文章编号: 1001-0920 (2016) 03-0528-07 DOI: 10.13195/j.kzyjc.2014.1969

基于ISM特性的检测器生成算法及模型

马占飞[1], 杨树英[2], 郭广丰[1]

(1. 内蒙古科技大学包头师范学院 计算机系，内蒙古 包头 014030；
2. 包头服务管理职业学校 信息技术系，内蒙古 包头 014030)

摘 要: 针对现有检测器生成算法存在效率低、自适应性差、生成的检测器集庞大且冗余等问题, 借鉴生物免疫系统中抗体的克隆机制和亲和度变异机制, 并融合小生境策略以及检测器的变异和优化等, 构建基于免疫软件人(ISM)特性的检测器生成算法及模型. 与传统算法相比, 该算法能够降低检测器的冗余度, 减少检测器集的规模, 保持检测器的多样性; 通过合理地改变其匹配阈值, 能够实现以较小的检测器集检测出更多的异常行为的目的. 实验结果表明, 所提出的算法具有较强的自适应性, 且拥有较高的检测效率和性能.

关键词: 网络安全；入侵检测；免疫软件人；检测器生成算法；小生境；变异；自适应

中图分类号: TP393.08 **文献标志码:** A

Detector generating algorithm and model based on immune-softman

MA Zhan-fei[1], *YANG Shu-ying*[2], *GUO Guang-feng*[1]

(1. Department of Computer，Baotou Teachers' College，Inner Mongolia University of Science and Technology，Baotou 014030，China；2. Department of Information Technology，Baotou Service Management Vocational School，Baotou 014030，China．Correspondent：MA Zhan-fei，E-mail：mazhanfei@163.com)

Abstract: Aiming at the low true-positive rate, poor adaptability, large size and redundant detector set of existent detector generation algorithms, a self-adaptive detector generation algorithm and model based on immune-softman(ISM) characteristics is proposed. The algorithm model draws lessons from the antibodies cloning mechanism and affinity variation mechanism of the biological immune system(BIS), and integrates the niching strategy and the variation and optimization operations of the detector. Compared with traditional detector generation algorithms, the proposed algorithm can decrease effectively the redundancy of detectors, minimize the size of detector set, and maintain the diversity of detectors. At the same time, by changing continuously the matching threshold of effective detectors, the algorithm can detect quickly abnormal behaviors in the scale of the non-self space by a smaller detector set. The experiment results show that the algorithm has better adaptability and higher detection efficiency and performance.

Keywords: network security; intrusion detection; immune-softman; detector generation algorithm; niching; variation; self-adaptive

0 引 言

检测器集是构建网络入侵检测与防御系统(NIDDS)的核心部分, 如何高效、快速地生成合格的检测器集, 对于网络安全系统的检测性能而言至关重要[1-3]. 早在1994年, Forrest等[4-5]在自然免疫学中自体(Self)/非自体(Non-self)划分的基础上, 提出了用阴性选择算法(NSA)生成有效的检测器集, 用以解决各种异常变化的检测问题, 并将其成功地应用于计算机病毒检测、故障检测与恢复以及时序异常检测. NSA的优点在于不需要先验知识就可以对未知的入侵行为进行较好地防御. 但是, NSA也存在明显不足: 一些Non-self字符串, 找不到与其匹配的有效检测器集, 检测效率会随字符串长度的变化呈线性变化, 由此导致Self集与检测器集的规模呈指数级代价关系. 为了克服这些局限性, D'haeseleer等[6]提出了两种改进的检测器生成算法: 线性检测法和贪婪法. 然而, 这

收稿日期: 2014-12-28；**修回日期:** 2015-03-20.

基金项目: 国家自然科学基金项目(61163025)；内蒙古自治区自然科学基金项目(2010BS0904)；内蒙古自治区高等学校科学研究基金重点项目(NJ10162)；包头市科学研究基金项目(2014S2004-3-1-26).

作者简介: 马占飞(1973−), 男, 教授, 博士, 从事计算机网络技术与信息安全、人工智能等研究；杨树英(1971−), 女, 副教授, 从事计算机网络技术与信息安全的研究.

两种算法在不同程度上也存在一些不足. 近年来, 一些研究者对这些算法进行了很多改进, 但在合格检测器的生成效率和检测性能上仍存在许多局限性[7-9].

