基于元学习的多头注意力时序卷积的入侵检测

王　明
（河北科技师范学院　网络技术中心，河北　秦皇岛　066001）

摘　要：为解决现有入侵检测方法在高阶依赖关系挖掘，处理时序特征和应对新型攻击手段检测等方面性能不足的问题，提出了一种基于元学习的多头注意力时序卷积的入侵检测方法。该方法引入了多头注意力机制，使模型能在不同尺度上捕捉网络数据的时序特征和高阶依赖关系。其次，结合多任务学习改进元学习算法对网络未知攻击进行识别，提升网络未知攻击的检测性能，此外，设计了一种自适应特征提取策略，动态调整特征提取粒度，以适应不同类型的网络攻击。在公开数据集实验对比表明，本文算法与主流算法相比，具有更高的准确率和 F 值。

关键词：入侵检测；元学习；多头注意力机制；时序卷积网络
中图分类号：TP393.08；TP18　　**文献标识码**：A　　**DOI**：10.19358/j.issn.2097-1788.2023.07.008
引用格式：王明. 基于元学习的多头注意力时序卷积的入侵检测［J］. 网络安全与数据治理，2023，42(7)：49-54.

Intrusion detection based on meta-learning with multi-head attention temporal convolution

Wang Ming
(Network Technology Center, Hebei Normal University Of Science & Technology, Qinhuangdao 066001, China)

Abstract: To address the performance limitations of existing intrusion detection methods in mining high-order dependency relationships, processing temporal features, and detecting new types of attack methods, this paper proposes an intrusion detection method based on meta-learning and multi-head attention temporal convolution. This method introduces a multi-head attention mechanism, allowing the model to capture the temporal features and high-order dependency relationships of network data at different scales. Secondly, this paper combines multi-task learning to improve the meta-learning algorithm for identifying unknown network attacks, thus enhancing the detection performance of unknown network attacks. In addition, this paper designs an adaptive feature extraction strategy that dynamically adjusts the feature extraction granularity to adapt to different types of network attacks. Experimental comparisons on public datasets show that the proposed algorithm has higher accuracy and F-score compared to mainstream algorithms.

Key words: intrusion detection; meta-Learning ; multi-head attention mechanism; temporal convolutional neural network

0　引言

网络入侵检测（Intrusion Detection System，IDS）[1]是一种用于监测网络活动的技术，旨在及时发现潜在的恶意行为、攻击以及系统安全策略的违规行为。根据检测方法的不同，网络入侵检测技术主要分为两类：基于签名的检测技术[2]（Signature-based Detection）和基于异常的检测技术[3]（Anomaly-based Detection）。随着深度学习为代表的人工智能技术的发展，近年来研究者们已经尝试利用深度学习技术[4]解决网络入侵检测中的一些挑战性问题，Gao[5]等人提出了基于深度信念网络（DBN）的入侵检测方法，实现了较高的检测准确率，但在处理大规模数据时计算复杂度较高。Kim[6]等人提出了基于长短时记忆网络（LSTM）的入侵检测方法，利用 LSTM 捕捉时序特征以提高检测性能，但对于未知攻击的检测能力有限。Tang[7]等人采用自编码器（AE）对网络流量进行特征提取，提高了异常检测的性能，但在处理时序相关性方面存在不足。Niyaz[8]等人利用卷积神经网络（CNN）进行入侵检测，实现了较高的检测准确率，但对高阶依赖关系的挖掘仍有改进空间。Vinayakumar[9]等人采用卷积神经网络（CNN）对恶意 URL 进行检测，实现了较高的检测准确率，但在实时性方面存在局限。Lo[10]等人提

