基于深度特征学习的网络流量异常检测方法

董书琴* 张 斌

(中国人民解放军战略支援部队信息工程大学 郑州 450001)

(河南省信息安全重点实验室 郑州 450001)

摘 要：针对网络流量异常检测过程中提取的流量特征准确性低、鲁棒性差导致流量攻击检测率低、误报率高等问题，该文结合堆叠降噪自编码器(SDA)和softmax，提出一种基于深度特征学习的网络流量异常检测方法。首先基于粒子群优化算法设计SDA结构两阶段寻优算法：根据流量检测准确率依次对隐藏层层数及每层节点数进行寻优，确定搜索空间中的最优SDA结构，从而提高SDA提取特征的准确性。然后采用小批量梯度下降算法对优化的SDA进行训练，通过最小化含噪数据重构向量与原始输入向量间的差异，提取具有较强鲁棒性的流量特征。最后基于提取的流量特征对softmax进行训练构建异常检测分类器，从而实现对流量攻击的高性能检测。实验结果表明：该文所提方法可根据实验数据及其分类任务动态调整SDA结构，提取的流量特征具有更高的准确性和鲁棒性，流量攻击检测率高、误报率低。

关键词：流量异常检测；深度学习；堆叠降噪自编码器；粒子群优化

中图分类号：TP393.08 **文献标识码**：A **文章编号**：1009-5896(2020)03-0695-09

DOI: 10.11999/JEIT190266

Network Traffic Anomaly Detection Method Based on Deep Features Learning

DONG Shuqin ZHANG Bin

(*PLA SSF Information Engineering University*, *Zhengzhou* 450001, *China*)

(*Henan Key Laboratory of Information Security*, *Zhengzhou* 450001, *China*)

Abstract: In view of the problems of low attack detection rate and high false positive rate caused by poor accuracy and robustness of the extracted traffic features in network traffic anomaly detection, a network traffic anomaly detection method based on deep features learning is proposed, which is combined with Stacked Denoising Autoencoders (SDA) and softmax. Firstly, a two-stage optimization algorithm is designed based on particle swarm optimization algorithm to optimize the structure of SDA, the number of hidden layers and nodes in each layer is optimized successively based on the traffic detection accuracy, and the optimal structure of SDA in the search space is determined, improving the accuracy of traffic features extracted by SDA. Secondly, the optimized SDA is trained by the mini-batch gradient descent algorithm, and the traffic features with strong robustness are extracted by minimizing the difference between the reconstruction vector of the corrupted data and the original input vector. Finally, softmax is trained by the extracted traffic features to construct an anomaly detection classifier for detecting traffic attacks with high performance. The experimental results show that the proposed method can adjust the structure of SDA based on the experimental data and its classification tasks, extract traffic features with a higher accuracy and robustness, and detect traffic attacks with high detection rate and low false positive rate.

Key words: Traffic anomaly detection; Deep learning; Stacked Denoising Autoencoders (SDA); Particle Swarm Optimization (PSO)

收稿日期：2019-04-18；改回日期：2019-10-09；网络出版：2019-10-16

*通信作者： 董书琴 dongshuqin377@126.com

基金项目：河南省基础与前沿技术研究计划基金(142300413201)，信息工程大学新兴科研方向培育基金(2016604703)，信息工程大学科研项目(2019f3303)

Foundation Items: The Foundation and Frontier Technology Research Project of Henan Province (142300413201), The New Research Direction Cultivation Fund of Information Engineering University (2016604703), The Research Project of Information Engineering University (2019f3303)

1 引言

网络流量异常检测是发现网络攻击的重要手段，随着流量特征维数以及噪声数据的增加，基于传统机器学习的流量异常检测方法面临流量特征提取准确性低、鲁棒性差等问题，在一定程度上降低了流量攻击检测性能。为此，基于深度学习的异常检测方法成为当前研究的热点[1]。

基于深度学习的异常检测方法主要有以下3种：(1)基于深度玻尔兹曼机的异常检测方法[2,3]，该类方法可通过对高维流量数据的学习提取其本质特征，从而提高对流量攻击的检测率，但该类方法提取特征的鲁棒性较差，当输入数据含有噪声时，其攻击检测性能变差；(2)基于堆叠自编码器(Stacked Auto Encoders, SAE)的异常检测方法[4,5]，该类方法可通过逐层对流量数据进行学习提取具有较高准确性的流量特征，但其提取特征的鲁棒性较差，当被测数据遭到破坏时，该类方法检测准确率降低；(3)基于卷积神经网络的异常检测方法[6,7]，该类方法提取的流量特征具有较强的鲁棒性，攻击检测性能较高，但需先将网络流量转换为图像，加大了数据处理负担，且未充分考虑网络结构信息对提取特征准确性的影响。

针对上述问题，文献[8]提出一种基于堆叠降噪自编码器(Stacked Denoising Autoencoders, SDA)[9]的异常检测方法，可有效提高大数据环境下提取流量特征的准确性和鲁棒性，同时避免了将流量转换为图像带来的额外处理负担；但其SDA仅有3个隐藏层且每层节点数相同，没有很好地发挥SDA的特征提取及降维能力，当训练数据较少时会在一定程度上降低提取特征的准确性，进而影响其流量攻击检测能力。

因此，本文提出一种基于两阶段寻优SDA的网络流量异常检测模型，在采用粒子群优化(Particle Swarm Optimization, PSO)算法对SDA结构进行两阶段寻优的基础上，对流量特征进行深度学习并构建异常检测分类器，从而实现对流量攻击的检测，主要内容如下：

(1) 设计一种基于PSO的SDA结构两阶段寻优算法。首先，采用PSO算法对SDA隐藏层层数进行寻优，确定最优隐藏层层数l_{gbest}；然后，在l_{gbest}维空间中对隐藏层每层节点数进行寻优，从而确定搜索空间中最优SDA结构，解决因SDA结构设置不当导致的提取特征准确性低的问题；

