

doi:10.3969/j.issn.1000-2162.2015.02.005

# 基于疫苗机制的垃圾邮件过滤模型

蒋亚平<sup>1</sup>,田月霞<sup>1</sup>,梅晓<sup>1</sup>,卢林<sup>2</sup>

(1. 郑州轻工业学院,计算机与通信工程学院,河南 郑州 450001;2. 安徽省财政信息中心,安徽 合肥 230061)

**摘要:**针对传统的垃圾邮件过滤方法不能有效阻止出现的新型垃圾邮件的问题,借鉴生物免疫系统中疫苗的概念和免疫记忆功能,提出了一种基于疫苗机制的垃圾邮件过滤模型 SFM-V(spam filtering model based on vaccine mechanism)。该模型详细描述了垃圾邮件检测器的演化和抗原提呈的过程,通过疫苗控制器实现疫苗提取和疫苗接种,新生成的免疫记忆细胞作为疫苗实现信息交互,共享抗体。并引入小生境免疫记忆与共享机制,增加检测器的多样性及稳定性,促进免疫记忆库及原始抗体库中优良个体的保存,有效地提取和记忆垃圾邮件的未知特征和变异特征。利用CCERT(china education and research network)的邮件样本集对该模型进行训练和检测,仿真实验结果表明该模型有效地提高了垃圾邮件的正确率、召回率等特性,降低了垃圾邮件的虚报率。

**关键词:**人工免疫;垃圾邮件;检测器;疫苗;免疫记忆

中图分类号:TP393

文献标志码:A

文章编号:1000-2162(2015)02-0024-08

## A improved spam filtering model based on vaccine mechanism

JIANG Ya-ping<sup>1</sup>, TIAN Yue-xia<sup>1</sup>, MEI Xiao<sup>1</sup>, LU Lin<sup>2</sup>

(1. School of Computer and Communication Engineering, Zhengzhou University of Light Industry, Zhengzhou 450001, China;  
2. The Center of Financial Information of Anhui Province, Hefei 230061, China)

**Abstract:** For the traditional spam filtering method could not effective prevent new spam problem, referenced to the concept of the vaccine in biological immune system and immune memory function, proposed a spam filtering model based on the mechanism of the vaccine. The model described the evolution of spam detector and the process of antigen presentation, the extraction and vaccination vaccine was achieved by vaccine controller, new generation of immune memory cells as a vaccine information interaction, sharing antibodies; and the introduction of immune memory niche and the sharing mechanism, increased the detector diversity and stability, promoted the preservation of the excellent individual of the original antibody immune memory library and the library, effectively extract and memory unknown and variation characteristics of the spam. Using CCERT mail sample set for training and testing the model, simulation results show that the models effectively improve the accuracy, the recall rate and other characteristics of spam, reducing the rate of false spam.

**Key words:** artificial immunity; spam; detector; vaccinum; immune memory

---

收稿日期:2014-10-12

基金项目:国家自然科学基金资助项目(61272038);河南省科技厅科技攻关资助项目(0624220084);河南省教育厅计划项目(2010A520044)

作者简介:蒋亚平(1970-),男,河南永城人,郑州轻工业学院副教授,研究生导师,博士。

在网络日益发达的今天,电子邮件因其快速、便捷、经济等特点已逐渐成为目前主要的通讯方式之一,但垃圾邮件<sup>[1]</sup>问题也日益严重,垃圾邮件不仅阻碍网络正常信息的传递,甚至有些垃圾邮件含有病毒木马,对网络安全构成极大的威胁.因此,有效地过滤垃圾邮件的收发成为目前技术研究的热点之一.

