

· 舆情研究 ·

面向舆情大数据的网民情绪演化机理及趋势预测研究^{*}

兰月新¹ 夏一雪¹ 刘冰月² 高扬¹ 李增¹

(1. 中国人民武装警察部队学院 廊坊 065000;

2. 天津交通职业学院 天津 300132)

摘要 [目的/意义] 大数据环境下,不同的网民情绪在舆情传播过程中相互影响,有些甚至产生“群体极化”现象,如何准确把握网民情绪演化机理并预测其演化趋势是大数据环境下网络舆情治理的重要内容。[方法/过程] 定性分析大数据环境下网民情绪特征和分类,构建网民情绪演化机理的微分方程模型,通过计算模型平衡点及其稳定性条件研究网民情绪竞争场景以及通过回归分析研究网民情绪演化趋势预测问题。[结论/结果] 经过理论建模和仿真分析得出不同类型情绪网民数量的变化规律,得出面向网络舆情的网民情绪趋势预测方法,最后提出网民情绪演化机理模型高维拓展问题、同时期舆情事件对网民情绪的冲击问题以及网络舆情信息内容和网络推手等对网民情绪的影响问题等一系列有待进一步研究的问题。

关键词 网络舆情 大数据 网民情绪 微分方程 演化机理

中图分类号 C912.6

文献标识码 A

文章编号 1002-1965(2017)11-0134-07

引用格式 兰月新,夏一雪,刘冰月等.面向舆情大数据的网民情绪演化机理及趋势预测研究[J].情报杂志,2017,36(11):134-140.

DOI 10.3969/j.issn.1002-1965.2017.11.021

Research on the Evolution Mechanism and Trend Prediction of Netizens for Public Opinion Big Data

Lan Yuexin¹ Xia Yixue¹ Liu Bingyue² Gao Yang¹ Li Zeng¹

(1. The Chinese People's Armed Police Force Academy Langfang 065000;

2. Transportation Vocational College, Tianjin 300132)

Abstract [Purpose/Significance] In big data network public opinion environment, different users' emotions interact in the process of public opinion dissemination, some even show a "group polarization" phenomenon, and how to accurately grasp the evolution mechanism for netizens' emotions and predict the evolution trend are the important content of network public opinion management under the big data environment. [Method/Process] Qualitative analysis is used to explore the netizens' emotional characteristics and classification under the big data environment, the differential equation model for emotion evolution mechanism of netizens is constructed, the netizen emotions competitive scene is studied via calculating the model balance point and discussing its stability conditions, and the prediction of the evolution trend of netizens' emotions is analyzed through regression analysis. [Result/Conclusion] Through theoretical modeling and simulation analysis, the change rules for the numbers of the netizens of different types of emotions are identified, and the emotions trend prediction methods oriented to Internet public opinion are summarized. The research also poses the issues for further study such as the netizen emotions evolution mechanism model of higher dimensional expansion, the impact of other same period public opinion events on the netizens' emotions, and the influence of network public opinion information content and Internet marketer on netizens' emotions.

Key words network public opinion big data netizens' emotions differential equations evolution mechanism

收稿日期: 2017-08-27

修回日期: 2017-09-28

基金项目: 河北省社会科学基金项目“面向突发事件的网民情绪风险建模与对策研究”(编号: HB16GL098)。

作者简介: 兰月新(ORCID: 0000-0002-4791-5094),男,1981年生,副教授,硕士生导师,研究方向: 网络舆情; 夏一雪(ORCID: 0000-0002-8044-0553),女,1983年生,博士,讲师,研究方向: 公共管理; 刘冰月(ORCID: 0000-0002-3874-7754),女,1989年生,硕士,助教,研究方向: 数据分析; 高扬(ORCID: 0000-0002-3221-7789),男,1987年生,硕士,讲师,研究方向: 数据分析; 李增(ORCID: 0000-0002-2822-0421),男,1982年生,博士,讲师,研究方向: 数据分析。

1 现状分析

截至 2017 年 7 月,中国网民规模 7.51 亿,互联网普及率 54.3%,其中手机网民比例为 96.3%^[1]。随着移动宽带互联网的普及,网络舆情在数据体量、复杂性和产生速度等方面发生巨大变化,已经呈现大数据环境。任何网络话题在网络上极易形成规模或大或小的网络舆情,不同的网民情绪在舆情传播过程中相互影响,有些甚至产生“群体极化”现象,如何准确把握网民情绪演化机理并预测其演化趋势是大数据环境下网络舆情治理的重要内容。

国内外研究网民情绪的文献主要包括两个方面:

①大部分从心理学、生理学、新闻学与传播学等视角研究网民情绪的影响因素。例如情绪与社会网络的关系^[2]、恐惧情绪感染时大脑皮质的变化^[3]、事件信息质量与隐含的情绪张力的关系^[4]、媒体信息传播与受众情绪感染的相关性^[5]、情绪分享与情绪性事件传播的关系^[6]、群体负面情绪的识别和测度^[7]、群体情绪的途径和影响因素等^[8]。②从系统动力学、统计学等视角通过建模与仿真研究情绪演化机理。例如通过数学方法建立突发事件群体情绪层次模型与情绪监控预警模型^[9]、通过心理实验验证群体性事件情绪动力模型^[10]、模仿—反馈机制的情绪感染机制^[11]、基于群际情绪理论(Intergroup Emotion Theory,简称 IET)分析群体情绪传播的基本过程^[12]等。

