基于 DBN 的网络安全态势评估和态势预测建模研究

熊中浩[1,2]，张　伟[1]，杨国玉[1]
(1.中国大唐集团科学技术研究院，北京 100040；2.大唐水电科学技术研究院有限公司，四川 成都 610031)

摘　要： 计算机通信网络技术高速发展，日新月异，随之涌现的网络攻击、破坏现象形态各异、层出不穷。态势感知系统为网络安全提供了全面保障，提高态势评估和态势预测建模的稳定性、精准性和快速性是态势感知系统研究的重要方向。深度信念网作为一种深度学习智能算法，为网络安全态势评估和态势预测的精确性、理论化带来新方向。考虑深度信念网算法采用受限玻尔兹曼机作为基础网络，逐层预训练和微调为网络核心部分。构建广义网络安全态势评估指标体系，并建立计算机通信网络安全的态势评估和态势预测数据驱动模型。通过入侵检测数据集CIC-IDS2017 进行实验仿真，验证了该模型的精准性和有效性。

关键词： 网络安全；态势评估；态势预测；深度信念网；建模仿真
中图分类号： TN03；TP393.0　　**文献标识码：** A　　**DOI：**10.16157/j.issn.0258-7998.200900

中文引用格式： 熊中浩，张伟，杨国玉. 基于 DBN 的网络安全态势评估和态势预测建模研究[J].电子技术应用，2021，47(5)：35-39，44.
英文引用格式： Xiong Zhonghao，Zhang Wei，Yang Guoyu. Research on network security situation assessment and situation prediction modeling based on DBN[J]. Application of Electronic Technique，2021，47(5)：35-39，44.

Research on network security situation assessment and situation prediction modeling based on DBN

Xiong Zhonghao[1,2]，Zhang Wei[1]，Yang Guoyu[1]
(1.China Datang Corporation Science and Technology Research Institute，Beijing 100040，China；
2.Datang Hydropower Research Institute Co.，Ltd.，Chengdu 610031，China)

Abstract： With the rapid development and rapid development of computer communication network technology，network attacks and destruction emerge in various forms and emerge in endlessly. Situation awareness system provides a comprehensive guarantee for network security. Improving the stability，accuracy and rapidity of situation assessment and situation prediction modeling is an important direction of situation awareness system research. As a deep learning intelligent algorithm，deep belief network brings new direction to the accuracy and theorization of network security situation assessment and situation prediction. Considering the deep belief network algorithm，the restricted Boltzmann machine is used as the basic network，and layer by layer pre-training and fine tuning are the core parts of the network. The generalized network security situation assessment index system is constructed，and the data-driven model of situation assessment and situation prediction of computer communication network security is established. Experimental simulation is carried out through the intrusion detection data set CIC-IDS2017 to verify the accuracy and effectiveness of the model.
Key words： network security；situation assessment；situation prediction；deep belief network；modeling and simulation

0 引言

计算机通信网络安全(网络安全)关乎国家安全和个人安全。建立一个安全、稳定、共享的网络环境是个人和国家的美好愿景。但网络建立初期到发展至今，恶意破坏网络安全的事件只增不减，且愈演愈烈，从非法入侵窃取隐私数据到入侵工控网络篡改运行参数，从经济损失到人员伤亡，危害国家安全。如 2011 年 12 月 21 日，CSDN 网站遭到黑客攻击，600 多万个明文注册邮箱被公布，造成了个人隐私数据泄露[1]。2010 年，一种针对工业控制网络系统的蠕虫病毒震网病毒大规模扩散，伊朗核设施遭到破坏，造成设备运行异常[2]。最近几年，又出现 NotPetya 勒索软件攻击，危害电网安全。传统的网络安全防护办法(如防火墙、漏洞扫描系统等)所提供的安全防御措施不能对网络安全状态进行实施评估，各种防御手段之间存在信息无法交互协同，缺乏整体性、动态性和持续性[3]。态势感知从上世纪 90 年代初发展以来，一直备受网络安全专家的重视和青睐[4]。态势感知具有全方位、全时段监测网络安全风险的能力，以网络安全大数据为基础，从全局视角监测安全威胁，既可以对当前网络安全进行评估，又可以预测将来时间的网络