在垃圾邮件过滤方法中,对邮件内容的过滤被认为是目前最有效的过滤方法之一.它主要包括朴素贝叶斯(Bayesian)<sup>[2]</sup>、K 邻近算法<sup>[3]</sup>、支持向量机<sup>[4]</sup>和神经网络<sup>[5]</sup>等文本过滤方法.这些方法在一定程度上有效地实现了垃圾邮件的检测和过滤功能,但对于邮件中变异的特征或新出现的特征则往往不能及时提取及发现.近年来,人工免疫系统 AIS(artificial immune system)作为生物智能计算的模仿已被成功运用到各领域<sup>[6]</sup>,采用人工免疫技术构造过滤效率高的系统也逐渐成为研究的热点. Andrew Secker 等将 AIS 应用于邮件分类,提出了 AISEC(artificial immune system for e-mail classification)算法模型<sup>[7]</sup>.该模型使用独立的记忆和成熟抗体集、基因库和自体库,具有良好的多样性、适应性和自学习性,但自体库和基因库更新速度不快.杨广明等实现了一种基于 ARTIS (artificial immune system) 人工免疫模型的反垃圾邮件模型,该模型将垃圾邮件看作侵入系统的抗原,模拟抗体消灭抗原的机理<sup>[8]</sup>,但检测器的分布性和合作性还有待提高.梁刚等借鉴生物免疫系统能有效地区别自体与非自体防御外部病原体入侵的工作机制,结合用户行为特征模式,提出了一种基于免疫机制的垃圾邮件过滤器模型<sup>[9]</sup>,该模型有效地提高了检测率,而且还表现出良好的自学能力与自适应性,但检测器中的抗体无法及时交互信息.

结合垃圾邮件过滤的实际需要并针对以上模型的不足,作者提出了一种基于疫苗机制的垃圾邮件过滤模型(SFM-V).该模型引入疫苗机制,使成熟检测器和记忆检测器中的抗体以疫苗作为媒介可以共享抗体,通过引入小生境技术,共享抗体,提高检测器的灵活性,有效的提取和记忆垃圾邮件的未知特征和变异特征.

## 1 过滤模型设计

在模型中共设计了3个阶段:检测器产生进化阶段(虚线部分)、抗原呈现阶段(实线)、疫苗引入阶段(点线).在模型中,设计了独立的疫苗模块,可在时间上控制疫苗接种和疫苗提取的时机,疫苗抗体可以共享信息.在模型中利用发生器传输运作信号,实现疫苗提取和疫苗接种.模型体系结构如图1.

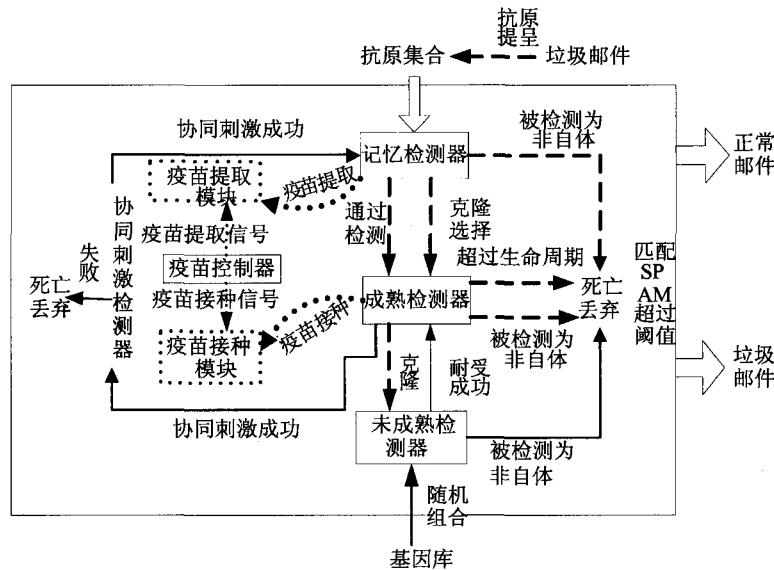


图 1 模型体系结构图

## 1.1 检测器演化及基因库生成

在论文的模型中,将检测器划分为:未成熟检测器(immature,简称  $I$ )、成熟检测器(mature,简称  $M$ )和记忆(ramaxel,简称  $R$ )检测器.  $D$  为检测器集合,且  $I, M, R \subset D$ .

垃圾邮件样本采集后,对其进行提呈,将其转换成数据集合向量  $E$ ,然后递交给检测器进行检测.如果数据被记忆免疫检测器识别,直接判断它为垃圾邮件,并向用户发出提示信息;如果记忆检测器匹配到了自体数据量集合  $SP$ ,则说明垃圾邮件特征出现变化,记忆检测器则不能再用来检测样本,将其删除.

