

树突细胞在入侵检测中的应用研究

代昆玉, 胡 滨, 李玉凯

(贵州大学 计算机科学与技术学院, 贵州 贵阳 550025)

摘 要: 提出了一种基于危险理论的网络攻击入侵检测算法. 该算法强调了树突细胞(Dendritic Cells)在人工免疫中的重要作用, 描述了在入侵检测过程中, DCs 的不同状态的转换和对信息的处理过程, 以及和 T 细胞的交互. 最后同几种经典的 AIS 算法进行比较, 仿真结果显示在网络攻击入侵的检测效率和误报率上都有较大的改善.

关键词: 人工免疫; 危险理论; 入侵检测; 树突细胞

中图分类号: TP393

文献标识码: A

文章编号: 1000-7180(2014)12-0114-04

Application of Dendritic Cells in Intrusion Detection

DAI Kun-yu, HU Bin, LI Yu-kai

(College of Computer Science and Technology, Guizhou University, Guiyang 550025, China)

Abstract: An Algorithm in Danger Theory field for intrusion detection is presented. The algorithm Stresses the important role of dendritic cells in artificial immunity and describes the conversion of the different states of DCs and the processing of information in intrusion detection and the interaction of T cells. Finally, we compared its detection method with classic AIS methods. These results show considerable improvements in terms of detection performance and false alarm rates.

Key words: Artificial Immune; Danger Theory; Intrusion Detection; Dendritic Cells

1 引言

人工免疫系统^[1]由两层防御构成: 先天性免疫系统和适应性免疫系统^[2]. 而 DCs 是先天性免疫和适应性免疫系统之间的重要纽带, 是危险理论的主要组成, 它提供病原入侵的初始检测, 通过收集由病原体和组织细胞产生的抗原来识别其是否是危险信号. 文献[3]中将人工免疫中的 DCs 用于异常入侵检测并首次提出树突细胞算法(Dendritic Cell Algorithm), 其思想是通过在形式上把抗原和输入信号表示的数据关联起来, 去识别抗原组是否正常. DCA 算法忽略适应性免疫系统的功能, 只模拟树突细胞能够通过代表其环境信号的相关性来激活或抑制免疫应答的能力, DCA 没有培训阶段, 因此误报率较高.

基于危险理论的另一个算法是由 Twycross 等

提出的 TLR(Toll-Like Receptor)算法^[4]. TLR 使用 DCs 和 T 细胞(T Cells), 同时模拟先天性和适应性免疫系统的功能. TLR 不考虑抗原结构, 利用精确匹配来完成绑定的过程, 因此, 在面对变化的入侵行为时的识别能力欠佳.

本文提出了一种新的改进算法. 和 TLR 算法一样, 改进算法也强调 DC 细胞和 T 细胞的交互, 但采用的是类似于 DCA 算法的信号处理过程. 相比现有的检测算法, 和 DCA, TLR 需要外部危险信号不同, 改进算法通过正常和异常结构培训产生信号, 信号的产生不需要外部源. 由于使用了培训数据, 本算法具有更低的误报率和更高的检测性能.

2 算法描述

2.1 树突细胞

当受到病原攻击时, 人体正常细胞就会产生危

收稿日期: 2014-02-18; 修回日期: 2014-04-09

基金项目: 贵州省科学技术基金(黔科合 J 字[2011]2198 号); 贵州大学青年教师科研基金(贵大自青基合字(2012)018 号)

险信号. 这些信号被 DCs 检测并呈现三种状态: 未成熟 (immature)、半成熟 (semi-mature) 和成熟 (mature) 状态. iDCs 位于组织中, 而其他 DCs 位于淋巴结中. 在未成熟状态下, DCs 收集来自本地环境的带有安全或是危险信号的抗原. DCs 通过整合这些信号来判断环境是否安全, 如果是安全的, DCs 就进入到半成熟状态, 同时传递抗原到 T 细胞, 并促使 T 细胞进行细胞耐受. 如果是危险的, DCs 就进入成熟状态, 并促使 T 细胞响应并杀死抗原. DCs 的状态转换如图 1 所示.