除此之外,网民情绪研究与网民情感研究在分类、识别及传播方面有共通之处:①分类研究。主要使用 EMCNN 模型^[13]、朴素贝叶斯^[14-15](NB)、最大熵^[16](maximum entropy,ME)、支持向量机^[17](support vector machines,SVM)等机器学习算法训练得到分类模型用于情感分类。②识别研究。主要通过条件随机场方法^[18]、卷积树核^[19]、全局变量 CRFs 模型^[20]等方法开展情感识别研究。③情感传播规律研究。国内学者纪雪梅等(2015)从 3 个方面对国内外在线社交网络情感传播的相关研究成果进行梳理并总结了目前研究存在的不足和未来的研究方向^[21]。

虽然学界已有较多网民情绪相关的研究,但目前仍存在许多不足之处,主要体现在:①现有研究大多集中于情绪分类和识别研究,较少涉及不同类型情绪网民之间的竞争和影响关系;②尚无大数据环境下网民情绪演化趋势预测方面的研究。基于此,本文在定性研究网民情绪特征和分类的基础上,通过建立微分方程模型,从定量研究的视角探讨面向舆情大数据的网民情绪演化机理以及趋势预测问题,以期政府治理网络舆情提供理论参考。

2 面向舆情大数据的网民情绪演化特征和分类

随着移动宽带互联网的普及,网络话题极易引发规模或大或小的网络舆情。由于政府、网络媒体、各类网民等舆情主体参与网络话题讨论时的情绪状态各有不同,导致各个主体发布反映各自情绪状态的文字、图片、视频、音频、表情包等信息。大量格式各异的情绪信息在微信、微博、论坛、新闻网站、视频网站等多个网络传播平台上扩散。由此不难看出:一方面,随着信息交互程度的加大,这些情绪信息体量巨大、产生速度快、信息格式多样;另一方面,由于信息异化,情绪信息中有价值的信息密度低下。综合考察这两个方面,网络舆情已经呈现大数据特征(见图 1),如何从网络舆情大数据中“提纯”,准确把握和预测网民情绪,感知风险,辅助政府决策,成为网络舆情治理的关键。

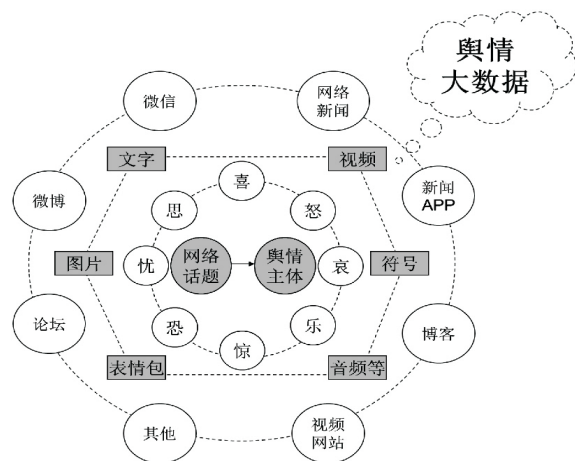


图 1 网络舆情大数据环境

a. 网民结构特征决定了整体网民的情绪特征,网民情绪状态多元化并且负面消极情绪比重大。我国网民的年龄、学历、收入、职业等结构复杂(表 1,数据由最新 CNNIC 报告^[1]整理),导致网民发表网络言论视角众多,网民心理多元化,娱乐时尚、减压宣泄、猎奇探究、彰显个性、渴求新知、跟风从众、追求平等、渴望创新和自我实现成为网民舆情心理的主要特征^[22],形成多元的网民情绪状态。同时,由于低年龄、低学历、低收入的网民所占比例较高,学生、自由职业者、失业等职业不稳定网民比例也很高,导致容易出现偏激、非理性等负面消极言论。纵观近几年的热点舆情案例,负面消极的网民情绪比重较大,这与我国当前网民结构特征是密不可分的。

表 1 我国网民结构统计数据

网民结构	大部分网民结构及数量比例		重点网民结构及数量比例	
年龄结构	10-29 岁	49.1%	20-29 岁	29.7%
学历结构	高中学历及以下	79.4%	初中学历以下	54.9%
收入结构	3000 元以下	58.4%	1000 元以下	26.6%
职业结构	学生、自由职业者、失业等	52.2%	学生	24.8%

b. 网民情绪分类是研究网民情绪特征的重点。关于网民情绪分类,在学术领域,宏观分法有二分法^[23](积极和消极)、三分法^[24](正面、中性和负面)等,微观分法有四分法^[25](愤怒、厌恶、高兴、悲伤)、七分法^[26](愤怒、厌恶、恐惧、高兴、喜好、悲伤、惊讶)等;而在应用领域,分类则相对简单,例如清博舆情、美亚舆情等监测软件通过情绪分类方法将网民情绪分为正面、中性和负面情绪,并实时将监测获取的网络舆情数据按网民情绪分为三类。本文致力于研究网民情绪演化机理,所以综合学术领域和应用领域,将网民情绪分为正面、中性和负面,来研究反映三种情绪的网民数量演化关系。

