

# 疫情期间舆情演化分析与热点事件模式挖掘 (ChinaVis 2020 数据可视分析挑战赛一等奖)

李泽宇, 陈诺, 王腾, 李思思, 贾世超, 张加万

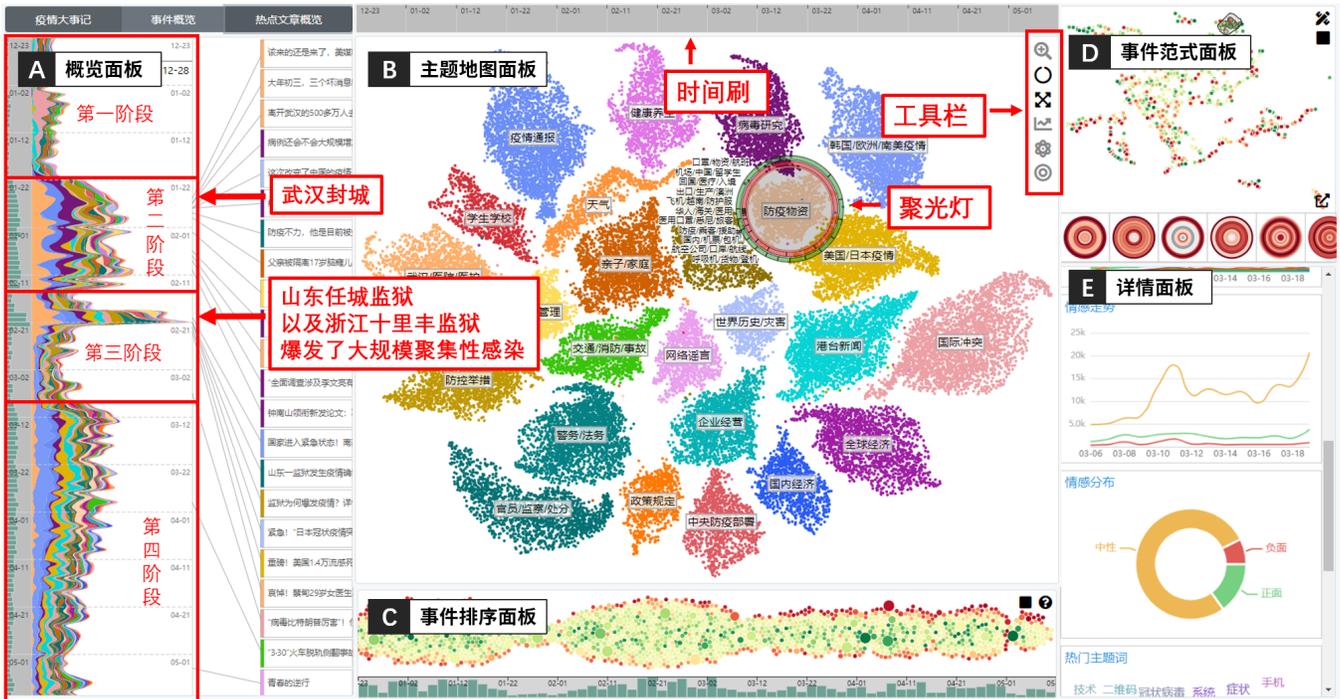


图 1: 系统界面。概览面板中的堆叠面积图表达各主题文章阅读总量随时间的变化; 主题地图面板描述主题的分布; 事件排序面板展现事件在时间维度和多种情感属性下的排序; 事件范式面板将具有相似演化范式的事件聚集在一起; 详情面板提供某个被关注事件的详细信息。

## 概述

掌握疫情发展趋势与关键节点、感知舆论话题与情感的动态演变、理解事件的热度演化范式对于政府和个人都具有重要意义。我们设计并实现了一个可视分析系统, 它能够帮助分析者从话题和词汇两个语义级别, 从描述词、人物、机构等多个角度完成上述任务。

**关键词:** 话题演化、情感分析、热度演化模式挖掘

### 1 数据、模型与分析任务

**数据:** 如下表, 我们收集的数据包括: 头条号热文数据、热点事件数据以及疫情新闻列表, 时间跨度为: 2019 年 12 月 23 日~2020 年 5 月 10 日

	热文数据	热点事件数据	疫情新闻列表
数量	42000	1781	690
来源	榜单: 清博大数据 文章: 今日头条	清博大数据	知微数据
核心字段	文章标题/内容 发布日期 阅读数	事件标题/概述/热词 发生日期 14天情感/热度趋势	新闻标题 发生日期
类型	时序文本	时序文本, 时间序列	时序文本

**模型:** 对于热文数据, 我们运用主题模型将它们分为 25 个主题, 接着运用 UMAP 算法投影在二维平面上 (主题地图面板)。对于热点事件数据, 我们使用 Tsfresh 提取热度序列的特征, 并使用 PCA 和 UMAP 算法将事件投影至二维平面 (图 1D)。此外, 我们根据事件的文本数据, 运用已经训练好的主题模型和降维模型, 将事件投影在主题地图中 (图 5)。

### 分析任务:

1. 明确疫情发展阶段, 识别重大热点事件
2. 从多个角度和多个尺度探究舆论关注点的转变
3. 找出在情感相关维度上特殊的热点事件
4. 挖掘热点事件的热度演化模式

- 李泽宇, 天津大学, lzytianda@tju.edu.cn
- 陈诺, 天津大学, 3016218109@tju.edu.cn
- 王腾, 天津大学, anianw@qq.com
- 李思思, 天津大学, sisilee@tju.edu.cn
- 贾世超, 天津大学, jsctju@gmail.com
- 张加万, 天津大学, jwzhang@tju.edu.cn

