基于 LSTM 与改进残差网络优化的异常流量检测方法

麻文刚，张亚东，郭进
（西南交通大学信息科学与技术学院，四川 成都 611756）

摘 要：传统的网络异常流量检测方法往往存在特征选择差与泛化能力较弱等缺陷，导致检测精度较低。为此，提出了一种基于长短记忆网络（LSTM）与改进残差神经网络优化的异常流量检测方法。首先分析网络流量特征，通过预处理来降低网络流量特征值的差异性；然后设计了一种三层堆叠 LSTM 网络来提取不同深度的网络流量特征；最后设计了一种带跳跃连接线的改进残差神经网络对 LSTM 进行优化，改善了深度神经网络中的过拟合与梯度消失等缺点，从而提高网络异常流量检测的准确率。实验表明，所提方法具有较高的训练准确率，数据处理的可视性效果较好，二分类和多分类下的分类准确率分别为 92.3%和 89.3%。与当前入侵检测方法相比，所提方法在精确率、召回率等参数最优时具有最低的误报率。在数据样本在遭到破坏时具有较强的稳健性，同时也具有较好的泛化能力。

关键词：异常流量检测；长短记忆网络；数据池化层；空洞卷积；改进残差神经网络
中图分类号：TP393.08
文献标识码：A
DOI: 10.11959/j.issn.1000−436x.2021109

Abnormal traffic detection method based on LSTM and improved residual neural network optimization

MA Wengang, ZHANG Yadong, GUO Jin
School of Information Science and Technology, Southwest Jiaotong University, Chengdu 611756, China

Abstract: Problems such as a difficulty in feature selection and poor generalization ability were prone to occur when traditional method was exploited to detect abnormal network traffic. Therefore, an abnormal traffic detection method based on the long short term memory network (LSTM) and improved residual neural network optimization was proposed. Firstly, the features and attributes of network traffic were analyzed, and the variability of the feature values was reduced by preprocessing of network traffic. Then, a three-layer stacked LSTM network was designed to extract network traffic features of different depths. Moreover, the problem of weak adaptability of feature extraction was solved. Finally, an improved residual neural network with skipping connecting line was designed to optimize the LSTM. The defects of deep neural network such as overfitting and gradient vanishing were optimized. The accuracy of abnormal traffic detection was improved. Experimental results show that the proposed method has higher training accuracy and better visibility of data processing. The classification accuracy rates under two classifications and multiple classifications are 92.3% and 89.3%. It has the lowest false positive rate when the parameters such as precision rate and recall rate are optimal. Moreover, it has strong robustness when the sample is destroyed. Furthermore, better generalization ability can be achieved.

Keywords: abnormal traffic detection, LSTM, data pooling layer, dilated convolution, improved residual neural network

收稿日期：2020−10−30；**修回日期**：2021−04−12
通信作者：张亚东，ydzhang@home.swjtu.edu.cn
基金项目：国家自然科学基金资助项目（No.61703349）；中央高校基本科研业务费专项资金资助项目（No.2682017CX101）；中国铁路总公司科技研究开发计划课题基金资助项目（No.N2018G062, No.K2018G011）
Foundation Items: The National Natural Science Foundation of China (No.61703349), The Fundamental Research Funds for the Central Universities (No.2682017CX101), China Railway Corporation Science and Technology Research and Development Project (No.N2018G062, No.K2018G011)

1 引言

互联网技术的出现使工控行业（如铁路、道路交通控制、制造及工业测量等）越来越智能化与标准化，各种行业之间的联系不再单一，互联网与物联网技术的发展为各个工业系统发展的合作通信以及业务交流带来一个很好的契机[1]。随着网络环境的不断更迭，工业网络将受到各种形式的安全威胁与网络攻击。从工业控制的各种信号系统[2-4]来讲，现有工业信号系统想要发展，就必须结合互联网技术不断更新信号数据与相关程序信息，然而信息技术的不断更新与发展，势必造成互联网技术被大量运用。但现有的网络安全防护技术依然存在很多不足，同时大数据量的需求也会使各个行业的网络配置相当复杂，各方面通信与连接的网络也较多，导致通信行业的网络安全防护存在极易被攻击的危险。因此各个国家与组织越来越注重对网络安全信息的攻击检测与防御，保障信号及其工业控制的网络安全，力争在网络空间安全中掌握主动。作为网络攻击中必不可少的一项工作，网络异常流量检测与分类成为当下工作的首要研究方向与重中之重。同时，近年来不断出现的异常流量、漏洞利用方法及深度学习方法的广泛应用，使网络异常检测又一次成为当下的研究热点。由于传统的检测方式在网络流量检测时泛化能力与特征选择较差，已经面临失效。因此寻找一种有效并且准确检测与区分异常流量的方法是目前工业与信号网络中需要解决的关键问题。在网络安全领域，异常流量检测一直是热点问题。本文的主要研究工作为分析当前网络异常流量检测的主体方法，针对现有处理网络流量过程中的不足进行改进。因此本文将分析最近几年网络异常流量检测的相关研究现状。

众所周知，对于网络流量特征的预处理能够使后续的异常流量检测更加高效，检测准确率更高。而现有方法都是将网络数据结合到一起，利用简单的人工分析、归一化及标准化完成，其处理效果一般。在利用机器学习进行异常流量检测时，要对网络数据中的大量信息特征进行文本分析。网络攻击流量特征具有高时序、乱排序及样本大的特点，存在不同方式的加密与替换，导致异常网络流量含有很多无用信息，因此需对数据进行预处理使后续检测效果更好。文献[5]通过流量特征之间的同义词转换，根据文本词库对流量信息进行不断扩展达到数据预处理的目的。Lu 等[6]为使支持向量机（SVM, support vector machine）具有更好的分类检测效果，根据传统的高斯随机方法构造一种半监督式的模型进行样本预处理，最终实现前端预处理的数据操作。Park 等[7]提出了一种字符级二值图像变换的卷积自动编码器，实现了对 Http 消息的异常检测，其中采用字符级分割的数据预处理方法，检测性能优于传统方法。Yu 等[8]将双向长短记忆网络（LSTM, long short-term memory）与深度神经网络结合用于网络中的恶意流量检测，通过将字符进行不断分割，从而完成对样本前期的处理与分析，该方法虽能取得较好结果，但时间复杂度较高。

而关于异常流量检测分类方法，Yang 等[9]设计了一种卷积门控递归单元神经网络，通过分析统一资源定位符（URL, uniform resource locator）流量特征搭建了一种检测的样本词库，并将该样本词库进行训练，最终用于文本分类特征 URL 检测。Choraś 等[10]通过模拟网络中不断发送的链接与获取的 Web 程序，从而生成网络中的正常流量样本信息。同时通过从各种网络协议与接收的网络数据包获取异常数据样本，最终利用一种图分割理论与卷积神经网络将其样本用于训练，实现不同的异常流量检测。Kruegel 等[11]搭建了一种多模型融合网络用于异常流量检测，该模型通过接收网络中协议的 Http 信息请求指令，分析指令中的多个网络数据特征来对异常流量进行分类检测，虽能取得一定效果，但泛化能力较差。Corona 等[12]通过训练网络中采集到的网络流量样本，设计完成了一种多分类网络异常流量检测模型。该模型在检测前就已设计好一个正常流量定义模型，同时模型将其用在网络协议中的各个连接段，最终结合统计分布与马尔可夫链完成了不同攻击流量的分类。该模型效果较好，但是稳健性一般。Ringberg 等[13]分析了来自 2 个 IP 骨干网（Abilene 和 Geant）和 3 个不同流量聚合（入口路由器、OD 流和输入链路）的一周全网流量测量，并对每个异常流量进行了特征时间序列的详细检查，得出了误报率对正常子空间中主分量数量的微小差异非常敏感的结论。Al-Obeidat 等[14]首先提出一种属性选择方法对其网络特征进行最优特征选择，然后提出混合机学习方法对其样本进行训练，最终结合模糊决策树方法完成异常流量分类，该方法效果较好。Erfani 等[15]采用信念网络（DBN, deep belief network）来提取通用的底层特征，用一个单分类 SVM 从 DBN

学习特征，最终完成了异常检测。但在利用 DBN 进行训练数据时会出现过拟合情况[16]，训练效果一般，因此分类性能一般。Zhang 等[17]将隐马尔可夫及 SVM 分类 3 种模型结合到一起，提出一种多模型网络异常流量检测方法。该方法通过对网络中接收到的不同网络字段内容进行异常检测，如果有一种模型显示异常，则总体异常。该方法虽能在二分类时取得较好的分类结果，但由于采用的模型皆为传统机器学习方法，因此在多分类任务中效果较差，稳健性一般。通过分析可知，上述研究内容都是采用一种机器学习方法或神经网络模型对网络流量特征进行训练，然后根据选择的分类器完成检测分类。但大部分研究都未采取检测分类前的数据预处理，而采取预处理的也只是利用相关文本分析与数据简单的归一化完成，因此得到的最优特征还是存在冗余性[18-21]。同时，只利用归一化等操作处理网络流量数据，会使原始网络数据样本的一些重要特征被直接滤除，不能达到数据最优清洗的目的。同样，仅仅使用样本增强也会使数据样本不断扩大，导致数据维度爆炸，最终使恶意流量特征及语义特征得不到保留。因此最终的检测方法特征选择较差，泛化能力一般。再者，现有的神经网络方法在训练数据时也会产生过拟合与梯度消失的缺点，导致模型训练丢失率过大，训练性能一般。

综上分析，本文的改进主要分为以下三点。1) 根据网络流量特征是否为数值型数据将其进行不同操作处理，数值型数据进行归一化，否则进行独热编码，最终将两部分的数据进行特征排序与特征清洗完成数据预处理；2) 设计一种三层堆叠 LSTM 网络来解决单层 LSTM 网络适应性弱的问题；3) 结合 Inception 结构与空洞卷积，设计一种带跳跃连接线的改进残差神经网络对 LSTM 进行优化，改善深度神经网络中的过拟合与梯度消失等缺点，最终利用 2 个公开的网络安全数据集来完成提出方法的性能验证。

2　相关理论

2.1　网络流量特征属性构建

在网络流量中，固定时间窗口 Δh 内，每一特定源 IP 地址将被抽象成为聚合流。根据集合 $X=\{x_1, x_2,\cdots,x_n\}$，$n$ 个源 IP 地址被设置，其中每个地址都有一个 w 维的统计特征属性 $\boldsymbol{V}=\{v_1,v_2,\cdots,v_w\}$。因此本文首先使用源 IP 地址的特征属性构建一个属性矩阵 $\boldsymbol{K}\in\Omega^{n\times w}$，其中 $K(:,i)\in\Omega^{w}$ 表示第 i 个源地址的特征属性。由于 IP 地址间的相似性与流量之间存在一定关系，因此根据 IP 地址的相似性构建邻接矩阵 $\boldsymbol{F}\in\Omega^{n\times n}$，当邻接矩阵中的某个元素 $F(i,j)=1$，则表示 x_i 与 x_j 相似，否则不相似。为得到属性矩阵主模式，本文建立了一种从属性矩阵中提取代表性的行与列的模型，从而组成新的属性矩阵 $\dot{\boldsymbol{K}}\in\Omega^{n\times w}$。该矩阵代表属性矩阵 $\boldsymbol{K}$ 的主模式，可以通过 $\boldsymbol{K}$ 与 $\dot{\boldsymbol{K}}$ 之间的差异进行网络异常流量检测。可以看出，网络异常流量检测可以描述为从大量 IP 地址中找出一个异常 IP 地址集合，该集合中流量特征与其他绝大多数 IP 地址的流量特征具有显著差异性。最后根据神经网络方法完成特征提取与训练，最终输入分类器完成各种异常流量分类。