结合大数据环境下网民情绪特征和分类,本文将围绕某个网络话题发表观点的网民分为三类:正面情绪的网民、负面情绪的网民和中性情绪的网民,不同情绪网民的数量变化规律是研究网民情绪演化机理的关键。当某个网络舆情发生后,随着三类情绪网民数量的增加,网络舆情信息量也在急剧增加,然而网络空间中可以容纳此类舆情的信息量是有上限的^[27-28],所以网民情绪演化机理的实质就是三类情绪网民竞争有限的网络空间信息量资源的问题(图2),这一点与生态科学中种群竞争有限资源演化机理相似。基于此,本文借鉴生态科学中种群竞争 Lotka-Volterra 模型^[29]的建模思路,通过微分方程模型研究网民情绪演化机理。

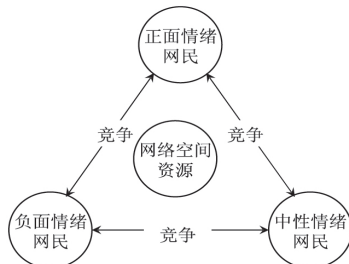


图2 网民情绪演化机理的实质

3 面向舆情大数据的网民情绪演化模型研究

3.1 建模基本假设 假设 $x_1(t)$, $x_2(t)$, $x_3(t)$ 分别代表正面情绪的网民、负面情绪的网民和中性情绪的网民的数量, r_1 , r_2 , r_3 分别代表三类网民对应的增长率, K_1 , K_2 , K_3 分别代表三类网民对应的上限, 初值分别为 x_{10} , x_{20} , x_{30} 。

当只有单一情绪类别网民时,即没有竞争关系时的,单类情绪的网民数量的变化量与数量基数 x_i 、“剩余”空间 $(1 - \frac{x_i}{K_i})$ 成正比,即

$$\begin{cases} \frac{dx_i(t)}{dt} = r_i x_i(t) (1 - \frac{x_i(t)}{K_i}) \\ x_i(0) = x_{i0} \end{cases}, i = 1, 2, 3.$$

这就是网络舆情发生后,描述网络空间单类情绪网民竞争的 logistic 模型。

3.2 面向舆情大数据的网民情绪演化模型 网络舆情发生后,正面情绪的网民、负面情绪的网民和中性情绪的网民等三类网民并存于网络空间,所以需要对单类情绪网民竞争模型进行拓展来研究三者并存的竞争模型。为详尽描述不同类别情绪网民相互影响、竞争的程度,根据前文研究,制定本文的竞争模式(图3)并同时引进描述竞争力的量化指标,即竞争参数 α , β , γ 。

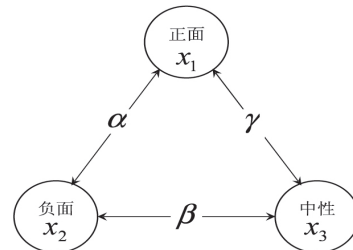


图3 不同类别情绪网民竞争模式

由于正面情绪网民和中性情绪、负面情绪网民的竞争作用,导致正面情绪网民数量“剩余”空间发生变化,即由 $(1 - \frac{x_1}{K_1})$ 变为 $(1 - \frac{x_1}{K_1} - \alpha \frac{x_2}{K_2} - \gamma \frac{x_3}{K_3})$, 从而正面情绪网民数量模型为

$$\frac{dx_1}{dt} = r_1 x_1 (1 - \frac{x_1}{K_1} - \alpha \frac{x_2}{K_2} - \gamma \frac{x_3}{K_3})$$

同理,由于三者竞争关系,负面情绪网民数量模型为

$$\frac{dx_2}{dt} = r_2 x_2 (1 - \frac{x_2}{K_2} - \alpha \frac{x_1}{K_1} - \beta \frac{x_3}{K_3});$$

中性情绪网民数量模型为

$$\frac{dx_3}{dt} = r_3 x_3 (1 - \frac{x_3}{K_3} - \gamma \frac{x_1}{K_1} - \beta \frac{x_2}{K_2}).$$

综合以上三个方程,得出面向舆情大数据的网民情绪演化机理模型为

$$\begin{cases} \frac{dx_1}{dt} = r_1 x_1 (1 - \frac{x_1}{K_1} - \alpha \frac{x_2}{K_2} - \gamma \frac{x_3}{K_3}) \\ \frac{dx_2}{dt} = r_2 x_2 (1 - \frac{x_2}{K_2} - \alpha \frac{x_1}{K_1} - \beta \frac{x_3}{K_3}) \\ \frac{dx_3}{dt} = r_3 x_3 (1 - \frac{x_3}{K_3} - \gamma \frac{x_1}{K_1} - \beta \frac{x_2}{K_2}) \end{cases},$$

模型初值分别为 x_{10} , x_{20} , x_{30} 。由于方程组没有解析解,所以需要研究模型平衡点及其稳定性、数值解的性质等来研究网民情绪演化机理。

3.3 网民情绪演化模型分析

3.3.1 模型平衡点及其稳定性 为研究方程解的性质及三类情绪网民的竞争结局,即 $t \rightarrow \infty$ 时, $x_1(t)$, $x_2(t)$, $x_3(t)$ 的趋向,需要研究平衡点及其稳定

