

doi:10.3772/j.issn.1000-0135.2012.12.010

## 微博舆情传播规律研究<sup>1)</sup>

钱颖<sup>1</sup> 张楠<sup>1</sup> 赵来军<sup>1</sup> 钟永光<sup>2</sup>

(1. 上海大学管理学院, 上海 200444; 2. 青岛大学管理科学与工程系, 青岛 266071)

**摘要** 随着用网络户数的增加, 微博已成为网民表达意愿的重要渠道, 成为当前最具影响力的传播工具之一。由突发事件导致的舆情在微博上传播迅速, 影响广泛。基于传统传染病模型(SIR模型), 本文对微博上舆情传播规律进行研究, 建立了微博舆情传播模型, 并用2011年甘肃校车事故在新浪微博上的传播进行实证研究。研究发现: 原创, 转发帖子爆发迅速, 第二天达到峰值, 衰退的速度较慢, 六天后帖子数稳定在较低水平。用微博传播模型对甘肃校车事件中舆情传播规律进行仿真, 仿真结果与历史数据拟合度高, 验证了此模型用来研究微博舆情传播规律的可行性。最后利用此模型, 研究了不同情景下的微博舆情传播规律。

**关键词** 新浪微博 突发事件 SIR模型 微博舆情模型

## The Spread of Public Sentiment on Micro-blogging under Emergencies

Qian Ying<sup>1</sup>, Zhang Nan<sup>1</sup>, Zhao Laijun<sup>1</sup> and Zhong Yongguang<sup>2</sup>

(1. Management School, Shanghai University, Shanghai 200444;

2. Management Science and Engineering department, Qingdao University, Qingdao 266071)

**Abstract** With the increasing number of users, micro-blogging has become the most influential communication channel. The public opinion generated by emergency spreads rapidly on the micro-blogging and has great impact. Based on the traditional rumor spreading model (SIR model), we studied the pattern of public opinion dissemination on the micro-blogging. The school bus accident happened in Gansu province in Nov 2011 is used as an empirical example. Data shows that the number of original posts and forwarded posts increased rapidly at the beginning, reaching peak on the second day, while reduced slowly, reaching a low stable level six days later. The model simulation result fits the historical data well, which is an evidence that the model could be used to study the public opinion spreading on micro-blogging. Based on the model simulation, we analyzed how public opinion spread under different scenarios.

**Keywords** Sina micro-blogging, emergency, SIR model, public sentiment spreading model on micro-blogging

## 1 引言

微博自2009年8月上线以来得到全面推广与应用。截止到2011年12月31日, 我国微博用户数

量已达到2.5亿。由于微博信息产生和获取具有很强的自主性, 内容短小精悍, 信息共享便捷迅速等特点, 微博已经成为网民们表达意愿、分享心情的重要渠道。人民网舆情分析专家预测, 在未来的突发事件中, 微博将成为最具影响力的传播工具。突发事

收稿日期: 2012年7月18日

作者简介: 钱颖, 女, 1976年生, 博士, 讲师, 主要研究方向: 信息资源管理、新媒体传播。张楠, 女, 1988年生, 硕士研究生, 主要研究方向: 网络舆情传播、新媒体传播。E-mail: zhangnan339@126.com。赵来军, 男, 1970年生, 博士, 教授, 高级工程师, 主要研究方向: 物流与供应链管理、安全管理、环境管理。钟永光, 男, 1972年生, 博士, 教授, 博士生导师, 主要研究方向: 系统动力学与运营管理。

1) 基金项目: 国家自然科学基金资助(90924030), 上海浦江人才项目资助(11PJJC075), 上海市教育委员会“曙光计划”项目(09SG38), 教育部人文社会科学研究一般项目(项目编号: 10YJA630021)。

件发生时,由于其突发性、不确定性,信息有限性,通常会引起相关舆情在网络上散布。错误的、片面的、过激的舆情会引起人们的恐慌行为,破坏社会稳定,给广大人民群众带来物质和精神损失。2011年日本“3.11”地震导致的“抢盐”事件,是一个典型的突发事件后由舆情引起的群体过激行为,扰乱了国家经济秩序。微博在此次事件中起了推波助澜作用。因此,了解微博舆论传播的规律,从而探索引导和控制微博舆情发展的政策和策略意义重大。

