一种基于数据融合的分布式入侵检测系统

郭文普,孙继银,任 俊
(第二炮兵工程学院,陕西 西安 710025)

摘 要:入侵检测是网络安全的一种重要手段,为提高入侵检测的准确性,文中提出了一种基于数据融合的分布式入侵检测系统,并详细论述了该系统的网络设计、包采集分析、局部判决、融合中心事件关联和数据融合等各个环节的具体设计与实现方法,分析表明文中提出的事件关联规则和加权表决法决策融合算法对分布式入侵检测系统是十分有效的。

关键词:入侵检测;数据融合;事件关联;加权表决法决策
中图分类号:TP393.08 **文献标识码**:A **文章编号**:1005-3751(2006)02-0217-03

A Kind of Distributed IDS Based on Data Fusion

GUO Wen-pu, SUN Ji-yin, REN Jun
(The Second Artillery Engineering College, Xi'an 710025, China)

Abstract: Intrusion detection is an important means of network security. This paper presents a kind of distributed IDS based on data fusion to improve the veracity of intrusion detection. It also discusses the design and realization of each step of this kind IDS in detail, including the net-designing, packet collecting and analyzing, local judging, event relating, data fusion in fusion center. What this paper put forward such as the regulation of event relating and the decision arithmetic of voting with weight in data fusion is useful.

Key words: intrusion detection; data fusion; event relating; decision arithmetic of voting with weight

0 引 言

入侵检测(Intrusion Detection)是网络安全的一个重要内容,是对入侵行为的发觉。它通过从计算机网络或计算机系统的关键点收集信息并进行分析,从中发现网络或系统中是否有违反安全策略的行为和被攻击的迹象。从1980年Anderson提出入侵检测概念、1987年Denning提出通用的入侵检测模型到目前一些常见的商用入侵检测系统,入侵检测技术的发展已经走过了20多个年头。目前入侵检测系统可分为基于主机的入侵检测系统和基于网络的入侵检测系统,而常用的入侵检测系统的检测方法有基于特征的检测和基于异常的检测。不管哪种检测方法,衡量其性能的指标都是准确率、误报率和发现新入侵行为的能力。

近几年随着网络带宽、网络用户的不断增长,网络攻击日益频繁,对入侵检测技术提出了更高的要求,新的检测方法逐渐被引入入侵检测系统中[1~3],而最为引人注意的就是基于数据融合的分布式入侵检测[4,5]。数据融合(Data Fusion)是将多个相异传感器得到的数据进行复合,以便对事件、活动、事态做出更为准确的推论的一种相对较新的技术。为了进行入侵检测信息的数据融合需要解决系统设计、信息表示和信息融合3个方面问题。

收稿日期:2005-05-03
作者简介:郭文普(1976—),男,河北河间人,博士研究生,研究方向为计算机网络安全;孙继银,教授,博士生导师,研究方向为作战指挥自动化。

1 基于数据融合的分布式入侵检测系统的设计

1.1 对网络攻击行为的分类

从本质上讲,入侵检测是一个分类问题,它是将网络中正常使用的数据流与攻击行为产生的数据流区分开来,将用户对计算机和网络设备的正常操作行为与黑客的攻击活动区分开来。无论网络还是主机上的活动可分为正常、异常和可疑这3种情况,而异常与可疑又可按照攻击的类型进行分类。目前对网络攻击的分类方法很多,本系统结合网络攻击的特征和攻击后果,将攻击行为分为:DDOS(Distributed denial of service,分布式拒绝服务攻击),Worm(蠕虫),Probe(刺探)和Control(获取权限)4类。

1.2 分布式入侵检测系统的网络设计

在图1所示的分布式入侵检测系统中,为了对入侵行为进行更准确地检测与识别,需要在网络的不同位置设置入侵检测探测器,这些位置包括防火墙外、防火墙内、受保护子网内、受保护主机上等。每个探测器根据统一的规则库对探测的数据包做出局部判决。在网络中设置一台主机为入侵检测融合中心(图1中的控制台),融合中心负责对入侵检测探测器传送来的局部判决信息做事件关联和信息融合,从而做出对入侵行为的判定。其逻辑连接如图2所示。