成熟检测器( $M$ )位于检测器演化的中间阶段.经记忆检测器检测通过数据,交给成熟检测器处理.在生命周期内,当  $num$  值达到匹配阈值且协同刺激成功时,则进化为记忆检测器;不能达到匹配阈值,协同刺激失败,则死亡丢弃.即

$$\begin{cases} M = M - d_M, & d_M \geq \gamma \wedge d_M, T_{\text{period}} < \delta. \\ R = R + d_M, & \end{cases} \quad (1)$$

为了提高检测器的多样性,选取一定比例的记忆和成熟检测器,通过克隆选择或者随机的方式产生一批新的未成熟免疫检测器  $d_I$  ( $d_I \in I$ ).  $d_I$  可能检测到正常的自体数据集合  $SP$ ,再对其进行自体耐受.

自体耐受过程如下

$$I_{\text{tolerance}}(d_I, SP) = \begin{cases} \text{success, } A_{\text{fitness}}(d_I, SP) < \delta, \\ \forall E_{sp} \in SP, d_I, T_{\text{period}} \geq \beta, \\ \text{fail, 其他,} \end{cases} \quad (2)$$

其中: $SP$  为自体数据集合; $\delta$  为匹配阈值; $\beta$  为耐受期阈值; $E_{sp}$  为自体数据集的元素.只有成功通过自体耐受的  $d_I$  才能继续存活,才有资格进化为成熟免疫检测器;否则,  $d_I$  将死亡.

邮件处理模型将邮件分为正常邮件和垃圾邮件,该模型主要是基于内容的垃圾邮件过滤.在模型中考虑文本单词的 MI(mutual information)互信息对于邮件分类的影响,在新的自体库和检测器生成过程中,通过计算 MI 互信息选取特征单词组成自体库和基因库,从而控制自体库和基因库规模<sup>[10]</sup>.对于每个单词,计算词和类别之间的平均互信息如下

$$MI_{\text{avg}}(W, K_i) = \sum_{i=1}^{|k|} P(K_i) \log \left| \frac{P(W | K_i)}{P(W)} \right|, \quad (3)$$

其中: $P(K_i)$  表示第  $i$  类文本在基因库中出现的概率; $P(W)$  表示特征词  $W$  在基因库中出现的概率; $P(W | K_i)$  表示当文档属于第  $i$  类的情况下出现特征词  $W$  的概率.

如果邮件中提取的某个特征项(如  $K_i$  类)提供的信息量远大于对其他类提供的信息量,说明单词对  $K_i$  类的区分能力比较强<sup>[11]</sup>,应加入到选取的特征信息中,所以利用互信息差值作为评价,如下式

$$W(t) = MI_{\text{max1}} - MI_{\text{max2}}. \quad (4)$$

利用式(4)改进后的互信息计算,其中: $MI_{\text{max1}}$  为互信息中的最大值, $MI_{\text{max2}}$  为次大值.设定一个阈值,大于这个阈值的词为特征值(阈值的选取和训练样本关),从而控制自体库和基因库规模.

## 1.2 抗原提呈

邮件样本中,数据经过抗原提呈,形成抗原集合,经过特定模式的过滤,选取一定数量的抗原进行检测.文本电子邮件的标准是由两部分组成:信封和内容.信封包括完成传递和投递所需的所有消息,内容则是由信头和信体两部分组成.消息的开始主要包含 sender、receiver、theme 和 send time 等信息.

设定抗体信息决策表  $F$ ,若满足条件的抗原对其进行提呈,符合条件的归类到自体库中,否则加入到非自体库中.  $F = (U, R, Y, f)$ ,其中: $U = \{x_1, x_2, \dots, x_n\}$  表示抗体对象的非空有限集合<sup>[12]</sup>; $R = C \cup D$ ,  $C = \{c_i | i = 1, 2, \dots, m\}$  是抗体对象的条件属性, $D = \{d\}$  表示抗体决策属性的非空有限集合,其中, $R = C \cup D; C \cap D = \emptyset$ ; $Y$  为抗体属性值的集合, $V_c$  是属性  $c \in C \cap D$  的值域; $f$  是信息函数表示每个抗体对象的属性值, $U \times R \rightarrow Y$ ,即  $\forall c \in R, x \in U$ ,有  $f(x, c) \in Y$ .

