

基于 TTL 阈值分析的检测伪造数据包方法

李 旻^{1,2}, 汪 泉¹, 刘建中¹, 严小燕¹, 许高建¹, 汪 阳¹

(1. 安徽农业大学 计算机系, 安徽 合肥 230036;

2. 安徽大学 人工智能研究所, 安徽 合肥 230039)

摘 要:在基于 TTL 分析的检测技术基础上提出了一种基于异常 TTL 阈值分析的源地址伪造数据包检测方法, 综合主动和被动两种检测技术, 设计了一种有效的混合型源地址伪造数据包检测方法, 同时用 2500 个正常数据包和 2500 个 TTL 值异常的伪造数据包来模拟遭受攻击, 通过模拟实验结果对该方法的优缺点进行了分析, 最后给出了该方法的优缺点并提出了改进策略。

关键词:TTL 异常; TTL 分析; 阈值; 伪造数据包

中图分类号:TP393.01

文献标识码:A

文章编号:1673-629X(2008)03-0157-04

Source Address Spoofed Packets Detecting Approach Based on TTL Threshold Value Analysis

LI Yang^{1,2}, WANG Quan¹, LIU Jian-zhong¹, YAN Xiao-yan¹, XU Gao-jian¹, WANG Yang¹

(1. Computer Department of Anhui Agricultural University, Hefei 230036, China;

2. AI Institute, Anhui University, Hefei 230039, China)

Abstract: Introduces a source address spoofed packets detecting approach based on TTL threshold value is proposed in this paper. And designs an efficient composite source address spoofed packets detecting approach. At the same time, the advantages and disadvantages of this approach is analyzed by simulating experiment which contains 2500 normal packets and 2500 TTL anomaly packets, and the corresponding improvement scheme is suggested.

Key words: TTL anomaly; TTL analysis; threshold value; packet spoof

0 引言

进入二十一世纪以后,随着“互联网冬天”的结束,网络又进入了一个快速发展的时期,随之而来的就是黑客的攻击和入侵。根据 US-CERT(美国国家计算机网络应急响应中心)的统计,在 2006 年最后的三个月中,接到了近 2 万次的安全事故报告,这一数字已经接近了此前 12 个月报告的总和。攻击事件与日俱增,网络安全领域面临巨大挑战。

1 常见的伪造源地址攻击技术

(1)较早期的 SYN Flooding 攻击,即最典型的 DoS 攻击技术是基于 TCP 连接的“三次握手”实现:客户 A 向目标机 B 发出访问同步请求(SYN)数据包, B 收到

后回应一个请求确认(SYN/ACK)数据包,并调整自己的连接处于半开状态;设定等待时间上限并在内存队列中加入该请求;正常时, A 回应一个确认(ACK)数据包即完成连接过程。但若攻击者始终不发 ACK 数据包,而不停送出 SYN 数据包,最后 B 因耗尽资源而中断一切访问请求^[1]。

(2)Smurf 攻击,DDoS 攻击技术的一种。攻击者伪造源 IP 地址,先利用某个中介网络向其“广播站”送出一个源地址为目标机的请求回应数据包,“广播站”自动将这一请求传播到该网络所有主机处。于是它们都送出各自响应数据包,发送至目标机并大量消耗其资源。攻击者还可以同时向其它中介网络进行类似操作,于是更大量的数据包涌向受害机处。这里中介网络起着放大器的作用,故称“反射机”^[2]。

(3)TCP 连接伪造攻击。该攻击的原理是:假设主机 B 信任主机 A,攻击者首先使用 DoS 攻击主机 A 使之下线,然后向主机 B 发送源地址为 A 的伪造数据包,主机 B 会向地址 A 发送包含初始序列号的确认

收稿日期:2007-06-17

基金项目:安徽省教委科研项目(2003-KJ-115;2005-KJ-086);
安徽省重点科研计划资助项目(2007ZD-7021010)

