

程俊霞¹, 李芝棠^{1,2}, 邹明光¹, 肖津¹. 基于SVM过滤的微博新闻话题检测方法[J]. 通信学报, 2013, (Z2): 74~78

基于SVM过滤的微博新闻话题检测方法

Novel topic detection method for microblog based on SVM filtration

投稿时间: 2013-09-08

DOI: 10.3969/j.issn.1000-436x.2013.Z2.015

中文关键词: [话题检测](#) [特征向量](#) [SVM](#)

英文关键词: [topic detecting](#) [characteristic vector](#) [SVM](#)

基金项目:

作者

单位

[程俊霞¹](#), [李芝棠^{1,2}](#), [邹明光¹](#), [肖津¹](#)

[1. 华中科技大学 计算机学院, 湖北 武汉 430074](#); [2. 下一代互联网接入系统国家工程实验室, 湖北 武汉 430074](#)

摘要点击次数: 77

全文下载次数: 41

中文摘要:

在基于聚类的话题检测方法上提出了一种基于SVM过滤的检测方法, 该方法在聚类前将微博文本特征抽象成用于输入向量机的向量, 对微博文本进行过滤, 降低了计算量。并针对微博聚类的长尾现象提出了基于高频词排序的改进单遍聚类方法, 能很好地检测孤立点的存在。实验表明, 该方法在海量微博数据中能有效地检测出新闻话题。

英文摘要:

A detection method based on SVM filtration was proposed. The method uses text feature as imported vectors to filtrate microblog news, reducing the amount of calculation greatly. A single-pass clustering algorithm based on the improvement of high-frequency words sorting was proposed, which can detect isolated points commendably. Experimental results show that the method can detect news topics from massive microblog data efficiently.

[查看全文](#) [查看/发表评论](#) [下载PDF阅读器](#)

关闭

版权所有: 《通信学报》

地址: 北京市丰台区成寿寺路11号邮电出版大厦8层 电话: 010-81055478, 81055479
81055480, 81055482 电子邮件: xuebao@ptpress.com.cn

技术支持: 北京勤云科技发展有限公司