出了一种基于图卷积神经网络的入侵检测方法，该方法利用图结构挖掘网络数据中的关联信息，但在处理大规模网络数据时效率有待提高。Cao[11]等人利用卷积神经网络和门控循环单元（GNU）进行网络入侵检测，有效地解决分类准确率低和类别不平衡的问题，但在处理未知攻击和动态环境下的泛化能力方面存在挑战。这些研究在网络入侵检测方面取得了一定的成果，但仍然存在一些问题。首先，许多现有方法对网络数据的时序特征和高阶依赖关系挖掘不足，导致检测性能有限。其次，处理未知攻击时，泛化性能有待提高。

为了解决上述问题，本文提出了一种基于元学习的多头注意力时序卷积的入侵检测方法。该方法充分挖掘网络数据的时序特征和高阶依赖关系，提高网络入侵检测的性能。本文的主要贡献如下：

（1）提出了一种基于多头注意力机制[12]的时序卷积神经网络[13]，能在不同尺度上捕捉网络数据的时序特征和高阶依赖关系，从而提高检测性能。

（2）结合多任务学习[14]和自适应元学习[15]，提出了一种基于多任务自适应的元学习算法，利用元学习[16]策略提高模型对未知攻击的适应能力，从而使网络入侵检测系统具有更强的实用性和鲁棒性。

（3）设计了一种自适应的特征提取策略，使得模型能够根据输入数据动态调整特征提取粒度，以更好地适应不同类型的网络攻击。

（4）引入类别敏感因子，改进了交叉熵损失函数，以解决网络入侵检测中的数据类别不平衡和误报率问题。

1　本文算法概述

1.1　本文算法的整体架构

本文算法模型主要包括输入层、多头注意力时序卷积层、自适应卷积层[17]、自适应池化层[18]、全连接层、基于多任务自适应元学习的未知攻击识别模块和输出层。本文算法整体框图如图 1 所示。

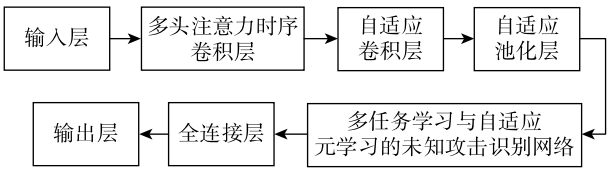

图 1　本文算法的整体框图

1.2　多头注意力时序卷积网络

多头注意力时序卷积网络（Multi-Head Attention Temporal Convolutional Network，MA-TCNN）主要由时序卷积网络和多头注意力机制两部分构成。时序卷积网络可以捕捉网络行为的时间依赖性，这有助于在网络流量中识别出潜在的异常行为和攻击模式。多头注意力机制可以捕捉网络数据的高阶依赖关系，从而提高入侵检测的性能。MA-TCNN 网络结构如图 2 所示。

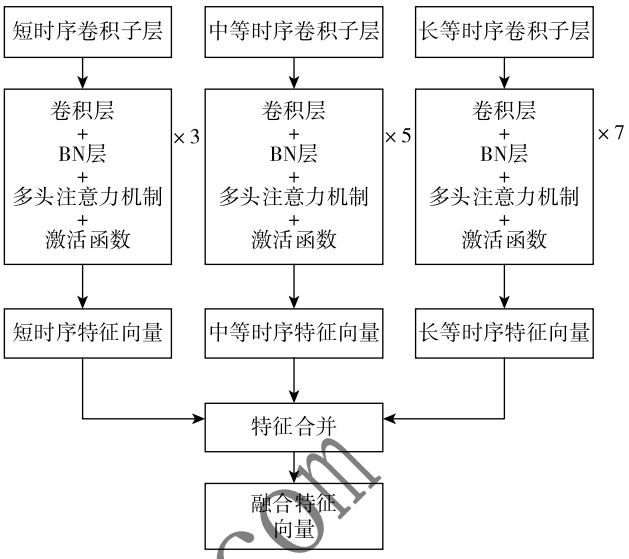

图 2　MA-TCNN 网络结构框图

1.2.1　时序卷积网络

时序卷积网络由短时序卷积子层、中等时序卷积子层和长时序卷积子层构成，其主要功能是从不同的局部感受野中提取时序特征。具体的实现方法是，在三个具有不同大小卷积核的卷积时序子网络并行处理输入数据，捕捉不同尺度的时序特征，分别生成了不同时序的特征向量。然后，将这三个特征向量融合在一起，形成一个融合特征向量。

