文章编号:1001 - 9081(2008)09 - 2255 - 04

基于 D-S 理论的入侵检测系统

赵晓峰
(河北工程大学 信息管理系,河北 邯郸 056038)
(qboy_best@163.com)

摘 要:单一的检测方法很难对所有的入侵获得很好的检测结果。所以,怎样将多种安全方法结合起来,为网络提供更加有效的安全保护,已经成为当前安全领域的研究热点之一。提出了一种基于数据融合的入侵检测系统,并将证据理论引入到网络安全中的入侵检测领域。该系统能够有效地解决单一检测算法无法对所有入侵都有很好检测效果的缺陷,并且相对于单一检测方法系统具有更好的可扩展性和鲁棒性。

关键词:入侵检测;证据理论;数据融合
中图分类号: TP309.2 **文献标志码**:A

D-S theory-based intrusion detection system

ZHAO Xiao-feng
(*Department of Information Management, Hebei University of Engineering, Handan Hebei* 056038, *China*)

Abstract: It is hard for single security measure to attain favourable detection result. Therefore, how to combine multiplicate security measures to provide the network system with more effective protection becomes one of the hot spots in current research. A data fusion based intrusion detection system was proposed in this paper. Multiplicate detection measures were "fused" in this system to solve the problem that single measures can not obtain good result for all intrusions, and the system has better scalabilities and robustness.

Key words: intrusion detection; evidence theory; data fusion

0 引言

入侵检测在 20 世纪 80 年代被提出后就成为计算机安全领域的研究热点之一。在过去的几十年里,人们一直把这一领域的研究重点放在寻找一个高效检测方法之上,并且也提出了多种入侵检测的方法。但是,经过了将近三十年的研究,至今还没有开发出一个在真正意义上能够投入实际运行的入侵检测系统。其主要原因是,与其他破坏性行为(比如:计算机病毒)不同,入侵手段具有多样性、复杂性和智能性的特点,而目前提出的入侵检测方法都只能对某些或某类入侵很有效,但是对另一些入侵的检测却存在很大问题,从而导致过高的漏报率和误报率。

针对目前单一检测方法存在的问题,我们希望寻找一种方法将多个入侵检测算法结合起来,共同完成检测的任务。

数据融合技术在上个世纪被提出后,经过几十年的飞速发展已经被广泛的应用于军事、地质、化工等领域。多传感器数据融合也成为了对复杂系统中大量异构数据进行分析的主要手段。在多传感器融合系统中,采用多个传感器可以获得更多的目标信息,合理利用这些信息并将数据进行整合,可以提高系统的测量精度,增强容错能力,提升其稳定性和可靠性,并最终提高系统的整体性能指标。所以,本文将多传感器数据融合的方法引入到入侵检测的研究中,对不同检测方法的结果或系统中的异构数据进行融合分析,并且从更高层面上提炼出黑客入侵场景和系统安全状况。

1 基于数据融合的入侵检测系统层次模型

对于多传感器融合层次的问题,人们存在着不同的看法[1]。本文采用普遍为学者们所接受的 3 层融合结构,即数据层、信息层和知识层。其总体结构如图 1 所示。

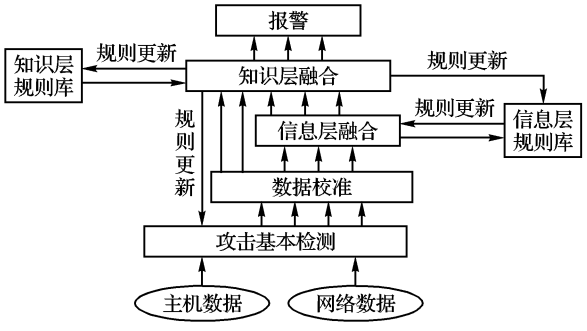

图 1 系统层次模型

1.1 攻击的基本检测层

在基本检测层布置多种基本检测器,每种检测器采用不同的检测方法(可以是一种检测算法,如支撑向量基,或是一个独立的检测系统,如 snort)对采集到的系统信息进行检测。并将最终检测结果传送给上层数据融合模块进行融合。这样做虽然每种基本检测方法只能在比较低的层次对系统的局部安全性作出判断,但是因为每个基本检测器的关注点不同,如:有的运用的是误用检测而另一些采用的是异常检测,或者,一些检测器采用的是基于主机的检测而另一些采用的是基于网络的检测,所以将所有的检测信息综合起来,可以为上层融合模块提供足够全面的系统安全信息。