作者简介:李 旻(1963-),男,安徽人,博士,副教授,研究方向为计算智能及其在计算机、网络工程、农业信息工程中的应用。

包,攻击者可以猜测该序列号,然后发送一个确认给 B,完成三次握手。此时攻击者拥有主机 A 对主机 B 的所有权限。该攻击最重要的步骤是攻击者必须能猜测到 B 发给 A 的确认包中的初始序列号。

2 TTL 的性质

2.1 TTL 简介

TTL,全称是 Time To Live,中文名为生存时间,它是 IP 报头中一个非常重要的参数;TTL 字段存在于 TCP/IP 层次模型的网络层的 IP 数据包中,该字段占 8 个 BIT,即最大值为 255,计数单位为跳。

IP 数据报结构如图 1 所示。

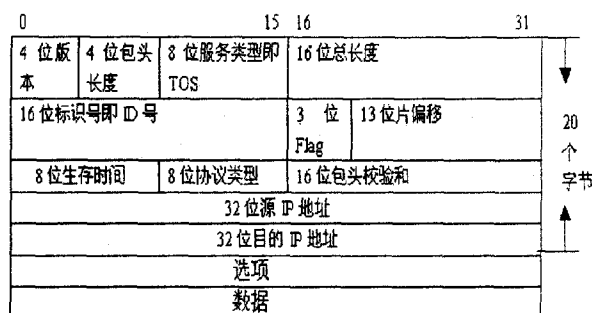


图 1 IP 数据报结构图

TTL 告诉网络中的路由器,数据包在网络中的时间是否太长而应被丢弃,TTL 的最初设想是确定一个时间范围,超过此时间就把包丢弃;由于数据包每经过一个路由器时,TTL 值都会至少被路由器减 1,所以 TTL 值通常表示包在被丢弃前还能最多经过的路由器个数。当 TTL 值为 0 时,路由器丢弃该数据包,并发送一个 ICMP 数据包给数据包的最初发送者^[3-6]。

2.2 TTL 值的可预测性

本机收到的 IP 数据包中的 TTL 字段与很多因素有关,例如 TTL 初始值,所经过的网络跳数等等,但是在一定程度上还是有规律可循的^[7]。

下面提出几个假设条件:

(1)当同一源主机发出的两个数据包经过相同的路由到达目标主机时,两个 IP 包的 TTL 值的减少量是一样的。

(2)在一个很短的时间间隔内发往同一目标的数据包通常选择同一条路径。

(3)相同任务条件下路由变化不频繁。

(4)路由即使变化,新路由的跳数和以前路由的跳数相比,相差不大。

在这三个假设中,假设(1)是显而易见的。假设(2)~(4)是在大多数情况下是可以成立的,所以说 TTL 值具有可预测性,至少是在一定范围内可以预测。

3 基于 TTL 分析的检测技术介绍

伪造源地址数据包攻击是近年来常见的网络攻击手段,包括 IP 欺骗,DoS 攻击等等都是建立在此基础上的;由于源地址伪造的隐蔽性很强,要防范是较为困难的。现在使用的防范手段主要是对数据包进行加密传输,或者用建立 VPN 隧道的方法来避免攻击。但是这些技术无一例外地加大了网络传输的数据量,影响了网络传输的速率^[8]。

在分析了大量的数据包实况后,提出了一种较为节省网络资源的简便易行的检测技术:基于 TTL 比较的源地址伪造数据包检测技术。

在上述 TTL 性质的基础上,把检测技术分为主动与被动两种。

(1)主动检测技术:当主机接收到一个数据包时,可以向源地址发送与该数据包采用相同协议的探测数据包。如果回应数据包 TTL 值与要检测的数据包 TTL 值相等,说明该数据包没有经过伪造。例如原数据包 TTL 值为 55,使用 ICMP 回送请求数据包探测,得到的回应数据包 TTL 值为 119,因为 $64 - 55 = 9$,且 $128 - 119 = 9$,两者相等,则可以认为 13 就是源主机和本机之间的路由跳数。则该数据包不是源地址伪造的。该技术的不足之处在于需要发送大量的探测数据包。

