# 面向入侵检测的集成人工免疫系统

陈岳兵[1]，冯超[2]，张权[2]，唐朝京[2]

(1. 总参第六十一研究所，北京 100141；2. 国防科技大学 电子科学与工程学院，湖南 长沙 410073)

**摘 要**：结合入侵检测的实际需求，提出了一种集成人工免疫系统(IAIS)。该系统结合了树突状细胞算法(DCA)和否定选择算法(NSA)，DCA 用于检测行为特征，NSA 用于检测结构特征。通过 KDD99 数据集实验对该系统进行验证，并与其他方法进行了比较。实验结果表明，IAIS 检测性能与经典分类算法相当。IAIS 具有不依赖明确标识的数据来训练检测器，可结合行为特征和结构特征进行实时入侵检测的特点。

**关键词**：人工免疫系统；集成人工免疫系统；树突状细胞算法；否定选择算法；入侵检测系统

**中图分类号**：TP391.4 **文献标识码**：B **文章编号**：1000-436X(2012)02-0125-07

# Integrated artificial immune system for intrusion detection

CHEN Yue-bing[1], FENG Chao[2], ZHANG Quan[2], TANG Chao-jing[2]

(1. The No.61 Research Institute of General Staff, Beijing 100141, China;

2. School of Electronic Science and Technology, National University of Defense Technology, Changsha 410073, China)

**Abstract:** According to the practical requirements of intrusion detection, an integrated artificial immune system (IAIS) was proposed. The system combined dendritic cell algorithm (DCA) and negative selection algorithm (NSA). DCA was used to detect behavioral features. NSA was used to detect structural features. IAIS was validated on KDD 99 dataset. Comparisons to other approaches were made. The experimental results show that the detection performance of IAIS is comparable to classic classification algorithm. IAIS does not rely on labeled data to train detectors. It combines behavioral features and structural features to detect intrusions in real-time mode.

**Key words:** artificial immune system; integrated artificial immune system; dendritic cell algorithm; negative selection algorithm; intrusion detection system

## 1 引言

人体免疫系统保护人体免受各种病原体的入侵，为入侵检测系统提供了丰富的启发。研究者应用免疫学原理来设计智能计算范例，逐渐发展成一个新的研究领域，称为人工免疫系统[1]。研究者分别基于否定选择机制和危险模型[2]提出了 NSA 和 DCA[3]。2 种算法都可以用于入侵检测，但前者可用性欠佳，后者检测性能偏低，在应用方面都有局限性。

在入侵检测研究中，追求高检测率、低误报率[4]和快速的检测速度。根据检测方法不同，分为误用检测和异常检测。误用检测又称为特征检测，它是通过对已知入侵的研究，提取攻击的特征形成特征集合，利用这些特征集合对当前的数据进行各种处理后，再进行特征匹配工作，如果匹配成功，则判定为入侵。该方法能够很好地检测已知的入侵，具有检测率高、误报率低和检测速度快等特点，缺点是它不能检测未知的入侵。当有未知的入侵出现的时候，模式库必须随之更新，给系统的维护管理带

**收稿日期**：2011-03-07；**修回日期**：2011-06-08

**基金项目**：国家自然科学基金资助项目（60872052）

**Foundation Item:** The National Natural Science Foundation of China (60872052)

来困难。异常检测又称为基于行为的检测，它是利用与系统行为相关的一些统计量来构造正常的行为模式[5]，如果待检测行为偏离正常行为，则判定为入侵。该方法存在统计量选取困难、异常阈值确定困难、误报率较高的问题，但是它能对未知的入侵进行检测。

本文对危险模型进行了分析，指出否定选择机制和危险模型两者之间并不矛盾：在危险模型中，否定选择机制和危险性判别机制在不同的位置发挥各自的作用，两者协作完成检测和应答。基于这种新的认知，结合实时 DCA 和 NSA 构建了一个集成人工免疫系统(IAIS)，用于入侵检测。实时 DCA 检测行为特征，NSA 检测结构特征，构建的 IAIS 可以同时利用结构特征和行为特征进行入侵检测。通过入侵检测基准数据集对 IAIS 进行验证，并与其他方法进行了比较。IAIS 检测性能与经典分类算法相当，且具有良好的可用性。

