

基于深度循环神经网络和改进 SMOTE 算法的组合式入侵检测模型

燕昺昊, 韩国栋

(国家数字交换系统工程技术研究中心, 河南 郑州 450002)

摘要: 已有入侵检测模型普遍只针对网络入侵行为的静态特征进行分析检测, 造成检测率低及误报率高等缺陷, 且无法有效应用低频攻击。为此提出一种新的基于深度循环神经网络 (DRNN) 和区域自适应合成过采样算法 (RA-SMOTE) 的组合式入侵检测模型 (DRRS)。首先, RA-SMOTE 对数据集中低频攻击样本进行自适应区域划分, 实现差别样本增量, 从数据层面提升低频攻击样本数量; 其次, 利用 DRNN 特有的层间反馈单元, 完成多阶段分类特征的时序积累学习, 同时多隐层网络结构实现对原始数据分布的最优非线性拟合; 最后, 使用训练好的 DRRS 模型完成入侵检测。实验结果表明, 相比已有入侵检测模型, DRRS 在改善整体检测效果的同时显著提高了低频攻击检测率, 且对未知新型攻击具有一定检出率, 适用于实际网络环境。

关键词: 网络安全; 深度学习; 入侵检测; 循环神经网络; 过采样算法

中图分类号: TP393.08

文献标识码: A

doi: 10.11959/j.issn.2096-109x.2018056

Combinatorial intrusion detection model based on deep recurrent neural network and improved SMOTE algorithm

YAN Binghao, HAN Guodong

National Digital Switching System Engineering and Technological Research Center, Zhengzhou 450002, China

Abstract: Existing intrusion detection models generally only analyze the static characteristics of network intrusion actions, resulting in low detection rate and high false positive rate, and cannot effectively detect low-frequency attacks. Therefore, a novel combinatorial intrusion detection model (DRRS) based on deep recurrent neural network (DRNN) and region adaptive synthetic minority oversampling technique algorithm (RA-SMOTE) was proposed. Firstly, RA-SMOTE divided the low frequency attack samples into different regions adaptively and improved the number of low-frequency attack samples with different methods from the data level. Secondly, the multi-stage classification features were learned by using the level feedback units in DRNN, at the same time, the multi-layer network structure achieved the optimal non-linear fitting of the original data distribution. Finally, the intrusion detection was completed by trained DRRS. The empirical results show that compared with the traditional intrusion detection models, DRRS significantly improves the detection rate of low-frequency attacks and overall detection efficiency. Besides, DRRS has a certain detection rate for unknown new attacks. So DRRS model is effective and suitable for the actual network environment.

Key words: network security, deep learning, intrusion detection, DRNN, oversampling algorithm

收稿日期: 2018-06-13; 修回日期: 2018-07-05

通信作者: 燕昺昊, ndscybh@qq.com

基金项目: 国家科技重大专项基金资助项目 (No.2016ZX01012101); 国家自然科学基金资助项目 (No.61572520); 国家自然科学基金创新群体资助项目 (No.61521003)

Foundation Items: The National Science Technology Major Project of China (No.2016ZX01012101), The National Natural Science Foundation Project of China (No.61572520), The National Natural Science Foundation Innovation Group Project of China (No.61521003)

1 引言

随着互联网应用不断向移动端扩展,网络智能技术已服务于各行各业。但同时针对网络发动的攻击愈加频繁,据国家互联网应急中心^[1]统计,仅2017年7月,我国境内便发生3 000余起针对银行、民航以及通信等重点服务领域的漏洞攻击事件,网络安全的重要性愈发凸显。入侵检测技术作为网络主动防御技术之一,其性能优劣直接决定受保护设备及领域的安全程度。