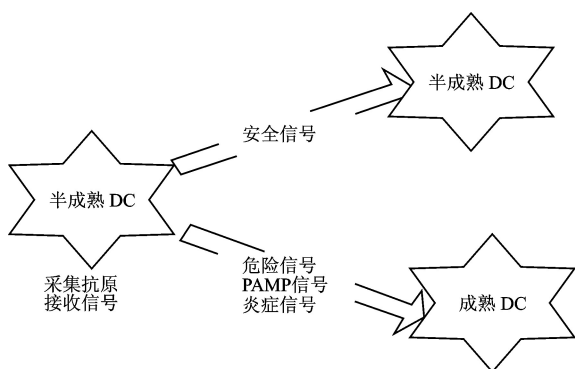


图 1 DCs 状态转换图

iDC 使用接收器采集抗原和警告信号. 警告信号分为四类, 分别^[5]如下:

(1) 危险信号. 细胞非正常死亡时产生, 表示可能有危险, 也可能没有危险, 但有危险的概率比没有危险的要高.

(2) 安全信号. 细胞正常死亡时产生, 表示没有危险发生.

(3) PAMP (Pathogenic Associated Molecular Patterns) 信号. 由细菌表示的蛋白质, 能够被 DCs 检测并激活免疫系统, PAMP 的存在通常表示一种异常情况.

(4) 炎症信号. 表示了各种免疫刺激分子受损的情况, 炎症信号不能单独刺激 DCs, 但能加强前面 3 种信号的影响.

DCs 通过整合不同输入信号产生相应的输出信号, 用输出信号的相对浓度来确定不同的状态, 用产生的两种分子来表示, 即 IL-12 和 IL-10.

2.2 特征抽取

要区分正常和异常行为, 首先对采集到的数据进行清理. 选取和网络行为相关度较高的几个属性: 服务类型, 源地址, 源端口, 目的地址, 目的端口, 时延作为有效子集成员. 这些特征构成一种模式并送入入侵检测系统中构成输入信号. 图 2 中显示了这

种模式结构.

服务类型	源地址	源端口	目的地址	目的端口	时延
------	-----	-----	------	------	----

图 2 输入模式结构

2.3 入侵检测

在检测中需对大量正常或异常的数据培训来产生不同的信号. 培训阶段可分两步: 产生存储在淋巴结的检查器、识别来自组织的信号源. 在第一步中, 产生一些幼稚 T 细胞 (Naive T Cells) 作为异常检测器. 这些 NTC 属于适应性免疫系统, 可将每一个 NTC 看成是一个来自自我域的球体. 因此在产生检测器前, 需定义自我域. 自我域由一些固定半径的球体并排组成, 其球心是正常训练样本. 式 (1) 定义自我域.

$$S_R = \bigcup_{1 \leq i \leq \Gamma} SG_i \quad SG_i = (\delta_i, r_s) \quad (1)$$

式中, r_s 是自我球体的恒定半径, δ_i 是正常训练样本中的第 i 个球体, SG_i 表示第 i 个自我球体的中心. 正常训练数据集中共有 Γ 个样本. 图 2 可知, 每一个样本有 6 个特征, 所以 δ_i 是一个 6 维向量.

定义了自我域后, 即可生成 NTC 并把它们存储在检测存储器中. 由于生成的检测器完全脱离自我域, 故根据式 (2) 定义 NTC 存储器.

$$NTC_S = \{ntc_j \mid ntc_j = (\eta_j, r_{\eta_j}, \cdot)\} \quad 1 \leq j \leq H \quad (2)$$

NTC 存储器包含了许多可变半径的球体, 其球心被随机定义或者用异常训练数据集定义. H 表示检测存储器的规模, η_j 和 r_{η_j} 分别表示第 j 个检测器的中心和可变半径. r_{η_j} 可根据式 (3) 求得.

$$r_{\eta_j} = (\min_{1 \leq i \leq \Gamma} dist(\delta_i, \eta_j)) - r_s \quad j = 1, 2, \dots, H \quad (3)$$

式中, $dist()$ 是计算曼哈顿距离的函数, r_{η_j} 表示了正常训练数据集中离 η_j 最近的球体到 η_j 的距离. 因为 η_j 总是在自我域外, 故对任意 δ_i 都有 $dist(\delta_i, \eta_j) > r_s$, 所以 r_{η_j} 恒大于 0.