## 2 启发

在免疫学中，感染非我模型[6]和危险模型[2]是 2 种主要的免疫学模型，两者都有各自的支持者和部分证据，都是具有一定合理性的解释。表面上看，2 种模型互相矛盾，但是危险模型的提出者 Matzinger 从未明言感染非我模型是错误的，也没有讲否定选择是不正确的。Matzinger 认为否定选择是一种避免自我免疫的前置条件，只是在免疫应答之前还需要进一步的确认机制以消灭有害抗原。危险模型在图 1 中给出，图中 APC(antigen presenting cells)、B、Tk 和 Th 分别表示抗原提呈细胞、B 细胞、毒性 T 细胞和辅助 T 细胞。APC 属于固有免疫系统；B 细胞、毒性 T 细胞和辅助 T 细胞属于适应性免疫系统。图中共有 3 类信号：抗原和 B 细胞受体，毒性 T 细胞以及 APC 之间的作用产生信号 1；辅助 T 细胞和 APC 释放信号 2；身体组织的正常细胞在受到损害的时候释放信号 0，或危险信号，用于激活本地 APC。感染非我模型和危险模型的唯一区别在信号 0。在感染非我模型中，APC 通过模式识别受体常见的非我分子，产生刺激信号。而在危险模型中，APC 感知组织中受损细胞释放的的危险信号。

感染非我模型和危险模型都包括抗原与 B 细胞以及毒性 T 细胞之间的结构匹配。虽然 2 种模型不可能同时正确，这并不表示这两者完全不相容。实际上，否定选择和自我/非我模型之间不能划等号。在图 1 中，否定选择是作为危险模型中的部分机制存在的，否定选择和危险性判别都是必要的。据此，提出结合危险性判别机制和否定选择机制，建立集成系统。否定选择机制负责检测结构特征，危险性判别机制负责检测行为特征。

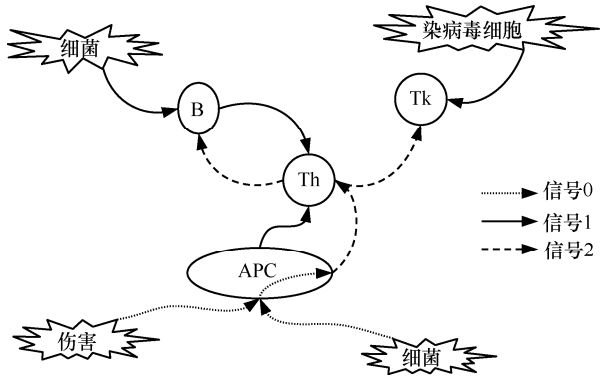


图 1 危险模型

## 3 集成人工免疫系统

本节根据对免疫系统的新的认识，提出结合 DCA 和 NSA 建立 IAIS，描述了系统结构及其实现，给出了系统的使用方法。

### 3.1 系统概述

IAIS 总体结构如图 2 所示，系统基本组件及其功能描述如表 1 所示。IAIS 集成的 2 个算法分别为实时 DCA 和二进制表示的 NSA。

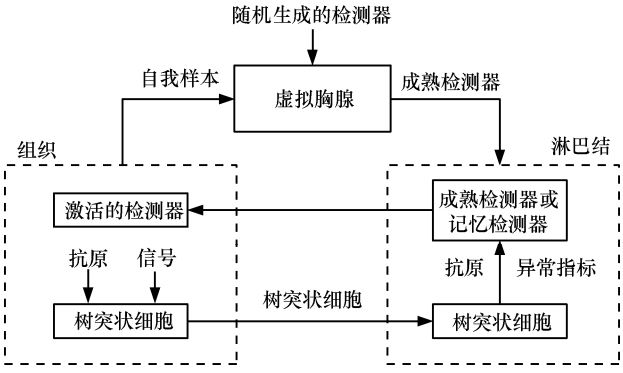


图 2 IAIS 结构

IAIS 的功能是检测目标系统中的异常实体，如图 2 所示。组织代表目标系统；抗原是组织中运行的实体；信号是与组织运行状态相关的观测值；树突状细胞负责将最新的抗原和信号从组织传送到淋巴结；淋巴结是一个分析中心，提呈的抗原和信号在此经过分析之后得到代表抗原异常程度的指