(2)被动检测技术:根据前面所做出的假设,在一段时间内,特定协议的数据包 TTL 值应该是稳定的,因此可以通过比较收到的数据包 TTL 值与预期的理论值的差异是否在一个可以接收的范围内,若超出了该范围,则认为该数据包可能是经过源地址伪造的。

4 混合型伪造源地址数据包检测技术

综合上述主动和被动两种检测技术,设计一种有效的混合型源地址伪造数据包检测方法。该方法的主要思路如下:

首先建立一个 TTL 期望值表,该表的关键字为 IP 地址,对应的值为来自该 IP 地址特定协议类型数据包的 TTL 期望值以及一个记录收到的该类型正常数据包的计数器。该表初始为空。

当收到一个数据包时:

(1)首先从 TTL 期望值表中查找该数据包的源 IP 地址所对应的 TTL 期望值,若表中没有该项,则转(2),否则转(3)。

(2)使用主动探测技术,发送该协议类型的探测数据包,比较回应数据包 TTL 值与该数据包 TTL 值。若相同,则在 TTL 期望值表中添加一项关键字为该数据包源 IP 与协议类型,值为该数据包 TTL 值的记录,

同时针对该记录的计数器加1;否则将该数据包标记为该地址伪造数据包,并写入日志文件,然后转(4)。

(3)比较该数据包 TTL 值与表项对应的值是否在正常范围内,若是,则认为该数据包正常,将表项的 TTL 期望值更新为(到达数据包的 TTL 值+期望值 \times 计数器值)/(计数器值+1),即对收到的所有数据包的 TTL 值求平均值,同时计数器值增加1;否则,将该数据包标记为源地址伪造数据包,并写入日志文件。

(4)继续监听网络,每收到一个新的数据包则转到(1)开始处理。

从上面的检测步骤可以看出,它是主动检测与被动检测的结合。其检测精度的关键就是收到的数据包 TTL 值与预期的理论值的差异是否在一个可以接收的范围内,也就是要确定一个阈值,来判定是否超出了合法范围。

这个阈值如何确定,确定为多少才是合适的?下面就将上面的步骤进行模拟实现来观察不同阈值对检测精度的影响,从而确定合理的阈值。

5 实现与结果分析

在 Windows XP 平台下,用 VB 生成 5000 个随机 TTL 值来模拟收到的混杂有伪造数据包的数据包组如图 2 所示。

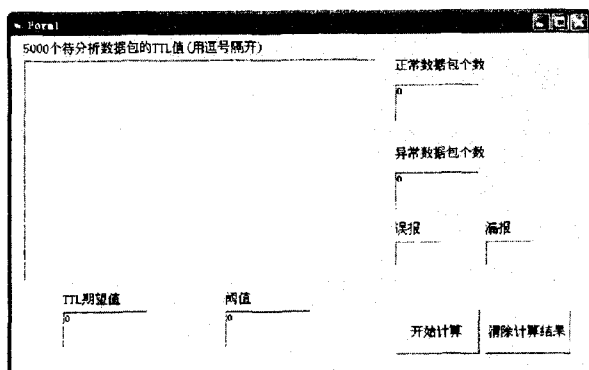


图2 TTL 数据分析器

现在用 2500 个正常数据包和 2500 个 TTL 值异常的伪造数据包来模拟遭受攻击的情况。这时,主机 A 与主机 B 正常通信。外界攻击者发送将源地址伪造成主机 A 的数据包,其 TTL 值随机生成,文中模拟的就是主机 B 检测这些数据包的过程如图 3 所示。

由于真实情况下,网络间的路由是可能发生变化的,但是与原路由相比,新的路由跳数与其相差不大,所以,文中设计了 2 种情况来分别考虑:一种是检测期间 A 与 B 之间路由未发生变化,设正常情况下 TTL 值为 59,一种是检测期间 A 与 B 之间路由发生变化,分别变化为 60(+1),61(+2),62(+3)。

