



Research on Overseas Public Opinion Anomaly
Recognition Based on News Sentiment Analysis
— Taking Huawei Enterprise as an Example

Li Xiong, Lulu Tang and Zhaoran Xu

EasyChair preprints are intended for rapid dissemination of research results and are integrated with the rest of EasyChair.

October 7, 2019

基于新闻情感分析的海外舆情异常识别研究——以华为企业为例

熊励, 唐露露, 许肇然

(上海大学 管理学院, 上海 200444)

摘要 在复杂多变的国际形势下, 针对国际舆论和海外舆情进行分析, 对中国跨国企业及时发现危机并做出正确决策具有重要意义。本文通过新闻情感挖掘, 研究如何利用情感分析方法构建跨国企业舆情指标, 并结合异常点检测算法识别企业的正面和负面舆情异常。以华为企业为例进行舆情异常识别, 结果显示: 在国际市场中华为正面舆情的出现与该企业新科技及新产品的发布相关; 华为负面舆情的出现与“国家安全”、“补贴”等国际政治经济因素相关。

关键词 企业舆情监测; 情感分析; 异常点检测

中图分类号 G35

Research on Overseas Public Opinion Anomaly Recognition Based on News Sentiment Analysis —— Taking Huawei Enterprise as an Example

XIONG Li, TANG Lulu, XU Zhaoran

(School of Management, Shanghai University, Shanghai 200444, China)

Abstract In the complex and ever-changing international situation, the analysis of international public opinion and overseas public opinion is of great significance to Chinese multinational enterprises to find out the crisis in time and make correct decisions. Through the news opinion mining, this paper studies how to use the sentiment analysis method to construct the public opinion index of multinational enterprises, and combines the abnormality detection algorithm to identify the positive and negative public opinion abnormalities of the enterprise. Taking Huawei enterprises as an example to identify abnormalities in public opinion, the results show that the emergence of Huawei's positive public opinion in the international market is related to the release of Huawei's new technology and new products; the emergence of Huawei's negative public opinion is related to international political and economic factors such as "national security" and "Subsidies".

Key words Enterprise public opinion monitoring; sentiment analysis; abnormality detection

在目前严峻的国际形势下，华为、中兴等中国跨国经营企业都面临着极大的挑战。特别是对于这些企业的国际化经营来说，由于涉及到国际政治、经济、文化、地理位置等因素的影响，舆情环境更加复杂，因此，中国跨国企业有必要进行海外舆情分析，及时发现危机并辅助企业做出正确的决策。随着社交媒体时代的到来，网络中信息观点的形成与传播给企业的管理带来深远影响，分析社交媒体中的公众偏好对于理解和反思公众观点至关重要（Halpern 等，2017），同时挖掘在线社交媒体内容已经成为昂贵且耗时的传统问卷调查方法的替代品（Brands 等，2018）。文本挖掘、情感分析、异常检测等数据挖掘技术的发展给企业舆情监测领域带来新的机会，因此，中国跨国企业如何运用这些数据挖掘技术来进行海外舆情的分析和预警并提出相应的解决方案成为一个值得研究的话题。

按照对企业发展进程产生的影响分类，企业舆情可以分为正面舆情和负面舆情（严波等，2019）。正面舆情有利于提升企业的良好形象和商誉价值，企业可以充分利用正面舆情的影响力助推企业发展。负面舆情可能通过企业突发事件爆发、经营管理不善等内部原因产生，也可能因为企业间不良竞争和个别人恶意散发言论形成网络热点而产生，并对企业的形象和商誉价值产生一定程度的破坏。前景理论认为，消费者对损失比对收益更加敏感（Kahneman 等，1979），与正面舆情相比，负面舆情更能影响消费者的行为，对企业造成严重的损失。因此，跨国企业需要及时关注国外媒体的报道尤其是企业的负面舆情，并迅速做出应对举措以减少损失。

