DOI:10.13203/j.whugis2008.10.018
第33卷 第10期 武汉大学学报·信息科学版 Vol.33 No.10
2008年10月 Geomatics and Information Science of Wuhan University Oct. 2008

文章编号:1671-8860(2008)10-0995-04 文献标志码:A

网络安全态势感知关键实现技术研究

王慧强[1] 赖积保[1] 胡明明[1] 梁 颖[1]

(1 哈尔滨工程大学计算机科学与技术学院,哈尔滨市南岗区南通大街145号,150001)

摘 要:建立了网络安全态势感知的分层实现模型,并针对每个层次提出了基于多分类器融合的安全态势提取方法、基于统计学习的分层态势评估方法以及基于遗传神经网络态势的动态预测方法。经仿真实验验证,每个层次的实现方法都是可行有效的。

关键词:网络安全;态势感知;要素提取;态势评估;态势预测

中图法分类号:TP393.3

网络安全态势感知是目前网络安全领域的一个研究热点,已经引起了相关科研机构和研究人员的足够重视[1]。当前主要是围绕网络安全态势的主动、实时评估和感知进行研究,采用的方法主要有多传感器数据融合方法[2]、分层分析法[3]、流(flow)分析法[4]等,而未见系统地对网络安全态势感知关键实现技术进行探讨。基于此,本文尝试建立一个网络安全态势感知分层实现模型,自底向上依次为网络安全态势要素提取层、态势评估层和态势预测层,分别研究探讨各层的实现方法,最后在所构建的实验环境中验证了每个层次实现方法的可行性和有效性。

1 网络安全态势感知分层模型

网络安全态势感知模型是开展该领域研究的前提和基础。在对典型态势感知模型JDL功能模型[5]和Endsely的态势感知认知模型[6]分析的基础上,提出了网络安全态势感知分层实现模型,如图1所示。通过多传感器监控和采集网络流量、IDS报警信息等安全状态数据,从中提取出影响网络安全态势的安全要素;采用"先局部后整体"的评估策略,分别对服务、主机以及网络系统所受到的安全威胁进行评估,并生成相应的安全威胁等级;依据已知 $T+1, T+2, \cdots, T+n$ 时刻的网络安全态势,预测 $T+(n+1)$ 时刻的网络安全态势,使决策者能据此掌握更高层的网络安全态势,为制定合理准确的决策提供依据

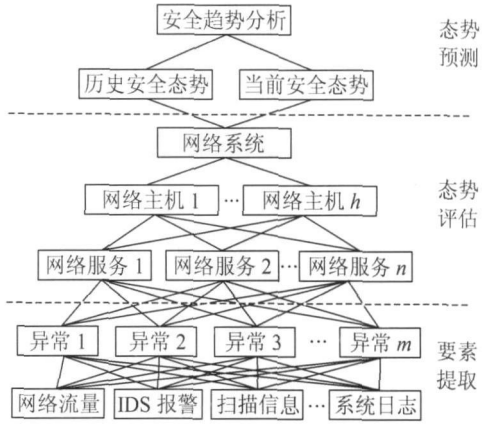

图1 网络安全态势感知分层实现模型

Fig.1 Hierarchical Implementation Model for Network Security Situation Awareness

2 基于多分类器融合的态势要素提取

由于网络安全态势易受攻击、病毒、漏洞以及人为等的影响,单纯使用一种方法或一类分类器很难在复杂的大规模网络环境中实现安全态势要素的提取,也很难保证检测效果。基于此,提出了一种基于多分类器融合的安全态势要素提取方法。如图2所示,该模型主要包含分类器和融合器两大部分,分别构建了遗传神经(BP neural network with genetic algorithm, GA-BPNN)分

