

# 基于移动代理的分布式入侵检测系统的研究

李 生, 邓一贵, 唐学文, 潘 磊, 林玉香

(重庆大学 计算机学院, 重庆 400044)

**摘 要:**目前的分布式入侵检测系统存在单点失效、网络带宽占用高、扩展性不强等缺点。文中针对这些缺点提出了一种基于移动代理的分布式入侵检测系统模型。该系统运用移动代理作为各种入侵的独立检测实体到主机监控代理上收集数据并聚集、关联这些数据,这使得系统减小了数据传输量、分散了计算,同时也提高了系统的可扩展性。着重讨论了系统部件失效的检测和自动恢复机制,使系统更加稳定可靠。对实验结果分析表明这种结构很好地减少了带宽的占用,容易对新的攻击进行扩展,失效的代理也能自动准确地恢复。

**关键词:**入侵检测;移动代理;网络安全;分布式系统

**中图分类号:**TP393.08

**文献标识码:**A

**文章编号:**1673-629X(2009)09-0132-04

## Research of Mobile Agent - Based Distributed Intrusion Detection System

LI Sheng, DENG Yi-gui, TANG Xue-wen, PAN Lei, LIN Yu-xiang

(College of Computer Science and Engineering, Chongqing University, Chongqing 400044, China)

**Abstract:** Currently, distributed intrusion detection models had some impediments as single point failure, overload parts of network and non-scalable. In this paper, a mobile agent-based distributed system is proposed. The new model uses a set of software entities called mobile agents that can move from one node to another node within a network, and perform the task of aggregation and correlation of the intrusion related data that it receives from another set of software entities called the host monitor agent. It reduces network bandwidth usage by moving data analysis computation to the location of the intrusion data, and offers more flexibility. In addition, particularly discusses how to enhance the stability and reliability of this system by accomplishing self-recovery from internal failures. Based on analysis and experimental values, conclude that this system can reduce network bandwidth usage, add agents for new attack easily and recover the internal failures with itself.

**Key words:** intrusion detection; mobile agents; network security; distributed systems

## 0 引言

伴随着网络技术的快速发展,网络安全问题也日益突出,以防火墙技术、加解密、身份认证和访问控制技术<sup>[1]</sup>、认证授权等为主的被动防御技术越来越不能满足人们对网络安全的需求。而以入侵检测技术为主的主动保护技术应运而生。它通过从计算机网络或计算机系统内的若干关键点收集信息,并对其进行分析,从而发现网络或系统中是否有违反安全策略的行为和遭到攻击的迹象<sup>[2]</sup>。入侵检测作为一种积极主动的安全防护技术,提供了对内部攻击、外部攻击和误操作的实时保护,在网络系统受到危害之前拦截和快速响应

入侵。进行入侵检测的软件和硬件的组合就是入侵检测系统(intrusion detection system, IDS)<sup>[3]</sup>。

基于主机的入侵检测系统(HIDS)用于保护关键应用的服务器,实时监视可疑的连接、系统日志和审计记录。基于网络的入侵检测系统(NIDS)主要用于实时监视所监控网段传输的数据包。分布式入侵检测系统一般指的是部署于大规模网络环境下的入侵检测系统,任务是用来监视整个网络环境中的安全状态,检测单个主机无法检测到的分布式攻击,也是NIDS和HIDS的结合。

## 1 现有入侵检测系统及其面临的问题

现有的分布式入侵检测系统主要存在两种方式,它们都有自己的缺陷。一是分布收集数据集中检测<sup>[4]</sup>:每个目标系统的数据都传送给中心处理模块,由中心处理模块分析网络的数据,判断是否发生了入侵。

