

## 基于改进人口模型的微博话题趋势预测

何炎祥<sup>1,2</sup>, 刘健博<sup>1,2</sup>, 刘楠<sup>3</sup>, 彭敏<sup>1,2</sup>, 陈强<sup>1,2</sup>, 何静<sup>4</sup>

(1. 武汉大学 计算机学院, 湖北 武汉 430072; 2. 武汉大学 软件工程国家重点实验室, 湖北 武汉 430072;  
3. 军事经济学院 军需系, 湖北 武汉 430035; 4. 肯尼索州立大学, 美国 30144)

**摘要:** 趋势预测问题是当前社会计算中的研究热点, 但微博平台的话题趋势预测还处在探索阶段。在分析、抽取和定义事件趋势的影响因素的基础上, 改进了人口模型使其适用于微博趋势预测, 然后将改进的人口模型映射在神经网络上, 并利用遗传优化的神经网络对事件的趋势进行预测。实验证明, 该方法对于预测网络中长期酝酿的事件发展趋势效果明显, 能够有效地预测事件的爆发点和发帖量, 而且适用于小样本预测问题。

**关键词:** 社会计算; 趋势预测; 神经网络; 人口模型; 社交网络

**中图分类号:** TP399

**文献标识码:** A

## Based on improved malthusian model microblogging topic trend forecast

HE Yan-xiang<sup>1,2</sup>, LIU Jian-bo<sup>1,2</sup>, LIU Nan<sup>3</sup>, PENG Min<sup>1,2</sup>, CHEN Qiang<sup>1,2</sup>, HE Jing<sup>4</sup>

(1. Computer School, Wuhan University, Wuhan 430072, China;  
2. State Key Lab of Software Engineering, Wuhan University, Wuhan 430072, China;  
3. Department of Quartermaster, Military Economic Academy, Wuhan 430035, China;  
4. Kennesaw State University, USA 30144)

**Abstract:** Although trend forecasting problem is the research focus of social computing, yet the topic trend forecast of micro logging platform is still at the exploratory stage. Based on the analyzing, extracting and defining the influencing factors of the trends of event, the improved Malthus model, which is more applicable to solve microblogging topic trend forecast problem is presented. Then the model is mapped to the neural network, and the genetic algorithms are used to predict the trend of events. Experiments show that this method can predict the trend of long-term events in the social network with good results, and can predict the outbreak point and posting volume effectively. It is worth nothing that the method is also suitable for small sample forecasting problems.

**Key words:** social computing; trend forecasting; neural network; malthus model; social networks

### 1 引言

微博, 即微型博客 (microblog), 是一个基于用户关系信息的分享、传播以及获取平台。根据《2011 上半年中国网络舆情指数报告》, 微博已成为仅次于新闻媒体报道的舆论源头, 2011 年上半年 80 个热点事件中, 有 15 个事件是在微博中首发, 如郭美美事件、故宫失窃案、随手拍解救乞讨儿童

行动等。微博以其碎片化、便捷性的特点逐渐成为热点事件的发源地, 同时也是推动事件发展的重要力量。新浪公司推出的新浪微博是极具影响力的微博平台, 拥有超过 5 亿的注册用户和超过 30 万的认证用户, 其中有 13 万多家企业与机构账户, 后文所提到的微博均代表新浪微博。

微博已经发展成长为中国最重要的网络舆论场, 热门信息在极短时间内被大量的评论和转发, 例如

收稿日期: 2014-01-16; 修回日期: 2014-07-08

基金项目: 国家自然科学基金资助项目 (61472290, 61472291); 国家自然科学基金青年基金资助项目 (61303115); 武汉市科技攻关基金资助项目 (201210421135)

**Foundation Items:** The National Natural Science Foundation of China(61472290,61472291);The National Science Foundation for Young Scholars of China(61303115); Key Technologies R&D Program of Wuhan(201210421135)

2011年“新浪体育”官方微博第一时间发布“李娜法网夺冠”的微博在10 min之内获得了5 000多次转发, 夺冠5 min之内祝贺微博数量超过30余万条, 平均每秒超过1 000条。受小世界效应<sup>[1]</sup>的影响, 热门话题在短时间内大范围扩散, 话题涉及各个领域, 话题相关的博文不仅包含了事件的具体描述, 而且还包括网民对该话题的观点和立场。因此尽早地预测微博上话题的走向和趋势, 不仅可以帮助政府部门了解社情民意并及时处理事件提供参考, 而且还可以为商业领域广告投放和危机公关提供帮助。

趋势预测在很多行业都有研究和应用, 如医疗、气象、环境等领域<sup>[2-4]</sup>, 计算机领域也有一些相关研究<sup>[5-7]</sup>, 但是基于互联网的趋势预测研究还处于探索阶段, 如文献[8]利用Google - Profile of Mood States算法可以提前2~6天预测道琼斯指数的波动, 文献[9]通过一段时间内用户对电影的评论来预测电影票房。