收稿日期:2008-08-28。

项目来源:国家863计划资助项目(2007AA01Z401);国家教育部高校博士点专项基金资助项目(20050217007)。

类器、支持向量机(support vector machine, SVM)分类器以及模糊聚类(fuzzy cluster, FC)分类器[7-9]。而对于同一事件,不同分类器的分类结果可能不同,因此,采用DS证据理论进行进一步融合推理,首先计算各个证据的基本概率赋值函数m、信任函数Bel和似然函数Pal;然后用DS组合规则计算所有证据联合作用下的基本概率赋值函数、信任度函数和似然函数;最后根据一定的决策规则,选择联合作用下支持度最大的假设。

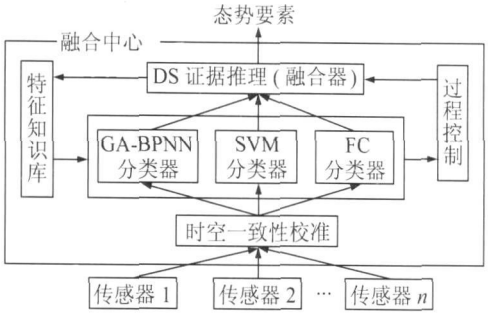

图2 多分类器融合的态势要素提取模型

Fig.2 Situation Element Extraction Model Based on Multi-Classifier Fusion

3 基于统计学习的分层态势评估

结合实际环境,按照文献[3]对网络安全态势评估所涉及到的服务威胁指数R_s、主机威胁指数R_H和网络系统威胁指数R_L给出相应的量化计算方法。

3.1 服务级

攻击对服务的安全威胁与攻击频率和攻击威胁严重程度相关。给定分析时间窗口Δt,定义t时刻服务S_j的威胁指数为:

$$R_{S_j}(t)=\overline{C_j}(t)\cdot 10^{\overline{C_j}(t)} \tag{1}$$

式中,$\overline{C_j}(t)$、$\overline{C_j}(t)$分别为t时刻攻击威胁严重程度和发生次数向量。

3.2 主机级

在时刻t主机H_k的威胁指数为:

$$R_{H_k}(t)=\mathbb{V}\cdot\mathbb{R}_s(t) \tag{2}$$

式中,$\mathbb{R}_s(t)$为t时刻主机H_k的服务安全威胁向量;V为服务在主机开通的所有服务中所占权重向量,其元素取值根据主机提供服务的重要性来确定。

3.3 网络系统级

在时刻t网络系统的威胁指数为:

$$R_L(t)=\mathbb{W}\cdot\mathbb{R}_H(t) \tag{3}$$

式中,$\mathbb{R}_H(t)$为t时刻网络系统内主机的安全威胁向量;$\mathbb{W}$为主机在被评估局域网中所占重要性的权重向量,其元素取值根据各主机在局域网中的地位来确定。

4 基于GA-BPNN的态势动态预测

在评估过去和当前网络安全态势的基础上,建立态势预测的神经网络模型,并采用改进的遗传算法对其进行优化,用于实现网络安全态势的非线性时间序列预测,具体步骤如下。

1) 依据历史和当前态势数据,定义态势预测的神经网络模型y_n^p和相应的误差函数E:

$$y_n^p=f\left(\sum_{k=1}^{K}v_{kn}\cdot f\left(\sum_{m=1}^{M}w_{mk}\cdot x_m^p-\theta_k\right)-\gamma_n\right) \tag{4}$$

$$E=\frac{1}{P}\sum_{p=1}^{P}\sum_{n=1}^{N}(y_n^p-T_n^p)^2 \tag{5}$$

式中,M、K、N分别表示输入层、隐含层、输出层的节点个数;w_{mk}表示输入层与隐含层之间的连接权值;v_{kn}表示隐含层与输出层之间的连接权值;θ_k和γ_n分别表示隐含层和输出层的阈值;f表示隐含层到输出层的Sigmoid函数,$f=\frac{1}{1+\exp(-x)}$;y_n^p和T_n^p分别代表第p个训练样本所对应的第n个实际输出和期望输出。

2) 采用遗传算法优化态势预测神经网络模型,使实际输出值与期望输出值一致,定义优化目标为:

$$f(t)=\frac{1}{1+E(t)} \tag{6}$$

式中,$t=1,2,\cdots$为种群的个体数;$f(t)$表示第t子代个体适应度值;$E(t)$表示第t子代个体误差情况。输出训练后的态势预测神经网络模型,并动态调整参数值,寻找出最优参数组合,输出预测结果。

