一种基于协议分析的入侵检测模型

燕振刚[1,2],罗进文[1]
(1.兰州交通大学 电子与信息工程学院,甘肃 兰州 730070;
2.甘肃农业大学 信息科学技术学院,甘肃 兰州 730070)

摘 要:对于入侵检测系统来说,选择好的入侵检测方法有利于提高检测效率,传统的入侵检测系统由于计算量大、漏报率和误报率高,已经不适应于当前网络系统的需求。协议分析是网络入侵检测中的一种关键技术,基于这种思想,介绍了协议分析的内容、过程、入侵特征的提取及协议分析在入侵检测中的应用,主要实现了对IP数据包内容分析,同时提出了一种与传统模式匹配算法相结合的可行入侵检测模型。经分析,该检测模型比传统的检测模型有着明显的优势。

关键词:入侵检测;协议分析;模型
中图分类号:TP393.08 **文献标识码**:A **文章编号**:1673-629X(2008)11-0146-03

An Intrusion Detection Model Based on Protocol Analysis

YAN Zhen-gang[1,2], LUO Jin-wen[1]
(1. School of Electronic and Information Engineering, Lanzhou Jiaotong University, Lanzhou 730070, China;
2. Institute of Information & Science Technology, Gansu Agricultural University, Lanzhou 730070, China)

Abstract: To intrusion detection system, it makes improving efficiency of intrusion detection by choosing better method of intrusion detection, traditional intrusion detection system because of large amount of calculation, the high rate of omissions and misstatements has not already adapted to the needs of the current network system. Protocol analysis is a kind of key technology for network intrusion detection. The paper which based on that idea will introduce content, process of protocol analysis, extraction of intrusion feature and application in intrusion detection, mainly implement IP packet analysis, and point out a feasible intrusion detection model which connect with pattern matching algorithm. Compared with other model, the model has obvious advantage by analysing.

Key words: intrusion detection; protocol analysis; model

0 引 言

对于入侵检测系统来说,要解决的核心问题是对高速率、大流量的网络进行无漏包监听和实时准确分析处理。在入侵检测系统中,传统模式匹配入侵检测技术把收集到的数据与特征库中特征码采用单模式或多模式进行匹配,主要缺点为计算量大、漏报率和误报率高,已经不适应于当前网络系统的需求。为此,人们提出了协议分析技术,它利用网络协议的高度规则性快速探测攻击的存在,是新一代IDS探测攻击手法的主要技术。

通过对协议内容和协议分析过程进一步分析,提出了一种与传统模式匹配算法相结合的可行入侵检测模型。分析其性能,表明该模型拥有检测速度快、漏报率低等特点。

收稿日期:2008-02-17
基金项目:教育部春辉计划科研项目(20567)
作者简介:燕振刚(1978-),男,甘肃定西人,硕士研究生,从事信息安全教育与研究;罗进文,教授,研究方向为无线通信和信息安全。

1 协议分析内容及基于协议分析的入侵检测模型

1.1 协议分析内容

协议分析的任务就是根据TCP/IP协议族的内容,对TCP/IP协议族中各种数据包的包头信息、数据、校验和以及网络会话的状态信息等进行分析,并实现数据的重组和还原。这就意味着校验许多层的协议字段不合法或是可疑的值,包括无效字段、非法值、不寻常的缺省、不合适的选择、没有顺序的顺序号、顺序号间隔、顺序号重叠、校验和及循环冗余校验修正值等的使用[1]。

1.1.1 协议类型分析

数据包内的信息必须遵循TCP/IP协议规范,该协议在标志和选项方面的内容非常丰富,它们监视和设置网络的当前状态。主要对协议的以下字段进行检

测:

(1)在 IP 数据包中,首先检查 IP 报头的版本号。如果版本号为 4,则检查报头长、报头长度、服务类型、数据报总长度、标识、标志、分割偏移、存活时间、上一层协议、校验和、源地址、目的地址及 IP 选项等。如果版本号为 6,则检查通信类型、流标记、有效载荷长度、下一个头部、跳数限制、源地址、目的地址等。

(2)在 TCP 数据包中,检查报头长、IP 源地址、IP 目的地址、TCP 源端口、TCP 目的端口及 TCP 标志等。

(3)在 ICMP 分组中,检查 IP 源地址、IP 目的地址、ICMP 类型字段、ICMP 标识及 ICMP 序列号等。

(4)在 UDP 数据包中,检查 IP 源地址、IP 目的地址、源端口号及目的端口号等。

(5)在 HTTP 报文中,可对报文的内容、URI 内容、请求报文名称、报文版本信息、报文响应代码等内容进行检测。

1.1.2　入侵特征提取

入侵特征提取如下:

(1)非正常活动分析。任何活动应遵守协议规范建立及关闭一个 TCP 连接。

(2)伪造的活动或信息分析。伪造的数据交换通常说明了一个问题或一个恶意活动。无数据交换的 TCP 会话、ICMP 分片包、带有数据的初始 TCP 三次握手包等属于这类。

(3)请求/响应比率分析。ICMP 和 ARP 定义了具有一对一的请求/响应原语,即一个请求对应一个响应。如果违反了该规定,则说明存在问题,如 Smurf 攻击、Dos 攻击。

(4)在 UDP 数据报中,检查 IP 源地址、IP 目的地址、源端口、目的端口。

(5)数据流的重组分析。如对分散在多个数据包中的 HTTP 请求进行分析处理。

(6)异常流量分析。在正常情况下,有一些类型的流量不应该频繁出现。例如"ICMP 端口和目的地址不可达"信息或带有 RST 标志的 TCP 包一般是很少出现的。这类信息的峰值或高比率出现指出了网络上存在着问题,需要进一步分析。

1.2　基于协议分析的入侵检测模型

1.2.1　模型的体系结构

基于协议分析的入侵检测模型如图 1 所示。

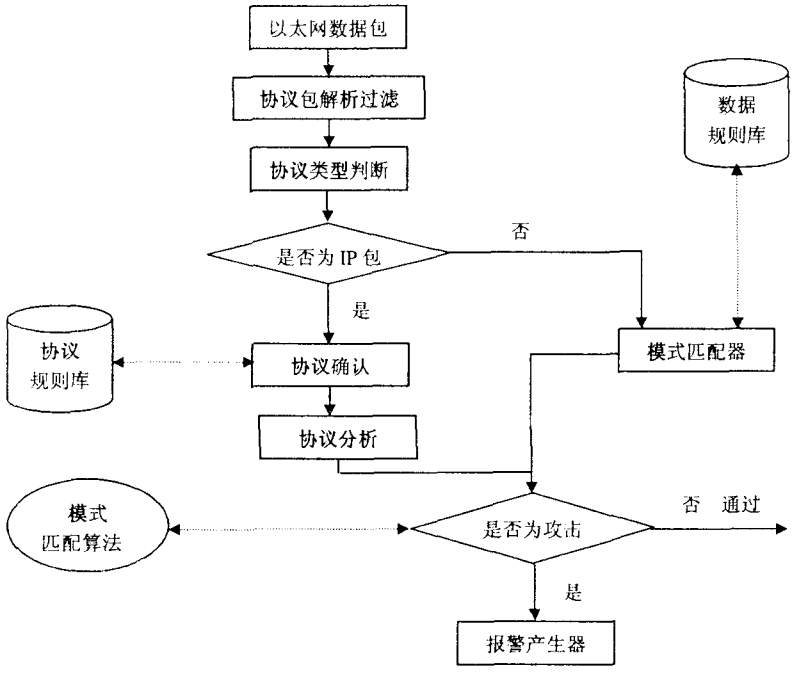

图 1　基于协议分析的入侵检测模型

1.2.2　协议分析过程描述

目前采用的网络协议有多种,但对局域网来说,用的最广的就是 TCP/IP 协议,在 RFC 的 0791[2] 和 0793[3] 文档中,分别定义了 TCP 数据包和 IP 数据包的格式。由于这种格式定义只与协议相关,与网络的结构、类型无关,所以协议分析具有很广泛的适用性。下面对图 1 进行说明。

利用开发包 Libpcap 捕获网络数据包后,在 Libpcap 使用的 BPF 机制中,有一个网络数据包过滤模块,使用它可以过滤各种各样的网络数据包。在捕获到了数据包后,就可以对网络数据包进行分析[4]。捕获到数据包后,首先判断是否为 IP 包(IPv4),如果不是,则利用模式匹配算法判断数据是否满足数据规则库的原则,如果满足,则根据以太网的帧结构的定义,如果为 0x0800,表示协议为 IP 协议。在分析 IP 协议时,根据协议类型的值判断传输层协议,在 IP 数据包的格式定义中,第 10 个字节为第四层协议标识,如:TCP 为 06,UDP 为 17,ICMP 为 01 等而 TCP 数据包的第 3、第 4 个字节为应用层协议标识(端口号)。如 80 为 HTTP 协议,21 为 FTP 协议,2 为 TELNET 协议等。其他协议类型的分析以此类推。