鉴于企业舆情分析和预警在企业管理中的重要作用，有必要对企业的正面舆情和负面舆情指标构建及异常识别进行专门研究。为此本文提出研究问题：（1）如何构建新闻舆情指标以评估每篇新闻隐含的媒体对跨国企业的情感倾向？（2）如何识别企业舆情的突发异常以便于跨国企业能够快速及时应对？本文以中国跨国经营企业华为的新闻舆情作为案例分析对象，利用情感分析方法构建企业新闻舆情的情感指标，并使用异常检测算法来识别企业舆情的异常点，从而进一步分析导致企业舆情异常出现的事件和原因。

1 文献综述

1.1 情感分析和企业舆情

企业网络舆情是指公众利用网络发表和传播对企业热点事件的不同认知、看法和态度的总和（肖丽妍等，2013）。目前国内舆情研究偏向于舆情指标体系构建、网民行为决策规律、网络舆情传播机制原理以及防控对策等。例如，朱舸（2015）构建了以舆情热度、危度、离散度为框架的企业危机事件网络舆情指标体系，并建立贝叶斯网络模型以预测评估企业网络舆情；林芹（2017）基于传播主体的特性，对情感距离、信息风险感知、企业形象和企业舆情事件处理效果进行敏感性分析和对比分析，并提出相应的企业网络舆情传播的防控对策；吴鹏（2018）通过基于网民群体负面情感的建模来预测在不同网络舆情阶段和不同政府应急响应措施下网民群体的行为决策规律。随着研究的深入，国内学者采用复杂网络理论、传播动力学等方法，从企业网络舆情的传播路径角度出发，构建企业舆情传播模型，探索企业舆情传播规律（王筱纶等人，2019；王家坤等人，2019）。

检测社交媒体中的情绪可以帮助企业理解相关产品和事件的观点（Howard 等，2013；Cheng 等，2017）。企业舆情监测的核心是通过网络爬虫获取海量的网络舆情数据，并通过对舆情信息的分析，按照一定的指标进行监测。其涉及的核心技术主要包括数据采集技术、自然语言处理技术、序列模式挖掘技术、情感分析技术等（张军玲，2016）。其中，情感分析可以被视为一个分类问题（Korayem 等，2016），旨在将文本或视频（Cambria 等，2014）分为正面，负面或中性（Chaturvedi 等，2018），通常包括三个主要步骤：预处理，特征选择和情绪分类。当前，情感分析主要分为基于语义的分析和基于机器学习的分析两类。基于语义的分析通过构造情感词典，借助 WordNet 或 HowNet 等语义分析工具进行情感倾向判别

(Giatsoglou 等, 2017)。基于机器学习的分析则通过对大量标注情感极性的训练构造一个二元分类器, 让分类器进行情感倾向判别。众所周知的分类器如支持向量机 (Cortes 等, 1995)、朴素贝叶斯分类器 (McCallum 等, 1998)、k-最近邻分类器 (Wang, 2015)、最大熵和人工神经网络等。譬如, Hichem 等人 (2019) 利用支持向量机对阿尔及利亚报纸网站的评论进行情感分类; Karin 等人 (2019) 训练朴素贝叶斯情感分类器来分析推文情感, 以预测英国退欧的可能性, 预测模型精度高达 97.98%。目前神经网络在情感分析应用中也颇为流行, 可被用来识别评论中单个单词的情绪甚至文本和视频数据的情感 (Tang, 2016; Chaturvedi 等, 2019)。本文通过基于语义的情感分析方法构建企业舆情情感指标, 帮助企业及时判别正面舆情以及负面舆情并作出正确应对。

1.2 异常点检测和企业舆情

异常点检测是指寻找那些偏离标准值或正常值的数据点。异常点也称离群点, 其类型有: 短期内产生的峰值, 包括最大值、最小值、以及零值。检测方法可以归类为两种: 第一种异常检测算法主要是对数据点进行分类, 将每个时间点标记为异常或者非异常, 常用算法有分类回归树 (Rahab 等, 2019)、神经网络等; 第二种算法主要是对未来数据走势做预测, 给出置信区间, 即预测某个时间点的信号, 衡量这个时间点的真实值与预测值的差值是否足够大, 如果差值过大则将其视为异常, 常用算法有 ESD 算法、指数平滑法、差分自回归移动平均模型等 (Rosner, 1983; Gardner, 1985; Zhang 等, 2003)。