在检测阶段, 每一个 NTC 都有三个接收器即, IL-12R, IL-10R 和 AgR 分别接 IL-12, IL-10 和抗原. 图 3 显示了 NTC 和自我域的关系.

培训阶段的第二步, 必须识别在组织中产生的一些信号源. 为简化模型, 选择危险信号、安全信号、PAMP 信号作为输入信号. 分别使用正常和异常训练样本作为安全和 PAMP 的信号源, 再用异常训练集定义一个危险区域来产生危险信号.

类似自我域的定义, 危险域由一些并排可变半

径的球体构成,其球心是异常训练样本.该区域由式(4)定义:

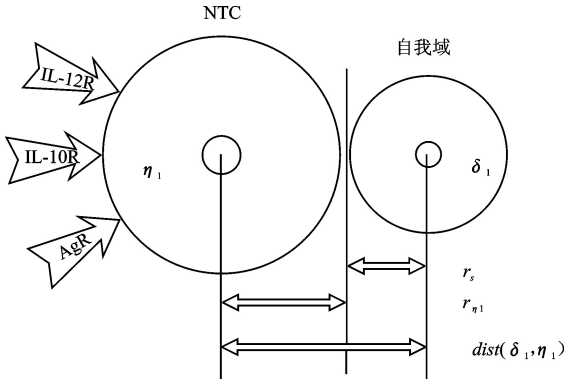


图3 NTC和自我域的关系

$$D_S = \bigcup_{1 \leq m \leq \Psi} DS_m; DS_m = (\psi_m, r_{dsm}) \quad (4)$$

式中, Ψ 表示异常训练数据集的规模. ψ_m 和 r_{dsm} 分别表示第 m 个危险域的中心和可变半径, ψ_m 是一个 6 维向量. 式(4)中的 r_{dsm} 根据 ψ_m 和异常训练数据集中离它最近的球体的距离计算得到, 式(5). 其中 MI 表示危险域和自我域之间的距离.

$$r_{dsm} = \begin{cases} 0, & \text{if } (\min_{1 \leq i \leq \Gamma} \text{dist}(\delta_i, \psi_m)) \leq r_s + MI \\ (\min_{1 \leq i \leq \Gamma} \text{dist}(\delta_i, \psi_m)) - r_s - MI, & \text{otherwise} \end{cases} \quad (5)$$

$$m = 1, 2, \dots, \Psi$$

入侵检查从组织开始,到淋巴结结束.在这个过程中,DCs 和 TCs 都至关重要并且表现为不同的状态. iDCs 负责从组织中收集抗原和信号,它们有 4 个接收器即: AgR, DR, SR 和 PR, 它们分别收集抗原信号、危险信号、安全信号和 PAMP 信号. mDCs 和 smDCs 则负责把抗原传递给 TCs, 同时在淋巴结分泌不同的细胞因子, 即 IL-12 和 IL-10. NTC 在淋巴结接收抗原和细胞因子, 如果接收到的是抗原和 IL-12 分子, NTC 就会被激活. 如果接收到的是 IL-10, NTC 就被抑制.

入侵检查过程如下: 一旦一个抗原进入组织中, iDCs 就朝抗原移动, 并且把自己的中心放到抗原上, 使用接收器捕获抗原和不同的信号. 然后 DR 检查 iDCs 的中心是否位于危险域, 如果 iDCs 在危险

域, DR 收到一个危险信号, 并把其值置为 1, 否则其值置 0.

DR 的计算公式见式(6).

$$DR = \begin{cases} 1, & \text{if } \exists_{1 \leq m \leq \Psi} \text{dist}(c_{iDC}, \psi_m) < r_{dsm} \\ 0, & \text{otherwise} \end{cases} \quad (6)$$

式中, c_{iDC} 表示定位到输入抗原的 iDCs 的中心, $\text{dist}(c_{iDC}, \psi_m) < r_{dsm}$ 检查 c_{iDC} 是否位于 m 个危险球体内.

