第40卷第10期 中国科学技术大学学报 Vol. 40, No. 10
2010年10月 JOURNAL OF UNIVERSITY OF SCIENCE AND TECHNOLOGY OF CHINA Oct. 2010

文章编号:0253-2778(2010)10-1096-05

基于优化 BP 算法的无线局域网入侵检测系统的设计

刘凤纯[1,2,3],周 颢[1,2,3],赵保华[1,2,3]

(1. 中国科学技术大学计算机科学与技术系,安徽合肥 230027;2. 网络与交换技术国家重点实验室,北京 100876;
3. 安徽省计算与通讯软件重点实验室,安徽合肥 230027)

摘要:针对现有 BP 神经网络算法存在的平坦曲面和局部极小点问题,提出了极小点跃迁遍历 BP 算法,以取得更好的收敛效果.将改进的 BP 算法应用于无线局域网入侵检测系统.实验表明,改进的算法提高了入侵检测的准确性和实时性.

关键词:无线局域网;入侵检测;神经网络;802.11 协议

中图分类号:TP393.08 **文献标识码**:A doi:10.3969/j.issn.0253-2778.2010.10.015

Intrusion detection system design in wireless LANs based on optimized BP algorithm

LIU Fengchun[1,2,3], ZHOU Hao[1,2,3], ZHAO Baohua[1,2,3]

(1. *Department of Computer Science and Technology, University of Science and Technology of China, Hefei* 230027, *China*;
2. *State Key Laboratory of Networking and Switching Technology, Beijing* 100876, *China*;
3. *Province Key Laboratory of Software in Computing and Communication, Hefei* 230027, *China*)

Abstract: In view of the existence of slightly curved surfaces and local minima problems with the application of the BP neural network algorithm, ergodic BP algorithm based on minimal point escape was proposed to achieve better convergence effect. The improved BP algorithm was applied to the wireless network intrusion detection system. Experiments show that, the improved BP algorithm enhances the accuracy and real time of intrusion detection systems.

Key words: wireless local area network; intrusion detection; neural networks; 802.11 protocol

0 引言

近年来无线局域网以其便捷性、易于扩展性和使用灵活性等特点,已经得到了飞速广泛地发展,但是其在安全方面的不足也限制了它的应用.在无线局域网中,数据是在空中传播,只要在无线接入点的覆盖范围内,终端都可以接收到无线信号,因此无线局域网的安全问题尤为突出.虽然无线局域网制定了相应的安全标准,但是并不能完全解决问题.将入侵检测系统应用于无线局域网中,是对其安全性的很好补充,也是解决当前无线网络安全问题的有效手段.无线局域网和有线网络的根本区别是其介质的开放性、物理链路的不稳定性以及 MAC 层协议的不同,这一特性使得在设计无线网络下的入侵检

收稿日期:2009-09-11;**修回日期**:2010-01-06

基金项目:国家自然科学基金(60872009,60602016),中国高技术研究发展(863)计划(2007AA01Z428,2009AA01Z148)和安徽高校省级自然科学研究计划重大项目(ZD2008005-2,ZD200904,JK2009A013,JK2009A025)资助.

作者简介:刘凤纯,男,1985 年生,硕士生.研究方向:通讯协议安全与测试. E-mail: kickct@mail.ustc.edu.cn

通讯作者:赵保华,教授. E-mail: bhzhao@ustc.edu.cn

测系统时,必须考虑其特殊性. 入侵检测系统是信息安全的重要组成部分,也是当前网络安全领域的研究热点. 传统的入侵检测系统主要是基于规则和专家知识库的方法,这些方法对于规则库和专家库的依赖比较强烈,不具有很好的自适应性[1]. 神经网络是一种模拟人的大脑的思维方法来对信息进行处理的技术,它作为数据挖掘的重要手段,具有自组织、自学习、自适应的特性,适合多变的无线网络环境下的入侵检测[2-3].

采用 BP 算法的多层感知器是应用比较广泛的一种反向传播神经网络,本文将这种神经网络引入到无线局域网的入侵检测系统中[4-5]. BP 神经网络由正向信号传播和反向信号传播两个过程组成. 正向传播时,信号从输入层到各个隐层,最后到输出层,如果输出结果和期望结果不符,则反向传播使用梯度法调整权值,反复此过程直到误差到达可以接受的程度或者预设次数为止. 图 1 显示了三层 BP 神经网络的结构图[3].