收稿日期:2008 - 04 - 21;**修回日期**:2008 - 06 - 30。
作者简介:赵晓峰(1978 -),男,浙江绍兴人,讲师,硕士,主要研究方向:计算机网络、信息安全。

1.2 数据校准层

因为系统中采用了多种检测方法,而不同的检测方法的输出格式可能不同,如:对于误用检测给出的是某种入侵发生的概率,但是对于异常检测给出的结果却是目前系统存在某种异常状态的概率,所以,要将这些不同的输出统一成融合模块能够处理的格式。另外,这一层还应包括一定的过滤功能,以去除那些不完整或不能被融合模块处理的信息。

1.3 信息层融合

不同的基本检测方法可能会从不同角度对某一入侵事件作出不同的判断,如:对于一次U2R攻击,基于状态转移的异常检测会因为用户的非法状态转移而报警,但是他只知道系统目前出现了异常状态,而不知道这个事件是由什么攻击行为造成的;基于规则的误用检测可能会正确的判断出这是一次U2R攻击;基于SVM的分类检测引擎可能会因为该入侵的特征不够明显而将其误认为是R2L入侵;而基于网络的检测模块根本就检测不到这样事件的发生,所以不会产生报警。信息层融合的最主要任务就是对这些不同IDS对同一事件的判断结果作出合理而有效的评估,从而给出正确的决策。

1.4 知识层融合

虽然信息层融合大大减少了原始报警数据的数量,并且决策结果的正确性较原始决策有了很大提高,从而能够正确的反映出目前系统遭受攻击的情况,但是其输出的数据量对于系统管理员来说还是太多,而这层给出的决策还是一个在比较低层面上对系统安全状况的反映。所以,还需要对其结果再进一步的提炼,以便能从更高的层面上对系统的安全态势和入侵者的入侵场景有一个更直观更全面的了解。所以,引入了知识层融合来对信息层融合的结果进行进一步融合处理,以获取当前系统的全局安全态势。

2 本文采用的信息层融合方法

整个系统最关键的部分是信息层融合,这一层融合方法的好坏直接影响到整个系统的检测效果。目前已经提出了多种融合底层检测结果的方法[2-4],其中主要包括:投票表决法、朴素贝叶斯决策方法、人工神经网络方法等。

D-S证据理论[5]已经被广泛应用于数据融合的各个领域,如:专家咨询系统、预测、医学、图像处理、决策分析、人工智能、故障诊断、指标体系、军事、地质测控、识别分类等。但是,目前将证据理论应用于入侵检测系统的研究还很少,显然在入侵检测研究中对于证据理论的重视度还不够。所以本文将D-S证据理论引入信息层融合方法中,并希望以此引起入侵检测研究领域中对证据理论的重视。

2.1 Dempster-Shafer 证据理论

假设有一个判决问题,对于该问题所能认识到的所有可能结果的集合用 Θ 表示。

定义1 设 Θ 为识别框架,Θ 上的基本置信指派(Basic Belief Assignment,BBA)定义为 $m:2^{\Theta}\rightarrow[0,1]$ 满足以下条件:

$$\sum\{m(A)\mid A\subseteq\Theta\}=1,\quad m(\varnothing)=0 \tag{1}$$

对于任意 $A\subseteq\Theta$,$m(A)$ 称为命题 A 的基本概率指派。$m(A)$ 表示指派给 A 本身的置信测度,即支持命题 A 本身发生,而不支持 A 的任何真子集的程度。$m(A)$ 也称为假设质量函数或mass函数。

