一种基于非单调逻辑理论的入侵检测系统

张 剑 龚 俭
(东南大学计算机科学与工程系 南京 210096)

摘 要 提出了用模糊默认理论改造传统的单调推理机制和响应引擎的方法,从而建立了基于人工智能的入侵检测系统.实验结果表明,改进后的系统不仅能适应高速主干网络的实时入侵检测需要,而且灵敏性有很大的提高;由于采用了响应回卷技术和面向代价的动态响应政策,从而大幅度降低了入侵检测和响应的代价.

关键词 模糊默认逻辑;入侵检测;单调逻辑;响应回卷
中图法分类号 TP393

Intrusion Detection System Based on Non-Monotonic Logic Theory

ZHANG Jian GONG Jian
(*Department of Computer Science and Technology, Southeast University, Nanjing* 210096)

Abstract This paper applies fuzzy default theory to transform reasoning mechanism and automated response system of intrusion detection system(IDS), and sets up an intelligent IDS——FDL-IDS. The experiment results show that FDL-IDS increases the detection speed and sensitivity and decreases the cumulative cost as compared with traditional intrusion detection expert system.

Keywords fuzzy default logic; intrusion detection; monotonic logic; response rollback

1 引 言

随着网络技术和网络规模的不断发展,网络入侵的风险性和机会也越来越多,网络安全已经成为人们无法回避的问题.因此为了保护现在越来越多的敏感信息,入侵检测系统(Intrusion Detection System,IDS)也成为了一种非常重要的技术,得到了越来越多的重视.

将人工智能技术应用到入侵检测中是目前国际上研究的热点,Dickerson[1]在 2001 年提出建立模糊入侵检测系统,其中主要技术是将普通规则改造成模糊规则,从而能更加精确地建立自然语言的知识与计算机知识之间的映射;Siraj[2]在 2001 年提出建立模糊认识图来支持智能入侵检测系统做出决策.这种模糊认识图反映了事件之间的模糊因果关系,并可用于计算事件的置信度.入侵决策引擎可以根据它作出更明智的决策;Geib[3]在 2001 年提出在入侵检测系统中结合人工智能方法来解决入侵意图识别.

为了实现一定的安全目标,通常要求 IDS:

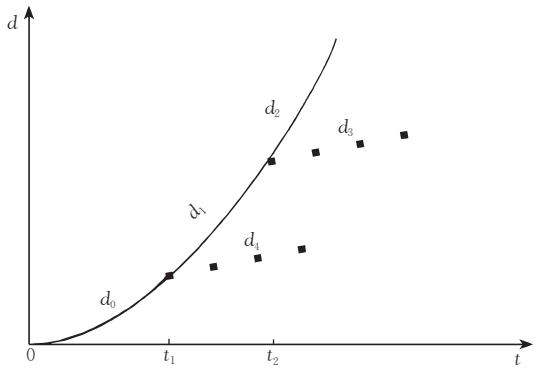
(1)在一定的准确度保证下能尽早地发现入侵意图并做出响应.随着入侵的进一步发展,对受保护对象(例如服务器)的破坏可能更加严重,因此要求检测和响应的灵敏性要高.

图 1 是华东(北)地区 CERNET 某次 Nimda 病毒发作时随时间推移造成的损失量化图.图中将受保护对象的损失假设为随入侵的不断深入而连续变化的可量化的指标.可以看到如果没有入侵检测系统的保护,损失将以指数级曲线上升;而如果检测与响应速度较慢时,它将在 t_2 时刻做出响应,这时损失将减缓,以线性速度上升到一定程度而停止;而如果

收稿日期:2002-10-08;修改稿收到日期:2003-03-13.本课题得到国家自然科学基金(90104031)资助.**张 剑**,男,1977 年生,博士研究生,主要研究方向为网络安全检测.E-mail:zhangj@njnet.edu.cn.**龚 俭**,男,1957 年生,教授,博士生导师,主要研究方向为网络安全、网络管理和网络体系结构.

检测和响应的速度较快时,它在 t_1 时刻就做出响应,从而降低了损失.

(2)在误报的情况下,应中止或取消原来的响应动作而采取正确的响应动作,以减少或消除由于错误响应带来的损失.这称为响应回卷.



$d_0+d_1+d_2$:入侵进度损失程度曲线
$d_0+d_1+d_3$:检测与响应进度损失程度曲线(慢速)
d_0+d_4:检测与响应进度损失程度曲线(快速)

图 1　检测与响应速度对受保护对象的影响

要做到这两点一般比较困难,因为入侵检测系统需要收集到足够的证据才能判断当前行为是否是入侵行为,而许多证据的收集不仅需要代价,而且往往需要贯穿于入侵的整个过程,因而无法较早地发现该行为并做出响应;由于现有的入侵检测系统采用经典逻辑推理方式,根据目前已掌握的证据来推导出的结论及相应的响应动作被视为永真结果,即使将随后获得的新知识或事实加入推理过程后表明该结论是错误的,也不能推翻原有的结论和消除其响应动作.