本文使用周期性混合 ESD (Seasonal Hybrid ESD, S-H-ESD) 算法来检测时间序列数据中的异常点 (Hochenbaum 等, 2017)。该算法在 ESD 算法的基础上, 将时间序列数据分解成三个要素: 趋势、周期、残差, 通过分析残差的背离程度, 引入一定的阈值, 然后将其作为预警依据。同时为了提高鲁棒性, S-H-ESD 算法使用中位数和绝对中位差 (Median Absolute Deviation, MAD) 来代替均值和标准差作为阈值。

目前, 国内外期刊文献在企业舆情领域依据时间序列数据进行异常点检测的研究较少, 大部分集中于决策树、神经网络等手段。

2 研究方法

2.1 研究设计

本文的研究思路主要分为三步: 数据收集和数据预处理、异常点检测以及企业舆情分析。首先, 本文将华为企业舆情作为研究对象, 并爬取国外新闻媒体发布文章作为数据来源。这个阶段的主要工作包括新闻数据采集、新闻文本数据预处理、构建 MPQA 情感词典, 并通过基于语义的情感分析技术构建新闻情感指标, 以便后续进行分析。其次, 在计算出具备时间序列数据特征的新闻综合情感得分之后, 通过 S-H-ESD 算法来检测时间序列中的异常点。最后, 通过检测出来的异常点判断该异常点反映的情绪是消极还是积极的, 并追溯异常点出现日期的新闻事件, 同时通过词云图来可视化分析新闻媒体的关注点, 以便企业能够及时获悉舆情并做出正确决策和应对。

本文的研究框架如图 1 所示, 其逻辑如下: 数据收集和预处理是本文研究的基础, 在该基础上通过 S-H-ESD 算法检测具备时间序列特征的舆情指标的异常值, 最后通过分析异常值对应的新闻和情绪来判断企业舆情并提出相应对策和建议。

