DOI:10.16644/j.cnki.cn33-1094/tp.2019.02.001

一种基于噪声与信号分离的网络攻击检测方法

杨 蕊，江浩巍，刘福泉

(浙江农林大学暨阳学院，浙江 诸暨 311800)

摘 要：网络攻击检测是防护网络空间安全的重要手段。检测的准确性和实时性是衡量网络攻击检测技术好坏的重要指标。分析了已有网络攻击检测技术在实时性和准确性存在的不足;提出了一种基于信号分离的网络攻击检测方法。实验表明,这种方法能够快速、准确地检测出网络攻击。

关键词：网络攻击；攻击检测；信号分离；噪声

中图分类号:G304 **文献标志码**:A **文章编号**:1006-8228(2019)02-01-04

A network attack detection method based on signal and noise separation

Yang Rui, Jiang Haowei, Liu Fuquan

(*Jiyang College of Zhejiang A&F University, Zhuji, Zhejiang 311800, China*)

Abstract: Network attack detection is an important means to protect cyber security. This paper analyzes the shortcomings of the existing network attack detection methods and proposes a new method which based on signal separation to detect the network attack quickly and accurately. Experiment shows that this method can detect the network attack efficiently.

Key words: network attack; attack detection; signal separation; noise

0 引言

一个复杂的网络攻击可能会经历以下几个阶段:侦察,扫描,获取访问权限,维护访问权限,进一步攻击和掩盖攻击轨迹[1]。相应地,对网络攻击防护机制大致可以分为三个阶段:预防、检测和对已经发生的攻击作出反应。预防通常用于控制或限制对系统的非法访问,常用的方法有防火墙、加密、认证、授权等。检测用于监视在系统上所进行的活动,试图识别对系统进行的攻击行为。反应用于控制攻击活动的进一步扩散,对攻击追踪和诊断,最终对系统进行恢复,纠正系统中存在的漏洞。

本文借鉴了物理空间中信号分离思想,在分析已有网络攻击检测方法存在的不足的基础上,提出一种基于信号分离的网络攻击检测方法。

1 传统的网络攻击检测方法存在的不足

网络攻击检测是一种主动的安全防护技术,传统的网络攻击检测方法主要采用误用检测和异常检测。

误用检测[2]方法根据已知的网络攻击提取出特征数据,根据这些特征数据,利用机器学习等方法训练出相应的攻击模型。将这些攻击模型存储在检测系统中,用于检查实际网络空间中是否存在与攻击模型匹配的数据,如果有,则发出攻击预警。然而,一方面,网络空间中每天都有新的攻击诞生,而新的攻击模式对于误用检测系统来说通常是未知的,因此,误用检测对新攻击是不起作用的;另一方面,网络攻击行为相对正常操作行为数量是很少的,很难快速准确地检测出来。

异常检测[3]方法根据正常的网络行为提取出特征数据,根据这些特征数据,利用机器学习等方法训练出正常网络空间行为模型。当实际网络空间行为偏离正常网络空间行为模型的距离超过一定阈值时,发出攻击预警。然而,如果攻击行为活动序列偏离正常行为模型的差值不超过阈值,则攻击就不能被检测出来。另一方面,如果一个网络空间活动是正常的,但却与常规的正常活动操作序列有很大区别,那么这个

收稿日期:2018-11-09

作者简介:杨蕊(1998-),女,浙江人,学生,主要研究方向:计算机网络。

通讯作者:刘福泉(1981-),女,湖南人,硕士,讲师,主要研究方向:计算机网络。

正常的网络活动也可能被误报为网络攻击。

为了减少误用检测的漏报和避免异常检测的误报率太高，本文提出基于信号分离的攻击检测方法，试图将两种方法结合起来，先通过异常检测方法过滤掉网络正常行为数据，减少正常网络行为数据对攻击检测的干扰，然后使用误用检测方法对剩下的数据进行误用检测。

