一种面向混合攻击的网络异常检测方法①

Network-Based Anomaly Detection Approach to Mixed Attacks

刘卫国 邹美群 (中南大学 信息科学与工程学院 湖南 长沙 410083)

摘 要: 考虑网络攻击的多样性，改进了异常检测中的特征提取及特征处理方法。该方法提取数据包的头部和应用层的数据，并将离散型和连续型特征分开进行处理。离散型特征采用基于时间的统计模型，连续型特征采用参数估计方法。实验采用 1999 DARPA 数据集，结果表明在保持低虚警率的情况下提高了多种攻击的综合检测率。

关键词: 网络异常检测 混合攻击 特征提取 特征处理

1 引言

入侵检测系统(IDS)是安全体系中的一种主动防御机制，按照数据来源和保护对象分为基于主机的 IDS (HIDS)和基于网络的 IDS(NIDS)[1]。HIDS 主要以主机的日志文件和审计记录作为数据源,检测针对主机的入侵事件，NIDS 捕捉网络数据包并检测针对网络的数据。

IDS 按照检测方法分为误用检测和异常检测。误用检测如 Snort[2]根据已知攻击的特征建立特征库,将待检测数据与已知攻击特征对照，如果匹配就表明发生了攻击。这种方法只能检测已知攻击，对未知攻击以及攻击变种无能为力，而且特征库的维护和更新代价很大。异常检测通过训练获取目标的正常模型，然后将目标的当前行为与之对照，如果有明显的偏差就表明发现了异常(或攻击)。这种方法能够检测已知攻击和未知攻击，但是容易产生误警。

基于网络的异常检测是将网络数据包或连接作为数据源，通过训练建立对象正常模型，用于对目标数据的检测。特征提取和特征处理是网络异常检测的关键，也是传统的流量模型和应用模型的区别所在。流量模型通过提取数据包头部字段建立正常流量模型以发现恶意流量，如 PHAD[3]、SPADE[4]及 NIDES[5]。这类模型能有效地检测流量型攻击，如 DoS 和 Probe，但是对应用型攻击的检测效果差。因为对于很多应用型攻击，仅通过分析数据包头部往往与正常连接并无差别，恶意数据潜藏在内容部分，必须采用应用模型分析数据包负载才能检测出来。应用模型如 PAYL[6]和 PCNAD[7]，为负载的全部或一部分建立正常字节频率分布轮廓，能有效地检测 U2R 和 R2L 攻击，尤其是针对 21 和 80 端口的攻击。缺点是需要大量的时间和计算空间,而且对DoS和Probe攻击的检测率很低。由此可见，单一的流量模型和应用模型只能检测特定类型的攻击，不能适应复杂的网络环境。

少量 IDS 试图检测各种类型的攻击，在特征提取方面包含了数据包头部字段，也包含了应用层信息。NETAD[8]采用了 48 个属性，包括从 IP 头开始的 40 个字节头部字段，以及负载的前 8 个字节。在特征处理时，把 48 个字节都看成离散型特征，为它们建立流量模型。此模型的缺点在于：提取的特征太多势必影响检测效率且耗费较大；48 个属性中有些是连续型特征(如数据包长度),采用与离散型特征相同的处理方法不尽合理，降低了检测率，提高了误报率。

本文主要通过改进传统异常检测的特征提取以及特征处理方法，以有效地检测各种类型的攻击。

2 方法描述

网络异常检测的流程如图 1 所示，数据源先经过预处理，识别出完整的请求报文，以服务类型划分网络连接；从连接中提取相关特征，用于训练和检测；训练阶段一般采用纯净数据集获得正常模型；检测阶段采用带有攻击的数据集，通过计算与正常模型的偏