同理, SR 和 PR 分别从正常或异常训练数据集中接收安全或 PAMP 信号. 可假设每个正常训练样本都发送一个安全信号给 iDC, 信号的强度用 δ_i 和 c_{iDC} 之间的曼哈顿距离来定义. SR 从不同的正常训练样本中收集所有的安全信号, 然后选择最强的信号作为 Safe_S, Safe_S 值可由式(7)计算.

$$Safe_S = \max_{1 \leq i \leq \Gamma} e^{-\text{dist}(c_{iDC}, \delta_i) / d_{\max}} \quad (7)$$

式中, $d_{\max}$ 是 6 维空间中 2 点间的最大距离. 对 PR 做同样的处理, PAMP_S 由式(8)给出.

$$PAMP_S = \max_{1 \leq j \leq U} e^{-\text{dist}(c_{iDC}, \delta_j) / d_{\max}} \quad (8)$$

根据计算出的 Safe_S 和 PAMP_S, SR 和 PR 的值可分别根据式(9)和(10)求得.

$$SR = \begin{cases} 1, & \text{if } Safe_S > PAMP_S \\ 0, & \text{otherwise} \end{cases} \quad (9)$$

$$PR = \sim SR \quad (10)$$

当 iDCs 通过接收器收到抗原或者其他信号时, 它在组织中接收到得信息传递到一个抗原提呈的上下文中. 如果 iDCs 显示的是 PAMP 或者危险信号, 它就会变成 mDC, 根据式(11), IL-12R 被置为 1. 如果 iDCs 显示的是安全信号, 它就会变成 smDC, 根据式(12), IL-10R 被置为 1.

$$IL-12R = DR \vee PR \quad (11)$$

$$IL-10R = \sim DR \wedge SR \quad (12)$$

经历上述过程后, 变形的 iDCs (mDC 或 smDC) 从组织迁移到淋巴结并提交其抗原给 NTC 存储器. 在该过程中, 每个 NTC 存储检测器通过收到抗原或者是 mDC 或者 smDC 的受体提供的信号来评估抗原的结合. 这个结合过程通过式(13)计算得到.

$$link = \begin{cases} 1, & \text{if } \exists_{1 \leq j \leq D} ((\text{dist}(Ag, d_j) < r_{dj}) \wedge TV_j = 1) \\ 1, & \text{else if } \exists_{\substack{1 \leq j_1, j_2 \leq D \\ j_1 \neq j_2}} \left\{ \begin{aligned} &(\text{dist}(Ag, d_{j_1}) < r_{dj_1}) \wedge \\ &(\text{dist}(Ag, d_{j_2}) < r_{dj_2}) \end{aligned} \right\} \\ 0, & \text{otherwise} \end{cases} \quad (13)$$

式中, $(\text{dist}(Ag, d_j) < r_{dj})$ 用于检查 NTC 存储器的第 j 个检测器是否跟抗原绑定. 根据式(13), 如果

一个 NTC 绑定抗原且有较高的信任值 ($TV=1$) 时, link 置为 1, 否则如果 NTC 绑定抗原时没有高

信任值,但至少要有 2 个低信任值时,link 也置为 1,否则 link 置为 0.

如果 $link=1$,NTC 接收器的信号就会被评估,然后,当抗原表现为 mDC 且 $IL-12R=1$ 时,NTC 被激活.激活的 NTC 能够杀死抗原.在这种情况下,系统分配异常标签给抗原. NTC 杀死抗原后,它会通知邻居结点并发送危险信号给他们.这里可通过增加一个危险球体(Danger Sphere)到危险域来模拟这个过程,DS 和其他危险域相似,是一个中心刚好在 Ag 上的一个球体,其可变半径通过式(5)计算得到,这里 $\psi=Ag$.

另一方面,当抗原表现为 smDC 且 $IL-10R=1$ 时,NTC 将被抑制.被抑制的 NTC 的信任值降低,故不会对抗原采取行动,此时,系统分配正常标签给抗原.

第三种情况是,当 $link=0$ 时,表示没有任何的抗原跟 NTC 绑定,NTC 在激活前不能杀死任何抗原,在这种情况下,系统分配正常标签给抗原.