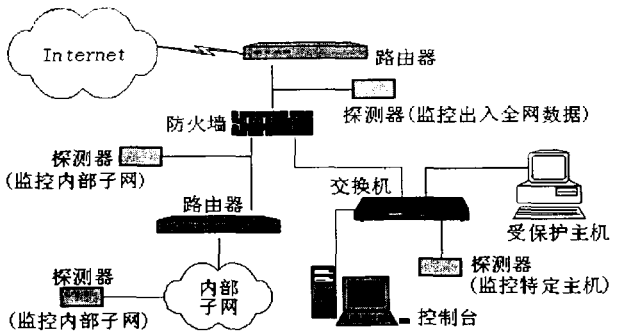

图 1 分布式入侵检测系统的网络设计

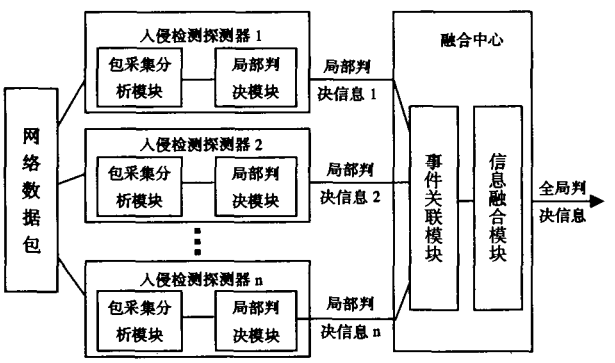

图 2 分布式入侵检测与决策融合

2 分布式入侵检测系统探测器的局部判决

2.1 探测器对数据包的采集分析

每个入侵检测探测器都必须完成对网络中数据包的采集和分析,这部分功能由包采集分析模块完成。首先,将探测器所在主机的网卡设成"杂听"模式,这样可以接收网络中而不是仅发到本机的数据包;其次,包采集分析模块完成对收到的数据帧的链路层解码和网络层解码,根据包头分析该数据包是建立连接过程的包,还是正常数据传送过程的包,为了减轻探测器和融合中心的负担,局部决策只对建立连接过程的包进行判决,而忽略正常数据传送过程的包;最后,包采集分析模块对包的数据内容进行分析,提取其特征字串,并将分析后的包传送到探测器判决模块。

2.2 探测器局部判决

由于网络攻击手段的处理,同一个数据包被处于网络中不同位置的探测器探测到的内容可能会发生变化,这样不同位置的探测器在参考同一个判别规则库进行判决时,其局部判决信息是不同的。为了对不同位置探测器的局部判决信息进行融合,必须为所有探测器设计统一的判决信息的传输格式。考虑到事件关联和信息融合两方面的要求,将每个探测器的局部判决信息输出格式定义为如图 3 所示的判决帧。

2	4	4	4	16	16	Byte
探测器编号	源 IP	目的 IP	时戳	对包分析得到折特征字串	μ_{DDoS},μ_{Worm},μ_{Probe},$\mu_{Control}$	

图 3 探测器局部判决信息输出的判决帧格式

其中,源 IP、目的 IP 分别表示探测器探测的数据包的源 IP 地址和目的 IP 地址,时戳表示该包到达网络的时间信息(注意,不是该包到达探测器的时间信息),μ_{DDoS},μ_{Worm},μ_{Probe},$\mu_{Control}$分别表示经探测器采用模式匹配得出的该包对于 4 种攻击形式的隶属度,即探测器的局部决策。

3 分布式入侵检测的信息融合

3.1 融合中心的事件关联

考虑到多个探测器送到融合中心的数据是一系列的局部判决信息流,融合中心首先要通过事件关联模块将这些局部判决信息流划分成一个个独立的事件后才能进行融合。划分事件应依据四个规则:

规则一:要将同一传感器对同一攻击行为引起的多个不同时序的数据包的局部判决信息流划分成一个事件;

规则二:要将来自不同传感器对同一数据包的局部判决信息划分成一个事件;

规则三:在前两个规则的基础上将这两类事件合并为同一事件;同时考虑到同一网络攻击行为有一定的时间上的连续性和相似性,从而探测器传送到融合中心的判决帧有一定的重复性,为简化事件关联和融合,提出规则四;

规则四:若一事件内的局部判决信息帧达到 k 个,则此事件不再接收新的判决信息帧。

依据上述四个规则,针对探测器传送到融合中心的各个判决帧,融合中心的事件关联可依据如图 4、图 5 所示流程实现(注:要求同时存在的事件数最多为 n 个)。

3.2 融合中心的信息融合

经过事件关联模块的关联后,每个事件由 k 个不同探测器对同一网络行为的局部判决帧组成,此时这些局部判决帧是对该事件对应的网络行为可能是哪种攻击的一个局部判决,通过 μ_{DDoS},μ_{Worm},μ_{Probe},$\mu_{Control}$表示其隶属某种攻击形式的可能性。融合中心必须经过信息融合后才能得出更准确的判决。针对网络攻击特征和每个探测器因位置不同所起的作用,融合中心采用加权的表决法进行融合。

该算法详细描述如下:

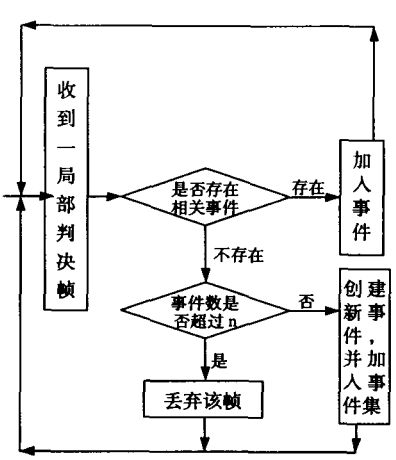

图 4 融合中心收到局部判决帧的处理

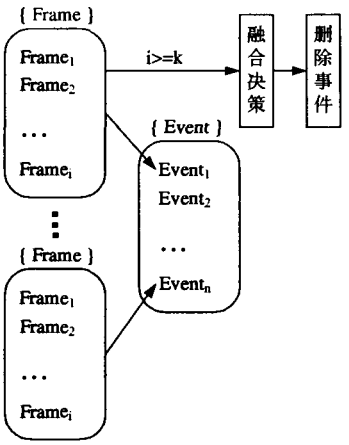

图 5　融合中心事件集及其处理

首先将入侵检测探测器按所处位置分为三类:一是被保护主机入侵检测探测器;二是某受保护子网入侵检测探测器;三是其它入侵检测探测器。并设每个探测器参预决策的权重为 m_i,该事件内不同探测器的总数为 N。

