一种基于免疫入侵检测的攻击路径标志技术研究

张凤斌，孙 刚，张 斌
（哈尔滨理工大学 计算机科学与技术学院，哈尔滨 150080）

摘 要：针对免疫入侵检测和攻击源追踪结合技术进行了研究。采用分布式免疫入侵检测系统与数据包标记理论，利用免疫入侵检测系统实时分析的网络数据特征指导路径标志技术动态处理，使路径标志方法能动态自适应不同网络数据特征，快速识别攻击路径，为免疫入侵检测系统针对攻击路径培养特征检测器提供路径信息。实验表明这一方案能快速重构出攻击路径信息，在收敛效率、误报率方面的表现优于目前的概率包标记算法，能为免疫入侵检测系统提供特征路径信息。

关键词：IP 包标记；攻击源追踪；自适应机制；人工免疫
中图分类号：TP393.04 **文献标志码**：A **文章编号**：1001-3695(2014)01-0217-05
doi:10.3969/j.issn.1001-3695.2014.01.051

Technology research of attack path identification based on immune intrusion detection

ZHANG Feng-bin, SUN Gang, ZHANG Bin
(*College of Computer Science & Technology, Harbin University of Science & Technology, Harbin* 150080, *China*)

Abstract: This paper did a research regarding the immune intrusion detection system and the IP traceback. By deploying distributed immune intrusion detection and data packet mark, this paper utilized immune intrusion detection system to analysis network packet's characteristic in real time to instruct path marking technique dynamically, making path marking method adaptable to different characteristics of network data, rapidly recognized attack path, and provided the path information to the immune intrusion detection system with the regard to develop characteristics checker for attack path. Experiments show that this method can reconstruct the information of attack path very fast, and is better than the-state-of-art probabilistic packet marking algorithm in terms of the convergence efficiency and false alarm rate, can provide the information of attack path to immune intrusion detection system.

Key words: IP packet marking; IP traceback; adaptive mechanisms; artificial immune

0 引言

计算机网络安全技术是一个多学科、跨专业的综合性学科，包括传统防火墙技术、访问控制技术、加密技术、入侵检测技术、攻击源追踪技术等[1]。网络安全技术的研究包含双重性，即攻击性和防御性，其中免疫入侵检测技术和攻击源追踪技术作为网络安全技术攻防两端的代表技术，虽然现阶段还不成熟，但作为网络安全技术的重要发展方向，持续受到专家学者的关注。免疫入侵检测系统的思想来源于生物免疫系统对非自体物质的识别和处理[2]，该系统并不依赖于大量特征库来判断入侵与否，而是将正常网络流特征建模，一旦当前网络特征不在正常范围内，则认为发现了潜在攻击行为。所以免疫入侵检测系统具有良好的动态自适应能力，对于未知攻击敏感度高，非常适合应用于当前多变的网络环境之中。另一方面，为了找到网络攻击的实施者并对其进行惩处，对潜在的攻击者产生威慑作用，攻击源追踪技术也稳步发展[3]。一直以来，攻击源追踪技术始终作为一种单独的网络安全技术被研究，但其研究成果一直难以应用，究其原因：a) IP 协议设计时的缺陷，当前追踪技术仅依靠更改 IP 协议，较难提高其效果；b) 追踪算法采取相对固定的算法追踪策略。为改进这一策略，本文提出一种将免疫入侵检测系统和攻击源追踪技术优势互补的免疫入侵检测的攻击路径标志技术。该技术通过免疫入侵检测系统实时分析的网络流特征指导一种改进的攻击源追踪算法，使算法收敛快；之后通过路径标志后续的数据包，为免疫入侵检测系统针对该路径培养特征检测器提供识别信息。

1 相关技术背景

1.1 免疫入侵检测技术现状

人工免疫入侵检测技术自 1994 年被提出以来，一直被当做典型的基于异常的入侵检测系统研究[4]。该技术不依赖于特征库，能够及时发现新型的、潜在的网络攻击，从而保护主机免受入侵。此后，美国孟菲斯大学提出的一种基于人工免疫的多代理入侵检测系统，使入侵检测系统不再局限于本地主机，各代理之间协同处理，提高了系统的鲁棒性，但存在覆盖不完全的问题[5]。Gonzalez 等人[6]提出使用实值来表示网络数据流特征，使得提取的特征方便人们理解，且更易于对数据分类。2010 年 Twycross 博士[7]提出了第二代人工免疫的概念，试图将先天性免疫与适应性免疫相结合，提出了评估测试免疫算法

收稿日期：2013-03-19；**修回日期**：2013-04-22

作者简介：张凤斌(1965-)，男，教授，博导，主要研究方向为网络与信息安全(zhangfb@hrbust.edu.cn)；孙刚(1986-)，男，硕士研究生，主要研究方向为网络与信息安全；张斌(1988-)，男，硕士研究生，主要研究方向为网络与信息安全.