2 基于信号分离的攻击检测

2.1 来自物理空间中信号分离技术的启示

信号分离技术在物理空间中的信号检测和增强应用中得到了广泛了应用[4]。网络空间的攻击检测与物理空间中信号检测具有一定相似性，可以从物理空间中的信号检测中得到一些启示。为了检测网络空间中的攻击活动，如果直接在混合数据中检测网络攻击，就跟直接在混有噪声的语音中检索机长的语音信号一样困难。如果先将混合数据中的噪声过滤掉，然后在过滤后的数据中使用模式识别方法就能、准确地快速检测出攻击数据。在物理空间中，可以在不同位置同时采集背景语音信号（噪声）和混合语音信号。然而，与物理空间中信号分离方法不同的是，在网络空间中无法同时采集正常操作数据（噪声）和混合操作数据。为此，文章中提出了一种噪声预测模型，将采集到的数据输入到噪声预测模型可以得到相应的预测噪声，从采集到的数据中过滤掉预测噪声，对过滤后的数据使用特征识别的方法进行攻击检测。

2.2 噪声预测模型

本文建立了一种基于统计方法的噪声预测模型，模型的建立过程如图1所示。

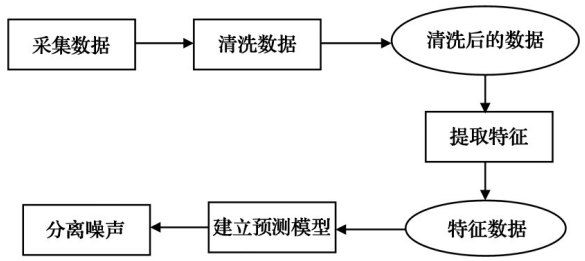

图1 噪声预测模型的建立过程

2.2.1 采集数据

在计算机网络入侵检测系统中广泛使用两类数据，即网络流和审计跟踪数据[5]。除此之外，我们还把计算机网络系统看成资源的集合，资源具有状态、性能等信息，在资源上进行的操作（不管是正常操作还是攻击操作）都会对资源产生影响。比如：内存资源、CPU资源、带宽资源等。因此，在我们提出的基于噪声信号分离技术的网络攻击检测方法中，采集的数据包括网络流量、审计跟踪数据和资源状态、性能数据。下面以Windows操作系统为例说明要采集的数据以及数据采集的方法。

网络流数据：单位时间流经网络接口的数据包数，包括发送与接收的数据包。使用WinDump[6]进行采集。

审计跟踪数据：审计跟踪数据包括系统活动和用户活动所产生的活动记录数据，保存在终端系统中。系统活动包括操作系统和应用程序进程的活动；用户活动包括用户在操作系统中和应用程序中的活动。使用Windows时间监控工具进行采集。

资源状态数据：可用的内存空间（字节）。使用Windows系统性能监控工具进行采集。

资源性能数据：单位时间产生的页面错误数。当线程引用不在主存储器中的虚拟存储页面的工作集中时，会发生页面错误。使用Windows系统性能监控工具进行采集。

采集数据的时间是离散的，只在发生网络事件发生时采集数据（比如接收到一次网络服务请求），可以把一次采集数据的事件看成是一个离散随机事件，用$X(t,\xi,\theta)$表示。其中，t表示采集数据的时间点，是一个时间序列，设T为时间域，采集数据的时间序列可以表示成：

$$t_0<t_1<\cdots<t_i<\cdots<t_n \in T$$

θ表示特征数据，是一个向量，$[\theta_0,\theta_1,\cdots,\theta_n]$表示提取出来的一组特征。$\xi$表示特征的权重，称为系数，也是一个向量，与特征向量对应。一个ξ的取值对应某种正常网络活动预测模型。所有ξ取值的集合为所有正常网络活动预测模型的状态空间。

2.2.2 清洗数据与提取特征

清洗数据是指去除那些与预测网络操作活动无关的数据，比如数据包的首部。提取特征是指将那些与正常网络操作活动相关的特征提取出来，比如：CPU的性能、内存的占用情况、单位时间流经网络接口的数据包数等。