人类社会, 人口数量是由人类自身的生育率与外界环境共同决定的。对于微博平台, 网民对话题的关注程度以及网络环境决定了话题的发帖量。人口数量与话题的发帖量的增长模式类似, 都是呈指数级增长的。因此, 本文改进了用于人口预测的Malthus人口模型, 使其适用于微博话题趋势预测, 同时考虑微博的参与人次、微博的影响力、微博增长率、微博参与度等影响微博发帖量增长的因素, 最后将改进的人口模型映射到神经网络上实现话题相关的微博趋势预测。

## 2 影响因素分析与定义

本文首先将实验数据集覆盖的时间段划分为若干个相等的单位时间片, 每个时间片的时间长度用 $\Delta T$ 表示。本文采用形式化的方法对第 $t$ 个时间片, 某一话题相关的微博内容 $MB(t)$ 定义为六元组,

$$MB(t) = (I_{mb}, I_u, N_{rt}, N_{cm}, N_{influence}, C) \quad (1)$$

其中,  $I_{mb}$ 表示微博内容的唯一标识的集合,  $I_u$ 表示微博内容发布用户的唯一标识的集合,  $N_{rt}$ 表示微博内容被转发的次数,  $N_{cm}$ 表示微博内容被评论的次数的集合,  $N_{influence}$ 表示微博内容的影响力的集合,  $C$ 表示微博文本的集合。因此这一话题的某一条微博内容可以定义为

$$mb(t) = (i_{mb}, i_u, n_{rt}, n_{cm}, n_{influence}, c) \quad (2)$$

其中,  $i_{mb} \in I_{mb}$ ,  $i_u \in I_u$ ,  $n_{rt} \in N_{rt}$ ,  $n_{cm} \in N_{cm}$ ,

$$n_{influence} \in N_{influence}, c \in C。$$

微博用户对于事件的关注程度直接的体现就是微博发帖的数量, 话题相关的微博数量在单位时间内越多就说明话题越热门。影响话题相关的微博发帖数量的因素很多, 在本文中考虑以下几个因素: 微博的参与人次、微博的影响力、微博增长率、微博参与度。

本文用 $N_{mb}(t)$ 表示第 $t$ 个时间片内微博趋势的微博发帖量。

微博的参与人次是指在第 $t$ 个时间片内所有用户参与微博话题的次数, 其值与微博的评论转发数量相同, 值越大表示话题的受关注的程度也就越大, 用 $N_{rtcm}(t)$ 表示。

$$N_{rtcm}(t) = \sum_{n_{rt} \in N_{rt}, n_{cm} \in N_{cm}} (MB(t)n_{rt} + MB(t)n_{cm}) \quad (3)$$

微博的影响力由发微博用户的粉丝数量来反映, 一条微博的影响力值为发布该条微博的用户的粉丝数量, 因此在第 $t$ 个时间片内的微博影响力 $N_{influence}(t)$ 为所有发布微博用户粉丝数量之和, 即

$$N_{influence}(t) = \sum_{n_{influence} \in N_{influence}} (MB(t)n_{influence}) \quad (4)$$

微博增长率反映当前时间段微博发帖数量和评论转发数量之和相对于前一时间段的增长情况, 在第 $t$ 个时间片内微博增长率 $N_{growth}(t)$ 为

$$N_{growth}(t) = \frac{(N_m b(t) + N_{rtcm}(t)) - (N_m b(t-1) + N_{rtcm}(t-1))}{N_m b(t-1) + N_{rtcm}(t-1)} \quad (5)$$

微博参与度是指当前时间段微博数量和微博参与人次的比值, 其值越大表示微博的参与话题的用户更倾向于发表新的微博, 其值越小表示参与话题的用户更倾向于参与微博讨论, 在第 $t$ 个时间片内微博参与度为

$$N_{rate}(t) = \frac{N_{mb}(t)}{N_{rtcm}(t)} \quad (6)$$

因此反映第 $t$ 个时间片内微博趋势的微博发帖量 $N_{mb}(t)$ 可以定义为

$$N_{mb}(t) = f(N_{mb}(t-1), N_{rtcm}(t-1), N_{influence}(t-1), N_{growth}(t-1), N_{rate}(t-1)) \quad (7)$$

其中,  $N_{mb}(t) = |MB(t)I_{mb}|。$

### 3 趋势预测方法

#### 3.1 改进的 Malthus 模型

1798 年，英国人口学家 Malthus 在分析了人口出生和死亡情况的数据后，发现人口增长呈指数变化规律，增长率基本是一个常数，因此他提出了一种人口增长模型<sup>[4]</sup>，即

$$\frac{dN}{dt} = rN \quad (8)$$

其中， $N$  表示  $t$  时刻的人口总量， $r$  是自然增长率。Malthus 模型只是一种增长模型，没有考虑自然环境的有限性。人口由于受物理环境的影响不会无限制的增长，随着人口的增加，自然条件的恶化，人与人之间出现了竞争，这种情况下该模型的预测结果会出现较大误差。

