文章编号: 1001-9081(2004)12Z-0131-03

一种基于 BP神经网络的异常检测系统的实现

蔡 坚,傅光轩,聂方彦
(贵州大学 计算机网络研究所,贵州 贵阳 550025)
(caijian330@tom.com)

摘 要:人工神经网络是当前的研究热点之一,它具有较强的学习归纳功能,将之应用于入侵检测,能有效地提高检测工具的各项性能。提出了一种基于 BP网络的异常检测系统,分析了几种典型的扫描和攻击技术,并由此选定了网络输入特征量,试验结果显示这种检测系统的检全率和误检率都十分令人满意。同时,由于网络的检测过程就是一些简单的数值计算,所以该系统检测速度快,实时性能好。

关键词:网络安全;入侵检测;BP网络;BP算法;异常检测
中图分类号: TP183 文献标识码: A

0 引言

入侵检测系统(Intrusion Detection System, IDS)是一种动态的网络攻击检测技术,能够在网络系统的运行过程中发现入侵者的恶行和踪迹,并作出适时的反应。其检测方法可分为两大类:误用检测和异常检测。误用检测使用已知的攻击模式或系统弱点来进行攻击识别,检测的精确度取决于模式库的完整性。通常,它只能检测出模式库中已有的攻击模式,而不能发现未知的(甚至是已知的变种)攻击模式,并且添加新的攻击模式是非常困难的。这种检测方法的检测率和误检率都比较低;异常检测通过提取网络流量或日志文件中的特征数据来描述用户行为,建立典型的网络活动轮廓模型。每当检测到一个新的行为模式,就会与建立的模型进行比较,如果超过已定义的阈值,则引发报警,表示一个可能的异常行为。与误用检测正好相反,这种方法能够发现许多未知的和已经攻击的变种,它的检全率相当高。但由于先前的经验不可能预知所有的未来模式,所以它的误检率也比较高。

目前,入侵检测工具最大的挑战之一就是对过去观察到的正常和异常行为进行总结归纳,并用于标识未知的攻击模式。为了解决这一问题,人们将人工神经网络引入了入侵检测系统,以期改善现有 IDS的性能。

1 人工神经网络

人工神经网络(ArtificialNeuralNetworks, ANNs),亦称为神经网络(NeuralNetworks, NNs),是由大量处理单元广泛互连而成的网络,是对人脑的抽象、简化和模拟,能够反映人脑的基本特性。它的研究始于 20世纪 40年代,经过三个阶段的曲折道路,其理论方法已经取得了长足的发展。近年来更是成为研究的热点,在时间序列分析、模式识别和控制等领域都得到了广泛的应用。

1.1 多层前向网络

前向网络分为单层前向网络和多层前向网络。多层前向网络是单层前向网络的推广,但它能解决单层前向网络所不能解决的非线性可分问题。其网络结构如图 1所示。

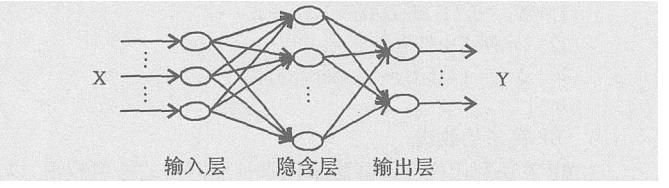

图 1 多层前向网络

X和 Y是网络的输入、输出向量。每个神经元用一个节点表示。网络由输入层、隐层和输出层节点组成。其中隐含层可以为一层或多层(图中是单隐层),前层到后层节点通过权连接,即拓扑结构为有向无环图。

输入层神经元的个数为输入信号的维数,隐含层个数以及隐节点的个数视具体情况而定,输出层神经元的个数为输出信号的维数。网络中每个神经元的激励函数必须是处处可导的,一般地,多数设计者都使用可微的 Sigmoid函数,如:

$$O_i = \frac{1}{1+\exp(-I_i)}$$

式中 I_i是第 i个神经元的输入信号,O_i是该神经元的输出信号。

1.2 BP学习算法

20世纪 80年代中期,以 Rumelhart和 McClelland为首,提出了多层前向网络(Multilayer Feedforward Neural Networks, MFNN)的反向传播(Back Propagation, BP)学习算法,简称 BP算法。它是一种有监督的学习算法,是梯度下降法在多层前向网络中的应用。

BP学习算法由正向传播和反向传播组成,整个过程描述如下:

(1)工作信号正向传播:输入信号从输入层经隐单元,传向输出层,在输出端产生输出信号,这是工作信号的正向传播。在信号的向前传递过程中网络的权值是固定不变的,第

收稿日期: 2004-05-14

作者简介:蔡坚(1980-),男,贵州安龙人,硕士研究生,主要研究方向:计算机网络与信息安全; 傅光轩(1944-),男,贵州松桃人,教授,主要研究方向:计算机网络体系结构、下一代互联网技术及网络与信息安全; 聂方彦(1977-),男,湖南麻阳人,硕士研究生,主要研究方向为计算机网络与信息安全.

一层神经元的状态只影响下一层神经元的状态。如果在输出层得到期望的输出，学习结束；否则，转入误差信号的反向传播。

(2)误差信号的反向传播：网络的实际输出与期望输出之间差值即为误差信号，误差信号由输出端开始逐层向前传播，这是误差信号的反向传播。在误差信号反向传播过程中，各层神经元的权值和偏差值由误差反馈进行调节，不断减小误差信号，使网络的实际输出更接近期望输出。

BP算法的可描述如下：

(1)初始化网络中所有权值和偏差；

(2)while停止条件不满足时{

(3) for samples中的每个训练样本 X {

(4) for每个隐含层和输出层{

(5) $O_j = \frac{1}{1+e^{-I_j}}$; $I_j = \sum_i \omega_{ij} O_i + \theta_j$; }

(6) for每个输出层单元 j

(7) $Err_j = O_j(1-O_j)(T_j - O_j)$;

(8) for每个隐含层单元 j

(9) $Err_j = O_j(1-O_j)\sum_k Err_k \omega_{jk}$;

(10) for网络中每个权重 ω_{ij} {

(11) $\Delta\omega_{ij} = (l) Err_j O_i$; $\omega_{ij} = \omega_{ij} + \Delta\omega_{ij}$; }

(12) for网络中每个偏差 θ_j {

(13) $\Delta\theta_j = (l) Err_j$; $\theta_j = \theta_j + \Delta\theta_j$; }

(14)}}

1.3 BP算法的改进

BP算法利用梯度下降法来搜索神经网络的一组权值，以使神经网络的类别预测与训练样本实际类别之间的均方差最小，从而成为相应分类问题的求解模型。由于梯度下降法本身的不足，在实际应用中存在两个重要问题：收敛速度慢和目标函数存在局部极小点。对此，提出了一些改善BP算法的方法：

(1)加入动量项。理论上，BP算法的训练是沿着误差曲面的斜面向下逼近的，但对一个复杂的网络来说，其误差曲面是一个高维的空间曲面，是非常复杂且不规则的，分布着许多局部极小点。解决这一问题最简单的方法是加入“动量项”，即令：

$$\Delta\omega_{ij}(n) = \alpha\Delta\omega_{ij}(n-1) + lErr_j(n)O_i(n)$$

α 为动量项，通常 $0 < \alpha < 1$。

引入动量项后，降低了网络对于误差曲面的局部细节的敏感性，不仅可以微调权值的修正量，也可以有效地抑制网络陷于局部极小。

(2)在BP算法中学习速率 l 的选取很重要，l 的值大网络收敛快，但过大会引起不稳定；l 值小虽然可以避免不稳定，但收敛速度就慢。较好的解决办法是设计一个自适应步长，使得权修改量能随着网络的训练而不断变化。在应用中，通常将学习速率设置为 l/t，t 为至今为止所处理的整个训练样本集的(循环)次数。

(3)尽可能使用顺序方式训练网络。顺序方式训练网络要比批处理方式更快，特别是在训练样本集很大，而且具有重复样本时，顺序方式的这一优点就更为突出。使用这种方式训练网络时需要注意的是，每一周期的训练样本其输入顺序应该尽是采用随机方式，使得连续输入的样本不属于同一类。

2 将NN用于异常检测

2.1 确定网络结构

要将BP神经网络用于异常检测，首先必须确定网络的结构，即输入层的节点数、输出层的节点数、隐层的层数及每一隐层的节点数。

应用于入侵检测，BP网络的输入节点数就是我们用于检测的特征数。文献[4]中，使用每一个网络数据包的一些域特征作为网络输入，而网络攻击事件通常是基于时间的序列模式，且如果对每一个数据包都进行一次分析计算，则在进行实时检测时必然会增加系统开销。所以本系统基于时间窗口，提取网络数据包的7个统计特征作为BP网络的输入，具体特征将在2.2节中详细说明。

输出层的节点数比较容易确定。在入侵检测领域，输出层的节点数就是欲分类的类别数。本系统输出层的节点数为2.当数据为正常样本时，节点1和节点2的期望输出分别为1.0和0.0.反之，当数据为异常样本时，它们的输出为0.0和1.0.