2.2.3 建立噪声预测模型

根据从网络空间中采集数据的方法和特点，把每次采集看成是一个离散随机事件$X(t,\xi,\theta)$。如果使用过去发生的离散随机事件预测将来的随机事件，贝叶斯定理是一个典型的预测模型[7]，用数学表示如下：

$$P\{X_{t+1}=i_{t+1}\mid X_t=i_t,X_{t-1}=i_{t-1},\cdots,X_0=i_0\} \tag{1}$$

其中，$X_t=i_t,t\in T,T=\{0,1,\cdots\}$。$i_t$表示在时间t上的状态，$X_t=i_t$是一个随机过程，表明在时间t上$X_t=i_t$的概率分布。

但是使用贝叶斯预测模型计算量太大，不能满足系统时效性的要求，于是采用了对贝叶斯预测模型进行简化了的马尔可夫模型[8]。这里假设在时间t+1上的状态$X_{t+1}=i_{t+1}$的概率分布只与时间t的状态$X_t=i_t$的概率分布有关，而与其他状态$X_{t-1}=i_{t-1},\cdots,X_0=i_0$的概率分布无关，用数学表示为：

$$P\{X_{t+1}=i_{t+1}\mid X_t=i_t\} \tag{2}$$

$p_{i,j}$表示系统在时间t时状态为i的前提下，在t+1时系统状态变为j的概率。如果系统的状态数是一个有限集(1,2,⋯,s)，马尔可夫预测模型可以表示成一个转换概率矩阵：

$$P=\begin{bmatrix} P_{1,1} & P_{1,2} & \cdots & P_{1,s} \\ P_{2,1} & P_{2,2} & \cdots & P_{2,s} \\ \vdots & \cdots & & \vdots \\ P_{s,1} & P_{s,2} & \cdots & P_{s,s} \end{bmatrix} \tag{3}$$

其中，$\sum_{j=1}^{s}p_{i,j}=1$，设在初始时间t=0时，系统处于状态i的概率，其中$i\in(1,2,\cdots,s)$。

根据马尔可夫链模型，可以得到状态序列X_1, $X_2,\cdots,X_T$的联合概率：

$$P\{X_1,X_2,\cdots,X_T\}=P(X_1).P(X_2)\cdots.P(X_T) \tag{4}$$

可以从过去的系统状态观察结果中获得马尔可夫链模型的转移概率矩阵和初始概率分布。设一组按时间序列的系统状态观察值为：$X_1,X_2,\cdots,X_{N-1}$，状态i的概率为：

$$q=\frac{N_i}{N} \tag{5}$$

其中，N为所有观察值的个数，N_i为状态i的观察值的个数。马尔可夫链的状态转换矩阵的转换概率为：

$$p_{i,j}=\frac{N_{i,j}}{N_i} \tag{6}$$

其中，N_i为观察到X_t的状态i到X_{t+1}的状态1,2,⋯,s的观察值的个数，$N_{i,j}$为观察到X_t的状态i到X_{t+1}的状态j的观察值的个数。

2.3 网络攻击检测过程

将噪声预测模块的输出作为基准输入，将从网络空间中实时采集到数据作为主输入，两者一起汇入信号分离模块，将主输入中的正常操作数据信号分离。将信号分离模块的输出结果即过滤后的数据，作为模式识别模块的输入。将模式识别模块的输出结果作为决策模块的输入，由决策模块决定是否发出网络攻击警报，如图2所示。