收稿日期:2008-12-18;修回日期:2009-03-02

**作者简介:**李 生(1983-),男,重庆人,硕士研究生,研究方向为信息网络安全技术;邓一贵,副教授,研究方向为移动代理技术;唐学文,高工,硕士生导师,研究方向为网络通信与网络安全。

这种方式的缺陷主要表现为:单一的中心处理模块所有数据无法适应网络规模的迅速扩张和网络速度的大幅度提高,可扩展性差;分布式程度不够,由于只采用了分布式的数据收集结构,而数据的分析,入侵的发现都采用单一主机完成,这势必会造成两个严重的安全隐患:单点失效和网络拥塞;系统缺乏灵活性,当增加新的功能模块时,整个系统需要升级。二是无控制中心协作式组件体系结构<sup>[5,6]</sup>,在大的网络中,每一个检测都可能要求大量的组件相互协作,这不利于检测规模的扩大。现有分布式入侵检测系统中失效恢复机制不完善,移动代理的安全性得不到保障。系统灾难恢复机制,如冗余、移动性、动态恢复等,很少在现行系统中采用,这使得系统健壮性得不到保证。

近年来,在入侵检测系统中引用移动代理的思想越来越受到人们的关注,移动代理是一个能在网络中自主地从一台主机迁移到另一台主机,并与其上面的代理或资源进行交互的程序。将之引入入侵检测系统,可以有效地弥补传统分布式入侵检测系统的不足,使系统具有智能性、灵活性、异步性、可扩展性等特点<sup>[7,8]</sup>。

针对以上问题,文中提出了一种基于移动代理分布式入侵检测模型,并对网络负载进行了详细分析,最后介绍了系统部件失效后的自动恢复机制。

## 2 系统模型结构描述

文中提出的基于移动代理的入侵检测系统(MADIDS)主要是为了降低网络负载,提高实时性、可靠性、可扩展性和平台无关性。MADIDS由系统管理器、用户界面、检测代理(IMA)、主机监控代理(HMA)等构成。其中系统管理器中包括系统控制器、报警代理(AA)、检测代理派遣器(MAD)和失效检测代理(FDA),如图1所示。

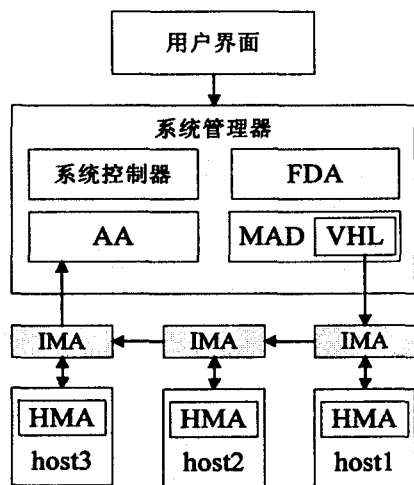


图1 MADIDS的体系结构

当主机监控代理检测到可疑行为时就产生一个可疑事件,如企图登录失败、可疑连接、端口扫描或者是可疑用户修改系统关键文件等等。监控代理把可疑事件封装成消息发送给检测代理派遣器,检测代理派遣器根据具体事件派遣相应检测代理完成检测任务。VHL(主机向量列表)是用来存储检测到有可疑行为主机地址的,每一个报告了可疑事件的主机地址都会被记录在VHL中。VHL中有多个列表,针对不同种类的可疑事件分别建立不同列表。具体检查代理会访问对应列表中的每一台主机的监控代理,从监控代理收集数据,然后聚集、关联从前面主机上收集来的数据,数据结果随检测代理一起移动到下一个监控代理继续上述过程。如此直到列表中的每个主机都访问完,如果检查代理检测到攻击便会产生一个警报,发送给报警代理。报警代理收到警报就通过用户界面显示给管理员。

MAD在派遣了检测代理之后需要间隔一段时间才能派发同种类型的检测代理,这可以避免MAD模块对同一任务派发两次代理。间隔时间可由管理员具体配置但至少是检测代理从MAD到AA的两倍时间,这样可以确保前次派发的检测代理访问完列表中的每台主机。下面详细描述MADIDS的每个模块。