性。令

$$\begin{cases} f_1(x_1, x_2, x_3) \equiv r_1 x_1 \left(1 - \frac{x_1}{K_1} - \alpha \frac{x_2}{K_2} - \gamma \frac{x_3}{K_3}\right) = 0 \\ f_2(x_1, x_2, x_3) \equiv r_2 x_2 \left(1 - \frac{x_2}{K_2} - \alpha \frac{x_1}{K_1} - \beta \frac{x_3}{K_3}\right) = 0 \\ f_3(x_1, x_2, x_3) \equiv r_3 x_3 \left(1 - \frac{x_3}{K_3} - \gamma \frac{x_1}{K_1} - \beta \frac{x_2}{K_2}\right) = 0 \end{cases}$$

得到 7 个非零平衡点

$$\begin{aligned} &P_1(K_1, 0, 0), P_2(0, K_2, 0), P_3(0, 0, K_3), \\ &P_4\left(0, \frac{K_2}{1+\beta}, \frac{K_3}{1+\beta}\right), P_5\left(\frac{K_1}{1+\gamma}, 0, \frac{K_3}{1+\gamma}\right), \\ &P_6\left(\frac{K_1}{1+\alpha}, \frac{K_2}{1+\alpha}, 0\right), P_7\left(\frac{\Delta_1 K_1}{\Delta}, \frac{\Delta_2 K_2}{\Delta}, \frac{\Delta_3 K_3}{\Delta}\right) \end{aligned}$$

其中

$$\begin{aligned} \Delta &= 1 + 2\alpha\beta\gamma - \alpha^2 - \beta^2 - \gamma^2, \\ \Delta_1 &= (1 - \beta) [1 + \beta - (\gamma + \alpha)], \\ \Delta_2 &= (1 - \gamma) [1 + \gamma - (\alpha + \beta)], \\ \Delta_3 &= (1 - \alpha) [1 + \alpha - (\beta + \gamma)]. \end{aligned}$$

根据文献^[30]的研究结果,整理获得本文模型的平衡点及稳定性条件见表 2。

表 2 模型平衡点稳定性

平衡点	稳定性条件
$P_1(K_1, 0, 0)$	$\alpha > 1, \gamma > 1$
$P_2(0, K_2, 0)$	$\alpha > 1, \beta > 1$
$P_3(0, 0, K_3)$	$\gamma > 1, \beta > 1$
$P_4\left(0, \frac{K_2}{1+\beta}, \frac{K_3}{1+\beta}\right)$	$0 < \beta < 1, 1 + \beta < \alpha + \gamma$
$P_5\left(\frac{K_1}{1+\gamma}, 0, \frac{K_3}{1+\gamma}\right)$	$0 < \gamma < 1, 1 + \gamma < \alpha + \beta$
$P_6\left(\frac{K_1}{1+\alpha}, \frac{K_2}{1+\alpha}, 0\right)$	$0 < \alpha < 1, 1 + \alpha < \gamma + \beta$
$P_7\left(\frac{\Delta_1 K_1}{\Delta}, \frac{\Delta_2 K_2}{\Delta}, \frac{\Delta_3 K_3}{\Delta}\right)$	$0 < \alpha, \beta, \gamma < 1,$ $1 + \alpha > \gamma + \beta$ $1 + \beta > \alpha + \gamma$ $1 + \gamma > \alpha + \beta$

3.3.2 模型分析 观察表 2 中 7 个平衡点的结构, 容易发现 7 个平衡点可以分为三个类型: 有两个分量为 0 的类型, 有 1 个分量为 0 的类型和 3 个分量均不为 0 的类型。为了便于讨论, 将这三个类型定义为极端场景(类型 1)、失衡场景(类型 2)和共存场景(类型 3)(表 3)。

表 4 情绪信息量统计数据

编号	舆情事件	信息总量	正面比例	负面比例	中性比例	近似归属类型
1	“一带一路”高峰论坛	1417686	86%	4%	10%	类型 2: 失衡场景
2	首个中国品牌日	20334	89%	1%	10%	类型 2: 失衡场景
3	雄安医疗	91586	86%	4%	10%	类型 2: 失衡场景
4	2017 年艺考	11471873	87%	2%	11%	类型 2: 失衡场景
5	杭州 G20 峰会	5963354	67%	5%	28%	类型 3: 共存场景
6	2017 春晚	658003	65%	8%	27%	类型 3: 共存场景
7	徐玉玉电信诈骗案	621316	20%	40%	40%	类型 3: 共存场景
8	宁波老虎咬人事件	44132	20%	45%	35%	类型 3: 共存场景

注: 表中数据由清博舆情监测软件(yuqing.gsdata.cn)获取。

表 3 模型平衡点场景归类

平衡点	竞争类型
$P_1(K_1, 0, 0)$	类型 1: 极端场景
$P_2(0, K_2, 0)$	
$P_3(0, 0, K_3)$	
$P_4\left(0, \frac{K_2}{1+\beta}, \frac{K_3}{1+\beta}\right)$	类型 2: 失衡场景
$P_5\left(\frac{K_1}{1+\gamma}, 0, \frac{K_3}{1+\gamma}\right)$	
$P_6\left(\frac{K_1}{1+\alpha}, \frac{K_2}{1+\alpha}, 0\right)$	
$P_7\left(\frac{\Delta_1 K_1}{\Delta}, \frac{\Delta_2 K_2}{\Delta}, \frac{\Delta_3 K_3}{\Delta}\right)$	类型 3: 共存场景