5 仿真实验和结果分析

为了验证网络安全态势感知关键实现技术的可行性和有效性,在局域网内搭建如图3所示的实验环境,在指定时间内采用攻击软件对3个受保护服务器发起各种攻击,攻击测试软件均属于Probe、DoS、U2R和R2L 4个大类别。

5.1 网络安全态势要素提取

在实验中将攻击分成4个大类别Probe、DoS、U2R和R2L,然后再针对各个类别的事件进行二次分类得到具体的攻击类型。下面以Probe攻击为例分析该过程,Probe类攻击细分为

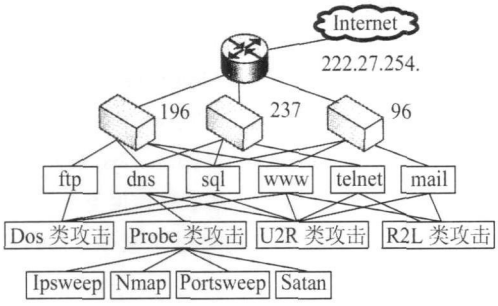

图 3　实验环境

Fig.3　Experimental Environment

Ipsweep、Nmap、Portsweep 和 Satan 4 种。表 1 给出了各个分类器的分类结果。对事件分类后的结果采用 DS 证据理论来进行融合推理,以事件分辨率作为各个分类器的信任度,即分类器分辨出来的事件个数与事件总数之比。计算可得,GA-BPNN分类器、SVM 分类器和 FC 分类器的信任度分别为 0.833 3、0.657 4 和 0.759 3。对 108 条事件进行融合后的结果,如图 4 所示,经融合后能够分辨出事件 104 个,分辨率为 0.963 0,远远高于任何单个分类器的分辨率。

表 1　各个分类器的分类结果

Tab.1　Result of Every Classifier

攻击类型	*I*	*N*	*P*	*S*
攻击数量	22	33	24	29
GA-BPNN 分辨攻击数	12	31	19	28
SVM 分辨攻击数	21	0	23	27
聚类分辨攻击数	21	31	20	10

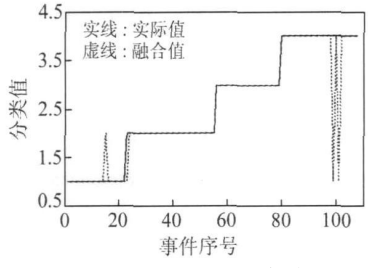

图 4　DS 融合结果

Fig.4　DS Fusion Result

5.2　网络安全态势评估

分别以实际攻击统计结果和态势要素提取结果作为态势评估的输入,评估出整个系统的安全态势。从图 5 可以直观地发现局域网系统的安全威胁状况,系统在 10:09～10:14,10:20～10:30,10:44～10:47 安全威胁指数很大,应引起管理员高度重视;同时,在 10:54～10:57 这段时间内有很明显的误报,究其原因是 196 服务器在这段时间内出现异常,而该服务器所占权值又很高,直接影响了整个网络安全态势。

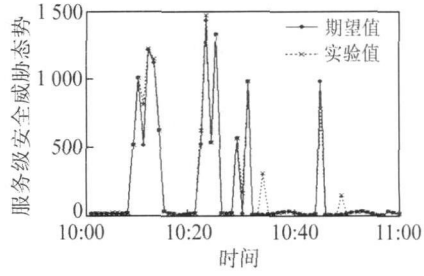

图 5　系统级安全威胁态势

Fig.5　System-Level Security Threat Situation

5.3　网络安全态势预测

依据态势评估结果,取连续的 90 个安全态势值,前 60 个作为样本,后 30 个作为测试样本。采用 GA 算法确定神经网络的权值和阈值。初始种群个数 popu=40,遗传代数 gen=100,目标误差 goal=0.001,学习速率 LP.lr=0.01,运行 33.922 s后,达到目标误差,训练步数为4 686,大约 45 代时染色体的平均适应度趋于稳定。采用训练后的态势预测模型和 BP 神经网络对测试样本进行实验,结果发现当训练步数为 200 000 时,仍未达到误差目标,均方误差为0.004 104 41。如图所示 6 所示,给出了 GA-BPNN 和 BP 算法的预测曲线。通过比较分析,GA-BPNN 算法无论是收敛速度还是运行时间都优于 BP 算法。