定义2 设 Θ 为识别框架,Θ 上由基本置信指派函数导出的置信函数(Belief Function)定义为 $Bel:2^{\Theta}\rightarrow[0,1]$ 且:

$$Bel(A)=\sum_{B\subseteq A}m(B) \tag{2}$$

$Bel(A)$ 表示给予命题 A 的全部置信程度,亦即 A 中全部子集对应的基本置信值之和。

定理1 设 Θ 为识别框架,集函数 $Bel:2^{\Theta}\rightarrow[0,1]$ 是置信函数当且仅当它满足:

1) $Bel(\varnothing)=0$;

2) $Bel(\Theta)=1$;

3) 任意自然数 $n,A_1,A_2,\cdots,A_n\subseteq\Theta$。

$$Bel(A_1\cup A_2\cup\cdots\cup A_n)\geqslant\sum_i Bel(A_i)-\sum_{i>j}Bel(A_i\cap A_j)+\cdots+(-1)^n Bel(A_1\cup A_2\cup\cdots\cup A_n) \tag{3}$$

定理2 设 Bel_1 和 Bel_2 是同一识别框架 Θ 上的两个不同置信函数,他们的基本置信指派是 m_1 和 m_2,焦元分别为 $A_1,A_2,\cdots,A_i$ 和 $B_1,\cdots,B_j$,如果

$$\sum_{A_i\cap B_j=A}m_1(A_i)m_2(B_j)<1 \tag{4}$$

那么,函数 $m:2^{\Theta}\rightarrow[0,1]$ 对于所有的非空集合 $A\subseteq\Theta$ 满足 $m(\varnothing)=0$ 且:

$$m(A)=\frac{\sum_{A_i\cap B_j=A}m_1(A_i)m_2(B_j)}{1-\sum_{A_i\cap B_j=\varnothing}m_1(A_i)m_2(B_j)}=\frac{1}{N}\sum_{A_i\cap B_j=A}m_1(A_i)m_2(B_j) \tag{5}$$

$A\neq\varnothing,m(A)=0,A=\varnothing$

其中 $N=1-\sum_{A_i\cap B_j=A}m_1(A_i)m_2(B_j)>0$。

2.2 信息层融合的输入形式

对应于不同的基本检测器,信息层融合可能包括以下几种输入。

1)0—1输入。基本检测器只能给出某事件是或不是某种入侵,基于规则的检测方法给出的大多都是这样的结果。对于数据融合来说这种输入最粗糙,其丢失的信息也最多。我们称之为第一类输入。

2)单值概率输入。基本检测器认为某事件是某种入侵(或异常),并且给出是该入侵的概率,基于状态转移的检测方法给出的大多是这样的结果。对于数据融合来说这种输入丢失了基本决策的某些信息,但是相对于第一种输入它保留的信息更多一些。我们称之为第二类输入。

3)多值概率输入。基本检测器并没有对某事件给出其最终决策,而是将该事件的所有可能分类以概率的形式给出。比如:某检测器给出某事件的所有可能分类为:neptune,smurf,pod,selfping,normal,unknown,其相应的概率为:30%,50%,5%,5%,1%,9%。这种输入最全面的保留了基本检测器的决策信息。我们称之为第三类输入。

我们采用D-S证据理论对基本检测的结果进行进一步的融合推理从而给出最终的决策。

假设在某一检测周期内对某IP有 M 个基本检测器对其进行检测,其中输出类型为第一、二、三类的检测器的数目为 m_1,m_2,m_3 个。在 M 个基本检测器中有 N 个产生了检测输出,即在融合层对于该IP有 N 个输入。其中属于第一、二、三类的输入数分别是 n_1,n_2,n_3 个($n_1+n_2+n_3=N$)。设所有产生输