其中，短时序卷积子网络包含 3 个卷积层，中等时序卷积子网络包含 5 个卷积层，长时序卷积子网络包含 7 个卷积层。每个卷积层后都跟有批量归一化层、多头注意力机制和激活函数。具体来说，对于第 n 个时序卷积子网络，第 m 个卷积层的卷积操作可以表示为：

$$C_{nm}=f\left(X*W_n+b_n\right) \tag{1}$$

其中，X 表示输入数据，n 表示时序卷积子网络的数量，m 表示每个时序卷积子网络中卷积层的数量，W_n 和 b_n 分别表示第 n 个卷积分支子网络的卷积核权重和偏置项，f 是激活函数，$*$ 表示卷积操作，C_{nm} 表示第 n 个卷积分支子网络第 m 个卷积层的输出。

1.2.2　多头注意力机制

多头注意力机制源自于自然语言处理领域的 Transformer[19]模型，已被证明对于学习输入特征之间的高阶依赖关系具有很好的效果。

多头注意力机制的计算公式：

$$\mathrm{MulHead}\left(\boldsymbol{Q},\ \boldsymbol{KV}\right)=\mathrm{Concat}\left(\mathrm{head}_1,\ \mathrm{head}_2,\ \cdots,\ \mathrm{head}_h\right)W^{\circ} \tag{2}$$

每个 Head 的计算方式如下：

$$\text{Head}_i = \text{Attention}\left(\boldsymbol{QW}_i^Q,\ \boldsymbol{KW}_i^K,\ \boldsymbol{VW}_i^V\right) \tag{3}$$

其中，$\boldsymbol{W}_i^Q$、$\boldsymbol{W}_i^K$、$\boldsymbol{VW}_i^V$ 是对应第 i 个 Head 的权重矩阵。注意力机制 Attention 的计算方式如下：

$$\text{Attention}\left(\boldsymbol{Q},\ \boldsymbol{K},\ \boldsymbol{V}\right) = \text{softmax}\left(\frac{\boldsymbol{QK}^{\mathrm{T}}}{\sqrt{d_k}}\right)\boldsymbol{V} \tag{4}$$

其中，$\boldsymbol{Q}$ 为查询（query）矩阵，$\boldsymbol{K}$ 为键（Key）矩阵，$\boldsymbol{V}$ 为值（value）矩阵，$\boldsymbol{K}^{\mathrm{T}}$ 是键矩阵的转置；d_k 表示键和查询的维度，在多头注意力中起到了缩放因子的作用，可以提高模型的稳定性。

1.2.3 时序卷积与多头注意力机制的结合

为了将时序卷积和多头注意力机制有效地整合，首先将输入特征经过卷积层和批量归一化层的处理，得到一系列特征向量。接着，将特征向量送入多头注意力机制，学习输入特征之间的相互关系，生成更具表达能力的特征表示。

具体来说，首先，对于每个时序卷积子网络中每个卷积输出 C_{nm}，将其作为查询矩阵 $\boldsymbol{Q}$、键矩阵 $\boldsymbol{K}$、值矩阵 $\boldsymbol{V}$ 的输入：

$$\boldsymbol{Q}_{nm} = C_{nm}\boldsymbol{W}_n^Q \tag{5}$$

$$\boldsymbol{K}_{nm} = C_{nm}\boldsymbol{W}_n^K \tag{6}$$

$$\boldsymbol{V}_{nm} = C_{nm}\boldsymbol{W}_n^V \tag{7}$$

其次，对于第 i 个 head，设权重矩阵 $\boldsymbol{W}_{i,n}^Q$、$\boldsymbol{W}_{i,n}^K$ 和 $\boldsymbol{W}_{i,n}^V$ 分别对应于第 i 个 head 和第 n 个时序卷积子网络的权重矩阵。首先计算 $\boldsymbol{Q}_{nm}$、$\boldsymbol{K}_{nm}$ 和 $\boldsymbol{V}_{nm}$ 的变换：