类型 1: 极端场景

极端场景中, 三类情绪的网民竞争的结果是只剩一类情绪的网民, 并且该情绪网民数量趋近于上限, 其他两类网民数量趋近于零, 这类极端场景在实际网络舆情传播过程中相对较为少见。

类型 2: 失衡场景

失衡场景中, 三类情绪的网民竞争的结果是剩余两类情绪的网民, 数量趋近于原始上限的一部分, 但两类网民的数量比值与竞争前一样, 另外一类情绪网民数量趋近于零。在网络舆情传播过程中, 竞争结果属于失衡场景的案例较多。例如表 4 中的前四个案例, 负面情绪比例极低, 正面和中性比例较高, 近似可以理解为属于此类场景。

类型 3: 共存场景

共存场景中, 三类情绪网民共存于网络舆情整个传播过程之中, 数量趋近于原始上限的一部分, 正面、负面和中性情绪网民数量比值受竞争影响呈现一定的比例

$$\frac{\Delta_1 K_1}{\Delta} : \frac{\Delta_2 K_2}{\Delta} : \frac{\Delta_3 K_3}{\Delta} = \Delta_1 K_1 : \Delta_2 K_2 : \Delta_3 K_3。$$

例如表 4 中的后四个案例, 正面、中性和负面情绪比例即为竞争结果, 均属于共存场景。

4 面向舆情大数据的网民情绪演化趋势预测研究

大数据的核心是预测。如何在舆情大数据环境中预测网民情绪趋势是网民情绪研究的关键。本文构建了网民情绪演化的微分方程模型, 据此可以研究网民

情绪演化趋势。然而在实际应用过程中,需要确定模型参数才能确定平衡点类型,进而研究某个具体舆情事件的网民情绪演化趋势。网络舆情数据是预测的基础,当网络舆情刚刚发生时,没有舆情监测数据或者数据量较少,此时难以通过数据预测网民情绪演化趋势。基于此,网民情绪演化趋势预测分为两类:定性预测和定量预测,具体预测思路见图 4。

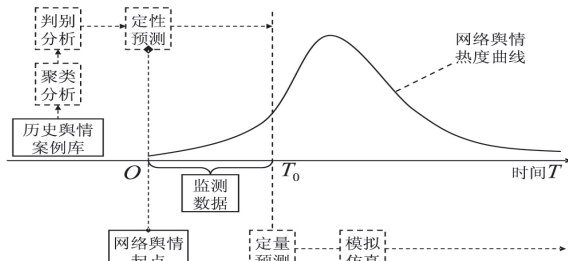


图 4 网民情绪演化趋势预测思路

a. 定性预测: 根据历史舆情案例库的数据信息,通过舆情事件属性开展聚类分析,然后计算每一类别中网民情绪演化模型对应的参数,当某个舆情事件发生后,通过判别分析确定归属,然后应用此类对应的参数,预测舆情传播初期的网民情绪演化趋势。

b. 定量预测: 在获取舆情传播初期的监测数据之后,通过定量分析方法确定本文构建模型的参数,然后在此基础上通过数值仿真预测网民情绪演化趋势。当获取新的舆情监测数据之后,动态修正预测结果,实现网民情绪演化趋势的动态预测。限于篇幅,本文重点研究根据舆情初期数据预测网民情绪演化趋势的方法。

4.1 面向舆情大数据的网民情绪演化趋势预测方法 将前文构建的网民情绪演化机理模型

$$\begin{cases} \frac{dx_1}{dt} = r_1 x_1 \left(1 - \frac{x_1}{K_1} - \alpha \frac{x_2}{K_2} - \gamma \frac{x_3}{K_3} \right) \\ \frac{dx_2}{dt} = r_2 x_2 \left(1 - \frac{x_2}{K_2} - \alpha \frac{x_1}{K_1} - \beta \frac{x_3}{K_3} \right) \\ \frac{dx_3}{dt} = r_3 x_3 \left(1 - \frac{x_3}{K_3} - \gamma \frac{x_1}{K_1} - \beta \frac{x_2}{K_2} \right) \end{cases}$$

变成其对应的差分方程组

$$\begin{cases} \Delta x_1(k) = r_1 x_1(k) \left(1 - \frac{x_1(k)}{K_1} - \alpha \frac{x_2(k)}{K_2} - \gamma \frac{x_3(k)}{K_3} \right) \\ \Delta x_2(k) = r_2 x_2(k) \left(1 - \frac{x_2(k)}{K_2} - \alpha \frac{x_1(k)}{K_1} - \beta \frac{x_3(k)}{K_3} \right) \\ \Delta x_3(k) = r_3 x_3(k) \left(1 - \frac{x_3(k)}{K_3} - \gamma \frac{x_1(k)}{K_1} - \beta \frac{x_2(k)}{K_2} \right) \end{cases}$$

其中 $\Delta x_i(k) = x_i(k) - x_i(k-1)$ $i = 1, 2, 3$ $x_i(k)$ 为某类情绪网民的数量(可通过舆情监测数据获取), k 为监测数据个数。

观察发现,在方程组第一个方程中,差分 $\Delta x_1(k)$ 与 $x_1(k)$ 、 $x_1^2(k)$ 、 $x_1(k)x_2(k)$ 、 $x_1(k)x_3(k)$