算法第一步:为每个入侵检测探测器进行参预决策的权重分配。

如果该事件中的目的 IP 地址与受保护主机的目的 IP 地址相同,则 $m_{主机探测器} = 0.6$,其余入侵检测探测器 $m_i = 0.4/(N-1)$;

如果该事件中的目的 IP 地址为某一受保护子网的 IP 地址,则 $m_{受保护子网} = 0.5$,其余入侵检测探测器 $m_i = 0.5/(N-1)$;

其它情况,则每个入侵检测探测器 $m_i = 1/N$。

算法第二步:加权表决法融合。

根据分配好的决策权重和局部判决进行如下计算:

$$\mu_{DDoS} = \sum m_i * \mu_{iDDoS} \quad \mu_{worm} = \sum m_i * \mu_{iworm}$$

$$\mu_{probe} = \sum m_i * \mu_{iprobe} \quad \mu_{control} = \sum m_i * \mu_{icontrol}$$

算法第三步:决策。

若融合后的 $\mu_{DDoS}, \mu_{Worm}, \mu_{Probe}, \mu_{Control}$ 的最大值小于门限值 $\mu_{阈值}$,则将此事件判断为正常,否则将此事件判断为该最大值对应的攻击行为。

4　结束语

文中提出了一种分布式入侵检测系统的设计与实现,提出的局部判决、事件关联和融合算法均具有较强的可操作性。文中未能讨论入侵检测响应问题,读者可参考文献[2]。

参考文献:

[1] de Boer R C. A Generic Architecture for Fusion - Based Intrusion Detection Systems[D]. [s.l.]:Erasmus University, Rotterdam School of Economics,2002.

[2] 郭代飞,杨义先,胡正名.基于大规模网络的自适应入侵响应模型研究[J].北京邮电大学学报,2004,27(1):79 - 83.

[3] 王　勇,王行愚,张瑞霞.基于 D - S 证据理论的分布式入侵检测方法研究[J].计算机工程与应用,2004(13):167 - 169.

[4] 吕慧勤,褚永刚,吕硕望,等.入侵检测系统融合机制研究[J].网络安全技术与应用,2003(12):10 - 14.

[5] 罗光春,卢显良,张　骏,等.基于多传感器数据融合的入侵检测机制[J].电子科技大学学报,2004,33(1):71 - 74.

(上接第 216 页)

```
exec xp_cmdshell'net user userB password /add'
```

这样数据服务器 A 就增加了预报员 B 用户,预报员 B 登录数据库服务器 A 后,在查询分析器中输入:

```
use master
go
exec xp_cmdshell'net localgroup administrators userB /add'
```

后再执行下面的 SQL 语句:

```
exec xp_cmdshell'net localgroup administrators'
```

发现 userB 已经在 administrators 组中,意味着预报员 B 已经拥有系统管理员的权限了,可以执行:

```
use master
go
exec sp_dropextendedproc'xp_cmdshell'
```

这个 SQL 语句是用来去掉 xp_cmdshell 存储过程,xp_cmdshell 存储过程是进入操作系统的最佳捷径,是数据库留给操作系统的一个大后门。另外,数据库服务器 A 的管理员可以在帐号管理中把系统帐号"BUILTIN \ Administrators"删除,不过这样的后果是一旦 sa 帐号忘记密码的话,就没有办法恢复了[5]。

4　结束语

尽管可以利用 SQL Server 气象资料数据库提供的许多工具来管理和配置数据库的安全性能,但在进行数据库的活动时仍旧可能忽略安全问题。进行数据库活动的过程中暴露出安全漏洞的可能性在很大程度上依旧存在,文中所讨论的扩展存储过程问题,它涉及到提升权限的安全问题。加深对安全机制的理解,可以让人们能够发现那些隐藏的安全漏洞。

参考文献:

[1] 何培英,刘雯琳,徐发山 .Web 数据库操作技术[J]. 微机发展,2003,13(2):80 - 82.

[2] 樊志平 . 数据安全性的实现方法[J]. 微机发展,2003,13(12):53 - 54.

[3] 罗远模,王　珊 .SQL Server 数据库系统基础[M]. 北京:高等教育出版社,2002.

[4] Nielsen P. Microsoft SQL Server 2000 宝典[M]. 刘　瑞,陈微,等译 . 北京:中国铁道出版社,2004.

[5] 周立柱,冯建华,孟小峰,等. SQL Server 数据库原理——设计与实现[M]. 北京:清华大学出版社,2004.