(2) 提出一种基于SDA的流量深度特征学习方法。采用小批量梯度下降算法对SDA进行训练，通过最小化含噪数据重构向量与原始输入向量间的差异，提取具有较强鲁棒性的流量特征，同时避免SDA训练速度慢及权重参数易陷入局部最优解的问题；

(3) 提出一种基于流量特征的softmax分类器构建方法。采用提取的流量特征对softmax进行训练构建异常检测分类器，同时通过对SDA参数的反向有监督微调进一步提高其提取特征的准确性，进而增强分类器的流量攻击检测能力。基于NSL-KDD数据集[10]的实验测试表明：采用本文所提方法提取的流量特征相比同类方法具有更高的准确性和鲁棒性，可有效提高流量攻击检测性能，且检测效果不受噪声数据影响。

2 基于两阶段寻优SDA的流量异常检测模型

基于两阶段寻优SDA的流量异常检测模型如图1所示，l为SDA隐藏层层数，该模型主要包含数据预处理、流量特征学习和流量异常检测3个组件，其中流量特征学习和异常检测组件是模型的重点，且流量特征学习阶段SDA隐藏层层数及每层节点数是影响其特征提取准确性的关键。因此，下面在设计SDA结构两阶段寻优算法的基础上，对流量特征学习方法与异常检测分类器构建方法进行描述。

2.1 SDA结构两阶段寻优算法

PSO算法被广泛应用于神经网络参数优化领域，可有效指导SDA结构寻优，但采用文献[11]所提一阶段寻优算法对SDA结构进行寻优时，其隐藏层每层节点数相同，若输入层破坏率σ较小，其编码函数学习为恒等变换的概率增大，会在一定程度上降低SDA的特征提取能力。为此，本文提出一种基于PSO的SDA结构两阶段寻优算法(PSO-SDA)，算法流程如图2所示，其中n表示隐藏层节点数，$t_{\max}$为粒子群最大迭代次数，$\boldsymbol{n}$表示l_{gbest}维空间中粒子的位置向量，$n_{\text{gbest}}^{(1)} - n_{\text{gbest}}^{(l_{\text{gbest}})}$表示SDA中第$1-l_{\text{gbest}}$隐藏层的节点数。