## 2 国内外研究现状分析

突发事件是指突然发生,造成或者可能造成重大人员伤亡、财产损失、生态环境破坏和严重社会危害,危及公共安全的紧急事件<sup>[1]</sup>。随着网络的普及,突发事件在网络上的传播也引起了许多学者的关注。张子利等提出互联网上突发事件的舆论传播具有伴随着舆情不断扩散、伴随内容不断演化以及舆论传播的发散性与多元性的特征<sup>[2]</sup>。突发事件的互联网传播动力学往往遵循“跳跃——衰退”模式<sup>[3]</sup>。近几年潘芳等对网络上舆情的传播特点以及网络舆情的传播主体进行研究,利用类似传染病传播的 SIS 模型建立危机事件网络舆情传播模型,将网民数目规模划分成小规模、中等规模以及大规模消极思潮人群,并根据划分采取相应的应急对策:针对小规模消极思潮,政府应将主要注意力集中在已有消极思潮的网民身上;对于中等规模,政府可通过电子政务将官方正确信息发布给网民并对其教育,使网民由消极态转化为非消极态;大规模消极思潮,政府有关部门可以通过网络媒体和传统媒体将正确的信息发布给网民<sup>[4]</sup>。但到目前为止,国内关于微博信息传播的研究大多都是基于定性的分析研究,定量的研究还很少。

Twitter 是国外非常流行的微博网站,国外的一些学者对 Twitter 的舆情传播也做了分析研究。Bernardo A. Huberman 等收集了 Twitter 上 309 740 个用户的数据,并对其分析,发现如果将两两有联系的人用连线连接起来,发现网络上的人与人之间的连线比现实生活中人与人之间的连线要多<sup>[5]</sup>。Amanda Lee Hughes 等利用 API (Application Programming Interface, 应用程序编程接口)对同一时间段美国发生的两个突发事件(Gustav 飓风和 Ike 飓风)和两个国家安全事件(民主党代表大会和共和党代表大会)在 Twitter 上的数据进行了收集并对

收集的数据进行分析,得出 Twitter 在突发事件发生后起了信息传播作用<sup>[6]</sup>。但是这些研究都是基于对微博上的数据进行数值分析研究,对微博舆情传播规律的模型研究却很少。

1927 年 Kermack 和 McKendrick 提出了传染病模型,即 SIR 模型,将人群分为以下三类:未染病的人,已染病的人和已经痊愈并对这种疾病具有免疫力的人。传染病传播的过程:未感染者与已染病者接触,并以一定的概率变成已染病者,已感染者又以一定的概率痊愈成为免疫者,传染过程以系统中出现已染病者为开始,以系统中已染病者消失为结束<sup>[10]</sup>。由于舆论在人际关系网络中的散布和病毒传播、扩散相似,因此现有的舆论传播模型大都借鉴了 SIR 模型<sup>[7]</sup>。

Sudbury 研究了信息在网络中传播的动力学机制,包括谣言在内的信息传播的形式为舆论传播,其动力学行为也服从 SIR 模型<sup>[11]</sup>。Zanette 将复杂网络理论应用于舆情传播的研究,在小世界网络上建立了舆情传播模型,得到传播临界值,并且采用 SIR 模型研究了舆情在小世界网络中的传播情况<sup>[8]</sup>。Moreno 等又在无标度网络上建立了舆情模型,同时把有计算机仿真和通过随机分析方法得出的结论进行了比较,结论指出舆情在均匀网络中的传播并不存在非零临界值<sup>[9]</sup>。

基于传统传染病模型(SIR 模型),本文对微博上舆情传播规律进行研究,建立了微博舆情传播模型,并用 2011 年甘肃校车事故在新浪微博上的传播进行实证研究。

## 3 微博舆情传播 SIR 模型

### 3.1 舆情传播过程分析

基于传染病传播模型,我们建立了微博舆情传播 SIR 模型:将系统中的全部人口分为不知道消息的人,即无知者(S 态);知道消息并传播消息的人,即传播者(I 态)和知道消息但不传播消息,即免疫者(R 态)。假设微博网络上有  $N$  个节点,每个节点代表可传播消息的个人,他们传播消息的行为规则如下:知道消息的人向他的邻居传播消息。如果邻居是无知者,则该邻居就得到此消息,并以一定的概率  $\alpha$  成为传播者。但如果他的邻居已经知道了此消息,那么传播消息的人就可能认为该消息失去了继续传播的价值,没有了传播该消息的兴趣,即以概率  $\beta$  转变为免疫者。与传染病模型相仿,舆情传播始