### 2.1 主机监控代理

主机监控代理负责检测本机活动,一旦检测到有可疑行为就产生可疑事件,可疑事件以消息形式发送给MAD。每一个可疑事件都要包含攻击的类型,例如主机监控代理把错误登录密码当作可疑行为时就会产生一个需要检测doorknob-ratting攻击的可疑事件,在短时内同一目标有大量连接建立就会产生一个需要检查DoS攻击的可疑事件。

主机监控代理可以拥有多个子代理以监视不同种类的攻击,其主要任务是收集日志,网络数据,检查、匹配与攻击相关的数据,把攻击相关数据和其他数据分开并格式化检测代理所需要的数据格式。当出现新的攻击时可以派遣新的子代理到主机监控代理。

### 2.2 检测代理派遣器

检测代理派遣器根据接收到的来自主机监控代理的可疑事件决定派遣哪种检测代理。所有的检测代理都从这里产生。MAD中还包含一个VHL,它包含多个列表,每个列表存储了所有报告了同种可疑事件的主机IP地址。例如所有产生了doorknob ratting攻击可疑事件的主机IP地址都将存储于VHL中的一个列表中。MAD一旦收到主机监控代理的可疑事件消息就会把这个主机监控代理主机的IP地址加入到和攻击类型对应的列表之中。VHL为检测代理在网络中

提供了行程路线。

### 2.3 检测代理

检测代理的任务是到每个受攻击的主机上收集攻击证据,并对收集到的数据做分析,检测出是否存在攻击,存在攻击则生成警报消息给 AA 警报子系统,其访问主机由 VHL 决定。每个检测代理都用于检测特定的一种攻击,所以代理的代码不多,大小也较小。这使得本入侵检测系统能很容易地针对新的攻击加入新的检测代理,也很容易修改现有检测代理以增强其检测能力,增强了系统的灵活性。检测代理会从 VHL 列表中的每台主机的检测代理那里读取所需数据,并聚集、关联数据。聚集、关联数据是在代理移动到下一台之前就进行的,并且去除多余的数据。所以检测代理携带的数据甚少,对网络负载影响很小,而且对数据的处理是在每台主机上处理的,这就避免了统一处理带来的瓶颈,检测代理也不需要其他主机和其相互协作,即使规模再大也是自己独立就能完成,不会造成网络间大量数据相互传输,系统的可扩展性明显增强。在检测代理被派出之后, MAD 的失效不会中断检测代理的执行,即是检测代理会在 MAD 失效的情况下正常完成检测任务,提高了系统的可靠性。

### 2.4 警代理和系统控制器

报警代理接收检测代理发来的警报,并把它通过用户界面显示给管理员。报警代理还有一个功能就是阻止重复报警,报警器也存储来自检测代理的信息以供以后分析使用。

系统控制主要用于管理员对各个模块的管理以及配置,当某个模块失效的时候及时恢复该模块。

## 3 系统网络带宽利用分析

在一个集中检测模型的人侵检测系统中,总的网络带宽占用可表示为:

$$H * C \quad (1)$$

其中  $H$  表示网络中被监控主机数量,  $C$  表示一个主机原始数据量。

在本系统中,总的网络带宽占用是检测代理携带收集数据在网络主机间移动所消耗带宽的总和。假设  $S_m$  是检测代理被派遣出来时的初始大小,  $S_0$  是检测代理访问第一台主机后收集到攻击数据的初始大小,  $S_1$  是检测代理访问第二台主机后携带数据的大小。在第二台主机上,把检测代理携带来的数据和第二台主机的数据进行关联、聚集之后,检测代理携带的数据增量是  $S_{inc} = S_1 - S_0$ 。为了分析简便,这里假设检测代理访问了每台主机过后携带数据的增量是相同的,都