标；虚拟胸腺负责生成成熟检测器。淋巴结中分析得到的异常指标用于辅助成熟检测器和记忆检测器对抗原的检测。

**表 1　　IAIS 组件**

| 组件 | 功能 |
|---|---|
| 组织 | 处理输入以得到信号和抗原 |
| 淋巴结 | 存储树突状细胞并检查其状态 |
| 虚拟胸腺[7] | 生成检测器 |
| 树突状细胞 | 在组织中采集抗原和信号，进行处理，迁移到淋巴结并执行抗原提呈 |
| 检测器 | 检测异常抗原 |

**算法 1**　IAIS 伪代码

输入：抗原流，经过预处理的信号流
输出：检测到的异常抗原
while 输入数据 do
　　if 组织状态正常 then
　　　　将输入抗原加入自我集合
　　end if
　　检查和生成 NSA 的检测器
　　通过实时 DCA 处理抗原和信号
　　执行实时分析以得到动态异常指标
　　根据动态异常指标计算匹配阈值
　　通过 NSA 的检测器检测抗原
end while

算法 1 描述了 IAIS 的运行过程。系统根据组织的当前状态维持一个动态的自我集合，组织状态异常时出现的抗原被禁止加入自我集合。然后检查 NSA 中记忆检测器和普通检测器的生命周期，清除过期检测器，生成新的检测器以填补空缺。实时 DCA 对输入抗原和信号进行处理以得到不同抗原类型的动态异常指标。

在 IAIS 中，提出结合抗原动态异常指标计算检测器和抗原之间的匹配阈值，如式(1)所示。通过汉明距离计算检测器和抗原之间的亲和度并与动态匹配阈值进行比较，亲和度在匹配阈值之上的判为异常，反之判为正常。此过程连续运行，不断处理新输入数据。

$$Y_\alpha = (L-\varepsilon)\mathrm{e}^{-a(X_\alpha-\delta)} \tag{1}$$

其中，$X_\alpha$ 是抗原类型 $\alpha$ 的动态异常指标，$Y_\alpha$ 是此类抗原的匹配阈值，$a$ 是常数，$\delta$ 是实时 DCA 中的异常阈值，$L$ 是检测器长度，$\varepsilon$ 是自我半径[8]，它决定检测器在二进制空间的覆盖范围。对记忆检测器取更大的 $a$ 值以降低器匹配阈值。根据式(1)，抗原类型的异常程度越高，匹配阈值越低；反之，抗原类型的异常程度越低，匹配阈值越高。

### 3.2　系统实现

本节描述了实时 DCA 和 NSA 实现。

1) 实时 DCA

实时 DCA 的基本功能是接收输入信号和抗原，输出每一类抗原的动态异常指标。多路信号源作为算法输入，分为 PAMP (pathogen-associated molecular pattern)、危险信号(danger signal)、安全信号(safe signal)和炎性信号(inflammation signal)，各信号的含义如下。

