

文章编号：1000-5641(2014)00—

# 通过信息技术提高高校网络舆情引导能力

王海涛，张福蒸，王义，梅涛

(对外经济贸易大学，北京 100029)

**摘要：**本文将传统舆情管理工作模式与现代信息技术紧密结合，应用信息化技术加强高校网络舆情引导，为开辟高校网络舆情管理工作新领域，维护校园和谐稳定奠定基础。

**关键词：**高校；网络舆情；网络舆情分析模型；舆情预警

中图分类号：G60 文献标识码：A

DOI：10.3969/j.issn.1000-5641.2014.00.0000

## Through information technology to improve the college network public opinion guiding ability

Haitao Wang, Fuzheng Zhang, Yi Wang, Tao Mei

(University of International Business and Economics, Beijing 100029)

**Abstract:** The paper will combine the traditional mode of management of public opinion work and the modern information technology closely, and apply information technology to enhance the network public opinion guidance, and lay the foundation to open up a new field of network public opinion management work in colleges and universities and maintain the stability of the harmonious campus.

**Key words:** Colleges and universities; network public opinion; analysis model of network public opinion; public opinion warning

### 0 引言

随着互联网的飞速发展，校园 BBS、贴吧、QQ 以及微博、微信等网络工具出现，高校大学生的信息来源呈现多元化，网络成为大学生表达观点的重要媒介。中国互联网络信息中心(CNNIC)第 34 次《中国互联网络发展状况统计报告》显示，截至 2014 年 6 月我国网民规模达到 6.32 亿，20 到 29 岁的青年是网民中规模最大的群体，占 30.7%；学生是网民中职业规模最大的群体，占 25.1%。<sup>[1]</sup>高校大学生是网络舆情的主要生成力量和影响对象，

当校园舆情聚集时就可以向舆论转化，作为社会舆情的激发者，直接影响着社会舆情的走向。

## 1 如何应用信息化技术加强高校网络舆情引导

我国高校网络舆情内容多元化、舆情主体集中化、舆情形成快速化、传播渠道多样化，对高校网络舆情引导工作增加了难度。采用传统的人工手段进行高校网络舆情引导工作存在一定程度的盲目性，决策拍脑袋，已无法满足网络和大数据时代的发展。迫切需要结合信息技术对舆情信息进行监测、分析和预警，为舆情引导工作提供管理依据和决策支持。

### 1.1 高校网络舆情分析模型的提出

应用信息化技术为高校网络舆情引导提供支持，使舆情引导变被动为主动。发挥和利用大数据技术对存储海量数据、处理非结构化数据的优势，应用高校网络舆情分析模型对舆情信息进行采集、处理、分析、预警和发布。

高校网络舆情分析模型<sup>[2]</sup>包括舆情信息采集、舆情信息处理、舆情数据分析、舆情预警和舆情发布几个部分组成,如图1。该模型能够有效分析高校网络舆情发生、发展和变化的规律,对舆情的未来趋势进行科学预测。

