神经网络在异常检测中的应用

宋　歌[1]　闫　巧[2]　喻建平[3]
[1] 西安电子科技大学通信工程学院,西安 710071)
[2] 西安电子科技大学电子工程学院,西安 710071)
[3] 深圳大学信息工程学院,深圳 518060)
E-mail:cindysong@263.net

摘　要　文章在简单介绍了入侵检测技术之后，在前人工作的基础上提出了一种在异常检测中用神经网络构建程序行为的特征轮廓的思想。文中给出了神经网络算法的选择和应用神经网络的两种网络设计方案,并对它们进行了比较。实验表明在异常检测中用神经网络构建程序行为的特征轮廓,能够大大提高检测系统对偶然事件和入侵变异的自适应性,特别是带有反馈的回归神经网络能更充分地利用数据信息,在保持系统的虚警率不变的情况下使检测率也有所提高。

关键词　入侵检测　异常检测　系统调用　神经网络　回归神经网络

文章编号　1002-8331-(2002)18-0146-03　　文献标识码　A　　中图分类号　TP183

The Application of Neural Networks in Anomaly Detection

Song Ge[1]　Yan Qiao[2]　Yu Jianping[3]
[1] (School of Communication Engineering,Xidian University,Xi'an 710071)
[2] (School of Electronic Engineering,Xidian University,Xi'an 710071)
[3] (School of Information Engineering,Shenzhen University,Shenzhen 518060)

Abstract: After giving a brief introduction of Intrusion Detection,this paper represents a method of using Neural Networks in anomaly detection to analyze the short sequences of system calls.The choice of algorithms used in Neural Networks is suggested and also two kinds of network design and their comparison are given in the paper.Experiments show that use NNet in Anomaly Detection to profile program behaviors can greatly improve the system's adaptability to new events and variance of intrusions.And using the Recurrent Neural Network with a feedback is especially better since it can improve the detection rate without increasing the false positives.

Keywords: Intrusion Detection,Anomaly Detection,systemcall,neural networks,recurrent NNet

1　引言

随着计算机网络的飞速发展，网络中的信息量越来越大，人们也越来越关注网络信息的安全问题。许多安全技术,如数据加密、防火墙、虚拟专用网、反病毒技术等都在快速发展并逐步走向成熟。近年来,对入侵检测技术的研究也越来越多,入侵检测系统也发展得越来越快，各种商用入侵检测系统相继出台,相关的理论也在逐步完善。

入侵检测(Intrusion Detection)顾名思义,便是对入侵行为的发觉。入侵检测系统在计算机网络或系统的若干关键点上收集信息并对这些信息进行分析,从中发现网络或系统中违反安全策略的行为或被攻击的迹象。1980 年 James Anderson's 在论文《Computer Security Threat Monitoring and Surveillance》中首次提出了入侵检测的概念。此后,入侵检测技术从主机到网络,从集中结构到分布式代理结构,经历了一个飞速的发展过程。入侵检测有两种最基本的方法:滥用检测(Misuse Detection)和异常检测(Anomaly Detection)。滥用检测是一种基于规则的检测技术,检测已知的入侵模式;而异常检测是一种基于行为的检测，它的检测是建立在"正常行为"的特征轮廓(Activity Profile)的基础上,通过比较被测活动的特征与正常特征轮廓的偏离程度来检测入侵，异常检测具有较强的适应性。

1996 年，墨西哥大学的 Forrest 等人提出了在异常检测中通过分析程序执行过程产生的系统调用序列来构建特征轮廓的方法[4],并用实验证明了程序执行轨迹的局部模式(系统调用的短序列)可以完全刻画程序行为的特征。这种基于程序行为的分析方法可以大大减少由用户行为的不可测性带来的系统虚警率。Forrest 对短序列的分析采用的是统计的方法,较为经典的是 tide 和 stide 两种方法。tide 方法存储所有唯一的短序列来构成特征数据库,检测时比较并记录被测轨迹的短序列与数据库中记录的不匹配,以此作为检测标准;stide 是在 tide 方法的基础上通过求序列的局部不匹配率来判断异常[4],这种方法认为局部区域的不匹配数目能够更好地表征异常行为。虽然采用了树状存储,但这些算法仍需要较大的空间来存储特征数据库,而且缺少对偶然事件和入侵变异的应变力,很容易产生