模糊默认逻辑是一种非单调逻辑,它是模糊逻辑和默认逻辑的有机结合.利用模糊逻辑理论可以改造传统的入侵检测知识库和支持模糊推理,利用默认逻辑理论可以支持入侵意图识别和响应回卷.因此它能较好地解决上述两个安全目标.

本文提出建立完整的基于模糊默认逻辑的入侵检测系统 FDL-IDS(Fuzzy Default Logic Based Intrusion Detection System)的方法,该系统包括基于模糊默认逻辑的知识库 FDL-KB、相对应的入侵推理引擎 FDL-IRE 和可回卷的响应引擎 RRE(Rollbackable Response Engine)3 个主要部件.FDL-IRE 在 FDL-KB 的支持下用基于模糊默认逻辑的推理方法取代经典逻辑推理来判断入侵行为;而 RRE 根据代价模型计算响应与否的代价比来决定如何响应,而且可以回卷不正确的响应动作.其优点在于:

(1)FDL-IRE 能在证据不充分的情况下判断当前行为是否是入侵行为,从而提高了检测的速度.

(2)当随后收集到的新知识表明原来的结论错误时,FDL-IRE 就将新知识加入到推理过程中,得出新的结论,然后向 RRE 请求回卷原来的响应动作和做出新的响应动作.

(3)衡量代价来决定响应动作,使响应更加智能和科学.

2　模糊默认逻辑的基本概念

2.1　入侵检测系统的逻辑非单调性

定义 1[4].　一种逻辑 L 可用一个四元组表示:$L=\{E,O,A,R\}$,其中 E 表示原子逻辑公式的集合;O 是定义在 E 及运算而得的合式逻辑公式之上的一组运算构成的集合;A 表示合式逻辑公式的一个子集,称为公理集;R 是由公理集合 A 或已证的合式逻辑公式推出的 L 中其它合式逻辑公式的推理规则的集合.由 E 开始利用 O 中运算所能得到的全体逻辑公式称为 L 中的合式逻辑公式的集合.由公理集 A 中的元素开始,通过有限次使用 R 中规则进行推理,所能推得的合式逻辑公式,称为定理.

定义 2[4].　设 $L=\{E,O,A,R\}$ 为一种逻辑,在 L 中所能证明的定理的全体记为 $T(L)$.若对任意的合式逻辑公式的子集 A',$A'\supseteq A$,在逻辑 $L'=\{E,O,A',R\}$ 中所能证明的定理的全体 $T(L')\supseteq T(L)$,则称 L 是一个单调逻辑,否则称为非单调逻辑.

定理 1.　存在误报的入侵检测系统是非单调逻辑系统.

证明.

反证法.入侵检测系统总是从某些已知的知识出发,通过某种检测算法(异常检测)或检测规则(滥用检测)判断入侵行为,因此可被看成是逻辑系统.设该逻辑系统为 $L=\{E,O,A(a,t),R\}$,A 是从任意时刻 a 开始的 t 时间内采集到的数据,相当于已知的知识;根据当前知识入侵检测系统检测到的入侵事件集合记为 $T(A(a,t))$,它相当于定理集合,即 $T(A(a,t))=T(L)$.当 $t_1\geq t$,则有 $A(a,t)\subseteq A(a,t_1)$;对于新逻辑系统 $L_1=\{E,O,A(a,t_1),R\}$,L_1 所能证明的定理集合为 $T(A(a,t_1))$.

入侵检测系统存在的误报可形式化描述为

$$\forall a\,\forall t\,\forall t_1\,\exists e((t\leq t_1)\wedge(e\in T(a,t))\rightarrow(e\notin T(a,t_1))) \tag{1}$$

假设 L 是单调逻辑系统,则

$$\forall a\,\forall t\,\forall t_1\,\forall e((t\leq t_1)\wedge(e\in T(a,t))\rightarrow$$

$$(e \in T(a,t_1))) \qquad (2)$$

式(2)与式(1)相矛盾,因此 L 是非单调逻辑系统.

证毕.