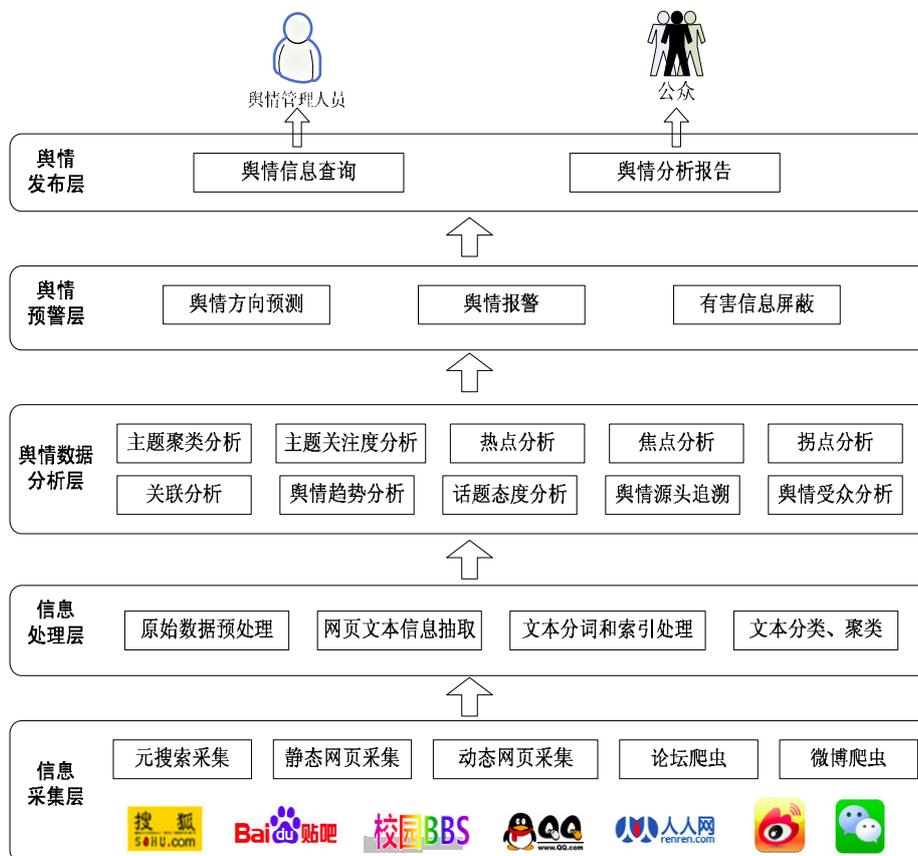


图1：高校网络舆情分析模型架构图

## 1.2 高校网络舆情分析模型工作原理

舆情信息采集层是高校网络舆情分析的基础和前提。通过对各类舆情渠道产生的舆情信息进行自动采集,统一的存储到网络存储系统。然后对舆情原始数据进行预处理、分类加工、整理,形成舆情信息库。通过对舆情信息的深入剖析进行舆情数据分析和预测预警,最后将舆情分析结果推送和展示给各用户。

### 2.2.1 舆情信息采集

高校网络舆情主要采集以下几类信息: 主流媒体网站类,包括各大新闻网站、百度搜索、贴吧等主流的信息渠道。 校内网站类,包括校园 BBS 及学生经常使用的各大平台系统。 交友互动类,包括人人网、开心网等交友网站,以及 QQ、微信等即时通讯工具,以及充分展现个人观点的微博、博客等信息。这些信息全覆盖文本、图片、音频和视频的各类综合数据。从中提取有效的舆情信息保存到舆情信息库中,并针对不同渠道的舆情信息制定相应的数据同步机制,进行舆情信息同步更新,保证舆情信息的准确性。

### 2.2.2 舆情信息处理

在舆情信息采集后,需要对收集到的原始数据进行基本信息预处理,如关键词提取、建立全文索引、自动去除重复信息、格式转换、数据清理、分区存储等。对于新闻评论,需要过滤掉无关信息,将新闻的标题、出处、发布时间、内容、点击次数、评论人、评论内容、评论数量等进行保存。对于论坛微博等,需要记录帖子的标题、发言人、发布时间、内容、回帖内容、回帖数量等信息。再通过文本分类、聚类技术进行舆情信息处理,为舆情主题分析提供基础。

### 2.2.3 舆情数据分析

舆情数据分析层对网络舆情信息进行分析 and 研判,从而为网络舆情引导奠定基础。根据不同的维度从不同的角度开展舆情分析,并通过统计图表实现对舆情分析结果的可视化展示。

#### (1) 主题聚类分析

将经过整理的舆情信息按照其内在的关联性自动划分成为不同的类别,从而发现分布于网络上各处与某一个主题相关的所有的新闻、帖子、博客以及相应的跟帖。对于聚类分析划分出来的主题,经过统计就可以发现目前网络上的舆情及其构成。