$$\boldsymbol{Q}_{i,nm} = \boldsymbol{Q}_{nm}\boldsymbol{W}_{i,n}^Q \tag{8}$$

$$\boldsymbol{K}_{i,nm} = \boldsymbol{K}_{nm}\boldsymbol{W}_{i,n}^K \tag{9}$$

$$\boldsymbol{V}_{i,nm} = \boldsymbol{V}_{nm}\boldsymbol{W}_{i,n}^V \tag{10}$$

第 n 个时序卷积子网络，第 m 个卷积的第 i 个 head 的注意力输出为 $\text{Attention}_{i,nm}\left(\boldsymbol{Q}_{i,nm},\ \boldsymbol{K}_{i,nm},\ \boldsymbol{V}_{i,nm}\right)$，如式（11）所示：

$$\text{Attention}_{i,nm}\left(\boldsymbol{Q}_{i,nm},\ \boldsymbol{K}_{i,nm},\ \boldsymbol{V}_{i,nm}\right) = \text{softmax}\left(\frac{\boldsymbol{Q}_{i,nm}\boldsymbol{K}_{i,nm}^T}{\sqrt{d_k}}\right)\boldsymbol{V}_{i,nm} \tag{11}$$

最后将 h 个 head 的输出连接起来，并通过输出权重矩阵 $\boldsymbol{W}_n^O$ 进行线性变换，以获得最终第 n 个时序卷积子网络第 m 个卷积层的多头注意力输出 MH_{nm}，如公式（12）所示：

$$\text{MH}_{nm} = \text{Concat}\left(\text{Attention}_{1,\text{nm}},\ \cdots,\ \text{Attention}_{h,nm}\right)\boldsymbol{W}_{nm}^O \tag{12}$$

式（12）可以简化为：

$$\text{MH}_m = \text{MultiHead}\left(C_{nm},\ C_{nm},\ C_{nm}\right) \tag{13}$$

令 MH_n 为第 n 个时序卷积子网络的最终输出，C_n 为第 n 个时序卷积子网络的最终输出，结合式（1）有：

$$C_n = f\left(\text{MH}_n * W_n + b_n\right) \tag{14}$$

然后，本文将三个时序卷积子网络输出结果进行融合，得到一个融合特征向量：

$$\boldsymbol{MHC} = \text{Concat}\left(C_1,\ C_2,\ C_3\right) \tag{15}$$

其中，$\boldsymbol{MHC}$ 表示融合后的多头注意力时序卷积网络输出特征向量。

1.3 自适应卷积层

自适应卷积层的主要作用是根据输入特征向量的尺寸自动调整卷积核大小和步长，从而实现对动态网络流量的有效建模。

设输入特征向量表示为 $\boldsymbol{A}$，长度为 P。输出特征向量为 $\boldsymbol{B}$，长度为 p。为了实现自适应卷积，需根据输入输出特征向量的尺寸，确定合适的卷积核大小 k_a 和步长 s_a，它们之间的关系可以采用以下公式表示：

$$k_a = \left\lfloor \frac{P - p + 1}{p} \right\rfloor \tag{16}$$

$$s_a = \left\lfloor \frac{P - k_a}{p - 1} \right\rfloor \tag{17}$$

为了更清晰地表示自适应卷积操作，可以将其表示为：

$$\boldsymbol{B} = F\left(\boldsymbol{A} \otimes_{k_a,s_a} W_a + b_a\right) \tag{18}$$

其中，$\boldsymbol{A}$ 表示输入的特征向量，W_a 和 b_a 分别表示卷积核权重和偏置项，F 是激活函数，$\otimes_{k_a,s_a}$ 表示为卷积核大小为 k_a、步长为 s_a 的卷积操作。

