数据库、信息处理

改进的一分类支持向量机的邮件过滤研究

秦 谊[1]，裴 峥[1]，杨霁琳[2]

1.西华大学 数学与计算机学院，成都 610039
2.西南交通大学 数学学院，成都 610031

摘要　　服务器端存在多个用户，且人们对邮件内容的理解和认可程度不同，因此邮件过滤中涉及到不确定信息的处理。就邮件内容来看，邮件过滤通常涉及到隐私，不利于大量收集样本并评价打分。因此提出了一种基于改进的一分类支持向量机的邮件过滤方法。该方法优点在于：（1）用户只需为不确定性很强的待区分邮件给出隶属度；（2）只需收集和训练一类邮件样本，便可以建立邮件分类模型；（3）把隶属度首次引入到1-SVM中，并且由隶属度的值的大小来确定惩罚因子的值。通过仿真实验验证了该方法的有效性。

关键词　　[一分类支持向量机](#)　[邮件过滤](#)　[隶属度](#)　[不确定性](#)　[有序加权平均算子](#)

分类号

Filtering e-mail based on improved One-class Support Vector Machine.

QIN Yi[1]，PEI Zheng[1]，YANG Ji-lin[2]

1.School of Mathematics & Computer Science，Xihua University，Chengdu 610039，China
2.School of Mathematics，Southwest Jiaotong University，Chengdu 610031，China

Abstract

  Because there are many users in server，and users have different understand or admitting degrees for the content of e-mails，uncertain information processing is dealt with in filtering e-mails.From the content of e-mails point of view，filtering e-mails always deals with privacy，this is disadvantage for largely collecting e-mails and evaluating them.Filtering e-mail based on improved one-class SVM is proposed，the advantages of the method are（1）users only give membership degrees for uncertain e-mails which will be dealt with；（2）classing e-mails model is constructed by a kind of e-mail samples；（3）membership degrees are discussed in one-class SVM，and membership degrees are also used to decide punish factors.Simulation shows that the method is effective.

Key words

[One-class Support Vector Machine（1-SVM）](#)　[e-mail filtering](#)　[membership degree](#)　[uncertainty](#)　[Ordered Weighted Averaging（OWA）　operator](#)

通讯作者　pqyz@263.net