PAMP：表明出现异常行为，此信号增强与高置信度异常行为的存在密切相关。

危险信号：表明可能存在异常行为，此信号增强表明异常行为存在的可能性增加。

安全信号：表明有正常行为发生，此信号增强表明正常行为的可能性增加。安全信号用于抵消 PAMP 和危险信号的影响。

炎性信号：表明总体状态异常。炎性信号用于放大其他信号。

实时 DCA 通过式(2)处理输入信号得到输出信号。输出信号包括协同刺激信号(csm)、半成熟信号(semi)和成熟信号(mat)。

$$O_j = (1+I)\sum_{i=0}^{2} w_{ij}S_i \tag{2}$$

其中，$S_i$ 是输入信号，$O_j$ 是输出信号，$I$ 是炎性信号，$w_{ij}$ 是从 $S_i$ 到 $O_j$ 的转换权值。

实时 DCA 是在 DCA[9]基础上进行改进提出的，主要包括 2 点改进。第一，实时 DCA 中增加了树突状细胞在淋巴结中的生命周期。为简化起见，让树突状细胞在淋巴结中存活一个系统周期，执行抗原提呈；第二，实时 DCA 提出了动态异常指标，在 MAC[9]基础上提出动态 MAC($C_\alpha$)，如式(3)所示。

$$C_\alpha = \frac{m_\alpha}{\sum_{i=1}^{A} A_i} \tag{3}$$

其中，$\alpha$ 是所有具有相同值的抗原结合，$m_\alpha$ 是抗原类型 $\alpha$ 被提呈为成熟抗原的数量，$A_i$ 是抗原类型 $i$ 被提呈的总数，$A$ 是抗原类型总数。$m_\alpha$ 和 $A_i$ 是根据当前淋巴结中存活的树突状细胞统计得到的，过期的树突状细胞不在统计之列。$C_\alpha$ 在 0 到 1 之间，$C_\alpha$ 越靠近 1，表明此类抗原异常程度越高。

实时 DCA 伪代码如算法 2 所示。组织和 T 细胞种群被实现为 2 个串行子系统，树突状细胞负责将信息从组织传送到淋巴结，T 细胞种群对信息进行分析。

**算法 2** 实时 DCA 伪代码

```
输入：抗原流，经过预处理的信号流
输出：各种抗原的动态异常指标
(组织)
初始化树突状细胞种群
while 输入数据 do
    更新组织抗原结构和信号矩阵
    for 每个树突状细胞 do
        获取并存储抗原
        获取并处理信号
        if cumulative csm≥迁移阈值 then
            if cumulative mat≥cumulative semi then
                cell context = 1
            else
                cell context = 0
            end if
            树突状细胞迁移到淋巴结
            组织中增加一个新树突状细胞
        end if
    end for
end while
(T 细胞种群)
for 每个树突状细胞 do
    计算每种抗原类型的异常指标
end for
计算并记录抗原类型及对应的动态异常指标
```

在组织中，首先初始化树突状细胞种群，为每个树突状细胞赋予一个迁移阈值，用于决定其在组织中的生命周期。迁移阈值是指定范围内的随机数，这使得实时 DCA 成为一个随机系统，自然地，IAIS 也是一个随机系统。当有输入数据的时候，组织处理数据得到信号和抗原，然后组织中每个树突状细胞经历一个累积信号并走向成熟过程[9]。与 DCA 不同的是，实时 DCA 中达到迁移阈值的树突状细胞要迁移到淋巴结中，并执行抗原提呈。

T 细胞种群对淋巴结中的树突状细胞进行实时分析，并生成异常阈值的时间序列，逐代记录与抗原类型对应的动态异常指标。

2) NSA

实现的 NSA 采用二进制表示法，算法包括 2 个阶段：检测器生成阶段和检测阶段。检测器生成阶段如算法 3 所示，随机生成候选检测器，匹配自我样本的候选检测器被清除，否则加入检测器集合，当检测器数目达到预设值的时候算法终止。

**算法 3** NSA 检测器生成伪代码

```
输入：自我集合，自我半径，预设检测器数量，检测器长度
输出：检测器集合
while 当前检测器数量<预设检测器数量 do
    随机生成指定长度的候选检测器
    for 每个自我样本 do
        计算候选检测器和自我样本之间的亲和度
        if 候选检测器不在自我样本的自我半径之内 then
            将候选检测器加入检测器集合
        end if
    end for
end while
return 检测器集合
```