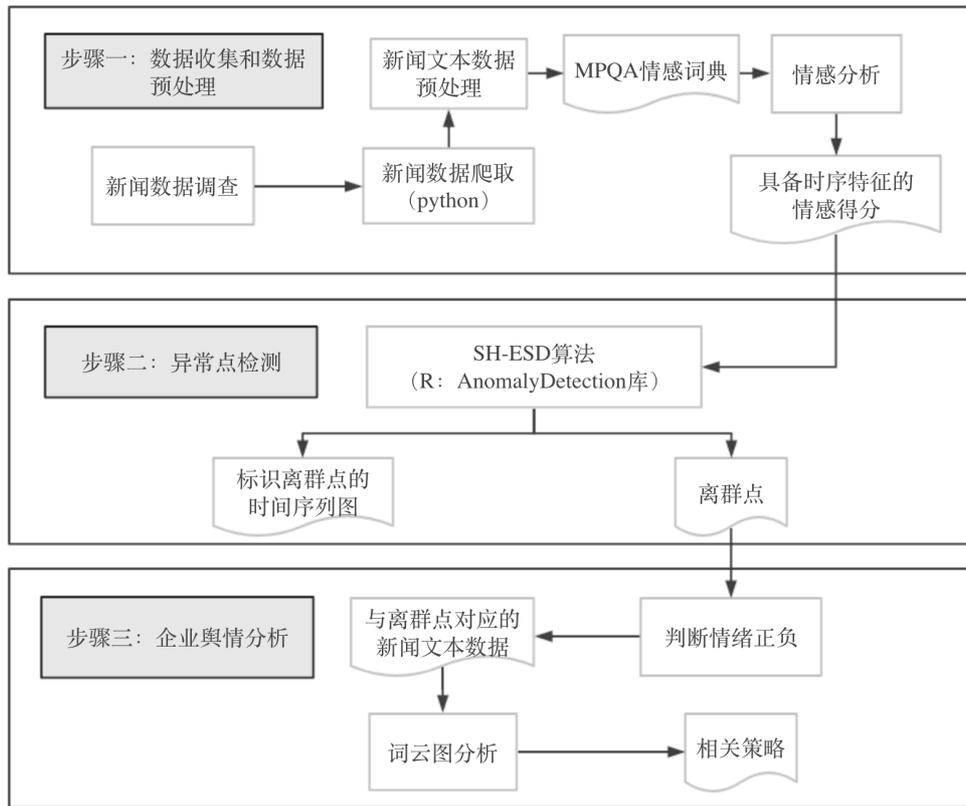


图1 研究框架

2.2 数据收集和预处理

在数据采集工具的选择上，本文通过 python 自建新闻爬虫框架来爬取国外新闻文本并存储到本地。在爬取内容方面，主要抓取了新闻的地址、标题、时间、作者以及新闻文章内容。在数据来源方面，本文主要选取《华盛顿邮报》、《今日美国》、《纽约每日新闻》、《纽约邮报》、《芝加哥论坛报》这五大美国较为权威且容易爬取新闻内容的新闻网站。以“Huawei”为关键词进行搜索，总共爬取了 2012 年至 2019 年 5 月期间 578 条新闻，删去了以视频等非结构化数据作为内容的新闻，最后留下了 538 条新闻。在新闻文本预处理方面，本文主要采用 python 的 Natural Language Tool Kit (NLTK) 库对新闻文本进行分词和词性标注，该库提供了分词、词干提取、解析与语义推理等功能，可用于工业级别的自然语言处理应用。

在进行情感分析时，不同领域的情感分析需要构建适用于该领域的情感词典。MPQA 语料库 (Multiple-Perspective Question Answering Corpus) 包含不同新闻来源的 535 个人工注解的新闻文章，因此，本文利用 MPQA 语料库里面的情感词典进行情感分析，该情感词典包含 8222 个人工注解的单词。每个单词包含 5 个属性：类型、单词长度、词性、词干以及单词情感极性。其中，类型有两个属性：强主观和弱主观。强主观在大多数情况下被认为具有强烈的主观意愿，而弱主观的主观意愿相对强主观来说没那么强但是仍然具有一定的主观性。本文根据强主观和弱主观两个属性，在计算新闻情感分值的过程中，将具备强主观属性的情感词赋予 2 的权重，将具备弱主观属性的情感词赋予 1.5 的权重。

本文根据 MPQA 情感词典整理出积极情感词典、消极情感词典以及否定词词典。积极情感词典收录 MPQA 情感词典里 2304 个情感极性为积极的单词。同样，消极情感词典收录了合计 4153 个消极情感单词。否定词词典收录了诸如“hardly”、“never”、“few”、“little”之类的英文否定词。接下来，本文根据整理出的新闻文本以及情感词典进行情感分析，计算每篇新闻的积极情感分值和消极情感分值。

2.3 综合情感得分指标构建

构建综合情感得分指标之前需要计算第 i 天每篇新闻的积极情感分值 Pos_i 和消极情感分值 Neg_i ，其运用的情感分析算法如下：

Step1. 读取每篇新闻文本，进行分词。

Step2. 导入积极情感词典、消极情感词典以及否定词情感词典。

Step3. 检测分词后的单词是否被情感词典收录，并且记录被收录的单词属于积极情感词 Pos 还是消极情感词 Neg 。如果该单词被情感词典收录，则该单词的初始情感分值 Pos 或 Neg 为 1；如果该单词不被情感词典收录，则默认该单词的初始情感分值 Pos 或 Neg 为 0。如果该单词属于积极情感词且属性为强主观，则分值赋予 2 的权重，即 $Pos*2$ ；若属性为弱主观，则赋予 1.5 的权重，即 $Pos*1.5$ 。消极情感分值赋予权重的方式同理。此外，记录该单词在新闻中的位置。

Step4. 在情感词前面查找英文否定词，如果存在英文否定词的数量为奇数，则该情感词的权重乘以-1；若为零或偶数，则不变。

Step5. 将某一篇新闻里面的积极情感分值 Pos 和消极情感分值 Neg 先分别求和再分别求均值，得到每篇文章的积极情感均值 Pos_i 和消极情感均值 Neg_i 。如果积极情感均值大于消极情感均值，则将该新闻的情感极性判定为 positive；反之为 negative。