2.2　LSTM 网络

深度学习中的循环神经网络（RNN, recurrent neural network）在处理长序列数据时非常有效，但是序列数据样本过大且环境较复杂时，模型训练容易出现梯度消失与爆炸及神经网络过于依赖训练轮数的缺点。因此本文利用 LSTM 模型中的长短期记忆模块来解决上述问题。LSTM 本质上为一个门限 RNN，在 RNN 模型的隐藏层增加 3 个门与一个细胞状态更新参数，使其检测网络具有变化的自循环权重。当出现模型参数不变状况时，在不同时间点可得到动态改变的积分尺度，从而解决网络梯度消失及梯度爆炸问题。

如图 1 所示，LSTM 网络由 4 个独立结构组合形成，包含一个状态为细胞状态，同时由 3 个门（遗忘门、输入门及输出门）组成的主体结构。其中，负责清除前一细胞 c_{t-1} 到当前细胞 c_t 信息的为遗忘门；通过计算决策更新值 i_L 与候选细胞 c_t 的为输入门；更新 c_t 通过遗忘门、输入门及 ϖ_t 三者进行计算得到，将 ϖ_t 进行不间断传递，由此到达下一状态 c_{t+1}；h_t 为 LSTM 网络的输出值，由输出门计算并决策。

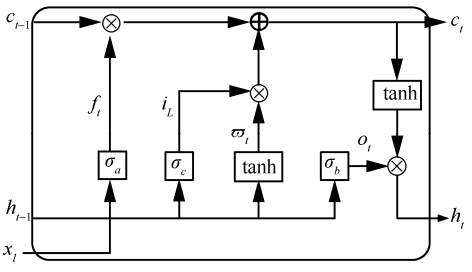


图 1　LSTM 网络结构

各个参数具体如下：f_t 为 LSTM 网络的更新遗忘门，其具体表达式为

$$f_t = \sigma_a(\boldsymbol{e}_f \cdot [h_{t-1}, x_l] + \boldsymbol{d}_f) \tag{1}$$

其中，$\cdot$ 为点乘，$\boldsymbol{e}_f$ 为遗忘门的权值矩阵，$\boldsymbol{d}_f$ 为偏置向量，$\sigma_a(x)$ 为遗忘门激活函数，用于计算 f_t。更新输入门输出 i_L 并计算候选状态 ϖ_t

$$i_L = \sigma_c(\boldsymbol{e}_i \cdot [h_{t-1}, x_l] + \boldsymbol{d}_i) \tag{2}$$

$$\varpi_t = \tanh(\boldsymbol{e}_c \cdot [h_{t-1}, x_l] + \boldsymbol{d}_c) \tag{3}$$

其中，$\boldsymbol{e}_i$ 为输入门权重矩阵，$\boldsymbol{d}_i$ 为偏置向量，$\sigma_c(x)$ 为输入门激活函数，用于计算 i_L。$\boldsymbol{e}_c$ 与 $\boldsymbol{d}_c$ 分别代表 tanh （激活函数）权重与系统偏置向量，通过这 2 个参数即可计算 ϖ_t，因此根据式(2)和式(3)，得到更新的 c_t 为

$$c_t = f_t c_{t-1} + i_L \varpi_t \tag{4}$$

更新输出门 h_t

$$o_t = \sigma_b(\boldsymbol{e}_o \cdot [h_{t-1}, x_l] + \boldsymbol{d}_o) \tag{5}$$

$$h_t = o_t \tanh(c_t) \tag{6}$$

其中，$\boldsymbol{e}_o$ 为输出门权重矩阵；$\boldsymbol{d}_o$ 为偏置向量；$\sigma_b(x)$ 为输出门激活函数，用于计算 o_t； tanh() 为 tanh 层的激活函数，用于计算输出 h_t。

在实际网络异常流量检测中，复杂的网络环境会导致LSTM在训练过程中出现训练集误差及新样本泛化能力弱的缺点。训练集偏差可通过扩大训练样本，选择适合的特征优化选择方法来解决，也可适当在LSTM网络中增加神经网络的训练深度来解决。但是深度选择不适时，方差依旧很大，因此综上分析后，本文将正则化处理与自适应学习率方法相结合来解决方差过大及目标函数优化问题。通过参数范数惩罚方法对 LSTM 模型进行正则化，直接对目标函数 G 优化。同时加入一个惩罚项 $\Omega(\lambda)$，$\Omega(\lambda)$ 可很大程度地改善方差过大问题，正则化后目标函数记为 $\dot{G}$，具体计算式为

$$\dot{G}(\lambda) = G(\lambda) + \beta\Omega(\lambda) \tag{7}$$

为了验证正则化过程及自适应学习率对 LSTM 网络优化的影响，图 2 分别给出了 LSTM 网络未正则化、L1 正则化、L2 正则化及本文方法优化后的方差对比。

从图 2(a)可知，随着时间及学习率的不断递增，LSTM 检测的方差也在增大。当学习率继续增大时，大部分数据的检测方差会减小，这是由于没有正则化从而导致数据的泛化能力差引起的方差过大。从图 2(b)可知，当 LSTM 检测模型在经过 L1 正则优化后，检测的方差减小；随着时间及学习率的不断递增，检测的方差逐渐减小。当学习率增加时，方差基本不变，但继续随着时间递增，方差又开始变大，这是由于 L1 正则化方法处理目标函数精度不高所导致的。从图 2(c)可知，当 LSTM 在经过 L2 正则优化后，检测数据的最高方差低于 L1 正则化优化，随着时间及学习率的不断递增，检测方差平均值基本不变，但低于 L1 正则化优化的方差。当学习率继续递增，检测方差又开始变大。从图 2(d)可知，当 LSTM 在经过 L2 正则化及自适应学习率优化后，检测数据的最高方差低于前述 2 种优化方法的方差，随着时间及学习率的不断递增，检测数据的方差平均值维持不变，但是低于前述 2 种优化方法的方差。当学习率继续递增，检测数据的方差变化不大，由此验证了 L2 正则化与自适应学习率方法对 LSTM 的优化性能。

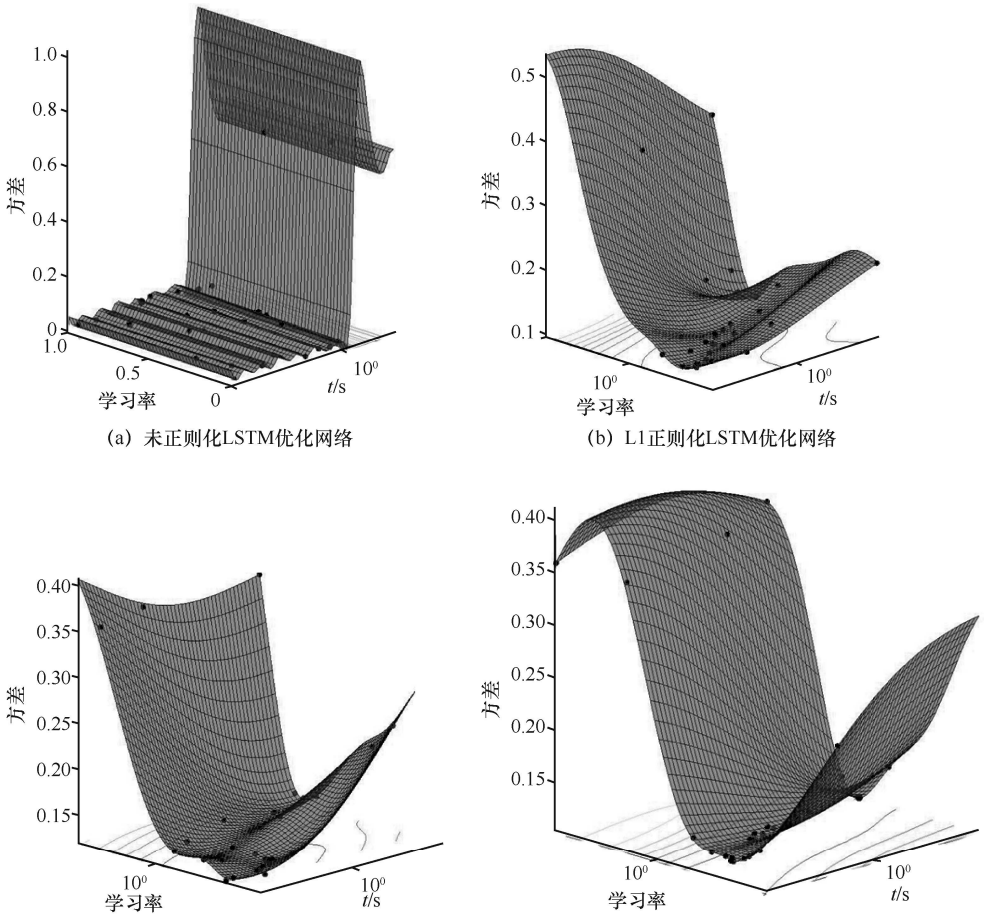
3 基于 LSTM 与改进残差网络优化的检测模型

在上述分析LSTM网络相关理论及流量检测原理之后，本文构建了基于 LSTM 与改进残差网络优化的异常流量检测模型，具体流程如图 3 所示。

模型检测的主要思路为：1) 网络数据集在预处理操作之后，得到输出 X； 2) 将此输出 X 作为三层堆叠 LSTM 网络的输入进行特征优化处理，从而得到输出 X_0，同时将此输出作为输入经过 Dropout 后得到 Y_0； 3) 得到的输出 Y_0 进入改进残差网络进行优化特征提取，且 Y_1 分别经过两条路径。第一条路径经过 Dense1 层，与输出权重相乘，利用批归一化（BN, batch normalization）与激活函数（ReLU）进一步优化，降低深层网络缺陷，之后进入 Dense2 与权重相乘得到 Y_2。第二条路径将改进残差神经网络的输出 Y_1 作为输入，最终两条路径输出相加，经过激活函数 ReLU 与 Softmax 层得到最终的输出结果 Y。因此，下述内容分别介绍各个优化策略的具体实现。

3.1 三层堆叠 LSTM 网络构建

本文首先构建了一种三层堆叠 LSTM（如图 4 所示），然后根据前述优化的 LSTM 网络，连续利用 3 个优化 LSTM 模型来设计一个三层堆叠结构，解决单层 LSTM 网络提取特征适应性弱的问题。