① 基金项目:国家自然科学基金(60773013)；湖南省自然科学基金(07JJ5078)
收稿时间:2009-01-19

差(异常分 A)大小来判定数据是否异常，当异常分大于给定的阈值时，认为连接记录为攻击，由告警单元发出告警信息。

训练和检测阶段均涉及到特征处理，并且依赖于提取的特征。因此特征提取和特征处理是检测的关键步骤，要建立一种能同时检测各种类型攻击的方法，必须在这两方面改进。

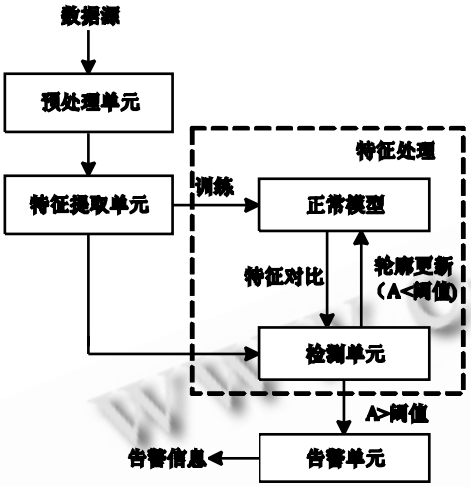

图 1　网络异常检测流程

2.1 特征提取

特征提取既要完备又要准确、简洁，才能提高检测率、降低误报率并达到较好的实时性。为了能有效检测各类攻击(包括流量型攻击和应用型攻击)，通过分析各类攻击的特点，考虑数据包头部的相关字段的同时结合应用层信息，提取了基本特征、流量特征及内容特征。这样弥补了传统流量模型仅提取数据包头部字段而应用模型仅对负载进行分析的缺陷。

(1) 基本特征

由于 TCP/IP 协议是目前使用最广泛的网络互联协议，TCP/IP 协议本身的安全性漏洞给了攻击者可乘之机，因此本文主要考虑 IP、TCP、UDP、ICMP 包。

很多攻击会在数据包头部留下痕迹，因此可以提取包头的若干有效字段。对于 IP 数据包，可以提取源地址、目的地址、协议类型、时间戳、标识。因为有些攻击者会利用伪 IP 来实施攻击，如 Land 攻击（攻击所产生的数据包源 IP 与目的 IP 相同）。提取协议类型是因为利用不同的协议可实施不同的攻击行为。有些攻击是通过频繁发送数据包以耗尽目标机资源，因此可提取时间戳。对于 TCP 包可以提取 TCP 标志位、源端口、目的端口；对于 UDP 包可以提取源端口、目的端口；对于 ICMP 包提取类型字段、代码字段。

(2) 流量特征

DoS 和 Probe 等攻击中，入侵特征主要体现在短时间内网络流量和主机的流量特征。如过去 2 秒内错误 SYN 数据包连接个数、同一主机与目标主机的连接数；对目标主机的不同端口进行的连接尝试等。

(3) 内容特征

U2R 和 R2L 攻击一般都潜藏在数据包的负载部分，且从单一的数据包分析和正常连接没有什么区别。特定类型的请求长度变化不大，有些攻击发生时会使请求长度明显增大，如缓冲区溢出。因些可将请求长度(连续发送的字节数)作为一个重要特征。另外可提取一个连接请求负载的若干字节。

本文以网络连接为单位，提取 22 个属性：DUR(连接持续时间)，LEN(请求长度)，COUNT(过去 2 秒内到同一主机的连接数)，Serror-rate(过去 2 秒内到同一目标机的错误 SYN 数据包连接比率)，SA3、SA2、SA1、SA0(源 IP 地址的 4 个字节)，DA1、DA0(目的 IP 的字节 1 和字节 0)，SP(源端口)、F1、F2、F3(第一个、倒数第二个、最后一个包的 TCP 标志)，W1~W8(连接发送的头 8 个字节)。前 4 个为连续型特征，后 18 个为离散型特征，用于建立正常行为模型和检测。