通过上面的算法生成每篇新闻文章的积极情感分值 Pos_i 和消极情感分值 Neg_i ，见表 2（以 2019 年数据为例），其中， i 代表对应的日期，polarity 代表新闻的情感极性。由于媒体不会每天都对同一企业进行报道，因此在媒体没有发布相关该企业新闻报道的那天，将缺失的情感分值自动默认为 0，如表中“2019-01-01”当天媒体没有发布关于华为企业的新闻报道，因此对应的积极情感分值和消极情感分值都默认为 0，新闻的情感极性默认为 neutral。此外，通过人工标注结果评估算法模型的准确率，发现总体分类的准确率达到 90%，表明本文的算法具备较高的准确度，能够正确地划分新闻的情感极性。

表 2 每篇新闻的情感分值和情感极性

News Source	Timestamp	Pos _i	Neg _i	Polarity
None	2019-01-01	0	0	Neutral
Daily News	2019-01-02	17.3	15.4	Positive
Chicago Tribune	2019-01-02	31.3	23.0	Positive
The Washington Post	2019-01-03	26.9	9.0	Positive
The Washington Post	2019-01-03	33.9	15.0	Positive
Daily News	2019-01-03	9.3	9.1	Positive
Chicago Tribune	2019-01-03	25.3	18.1	Positive
Chicago Tribune	2019-01-03	38.1	21.1	Positive
New York Post	2019-01-03	5.4	7.1	Negative
None	2019-01-04	0	0	Neutral
None	2019-01-05	0	0	Neutral
The Washington Post	2019-01-06	6.3	7.5	Negative
.....
USA Today	2019-03-03	61.1	4.8	Positive
.....
The Washington Post	2019-05-13	29.5	16.5	Positive

考虑到不同媒体可能在同一天都发布了对于该企业的报道(比如表 2 中 2019 年 1 月 3 日分别有 6 家媒体都对华为企业做了新闻报道), 鉴于此, 将当日的积极情感均值等同于多篇新闻积极情感值 Pos_i 的均值。因此, 本文建立了第 i 天当天所有新闻文章的积极情感均值指标 $AvgPos_i$, n 表示当日某企业新闻的篇数, j 表示不同的新闻报道:

$$AvgPos_i = \frac{\sum_1^n Pos_{ij}}{n}$$

同理, 第 i 天当天所有新闻文章的消极情感均值指标 $AvgNeg_i$ 为:

$$AvgNeg_i = \frac{\sum_1^n Neg_{ij}}{n}$$

以这两个指标为基础构建第 i 天当天所有新闻文章的综合情感得分指标 $NetPos_i$, 该指标为前两者的差值, 公式如下:

$$NetPos_i = AvgPos_i - AvgNeg_i$$

新指标反映了第 i 天当天所有媒体对于企业评价的总体情感倾向。如果该指标为正, 说明第 i 天媒体对于企业的总体评价属于积极乐观的; 相反, 如果该指标为负, 表明当前企业正面临或即将面临威胁。研究通过构建综合情感得分指标, 简化了时序数据, 有利于之后的企业舆情异常检测。基于时间序列的综合情感得分指标见表 3 (以 2019 年数据为例), 其中 $NetPos_i$ 作为 Count 的值。

表 3 基于时间序列的综合情感得分指标

Timestamp	Count
2019-01-01	0
2019-01-02	5.05
2019-01-03	9.92
2019-01-04	0
2019-01-05	0
2019-01-06	-1.2
.....
2019-03-03	56.3
.....
2019-05-13	13

3 企业舆情异常识别与分析

3.1 企业舆情异常识别

本文使用 S-H-ESD 算法来检测综合情感得分时序数据中的异常点, 该算法不仅可以检测局部异常, 也可以检测全局异常, 对峰值异常较敏感。该算法目前已在 Twitter 开源的异常检测库 Anomaly Detection 实现, 因此本文通过 R 导入 Anomaly Detection 库来检测时间序列数据中的异常点, 为进一步企业舆情分析作铺垫。Anomaly Detection 是一个用于检测异常情况的开源库, 其使用场景包括: 在新软件发布后检测系统指标中的异常, 在 A/B 测试后发现用户参与的异常, 以及在计量经济学、金融工程、政治和社会科学领域中的异常, 此外, 它还提供丰富的可视化界面以支持用户进行后续数据分析。本文导入表 3 的时间序列数据并运行程序后, 得到表 4 和图 2。图 2 的点和表 4 的数据相吻合, 可以看到 2019 年 1 月 1 日至 2019 年 5 月 13 日期间出现了三个负值异常点和两个正值异常点。