在微博平台中，只要有足够的存储介质，微博发帖数量的增长就不会受到物理环境的限制，而且微博发帖量的增长模式和人口增长模式类似，都是随着时间呈指数增长的，所以可以将 Malthus 模型用于解决微博趋势预测的问题。本文针对微博趋势预测问题对 Malthus 模型进行了改进，即

$$\frac{dQ}{dt} = r(Q + U) \quad (9)$$

$$U = a_1x_1(t) + a_2x_2(t) + \dots + a_nx_n(t) \quad (10)$$

其中， $Q$  表示  $t$  时刻的微博发帖总数量， $U$  表示  $t$  时刻的影响基数， $x_1(t), \dots, x_n(t)$  表示  $t$  时刻影响因素的值， $a_1, \dots, a_n$  表示影响系数，此模型中  $Q$  和  $x_1, \dots, x_n$  都需要做归一化处理。求解式(9)可得

$$Q(t) = (Q_0 + U)e^{r(t-1)} - U \quad (11)$$

其中， $Q_0$  为初始时刻的微博数量。因此时间段  $t$  到  $t + \Delta T$  内微博发帖数量为

$$N_{mb}(t + \Delta T) = Q(t + \Delta T) - Q(t) \quad (12)$$

#### 3.2 模型映射

根据式(11)可以得到第  $t + 1$  时间段内  $Q(t + 1)$  为

$$Q(t + 1) = (Q_0 + U)e^{rt} - U = \left( (Q_0 + U) - Q_0 \frac{1}{1 + e^{rt}} - 2U \frac{1}{1 + e^{rt}} \right) (1 + e^{rt}) \quad (13)$$

为了确定参数  $r, a_1, \dots, a_n$  的值，将式(10)、式(13)映射到一个如图 1 所示的 BP 神经网络中。BP 神经网络是一种按误差逆传播算法训练的多层前馈网

络<sup>[11]</sup>，是目前应用最为广泛的神经网络模型之一，数学理论已经证明 BP 神经网络可以实现任何复杂非线性映射的功能<sup>[11]</sup>，BP 神经网络特别适合求解内部机制复杂的问题。因此 BP 神经网络学习的过程就是对各连接权值的动态调整的过程。

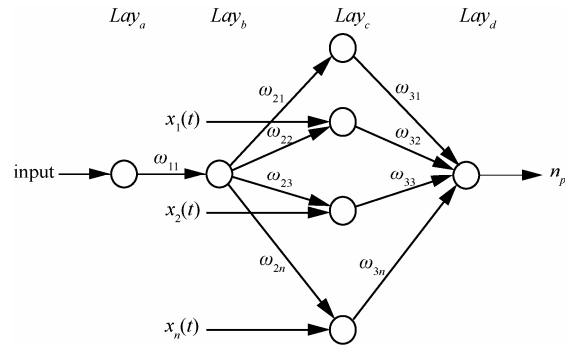


图 1 神经网络结构

$Lay_a, \dots, Lay_d$  表示神经网络的 4 层， $x_1(t), \dots, x_n(t)$  为神经网络的输入参数， $\omega_{ij}$  为网络权值， $n_p$  为网络预测值，其中  $\omega_{11} = -r$ ， $\omega_{21} = -Q_0$ ， $\omega_{22} = -a_1$ ， $\omega_{23} = -a_2$ ， $\dots$ ， $\omega_{2m} = -a_n$ ， $\omega_{31} = \omega_{32} = \dots = \omega_{3m} = 1 + e^{rt}$ 。因此，该 BP 神经网络的学习流程如下。

- 1) 初始化参数  $r, a_1, \dots, a_n$ 。
- 2) 根据初始化参数计算每层输出，即有  $Lay_a$  层的输出为

$$output_a = \omega_{11}t \quad (14)$$

$Lay_b$  层选取 Sigmoid 函数作为激励函数，则输出为

$$output_b = \frac{1}{1 + e^{-\omega_{11}t}} \quad (15)$$

$Lay_c$  层输出为

$$\left\{ \begin{array}{l} output_{c1} = output_b \omega_{21} \\ output_{c2} = 2output_b x_1(t) \omega_{21} \\ \dots \\ output_{cm} = 2output_b x_n(t) \omega_{2m} \end{array} \right\} \quad (16)$$

$Lay_d$  层的输出为

$$output_d = \omega_{31}output_{c1} + \omega_{32}output_{c2} + \dots + \omega_{3m}output_{cm} - \theta \quad (17)$$

其中， $\theta = -(Q_0 + U)(1 + e^{rt})$  为  $Lay_d$  层的输出阈值。

- 3) 计算各层误差。其中， $Lay_d$  层误差为

$$\delta_d = -U - n_p \quad (18)$$

Lay<sub>c</sub> 层误差为

$$\delta_{c1} = \delta_{c2} = \dots = \delta_{cm} = \delta_d (1 + e^{-\alpha_1 t}) \quad (19)$$