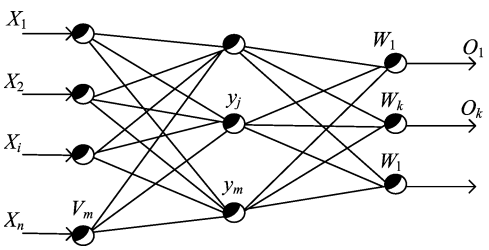

图 1　三层 BP 神经网络结构图

Fig. 1　Three-layer BP neural network structure

基于 BP 算法的多层感知神经网络能够学习和存储大量的输入和输出模式映射关系而无需事先了解描述这种映射关系的数学方程. 通过训练就可以达到从 n 维输入空间到 m 维输出空间的映射. 下面对 BP 神经网络的标准算法给出简单的描述:

(Ⅰ) 初始化权值矩阵 $\boldsymbol{W}$,$\boldsymbol{V}$ 和变量 $E_{\min}$,η,E.

(Ⅱ) 样本 $\boldsymbol{X}$ 作为输入,通过变换函数 $f(x)=\frac{1}{1+e^{-x}}$,计算出隐层输出向量 $\boldsymbol{Y}$ 和输出向量 $\boldsymbol{O}$ 的各个分量.

(Ⅲ) 计算误差值 E,若 $E<E_{\min}$,算法结束,否则继续执行.

(Ⅳ) 根据(Ⅱ)计算的结果,利用梯度下降算法,修正权值矩阵 $\boldsymbol{W}$ 和 $\boldsymbol{V}$,取下一个样本并转入步骤(Ⅱ).

由于标准算法是采用具有局部性的梯度下降方法以及变换函数具有饱和性的固有特性,使得 BP 算法天生带有平坦曲面和局部极小点问题. 本文在采用可变的变换函数的基础上,引入基于局部极小点跃迁的遍历方法,并将改进后的算法引入到无线局域网的入侵检测系统中来. 入侵检测系统首要考虑的问题是准确性,面对无线网络传播介质的不稳定性特点,这一特性的提高尤其显得重要,减小 BP 算法的误差是解决这一问题的有效途径. 下文中节 1 主要论述了改进的 BP 的缺点和改进方法;节 2 主要是入侵检测系统的设计和相关要点;最后是实验和总结.

1　改进 BP 算法

由于 BP 算法是基于梯度下降的方法来进行权值的调整,所以它天生有以下几个方面的不足:

(Ⅰ) 由于变换函数存在饱和区域,误差曲面会因为这些饱和区域出现较为平坦的部分,由于这些区域的梯度比较小,使得算法的速度比较慢,迭代次数增加;

(Ⅱ) 由于误差曲面是极其复杂的曲面函数,它存在着多个极小点,基于梯度下降的调整方法,使得算法极有可能陷入局部极小点,收敛的结果不是最优的全局极小点,最终将反映为误报率和漏报率的升高.

神经网络系统应用于入侵检测系统,最关键的部分是算法的准确性,它直接关系到入侵检测的误报率和漏报率. 针对 BP 算法的两个缺点,本文提出了改进的 BP 算法来提高算法的准确性. 由输出误差梯度表达式:

$$\Delta\omega_{jk}=-\eta\frac{\partial E}{\partial w_{jk}}=-\eta\delta_k^o y_j \tag{1}$$

可知,当误差曲面出现平坦区域的时候,误差梯度趋于零,也就是 δ_k^o 趋于零. 从文献[3]中,我们知道 δ_k^o 的公式如下:

$$\delta_k^o=(d_k-o_k)o_k(1-o_k) \tag{2}$$

式中,d_k 和 o_k 分别表示期望输出值和样本训练输出值. 从式(2)中可以看出有三种情况使得 δ_k^o 趋于零. o_k 逼近 d_k 时说明这时候应该是趋于收敛的极小点. o_k 逼近 0 或者 1 是由于变换函数具有饱和特性,所以我们在算法中引入了可调的变换函数 $f(x)=\frac{1}{1+e^{-x/\lambda}}$,$\lambda$ 是可调因子. 当 $\lambda>1$ 时,可变函数的 x 坐标被压缩,从而提高了非饱和区域的范围,跳出平

坦区域. 可调因子的调整原则如下：

①当 $x \geqslant \theta$ 时，调整 $\lambda=\theta$，其中 $\theta>1$；

②当 $x<\theta$ 时，调整为原值 $\lambda=1$.

为了解决多个极小值局部收敛问题，本文提出基于逃逸的遍历极小点的方法. 这种方法要求我们可以判断一个极小值的出现，同时能够逃逸出极小值，以达到遍历所有极小点并在所有极小点中选择最优解的目的.