的模型,但该模型依旧难以实际应用。当前免疫入侵检测系统仍存在自体集不完全,本地入侵检测系统负担过重等问题,但其对未知的新型攻击具有高度敏感性,对已知攻击的快速反应能力及较强适应网络特征变化等属性,具有应用价值。

1.2 攻击源追踪现状

攻击源追踪技术在2000年左右开始受到学术界的重视[8],Stone在这年提出一种通过入口调试的方法递归地查找攻击源的技术,但需要得到路由器管理权限,而且只能在攻击发动中进行攻击,没有应用价值。Burch则提出了一种基于拒绝服务方法的链路测试追踪策略,但该方法本身也给网络带来巨大负担。此后,Bellovin提出的基于ICMP协议的攻击源追踪技术中应用了一种新的iTrace报告来跟踪攻击源,通过路由器协调发送此类消息来帮助受害主机识别攻击路径,但该方法会使路由器负载过高,影响其正常性能,而且ICMP报文容易被安全策略过滤[9]。当前研究的攻击源追踪技术主要还是对概率包标记方法的改进,该方法由Savage提出,要求路由器以固定的概率 p 对为每一个通过的数据包标记上自己的地址信息,当攻击主机向受害主机发送大量攻击数据包时,受害主机可以通过收到的数据包上的信息重构出攻击路径[10],但该方法不但存在最弱链问题[11],且重构需要大量数据包使得效率偏低,而且算法的误报率高,抗干扰性差。以包标记为代表的被动追踪算法虽然一直在进行改进,但是由于IP协议的固有缺陷和固定格式,使得单独使用包标记算法或改进的包标记算法存在着各种限制,而基于ICMP协议实现的iTrace方法提供了良好思路,但是基于路由器的实现方法需要进行改进。另外,目前追踪算法都不具有动态适应性。包标记算法并不根据数据包特征,或者数据流量的特征动态检测与追踪,而是始终采用固定的标记策略,这就导致了其一方面标记了大量的正常数据包,另一方面在追踪攻击主机时,显得效率不足。

2 基于分布开放式免疫入侵检测的攻击路径标志模型与算法

将分布式免疫入侵检测系统与攻击源追踪技术相结合,将一分布开放的免疫响应服务器放置于交换网络中的关键路径位置,为本地免疫入侵检测系统提供服务。一方面可以利用免疫入侵检测系统实时分析的网络数据特征情况动态地指导攻击源追踪算法的展开,使算法具有动态自适应性,快速找到攻击的具体路径;另一方面,由于单次攻击特征在形态空间中相对连续,如图1所示,将后续数据包标记上特定路径标志,使免疫入侵检测系统可以针对特定路径的攻击数据采用肯定选择方法培养生成免疫检测器,对该路径后续的攻击数据包进行响应处理。

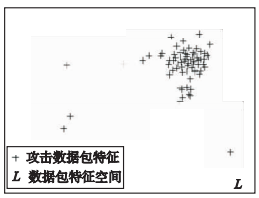

图1 单次攻击特征在形态空间中的映射

为了方便算法和模型的讨论及研究,现规定前提条件如下:

a)数据包在传输过程中可能丢失也可能失序。

b)相同源主机和目的主机传输的数据包路径基本稳定。

c)攻击者可以产生任意想要的数据包。

d)路由器是可信的,但是资源有限。

基于分布开放式免疫入侵检测的攻击路径标志模型与算法的设计主要分为以下步骤:

a)构建总体模型与框架,确定免疫响应服务器位置。

b)确定模块和具体工作步骤。

c)重新定义IPv4数据包头,用于数据包标记。

d)设计各模块算法。

2.1 系统总体模型

基于免疫入侵检测系统的路径标志技术模型构建和核心思想是将免疫入侵检测系统的分析与响应模块置于传输网络中,实时分析网络数据后,根据数据特征决策动态调用攻击源追踪算法模块。通过参考分布式入侵检测系统及分布开放式的入侵检测与响应架构IDRA技术[12]和DDoS分布式处理思想[13],在传输网络关键路径(如网络边界路由器)处部署一种开放式的免疫追踪响应服务器,如图2所示。该服务器不属于某一单独的入侵检测系统所有,而是能够与任意免疫入侵检测系统提供服务并协同工作,该服务器用免疫检测器与通过它的数据包进行匹配,用于判断数据包是否需要进行响应处理或攻击路径追踪,并协同现存的各种路由器实现路径标记,使追踪算法能快速收敛,及时追踪到理想的位置。另外,该服务器的设置可以释放之前各类算法对于路由器资源的大量依赖,使得路由器仅处于协同工作,而非主动、高负载的追踪状态,可以对网络状况有较好的保证。