(a) 未正则化LSTM优化网络　(b) L1正则化LSTM优化网络

(c) L2正则化LSTM优化网络　(d) 正则化及自适应学习率优化LSTM网络

图 2　LSTM 优化效果对比

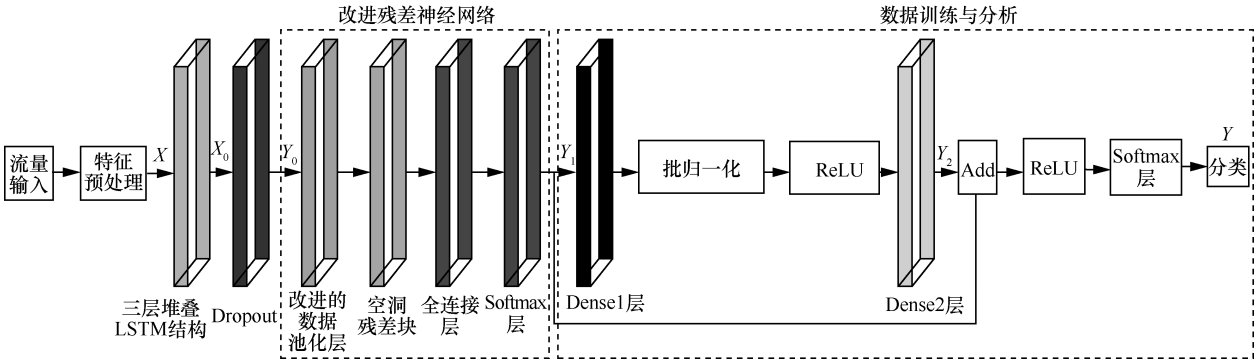


图 3　基于 LSTM 与改进残差网络优化的异常流量检测模型

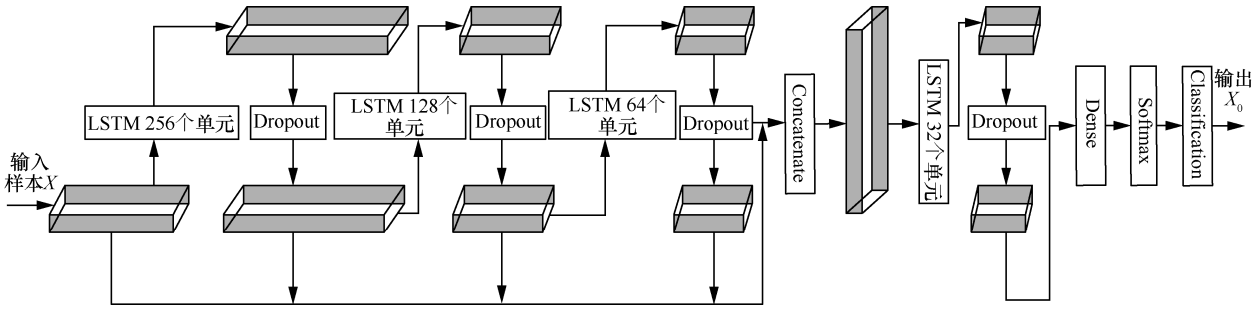


图 4　三层堆叠 LSTM 网络结构

图4输入为前期特征优化与预处理的网络数据集样本 X，分别利用三层堆叠 LSTM 从数据中提取不同深度特征，然后将处理的样本特征进行融合分析，继续传递到下一LSTM网络中训练分析，利用 Dropout 与 Dense 层来适当削弱梯度消失的弱点，将输出结果作为输入传递到 Softmax 层进行分类。具体过程为：经过预处理之后的样本特征 X 首先通过含有256个单元的第一层LSTM得到一阶特征；特征在传入堆叠的第二层 LSTM 前，先经过一层 Dropout 层对神经元进行优化，优化后的数据传入含有128个单元的第二层 LSTM 得到二阶特征；特征在传入堆叠的第三层 LSTM 前，再经过一层 Dropout 层对神经元进行优化，优化后的二阶特征数据传入一个含有 64 个单元的第三层 LSTM 得到三阶特征，再经过一层 Dropout 层对神经元进行优化。将这些不同深度的特征与数据合并（Concatenate）得到新的不同深度特征的数据。最后，将不同深度特征的数据输入一层含有 32 个单元的LSTM中，得到最后时刻隐藏层状态的数据。经过LSTM最后时刻隐藏层状态的数据传入一层全连接层得到32维特征，最终使用一层Softmax进行分类，提取等于目标类数量的一维5个元素的向量 X_0，经过 Dropout 后得到 Y_0。

3.2 基于 Inception 的改进数据池化层

LSTM 网络除梯度消失与过拟合弱点，也存在如何确定网络深度的问题。传统方式为人工实验确定，所需时间较长，因此本文采用残差神经网络（ResNet）来构造模型中的全连接神经网络。利用残差神经网络的结构特性选择一条有效路径来降低层数选择的难度，相比于 BN 方法，本文方法能够使神经网络具有更深的网络结构，同时也不会出现明显的过拟合以及梯度消失问题。残差网络结构如图5所示，三层堆叠网络输出 Y_0 与 $H(Y_0)$ 分别为残差输入与输出，此时输出为 $H(Y_0)=F(Y_0)+Y_0$。为了进一步优化 LSTM 网络，本文根据 $H(Y_0)=F(Y_0,[W_i])+Y_0$ 建立模型，形成一个恒等映射函数，用于分析从 LSTM 网络中的各个特征信息。其中 $[W_i]$ 为在处理输入样本数据时，传递信息到达第 i 个卷积层所设计的权重，卷积层的核大小都为 3×3。因此整个残差网络的具体优化原理为：通过预先设计好的快捷链接拟合 $F(Y_0)$ 这一残差映射函数，使输入数据 Y_0 与 $F(Y_0)$ 所获得的尺寸一致。在此基础上，利用模型的输入与输出继续重构样本进行不断学习分析，进而将模型训练时最底层的误差传递到上层，通过重构学习来减小误差，有效地解决残差网络训练时导致的梯度消失问题。在利用传统残差网络结构进行优化模型时，其分析样本所设计的数据池化层包含 7×7 原始池化层。

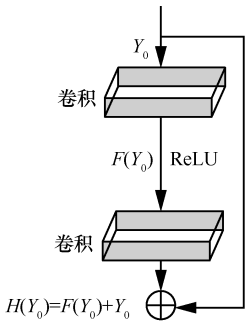


图5 残差网络结构

当利用图5模型训练数据时，由于输入样本过大，导致所需的硬件 GPU 必须不断更迭才能完成模型数据训练，因此时间复杂度与成本过高。同时网络流量数据集特征具有时变性强的特点，且有效样本有限，会导致搭建的模型提取特征信息不完善与信息丢失等问题。因此本文在分析上述情况后，根据 Inception 模块对原先数据池化层进行优化。同时，综合考虑到 Inception-V1～Inception-V4 的变化，在残差网络的改进过程中，为了能够保留更多目标的原始细节特征，本文继续对 Inception 网络结构进行适应性优化。将 5×5 卷积核替换成 2 个 3×3 卷积核以提升计算速度，同时减少了 Inception 网络中每层特征数量，保持特征总和与原残差神经网络的额外层特征总数相同。

为了能够反映引入的 Inception 结构对不同尺度卷积核的重要性，本文对结构中现存的2种卷积核（1×1 与 3×3）附以加权，比值为 1:2。在靠近输出端添加 1×1 Conv 来降低参数数量，加快计算速度。最后将整个结构辅以残差结构连接，综合提升算法的训练与测试性能。经过 Inception 改进后的数据池化层更加具有特征提取能力，进而在训练数据时，达到最好的训练效果，具体结构如图6所示。经过上述的 Inception 结构改进后，最终设计的改进数据池化层由原先的 7×7 卷积层变为 3 个 3×3 小卷积层，分别为1个 $3\times3\times8$ 与2个 $3\times3\times16$。同时为了使网络提取效率更高与提取特征信息能力更强，分别在3个卷积层后加入 BN 与 ReLU 激

活函数，之后利用一个 Concat 层将分析后的特征融合到一起再次通过一个1×1卷积层继续优化提取特征。由于增加了网络层数，因此可能会随着网络层数的加大，使模型在训练时出现严重的过拟合问题，导致网络深度不好，分类效果不佳。因此本文再次加入残差连接，利用一个最大池化层最终输出提取的特征信息。

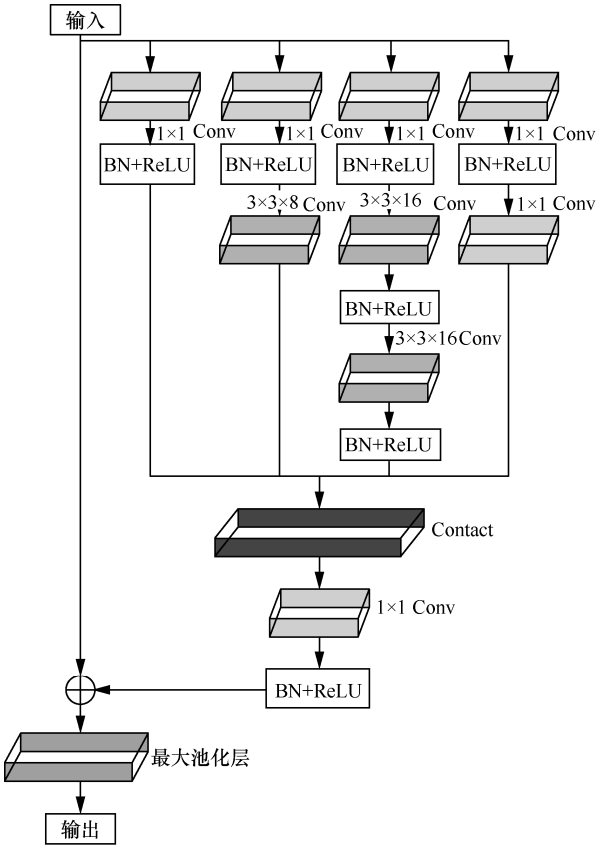


图 6　改进的数据池化层结构

3.3　改进的空洞残差块

传统残差神经网络中的残差块内部第二个卷积层仅对第一个卷积层的特征向量进行卷积运算，无法有效利用残差块输入向量与第二个卷积层之间的相关性，从而限制了残差块对特征的学习效率。因此本文将普通卷积替换为空洞卷积，空洞卷积在普通卷积中添加零填充，是一种扩展卷积核感受野的方法。其优势是在不改变特征分辨率前提下使感受野更大，感知信息范围更广，进而改善下采样带来的特征信息丢失问题。假设等效卷积核大小为 r^*，真实卷积核尺寸为 r，扩张率为 d，则等效卷积尺寸为

$$r^* = (d-1)(k-1)+k \tag{8}$$

当扩张率较小时，感受野与卷积核尺寸较小。当扩张率变大时，感受野也会增加，具体计算式为

$$H_{i+1} = H_i + (r^*-1)S \tag{9}$$

其中，H_i 表示当前层的感受野，H_{i+1} 表示下一层的感受野，S 表示从第一层到第 $i-1$ 层步长的乘积。因此可知，当空洞卷积级联时，其感受野面积呈指数级增长，获得的特征信息较好。此外，相较于普通卷积，空洞卷积不会因增大感受野而造成训练参数增加，使网络在获取更大范围内的特征信息时更加高效与便捷。然而空洞卷积虽可以增大感受野，但当空洞率较大时，读取的数据更稀疏，从而导致特征信息连续性被破坏，空洞残差块不能有效提取信号特征中的异常信息。