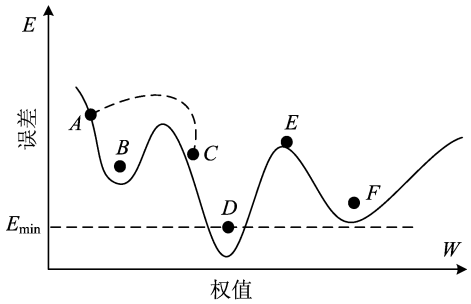

图 2 多个极小点误差图

Fig. 2 Multiple minima deviation chart

从图 2 中，我们可以看出，极小点 B,D,F 的共同点都是梯度等于零，所以如果初始权值落入 A 点，那么最终的收敛将落入 B 点，而得不到最优的全局极小点 D. 为了解决局部最小值问题，当收敛到达局部最小值 B 时，我们再返回到收敛点 B 的前一步，增大学习步长 η，使得下次收敛越过 B 点，到达 C 点，继续训练过程，最终把所有的极小点遍历，得到全局最小点.

我们知道极小点处梯度趋于零，而从上文中知道平坦区域的梯度也是趋于零，但是由于我们在算法中引入可调因子，所以不存在连续多个样本输入时都是陷入平坦区的情况. 所以我们给出极小点的判定规则：

（Ⅰ）极小点判断规则

$$\Delta E_k = E_k - E_{k-1},\ k=1,2,3,\cdots,n \tag{3}$$

$$\Delta E = \sqrt{\frac{1}{n}\sum_{k=1}^{n}(E_k-\overline{E})^2} < \theta \tag{4}$$

式中，E_k 表示第 k 个样本的误差值，ΔE_k 表示连续两个样本误差的差值. 当式(4)成立时，判断 n 个样本中，方差 E 最小的即为极小点.

（Ⅱ）极小点逃逸成功判定规则

$$|E-E_{\min}|>\varepsilon \tag{5}$$

这表明逃逸成功.

当然该规则具有失败的可能性. 如果在利用该规则后，得到的是先前跳出的极小点，表明该规则失败，执行结束算法或者其他处理. 根据以上规则，下面给出改进 BP 算法的伪代码描述：

Step 1 初始化权值矩阵，训练期望值，可调因子数组以及其他参数.

Step 2 取下一个样本对 (X,D) 做为输入样本，根据变换函数

$$f(x)=\frac{1}{1+e^{-x/\lambda}} \tag{6}$$

计算出输出 $\boldsymbol{X}$ 和 $\boldsymbol{O}$. 这里 λ 的取值按照上文可调因子的调整原则来调整，达到解决平坦区的效果.

Step 3 计算本次样本的误差 E_k，根据极小点判定规则判定当前是否达到极小点. 如果判定是前一个极小点，则退出算法；如果不是，调整学习步长 η，对极小点前一个样本对 $(X_{\text{pre}},D_{\text{pre}})$ 重新训练，并且调整权值矩阵，利用极小点跳出判定原则来跳出极小点，并转入 Step 6. 如果不是极小点则继续执行.

Step 4 利用梯度计算公式[3]

$$\Delta\omega_{jk} = \eta\delta_k^o y_j = \eta(d_k-o_k)o_k(1-o_k)y_j \tag{7}$$

$$\Delta v_{ij} = \eta\delta_j^y x_i = \eta\Big(\sum_{k=0}^{l}\delta_k^o\omega_{jk}\Big)y_i(1-y_i)x_i \tag{8}$$

计算出隐层和输出层的梯度.

Step 5 反向利用梯度来调整权值矩阵，输出层调整公式为 $\omega_{jk}=\omega_{jk}+\eta\delta_k^o y_j$，隐层调整公式为 $v_{ij}=v_{ij}+\eta\delta_j^y x_i$，同时样本计数增加 1，继续执行.

Step 6 判定是否所有样本执行完毕，如果执行完毕，转入 Step 7，否则转入 Step 2.

Step 7 计算总误差 E，如果 E 满足期望误差或者达到算法训练次数的最大值，算法结束，否则设置样本次数为 1，转入 Step 2，开始新一轮的样本训练.

2 WLAN 入侵检测系统设计

将神经网络算法应用于无线局域网，主要是基于协议数据包的分析，通过无线网卡来捕获无线数据包，然后根据 802.11 协议[6] 对报文进行解码. 由于解码后的报文并不能直接用来进行入侵检测和神经网络的训练，因此需要一个预处理的过程，提取攻击的相关特征量，转换为神经网络的输入，系统模型见图 3. 无线局域网下的攻击主要是针对 802.11 协议漏洞[7] 的攻击，而且攻击具有时间相关性的特点，所以我们将以一段时间捕获的 802.11 协议报文为基本单位来作为训练输入和检查入侵的输入.