出的基本检测器发现的所有入侵事件为 $A_1,\cdots,A_s,A_s$ 对应于检测器未确认类型(UnIdentified)的事件输出。

所以我们确定的识别框架 Θ 为($A_1,A_2,\cdots,A_s$,Normal)。

对于第一类输入,有 $n_{1i}(\sum_{1\leqslant i\leqslant S} n_{1i} = n_1)$ 个检测器确认其为 A_i 类,则该事件属于第 A_i 类的基本置信指派为 $m1(A_i) = \frac{n_{1i}}{m_1}$,而事件属于 Normal 的基本置信指派为:

$$m1(\text{Normal}) = \frac{m_1 - n_1}{m_1} \tag{6}$$

显然,第一类输入满足定理 1,所以其集函数 *Bel* 是置信函数。

对于第二类输入,假设第 j 个检测器以概率 P_{ji} 认为该事件为 A_i,则设 $\bar{A_i} = \Theta - A_i$,则基本置信指派 $mj(A_i) = P_{ji}$,$mj(\bar{A_i}) = 1 - P_{ji}$,$2 \leqslant j \leqslant n_2 + 1$。显然,第二类输入满足定理 1,所以其集函数 *Bel* 是置信函数。

对于第三类输入,设第 k 个检测器输出的概率分布为:

$A_1,A_2,\cdots,A_s$,Normal

$P_{k1},P_{k2},\cdots,P_{ks},P_{\text{kNormal}}$

其对应的基本置信指派为 $mk(A_i) = P_{3i}$,$n_2 + 2 \leqslant k \leqslant n_2 + 1 + n_3$。显然,第三类输入满足定理 1,所以其集函数 Bel 是置信函数。

综合三类输出的基本置信指派可以得到用于融合的输入矩阵。

	A_1	A_2	…	A_j	…	A_s	$\overline{A_j}$	Normal
$m1$	$m1(A_1)$	$m1(A_2)$	…	$m1(A_j)$	…	$m1(A_s)$		$m1$(Normal)
$m2$			…	$m2(A_j)$	…		$m2(\overline{A_j})$	
⋮	⋮	⋮		⋮		⋮	⋮	⋮
mi		$mi(A_2)$	…		…		$mi(\overline{A_j})$	
$m(i+1)$	$m(i+1)(A_1)$	$m(i+1)(A_2)$	…	$m(i+1)(A_j)$	…	$m(i+1)(A_s)$		$m(i+1)$(Normal)
⋮	⋮	⋮		⋮		⋮	⋮	⋮
mn	$mn(A_1)$	$mn(A_2)$	…	$mn(A_j)$	…	$mn(A_s)$		mn(Normal)

其中:$i = n_2 + 1$,$n = n_2 + n_3 + 1$,m_2 到 m_i 的取值视具体情况而定。

对于输入矩阵的各行采用 Dempster-Shafer 合成公式,将其合并成新的证据体。根据新证据体作出最终决策,决策的不确定性可以由证据体的不确定区间给出。

2.3 入侵互斥问题分析

Shafer 指出[5]各个假设相互排斥,并且完整地描述了问题的所有可能。我们要求,在系统进行入侵的分类时,要尽可能地按入侵方法将入侵进行细分。这样会导致整个系统中定义的总入侵数量大大增加。但是,因为我们将识别框架定义为:某一检测周期内,基本检测层所有不同检测方法产生的报警总集。而相对于系统定义的全部入侵种类,在某一检测周期内(如:30 s)入侵的种类和数量要少很多,所以整个识别框架中假设的数量不会太大,这就保证了融合的效率。以这样的方法,我们只能够发现系统中最基本的底层入侵。而一次完整的入侵过程要由多个基本入侵共同组成。为了发现更加完整的入侵场景,我们在知识层融合时要对信息层的报警进行入侵场景再现的处理。

3 实验及分析

实验采用的是 KDD CUP'99[6]提供的数据集,该数据集包含 500 万条连接记录,每条连接包括 41 个属性。共 4 大类入侵:DoS,Probing,R2L,U2R。

基本检测算法采用:C4.5,Naïve Bayes,三层神经网络,MDT[7],KNN(K =10),KNN(K =5),SVM。

实验中融合算法采用 Murphy 平均法[8]进行。研究表明这种方法可以更好地处理证据的冲突问题。

3.1 算法融合效果测试

为了测试 D-S 算法的融合效果,共进行了三组实验。每一组实验分别从 KDD CUP'99 提供的训练数据集中随机抽取 30000、20000、10000 条对每个基础检测器进行训练。从测试数据集中随机抽取与训练数据不同的 30000 条数据作为三组实验的测试数据(在所有的实验中都使用这 30000 条数据作为测试数据)。每组实验先用基础检测器对测试数据进行检测,然后,采用本文提出的融合方法对各基础检测器对于测试数据的检测结果进行融合。以下是三组实验的检测结果对照表,表中数据为每种方法对于不同入侵(DoS、Probe、U2R、R2L)和正常数据(Normal)的检测正确率。

