基于融合模型的网络安全态势感知方法

郭尚伟，刘树峰，李子铭，欧阳德强，王宁*，向涛
（重庆大学计算机学院，重庆 400044）

摘　要：伴随着网络技术的飞速发展，网络安全面临的风险也日益增加，网络攻击呈现复杂化、多样化的特征，给现有网络攻击应对措施带来了巨大挑战。态势感知技术作为一种新兴概念，为网络安全领域带来了新的思路。针对现有网络安全态势感知方法存在数据特征提取及较长时间序列数据处理能力不足的问题，提出一种融合堆栈稀疏自编码器（SSAE）、卷积神经网络（CNN）、双向门控循环单元（BiGRU）和注意力机制（AM）的模型。通过 SSAE 和 CNN 提取数据特征，利用 AM 强化 BiGRU 对关键信息的关注度，实现对异常流量的攻击类别判定，并结合网络安全态势量化指标，对网络安全态势进行量化评分并划分等级。实验结果表明，融合模型在各项指标上均优于传统深度学习模型，能够准确感知网络态势。

关键词：态势感知；威胁检测；堆叠稀疏自编码器；卷积神经网络；双向门控循环单元；注意力机制
中图分类号：TP309　　**文献标志码：**A　　**DOI：**10.19678/j.issn.1000-3428.0069758

Network Security Situation Awareness Method Based on Fusion Model

GUO Shangwei, LIU Shufeng, LI Ziming, OUYANG Deqiang, WANG Ning*, XIANG Tao
(College of Computer Science, Chongqing University, Chongqing 400044, China)

【Abstract】 Cybersecurity threats are becoming increasingly prevalent with the rapid advancement of Internet technologies. Cyberattacks exhibiting high complexity and diversity, are posing significant challenges to existing defense mechanisms. As an emerging concept, situation awareness technology offers new approaches to enhancing cybersecurity defense. However, the current cybersecurity situation awareness methods suffer from limited data feature extraction capabilities and inadequate handling of long-term sequential data. To address these issues, this study proposes a fusion model that integrates Stack Sparse Auto-Encoder (SSAE), Convolutional Neural Network (CNN), Bidirectional Gated Recurrent Unit (BiGRU), and Attention Mechanism (AM). By utilizing SSAE and CNN to extract data features and enhancing the focus on critical information through the AM in the BiGRU model, the proposed model aims to classify the attack categories of abnormal traffic. In conjunction with the network security situational quantification indicators proposed in this study, the network security situation is quantitatively evaluated and classified. The experimental results demonstrate that the proposed fusion model outperforms traditional deep learning models in various metrics, enabling an accurate perception of the network situation.

【Key words】 situation awareness; threat detection; Stack Sparse Auto-Encoder (SSAE); Convolutional Neural Network (CNN); Bidirectional Gated Recurrent Unit (BiGRU); Attention Mechanism(AM)

0　引言

随着互联网技术的飞速发展和各类网络应用的广泛普及，网络安全问题已经成为全球重点关注的热点议题。在此背景下，网络安全面临的挑战日益严峻，网络攻击呈现出攻击形式多样化、攻击手段复杂化等特征，由此引发的安全事件频频发生，给经济发展乃至国家安全带来了巨大的挑战[1-2]。面对技术日臻成熟、形式日渐复杂的网络攻击，防火墙、漏洞扫描、入侵检测系统等传统的网络安全防护措施已经难以全面有效地应对这些复杂多变的安全威胁。

为了应对上述挑战，网络安全态势感知作为一种新兴的概念和技术应运而生，逐渐成为网络安全研究的重要课题[3]。网络安全态势感知技术旨在通过综合利用现有网络环境的各类信息资源，包括网络流量数据、安全事件日志、威胁情报等，通过数据融合、分析和建模，实现对网络安全状态的全面评估和对未来态势的准确预测。这有助于网络管理员实时掌握网络安全状况，根据分析结果预测潜在的安

收稿日期：2024-04-17　　**修回日期：**2024-07-12
基金项目：国家重点研发计划（2022YFB3103501）；国家自然科学基金（62101079）；重庆市自然科学基金面上项目（cstc2021jcyj-msxm0465）；重庆市留学人员回国创业创新支持计划（cx2021012）。
通信作者 E-mail：nwang5@cqu.edu.cn

计算机工程 www.ecice06.com

全威胁，从而提前做好防范准备，有效降低安全事件带来的各项损失。

考虑到网络安全态势的复杂性和动态变化性等特征，传统的基于规则的安全评估方法已经难以适应高度多样化、复杂化的网络环境。因此，研究人员开始将机器学习等先进的技术方法引入网络安全态势感知领域。

胡庆爽等[4]提出一种优化置信规则库分级的网络安全态势预测方法，能够有效避免过拟合现象，提高网络安全态势预测精度。李欣等[5]将人群搜索算法与 Baum-Welch 参数优化算法相结合，优化隐马尔可夫模型的参数，提高网络安全态势评估的准确性。丁华东等[6]提出一种信息融合模型，通过对网络环境中收集到的态势指标数据进行预处理，并通过贝叶斯网络模型进行层次化融合，实现网络安全态势的准确评定，提高网络环境的稳定性和可靠性。ZHANG 等[7]采用基于威胁情报与随机游戏算法构建的态势感知模型，通过比较外来威胁情报与系统内部安全事件，准确反映和预测网络安全态势变化。王金恒等[8]通过遗传算法优化概率神经网络的修正因子，建立网络安全态势评估模型，有效提高了评估准确率和训练速度。CHEN 等[9]通过模拟退火和混合层次遗传算法优化径向基函数（RBF）神经网络，构建了高预测准确性的网络安全态势感知模型，有效应对多维非线性数据挑战，提升了网络安全防护能力。

