

金牌邮局智能反垃圾

截至 2007 年 12 月，中国网民收到的垃圾邮件总量为 694 亿封，与去年的 500 亿封相比增长率为 38.8%。网民收到垃圾邮件的比例为 55.65%，与去年同期的 58.98%相比下降了 3.33%。2007 年垃圾邮件给中国造成的损失达 188.4 亿人民币，与去年的 104.315 亿相比增长了 80.6%。

金牌邮局智能反垃圾系统采用国际领先的人工智能模拟技术，用电脑模拟人脑判别垃圾邮件的“模糊判断”过程，帮助用户智能的识别垃圾邮件。

它是一个基于规则判断的系统，利用一个庞大的规则集对电子邮件的各个组成部分进行比对测试，判断一封电子邮件与“垃圾邮件”的相似程度。每个规则都会增加或减少该电子邮件的“垃圾邮件相似分数”。如果一封电子邮件具有很高的“垃圾邮件相似分数”的话，它就会判断这是一封垃圾邮件。

工作原理：

- ◆ 使用大量的、不同类型的规则和权重来判断垃圾邮件，那些在实践中已经被证实有效的规则会被赋予较高的权重用以区分垃圾邮件和非垃圾邮件。
- ◆ 通过建立一个俗称“垃圾邮件陷阱”（蜜罐）的电子邮件地址，随时向垃圾邮件信息交换中心报告新的垃圾邮件信息，保证反垃圾规则的及时有效。
- ◆ 检查电子邮件的首部信息是否符合各种Internet标准（例如：数据的格式是否正确？）
- ◆ 检查电子邮件的首部和内容部分的信息是否包含一些垃圾邮件中常见的、用各种语言写成的短语或是句子（例如：“快速致富”或是退定该邮件的一些方法描述）。
- ◆ 将首部和内容部分的“校验和”与多个垃圾邮件信息在线数据库中的“校验和”作比较来确认垃圾邮件。
- ◆ 检查邮件发送者的IP地址是否在一些在线的黑名单站点列表中（这是一些已经被垃圾邮件发送者利用，或是怀疑被垃圾邮件发送者利用的站点）。
- ◆ 自定义地址、主机名和域名的白名单（白名单中的邮件不会作为垃圾邮件）和黑名单（上了黑名单的邮件会被当作垃圾邮件）。
- ◆ 使用管理员/用户指定的一些垃圾邮件和非垃圾邮件的例子来训练垃圾库识别用户收到的各种不同的垃圾邮件。

操作说明：

垃圾邮件的培训分为“垃圾邮件”培训和“正常邮件”培训两个部分，两个部分相互独立分开，而又相互影响相互作用。

邮箱级过滤：

进行垃圾邮件培训前，要先清空垃圾培训库。在邮箱的收件箱中选择“典型”的垃圾邮件，点击【作为垃圾邮件】，对“垃圾邮件”进行培训。

在邮箱的收件箱中选择“典型”的正常邮件，将它们【转移】到【不明邮件箱】，选择它们，并点击【不作为垃圾邮件】，对“正常邮件”进行培训。

通过以上两步，就完成了邮箱级的垃圾邮件库培训。

邮局级过滤:

邮局管理员可以先针对一个邮箱做标准化的过滤培训。点击左菜单的【邮局参数设置】，再选择【垃圾邮件设置】，选择“标准”邮箱作为整个邮局的过滤标准库，点击【确定】按钮设定邮局级垃圾邮件过滤库；点击【清空过滤库】，清空邮局级垃圾邮件过滤库，不做邮局级过滤，或者重新培训

系统级过滤:

系统管理员可以根据【蜜罐】和用户举报收集到的垃圾邮件信息，针对一个邮箱做标准化的过滤培训。然后登录目标服务器的系统管理页面，点击左菜单的【垃圾邮件管理】，选择培训好的过滤库文件，点击【上传】，文件自动更新原有的垃圾邮件过滤库

注意事项

“垃圾邮件”培训和“正常邮件”培训都是必不可少的，用户一般会忽略“正常邮件”的培训导致过滤效果不好，正确的方式是培训同样数量的邮件，一般以各 100~200 封为宜。培训时，注意邮件的“典型”性，相类似的邮件不必重复多次培训，注意针对不同的样本，区分中英文，大面积多类型的进行培训

培训时，切忌反复，不要一会儿培训一封邮件为垃圾邮件，一会儿又培训它或者类型相似的邮件为正常邮件，以免导致系统的分值评判标准紊乱

一个垃圾邮件库并不能持续使用太长时间，如果觉得当前库使用不理想或者已经出现分值评判标准紊乱，应该清空重新培训。一个原因是垃圾邮件技术在不断的进步，花样翻新，需要样本库不断的跟进更新；另一个原因是样本库过大时，再持续培训，效果不明显而且容易导致分值评判标准紊乱