安全指数，为安全威胁处理决策和行动提供依据，真正地做到防患于未然。发展至今，网络安全态势评估和态势预测是态势感知的重要研究部分，主流的研究方法有：数学理论、知识推理和模式识别，其中基于模式识别的态势评估和态势预测方法是近十年研究的热点[5]。文献[6]、[7]利用粒子群优化算法和灰色关联分析法的优点，相应地提出基于粒子群优化指标的SVM(Support Vector Machine)态势评估模型和基于灰色关联分析的SVM态势评估模型；文献[8]、[9]提出基于径向基函数和基于灰色理论的BP(Back Propagation)神经网络的网络安全态势评估模型，解决了态势要素与评估结果中的不确定性和模糊性问题，解释了态势要素间非线性映射的理论原因；文献[10]构建多维度的评价指标体系，结合卷积神经网络算法并对比验证其有效性。由于BP神经网络具有极强的非线性映射和自组织、自学习以及强泛化等特性，被众多学者青睐并提出多种改进算法的态势感知和态势预测模型[11-13]。近十年，深度学习算法研究迅猛进步，应用在网络安全态势评估和态势预测的研究也逐步显现，文献[14]提出深度自编码网络作为基分类器，改善态势要素提取机制；文献[15]、[16]较早地提出基于深度学习算法的网络安全态势评估和态势预测模型。

通过广泛的文献搜索，深度学习算法在网络安全态势评估和态势预测模型建立方面的研究不够深入，如模型架构简单、指标选取不全面和数据集陈旧、单一等问题。本文分析了深度信念网(Deep Belief Network，DBN)的特点以及在网络安全态势评估和态势预测方面应用的可行性。根据DBN在预测模型中具有非监督学习的特点，构建广义网络安全态势评估指标体系。创新性地提出DBN在训练集的动态过程和输出结果是网络安全态势评估模型的相关表达，在校验集的动态过程和输出结果是网络安全态势预测的相关表达，为后续深度学习算法支持、论证网络安全态势评估和态势预测的理论化提供思路。

1 基于DBN的态势评估和态势预测建模

1.1 DBN建模可行性分析

DBN属于深度学习算法，是机器学习和智能算法的一个重要方向。2006年，人工智能领域领军人物HINTON G E在《Science》期刊提出基于玻尔兹曼机的深度信念网，完美解决了神经网络训练时出现的梯度弥散问题[17]。HINTON G E提出无监督的贪心逐层网络参数算法通过受限玻尔兹曼机的堆叠，克服或减弱了BP神经网络算法中出现的梯度弥散现象，首次成功地训练了3层隐含层的深度神经网络，在众多测试集上均取得满意结果。DBN所具有多层网络架构能提取数据中隐含的更多特征值，训练数据的无标签化更好地展现了数据输入与输出间的关系表达，使用的对比散度(CD)算法能保证网络计算的快速收敛，满足工业上输入响应快速性的要求[18]。DBN具有的这些特点和优点满足了网络完全态势评估和态势预测的全局性、精确性和实时性的要求。

1.2 DBN概述

DBN的网络架构上为深层-前馈型-神经网络(Deep-Feedforward-Neural Networks，DFNN)，DBN建模算法的核心部分是逐层预训练(layer-wise training)和微调(fine tune)。图1详细展示了DBN训练中逐层预训练和微调部分。

从图1可以看出，逐层预训练策略就是对深度神经网络的训练参数进行剖分式学习，相近层级视为一个浅层神经网络，可以发挥浅层神经网络的快速学习得到特征值的优点，每组输出层级获取初始化参数后通过堆栈形成深度神经网络，既可以得到更多的隐含特征值表达，也可以提高网络计算速度，提高训练模型的泛化能力。DBN构架由多个RBM堆栈组成，各个层级间的参数初始化利用RBM的学习方式获得，即将RBM中的隐含层乘性偏置和权值连接矩阵直接赋予给相应层级的权值矩阵和偏置。每一个RBM得到自身最优参数的过程就是DBN无监督预训练的过程。在反向通道中，通过有监督的算法(如BP算法、wake-sleep算法)和少量带有标签的样本对整个网络进行微调。一个典型的DBN结构图如图2所示。

