

面向时序数据的多维度网络舆情演化分析研究

李旻¹, 王志华², 李大宇¹, 赵鑫³, 詹雅慧¹, 王素格²

1. 山西财经大学金融学院, 山西 太原 030006;

2. 山西大学计算机与信息技术学院 (大数据学院), 山西 太原 030006;

3. 蜜度科技股份有限公司上海浦东微热点大数据研究院, 上海 201204

摘要

针对当前网络舆情演化研究存在的视角单一、主题挖掘不全、情感分析不深、群体行为发现不准等局限, 提出了一项多维度网络舆情演化分析框架。首先, 依据社会影响力, 划分了舆情演化的4个生命周期阶段。其次, 基于BERTopic模型和相似性计算分析了主题演化过程, 利用大语言模型分析了情感波动, 基于交互关系划分了用户群体并识别了意见领袖。最后, 对不同周期和地域的社会影响力与情感差异进行了可视化呈现。以时序舆情数据中“印花税调整”事件为例, 研究发现在舆情的4个生命周期, 民众关注的话题、情感倾向、用户群体以及意见领袖均发生了变化, 东部地区对“印花税调整”事件的情感反应更积极, 且舆情主题和情感持续时间更长。该研究可为揭示舆情演化规律、实施有效舆情监控提供技术支持。

关键词

时序数据; 网络舆情; 多维度演化; 主题演化; 情感波动; 用户群体

中图分类号: N99; TP391

文献标志码: A

doi: 10.11959/j.issn.2096-0271.2026012

Research on online public opinion evolution analysis of multi-dimensional for time series data

Li Yang¹, Wang Zhihua², Li Dayu¹, Zhao Xin³, Zhan Yahui¹, Wang Suge²

1. School of Finance, Shanxi University of Finance and Economics, Taiyuan 030006, China

2. School of Computer and Information Technology (School of Big Data), Shanxi University, Taiyuan 030006, China

3. Micro Hotspot Big Data Research Institute, Midu Technology Co., Ltd., Shanghai 201204, China

Abstract

A multi-dimensional framework for analyzing the evolution of online public opinion was proposed, aiming to address the limitations in current research, such as a single perspective, incomplete topic mining, shallow sentiment analysis, and inaccurate identification of group behavior. Firstly, the four life cycle stages of public opinion evolution were divided based on social influence. Secondly, the topic evolution process was analyzed using the BERTopic model and similarity calculation. Sentiment fluctuations were studied by leveraging large language models. User groups were classified and opinion leaders were identified based on interaction relationships. Finally, the differences in social influence and sentiment across different periods and regions were visualized. Taking the event of "adjustment of stamp

duty" in temporal public opinion data as an example, the study discovered that the topics of public concern, sentiment tendencies, user groups, and opinion leaders changed during the four life cycle stages of public opinion. The sentiment response in eastern regions was more positive, and the duration of public opinion topics and sentiments was longer. This study can provide technical support for revealing the laws of public opinion evolution and implementing effective public opinion monitoring.

Key words

time-series data, online public opinion, multidimensional evolution, topic evolution, sentiment fluctuation, user group

0 引言

随着媒体的迅猛发展，微博数据中蕴含着大量公众较为关注的话题，特别是与人们息息相关的民生问题。例如，一项经济政策从发布到具体的落地实施，经媒体报道和传播后，可能在短时间内引起公众的广泛关注和讨论，公众会在网络平台上围绕该话题发表个人情感态度和意见观点，并通过转发和@他人等行为进行传播扩散，这极易引发网络舆情。网络舆情是以网络为载体、以事件主题为核心，广大公众对情感、态度、意见、观点进行表达、传播与互动，以及由此产生的后续影响力的集合^[1]。与传统舆情相比，网络舆情的传播速度更快，影响范围更广，引发的后果更严重。特别是对一些经济政策的热议，如果不能采取有效的措施加以引导控制，将会导致舆情爆发，进而使人们对经济失去信心，引发恐慌，影响社会稳定。对于演化研究，早在1986年，Fink^[2]就提出了危机生命周期理论，将危机分为4个阶段，依次为潜伏期、爆发期、蔓延期、衰退期。现有的网络舆情演化分析采用社会影响力、舆情主题、舆情热度，以及负面情绪和负面舆论等因素^[3-6]，但缺乏从主题、情感、社会影响力、生命周期、用户群体行为和地域差异等多个维度展开的研究。同

时，现有的生命周期划分方法多依赖于信息发布数、搜索指数等单一或双重标准，未能实现多维指标的整合^[7]。本文基于蜜度科技股份有限公司上海浦东微热点大数据研究院提供的第六届传播数据挖掘竞赛数据集，以A股市场“印花税调整”事件为例，针对前述问题展开分析。首先，从舆情的社会影响力出发，从舆情传播的速度、广度和强度等多个维度划分舆情演化的生命周期；随后，采用BERTopic工具挖掘公众关注的主题，并利用基于提示学习的大语言模型（简称大模型）生成主题名称；进而，基于大模型分析不同周期内公众的情感，并通过设计用户行为的传播作用机制，识别关键用户群体与意见领袖。最后，本文通过分析该事件在不同周期与地域下的社会影响力与情感演变，为相关部门有效监测舆情趋势、制定精准管控政策提供支持。

1 相关研究现状

1.1 舆情主题与情感演化研究

对于舆情主题演化，研究者们关注了主题信息的获取和分析。2024年，杨洋洋^[4]运用潜在狄利克雷分布（latent Dirichlet allocation, LDA）模型对不同风险的4类企业突发舆情的主题进行分析，并进行主题内容可视化；高春玲等^[5]采用BERTopic模