设某入侵方式为 U,它可由以下析取范式描述:

$$U = F_1 \cup F_2 \cup \cdots \cup F_n,$$

其中 $F_i \Leftrightarrow f_1 \cap f_2 \cap \cdots \cap f_{m_i}, i=1,2,\cdots,n.$

在这里 F_i 是判断 U 的规则,规则与规则之间在判断 U 上是相互独立的;每条规则由一组特征 $f_1, f_2, \cdots, f_m$ 的交组成前提,只有满足某条规则的所有特征才能触发该规则.

设规则 F 的特征集合为 $S(F)$,$S(F)=\{f_1, f_2, \cdots, f_m\}$. 入侵检测系统的逻辑非单调性是由于:

(1) 对规则的描述不完整

设规则 F 被不完整地描述为 F',即 $S(F') \subset S(F)$. 规则描述得不完整就会将正常行为误判为入侵行为,尽管这些特征在正常行为中出现的概率比较小.

(2) 许多规则的本质是定性的、模糊的知识

这个原因包含两层意思,第一,规则有一定的置信度,通常是有很大的置信度,而不是绝对的正确;第二,规则前提的定义是模糊的,通常只适合用自然语言描述,通常这也是造成漏报的一个原因. 具有这两类特征的规则称为模糊规则,它是普通规则的推广. 由于规则的模糊性,用一般的规则表示方法和推导方法就会使问题绝对化,从而造成误报.

2.2 模糊默认逻辑的基础理论

默认推理逻辑(default reasoning logic)是 Reiter 在 1980 年提出的一种非单调逻辑,他将人类在进行不充分证据下推理的某种情况提炼出来并加以形式化,对人工智能及其它领域很有启发意义. 在入侵检测技术中运用默认推理逻辑可以在掌握不充分证据的情况下判断入侵行为,因此比普通的入侵检测速度要迅速和灵敏. 关于默认逻辑的完整理论请参见参考文献[5].

默认推理的缺点是缺乏可信度参数,而对于 IDS 等科学考察,最为关心的是如何得出可靠性较高的结论以及通过什么途径能够进一步提高结论的可靠性. 模糊逻辑能较好地弥补默认推理逻辑的这个缺点. 为了满足在 IDS 中应用的需要,本文提出一种基于模糊逻辑的默认逻辑理论:模糊默认逻辑理论.

定义 3. 模糊默认规则的一般形式为

$$\alpha(\vec{x}) : M\beta_1(\vec{x}), M\beta_2(\vec{x}), \cdots, M\beta_m(\vec{x}) \rightarrow \omega(\vec{x}) \quad CF, \tau.$$

其中 $\vec{x}=\langle x_1, x_2, \cdots, x_n\rangle$,$\alpha(\vec{x})$,$M\beta_1(\vec{x})$,$M\beta_2(\vec{x})$,$\cdots$,$M\beta_m(\vec{x})$,$\omega(\vec{x})$ 为模糊逻辑的合式公式;$\alpha(\vec{x})$ 称为规则的先决条件,它是已知的事实或模糊事实;各个 $M\beta_i(\vec{x})$ 称为规则的默认条件,它是一种概率事件,其出现通常与先决条件有很大的关联;M 表示"可能"的意思;$\omega(\vec{x})$ 叫做规则的结论;τ 称为应用阈值. τ 是规则先决条件的应用阈值,只有 $T(\alpha(\vec{x})) \geq \tau, T(M\beta_i(\vec{x})) > 0.5, i=1,2,\cdots,m$ 时,该规则才被激活. CF 是规则的可信度,通常为[0,1]区间的一个数.

定义 4. 设 $L=\{E,O,A,R\}$ 为模糊默认逻辑,E 是原子模糊逻辑公式的集合,$O \subseteq \{\neg, \vee, \wedge, \rightarrow\}$,$A$ 是模糊命题的集合,R 是模糊默认逻辑的推理规则,表示为模糊默认逻辑的三段论.

模糊默认逻辑的三段论是模糊逻辑三段论的推广,它表示为

$$\begin{array}{lll}
\alpha : M\beta_1, M\beta_2, \cdots, M\beta_m \rightarrow \omega & & CF, \tau \\
\alpha & & T(\alpha) \\
\beta_i & & T(\beta_i), i=1,2,\cdots,m \\
 & CF+\min(T(\alpha), T(\beta_1), \cdots, T(\beta_m)) > 1 & T(\alpha) \geq \tau \\
\hline
\omega & T(\omega)=\min(T(\alpha), T(\beta_1), \cdots, T(\beta_m))+CF-1 &
\end{array}$$

闭模糊默认规则、闭模糊默认理论和规范模糊默认理论的定义与默认逻辑的定义相似.