隐层层数的增加不一定能够提高网络的精度和表达能力，在多数情况下，BP网络一般都选用二级网络。同样，本系统中选用二级BP网，即隐层层数为1.由于人工神经网络是一个极为复杂的非线性动态系统，很难找到一个有关隐层节点数的简洁表达式。一般地，节点过少无法产生足够的连接权组合数来满足若干样本对的学习；反之，节点过多则学习以后网络的泛化能力变差。试验结果显示，本系统的隐层节点数为22时，网络的泛化能力(即正确分类的能力)最好。

2.2 选取统计特征

在本系统中，分析了几种典型的网络攻击和网络扫描的实质，并由此确定了网络的输入统计特征量，简单的介绍如下：

(1) SYN flood：该攻击以多个随机的源主机地址向目的主机发送SYN包，而在收到目的主机的SYN ACK后并不回应，这样，目的主机就为这些源主机建立了大量的连接队列，而且由于没有收到ACK一直维护着这些队列，造成了资源的大量消耗而不能向正常请求提供服务。

(2) FIN扫描：对某端口发送一个TCP FIN数据报给远端主机。如果主机没有任何反馈，那么这个主机是存在的，而且正在监听这个端口；主机反馈一个TCP RST回来，那么说明该主机是存在的，但是没有监听这个端口。

(3) DDoS(反射式分布拒绝服务攻击)：将伪造了源地址的SYN请求包发送到许多被欺骗的计算机上，根据TCP三次握手的规则，这些计算机会向源IP发出SYN+ACK或RST包来响应这个请求，受害机忙于处理这些回应而被拒绝服务攻击。

(4) FIN+URG+PUSH扫描：向目标主机发送一个FIN、URG和PUSH分组，根据RFC793，如果目标主机的相应端口是关闭的，那么应该返回一个RST标志。

(5) NULL扫描：给目标主机发送一个没有任何标志位的TCP包，根据RFC793，如果目标主机的相应端口是关闭的，应该发送回一个RST数据包。

(6) UDP flood：目前在互连网上提供的WWW、Mail等服务一般为使用UNIX操作系统，默认情况下它们开放了一些UDP服务，如：原本作为测试功能的chargen服务会在收到每一个数据包时会随机反馈一些字符，如果恶意攻击者将两个UDP服务互指，则网络可用带宽会就会很快耗尽。

(7) Smurf向一个子网的广播地址发一个带有特定请求(如 ICMP回应请求)的包，并且将源地址伪装成想要攻击的主机地址。子网上所有主机都回应广播包请求而向被攻击主机发包，使之受到攻击。

基于上述七种入侵形式，选取了以下七个统计特征：SYN、FIN、RST、FIN＋URG＋PUSH、NULL、UDP以及 ICMP，在2秒的时间窗口内，对 SYN、FIN、RST、FIN＋URG＋PUSH、NULL五种标志和 UDP数据包及代码为 0和 8的 ICMP数据包计数。

3　试验及结果分析

3.1　试验系统设计

本系统的研究和测试过程都是在我们实验室的局域网环境中完成的。PC机的硬件配置为：CPU Celeron 1.7G，内存256MB RAM；操作系统：Redhat9.0 linux 内核版本 2.4.20—8；开发工具：XEmacs；编程语言：C。

为了提高程序的抓包性能，在试验中采用了由美国洛伦兹伯克利国家实验室所编写的专用于数据包截获功能的 API函数库“libpcap”。试验采用三层 BP神经网络：输入层 7个节点，隐含层 22个节点，输出层 2个节点。停机条件为(1条满足即可)：①被错误分类的样本占总样本数的比例小于3%；②执行了 15000次处理循环。

整个系统由四个部分组成：数据包解析器、数据库、网络参数生成器和实时检测引擎(如图 2所示)。解析器基于 libpcap，将训练所用的统计特征量存入数据库，网络参数生成器读入这些特征数据并训练 BP网络，最后把得到的权值和偏差输入实时检测引擎进行实时检测。