下面分别对PSO-SDA算法的隐藏层层数寻优算法和隐藏层每层节点数寻优算法进行描述。

(1) 隐藏层层数寻优算法

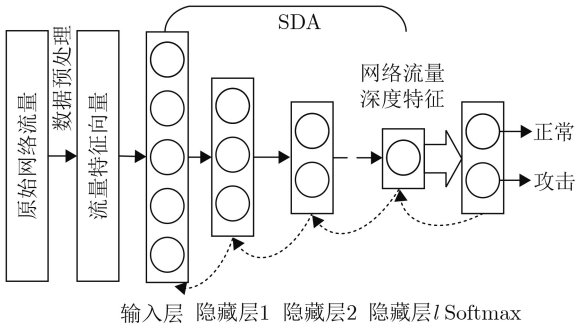

图 1 基于两阶段寻优SDA的流量异常检测模型

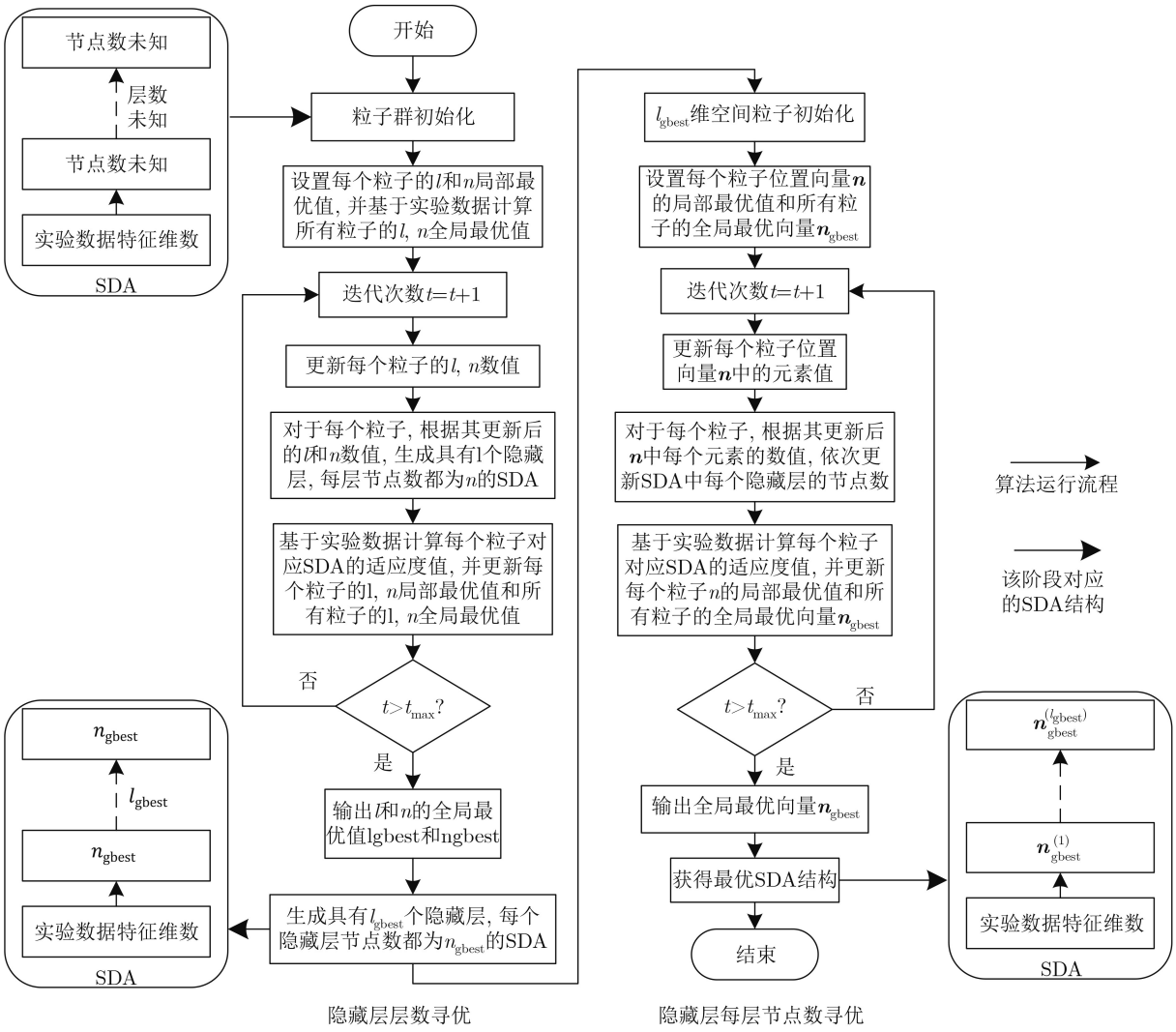

图 2 基于PSO的SDA结构两阶段寻优算法流程

对于由NP个粒子构成的粒子群而言，粒子群中第i个粒子在第t次迭代时$(i,t=1,2,\cdots,N$，且$i\in[1,\mathrm{NP}]$, $t\in[1,t_{\max}])$，粒子的速度和位置更新规则可用式(1)—式(4)表示。

$$v_{l_i}(t)=wv_{l_i}(t-1)+c_1r_1(l_{i,\mathrm{pbest}}-l_i(t-1))+c_2r_2(l_{\mathrm{gbest}}-l_i(t-1)) \tag{1}$$

$$l_i(t)=l_i(t-1)+v_{l_i}(t) \tag{2}$$

$$v_{n_i}(t)=wv_{n_i}(t-1)+c_1r_3(n_{i,\mathrm{pbest}}-n_i(t-1))+c_2r_4(n_{\mathrm{gbest}}-n_i(t-1)) \tag{3}$$

$$n_i(t)=n_i(t-1)+v_{n_i}(t) \tag{4}$$

其中，v_{l_i}和v_{n_i}分别表示粒子i对应l和n的速度，$l_{i,\mathrm{pbest}}$和$n_{i,\mathrm{pbest}}$分别表示粒子i对应l和n的局部最优值，l_{gbest}和n_{gbest}分别表示l和n的全局最优值，w为惯性参数，c_1和c_2为加速因子，r_1–r_4为0和1之间的随机小数。当$t=1$时，粒子i对应l和n的初始值及其速度初始值由式(5)—式(8)确定。

$$l_i(0)=l_{\min}+r\cdot(l_{\max}-l_{\min}) \tag{5}$$

$$v_{l_i}(0)=v_{l,\min}+r\cdot(v_{l,\max}-v_{l,\min}) \tag{6}$$

$$n_i(0)=n_{\min}+r\cdot(n_{\max}-n_{\min}) \tag{7}$$

$$v_{n_i}(0)=v_{n,\min}+r\cdot(v_{n,\max}-v_{n,\min}) \tag{8}$$

其中，$l_{\max}$和$l_{\min}$表示l的最大值和最小值，$v_{l,\max}$和$v_{l,\min}$表示l速度的最大值和最小值，$n_{\max}$和$n_{\min}$表示n的最大值和最小值，$v_{n,\max}$和$v_{n,\min}$表示n速度的最大值和最小值，r为0和1之间的随机小数。

算法迭代过程中$l_{i,\mathrm{pbest}}$, $n_{i,\mathrm{pbest}}$, l_{gbest},n_{gbest}依据适应度函数进行更新。在异常检测领域，准确率(Acc)是衡量检测算法性能优劣的关键指标，为此设计适应度函数如式(9)所示，$\mathrm{fit}(l_i(t),n_i(t))$越小，$l_i(t),n_i(t)$越优。

$$\mathrm{fit}(l_i(t),n_i(t))=1-\mathrm{Acc} \tag{9}$$

隐藏层层数寻优算法如表1所示。

(2) 隐藏层每层节点数寻优算法

表 1 隐藏层层数寻优算法

输入：流量异常检测数据集，NP,$t_{\max}$, w, c_1, c_2, $l_{\max}$, $l_{\min}$, $v_{l,\max}$, $v_{l,\min}$, $n_{\max}$, $n_{\min}$, $v_{n,\max}$, $v_{n,\min}$

输出：具有l_{gbest}个隐藏层且每层节点数为n_{gbest}的SDA

for $i=1$ to NP do

采用式(5)—式(8)对粒子群进行初始化，并分别将$l_{i,\text{pbest}}$和$n_{i,\text{pbest}}$初始化为$l_i(0)$和$n_i(0)$；

基于实验数据，采用式(9)计算粒子i的适应度值；

将最小适应度值对应的l和n设置为l_{gbest}和n_{gbest}初始化值；

for $t=1$to $t_{\max}$ do

for $i=1$to NP do

采用式(1)—式(4)更新粒子i的$l_i(t)$速度和数值，以及$n_i(t)$的速度和数值；

if $v_{l_i}(t)$, $l_i(t)$,$v_{n_i}(t)$ or $n_i(t)$超过其搜索范围

对$v_{l_i}(t)$, $l_i(t)$,$v_{n_i}(t)$ or $n_i(t)$再次进行随机初始化；

生成具有$l_i(t)$个隐藏层且每层节点数为$n_i(t)$的SDA；

基于实验数据，采用式(9)计算粒子i的适应度值；

if($\text{fit}(l_i(t),n_i(t))<\text{fit}(l_{i,\text{pbest}},n_{i,\text{pbest}})$)//若粒子$i$的适应度值小于局部最优值对应的适应度值，则对局部最优值进行更新