由此可以看出,协议分析的分析过程就是利用网络协议的高度规则性来分析判断协议类型,确定协议类型后,再采用模式匹配算法来检测攻击,因此大大减少了计算量,提高了算法的效率。

1.2.3　协议分析的算法实现

利用 Libpcap 提供的库函数进行数据采集,通过

函数 pcap_open_live 设置网卡的状态为混杂模式，利用 Libpcap 可以提供直接从链路层捕获数据包的功能，并可以设置数据包的过滤器以捕获指定的数据[5]。

在协议分析过程中，主要实现对 IP 数据包内容的分析，具体实现如下：

```
void ip_protocol_packet (struct pcap_pkthdr * packet_head, u_
char * packet_content)
//packet_head 表示捕获到的数据包头信息
// packet_content 表示捕获到的数据包内容
{ struct ip_head * ip_protocol //ip 协议变量，根据 ip 协议格式可
以自己定义 ip_head
ip_protocol = (struct ip_head * )(packet_content + 14)
if (ip_source_add = = ip_local_add)//判断源 ip 是否为本机 ip
{ if (ip_ destination_ add&ip_ add_ broakcast = = ip_ add_ broak-
cast)//判断 smurf 和 land 攻击
.....//smurf 攻击分析程序
else
if(ip_ destination_ add = = ip_ source_ add)
.....//land 攻击程序
}
else
{switch(ip_ protocol - >ip_ protocol)
{ case 6:
//tcp 协议内容分析
tcp_ protocol = (struct tcp_ head * )(packet_ content + 14 + 20);
source_ port = ntohs(tcp_ protocol - > tcp_ source_ port);//获取源
端口
destination_ port = ntohs ( tcp_ protocol - > tcp_ destination_
port);//获取目的端口
.....
switch(destination_ port)
{ case 80:
.....//HTTP 协议分析程序
case 21:
.....//FTP 协议分析程序
case 23:
.....//TELNET 协议分析程序
case 25:
.....//SMTP 协议分析程序
}
case 17:
.....//UDP 协议内容分析
case 01:
.....//ICMP 协议内容分析
}}}
```

在算法的实现上不一定要局限于上述方法，可以针对不同的协议来定义不同的函数，而在主函数中采用循环捕获网络数据包函数 pcap_loop( )来捕获数据包，然后层层去调用协议分析函数。

2 协议分析在入侵检测中的应用

网络攻击行为可以分成发生于头部的攻击和发生于数据部分的攻击[1]。对于发生于头部的攻击，只需分析单个或多个数据包的头部信息即可检测出是否有攻击发生。例如 Land 攻击，Land 攻击利用了 TCP 连接建立的三次握手过程，通过向一个目标计算机发送一个 TCP SYN 报文(连接建立请求报文)而完成对目标计算机的攻击。与正常的 TCP SYN 报文不同的是，Land 攻击报文的源 IP 地址和目的 IP 地址是相同的，都是目标计算机的 IP 地址。这样目标计算机接收到这个 SYN 报文后，就会向该报文的源地址发送一个 ACK 报文，并建立一个 TCP 连接控制结构(TCB)，而该报文的源地址就是自己，因此，这个 ACK 报文就发给了自己。这样如果攻击者发送了足够多的 SYN 报文，则目标计算机的 TCB 可能会耗尽，最终不能正常服务。在此类攻击中，只需判断数据包的头信息传输前后是否一致。

对于发生于数据部分的攻击，则需要用模式匹配算法进行分析，将获得的特征值与特征库内容比较，如与特征吻合，则访问被判断为攻击行为，产生报警，检测效果的好坏取决于模式匹配算法。

3 模型特性分析

协议分析技术利用协议规则寻找攻击，极大地减少所需的计算量，即便在高负载的网络上也可以完全探测出各种攻击，并对其进行更详细的分析而不会丢包。该模型的特征在于：

1)该模型中结合模式匹配技术应用到入侵检测中，可以提高检测效率。

2)针对不同的应用协议形成协议规则库，提高检测的准确性。

3)在通信中如果出现 IP 碎片，可以重组数据报，然后进行协议分析探测碎片攻击。

4)系统资源的极低消耗。

5)由于单纯采用模式匹配技术误报率高，而采用该模型误报率大大降低。

4 结束语

提出了一种基于协议分析的入侵检测模型，通过分析其原理，表明使用协议分析方法和模式匹配相结合的入侵检测系统比单纯使用模式匹配的方法在性能上有很大提高。在以后的工作中，将继续完善该检测系统的功能，如入侵特征规则库动态建立、入侵特征的

(下转第 155 页)