在网络实际环境中,未知的源地址 TTL 期望值是用主动检测的方法发送探测报文来测试,这里用直接给出的方式代替。检测结果如图 4~7 所示。

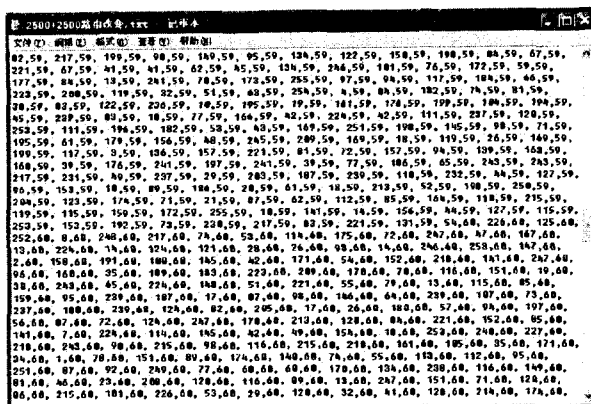


图3 5000 个 IP 包的 TTL 值(路由改变+1)



图4 路由未发生变化(+0)情况下的条状图

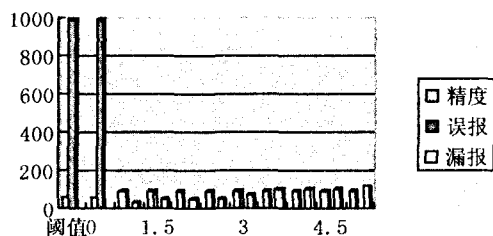


图5 路由发生(+1)变化情况下的条状图

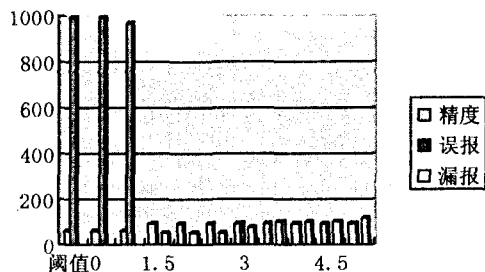


图6 路由发生(+2)变化情况下的条状图

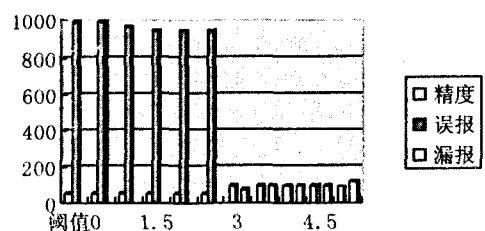


图7 路由发生(+3)变化情况下的条状图

可以看到,该方法的精度和效率均比较高,但会出现漏报和误报现象;阈值设置过低时,路由变化一旦超出阈值范围,会产生误报现象;设置过高,则会产生漏

报现象。在实际应用中,可以根据网络状况相应地设置阈值,如果网络的路由变化相对频繁,则可以将阈值设置稍高;反之,则设置稍低。而在路由不可能变化的情况下,将阈值设为零,可最大程度地提高精度。

将各个阈值下的精度做个平均值(见表 1)。从平均值表上可以看出,当阈值较小时候由于误报较多影响了精度,而阈值高了,精度受漏报增多的影响也不高。当阈值取 3 的时候,漏报与误报的影响达到一个平衡值,取得一个相对较好的精度。在实际应用的时候,可以先对欲保护的主机的日常流量进行一段时间的监视(如 24 小时),然后将其经常访问的若干节点的 TTL 值的变动进行分析,用上述方法进行各个阈值下的精度比对,从而确定最佳阈值,并可定期进行调整。

表 1 所取阈值与精度关系

阈值	精度(+0)	精度(+1)	精度(+2)	精度(+3)	精度(avg)
0	99.68%	60.32%	60.32%	60.32%	70.16%
0.5	99.68%	60.32%	60.32%	60.32%	70.16%
1	98.72%	98.72%	61.28%	61.28%	80.00%
1.5	97.84%	97.84%	97.84%	62.16%	88.92%
2	97.84%	97.84%	97.84%	62.16%	88.92%
2.5	97.84%	97.84%	97.84%	62.16%	88.92%
3	96.92%	96.92%	96.92%	96.92%	96.92%
3.5	95.88%	95.88%	95.88%	95.88%	95.88%
4	95.88%	95.88%	95.88%	95.88%	95.88%
4.5	95.88%	95.88%	95.88%	95.88%	95.88%
5	95.16%	95.16%	95.16%	95.16%	95.16%