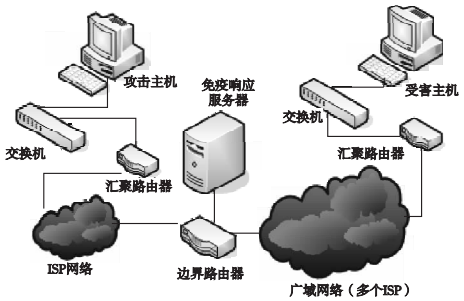

图2 免疫响应服务器位置示意图

2.2 模块设计

基于免疫入侵检测系统的攻击源追踪算法主要分为追踪/响应服务器算法、路由器算法和路径重构算法。模块设计如图3所示 。

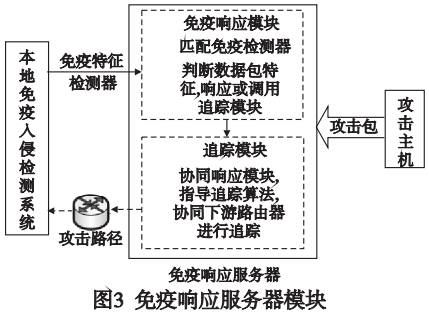

图3 免疫响应服务器模块

追踪算法仅针对异常的数据包进行追踪,虽然看似路由器需要实现的协同算法比较复杂,但网络中绝大多数传送的正常数据包并不处理,相对其他追踪算法能减少路由器及网络负担,但需要免疫响应服务器有较好的性能,并能在辨识攻击路径之后,为免疫入侵检测系统提供该路径数据包信息。

下面描述基于免疫的分布开放式系统模型的整体步骤:

a)免疫响应服务器利用节点内的免疫记忆检测器和特征检测器对经过的网络数据包进行分析,对于没有被两者匹配的数据包进行正常转发,删除与特征检测器匹配的数据包,而对于记忆检测器匹配的数据包开始追踪算法。

b)免疫响应服务器向受害主机申请路径标志 F,同时提交判断为异常的数据包信息。

c)免疫响应服务器向受害主机发送 64 个追踪数据包,供下游路由器标记,且包中数据为申请的路径标志。

d)下游路由器识别自己应该标记的数据包,并将自己的 IP 记录其两个数据包中。

e)受害主机验证数据包,并重构出从免疫响应服务器到自己的路径。

f)受害主机通知免疫响应服务器,重构完成,接收由改进的概率包标记追踪的控制区域外的攻击路径。

g)受害主机通知重构路径中中间的免疫响应服务器,对标记为 F 的数据包不再重复标记。

h)此后,免疫响应服务器将经过自己并发往受害主机的所有数据包通过路径标志符 F 进行标记,以便于本地免疫入侵检测系统对该路径进行免疫特征检测器生成。

i)受害主机免疫入侵检测系统区分标记数据包为自体(正常)或者非自体(攻击)数据。

j)受害主机免疫入侵检测系统通过否定选择算法对非自体攻击数据包特征培养生成免疫特征检测器。

k)将生成的攻击免疫特征检测器传送回免疫响应服务器的特征库,在其生命周期内识别并处理后续的攻击数据包。免疫特征检测器生命周期结束后则放入记忆异常检测器,用于发现潜在攻击。

其中,步骤 a) ~ h)为基于免疫入侵检测系统的路径标志算法需要设计实现的,是本文讨论的重点。

2.3　IPv4 包头格式重定义

为了实现基于免疫入侵检测系统的路径标志算法,需要对 IPv4 的包头结构进行重新定义,以便适应算法要求。当前网络主要通过 IPv4 传输数据包,每一个数据包含有包头和数据区两部分,其包头结构如图 4 所示。路由器是一种工作在第三层的网络设备,它可以对 IP 数据包的包头部分进行操作以完成数据包的寻路和转发。为了尽量不对现存网络传输协议进行改进,同时得到路由器的地址信息。如图 4 中灰色部分,在 IPv4 包头中找到了 25 位可以在追踪算法中进行利用,包括 8 位服务类型字段(图中①)、16 位标志字段(图中②)和 1 位保留字段(图中③)。在当前 IPv4 协议中,服务类型字段没有统一格式标准,保留字段没有被利用,标志字段则用于数据包分段,通过分析目前网络中 IPv4 的传输情况可知,需要进行分段处理的数据包不足总量的 0.25%,因此在各种攻击源追踪算法中,这 16 位标志字段是利用的重点。

图4 IPv4包头结构

为配合基于免疫入侵检测系统的攻击源追踪模型的实现,将 IPv4 中 25 位重新定义:将保留位重定义为控制区域位;将 8 位服务类型字段重新定义为 1 位追踪标记位、1 位 IP 偏移位、1 位 IP 校验位和 5 位的标记路由器距离字段;16 位的标志字段作为一个整体空间使用,并根据算法不同阶段重新定义成了三种互斥的意义,分别是 TTL 比对字段、路由器标记 IP 字段和路径标志字段。