Lay<sub>b</sub> 层误差为

$$\delta_b = \frac{1}{1 + e^{-\alpha_1 t}} \left( 1 - \frac{1}{1 + e^{-\alpha_1 t}} \right) (\omega_{21} \delta_{c1} + \dots + \omega_{2m} \delta_{cm}) \quad (20)$$

4) 根据误差调整权值和阈值。调整 Lay<sub>a</sub> 到 Lay<sub>b</sub> 的连接权值

$$\omega_{11} = \omega_{11} - rt\delta_b \quad (21)$$

调整 Lay<sub>b</sub> 到 Lay<sub>c</sub> 的连接权值

$$\left\{ \begin{array}{l} \omega_{21} = -Q_0 \\ \omega_{22} = \omega_{22} - \mu_1 \delta_{c2} output_b \\ \dots \\ \omega_{2m} = \omega_{2m} - \mu_n \delta_{cm} output_b \end{array} \right\} \quad (22)$$

其中,  $\mu_1, \dots, \mu_n$  为学习速率。调整阈值

$$\theta = -(Q_0 + \omega_{22} x_1(t) + \dots + \omega_{2m} x_n(t))(1 + e^{-\alpha_1 t}) \quad (23)$$

5) 判断学习是否结束, 如果没有结束, 那么继续执行第 3) 步。

### 3.3 参数初始化方法

由于 BP 神经网络权值和阈值是随机初始化的, 最终的结果很容易就陷入局部最优, 因此为了提高预测精度, 本文使用遗传算法来优化参数  $r, \alpha_1, \dots, \alpha_n$ 。

遗传算法是模仿自然界生物进化机制发展起来的随机全局搜索和优化方法<sup>[12]</sup>, 它借鉴了达尔文的进化论和孟德尔的遗传学说。其本质是一种高效、并行、全局搜索的方法, 它能在搜索过程中自动获取和积累有关搜索空间的知识, 并自适应的控制搜索过程以求得最优解。遗传算法操作使用适者生存的原则, 在潜在的解决方案种群中逐次产生一个近似最优解的方案, 在遗传算法的每一代中, 根据个

体在问题域中的适应度值和从自然遗传学中借鉴来的再造方法进行个体选择, 产生一个新的近似解。这个过程导致种群中个体的进化, 得到的新个体比原来个体更能适应环境, 就像自然界中的改造一样。

算法流程如图 2 所示, 本文所有实验的初始种群数为 50。

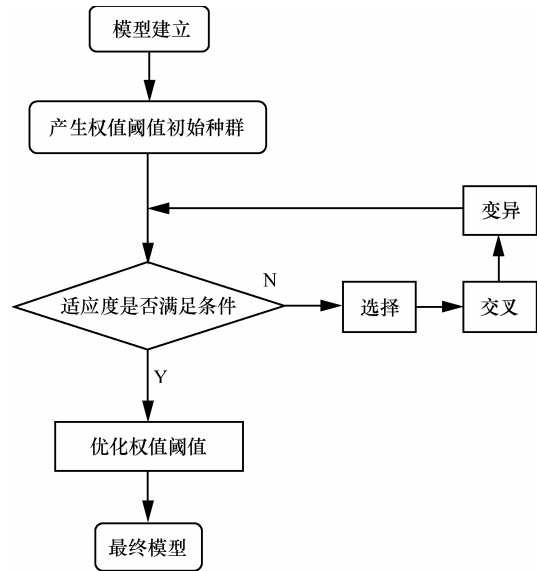


图 2 遗传算法流程

## 4 实验

### 4.1 实验数据

本文爬取了新浪微博 2013 年上半年 7 个热门话题共计 35 万条微博, 微博数据包含的字段有: 微博 id、微博用户名、微博发布时间、微博的评论数量、微博的转发数量、微博发布者粉丝数以及微博内容等, 表 1 为各话题的详细信息, 图 3 是实验数据中 7 个话题单位时间 (单位时间为一天) 发帖量的折线图, 横坐标为日期, 纵坐标为发帖量。各个话题的发帖量反映了事件的实际发展趋势。

表 1 实验数据的详细信息

话题编号	话题名称	开始时间	结束时间	持续时间/天	发帖量	爆发时间
1	2013 年两会	3 月 2 日	3 月 14 日	12	73 310	3 月 6 日
2	H7N9 流感	3 月 31 日	4 月 13 日	13	78 813	4 月 8 日
3	北京雾霾天气	1 月 5 日	1 月 20 日	15	32 600	1 月 17 日
4	国五条颁布	3 月 1 日	3 月 30 日	29	26 509	3 月 29 日
5	黄浦江死猪事件	3 月 5 日	3 月 23 日	18	19 279	3 月 13 日
6	撒切尔夫人去世	4 月 8 日	4 月 21 日	20	20 418	4 月 9 日
7	雅安地震	4 月 20 日	5 月 4 日	15	101 967	4 月 24 日