3 FDL-KB 和 FDL-IRE

3.1 模糊默认知识库 FDL-KB 的原理

传统入侵检测专家系统的知识库包括事实和普通规则,这种知识库存在以下的缺点:

(1)知识是静态的. 知识一旦输入就很少更新,特别是知识的修改和删除.

(2)普通事实和普通规则不能完整地描述知识. 一些模糊概念、信念和经验等无法用普通事实和普通规则表示.

(3)知识库需要人工维护. 新知识的补充主要依赖人工的输入,知识的不一致也是人工检查.

鉴于传统知识库的上述缺点,本文提出建立模糊默认知识库 FDL-KB 取代传统知识库. FDL-KB 相当于闭规范模糊默认逻辑系统 $\{E,O,A,R\}$ 中的 A,它由模糊默认规则子库和关联规则子库构成.

设某入侵检测系统的测度集合为 C. 某测度的收集代价是指根据其占用系统资源(如存储和计算)和收集消耗时间来计算的. 根据某种收集代价计算规则,可以将 C 分为 n 个不相交的子集,称收集代价最小的测度集合为第一类测度,而收集代价最大的测度集合为第 n 类测度. 其余的以此类推.

高类的测度的收集不但耗费大量的系统资源(例如一些需要在报文中匹配某个字符串的测度,需要在整个报文中搜索该字符串),而且严重影响检测的速度,因为这些测度往往在入侵完成后才能收集到,这时候入侵造成的损失已经产生甚至最大化了.

利用模糊默认规则来改造普通的入侵检测规则,可以达到避免收集高类测度的目的.这时高类测度作为模糊默认规则的默认条件,而低类测度作为规则的先决条件.这样的规则不但使得收集入侵证据更加简单、快速,而且使得检测更加迅速和灵敏.

设某入侵检测系统的测度集合为 C,$c\in C$. $D(c,a)$是谓词,表示 c 的值为 a. 设 $c_1,c_2,\cdots,c_n\in C$,用数据挖掘技术发现存在关联规则:$\bigwedge_{i=1}^{n-1} D(c_i,a_i)\rightarrow D(c_n,a_n)[Supp,Conf]$,$Supp$ 是该关联规则的支持度,它相当于 $P(D(c_1,a_1)\wedge D(c_2,a_2)\wedge\cdots\wedge D(c_n,a_n))$,即 c_i 的值为 a_i 同时出现的概率,$i=1,2,\cdots,n$;$Conf$ 是该关联规则的可信度,它相当于 $P(D(c_n,a_n)|D(c_1,a_1)\wedge\cdots\wedge D(c_{n-1},a_{n-1}))$,即 c_i 的值为 a_i 同时出现时 c_n 的值为 a_n 的概率,$i=1,2,\cdots,n-1$. 当 $Conf>0.5$ 时,可以建立闭规范模糊默认规则:

$$\bigwedge_{i=1}^{n-1} D(c_i,a_i):MD(c_n,a_n)\rightarrow D(c_n,a_n)CF,\tau,$$

其中 $CF=1$,τ 由领域专家给出.

定义 5.　从普通规则到闭模糊默认规则的转换步骤是:

(1)将普通规则的前提转换成析取范式,然后分解成 n 个子规则.表示为

Ⅰ. $P\rightarrow Q\Rightarrow\bigvee_{i=1}^{n} P_i\rightarrow Q$,其中 P_i 是原子逻辑公式的交.

Ⅱ. $\bigvee_{i=1}^{n} P_i\rightarrow Q\Rightarrow\begin{cases}P_1\rightarrow Q\\ \vdots\\ P_n\rightarrow Q\end{cases}$.

(2)将子规则中语义是模糊的原子逻辑公式转化为模糊原子逻辑公式.

设 x 是形式为"a 是 A"的模糊命题,其中 A 是用模糊集合表示的模糊概念,则 $T(x)=\mu_A(a)$.

(3)将子规则转化为闭规范模糊默认规则.

Ⅰ. $P_i\rightarrow Q\Rightarrow P_i\rightarrow Q\quad CF,\tau$

Ⅱ. $P_i\rightarrow Q\quad CF,\tau\Rightarrow P_i:MQ\rightarrow Q\quad CF,\tau$

其中Ⅰ中的转化结果是模糊规则,CF 和 τ 由领域专家指定.

定义 6.　设 FDL-KB 的模糊默认规则子库和关联规则子库分别为集合 θ 和 ξ,模糊默认知识库的创建步骤是:

(1)$\theta=\{$普通规则$\}$,$\xi=\varnothing$.

(2)利用定义 5 将 θ 中所有普通规则转化为模糊默认规则,这时 $\theta=\{$模糊默认规则$\}$.