在检测阶段，计算输入样本和所有检测器的亲和度，如果样本在某一个检测器的自我半径之内，则将此样本判为异常。

### 3.3 使用过程

在实验之前，首先介绍如何使用 IAIS，其一般过程包括 4 个部分。

1) 定义抗原：抗原就是目标系统中运行的实体。IAIS 需要身份标识来表示这些实体，通常用数字和字符串形式的唯一标识符表示一个抗原或者一类抗原。

2) 抽取信号：通过各种信号表示系统状态。用户需要根据先验知识和分析抽取有用信号来代表整个系统状态，包括 PAMP、危险信号、安全信号和炎性信号。

3) 数据预处理：将原始信号转化为无量纲的数字，再对每类信号进行归一化处理，例如，将指定信号区间映射到[0,100]，常用的归一化函数包括线性函数，阶梯函数和 Sigmoid 函数，最后将抗原流和信号流提交给 IAIS。

4) 数据处理：IAIS 处理提交的抗原流和信号流，并输出每类抗原的动态异常指标。

## 4　实验测试

在 KDD99 数据集上进行实验，研究 IAIS 的属性，并通过与其他方法的比较来评估系统性能。

### 4.1　测试集

KDD99 数据集是一个用于入侵检测领域的基准数据集。实验采用 10%子集，它包括 494 021 个数据项（实例或样本）。10%子集与完整数据集具有类似统计特性，保持了类似的正常连接和攻击比。KDD99 数据集本身没有时间戳，为了测试 IAIS，每秒采样 10 个抗原来模拟实际情况，在数据集中加入时间戳，假设数据集以 1s 为间隔采集。KDD99 数据集的数据项是 41 维（特征或属性）向量。根据文献[10]介绍的数据域含义，对数据集进行预处理，得到抗原和信号。抗原代表结构特征，信号代表行为特征。

1) 抗原

将选择的数据域转换为 24bit 的二进制串，用来表示抗原，如表 2 所示。

**表 2　　抗原构造**

| 数据域 | 转换 | 比特数 |
|---|---|---|
| 25 | 如果值为 0，则为 0，否则为 1 | 1 |
| 26 | 如果值为 0，则为 0，否则为 1 | 1 |
| 2 | TCP，UDP 和 ICMP 分别用 00，01 和 10 表示 | 2 |
| 3 | 按首字母顺序编为 1 到 66，再将序号转换为二进制格式 | 7 |
| 4 | 按首字母顺序编为 1 到 11，再将序号转换为二进制格式 | 4 |
| 29 | 如果值为 1，则为 1，否则为 0 | 1 |
| 30 | 如果值为 0，则为 0，否则为 1 | 1 |
| 31 | 原始值乘以 100，再转换为二进制格式 | 7 |

2) 信号

选择作为信号的数据域与文献[11]相同。10 个数据域分为 3 类（不包括炎性信号）。

PAMP：数据域 25、26、29、38 和 40。

危险信号：数据域 23 和 24。

安全信号：数据域 12、31 和 32。

设 $x$ 为数据域的值，如果 $x\in[m,n]$时表现异常，这个域为 PAMP 或者危险信号；如果表现正常，则为正常信号。将这些数据域的指定区间按照式(4)归一化到[0,100]。

$$f(x)=\begin{cases}0, & x\in[0,m)\\ 100\dfrac{x}{n-m}, & x\in[m,n]\\ 100, & x\in(n,+\infty)\end{cases} \tag{4}$$

对数据域 12，感兴趣的指定区间为[0, 0.99]，其他数据域的指定区间按照[*min*, *max*]构造。每类信号的均值作为此类信号的值。

### 4.2　实验设置

IAIS 代码通过 MATLAB R2009a 实现，所有实验在 Windows 7(Intel Pentium Dual CPU T2330, 4GB RAM)平台下运行。除非特别指出，实验使用表 3 中给出的参数。树突状细胞数量的选择是精度和代价之间的平衡，考虑计算开销，设为 10；细胞周期率(cell cycle rate)根据数据集中信号采样率设置，设为 1 采样/s；检测器长度($L$)为抗原位数，$L$=24。实验对系统性能有显著影响的主要参数进行了敏感性分析。信号处理通过式(2)进行，权值与 DCA[3] 中使用的一样，如表 4 所示。为了模拟实时运行环境，通过一段代码读取数据集并提交给 IAIS。因为 IAIS 是随机系统，每个实验运行 10 次，结果取平均值。利用 ROC 曲线来评估 IAIS 的性能。