DBN建模中所使用的数据集分为：有类标数据集(训练集)和无类标数据集(校验集)。记为：

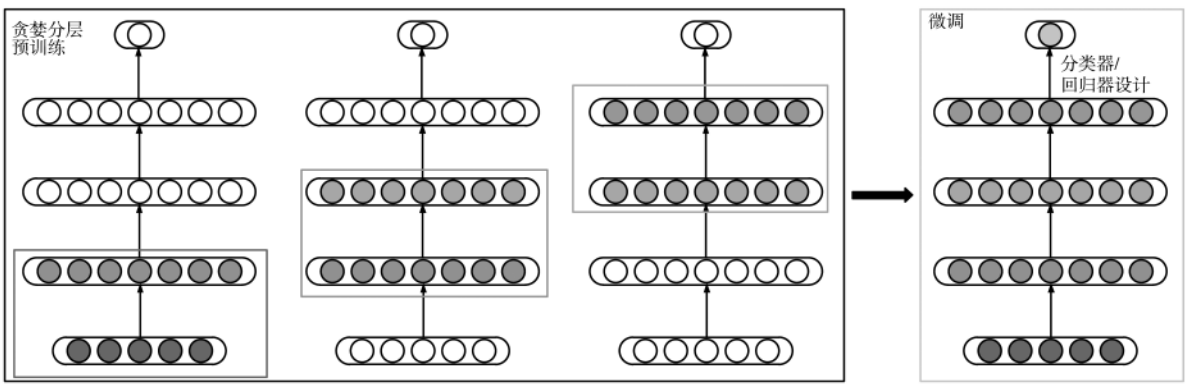

图1 DBN逐层预训练和微调

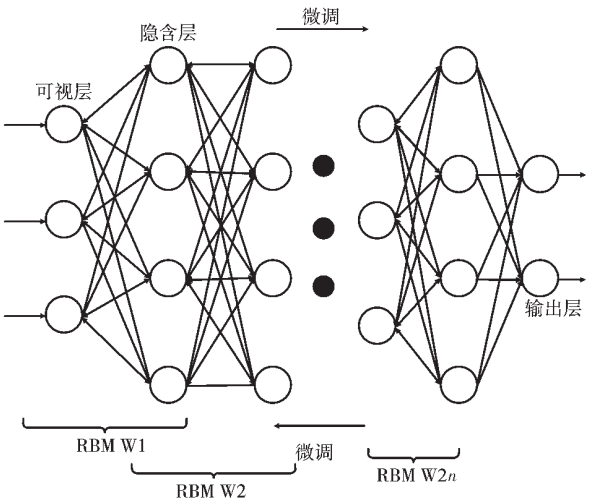

图 2 DBN 网络结构

$$\begin{cases}\{x^{(n)},y^{(n)}\}_{n=1}^{N}\rightarrow\text{训练集}\\ \{\tilde{x}^{(t)}\}_{t=1}^{T}\rightarrow\text{校验集}\end{cases}\tag{1}$$

其中，数据集的个数为 $N+T$。

2 广义网络安全态势评估指标体系建立

网络安全态势评估指标体系的建立是态势评估和态势预测的重要前提，是网络安全的基本要素表现。它作为网络安全态势评估和态势预测模型的输入部分，决定模型建立的合理性、架构的完整性以及输出结果的精确性。

