基于对手思维建模的分布式入侵检测模型*

陆 俊[1,2], 王崇骏[1,2], 王 珺[1,2], 陈世福[1,2]
(1. 南京大学 计算机软件新技术国家重点实验室, 江苏 南京 210093; 2. 南京大学 计算机科学与技术系, 江苏 南京 210093)

摘 要: 研究网络入侵和入侵检测系统的现状和发展趋势,将对手思维建模和意图识别技术引入入侵检测系统,提出了一个基于对手思维建模的入侵检测模型(IRAIDS),为解决大规模、分布式、智能化入侵提供了解决方法。

关键词: 思维建模; 入侵检测; 网络入侵
中图分类号: TP393 **文献标志码**: A **文章编号**: 1001-3695(2007)05-0115-04

Intrusion Detection Model Based on Intention Modeling

LU Jun[1,2], WANG Chong-jun[1,2], WANG Jun[1,2], CHEN Shi-fu[1,2]
(1. National Key Laboratory for Novel Software Technology, Nanjing University, Nanjing Jiangsu 210093, China; 2. Dept. of Computer Science & Technology, Nanjing University, Nanjing Jiangsu 210093, China)

Abstract: Through researching the past and resent network intrusion and intrusion detection system, intention modeling and intention recognition to intrusion detection system and propose an intrusion detection system model IRAIDS were introduced, which helps to solve distributed intelligent network intrusions.

Key words: intention modeling; intrusion detection; network intrusion

0 引言

伴随着网络技术的不断发展,网络安全已经成为一个至关重要的问题,也是计算机领域的研究热点之一。为了达到当场检测出恶意的网络入侵行为并马上采取防范反击措施的目的,实时监测黑客入侵行为并以程序自动产生响应的网络入侵检测系统(Intrusion Detection System, IDS)产生了。入侵检测被认为是防火墙之后的第二道安全闸门,可以在不影响网络性能的情况下能对网络进行监听,从而提供对内部攻击、外部攻击和误操作的实时保护,大大提高了网络的安全性[1]。

自从 1980 年 4 月 James P. Anderson 第一次详细阐述了入侵检测的概念以来,入侵检测系统经历了从集中式系统向分布式、智能化系统的发展历程。与此同时,高级入侵活动也变得越来越呈现出分布性和协调性的特点,具体表现在:

(1) 一次入侵可能分布在网络中的多个机器上[2]。

(2) 一次攻击可能只是一个更大规模入侵的一个部分,它只是使用当前被攻陷的网络作为跳板,最终的目标可能是攻击别的系统或非法得到其他资源[3]。

(3) 多次简单攻击可以组合成为一次更复杂的长时间协同入侵[4]。

这使得传统的基于分布式数据采集和集中分析的分布式入侵检测系统很难检测出大规模的分布式智能协作攻击,并且不能对这些攻击作出实时反应。然而,Agent 技术的发展为解决这些问题提供了契机。一些学者提出,由于 Agent 本身具有协同工作、智能化、自治性和移动性等特点,将其引入入侵检测系统可以弥补传统分布式入侵检测的不足,也为入侵检测技术提供了很多新的思路。文献[5]提出利用免疫学的思想设计出一个基于 Agent 的入侵检测模型。这个模型是层状结构,能动态学习、检测出已知和未知的入侵、检测出不同层次的入侵。文献[6]提出了一种采取无控制中心的多 Agent 结构,每个检测部件都是独立的检测单元,尽量降低各检测部件间的相关性。文献[7]提出了一种层次结构的基于自治 Agent 的入侵检测框架 AAFID。其中监视器是系统的单一失效点。尽管如此,入侵者出于自身利益的考虑会对检测 Agent 有意回避甚至对关键节点 Agent 主动攻击,而且协作入侵通常会故意采取一些行动以隐藏其意图或掩盖其行动轨迹,如:

(1) 能够检测到的入侵行动可能只是一部分,甚至检测到的入侵行动可能是误导的。

(2) 入侵者可能采用灵活的计划以同时完成多个目的。

(3) 入侵者也许多次重复一些步骤。

这些都极大地增大了入侵检测的难度,仅仅对检测系统自身采用 Agent 技术的分布式入侵系统很难达到实时检测复杂入侵的目标。在更高的层次上,笔者将入侵者和入侵检测方抽象为是以“局部运作、全局共享”为核心的多 Agent 系统,对抗双方都是一组自治的 Agent 通过协调它们的知识完成入侵和

收稿日期: 2006-03-19; 修返日期: 2006-05-08 基金项目: 国家自然科学基金资助项目(60503021);江苏省自然科学基金资助项目(BK2005075)

作者简介:陆俊,硕士研究生,主要研究方向为多 Agent 与计划识别(lujun@ai.nju.edu.cn);王崇骏,博士,主要研究方向为机器学习与分布式人工智能;王珺,博士研究生,主要研究方向为人工智能与信息安全;陈世福,博导,主要研究方向为人工智能.

入侵检测。在求解的过程中,各 Agent 之间达成了协作、协调、协商、理性、对抗、交互等各种关系。基于这种抽象以及入侵检测的本质(根据入侵者的行动及时推断出入侵者的意图),本文提出了以对手思维建模和对手意图识别技术为基础的多 Agent分布式入侵检测系统模型 IRAIDS,为解决大规模、分布式、智能化入侵提供了解决方法。

1 IRAIDS 模型

1.1 对手思维建模

对手思维模型的核心在于意图的识别,因为意图对应于实际的行为规划,这也正是入侵检测的目标。本文提到的意图识别包含两个层次的含义:

(1) 单个对手意图的识别

在该模型中使用 TA(Tracer Agent) 针对单个对手的思维状态建模,目标是通过分析单个对手的行动序列推测其可能的入侵行为。具体请见 TA 部分。

(2) 对手群体意图的识别

单个对手的思维建模只能检测出简单的个体入侵意图,对于大规模的分布式网络入侵就无能为力了。为了解决这个问题,笔者在模型中提出通过检测 Agent 之间的协作方式,分析群体对手的意图以找出其入侵计划。这些工作主要是通过 BA (Basic Agent) 内部 TA 之间的合作以及 BA 之间的协作完成的。

1.2 IRAIDS 模型描述

IRAIDS 模型由 TA、BA、SA(Supervise Agent) 和 MA(Manage Agent) 组成,如图 1 所示。在一个网络中 BA、SA、MA 通过相互协作监督组成了一个严密安全的入侵检测系统。其中 TA 是 BA 内根据检测到的对手主机的访问情况对对手思维建模的 Agent,主要用来识别单个对手的入侵及其意图。BA 是执行某些检测任务的 Agent,它可以分布在主机或网络上,将多个可疑对手归结为一个对手群,对对手群体入侵目的进行意图识别。SA 是某一逻辑网段的监督 Agent,它监督网段内的 BA 的运行状态并对网段内的流量和访问等信息进行统计。MA 是整个系统的管理者,处于整个网络与 Internet 接口处(通常是网关),对整个系统的流量、网段内的 SA 的状态进行监督管理。

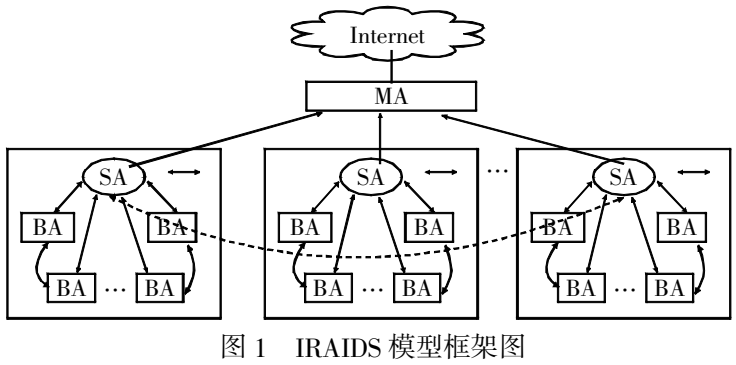

图 1 IRAIDS 模型框架图

2 模型结构分析

2.1 TA 的结构

由于网络中需要检测的对方主机的数量可能会很多,实现对手 Agent 要求所占资源必须足够小,在模型上就需要足够精简。本文对对手建立一个简洁的轻型 Agent 模型,如图 2 所示。对手思维模型由三层组成,分别是:

(1) 交互层。当某个 BA 判断网络中某台主机的行为超过了可能入侵的阈值时,就根据下面的模型对对手建立一个轻型 Agent,以后获取的这个对手的行为就交由该 Agent 处理,由 TA 对对手的操作进行匹配、识别其入侵意图。交互层主要与产生它的 BA 进行交流,由 BA 将它所“关心”的对手情况传送给它,处理后由交互层反馈给 BA。

(2) 处理层。利用所归属的 BA 检测数据库对交互层传来的信息进行入侵规则匹配与意图识别,在单个对手的层次上检测可能的入侵。

(3) 存储层用来存放处理层的中间过程和结果。

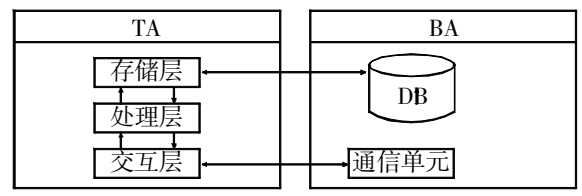

图 2 TA 的结构图

2.2 BA的主要功能和结构

BA 是运用对手群意图识别技术,通过相互协商方式进行入侵检测的 Agent。每个 BA 负责一定的检测任务,检测入侵的某个方面。BA 由下面几个部分组成: DB(数据库) 、通信单元、加/解密单元、数据采集单元、数据预处理单元、分析引擎、状态分析引擎、入侵处理单元、用户接口和对对手思维建模的 Agent(TA) ,如图 3 所示。

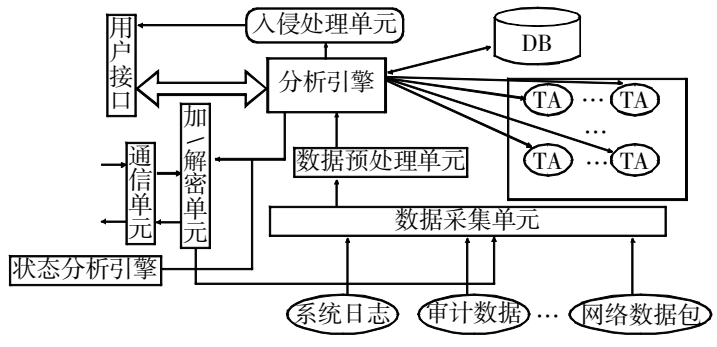

图 3 BA 的结构组织图

下面按照分类详细介绍各部分的功能:

(1) 通信单元。提供 BA 与逻辑网段内其他 BA 以及上层 SA 进行加密通信的能力,也是 BA 的数据来源之一。

(2) 加/解密单元。为了保障 Agent 之间通信的安全化而设立的一个对消息进行加密和解密的单元,是维护系统安全运行的重要单元之一。

(3) 数据采集单元。系统数据的另一来源,从系统日志、审计文件以及网络数据包中获取检测需要的数据。

(4) 数据预处理单元。对从采集单元获得的数据进行过滤、抽象化和标准化操作,便于分析引擎对数据进行分析。

(5) 分析引擎。BA 的核心单元,利用内部 TA 反馈的可疑入侵信息,对可疑入侵群体进行意图识别,同时与逻辑网段内其他 BA 协作检测并向上层 SA 报告。

(6) 状态分析引擎。它是系统进行自我状态进行监督以及对上层 SA 观察的单元,一旦发现 SA 存在问题将及时发动

一次选举,保证系统运行的完整性。

(7) 入侵处理单元。发现入侵时的处理模块,通常的做法是向用户报警并根据其危害性采取限制登录范围、锁定用户账号甚至切断网络连接等措施。

(8) 用户接口。它是用户与 BA 交互的单元,用户可以向 BA 中添加新的检测模型、规则等信息,也可以对可疑入侵进行判断分析并给出结论。

(9) DB。数据库是 BA 存储系统信息、检测模型、经预处理后的数据和中间数据的单元。数据库中的内容主要有下面三部分:

①网络的拓扑信息,包括当前 BA 所处网络的其他 BA 的地址、上层 SA 的地址、MA 地址等描述网络拓扑结构的信息。

②知识库,主要包括入侵检测方法、单个对手意图识别算法、对手群体意图识别算法。其中检测方法数据库包括正常模式库和异常模式库(分别用于异常检测和误用检测方法)。这两种数据库可以根据数据来源分为不同类别子数据库。这样,异常数据可以根据其来源在相应的子数据库中得到迅速的匹配。

③经预处理的数据。所有经过预处理得到的标准化和格式化的数据都保存在数据库中,用来保存对手的入侵证据。

2.3　SA的主要功能和结构

SA 的主要作用是对所在逻辑网段内其他 BA 的状态进行监督管理、接收 BA 发送的可疑报告、同其他 SA 协商以及向上层 MA 提交入侵和可疑入侵情况。由于 SA 是从一个逻辑网段的所有 BA 中选举出来的,与 BA 的结构基本相同。不同的是,SA 数据库中的网络拓扑结构除了包括本逻辑网段内的其他主机的地/址信息外,还保存其他网段的 SA 和部分 BA(便于发现某个 SA 失效时可以通知其网段内的 BA)的地址信息。

2.4　MA的主要功能和结构

MA 是处于整个网络与 Internet 接口处,是对整个网络进行监控和检测入侵的核心单元,常常处于与 Internet 接口的网络设备中,对全网的数据流量进行统计跟踪。MA 可以通过大量的学习,统计正常和异常情况下网络中的数据流量、各个局域网段的流量以及正常和异常的访问信息等知识。通过 MA 在高层监控,一旦发现流向某个网段或者整个系统的流量发生剧烈变化,可以通过学习的结果判断是否出现入侵,并及时向可能发生的入侵网段的 SA 发出警告。

MA 的结构与 BA、SA 十分相似,但是它只需要对网络中的数据包进行分析,在数据采集单元中只接收通信单元和网络中的数据包的信息。其他部分类似,需要详细解释的是 MA 的数据库。MA 的数据库中存放着下面三种数据:

(1) 整个网络主机和 Agent 位置的实时信息,尤其是网络中 SA 的实时位置信息。这个目标主要通过定期检测和接收更新的方式达到。

(2) 知识库,其中包括了 MA、SA、BA 检测需要使用的所有知识。也就是说,MA 的数据库中存放了整个系统需要的知识的总和。

(3) MA 从网络中获取的数据流量和统计信息,以及入侵检测的中间信息都存储在数据库中,便于进一步的统计学习。

MA 为用户提供了观测整个系统的运行状态、网络中的数据流量以及动态配置系统的功能。更重要的是为用户提供了一个向系统添加修改检测模式的入口,可以随时更新系统的检测模式库,增强了系统检测入侵的能力。

2.5　系统运行流程

系统工作的流程主要包括两部分,即自底向上过程和自顶向下过程。

2.5.1　自底向上处理过程

(1) BA 获取系统日志、审计数据以及抓取网络中的需要检测数据包,经过预处理后根据源 IP 地址进行以下分类处理:

①对未建立的对手模型的数据包通过定义的疑似度函数度量入侵的可能性,对超过阈值的包建立对手思维追踪 Agent 即 TA,并将数据交给 TA 处理。

②对已经建立起 TA 的对手直接将数据传给 TA。

(2) 获取数据后的 TA 进行误用检测、异常检测、意图识别等一系列操作,如检测到可疑入侵则根据异常检测设定的关联阈值判断是否需要向 BA 报告,若发现入侵则立即报告。

(3) BA 接收到 TA 发送的关联检测报告时,将其内部 TA 的关联阈值合并起来,看其是否超过联合攻击阈值,一旦超过立即进行联合入侵检测和群体意图识别,即将各个 TA 在数据库中存储的数据合并起来,先检测是否为已知的联合攻击,再对对手群体意图进行识别。对于对手群的行动不符合任意已知的联合入侵方式但入侵可疑度很高的,将利用数据挖掘技术从存储的数据中挖掘出新的入侵模式并向系统中其他 Agent 广播。