因此，本文分别分析验证了空洞率为 2 与 3 时的效果。当空洞率为 2 时，本文方法具有更高的检测准确率，所以将普通卷积核替换为空洞率为 2 的空洞卷积核，此时，获得的感受野等价于 2 个空洞卷积核的结果。对空洞卷积核中没有参数的位置填充 0。通过空洞卷积的操作后，残差块可以获取更多的信息，空洞残差块结构如图 7 所示，其中 λ 为空洞率。同理，为使网络提取效率更高与提取特征信息能力更强，分别在 2 个空洞卷积后加入 BN 与 ReLU 激活函数来优化模型。

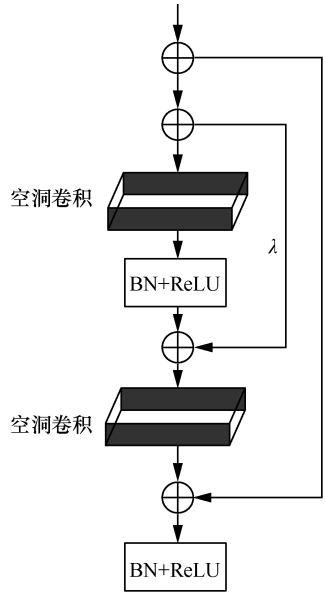


图 7　空洞残差块结构

3.4　最终的残差神经网络结构

经过上述的改进数据池化层与空洞残差块后，最终改进的残差网络结构如图 8 所示。为使优化后的训练模型具有更好的训练性能，需控制设计的网

络深度与残差块个数。本文根据实验分析，选择使用 5 个残差块来设计改进残差神经网络。这 5 个残差块按照端到端方式形成完整的残差神经网络用于训练模型优化。更进一步地，为使残差网络优化效果更好，将第一个与第三个残差块利用跳跃连接线，设计为空洞残差块；第二个、第四个及第五个残差块维持不变。5 个残差块中的卷积核大小均为 3×3，卷积核空洞率 $d=2$，跳跃连接线倍数 $j=0.2$，加入一个 Dropout 层防止过拟合情况，进一步提升训练性能，使分类效果更佳。同时，为了使本文在后述的对比结果更加清晰，将改进残差神经网络优化后的 LSTM 标记为 LSTM-ResNet 网络。

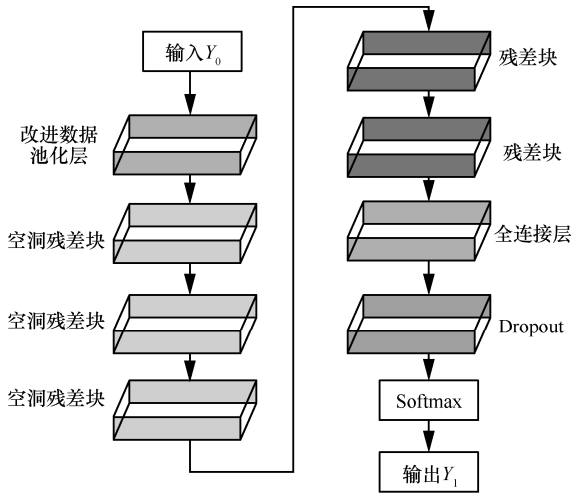


图 8　改进的残差网络结构

4　实验分析

4.1　实验数据

仿真实验数据选自 NSL-KDD 数据集与来自 Fsecrurify 的基于 Http 的开源 WAF 请求数据集，然后将这 2 个数据集分别选取子数据集后混合成为 NSL-KDD 混合数据集。NSL-KDD 数据集之所以能被广泛用于异常流量检测模型的验证，是因为 NSL-KDD 中的训练集与测试集中的记录数目合理，同时也添加了一个 difficulty level 属性。利用该属性可使分析网络中的各个连接时，与网络连接记录在 NSL-KDD 数据集中的比重形成反比例关系，这就更加突出了不同机器学习方法的分类差异，评估各种方法的分类性能与效率。本文选取的训练集为数据集前 7 周的网络连接记录，是二进制 TCPdump 压缩样本数据。数据集总数据量超过 4 GB，测试集包括最后 2 周的网络连接记录，网络数据按照 src 流向 dst。具体分类为：若为正常连接记录，则被标记为 Normal；否则被标记为一个明确类型的入侵（Attack）。流量数据处理后，共有 4 类攻击，具体如表 1 所示。

表 1　　NSL-KDD 数据集

标签类别	训练集/个	测试集/个
Normal	67 343	9 711
Probe	11 656	2 421
DoS	45 927	7 458
R2L	995	2 754
U2R	52	200
合计	125 973	22 544

4.2　数据预处理

训练集与测试集在应用到异常检测方法之前，需对数据进行预处理，具体流程如图 9 所示。本文首先对数据进行数值化操作，若不是数值型数据，将利用 One-Hot 编码进行重新分析，使其为数值型。One-Hot 编码主要采用 N 位状态寄存器对 N 个状态进行编码。当表示某一个状态时，只需将该状态的位置 1，其他位置 0 即可。本文利用 One-Hot 编码对网络连接中的每个属性进行处理时，将所有离散型数据进行量化，使特征更易分析处理，最后将得到的流量数据特征通过 LSTM 进行特征提取，得到最优的特征字段。选取的流量特征被分为 4 类：每个 tcp 连接的基本属性、连接中内容特征、时间窗口内的流量信息及连接窗口内的流量信息。其中“相同主机”仅统计 2 s 内与当前 connection 具有相同 dst 的连接，并计算与协议、服务相关的统计数据。本文选取的 NSL-KDD 混合数据集特征数为 41，其中只有 3 个特征为非数值型，其余均为数值型，同时类标签也为非数值型。4 个特征分别为 protocol_type、service、flag 及 class。其中 protocol_type 特征有 3 个属性：tcp、udp 及 icmp。首先需对非数值型数据进行标签编码，再对经过标签编码后的数据进行 One-Hot 编码，即可得到 3 个 3 bit 的向量[1,0,0]、[0,1,0]及[0,0,1]，分别代表 tcp、udp 及 icmp。最后利用同样方法对有 70 个属性的 service、11 个属性的 flag 及 5 个属性的 class 进行 One-Hot 处理，最终与其他 38 个特征拼接成预处理后的数据集。

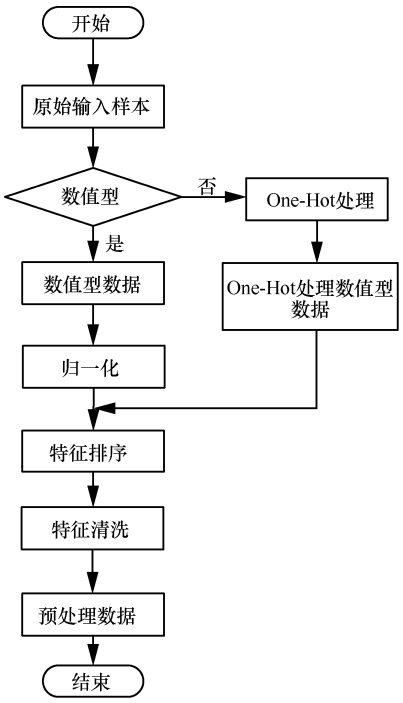


图 9　数据预处理过程

更进一步地，为降低特征值的差异性，需要对特征进行归一化，即将特征值映射到[0,1]区间。通过归一化和独热编码将原始数据全部转化为无量纲的数值数据。设有 n 条数据，第 i 条数据中统计型字段的第 j 个子属性值为 a_{ij}，则归一化为

$$I_{ij}=\frac{a_{ij}}{\max\limits_{1\leqslant i\leqslant n} a_{ij}}, j\in[1,m] \tag{10}$$

其中，I_{ij} 为归一化结果，m 为属性总数。归一化具体目的是使各特征数据分布尽可能接近，便于统一度量与分析。根据网络连接 IP 地址与 MAC 地址等信息，可得到多种比例型数据。其中一部分数据处于[0,1]，数值相近，因此保留；而大于 1 的数据，由于分布广泛导致难以把握规律，因此将其与上述数据同步分析。考虑此情况，本文利用 sigmoid 函数将其映射到(0,1)，然后再与原来属性值在[0,1]的数据进行联合分析，继续进行特征排序与特征清洗后，最终得到预处理数据。

4.3　实验评估指标

实验采用准确率（A, accuracy）、精确率（P, precision）、召回率（R, recall）、误报率（FPR, false positive rate）及调和平均值（F-measure）对检测方法进行验证，其参数计算方法与各部分定义参照文献[24]。其中，A 定义如式(11)所示，其值越高，分类器总体性能越好；P 定义如式(12)所示，其值越高，分类器的误报率越低；R 定义如式(13)所示，其值越高，分类误报率越低，分类效果越好；FPR（式(14)）反映了误报率，FPR 值越大，分类性能越差；F-measure（式(15)）反映 P 与 R 的调和平均值，其值越大，表明 P 与 R 值越接近 1，分类器对正类流量的检测性能越好。

$$A=\frac{\mathrm{TP+TN}}{\mathrm{TP+TN+FP+FN}} \tag{11}$$

$$P=\frac{\mathrm{TP}}{\mathrm{TP+FP}} \tag{12}$$

$$R=\frac{\mathrm{TP}}{\mathrm{TP+FN}} \tag{13}$$

$$\mathrm{FPR}=\frac{\mathrm{FP}}{\mathrm{TN+FP}} \tag{14}$$

$$\text{F-measure}=\frac{2PR}{P+R} \tag{15}$$

4.4　实验结果分析讨论

4.4.1　数据可视化分析讨论

为了验证数据预处理过程中每个步骤的必要性，实验采用 NSL-KDD 混合数据集中的一部分数据，分别针对未处理、独热编码、独热编码+归一化处理及独热编码+归一化处理+特征映射 4 种方式的处理做了数据散点可视化分析，主要对比数据集在不同预处理步骤下 Normal 流量与 Attack 流量的分布情况，4 种处理方式的可视化结果如图 10 所示。其中大部分中心区域的样本为 Normal 样本，大部分边缘区域的样本为 Attack 样本。

图 10(a)是 NSL-KDD 混合数据集未处理时的可视化结果，在未处理过的数据集中，原样本中的 Normal 与 Attack 流量数据是混乱离散的，无法对数据集进行有效的训练。图 10(b)是经过独热编码处理后的输出结果，相较于 10(a)，独热编码后的数据聚合程度较好，某些 Normal 与 Attack 流量数据点能够得到有效分离，但是大部分数据点还是混乱离散，这是由数据维度过高导致数据中存在一些无用数据点引起的。图 10(c)是在图 10(b)基础上加了归一化的可视化结果，该处理方式得到的 Normal 流量数据点中夹杂的 Attack 数据点较少，但是某些区域还是混乱离散，数据点没有得到很好的分离，这是由数据分布广泛导致难以把握其规律引起的。图 10(d)为本文最终提出

