面向入侵检测的网络安全监测实现模型

龚俭 董庆 陆晟

（东南大学计算机系 南京 210096）

摘 要：本文提出了一种面向入侵监测的网络安全监测模型，它由数据采集、决策分析和分析机三个独立的部分以层次方式构成，能够对已知的网络入侵方式进行有效地实时监测。文章给出了基于安全分析机概念的安全知识表达方法，并对扫描(scan)，teardrop，land 等常见攻击方式进行了特征刻划。此外，论文还对安全监测系统设计中应当考虑的问题，如报警问题，进行了讨论。

关键词：网络安全、安全监测、入侵检测、扫描、teardrop、land、报警机制

1.引言

网络安全监测系统是保障网络正常运行的重要工具，它与网络运行管理系统相结合，可有效地监察网络用户的使用行为。网络安全监测系统的基本功能包括检测出正在发生的攻击活动；发现(track)攻击活动的范围和后果；诊断并发现攻击者的入侵方式和入侵地点，并给出解决建议；以及收集并记录入侵的活动证据。

网络安全监测系统可分为异常检测(anomaly detection)和入侵检测(misuse detection 或 rule-based)两类，前者通过采集和统计发现网络或系统中的异常行为，向管理员提出警告；后者通过对采集的信息按已知的知识进行分析，发现正在发生和已经发生的入侵行为。异常检测系统通常具有较强的检测能力，可发现新的安全问题。但是由于异常行为是相对而言的，因此这类系统存在误报问题，而且配置和管理比较复杂，需要较高的技术水平。入侵检测系统依赖所谓的攻击知识数据库，因此有较高的处理效率，管理简单（类似于现有的计算机病毒检测程序），但这类系统要求攻击知识数据库必须及时更新，而且不能处理未知的问题或现有问题的变形[1][2]。

网络安全监测系统可以根据数据采集方式的不同分为基于网络和基于主机两类[3][4]，而它们在处理上又可以分别采用实时处理和批处理两种不同的方式。一般而言，基于网络的安全监测系统具有较好的实时性，获得的信息也更全面，反应更灵敏，从而在系统的设计和实现难度上也更大一些。所以，面向网络的实时安全监测系统是目前研究的一个热点。

本文提出了一个面向实时入侵监测的网络安全监测实现模型，其主要特点是采用了层次化的集成框架，具有良好的适应性和可扩充性，这一点对 CERT 每年要报告 250 个左右的新漏洞这一事实来说是十分重要的[10]。

2. 系统的模型框架

这个实现模型将入侵检测分为三层来处理，分别是数据采集层，分析决策层和分析机层(见图 1)。

数据采集层的主要功能是获取分析所需的数据。因为系统是基于网络的，所以决定了

数据采集的方法是从网络上直接抓取数据包。数据采集部分的另一个任务是把获取的数据转换成标准形式，准备用于下一层分析。抓取的数据主要是 TCP/IP 协议的报头信息。

分析决策层是对数据采集层送来的数据进行初次分析。当有一个报文由采集部分送来，分析决策层就匹配各个分析机的启动条件，启动满足条件的分析机。因此它具有协调各个分析机工作的任务，是本系统的主控部分。

分析机层是一系列针对某种攻击设计的一组分析程序。分析决策层进行上下文无关分析，而分析机层负责具体的，上下文有关的分析工作。每个分析机可以根据需要拥有各自的存储空间，存储分析所需的历史数据。本系统目前实现了三个分析机：扫描分析机，teardrop 分析机，和 land 分析机。分析机层的另一个重要的部分是报警。

系统的工作方式为：位于第一和第二层的各进程始终处于工作状态，而位于第三层的各分析机组件则处于睡眠状态，这些组件由系统设在第二层主控部分根据不同的条件分别来启动。各个分析机自动完成分析工作，然后重新进入睡眠状态。

层次间的关联关系通过自定义的数据接口标准实现，以保证良好的可扩展性。在实现中这个特性主要体现在第二层和第三层。

把分析部分细分为分析决策层和分析机层的主要原因是基于通常攻击有上下文有关和上下文无关两种特征这个事实，而对这两者的特征分析和处理在难度方面有着很大的差别。由此带来的好处在于：一方面缩小上下文有关分析的范围，另一个方面是考虑到将来的分析功能扩充。