实验分为 3 部分进行。

E1：自我半径实验。

E2：异常阈值实验。

E3：与其他方法的比较。

E1，E2 使用 10%子集的前 10 000 个项，记为 $S_1$；E3 使用完整 10%子集，记为 $S_C$。

**表 3　　IAIS 参数**

| 算法 | 参数 | 值 |
|---|---|---|
| 实时 DCA | 树突状细胞数量 | 10 |
| | 迁移阈值($M_{th}$) | [100, 500] |
| | 异常阈值($\delta$) | 0.35 |
| NSA | 自我半径($\varepsilon$) | 4 |
| | 记忆检测器数量 | 100 |
| | 普通检测器数量 | 1 000 |
| | 记忆检测器 | 1 |
| | 普通检测器 | 0.5 |

**表 4　　式(2)推荐权值**

| $w$ | PAMP($S_0$) | 危险信号($S_1$) | 安全信号($S_2$) |
|---|---|---|---|
| 协同刺激信号 ($O_0$) | 2 | 1 | 2 |
| 半成熟信号 ($O_1$) | 0 | 0 | 3 |
| 成熟信号 ($O_2$) | 2 | 1 | –3 |

## 5　结果和分析

### 5.1　E1：自我半径实验

自我半径($\varepsilon$)决定单个检测器的覆盖范围，如果

$\varepsilon$=0，则要求检测器和抗原完美匹配，即结构完全相同。当检测器和抗原之间的亲和度大于 $L-\varepsilon$ 时，认为两者匹配。为研究自我半径对 IAIS 性能的影响，在子集 $S_1$ 上测试了如下自我半径 $\varepsilon$ = 0, 1, 2, …, 12。图 3 给出改变自我半径时的 ROC 曲线。

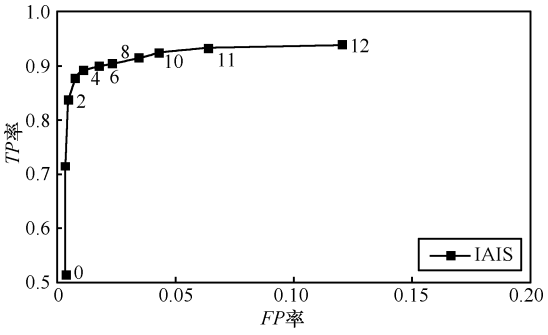


图 3　E1 的 ROC 曲线

由图可知，自我半径的取值对 *TP* 率和 *FP* 率有影响，随着自我半径增加，*TP* 率和 *FP* 率都随之增加。ROC 曲线上越靠近左上角表示性能越好，即 *TP* 率高，*FP* 率低。在 *TP* 率和 *FP* 率之间进行平衡，下面的实验将自我半径设为 $\varepsilon$=4。

### 5.2　E2：异常阈值实验

异常阈值($\delta$)用于决定抗原是正常还是异常，动态 MAC 值高于异常阈值的抗原认为是异常抗原。在子集 $S_1$ 上应用不同的异常阈值来计算检测率和误报率，测试的异常阈值包括 $\delta$=0.1,0.2,…,1，实验结果如表 5 所示。

**表 5　　E2 的检测率和误报率**

| 异常阈值 | 检测率 | | 误报率 | |
|---|---|---|---|---|
| | 均值 | 方差 | 均值 | 方差 |
| 0.1 | 0.901 6 | 0.001 7 | 0.024 1 | 0.001 9 |
| 0.2 | 0.898 8 | 0.001 9 | 0.018 8 | 0.001 6 |
| 0.3 | 0.896 8 | 0.001 1 | 0.015 8 | 0.002 1 |
| 0.4 | 0.888 3 | 0.001 5 | 0.012 2 | 0.001 7 |
| 0.5 | 0.867 7 | 0.001 5 | 0.004 8 | 0.002 3 |
| 0.6 | 0.761 7 | 0.003 1 | 0.002 5 | 0.001 5 |
| 0.7 | 0.456 9 | 0.003 6 | 0.001 2 | 0.000 4 |
| 0.8 | 0.175 7 | 0.001 8 | 0.000 4 | 0.000 1 |
| 0.9 | 0.032 6 | 0.001 8 | 0 | 0 |
| 1.0 | 0.030 1 | 0.001 2 | 0 | 0 |