基金项目:国家 863 应急项目信息安全技术(项目号 301-6-6);"十五"计划子课题:入侵检测预警和安全管理技术(编号 863-104-02)
作者简介:宋歌,硕士,研究方向:网络安全与入侵检测技术。闫巧,女,博士,主要研究方向有信息安全,计算机通信等。喻建平,博士,深圳大学副教授。

虚报。为了得到一个适应性更强,虚警率更低的系统,研究人员开始将各种智能方法运用到入侵检测技术中。例如,LEE[8]等人用数据挖掘的方法研究系统调用数据的采样,他们使用一个称为“RIPPER”的程序,用一个较小的规则集合来描述正常数据的模式特征,在监控时,违反这些特征的序列被视为异常;在我国李之棠[10]等人建立了一种基于模糊专家系统的入侵检测框架模型。该模型吸收了滥用检测和异常检测的优点,能较好地降低漏警率和虚警率。

神经网络当然也被用于入侵检测中。例如,AnupK.Ghosh等人提出了一种用神经网络对程序行为的特征进行分析的方法[9],他们将系统调用短序列变换到一个从样本中随机选取的X维空间上,并以每个短序列在X维空间的坐标为神经网络的输入进行异常检测。论文在前人工作的基础上,给出了两种用神经网络构建的特征轮廓的方案,一种运用简单的BP网络实现虚警率的降低,另一种用带有反馈的回归网络实现了在不增大虚警率的基础上对系统检测率的提高。文中详细说明了运用神经网络技术分析系统调用短序列的方法。

2 用神经网络构建程序行为的特征轮廓

异常检测的关键就是如何形成一个用户或系统正常活动的特征轮廓。神经网络的自学习性和自适应性吸引越来越多的学者研究如何将其运用到入侵检测中,并且逐渐成为分析特征数据的工具的首选。它可以通过训练,不断地学习、调整主体的特征模式,从而构建出一种具有自适应特性的特征轮廓。这个轮廓对于一些偶然发生但仍属于正常范围的特征量或未知入侵行为的变异都可以做出正确反映,降低了系统虚警率,更加适应网络环境的千变万化。而且,神经网络本身就可以将所描述的行为特征“记忆”下来,因此不需要另外的存储空间来存储特征数据库。

2.1 神经网络算法的选择

选择神经网络的目的就是要利用它的自学习性,使得系统在数据不完整和易变的情况下,仍然能得到一个全面的特征轮廓。算法的选择与网络功能设计息息相关。当要完成分类的功能时,最简单的考虑就是用反向误差传播网络(BP网络)。

BP算法的基本思想是正向传播数据,计算每一层的输出;计算实际输出与期望输出的误差,并反向传播这个误差来调整各层的权值和阈值,使整个网络的输出误差达到最小。网络正向传输数据时,用S型激活函数来计算每一层的正向输出。典型的S型函数如下式:

$$F(x)=\frac{1}{1+\exp(-(x+b))}$$

反向传播误差时采用梯度下降的方法来调整每一层的权值和阈值。权值调整的基本数学式如下:

$$w_{ij}(t+1)=w_{ij}(t)+\eta*\delta_{pj}*y_j+\alpha*(w_{ij}(t)-w_{ij}(t-1))$$

其中$w_{ij}(t+1)$为下一时刻的权值,$w_{ij}(t)$为当前的权值,$w_{ij}(t-1)$为上一次的权值,η为网络的学习速率,δ_{pj}为j节点p模式的误差,y_j为j节点的输出,α为动量因子。网络训练的过程就是不断计算输出误差,并依次调整网络权值直到输出达到一定的误差标准为止。

为了能够更充分地反应系统的动态特征,采用带有反馈的回归神经网络进行数据分析。在回归神经网络中,一部分输出被反馈回输入并参与下一步的训练。这样就在网络中形成了一个全局记忆,这些记忆信息可以自动地适应网络预测的需要。文章采用了一种完全连通的自适应网络,并且允许各层之间保持误差的反向传播用于权值的调整。反馈数取为1。网络中始终保持着整个过程的记忆信息:各神经元之间的连接权值反应了网络的长期记忆,存储着行为的规律;而每个神经元的激活函数则反应了网络的短期记忆,它存储的是当前序列的信息。作者应用这种网络来预测下一个可能的序列,以此进行异常分析。