#### (2) 主题关注度分析

分析在过去某一段时间内舆情主题被关注的程度,用该主题的相关帖子回复数或与该主题的相关网页数进行衡量。通过设定舆情主题、起始时间、终止时间以及统计周期,检索得到舆情主题在某个时间段内的关注度。

以对外经济贸易大学小天鹅论坛为例,从总共 180 个版块中选取比较活跃的版块进行信息采集,将采集对象限定于发帖时间在 2014 年 2 月 24 日到 2014 年 6 月 30 日(一个完整学期)之间的帖子。经过数据分析总结出 Top10 话题排行榜,如下表。其中 5 个为情感话题,2 个为课业交流话题,1 个为就业创业话题,1 个为媒体热议话题,1 个为其他话题。

表：对外经济贸易大学小天鹅论坛 Top10 话题排行榜

序号	主题	回帖量	发帖时间	最后回复时间	版块	所属话题类别
1	我为舍友找对象--之西门盖燧征友!	254	2014-03-29	2014-04-20	缘来是你	情感话题
2	八一八咱本科毕业生就业起薪? 开贴需勇气	126	2014-05-22	2014-06-17	惠园特快	就业创业话题
3	还未跟你饮过冰,零度天气看风景。祝大家生活愉快!	122	2014-03-17	2014-03-23	惠园特快	其他
4	在(部分)贸大女生眼里,(部分)贸大男生存在哪些不够成熟的	118	2014-05-28	2014-06-03	惠园特快	情感话题
5	是不是男生都不喜欢弱弱的女生	107	2014-06-15	2014-06-18	水晶之恋	情感话题
6	我想知道哪里不对?	105	2014-04-02	2014-06-18	水晶之恋	情感话题
7	【我就不信我们贸大没有好老师】想去蹭课,求推荐啊	76	2014-04-20	2014-05-16	课业信息交流	课业交流话题
8	重修狗 跪求微积分2期末题库或过往题目	76	2014-05-23	2014-06-16	课业信息交流	课业交流话题
9	谈恋爱要不要考虑门当户对啊	74	2014-06-11	2014-06-16	水晶之恋	情感话题
10	【可怜母子】刘强东有婚史,儿子已8岁。	71	2014-04-15	2014-04-18	惠园特快	媒体热议话题

### (3) 主题热度分析

分析在某一时间段内相对更加被关注或集中关注的舆情主题,用该主题的关注度进行衡量。设定热度阈值,计算所有舆情主题的关注度,比较其与热度阈值的大小,将关注度大于阈值的主题设为热点主题。在某一时间段内,对所有设为热点的主题按关注度的降序进行排列,生成该时间段内的热点主题排行榜。排在榜首的主题关注度最高,也就是热点问题所在。

### (4) 焦点分析

分析过去某一时间段内被各类网络舆情渠道产生频次较高的舆情信息。设定焦度阈值,计算各条舆情信息的焦度,分别比较其与焦度阈值的大小,将焦度大于阈值的舆情信息设为焦点信息。在某一时间段内,对各条舆情信息的焦度进行降序排列,生成该时间段内的舆情焦点排行榜。