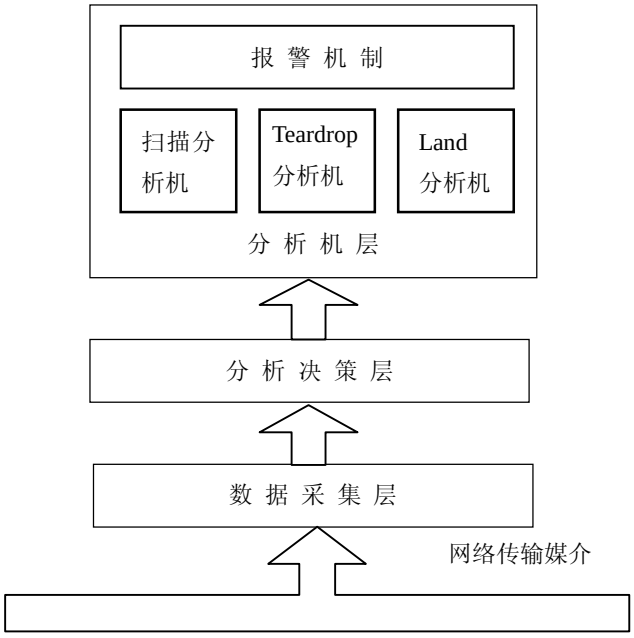

图 1　安全监测系统的模型框架

3. 监测数据的采集

监测数据的采集由数据采集层实现，包括数据采集和预处理两部分的功能，前者使用 tcpdump 工具获取网上报文，而后者对获取的报文进行预处理，使之具有符合定义的标准形式。

tcpdump 作为一个单独的子进程和分析父进程并行执行,在两个进程之间通过管道进行通信。tcpdump 的执行结果采用二进制形式输出，输出被重定向到管道中，因为二进制形式比可读的形式更便于程序分析。在管道的另一端是是父进程，它从管道中读取数据并进行有关的分析和处理。

预处理部分将 tcpdump 以二进制方式输出的数据转化成统一的内部格式，包括 IP 地址、端口、协议、状态、时标、长度、偏移等内容，并滤去与分析无关的信息[9]。

4. 上下文无关的安全分析

模型的第二层是分析决策层，它是系统的总控部分，完成对报文 What-If 方式的上下文无关的分析。分析依据是各种攻击的启动条件，通常也就是攻击发生的前提条件。当有一个报文和某个启动条件相匹配时，就启动分析机层中相应的进程。由此可看出，分析决策层实际上可被看作为一个规则库，每个启动条件对应规则库一条或一组规则。在本系统中，规则采用逻辑表达式（见表 1）的方法实现，其优点在于它缩小了下一步分析的范围，这个范围越精确，上下文有关分析工作的负担就越小。

这种规则库的模型还有利于系统功能的扩充，即可以看作是在规则库中增加规则和相应的动作。

表 1　系统分析决策层的例子

规则	规则匹配后的动作
（源地址＝目的地址）＆＆〔目的端口＝源端口〕	启动 Land 分析机
报文是一个 TCP 请求连接报文	启动扫描分析机
报文是一个分段的报文	启动 Teardrop 分析机

5. 上下文相关的安全分析

所谓上下文相关的安全分析是指对当前安全事件的分析要与过去所了解的相关历史联系起来，从而有可能作出比 What-If 方式更准确的判断。在本文所建议的实现模型中，上下文相关的安全分析分别由不同的分析机完成。从功能上讲，这些分析机是独立的，可以单独设计，只要遵从监测系统定义的统一数据格式，就可以通过在分析决策层增加规则和相应动作的方法，方便地加入监测系统。在一般情况下，分析机之间没有关联，可以并行工作。

准确地选择攻击特征并对其进行描述是分析机设计最重要的环节。一方面，特征的选择是否适当主要是看漏报（false positives）和误报（false negatives）的数量是否在允许的范围内，要在效率与准确性之间有一个平衡。其次，选择特征描述还要依据具体应用，通常一个好的特征描述应该比较容易转化为算法语言。