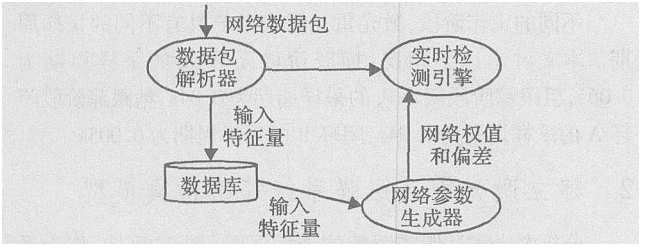

图 2　系统体系结构

3.2　样本收集

试验中，我们收集了 10组训练数据，其中正常和异常各5组，每组样本 25个。也收集了 20组测试数据，其中正常 10组，每组 20个；异常 10组，每组 30个。这样就使得我们的训练样本中，既包括了正常的行为信息，也包括了攻击行为信息。

在异常数据的收集过程中，我们使用了 toast[8]入侵工具和 Linux自带的 nmap主机扫描程序。

3.3　结果分析

使用 10组训练数据以随机的顺序输入对网络进行训练，将最终得到的连接权值和各神经元的偏差值存入文件，然后启动测试程序装入权值和偏差对测试数据进行计算判别。结果如表 1所示。

由表 1可得：

检全率：293 /300＝97.7%；

误检率：9 /500＝1.8%；

漏报率：7 /500＝1.4%；

检测正确率：1 － 9 /500 － 7 /500＝96.8%。

表 1　测试数据的检测结果

实际	结果	
	正常	异常
正常	191	9
异常	7	293

4　结语

人工神经网络通过对大量训练样本的学习，可以获得正常和异常数据的分类知识，从而能够对入侵的异常数据进行识别。并且在识别过程将对模式的分类转换为完全抽象的数值计算，无需假设数据的分布，也无需向神经网络解释知识的细节。另外，神经网络在学习阶段需要花费较长的时间和较多的系统资源，但是在测试和检测阶段它的速度都比较快，而且资源的占用率也比较低，完全可以用于实时检测。

从试验数据我们发现，这种基于 BP神经网络的异常检测系统，不仅是在测试阶段的检全率和误检率达到了令人满意的效果。而且在实时检测中，由于计算量不大，对于攻击和扫描的反应速度快，只要建立相应的报警机制，一旦检测到可能的入侵行为，系统就会立即通知管理员采取适当的措施，保护系统安全。

不过，现阶段我们的系统仅仅是基于时间窗口内的统计量进行检测，这样的系统对于一些高级的攻击方式而言是无效的，如：慢扫描、碎片攻击等。而且 BP神经网络有着一些自身不可克服的缺点，应该尝试使用其他类型的神经网络，甚至将各种方法结合使用，这些都是下一步即将展开的工作。

参考文献：

[1] RYAN J, LIN M, MIIKKULAINEN R. Intrusion Detection with Neural Network[A]. Proceedings of the 1997 conference on Advances in neural information processing systems 10[C]. MIT Press, 1998. 943－949.

[2] GHOSH AK, SCHWARTZBARD A, SCHATZ M. Learning Program Behavior Profiles for Intrusion Detection[A]. Proceedings of the 1st USENIX Workshop on Intrusion Detection and Network Monitoring[C]. USENIX Association, 1999.

[3] GHOSH AK, SCHWARTZBARD A. A Study in Using Neural Networks for Anomaly and Misuse Detection[A]. Proceedings of the 8st USENIX Workshop on Intrusion Detection and Network Monitoring[C]. USENIX Association, 1999.

[4] CANNADY J. Artificial Neural Networks for Misuse Detection[A]. Proceedings of the 1998 National Information Systems Security Conference (NISSC'98)[C]. Arlington, VA, 1998.

[5] 朱明. 数据挖掘[M]. 第一版. 合肥：中国科学技术大学出版社，2002.

[6] 高隽. 人工神经网络原理及仿真实例[M]. 第 1版. 北京：机械工业出版社，2003.

[7] 靳蕃. 神经计算智能基础：原理·方法[M]. 第一版. 成都：西南交通大学出版社，2000.

[8] Gridmark, toast[EB/OL]. http://packetstormsecurity.nl/DoS/indexsize.html, 2000－02－29.