(3)建立第 i 类测度到第 j 类测度的关联规则,其中 $i<j$.

(4)将 $Conf\leq 0.5$ 的关联规则过滤掉,其余的关联规则输入 ξ,$\xi=\{$关联规则 $A|Conf_A>0.5\}$.

(5)根据 ξ 建立相应的闭规范模糊默认规则集合 P. 这时 $\theta=\theta\cup P$.

(6)当 ξ 的更新时间来到,将 ξ 清空以及将 P 从 θ 中删除,即 $\theta=\theta-P$,$\xi=\varnothing$.

(7)回到(3)开始循环.

3.2　模糊默认推理引擎 FDL-IRE 的原理

传统的入侵检测专家系统通常默认使用经典逻辑的推理规则,即

$$\frac{A,A\rightarrow B}{B},$$

这种推理引擎的优点是简单、易于实现.但它是基于经典逻辑知识库的推理方法,当这种知识库转化为模糊默认知识库时,就应该改造原有的推理引擎,相应地建立基于模糊默认逻辑的推理引擎.

定义 7.　命题表是模糊默认推理日志,该表记录了每次推理或收集到新知识后应该添加和删除的事实.每行的内容包括推理执行的先决条件、默认条件、结论、结论的可信度和删除的事实.

命题表是 FDL-IRE 与响应引擎的接口.当某个推理动作执行后,FDL-IRE 将相关的信息记录在命题表中,然后继续处理下一个推理动作,而响应引擎则查询命题表有无新的结论产生,如果有,就提取出来进行相应的响应决策.

FDL-IRE 的推理过程可以分为以下步骤:

1. 根据数据采集器输出的数据和 FDL-KB 的模糊默认规则进行推理.

2. 如果能推导出入侵行为,则将该推理过程的信息(包括规则的先决条件、默认条件和结论的可信度等)记录到命题表中,同时检查推理结论与命题表前面的结论、先决条件与命题表前面的默认条件的一致性,如果发生不一致,则撤销原有的默认条件和结论,将它记录到命题表的删除事实中.

3. 如果不能推导出入侵行为,则检查采集到的数据与命题表前面的默认条件记录的一致性,如果发生不一致,则撤销原有的默认条件和结论,将它记录到命题表的删除事实中.然后继续处理下面的数据,返回步 1.

4 支持响应回卷的响应引擎 RRE 的原理

RRE 有两个特点：

(1)它能支持响应回卷.

(2)它基于响应代价模型做出响应决策. RRE 首先根据响应代价模型，计算出采用某种响应方式与否之间的代价比，然后根据该代价比决定如何响应.

RRE 是 FDL-IDS 不可缺少的组成部分，这是因为：

(1)非单调推理可能导致错误的结论，必须删除该结论和取消根据该结论做出的响应动作. 传统的响应引擎不能支持这一点.

(2)FDL-IRE 向响应引擎提交当前发生的入侵事件及其置信度，而置信度是代价敏感模型的重要参数，RRE 的响应是以代价敏感模型为基础的.

4.1 支持响应决策的入侵检测加权代价敏感模型

4.1.1 入侵检测及响应代价分析

(1)检测代价

指检测所消耗的资源数量，记为 $ICost$.

(2)响应代价

指响应所消耗的资源数量，记为 $RCost$.

(3)损失代价

分为三种情况，第一是不采取响应时入侵造成的损害，记为 $DCost$；第二是采取某种响应后入侵造成的损害，记为 $DICost$；第三是采取某种响应后响应动作造成的损失，记为 $DRCost$，例如关闭服务器这类响应动作对其用户的损失.

4.1.2 代价敏感模型

建立代价敏感模型前需要确定代价权重.

代价权重是每种代价类型在代价敏感模型中的重要性指标. 设代价类型的集合 $C=\{c_1,c_2,\cdots,c_n\}$，权重向量 $\boldsymbol{\omega}=[\omega_1,\omega_2,\cdots,\omega_n]$，其中 $\boldsymbol{\omega}_i$ 是 c_i 的权重，$i=1,2,\cdots,n$. $\omega_1+\omega_2+\cdots+\omega_n=1$.

代价权重能反映了某个特定环境的安全目标，它一般通过专家来指定. 可以采用模糊数学中的二元对比排序法确定代价权重，该方法先在二元对比中建立比较级，然后通过一定的算法转化为总体的顺序.

为了支持响应决策，该模型首先计算对于某种响应方式在响应和不响应两种情况下承受的代价，然后比较两者大小来确定如何响应.