2.1 指标选取

网络安全态势评估指标选一般遵循 4 个原则：独立性、完备性、科学性和主成分性原则[19]。随着网络安全态势感知系统的发展和网络攻击、威胁的升级，更多的态势评估指标被选择，发现部分指标间存在相容关系，如各主要数据包分布、子网数据流量和子网流量变化等。所提出的部分相容性原则能更好地解释指标间的内在联系和使评估、预测结果更准确。参考 GB/T 20984-2007 网络信息安全评估规范[20]并结合文献[10]建立的威胁子态势、基础运行子态势和脆弱子态势包含的 17 个二级指标以及文献[21]建立的 33 个一级指标，构建脆弱性子态势、容灾性子态势、威胁性子态势和稳定性子态势 4 个一级指标和 38 个二级指标的广义网络安全态势评估指标体系，如图 3 所示。

2.2 指标量化

部分二级指标可以直接数据集或产品资料中获取，如子网数据流量、带宽使用率等。其他二级指标不能直

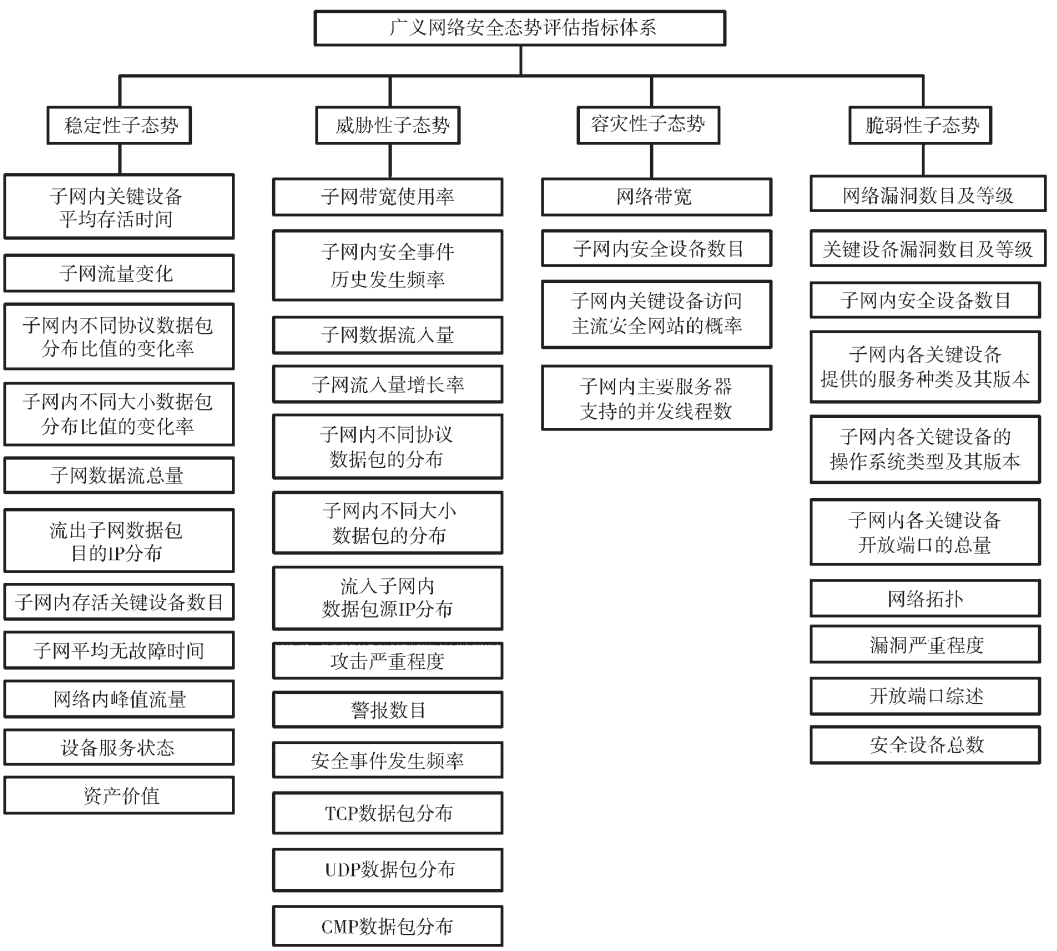

图 3 广义网络安全态势评估指标体系

接得到数据资料，需要进行数学量化转换。参考 CVSS (Common Vulnerability Scoring System)标准，部分二级指标的量化公式如下。