**定义 1** 其中的每一个抗体属性子集  $P \in R$  决定一个二元等价关系  $IND(P)$ ,用公式表示为

$$IND(P) = \{(x, y) \in U \times U | \forall b \in P, f(x, b) = f(y, b)\}.$$

$IND(P)$  将  $U$  划分为  $k$  个类  $X_1, X_2, \dots, X_k$ ,其代表不同的抗体类型,即

$$IND(P) = [X]IND(P) = [X_1, X_2, \dots, X_k]. \quad (5)$$

集合  $X$  关于  $P$  的下近似定义为

$$\underline{PX} = U \{ Y_i \mid Y_i \in U / \text{IND}(P) \wedge Y_i \in X \}. \quad (6)$$

Q 的 P 正域记为

$$\text{POS}(Q) = \bigcup_{X \in U/Q} \underline{PX}. \quad (7)$$

P 与 Q 为等价关系族,用约简定义表示为:  $P = \{C - \{c_i\}\}$ , 如果

$$\text{POSIND}(P)(\text{IND}(Q)) = \text{POSIND}(P) - \{c_i\}(\text{IND}(Q)),$$

则称  $c_i$  为 P 中 Q 不必要部分,归类到非自体规则库;否则  $c_i$  为 P 中 Q 必要部分,归类到自体规则库.

算法描述如下:

输入:抗体形成的决策表;输出:自体规则库,非自体规则库.

1) 由式(5)计算抗体条件属性  $X\{x_1, x_2, \dots, x_n\}$  的等价集;

2) 计算抗体决策属性 D 的等价集;

3) for 每一等价集 do;

4) 由式(6)计算抗体决策属性的下近似集;

5) 由式(7)计算  $\text{POSP}(X, D)$ ;

6) end for;

7) for 每个属性  $x_i$  do //属性约简;

8) 计算条件属性  $X - x_i$  的等价集;

9) 由式(6)计算抗体属性的各等价集的下近似集;

10) 由式(7)计算  $\text{POS}(X\{x_i\}, D)$ ;

11) If  $(C - \{c_i\}, \text{POS}(C - \{c_i\}), Q) = \text{POSC}(Q)$ ;

12) Then  $C' = \{C - \{c_i\}\}$  //删除  $c_i$  所在的列且合并 U 中重复的行;

13) Else  $C' = C$ .

将不必要的信息删除后,则把剩余关键词组成新的向量放入到抗原特征向量中,形成自体规则库.

这样既保证了选择的特征向量对邮件内容的代表意义,同时又限制了特征向量长度的过度膨胀.

### 1.3 疫苗引入

在疫苗引入阶段,主要包含疫苗提取模块与疫苗接种模块,两模块由疫苗控制器发出指令协调配合实现.在模型中疫苗提取选用动态提取与接种,即在检测器需要的时候疫苗随时接种或注入.模块的工作周期是从开始收到接种或提取疫苗的信号开始,直到疫苗提取或接种周期结束.

#### 1.3.1 疫苗提取

疫苗提取是指抽取抗原刺激的数据特征转化成有抵抗能力的检测器.在模型中,疫苗提取的过程即是对垃圾邮件特征信息的提取过程.

在模型中,疫苗提取采用动态方式,即从记忆检测器中的抗体提取数据特征转化为疫苗.模型中疫苗提取步骤:

步骤 1 疫苗控制器传输疫苗提取信号,模块开始工作(疫苗提取源来自记忆检测器集合).

步骤 2 对记忆检测器集合进行分类加入进入分类器,分成若干份;然后,在每份分类器中随机抽取一定数目的抗体因子进入未成熟疫苗种群集合.

步骤 3 对未成熟疫苗集合进行特征信息提取,计算疫苗亲和力,若亲和力不符合要求,进入分类器,符合要求,转化成为成熟疫苗.

$f(x)$  表示疫苗之间亲和力函数.亲和力的值用  $[0, 1]$  之间的实数表示,设  $A_1, A_2$  为两个不同的疫苗,则  $A, A_2$  之间的亲和力定义为

$$f(A_1, A_2) = \frac{A_1 \cap A_2}{\min(A_1, A_2)}. \quad (8)$$

**定义 2**  $M_k = \{J_{i,k} \mid i=1, 2, \dots, n\}$  表示为  $k$  代未成熟疫苗种群,  $n$  代表成熟疫苗种群数量,  $T_b$  表示成熟疫苗集合,使用  $A_k$  表示  $k$  代未成熟疫苗种群中最佳个体疫苗.