随着网络安全挑战日益复杂，研究人员开始利用深度学习方法探索网络安全态势评估问题。杨宏宇等[10]提出一种基于深度自编码器（AE）和深度神经网络的网络安全态势评估方法，提高了网络攻击检测的准确性和效率。陈解元[11]提出一种基于卷积神经网络（CNN）和长短期记忆（LSTM）网络的异常流量检测方法，充分发掘攻击流量的结构化特点，提取流量数据的时空特征，提高了异常流量检测性能。杨宏宇等[12]设计改进并行特征提取网络和注意力机制的双向门控循环单元模型，提高网络威胁检测的精确率和召回率。WANG 等[13]基于贝叶斯网络和 LSTM 神经网络算法，对网络信息传输的安全态势进行智能计算和预测，实现了较高的检测率和召回率。SOKOL 等[14]通过比较统计模型和神经网络模型，发现神经网络方法在网络安全态势感知预测中更为准确。LU 等[15]开发了一种基于深度自编码的网络安全评估系统，使用深度强化学习算法，显著提升了网络攻击的检测效率和准确性，有效优化了网络安全评估方法。杨志鹏等[16]等通过自注意力机制和 RBF 神经网络，使模型能够有效识别并强调关键节点，增强模型对全局安全态势的感知能力。

尽管近年来基于深度学习等技术的网络安全态势感知方法取得了一定的进展，但现有研究仍面临着一些挑战。首先，在数据特征提取方面，现有的深度学习模型存在一定的缺点，尤其是提取多种类型的网络攻击数据特征的能力不足，限制了它们在网络安全态势感知中的应用效果。其次，对于较长的时间序列数据，现有方法无法有效处理时间序列数据中的时间依赖性，难以充分利用上下文信息，例如网络流量的长期依赖性和模式变化，影响了模型识别基于时间的复杂攻击活动（如横向移动或多阶段攻击）的精确性。

基于上述问题，本文提出一种融合堆栈稀疏自编码器（SSAE）、CNN、双向门控循环单元（BiGRU）和注意力机制（AM）的模型，以解决现有网络安全态势感知方法在数据特征提取方面以及长时间序列处理方面的不足。本文提出的基于融合模型的网络安全态势感知方法结合 SSAE 和 CNN 的局部特征提取能力、BiGRU 在处理时间序列数据上的优势，以及注意力机制在强化关键信息上的作用，显著提高了数据特征提取的能力，并且能够充分利用上下文信息，有效处理较长的时间序列数据。

1 网络安全态势感知方法框架

本文提出的网络安全态势感知方法由特征提取、威胁分析、态势评估 3 个模块组成，如图 1 所示。

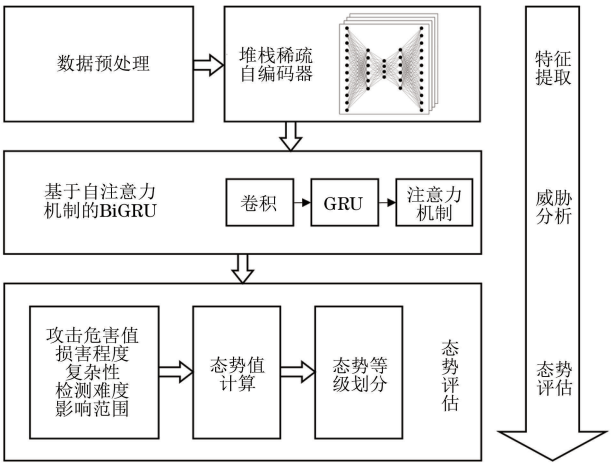


图 1 网络安全态势感知方法框架

Fig. 1 Framework of network security situation awareness method

1）特征提取模块：对网络流量信息进行归一化等预处理后，使用 SSAE 对网络流量进行特征提取，将结果输出至威胁分析模块。

2）威胁分析模块：融合 CNN、BiGRU 和 AM，

计算机工程 www.ecice06.com

判断流量数据是否为网络攻击异常流量，并判定其攻击类别。

3)态势评估模块：对各类别网络攻击的危害性进行评估，根据威胁分析模块的输出结果计算网络安全态势值，并进行态势等级划分。

本文主要贡献如下：

1)提出一种融合 SSAE、CNN、BiGRU 和 AM 的模型，显著提升了网络攻击的检测效率和准确性。

2)提出一种基于攻击严重度、攻击影响度、节点中心度和节点重要度的网络安全态势评估量化方法，能够准确地反映网络中不同节点的安全情况。

2 基于 SSAE 的特征提取方法

SSAE 是一种深度神经网络模型，通过多层自编码器，学习数据的高层次特征表示，捕捉复杂的抽象特征[17]。它通过在自编码器的隐藏层引入稀疏性约束，促使神经网络学习到重要性高但数量较少的激活神经元，从而获取更加有效的数据表示。

AE 的训练过程涉及前向传播和后向传播。在前向传播过程中，数据通过网络的每一层进行转换，并通过最后一层生成重构数据。在后向传播过程中，比较原始输入和重构输出，根据损失函数对网络的权重进行更新，以最小化误差。常用的损失函数是均方误差，它量化了输入数据和重构数据之间的差异，计算方法如下：

$$A_{\mathrm{AE}}=\frac{1}{2n}\sum_{i=1}^{n}\left\|x^{(i)}-x'^{(i)}\right\|^2 \tag{1}$$

式中：x 是原始数据点；x'是对应的重构数据点。

SSAE 的目标函数由重构误差项和稀疏性惩罚项两部分构成。重构误差项用来度量输入数据和重构数据之间的差异，稀疏性惩罚项用来确保隐藏层表示的稀疏性。此惩罚项使用 Kullback-Leibler(KL)散度，度量实际激活度和目标激活度之间的差异。具体的目标函数如下：

$$A_{\mathrm{SSAE}}=A_{\mathrm{AE}}+\beta\sum_{i=1}^{s}\mathrm{KL}(\rho\|\hat{\rho}_i) \tag{2}$$