鉴于此, 本文在基于免疫软件人(ISM)技术的网络入侵检测与防御系统的研究成果基础上, 通过对典型检测器生成算法的深入研究, 借鉴生物免疫系统(BIS)中抗体的克隆机制和亲和度变异机制, 并融合小生境(Niching)策略中“抗体”向“抗原”进化的思想[10-11], 以及对重叠检测器的变异和优化等操作, 构建了基于ISM特性的检测器自适应生成算法模型. 该算法的核心思想是:

1) 成熟检测器集通过训练-变异-检测-变异等几个阶段, 减少了检测器的冗余度, 增强了检测器集的完备性, 保持了检测器的多样性;

2) 记忆检测器集采用了变异和优化策略, 使检测器维持在一个较为适中的数量, 便于动态更新和快速检测, 进而增强其对曾经检测过的异常行为的识别能力.

1　算法参数的定义

定义 1　将检测器集(U)定义为一个有限符号表上的字符串集合. U 被划分为两个子集 S 和 N, 其中 S 表示“自体”(Self)集, N 表示“非自体”(Non-self)集. Self集合代表了合法的事件(即被保护的数据或活动), 而Non-self集合则代表了非法事件(即不可接受的或非法的数据或活动). 在算法中, U 是一个封闭和有限的集合, 即 $S \bigcup N = U$, $S \bigcap N = \varnothing$.

定义 2　在有限的匹配规则(Rule)下, 若两个字符串 a 和 b 的相似度超过匹配阈值 β(即为检测异常事件或数据而设定的参数值), 则称 a 与 b 匹配, 记作 $\mathrm{Match}\,(a, b)$.

定义 3 [13]　对于任意两个字符串 a 和 b, 如果 a 和 b 在长度大于或等于 r-连续位上的对应字符相同, 则称字符串 a 与字符串 b 在 r-连续位上匹配.

由此定义两个随机的字符串 a 和 b 在 r-连续位规则下匹配的概率为

$$P(\mathrm{Match}\,(a,b)) = 2^{-r}\left[\frac{l-r}{2}+1\right], \tag{1}$$

其中 l 表示位串的总长度.

定义 4 [14]　若某一字符串 $w \in N$ (Non-self)与任一给定的字符串 $z \in S$ (Self)匹配, 即 $\mathrm{Match}\,(z, w) = \mathrm{True}$, 其中 w 和 $z \in U$, 则称 w 是一个检测“黑洞”.

定义 5　针对基于Multi-ISM联盟的网络入侵检测与防御系统问题, 设定一个有限资源集为 U, 对于一个分类模式 $s \in U$, 如果 s 是Normal(正常), 则对应于ISM的Self集; 如果 s 是Abnormal(异常), 则对应于ISM的Non-self集. 假定一个网络入侵检测系统 D 由两部分构成, 即

$$D = (f, M).$$

其中: f 为分类函数; M 为从 U 中选取的表示Self模式集, 即 $M \in U$. 则 f 到 M 的映射与分类模式 $s \in U$ 可能是Normal或Abnormal, 即

$$f : U' \times U \to \{\mathrm{Normal}, \mathrm{Abnormal}\},$$

其中 U' 是 U 的幂集. 如果 s 属于Normal集, 则称该分类模式 s 是Normal模式; 否则, 称其为Abnormal模式. 即

$$f(M,s) = \begin{cases} \mathrm{Normal},\ s \in M; \\ \mathrm{Abnormal},\ s \notin M. \end{cases} \tag{2}$$

其中: f 对于 s 的分类可能会产生两种错误, 即假阳性(δ^+)和假阴性(δ^-).

定义 6　一个Self模式被检测系统 D 判定为Abnormal, 即为假阳性, 记为

$$\delta^+ = \{s \in S | f(M,s) = \mathrm{Abnormal}\}. \tag{3}$$

这类错误一般被称为误检(即将正常的事件误检为异常事件), 其对应的比率称为误检率.

定义 7　一个Non-self模式被检测系统 D 判定为Normal, 即为假阴性, 记为

$$\delta^- = \{s \in S | f(M,s) = \mathrm{Normal}\}. \tag{4}$$

这类错误一般被称为漏检(即将异常的事件误检为正常事件), 其对应的比率称为漏检率.

在一个分布式网络环境中, 假定 V 表示基于Multi-ISM联盟的网络入侵检测与防御系统中所有ISM的节点集, 其定义如下.