型对有关老年人健康的微博数据进行主题聚类分析，聚类效果明显。2025年，Zhang等^[6]从情感分析和主题分析两个维度给出了微博数据的特征，发现灾难发生的时期不同，网络舆情讨论的焦点随情感的变化也有所不同；许根玮等^[8]基于电信网络诈骗案件数据，构建时间连续性潜在狄利克雷分布（temporal continuity latent Dirichlet allocation, TC-LDA）模型，挖掘电信网络诈骗案件的主题特征并揭示其演化趋势。以上这些方法均未结合网络舆情的生命周期考虑主题的变化。舆情情感演化的主要任务是对舆情文本进行情感识别、对在线评论进行情感判断等。

对于情感分析研究方法，2022年，李爱黎等^[9]通过将情感信息融入基于Transformer的双向编码器表征（bidirectional encoder representations from transformers, BERT）模型，有效提升了网络舆情情感分析的性能。2024年，韩小伟等^[10]对“重庆星巴克驱赶民警”事件情感演化进行研究，并构建了突发公共事件网络舆情引导模型；Kumar等^[11]运用BERT和卷积神经网络（convolutional neural network, CNN），分析了Twitter平台上与新型冠状病毒感染相关的公众情感。2025年，惠调艳等^[12]基于词典-TextCNN-Word2Vec组合模型对在线评价细粒度情感进行了分析。随着大模型的问世，研究者发现在标注资源有限时，其具有一定的潜力^[13]。

1.2 用户群体行为分析研究

用户在舆情演化过程中扮演着重要的角色，从计算传播学的视角看，用户在舆情传播中同时承担着传播者与受众的双重身份^[14]。为此，需要对用户群体行为进行

分析，发现不同用户群体在舆情传播中的作用差异。2022年，黄仕靖等^[15]通过分析舆情演化时空差异，发现普通网民群体相较于高影响力群体在舆情演化阶段的负面情绪更为严重；林国英等^[16]基于用户交互行为构造了用户交互矩阵，并采用louvain模块度算法（louvain modularity algorithm, LMA）进行了用户社区识别，发现不同社区的用户对同一事件往往具有不同的注重点。2023年，Liu等^[17]利用新浪微博平台上的数据，采用社会网络分析方法，分析网络舆情各阶段的空间和拓扑结构。

1.3 坚持应用需求引导数据开放

关于舆情的地域影响差异，2016年，曹丽娜等^[18]通过分析网络论坛数据的地域分布，发现各省热门话题存在显著差异，这些话题分布也反映了当地的环境特征与现实问题。2022年，边晓慧等^[19]研究了重大突发公共卫生事件，发现公众在社交媒体上的情感表达及演化趋势受时空差异和主题差异等影响，并存在地域扩散的规律。2024年，Xu等^[20]采用LDA主题挖掘方法对舆情事件的主题进行挖掘，并分析了舆情热点事件的传播周期和区域传播模式，揭示了其形成和演化规律。

综合上述研究成果可知，现有研究还未从主题、情感、社会影响力、生命周期、用户群体行为和地域差异等维度进行全面的网络舆情分析。

2 研究方法

本文针对微博时序数据，从影响力、主题、情感、用户群体行为和地域差异

5个维度，构建了多维度网络舆情演化分析研究框架，如图1所示。该框架包括基于影响力的舆情生命周期划分、基于BERTopic模型和相似度计算的主题演化分析、基于提示学习的大模型情感演化分析、基于交互关系的用户群体划分与意见领袖发现、基于舆情主题与情感的地域差异分析这5个部分。

标体系，本文从影响力广度、强度和速度3个维度进行建模，涉及地区覆盖度为RR，传统和网络媒体的覆盖度分别为RS和RW，原创、转发、新增原创、新增转发、发表正负情感的博主数分别为SO、SS、TO、TS、SP、SN，影响力持续时间为 DT_t 。为了将这些不同维度的数据映射到同一空间，对它们进行离散化处理，设定离散值范围为[1,5]，离散化值由基于DeepSeek大模型的专家系统给定。定义第t时刻的影响力广度 R_t 、强度 S_t 、速度 T_t 以及社会影响力 LP_t ，分