$T_{\text{mature}}$  表示成熟疫苗的所用时间,  $T_{\text{pick-up}}$  表示疫苗提取周期,  $T_b(\text{size})$  表示成熟疫苗集合的大小. 提取成熟疫苗算法流程如图 2 所示.

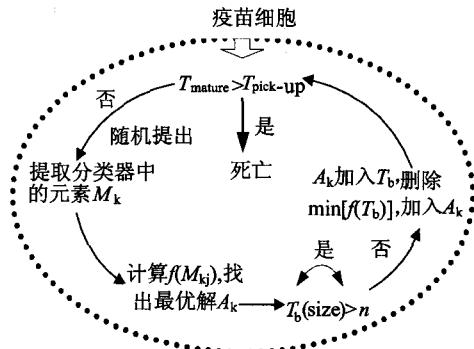


图 2 提取成熟疫苗算法流程图

Fig. 2 Mature vaccine algorithm flow chart

**定义 3** 疫苗  $V_a \in H, H = \{0, 1, *\}^l (l \in N, l > 0)$ , 定义  $V_a$  由  $0, 1, *$  组成的长度为 1 的二进制串,  $V_a^m$  表示  $V_a$  的第  $m$  个基因位上的编码. 设抗体种群  $A$  的个体为  $a_1, a_2, \dots, a_n, a_i^m$  表示第  $i$  个抗体第  $m$  个基因位上的编码.

假设  $a_1, a_2, \dots, a_i (i \leq n)$  是抗体种群  $A$  在特定判断标准下的优良个体, 疫苗提取公式定义为

$$V_{a_i^m} = \begin{cases} 1, & \frac{1}{H} \sum_{i=1}^t a_i^m(j) > \theta_1, \\ 0, & \frac{1}{H} \sum_{i=1}^t a_i^m(j) < \theta_2, \\ *, & \text{其他}, \end{cases} \quad (9)$$

其中: 参数  $\theta_1, \theta_2$  的取值根据实际情况确定. 从定义 3 可以得出: 一个疫苗代表一种优良模式. 疫苗中除去取值为 \* 的基因, 称其为优良基因位.

### 1.3.2 疫苗接种

疫苗接种: 用提取的疫苗(包含合格疫苗和优良疫苗)来更改抗体的某些基因位, 将其优良基因传递给下一代, 提高优良模式繁殖的概率, 修复在交叉和变异操作中被损坏的优良基因.

当疫苗接种模块从接收到信号后开始工作, 疫苗接种目标是把抗体加入到成熟检测器集合. 疫苗接种过程描述如下:

步骤 1 疫苗控制器传输疫苗接种的信号, 模块开始工作(疫苗提取源来自成熟检测器集合).

步骤 2 随机抽取的一定数量的成熟检测器为待接种目标, 与注入的成熟疫苗集合, 在疫苗接种器发出信号后进行接种.

步骤 3 计算疫苗接种器中疫苗阈值, 若达到阈值, 进入待接种成熟检测器集合, 并添加到成熟检测器集合; 未达到匹配阈值, 排除删除.

**定义 4** 设疫苗  $V_a$ , 抗体为  $a_i$ , 疫苗接种操作为  $a_i = a_i \Theta V_{a_i}$ , 用  $a_i$  表示抗体  $a_i$  接种疫苗后的编码形式. 疫苗接种公式定义为

$$V_{a_i^m} = a_i^m \Theta V_{a_i^m} \begin{cases} V_{a_i^m}, & V_{a_i^m} = 0 \text{ 或 } 1, \\ a_i^m, & V_{a_i^m} = *. \end{cases} \quad (10)$$

### 1.3.3 小生境免疫记忆与共享机制

为提高检测器中抗体之间的共享性, 利用共享适应度小生境技术, 设计一种获取多种免疫记忆细胞的共享机制算法, 其目的是增强检测器中抗体的多样性并保存优良个体, 为其创造出小生境的进化环境. 通过小生境技术将每种类型中的抗原的优良个体保存于免疫记忆库中, 通过共享选择策略, 用抗体群及免疫记忆抗体群中个体间相似度的共享函数来调整各个体的适应度, 并分别进行进化, 增加检测器

的多样性及信息共享.