分别将$l_i(t)$和$n_i(t)$赋值给$l_{i,\text{pbest}}$和$n_{i,\text{pbest}}$；

if($\text{fit}(l_i(t),n_i(t))<\text{fit}(l_{\text{gbest}},n_{\text{gbest}})$)//若粒子$i$的适应度值小于全局最优值对应的适应度值，则对全局最优值进行更新

分别将$l_i(t)$和$n_i(t)$赋值给l_{gbest}和n_{gbest}；

迭代结束后，生成具有l_{gbest}个隐藏层且每层节点数为n_{gbest}的SDA；

return 具有l_{gbest}个隐藏层且每层节点数为n_{gbest}的SDA。

以隐藏层层数寻优算法生成的SDA为基础，在l_{gbest}维空间对隐藏层每层节点数进行寻优。第t次迭代时，粒子i在第$h(h=1,2,\cdots,N$，且$h\in[1,l_{\text{gbest}}])$维空间的速度$v_i^{(h)}(t)$和位置$n_i^{(h)}(t)$更新规则如式(10)、式(11)所示。

$$v_i^{(h)}(t)=wv_i^{(h)}(t-1)+c_1r_5\left(n_{i,\text{pbest}}^{(h)}-n_i^{(h)}(t-1)\right)+c_2r_6\left(n_{\text{gbest}}^{(h)}-n_i^{(h)}(t-1)\right) \tag{10}$$

$$n_i^{(h)}(t)=n_i^{(h)}(t-1)+v_i^{(h)}(t) \tag{11}$$

式(10)中，$n_{i,\text{pbest}}^{(h)}$表示粒子i的局部最优值在第h维空间的数值，$n_{\text{gbest}}^{(h)}$表示全局最优值在第h维空间的数值，r_5,r_6含义同r_1。当$t=1$时，粒子i在第h维空间的位置初始值$n_i^{(h)}(0)=n_{\text{gbest}}$，速度初始值由式(12)确定。

$$v_i^{(h)}(0)=v_{\min}+r\cdot(v_{\max}-v_{\min}) \tag{12}$$

其中，$v_{\max}$和$v_{\min}$分别表示速度最大值和最小值。

算法迭代过程中$n_{i,\text{pbest}}^{(h)}$和$n_{\text{gbest}}^{(h)}$根据式(13)所示适应度函数进行更新，且$\text{fit}(\boldsymbol{n}_i(t))$越小，$\boldsymbol{n}_i(t)$越优。

$$\text{fit}(\boldsymbol{n}_i(t))=1-\text{Acc} \tag{13}$$

隐藏层每层节点数寻优算法如表2所示。

2.2 基于SDA的流量深度特征学习方法

在确定最优SDA结构的基础上，采用小批量梯度下降算法对SDA进行逐层无监督预训练[12]，进而提取流量深度特征。具体地，在对第1个降噪自编码器(Denoising Autoencoders, DA)进行训练时，可采用式(14)、式(15)求得每个小批量中第a个样本对应隐藏层向量$\boldsymbol{y}$的第k'个元素和重构向量z的第k个元素。

$$y_{ak'}=s\left(\sum_{k=1}^{d}W_{k',k}\overline{x}_{ak}+b_{k'}\right) \tag{14}$$

$$z_{ak}=s\left(\sum_{k'=1}^{d'}W'_{k,k'}y_{ak'}+b'_k\right) \tag{15}$$

其中，$s(x)=\tanh(x)=(\text{e}^x-\text{e}^{-x})/(\text{e}^x+\text{e}^{-x})$, $W_{k',k}$为编码权重矩阵$\boldsymbol{W}$的第k'行第k列元素，$\overline{x}_{ak}$为第a个样本对应输入向量$\boldsymbol{x}$采用masking加噪所得$\overline{\boldsymbol{x}}$的第$k$个元素，$b_{k'}$为编码偏置向量$\boldsymbol{b}$的第$k'$个元素；$W'_{k,k'}$为解码权重矩阵$\boldsymbol{W}'$($\boldsymbol{W}'=\boldsymbol{W}^{\text{T}}$)的第$k$行第$k'$列元素，$b'_k$为解码偏置向量$\boldsymbol{b}'$的第$k$个元素。

DA训练过程中，参数$W_{k',k},b_{k'},b'_k$的最优解可在随机初始化$b_{k'},b'_k$并采用Xavier方法[13]初始化$W_{k',k}$的基础上，通过最小化式(16)所示损失函数获得，参数更新规则如式(17)所示。

$$L(\boldsymbol{x},\boldsymbol{z})=-\frac{1}{m}\sum_{a=1}^{m}\sum_{k=1}^{d}[x_{ak}\ln z_{ak}+(1-x_{ak})\ln(1-z_{ak})] \tag{16}$$

表 2　隐藏层每层节点数寻优算法

输入：流量异常检测数据集，NP, $t_{\max}$, w, c_1, c_2, $v_{\max}$, $v_{\min}$, l_{gbest}, n_{gbest}

输出：最优SDA结构

for $i=1$ to NP do

　for $h=1$ to l_{gbest} do

　　初始化粒子位置$n_i^{(h)}(0)=n_{\text{gbest}}$，采用式(12)初始化粒子速度，并将局部最优向量$\boldsymbol{n}_{i,\text{pbest}}$中的$n_{i,\text{pbest}}^{(h)}$初始化为$n_{\text{gbest}}$；