式中：$\mathrm{KL}(\rho\|\hat{\rho}_i)$ 是稀疏性惩罚项，ρ 是目标激活度，$\hat{\rho}_i$ 是实际激活度。

ρ 和$\hat{\rho}_i$ 之间的 KL 散度定义如下：

$$\mathrm{KL}(\rho\|\hat{\rho}_i)=\rho\ln\left(\frac{\rho}{\hat{\rho}_i}\right)+(1-\rho)\ln\left(\frac{1-\rho}{1-\hat{\rho}_i}\right) \tag{3}$$

在威胁检测任务中，攻击类型的多样性对于模型的特征提取能力提出了更高的要求。单个 SSAE 无法充分捕获到各个不同的攻击类型独特的数据分布特征。为了有效应对这种多样性，采用并行的多个 SSAE，如图 2 所示。针对每种攻击类型，分别训练一个独立的 SSAE，使得每个 SSAE 都能够专注于识别其对应攻击类型的特有特征和分布。这不仅提高了模型的特征提取效率，而且增强了模型对于不同攻击类型的辨识能力。

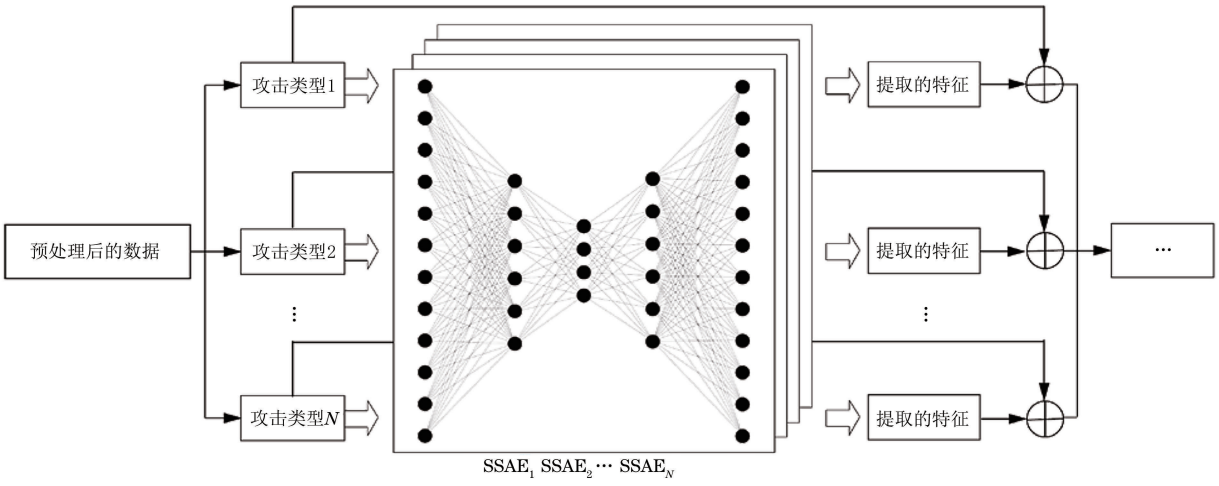


图 2 并行 SSAE 结构

Fig. 2 Parallel SSAE structure

具体实施过程如下：

1)对网络流量数据进行特征数值化、独热编码、数值归一化等预处理，确保这些数据能被并行特征提取网络进行有效处理。

2)创建一个由多个 SSAE 组成的并行特征提取网络，并初始化网络参数。

3)将训练数据按攻击类型分别输入相应的特征提取单元进行处理，对数据进行编码和解码，计算输入数据和重构数据之间的差异。

4)重复上述过程，利用梯度下降法优化网络参数，对并行特征提取网络进行训练。

5)在测试过程中，面对未知攻击类型的测试数

计算机工程 www.ecice06.com

据，将其输入到训练完成后的并行特征提取网络中的每个特征提取单元进行分析，输出一系列重构误差。选取其中误差最小的特征，将提取的特征与原始数据进行特征融合，传递至威胁分析模块。

3 基于融合模型的威胁检测方法

BiGRU 是一种有效处理序列数据的网络，能够通过双向结构捕获前向和后向的依赖关系[18]。然而，BiGRU 存在一定的局限性。虽然它能够关注到周围的信息，但它对于较长序列中的关键信息可能存在敏感性不足的问题，特别是当序列足够长时，模型可能会丧失对早期或者远端信息的关注。

为了克服这一问题，自注意力机制被引入模型[19]。自注意力机制使模型能够动态地关注输入序列中的不同部分，为不同的时间步赋予不同的权重，从而更有效地捕获长距离的依赖关系，即使关键信息位于序列的远端。自注意力机制通过对序列中各个部分的重要性进行自我学习，显著提高了模型对关键信息的捕捉能力，特别是在处理较长的输入序列时。

本文提出了一个结合卷积层的特征提取能力、门控循环单元(GRU)的时序信息处理能力以及自注意力机制的信息突出能力的融合模型，如图 3 所示。模型通过卷积层处理输入，捕获局部特征；通过 GRU 层，捕获序列的时间动态；通过自注意力模块对序列中的每个部分赋予不同的权重，以突出最关键的信息。此设计使得模型不仅可以有效捕获数据的局部结构和全局依赖，而且可以自适应地关注序列中的关键信息，有效处理网络流量信息中复杂的序列数据。