入侵检测技术根据其检测方式的差异,可分为误用检测和异常检测。误用检测通过维护特定的规则库,基于特征匹配实现对已知攻击行为的检测,检测速度快且准确度高,但其检测性能依赖于规则库大小及更新频率,且无法检测未知攻击。异常检测则通过对网络数据记录进行分析,建立系统正常行为模型,将任何偏离模型正常值的异常行为标记为攻击行为。异常检测优点在于可检测出未知攻击,但存在误报率较高等缺陷。随着人工智能技术的兴起,基于机器学习的入侵检测技术,如神经网络^[2]、支持向量机(SVM, support vector machine)^[3]、仿生算法^[4]等,成为领域研究热点。机器学习模型将检测问题转化为分类问题,通过大量原始数据样本训练,自适应学习正常行为与异常行为之间的差异,避免复杂的人工干预及领域专家知识,有效增强了入侵处理实时性。但传统的机器学习模型仍存在以下问题。

1) 低频攻击检测困难。机器学习模型通常以最大化整体分类精确度为目标,低频攻击由于实际样本数量稀少,分类特征难以捕捉,故对模型无法形成有效训练,导致低频攻击检测难度大且检测率低。

2) 模型拟合能力差。传统机器学习模型结构简单,特征提取及学习能力有限,仅对小批量数据具有良好的拟合效果,当面临大规模数据集时无法对数据分布形成有效的非线性映射。

3) 动态特性不足。已有机器学习模型普遍忽视入侵事件之间的关联性,即假定入侵行为之间具有独立不相关性并据此进行建模,造成模型缺乏必要的动态特性,检测率偏低,同时对于未知

攻击无法形成有效检测^[5]。

针对存在问题,文献[6]将合成过采样算法(SMOTE, synthetic minority oversampling technique algorithm)与最近邻聚类算法(CCNN, cluster center and nearest neighbor)结合进行入侵检测,同时引入特征选择机制,从数据层面提高了低频样本检测率。文献[7]则从算法层面出发,提出了一种新的基于支持向量机和遗传规划(GP, genetic programming)的组合式GPSVM分类器,同样对低频样本取得不错的检测效果。文献[8-9]分别利用基于深度学习理论的自编码网络(AEN, auto encoder network)和深度神经网络(DNN, deep neural network)中多隐层结构,有效实现了输入特征提取与模型非线性映射,提高检测速率的同时对未知攻击具有一定检出效果。但上述解决方法并未考虑入侵行为之间存在时序维度上的关联性。文献[10]采用基于样本序列分析的层次化隐马尔可夫模型(LHMM, layered hidden markov model),根据输入不同,该模型更新状态概率分布值并计算转移概率,实现时序状态预测并最终完成入侵检测。但模型更新过程需要对每一项输入计算其多个概率分布值,且需要特定算法来实现特征维度缩减,复杂度较高。

为解决上述问题,本文提出一种基于深度循环神经网络(DRNN, deep recurrent neural network)和区域自适应合成过采样算法(RA-SMOTE, region adaptive SMOTE)的组合式入侵检测模型DRRS。首先,利用RA-SMOTE算法对已有非平衡数据集中低频样本实现增量处理与时序注入。其次,在新数据集上对DRNN网络进行时序权重训练,直到网络稳定且误差达到预定要求。最后,使用训练好的DRNN完成入侵检测,同时将检测出的低频攻击样本返回低频样本处理单元进行再处理,输入网络进行权值再调优,增强模型稳定性与稳健性。

与已有方法相比,本文所提模型优势如下。首先,RA-SMOTE算法解决了低频攻击检测率较低的问题,且其特有的自适应性可以更高效地合成新样本,加速后续模型训练过程;其次,DRNN特有的时序循环结构可充分挖掘并利用样本之间的时序相关性实现入侵检测,同时多隐层网络结

构提升模型数据刻画能力,提升模型检测性能。仿真结果验证了模型有效性,适用于实际网络。

2 DRRS 组合式入侵检测模型

DRRS 入侵检测模型主要由低频样本处理单元和时序学习单元组成,其中,低频样本处理单元基于改进后的区域自适应合成过采样算法(RA-SMOTE)实现,完成对低频样本的过采样增量处理;时序学习单元基于深度循环神经网络(DRNN)实现,完成对输入样本的时序积累建模与异常检测。