2.1 多维度舆情影响力计算

依据文献[3]中的舆情影响力评价指

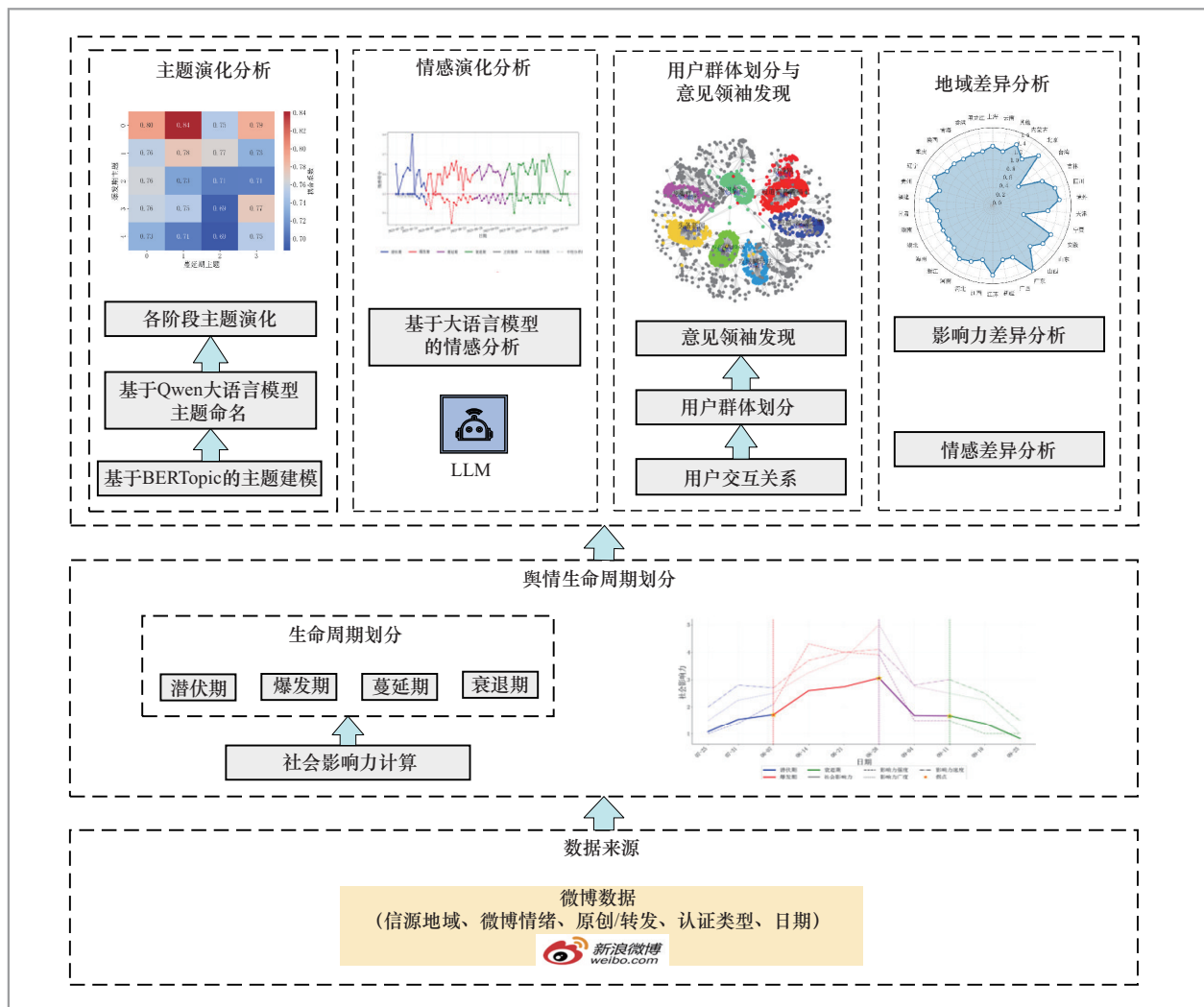


图1 多维度网络舆情演化分析研究框架

别如下所示:

$$R_t = a_1 RR_t + a_2 RS_t + a_3 RW_t \quad (1)$$

$$S_t = b_1 SO_t + b_2 SS_t + b_3 SN_t + b_4 SP_t \quad (2)$$

$$T_t = c_1 TO_t + c_2 TS_t + c_3 DT_t \quad (3)$$

$$LP_t = \alpha R_t + \beta S_t + \gamma T_t \quad (4)$$

为了准确地划分4个生命周期,本文利用式(1)~式(4),以周为单位,分别计算各周第7天与第1天的4个指标的差值,差值绝对值越大,说明4个指标值在本周内变化越大,越有可能作为生命周期的分割点。4个指标的差值表示为:

$$P_t = M_{7t} - M_{1t} \quad (5)$$

其中, M_{1t} 与 M_{7t} 分别表示第 t 周的第1天与第7天的 LP_t 、 R_t 、 S_t 和 T_t 的指标值。

2.2 主题特征获取及其相关性度量

为了分析生命周期中的舆情主题及其相关性,本文利用BERTopic模型获取主题,再根据主题词,利用通义千问大模型QwenLLM生成主题名称。在此基础上,设计Cos相似性计算对主题A和B包含的文档集合的相关性进行分析,其中文档集中的每个文档采用BERT向量化表示。

2.3 基于提示学习的大模型情感分类

假设数据集 $DL = \{d_1, d_2, \dots, d_k\}$, 带标签数据集 $DLS = \{d_1^s, d_2^s, \dots, d_q^s\}$, $DLS \subseteq DL$ 。初始的提示样例 $S^s = \{s_1^s, s_2^s, \dots, s_k^s\}$ 。基于提示学习的大模型的情感分析过程如下。

步骤1 将带标签的数据 d_i^s 和初始提示样例 d_j^s 作为输入,利用通义千问大模型QwenLLM生成提示示例集

$$S = \{s_1, s_2, \dots, s_m\}。$$

步骤2 将生成的提示示例 s_i 与带标签的数据 DLS 进行拼接,作为 ChatGLM-6b 的输入数据,对 ChatGLM-6b 进行微调,利用微调后的 ChatGLM-6b 对未标注的数据进行情感极性评分 $S(d_i)$ 。

步骤3 对于数据标注评分 $S(d_i)$ 进行评估,若 $S(d_i) \geq 0.5$,则人工进行校验并保存;对于评分较低的数据,即 $S(d_i) < 0.5$,重新利用通义千问大模型 QwenLLM 更新提示样例。

步骤4 不断重复步骤2~步骤3,直到对所有的数据进行标注。

通义千问大模型 QwenLLM 的实现如下所示:

$$S_i = \text{QwenLLM}(d_i^s \oplus s_i^s) \quad (6)$$

基于微调的 ChatGLM-6b 的数据评分如下所示:

$$S(d_i) = \text{Score}(\text{ChatGLM}(d_i \oplus s_i)) \quad (7)$$

2.4 基于用户交互的用户群体意见领袖发现

为了对用户群体进行划分并发现意见领袖,本文参考文献[16],利用用户之间的交互关系,即转发和@他人,设计直接交互度和间接交互度,构成用户交互度。其中,直接交互度是指用户之间的转发次数和@次数的总和 $DInter(u_i, u_j)$;间接交互度是指用户之间共同转发或@的第三方微博次数总和 $InDInter(u_i, u_j)$ 。 u_i 和 u_j 的交互度定义为两种和的加权,如下所示:

$$I(u_i, u_j) = \alpha DInter(u_i, u_j) + (1 - \alpha) InDInter(u_i, u_j) \quad (8)$$