### (5) 拐点分析

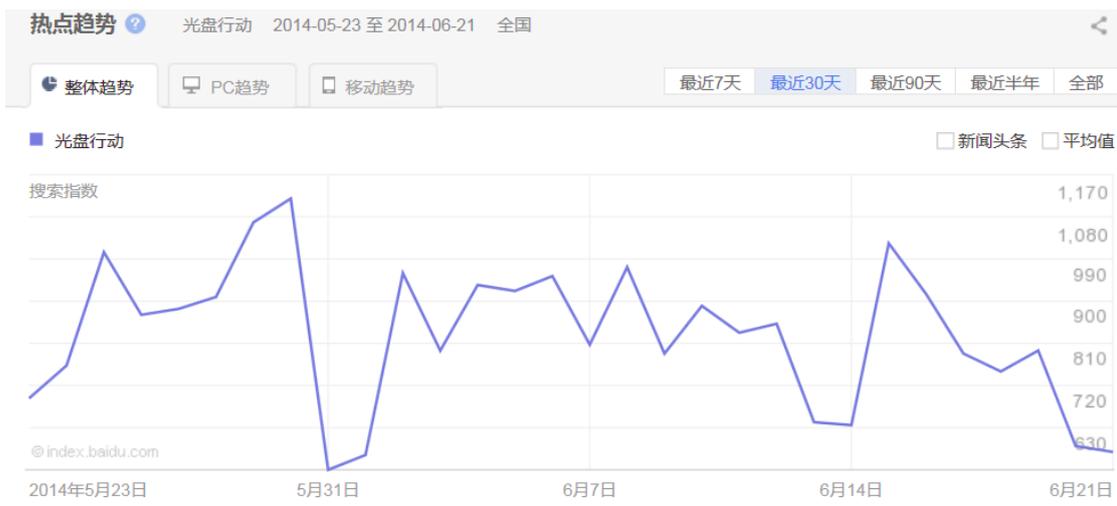
分析过去某一统计周期内关注度剧烈变化的舆情主题。计算各个舆情主题在不同统计周期内的拐度,从而得到拐度矩阵。将拐度矩阵的列数据按降序排列,得到所有舆情主题在某一统计周期内的拐度排行榜,排在最前面的舆情主题即为拐点主题。将拐度矩阵的行数据按降序排列,排在最前面的统计周期即为该舆情主题的拐点统计周期。<sup>[3]</sup>

### (6) 关联分析

通过数据对比分析,寻找舆情主体与舆情信息之间、不同时间段的舆情信息之间的关联关系。例如,从上述对外经济贸易大学小天鹅论坛的分析数据中,发现某个时间段对某类主题的关注度明显增加,与该时期的特定事件或热点事件有着密不可分的关系。排名Top10的话题中50%为情感话题,可以充分体现对外经济贸易大学文科院校、女生人数多、较多注重感性认识等特点。课业交流话题集中在5月-6月,说明临近期末考试大家会积极进行学习讨论和课业交流。就业创业话题也在5月-6月的毕业季,同时实习信息也明显增多。

### (7) 趋势分析

分析一定时间段的舆情信息,得出舆情变化和发展的趋势。把时间作为横坐标,把发帖量、回帖量等数值作为纵坐标,通过折线统计图展示某时间范围内的舆情发展趋势。以“光盘行动”热点事件为例,分析2014年5月23日至2014年6月21日一个月内的舆情变化和发展趋势,如图2。



数据来源：百度指数

图2：“光盘行动”热点走势图

上图“光盘行动”热点趋势5月30日为该月最热,为响应中宣部、国家发改委开展节俭节约活动的要求,北京多家高校发起“光盘行动”推出“半价菜”,食堂路上贴海报,食堂里LED显示屏宣传“光盘行动”,减少高校餐桌的浪费现象。第二个高点为6月16日,为倡导“光盘行动”,南昌师范学院开展了“文明校园 从我做起”的快闪活动。至此各地积极深化“光盘行动”,倡导“舌尖上的文明”。

### (8) 话题态度分析

从舆论的态度上,可以根据主帖和回复帖的赞成或反对态度,划分为正面、中性、负面的态度,经过加权平均后得出大学生群体对某一话题的态度。

### (9) 舆情源头追溯

根据学生在网络上发布和传播舆情信息时留下的踪迹,如页面相互之间的链接,帖子上注明的转帖地点,发帖时间等,整理出各个帖子之间的链接结构和发布的先后顺序,发现原始帖,追溯到舆情源头。