表 4 2019 年 1 月 1 日-2019 年 5 月 13 日期间企业舆情异常点

Timestamp	Count	News title	polarity
2019-01-03	-1.7	State Department issues China travel warning	negative
2019-01-28	-3.3	U.S. charges Chinese tech giant Huawei, top executive with bank fraud	negative
2019-02-21	-3.5	Here are 5 cases where the U.S. says Chinese companies and workers stole American trade secrets	negative
2019-03-03	56.3	The new phones are coming: The real-world impact of foldable screens and 5G	positive
2019-03-06	33.8	Fitbit introduces budget-friendly line of wearables for kids and adults	positive

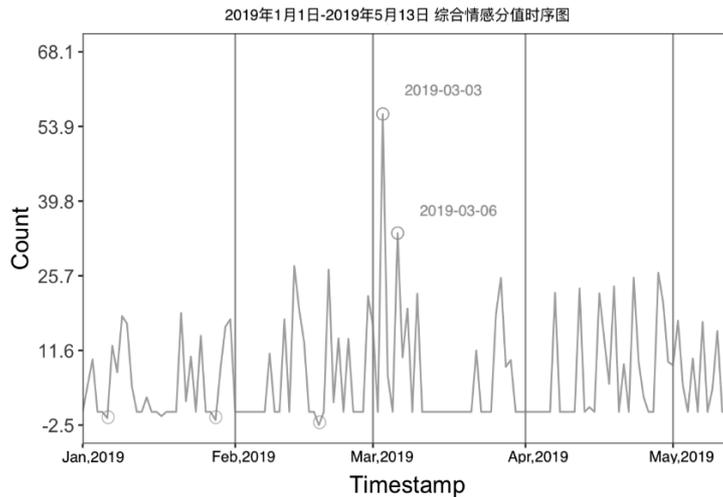


图 2 2019 年 1 月 1 日至 2019 年 5 月 13 日综合情感得分异常点检测

3.2 企业舆情异常分析

企业可以通过查看异常点对应的日期、相关新闻文章及词云图，从而及时了解媒体情绪异常的原因，不仅有利于帮助企业及时正确应对舆情危机，也能帮助企业利用机会加强宣传。异常点既包括正值，又包括负值。前文提过，综合情感得分为正值代表新闻媒体的情感倾向是积极的，综合情感得分为负值代表新闻媒体的情感倾向是消极的。而当积极的综合情感得分成为一个异常点时，表明与企业相关的某一件事情或几件事情的出现将新闻媒体的正面情绪推向了高潮，即出现了企业正面舆情。与此相同，当消极的综合情感得分成为一个异常点时，表明某一件事情或几件事情的出现使得新闻媒体对于企业产生了极大的消极情绪，即出现了企业负面舆情。接下来，本文分别以负值异常点和正值异常点为例，以华为作为企业主体对其企业舆情异常进行分析。

3.2.1 企业正面舆情分析

以表 4 中的某一正值异常点为例，该正值异常点出现在 2019 年 3 月 3 日，情感分值达到 56.3，找到该日的新闻文章，发现这篇新闻来自于《今日美国》，报道的是“华为首款 5G 折叠屏手机——华为 MateX 全球发布，号称‘全球最快的可折叠 5G 手机’”，由此可见，华为首款 5G 折叠屏手机的出现带动了媒体的高涨情绪。从绘制的词云图（图 3）能够看出，“5G”、“foldable”、“smartphone”、“design”等与 5G 折叠屏手机密切相关的词汇受到了媒体的关注和重视。正值异常点的出现一方面表明该企业舆情是正面舆情，另一方面说明企业的新产品或新技术的发布受到媒体的强烈关注，推动媒体气氛到了高潮，可以借此机会加强企业的品牌宣传，提高消费者的品牌形象感知。