其中, α 为超参数,设定 α 值为 0.7。进一步采用 Louvain 算法进行用户群体划分。

3 实例分析

3.1 数据来源

本文使用的是上海蜜度科技股份有限公司上海浦东微热点大数据研究院提供的第六届传播数据挖掘竞赛数据集，选用发生在2023年7月24日—2023年10月2日的有关“印花税调整”事件的数据，该事件讨论持续时间78天，共计6 636条数据。依据此数据，本文从舆情生命周期划分、主题演化、情感波动、用户群体划分与意见领袖发现和地域差异5个方面进行实例分析。

3.2 舆情生命周期划分

本文将微博用户认证为“机构认证-学校”和“机构认证-政府”的归为传统媒体，认证为“机构认证-媒体”“机构认证-团体”和“机构认证-网站”的归为网络媒体。本文根据式(1)~式(4)，选取 a_1 、 a_2 和 a_3 分别为0.5、0.25和0.25，

b_1 、 b_2 、 b_3 和 b_4 分别为0.3、0.2、0.2和0.3， c_1 、 c_2 和 c_3 分别为0.2、0.3和0.5， α 、 β 和 γ 分别为0.25、0.2和0.25，进而得到舆情的4个生命周期。基于社会影响力的“印花税调整”事件的生命周期划分如图2所示。

从图2可以看出，舆情广度、强度、速度以及社会影响力的值变化趋势基本一致。由于社会影响力是由多个维度构成的，因此可以更全面地表达舆情生命周期，而影响力速度是由新增原创博主数、新增转发博主数以及影响力持续时间决定的。2023年8月7日—2023年8月14日的影响力差值和影响力速度差值均是最大值，说明其上升是最快的，因此，将2023年7月24日—2023年8月7日作为潜伏期。2023年8月28日—2023年9月4日的影响力差值和影响力速度差值均是最小值，说明其下降是最快的，因此，将2023年8月7日—2023年8月28日划分为爆发期。2023年9月11日—2023年9月18日的4个指标的差值全部下降，因此，将2023年9月11日—2023年9月25日划分为衰退期。

3.3 主题生成与主题演化分析

为了分析不同生命周期的“印花税调整”事件中主题演化情况，本文利用3.2节给出的4个生命周期，采用2.2节方法获得每个主题的关键词，并按照关键词建立提示模板，输入QwenLLM，生成各个主题的名称。4个生命周期的“印花税调整”事件主题名称及其关键词见表1。

由表1可以看出，在舆情的不同生命周期，群众的讨论关注度也在发生着变化。在舆情潜伏期，公众关注度较低，引发了3个主题，在此期间关键词偏中性和积极，主要围绕“信心”“央广网”“利好”等话

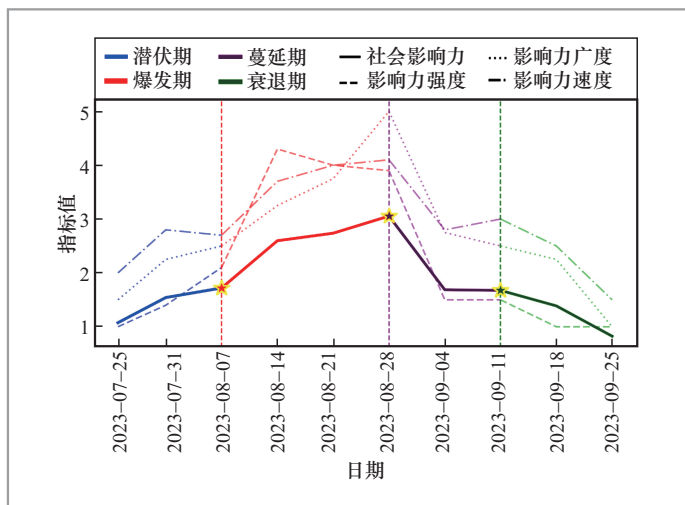


图2 基于社会影响力的“印花税调整”事件的生命周期划分

表1 4个生命周期的“印花税调整”事件主题名称及其关键词

周期	编号	主题名称	主题关键词
潜伏期	0	央广网印花税相关新闻	a股、呼声、央广网、印花税
	1	印花税传闻下股民讨论	信心、经济、政策、印花税
	2	印花税传闻刺激下股市状态	印花税、股市、政策、利好
爆发期	0	政策落地前股民讨论	印花税、市场、下调、A股
	1	政策落地前股民诉求	市场、政策、利好、机会
	2	证监会回应股民呼声	证监会、会议、措施、降费
	3	政策预期降温下股市状态	市场、政策、预期、因素
	4	印花税调整负面观点	印花税、没用、作假、土壤
蔓延期	0	政策落地后股民讨论	涨停、印花税、政策、利好
	1	印花税下调政策解读	利好、经济学家、历次、环比
	2	印花税下调历史回顾	改革开放、国务院
	3	印花税调整部分股民态度	依然、缺口、高开低走、翻身
衰退期	0	政策未达预期效果下股市状态	遇阻、反弹、下跌、回落
	1	政策未达预期效果下股民观点	降准、熊市转身、底部、预期

题开展讨论；在爆发期，公众讨论热度急剧升高，引发了5个主题，并且情感态度出现争议，主要围绕“机会”“利好”“作假”等话题开展激烈讨论；在蔓延期，公众的讨论热度开始下降，转化为4个主题，重点围绕“高开低走”“涨停”“缺口”等话题开展讨论；在衰退期，公众的讨论热度降至最低，仅引发了2个主题，围绕“反弹”“预期”“底部”等话题开展讨论。

3.4 基于大语言模型的舆情情感波动分析

本文利用2.3节基于提示学习的大语言模型情感分类方法，得到2023年7月25日—2023年10月4日的时序正负情感波动图。基于大语言模型的不同生命周期的时序情感波动如图3所示。