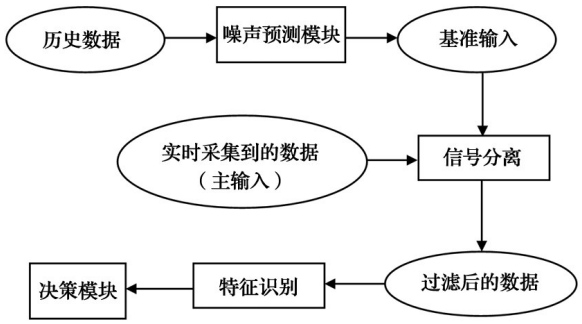

图2 网络攻击检测过程

2.4 cuscore信号分离模型

本文运用cuscore模型[9]，对主输入数据和基准输入数据的对应特征数据之差进行累积运算，实现对用户正常操作事件进行过滤。

基准输入用公式表示为：

$$y_{t0}=T+a_{t0} \tag{7}$$

其中，T为用户正常操作事件的特征向量值。a_{t0}为产生基准输入过程中的白噪声，服从正态分布。

主输入用公式表示为：

$$y_{t1}=T+\delta\sin x+a_{t1} \tag{8}$$

其中，T为用户正常操作事件的特征向量值。a_{t1}为采集主输入过程中的白噪声，服从正态分布，δ sinx为信号模型。公式(8)只是简单地将主输入数据看成是用户正常操作数据与攻击操作数据叠加而成的，然而实际过程中，正常数据可能会受到攻击数据的影响而变形，不那么容易区分出来。于是，本文采用cuscore累积模型来实现对攻击数据与正常数据的分离。该用数学表达式表示如下。

$$Q=\sum_t(y_t-T)(\sin x_t) \tag{9}$$

其中，y_t为采集到的混合数据，T为用户正常操作数据，$\sin x_t$为信号。

3 实验

为了说明基于信号/噪声分离检测网络攻击的方法，选择了Ettercap[10]攻击进行实验。Ettercap是一种地址解析协议(ARP)攻击。ARP攻击通过向当前子网上的每个IP地址发送一系列ARP请求来确定当前网络上有哪些计算机。然后，攻击者向受害者发送欺

骗性ARP响应。响应信息中填写的是该网络中的IP地址和攻击者的物理(MAC)地址。一旦受害者计算机用该响应信息更新其ARP表,受害者计算机以后发送的数据都会到达攻击者的机器。图2对Ettercap攻击的活动及对资源的影响进行了分析。对网络接口单位时间接收/发送的数据包、单位时间进程IO读写的字节数、单位时间页面错误数进行了采集。利用小波变换将时域分析变换成频域分析。然后计算样本概率,根据马尔可夫链得到噪声模型,结合cuscore算式得到信号模型。实验过程中,用户打开IE浏览器,向www.baidu.com网站发起搜索。在发起Ettercap攻击之前我们采集了10分钟数据用来建立噪声模型。在发起攻击的过程中,采集了10分钟数据用来建立信号模型。然后采用混合数据(既有正常活动也有攻击活动的数据)进行测试,将网络接口流量、进程IO读写数据量和进程处理过程中的页面错误数等的观察值作为cuscore模型的y_t进行计算,得到的结果分别对应于图3、图4和图5。在实验中,第0到300次观察过程中,我们只是简单的打开浏览器进行预定的搜索,从301次观察开始实施了攻击活动,从实验结果可以看出我们的模型能够快速、准确地检测出这种网络攻击。