第二，华为企业负面舆情的出现大部分与“国家安全”、“监控”、“政府补贴”等国际政治、经济因素有关。这些代表负面舆情的负值异常点在 2012 年出现大量聚集情况，该年开始出现关于美国指控华为、中兴通讯等企业威胁美国安全的新闻报道，当时也影响了韩国、英国等发达国家对华为技术产品的购买；同年欧盟指控中国政府通过补贴华为、中兴通讯等企业向海外倾销中国产品，这两个事件发生导致海外媒体对华为产生了负面情绪。2012 年到 2018 年间断断续续的出现的负值异常点多与这两个事件相关。而在 2018 年孟晚舟事件发生，中美贸易摩擦升级，随着海外大量新闻媒体的报道，负值异常点又出现大量聚集。企业可以通过加强对这些关键词的重视程度，提高警惕并做好应急方案以减少损失。

4 结论与讨论

随着社交媒体的发展，企业需要及时获悉舆情以便于做出正确决策和应对，因此网络舆情的管理对于企业来说有着重要的价值。对于中国跨国经营企业来说，由于涉及到国际经济、政治、文化等因素的影响，跨国经营企业的舆情管理更为复杂。本文将中国跨国经营企业华为作为案例研究对象分析其海外网络舆情，通过情感分析技术构建新闻舆情指标，并结合 S-H-ESD 算法来识别企业舆情的异常，然后以这些异常为线索追溯背后新闻事件，同时通过词云图来可视化分析新闻媒体的关注热点，最后得出结论：华为企业正面舆情的出现大部分与华为新技术或者新产品的发布有关；华为企业负面舆情的出现大部分与“国家安全”、“监控”、“政府补贴”等国际政治、经济因素有关。

从理论意义上来说，本文使用基于新闻情感分析构建的企业舆情指标以及采用的时间序列数据挖掘算法为企业舆情的监测和预警提供了补充。从实践意义来说，本文的方法不仅能够准确地识别企业舆情的异常，而且不论是对华为还是其他跨国经营企业的舆情管理都有所启示。

第一，企业舆情监测中出现正值异常点时，表明企业出现了正面舆情异常，可以借此机会加强企业的品牌宣传，提高消费者的品牌形象感知，生产销售更多的产品提高市场份额。

第二，企业舆情监测中出现负值异常点时，表明企业出现了负面舆情异常，企业需要提升一定的警惕性并做好危机预警。比如，企业需要重点关注导致负面舆情的事件的关键词，一旦新闻媒体的报道有涉及这些关键词需立刻加以重视并采取相关举措减少负面舆情带来的危害影响。

本文也存在一定的局限性。首先，在新闻的采集过程中，没有将华为作为新闻的单一主体，一定程度上影响了情感分析的准确率。其次，基于机器学习的情感分析方法相对本文使用的基于语义的情感分析方法准确性更高，后续可以使用基于机器学习的情感分析方法改进。最后，利用话题跟踪监测以及知识图谱等技术深入挖掘企业舆情也可以作为未来研究方向之一。

参 考 文 献

- [1] Halpern D , Valenzuela, Sebastián, Katz J E . We Face, I Tweet: How Different Social Media Influence Political Participation through Collective and Internal Efficacy[J]. Journal of Computer-Mediated Communication, 2017.
- [2] Schwanholz, Julia, Graham, Todd, Stoll, Peter-Tobias. Managing Democracy in the Digital Age || Social Media Sourcing Practices: How Dutch Newspapers Use Tweets in Political News Coverage[M]. 2018.
- [3] 严波, 厉彦柏, 卢明. 大型企业互联网舆情监测与应对分析[J]. 中国管理信息化, 2019, 22(05):81-82.
- [4] Tversky K A . Prospect Theory: An Analysis of Decision under Risk[J]. Econometrica, 1979, 47(2):263-292.
- [5] 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):6-7.
- [6] 朱舸, 齐佳音. 企业危机事件网络舆情态势评估[J]. 情报科学, 2015, 33(06):48-53+57.
- [7] 吴鹏, 强韶华, 高庆宁. 基于 SOAR 模型的网民群体负面情感建模研究[J]. 中国管理科学, 2018.