较强的抵抗图像缩放攻击和一定的剪切攻击等几何攻击的能力，尤其是抗 JPEG 攻击的能力很强，能够完全抵御 JPEG 攻击，具有较好的实用性。

表 1　对嵌入水印后的图像进行几种常见的攻击后的检测结果

攻击方法		PSNR(dB)	NC
JPEG 压缩	保持品质因子 90	39.76	1
	保持品质因子 60	34.52	1
	保持品质因子 50	32.60	0.99
	保持品质因子 30	29.54	0.93
	保持品质因子 10	24.68	0.85
加噪声	强度为 0.01 的高斯噪声	23.38	0.98
	强度为 0.01 的乘性噪声	30.17	1
	强度为 0.01 的椒盐噪声	28.71	1
	强度为 0.05 的椒盐噪声	20.88	0.83
平滑滤波	均值滤波	25.7662	1
	中值滤波	30.65	1
	高斯低通滤波	35.89	1
	维纳滤波	34.08	1
图像缩放	缩小 4 倍	30.54	1
马赛克效果	2×2 块	29.16	1
	3×3 块	26.77	0.94
锐化	锐化	31.25	1
剪切	左上剪去 1/4	14.27	0.86

参考文献：

[1] Lee Chang - hsing, Lee Yeuan - Kuen. An adaptive digital image watermarking technique for copyright protection[J]. IEEE Trans on Consumer Electronics, 1999, 45(4): 1005 - 1015.

[2] Podilchuk C I, Zeng W. Image - adaptive watermarking using visual models[J]. IEEE Journal on Selected Areas in Communications, 1998, 16(4): 525 - 539.

[3] Shapiro J M. Embedded image coding using zerotrees of wavelet coefficients[J]. IEEE Trans Singal Processing, 1993 (41): 3445 - 3462.

[4] Ntalian K S, Doulam A D, Doulam N D. An automatic scheme for stereoscopic video object based watermarking using qualified significant wavelet trees[J]. Image Processing, 2002(3): 501 - 504.

[5] Innous H, Miyazaki A, Yamamoto A, et al. A digital watermark based on the wavelet transform and its robustness on image compression[C]// Proceedings of the IEEE International Conference on Image Processing, ICIP'98. Chicago, USA: [s. n.], 1998.

[6] Cox I, Kilian J, Leighton T, et al. Secure spread spectrum watermarking for multimedia[J]. IEEE Transactions on Image Processing, 1997, 6(12): 1673 - 1687.

[7] 孙圣和，陆哲明，牛夏牧，等. 数字水印技术及应用[M]. 北京：科学出版社，2004.

(上接第 148 页)

快速提取、模式算法改进，力争开发出基于协议分析的通用入侵检测系统。

参考文献：

[1] 杜建国. 协议分析和命令解析在入侵检测中的应用[J]. 计算机工程与应用，2004(18): 2 - 3.

[2] Postel J. Internet Protocol DARPA Internet Program Protocol Specification [S/OL]. 1981. http://www.ietf.org/rfc/rfc0791.txt? number=791.

[3] Postel J. Transmission Control Protocol DARPA Internet Program Protocol Specification[S/OL]. 1981. http://www.ietf.org/rfc/rfc0793.txt? number=793.

[4] 刘文涛. 网络安全开发包详解[M]. 北京：电子工业出版社. 2005.

[5] 田　伟. 基于协议分析的网络入侵检测系统研究[D]. 南京：南京信息工程大学，2007.

(上接第 151 页)

[6] Lee. A Data Mining Framework for Constructing Features and Models for Intrusion Detection System[D]. USA: Columbia University, 1999.

[7] Kumar S, Spafford E H. An Application of Patter Matching in Intrusion Detection[R]. USA: Department of Computer Science, Purdue University, 1994.

[8] Doak J. Intrusion Detection: The Application of a Feature Selection - A Comparison of Algorithms and the Application of Wide Area Network Analyzer[R]. USA: Department of Computer Science, University of California, 1992.

[9] Lee W, Stroifo S J. Data mining approaches for intrusion detection[C]//Proc of the 7th USENIX Security Symposium. San Antonio, TX: [s. n.], 1998.

[10] Joshi M, Karypis G. A Universal Formulation of Sequential Patterns[R]. USA: Department of Computer Science, University of Minnesota, 1999.

[11] Bace R G. Intrusion Detection[M]. USA: Macmillan Technical Publishing, 1999.

[12] Mannila H, Tovonen H. Discovering generalized episode using minimal occurrences[C]//Proc. of the 2nd Intl. Conf. on Knowledge Discovery in Database and Data Mining. Portland, Oregon: [s. n.], 1996.