2.2 网络的设计方案

在对系统调用序列分析时,作者仍然采用tide方法中滑窗的办法来生成短序列,设序列长度为k。针对神经网络的功能,这里提出了两种网络设计方案。

设计方案[I]:

在方案一中,主要思想是将入侵检测看作一个简单的分类问题,即将所有程序行为分为“正常”和“入侵”两类。神经网络的主要作用就是要找到正常行为系统调用短序列的特征。训练时,短序列的每一位对应一个输入节点,正常序列的期望输出为1,异常序列的期望输出为0。每次输入一个短序列,计算输出并调整权值,循环进行直到达到预定的误差标准。训练过程中还要产生不匹配的判决门限和异常的判决门限。即一方面,要确定当被测短序列偏离正常模式到什么程度时记为一次不匹配;另一方面,要确定当被测行为产生的不匹配率达到什么范围时就认为发生了异常。

这种方案的设计较为简单,但是它没有充分利用神经网络的学习预测性能和数据的先验知识,训练时间长且不易收敛。试验表明,采用方案一设计的网络灵敏性较低,对于正常和入侵两种行为得出的不匹配率相差很小。它对系统的虚警率改善很小并且没有兼顾到系统的检测率。

设计方案[II]:

在方案二中,主要是利用回归网络,通过预测下一个可能出现的短序列来进行检测。网络的大体框架如图1。

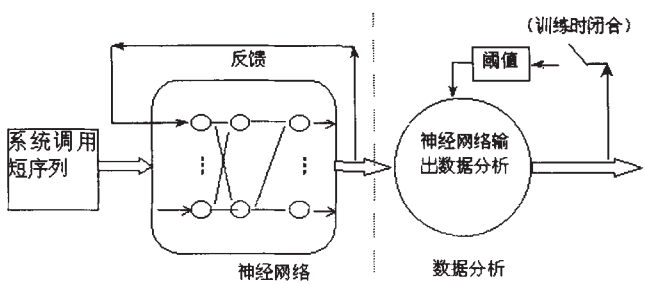

图1 用带反馈的回归网络构建特征轮廓的设计方案

其设计思想如下:

第一,输入层节点数取为$n+1$。其中n个输入节点分别对应n个程序中可能出现的系统调用;另外1个节点则对应于系统的反馈。输入节点按照系统调用的代码编号(反馈节点标为0号),当短序列样本输入时,将序列中每一位的系统调用对应的输入节点置1,其余节点输入为零。

第二,输出层为n个神经元,每一个神经元也都对应一个系统调用,输出值代表着该调用在下一个序列中出现的可能性。

第三,采用完全连通的反馈网络,并且网络中仍然采用误

差的反向传播来帮助修正权值。

第四,检测时,将预测的短序列与被测数据中实际应出现的短序列进行比较,计算其距离 d,并求出其中的最大值 S 来检测不匹配。然后同前面一样,由反应系统分析不匹配率,得出检测结果。

方案二充分利用了神经网络的学习预测性能和数据的先验知识,它在保持系统虚警率基本不变的情况下提高了系统的检测率。但是这种方法设计的网络规模比较庞大,且输入层用于接受反馈的神经元的个数 N 很难确定。从利用数据先验知识的角度出发,作者希望 N 越大越好;但是为了减少不必要的计算,又希望减小 N。目前还只能通过多次实验分析比较来选取 N 值。

3 实验结果

实验主要针对 sendmail 特权程序进行了测试和分析。正常数据是在 Sun SPARC 工作站上运行 SunOS 4.1.1 and 4.1.4 中运行 sendmail 程序产生的系统调用的数据。测试用的攻击数据有两种,一种是对 sendmail 的"decode intrusion"攻击,另一种是"sunsendmailcp intrusion"攻击。这些数据都是由新墨西哥大学计算机技术学院的研究人员在运行 SunOS 4.1.1 和 4.1.4 的 Sun SPARC 工作站上仿真模拟合成的(http://www.cs.unm.edu)。为了方便分析,将所有的系统调用命令按统一的编码表进行了编码,用十进制序号取代了原始命令,并对数据进行预处理,将系统调用序列以 7*x 的矩阵存储与数据库中以便使用。