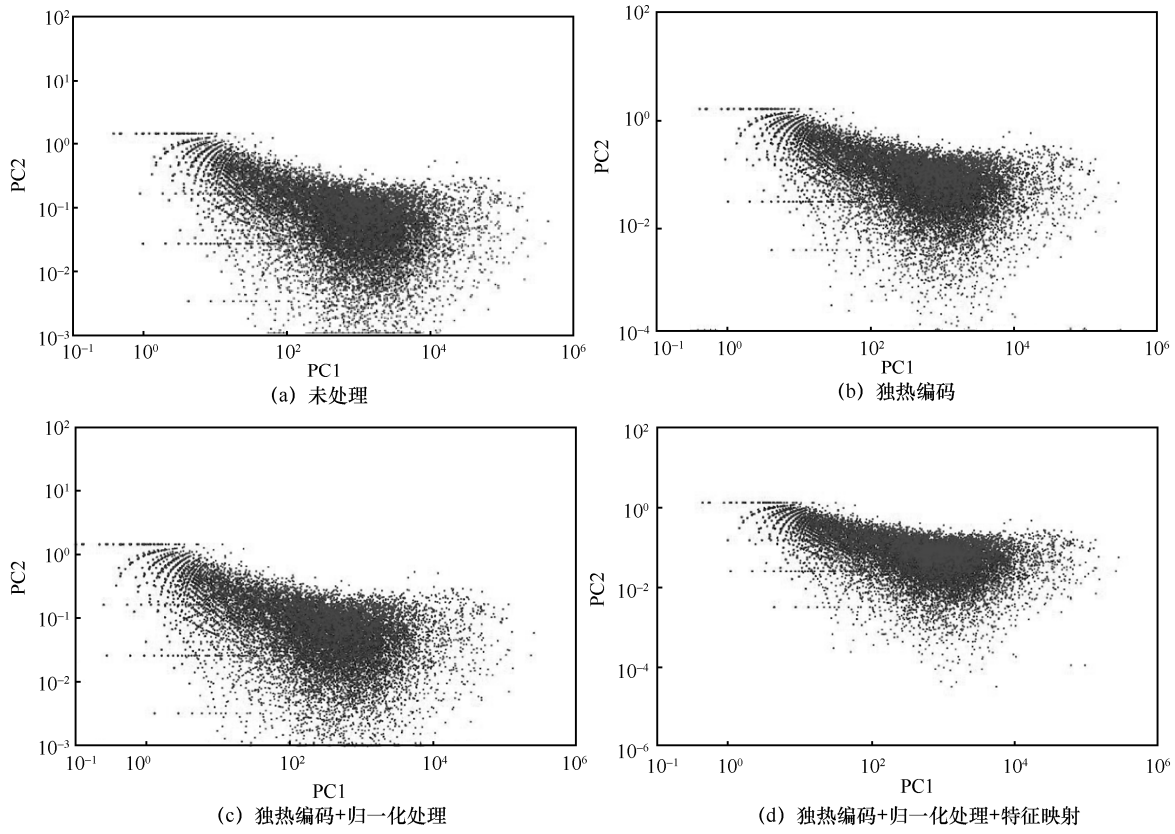


图 10　数据集预处理可视化分析

的处理方式，该方式能把原来数据样本中的大部分 Normal 与 Attack 流量数据点分离，且 Normal 与 Attack 流量数据点各自都变得比较收敛聚拢，特征较清晰，由此验证了本文预处理方法每个步骤的必要性。

4.4.2　训练数据集参数最优验证

为了验证时间与训练轮数对本文检测模型的影响，实验根据搭建好的检测模型，在时间与训练轮数不断递增的情况下，利用本文模型进行数据样本训练。每轮训练完成后使用训练集来做交叉验证，图 11 给出了 NSL-KDD 混合数据集中的一部分数据进行训练输出数据。从图 11 可以看到，随着时间的不断递增（训练轮数也在不断递增），目标数据个数始终在一个比较平稳的范围内进行输出目标数据，且训练集的输出个数与测试输出个数相差不大，错误率为 0.35%～0.5%。但随着时间继续递增，目标输出数据减少，且此时的训练输出与测试相差较大。这是由于此时训练数据已经出现了过拟合情况，导致输出减少，错误率也增大到了 1.5%～2%。当继续递增训练轮数进行训练时，虽然训练数据输出增大，但是测试数据输出减小，此时的数据分类效果一般。

因此，本文将训练轮数设置为 50 轮，此时的训练输出结果最好，错误率最低。同时为使训练效果更优，本文将训练集与测试集分成不同组合进行训练。经过实验验证后得知，当训练集 75%与测试集 25%组合时，分类获得的准确率较高，因此更加优化了训练模型的输出。

为了进一步验证每个设计步骤的有效性，建立了训练准确率与丢失率的对比实验，如图 12 所示。图 12 分别对比了不同方法下 Normal 与 Attack 样本的训练情况，所选方法分别为 RNN、LSTM、三层堆叠 LSTM、RNN+改进残差网络、LSTM+改进残差网络及本文方法。所有实验在 Windows 系统下使用 Tensorflow 作为后端的 Keras 深度学习框架下完成。实际应用中根据实验调整相应参数，Dropout 的 dropout rate 为 0.5。由图 12 可知，虽然 LSTM 网络基于 RNN，但是处理流量序列数据，LSTM 网络的性能优于 RNN；当 LSTM 网络被设计为三层堆叠 LSTM 之后，性能会进一步提高，训练性能优于前两者。当最终使用改进残差网络优化各个网络后，训练性能会进一步提高，且同时用改进残

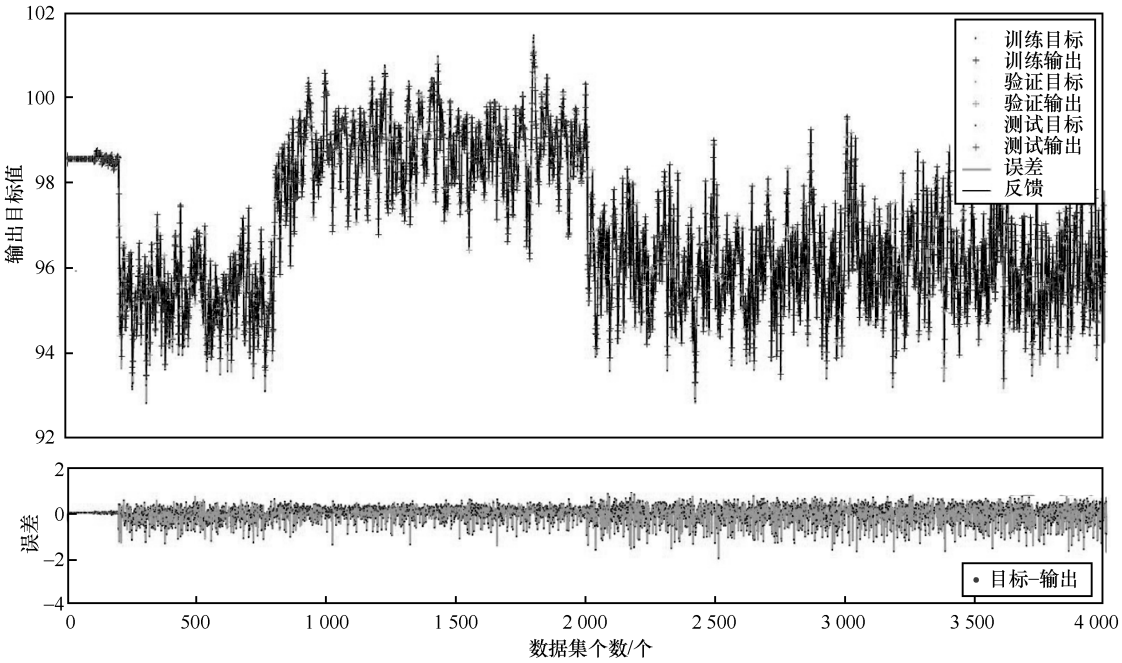
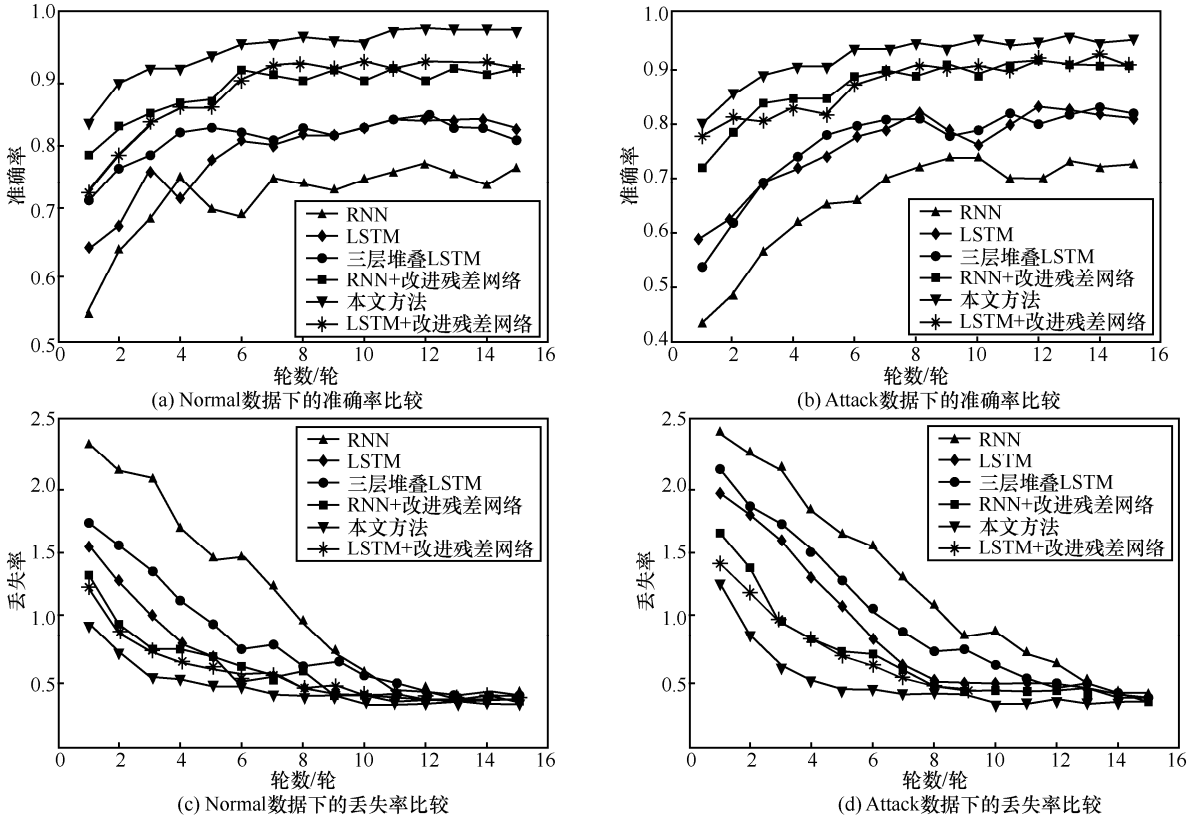


