首页　　|期刊简介|编委会|投稿须知|在线订阅|资料下载|编委论坛

基于依存连接权**VSM**的子话题检测与跟踪方法

Sub-topic detection and tracking based on dependencyconnection weights for vector space model

作者　　　　　　　　　　　　　　　　　　　　单位

周学广，高飞，孙艳　　　　　　　　　　　　海军工程大学 信息安全系，湖北 武汉430033

中文摘要**：**

　　针对在新闻话题中报道突发、热点相似且子话题层次丰富的现象，依据增量TF-IDF值构造特征维，生成全局向量；然后在时间窗内生成特征连接权的局部邻接图，利用依存句法进行分析降维；最后采用领域词典加权，时间阈值衰减；从而构造出利用依存连接权VSM进行关联分析的子话题检测与跟踪（sTDT）计算方法。实验表明，利用依存关联分析使文本表示由线性变为平面结构，能够有效地提取描述子话题；在人工标注的测试语料下，其最小DET代价比经典方法至少降低2.2%。

英文摘要**：**

　　Aiming at the phenomenon that there are abrupt reports, similar topics and abundant levels of subtopics in the news, a novel method based on relationship analysis using dependent sentence pattern was proposed for sub-topic detection and tracking (sTDT), which constructed feature dimensions to generate the global vectors according to the increment of TF-IDF, and then created the partial adjoin map based on the connection weights within the time window and decreased the dimensions through dependent sentence pattern. Finally, a novel method for sTDT computing was built with adjoins dictionary weights and time threshold attenuation. Experiments show that the proposed method transferrs the text from linear to plane structure, and extracts the subtopics effectively, of which the minimum DET cost is reduced by at least 2.2 percent than that of classical methods.

查看全文　查看/发表评论　下载PDF阅读器

关闭