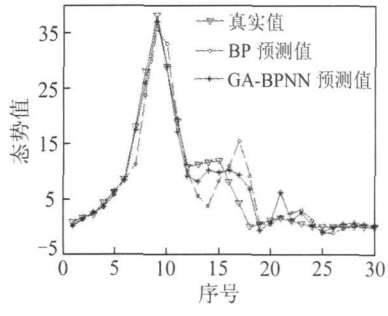

图 6　GA-BPNN 与 BP 算法的预测曲线

Fig.6　Prediction Curve of GA-BPNN and BP

参　考　文　献

[1]　王慧强,赖积保,朱亮 等. 网络态势感知系统研究综述[J]. 计算机科学, 2006, 33(10):5-10

[2]　Bass T. Intrusion Detection Systems and Multi-sensor Data Fusion: Creating Cyberspace Situational Awareness [J]. Communications of the ACM, 2000, 43(4):99-105

[3]　陈秀真,郑庆华,管晓宏,等. 层次化网络安全威胁态势量化评估方法[J]. 软件学报, 2006, 17(4):885-897

[4]　Yin Xiaoxin, Yurcik W, Slagell A. The Design of

VisFlowConnec-IP: A Link Analysis System for IP Security Situational Awareness[C]. IWIA'05, Baltimore, USA, 2005

[5] Steinburg A N, Bowman C L, White F L. Revisions to the JDL Data Fusion Model[C]. NATO/IRIS Conference, Quebee City, Canada, 1998

[6] Endsley M R. Toward a Theory of Situation Awareness in Dynamic Systems [J]. Human Factors, 1995, 37(1): 32-64

[7] Wang Huiqiang, Liu Xiaowu, Lai Jibao, et al. Network Security Situation Awareness Based on Heterogeneous Multi-sensor Data Fusion and Neural Network[C]. IMSCCS'07, USA, 2007

[8] Liu Xiaowu, Wang Huiqiang, Lai Jibao, et al. Multiclass Support Vector Machines Theory and Its Data Fusion Application in Network Security Situation Awareness[C]. WICOM'07, Shanghai, 2007

[9] 朱卫末, 王卫平, 梁楑. 基于模糊聚类分析的入侵检测方法[J]. 系统工程与电子技术, 2006(28): 474-477

第一作者简介: 王慧强, 教授, 博士生导师。主要研究方为向网络安全、自律计算。

E-mail: wanghuiqiang@hrbeu.edu.cn

Research on Key Technologies for Implementing Network Security Situation Awareness

WANG Huiqiang[1] *LAI Jibao*[1] *HU Mingming*[1] *LIANG Ying*[1]

(1 College of Computer Science & Technology, Harbin Engineering University, 145 Nantong Street, Harbin 150001, China)

Abstract: Network security situation awareness technology is a novel technology to defend attacks and intrusions and provide global network security situation in an active and real-time style. The hierarchical realization model of network security situation awareness is built. The corresponding realization method of each layer is put forward respectively, including network security situation element extraction based on multi-classifier fusion, hierarchical situation assessment based on statistical learning and dynamic situation prediction based on back propagation neural network with genetic algorithm. Experimental results show that the proposed realization methods are feasible and reasonable.

Key words: network security; situation awareness; element extraction; situation assessment; situation prediction

About the first author: WANG Huiqiang, professor, Ph.D supervisor, majors in network security and autonomic computing.
E-mail: wanghuiqiang@hrbeu.edu.cn

下期主要内容预告