响应时的代价记为 RC，表示为

$$RC=\omega_I\times ICost+\omega_R\times RCost+\omega_{DI}\times\mu\times DICost+\omega_{DR}\times DRCost.$$

不响应时的代价记为 NRC，表示为

$$NRC=\omega_I\times ICost+\omega_D\times\mu\times DCost.$$

记代价敏感模型为 M，表示为

$$M=\frac{RC}{NRC}.$$

上几式中 μ 代表该入侵事件的可信度，$\omega_I+\omega_R+\omega_D+\omega_{DI}+\omega_{DR}=1$. 当 $M\geq 1$ 时，不做响应；当 $M<1$ 时，采取 RC 最小的响应方法.

4.2 响应回卷

响应回卷是一种特殊的响应动作，它针对由于误报而产生的响应动作，响应回卷中断或撤销其动作，消除响应带来的负面影响.

在 Curtis 对 IDS 响应技术的调查的基础上，根据响应动作是否可回卷和可消除影响将其分为三类：

(1)可撤销的响应动作，包括锁住用户帐户、阻塞攻击源地址、关闭主机、从网络中断开等. 对于这些响应动作，响应回卷采取相反的动作.

(2)不可撤销但可以消除影响的响应动作，包括产生一份报告、产生警告、创建备份等. 对于这些响应动作，响应回卷可以采用中断或产生回卷日志的方法.

(3)不可撤销和消除影响的响应动作等，包括警告入侵者、中断用户会话. 对于这些响应动作，响应回卷报告安全管理员或写入日志.

响应回卷技术需要保存响应日志，响应日志记录了响应的全部动作. 一旦 RRE 接到响应回卷请求，RRE 将查询响应日志以进行响应回卷.

5 试验及其结论

5.1 试验目的

本试验的目的是测试 FDL-IDS 的检测速度和灵敏性以及评估 FDL-IDS 的代价. 单位时间平均处理报文数(Packet Per Second，PPS)作为检测速度的指标，而单位时间平均检测到的入侵事件数(Event Per Second，EPS)作为检测灵敏性的指标. 为了评估 FDL-IDS 的代价，我们采用 Lee[6] 的代价敏感模型，而检测和响应每类事件的平均代价(Cost Per type of Eevent，CPE)作为评估的指标.

Snort 是一个开放源代码的网络入侵检测系统，它用以对 IP 网络执行实时流量分析和产生可疑报文日志. 其重要功能包括协议分析、报文内容检

索/匹配和检测各种入侵类型事件.但是 Snort 没有自动响应系统,于是我们将 Snort 做如下改造:

(1)为 Snort 增加管理器,它实际是一个自动响应系统.

(2)Snort 直接向管理器报告入侵事件,而不是记录它们.

改造后的 Snort 被称为 Snort-IDS,然后 Snort-IDS 被进一步改造成 FDL-IDS:

(1)将 Snort 的入侵检测规则库改造为 FDL-KB.

(2)将 Snort 的报文匹配模块改造成 FDL-IRE.

(3)将响应回卷功能加入管理器中.

本试验将比较 FDL-IDS 与 Snort-IDS 的性能.

5.2　系统结构和试验数据

FDL-IDS 包括采集器、分析器和后处理器.采集器负责从高速信道中采集网络报文并承担简单的报文过滤工作①,然后将报文传递给分析器;分析器是该系统的核心,它根据其内部的知识库来匹配报文,如果匹配成功,则将入侵事件的相关信息报告后处理器;后处理器负责对入侵事件进行响应,它还维护着入侵事件数据库,该库记录了检测到的入侵事件的详细信息,例如源地址、目的地址、事件名称等.系统结构见图 2.

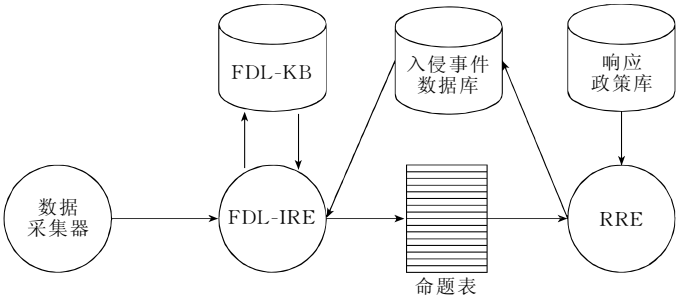


图 2　FDL-IDS 系统框架

Snort-IDS 与 FDL-IDS 有着相似的系统结构,它们的区别可归纳为表 1.