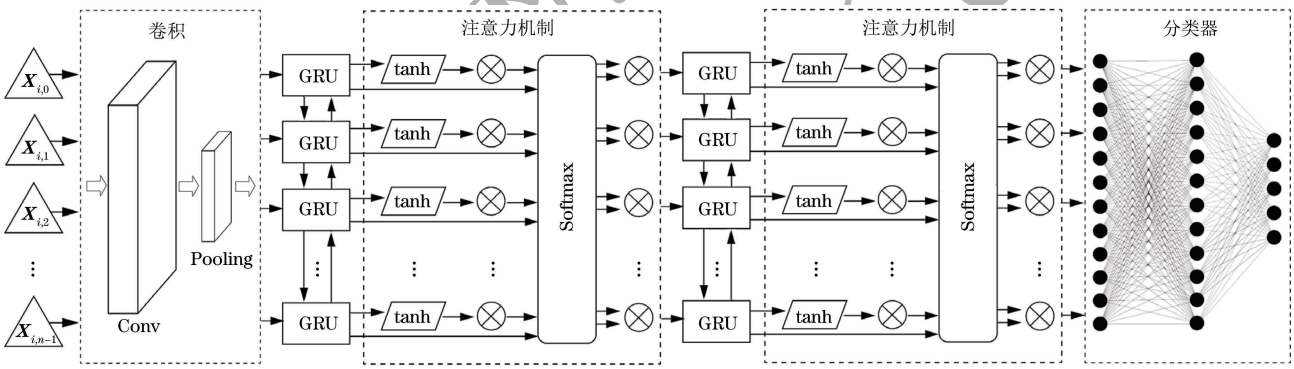


图 3 威胁检测模型结构

Fig. 3 Threat detection model structure

具体流程如下：

1)使用卷积层(Conv)处理输入数据，通过学习数据的局部模式来提取特征。利用池化层(Pooling)降低这些特征的维度，以减少后续计算的复杂性。

2)将卷积层输出的特征输入到一系列的双向 GRU 层中。每一个 GRU 单元都更新其隐藏状态 $\boldsymbol{h}_t$，以便捕获输入序列的长期依赖关系：

$$\boldsymbol{h}_t=\mathrm{GRU}(\boldsymbol{x}_t,\boldsymbol{h}_{t-1}) \tag{4}$$

式中：$\boldsymbol{h}_t$ 是前一时间步的隐藏状态；$\boldsymbol{x}_t$ 是当前时间步的输入特征。

3)通过自注意力模块，计算序列中各个时间步的注意力权重，并产生加权的输出。

首先，利用权重矩阵 $\boldsymbol{W}$ 对每个时间步的特征进行线性变换，并通过 tanh 激活函数计算变换后的特征 $\boldsymbol{u}_t$：

$$\boldsymbol{u}_t=\tanh(\boldsymbol{W}\cdot\boldsymbol{x}_t) \tag{5}$$

接着，将得到的变换特征 $\boldsymbol{u}_t$ 与另一个权重矩阵 $\boldsymbol{U}$ 进行矩阵乘法来计算注意力得分 $\boldsymbol{e}_t$：

$$\boldsymbol{e}_t=\boldsymbol{u}_t^{\mathrm{T}}\boldsymbol{U} \tag{6}$$

然后，使用 Softmax 函数对所有时间步的注意力得分进行归一化，以获得注意力权重矩阵 $\boldsymbol{a}_t$：

$$\boldsymbol{a}_t=\frac{\exp(\boldsymbol{e}_t)}{\sum_{j=1}^{T}\exp(\boldsymbol{e}_j)} \tag{7}$$

最后，加权输出 $\boldsymbol{o}_t$ 是输入 $\boldsymbol{x}_t$ 与其对应的注意力权重 $\boldsymbol{a}_t$ 的元素乘积：

$$\boldsymbol{o}_t=\boldsymbol{x}_t\cdot\boldsymbol{a}_t \tag{8}$$

4)将结合自注意力机制的 BiGRU 模块的输出结果传递至分类器，得到所有正常网络流量以及各类型网络攻击的概率。

4 网络安全态势量化评估方法

传统态势评估方法主要考虑节点面临的网络攻击威胁，而忽略了节点自身的重要性和节点中心度。因此，本文提出一种态势量化评估方法，通过威胁检测模块提供的检测结果，结合各类型网络攻击的严重度、影响度、节点重要程度和节点中心度，计算节点网络安全态势值，并划分态势等级。

4.1 网络攻击严重度计算

1)将数据集划分为训练集、测试集、验证集。

2)将训练集输入训练好的 SSAE+CNN+BiGRU+AM 网络进行攻击类型判定。统计各类攻击预测出现的次数 $P=\{P_1,P_2,\cdots,P_n\}$ 与各类攻击实际出现的次数 $p=\{p_1,p_2,\cdots,p_n\}$。

3)为了减少模型预测攻击出现次数与实际攻击出现次数的差异,根据式(9)计算概率误差矩阵:

$$\boldsymbol{e}_{\mathrm{err},i}=\frac{p_i}{P_i} \tag{9}$$

4)输入验证集,测试模型实际性能,根据式(10)修正模型预测攻击出现次数:

$$P_{\mathrm{Mod}}=[N_1,N_2,\cdots,N_n]\circ[\boldsymbol{e}_{\mathrm{err},1},\boldsymbol{e}_{\mathrm{err},2},\cdots,\boldsymbol{e}_{\mathrm{err},n}] \tag{10}$$