(4) 当 BA 没有发现大量可疑入侵,但 SA 监测网段内的数据流量等信息反映入侵可能发生时,BA 之间通过相互协作检测入侵。

(5) 当 MA 根据统计分析判断系统可能被入侵时,SA 之间相互协作,完成共同检测复杂入侵的任务。

2.5.2　自顶向下处理过程

(1) 用户通过 MA 提供的接口向全局数据库中添加能检测新的入侵方式的模型,随后 MA 将模型发送到下层的 SA 中。SA 接收到新的检测模型后,根据网络情况选择要扩充此类检测类型的 BA,并将模型发送到 BA 的数据库中。

(2) BA 接收到新的入侵检查模型后,立即更新 TA 的检测方式,即 TA 在检测时对这个新模型的规则也进行匹配。

3　关键技术

3.1　检测方法

文献[8]中提到了用 IP 陷阱的方式,利用 IP 陷阱和流量标本不仅能够识别已知的异常行为,还能够不断地学习和积累。文献[9]中提到了一种追踪可疑用户入侵链的方式获取入侵者的访问信息。在检测方法中,不管是误用检测还是异常检测,基于主机型或是基于网络型,各种检测方法和技术手段都各有利弊。目前还没有通用的检测方法出现。比较恰当的

做法是: 综合分析各种检测方法的利弊以及最优使用场所, 在复杂的网络环境中根据各个网段的情况, 在每个网段中部署多个最优检测方法的 Agent, 这些 Agent 之间相互合作共同完成检测任务。这样做的好处是可以利用多种检测方法的优点, 避免其缺点, 增强系统的检测能力和检测效率。

使用误用检测和异常检测相结合的方法进行入侵检测, 即首先对于获取的数据进行误用检测; 如不能判断是否为入侵(有可能是未曾识别出的入侵类型) 再使用异常检测方法获取一个可疑度, 将此可疑度与阈值相比判断是否为入侵。

3.2 Agent 状态监测及恢复策略

系统中的 BA、SA、MA 均存在状态分析引擎, 用于监控自身所处的状态(如负载重、轻、没有负载、是否被入侵等) 、检测其他 Agent 的状态和回复其他 Agent 发送的询问请求。Agent 之间状态的监督包括下面几种: BA 内的监督、SA 内的监督、BA 与 SA 之间相互监督, 以及 MA 与 SA 之间相互监督。其中, BA 每隔一定时间检测所在网段的 SA 是否正常工作, 并在需要时报告入侵检测情况, SA 也采用同样的方法检测 BA 的状态; BA 内、SA 内采用被动监督方式, 即在需要协同工作时才检测对方状态; MA 与 SA 的监督方式和 BA 与 SA 的相似。

状态监督可以保证系统整体的正常运行, 当系统中某些 Agent 出现问题时按照下面的恢复策略进行恢复:

(1) 如果 SA 发现 BA 没有响应状态请求, 判定失效后立即调整网内其他 BA 的检测工作, 让其他 BA 合作完成失效 BA 的检测任务。如果在一定时间段内失效的 BA 仍没有恢复, 在确定所在主机正常工作的情况下复制一个新的 BA 重新开始检测。

(2) 如果 BA 发现 SA 停止响应, 则立即发动一次选举, 从网络中正常工作的 BA 中选举一个负载最轻的担任 SA 任务; 被选中的 SA 向 MA 报告网段内 SA 的更换信息, 确保 MA 的及时更新。