表 1　FDL-IDS 与 Snort-IDS 系统结构的区别

	FDL-IDS	Snort-IDS
分析器	包含 FDL-IRE 和 FDL-KB	包含报文匹配模块和普通规则库,没有命题表
管理器	RRE	自动响应系统,但是没有响应回卷功能

此外,FDL-KB 在系统初始运行阶段与 Snort-IDS 的规则库有相同的检测规则(尽管形式不同),即大约有 600 余条规则.这些规则包括堆栈溢出、端口扫描、CGI 攻击等攻击类型的检测,其中许多规则需要匹配报文内容中的特征,而这种匹配动作是最耗费 CPU 资源和时间的.然而正是这些规则充分显示了模糊默认规则的优势.

试验数据从 http://iris.cs.uml.edu:8080/network.html 中获得,这些通过 tcpdump 收集的数据在本试验中作为采集器的输入数据.

5.3　分析器的工作原理

FDL-KB 分为两个子库,一个包含永久性的规则,它们通常是以模糊默认规则形式表达的专家知识;另一个包含通过数据挖掘产生的临时规则,它们以 10min 的周期从入侵事件库中产生并更新,本试验采用 RIPPER(一个著名的规则发现工具)来执行数据挖掘任务,然后将发掘的规则通过 3.1 节的方法转化为模糊默认规则.这些临时规则是提高系统检测速度和灵敏性的关键,因为这些规则只需要计算低类的测度.

当系统初始运行时,FDL-KB 只包括永久性规则,因为这时在入侵事件库中没有最近的入侵事件记录.而且由于 FDL-IRE 需要计算检测到的入侵事件的置信度,因此这时候 FDL-IRE 的速度要比 Snort-IDS 的稍微慢些.但随着临时规则的产生,FDL-IRE 的速度大大提高了.

FDL-IRE 将入侵事件推理过程的相关信息记录到命题表中,当随后的数据增加到推理过程中证明原来的结论错误时,FDL-IRE 就请求 RRE 回卷原来的响应.

5.4　RRE 的工作原理

RRE 的结构如图 3 所示.

RRE 由分析引擎、回卷逻辑层和回卷物理层组成.当分析引擎接收到一个响应请求,它首先查询响应政策库,获得对当前入侵事件的各种响应方法;然后根据 4.1.2 节计算各种响应方法的代价,选取分析引擎选取代价最低的响应方法;然后分析引擎命令回卷逻辑层执行该响应方法.对于代价类型集合

① 例如根据目标端口或报文标识来过滤报文.

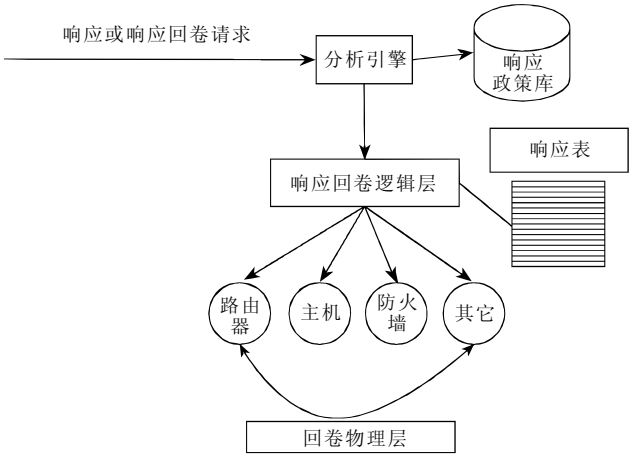


图 3 RRE 的结构

$C=\{ICost, RCost, DCost, DICost, DRCost\}$，本试验中取代价权重向量$\boldsymbol{\omega}_C=[1,1,10,10,1]$. 从代价权重向量可以看出我们更重视入侵带来的损失(包括 *DCost* 和 *DICost*).

回卷逻辑层维持一张响应表，该表包括事件编号、响应请求和响应动作等. 当它接收到响应命令，它将响应命令编译为某种自动响应脚本①，并且将响应命令记入响应表中，然后命令回卷物理层执行自动响应脚本；而当回卷逻辑层接收到响应回卷命令，它首先查找响应表，将其回卷响应命令编译为自动响应脚本，然后将其传递给回卷物理层.

回卷物理层负责具体执行自动响应脚本.

5.5 试验结果

我们在采集器(Intel ISP4400 server，Redhat 6.2)、分析器(Intel ISP2150 server，Redhat 6.2)和管理器(Sun Sparc20，Solaries 7)的硬件平台上进行试验，而入侵事件库采用 SYBASE 数据库.

表 2 FDL-IDS 与 Snort-IDS 的性能比较

	初始 PPS (M/s)	初始 EPS (M/s)	PPS(M/s)	EPS(M/s)
Snort-IDS	0.21	0.042	0.21	0.042
FDL-IDS	0.167	0.033	3.71	0.72

根据 Lee 的攻击分类方法以及代价计算思想，我们设计了以下步骤来计算 Snort-IDS 和 FDL-IDS 的累计代价.