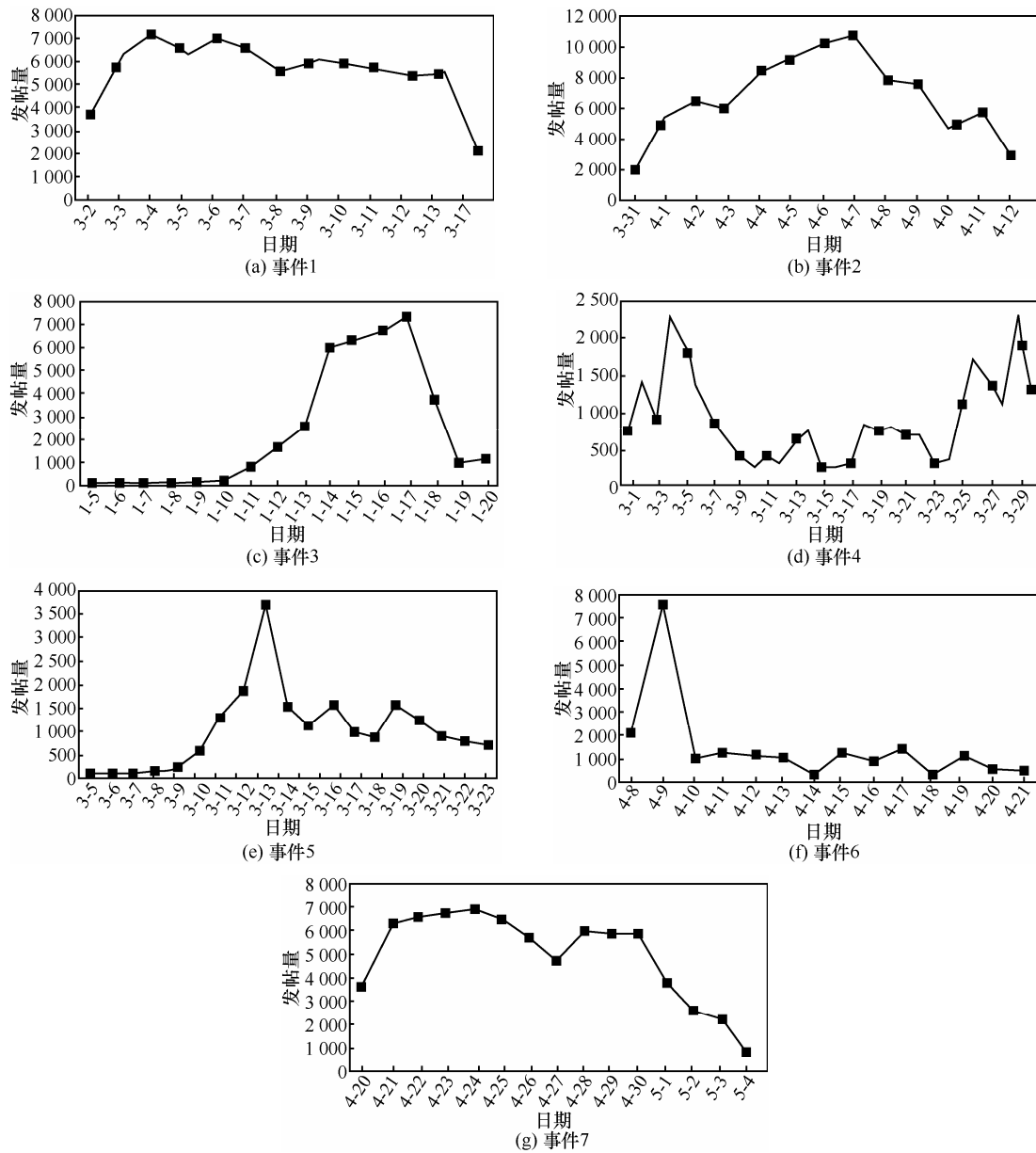


图 3 发帖量折线图

根据表 1 和图 3 可以看出，各个事件的时间跨度不同，在实际解决话题预测问题的时候，重点需要预测话题的爆发时间点即发帖量的峰值点，爆发点之前和之后发帖量的上下波动的预测不是本文讨论的重点，因此本文选取事件包括爆发点当天以及爆发点前后共计 5 天的发帖量作为测试数据，实验结果以这 5 天的预测准确度作为评价标准。以事件 1 为例，话题的峰值出现在 3 月 6 日，因此选取 3 月 5 日到 3 月 9 日的的数据为测试集，3 月 2 日到 3 月 4 日的的数据为训练集。各话题的训练集和测试集情况如表 2 所示。

表 2 训练集和测试集分布

话题编号	训练集时间段	测试集时间段
1	3 月 2 日~3 月 4 日	3 月 5 日~3 月 9 日
2	3 月 31 日~4 月 5 日	4 月 6 日~4 月 10 日
3	1 月 5 日~1 月 14 日	1 月 15 日~1 月 19 日
4	3 月 1 日~3 月 25 日	3 月 26 日~3 月 30 日
5	3 月 5 日~3 月 10 日	3 月 11 日~3 月 15 日
6	4 月 1 日~4 月 1 日	4 月 2 日~4 月 6 日
7	4 月 20 日~4 月 22 日	4 月 23 日~4 月 27 日