图 11　数据集训练



(a) Normal数据下的准确率比较　(b) Attack数据下的准确率比较

(c) Normal数据下的丢失率比较　(d) Attack数据下的丢失率比较

图 12　不同方法下的训练指标对比

差网络优化 RNN、LSTM 及本文的三层堆叠 LSTM 时，本文方法的训练性能指标最好，能够解决其单层 LSTM 适应性弱的问题。由此验证了本文设计步骤的有效性。

4.4.3　二分类性能对比分析

为了验证本文方法的检测性能，实验将数据集 4 类攻击合并为 Attack，正常流量记为 Normal，进行二分类对比实验。图 13 给出了数据集在未使用

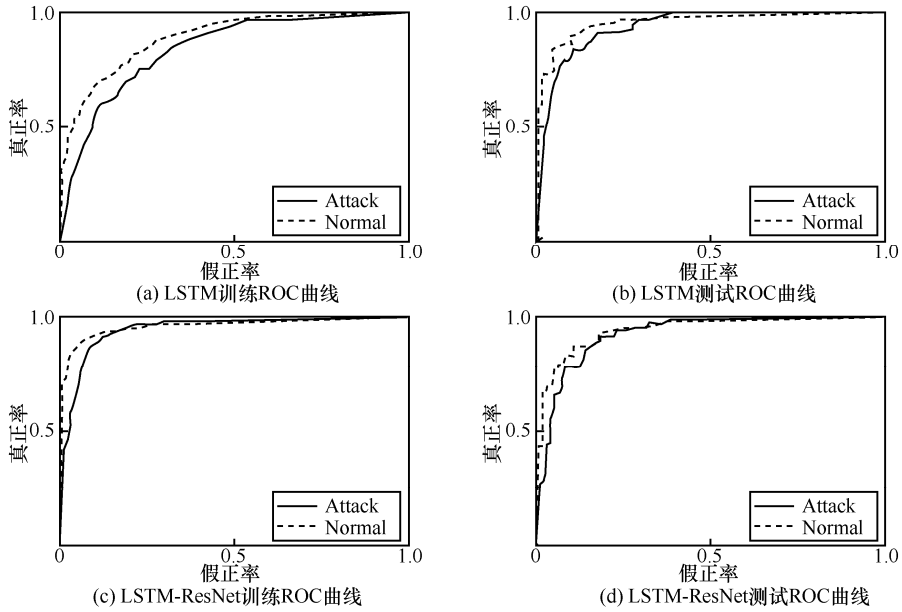


图 13　LSTM 与 LSTM-ResNet 的输出 ROC 曲线

残差网络优化与使用改进残差网络优化下的训练 ROC 曲线与测试 ROC 曲线。

AUC 曲线下的面积被称为 AUC 值，用来衡量预测准确性，其值越大，预测准确率越高。从图 13 可以看到，在仅使用 LSTM 网络训练数据时，无论是训练 ROC 曲线还是测试 ROC 曲线，AUC 值都比较低，其中训练 ROC 曲线的 AUC 值与测试 ROC 曲线的 AUC 值相差较大，导致在分割验证数据集时出现的误差较大，不能很好地提取与分析网络流量数据特征。由图 13 可知，在使用改进的残差网络优化 LSTM 网络后，训练数据时，无论是训练 ROC 曲线还是测试 ROC 曲线，取得的 AUC 值都比较高，其中训练 ROC 曲线的 AUC 值与测试 ROC 曲线的 AUC 值相差不大，同时无论检测 Normal 还是 Attack 数据，本文方法的 AUC 值均较优，突出了引入改进残差网络的必要性。

表 2 给出了各方法的准确率对比结果。7 种机器学习[20]方法分别为反向传播（BP, back propaga-tion）、逻辑回归（LR, logistic regression）、最近邻（KNN, k-nearest neighbor）、决策树（DT, decision tree）、随机森林（RF, random forest）、随机决策森林（RDF, random decision forest）、SVM。5 种深度神经网络分类方法分别为 2 层卷积神经网络（CNN2）、2 层深度神经网络（DNN2）、DBN、LSTM 与本文方法（LSTM-ResNet）。

由表 2 可知，本文提出的 LSTM-ResNet 检测模型取得的准确率最高，为 92.3%，而 LSTM 网络在未经 ResNet 结构优化之前，取得的检测准确率只有 87.6%，与传统 DNN2 模型取得的准确率相近。其余各检测方法取得的准确率都比较低，这是由数据集庞大且算法本身存在的过拟合问题导致的，本文通过改进残差神经网络进行优化 LSTM，因此分类效果最好。

表 2　　各方法的准确率对比结果

方法	Accuracy	方法	Accuracy
BP	78.2%	SVM	76.1%
LR	73.6%	CNN2	86.7%
KNN	82.4%	DNN2	87.2%
DT	83.6%	DBN	89.4%
RF	79.3%	LSTM	87.6%
RDF	83.5%	本文方法	92.3%

除准确率外，表 3 给出了精确率、召回率、误报率、F-measure 值及 AUC 值对二分类进行实验对比。由表 3 可知，对于测试精确率对比，LSTM-ResNet 检测模型的性能远优于 BP、LR、KNN、DT、RF、RDF 及 SVM。这是由于这几种检测方式为传统机器学习方法，对网络流量数据处理的效果整体一般，同时传统方法在训练过程中，学习深度不够会导致检测精确率较低；而 CNN2、DNN2、DBN 及 LSTM 这 4 种检测方法取得的检测精确率虽比前述几种方法高，但对 Normal 流量数

表 3　　**各方法的性能评价指标**

方法	类别	Precision	Recall	FPR	F-measure	AUC
BP	Normal	44.362%	76.342%	22.363%	56.116%	0.695
	Attack	78.325%	79.427%	19.635%	78.872%	0.687
LR	Normal	65.437%	73.635%	11.258%	69.294%	0.542
	Attack	79.654%	76.894%	25.634%	78.250%	0.593
KNN	Normal	75.724%	69.358%	9.365%	72.401%	0.753
	Attack	82.359%	72.417%	15.985%	77.069%	0.782
DT	Normal	73.612%	80.364%	4.952%	76.840%	0.714
	Attack	85.627%	82.348%	12.305%	83.956%	0.714
RF	Normal	58.952%	79.459%	6.654%	67.686%	0.728
	Attack	88.671%	81.872%	9.258%	85.136%	0.764
RDF	Normal	68.425%	82.328%	5.369%	74.736%	0.792
	Attack	85.654%	84.759%	8.258%	85.204%	0.792
SVM	Normal	52.872%	79.125%	18.675%	63.388%	0.826
	Attack	79.361%	82.388%	22.464%	80.846%	0.834
CNN2	Normal	70.821%	85.696%	3.254%	77.552%	0.865
	Attack	88.657%	87.397%	6.272%	88.022%	0.854
DNN2	Normal	74.258%	86.965%	2.359%	80.111%	0.756
	Attack	87.696%	88.564%	4.872%	88.128%	0.793
DBN	Normal	80.920%	89.781%	3.695%	85.121%	0.899
	Attack	89.471%	96.265%	4.259%	92.744%	0.863
LSTM	Normal	81.586%	89.820%	2.215%	85.505%	0.798
	Attack	89.214%	91.652%	3.269%	90.417%	0.779
本文方法	Normal	87.932%	93.587%	1.454%	90.671%	0.893
	Attack	93.625%	95.634%	2.651%	94.619%	0.883

据的检测精确率较低，因此泛化性能一般。本文的 LSTM-ResNet 检测模型取得的精确率平均为 90.772%，相比其他几种方法取得了一定程度的提高，性能较好。对于召回率的对比，基于传统机器学习的分类方法其召回率都较低，因此分类性能一般。而基于深度学习的几种检测方法召回率基本都在 85%以上，其中本文方法虽在检测 Attack 流量数据时取得的召回率略低于 DBN 方法，但在检测 Normal 数据时取得的召回率却高于 DBN 方法，平均召回率在所有方法中最高，为 94.610%。对于误报率的对比，每种方法在检测 Attack 流量时取得的误报率都高于 Normal 流量数据。这是由于测试集是随机选取的数据组合，其 Attack 类型数据所占比重较大，因此误报率较高。而相比其他检测方法，本文方法检测取得的误报率最低，平均为 2.053%，远低于 7 种传统的机器学习方法，同时也比 CNN2 低 2.71%，比 DNN2 低 1.563%，比 DBN 低 1.924%，比 LSTM 低 0.689%。因为本文使用改进残差神经网络对模型做了优化，所以总体误报率最低。对于 F-measure 的对比，几种传统的机器学习方法取得的值都较小，其中 LR 方法最差，平均仅为 67.494%，本文模型取得的 F-measure 值最好，在 4 种深度学习方法中平均调和值是最大的，平均为 92.645%，比 CNN2 与

DNN2 方法分别高 9.858%与 8.525%，比 DBN 与 LSTM 方法分别高 3.712%与 4.684%。由表 3 可知，无论检测 Normal 还是 Attack 流量数据，LSTM-ResNet 的 AUC 值均较优，但比 DBN 的正常数据检测差 0.006。在大多数情况下，LSTM 的 AUC 值在 4 种深度学习方法中是最低的，而且平均 AUC 值比 SVM 方法低 0.083，这充分说明，LSTM 在经过网络异常流量检测时存在网络过拟合与梯度消失的缺点，突出了改进残差神经网络优化的必要性。通过上述参数对比，验证了本文方法网络异常流量二分类下的良好性能。

4.4.4　多分类性能对比分析

更进一步地，为了验证本文在多分类任务中的分类性能，实验将 NSL-KDD 混合数据集中的 Normal、Probe、DoS、R2L 及 U2R 分为一类，首先分析 Normal 分别与 4 种攻击流量数据的 ROC 曲线对比，如图 14 所示。

由图 14(a)和图 14(b)可知，用 LSTM-ResNet 在对 Probe 流量数据进行分类时，ROC 曲线中的 AUC 值较好。因为 Normal 与 Probe 数据在样本数据集中占的比重较大，同时 2 种流量数据类型特征差异也较大，所以分类较准确；而采用 LSTM-ResNet 方法在对 DoS 数据进行分类时，ROC 曲线中的 AUC 值较 Probe 类型差了一点，但是相差不大，因为 DoS 与 Probe 数据的特征类型较为相近，所以有时会导致错误分类。

由图 14(c)和图 14(d)可知，采用 LSTM-ResNet 方法在对 R2L 与 U2R 流量数据进行分类时，ROC 曲线中的 AUC 均一般，这 2 种流量数据在样本数据集中占的比重较小，且 2 种流量数据类型特征相近，所以有时会导致错误分类。但是综合下来，LSTM-ResNet 方法对 4 种攻击类型流量取得的 AUC 值都比较好，错误分类的概率在可接受范围之内，由此验证了本文方法在 ROC 曲线上的优势。

表 4 给出了几种检测模型的准确率在多分类任务下的平均准确率对比。从表 4 可以看出，LSTM-ResNet 的准确率最高，为 89.3%。几种传统的机器学习方法中性能最差的是 LR 方法，准确率仅为 72.6%。这是由于对于多分类任务来说，LR 方法容易欠拟合，在数据特征有缺失或者特征空间很大时分类效果不好。

