

观察TikTok平台有关“科目三”网络模因相关的标签内容，最终确定以“#科目三”“#一笑江湖”“#dieunhaylangla”（出自越南语，部分TikTok用户使用该标签代指“科目三”）为搜索关键词。根据TikTok平台的算法热度推荐，爬取了截至2024年7月25日有关“科目三”网络模因的600条视频内容，去除广告、重复、无意义符号、播放数小于一万等无效内容，最终提取了479条有效视频样本及视频标题内容，累计产生10.91亿次播放、4504.38万点赞、63.69万评论。相关数据表明，“科目三”短视频模因在TikTok平台具备较强的传播力与影响力。为确保有关评论具有一定代表性，研究者选取每条有效视频样本中算法热度排名前50的评论，最终得到24966条有效评论数据，共计263154个评论字数，最后将其全部转换为中文数据输出为EXCEL表格内容。

（二）研究方法与数据处理

为进一步了解短视频网络模因“科目三”在跨文化语境下的传播与解码效果，分析运用TikTok用户对该模因的关注焦点以及情感态度相结合的研究方法对数据进行处理。

1、聚类与可视化

语料库语言学是一种基于大量文本进行量化分析的研究方法，它能够揭示给定文本中的关键语言现象。^[18]该方法可以识别和描述文本的总体特征，以词频、关键词、词簇和词组的形式，为话语分析提供基于数据的支持和深入理解。^[19]在本研究中，主要通过语料库语言学方法来观察TikTok平台有关“科目三”网络模因受众反馈内容的两个特征：关键词以及词组。为此，研究对爬取到的用户评论数据以及视频标题数据进行可视化处理，通过选定的视觉媒介以直观方式传达数据信息。研究使用KH Coder软件绘制TikTok“科目三”网络模因479条视频标题数据的词频共现网络图，该图能够识别出具有相似出现模式的词汇。其次，研究借助Python中的WordCloud库来生成用户24966条评论数据的词云图，以视觉形式展示用户评论中出现的高频词汇。

2、基于LDA的网络评论主题发现

潜在狄利克雷分配（Latent Dirichlet Allocation, LDA）是一种基于概率的主题发现模型，也是较为经典并被广泛使用的网络文本分析方法。该模型包含词、主

题、文档三层结构，能够在不预设特定规则的情况下，有效发现评论文本中隐含、有价值的主题，运用该方法可以有效处理大规模语料库资料。^[20]基于此，本文将LDA主题发现模型运用到TikTok平台有关“科目三”网络模因参与用户的评论主题挖掘中，结合词项语义分析，实现对网络评论的主题发现。具体操作步骤如图1所示。

3、情感分析

情感分析，又称意见挖掘（opinion mining），是指人们对某一实体的观点、态度和情感的计算研究，这个实体可以代表个人、事件或话题。^[21]现有的情感分析方法主要有字典法和机器学习方法两种，本文在Python编程语言的所爬取的评论数据基础上，以TikTok用户对“科目三”网络模因所发布的24966有效评论为分析样本，使用自然语言处理（NLP）的方法进行情感分析，了解TikTok用户对于该模因的情感态度。

研究发现

（一）TikTok平台“科目三”模因的内容特征

研究使用Python软件爬取了TikTok平台算法热门推荐的479条“科目三”短视频内容的标题数据，包含创作者所使用的标签信息（TAG）以及对视频的描述信息，使用KH Coder软件将相关数据绘制成词频共现网络图（见图2）

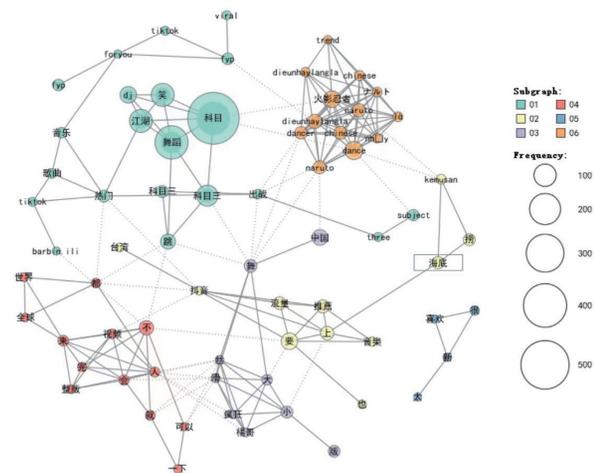


图2 TikTok平台“科目三”短视频标题数据的词频共现网络图

士弗曼提出分析网络模因的框架，认为人们可能会从“内容”“形式”和“立场”三个维度对模因进行模仿。“内容”维度主要涉及模因特定文本的内容，包含模因的具体信息和思想观点；“形式”维度是由人们感官系统所感知到的模因物理形态，包括对模因听觉和视觉的感知等；“立场”维度是指模因传达与传播者自身沟通有关的信息，包括参与结构、定调和传播功能三个方面。^[22]本研究通过对TikTok平台“科目三”相关短视频

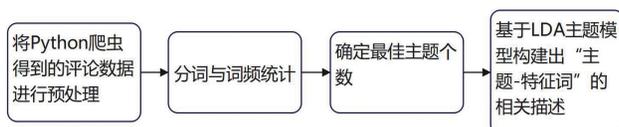


图1 基于LDA模型分析用户评论