表中给出的检测率和误报率的方差相对较小，表明算法性能稳定。从均值来看，异常阈值的取值对 IAIS 的性能具有显著的影响，高异常阈值导致低检测率和低误报率。实际应用的时候要在检测率和误报率之间进行平衡。当异常阈值介于 0.3 和 0.4 之间时，检测率介于 0.896 8 和 0.888 3 之间，误报率介于 0.015 8 和 0.012 2 之间。下面的实验将异常阈值设为 $\delta$=0.35。

### 5.3　E3：比较

在数据集 $S_C$ 上进行实验，测试 IAIS 的性能，并与其他方法进行比较，如表 6 所示。IAIS 的实验结果是 10 次实验的平均值。表中包括 NSA 和 DCA 的实验结果，此外还列出了其他智能方法的实验结果，包括人工神经网络，进化计算，支持向量机等。表中各种方法具有各自的特性，实验设置不尽相同，具体细节可参见各种方法的相关文献。

**表 6　　KDD99 数据集上的性能比较**

| 序号 | 方法 | 检测率 | 误报率 |
|---|---|---|---|
| 1 | NSA[12] | ≈0.95 | ≈0.01 |
| 2 | DCA[11] | ≈0.75 | ≈0 |
| 3 | SOM[12] | 0.906 0 | 0.015 7 |
| 4 | K-NN[13] | 0.910 0 | 0.080 0 |
| 5 | GP[14] | 0.910 2 | 0.004 3 |
| 6 | PNrule[15] | 0.911 0 | 0.004 0 |
| 7 | LGP[16] | 0.944 0 | 0.035 0 |
| 8 | SVM[13] | 0.980 0 | 0.100 0 |
| 9 | IAIS | 0.969 1 | 0.032 1 |

首先，将 IAIS 与 NSA 和 DCA 进行比较。与 DCA 进行比较，IAIS 的检测率比 DCA 有明显提高，但与此同时，误报率也提高了，但是幅度不大。相对而言，DCA 的检测率太低，因此 IAIS 更可取。与 NSA 进行比较，IAIS 的检测率和误报率都比 NSA 高，但幅度有限。需要指出的是，文献[12]中 NSA 有个训练过程，需随机选择 80%的正常样本进行训练，剩余 20%的正常样本和异常样本一起作为测试集。而 IAIS 定义各种信号之后，不需要训练集来训练检测器，训练检测器的自我集合是根据危险信号在线生成的，也就是说，IAIS 可以以实时或近实时的模式运行。

其次，将 IAIS 与其他方法进行比较。IAIS 具有次于 SVM 的检测率，但是 SVM 误报率太高，综合检测率和误报率，IAIS 性能更好。IAIS 误报率与 LGP 相当，检测率比后者高。其他方法虽然误报率低，检测率比 IAIS 要低超过 5 个百分点。通过比较可知，IAIS 具有良好的综合性能。

## 6　结束语

本文对危险模型进行了分析，根据新的关于危险模型和否定选择机制之间关系的理解，提出结合实时 DCA 和 NSA 来建立 IAIS，用于执行实时入侵检测。系统中实时 DCA 利用行为特征，NSA 利用结构特征。在 KDD99 数据集上对 IAIS 进行验证，并与其他方法进行比较。IAIS 检测性能与经典分类算法相当，其特点包括：可结合结构特征和行为特征进行检测，不依赖明确标识的数据集来训练检测器，以实时或者近实时的模式运行。

当前 IAIS 参数较多，文中仅对主要参数进行了敏感性分析。为了简化起见，部分参数设为常数。人工免疫系统的优点之一就是它能够动态地维持细胞种群数量，因此有待改进，后续工作中可以引入反馈机制让参数根据外部环境的变化进行自适应调整。

## 参考文献：


[1] DASGUPTA D. Advances in artificial immune systems[J]. Theoretical Computer Science, 2006, 403(1):11-32.