定义 8　某个ISM节点 i 的分类函数 f_i 定义如下:

$$f_i(M_i,s) = \begin{cases} \mathrm{Normal},\ s \in M_{v \in V_c}; \\ \mathrm{Abnormal},\ s \notin M_{v \in V_c}. \end{cases} \tag{5}$$

其中: V_c 表示进行分类判断时与该ISM节点通信的其他所有ISM节点; M 是从 U 中选取的, 表示Self集; s 表示一个分类模式.

定义 9　全局分类函数的定义如下:

$$g(\{M_v\}, s) = \begin{cases} \mathrm{Normal},\ \exists v \in V | s \in U_v \bigwedge f_v(M_v,s) = \mathrm{Normal}; \\ \mathrm{Abnormal},\ \exists v \in V | s \in U_v \bigwedge f_v(M_v,s) = \\ \qquad\qquad \mathrm{Abnormal}. \end{cases} \tag{6}$$

依据各ISM节点性质, 对于被检测的每一个模式串, 若发现任意一个ISM节点将它检测为Abnormal, 则该全局分类函数就被认定为Abnormal.

定义 10　全局误差的定义如下: 若 $s \in S$, 且 $\exists v$

$\in V|s \in S_v \bigwedge f_v(M_v, s) = \text{Abnormal}$, 则 s 是一个全局假阳性 (误检); 若 $s \in N$, 且 $\exists v \in V|s \in N_v \bigwedge f_v(M_v, s) = \text{Normal}$, 则 s 是一个全局假阴性 (漏检).

2 算法模型的构建与分析

针对现有检测器生成算法存在检测效率低、自适应性差, 且在生成检测器集合时存在大量冗余和无效的检测器等问题[15-16], 借鉴生物免疫系统中抗体的克隆机制和亲和度变异机制, 并融合小生境策略中“抗体”向“抗原”进化的思想[17-18], 以及对重叠检测器的变异和优化等操作, 在此基础上构建一种新型的基于ISM特性的检测器自适应生成算法模型, 如图1所示.