但是,还有一种情况,即当 DR 和 PR 都为 1 时显示抗原的上下文极为异常,这时如果给其一个正常标签显然是不合理的.在这种情况下,我们将更新 NTC 存储器,用一些新的检测器去替换信任值低的 NTC.这些新的 NTC 和其他检测器一样,是一个中心正好在 Ag 上的一个球体,其可变半径用式(3)计算,这里 $\eta=Ag$.现在,被绑定的参数值变为 1,同时对 NTC 的信号接收器进行评估.

综上所述,可由式(14)给抗原分配正常或者异常标签.

$$Input_Label = \sim link \vee (link \wedge IL-10R) \quad (14)$$

式中, $Input_Label=1$, 抗原正常; $Input_Label=0$, 抗原异常.

3 仿真实验及结果分析

为了实验的进行,需要在数据中采集正常和异常模式.实验使用测试数据 KDDCUP 1999 作为数据样本.为了提供异常模式,选取了 3 种典型的网络攻击入侵数据:

(1)远端未授权访问(Remote To Local)

(2)分布式拒绝服务(Distributed Denial of Service)

(3)未经授权提升访问(User To Super User or Root Attack).

实验中从 KDDCUP99 中选取两类数据:培训数据和测试数据.培训数据用于对算法的培训阶段,

测试数据在检测阶段中充当抗原的角色.每一个测试数据都有一个确切的分类:正常或异常,同时,本算法给每一个数据分配一个检测后的类别,如果检测类别和实际类别相同,则证明检测过程成功;否则表示检测过程失败.实验中,用 Sen 表示入侵检测的检测率, FPR 表示入侵检测的误报率.

$$Sen = TP / (TP + FN) \quad (15)$$

$$FPR = FP / (TN + FP) \quad (16)$$

式(15)、式(16)中,FP 表示没有攻击发生时假报警的数量, FN 表示有攻击发生时显示正常状态, TP 表示攻击发生时显示故障状态, TN 表示没有攻击发生时的正常状态.

用上述 2 个检测指标评估算法的实验结果.在培训阶段,根据式(1)和式(4)定义的自我域和危险域,需事先确定 r_s 和 MI 的值.根据文献[6]使用学习自动机来优化这些参数,可令 $r_s=0.0003$, $MI=0.002$.同时,定义 NTC 存储器的数量,这是一个可变值.表 1 中给出了不同数量 NTC 存储器的实验结果.

表 1 不同数量 NTC 存储器的实验结果

NTC 数量	Sen/%	FPR/%
50	38.02	3.61
100	57.31	5.42
200	84.23	6.42
1 000	90.18	6.21
5 000	90.42	5.89
10 000	90.87	6.13

由表 1 可知,随着 NTC 数量的不断增加,TP 和 FP 的值也相应的增加,但是由于采用合适的信号处理函数导致 TP 比 FP 增加更多,导致 Sen 比 FPR 增加更快.

将本文改进的算法和 DCA、TLR 算法进行比较.图 4 中显示三种算法随着 NTC 数量增加检查准确性的比较.

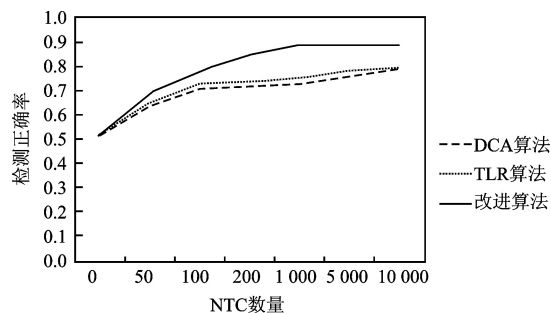


图 4 改进算法和 DCA、TLR 算法检测准确性比较

由于本文提出的算法考虑抗原结构,并通过式

(下转第 122 页)

后的连通区域进行伤痕判定,实现对整个磁痕图像的缺陷识别,完成磁痕图像缺陷的高精度检测.实验结果表明,利用改进算法进行磁痕图像缺陷的高精度检测,极大的提高了检测的效率与准确性,满足了人们的实际需求.