2.2 特征处理

目前大部分入侵检测要么只选取离散型属性，要么同时选取连续型属性和离散型属性但是将所有属性均用仅适合离散型属性的方法建立检测模型，这存在其不合理性。因为离散型属性的取值有限，可以通过正常训练集获得，检测阶段一旦发现新值，可以认为是异常数据；而连续型属性的取值个数可能有无数个(如数据包的负载长度)，无法像离散型属性在训练阶段收集所有合理的取值。本文对连续型特征与离散型特征采用不同的处理方法。

(1) 连续型特征的异常分计算

把每个连续型特征看成一个自由变量，训练中出现的属性值看成样本，使用统计推论和参数估计方法来建立正常模型。连续型特征 Xi 的值符合正态分布 $N(\mu,\sigma 2)$，目标数据偏离均值 μ 越大，异常的可能性就越大，相应地异常分就越大。检测阶段将目标数据与正常模型的偏差大小作为判定异常的标准，连续型

特征 Xi 的异常分 Ai 计算公式为：

$$A_i = m^{\frac{|l_i-\mu_i|}{n\sigma_i}}$$

这里 μ_i、σ_i 分别为训练阶段得到的属性 X_i 的均值和标准差，l_i 为 X_i 在检测阶段的实际取值。为了容忍连续型特征有一个合理偏差，这里设置了两个常数因子 m、n，这两个因子可根据请求类型以及属性调整，但在实现时必须依据所有连续型属性的 A 值在一个数量级的原则下设定。

(2) 离散型特征的异常分计算

本文对离散型特征的处理使用基于时间的统计模型，即一个事件的发生概率依赖于它从最后一次发生异常以来的时间。对每个离散型属性，我们收集它在训练阶段出现的所有合法值，检测时若出现新值，就赋一个异常分。离散型特征 X_i 的异常分 A_i 计算公式为：

$$A_i = t_i n / r_i$$

这里 t_i 为离散型属性 X_i 在训练或检测阶段自最后一次出现异常以来模型子集中的连接数，而非真正意义上的时间。t_i 越大，事件的发生概率就越小，因此异常值越大。n 为同一个服务类型的训练样本数，r_i 为属性 X_i 的合理取值个数，由训练得到。r_i/n 为属性 X_i 在训练阶段异常的平均概率，因此 n/r_i 越大的属性越不可能发生异常，则在检测阶段当该属性出现新值时应该赋一个更大的异常分，越长时间未出现异常的属性一旦出现新值，也应赋一个更大的异常分。

一条连接记录的异常分为各个属性的异常分之和，本文提取 4 个连续型属性和 18 个离散型属性时的异常分 A 为：

$$A=\sum_{i=1}^{22}A_i=\sum_{i=1}^{4}m^{\frac{|l_i-\mu_i|}{n\sigma_i}}+\sum_{i=5}^{22}t_i n/r_i$$

考虑到提取到的特征对检测效果的贡献率不等，因此可要据通过分析各种攻击的特点及协议类型为各个特征设置权值，越重要的特征的权值越大。为特征 X_i 设置权值 β_i，则修改后的连接记录的异常分为：

$$A=\sum_{i=1}^{22}\beta_i A_i=\sum_{i=1}^{4}\beta_i m^{\frac{|l_i-\mu_i|}{n\sigma_i}}+\sum_{i=5}^{22}\beta_i t_i n/r_i$$

设定一个阈值 ε，若 $A>\varepsilon$，则认为该网络连接为攻击。

3 实验分析

为了验证本文所提方法的性能，采用与 PHAD、PCNAD 及 NETAD 相同的数据集 1999 DARPA 进行测试评估。1999 DARPA 数据集中第 3 周数据为带标记的纯净数据，第 4，5 周的数据为带有攻击的数据。我们用第 3 周的数据作训练数据集，将第 4，5 周的数据作为检测数据集。先过滤数据集，保留流入方向的数据，则检测数据集中含有 185 个可检测的攻击，其中有 36 个 Probe 攻击，63 个 DoS 攻击，49 个 R2L 攻击，33 个 U2R 攻击，4 个其它攻击。