- [8] 王筱纶,顾洁.企业危机网络舆情的传播路径及其在供应链中的纵向溢出效应研究[J].管理科学,2019,32(01):42-55.
- [9] 林芹,郭东强.企业网络舆情传播的系统动力学仿真研究——基于传播主体特性[J].情报科学,2017(04):57-63+70.
- [10] 王家坤,王新华.双层社交网络上的企业舆情传播模型及控制策略研究[J].管理科学,2019,32(01):28-41.
- [11] Cheng K, Li J, Liu H. Unsupervised Feature Selection in Signed Social Networks[C]// Acm Sigkdd International Conference on Knowledge Discovery & Data Mining. 2017.
- [12] Howard N, Erik Cambria. Intention awareness: improving upon situation awareness in human-centric environments[J]. Human-Centric Computing and Information Sciences, 2013, 3(1):1-17.
- [13] 张军玲.我国网络舆情信息挖掘研究综述[J].情报科学,2016,34(11):167-172.
- [14] Korayem M , Aljadda K , Crandall D . Sentiment/Subjectivity Analysis Survey for Languages other than English[J]. Social Network Analysis & Mining, 2016, 6(1):75.
- [15] Cambria E , Wang H , White B . Guest Editorial: Big Social Data Analysis[J]. Knowledge-Based Systems, 2014, 69:1-2.
- [16] Chaturvedi I , Ragusa E , Gastaldo P , et al. Bayesian Network based Extreme Learning Machine for Subjectivity Detection[J]. Journal of the Franklin Institute, 2017:S0016003217303009.
- [17] Giatsoglou M , Vozalis M G , Diamantaras K , et al. Sentiment analysis leveraging emotions and word embeddings[J]. Expert Systems with Applications, 2017, 69:214-224.
- [18] Cortes C, Vapnik, Vladimir. Support-vector networks[J]. Machine Learning, 1995, 20(3):273-297.
- [19] Nigam K, McCallum A, Thrun S, et al. Learning to Classify Text from Labeled and Unlabeled Documents[C]// Fifteenth National/tenth Conference on Artificial Intelligence/innovative Applications of Artificial Intelligence. 1998.
- [20] Wang X , An K , Tang L , et al. Short Term Prediction of Freeway Exiting Volume Based on SVM and KNN[J]. International Journal of Transportation Science and Technology, 2015, 4(3):337-352.
- [21] Sabatovych. Use of Sentiment Analysis for Predicting Public Opinion on Referendum: A Feasibility Study[J]. The Reference Librarian,2019,60(3).
- [22] Tang D , Qin B , Liu T . Aspect Level Sentiment Classification with Deep Memory Network[J]. 2016.
- [23] Iti Chaturvedi, Ranjan Satapathy, Sandro Cavallari, Erik Cambria. Fuzzy commonsense reasoning for multi-modal sentiment analysis[J]. Pattern Recognition Letters,2019,125.
- [24] Hichem Rahab, Abdelhafid Zitouni, Mahieddine Djoudi. SANA: Sentiment Analysis on Newspapers comments in Algeria[J]. Journal of King Saud University - Computer and Information Sciences,2019.
- [25] Rosner B . Percentage Points for a Generalized ESD Many-Outlier Procedure[J]. Technometrics, 1983, 25(2):165-172.
- [26] Zhang G P . Time series forecasting using a hybrid ARIMA and neural network model[J]. Neurocomputing, 2003, 50(none):159-175.
- [27] Gardner E S. Exponential smoothing: The state of the art—Part II[J]. International Journal of Forecasting, 2006, 22(4):637-666.
- [28] Hochenbaum J , Vallis O S , Kejariwal A . Automatic Anomaly Detection in the Cloud Via Statistical Learning[J]. 2017.

作者简介:

熊励(1966-),女,上海大学管理学院教授,研究方向包括信息决策、跨境电商, xiongli8@shu.edu.cn。
唐露露(1995-),女,上海大学管理学院硕士研究生,研究方向包括企业舆情, luckylulu@shu.edu.cn。
许肇然(1985-),女,上海大学管理学院博士,研究方向包括跨境电商、信息管理, xuzhaoran@163.com。