式中:N_i 表示模型检测到的第 i 种攻击类型出现的次数。

5)在现有研究中,层次分析法(AHP)[20]过于依赖专家经验,且获取难度大,因此本文参考权重系数理论[21],将攻击划分为 c 个等级,根据式(11)计算得到每种攻击的威胁因子,其中,d_i 表示第 i 种攻击类型的攻击等级,T_i 表示第 i 种攻击的威胁因子。如表 1 所示,攻击威胁因子越大,表示攻击越严重,对网络的危害越大。

$$T_i=\begin{cases}\dfrac{1}{2}+\dfrac{\sqrt{-2\ln\dfrac{2d_i}{n}}}{6}, & 1\leqslant d_i<\dfrac{n}{2}\\[2ex]\dfrac{1}{2}, & d_i=\dfrac{n}{2}\\[2ex]\dfrac{1}{2}-\dfrac{\sqrt{-2\ln\dfrac{2d_i}{n}}}{6}, & \dfrac{n}{2}<d_i\leqslant n\end{cases} \tag{11}$$

表 1 攻击严重等级和攻击威胁因子

Table 1 Attack severity levels and attack threat factors

攻击类型	攻击严重等级	攻击威胁因子
DoS	1	0.696
U2R	2	0.500
R2L	2	0.500
Probe	3	0.304

6)攻击严重程度量化,根据式(12)计算得到每种攻击的严重程度 $R_{\mathrm{AL},i}$。

$$R_{\mathrm{AL},i}=P_{\mathrm{Mod},i}\cdot 10^{T_i} \tag{12}$$

4.2 网络攻击影响度计算

根据通用漏洞评分系统(CVSS)[22]对保密性(C)、完整性(I)、可用性(A)3 个属性的影响程度进行评定,并根据各种攻击对 C、I、A 的影响程度进行等级划分。将 C、I、A 按照无影响、低影响、高影响划分为 3 个等级,分别按照 0、0.22、0.56 取值。根据式(13)计算每种攻击的影响度 $R_{\mathrm{AI},i}$:

$$R_{\mathrm{AI},i}=\mathrm{lb}\left(\frac{W_1\cdot 2^{C_i}+W_2\cdot 2^{I_i}+W_3\cdot 2^{A_i}}{3}\right) \tag{13}$$

式中:W_1、W_2、W_3 为 C、I、A 的 Softmax 归一化权重;C_i、I_i、A_i 为第 i 类攻击的 C、I、A 取值。

4.3 网络安全态势值计算

传统态势评估方法主要考虑节点面临的网络攻击威胁,而忽略了节点自身的重要性和节点中心度。为了有效评估特定时间段内的网络安全态势,本文提出的态势量化评估方法通过威胁检测模块提供的检测结果,结合各类型网络攻击的严重度、影响度、节点重要程度和节点中心度对网络进行细粒度评估。该方法的核心在于利用每种网络攻击类型的危害值和该时间段内每条流量对应各种攻击类型的概率,结合节点重要性和节点中心度,综合反映网络安全的潜在威胁程度。

本文提出的网络安全态势评估方法基于式(14)进行计算:

$$S_j=\sum_{i=1}^{n}(R_{\mathrm{AL},i}\cdot R_{\mathrm{AI},i})\cdot C_j\cdot I_j \tag{14}$$

式中:S_j 表示第 j 个节点的网络安全态势值;n 表示检测到网络攻击数;C_j 表示第 j 个节点的中心度,根据式(15)计算紧密中心度,紧密度中心度用于评价一个节点到其他所有节点的紧密程度,节点的紧密中心度越高,节点在网络中越重要,攻击者攻陷该节点造成的危害也越大;I_j 表示第 i 个节点的重要度,代表节点上运行的服务的重要程度。

$$C_v=\frac{N-1}{\sum_{i\neq v}d_{vi}} \tag{15}$$

如表 2 所示,不同的节点具有不同的重要程度,例如,数据库服务节点承载着用户信息、交易记录等敏感数据,一旦受到攻击或者失效,可能会导致数据泄露、丢失或者服务不可用等严重后果。这类节点被攻击的风险以及攻破造成的影响较普通节点来说更大。

表 2 节点类型重要程度

Table 2 Importance of node types

节点类型	节点重要程度
普通主机	2
交换机	5
路由器	5
服务器	6
数据库服务器	10

计算机工程 www.ecicon.com

本文方法旨在通过综合考虑网络流量被识别为各种攻击类型的概率与相应的危害值，结合节点重要性和节点中心度，细粒度地评估特定时间段内的网络安全态势。本文方法将一段时间内所有检测到的网络流量中的潜在攻击经修正后，结合每种攻击类型的严重度、影响度、节点重要度和中心度，能够全面考虑各种潜在威胁对网络安全的总体影响，增强模型对低概率高危害事件的敏感性，使安全态势的评估更为准确。

在获取一段时间内的网络安全态势值后，参考《国家突发公共事件总体应急预案》[23]，将态势划分为 5 个等级及区间，如表 3 所示。

表 3 网络安全态势评估等级及区间

Table 3 Network security situation assessment levels and ranges

态势等级	态势值区间
安全	[0, 0.25]
低风险	(0.25, 0.50]
中风险	(0.50, 0.75]
高风险	(0.75, 1.00]

5 实验与结果分析

5.1 实验环境与数据集

实验基于 Ubuntu 系统进行，利用 PyTorch 框架搭建模型，使用 RTX 3090 显卡进行模型训练。采用 NSL-KDD 数据集[24] 和 CIC-IDS2017 数据集[25]作为实验数据集。

NSL-KDD 数据集改进了 KDD99 数据集中存在的问题[26]，在测试集中加入了一些在训练集中未出现的攻击类型，更加贴合实际的网络环境。该数据集已经提前划分好了训练集与测试集，被广泛应用于评估网络威胁检测模型的性能，各类流量数据信息如表 4 所示。

表 4 NSL-KDD 数据集信息

Table 4 NSL-KDD dataset information 单位：条

数据集	训练集数据量	测试集数据量
Normal	67 343	9 710
DoS	45 927	7 456
Probe	11 656	2 421
R2L	995	2 754
U2R	52	202
总计	125 973	22 543