表 4　几种检测模型的准确率在多分类任务下的平均准确率对比

方法	Accuracy	方法	Accuracy
BP	73.4%	SVM	82.1%
LR	72.6%	CNN2	84.7%
KNN	79.4%	DNN2	85.5%
DT	81.2%	DBN	88.4%
RF	83.3%	LSTM	83.7%
RDF	84.5%	本文方法	89.3%

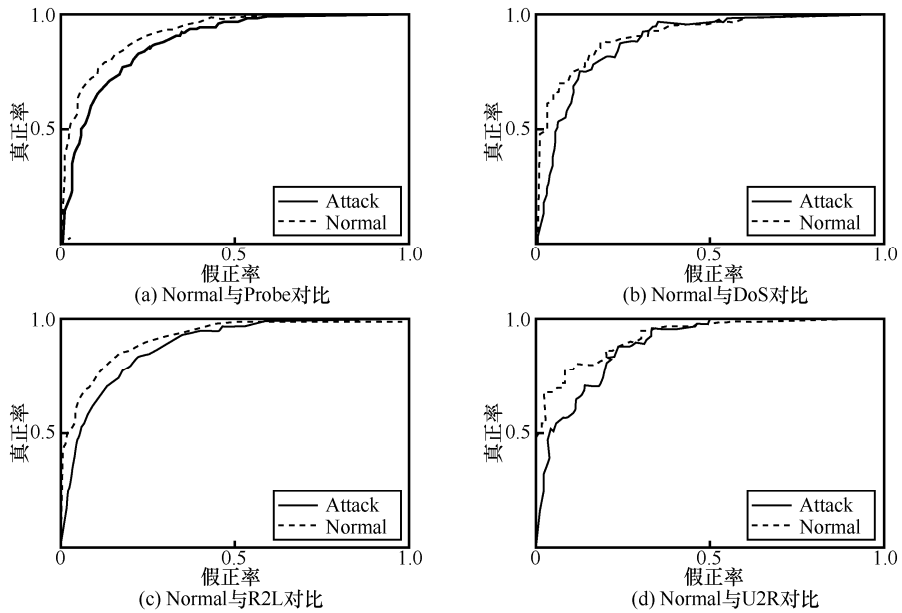


图 14　数据训练 ROC 曲线对比

为了进一步验证本文方法的多分类任务性能，图 15～图 19 给出了 5 种流量在精确率、召回率、误报率、F-measure 值下的实验对比结果。

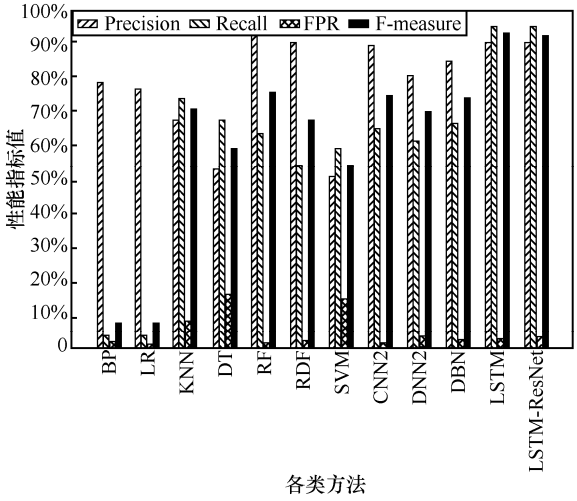


图 15　各算法在 Normal 上的性能评价指标

由图 15 可知，BP 和 LR 对于 Normal 流量类型数据综合检测性能最差，召回率仅为 4.357%和 4.723%，召回率与 F-measure 值等参数也较差。在多分类任务中，LSTM-ResNet 的分类性能在 4 种深度学习方法中是最好的，同时相比较于其他几种传统机器学习方法，本文方法取得的参数值均有所提升，但由于 LSTM-ResNet 将大量正常数据误判为攻击类型，会导致精确率与误报率略差。

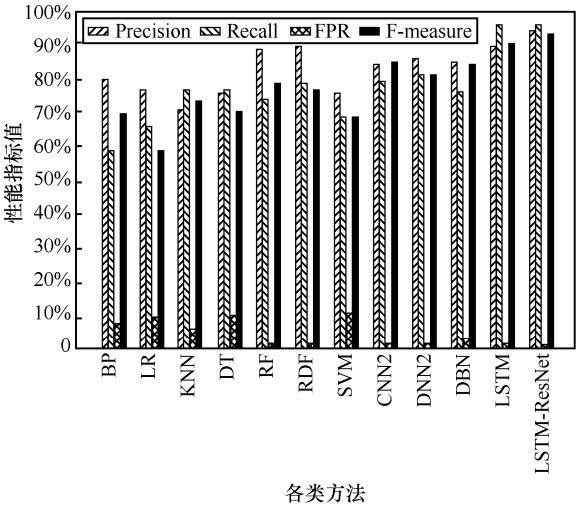


图 16　各算法在 Probe 上的性能评价指标

由图 16 可知，对于 Probe 类型攻击，BP、LR、KNN 及 DT 这 4 种方法的整体检测性能都比较差，其检测的精确率都低于 75%。这是由于传统的机器学习方法提取异常流量特征时易造成信息丢失；而 RF、RDF 及 SVM 这 3 种机器学习方法取得的性能参数值与深度学习方法比较相近，但还是比 LSTM-ResNet 检测模型的性能差。

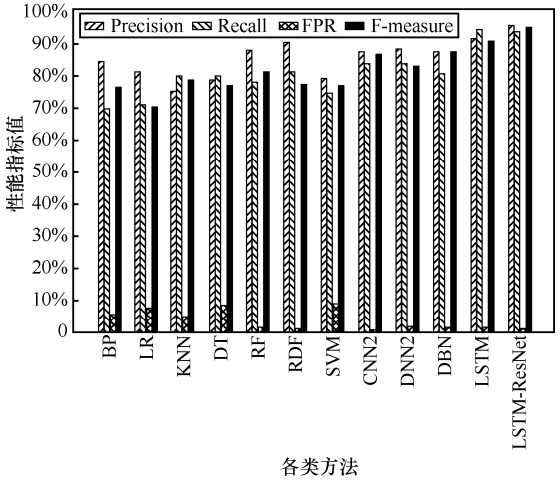


图 17　各算法在 DoS 上的性能评价指标

由图 17 可知，对于 DoS 类型攻击，LSTM-ResNet 的检测性能最优，LR 检测性能最差，KNN 和 DT 的召回率较差，仅为 24.273%和 56.109%，其余各方法的性能评价指标均较优。这是由于 DoS 攻击数目最多，各种方法在检测时能够识别的数据样本特征较丰富，因此检测的性能指标参数都比较良好。本文方法取得的精确率为 85.642%，召回率为 89.321%，误报率为 3.127%，F-measure 值为 87.443%，AUC 为 0.849，相较其他方法，本文方法最优。

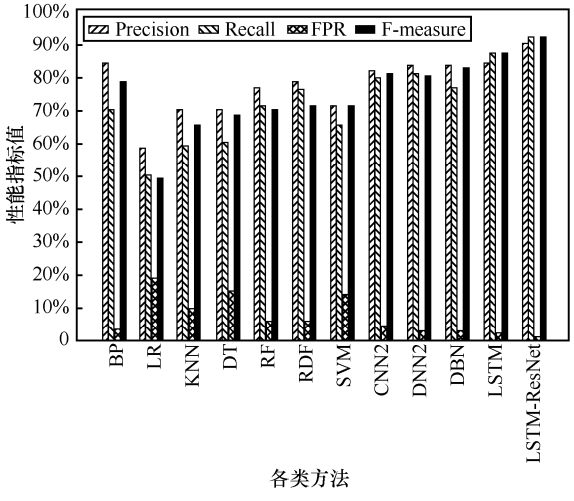


图 18　各算法在 R2L 上的性能评价指标

由图 18 可知，对于 R2L 攻击类型，在大多数情况下，BP、LR、KNN、DT 的性能相近，效果均较

差。而 RF、RDF、SVM 取得的精确率最低，基本上无法正确识别 R2L 攻击类型。这是由于 R2L 入侵是一种利用自身伪装，将其变为合法用户，使其特征与正常数据包类似，造成检测精度较低。4 种深度学习方法虽比前述几种方法有一定提高，但是误报率较高。本文方法性能较优，相比于 DBN 与 LSTM，精确率分别有一定的提高，这是由于本文采用的残差网络结构很好地学习了 R2L 的特征，将其正确分类。

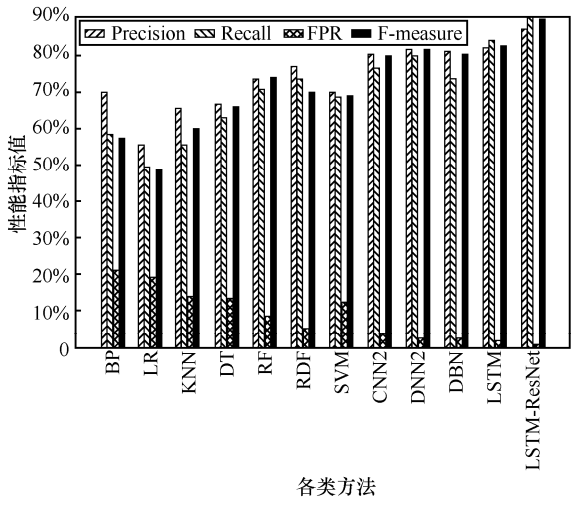


图 19　各算法在 U2R 上的性能评价指标

由图 19 可知，对于 U2R 类型攻击，RDF、SVM 及 4 种深度学习方法的性能相近，效果较优。BP、LR、KNN、DT 及 RF 的性能相近，效果较差，且本文方法相比于 RDF、SVM 及 4 种深度学习方法的分类性能均有所提高，AUC 分别提高了 0.247、0.387、0.087 2、0.067 1、0.033 4 及 0.029。

从图 15～图 19 的各项性能参数对比来看，本文检测模型无论是对 Normal 数据还是 4 种 Attack 数据，取得的性能指标值都比较好。虽在一些参数对比上低于 DBN 方法，但是总体对比，本文方法检测及其分类性能最好，能够对 NSL-KDD 混合数据集进行有效分类。由此验证了本文方法对于多分类任务的有效性与优越性。

4.4.5　总体分类性能对比

通过与 7 种传统机器学习方法、4 种深度神经网络方法进行分类对比后，本节选取了当前较流行的 4 种网络入侵检测方法[22-25]，通过整体性能分析来更进一步验证本文方法的性能。同时为了全面分析各种方法，在 NSL-KDD 混合数据集的测试集与验证集分别进行实验，具体结果如表 5 所示。由表 5 可知，在验证集中，所有方法都达到了不错的效果，本文方法达到了 94.09%的精确率，虽比文献[25]低，但是其他指标均优于文献[25]。同时与其他 3 种方法相比，本文方法的性能也均是最好的。但是在测试集中，所有方法性能都有所下降。因为测试集中存在训练集不曾出现过的攻击特征，NSL-KDD 与 Fsecrurify 的基于 Http 的开源 WAF 请求数据集混合后，数据分布仍存在一定差异。但是总体对比下来，本文方法取得的各项性能指标还是最优的，与其他方法相比还是具有很好的检测性能。