从图3可以看出，日情感平均得分围绕0.5正面向上、负面向下波动。在潜伏期，日

情感平均得分达到了正面最高值0.8分，说明公众预期此次印花税调整会产生利好收益，公众对此次事件产生喜悦的情感，而由于具体政策内容的不确定，公众对此次事件的喜悦情感达到了最高。在爆发期，证监会对印花税调整做出回应，公众广泛参与到此事件的讨论中，导致舆情的情感波动频率较大，公众在此段时间表现为喜悦的正面情绪和焦虑的负面情绪混杂。在蔓延期，印花税调整政策被经济学家和财经博主解读，公众对此次事件呈现不同的态度，情感值继续波动，极端的负面情绪减少；在衰退期，部分群众发现在股票市场套利操作并没有达到预期效果，表现出消极态度，而有的群众则充满希望，正面情感上升。

3.5 基于用户交互的用户群体意见领袖分布分析

为了分析不同生命周期参与讨论的公

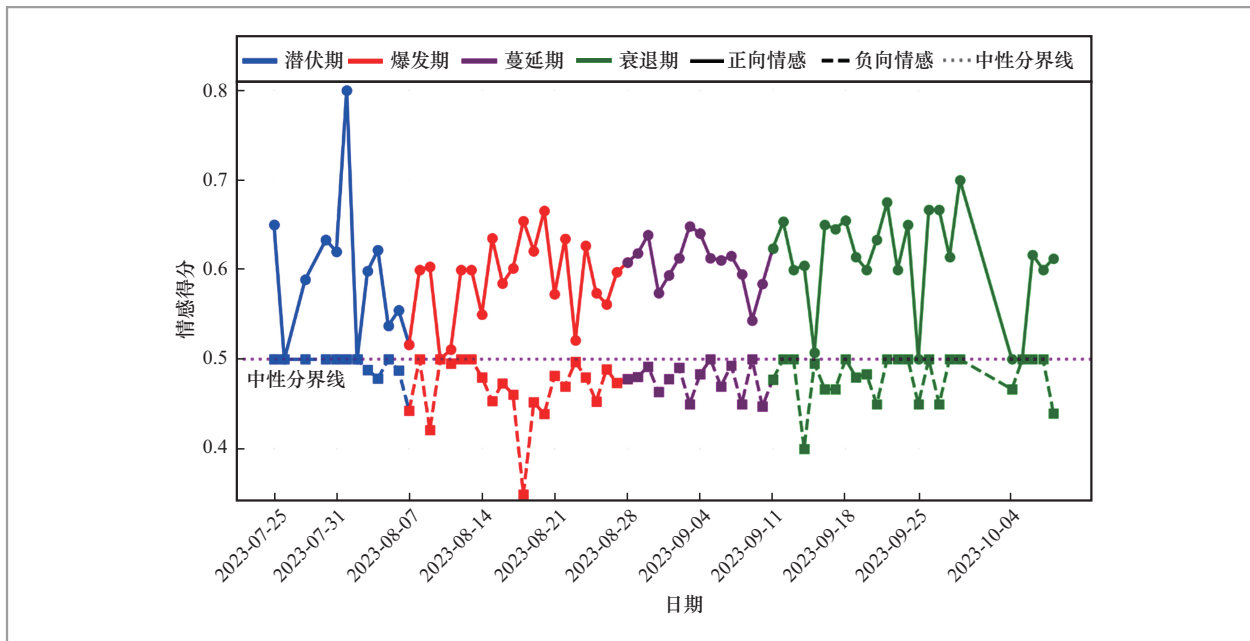


图3 基于大语言模型的不同生命周期的时序情感波动

众群体和意见领袖的影响范围，本文利用2.4节方法对“印花税调整”事件在不同生命周期的群体数、总体影响力和平均影响力广度、用户总数以及平均用户数计算的具体情况进行分析。不同生命周期的群体数、用户总数和影响力广度见表2。

由表2可知，潜伏期参与讨论的用户数较少，群体影响力广度较小；随着爆发期话题的讨论度激增，群体数、总体和平均影响力广度、用户总数和平均用户数均出现大幅上涨，说明在爆发期官方发布的

信息吸引了大量用户参与该话题的讨论，形成了众多群体。相较于爆发期，蔓延期的5项指标均出现一定下降，说明政策确定后群体中部分人对此失去兴趣，逐渐淡出讨论；在衰退期，大量用户退出对该话题的讨论，群体影响力广度明显下降，群体数量同时减少，但平均社会影响力广度并未发生明显变化，并且平均群体用户数激增，说明关于该话题的讨论更加集中，声音更加统一。

为了进一步分析用户群体及其意见领

表2 不同生命周期的群体数、用户总数和影响力广度

时期	群体数	总体影响力广度	平均影响力广度	用户总数	平均用户数
潜伏期	13	16.75	1.28	65	5
爆发期	106	148.75	1.39	1 853	17.48
蔓延期	102	138.25	1.35	817	8
衰退期	16	24.50	1.36	215	13.4

袖的影响范围，本文将用户身份认证划分为普通用户、微博大V、政府机构、新闻媒体和企业，同时，通过散点图直观呈现用户群体及其意见领袖的分布情况。不同生命周期的群体及意见领袖分布如图4所示。

从图4可以看出，在各个生命周期，意见领袖以微博大V为主，其他群体分析如下：在潜伏期，如图4(a)所示，用户群体稀疏，由新闻媒体的“爱股票App”形成的群体在潜伏期也有一定的影响力；在爆发期，如图4(b)所示，用户群体数急剧上升，出现了不同用户类型的意见领袖，但此周期新闻媒体仍占据一定比例；