1.4 自适应池化层

自适应池化层的主要功能是降低卷积层输出的尺寸，用于降低特征维度，减小模型的参数数量与计算复杂度。

为了实现自适应池化，需要根据输入输出特征向量的尺寸确定池化窗口的大小 w 和步长 z。设输入特征向量为 $\boldsymbol{G}$，长度为 I，输出特征向量为 $\boldsymbol{L}$，长度为 J。给定输入特征向量 $\boldsymbol{G}$ 中的元素 G_j，输出特征向量为 $\boldsymbol{L}$ 中元素 L_j，则池化窗口的大小 w 和步长 z 的关系可以表示为：

$$w = \text{ceil}\left(\frac{I}{J}\right)z = \text{floor}\left(\frac{(I - w)}{(J - 1)}\right) \tag{19}$$

自适应最大池化操作可以表示为：

$$L_j = \max\left(G_{(j*z)},\ G_{(j*z+1)},\ \cdots,\ G_{(j*z+w-1)}\right) \tag{20}$$

$$j = \text{floor}\left(\ (i-1) * (J-1)/(I-1)\right) \tag{21}$$

其中，celi 表示向上取整，floor 表示向下取整。

1.5 多任务自适应元学习的未知攻击识别

为了提升网络入侵检测模型在识别未知攻击方面的性能，本文结合多任务的优点改进了元学习方法，增强

模型的泛化能力和鲁棒性。

1.5.1 多任务学习

多任务学习通过在多个相关任务间共享信息来提高模型的泛化能力。在网络入侵检测领域，不同类型的攻击可以被视为不同的任务。在一个统一的模型中同时学习多个任务可以有效地利用任务间的相互关系，提高模型对未知攻击的识别能力。设有 M 个任务，每个任务的损失函数为 $L_{\mathrm{MTL}i}$，多任务损失函数 L_{MTL} 定义如下：

$$L_{\mathrm{MTL}} = \sum_{i=1}^{M} \alpha_i * L_{\mathrm{MTL}i} \tag{22}$$

其中，α_i 为第 i 个任务的权重，取值范围为［0，1］，满足 $\sum_{i=1}^{M} \alpha_i = 1$。任务权重可根据训练过程中的性能表现确定，以确保模型在所有任务上达到最优性能。

1.5.2 自适应元学习

自适应元学习通过调整元参数来适应不同任务的学习过程。在网络入侵检测任务中，自适应元学习有助于模型在遇到新型攻击时快速适应。具体来说，可以将元学习过程建模为一个条件概率分布 $P(y \mid x, \Phi)$，其中 x 表示输入数据，y 表示对应的标签，Φ 表示元参数。元参数 Φ 的学习可以通过最大化以下对数似然函数实现：

$$\Phi^* = \mathrm{argmax}\_\ \Phi \sum_{i=1}^{M} \log P(y_i \mid x_i, \Phi) \tag{23}$$

1.5.3 多任务自适应元学习的未知攻击识别

本文将多任务学习与自适应元学习相结合，首先，通过多任务学习共享信息，然后利用自适应元学习为每个任务适应不同的元参数。为实现这两种方法的融合，可以将 MTL 损失函数与 AutoMeta 对数似然函数结合，构建联合优化目标：

$$L_{\mathrm{total}} = L_{\mathrm{MTL}} - \zeta * \Phi^* \tag{24}$$

其中，ζ 是一个超参数，用于平衡多任务学习和自适应元学习的影响。通过优化 L_{total}，实现多任务学习与自适应元学习的融合。

1.6 损失函数与优化

本文通过引入了一个类别敏感因子，改进了交叉熵损失函数，称为类别敏感交叉熵损失，以更好训练本文算法模型。

$$L_{\varepsilon} = -\sum_{\chi=1}^{\Omega} \sum_{\mu=1}^{C} \psi_{\chi\mu} * \log(\Theta_{\chi\mu}) * \sigma(c_{\mu}) \tag{25}$$