5.1 扫描分析机

扫描攻击是目前困扰我国互联网络最大问题之一。当黑客对要攻击的网络一无所知时，首先会用扫描器其进行扫描，以便获得被攻击网络的有关信息。因此，扫描攻击的发生，常常是更严重入侵的前奏。

从直观上看，扫描攻击的一般特征是：***在较短的时间里，攻击者从同一个源地址，试图连接许多的目的端口，通常包括一些不处于监听状态的端口。***[5] [8]经过效率和准确度的折衷，选择适合于算法的扫描器特征描述如下：

“某一源地址扫描了至少 COUNT 个端口；每 REFRESH-RATE 时间减少一个定值的 COUNT。”

当有一端发起连接请求时，该端会发送一个有 SYN 标志的 TCP 报文，扫描分析机据此做一次计数。在一段时间内，统计请求连接到一台主机的不同端口的这类报文的数量，并定时刷新

统计结果。该描述比较简单，容易转化成算法语言，适合对效率要求较高的实时系统。特征中的两个量 COUNT(单位时间内允许的连接请求个数)，和 REFRESH-RATE(单位刷新时间)是可配置的，可以视网上的情况而定，以达到较高的准确度。

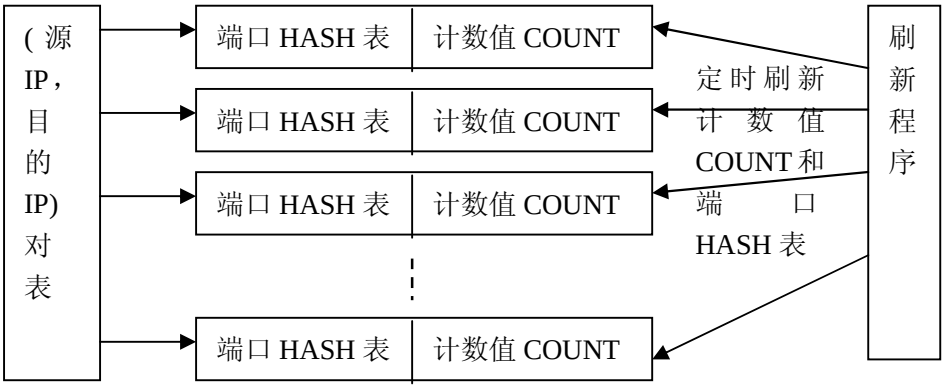

图 2　　扫描分析机根据特征确定的算法示意图

地址表中存放请求连接的源和目的地址，每个地址对对应一个哈希表和一个计数值。当有数据包到达时，与地址表中已有的记录匹配。如果匹配成功，就根据哈希函数确定端口的哈希表位置是否被置位；如果没有则该连接的 COUNT 增量并把该位置置位；当 COUNT 达到某一门限时报警。刷新程序根据 REFRESH－RATE 定期刷新所有的哈希表和 COUNT，刷新时 COUNT 减去一个常数，直至 COUNT 小于或等于 0 ，则该项将从表中删除。这样事实上扫描分析机只对从一个源地址到某个目的地址发起的大密度的请求连接敏感，对正常的访问没有影响。

具体使用的哈希函数是目的端口的低 8 位。哈希表是有 256 bit 组成的线性表，每一位对应一个端口，该位置 1 表示该端口被访问，否则表示未被访问。这样可以较好地解决查找端口的效率问题。

算法的另一个难点是两个参数 COUNT，REFRESH－RATE 的确定，它们与网络的实际情况有关。在系统中 COUNT 是根据 SATAN 中度扫描的端口数确定的，系统设定的默认值为 10。刷新的频率和网络运行情况有关，系统目前采用的方法是每隔 N 个报文刷新一次。为了确定 N 的取值，我们在网上以 100000 为容量，对报文进行 2 次采样分析，结果如下(见表 2)。从表中的数据可以看出 SYN 报文的数量比均在 1%以下，因此本系统选择的 N 为 50 个请求连接报文，实际的刷新频率大约是 5000 左右的普通报文。经实测，可以感知 SATAN，ISS，STROBE，NMAP，CYBERCOP SCAN 等常见扫描器的 TCP 扫描。