设置全局最优向量$\boldsymbol{n}_{\text{gbest}}=\min\{\boldsymbol{n}_{1,\text{pbest}},\boldsymbol{n}_{2,\text{pbest}},\cdots,\boldsymbol{n}_{\text{NP,pbest}}\}=[n_{\text{gbest}}\,n_{\text{gbest}}\cdots n_{\text{gbest}}]^{\text{T}}$；

for $t=1$to $t_{\max}$ do

　for $i=1$to NP do

　　for $h=1$tol_{gbest} do

　　采用式(10)和式(11)更新粒子i位置向量$\boldsymbol{n}_i(t)$中元素$n_i^{(h)}(t)$的速度和数值；

　　if $v_i^{(h)}(t)$ or $n_i^{(h)}(t)$超过其搜索范围

　　对$v_i^{(h)}(t)$ or $n_i^{(h)}(t)$再次进行随机初始化；

　　根据更新后的$\boldsymbol{n}_i(t)$，将SDA每个隐藏层的节点数分别更新为$n_i^{(1)}(t),n_i^{(2)}(t),\cdots,n_i^{(l_{\text{gbest}})}(t)$；

　　基于实验数据，采用式(13)计算粒子i的适应度值；

　　if($\text{fit}(\boldsymbol{n}_i(t))<\text{fit}(\boldsymbol{n}_{i,\text{pbest}})$)//若粒子$i$的适应度值小于局部最优向量对应的适应度值，则对局部最优向量进行更新

　　$\boldsymbol{n}_{i,\text{pbest}}\leftarrow\boldsymbol{n}_i(t)$；

　$\boldsymbol{n}_{\text{gbest}}\leftarrow\min\{\boldsymbol{n}_{1,\text{pbest}},\boldsymbol{n}_{2,\text{pbest}},\cdots,\boldsymbol{n}_{\text{NP,pbest}}\}$；//采用局部最优向量中的最小值更新全局最优向量

迭代结束后，根据最终$\boldsymbol{n}_{\text{gbest}}$分别将SDA的隐藏层每层节点数更新为$n_{\text{gbest}}^{(1)},n_{\text{gbest}}^{(2)},\cdots,n_{\text{gbest}}^{(l_{\text{pbest}})}$；

return 最优SDA结构。

$$\left.\begin{aligned}
W_{k',k}&=W_{k',k}-\frac{\varepsilon}{m}\left[\sum_{a=1}^{m}\sum_{k=1}^{d}\frac{z_{ak}-x_{ak}}{z_{ak}(1-z_{ak})}\right](1-z_{ak}^2)\\
&\cdot\left[\sum_{k'=1}^{d'}W'_{k,k'}(1-y_{ak'}^2)\right]\left(\sum_{k=1}^{d}\overline{x_{ak}}\right)\\
b_{k'}&=b_{k'}-\frac{\varepsilon}{m}\left[\sum_{a=1}^{m}\sum_{k=1}^{d}\frac{z_{ak}-x_{ak}}{z_{ak}(1-z_{ak})}\right](1-z_{ak}^2)\\
&\cdot\left[\sum_{k'=1}^{d'}W'_{k,k'}(1-y_{ak'}^2)\right]\\
b'_k&=b'_k-\frac{\varepsilon}{m}\left[\sum_{a=1}^{m}\sum_{k=1}^{d}\frac{z_{ak}-x_{ak}}{z_{ak}(1-z_{ak})}\right](1-z_{ak}^2)
\end{aligned}\right\}\tag{17}$$

其中，m为每个小批量的样本数，x_{ak}为第a个样本对应输入向量$\boldsymbol{x}$的第k个元素，ε表示学习率。

在训练完所有无标签训练样本并进行设定的迭代次数后，隐藏层向量$\boldsymbol{y}$即为原始流量经DA学习到的网络流量特征。然后将第1个DA的隐藏层向量$\boldsymbol{y}$作为第2个DA的输入向量，按上述过程采用式(17)对SDA进行逐层训练，最终依次确定SDA中各个DA的权重和偏置参数，并输出流量深度特征$\boldsymbol{y}^{(l)}$。

2.3　基于流量特征的softmax分类器构建方法

异常检测分类器构建过程中，首先基于SDA提取带标签验证样本V的流量深度特征$\boldsymbol{y}_V^{(l)}$并输入softmax，$\boldsymbol{y}_V^{(l)}$经softmax处理后输出样本V的预测类别V_{c}，

$$\begin{aligned}
V_{\text{c}}&=\arg\max_q[\text{softmax}_{\text{q}}(\boldsymbol{W}_V\boldsymbol{y}_V^{(l)}+\boldsymbol{b}_V)]\\
&=\arg\max_q\frac{1}{\sum_{q=1}^{Q}\text{e}^{\boldsymbol{W}_V(q)\boldsymbol{y}_V^{(l)}+b_V(q)}}\\
&\cdot\begin{bmatrix}\text{e}^{\boldsymbol{W}_V(1)\boldsymbol{y}_V^{(l)}+b_V(1)}\\ \text{e}^{\boldsymbol{W}_V(2)\boldsymbol{y}_V^{(l)}+b_V(2)}\\ \vdots\\ \text{e}^{\boldsymbol{W}_V(Q)\boldsymbol{y}_V^{(l)}+b_V(Q)}\end{bmatrix}
\end{aligned}\tag{18}$$

其中，$\boldsymbol{W}_V,\boldsymbol{b}_V$分别表示SDA隐藏层$\boldsymbol{l}$与softmax间的权重矩阵及偏置向量，$\boldsymbol{W}_V(q)$表示权重矩阵$\boldsymbol{W}_V$中第$q$行，$b_V(q)$表示偏置向量$\boldsymbol{b}_V$中的第$q$个元素，$q=1,2,\cdots,Q$。