其中，L_{ε} 表示类别敏感交叉熵损失，$\psi_{\chi\mu}$ 表示第 χ 个样本的第 μ 个类别的实际标签（0 或 1）；$\Theta_{\chi\mu}$ 表示第 χ 个样本的第 μ 个类别的模型预测概率分布，Ω 表示样本数量，C 表示类别数量，c_{μ} 表示第 μ 个类别，$\sigma(c_{\mu})$ 为类别敏感因子。

2 实验结果分析

2.1 实验设置介绍

2.1.1 数据集介绍

（1）NSL－KDD 数据集是 KDD99 数据集的一个改进版本，该数据集由四个子数据集组成：KDDTest＋、KDDTest－21、KDDTrain＋、KDDTrain＋_ 20Percent。与 KDD99 数据集相比，NSL－KDD 数据集在数据分布和类别不平衡方面有所改进。

（2）UNSW－NB15 数据集是一个广泛应用于网络入侵检测研究的数据集。该数据集包含了大量正常流量与恶意流量数据，其中恶意流量涵盖了 9 种攻击类别，总共包含约 250 万条数据记录。

2.1.2 评价指标介绍

本文使用准确率（Accuracy，Ac）、查准率（Precision，Pr）、查全率（Recall，Re）和 F1 值（FScore，F1）进行评价。评价指标定义如下：

（1）准确率：正确分类的样本占总样本的比例。

$$\mathrm{Ac} = \frac{(\mathrm{TP} + \mathrm{TN})}{(\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN})} \tag{26}$$

（2）查准率：实际为正例的样本被正确分类的比例。

$$\mathrm{Pr} = \frac{\mathrm{TP}}{(\mathrm{TP} + \mathrm{FP})} \tag{27}$$

（3）查全率：被正确分类的正例占所有正例的比例。

$$\mathrm{Re} = \frac{\mathrm{TP}}{(\mathrm{TP} + \mathrm{FN})} \tag{28}$$

（4）F1 值：表示查准率和查全率的调和平均值，用于综合评价模型的性能。

$$\mathrm{F1} = \frac{2 \cdot \mathrm{Pr} \cdot \mathrm{Re}}{\mathrm{Pr} + \mathrm{Re}} \tag{29}$$

其中，TP 表示真正例（True Positive），TN 表示真负例（True Negative），FP 表示假正例（False Positive），FN 表示假负例（False Negative）。

2.2 实验结果分析

为了验证本文所提算法的有效性，本文将其与其他几种主流入侵检测方法进行了对比，如表 1、表 2 所示。本节将对实验结果进行详细分析。

表 1 KDDTest＋数据集中 5 种算法多分类对比（%）

方法	准确率	查准率	查全率	F1 值
AlertNet[20]	78.50	81.00	78.50	76.50
CNN[21]	81.75	82.43	82.71	82.57
MDNN[22]	77.55	81.23	77.55	75.43
CNN-BiLSTM[23]	83.58	85.82	84.49	85.14
本文算法	85.23	86.47	85.79	86.13

表 2　UNSW-NB15-test 数据集中 5 种算法多分类对比（%）

方法	准确率	查准率	查全率	F1 值
AlertNet [20]	66.00	62.30	66.00	59.60
CNN[21]	74.61	81.01	75.65	78.24
MDNN[22]	62.87	76.01	63.10	64.15
CNN-BiLSTM[23]	77.16	82.63	79.91	81.25
本文算法	79.23	82.20	79.50	81.55

AlertNet [20] 与 MDNN[22] 方法采用 DNN 作为主体网络结构，处理时序特征能力稍有不足。CNN[21] 方法可以实现较高的检测准确率，但对高阶依赖关系的挖掘仍有改进空间。CNN-BiLSTM[23] 结合了 CNN 和 LSTM 的优势，对高阶依赖关系和时序特征的处理有了较大提高，但对未知攻击的检测能力有限。本文算法在高阶依赖关系挖掘、处理时序特征和对新型攻击手段检测等方面进行了针对性的网络设计，在检测性能上具有较好的表现。