### 4.2 实验分析

实验机器的系统为 UBUNTU12.04，通过 MATLAB 实现第 3 节内容，通过实验发现当遗传算法初始种群数为 50 时，各组数据在 100 代以内都可以得到理想结果。

支持向量机的回归模型<sup>[13]</sup>是一种应用十分广

泛的预测模型，图 4 是本文方法与支持向量机的回归模型的预测结果对比。7 个热门事件中，除了事件 6 以外其他事件都在网络中长期的酝酿和发展才爆发，事件 6 的发生和爆发几乎是同时发生，导致了训练集不足，所以 2 种方法都未能得到结果。根据图 4 的预测结果表明，虽然支持向量机的回归模

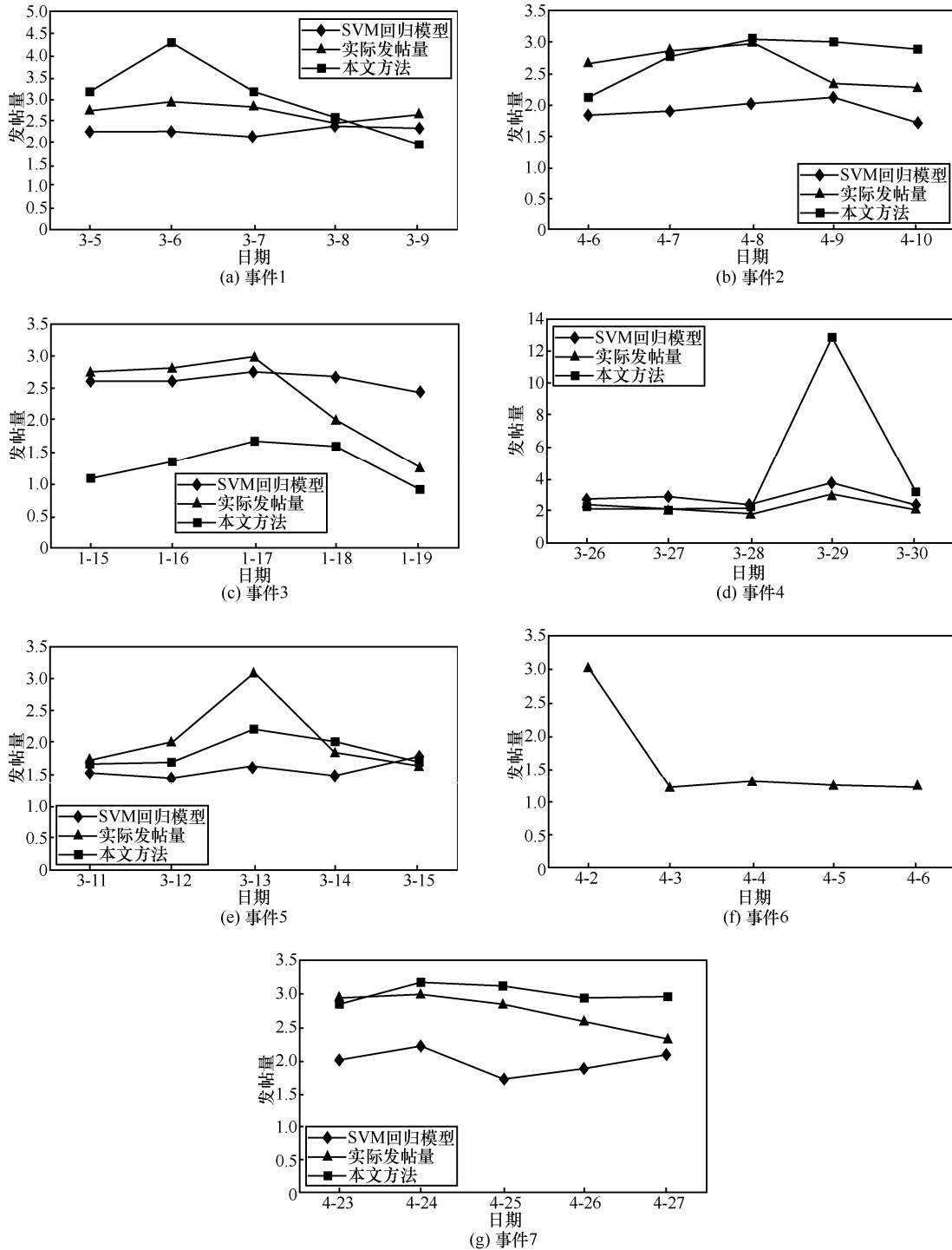


图4 预测结果

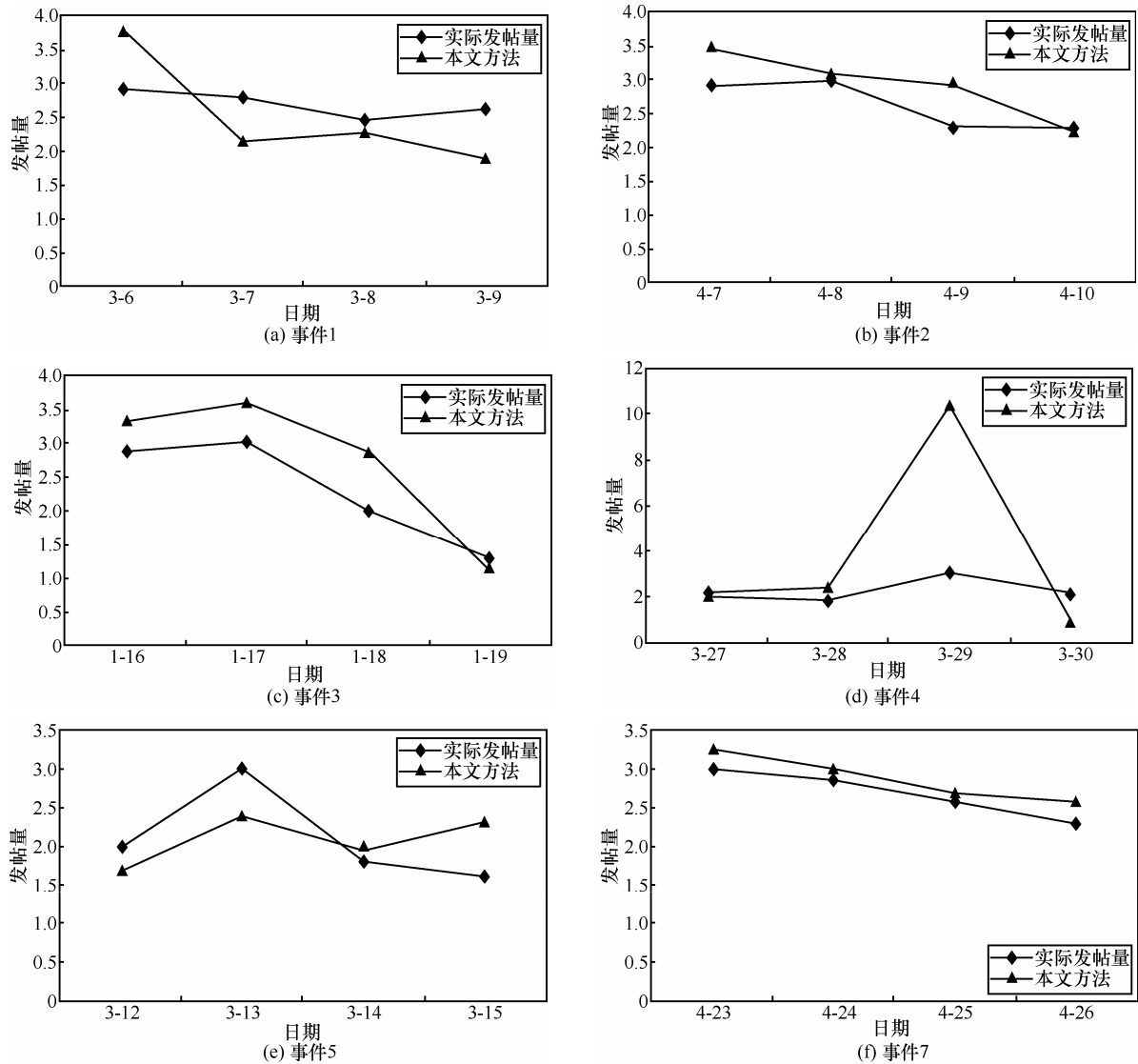


图 5 优化的预测结果

型在很多领域都有应用，但是在微博话题趋势预测的应用场景下由于训练集规模太小，所以准确率普遍较低。

而本文的方法是根据微博话题趋势增长的特征提出的，在训练集规模很小的环境下，本文的预测方法能够反映事件的增长情况，而且准确地预测了 7 个热门事件中 6 个事件爆发点出现的时间，但是预测结果的数值准确度不高，因此以预测结果的增长率来修正预测结果的数值以提高预测准确率，结果如图 5 所示。为更直观反映实验结果，图 4 和图 5 中都对实际发帖的数值进行了归一化处理，归一区间为[0,3]，预测发帖量是基于归一化数据的预测结果。

根据图 5 的修正结果，各个事件预测发帖量的误差率如表 3 所示，修正后的预测结果比较准确的

预测了实际的发帖量，事件 4 平均误差率较高是由于 3 月 29 日当天的实际发帖量并未像预测的那样剧烈增长，但预测结果正确地反映了事件的实际发展趋势。

表 3 预测结果的平均误差率

事件编号	平均误差率
1	0.215
2	0.139
3	0.211
4	0.864
5	0.149
6	—
7	0.075

因此,在训练数据有限的情况下,本文提出的方法对于预测微博平台上长期酝酿的事件发展趋势效果明显,能够有效地预测事件的爆发点和发帖量。

此外,在时间片长度更短,训练数据足够的情况下,本文的方法也适用于突发事件的预测,但是预测效果还有待验证。

## 5 结束语

本文针对微博社交平台提出了一种事件趋势预测的方法,这种方法结合了人口论和神经网络相关领域的内容,对事件发展的趋势进行量化的分析和建模,并利用实际的数据对该方法进行了验证,为社交网络平台的趋势预测问题提供了一种新的思路。在以后的研究工作中,会考虑更多影响事件发展的因素,发现微博文本对事件发展趋势的影响,从而提高预测准确率。