4.5　模型稳健性与运行时间分析

为了验证本文方法的稳健性与复杂环境适用性，实验分别在被测数据属性特征破坏率为 0.15、0.25 及 0.35 时，对比 4 种深度网络与文献[22-25]（由于篇幅原因，不再对比传统机器学习方法）在分类场景下的检测准确率。同时也对比了不同方法的训练与测试时间，每个实验做 20 次取平均值，结果如表 6 所示。

由表 6 可知，当输入的被测样本数据遭到噪声破坏时，基于 CNN2 方法的异常检测模型准确率最差，随破坏率增大，其准确率也会持续下降；基于 DNN2 与 DBN 的异常流量检测模型准确率下降幅度较前者小，但加大特征破坏率会进一步降低准确率，且降低幅度较大；文献[22-25]方法在遭到含噪声流量破坏时，性能较相近，但是性能相对较好的是文献[25]；而基于 LSTM 的异常检测模型受特征破坏率

表 5　　**与当前方法的总体分类性能对比**

方法	验证集					测试集				
	Precision	Recall	FPR	F-measure	AUC	Precision	Recall	FPR	F-measure	AUC
文献[22]	89.34%	87.58%	6.96%	88.45%	0.885	86.12%	85.48%	7.43%	85.80%	0.871
文献[23]	85.68%	84.63%	5.27%	85.15%	0.876	83.47%	84.02%	6.84%	83.74%	0.852
文献[24]	92.41%	91.27%	4.65%	91.84%	0.913	89.68%	88.79%	5.02%	89.23%	0.897
文献[25]	94.57%	93.81%	2.58%	94.18%	0.934	91.59%	92.36%	3.36%	91.98%	0.911
本文方法	94.09%	94.19%	2.86%	94.14%	0.946	92.73%	93.28%	2.41%	93.01%	0.928

表 6　　不同方法检测含噪流量的准确率及训练与测试时间

方法	Accuracy			时间/s	
	0.15	0.25	0.35	训练	测试
CNN2	81.971%	75.356%	71.642%	43.151	6.347
DNN2	83.594%	81.277%	79.706%	36.273	7.652
DBN	85.383%	82.922%	80.525%	39.482	8.526
LSTM	87.259%	85.934%	83.807%	27.496	4.258
文献[22]	85.417%	80.336%	79.634%	33.429	7.413
文献[23]	84.264%	81.228%	80.217%	35.472	7.231
文献[24]	87.678%	84.364%	81.369%	28.553	4.209
文献[25]	88.692%	85.697%	84.442%	30.664	5.214
本文方法	89.686%	89.326%	88.985%	29.147	4.384

影响较小，随着破坏率增大其准确率虽持续下降，但下降幅度不大，尤其是在经过改进残差网络优化后，其下降幅度继续减小。因此说明改进残差网络可有效降低噪声数据对提取流量特征准确性的影响，模型稳健性最强，其检测准确率最高，具有不错的泛化能力。LSTM-ResNet 网络的训练时间与测试时间虽比 LSTM 较长，但是相差不大，在可接受范围之内。本文方法比文献[22-25]方法的测试时间都短，说明本文方法能够实时进行网络异常流量入侵检测，且效率较高。

5　结束语

本文提出了一种基于 LSTM 与改进残差网络优化的异常流量检测方法，有效解决了现有方法存在的准确率、精确率及误报率等参数较差的问题。检测模型在公开的网络数据集 NSL-KDD 与开源 WAF 请求数据集进行实验验证，主要结论如下：三层堆叠的 LSTM 网络相比单层 LSTM 网络，特征选择性能更好，能够解决其单层网络的适应性弱问题；基于 Inception 结构与空洞残差块设计的改进残差神经网络能够解决 LSTM 的缺陷，使检测模型泛化能力更佳；无论是二分类实验还是多分类实验，本文方法在各种评价指标上取得的性能最好，与传统的机器学习方法、现有的深度神经网络及当前的网络入侵检测方法对比，本文方法都具有明显优势，稳健性较好，且测试时间最短。

参考文献：

[1] 张定华, 胡祎波, 曹国彦, 等. 面向工业网络通信安全的数据流特征分析[J]. 西北工业大学学报, 2020, 38(1): 199-208.
ZHANG D H, HU Y B, CAO G Y, et al. Dataflow feature analysis for industrial networks communication security[J]. Journal of Northwestern Polytechnical University, 2020, 38(1): 199-208.

[2] 李赛飞, 闫连山, 郭伟, 等. SD-SSDN: 基于 SDN 架构的高速铁路信号系统安全数据网的安全管控研究[J]. 铁道学报, 2018, 40(12): 81-92.
LI S F, YAN L S, GUO W, et al. SD-SSDN: software-defined signal safety data network for high-speed railway systems[J]. Journal of the China Railway Society, 2018, 40(12): 81-92.

[3] 丁建文, 宋甲英, 林思雨, 等. 基于 GPRS 分组交换网络的 CTCS-3 级列控系统车地安全数据传输的可行性[J]. 中国铁道科学, 2015, 36(3): 119-126.
DING J W, SONG J Y, LIN S Y, et al. Feasibility of train-ground safety data transmission for CTCS-3 train control system based on GPRS packet switching network[J]. China Railway Science, 2015, 36(3): 119-126.

[4] 李赛飞, 闫连山, 李洪赭, 等. 铁路通信网络安全的分析测试与可信防御研究[J]. 西南交通大学学报, 2018, 53(6): 1130-1136,1149.
LI S F, YAN L S, LI H Z, et al. Analysis and testing of network security for China railway communication networks and proposed architecture based on trusted computing[J]. Journal of Southwest Jiaotong University, 2018, 53(6): 1130-1136,1149.

[5] ZHANG X, ZHAO J B, LECUN Y. Character-level convolutional networks for text classification[C]//Advances in Neural Information Processing Systems. Massachusetts: MIT Press, 2015: 649-657.

[6] LU X H, ZHENG B, VELIVELLI A, et al. Enhancing text categorization with semantic-enriched representation and training data augmentation[J]. Journal of the American Medical Informatics Association, 2006, 13(5): 526-535.

[7] PARK S, KIM M, LEE S. Anomaly detection for HTTP using convolutional autoencoders[J]. IEEE Access, 2018, 6: 70884-70901.

[8] YU Y Q, LIU G, N YAN H B, et al. Attention-based Bi-LSTM model for anomalous HTTP traffic detection[C]//2018 15th International Conference on Service Systems and Service Management. Piscataway: IEEE Press, 2018: 1-6.

[9] YANG W C, ZUO W, CUI B J. Detecting malicious URLs via a keyword-based convolutional gated-recurrent-unit neural network[J]. IEEE Access, 2019, 7: 29891-29900.

[10] CHORAŚ M, KOZIK R. Machine learning techniques applied to detect cyber attacks on web applications[J]. Logic Journal of the IGPL, 2015, 23(1): 45-56.

[11] KRUEGEL C, VIGNA G. Anomaly detection of Web-based attacks[C]//Proceedings of the 10th ACM conference on Computer and Communications security. New York: ACM Press, 2003: 251-261.

[12] CORONA I, TRONCI R, GIACINTO G. SuStorID: a multiple classifier system for the protection of Web services[C]//Proceedings of the 21st International Conference on Pattern Recognition. Piscataway: IEEE Press, 2012: 2375-2378.

[13] RINGBERG H, SOULE A, REXFORD J, et al. Sensitivity of PCA for traffic anomaly detection[C]//ACM SIGMETRICS Performance Evaluation Review. New York: ACM Press, 2007, 35(1): 109-120.

[14] AL-OBEIDAT F, EL-ALFY E S M. Hybrid multicriteria fuzzy classification of network traffic patterns, anomalies, and protocols[J]. Personal and Ubiquitous Computing, 2019, 23(5/6): 777-791.

[15] ERFANI S M, RAJASEGARAR S, KARUNASEKERA S, et al. High-dimensional and large-scale anomaly detection using a linear one-class SVM with deep learning[J]. Pattern Recognition, 2016, 58: 121-134.

[16] DU M, LI F F, ZHENG G N, et al. DeepLog: anomaly detection and diagnosis from system logs through deep learning[C]//Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security. New York: ACM Press, 2017: 1285-1298.

[17] ZHANG M, LU S B, XU B Y. An anomaly detection method based on multi-models to detect web attacks[C]//2017 10th International Symposium on Computational Intelligence and Design. Piscataway: IEEE Press, 2017: 404-409.

[18] 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.
GAO N, GAO L, HE Y Y, et al. A lightweight intrusion detection model based on autoencoder network with feature reduction[J]. Acta Electronica Sinica, 2017, 45(3): 730-739.

[19] ALRAWASHDEH K, PURDY C. Toward an online anomaly intrusion detection system based on deep learning[C]//2016 15th IEEE International Conference on Machine Learning and Applications. Piscataway: IEEE Press, 2016: 195-200.

[20] 李艳霞, 柴毅, 胡友强, 等. 不平衡数据分类方法综述[J]. 控制与决策, 2019, 34(4): 673-688.
LI Y X, CHAI Y, HU Y Q, et al. Review of imbalanced data classification methods[J]. Control and Decision, 2019, 34(4): 673-688.

[21] 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.
CHEN J T, XIANG Y. Survey of unstable gradients in deep neural network training[J]. Journal of Software, 2018, 29(7): 2071-2091.

[22] DAS T K, ADEPU S, ZHOU J Y. Anomaly detection in industrial control systems using logical analysis of data[J]. Computers & Security, 2020, 96: 101935.

[23] 宋勇, 侯冰楠, 蔡志平. 基于深度学习特征提取的网络入侵检测方法[J]. 华中科技大学学报(自然科学版), 2021, 49(2): 115-120.
SONG Y, HOU B N, CAI Z P. Network intrusion detection method based on deep learning feature extraction[J]. Journal of Huazhong University of Science and Technology (Natural Science Edition), 2021, 49(2): 115-120.

[24] 张兴兰, 尹晟霖. 可变融合的随机注意力胶囊网络入侵检测模型[J]. 通信学报, 2020, 41(11): 160-168.
ZHANG X L, YIN S L. Intrusion detection model of random attention capsule network based on variable fusion[J]. Journal on Communications, 2020, 41(11): 160-168.

[25] YANG J, LIANG G, LI B B, et al. A deep-learning- and reinforcement-learning-based system for encrypted network malicious traffic detection[J]. Electronics Letters, 2021, 57(9): 363-365.

[作者简介]

麻文刚（1993− ），男，甘肃天水人，西南交通大学博士生，主要研究方向为通信系统、信息安全等。

张亚东（1983− ），男，河南商丘人，博士，西南交通大学讲师、硕士生导师，主要研究方向为系统可靠性与安全性理论、系统仿真测试等。

郭进（1960− ），男，四川成都人，博士，西南交通大学教授、博士生导师，主要研究方向为系统安全理论、安全苛求系统设计与验证等。