在蔓延期，如图4(c)所示，用户群体数逐渐减少，用户群体分布相对分散，以新闻和财经博主为意见领袖的群体呈团状；在衰退期，如图4(d)所示，用户群体数急剧减少，在群体上的用户更为集中，以“化股绵涨解析”为意见领袖形成的群体聚集了衰退期的大多数用户。

3.6 舆情地域差异分析

为了分析“印花税调整”事件在不同生命周期的地域差异，本文从舆情社会影响力和舆情情感两个方面对我国部分省区以及境外的相关舆情进行分析。

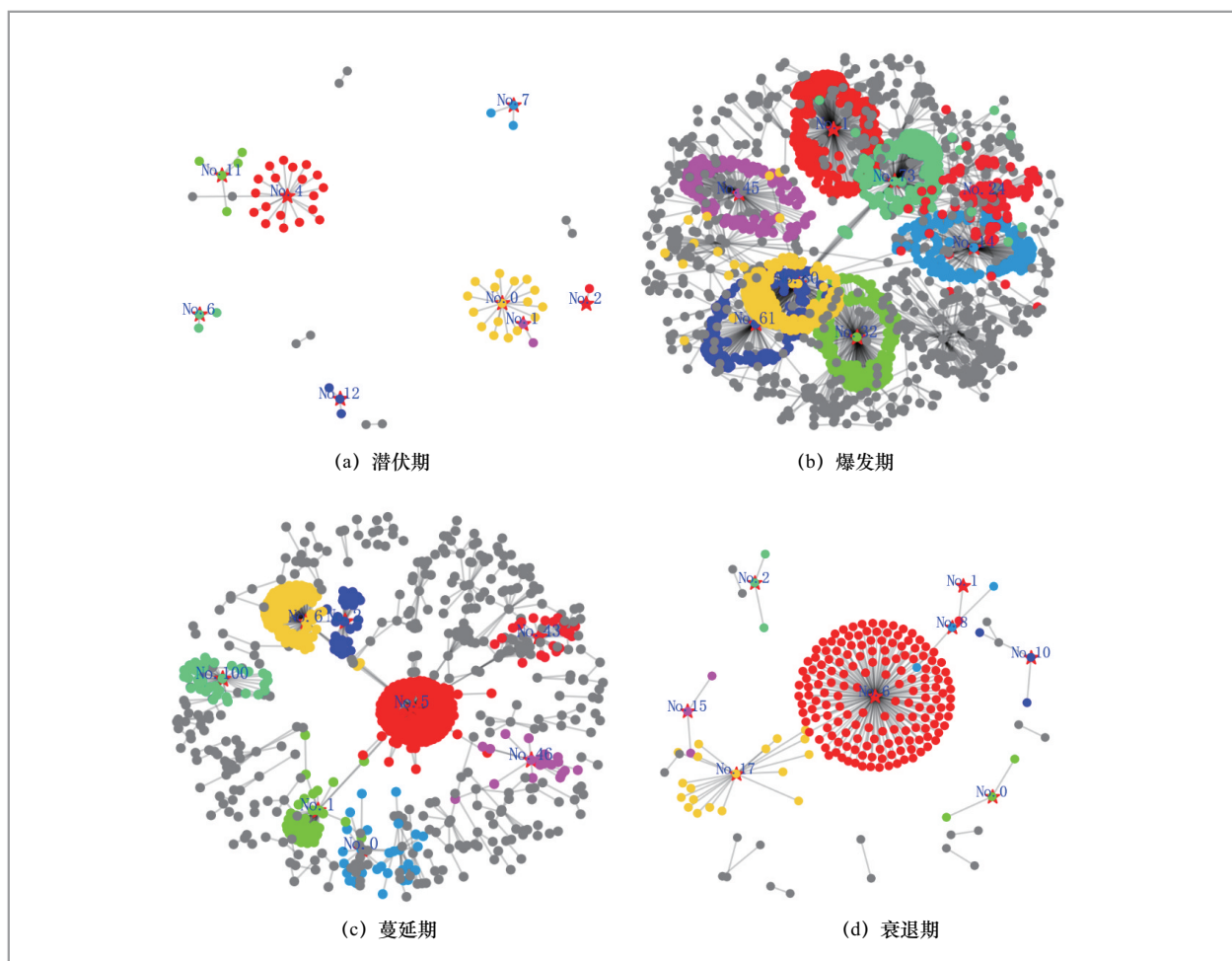


图4 不同生命周期的群体及意见领袖分布

(1) 舆情社会影响力变化

本文利用2.1节的舆情社会影响力计算方法，按照舆情发展的4个生命周期计算社会影响力，不同地域的社会影响力变化情况如图5所示。

通过对“印花税调整”这一事件社会影响力的分析发现，北京在4个时期均显著高于全国平均水平，这与其作为政治经济中心的敏感性和信息传播效率较高有关。广东作为经济大省，爆发期和蔓延期也具有较高影响力，凸显了其对政策变动的快速反应能力。相比之下，宁夏、青海和新疆等地区在爆发期影响力普遍停留在基准水平，且衰退迅速。台湾省仅在爆发期的社会影响力为0.7，其余时期均为0，说明舆情的传播受阻。整体上，我国地域的舆情演化呈现“东高西低、南快北缓”的特点。例如，安徽、甘肃、贵州等中西部省份多为“零潜伏期”，舆情直接进入爆发阶段，可能与信息传递的滞后有关。进一步分析发现，经济发达地区（如北京、广东、上海）的表现优于天津和浙江等地，这可能与金融业较为集中有关。总之，“印花税调整”舆情的影响力与区域经济水平、信