表 1　D-S 算法与基本检测算法结果对照表 1(第一组)%

检测算法	Normal	DoS	Probe	U2R	R2L
C4.5	98.9	93.5	74.7	40.0	7.0
Bayes	96.7	73.9	89.1	40.0	24.1
神经网络	99.2	91.5	85.6	20.0	7.0
MDT	86.3	89.4	42.8	20.0	39.0
D-S	99.1	93.7	90.0	20.0	10.3

表 2　D-S 算法与基本检测算法结果对照表 2(第二组)%

检测算法	Normal	DoS	Probe	U2R	R2L
C4.5	97.8	91.6	55.0	37.4	12.5
Bayes	96.7	73.8	81.2	20.5	24.4
神经网络	99.5	74.1	59.0	20.0	8.0
MDT	85.7	89.8	43.7	20.0	27.6
D-S	98.9	90.7	69.7	20.0	10.3

在第一组实验中各基本检测器经过了充分训练。其中,对于 DoS 入侵,C4.5 具有最好的检测效果,但是 Naive Bayes、神经网络和 MDT 的检测效果都不理想,而经过融合后的检测率也更加接近于 C4.5 的结果。对于 Probe 入侵,四种基础检测器的检测结果都没有超过 90%,而经 D-S 算法融合后的结果超过了 90%。对于后两种入侵因为在训练数据中出现的

概率很小,比如R2L入侵在KDD CUP'99提供的490 000条训练数据中只出现了几百次,而为了保证训练结果的普遍性我们采用随机的方式选取训练数据,所以在训练数据中后两种入侵出现的频率很低,这就使所有检测器对他们的检测效果都很不理想,并且检测结果的随机性很强。

表3 D-S算法与基本检测算法结果对照表3(第三组)%

检测算法	Normal	DoS	Probe	U2R	R2L
C4.5	90.8	87.5	36.7	21.0	14.6
Bayes	89.5	66.6	9.6	15.0	13.1
神经网络	88.1	92.6	0.0	15.6	21.9
MDT	75.9	73.2	43.7	29.7	8.5
D-S	93.2	88.7	39.2	21.0	11.0

可以看出每种基础检测器都对某(几)种入侵具有较高的检出率,但是对其他入侵的检测效果并不理想,比如C4.5算法对Normal和DOS入侵具有较好的检测正确率但是对于Probe入侵的检出率却很低。而经过融合的结果都超过或接近各基础检测器对每个分类的最好检出率值。这使得系统对所有入侵都具有相对较高的检出率,从而弥补了单一检测算法无法对所有入侵都有很好检测效果的缺陷。另外,三组实验经过融合后的Normal类的检测正确率都接近或超过各基础检测器正确率的最高值,尤其是第三组实验,各基础检测器的检测效果都不理想,但是经过融合后的检测正确率却达到了93.2%。也就是说经过融合,系统具有较低的误报率。

当少数基础分类器的分类结果不理想时,不会对融合结果造成很大影响(如:第一次实验时Bayes方法对于Dos入侵的检出率较低,但是融合后的检出率却超过了所有基础检测器的检出率),这就使整个检测系统更加稳定,少数检测器的损坏和误检不会对整个系统的总体检测效果造成很大影响。但是其不足之处是如果大多数基础检测器的检测结果都不理想,也很难获得较好的融合结果(比如,第三次实验的Probe入侵检测结果)。总之,本文提出的方法能够相互弥补各基础检测器的缺陷,并且检测的结果也更加稳定。但是,融合结果还是要依赖于基本检测器的好坏,当系统中存在大量低检测率的检测器时,即使采用融合算法也很难得到满意的结果。这就要求在实际应用时,对于每种入侵都有多数检测器对其具有较高的检出率。