(1)攻击严重程度：考虑主机总数 N、攻击总数 A，量化公式如下：

$$Y=\frac{\sum_{j=1}^{N}\sum_{i=1}^{A}10^{P_{ji}}Q_jC_{ji}}{A} \tag{2}$$

$$I_j=\begin{cases}1.0 & \text{机密}\\0.7 & \text{重要}\\0.3 & \text{普通}\end{cases} \tag{3}$$

$$Q_j=\frac{I_j}{\sum_{j=1}^{N}I_j} \tag{4}$$

其中，Y 表示攻击的严重程度，数值越大，受攻击越严重；C_{ji} 为第 j 个主机受攻击 i 的次数；P_{ji} 为第 j 个主机受第 i 种攻击时的攻击等级因素；Q_j 表示第 j 个主机的重要程度；I_j 表示主机资料的重要性。

(2)漏洞严重程度：考虑漏洞总数 A、漏洞种类数 M、漏洞等级因素 W_{ji}，量化公式如下：

$$L=\frac{\sum_{j=1}^{N}\sum_{i=1}^{M}10^{W_{ji}}Q_jD_{ji}}{A} \tag{5}$$

其中，L 表示漏洞严重程度，D_{ji} 表示第 i 种漏洞在第 j 个主机上的个数。

为了消除数据样本中存在的奇异数据和消除由于量纲不同而带来的影响，对所有指标数据进行离差标准归一化处理，使所有指标数据在(0，1)范围内：

$$x'=\frac{x-\min(x)}{\max(x)-\min(x)} \tag{6}$$

其中，$\min(x)$为需要处理数据中最小的数据，$\max(x)$为需要处理数据中最大的数据，x'为归一化后的数据，x 为需处理的数据集。

3 模型建立及实验分析

DBN 具有深层堆栈式 RBM 网络架构，对复杂函数的逼近表现出快速性、简洁性。因为使用 RBM 作为核心基础网络层，更好地体现出态势评估指标独立性、相容性、主成分性原则。DBN 属于无监督深度学习网络，其训练过程是网络安全态势评估过程的体现；使用已训练的权值和偏置应用在校验过程，是态势预测过程的体现。

3.1 态势评估模型建立

本文所建立的网络安全态势评估 DBN 模型如图 4 所示。

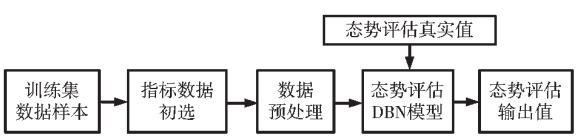

图 4 态势评估 DBN 模型

训练集数据样本：选用经过数据处理的 CIC-IDS2017 入侵检测数据集，选取 90 000 组数据作为训练集数据。

指标数据初选：根据所建立的广义网络安全态势评估指标体系，选取 33 个二级指标，考虑态势评估具有时序效应，每个指标再多选取 3 个采样时间的数据作为输入。一共选取输入维数 33×4=132 个。

数据预处理：对所选取的 132 位输入数据进行数学公式化和离差归一化处理，处理后均为在(0，1)区间上的有效数据。

态势评估 DBN 模型：采用 132-500(40 层)-4-1 结构的 DBN 架构，即 1 个输入层包含 132 个神经元；41 个隐含层，前 40 个隐含层每层包含 500 个神经元，最后一个隐含层包含 4 个神经元，体现出对二级指标具有 1 级指标的分类效应；1 个输出层，输出态势评估值。考虑网络计算的快速性，激活函数选用 ReLU 函数，初始权值均设为满足正态分布 $N(0，0.1)$的随机数，可见层和隐含层的初始权值均设为 0，采用一步 CD 算法，即 CD-1 算法。加入态势评估真实值数据对采用 weak-sleep 算法对输出结果进行微调，对整体网络的权值和偏置进行优化。保存训练好的权值 w 和偏置 b。