息基础设施和政策关联度呈正相关，其时空分异特征为政策制定者实施差异化舆情引导提供了实证依据。

(2) 不同生命周期的地域舆情情感演化差异

本节分析舆情在不同生命周期下各地区的多维度舆情情感差异。在潜伏期，大部分地区未有明显情绪反应，而北京、广东等地及周边呈较为正向情绪，其他地区则以中性态度为主；在爆发期，随着政策消息释放，全国范围内形成以正面或中性情绪为主导的态势；在蔓延期，在政策落地和配套宣传作用下，舆情情绪显著转向积极；在衰退期则出现区域分化，部分经济欠发达地区（如甘肃、内蒙古等地）情绪回落至中性偏负面水平。整体来看，东部发达地区（如北京、广东等）情绪稳定性较高，而西部地区（如新疆、甘肃等）则表现出较强的政策敏感性，情绪波动较为剧烈。这些发现不仅揭示了政策舆情传播的动态规律，也为精准化舆情管理提供了重要依据。建议重点加强对情绪波动显著区域的实时监测和定向引导，同时对情绪稳定地区保持常态化关注，以实现舆情管理的差异化和最优化。

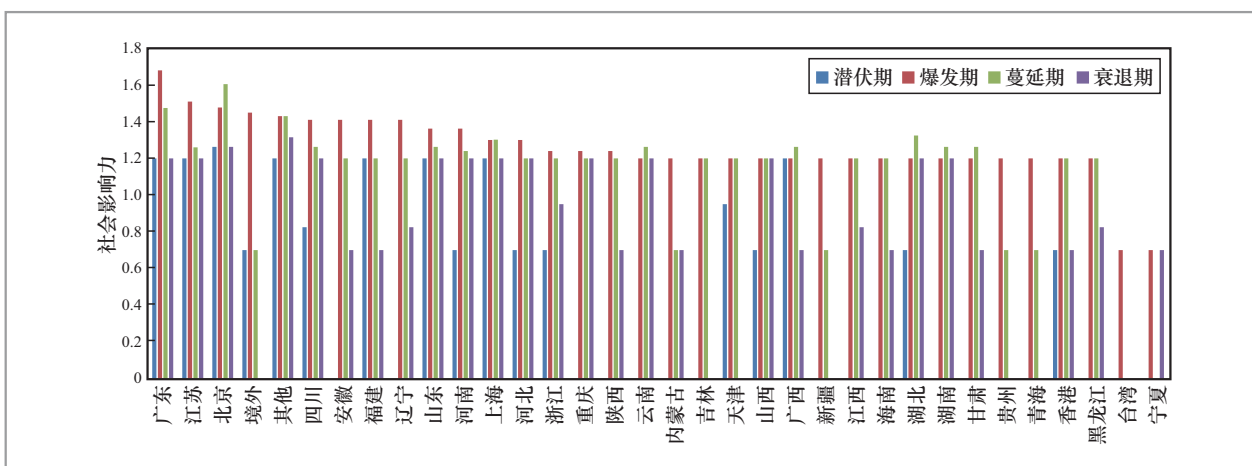


图5 不同地域的社会影响力变化情况

4 结束语

针对微博时序数据中的“印花税调整”事件，本文从主题、情感、社会影响力、生命周期、用户群体行为和地域差异等维度，建立了网络舆情演化框架。研究发现，从主题和情感角度来看，微博用户关注点和情感倾向往往与个人利益紧密相关，在印花税政策尚未确定时，众多用户高度关注此话题，希望从中获益，而一旦政策落地，经过一段时间的市场适应后，用户对话题的兴趣逐渐消退，关注度随之降低。从用户群体的舆情传播效力来看，新闻媒体在扩大话题影响力方面发挥了重要作用，专业领域微博大V在特定群体中具有极高的影响力；在舆情的爆发期和蔓延期，话题影响力范围较大，而在衰退期，舆情热度下降但讨论更为集中，意见领袖引导效果显著。从地域差异来看，东部地区的社会影响力最为显著，西部欠发达地区的影响力相对较小，东部地区对此事件的态度更为积极，广东地区尤为突出。本文的研究可以为有关部门及时监控舆情演化趋势、制定管控的有效政策和保障社会稳定提供技术支持。

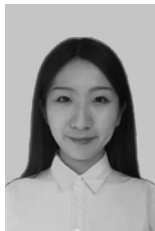
参考文献：

- [1] 迪路阳, 钟寒, 施水才. 网络舆情预警研究综述[J]. 数据分析与知识发现, 2023(8): 17-29.
Di L Y, Zhong H, Shi S C. Review of early warning for online public opinion[J]. Data Analysis and Knowledge Discovery, 2023 (8): 17-29.
- [2] Fink S. Crisis management: planning for the inevitable[M]. New York: American Management Association, 1986.
- [3] 肖丽妍, 齐佳音. 基于微博的企业网络舆情社会影响力评价研究[J]. 情报杂志, 2013, 32(5): 5-10, 19.
Xiao L Y, Qi J Y. On the evaluation system of the social influence of enterprise public opinion on Internet based on microblog[J]. Journal of Intelligence, 2013, 32(5): 5-10, 19.
- [4] 杨洋洋. 风险感知视角下企业突发舆情主题分异研究[J]. 情报理论与实践, 2024, 47(5): 158-166.
Yang Y Y. Research on the theme differentiation of sudden public opinion in enterprises from the perspective of risk perception[J]. Information Studies (Theory & Application), 2024, 47(5): 158-166.
- [5] 高春玲, 姜莉媛, 董天宇. 基于BERTopic模型的老年人健康信息需求主题演化研究: 以新浪微博平台为例[J]. 情报科学, 2024, 42(4): 111-118.
Gao C L, Jiang L Y, Dong T Y. Thematic evolution of health information needs in the elderly based on BERTopic model[J]. Information Science, 2024, 42 (4): 111-118.
- [6] Zhang P, Zhang H, Kong F. Study on the evolution of online public opinion and government response strategies for the “7-20” extraordinary rainstorm and flooding disaster in Zhengzhou, China[J]. Natural Hazards, 2025, 121(3): 2849-2872.
- [7] 王姝雅, 张博. 基于生命周期理论的舆情演化研究综述[J]. 新媒体研究, 2022, 8(23): 6-10, 32.
Wang S Y, Zhang B. A review of public opinion evolution research based on life cycle theory[J]. New Media Research,