表 2　对网上数据包的类型的统计[1]

	第一次测试		第二次测试	
数据包类型	数据包的数量	百分比	数据包的数量	百分比
SYN	753	0.8%	594	0.6%
UDP	14470	14.5%	6950	7.0%
TCP	40507	40.5%	59092	59.1%
FRAG	0	0.0%	165	0.2%
总计		100000		100000

[1] 注：SYN 表示有 SYN 标志的报文的数量；UDP 表示 UDP 协议报文的数量； TCP 表示 TCP 协议报文的数量； FRAG 表示分片的 IP 报文的数量。其中的各统计项是非互斥的，所以百分比相加可能不为 1。

5.2 teardrop 分析机和 land 分析机

teardrop 和 land 都属于 IP 服务失效(Denial-of-Service)攻击，有关它们的报告都出自[CERT* Advisory CA-97.28][10]。 teardrop 攻击主要是利用了 IP 报文重组上的一个漏洞。正常的 IP 报文重组是各个报文首尾衔接的，而当两个分片报文的地址发生重叠（不是交叉），在某些操作系统的重组过程中可能会出现服务失效。teardrop 攻击的特征可以描述为：

“一个分片的 IP 报文在重组过程中，出现重叠（overlap）。”

这是一个较为精确的描述，因为其他原因出现这种情况的概率非常小。这个描述可以监测这一类的攻击行为，包括 teardrop 的改进版本 bonk 和 boink 以及 teardrop 的 TCP 版本 syndrop 都可以监测。

监测 teardrop 的方法是模拟 IP 报文的重组过程（只存储报头地址信息），从中发现重叠现象。根据表 2 的统计结果可知分片报文的比例比 SYN 报文还小，因此，实时模拟分片报文的重组是可行的。

Land 是一种简单而有效的攻击，它利用 TCP 的“三次握手”协议的漏洞：它每一次攻击时发一个经过地址欺骗（spoofing）的报文，报文的目的地址和源地址相同，目的端口和源端口相同的 TCP 报文，报文到达目的主机，使主机总处于等待状态。它的特征描述非常简单：

“一个报文的源地址和目的地址相同，并且源端口目的端口也相同。”

因为该攻击特征简单，且是一种上下文无关的描述，所以分析工作在第二层中就可以完成，而分析机所做的工作就是送去报警。

6. 报警机制

报警部分是每一个安全监测系统必不可少的部分，不合适的报警方式会严重影响系统的效率。报警的方法可以大致分为被动方式和主动方式，前者包括记录日志、通知管理员等，后者则包括切断攻击者的连接，甚至调整防火墙配置以阻止攻击者的其它动作。

由于攻击者常常使用地址欺骗(spoofing)，因此安全监测系统仅通过一次查看报警信息往往无法确定攻击的真正源地址，这时就冒然采取行动是不妥当的，很可能会被精明的黑客利用，因此在一般情况下只有对安全要求非常高的站点才需要采取主动的方式。

报警最常用的方式是记录日志，可以详细记录攻击发生的时间地点类型等信息，以便以后分析查阅。然而对一个实时监测系统而言，只做日志记录可能是不够的，还应该在第一时间通知管理员，并在主控台实时显示告警，如有必要还应该通知异地主机和管理员。本系统采取了日志记录和主控台屏幕显示的两种方式，在显示中并伴有蜂鸣。在报警时提供的各种信息中，时间是运行监测系统主机提供的，是可靠的。攻击类型符合一定的模式也是可靠的，而源地址和目的地址就存在欺骗的可能，需要管理员作进一步的分析和确认。

报警频率是另一个需要考虑的问题。因为运行监测系统的内存和硬盘空间等资源都是有限的，如果对报警系统频率和数量不做限制的话也有问题：例如黑客先利用非真实的信息将主机有限的资源塞满，那么再进行真正的攻击，这样监测系统将无法记录真正的攻击信息。解决这个问题的办法之一是限制某类攻击在一定时间内的报警次数，但这样有可能发生漏报；另外一种方法是给每种攻击告警固定各自的空间。本系统采用了前者。