1)控制区域位 当数据包通过第一个追踪响应服务器后,该位置为 1,表示进入控制区域。所谓控制区域是指数据包进入了可以被追踪响应服务器分析的网络传输部分。

2)追踪标记位 该位默认为 0,表示正常数据包;当该位置位时,表示当前数据包是由追踪响应服务器产生的,供下游路由器标记其地址信息的追踪专用追踪数据包。

3)IP 偏移位 该位用于告知路由器需要标记其 IP 地址的前 16 位还是后 16 位。

4)IP 校验位 用于 IP 地址校验。

5)标记路由器距离字段 标记路由器距离字段默认值为 0,路由器在标记自身 IP 到路由器标记 IP 字段的同时置为 1,其他路由器则读取该字段,若不为 0 则加 1,即可统计经过的路由器数量。

6)TTL 比对字段 算法利用 TTL 自减的特点,通过 TTL 比对字段和数据包 TTL 进行比较,使下游路由器判断是否需要标记。即,将追踪数据包的 TTL 初始设置为 33,每过一个路由器则自动减 1,将 TTL 比对字段设置为 32 ~ 1 依次递减地标记到标志段。下游路由器一旦发现追踪数据包的 TTL 比对字段和当前数据包 TTL 值相同,则标记自己的地址信息。

7)路由器标记 IP 字段 由于 TTL 比对字段和路由器标记 IP 字段在标记时间上互斥,所以 TTL 比对字段完成标志识别后,可以用于路由器标记 IP,两者可以根据标记路由器距离字段时为 0 进行区别。

8)路径标志字段 该字段同样与 TTL 比对字段和路由器标记 IP 字段存在标记时间的互斥性,可以定义在同一包头区域内。设置该字段是为了防止攻击者伪造追踪数据包,干扰追踪算法,控制追踪算法在产生追踪数据包之前需要首先通过一个同样为 TTL = 32 的申请标志数据包向受害主机申请一个 16 位长度和随机标志符,受害主机记录收到的申请标志数据包的 TTL 值,并返回随机标志符,追踪响应服务器收到随机标志符后,将它放入 64 个追踪 IP 数据包的数据域内,并将 TTL 设置成 33。受害主机收到自称为追踪数据包时,首先检查其携带的随机标志符和 TTL 是否与其给予的随机标志符和记录的 TTL 相同,若不相同则说明该数据包为恶意伪造的,并将其丢弃。当追踪响应服务器发送完追踪数据包后,它就使用申请到的随机标志符标记到之后经过它的每一个数据包的标志字段,受害主机通过该标志符便可了解数据包来源的路径。

2.4　算法设计

2.4.1　免疫响应算法

该算法运行于免疫响应服务器的免疫响应模块中。算法利用免疫节点中的记忆免疫检测器和特征免疫检测器对通过的数据包进行比对。其中记忆检测器集合来自于对生命周期结束的特征免疫检测器的进一步遗传变异,而特征检测器则来自于受攻击主机端的私有免疫入侵检测系统对单一攻击特征培养和提交。数据包一旦与特征免疫检测器匹配,则认为该包属于攻击行为,可以进行删除;而如果与记忆检测器匹配,则认为当前数据包与之前的某次攻击行为相似,属于潜在的攻击,从而对其进行追踪。

当然,为了进一步提高免疫响应节点对潜在攻击的识别能

力,记忆检测器集合能接收该免疫响应服务器对通过本地的数据特征长期建立正常模型后通过肯定选择算法生成的非自体免疫特征,这可以由免疫响应服务器的处理性能和特征建模的情况具体决定,在此不作进一步讨论。

免疫响应算法如下:

a)处理下一个数据包,将数据包控制区域位置1。

b)对该数据包进行特征提取。

c)将数据包特征与其目的IP一致的免疫入侵检测系统提交的免疫特征检测器匹配,成功则将数据包删除,并转算法a),否则继续。

d)将数据包特征与记忆检测器集合匹配,失败转算法e),否则转算法g)。

e)检查是否存在于其目的IP一致的路径标志符F,存在则将F标记入路径标志字段并转发,否则直接转发。

f)回算法a)。

g)记录该数据包信息,并提交给追踪模块开始通过控制追踪算法进行路径标志。

h)回算法a)。

2.4.2 控制追踪算法

该算法运行于追踪模块,被免疫响应模块调用。出于安全性、可控性和可行性的考虑,在控制区域内的路由器并不主动标记地址,而是由免疫响应服务器生成一种特定的供下游路由器协同操作的数据包,称其为路径追踪数据包,由路由器识别并依次标记自身地址。由于每个路径追踪数据包中只有16位的空间可以携带地址信息,则一个路由器标记自己的地址需要两个数据包,且通过对目前网络数据包的监控了解到,从任意主机到达受害主机的路径上的路由器不会超过32个,所以免疫响应服务器只需要产生64个路径追踪数据包。为了使下游路由器能够判断应该将自己的IP地址标记到哪个追踪数据包,追踪数据包中必须有一个对比标记,又因为标记IP后则不再需要对比标记,因此可将对比标记和标记IP共用一个头部,即包头标志段。控制追踪算法利用TTL自减的特点,通过在报头标志段设置TTL比对和TTL进行比较,判断是否需要标记。