#### (10) 舆情受众分析

对每个学生在登录或访问校园网时在服务器上留下的痕迹,包括进入每一个网站、板块的时间,访问的主题,是否发言跟帖等,进行统计分析得出每个学生的访问偏好、意见倾向性以及以往在网上的表现。

### 2.2.4 舆情预警

在对舆情信息进行数据挖掘、分析和研判的基础上,通过发展趋势和数据关联分析结果对网络舆情进行预测预警,对负面突发敏感舆情的及时预警,使网络舆情引导由被动变为主动。

#### (1) 舆情方向预测

将历史数据按周、月、年进行对比,获得趋势预测分析走势,可以预测大学生舆情方向以及对重大热点事件可以进行及时有效的舆情导向,为舆情引导工作提供参考。

#### (2) 舆情报警

根据预先设定的一组关键词定时搜索本地索引和舆情数据库。若出现关键词,以短信或邮件的形式自动发送给舆情管理人员。高校学生重点关注的热点事件在出现关注度增长时以舆情报警的形式发送给舆情管理人员。

#### (3) 有害信息屏蔽

通过对网络有害信息过滤,实现对负面舆情的有效控制。根据需要预先设定过滤的指标,从动态的舆情信息中剔除恶意信息、负面信息。过滤机制包括基于IP地址的过滤、基于关键词的过滤和基于语义的信息过滤。

### 2.2.5 舆情发布

舆情发布层面向舆情管理人员及社会公众。可通过短信、邮件等形式将舆情分析和预测结果定制推送给舆情管理人员和领导。舆情管理人员可通过关键词进行舆情信息检索,也可自行配置搜索条件,如设置时间范围、地点、类型、网站、在页面中出现的位置等进行舆情综合查询。

还可将舆情分析结果生成舆情分析报告,公布给社会公众。舆情分析报告内容可包括分类舆情信息、负面舆情、舆情事件、舆情简报等,通过统计图表直观的展现高校网络舆情现状和发展情况。

## 1.3 关键技术在校网络舆情分析模型中的应用

在校网络舆情分析模型各层中,应用关键技术为舆情分析提供技术支持。舆情信息采集使用网络爬虫技术,按照设定种子URL集合自动采集相关舆情信息。还可设定主题目标进行信息自动采集。对于静态数据可以采用路径分析技术;对于动态信息,如用户访问日志等信息可采用关联规则挖掘技术。

信息处理层对舆情信息采集层获取的网页文本信息进行抽取, 然后对抽取的文本信息进行文本分词和索引处理。采用文本分类技术对存入索引的文本进行过滤, 以去除和检测目标不符的信息, 最后采用文本聚类技术对相似度尽可能大的文本集合进行聚类, 为舆情主题的分析做好准备。

舆情数据分析和预警过程应用数据挖掘、敏感识别、热点分析等技术, 对网络舆情信息进行分析 and 展示。<sup>[4]</sup>

### 3 总结

加强高校网络舆情引导机制的研究, 通过高校网络舆情分析模型, 应用信息化技术加强高校网络舆情引导, 为大学生营造积极向上的网络舆情氛围和学习环境。随着数据挖掘技术等大数据应用越来越多的融入到高校各管理领域, 以及高校网络舆情管理的深入需要, 今后高校还应在网络舆情预警分析的研究和应用方面进行进一步的探索。(对外经济贸易大学 2014 年党建研究课题资助项目)

#### [参 考 文 献]

- [1] 中国互联网信息中心. 第 34 次中国互联网络发展状况统计报告[R]. 中国互联网信息中心, 2014-7.
- [2] 尹培培. 大数据时代的网络舆情分析系统[J]. 广播与电视技术, 2013(07).
- [3] 钱爱兵. 基于主题的网络舆情分析模型及其实现[J]. 情报分析与研究, 2008(04).
- [4] 吴晓倩, 陈诚. 高校网络舆情监测关键技术研究[J]. 电脑知识与技术, 2013, 9(16).