是  $S_{inc} \circ N$  是 VHL 中主机 IP 地址数量。那么,检测代理消耗的带宽可表示为:

$$N * (2 * (S_m + S_0) + (N - 1) * S_{inc}) / 2 \quad (2)$$

为了确定检测代理在时间范围内从 MAD 到 AA 的最后大小,笔者在多种攻击情况下做过多次实验。在到达第一台主机之前,由于检测代理还没有收集任何数据,所以初始大小  $S_m$  几乎可以忽略不计。代理在访问完所有主机后的最后大小取决于增量  $S_{inc}$ ,而  $S_{inc}$  的值取决于从当前主机审计数据收集到的攻击踪迹数据和在之前主机收集的数据的聚集程度。通常情况下,从主机收集来的数据一部分会被聚集,另一部分作为新的记录,所以  $S_{inc}$  的值不会太大。最好情况下,极大部分从主机收集到的数据都可以和前面收集的数据聚集,这种情况下,  $S_{inc}$  的值可以忽略。当然,在最坏情况下,从主机收集来的任何数据记录都不能和代理在之前收集的数据聚集,这时  $S_{inc}$  的值则会较大。图 2 显示了在最坏情况、通常情况、最好情况下检测代理大小和访问主机数量关系图。

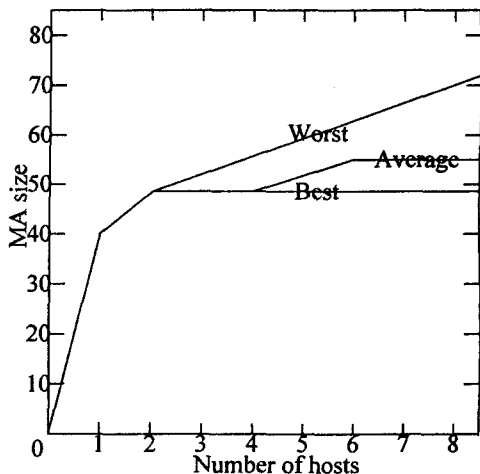


图 2 MA 大小(bytes)与主机数量关系图

由于大多数攻击常常是少部分主机受到侵害,所以检测代理要访问的主机数量并不多,因此可以认为  $N \ll H$ 。检测代理为检测特定入侵而收集的与攻击有关的数据只是原始数据中的很小一部分,所以即使是在最坏的情况下,增量  $S_{inc}$  也要比原始审计数据量  $C$  要小。

从上面的分析可以看出本系统对网络带宽的消耗要小于集中检测的分布式入侵检测系统。检测代理生命周期决定了系统入侵检测响应时间。检测代理的生命周期取决于访问主机的数量和聚集、关联从主机收集到数据的时间。当 VHL 很长时周期会有所增加,但是聚集、关联函数所处理的数据是从主机监控代理收集来的,它要比原始数据小。在集中检测的分布式入侵检测系统中,所要分析的数据是从全部主机收集来

的原始数据,这就使得响应时间受网络主机总数量和这些原始数据的影响而增加了响应时间。所以本系统的响应时间也要优于集中检测的入侵检测系统。

## 4 系统失效恢复机制

系统失效的部件主要包括检测代理派遣器、主机监控代理、报警代理、系统控制器以及检测代理。这些部件的失效就会导致系统无法正常运行,在这期间入侵就可能完全无法检测到,所以要尽快将失效的部件恢复正常。本系统采用静态失效检测代理和移动失效检测代理来检测系统部件的失效并在较短时间内恢复其功能。

### 4.1 系统管理器内各模块失效恢复

系统管理器主要包括报警代理和检测代理派遣器,针对这两个模块静态失效检测代理会周期( $T$ )地向他们发送 ACTIVE\_TEST 消息,当在规定时间内 $t$ 内没收到 ACTIVE\_TEST\_RSP 消息可再发消息,若 $N$ 次没收到消息则认为其失效了。静态失效检测代理立刻报告系统控制器重新启动或建立新的代理。如果系统控制器失效则由静态失效检测代理负责恢复。 $T, t, N$ 的值可自行决定。静态失效检测代理也有可能自身失效,所以要建立一个包含同样内容的备份静态失效检测代理,当静态失效代理更新内容时同样内容也要发给备份。备份代理和原代理相互检测,任何一方失效,另一个立即建立一个新的代理替换失效代理以保证正常工作。