于系统中出现一个或若干个传播者,终止于传播者的消失。这种舆情传播的过程如图1所示。

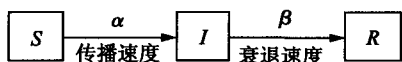


图1 舆情传播过程

在新浪微博用户中,用 $S$ 、 $I$ 和 $R$ 代表无知者、传播者和免疫者,它们的总和是新浪微博的用户总数 $N$ ,即: $N = I + S + R$ 。在传播过程中无知者和传播者的人数不为0,我们利用谣言传播模型研究微博上舆情传播的规律。模型的描述如下:

传播速度: $S * (I/N) * \alpha$ ,

衰退速度: $I * (I + R)/N * \beta$ 。

### 3.2 传播阈值分析

当传播速度大于衰退速度时,传播呈增长趋势。

即: $S * (I/N) * \alpha > I * (I + R)/N * \beta$ ;

$S * \alpha > (I + R) * \beta$ ;

得: $(I + R)/S < \alpha/\beta$

即当传播者和免疫者人数之和与无知者人数之比小于传播速度与衰退速度之比时传播呈增长趋势,反之则舆情传播呈消退趋势。

## 4 案例分析——甘肃校车事件

2011年11月16日上午9时许,在甘肃省庆阳市正宁县榆林子镇西街道班门“小博士”幼儿园校车与一辆陕西咸阳号牌大翻斗运煤货车(陕D—72231)迎面相撞,造成19名幼儿、1名司机及1名陪护教师死亡,此次车祸还造成车上的44人受伤,伤者多为幼儿园的儿童,其中重伤十余人。该车严重超载,核载9人车辆实际上载64人,且事发时在逆向行驶。

甘肃校车事件是一个突发事件,事件发生后,在微博上引起了热议,在短短的几天里,很多网民转发关于此事故的微博,同时进行评论。因此,这是一个比较典型的微博舆情传播的案例。

### 4.1 数据收集

我们利用新浪微博里“高级搜索”功能收集了11月16~25日这一事件相关的微博数据。我们所收集的是公开的微博。如果用户将他们的微博设置成私密,则只有他们自己才能够看到他们的微博,我们无法追踪这些微博。但由于私密微博不参与传播

过程,因此缺失的这一部分数据不会影响模型的构建与分析。

由于关键词选择的不同会导致数据搜集结果的不同。我们尝试了“甘肃校车”、“校车事故”、“幼儿园校车”、“安全条例”等跟这次突发事件有关的关键词,但是经过搜索,发现这些关键词都过于片面,会遗漏部分数据。例如,若使用“甘肃校车”作为关键词,由于事件发生在庆阳市正宁县榆林子镇,就会丢失微博中使用“庆阳校车”或“正宁校车”这样字眼的微博数据。所以最后我们决定用“校车”作为搜索关键词。校车这样的关键词范围较广,但是考虑到在16~25日绝大部分微博用户提到“校车”时都是在谈论关于这起校车事件的,提到校车而非在谈论甘肃校车事件的用户极少,我们认为可以忽略不计,使用“校车”作为关键词可以较准确地反应微博上甘肃校车事件舆情传播情况。

为了研究微博上甘肃校车事件的舆情传播规律,我们针对这起事件对每日的原创微博(即作者自己首创的或简短文字,或图片,或音频,或视频的微博),转发微博(非首创,转载别人所发的微博),总微博(原创微博加转发微博),带有视频的微博(即微博中有插入视频),以及带有URL(即微博中有短链可以链接到其他网页)的微博进行了统计。表1为收集的甘肃校车事件中16~25日的微博量。