其具体算法如下:

a)向追踪的目的IP地址的免疫入侵检测系统申请随机路径标志符F,同时提交引起该次追踪算法的异常数据包信息。

b)等待申请的路径标志符,超时或拒绝则结束,成功则继续。

c)生成64个路径追踪数据包,每两个1组,一共32组;每组内的两个数据包IP包头的TTL比对字段相同,32组分别从33~1依次递减;每组内的两个数据包的IP偏移位不同,一个为0,另一个为1;所有64个数据包的追踪标记置1,距离字段置0;数据包数据部分载入路径标志符F。

d)将64个数据包发往目的IP。

2.4.3 路由器协同算法

下游路由器的协同算法是根据通过数据包的不同状态而做出的不同动作。路由器首先判断控制区域位,如果该位等于0(暂不考虑伪造情况),则说明该数据包没有进入控制区域,无法被控制追踪算法追踪。为了弥补此段(一般为完整路径的前几跳,不超过5跳)追踪算法的缺失,路由器配合采用改进的概率包标记的算法标记数据包[14],该段追踪算法将在通过的第一个免疫响应服务器构造,并可以发送至受害主机,由受害主机构造完整路径。一旦确定数据包P进入了控制区域,协同路由器则通过追踪标记位判断P是普通数据包还是需要协同处理的追踪数据包,当P为追踪数据包时,读取P的TTL比对字段(标志字段)和P的TTL对比,如果不相同则判断其距离字段是否不为0,不为0时说明该包已经被上游路由器标记,需要距离加1后转发,否则直接转发。当P的TTL比对字段和P的TTL相同,则路由器通过IP偏移位将自己IP地址的前16位或者后16位标记到TTL比对字段,即标志字段,一旦IP标记到标志字段,标志字段的含义则从TTL比对变成IP标志。同时,P的奇偶校验位根据填入的标记IP置位。至此,路由器完成其协同算法。由以上描述可知,标志字段的值小于32时,应该作为TTL比对值,为了防止冲突,所以受害主机分配的16位随机标志符的10进制值应该大于32。

路由器协同算法如下:

a)处理下一个数据包。

b)若数据包控制区域位为0,则通过改进的概率包标记算法追踪该数据包并转算法a),否则继续。

c)若数据包的追踪标记位为0,则转发该数据包,否则继续。

d)若数据包的距离字段不为0,则将距离字段加1后转发该数据包,否则继续。

e)若数据包的TTL比对字段与TTL不同,则转发该数据包,否则继续。

f)若数据包中的IP偏移字段为0,则路由器将自己IP的高16位标记入数据包的标记IP字段,并将其距离字段置1,IP校验字段填入校验位后转发,并回算法a)。

g)路由器将自己IP的低16位标记入数据包的标记IP字段,并将其距离字段置1,IP校验字段填入校验位后转发,并回算法a)。

受害主机收到64个追踪数据包后,首先验证其包内数据字段的随机标志是否是自己发送的,随后可以迅速根据数据包的内容构建出该路径,并向免疫响应服务器发送一个确认报文,告知其追踪完成。

当免疫响应服务器收到确认报文后,说明控制追踪算法完成。此后,免疫响应服务器可以将控制区域外的通过概率包标记生成的路径告知受害主机,使受害主机构造完整的路径。更重要的是,免疫响应服务器将之后通过它发往受害主机的数据包的标志符中均写入之前获取的16位标志符,使得受害主机知道此后哪些数据包是来源于该路径的,方便进行处理,或者训练特定的免疫检测器。由于一条路径上可能存在一个或者几个免疫响应服务器,如果都将自己的随机标志符标记到数据包上,那只有离受害主机最近的服务器标志能保留。为了解决这一问题,需要受害主机在本地根据收到的数据包构建一个完整的路径树,一旦发现一条路径上存在多个免疫响应服务器,则发送数据包通知中间的服务器,不再重复标记最远的服务器随机标志。该方法可以有效防止攻击者伪造随机标志而造成路径中所有免疫响应服务器均不标记数据包的情况,也能使追踪的路径最长,标记包的特征性更为明显。另外,当某个随机标志符在一定生存时间内没有再使用,则受害主机清空该标志符的所有信息,可以重新将该标志符分配出去。