共享函数是关于两个个体之间关系密切程度的函数<sup>[13]</sup>. 将抗原视为问题, 抗体  $A_b$  为此问题的候选解, 抗原对抗体  $A_b$  的亲和力为  $f(A_b)$ , 群体规模指定为  $A'$ . 当两个抗体之间的关系比较密切时, 共享函数值较大, 其值接近 1; 否则, 其值接近 0. 在免疫算法中, 亲和力函数等值于适应值函数.

**定义 5** 假设群体为, 令  $S(d_{ij})$  表示抗体  $A_{bi}$  和  $A_{bj}$  之间的共享函数

$$S(d_{ij}) = \begin{cases} 1 - \left(\frac{d_{ij}}{r_s}\right)^\alpha, & d_{ij} < r_s, \\ 0, & d_{ij} \geq r_s, \end{cases} \quad (11)$$

其中:  $d_{ij}$  为个体间的距离测度, 采用欧式距离;  $r_s$  为小生境半径;  $\alpha$  为共享函数调整参数.

**定义 6** 对于群体, 个体  $A_{bi}$  在群体中的适应度为

$$S_i = \sum_{j=1}^n S(d_{ij}). \quad (12)$$

**定义 7** 基于个体的共享适应度值的调整公式为

$$f^*(A_{bi}) = \frac{f(A_{bi})}{S_i} (i = 1, \dots, n). \quad (13)$$

设抗体群  $A' = \{A_{b1}, A_{b2}, \dots, A_{bn}\}$  的亲和力函数表示为  $f(A_b)$ , 由欧几里德距离计算抗体间的距离测度. 根据  $d_{uv} < r_s$ , 确定局部小生境优良抗体库  $A'_s$ , 并初时化记忆抗体库  $A'_M$ . 事实上, 记忆单元也是一类抗体群, 定义记忆单元如下

$$A'_M = \{A_{bm1}, A_{bm2}, \dots, A_{bmM}\}. \quad (14)$$

如果  $A'_s$  中包含种群  $A'$  中适应度最高的抗体, 则保留其中一个最高亲和力的抗体, 并按公式(13)更新其余抗体亲和度. 利用更新后的亲和度及处罚函数对子种群中浓度高、亲和度低的抗体进行处罚. 按亲和力大小进行各抗体排序, 选择前  $M$  个抗体更新记忆抗体集, 记为  $A'_M$ .

模型通过采用小生境技术, 每种类型中的抗原的优良个体保存于免疫记忆库中, 使检测器之间可以互通信息, 共享抗体信息, 增加检测器的多样性, 使邮件特征信息及时共享.

## 2 仿真实验与分析

### 2.1 实验评测标准

模型使用准确率、召回率、精准率和虚报率作为衡量该模型效率的主要标准. 测试使用中国教育和科研网中文邮件样本集, 正常邮件 9 272 封, 垃圾邮件 25 088 封. 仿真计算模型的查准确率与查全率, 并与其他模型(如朴素贝叶斯、人工免疫)比较.

定义变量: 垃圾邮件(Spam), 合法邮件(Ham).  $N_{S \rightarrow S}$  表示 Spam 判断为 Spam 的数目;  $N_{H \rightarrow S}$  表示 Ham 判断为 Spam 的数目;  $N_{S \rightarrow H}$  表示 Spam 判断为 Ham 的数目;  $N_{H \rightarrow H}$  表示 Ham 判断为 Ham 的数目, 设测试集中共有  $N$  封邮件,  $N = N_{H \rightarrow S} + N_{H \rightarrow H} + N_{S \rightarrow S} + N_{S \rightarrow H}$ .

进而可以定义出准确率、精准率、虚报率和召回率等评价指标, 对邮件分类系统性能进行评估.

召回率:  $Recall = N_{S \rightarrow S} / (N_{S \rightarrow S} + N_{S \rightarrow H})$ , 即垃圾邮件被重新检测出的概率. 在模型中, 召回率越高, 检测出的垃圾邮件越多.

准确率:  $Precision = N_{S \rightarrow S} / (N_{S \rightarrow S} + N_{H \rightarrow S})$ , 即邮件为垃圾邮件, 对其判断正确的概率.

精准率:  $Accuracy = (N_{S \rightarrow S} + N_{H \rightarrow H}) / N$ , 即对所有邮件进行判断且判断正确的概率.