先过滤数据流，仅保留从外网发往内网的数据流，并且实现 IP 碎片和 TCP 流的重组，识别出特定网络服务完整的请求报文，将报文按服务类型划分网络连接(FTP 服务、Telnet 服务、HTTP 服务、SMTP 服务和其它服务等 5 类)。然后对 22 个特征进行处理，4 个连续型特征采用参数估计方法计算异常分，18 个离散型特征使用基于时间的统计模型。4 种攻击类型的检测结果为：检测出 31 个 Probe 攻击，45 个 DoS 攻击，39 个 R2L 攻击，21 个 U2R 攻击，一共检测出 136 个，总的检测率为 74%。

PHAD 只提取包头 33 个字段，而 PCNAD 使用 62.64% 的负载及 TCP 连接标志字段，NETAD 提取 48 个字节(包括负载的前 8 个字节)，将这 3 种模型与本文所提出的方法比较，在虚警个数为 100 时，它们的性能如表 1 所示。

表 1 性能对比

	Probe	DoS	R2L	U2R	总检测数(总检测率)
PHAD	31(86%)	45(72%)	10(20%)	6(18%)	92(50%)
PCNAD	10(28%)	9(15%)	44(89%)	28(85%)	91(49%)
NETAD	32(89%)	43(68%)	38(78%)	18(55%)	131(71%)
本文方法	31(86%)	45(72%)	39(79%)	21(64%)	136(74%)

由表 1 可见，PHAD 对于 Probe 和 DoS 攻击的检测率比较高，但对 R2L 和 U2R 攻击的检测效果很差；PCNAD 刚好相反，对后两种攻击检测率比较高，因为它着重在检测数据包负载；NETAD 除提取包头信息外还加入了 8 个字节的负载，故对基于内容的攻击

比 PHAD 大大提高了；本文所提出的方案结合了流量模型与应用模型，对 4 种类型攻击的检测率都较高，总检测率略高于 NETAD，原因在于为不同类型的特征建立不同的检测模型，实验表明这种处理方法较传统方法更合理，提高了混合攻击的综合检测率。

4 总结

本文所介绍的检测方法通过同时分析包头和负载，以发现流量异常和数据包内容异常。本文的另一个特点在于对特征的处理方法上，将连续型特征和离散型特征分开进行处理，并且加入了权值计算，改变了传统方法将所有特征均按照字符特征进行统计处理或全部转化成数值计算偏差的做法。通过这两方面的改进提高了系统的检测率及实时性。方法还可进行改进，对负载的处理，可多提取一些内容分析负载字节频率分布，从而更加有效地检测出潜藏在负载中的恶意攻击。

参考文献

1 蔡龙征,余胜生,周敬利.一种非纯净训练数据异常入侵检测方法.小型微型计算机系统, 2006,27(3):437－441.

2 Roesch M. Snort-Lightweight Intrusion Detection for Networks. Proc of LISA 99, 1999.

3 Mahoney M, Chan PK. PHAD: Packet Header Anomaly Detection for Identifying Hostile Network Traffic. Florida Tech. Technical Report 2001-04, http://cs. fit.edu/~tr/

4 SPADE, Silicon Defense, http://www.silicondefense.com/software/spice/

5 Debra A, Lunt TF, *et al*. Detecting unusual program behavior using the statistical component of the Next-generation Intrusion Detection Expert System (NIDES). Computer Science Laboratory SRI-CSL95-06, May 1995. http://www.sdl.sri.com/papers/5/s/5sri/5sri.pdf

6 Wang K, Stolfo S. Anomalous payload-based network intrusion detection. In Recent Advances in Intrusion Detection, RAID 2004, September 2004:203－222.

7 Thorat SA, Khandelwal AK, Bezawada Bruhadeshwar, et al. Payload Content based Network Anomaly Detection. Applications of Digital Informa- tion and Web Technologies, 2008. ICADIWT 2008, First International Conference, 2008, 4－6 Aug:127－132.

8 Mahoney MV. Network traffic anomaly detection based on packet bytes. ACMSAC, 2003,1(13):21－38.