3 算法的理论分析

当前主流的追踪算法中,路由器的性能和行为造成了追踪

算法的各种缺陷,如重构需要的数据包数量太大、误报率高等问题。由于路由器需要转发大量数据,因此不适合于大量计算和存储,且路由器仅能对 IP 数据包的包头进行操作,这就无法让所有的路由器地址标记在有限的包头之中,甚至标记一个路由器地址都需要大量数据包(包头选项不实用)。而免疫响应服务器的存在可以改良或解决以上问题:该算法中并不对正常的数据包进行路径标记,而是通过免疫响应服务器实时分析网络特征后,视情况主动产生路径追踪数据包发往下游,供路由器识别和标记。因此可以解决概率包标记收敛速度慢、路由器算法复杂、追踪策略不具有动态性,以及最弱链等问题。而且免疫响应服务器中的记忆检测器可以根据网络特征的变化而进化,使追踪算法对网络特征具有自适应性。

评价一个攻击源追踪技术优劣的主要指标包括算法的收敛速度、误报率和抗干扰性三个方面进行衡量。一个优秀的攻击源追踪算法应尽量做到快速收敛、较低误报且能抵挡攻击者的恶意干扰性攻击。下面从理论上分析基于免疫入侵检测系统的攻击源追踪算法的收敛性、误报率和抗干扰性。

1)收敛性　该算法的收敛需要控制区域内路径追踪算法和控制区域外路径追踪两部分算法都收敛。在控制区域内,收敛于算法开始后发送的第 64 个数据包。追踪算法开始于某一数据包与正常数据包的特征相似度较小,即与追踪响应服务器中的某个自体检测器的亲和力小于给定的阈值。在控制区域外,如果采用改进的概率包标记算法,由于攻击者离第一个边界路由器一般不超过 5 跳,则可以设标记概率 $p = 1/5$。对于一个控制区域外 5 跳的攻击路径,仅仅需要几十个数据包,就可以构造出来。另一方面,控制区域外的路由器应该属于唯一的 ISP 所有,通过 ISP 方记录日志文件,也可以较容易地找出数据包发送者。

2)误报率　该算法中用一个 16 位的随机标志符唯一地标志一条追踪响应服务器到受害主机的路径,只要攻击路径上不多于 $2^{16} - 32$ 个追踪响应服务器,即 65 504 个时,不会出现主动误报或者漏报的情况。但一旦主机遭受超过 65 504 个追踪响应服务器的攻击,会导致受害主机不能分配随机标志符无法完成追踪,出现漏报情况。根据实际攻击经验,一般即使是 DDoS 网络攻击,攻击主机也不会超过几百台,本文方案完全可以避免误报路径的出现。

3)抗干扰性　基于免疫入侵检测的攻击源追踪算法不使用普通的数据包作为路径信息的载体,而是以追踪响应服务器产生的路径追踪数据包进行 IP 传送,给攻击者伪造路径追踪信息增加了难度,基于免疫的路径标志算法可以在受害主机本地记录对应的申请路径标志的免疫响应主机数据包的 TTL 值,这样攻击者想干扰一个路径构建需要满足随机标志符和对应的 TTL 跳数完全与受害主机记录的相同。这对攻击者来说是很难做到的,其成功仿造一个路径追踪数据包的概率约为 $\frac{1}{2^{16} \times 32}$,而且即使伪造成功,攻击者也难以知道。所以算法的抗干扰性较强。

4　仿真实验

本文通过在 Linux 下用 NS-2 和 C 语言等仿真手段对算法中的核心部分进行实验分析,观察算法在收敛时间、误报率等方面的表现。

4.1　收敛性实验

实验 1 中对在单一攻击路径下,攻击主机与受害主机距离从 1 ~ 30 的不同情况对追踪算法收敛时间进行比较。其中假设免疫响应服务器放置于第 5 跳的位置当中,在控制区域外采用改进的概率包标记算法进行追踪。实验结果如图 5 所示。

从实验结果可以看出,由于免疫响应服务器的存在,能够协助对攻击路径进行标记,使本文提出的免疫路径标志算法在追踪路径的时候,相对其他两种包标记算法的算法收敛时间降低到常数级,只要攻击路径上第一个免疫响应服务器与受害主机的距离不超过 32 跳,则都能很快地找到攻击路径。控制区域外的收敛算法虽然仍采用概率包标记的改进方案,但由于跳数相对很少,从实验结果中也可以看出其算法也能够快速收敛。而两种概率包标记的追踪方案在算法的收敛时间上和跳数成正比,基本概率包标记在攻击主机与受害主机距离增大时,最弱链的问题慢慢凸显出来;而高级概率包标记虽然收敛时间相对基本概率包标记方案有较大改进,但其要求实现知道网络拓扑结构,在实际情况中难以应用。