(1)将入侵事件库中每类事件的累计代价相加. 设入侵事件集合为 E，则

$$CPE = CumulativeCost(E)/|E|.$$

(2)分析 FDL-IDS 的误报率. 这通过比较入侵事件库和真实发生的入侵事件来完成. 即若入侵事件库的事件集合为 ES，真实发生的为 ES_2，则误报率为

$$\lambda = 1 - \frac{|ES_1 \cap ES_2|}{|ES_1|}.$$

此外，由于 DICost 和 DRCost 难以量化，我们请专家来评估这些损失.

表 3 FDL-IDS 与 Snort-IDS 的 CPE 比较

	CPE	响应回卷比率(%)	不响应比率(%)
Snort-IDS	247.73	—	—
FDL-IDS	73.43	5.12	10.32

结果显示 FDL-IDS 的 PPS 与 EPS 与 Snort-IDS 的比有很大的提高. 这是因为 FDL-IDS 推导一个结论需要较少的证据. 另外，FDL-IDS 的 CPE 也比 Snort-IDS 的要低，这是因为 FDL-IRE 在许多情况下仅需要计算第一类测度，而且 RRE 大大降低了 *DRCost*. 响应回卷率意味着需要回卷的事件占所有事件的 5.12%，从而这部分事件的 *DRCost* 降低了. 不响应意味着对于某些事件不响应是代价最优的.

6 总 结

目前国际上实现的入侵检测系统大多存在检测速度和灵敏性问题，检测速度涉及到 IDS 在高速数据信道的数据处理性能，而灵敏性问题涉及到 IDS 能否做到入侵意图识别. 本文讨论利用模糊默认逻辑理论改造入侵检测系统的推理引擎，特别是入侵检测专家系统的推理引擎，试验证明该思想对提高 IDS 检测速度和灵敏性有很好的效果.

响应回卷是本文提出的另一概念，它主要针对

① 本试验将响应命令转化为 Expect 脚本，Expect 是一个交互程序自动执行脚本的命令解释器.

做出响应的误报的入侵事件.通常这也会造成损害,例如关闭服务器这种响应动作对其用户的损失.因此响应回卷是很有必要的,它能降低错误的响应带来的损失.试验证明通过响应回卷能有效地降低IDS付出的代价.

通过衡量响应与否的代价来决定是否响应是本文制定的响应策略.IDS应该以最小的代价换取最大的安全目标,这也是IDS的发展趋势.本文建立了入侵检测代价敏感模型来指导响应,该模型的先进之处在于它为每种代价分配权重,因此可以灵活地利用于各种环境中.

将上述技术有机地结合就形成了FDL-IDS,该系统以模糊默认逻辑理论为基础,通过代价敏感模型来指导响应,从而使IDS具有了一定的人工智能.

参　考　文　献

1　Dickerson J E, Juslin J, Koukousoula O, Dickerson J A. Fuzzy intrusion detection. In: Proceedings of IFSA World Congress and 20th NAFIPS International Conference, Vancouver, British Columbia, 2001. 1506～1510

2　Siraj A, Bridges S M, Vaughn R B. Fuzzy cognitive maps for decision support in an intelligent intrusion detection system. In: Proceedings of IFSA World Congress and 20th NAFIPS International Conference, Vancouver, British Columbia, 2001. 2165～2170

3　Geib C W, Goldman R P. Plan recognition in intrusion detection system. In: Proceedings of DARPA Information Survivability Conference & Exposition II, Hilton Anaheim, California, 2001. 46～55

4　He Xin-Gui. Fuzzy theories and fuzzy techniques in knowledge processing. 2nd Edition. Beijing: National Defense Industry Press, 1998(in Chinese)
(何新贵.模糊知识处理的理论与技术.第2版.北京:国防工业出版社,1998)

5　Yang Bing-Ru. Knowledge engineering and knowledge discovery. Beijing: Metallurgy Industry Press, 2000(in Chinese)
(杨炳儒. 知识工程与知识发现.北京:冶金工业出版社,2000)

6　Lee Wenke, Fan Wei, Miller Matthew, Stolfo Sal, Zadok Erez. Toward cost-sensitive modeling for intrusion detection and response. Journal of Computer Security, 2002, 10(1): 318～336

ZHANG Jian, born in 1977, Ph. D. candidate. His research interests include network security monitoring and intrusion detection.

GONG Jian, born in 1957, Ph. D., professor, Ph. D. supervisor. His research interests include network security, network management and network system architecture.