7．结论

网络安全监测是当前的一个研究与开发的热点，但多数系统仍然停留在以 What-If 方式为主的检测方式阶段，因此适应性和准确性都不理想。本文提出了一种面向实时入侵检测的网络安全检测实现模型，通过讨论这个模型的一些实现细节来介绍系统的基本思路。作为一种范例，这个实现模型具体探讨了如何高效地实现基于上下文的入侵检测，重点在于安全分析知识的实现方式，即通过分析机来表达对某一类入侵事件的检测知识和相关的上下文。然而，虽然基于网络的安全监测系统能够弥补防火墙的不足，但是它也有自己的适用范围，譬如：无法监测未知类型的攻击；无法感知抽象程度较高的安全事件（非法删除文件，非法用户登录等）。

本系统作为一个安全工具，适用于小规模的网络，并具有灵活、方便的特点，且效率较高。如果要运用到大规模的计算机互联网中，需要在系统中增加安全协同能力，即需要采用分布式的体系结构[6][7]，在网络中不同位置的安全监测系统之间进行观测数据和安全知识的交换，从而可以更全面地掌握情况，作出正确的安全结论。

参考文献：

[1] Herringshaw, C., “Detecting attacks on networks”, “Computer” p16-17, 1997.12.

[2] Shiuh-Pyng Shieh; Gligor, V.D., “On a pattern-oriented model for intrusion detection”, “IEEE Transactions on Knowledge and Data Engineering” Vol. 9 No. 4, 1997.8.

[3] L.T. Heberlein, G.V. Dias, K.N. Levitt, B. Mukherjee, J. Wood and D. Wolber, “A network security monitor”, Proceedings of the IEEE Symposium on Security and Privacy, IEEE press, 1990

[4] Eric Maiwald, “Deploying Intrusion Detection System”, White paper of Fortrex technologies Inc., 1999

[5] Fyodor, “The Art of Port Scanning”, “Phrack” Issue 51 Vol. 7 article 11, 1997.9.

[6] Steven Cheung, Rick Crawford, Mark Dilger, Jeremy Frank, Jim Hoagland, Karl Levitt, Jeff Rowe, Stuart Staniford-Chen, Raymond Yip, Dan Zerkle, “The Design of GrIDS: A Graph-Based Intrusion Detection System”, 1999.1.

[7] S.R. Snapp, J. Brentano, G.V. Dias, T.L. Goan, L.T. Heberlein, C. Ho, K.N. Levitt, B. Mukherjee, S.E. Smaha, T. Grance, D.M. Teal, D. Mansur, “DIDS (Distirbuted Intrusion Detection System) – Motivation, Architecture, and an early Prototype”, Proceddings of the 14th National Computer Security Conference, p167-176, 1991.10.

[8] 匿名作者， 前导工作室 译， “网络安全技术内幕”， 机械工业出版社， 西蒙与舒斯特国际出版公司， p105-127， 1999.4.

[9] W.RICHARD STEVENS （美）著， 《UNIX 网络编程卷 1》 p703-726

[10] http://www.cert.org/

An Implementation Model of Network Monitoring for Misuse Detection

GONG Jian　DONG Qing　LU Sheng

（Computer Department of Southeast University, Nanjing 210096）

Abstract: An implementation model of network monitoring for misuse detection is proposed in this paper. The model contains three hierarchically related functional components: data collecting, analysis-decision, and analyzer, which can be effectively used to detect known misuses in a real-time way. A security knowledge expression method based on the concept of analyzer is introduced, and is applied to three well known attacks, scan, teardrop, and land as examples. Some other implementation issues like response mechanism are mentioned as well.

Key words: Network security, Network monitoring, Misuse detection, Scan, Teardrop, Land, Response mechanism.

作者简介
龚俭，教授、博导，主要研究方向为网络安全、网络管理、网络体系结构。
董庆，东南大学计算机系毕业生，现在深圳华为工作。
陆晟，博士研究生，主要研究方向为网络安全监测。