4.2　误报率实验

实验 2 中对在多攻击路径下三种路径标记算法的误报率情况进行了模拟。根据实际攻击的情况,即使是大规模分布式拒绝服务攻击,其攻击路径也不会超过 1 000 条,所以实验中对 1 000 条攻击路径内的情况进行了仿真,其结果如图 6 所示。

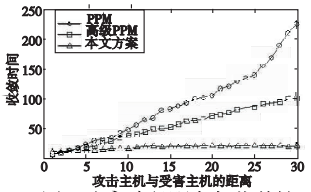

图5 攻击路径距离与收敛性

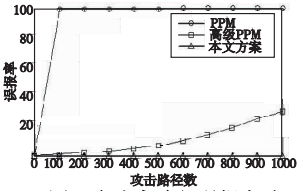

图6 多攻击路径误报实验

从实验结果可以看出,由于没有针对误报的问题采取办法,基本 PPM 的误报情况非常严重,面对多条路径的攻击时完全不能正确地重构出攻击路径。高级 PPM 对误报率方面有较大改进,但在攻击路径增加时,其误报率也快速上升。本文提出的基于免疫的路径识别算法中,由于免疫响应服务器和受害主机进行联系并确定了唯一的路径标志符,所以在控制区域内不存在误报的情况,控制区域外的误报率则需要根据采用的不同方法进行不同的讨论。低误报率使得追踪的路径信息的信息熵更高,对于后续的免疫入侵检测系统的针对性处理更具有参考价值。

4.3　实验结论

通过以上讨论和实验可以看出,基于免疫入侵检测的路径标志技术在查找攻击包来源时的效率比现存的概率包标记等方法更高。其能够在一个相对固定的时间内找到攻击数据包的来源,且保证路径信息的真实可信性,与其他两种追踪算法相比更具应用价值。

5　结束语

本文通过参考 IDRA 和 DDoS 分布式处理思想,结合分布式免疫入侵检测系统理论,设计实现了一种基于分布开放式免疫入侵检测系统的路径标志技术。技术实现依赖于部署于交换网络中的免疫响应服务器,该服务器一方　(下转第 235 页)

3.2　结果分析

从测试数据及结果分析中可以看出本文提出的安全模型具有如下几个特点:

a)安全性能高。在安全有效性中,本文对虚拟机管理、事件通道、资源管理以及内存共享上的访问控制进行了安全性测试。测试结果显示:虚拟机管理控制符合 CW 的利益冲突集模型要求;资源管理和事件通道的控制能得到冲突集扩充定义的有效控制;对内存共享上的控制遵循本文 VBAC 模型的多级安全控制规则。

b)空间开销少。在访问控制模块中,安全级与冲突集类型在映射后存储于数组中,策略信息缓存数组大小由安全类型数量 M 和虚拟机数量 N 决定。假设每个数组元素所占空间大小为 1 Byte,那么策略信息缓存数组所占空间为 $M \times N$ Byte。在资源控制上,内存页的安全控制空间开销最大,为了减少空间的开销,本文采用了 Bit-map 的存储方式,大大地缩减了空间开销。

c)存在一定的时间性能损耗。从测试过程中可以看出,访问控制所带来的性能损耗主要是内存分配,因为每次系统启动都将分配大量的内存页,每一页内存的分配都将产生决策信息的读取和判定,从而导致了一定的时间损耗。

4　结束语

虚拟机是云计算的底层关键技术之一,但是目前存在很多对于虚拟机的攻击。针对虚拟机的内部安全隐患,本文提出了一种适用于虚拟机环境的访问控制模型,在 PCW 安全模型的基础上,引入了 BLP 多级安全模型,并对 BLP 进行了相应改进,本文提出的模型同时使用了 CW、PCW 和 BLP,具有安全性能高、空间开销小等优点。仿真实验证明本文所提模型有较强的可行性及安全性。

参考文献:

[1] 冯登国,张敏,张妍,等. 云计算安全研究[J]. 软件学报, 2011,22(1):71-83.

[2] WANG Zhi,JIANG Xu-xian. HyperSafe:a lightweight approach to provide lifetime hypervisor control-flow integrity[C]//Proc of IEEE Symposium on Security and Privacy. 2010:380-395.

[3] SALAUM M. Practical overview of a Xen covert channel[J]. Journal in Computer Virology,2010,6(4):317-328.

[4] LIU Qian,WNAG Guan-hai,WENG Chu-liang,*et al*. A mandatory access control framework in virtual machine system with respect to multi-level security Ⅱ: implementation [J]. China Communications, 2011,8(2):86-94.

[5] RANJITH P,PRIYA C,SHALINI K. On covert channels between virtual machines[J]. Journal in Computer Virology,2012,8(3):85-97.

