨丽英,李红娟,张永奎. 突发事件新闻语料分类体系研究[C] // 中文信息处理前沿进展——中国中文信息学会 25 周年学术会议论文集. 北京:中国中文信息学会,2006.
- [31] 刘华. 超大规模分类语料库构建[J]. 现代图书情报技术, 2006,1:70,71-73.
- Liu H. Construction of a super classed and denoted corpus [J]. New Technology of Library and Information Service, 2006,1:70,71-73.
- [32] 李培峰,朱巧明,钱培德. 基于 Web 的大规模语料库构建方法[J]. 计算机工程, 2008,34(7):41-43,46.
- Li P F, Zhu Q M, Qian P D. Construction approach of large-scale corpus based on web[J]. Computer Engineering. 2008,34(7):41-43,46.