态势评估输出值：输出在(0，1)间的态势评估值。

3.2 态势预测模型建立

将态势评估模型训练好的权值 w 和偏置 b 应用在态势预测模型上，所建立的网络安全态势预测模型如图5 所示。

图 5 态势预测 DBN 模型

校验集数据样本：选用经过数据处理的 CIC-IDS2017 入侵检测数据集，选取 10 000 组数据作为校验集数据。

指标数据初选方法和数据预处理方法与态势评估建模过程选用方法相同。

态势预测模型：使用态势评估模型训练好的权值 w 和偏置 b 作为态势预测模型所使用的权值和偏置，预测建模过程无需使用态势评估真实值进行微调。其他 DBN 网络设置与态势评估 DBN 模型参数一致。

态势预测输出值：输出在(0，1)间的态势预测值。

3.3 CIS-IDS2017 数据集介绍及处理

CIC-IDS2017 数据集共包含周一至周五 5 天的攻击和正常活动，总量为 55 GB，具有完整的网络配置、完整的流量统计、标签数据集、完整的交互、完全捕获、多可用协议、攻击多样性和异构性。需要处理 PCAP 和 CSV 格式的文件。

为了保证实验的有效性，将数据集分割成 5 200 个时间片，在时间片中进行数据处理，进行相关态势要素提取，对所提取的二级指标数据进行数学公式量化处理

成 DBN 模型可以使用的训练集和校验集数据。PCAP 文件使用 Wireshark 工具进行文件回放，CSV 文件使用 Excel 相关函数进行处理，对复杂数据进行预处理。经过预处理后的二级指标输入数据部分如表 1 所示。

表 1 部分二级指标样本输入值

部分二级指标	样本 30	样本 300	样本 3 000
供给严重程度	0.478	0.799	0.237
报警数目	0.246	0.483	0.504
子网数据流量	0.637	0.275	0.628
漏洞严重程度	0.422	0.688	0.592

3.4 实验结果及分析

共制作 5 200 个样本，其中划分 3 000 个为训练组样本(即态势评估模型数据样本)、2 200 作为校验组样本(即态势预测模型数据样本)。网络安全态势评估模型训练结果如图 6 所示。可以看出模型训练是成功的，总体效果是不错的，在峰值、拐点处存在极少评估失准的情况。DBN 态势评估模型训练时间为 65.32 s。

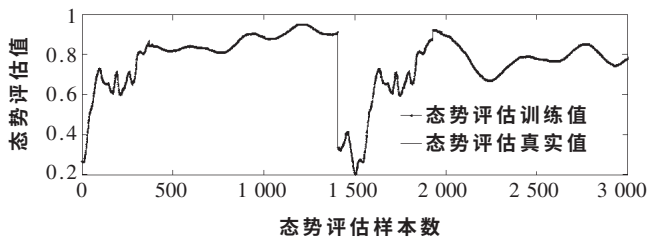

图 6 DBN 网络安全态势评估模型训练图

网络安全态势预测模型使用网络安全态势评估模型训练成功的 DBN 的权值和偏置。网络安全态势预测模型校验结果如图 7 所示。可以看出模型校验是成功的，总体效果是优良的，未出现预测失准的情况。DBN 态势预测模型训练时间 25.45 s。

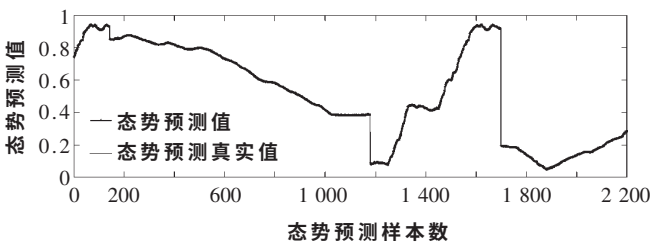

图 7 DBN 网络安全态势预测模型校验图

预测曲线直观表现了 DBN 网络安全态势预测模型的精准性和泛化能力。为了在数据上量化表现模型的精确程度，采用 R-Square 决定系数来衡量，公式如下：

$$R^2=\frac{\sum(Y_{预测值}-Y_{平均值})^2}{\sum(Y_{真实值}-Y_{平均值})^2} \tag{7}$$

其中，$Y_{预测值}$表示态势预测的预测值，$Y_{平均值}$表示态势预测的均值，$Y_{真实值}$表示态势预测的真实值，R^2 表示 R 方值。根据公式所示，R 方值越接近 0，表示曲线拟合越精准，效果越好。该 DBN 模型的 R 方值为 0.001 844 2。