[6] OKAMURA K,OYAMA Y. Load-based covert channels between Xen virtual machines[C]//Proc of the 25th Annual ACM Symposium on Applied Computing. Sierre: Association for Computing Machinery, 2010:173-180.

[7] WU Jing-zheng, DING Li-ping, WANG Yong-ji, *et al*. Identification and evaluation of sharing memory covert timing channel in Xen virtual machines[C]//Proc of the 4th IEEE International Conference on Cloud Computing. Los Alamitos, CA: IEEE Computer Society, 2011: 283-291.

[8] SAILER R, VALDEZ E, JAEGER T, *et al*. sHype: secure hypervisor approach to trusted virtualized systems, RC23511[R]. [S. l.]: IBM, 2005.

[9] 程戈,金海,邹德清,等. 基于动态联盟关系的中国墙模型研究[J]. 通信学报,2009,30(11):93-100.

[10] CHENG Ge,JIN Hai-jin,ZOU De-qing,*et al*. A prioritized Chinese wall model for managing the covert information flows in virtual machine systems[C]//Proc of the 9th International Conference for Young Computer Scientists. Los Alamitos:IEEE Computer Society,2008.

[11] 石磊,邹德清,金海. Xen 虚拟化技术[M]. 武汉:华中科技大学出版社,2009.

[12] FOLEY S N. Building Chinese Walls in standard UNIX™[J]. Computers & Security,1997,16(6):551-563.

[13] ZHAO Gan-sen,CHADWICK D W. On the modeling of Bell-LaPadula security policies using RBAC[C]//Proc of the 17th IEEE Workshops on Enabling Technologies:Infrastructure for Collaborative Enterprises. Piscataway,NJ:IEEE Press,2008:257-262.

(上接第 221 页)面可以对经过的攻击数据包进行免疫处理,同时可以对可疑的潜在攻击发起路径标志,使攻击源追踪算法具有了动态性和自适应性,节省资源。此外,该路径标志技术可以进一步对通过的数据包进行标记,为之后免疫入侵检测系统针对单一攻击培养生成免疫特征检测器提供了技术基础,为在攻击路径中处理攻击数据包,避免受害主机遭受大规律攻击数据包汇聚而造成的拒绝服务情况。

实验证明,本文提出的方案在收敛效率、误报率等方面均优于传统追踪算法。虽然需要免疫响应服务器的支持,但该服务器并非仅针对路径标志而放置,其主要功能是作为免疫节点在交换网络中处理攻击数据包。

参考文献:

[1] 彭沙沙,张红梅. 计算机网络安全分析研究[J]. 现代电子技术, 2012, 35(4):109 - 116.

[2] 王大伟. 基于免疫的入侵检测系统中检测器性能研究[D]. 哈尔滨:哈尔滨理工大学, 2010.

[3] 闫巧, 雷琼钰. IP 追踪新进展[J]. 小型微型计算机系统,2012,33(9):2027 - 2032.

[4] FORREST S,PERELSON A S,ALLEN L. Self-nonself discrimination in a computer[C]//Proc of IEEE Society Symposium on Research in Security and Privacy and Privacy. 1994: 202 - 212.

[5] 李订芳. 一种提高检测率的免疫优化模型[J]. 微电子学与计算机, 2007, 23(5): 195-197.

[6] GONZALEZ F, DASGUPTA D, NINO L F. A randomized real-valued negative selection algorithm[C]//Proc of the 2nd International Conference on Artificial Immune Systems. 2003: 261 - 272.

[7] TWYCROSS J. Stochastic and deterministic multiscale models for systems biology: an auxin-transport case study[J]. BMC Systems Biology, 2010, 12(9): 29-41.

[8] 闫巧,吴建平,江勇. 网络攻击源追踪技术的分类和展望[J]. 清华大学学报:自然科学版,2005,45(4):497-500.

[9] 张婵. 一种改进的 iTrace 技术的研究[J]. 科学技术与工程, 2007,12(7): 3013-3016.

[10] SAVAGE S, WETHERALL D, KARLIN A, *et al*. Practical network support for IP traceback[C]//Proc of ACM SIGCOMM Conference. [S. l.]:ACM Press,2000:118-128.

[11] 揭摄,孙乐昌. 一种新的非重复性包标记 IP 追踪方案[J]. 计算机工程, 2007, 33(10): 105-107.

[12] 杨海松,李津生. 分布开放式的入侵检测与响应框架——IDRA[J]. 计算机学报,2003,26(9):1177-1182.

[13] 张永铮,肖军,云晓春,等. DDoS 攻击检测和控制方法[J]. 软件学报,2012,23(8):2058-2072.

[14] 宋东辉, 李丽娟. 改进的概率包标记 IP 追踪方法[J]. 计算机工程,2011,37(4):147-149.