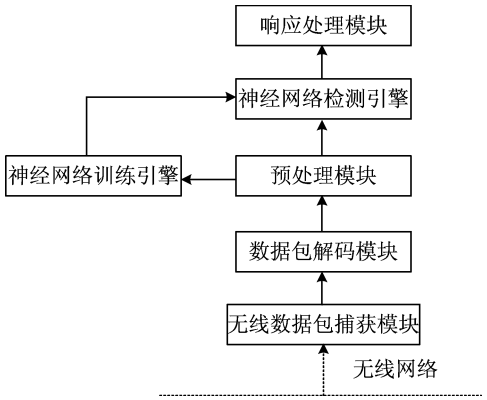

图 3　无线局域网入侵检测系统模型

Fig. 3　Wireless LAN intrusion detection system model

通过对 802.11 协议已经存在的攻击的分析，我们发现针对 802.11 协议的攻击主要是集中在管理帧上. 为了简化我们的实验，我们只是针对管理帧的特征进行提取，所以我们提取了以下一些特征：不同源地址 Association 帧的数量，Disassociation 帧的数量，Deauthentication 帧的数量，不同源地址 Authentication 帧的数量，不同源地址 Beacon 帧的数量，具有不同源地址 Reassociation 帧的数量. 为了更好地收敛，我们需要对样本数据进行归一化处理，假设输入样本 $X=[x_1,x_2,x_3,\cdots,x_i,\cdots x_n]$，那么 x_i'的计算公式为

$$x_i'=\frac{x_i-\bar{x}}{\sigma(x)} \tag{9}$$

式中，$\bar{x}$ 表示平均值；$\sigma(x)$表示标准差. 在用 BP 算法的神经网络中，根据输入特征可以确定输入节点数；根据输出的期望结果类别，可以确定输出节点数；对于隐层节点数，一般存在一些常用的经验公式：

$$m=\sqrt{n+l}+\alpha,\ m=\log_2 n,\ m=\sqrt{nl},$$

其中，m 表示隐层节点数，n 表示输入节点数，l 表示输出节点数，α 表示 1～10 之间的常数. 实验中也可以对不同隐层节点数测试，选出最优值.

3　实验及分析

本实验是以国家 863 项目无线局域网入侵检测系统为平台，分别将数据采集模块、标准 BP 算法和改进后的 BP 算法三个模块加入到系统中来进行实验.

实验中我们在无线局域网入侵检测平台的 Agent 端执行以上六个特征相关的 tcl 攻击脚本，使用 Ominipeek 工具来抓取攻击数据包和正常数据包，并且通过数据采集模块来提取训练样本和采集样本. 实验中我们对每个输入特征提取与它相关的攻击样本 20 个，正常样本 120 个，将所有样本乱序，作为训练样本. 图 4 给出了分别使用标准算法和改进算法对于不同期望误差的训练次数比较，可以看出改进后的训练次数比标准算法训练次数低.

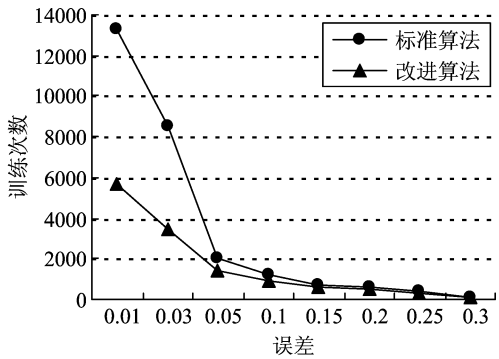

图 4　标准算法和改进算法训练次数比较

Fig. 4　Training times compared between standard algorithm and improved algorithm

样本测试中我们提取每种攻击的1 000个样本，并且混合1 000份正常样本，在训练误差为 0.05 的情况下对样本进行检测. 我们规定输出对(x_1,x_2)满足条件：

①$0.85<x_1<1$ 且 $0<x_2<0.15$ 为正常样本；

②$0<x_1<0.15$ 且 $0.85<x_2<1$ 为异常样本.

其他类型值，表示不可判断. 实验结果见表 1.