参考文献:

- [1] 王勇,郭慧.基于形态学的焊缝 X 射线图像缺陷检测[J]. 东华大学学报:自然科学版, 2013, 39(4): 460-463.
- [2] 张岩,李涛,李庆领.轮胎激光错位散斑干涉图像缺陷检测及定位方法[J]. 橡胶工业, 2013, 60(4): 247-250.
- [3] 龚聪,徐杜.光源强度变化对图像检测精度的影响及其解决方法[J]. 科学技术与工程, 2014(13): 236-239.
- [4] 罗桓,田翔.基于改进 Canny 算子的电力设备图像检测研究[J]. 电测与仪表, 2014, 51(10): 77-81.
- [5] 赵大伟,陈刚.机载不定视觉下的图像小目标定位算法仿真[J]. 计算机仿真, 2013, 10(30): 108-111.
- [6] 李丹,张鑫.改进小波阈值去噪在 PCB 图像检测中的应用研究[J]. 现代电子技术, 2014, 37(2): 34-37.

(上接第 117 页)

(13)来完成绑定的过程.因此,该算法在面对变化的入侵行为时有更好的识别能力.实验结果显示,改进的算法在检测性能和误报率上有很大的改善.

4 结束语

本文提出了一种新的基于危险理论的异常检测算法,由于使用基于安全信号、危险信号和 PAMP 信号的适当的信号处理方法,该改进的算法与传统的危险理论算法相比,具有更低的误报率和更高的检测率.但是,为简化模型,本算法对炎症信号尚未考虑进来.如何在算法中考虑所有的警告信号以及如何有效提高本检测算法的适应能力和检测效率是下一步的工作.

参考文献:

- [1] S Forrest, A S Perelson, L Allen, et. al, Self-non-self discrimination in a computer[C]// Proceedings of 1994 IEEE Computer Society Symposium on Research in Security and Privacy, Oakland, CA, USA: IEEE, 1994, 202-212.
- [2] J P Twycross. Integrated innate and adaptive artificial immune systems applied to process anomaly detection

- [7] 胡鹏博,刘晓利,白宏阳.基于改进梯度向量流动态轮廓模型的图像检测[J]. 探测与控制学报, 2013, 35(5): 26-30.
- [8] Liping Liu, Ye Tian, Xiudong Sun, et al. Influence of laser mode on splitting beam illumination effect of Dammann grating[J]. Chinese Optics Letters, 2009, 7(6): 458-461.
- [9] 徐亮,魏锐.基于 Canny 算子的图像边缘检测优化算法[J]. 科技通报, 2013, 29(7): 127-131.
- [10] Katia Stankov, Dong-Chen He. Using the spectral similarity ratio and morphological operators for the detection of building locations in very high spatial resolution images[J]. Journal of Communication and Computer, 2013, 10(3): 309-324.

作者简介:

文政颖 女, (1979-), 硕士, 讲师. 研究方向为计算机应用与图像处理.

王佳欣 男, (1983-), 硕士, 讲师. 研究方向为计算机应用.

[D]. Nottingham: University of Nottingham, 2007.

- [3] J Greensmith, U Aickelin, J P Twycross, Detecting danger: applying a novel immunological concept to intrusion detection systems[C]// Poster Proceedings of ACDM, Bristol, UK: IEEE, 2004, 1135-1142.
- [4] J P Twycross, U Aickelin. An immune-inspired approach to anomaly detection[J]. Handbook of research on information assurance and security, information science reference 2007, 6(4): 109-21.
- [5] J Greensmith, The dendritic cell algorithm[D]. Nottingham: University of Nottingham, 2007.
- [6] Afzali N, Azmi R, Pishgoo B. A new clonal selection algorithm based on radius regularization of anomaly detectors[J]. Accepted in the 16th CSI international symposium on Artificial intelligence and signal processing, 2012(3): 497-502.

作者简介:

代昆玉 女, (1979-), 硕士, 讲师. 研究方向为数据挖掘, 计算机网络安全. E-mail: 10618709@qq.com.

胡滨 男, (1977-), 博士研究生, 副教授. 研究方向为人工智能, 计算机网络安全.

李玉凯 男, (1975-), 硕士, 讲师. 研究方向为计算机网络安全.