虚报率:  $Fallout = N_{H \rightarrow S} / (N_{H \rightarrow S} + N_{H \rightarrow H})$ , 即模型将正常邮件判为垃圾邮件的概率, 虚报率越大, 则模型将正常邮件判为垃圾邮件的可能性就越大.

### 2.2 仿真实验结果及分析

对该模型进行仿真实验, 将数据集分为训练集和测试集, 选择 2 360 封邮件(1 272 封正常邮件和 1 088 垃圾邮件样本)作为训练集, 得到初始检测器. 使用剩余的 8 000 封合法邮件和 24 000 封垃圾邮件平均分为 10 组, 组成测试集, 进行测试, 最后取 10 次实验的平均值作为最后的实验数据, 其平均召回率

与平均精准率即为模型的准确率与平均精准率. 图3是在阈值选取0.85、该模型与贝叶斯模型在仿真环境下进行垃圾邮件过滤实验所得的统计数据.

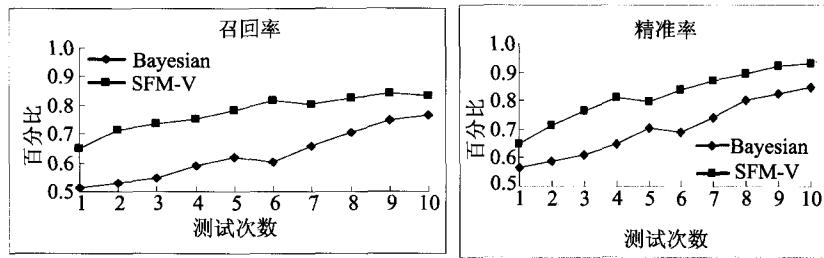


图3 精准率与召回率比较图

Fig. 3 Precision and recall comparison chart

从图3中可以看出,随着测试次数的增加,精准率和召回率在逐渐提高,这是因为随着检测器的增多,模型中抗体的自我学习,判别垃圾邮件的能力在逐渐提升,免疫细胞能力逐渐增强,训练出的检测器能更加精准地检测垃圾邮件;当测试到某一值时,模型的精准率与召回率均出现下降现象,这是因为在模型中出现了新的垃圾邮件,免疫记忆库中的记忆细胞无法及时识别邮件特征,模型需要重新学习和记忆垃圾邮件特征;到后期,由于抗体受基因库大小的影响,邮件的召回率与精准率基本保持稳定,但整体基于疫苗机制模型的效率要高于基于贝叶斯模型的效率.

为了进一步验证模型的效率,论文在同等条件下(阈值选取0.85)与基于贝叶斯的垃圾邮件过滤模型<sup>[14]</sup>和基于AIS的垃圾过滤模型<sup>[15]</sup>进行对比实验,如图4所示. SFM-V较基于贝叶斯方法的模型和AIS模型的正确率有大幅提高,且趋于稳定,误报率降低且波动较小.