6 优缺点和改进方向

在实际应用过程中,当目标主机是存在于同一个子网中的主机时,可以对这个子网中任意一台主机发

出主动检测数据包,用其返回的 TTL 值作为该子网所有主机的 TTL 期望值,这样可以大幅减少主动检测数据包的发送量,提高检测效率。但是当发送伪造数据包的主机与目标之间的条数若恰好等于源主机与目标机之间的条数时,该方法就会失效。如何克服这一弱点,是今后实际应用研究工作的重点。

今后将结合其他的伪造报文检测手段,以期弥补该方法的不足,使之精度更高,误报和漏报率更低。

参考文献:

- [1] Bellovin S. Security Problems in the TCP/IP Protocol Suite [C]//ACM SIGCOMM Computer Communications Review. New York: ACM Press, 1989: 32-48.
- [2] 母军臣 朱长江. 基于概率 TTL 终值的 IP 欺骗 DDOS 防御策略[J]. 河南大学学报: 自然科学版, 2006(4): 96-99.
- [3] Kurose J F, Ross K W. Computer Networking: A top-down Approach Featuring the Internet[M]. 北京: 机械工业出版社, 2006.
- [4] Doyle J. CCIE #1919. Routing TCP/IP Volume 1[M]. 北京: 人民邮电出版社, 2003.
- [5] Stevens W R. TCP/IP 详解. 卷一. 协议[M]. 范建华等译. 北京: 机械工业出版社, 2004.
- [6] Tanenbaum A S. 计算机网络[M]. 第 4 版. 潘爱民译. 北京: 清华大学出版社, 2004.
- [7] 包怀忠. 马季. IP 网络路由追踪技术研究[J]. 微电子学与计算机, 2004(8): 59-62.
- [8] 佚名. IP 欺骗原理精解和防范手段综述[EB/OL]. 2005-04-03. <http://www.gipsky.com/modules/wfsection/article.php?articleid=5>.

(上接第 147 页)

UML 的静态结构类图,通过 UML 对 FIFO 的建模并转换成 SystemC 语言,来说明 UML 与 SystemC 结合进行 SoC 建模的特性。

5 结论

通过 UML profile for SoC,在传统的设计流程中结合 UML 的可视化建模能力,将不可见的代码对系统进行功能描述建模的能力,提升到了一种以可视化的图形方式对系统的功能进行建模,不仅有利于系统设计人员与用户之间的沟通,而且也方便了系统设计人员之间的交流沟通。UML profile for SoC 的扩展能力,在系统级描述语言(比如 SystemC)进行系统描述的过程中,通过静态结构类图对 SoC 的系统结构进行建模,并方便地将 UML 的图形化描述方式转换成系统级语言描述。在模型驱动开发(MDA)这种新的开发模式下,将更有利于系统的模型建立、系统的文档维护,提

高系统的开发效率。

参考文献:

- [1] IEEE-SA Standards Board. IEEE Standard SystemC Language Reference Manual[EB/OL]. 2005-12-06. <http://www.systemc.org>.
- [2] Riccobene E, Scandurra P, Rosti A, et al. A SoC Design Methodology Involving a UML 2.0 Profile for SystemC[M]. [s.l.]: IEEE, 2005.
- [3] OMG. UML Profile for System on a Chip (SoC)[EB/OL]. 2006-08-01. <http://www.omg.org>.
- [4] 陈燕. 基于 UML 的嵌入式系统系统级设计方法研究[D]. 上海: 复旦大学, 2005.
- [5] 石柯. 基于 UML 和 SystemC 的嵌入式系统集成开发方法的研究[J]. 高技术通讯, 2003(11): 44-47.
- [6] 王建新, 姚放吾. 基于 UML 的软硬件协同设计方法[J]. 计算机技术与发展, 2006, 16(1): 96-98.