表 1 展示了在 KDDTest + 数据集上 5 种算法的对比结果。从表中可以看出，本文算法在准确率、查准率、查全率和 F1 值方面均优于其他方法，相较于其他算法有显著的性能提升。

表 2 展示了在 UNSW-NB15-test 数据集上 5 种算法的对比结果。本文算法的准确率高于其他算法，查准率和查全率略低于 CNN-BiLSTM 算法，但 F1 值高于其他算法，证明了本文算法的有效性。

2.3　消融实验

为了验证本文所提算法各部分对算法性能的贡献，本文进行了消融实验。消融实验分别去掉了多头注意力机制（MA-TCNN no Attention）和元学习模块（MA-TCNN no Meta），并将其结果与本文算法的结果进行对比。实验结果如表 3、表 4 所示。

表 3　KDDTest + 数据集中消融实验（%）

方法	准确率	查准率	查全率	F1 值
本文算法	85.23	86.47	85.79	86.13
MA-TCNN no Meta	82.65	83.94	83.24	83.59
MA-TCNN no Attention	80.32	81.57	80.85	81.21

表 4　UNSW-NB15-test 数据集中消融实验（%）

方法	准确率	查准率	查全率	F1 值
本文算法	79.23	82.20	79.50	81.55
MA-TCNN no Meta	75.45	78.31	76.22	77.26
MA-TCNN no Attention	72.19	75.01	73.04	74.02

表 3 展示了在 KDDTest + 数据集上的消融实验结果。通过对比，可以看出本文算法在准确率、查准率、查全率和 F1 值方面均优于去掉元学习模块和去掉多头注意力机制的版本。表明多头注意力机制和元学习模块对本文算法性能提升有显著贡献。

表 4 展示了在 UNSW-NB15-test 数据集上的消融实验结果。同样地，本文算法在准确率、查准率、查全率和 F1 值方面均优于去掉元学习模块和去掉多头注意力机制的版本。这进一步证实了多头注意力机制和元学习模块对本文算法性能的重要贡献。

3　结论

本文提出了一种基于元学习的多头注意力时序卷积的入侵检测方法。本文算法引入多头注意力机制以捕捉不同尺度的时序特征，充分挖掘网络数据的高阶依赖关系。同时，提出了一种多任务自适应地元学习未知攻击识别方法，以提升入侵检测对未知攻击的防护能力。此外，本文还设计了一套自适应的特征提取策略，通过自适应的调整卷积操作，对不断变化的数据进行高效的建模。实验结果表明，本文算法在检测性能方面优于其他主流入侵检测算法，证明了本文算法的优越性。

参考文献

[1] LIAO H J, LIN C H, LIN Y C, et al. Intrusion detection system: a comprehensive review [J]. Journal of Network and Computer Applications, 2013, 36 (1): 16 - 24.

[2] KUMAR V, SANGWAN O P. Signature based intrusion detection system using SNORT [J]. International Journal of Computer Applications & Information Technology, 2012, 1 (3): 35 - 41.

[3] JYOTHSNA V, PPASAD R, PRASAD K M. A review of anomaly based intrusion detection systems [J]. International Journal of Computer Applications, 2011, 28 (7): 26 - 35.

[4] GURUNG S, GHOSE M K, SUBEDI A. Deep learning approach on network intrusion detection system using NSL-KDD dataset [J]. International Journal of Computer Network and Information Security, 2019, 11 (3): 8 - 14.

[5] GAO N, GAO L, GAO Q, et al. An intrusion detection model based on deep belief networks [C] // Proceedings of the 2th International Conference on Advanced Cloud and Big Data, Huangshan, China, 2014: 247 - 252.