CIC-IDS2017 数据集基于真实网络环境中捕获的流量数据，具有较高的真实性和可信度，包含多种类型的网络流量，覆盖了各种常见的网络攻击和恶意活动。该数据集提供了丰富的标记信息，包括攻击类型、攻击流量的特征等有助于评估和改进态势评估系统的性能。本文按照 CIC-IDS2017 数据集中的攻击类型将攻击归类为 7 类，按照 8∶2的比例划分了训练集和测试集，删除了 10 个没有方差的特征和 17 个高度相关的特征，因为它们给模型带来了冗余信息，各类流量数据信息如表 5 所示。

表 5 CIC-IDS2017 数据集信息

Table 5 CIC-IDS2017 dataset information 单位：条

攻击类型	标签	数据量
Benign	Benign	1 807 787
DoS/DDoS	Heartbleed, DDoS, DoS Hulk, DoS GoldenEye, DoS Slowloris, DoS Slowhttptest	320 269
PortScan	PortScan	57 305
Brute Force	FTP-Patator, SSH-Patator	8 551
Web Attack	Web Attack Brute Force, Web Attack XSS, Web Attack SQL Injection	2 118
Botnet	Bot	1 943
Infiltration	Infiltration	34

5.2 数据预处理

针对 NSL-KDD 和 CIC-IDS2017 数据集进行预处理时，需要将部分文本特征（如协议类型和服务类型）通过编码转换为数值。针对数据集中个别流量数据较少导致的类别不平衡问题，首先采用 ADASYN 重采样方法进行重采样，然后通过独热编码将分类特征转换为数值特征，最后将特征数据进行归一化处理，归一化计算公式如下：

$$X_{\text{norm}}=\frac{X-X_{\min}}{X_{\max}-X_{\min}} \tag{16}$$

式中：X 是原始数据值；$X_{\min}$ 是数据集中的最小值；$X_{\max}$ 是数据集中的最大值；X_{norm} 是归一化后的值。

5.3 性能评估指标

为了详细评估模型的性能，本文采用了精确率（P）、召回率（R）以及 F1 值（F_1）3 个评估指标。

精确率是评估模型在预测正类时的准确性，表示被正确标记为正类的样本数占模型标记为正类的总样本数的比例，计算公式如下：

$$P=\frac{N_{\text{TP}}}{N_{\text{TP}}+N_{\text{FP}}}\times 100\% \tag{17}$$

式中：N_{TP} 是真正例数量，即模型正确预测为正类的数量；N_{FP} 是假正例数量，即错误标记为正类的数量。

召回率衡量了模型捕捉正类样本的能力，指的是被正确标记为正类的样本数占所有实际正类样本

计算机工程 www.ecice06.com

的比例，计算公式如下：

$$R=\frac{N_{TP}}{N_{TP}+N_{FN}}\times 100\% \tag{18}$$

式中：N_{FN} 是假负例数量，即错误标记为负类的正类样本数量。

F1 值是精确率和召回率的调和平均值，用于综合反映模型性能，适用于不平衡类别数据的场景，计算公式如下：

$$F_1=2\times\frac{P\times R}{P+R}\times 100\% \tag{19}$$

5.4 威胁检测实验结果

将本文提出的融合模型与卷积神经网络结合长短期记忆网络(CNN+LSTM)的模型[11]、自编码深度神经网络(AEDNN)模型[10]、并行特征提取结合改进 BiGRU 的模型 PEEN+ABiGRU[12]、BiGRU、CNN+BiGRU+AM 模型进行对比，以证明融合模型的有效性。所有这些模型均采用了统一的网络参数配置，在 NSL-KDD 数据集的训练集和测试集上进行了五分类训练和评估，在 CIC-IDS2017 数据集的训练集和测试集上进行了七分类训练和评估。6 种模型的精确率、召回率和 F1 值结果对比如表 6、表 7 所示。

表 6 NSL-KDD 上多种威胁检测模型的五分类实验结果

Table 6 Experimental results of five classification of multiple threat detection models on NSL-KDD

模型	精确率	召回率	F1 值
融合模型	0.850 5	0.801 2	0.815 2
CNN+BiGRU+AM	0.818 5	0.783 7	0.790 1
BiGRU	0.807 4	0.784 6	0.776 0
CNN+LSTM	0.829 7	0.777 3	0.797 9
PEEN+ABiGRU	0.798 6	0.709 6	0.725 4
AEDNN	0.815 0	0.769 7	0.774 5

表 7 CIC-IDS2017 上多种威胁检测模型的七分类实验结果

Table 7 Experimental results of seven classification of multiple threat detection models on CIC-IDS2017

模型	精确率	召回率	F1 值
融合模型	0.997 8	0.997 0	0.997 3
CNN+BiGRU+AM	0.995 3	0.995 8	0.995 4
BiGRU	0.994 0	0.991 0	0.992 4
CNN+LSTM	0.996 8	0.992 8	0.994 6
PEEN+ABiGRU	0.983 4	0.961 7	0.971 5
AEDNN	0.986 0	0.980 5	0.982 9

根据表 6、表 7 的数据可以观察到，融合模型在 NSL-KDD 数据集上的精确率较 BiGRU、CNN+BiGRU+AM 提升了 4.31 和 3.20 个百分点，较 PEEN+ABiGRU、AEDNN、CNN+LSTM 也提升了 5.19、3.55 和 2.08 个百分点。在 CIC-IDS2017 数据集上，Benign 类别数量占比过大导致各模型实验结果差距较小，但融合模型精确率较 PEEN+ABiGRU、AEDNN 有所提升。此外，在 NSL-KDD 和 CIC-IDS2017 数据集上，融合模型在精确率、召回率和 F1 值 3 个评估指标上的表现均优于其他模型。