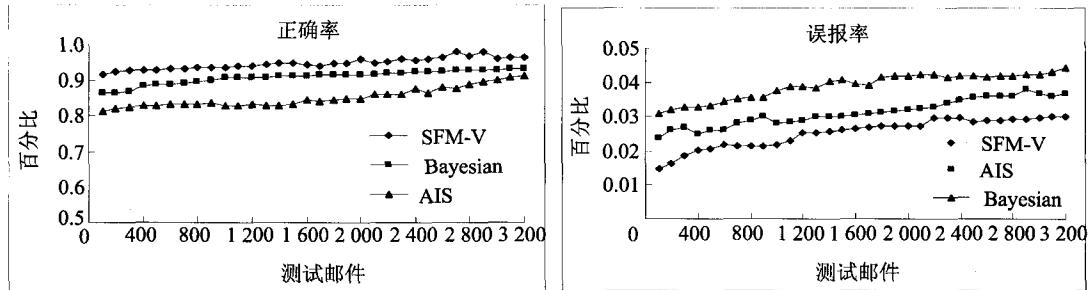


图4 模型效率比较图

Fig. 4 Model efficiency comparison chart

基于Bayesian的模型采用先验概率的规则来测试邮件,如果在检测阶段检测的邮件包含了很多在训练阶段邮件没有出现过的新词,即新特征信息,则该模型需要一段较长时间的学习适应,学习记忆能力较低,对邮件的正确率不是很高,在虚报率方面不稳定且垃圾邮件虚报率不精准;基于AIS的模型则可以分布式方式识别垃圾邮件,并能学习和记忆邮件的特征,对邮件判断正确的提升,虚报率有所降低,但是存在检测器的灵活性较差,在测试阶段不能精确地对垃圾邮件进行分类,记忆抗体特征;论文提出的模型SFM-V,在训练阶段产生的初始抗体,通过疫苗控制器实现疫苗提取和疫苗接种,不仅可以新生成的免疫记忆细胞作为疫苗实现信息交互,而且可以共享抗体,呈现出较高效率,且虚报率较低,又由于模型具有协同刺激机制,检测出的邮件不会被误删.

### 3 结束语

借鉴生物免疫系统中疫苗的概念和免疫记忆功能,提出一种基于疫苗机制的垃圾邮件过滤模型SFM-V.该模型通过引入MI互信息,控制自体库和基因库规模;在抗体提呈过程中构造抗原决策表,对抗原进行提呈,可以有效地对垃圾邮件进行分类并提取和记忆邮件特征;在疫苗进入阶段,通过疫苗接种和提取增加了检测器的灵活性,并可以使抗体共享,有效识别未知特征及变异特征等特点.实验结果表明该模型较其他模型具有更好的性能,而且有效地提高了垃圾邮件过滤效率.

**参考文献:**

- [1] Gansterer W, Ilger M, Neumayer P, et al. Anti-spam methods state-of-the-art[D]. Vienna: Faculty of Computer Science, University of Vienna, 2005: 1-99.
- [2] Marsono M N, El-Kharash M W, Gebali F. Targeting spam control on middle boxes: spam detection based on layer-3 e-mail content classification[J]. Computer Networks, 2009, 53(6): 835-848.
- [3] Mehmet A, Cigdem I, Mutlu A. Bayesian methods and genetic algorithm[J]. Expert Systems With Applications, 2010, 37(7): 5061-5067.
- [4] Yu B, Xu Z B. A comparative study for content-based dynamic spam classification using four machine learning algorithms[J]. Knowledge-Based Systems, 2008, 21(4): 355-362.
- [5] Clark J, Koprinska I, Poon J. A neural network based approach to automated e-mail classification[C]. Web Intelligence: Proceedings of the 2003 IEEE/WIC International Conference on Web Intelligence, 2003: 13-17.
- [6] Qing J J, Mao R L, Bie R F, et al. An AIS-based e-mail classification method[C]//The 2009 International Conference on Intelligent Computing, Ulsan, Korea, 2009: 492-499.
- [7] Secker A, Freitas A, Timmis J. AISEC: an artificial immune system for e-mail classification[C]//The 2003 Congress on Evolutionary Computation, California, USA, 2003: 131-138.
- [8] 刘凤玲,杨广明,王欣艳,等. ARTIS 人工免疫模型在邮件过滤中的研究与应用[J]. 小型微型计算机系统, 2007, 28(7): 1293-1296.
- [9] 梁刚,刘晓洁,李涛,等. NSC:一种新型的垃圾邮件过滤器[J]. 小型微型计算机系统, 2008, 29(1): 158-161.
- [10] 黄珏,陈兵,廖常武. 改进的人工免疫垃圾邮件过滤算法[J]. 计算机工程与应用, 2011, 47(30): 72-74.
- [11] 李霞,蒋盛益. 一种垃圾邮件快速识别方法[J]. 小型微型计算机系统, 2013, 34(3): 498-502.
- [12] 张玲,白中英,罗守山,等. 基于粗糙集和人工免疫的集成入侵检测模型[J]. 通信学报, 2013, 34(9): 166-176.
- [13] 徐雪松. 基于免疫记忆共享机制的工业数据约简方法[J]. 计算机集成制造系统, 2013, 19(11): 2864-2870.
- [14] Zhang L, Zhu J B, Yao T S. An evaluation of statistical spam filtering techniques[J]. ACM Transactions on Asian Language Information Processing (TALIP), 2004, 3(4): 243-269.
- [15] Guzella T S, Mota-Santos T A, Uchôa J Q, et al. Identification of SPAM messages using an approach inspired on the immune system[J]. Biosystems, 2008, 92(3): 215-225.

(责任编辑 朱夜明)