

软件技术与数据库

基于文本挖掘的话题发现技术

高 妮¹, 周明全², 耿国华¹, 王学松², 贺毅岳¹

(1. 西北大学信息学院计算机科学系, 西安 710069; 2. 北京师范大学信息科学与技术学院, 北京 100875)

收稿日期 修回日期 网络版发布日期 接受日期

摘要 在分析灾害新闻特点的基础上, 提出一种基于文本挖掘的话题发现技术, 采用基于平均分组的层次聚类算法, 对灾害新闻资料进行组织, 从而生成新闻专题, 为用户提供个性化服务, 并形成专题检测系统, 同时介绍基于时间和地点权值向量的相似度计算模型以及基于时间的动态阈值模型。实验结果表明, 该算法能够获得较好的性能。

关键词 [话题发现与跟踪](#); [层次聚类](#); [文本挖掘](#); [动态阈值](#)

分类号 [TP311.52](#)

DOI:

通讯作者:

作者个人主页: [高 妮¹](#); [周明全²](#); [耿国华¹](#); [王学松²](#); [贺毅岳¹](#)

扩展功能

本文信息

- ▶ [Supporting info](#)
- ▶ [PDF\(147KB\)](#)
- ▶ [\[HTML全文\]\(0KB\)](#)
- ▶ [参考文献\[PDF\]](#)
- ▶ [参考文献](#)

服务与反馈

- ▶ [把本文推荐给朋友](#)
- ▶ [加入我的书架](#)
- ▶ [加入引用管理器](#)
- ▶ [引用本文](#)
- ▶ [Email Alert](#)
- ▶ [文章反馈](#)
- ▶ [浏览反馈信息](#)

相关信息

- ▶ [本刊中 包含“话题发现与跟踪; 层次聚类; 文本挖掘; 动态阈值”的 相关文章](#)
- ▶ [本文作者相关文章](#)