4 结论

本文从网络安全态势评估和态势预测的深度信念网建模方法展开深入研究，具体如下：

(1)研究了深度信念网的数据驱动建模方法和其在网络安全态势评估和态势预测建模的可行性分析。DBN 所使用的基础网络受限玻尔兹曼机以及逐层预训练和微调核心算法在理论上符合态势评估和态势预测的动力学表达，使建立的模型更稳定、精准和快速。

(2)分析了网络安全态势评估指标选取的 4 个原则，构建了 4 个一级指标和 38 个二级指标的广义网络安全态势评估指标体系。

(3)建立基于 DBN 的网络安全态势评估和态势预测模型。对 CIC-IDS2017 数据集进行预处理，使用 5 200 组数据样本，其中3 000 组作为训练组，2 200 组作为校验组。根据实验仿真结果，模型建立成功，其精度、速度令人满意。

本文使用 DBN 对网络安全态势评估和态势预测进行建模仿真，在研究过程中，发现有两点可以进一步研究：

(1)DBN 模型算法的优化，如增加学习动量项、考虑模型稀疏性等。

(2)网络安全态势评估指标体系的优化，如建立三级指标体系、考虑各指标间更加复杂的联系关系等。

参考文献

[1] 甘利杰，孔令信，马亚军.大学计算机基础教程[M].重庆：重庆大学出版社，2017.

[2] 李东.震网病毒事件浅析及工控安全防护能力提升启示[J].网络安全技术与应用，2019(1)：9-10，24.

[3] 胡国良，肖刚，张超.新形势网络安全管理存在的问题及应对建议浅析[J].网络安全技术与应用，2020(8)：7-8.

[4] ENDSLEY M R.Design and evaluation for situation awareness enhancement[C].Proceedings of the Human Factors Society annual meeting.Sage CA：Los Angeles，CA：SAGE Publications，1988：97-101.

[5] 龚正虎，卓莹.网络态势感知研究[J].软件学报，2010，21(7)：1605-1619.

[6] 汪材印.灰色关联分析和支持向量机相融合的网络安全态势评估[J].计算机应用研究，2013，30(6)：1859-1862.

[7] 陈善学，杨政，朱江，等.一种基于累加 PSO-SVM 的网络安全态势预测模型[J].计算机应用研究，2015，32(6)：1778-1781.

[8] 谢丽霞，王亚超，于巾博.基于神经网络的网络安全态势感知[J].清华大学学报(自然科学版)，2013，53(12)：1750-1760.

(下转第 44 页)