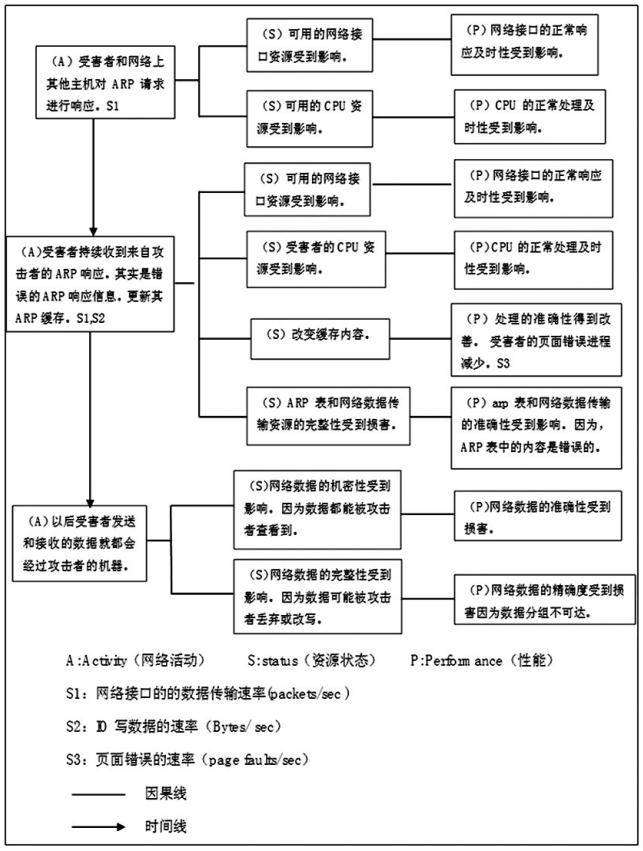

图3 Ettercap攻击的过程及对资源的影响

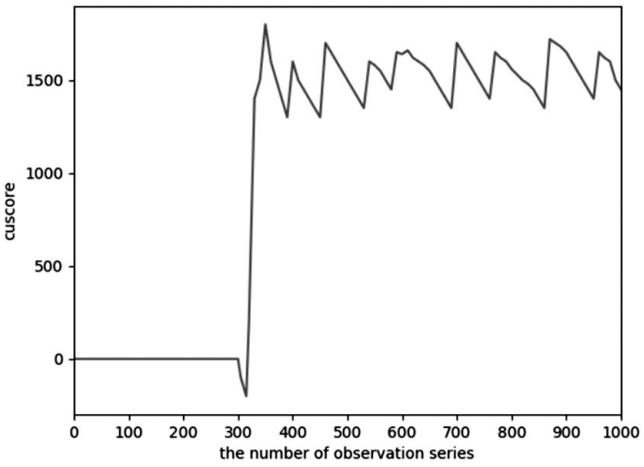

图4 数据包的cuscore值

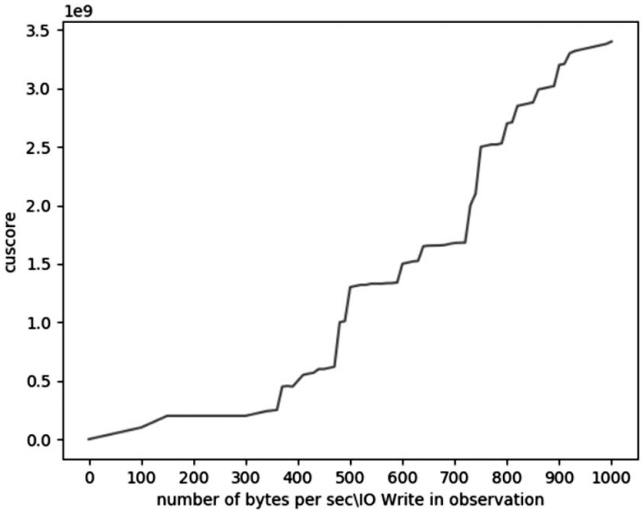

图5 I/O字节的cuscore值

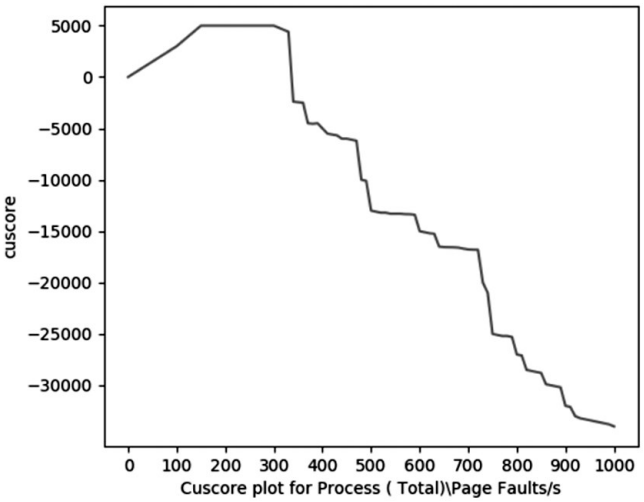

图6 页面错误数的cuscore值

4 结束语

使用基于噪声与信号分离的网络攻击检测方法

(下转第8页)