2.1 低频样本处理算法

尽管当前网络攻击常有发生且出现日益增长的趋势,但相比网络中存在的海量正常数据,其数量依然微乎其微。因此,由于攻击数据流或行为样本与正常数据流量相比处于低频范畴,导致其分类特征难以捕捉与记录,从而进一步造成机器学习模型无法对攻击样本建立准确的检测模型并实现有效的训练;另外,大多数机器学习模型及算法以最大化样本整体分类准确率为目的,致使低频攻击样本在训练及分类过程中被忽视,进而导致训练完成的模型具有明显的分类偏向性,在实际应用中产生错误检测结果。本文基于已有的非平衡数据处理方式进行改进,提出新的区域自适应合成过采样算法,主要分为低频样本增量与新样本时序关系处理 2 个步骤。

2.1.1 低频样本增量

根据处理层面不同,低频样本处理方法可分为算法层面和数据层面。算法层面方法主要通过为分类决策函数添加调控系数或惩罚因子,使低频样本具有更高分类权值,提高检测率。相比算法层面,数据层面处理方法复杂度更低,应用更为广泛。数据层面方法主要包括过采样方式和减采样方式,但传统的过采样方式只是对已有数据样本的简单重复复制,造成特征多样性不足;而减采样通过对高频样本抽样,易损失样本内在特性。

针对上述问题,文献[11]提出 SMOTE 算法,通过在已知低频样本与其最近邻之间随机线性插值,实现样本增量的同时有效解决特征不足的问题。但 SMOTE 算法并未考虑样本分布边缘化及

离群点处理等问题,同时对所有低频样本采用统一增量方法,造成部分计算资源浪费且分类效果较差。为此,本文提出了改进的 RA-SMOTE 算法。首先,设置低频样本最近邻半径,半径大小用最近邻样本个数表示。其次,根据半径内所包含的高频样本数量的差异,将低频攻击样本划分为不同区域集合,包括离群点域(IPR, independence point region)、安全点域(SPR, safety point region)和危险点域(DPR, danger point region)。区域不同表示分类难度不同,因此对不同区域内低频样本采用不同的样本增量方式,在保证分类准确度的前提下,减少了样本增加数量,从而降低后续 DRNN 训练量及训练时间。RA-SMOTE 算法样本增量过程伪代码如下所示。

算法 RA-SMOTE 算法样本增量过程

输入算法参数 最近邻半径 r

采样率 S

输入样本集合 原始样本集合 X

低频样本集合 X_{mi} // $X_{mi} \in X$

高频样本集合 X_{ma} // $X_{ma} \in X$

输出样本集合 新样本集合 X_{new}

1) For each sample $x \in X_{mi}$ do

2) $k = |X_r \cap X_{ma}|$ // 计算最近邻半径内高频样本数量, X_r 为最近邻样本集合

3) If $k = r$ then

4) $x \in IPR$ // 将 x 划分为离群点域

5) $x_{new} = x$ // 生成 S 个新样本

6) Else If $0 \leq k \leq \frac{r}{2}$ then

7) $x \in SPR$ // 将 x 划分为安全点域

8) do nothing for x // 不生成新样本

9) Else $\frac{r}{2} \leq k \leq r$ then

10) $x \in DPR$ // 将 x 划分为危险点域

11) $x_{mean} = \frac{1}{r-k} \sum_{i=1}^{r-k} x_i$ // 计

算 r 内低频样本均值

12) $\mathbf{x}_{\text{new}} = \mathbf{x} + u_{(0,1)}(\mathbf{x}_{\text{mean}} - \mathbf{x})$
 // 生成 S 个新样本, $u_{(0,1)}$ 为 $[0,1]$ 上随机数
 13) End If
 14) End If
 15) End For
 16) Return $\mathbf{x}_{\text{new}}$ // 返回新样本集合