[4] VIOLA P A,JONES M J.Robust real-time face detection[C].ICCV 2001.Proceedings of Eighth IEEE International Conference on Computer Vision,2001.
[5] REN S,HE K,GIRSHICK R,et al.Faster R-CNN:towards real-time object detection with region proposal networks[J].IEEE Transactions on Pattern Analysis and Machine Intelligence,2015,39(6):1137-1149.
[6] REDMON J,FARHADI A.YOLOv3:an incremental improvement[R].arXiv e-Prints,2018.
[7] LIU W,ANGUELOV D,ERHAN D,et al.SSD:single shot multibox detector[C].European Conference on Computer Vision,2016.
[8] CAI Z,VASCONCELOS N.Cascade R-CNN:delving into high quality object detection[C].2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition,2018:6154-6162.
[9] ZHANG K,ZHANG Z,LI Z,et al.Joint face detection and alignment using multitask cascaded convolutional networks[J].IEEE Signal Processing Letters,2016,23(10):1499-1503.
[10] 孙贵华,陈淑荣.一种改进的RefineDet多尺度人脸检测方法[J].电子技术应用,2019,45(8):34-39.
[11] 王静波,孟令军.卷积神经网络人脸检测算法[J].电子技术应用,2020,46(1):34-38.
[12] 林志文,林志贤,郭太良,等.基于FPGA加速的卷积神经网络识别系统[J].电子技术应用,2020,46(2):24-27.
[13] 张雷,王越.嵌入式平台下的车辆跟踪系统设计[J].电子技术应用,2019,45(11):13-16.
[14] 陈辰,严伟,夏珺,等.基于FPGA的深度学习目标检测系统的设计与实现[J].电子技术应用,2019,45(8):40-43,47.
[15] 柳永翔,付晓峰,付晓鹃,等.深度可分离CNN在表情识别中的应用研究[J].工业控制计算机,2020,33(10):71-73,76.
[16] 童星,张激.基于SSD-MobileNet模型的ROS平台目标检测[J].计算机系统应用,2019,28(1):94-99.
[17] 葛雯,张雯婷,孙旭泽.基于Jetson TX1的目标检测系统[J].沈阳工业大学学报,2019,41(5):539-543.
[18] 许喜斌.结合R-DAD和KCF的行人目标跟踪改进算法[J].智能计算机与应用,2019,9(4):263-266,270.
[19] HOWARD A,ZHU M,CHEN B,et al.MobileNets:efficient convolutional neural networks for mobile vision applications[J].ArXiv abs/1704.04861,2017.

(收稿日期:2020-11-11)

作者简介:
祁星晨(1995-),男,硕士,主要研究方向:深度学习、目标检测。
卓旭升(1967-),男,博士,副教授,主要研究方向:智能感知、智能控制等。

扫码下载电子文档

(上接第39页)

[9] 邓勇杰,文志诚,姜旭炜.基于灰色理论和BP神经的网络安全态势预测[J].微型机与应用,2015,34(20):1-3,8.
[10] 朱晨飞.基于神经网络的网络安全态势评估与预测方法研究[D].北京:中国人民公安大学,2019.
[11] 陈维鹏,敖志刚,郭杰,等.基于改进的BP神经网络的网络空间态势感知系统安全评估[J].计算机科学,2018,45(S2):335-337,341.
[12] 郭文忠,林宗明,陈国龙.基于粒子群优化的网络安全态势要素获取[J].厦门大学学报(自然科学版),2009,48(2):202-206.
[13] 李天骐.基于神经网络的网络安全态势评估与预测技术研究[D].北京:华北电力大学(北京),2016.
[14] 朱江,明月,王森.基于深度自编码网络的安全态势要素获取机制[J].计算机应用,2017,37(3):771-776.
[15] 周长建,司震宇,邢金阁,等.基于Deep Learning网络态势感知建模方法研究[J].东北农业大学学报,2013,44(5):144-149.
[16] 俞中华,杨晓东.基于深度自编码网络的网络安全态势感知与预警机制[J].广播电视网络,2020,27(6):63-65.
[17] HINTON G E,SALAKHUTDINOV R R.Reducing the dimensionality of data with neural networks[J].Science,2006,313(5786):504-507.
[18] 乔俊飞,潘广源,韩红桂.一种连续型深度信念网的设计与应用[J].自动化学报,2015,41(12):2138-2146.
[19] LI F C,JIANG Y X,ZHOU D Q.The building model of decision on the core competitive capacity of enterprises and evaluating demonstration[J].Business Economics and Administration,2006(6):42-46.
[20] 中国国家标准化管理委员会.GB/T20984-2007信息安全技术信息安全风险评估规范[S].北京:中国标准出版社,2007.
[21] 王娟,张凤荔,傅翀,等.网络态势感知中的指标体系研究[J].计算机应用,2007(8):1907-1909,1912.

(收稿日期:2020-09-11)

作者简介:
熊中浩(1994-),男,硕士研究生,助理工程师,主要研究方向:智能算法、优化控制。
张伟(1976-),男,硕士研究生,高级工程师,主要研究方向:网络安全。
杨国玉(1980-),男,硕士研究生,高级经济师,主要研究方向:信息化与网络安全管理。

扫码下载电子文档

版权声明