呈现四元线性关系,应用四元线性回归分析可以获取回归系数

$$r_1 \frac{r_1}{K_1} \frac{r_1 \alpha}{K_2} \frac{r_1 \gamma}{K_3}。$$

同理,第二个方程、第三个方程也可应用回归分析得到回归系数

$$r_2 \frac{r_2}{K_2} \frac{r_2 \alpha}{K_1} \frac{r_2 \beta}{K_3} \text{ 和 } r_3 \frac{r_3}{K_3} \frac{r_3 \gamma}{K_1} \frac{r_3 \beta}{K_2}。$$

通过简单计算便可以得到网民情绪演化机理模型的全部参数

$$r_i, K_i, \alpha, \beta, \gamma, i = 1, 2, 3。$$

根据这些参数便可以通过 MATLAB 开展数值仿真来预测网民情绪演化趋势。

4.2 基于数值仿真的网民情绪演化趋势预测 根据前文研究,由于不同的参数确定的竞争场景各有不同,当根据部分网络舆情数据通过回归分析确定模型参数后,可以根据平衡点稳定性条件应用数值仿真预测三种场景的网民情绪演化趋势。

类型 1: 极端场景

当确定参数后(表 5),通过 MATLAB 进行数值仿真并绘制网民情绪演化趋势图(图 5)。观察图像发现,三类情绪网民数量趋近于 $P_3(0, 0, 14000)$,其中正面和负面情绪网民数量短时间增加后达到峰值,此后在 10 个单位后很快趋近于 0,而中性情绪网民数量增长呈现近似的“S”型曲线结构,并在 15 个单位后趋近于 14000。同理,亦可设置参数获取 P_1 和 P_2 两个平衡点的情况。在网络舆情传播过程中,极少出现这种极端场景。

表 5 极端场景参数表

情绪类型	初值	上限	增长率	参数组
正面	$x_{10} = 100$	$K_1 = 10000$	$r_1 = 0.5$	$\alpha = 2$
负面	$x_{20} = 60$	$K_2 = 6000$	$r_2 = 0.8$	$\beta = 1$
中性	$x_{30} = 140$	$K_3 = 14000$	$r_3 = 1$	$\gamma = 3$

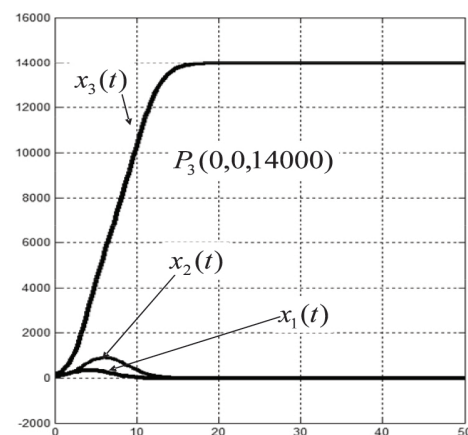


图 5 极端场景趋势预测图

类型 2: 失衡场景

设置三组参数(表 6),研究 P_4 、 P_5 、 P_6 三个平衡点

的情况,分别通过 MATLAB 进行数值仿真并绘制网民情绪演化趋势图(图 6)。观察图像发现,三类情绪网民数量分别趋近于 $P_4(0, 4000, 9333)$ 、 $P_5(7692, 0, 10769)$ 和 $P_6(9091, 5455, 0)$ 。三个平衡点分别于 10 个、20 个、35 个时间单位之后开始趋于稳定,分别与参数 $\beta = 0.5$, $\gamma = 0.3$, $\alpha = 0.1$ 有关,关键的竞争参数越大,网民数量趋于稳定的时间越快,反之越慢,且中间会出现较多波动(第三组参数情形)。

表 6 失衡场景参数表

情绪类型	初值	上限	增长率	参数组 1	参数组 2	参数组 3
正面	$x_{10} = 100$	$K_1 = 10000$	$r_1 = 0.5$	$\alpha = 1$	$\alpha = 1$	$\alpha = 0.1$
负面	$x_{20} = 60$	$K_2 = 6000$	$r_2 = 0.8$	$\beta = 0.5$	$\beta = 2$	$\beta = 1$
中性	$x_{30} = 140$	$K_3 = 14000$	$r_3 = 1$	$\gamma = 2$	$\gamma = 0.3$	$\gamma = 2$

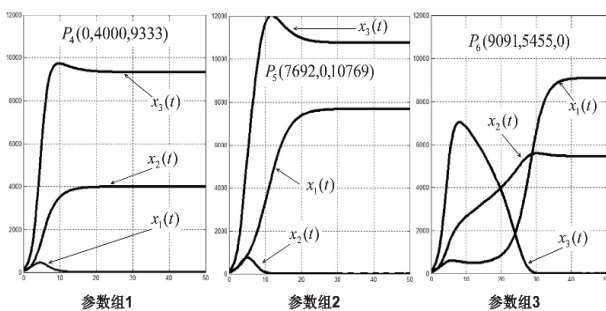


图 6 失衡场景(三个参数情况)趋势预测图

类型 3: 共存场景

设置三组参数(表 7),研究平衡点 P_7 的情况,通过 MATLAB 进行数值仿真并绘制网民情绪演化趋势图(图 7)。观察图像发现,三类情绪网民数量趋近于 $P_7(5000, 3375, 4375)$ 。其中正面情绪网民数量于 55 个单位之后逐步趋于稳定,整体单调递增;负面情绪网民数量变化较快,于 15 个时间单位之后趋于稳定,整