表 1　神经网络方法检查效果

Tab. 1　Neural network detection effect

攻击类型	正确检测	错误检测	检测率	误报率
断开认证	998	4	99.8%	0.4%
连接攻击	972	12	97.2%	1.2%
断开连接	993	6	99.3%	0.6%
认证攻击	996	23	99.6%	2.3%
信标攻击	964	13	96.4%	1.3%
重连攻击	987	9	98.7%	0.9%

通过实验我们发现将神经网络引入到入侵检测系统中，具有较高的检测率和较低的漏报率.

4　结论

针对 BP 算法提出逃逸出极小点和可变的变换函数的方法来改进标准算法的不足，同时将 BP 神经网络算法引入到了基于 802.11 协议的无线局域网入侵检测系统，具有较高的入侵检测率和自适应性. 由于本算法提出的极小点逃逸成功判定原则的

不可靠性,所以针对这个问题,将来需要进一步研究出更为合理有效的方法.

参考文献(References)

[1] 杨义先,钮心忻. 入侵检测理论与技术[M]. 北京:高等教育出版社,2006.

[2] Lee W, Stolfo S J. Data mining approaches for intrusion detection [C]// Proceedings of the 7th USENIX Security Symposium. San Antonio, USA: USENIX Assoc, 1998: 79-93.

[3] 韩力群. 人工神经网络教程[M]. 北京: 北京邮电大学出版社,2006.

[4] Lippmann R P, Cunningham R K. Improving intrusion detection performance using keyword selection and neural networks[J]. Computer Networks, 2000, 34: 597-603.

[5] Zhang Z, Li J, Manikopoulos C N, et al. HIDE: A hierarchical network intrusion detection system using statistical preprocessing and neural network classification[C]// Proceedings of the 2001 IEEE Workshop on Information Assurance and Security. West Point, NY: United States Military Academy, 2001:85-90.

[6] IEEE802.11 Working Group. 2007. IEEE standard for information technology- Telecommunications and information exchange between systems-Local and metropolitan area networks-Specific requirements- Part 11: Wireless LAN medium access control (MAC) and physical layer (PHY) specifications 2007[EB/OL]. [2009-08-28]. http://standards.ieee.org/getieee802/download/802.11-2007.pdf.

[7] Welch C D J, Lathrop M S D. A Survey of 802.11a Wireless Security Threats and Security Mechanisms [R]. West Point, New York: United States Military Academy, 2003.

(上接第 1 086 页)

参考文献(References)

[1] Qian Jingren, Zhang Zhichun, Zhang Hua. Twisted fiber ring rsonator of keeping Faraday rotation[J]. Acta Optica Sinica, 1998, 13(3): 324-329.
钱景仁,张志淳,张华. 保持法拉第效应的光纤环形腔结构[J]. 光学学报,1998,13(3):324-329.

[2] Kopp V I, Churikov V M, Singer J, et al. Chiral fiber gratings[J]. Science, 2004, 305: 74-75.

[3] Kopp Victor I, Churikov Victor M, Zhang G, et al. Single- and double-helix chiral fiber sensors[J]. J Opt Soc Am B, 2007, 24(10): A48-A52.

[4] Kopp V I, Churikov V M, Genack A Z. Synchronization of optical polarization conversion and scattering in chiral fibers[J]. Optics Letters, 2006, 31(5): 571-573.

[5] Qian Jingren, Guo Qing, Li Lusun. Spun linear birefringence fibres and their sensing mechanism in current sensors with temperature compensation[J]. IEE proc-Optoelectron, 1994, 141(6): 373-380.

[6] Qian Jingren, Huang Weiping. LP modes and ideal modes on optical fibers[J]. Journal of Lightwave Technology, 1986, 4(6): 626-630.

[7] Qian Jingren, Huang Weiping. Coupled-mode theory for LP modes[J]. Journal of Lightwave Technology, 1986, 4(6): 619-625.

[8] Qian Jingren, Wang Xuxu. Coupled-mode theory for spun multi-lobe stress region fibers[J]. Acta Optica Sinica, 2007, 27(3): 550-554.
钱景仁,王许旭. 多叶应力区扭转光纤的耦合模理论[J]. 光学学报,2007,27(3):550-554.

[9] Erdogan T. Cladding-mode resonances in short- and long-period fiber grating filters[J]. J Opt Soc Am A, 1997, 14(8): 1 760-1 773.

[10] Qian Jingren. On coupling coefficients of coupled-wave equations[J]. Acta Electronica Sinica, 1982(2): 46-54.
钱景仁. 论耦合波方程中的耦合系数[J]. 电子学报, 1982(2): 46-54.

[11] Qian Jingren, Li Lusun. Spun highly linearly birefringent fibres for current sensors[J]. Science in China(Series A), 1990, 33(1): 99-107.