故整体效果不如基于层次聚类的改进贝叶斯算法。

3 总结

为了提高定位判别正确率，以及提高定位速度，本文提出了一种基于层次聚类质心距离的聚类方法，通过计算每类质心所包含的无线信号强度与实时信号强度的欧氏距离大小，选取距离最近的类作为预匹配的结果，并在该区域内利用改进贝叶斯进行进一步判定，经过在相关区域进行实验，验证了该方法的可行性。与传统贝叶斯算法和传统WKNN算法进行了比较，结果显示，本文方法可以在一定程度上提高定位判别率，并且在定位时间方面有了很大的降低，提高了实时性。

本文研究的方法主要针对二维平面位置，大多数情况下不仅需要知道平面位置，还需要知道楼层号码，所以接下来，主要针对楼层判别方法进行研究。进一步的研究将对救援、根据位置提供服务、防止人员走丢等方面都有巨大的作用。

参考文献(References):

[1] Winternitz L M B,Bamford W A,Heckler G W.A. GPS receiver for high satellite navigation. IEEE Journal of Selected Topics in Signal Processing,2009.3(4):541–556

[2] 汪伦杰,廖兴宇,潘伟杰等.基于信号均值滤波+ k–means+ WKNN的Wifi指纹定位算法研究[J].微电子学与计算机,2017.34(3):30–34

[3] 高仁强,张晓昉,熊艳,吴水平,晏磊.模糊数学的WiFi室内定位算法[J].测绘科学,2016.41(10):142–148

[4] 李军,何星,蔡云泽,徐琴.基于K–means和Random Forest的WiFi室内定位方法[J].控制工程,2017.24(4):787–792

[5] 曹晓祥,陈国良.一种改进的组合定权的指纹定位算法[J].测绘通报,2018.2:6–10

[6] 王怡婷,郭红.基于层次聚类的WiFi室内位置指纹定位算法[J].福州大学学报(自然科学版),2017.45(1):8–15

(上接第4页)

在进行了攻击检测实验，从实验结果可以看出该方法能够快速、准确地检测出这种网络攻击。但实验设计上尚存不足，目前只验证了方法的可行性，尚未实现与其他方法比较。

参考文献(References):

[1] DING Derui, HAN Qing–Long, XIANG Yang, etal. A survey on security control and attack detection for industrial cyber–physical systems[J]. In:Neurocomputing, 2018.275:1674–1683

[2] 唐正军.网络入侵检测系统的设计与实现[M].电子工业出版社,2002.

[3] George Chin Jr. Predicting and Detecting Emerging Cyberattack Patterns[R]. In: 2014 9th Cyber and Information Security Research Conference, ACM Press, 2014:95–105

[4] 赵力.语音信号处理[M].机械工业出版社,2009.

[5] KUKREIA Kashish, KARAMCHANDANI Yugal, KHANDELWAL Niraj, etal. Intrusion Detection System[J]. International Journal of Scientific and Research Publications,2015.5:709–711

[6] WinDump[EB/OL].https://www.winpcap.org/windump/

[7] 陈伟,陈继明.基于贝叶斯模型的云服务服务质量预测[J].计算机应用,2016.36(4):914–917,926

[8] 尹清波,张汝波,李雪耀等.基于线性预测与马尔可夫模型的入侵检测技术研究[J].计算机学报,2005.28(5):900–907

[9] Giovanni Radaelli. Using the Cuscore technique in the surveillance of rare health events[J].Journal of Applied Statistics,2006.7:75–81

[10] Ettercap Homepage[EB/OL].http://www.ettercap–project.org/ettercap/