5.5 态势评估实验结果

为验证本文提出的态势评估量化方法的可行性与有效性，使用 VMware 软件模拟了如图 4 所示的网络实验环境。该网络主要包含了 4 台普通电脑、2 台交换机、1 台路由器、1 台服务器、1 台数据库服务器。

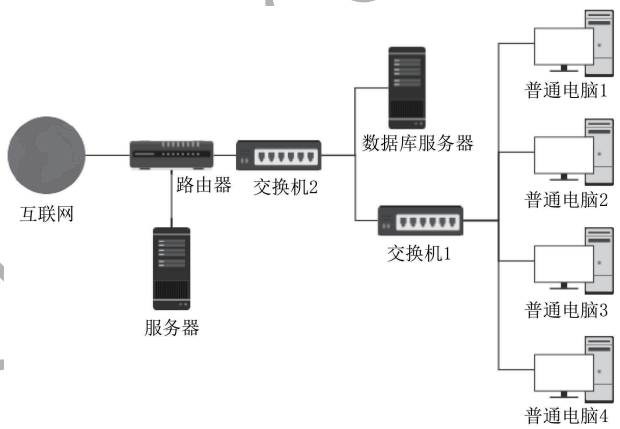


图 4 网络实验环境

Fig.4 Network experimental environment

为模拟真实网络环境，将 NSL-KDD 测试集按照 5∶5的比例重新划分为测试集和验证集。首先，将验证集输入威胁检测模型中预测攻击类型，统计攻击数量，计算得到概率误差矩阵。然后，将测试集按照网络节点数划分，模拟真实环境中的节点流量数据，将流量数据输入训练好的威胁检测模型中，统计模型预测攻击数量，并用概率误差矩阵修正模型预测结果，得到修正结果 P_{Mod}。最后，结合攻击严重度、攻击影响度、节点中心度、节点重要度计算得到节点态势值，态势评估结果如图 5 所示。

从图 5 可以看出，6 种模型的态势评估结果与真实结果大致相同，但本文所提融合模型的态势评估结果与真实值最为接近。此外，对于交换机 1，真实态势值为中风险，AEDNN、PEEN+ABiGRU 和 CNN+LSTM 评估结果为低风险，BiGRU、CNN+BiGRU+AM 和融合模型评估结果为中风险。这表明本文所提融合模型的态势评估结果更贴合实际的网络态势情况。

计算机工程 www.ecj.com

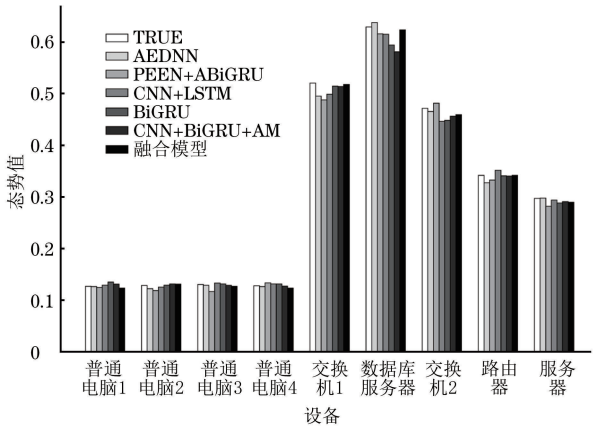


图 5 网络安全态势感知实验结果

Fig.5 Network security situation awareness experimental results

6 结束语

随着网络技术的发展,网络安全形势愈发严峻。传统的网络安全态势感知方法在数据特征提取和长时间序列数据处理上表现不佳。本文提出一个用于网络安全态势感知的融合 SSAE、CNN、BiGRU 和 AM 的混合深度学习模型。通过 SSAE 提取关键数据特征,利用注意力机制增强 BiGRU 的关键信息处理能力,有效识别网络流量数据的网络攻击类型,获得网络安全态势。实验结果表明,该模型在各项评估指标上优于传统模型,能够通过流量信息准确识别网络攻击类型,有效感知网络态势。后续研究一方面将整合多种数据源,如网络资产和用户行为,提升态势感知模型的全面性;另一方面将优化深度学习模型的结构和算法,提高模型的泛化能力,增强其对新型未知网络攻击的识别能力。

参考文献

[1] 周泽岩. 论网络空间安全与治理[J]. 信息网络安全, 2020(S1):153-155.
ZHOU Z Y. On Cyberspace security and governance [J]. Information Network Security, 2020(S1): 153-155. (in Chinese)

[2] 孙理理. 网络安全治理对策研究[J]. 信息网络安全, 2023, 23(6): 104-110.
SUN C C. Research on countermeasures for network security governance[J]. Netinfo Security, 2023, 23(6): 104-110. (in Chinese)

[3] CHEN C, YE L, YU X Z, et al. A survey of network security situational awareness technology [M]. Berlin, Germany: Springer International Publishing, 2019.

[4] 胡庆爽,李成海,路艳丽,等. 基于分级优化置信规则库的网络安全态势预测方法[J]. 计算机工程,2020,46(12):127-133.
HU Q S, LI C H, LU Y L, et al. Network security situation prediction method based on hierarchical optimization of belief rule base[J]. Computer Engineering, 2020, 46(12): 127-133. (in Chinese)

[5] 李欣,段詠程. 基于改进隐马尔可夫模型的网络安全态势评估方法[J]. 计算机科学,2020,47(7):287-291.
LI X, DUAN Y C. Network security situation assessment method based on improved hidden Markov model [J]. Computer Science, 2020, 47(7): 287-291. (in Chinese)

[6] 丁华东,许华虎,段然,等. 基于贝叶斯方法的网络安全态势感知模型[J]. 计算机工程, 2020, 46(6): 130-135.
DING H D, XU H H, DUAN R, et al. Network security situation awareness model based on Bayesian method [J]. Computer Engineering, 2020, 46(6): 130-135. (in Chinese)

[7] ZHANG H B, YIN Y, ZHAO D M, et al. Network security situational awareness model based on threat intelligence[C]// Proceedings of International Conference on Mobile Multimedia Communications. Berlin, Germany: Springer, 2021: 526-536.