2.1.2 新样本时序关系处理

DRNN 中训练样本需根据时序特征串行输入,使网络可以学习并记忆样本间时序连接关系。而 RA-SMOTE 算法生成新样本为随机模拟样本,样本间不存在特定时序关系。因此,为使新样本适用于 DRNN,需要对新样本进行时序添加。首先进行如下定义。

定义 1 如果新样本 $\mathbf{x}_{\text{new}}$ 由原始样本 $\mathbf{x}$ 生成,称 $\mathbf{x}$ 为 $\mathbf{x}_{\text{new}}$ 的元样本。

因此时序添加过程表示为

$$t_{\mathbf{x}_{\text{new}}} = t_{\mathbf{x}} + 1 \quad (1)$$

其中, $t_{\mathbf{x}}$ 表示样本 $\mathbf{x}$ 在原始样本集中时序位置, $t_{\mathbf{x}_{\text{new}}}$ 表示新样本时序位置。由式(1)可知,在新数据集中,新样本时序位置位于其生成样本之后,即输入数据时,新样本在其元样本之后输入。

2.2 深度时序网络模型

2.2.1 深度神经网络

深度神经网络是指区别于传统浅层神经网络

三层架构的、具有多隐层的人工神经网络(ANN, artificial neural network)。对于人工神经网络的研究最早可追溯至 20 世纪 40 年代,但由于硬件水平受限且无法克服多层结构带来的大数据量与高维参数训练等问题,ANN 一直处于浅层阶段。2006 年, Hinton 等^[12]在《Science》上发表著作,重新掀起了 DNN 研究浪潮。文章指出, DNN 特有的多隐层架构具有出众的非线性学习能力与特征表达能力,可有效提高神经网络分类结果。同时 Hinton 等提出分层预训练与微调机制,很好地解决了 DNN 训练缺陷。

DNN 通常包括输入层、输出层及三层以上隐层,如图 1 所示。DNN 中各节点表示网络激活函数,连接权值表示信息传递激励强度。各层神经元之间为全连接或稀疏连接状态,并在学习过程中不断调整连接权值使网络收敛,存储数据分布特征。输入层完成预处理数据的读取,多隐层架构实现对输入向量的逐层特征提取与样本学习,输出层则根据需求完成回归或分类任务。

2.2.2 深度循环神经网络

虽然 DNN 在分类与特征提取方面突破浅层网络限制,拥有极强的非线性模拟能力,但主流 DNN 架构并未考虑分类样本之间存在的时序连接关系,造成部分关联信息损失。近年来,结合深度学习理论的循环神经网络