表 7 共存场景参数表

情绪类型	初值	上限	增长率	竞争参数
正面	$x_{10} = 100$	$K_1 = 10000$	$r_1 = 0.5$	$\alpha = 0.5$
负面	$x_{20} = 60$	$K_2 = 6000$	$r_2 = 0.8$	$\beta = 0.6$
中性	$x_{30} = 140$	$K_3 = 14000$	$r_3 = 1$	$\gamma = 0.7$

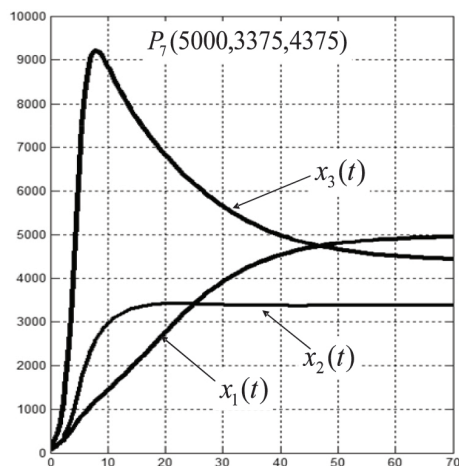


图 7 共存场景趋势预测图

体单调递增且呈现“S”型曲线结构;中性情绪网民则在 10 个单位内快速单调增长至峰值 9200 左右,然后单调递减于 55 个时间单位之后趋于稳定。通过图 7 也可以看出,当通过部分舆情监测数据获取模型参数后,应用 MATLAB 数值仿真便可以预测网民情绪演化趋势,并且随着数据量的增多,预测结果也越来越精确。

5 结论与展望

5.1 研究结论 结合大数据环境下网民情绪特征和分类,围绕某个网络话题发表观点的网民分为三类:正面情绪的网民、负面情绪的网民和中性情绪的网民,不同情绪网民的数量变化规律是研究网民情绪演化机理的关键。基于此,本文借鉴生态科学中种群竞争 Lotka-Volterra 模型的建模思路,构建了网民情绪演化机理模型,通过本文的研究可以得出以下结论:

a. 构建微分方程组研究网民情绪演化机理是可行的。通过分析模型平衡点及其稳定性可以深入研究不同类型情绪网民数量的变化规律,并据此可以将网民情绪竞争结果分为极端场景、失衡场景、共存场景等三种情况,便于清晰研究网民情绪演化机理。

b. 通过网民情绪演化模型可以开展大数据环境下网民情绪演化趋势预测。本文研究了面向网络舆情的网民情绪趋势预测方法,并重点研究通过少量舆情数据确定模型参数的定量预测方法,在此基础上,通过 MATLAB 开展数值仿真预测网民情绪演化趋势。

5.2 研究展望

a. 网民情绪演化机理模型高维拓展问题。参考文献中关于网民情绪分类除了三分法以外,还有四分法、七分法等,如何基于本文的建模思想构建多维度的网民情绪演化机理模型还需进一步研究。

b. 同时期舆情事件对网民情绪的冲击问题。通过分析近几年的网络舆情案例发现,在同一时期经常会出现多个网络舆情事件,在网络空间中,多个舆情事件之间的相互干扰和冲击对网民情绪影响很大,如何在本文建模的基础上考虑舆情事件冲击等外部因素是下一步研究的重点。

c. 网络舆情信息内容以及网络推手等对网民情绪的影响问题。研究显示,互联网对于人们情绪有着较为显著的影响。Facebook 曾做过一个备受争议的实验,结果表明:网民情绪完全可以在不知不觉中被操控。在构建网民情绪演化机理时,需要进一步研究网络舆情信息内容以及网络推手等对网民情绪的影响程度。