表1 甘肃校车事件微博数据

日期	总微博量	原创 微博量	转发 微博量	带视频 微博量	带URL 微博量
11.16	492 100	68 628	423 472	46 284	104 272
11.17	1 078 896	96 824	982 072	83 524	245 252
11.18	573 496	95 760	477 736	22 876	163 856
11.19	332 500	37 408	295 092	7 221	65 436
11.20	221 312	33 516	187 796	2 858	50 008
11.21	101 080	26 068	75 012	1 908	23 408
11.22	90 440	18 088	72 352	10 762	38 836
11.23	47 880	8 733	39 147	377	7 980
11.24	31 388	8 512	22 976	254	9 576
11.25	22 344	5 320	17 024	303	7 448

### 4.2 数据分析

分析表1数据,发现以下规律:

1) 每日关于校车事件的微博量以及转发微博的微博量如图2,校车事件发生后的前两天,关于该事件的微博量以非常快的速度激增,第二天的微博量达到峰值。从第三天开始关于该事件的微博量就逐渐减少。从图3中可以观察到无论原创微博、带有视频的微博,还是带有URL的微博量都基本遵循这个规律。

2) 从图2和图3中发现校车事件在微博上的传播过程中,微博舆情的爆发速度快,衰退的速度则稍缓,在达到峰值的六天后相关微博数稳定在较低的水平。

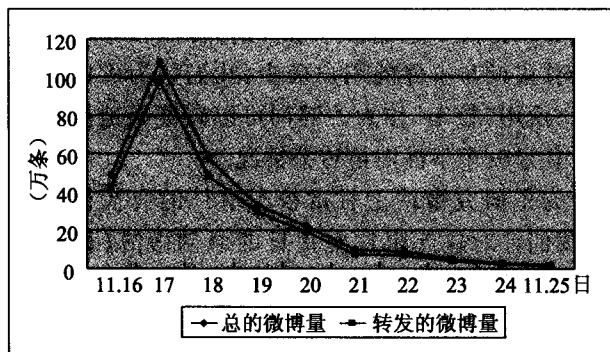


图2 总微博量及转发微博量

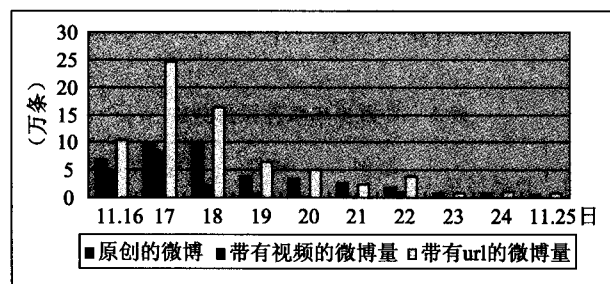


图3 原创、带视频、带URL微博量

3) 在图3中我们观察到,带有URL的微博量比带有视频的微博量明显要多。微博功能限制每条微博字数最多140字,用140字可能不足以把事件说明清楚,所以,很多用户都需要通过URL链接到其他网页,获得很多相关信息。因此,带有URL的微博得到了更广泛的传播。

4) 带有视频的微博量并不大,小于原创的微博量,也小于带URL的微博量。这说明只有部分原创微博是带有视频的。通常人们认为图片和视频更容易引起注意。但是从我们收集的数据来看,校车事件中带有视频的微博并没有得到广泛的传播。我们发现校车事件中的视频基本都是电视台的新闻节

目,没有目击者上传事件发生的实时情况。因为新闻视频已经出现在电视媒体,所以转载率可能大大降低。可见图片和视频虽容易吸引眼球,但最终是否能得到广泛传播内容才是关键因素。

### 4.3 模型仿真

在甘肃校车事故发生时,新浪微博的总用户数大约是2亿。新浪总裁曹国伟在与投资者进行沟通时称,新浪微博“每日登录用户所占比例与Twitter相似,大概是10%”,即约2000万。剩余的90%用户为不活跃用户,他们对事件传播的影响可以忽略不计,我们将这部分不活跃用户去除,所以我们考虑的总人数为2000万活跃用户。数据显示从2011年11月16日下午13点49分,微博上出现关于校车事件的第一条微博,此时距离14点仅仅只有11分钟,此间的数据只有19条微博。考虑到在微博网络中可能存在多个传播源头,所以决定采用从13点49分到15点之间出现的数据,此间原创微博共3480条,我们以此数据为初始的传播用户数 $I(0)$ 。 $S(0)$ 和 $R(0)$ 分别代表当日15点时无知者人数和免疫者人数。因此,根据收集的甘肃校车事件数据,我们对舆情传播模型参数经行如下估计:

$$N = 2000 \text{ 万人};$$

$$I(0) = 0.348 \text{ 万人};$$

$$S(0) = 1999.6520 \text{ 万人};$$

$$R(0) = 0$$

通过模型拟合,我们将传播率和衰退率分别设置为 $\alpha = 0.17$ 和 $\beta = 0.04$ 。

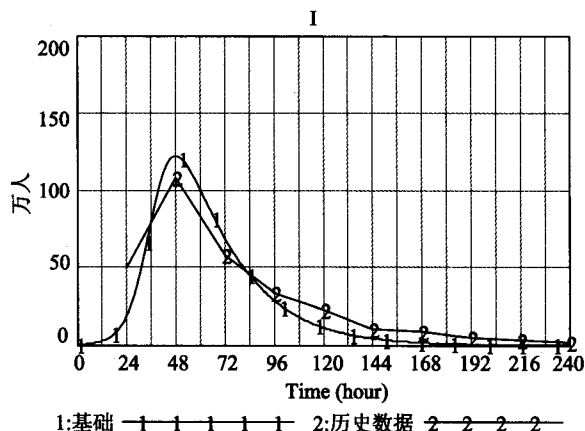


图4 SIR模型和案例参考模型传播过程

模型的仿真结果(1曲线)与案例中的舆情传播数据(2曲线)基本吻合:在第二天传播者人数达到峰值。第二天后传播开始衰退,但是衰退速度相对

于爆发速度稍缓,同样也是在达到峰值后六天传播者才逐渐消失。这样的结果说明微博舆情传播和传统的谣言传播有相似之处,可以用 SIR 模型来仿真微博舆情传播。

#### 4.4 情景分析

利用微博舆情传播模型,我们仿真几种不同情景下微博舆情传播的情况。为了增加各个情景模拟结果的对比性,我们决定选择以 50% 的幅度调整各参数:情景一是基于初始参数设置的基础情景;情景二将活跃用户人数增加 50%,达到 3000 万,其他参数均不变;情景三将传播率增加 50%,达到 0.255,其他参数均不变;情景四将衰退率增加 50%,达到 0.06,其他参数均不变。这四个情境的具体参数设置见表 2。

表 2 四个情境的具体参数设置

情景	活跃用户数(万人)	传播率	衰退率
基础	2000	0.170	0.040
增加活跃用户	3000	0.170	0.040
增加传播率	2000	0.255	0.040
增加衰退率	2000	0.170	0.060

上述四种情景下微博舆情的传播过程如图 5 所示。观察四种情景的仿真结果发现:

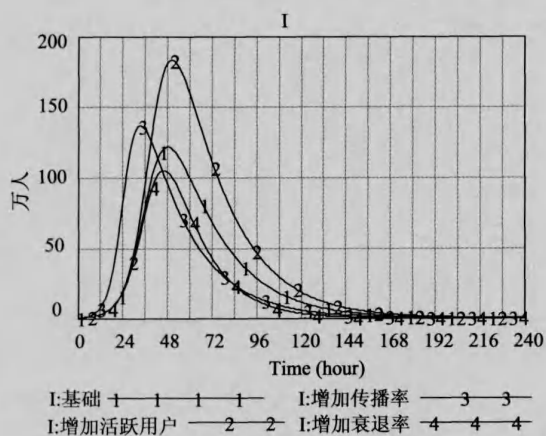


图 5 四种情景的仿真结果

1) 当活跃用户增加 50% 时,其传播过程与基础情景相似,第二天出现峰值,六天后基本消退。但其达到的峰值比基础情景下增加 30%。这就是说舆情传播的范围更加广泛了。

2) 当传播率增加 50% 时,传播者达到顶峰的速度相对于基础情景时快,在 1 天半时就达到峰值,峰值比基础情景时高近 10%。当传播速度加快时,舆情爆发的更快,但舆情传播的范围的增加有限。由此可见,爆发速度快的舆情未必影响力更大。