然后通过最小化式(19)所示分类代价函数[14]对参数$\boldsymbol{W}_V(q)$和$b_V(q)$进行训练，其更新规则如式(20)所示

$$\begin{aligned}
J(\boldsymbol{W}_V,\boldsymbol{b}_V)=&-\frac{1}{n_V}\left[\sum_{V=1}^{n_V}\sum_{q=1}^{Q}1\{V_{\text{t}}=q\}\right.\\
&\left.\cdot\ln\frac{\text{e}^{\boldsymbol{W}_V(q)\boldsymbol{y}_V^{(l)}+b_V(q)}}{\sum_{q'=1}^{Q}\text{e}^{\boldsymbol{W}_V(q')\boldsymbol{y}_V^{(l)}+b_V(q')}}\right]
\end{aligned}\tag{19}$$

$$\left.\begin{aligned}\boldsymbol{W}_V(q)&=\boldsymbol{W}_V(q)+\frac{\varepsilon}{n_V}\left[\sum_{V=1}^{n_V}\sum_{q=1}^{Q}1\{V_{\rm t}=q\}\right.\\&\left.\cdot\boldsymbol{y}_V^{(l)}\sum_{q'=1}^{Q}{\rm e}^{\boldsymbol{W}_V(q')\boldsymbol{y}_V^{(l)}+b_V(q')}\right]\\b_V(q)&=b_V(q)+\frac{\varepsilon}{n_V}\left[\sum_{V=1}^{n_V}\sum_{q=1}^{Q}1\{V_{\rm t}=q\}\right.\\&\left.\cdot\sum_{q'=1}^{Q}{\rm e}^{\boldsymbol{W}_V(q')\boldsymbol{y}_V^{(l)}+b_V(q')}\right]\end{aligned}\right\}\tag{20}$$

式(19)中，n_V表示验证样本总数，$V_{\rm t}$表示样本V的实际类别，当$V_{\rm t}=q$时，$1\{V_{\rm t}=q\}=1$，否则，$1\{V_{\rm t}=q\}=0$。

$\boldsymbol{W}_V,\boldsymbol{b}_V$训练完成后即构建起流量异常检测分类器。最后对SDA权重与偏置参数进行反向有监督微调，在提高SDA提取特征准确性的基础上，进一步增强softmax分类器的流量攻击检测能力。

3 实验及结果分析

本文基于windows10平台和广泛用于异常检测算法性能测试的NSL-KDD数据集，采用Python3.6.5和Tensorflow1.8.0进行实验，对PSO-SDA算法进行验证，并对流量异常检测模型进行测试。

3.1 实验数据

选择NSL-KDD的“KDDTrain+”作为训练集，并去掉其类别标签；选择“KDDTrain+_20Percent”作为验证集；选择“KDDTest+”作为测试集；每个数据集包含5类流量：Normal, DoS, Probe, R2L和U2R。

由于实验主要关注从流量方面检测网络异常的效果，因此只保留数据集中与流量相关的28个特征[15]，同时将符号型特征转变为数值型特征，并采用最小-最大规范化方法对28维特征的属性值进行归一化处理[2]，从而构建流量特征向量用于异常检测模型的训练，进而实现流量异常检测。

3.2 评价标准

采用分类器训练时间$T_{\rm tr}$、检测时间$T_{\rm te}$、准确率[2](Acc)、检测率[2](Detection Rate, DR)、召回率[6](Rec)和误报率[2](False Positive Rate, FPR)作为评价指标对PSO-SDA算法和异常检测模型进行测试。

3.3 参数设置

由于在流量异常检测领域，当前用到的深度学习模型通常不超过20层，且SDA在提取流量特征的同时还可对流量特征进行降维，故设置$l_{\max}=20$, $n_{\max}=28$，并设置$w=0.8$, $c_1=c_2=2$, NP$=10$, $t_{\max}=10$, $l_{\min}=1$, $n_{\min}=1$[11]，则$v_{l,\max}=1.9$, $v_{l,\min}=-1.9$, $v_{n,\max}=v_{\max}=2.7$, $v_{n,\min}=v_{\min}=-2.7$。设置$\sigma=0.1$, $m=50$，无监督训练和有监督微调阶段迭代次数分别为50和200，采用Adam算法[16]对ε进行自适应优化。

3.4 实验结果分析

3.4.1 SDA结构寻优

根据分类目标不同，实验主要从二分类和多分类场景出发，验证PSO-SDA算法性能，同时给出搜索空间中的最优SDA结构。二分类和多分类场景下SDA结构寻优过程中适应度值变化情况如图3、图4所示。

由图3可知，二分类场景下隐藏层层数寻优算法的适应度值最终收敛于1.63×10^{-1}，其对应的隐藏层层数$l_{\rm gbest}=8$，隐藏层每层节点数$n_{\rm gbest}=2$。隐藏层每层节点数寻优算法最终收敛于1.46×10^{-1}，其对应的隐藏层结构为[3, 2, 2, 2, 1, 3, 3, 3]，则二分类场景下给定搜索空间中的最优SDA结构为[28, 3, 2, 2, 2, 1, 3, 3, 3]。

由图4可知，多分类场景下隐藏层层数寻优算法的适应度值最终收敛于2.17×10^{-1}，其对应的隐藏层层数$l_{\rm gbest}=1$，隐藏层每层节点数$n_{\rm gbest}=25$；隐藏层每层节点数寻优算法最终收敛于2.15×10^{-1}，其对应的隐藏层节点数为24，则多分类场景下给定搜索空间中的最优SDA结构为[28, 24]。

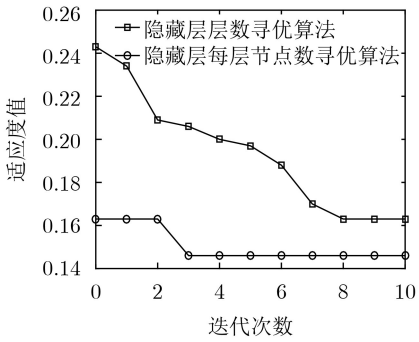

图 3 二分类场景下SDA结构寻优过程

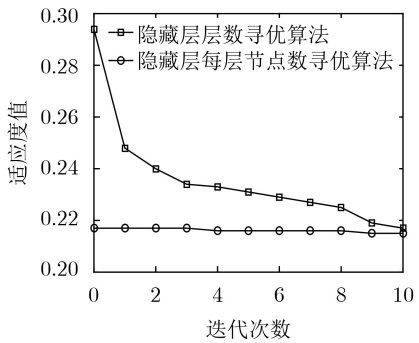

图 4 多分类场景下SDA结构寻优过程

综上，PSO-SDA算法可有效根据实验数据及其分类任务确定搜索空间中的最优SDA结构。

3.4.2 模型检测性能测试

在上述SDA结构的基础上，从二分类和多分类场景出发，测试SAE、传统SDA[8]、一阶段寻优SDA[11]和两阶段寻优SDA应用于异常检测模型时的检测性能。为说明SDA优势，实验中设置SAE结构与两阶段寻优SDA一致。每个实验独立重复10次并取平均值，实验结果如表3、表4所示。