[6] KIM G, YI H, LEE J, et al. LSTM-based system-call language modeling and robust ensemble method for designing host-based intrusion detection systems [EB/OL]. (2016 - 11 - 06). https://arxiv.org/abs/1611.01726.

[7] TANG T, MHAMDI L, MCLERNON D, et al. Deep learning approach for network intrusion detection in software defined networking [C] // Proceedings of the 2016 International Conference on Wireless Networks and Mobile Communications, Fez, Morocco,

2016：258 – 263.

[8] NIYAZ Q, SUN W, JAVAID A, 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, New York City, NY, United states, 2016：21 – 26.

[9] VINAYAKUMAR R, SOMAN K, POORNACHANDRAN P. Evaluating deep learning approaches to characterize and classify malicious URL's [J]. Journal of Intelligent & Fuzzy Systems, 2018, 34 (3)：1333 – 1343.

[10] LO W W, LAYEGHY S, SARHAN M, et al. E-graphsage：a graph neural network based intrusion detection system for iot [C]// Proceedings of the 2022 – 2022 IEEE/IFIP Network Operations and Management Symposium. Budapest, Hungary, 2022：1 – 9.

[11] CAO B, LI C H, SONG Y F, et al. Network intrusion detection model based on CNN and GRU [J]. Applied Sciences, 2022, 12 (9)：4184.

[12] XI C, LU G, YAN J. Multimodal sentiment analysis based on multi-head attention mechanism [C] //Proceedings of the 4th International Conference on Machine Learning and Soft Computing, Haiphong, Vietnam, 2020：34 – 39.

[13] PANDEY A, WANG D L. TCNN：temporal convolutional neural network for real-time speech enhancement in the time domain [C]//Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, Brighton, United kingdom, 2019：6875 – 6879.

[14] KOLLIAS D. Abaw：learning from synthetic data & multi-task learning challenges [C] // Proceedings of the European Conference on Computer Vision. Tel Aviv, Israel, 2023：157 – 172.

[15] YU R S, GONG Y, HE X, et al. Personalized adaptive meta learning for cold-start user preference prediction [C] //Proceedings of the AAAI Conference on Artificial Intelligence, Virtual, Online, 2021, 35 (12)：10772 – 10780.

[16] ZHANG J, SONG J, GAO L, et al. Progressive meta-learning with curriculum [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2022, 32 (9)：5916 – 5930.

[17] ZHANG Y, ZHAO D, SUN J, et al. Adaptive convolutional neural network and its application in face recognition [J]. Neural Processing Letters, 2016, 43 (2)：389 – 399.

[18] MCFEE B, SALAMON J, BELLO J P. Adaptive pooling operators for weakly labeled sound event detection [J]. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2018, 26 (11)：2180 – 2193.

[19] HAN K, WANG Y, CHEN H, et al. A survey on vision transformer [J]. IEEE transactions on pattern analysis and machine intelligence, 2022, 45 (1)：87 – 110.

[20] VINAYAKUMAR R, ALAZAB M, SOMAN K, et al. Deep learning approach for intelligent intrusion detection system [J]. IEEE Access, 2019, 7：41525 – 41550.

[21] Wu Kehe, Chen Zuge, Li Wei. A novel intrusion detection model for a massive network using convolutional neural networks [J]. IEEE Access, 2018 (6)：50850 – 50859.

[22] ALTWAIJRY N, ALQAHTANI A, ALTURAIKI I. A deep learning approach for anomaly-based network intrusion detection [C]// Proceedings of the Big Data and Security：First International Conference, Nanjing, China, 2020：603 – 615.

[23] JIANG K, WANG W, WANG A, et al. Network intrusion detection combined hybrid sampling with deep hierarchical network [J]. IEEE Access, 2020 (8)：32464 – 32476.

（收稿日期：2023 – 05 – 07）

作者简介：

王明（1994 – ），男，硕士，助教，主要研究方向：深度学习、网络安全。