3) 当衰退率增加 50% 时,传播过程中,传播速度与基础情景基本一致,也是第二天出现峰值,六天后基本消退。但峰值较基础情景时低近 10%。当衰退率增加时,舆情传播范围有所减小,但减小力度有限。这说明增加衰退率对于控制舆情发展的作用有限。

当活跃用户、传播率和衰退率这三个参数各增加 50% 的情况下,活跃用户的增加对舆情传播的影响最大。所以可以得出影响舆情传播范围的主要因素是总人数,也就是舆情传播初期无知者人数的多少。随着新浪等微博被越来越多的人应用,无知者人数会增加,舆情的传播范围也会更广。

## 5 结 论

本文建立了微博舆情传播 SIR 模型,并以“11.16”甘肃校车事件为例进行了实证分析,得到了以下结论:

1) 通过微博舆情传播 SIR 模型可以对微博上舆情传播规律进行理论分析和预测。

2) 甘肃校车事件在微博上传播的规律为舆情初期爆发快,后期消退慢;影响用户传播微博的主要因素是内容:所有微博中带有 URL 的微博量最大,因为用户可以通过 URL 获得更详尽的信息。带有新闻视频的微博并没有得到广泛传播,这是因为那些视频在新闻中已经播出,并没有特别新鲜的内容。

3) 微博舆情传播 SIR 模型仿真结果与校车事件在微博上的传播过程基本吻合。

4) 情景仿真结果发现相对于传播率和衰退率同比例的变化,活跃用户人数的变化对舆情传播的影响最大。

本文对突发事件在微博上传播的行为进行了分析,探究其传播规律。但此研究仅以甘肃校车事件为案例,其他不同案例的传播过程可能会有不同。在未来的研究中可以多收集不同案例的数据,进行研究。目前,我们仅考虑每天微博数量变化,未来会考虑更多的信息。比如,通过微博语义的变化趋势,可以了解民众情绪的转变过程等。本文的研究可以



为科学的调节和控制突发事件舆情传播,降低突发事件中谣言传播,为突发事件应急管理提供理论支持。

### 参 考 文 献

- [1] 朱力. 突发事件的概念、要素及类型[J]. 社会研究, 2007(11):81-88.
- [2] 张子利,蒋远翔. 互联网突发事件的舆论传播特征及引导研究[J]. 电脑知识与技术, 2011, 13(7): 3045-3047.
- [3] 宋晓龙. 突发事件的互联网信息传播规律研究[D]. 哈尔滨:哈尔滨工业大学,2011.
- [4] 潘芳,卞艺杰,潘郁. 危机事件网络舆情传播模型及消极思潮应急对策[J]. 图书馆学理论研究, 2010, 54(15):40-43.
- [5] Bernardo A H, Daniel M R, Wu Fang. Social networks that matter: Twitter under the microscope [J]. Physical Review, 2009(14):1-5.
- [6] Amanda L H, Leysia P. Twitter Adoption and Use in Mass Convergence and Emergency Events[C]. Proceedings of the 6<sup>th</sup> International ISCRAM Conference. Sweden, 2009.
- [7] 张芳,司光亚,罗批. 谣言传播模型研究综述[J]. 复杂系统与复杂性科学, 2009, 6(4):1-11.
- [8] Zanette D H. Dynamics of rumor propagation on small-world networks[DB/OL]. [2009-03-01]. [http://arxiv.org/PS\\_cache/cond-mat/pdf/0110/0110324v1.pdf](http://arxiv.org/PS_cache/cond-mat/pdf/0110/0110324v1.pdf).
- [9] Moreno Y, Nekovee M, Pacheco A F. Dynamics of rumor spreading in complex networks[J]. Physical Review, 2004(69):6-30.
- [10] Florence D\_ebarre. SIR models of epidemics[DB/OL]. [2012-09-02]. <http://www.tb.ethz.ch/education/model/SIR/sir.pdf>.
- [11] Sudbury A. The proportion of the population never hearing a rumour[J]. Journal of Applied Probability, 1985, 22(2): 443-446.

(责任编辑 化柏林)