由表3可知，二分类场景下虽然基于两阶段寻优SDA的异常检测模型T_{tr}略有提升，但其检测性能最好，相比基于SAE、传统SDA和一阶段寻优SDA的异常检测模型，准确率分别提高了7.41%, 7.12%和7.05%，检测率分别提高了4.25%, 0.7%和2.16%，召回率分别提高了4.93%, 1.94%和5.85%，且误报率分别降低了82.35%, 24.26%和29.04%。

由表4可知，多分类场景下虽然基于两阶段寻优SDA的异常检测模型对R2L和U2R流量的检测性能适中，但其流量检测准确率最高，同时对Normal, DoS和Probe流量的检测率和召回率较高、误报率较低。相比基于SAE、传统SDA和一阶段寻优SDA的异常检测模型，准确率分别提高了1.49%, 1.26%和0.48%；对Normal流量的检测率分别提高了2.08%, 1.14%和0.55%，召回率分别提高了

表 3　二分类场景不同模型检测性能

模型类型	基于SAE的异常检测模型	基于传统SDA的异常检测模型	基于一阶段寻优SDA的异常检测模型	基于两阶段寻优SDA的异常检测模型
模型结构	[28, 3, 2, 2, 2, 1, 3, 3, 3, 2]	[28, 28, 28, 28, 2]	[28, 2, 2, 2, 2, 2, 2, 2, 2, 2]	[28, 3, 2, 2, 2, 1, 3, 3, 3, 2]
Acc (%)	86.29	86.52	86.58	92.68
DR (%)	92.85	96.10	94.75	96.80
Rec (%)	90.04	92.68	89.26	94.48
FPR (%)	4.96	3.38	3.51	2.72
$T_{\mathrm{tr}}(m)$	8.24	8.52	7.45	8.50
$T_{\mathrm{te}}(s)$	0.18	0.18	0.18	0.18

表 4　多分类场景不同模型检测性能

模型类型		基于SAE的异常检测模型	基于传统SDA的异常检测模型	基于一阶段寻优SDA的异常检测模型	基于两阶段寻优SDA的异常检测模型
模型结构		[28, 24, 5]	[28, 28, 28, 28, 5]	[28, 25, 5]	[28, 24, 5]
Acc (%)		84.12	84.31	84.96	85.37
Normal	DR (%)	84.58	85.37	85.87	86.34
	Rec (%)	96.74	96.88	97.01	97.28
	FPR (%)	17.98	18.89	18.06	17.25
DoS	DR (%)	94.08	94.74	94.92	95.59
	Rec (%)	83.65	84.51	82.63	85.88
	FPR (%)	2.05	2.04	2.02	1.72
Probe	DR (%)	79.42	75.58	79.71	83.27
	Rec (%)	65.14	67.29	63.78	68.28
	FPR (%)	1.78	2.21	1.70	1.34
R2L	DR (%)	90.96	92.06	83.78	90.50
	Rec (%)	58.23	60.99	58.34	60.23
	FPR (%)	0.27	0.21	0.57	0.30
U2R	DR (%)	88.05	28.60	72.58	76.19
	Rec (%)	2.50	2.00	4.50	3.00
	FPR (%)	0.01	0.03	0.01	0.01
$T_{\mathrm{tr}}(m)$		3.94	6.32	6.54	5.36
$T_{\mathrm{te}}(s)$		0.20	0.40	0.41	0.26

0.56%, 0.41%和0.28%，误报率分别下降了4.23%, 9.51%和4.70%；对DoS攻击流量的检测率分别提高了1.61%, 0.90%和0.71%，召回率分别提高了2.67%, 1.62%和3.93%，误报率分别下降了19.19%, 18.60%和17.44%；对Probe攻击流量的检测率分别提高了4.85%, 10.17%和4.47%，召回率分别提高了4.82%, 1.47%和7.06%，误报率分别下降了32.84%, 64.93%和26.87%；且其T_{tr}和T_{te}优于基于传统SDA和一阶段寻优SDA的异常检测模型。

综上，两阶段寻优SDA可有效提高提取流量特征的准确性，进而提高异常检测模型的检测性能。

3.4.3 模型噪声鲁棒性测试

当被测数据属性特征破坏率为0.1, 0.2和0.3时，对比多分类场景下不同模型(模型结构同表4)的流量检测准确率，每个实验独立重复10次并取平均值，实验结果如表5表示。

表 5 多分类场景不同模型检测含噪流量的准确率

模型类型	Acc (%)		
	0.1	0.2	0.3
基于SAE的异常检测模型	81.57	79.31	76.69
基于传统SDA的异常检测模型	83.63	83.54	83.48
基于一阶段寻优SDA的异常检测模型	84.71	84.52	84.23
基于两阶段寻优SDA的异常检测模型	85.08	85.01	85.02

由表5可知，当被测数据特征遭到破坏时，基于SAE的异常检测模型流量检测准确率降低，且随着破坏率的增大其准确率持续下降；而基于SDA的异常检测模型流量检测准确率下降幅度较小，且加大特征破坏率并不会进一步降低其准确率，表明SDA具有较强的噪声鲁棒性，可有效降低噪声数据对提取流量特征准确性的影响，且两阶段寻优SDA的鲁棒性最强，其异常检测模型流量检测准确率最高。