- 2022, 8(23): 6-10, 32.
- [8] 许根玮, 冯文刚. 基于TC-LDA模型的电信网络诈骗案件主题挖掘与演化[J]. 情报杂志, 2025, 44(10): 86-96.
- Xu G W, Feng W G. Topic mining and evolution of telecommunications fraud based on the TC-LDA model[J]. *Journal of Intelligence*, 2025, 44(10): 86-96.
- [9] 李爱黎, 张子帅, 林荫, 等. 基于社交网络大数据的民众情感监测研究[J]. 大数据, 2022, 8(6): 105-126.
- Li A L, Zhang Z S, Lin Y, et al. Research on emotion monitoring of public based on social network big data[J]. *Big Data Research*, 2022, 8(6): 105-126.
- [10] 韩小伟, 张传洋, 张起超, 等. 大数据背景下突发公共事件网络舆情情感演化及舆情引导策略研究[J]. 情报科学, 2024, 42(2): 56-63.
- Han X W, Zhang C Y, Zhang Q C, et al. The emotional evolution and guidance strategies of network public opinion event under the background of big data[J]. *Information Science*, 2024, 42(2): 56-63.
- [11] Kumar G, Agrawal R, Sharma K, et al. Combining BERT and CNN for sentiment analysis a case study on COVID-19[J]. *International Journal of Advanced Computer Science and Applications*, 2024, 15(10): 676-686.
- [12] 惠调艳, 王智, 何振华, 等. 基于词典-TextCNN-Word2Vec组合模型的在线评价细粒度情感分析[J]. 情报理论与实践, 2025, 48(2): 168-177.
- Hui T Y, Wang Z, He Z H, et al. A fine-grained sentiment analysis of online reviews based on the dictionary-TextCNN-Word2Vec combined model[J]. *Information Studies (Theory & Application)*, 2025, 48(2): 168-177.
- [13] Deng Y, Zhang W X, Pan S, et al. SOUL: towards sentiment and opinion understanding of language[C]//Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing. Stroudsburg, PA, USA: ACL, 2023: 8704-8711.
- [14] 王续澎, 何洪波, 王闰强. 基于5W传播模型的技术体系: 计算传播技术综述[J]. 大数据, 2025, 11(3): 139-166.
- Wang X P, He H B, Wang R Q. Technology system based on the 5W communication model: a review of computational communication technology[J]. *Big Data Research*, 2025, 11(3): 139-166.
- [15] 黄仕靖, 吴川徽, 袁勤俭, 等. 基于情感分析的突发公共卫生事件舆情时空演化差异研究[J]. 情报科学, 2022, 40(6): 149-159.
- Huang S J, Wu C H, Yuan Q J, et al. Spatiotemporal evolution of public opinion in public health emergencies based on sentiment analysis[J]. *Information Science*, 2022, 40(6): 149-159.
- [16] 林国英, 汪明艳. 基于交互行为的突发事件微博用户社区识别及研究[J]. 网络安全与数据治理, 2022, 41(9): 54-59.
- Lin G Y, Wang M Y. Identification and research of microblog user community in emergencies based on interactive behavior[J]. *Cyber Security and Data Governance*, 2022, 41(9): 54-59.
- [17] Liu Z X, Wu X W. Structural analysis of the evolution mechanism of online public opinion and its development stages based on machine learning and social network analysis[J]. *International Journal of Computational Intelligence Systems*, 2023, 16(1): 99.
- [18] 曹丽娜, 唐锡晋. BBS话题的地理分布分析[J]. 系统科学与数学, 2016, 36(5): 671-682.
- Cao L N, Tang X J. Analysis of topics

- distribution in geography based on BBS[J]. Journal of Systems Science and Mathematical Sciences, 2016, 36(5): 671-682.
- [19] 边晓慧, 徐童. 重大突发公共卫生事件下的公众情感演进分析: 基于新冠肺炎疫情的考察[J]. 数据分析与知识发现, 2022, 6(7): 128-140.
- Bian X H, Xu T. Evolution of public sentiments during COVID-19 pandemic[J]. Data Analysis and Knowledge Discovery, 2022, 6(7): 128-140.
- [20] Xu Z Z, Ye Z, Ye H Y, et al. Public opinion evolution law and sentiment analysis of campus online public opinion events[J]. Journal of Advanced Computational Intelligence and Intelligent Informatics, 2024, 28(4): 990-1004.

作者简介



李旸 (1988-), 女, 博士, 山西财经大学金融学院副教授, 主要研究方向为情感计算和舆情分析。



王志华 (2000-), 男, 山西大学计算机与信息技术学院 (大数据学院) 硕士生, 主要研究方向为数据挖掘。



李大宇 (1994-), 男, 博士, 山西财经大学金融学院讲师, 主要研究方向为情感计算。



赵鑫 (1983-), 女, 蜜度科技股份有限公司上海浦东微热点大数据研究院研究员, 主要研究方向为舆情分析。



詹雅慧（2001-），女，山西财经大学金融学院硕士生，主要研究方向为金融工程和风险管理。



王素格（1964-），女，博士，山西大学计算机与信息技术学院（大数据学院）教授，主要研究方向为自然语言处理、情感计算、舆情分析。

收稿日期: 2025-11-20

通信作者: 王素格, wsg@sxu.edu.cn

基金项目: 国家自然科学基金项目(No.62376143);山西省基础研究计划项目(No.202503021211239, No.202203021212499)

Foundation Items: The National Natural Science Foundation of China (No.62376143), The Shanxi Provincial Basic Research Program Project (No.202503021211239, No.202203021212499)