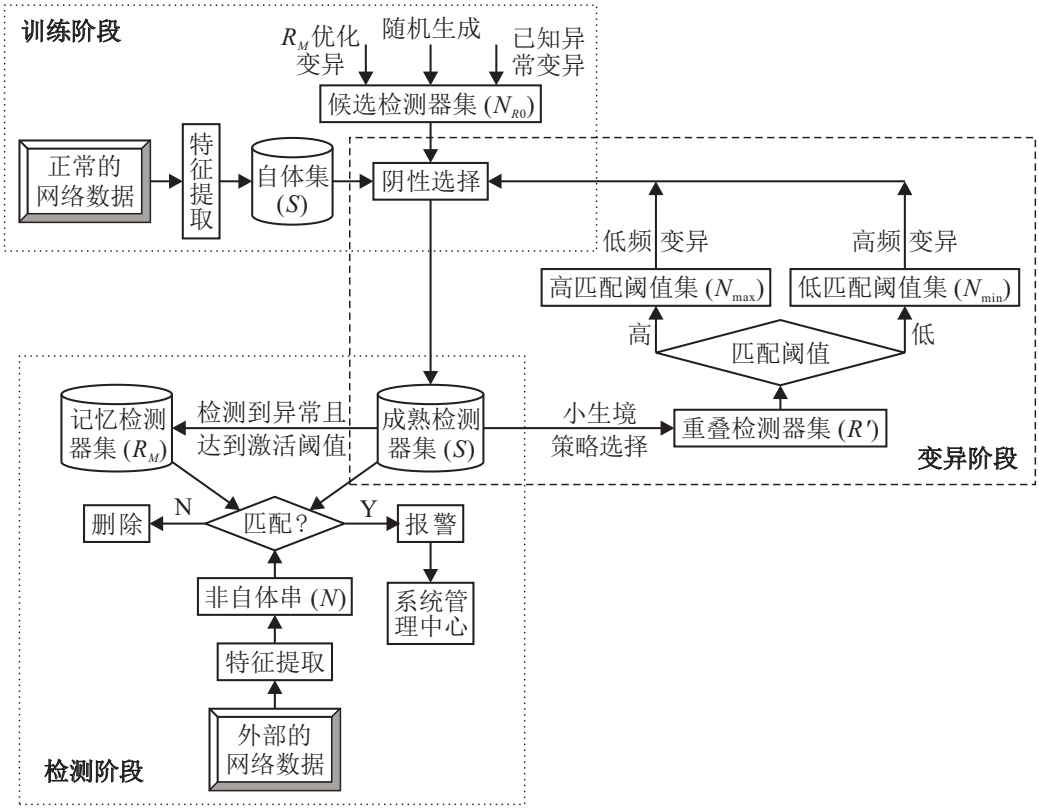

图 1 基于ISM特性的检测器自适应生成算法模型

2.1 算法工作机理

首先对正常的网络数据进行特征提取, 组成ISM的自体数据集 S, 通过与候选检测集 N_{R0} 进行阴性选择, 生成ISM的初始检测器集合 (成熟检测器集的初态), 并对其进行亲和度变异, 生成成熟检测器集 R. 由于匹配规则的限定, 在生成的成熟检测器集合 R 中会产生大量重叠信息, 为此采用小生镜策略对其进行筛选, 将重叠信息严重的组成新的检测器 R'. 在检测器 R' 集合中, 根据预先设定的匹配阈值对 R' 进行分类, 即对匹配程度较高的 R' 采取低频变异 (LFV); 对匹配程度较低的 R' 采取高频变异 (HFV)[19]. 高频变异的主要目的是使生成的检测器能够识别未知的入侵; 低频变异的主要目的是使生成的检测器能够识别已知入侵的变异. 最后采集外部网络数据, 通过特征提取后与已生成的记忆检测器集 R_M 和成熟检测器集 R 进行模式匹配, 进而生成ISM的记忆检测器集 R_M 和成熟检测器集 R. 在生成的检测器 R 中, 由于一些规则和参数的限定, 其中仍有大量重叠信息, 还需对这些重叠信息进行变异处理, 进而优化ISM的检测器集 R 和 R_M, 以增强检测器的检测效率.

简言之, 该算法主要经历训练-变异-检测-变异等阶段, 其主要目的是: 通过合理地改进有效检测器集合 R 的匹配阈值, 较好地解决传统检测器生成算法在生成检测器集 R 过程中存在的不规律和跳变等现象, 实现以较小的代价达到检测较大范围的非自体 (Non-self) 行为的目的, 从而提高系统的检测性能和效率.

2.2 算法的执行过程

1) 训练阶段算法描述如下.

Step 1: 初始化ISM检测器. ISM的Self集定义为 S, 并将正常网络数据提取的Self模式放入ISM的 S 中; 初始检测器集 (成熟检测器集) 定义为 R, 并初始化 R 为空集.

Step 2: 生成ISM候选检测器集 N_{R0}. 通过随机

发生器生成一串字符串, 其长度为l, 经格式化处理后放入ISM的候选检测器集N_{R0}; 同时, 抽取ISM已有检测器中的部分规则, 进行相应编码、交叉变异后放入N_{R0}; 此外, 采用变异和优化策略从ISM的记忆检测器集R_M中挑选出一部分检测器, 经变异处理后放入N_{R0}.

Step 3: 生成ISM成熟检测器集R. 将产生的ISM自体集S与候选检测器集N_{R0}进行阴性选择, 生成一个有效的检测器集R作为ISM的初始检测器, 即成熟检测器集.

Step 4: 判断ISM成熟检测器集R中的检测器数目. 如果ISM成熟检测器集R中的检测器数目达到设定的阈值, 则停止此阶段的训练, 转Step 5, 否则, 转Step 2继续运行.

Step 5: 优化ISM成熟检测器集R. 该算法虽然在选取ISM的候选检测器集N_{R0}的数据来源上进行了改进, 但由于匹配规则的限制, 生成ISM的成熟检测器集R中仍有大量重叠信息, 这将消耗系统的大量内存空间, 同时也会降低系统的检测效率. 为此, 需要对该R中的重叠信息进行“变异”处理, 进一步优化该检测器集, 从而提高系统的检测性能与效率.

2) 变异阶段算法描述如下.

Step 1: 初始化参数. 成熟检测器集R的匹配阈值定义为β, 上限值定义为$\beta_{\rm up}$, 下限值定义为$\beta_{\rm down}$; 检测器的相似度定义为ξ. 依据生成的成熟检测器集R的规模初始化这些参数.

Step 2: 筛选ISM成熟检测器集R. 借鉴进化论中小生镜策略, 筛选R中重叠部分较为严重的检测器, 组成新的重叠检测器集R'.