4 结束语

本文提出一种基于深度特征学习的网络流量异常检测方法，该方法可根据实验数据及其分类任务的不同动态寻优搜索空间中的SDA结构；相比SAE、传统SDA和一阶段寻优SDA而言，采用该方法提取的流量特征具有更高的准确性和鲁棒性，可有效提高异常检测模型在二分类场景下的检测性能，以及多分类场景下的流量检测准确率及Normal流量和DoS, Probe攻击检测性能。下一步将提高R2L和U2R流量攻击检测性能，并对PSO-SDA算法和SDA训练方法作并行化处理，加快SDA结构寻优和训练过程。

参 考 文 献

[1] KWON D, KIM H, KIM J, *et al*. A survey of deep learning-based network anomaly detection[J]. *Cluster Computing*, 2019, 22(Suppl 1): 949–961.

[2] 高妮, 高岭, 贺毅岳, 等. 基于自编码网络特征降维的轻量级入侵检测模型[J]. 电子学报, 2017, 45(3): 730–739. doi: 10.3969/j.issn.0372-2112.2017.03.033.
GAO Ni, GAO Ling, HE Yiyue, *et al*. A lightweight intrusion detection model based on autoencoder network with feature reduction[J]. *Acta Electronica Sinica*, 2017, 45(3): 730–739. doi: 10.3969/j.issn.0372-2112.2017.03.033.

[3] ALRAWASHDEH K and PURDY C. Toward an online anomaly intrusion detection system based on deep learning[C]. The 15th IEEE International Conference on Machine Learning and Applications, Anaheim, USA, 2016: 195–200. doi: 10.1109/ICMLA.2016.0040.

[4] JAVAID A, NIYAZ Q, SUN Weiqing, *et al*. A deep learning approach for network intrusion detection system[C]. The 9th EAI International Conference on Bio-inspired Information and Communications Technologies, New York, USA, 2015: 21–26. doi: 10.4108/eai.3-12-2015.2262516.

[5] YOUSEFI-AZAR M, VARADHARAJAN V, HAMEY M, *et al*. Autoencoder-based feature learning for cyber security applications[C]. The 2017 International Joint Conference on Neural Networks, Anchorage, USA, 2017: 3854–3861. doi: 10.1109/IJCNN.2017.7966342.

[6] WANG Wei, ZHU Ming, ZENG Xuewen, *et al*. Malware traffic classification using convolutional neural network for representation learning[C]. 2017 International Conference on Information Networking, Da Nang, Vietnam, 2017: 712–717. doi: 10.1109/ICOIN.2017.7899588.

[7] 王勇, 周慧怡, 侯皓, 等. 基于深度卷积神经网络的网络流量分类方法[J]. 通信学报, 2018, 39(1): 14–23. doi: 10.11959/j.issn.1000-436x.2018018.
WANG Yong, ZHOU Huiyi, FENG Hao, *et al*. Network traffic classification method basing on CNN[J]. *Journal on Communications*, 2018, 39(1): 14–23. doi: 10.11959/j.issn.1000-436x.2018018.

[8] YU Yang, LONG Jun, and CAI Zhiping. Session-based network intrusion detection using a deep learning architecture[C]. The 14th International Conference on Modeling Decisions for Artificial Intelligence, Kitakyushu, Japan, 2017: 144–155. doi: 10.1007/978-3-319-67422-3_13.

[9] VINCENT P, LAROCHELLE H, LAJOIE I, *et al*. Stacked Denoising Autoencoders: Learning useful representations in a deep network with a local denoising criterion[J]. *The Journal of Machine Learning Research*, 2010, 11: 3371–3408.

[10] Canadian Institute for Cybersecurity. NSL-KDD dataset[EB/OL]. https://www.unb.ca/cic/datasets/nsl.html, 2018.

[11] QOLOMANY B, MAABREH M, AL-FUQAHA, *et al*. Parameters optimization of deep learning models using particle swarm optimization[C]. The 13th International Wireless Communications and Mobile Computing Conference, Valencia, Spain, 2017: 1285–1290. doi: 10.1109/IWCMC.2017.7986470.

[12] WANG Yao, CAI Wandong, and WEI Pengcheng. A deep learning approach for detecting malicious JavaScript code[J]. *Security and Communication Networks*, 2016, 9(11): 1520–1534. doi: 10.1002/sec.1441.

[13] 陈建廷, 向阳. 深度神经网络训练中梯度不稳定现象研究综述[J]. 软件学报, 2018, 29(7): 2071–2091. doi: 10.13328/j.cnki.jos.005561.

CHEN Jianting and XIANG Yang. Survey of unstable gradients in deep neural network training[J]. *Journal of Software*, 2018, 29(7): 2071–2091. doi: 10.13328/j.cnki.jos.005561.

[14] 谷丛丛, 王艳, 严大虎, 等. 基于自编码组合特征提取的分类方法研究[J]. 系统仿真学报, 2018, 30(11): 4132–4140. doi: 10.16182/j.issn1004731x.joss.201811011.

GU Congcong, WANG Yan, YAN Dahu, *et al*. Research on classification based on autoencoder combination features extraction method[J]. *Journal of System Simulation*, 2018, 30(11): 4132–4140. doi: 10.16182/j.issn1004731x.joss.201811011.

[15] FIORE U, PALMIERI F, CASTIGLIONE A, *et al*. Network anomaly detection with the restricted Boltzmann machine[J]. *Neurocomputing*, 2013, 122: 13–23. doi: 10.1016/j.neucom.2012.11.050.

[16] KINGMA D and BA J. Adam: A method for stochastic optimization[C/OL]. https://arxiv.org/abs/1412.6980, 2017.

董书琴：男，1990年生，博士生，研究方向为网络安全态势感知.

张　斌：男，1969年生，教授，博士生导师，研究方向为网络空间安全.