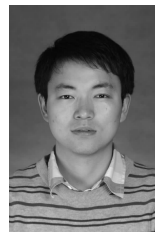
### 参考文献:

- [1] Six degrees of separation[EB/OL]. <http://en.wikipedia.org/wiki/Sixdegreesofseparation>, 2015.
- [2] 倪允琪,周秀骥. 中国长江中下游梅雨锋暴雨形成机理以及监测与预测理论和方法研究[J]. 气象学报, 2004, 62(5): 647-662.  
NI Y Q, ZHOU X J. Study for formation mechanism of heavy rainfall within the meiyu front along the middle and downstream of yangtze river and theories and methods of their detection and prediction[J]. Acta Meteorologica Sinica, 2004, 62(5):647-662.
- [3] 张志华,陈乾安,吕丽红等. 浙江东南沿海地区幽门螺杆菌的耐药趋势预测[J]. 中华消化内镜杂志, 2013,30(6):328-331.  
ZHANG Z H, CHEN Q A, LV L H, et al. Prediction of drug resistance trend of helicobacter pylori in southeast coastal area of Zhejiang province[J]. Chinese Journal of Digestive Endoscopy, 2013,30(6): 328-331.
- [4] 陈淑燕,王炜. 交通量的灰色神经网络预测方法[J]. 东南大学学报:自然科学版, 2004, 34(4): 541-544.  
CHEN S Y, WANG W. Grey neural network forecasting for traffic flow[J]. Journal of Southeast University(Natural Science Edition), 2004, 34(4):541-544.
- [5] 温祥西,孟相如,马志强等. 小时间尺度网络流量混沌性分析及趋势预测[J]. 电子学报, 2012, 40(8):1609-1616.  
WEN X X, MENG X R, MA Z Q, et al. The chaotic analysis and trend prediction on small-time scale network traffic[J]. Acta Electronica Sinica, 2012, 40(8):1609-1616.
- [6] 钱爱玲,瞿彬彬,卢炎生等. 多时间序列关联规则分析的论坛舆情趋势预测[J]. 南京航空航天大学学报, 2013, 44(6): 904-910.  
QIAN A L, QU B B, LU Y S, et al. Forum sentiment trend prediction based on multi time series association rule analysis[J]. Journal of Nanjing University of Aeronautics and Astronautics, 2013, 44(6): 904-910.
- [7] 高辉,王沙沙,傅彦. Web 舆情的长期趋势预测方法[J]. 电子科技大学学报, 2011, 40(3): 440-445.  
GAO H, WANG S S, FU Y. Prediction model for long-term development trend of Web sentiment[J]. Journal of University of Electronic Science And Technology of China, 2011, 40(3): 440-445.
- [8] JOHAN B, MAO H N, ZENG X J. Twitter mood predicts the stock market[J]. Journal of Computational Science, 2011,2(1):1-8.
- [9] SITARAM A, BERNARDO A H, GABOR S W. Trends in social media: persistence and decay[A]. ICWSM 2011[C]. 2011.8.
- [10] MALTHUS T R. 人口论[M]. 北京: 北京大学出版社, 2008.  
MALTHUS T R. Population Theory (hardcover)[M]. Beijing: Peking University Press,2008.
- [11] DAVID E R, G E H, RONALD J W. Learning representations by back-propagating errors[J]. Nature, 1986, 323(6088):533-536.
- [12] RICCARDO P, WILLIAM B L, NICHOLAS F M, et al. A field guide to genetic programming[EB/OL]. www. Lulu. com, 2008.
- [13] KLAUS-ROBERT M, ALEX J, SMOLA G R, et al. Predicting time series with support vector machines[A]. ICANN[C]. 1997. 999-1004.

### 作者简介:



何炎祥(1952-),男,湖北应城人,武汉大学教授、博士生导师,主要研究方向为可信软件、自然语言处理、并行分布处理和软件工程。



刘健博(1986-),男,湖北黄石人,武汉大学博士生,主要研究方向为社交网络、自然语言处理、人工智能。

刘楠(1983-),男,河南信阳人,军事经济学院讲师,主要研究方向为自然语言处理、社交网络和软件工程。

彭敏(1973-),女,青海西宁人,武汉大学教授、博士生导师,主要研究方向为网络服务、自然语言处理、信息检索。

陈强(1988-),男,安徽马鞍山人,武汉大学博士生,主要研究方向为自然语言处理和社交网络。

何静(1977-),女,湖北武汉人,美国肯尼索州立大学助理教授,主要研究方向为有无线网络、人工智能和社会网络。