Step 3: 分类重叠检测器集R'. 依据匹配阈值的上下限值$\beta_{\rm up}$和$\beta_{\rm down}$对R'作进一步筛选, 生成两个检测器集$R_{\rm up}$和$R_{\rm down}$. $R_{\rm up}$表示匹配程度高且重叠部分较多, $R_{\rm down}$表示匹配程度低且重叠部分较多.

Step 4: 变异重叠检测器集. 对检测器集$R_{\rm up}$进行低频变异, 对检测器集$R_{\rm down}$进行高频变异.

Step 5: 对变异后的检测器集进行阴性选择. 将经过变异后的检测器集再次与ISM的自体集S进行阴性选择处理, 同时将处理后的结果存储到ISM的成熟检测器集R.

3) 检测阶段算法描述如下.

Step 1: 初始化参数. ISM的记忆检测器集定义为R_M, 并初始化R_M为空集; ISM的Non-self集定义为N, 并初始化N为空集; 初始化ISM成熟检测器集R的生命周期T、激活阈值δ和检测时间T_x.

Step 2: 生成ISM的Non-self集N. 将所获得的外部网络数据预处理为ISM检测系统可以识别的模式串, 并存储到ISM的Non-self集N中.

Step 3: 匹配ISM的记忆检测器集R_M和Non-self集N. 如果二者匹配, 则产生报警并转Step 7, 否则, 转Step 4运行.

Step 4: 匹配ISM的成熟检测器集R和Non-self集N. 如果二者匹配, 则产生报警并转Step 7, 否则, 转Step 2运行.

Step 5: 优化ISM成熟检测器集R. 在R的生命周期T内, 如果该检测器没有检测到任何异常, 则将其删除.

Step 6: 优化ISM记忆检测器集R_M. 在R的生命周期T内, 如果其匹配阈值β达到设定的激活阈值δ, 则将R标注时间标志; 若此时又收到了协同刺激信号, 则将R进化为记忆检测器, 并存储到ISM的记忆检测器集R_M. 在此基础上, 还要从R_M中选取部分对Non-self模式适应度较高的数据, 经克隆交叉、变异后, 存储到ISM的基因数据库(GD)中. 这样, 可以缓解成熟检测器集R和记忆检测器集R_M的存储性能.

Step 7: 判断报警的真实性. 如果系统接收的报警信息是真实的, 则将该报警信息反馈到ISM系统的管理中心作进一步处理, 并修改对应R的匹配阈值β, 然后转Step 2运行.

Step 8: 判断R的检测时间. 如果该检测器达到了其检测时间T_x, 则结束此阶段检测, 否则, 转Step 2继续运行.

2.3　记忆检测器集R_M的优化

在系统检测过程中, 随着越来越多的成熟检测器被激活, 并成为记忆检测器, 必将导致ISM的记忆检测器集R_M过于庞大, 从而影响整个系统的检测性能. 为此, 笔者借鉴生物免疫系统中抗体的克隆选择和亲和度变异机制[20-21], 给出了记忆检测器集R_M的优化算法. 该算法的具体思路如下.

假定: r_m是ISM的记忆检测器集R_M中的一个记忆检测器, 对于其中任意一个r_m, 均有一个亲和度$f(r_i)$与之对应. f_0为预先设定的各记忆检测器r_i对应的亲和度$f(r_i)$的初值, x是增长因子, y是衰减因子, x和y都是预先设定的非常小的正实数, C也是预先设置的正实数.

Step 1: 为每一个记忆检测器的亲和度赋初值, 即令$f(r_i)=f_0$, 其中$i=1,2,\cdots,m$, m是当前记忆检测器的数目.

Step 2: 在每一次检测过程中, 根据下式修正该记

忆检测器的亲和度 $f(r_i)$:

$$f(r_i)=\begin{cases} f(r_i)+x, \text{ 已被激活;} \\ f(r_i)-y, \text{ 未被激活.} \end{cases} \tag{7}$$

Step 3: 优化记忆检测器集 R_M. 如果记忆检测器的亲和度 $f(r_i)<C$, 则删除该记忆检测器, 并依据检测系统的实际情况适时修改各参数值, 否则, 转 Step 2 继续运行. 对于被删除的记忆检测器, 系统将进行优化变异, 并存储到 N_{R0}, 再与 S 进行阴性选择, 生成新的检测器 R'', 最后存储到 ISM 成熟检测器集 R 中.

记忆检测器集 R_M 经过优化处理后, 能够使其维持在一个适中的数目, 以便于进行快速检测和动态更新, 从而对常见的网络入侵行为具有较强的自适应能力和较高的检测性能.