### 4.2 主机监控代理失效恢复

本系统利用移动失效检测代理来查询各个主机上 HMA 的运行状态。静态失效检测代理周期性地发送一个移动失效检测代理(MFDA)使他按照一定路线在各主机间移动。每到一个主机便向本地 HMA 发送 ACTIVE\_TEST 消息,当 $t$ 时间没收到 ACTIVE\_TEST\_RSP 消息便再次发送消息收到消息或 $N$ 次后仍未收到消息后继续到下一个主机,最后把数据一起交给系统控制器。这里 ACTIVE\_TEST\_RSP 消息包含了一个列表,列出了当前 HMA 中所有子代理状态(运行、暂停)和 ID。系统控制器查看每个主机上收集的数据结果,若 MFAD 一直没收到某个 HMA 的 ACTIVE\_TEST\_RSP 消息则认为当前主机上的 HMA 失效了。这时系统控制器就销毁它并重新创建一个 HMA,将它发生到原来的主机上。若收到了 ACTIVE\_TEST\_REP 消息,则把消息和数据库中的配置消息进行比较,看是否和配置信息中列出的子代理的

ID 和状态一致。如果不一致则认为子代理失效,系统控制器就销毁该子代理,再重新创建一个子代理并发送到原 HMA 中。

如果某段时间内某个部件模块频繁失效,则控制器会通过用户界面报告给管理员。

### 4.3 检测代理失效恢复

检测代理在网络中移动,容易发生错误而失效,检测代理在检测完成后都会向 MAD 发送一个检测完毕的消息,如果 MAD 在指定时间内没收到检测代理检测完毕的消息,则立即创建一个一样的检测代理发送出去。尽管这样可能有两个同样的移动代理在网络中移动,但这并不会对系统有什么不良影响。

## 5 结束语

文中所述的基于移动代理的入侵检测系统运用主机监控代理在主机上监控网络可疑事件。针对每种攻击类型用不同的检测代理到网络中收集攻击证据,检测攻击,降低了网络负载。当出现新的攻击时,可以很容易地添加新的检测代理,具有很好的扩展性。同时失效恢复机制也使得系统具有良好的容错性和健壮性。然而,移动代理本身的安全问题也是很重要的,恶意代理可以入侵主机,主机也可破坏代理,因此入侵检测系统本身的安全是以后研究中需要解决的问题。

### 参考文献:

- [1] Stalls W, 计算机密码学与网络安全—原理与实践[M]. 第3版. 北京:电子工业出版社,2004:2-9.
- [2] 熊 焰,苗付友,张泽明. 一个基于移动代理的分布式入侵检测系统[J]. 小型微型计算机系统,2004,25(2):192-194.
- [3] 朱长棣,刘方爱. 基于 P2P 和移动代理的入侵检测系统研究[J]. 计算机技术与发展,2007,17(1):170-172.
- [4] 李守国,李 俊. 基于数据挖掘的入侵检测技术研究[J]. 计算机技术与发展,2006,16(4):218-220.
- [5] Jansen W, Mell P, Karygiannis T, et al. Applying Mobile agents to Intrusion Detection and Response[R]. NIST Interim report (IR)-6416, USA: National institute of Standards and Technology, 1999.
- [6] White G, Fisch E, Pooch U. Cooperating security managers: A peer-based intrusion detection system[J]. IEEE Network, 1994,10(1):20-23.
- [7] 危胜军. 模糊 petri 网知识表示方法在入侵检测中的应用[J]. 计算机工程,2005,31(2):130-132.
- [8] 罗光春,卢显良. MADIDS: 一种基于移动代理的新型分布式入侵检测系统[J]. 计算机科学,2003,30:69-71.