在第一种方案的测试中,采用了 6020 个短序列样本,隐层的神经元数分别取 24、30、50、60、90、120。仿真结果表明,当隐层神经元取为 60 时,网络有较好的收敛速度及仿真结果。网络的检测率基本在 85%~86%之间,而虚警率控制在 12%左右。

对第二种方案,网络设计为 182 个输入节点对应于每一个调用(共有 181 种系统调用)和一个反馈,181 个输出节点。对每一个短序列样本,将相应的输入节点置 1。输出层神经元的值在 0 和 1 之间变化,表示着下一个序列中将要出现的新调用的概率(采用滑窗方法时,每一个序列都只有一位是新的调用)。然后由反应系统合成对一个短序列的预测结果,并进行分析。在这里作者只关心神经网络的输出。每一个样本的输出都可能有三种情况:

(1)预测结果十分明确。此时,输出的最大值明显大于 0.5,其它值都明显小于 0.5。在这种情况下,网络很容易做出判决,但是如果预测出现错误,这种情况就会产生很大的问题;

(2)训练不足。如果最大输出值小于 0.1,并且所有的输出都在同一范围内并且很小,那么网络将无法从给定的先验知识中确定下一个将要出现的调用,这种情况可能是由于缺少训练样本造成的;

(3)模糊不清。此时最大输出值在 0.1 和 0.5 之间,并且第二最大值跟它之间的差距很小。此时网络的输出只能给出下一个调用的可能趋势,但网络往往要面对两个或更多十分接近的值而无从选择。

作者首先用 6020 个样本的 40%对网络进行预训练,然后再用整个样本集完成对网络的训练过程,可以看到第二种方案的检测率,最好的时候达到 92%。表 1 给出了几种对系统调用进行分析的入侵检测的检测率。

表 1 几种检测方法的结果比较

	TIDE 法	BP 网络法	回归网络法
检测率	86.4%	86.5%	92%
虚警率	50%	22%	21%

4 小结

入侵检测是一种通过监控计算机或网络中发生的事件,来分析并检测对系统安全造成威胁的非法行为的技术。与传统的密码技术和防火墙技术等安全措施相比,它对系统提供的不仅仅是一种静态、被动的保护,它可以主动地发现入侵行为并采取相应措施。网络中信息安全的威胁越来越多,黑客的入侵手段也变化多端,因此网络对于入侵检测也必然会提出更高的要求。将神经网络这种智能的手段运用到入侵检测中,必然会大大提高入侵检测系统的性能,使其能够适应于网络信息的多样性,适应于不断扩大的网络产生的不断增多的安全威胁。

(收稿日期:2002 年 1 月)

参考文献

1.Steven A Hofmeyr,Stephanie Forrest,Anil Somayaji.Intrusion Detection using Sequences of System Calls.Dept of Computer Science University of New Mexico Albuquerque,NM 87131-1386

2.Christina Warrender,Stephenie Forrest,Barak Pearlmutter.Detecting Intrusions Using System Calls:Alternative Data Models.University of New Mexico (Albuquerque,NM 87131-1386)

3.Hervé DEBAR,Monique BECKER,Didier SIBONI.A Neural Network Component for an Intrusion Detection System.CSEE/DCI 6,Avenue des tropiques BP80 91943 Les Ulis Cedex France

4.Stephenie Forrest,Steven A.Hofmeyr,Anil Somayaji.A Sense of Self for Unix Processes.Dept of Computer Science University of New Mexico,Thomas A Longstaff,CERT Coordination Center Software Engineering Institute Carnegie-Mellon University

5.金波,林家骏,王行愚.入侵检测技术评述[J].华东理工大学学报(自然科学版),2000,26(1):191~197

6.赵海波,李建华,杨宇航.网络入侵智能化实时检测系统[J].上海交通大学学报,1999,(1)

7.Stephen Northcutt 著.余青霓,王晓程,周钢等译.网络入侵检测分析员手册[M].人民邮电出版社

8.W Lee,S J Stolfo.Data mining approaches for intrusion detection[C].In:Proceedings of the 7th USENIX Security Symposium,1998

9.Anup K Ghosh,J Wanken,F Charron.Detecting anomalous and unknown intrusions against programs[C].In:Proceeding of the 1998 Annual Computer Security Applications Conference (ACSAC'98),1998

10.李之棠,杨红云.模糊入侵检测模型[J].计算机工程与科学,2000,22(1)