3 算法模型优点

在检测器的生成过程中, 该算法模型通过训练-变异-检测-变异等阶段, 与其他的检测器生成算法相比, 不仅加强了系统对未知入侵和已知入侵的变异的检测分析能力, 而且降低了检测器集的冗余度, 增强了其完备性, 减少了虚警的发生[22-23]. 该算法的主要特色体现在以下几个方面.

1) 数据来源的多样性. 在选取 ISM 候选检测器集 N_{R0} 的数据来源上, 除随机生成外, 还引入了已知异常变异和记忆检测器集 R_M 的优化变异等, 这样不仅保证了 N_{R0} 的多样性和灵活性, 而且强化了检测器对已知入侵行为的识别能力, 从而较好地提高了系统的检测性能.

2) 记忆检测器的优胜劣汰. 对 ISM 记忆检测器集 R_M 采用克隆抑制和亲和度变异, 不仅使优秀抗体 (匹配规则) 得以保存, 而且减少了 R_M 的冗余计算, 避免了因 R_M 数量过于庞大而影响系统的检测效率.

3) 引入小生境策略. 小生镜策略[24-25]的引入较好地减轻了检测器的冗余度, 提高了检测器的生成效率和搜索性能.

4) 采用多种变异和优化机制. 通过对各类检测器的变异 (包括高频变异和低频变异) 和优化等, 不仅降低了检测器的冗余度, 而且增强了系统的检测分析能力和自适应性.

5) 加强了协同信号管理机制. 通过协同信号管理机制[26]的引入, 不仅避免了 ISM 检测器产生的自免疫, 而且强化了检测器的免疫应答机制, 并大大降低了虚警率的发生.

6) 成熟检测器标注时间. 对 ISM 成熟检测器集 R 中的数据标注时间, 不仅优化了检测器的数据集, 而且强化了检测器的检测效率和性能.

4 算法性能分析和实验验证

为验证基于 ISM 特性的检测器生成算法的有效性和自适应性, 该实验选取用特征提取的字符串构成不同大小的 Self 集 S, 在相同实验环境 (CPU: Intel 英特尔酷睿 2 双核 E7500, 内存 2GB, 操作系统 Windows Server 2000, 编程语言 Visual C++ 6.0) 下运行, 然后分别采用基于 ISM 特性的检测器生成算法、穷举法、线性法和贪心法生成相应的成熟检测器集 R, 分析和比较各种算法在生成 R 的效率和性能上的差异. 预先设定参数: 字符串中所含的符号数目 $m=2$, 字符串的长度 $l=32$, 匹配长度 $r=12$, 检测失败概率 $P_f=0.100$, 记录对于不同规模的 Self 集各算法的运行时间和失败概率. 计算 1 000 次实验的 P_f 均值, 实验结果如表 1 所示.

表 1 各种检测器生成算法的测试性能对比

N_s (Self 规模)	穷举算法		线性算法		贪心算法		ISM检测器算法	
	时间 / ms	失败概率 P_f	时间 / ms	失败概率 P_f	时间 / ms	失败概率 P_f	时间 / ms	失败概率 P_f
8	2	0.100	9	0.099	31	0.085	8	0.058
16	4	0.087	12	0.101	53	0.078	10	0.052
32	11	0.098	30	0.106	88	0.076	27	0.063
64	50	0.112	49	0.111	133	0.082	48	0.071
128	1 268	0.139	90	0.128	201	0.089	76	0.079
256	3 986	0.168	158	0.156	312	0.095	142	0.086

从表 1 中的测试结果可以看出: 随着 Self 集规模 N_s 的增大, 穷举算法所耗费的时间呈指数级递增, 而线性算法、贪心算法和基于 ISM 特性的检测器生成算法所耗费的时间基本上呈线性递增, 特别是当 N_s 较大时, 本文提出的算法在时间代价上明显优于其他 3 种算法; 从所生成的成熟检测器集 R 的性能看, 本文检测器生成算法的失败概率 P_f 明显优于其他 3 种算法, 这是因为本文算法在生成成熟检测器过程中经过了一系列变异和优化等操作, 它能够覆盖更多 Non-self 空间; 在整个测试过程中, 本文算法的失败概率 P_f 始终小于预先设定的 0.100, 由此可以看出, 基于 ISM 特性的检测器自适应生成算法不仅在合格

检测器的生成质量上有了较大改善, 而且在合格检测器的生成效率上也有了明显提高, 算法的有效性达到了预期目标.