因为系统融合采用的输入矩阵考虑到了基础检测器可能产生的所有输出形式,所以新的检测方法可以直接加入到系统中,而不需对融合层进行任何修改,这就使系统具有较高的可扩充性。

3.2 节点数对融合结果的影响

为了测试基本检测方法的数量对融合结果的影响,设计了6种基本检测算法:C4.5,Naïve Bayes,三层神经网络,MDT,KNN($K=10$),KNN($K=5$),SVM。

共进行了5次实验,分别选取2,3,4,5,6种算法(每次增加一个基本检测器)作为基本检测器进行融合(训练和测试集与实验1一致)。实验结果如图2所示。从图2中可以看出,当只采用两个基本检测器时融合效果并不理想,随着检测器的增加融合效果逐渐增加,其中对于Probe入侵的第三次实验因为,增加的MDT算法对Probe入侵的检测效果并不理想,所以融合后的检出率并没有增加,这就又证明了融合对基本检测器的依赖性。

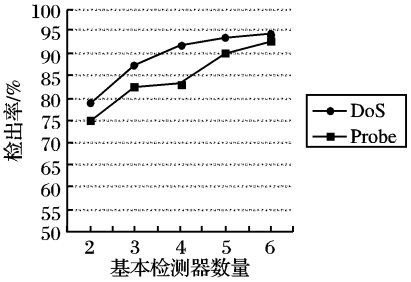

图2 基本检测器数量对融合的影响

10 000次入侵的平均融合时间为:2节点0.02 s,3节点0.03 s,4节点0.04 s,5节点0.06 s,6节点0.09 s。但是,因为采用模拟实验入侵种类只有4种(和Normal),对时间性能的测试并不具有实际意义,而在实际情况(入侵分类更多)下,融合的时间性能还有待进一步证明。

4 结语

本文提出了一种基于数据融合的入侵检测模型,将入侵检测的问题转化成为一个将系统信息从低级到高级逐层抽象的过程,指出整个检测系统应该充分融合和利用底层不同基本检测方法的检测信息,以得到事件的更加准确的判定。并且将证据理论引入到模型的信息层融合之中。

实验结果证明:通过融合能够使整个系统对所有入侵的检出率超过或接近各基础检测器对每个分类的最好检出率值。这使得系统对所有入侵都具有相对较高的检测率,从而弥补了单一检测算法无法对所有入侵都有很好检测效果的缺陷。但是,通过分析还发现,现有的D-S融合算法,对基本检测器的检测效果存在一定程度的依赖,而寻找一种更加稳定的D-S融合方法也是我们下一步的主要工作方向。

参考文献:

[1] 何友,王国宏;多传感器信息融合及应用[M].北京:电子工业出版社,2000:11.

[2] KLEIN LA. A boolean algebra approach to multiple sensor voting fusion[J]. IEEE transactions on aerospace and electronic systems, 2004,29(1):317-327.

[3] CHAN A P F, NG W W Y, YEUNG D S, *et al*. Multiple classifier system with feature grouping for intrusion detection: Mutual information approach[C]// Proceeding of the 9th International Conference on Knowledge-Based Intelligent Information & Engineering Systems. Melbourne, Australia:[s.n.], 2005:215-221.

[4] NG W W Y, CHAN A P F, YEUNG D S, *et al*. Quantitative study on the generalization error of multiple classifier systems[C]// Proceeding of International Conference on Systems, Man and Cybernetics. Hawaii, USA: IEEE Press, 2005:405-416.

[5] SHAFER G. A mathematical theory of evidence[M]. Princeton: Princeton University Press, 1976.

[6] KDD Cup 1999 Data[EB/OL].[2008-01-01]. http://www.ics.uci.edu/~kdd/databases/kddcup99/kddcup99.html.

[7] 赵晓峰,叶震. 基于多随机决策树的入侵检测系统[J].计算机应用,2007,27(5):1041-1043.

[8] GRUNDEL D, MURPHEY R, PARALOS P. Theory and algorithms for cooperative systems[M]. Singapore: World Scientific, 2005: 239-310.