[2] MATZINGER P. Tolerance, danger and the extended family[J]. Annual Review of Immunology, 1994, 12: 991-1045.

[3] GREENSMITH J, AICKELIN U, CAYZER S. Introducing dendritic cells as a novel immune-inspired algorithm for anomaly detection[A]. International Conference on Artificial Immune Systems[C]. Banff, Canada, 2005.153-167.

[4] LI M. An approach to reliably identifying signs of DDOS flood attacks based on LRD traffic pattern recognition[J]. Computers & Security, 2004, 23(7):549-558.

[5] LI M. Change trend of averaged hurst parameter of traffic under DDOS flood attacks[J]. Computers & Security, 2006, 25 (3):213-220.

[6] JANEWAY C. The immune system evolved to discriminate infectious nonself from non-infectious self[J]. Immunology Today, 1992, 13:11-16.

[7] SARAFIJANOVIĆ S, BOUDEC J L. An artificial immune system for misbehavior detection in mobile ad-hoc networks with virtual thymus, clustering, danger signal, and memory detectors[A]. International Conference on Artificial Immune Systems[C]. Catania, Italy, 2004.342-356.

[8] CHEN Y B, FENG C, ZHANG Q, *et al*. Negative selection algorithm with variable-sized r-contiguous matching rule[A]. International Conference on Progress in Informatics and Computing[C]. Shanghai, China, 2010.150-154.

[9] GREENSMITH J. The Dendritic Cell Algorithm[D]. Nottingham, UK: School of Computer Science, University of Nottingham, 2007.

[10] KAYACIK H G, ZINCIR-HEYWOOD A N, HEYWOOD M I. Selecting features for intrusion detection: a feature relevance analysis on KDD 99 intrusion detection datasets[A].Third Annual Conference on Privacy, Security and Trust[C]. New Brunswick, Canada, 2005.

[11] GU F, GREENSMITH J, AICKELIN U. Further exploration of the dendritic cell algorithm[A]. International Conference on Artificial Immune Systems[C]. Phuket, Thailand, 2008.142-153.

[12] GONZ LEZ F A, DASGUPTA D. Anomaly detection using real-valued negative selection[J]. Genetic Programming and Evolvable Machine, 2003, 4(4): 383-403.

[13] ESKIN E, ARNOLD A, PRERAU M, *et al*. Applications of Data Mining in Computer Security[M]. Boston, Kluwer, 2002.

[14] FOLINO G, PIZZUTI C, SPEZZANO G. GP ensemble for distributed intrusion detection systems[A]. Third International Conference on Advances in Pattern Recognition[C]. Bath, UK, 2005. 54-62.

[15] AGARWAL R, JOSHI M V. PNrule: a new framework for learning classifier models in data mining (a case-study in network intrusion detection)[A]. First SIAM Conference on Data Mining[C]. Chicago, 2001.1-17.

[16] SONG D, HEYWOOD M I, ZINCIR-HEYWOOD A N. Training genetic programming on half a million patterns: an example from anomaly detection[J]. IEEE Trans on Evolutionary Computation, 2005,9(3):225-239.


## 作者简介：


**陈岳兵**（1980-），男，江苏江都人，总参第六十一研究所工程师，主要研究方向为人工免疫系统和信息安全。

**冯超**（1983-），男，江苏盱眙人，国防科技大学博士生，主要研究方向为安全协议分析。

**张权**（1974-），男，上海人，国防科技大学副教授、硕士生导师，主要研究方向为量子通信和信息安全。

**唐朝京**（1962-），男，江苏武进人，国防科技大学电子科学与工程学院院长、教授、博士生导师，主要研究方向为多媒体通信、网络攻防与对抗。