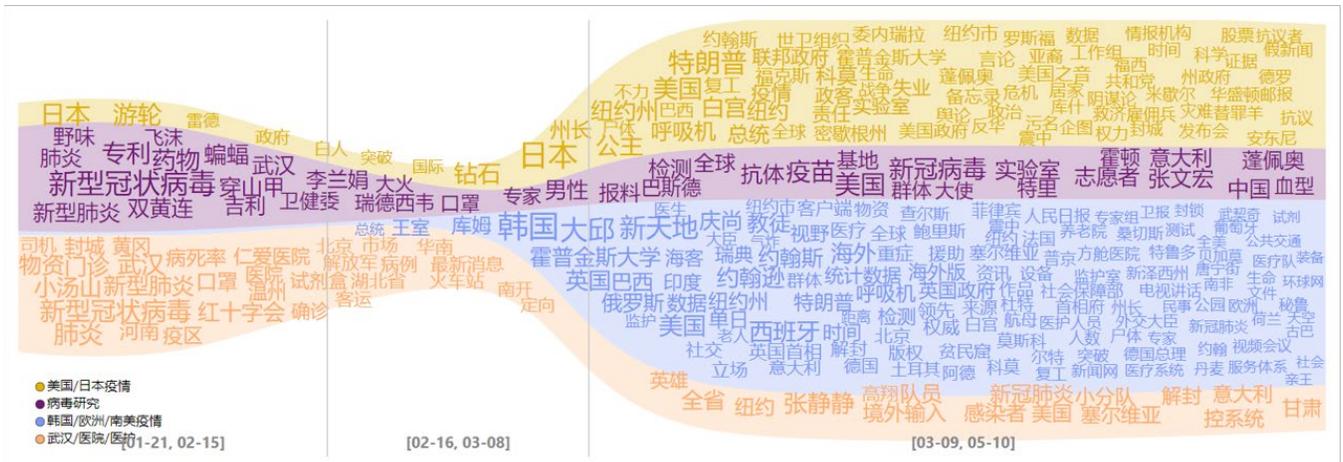


图3: WordStream揭示舆论关注点在词级别的演化

## 2 可视分析系统简介

如图1, 我们的系统界面包括五个部分。Galex[1] (图2) 和WordStream[2] (图3) 分别用于揭示舆论热点在主题级别和词级别的演化, 聚光灯支持检视主题地图各处的语义含义。事件排序面板 (图4) 中节点颜色和大小分别表示事件的情感倾向和热度。

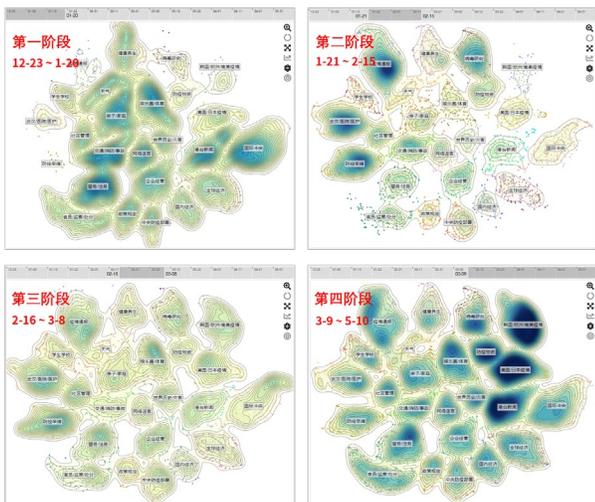


图 2: 主题地图揭示舆论关注点在主题级别的演化

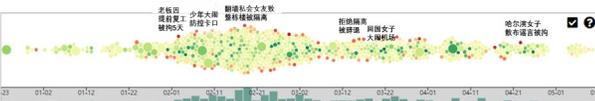


图 4: 疫情相关事件在时间维度上的分布

## 3 发现

图1和图2揭示了疫情舆论的四个阶段:

**第一阶段**, 12月23日至1月20日 (国内疫情舆论“密切关注期”)。疫情尚未引起公众广泛的关注。

**第二阶段**, 1月21日至2月15日 (国内疫情舆论“爆

发期”)。1月23日, “武汉封城”的新闻引爆了公众关注度, 疫情引起了全国范围的高度重视。

**第三阶段**, 2月16日至3月8日 (国内疫情舆论“平缓期”)。但在2月19、20日出现了一个爆发式增长, 这是因为山东任城监狱以及浙江十里丰监狱发生了规模较大的聚集性感染。

**第四阶段**, 3月9日至5月10日 (国际疫情舆论“爆发期”)。此阶段, 疫情在全球爆发。

在图3中, “韩国/欧洲/南美疫情”话题和“美国/日本疫情”话题中热词的转变与疫情在国际的传播过程高度吻合, 即从日韩开始, 再到欧洲, 再到美国、南美。图4揭示与疫情相关事件中的负面事件大多与“不遵守防疫规定、散布谣言”等行为相关。



图5: 事件嵌入在主题地图中

## 4 结论

我们设计并实现了一个用于分析疫情期间舆情演化与热点事件模式的可视分析系统, 它整合了一些新颖的可视设计并提供了大量灵活的交互。实际上, 它支持多种类型的文本数据和时序数据的分析。

### 主要参考文献

[1] Li Z, Zhang C, Jia S, et al. Galex: Exploring the Evolution and Intersection of Disciplines[J]. IEEE transactions on visualization and computer graphics, 2019, 26(1): 1182-1192.  
 [2] Dang T, Nguyen H N, Pham V, et al. WordStream: Interactive Visualization for Topic Evolution[C]//EuroVis. 2019.