在上述实验的基础上, 笔者还从经典KDD CUP 99测试数据集中抽取了8 000条数据作为本实验的测试训练集, 其中4 500条作为Normal数据, 3 500条作为Abnormal数据. 训练集中的Normal数据存放于Self集S, Abnormal数据包括DoS和R2L两类攻击. 预先设定成熟检测器的个数为200, 进化代数为9. 对于不同变异次数n, 记录所提出算法的检测率、漏报率与变异次数n的对应关系, 实验结果如图2所示.

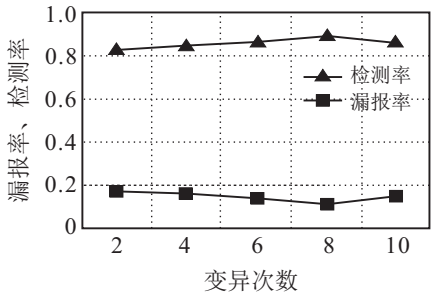

图 2　检测率、漏报率与变异次数的对应关系

从图2的测试结果可以看出: 基于ISM特性的检测器自适应生成算法的检测率和漏检率与变异次数n的对应关系均达到了预期目标; 随着n的增大, 所提出算法的检测率和漏报率均得到了很好的改善, 但是, 当n的值增长到一定数值时, 二者均趋于稳定, 此时如果再增加n的值, 则必将给系统带来额外负担, 检测器的合理分布也将遭到破坏, 最终导致系统的检测性能下降. 因此, 在对ISM的成熟检测器进行变异操作时, 应根据其成熟检测器的规模和进化代数等参数合理选定变异次数. 这样, 一方面可以较好地减少冗余的检测器, 增强ISM的检测器集的完备性, 降低虚警率的发生, 进而提高系统的检测效率; 另一方面, 可以有效地提高系统资源的利用率, 增强基于Multi-ISM联盟的网络入侵检测与防御系统的检测性能和自适应性.

5　结　论

本文针对传统检测器生成算法在检测效率和性能上存在的局限性, 在剖析这些算法的基础上, 借鉴小生境策略中“抗体”向“抗原”进化的思想, 以及对重叠检测器的变异和优化等操作, 构建了新型的基于ISM特性的检测器自适应生成模型及算法. 实验结果表明: 所提出的算法在检测性能上明显优于其他的检测器生成算法, 较好地抑制了“黑洞”的产生, 尤其在检测未知入侵和已知入侵变异的能力上得到较大幅度的提升. 同时, 该算法能够以较小的检测器集来检测较大范围的Non-self行为, 具备了一定的自学习和自适应能力, 是一个高效的检测器自适应生成算法.

参考文献(References)

[1] Behjat A R, Vatankhah N, Mustapha A. Feature subset selection using genetic algorithm for intrusion detection system[J]. Advanced Science Letters, 2014, 20(1): 235-238.

[2] Hu W, Gao J, Wang Y, et al. Online adaboost-based parameterized methods for dynamic distributed network intrusion detection[J]. IEEE Trans on Cybernetics, 2014, 44(1): 66-82.

[3] Silva G C, Palhares R M, Caminhas W M. Immune inspired fault detection and diagnosis: A fuzzy-based approach of the negative selection algorithm and participatory clustering[J]. Expert Systems with Applications, 2012, 39(16): 12474-12486.

[4] Forrest S, Perelson A S, Allen L, et al. Self-nonself discrimination in a computer[C]. Proc of the IEEE Symposium on Research in Security and Privacy. Oakland: IEEE Computer Society Press, 1994: 202-212.

[5] Esponda F, Forrest S, Helman P. Negative representations of information[J]. Int J of Information Security, 2009, 8(5): 331-345.

[6] D'haeseleer P, Forrest S, Helman P. An immunological approach to change detection: Algorithms, analysis and implications[C]. Proc of the IEEE Symposium on Computer Security and Privacy. Oakland, 1996: 110-119.

[7] Dasgupta D, KrishnaKumar K, Wong D, et al. Negative selection algorithm for aircraft fault detection[M]. Artificial Immune Systems. Berlin: Springer, 2004, 3239: 1-13.

[8] Choi Y H, Jung M Y, Seo S W. A fast pattern matching algorithm with multi-byte search unit for high-speed network security[J]. Computer Communications, 2011, 34(14): 1750-1763.