[8] 王金恒,单志龙,谭汉松,等. 基于遗传优化 PNN 神经网络的网络安全态势评估[J]. 计算机科学, 2021, 48(6): 338-342.
WANG J H, SHAN Z L, TAN H S, et al. Network security situation assessment based on genetic optimized PNN neural network[J]. Computer Science, 2021, 48(6): 338-342. (in Chinese)

[9] CHEN Z H. Research on Internet security situation awareness prediction technology based on improved RBF neural network algorithm[J]. Journal of Computational and Cognitive Engineering, 2022, 1(3): 103-108.

[10] 杨宏宇,曾仁韵. 一种深度学习的网络安全态势评估方法[J]. 西安电子科技大学学报,2021,48(1):183-190.
YANG H Y, ZENG R Y. A deep learning-based network security situation assessment method[J]. Journal of Xidian University, 2021, 48(1): 183-190. (in Chinese)

[11] 陈解元. 基于 LSTM 的卷积神经网络异常流量检测方法[J]. 信息技术与网络安全, 2021, 40(7): 42-46.
CHEN J Y. Network intrusion detection based on convolutional neural networks with LSTM[J]. Information Technology and Network Security, 2021, 40(7): 42-46. (in Chinese)

[12] 杨宏宇,张梓锌,张良. 基于并行特征提取和改进 BiGRU 的网络安全态势评估[J]. 清华大学学报(自然科学版), 2022, 62(5): 842-848.
YANG H Y, ZHANG Z X, ZHANG L. Network security situation assessments with parallel feature extraction and an improved BiGRU [J]. Journal of Tsinghua University (Science and Technology), 2022, 62(5): 842-848. (in Chinese)

[13] WANG J, JIANG X L, ZHU J. Security situation awareness algorithm of network information transmission based on intelligent computing [C] // Proceedings of the 2nd International Conference on Electronic Information Engineering and Computer Communication (EIECC 2022). Xi'an, China: SPIE, 2023: 564-570.

[14] SOKOL P, STAŇA R, GAJDOŠ A, et al. Network security situation awareness forecasting based on statistical approach and neural networks[J]. Logic Journal of the IGPL, 2023, 31(2): 352-374.

[15] LU Y, KUANG Y X, YANG Q F. Intelligent prediction of network security situations based on deep reinforcement learning algorithm [J]. Scalable Computing: Practice and Experience, 2024, 25(1): 147-155.

[16] 杨志鹏,刘代东,袁军翼,等. 基于自注意力机制的网络局域安全态势融合方法研究[J]. 信息网络安全, 2024, 24(3): 398-410.
YANG Z P, LIU D D, YUAN J Y, et al. Research on network local security situation fusion method based on self-attention mechanism[J]. Netinfo Security, 2024, 24(3): 398-410. (in Chinese)

[17] SHIN H C, ORTON M R, COLLINS D J, et al. Stacked autoencoders for unsupervised feature learning and multiple

计算机工程 www.ecice06.com

organ detection in a pilot study using 4D patient data[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(8): 1930-1943.
[18] CHO K, VAN MERRIËNBOER B, GULCEHRE C, et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation[EB/OL]. [2024-03-17]. https://arxiv.org/abs/1406.1078.
[19] LUONG M T, PHAM H, MANNING C D. Effective approaches to attention-based neural machine translation[EB/OL]. [2024-03-17]. http://arxiv.org/abs/1508.04025.
[20] HU W, LI J, JIANG X. A hierarchical algorithm for cyberspace situational awareness based on analytic hierarchy process[J]. High Technology Letters, 2007, 13(3): 291-296.
[21] 刘效武,王慧强, 吕宏武, 等. 网络安全态势认知融合感控模型[J]. 软件学报, 2016, 27(8): 2099-2114.
LIU X W, WANG H Q, LÜ H W, et al. Fusion-based cognitive awareness-control model for network security situation[J]. Journal of Software, 2016, 27(8): 2099-2114. (in Chinese)
[22] Forum of Incident Response and Security Teams (FIRST). Common Vulnerability Scoring System (CVSS) specification document[EB/OL]. [2024-03-17]. https://www.first.org/cvss/v3.1/specification-document.
[23] 国务院. 国家突发公共事件总体应急预案[M]. 北京: 中国法制出版社, 2006.
The State Council. National emergency response plan for public emergencies [M]. Beijing: China Legal Publishing House, 2006. (in Chinese)
[24] FERRAG M A, MAGLARAS L, MOSCHOYIANNIS S, et al. Deep learning for cyber security intrusion detection: approaches, datasets, and comparative study[J]. Journal of Information Security and Applications, 2020, 50: 102419.
[25] ROSAY A, CHEVAL E, CARLIER F, et al. Network intrusion detection: a comprehensive analysis of CIC-IDS2017 [C]// Proceedings of the 8th International Conference on Information Systems Security and Privacy. [S. l.]: Science and Technology Publications, 2022: 25-36.
[26] JAVAID A, NIYAZ Q, SUN W Q, et al. A deep learning approach for network intrusion detection system [C]// Proceedings of the 9th EAI International Conference on Bio-inspired Information and Communications Technologies (formerly BIONETICS). New York, USA: ACM Press, 2016: 21-26.

编辑 陆燕菲

计算机工程 www.ecice06.com