(3) 如果 MA 或 SA 发现 SA 停止工作, 则发送消息给同网段的其他 BA 发动一次选举。具体的选举算法可以使用 Bully 或者是 Ring 算法。

此外, Agent 之间还存在着相互协作的关系。具体而言, 这些协作包括单个 BA 内 TA 之间的协作、BA 之间的协作、SA 之间的协作。

3.3 Agent 安全问题

IDS 的引入使得系统的安全得到了一定层次上的保证, 然而, 处于主机或网络中的 Agent 也有可能被对手攻破, 成为操纵系统的工具。为保证 Agent 的安全必须考虑两个问题, 即 Agent自身的安全和 Agent 之间的通信。文献[10] 使用 BEEP (Blocks Extensible Exchange Protocol) 来进行用户验证, 确保传输数据的完整性和机密性。本文设计的系统中, 对 Agent 之间的通信采取了加密的方式, 这样即使对手捕获了数据包也不容易获取其中的信息; Agent 自身的安全是由 Agent 内部的状态分析引擎以及其他 Agent 完成的。状态分析引擎实时监督 Agent所处的状态, 并在异常状态时立即向上层 Agent 报告, 同时其他 Agent 如发现其状态异常也可以向上层 Agent 报告。

3.4 Agent 学习

Agent 具有学习能力是入侵检测系统能够适应网络环境变化的重要条件之一, 也是当前对入侵检测进行研究的一个热点。在本文的系统中, Agent 的学习主要是通过异常检测和入侵模式挖掘的方法实现的。当 Agent 通过异常检测或者高层的检测发现新的入侵时, 通过使用数据挖掘技术从访问数据中挖掘出新的入侵规则并在检测系统中广播, 使得其他 Agent 也立即具备检测这种新的入侵方式的能力。

4 结束语

本文提出了一种基于对手思维建模的多 Agent 分布式入侵检测系统模型 IRAIDS, 可以识别单个对手的意图和对手的群体意图, 适用于检测大规模的分布式协作攻击。另外, 系统还具备了学习能力, 可以识别并检测出未知的入侵方式。本模型没有采用集中式的控制方式, 主要通过 Agent 之间的协作完成检测任务, 但考虑到集中式控制方式的优点, IRAIDS 模型中设计了处于中上层的 SA 和 MA, 用于统计入侵情况和为用户提供扩展接口和报警措施。但是 SA 和 MA 本质上是和 BA 与 SA 相同的 Agent, 只是分工的不同而已。

参考文献:

[1] UNT T. Detecting intruders in computer systems: proceedings of the 6th Annual Symposium and Technical Displays on Physical and Electronic Security [ED/OL]. http://citeseer.ist.psu.edu/lunt93detecting.html.

[2] NORTHCUTT S. Network intrusion detection: an analyst's handbook [M]. Indianapolis: New Riders Publishing, 1999.

[3] VIGNA G, KEMMERER R A. NetSTAT: a network-based intrusion detection system: proc. of the 14th Annual Computer Security Applications Conference[C]. Scottsdale: [s. n.], 1998.

[4] TSENG C Y, BALASUBRAMANYAM P, KO C, *et al*. A specification-based intrusion detection system for AODV: proc. of the 2003 ACM Workshop on Security of Ad hoc and Sensor Networks (SASN'03) [C]. Farirfax: Virginia, 2003: 125-134.

[5] DASGUPTA D. Immunity-based intrusion detection systems: a general framework: proceedings of the 22nd National Information Systems Security Conference (NISSC) [C]. [S. l.]: [s. n.], 1999: 147-160.

[6] 马恒太, 蒋建春, 陈伟锋, 等. 基于 Agent 的分布式入侵检测系统模型[J]. 软件学报, 2000, **11**(10): 1312-1319.

[7] JAI S B, JOSE O G, DAVID I, *et al*. An architecture for intrusion detection using autonomous agents[D]. [S. l.]: Department of Computer Sciences, Purdue University, 1998.

[8] 陈硕, 安常青, 李学农. 分布式入侵检测系统及其认知能力[J]. 软件学报, 2001, **12**(2): 225-232.

[9] 张勇, 张德运, 李胜磊. 基于分布协作式代理的网络入侵检测技术的研究与实现[J]. 计算机学报, 2001, **24**(7): 736-741.

[10] SHI Zhicai, JI Zhenzhou, HU Mingzeng. A novel distributed intrusion detection model based on mobile agent[J]. **ACM SIGOPS Operating Systems Review**, 2003, **37**(1): 46-53.