工程与应用

基于SVM的公安情报自动分类系统的研究与设计

司志刚, 牛 琳, 常朝稳

信息工程大学 电子技术学院 404教研室，郑州 450004

摘要　依据公安情报文本中不同位置的词条对区分文本类别的贡献显著不同的特点，引入位置权重系数，改进了经典的文本特征权重计算方法（TF-IDF），使文本的权重能够更加全面地反映文本的类别信息。根据公安情报分类系统的需求，设计了基于支持向量机（SVM）的公安情报分类系统，该系统不仅能够实现情报文本的自动分类，而且能够保留在情报文本分类的不同阶段语料的特征信息，为情报信息的进一步加工处理提供支持，同时系统中各模块间采用松耦合的方式衔接，提高了系统的适应性和灵活性。通过实验验证了系统设计的合理性和有效性。

关键词　文本分类　支持向量机　向量空间模型　公安情报

分类号

Reserch and design of automatic text categorization system of public security information based on SVM

SI Zhi-gang,NIU Lin,CHANG Chao-wen

Institute of Electronic Technology，Information Engineering University，Zhengzhou 450004，China

Abstract

In the public security information text，the vocabulary entries in different positions own obviously different ability to distinguish the text categorization.According to the features of the public security information text，this paper introduces the regional weight modulus，improves the classic TF-IDF.It makes the weight of vocabulary entry can full-scale express the contributions on the text's category.This paper designs the automatic text categorization system of public security information based on Support Vector Machine（SVM），according to the requirements of public security information.This system not only distinguishes the text classification automatically，but also retrains the features information at different stages of the categorization process.It will support the further information procession.Meanwhile，this system reduces the coupling degree of the subcomponents，provides the flexible system structure，and improves the adaptability of system.At last，the rationality of the design for this system is verified by experiment.

Key words　text categorization　Support Vector Machine（SVM）　Vector Space Model（VSM）　public security information

通 讯 作者　司志刚 changchaowen@hnsl.gov.cn

扩展功能

本文信息

- Supporting info
- PDF(905KB)
- [HTML全文](0KB)
- 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
- 复制索引
- Email Alert
- 文章反馈
- 浏览反馈信息

相关信息

- 本刊中 包含"文本分类"的相关文章

本文作者相关文章

- 司志刚
- 牛 琳
- 常朝稳