[9] Gorbenko A, Popov V. Self-learning algorithm for visual recognition and object categorization for autonomous mobile robots[M]. Computer, Informatics, Cybernetics and Applications. Netherlands: Springer, 2012: 1289-1295.

[10] Ji Z, Dasgupta D. V-detector: An efficient negative selection algorithm with“probably adequate”detector coverage[J]. Information Sciences, 2009, 179(10): 1390-1406.

[11] Wong N, Ray P, Stephens G, et al. Artificial immune systems for the detection of credit card fraud: An architecture, prototype and preliminary results[J]. Information Systems J, 2012, 22(1): 53-76.

[12] 马占飞. 基于免疫机理与“软件人”技术的网络安全系统研究[D]. 北京: 北京科技大学计算机与通信工程学院, 2008.
(Ma Z F. Study of network security systems based on immunity and SoftMan technology[D]. Beijing: School of Computer & Communication Engineering, University of Science and Technology Beijing, 2008.)

[13] 张衡, 吴礼发, 张毓森, 等. 一种 r 可变阴性选择算法及其仿真分析[J]. 计算机学报, 2005, 28(10): 1614-1619.
(Zhang H, Wu L F, Zhang Y S, et al. An algorithm of r-adjustable negative selection algorithm and its simulation analysis[J]. Chinese J of Computers, 2005, 28(10): 1614-1619.)

[14] Dasgupta S, Das S, Abraham A, et al. Adaptive computational chemotaxis in bacterial foraging optimization: An analysis[J]. IEEE Trans on Evolutionary Computation, 2009, 13(4): 919-941.

[15] Hofmeyr S, Moore T, Forrest S, et al. Modeling internet-scale policies for cleaning up malware[M]. Economics of Information Security and Privacy III. Berlin: Springer, 2013: 149-170.

[16] Allen L V, Tilbury D M. Anomaly detection using model generation for event-based systems without a preexisting formal model[J]. IEEE Trans on Systems, Man and Cybernetics, Part A: Systems and Humans, 2012, 42(3): 654-668.

[17] Arshad J, Townend P, Xu J. An automatic intrusion diagnosis approach for clouds[J]. Int J of Automation and Computing, 2011, 8(3): 286-296.

[18] Abadeh M S, Habibi J, Daneshi M, et al. Intrusion detection using a hybridization of evolutionary fuzzy systems and artificial immune systems[C]. Proc of the IEEE Congress on Evolutionary Computation. Singapore, 2007: 3547-3553.

[19] 姜恩龙. 基于否定选择的检测器生成算法研究[D]. 哈尔滨: 哈尔滨理工大学计算机科学与技术学院, 2007.
(Jiang E L. Research of detector generation algorithm based on negative selection[D]. Harbin: School of Computer Science and Technology, Harbin University of Science and Technology, 2007.)

[20] Hoque M S, Mukit M, Bikas M, et al. An implementation of intrusion detection system using genetic algorithm[J]. Int J of Network Security & Its Applications, 2012, 4(2): 109-120.

[21] Weimer W, Forrest S, Le Goues C, et al. Automatic program repair with evolutionary computation[J]. Communications of the ACM, 2010, 53(5): 109-116.

[22] Lin S W, Ying K C, Lee C Y, et al. An intelligent algorithm with feature selection and decision rules applied to anomaly intrusion detection[J]. Applied Soft Computing, 2012, 12(10): 3285-3290.

[23] Mailloux L O, Grimaila M R, Hodson D D, et al. Performance evaluations of quantum key distribution system architectures[J]. IEEE Security & Privacy, 2015, 13(1): 30-40.

[24] Hofmeyr S. The implications of immunology for secure systems design[J]. Computers & Security, 2004, 23(6): 453-455.

[25] Pereira C M N A, Sacco W F. A parallel genetic algorithm with niching technique applied to a nuclear reactor core design optimization problem[J]. Progress in Nuclear Energy, 2008, 50(7): 740-746.

[26] Holme P, Karlin J, Forrest S. An integrated model of traffic, geography and economy in the internet[J]. ACM SIGCOMM Computer Communication Review, 2008, 38(3): 5-16.

[27] Papadogiannakis A, Vasiliadis G, Antoniades D, et al. Improving the performance of passive network monitoring applications with memory locality enhancements[J]. Computer Communications, 2012, 35(1): 129-140.

(责任编辑：曹洪式)