参考文献

[1] 中国互联网络信息中心.第 40 次中国互联网络发展状况统计

- 报告 [R]. http://cnnic.cn/gywm/xwzx/rdxw/201708/t20170804_69449.htm 2017-8-4.
- [2] Curci A, Bellelli G. Cognitive and social consequences of exposure to emotional narratives: two studies on secondary social sharing of emotions [J]. *Cognition and Emotion*, 2004, 18(7): 881-900.
- [3] Hadjikhani N. Body expression of emotion do not trigger fear contagion in autism spectrum disorder [J]. *Social Cognitive & Affective Neuroscience* 2009 4(1): 70-78.
- [4] 张淑华. 节点与变量: 突发事件网络“扩音效应”产生的过程考察和一般模式 [J]. *新闻与传播研究* 2016 (7): 60-128.
- [5] 杨 娴. 媒体信息传播与受众恐惧相关性研究 [D]. 南昌: 南昌大学 2009.
- [6] 孙俊才. 情绪社会分享的研究现状与发展趋势 [J]. *心理科学进展* 2007, 15(5): 816-821.
- [7] 李从东, 洪宇翔. 基于社会网络的农村群体负面情绪传播机制 [J]. *计算机应用研究* 2015, 32(1): 85-88.
- [8] 赖安婷. 群体情绪传播途径及其影响因素 [D]. 北京: 首都师范大学 2013.
- [9] 刘志明, 刘 鲁. 面向突发事件的情绪监控预警 [J]. *系统工程*, 2010, 28(7): 66-78.
- [10] 王 雷, 方 平. 基于系统动力学的群体情绪传播模型 [J]. *心理科学* 2014, 37(3): 678-682.
- [11] Smith E R, Henry S. An in-group becomes part of self: Response time evidence [J]. *Personality and Social Psychology Bulletin*, 1996, 22: 635-642.
- [12] Maitner A T, Mackie D M, Smith E R. Evidence for the regulatory function of intergroup emotion: Implementing and impending intergroup behavioral intentions [J]. *Journal of Experimental Social Psychology* 2006, 42: 720-726.
- [13] HE Yan-xiang, SUN Song-tao, NIU Fei-fei, LI Fei. A Deep Learning Model Enhanced with Emotion Semantics for Microblog Sentiment Analysis [J]. *Online Publishing* 2016, 39(144).
- [14] Claster W B, Hung D, Cooper M. N B and unsupervised artificial neural nets for Cancun tourism social media data analysis [C] // *Proceedings of the 2nd World Congress on Nature and Biologically Inspired Computing (NaBIC)*. Piscataway: IEEE 2010: 158-163.
- [15] Tseng C, Patel N, Paranjape H et al. Classifying Twitter data with N B classifier [C] // *Proceedings of 2012 IEEE International Conference on Granular Computing (GrC)*. Piscataway: IEEE, 2012: 294-299.
- [16] Jung J J. Maximum entropy-based named entity recognition method for multiple social networking services [J]. *Journal of Internet Technology* 2012, 13(6): 931-937.
- [17] Escalante H J, Montes-Y-Gomez M, Solorio T. A weighted profile intersection measure for profile-based authorship attribution [C] // *Proceedings of the 10th Mexican International Conference on Advances in Artificial Intelligence*. Heidelberg: Springer-Verlag Berlin 2011: 232-243.
- [18] Jakob N, Gurevych I. Extracting opinion targets in a single and cross domain setting with conditional random fields [C] // *Cambridge, Massachusetts: Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics 2010.
- [19] 陈 锋, 巢文涵, 周 庆 等. 基于卷积核的中文微博情感要素识别 [J]. *计算机科学* 2014, 41(12): 133-138.
- [20] 郝志峰, 杜慎芝, 蔡瑞初 等. 基于全局变量 CRFs 模型的微博情感对象识别方法 [J]. *中文信息学报* 2015, 29(4): 50-59.
- [21] 纪雪梅, 王 芳. SNA 视角下的在线社交网络情感传播研究综述 [J]. *情报理论与实践* 2015, 38(7): 139-144.
- [22] 毕宏音. 网民心理特征分析 [J]. *社科纵横* 2006, 21(9): 38-39.
- [23] 赵妍妍, 秦 兵, 刘 挺. 文本情绪分析 [J]. *软件学报* 2010, 21(8): 1834-1848.
- [24] Zhang Jiaming, Wang Bo, Tang Haohao et al. Unsupervised Sentiment Orientation Analysis on Microblog Based on Biterm Topic Model [J]. *Computer Engineering* 2015, 41(7): 219-223.
- [25] Jichang Zhao, Li Dong, Junjie Wu et al. Moodlens: An Emoticon-Based Sentiment Analysis System for Chinese Tweets [C] // *Proceedings of the Eighteenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)* 2012: 1528-1531.
- [26] 贺飞艳, 何炎祥, 刘楠. 面向微博短文本的细粒度情绪特征抽取方法 [J]. *北京大学学报* 2014, 50(1): 48-54.
- [27] 陆 征一, 王稳地. 生物数学前沿 [M]. 北京: 科学出版社, 2008, 1-2.
- [28] 兰月新, 曾润喜. 突发事件网络舆情传播规律与预警阶段研究 [J]. *情报杂志* 2013, 32(5): 16-19.
- [29] 蒋启源, 谢金星, 叶俊. 数学模型 [M]. 北京: 高等教育出版社, 2011, 1.
- [30] 王进良, 张令元. 三种竞争种群的古典 Gause-Lotka-Volterra 系统 [J]. *生物数学学报* 1995, 10(4): 164-173.

(责编: 王平军; 校对: 白燕琼)

(上接第58页)

- [13] Allan J, Papka R, Lavrenko V. On-line new event detection and tracking [A]. *Proceedings of the 21st annual international ACM SIGIR conference on Research and development in information retrieval* [C]. New York: ACM Press 1998: 37-45.
- [14] Yang Y, Zhang J, Carbonell J et al. Topic-conditioned novelty detection [A]. *Proceeding of the 8th ACM SIGKDD International Conference* [C]. New York: ACM 2002: 688-693.
- [15] 贾自艳, 何 清, 张俊海 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(7): 1273-1280.
- [16] 唐晓波, 王中勤, 钟林霞. 基于维基语义扩展的微博话题追踪模型研究 [J]. *情报科学* 2017(2): 80-85.
- [17] 吴小兰, 章成志. 基于突发事件特征网络的用户社区发现与社区主题演化研究——以新浪微博 H7N9 事件为例 [J]. *情报理论与实践* 2017(5): 94-98.
- [18] 吴树芳, 徐建民. 动态话题追踪中的误报检测 [J]. *计算机应用研究* 2015(2): 547-551.
- [19] Salton G, Yang C S. On the specification of term values in automatic indexing [J]. *Journal of Documentation* 1973(4): 351-372.

(责编: 王 菊; 校对: 贺小利)