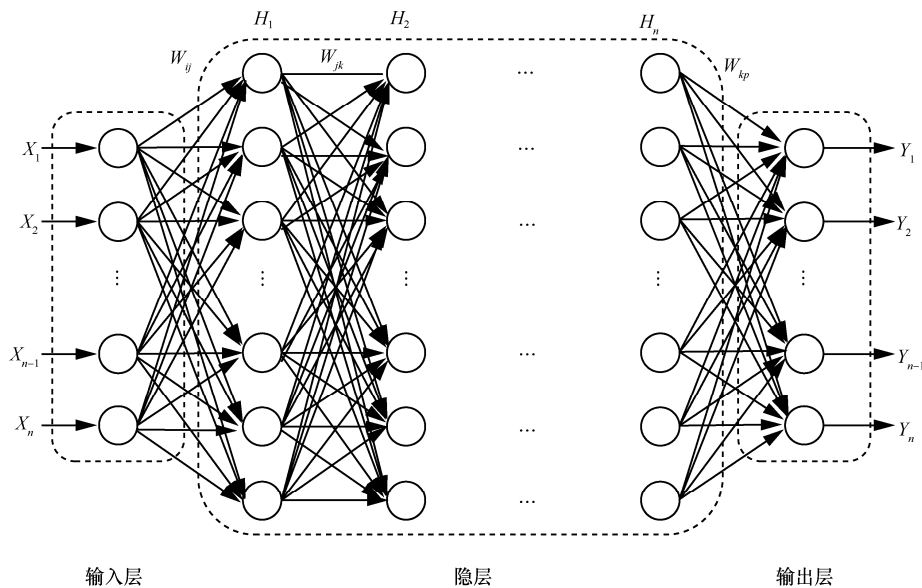


图 1 DNN 结构示意图

(DRNN) 在图像目标提取^[13]、语音识别^[14]、机器翻译^[15]中都取得不错效果。循环神经网络(RNN, recurrent neural network)思想起源于 Jordan 等^[16]提出的连接序列机, 通过在神经网络中引入自反馈神经元, 使网络对输入数据产生记忆功能, 充分挖掘样本之间存在的相关性特征。文献[17]曾捕获 3 年内某城市主干网络中入侵事件并进行统计, 发现不同入侵行为具有明显的时间分布特性, 且同一入侵行为也具有持续性等特点, 如拒绝服务攻击 (DoS, denial of service), 攻击者在短时间内反复高频地发送服务请求, 占用大量网络带宽资源, 造成用户合法请求无法完成。针对此类攻击, 利用已有事件与当前事件之间时序相关性进行入侵检测及预测至关重要。

单隐层 RNN 时序结构及展开形式如图 2 所示, $t-1$ 、 t 、 $t+1$ 表示网络训练或测试过程中连续 3 个时间状态点。 $X_t \in \mathbf{R}^{N+1}$ 、 $h_t \in \mathbf{R}^{M+1}$ 、 $Y_t \in \mathbf{R}^{L+1}$ 分别表示 t 时刻的输入向量、隐层向量和输出向量, N 、 M 、 L 分别为输入向量维度、隐层维度及输出向量维度。 $W_{xh} \in \mathbf{R}^{M \times N}$ 、 $W_{hh} \in \mathbf{R}^{M \times M}$ 、 $W_{hy} \in \mathbf{R}^{L \times M}$ 分别表示输入层与隐层、隐层与隐层、隐层与输出层之间权重矩阵, 且不同时刻连接权值共享, 缩减网络参数量。

t 时刻隐层值 h_t 由当前状态输入值 X_t 和 $t-1$ 时刻隐层神经元激活输出值共同决定。

$$h_t = f(X_t, h_{t-1}, b) = f(W_{xh}X_t + W_{hh}h_{t-1} + b) \quad (2)$$

一般无初始值情况下

$$h_{-1} = (0, \dots, 0)_{M \times 1}^T \quad (3)$$

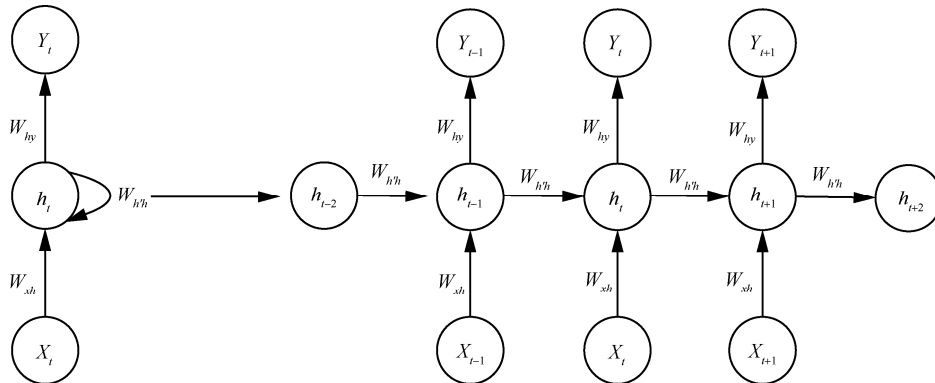


图 2 RNN 结构示意图

$$h_0 = f(W_{xh}X_t + b) \quad (4)$$

其中, b 为偏置向量, $f(\cdot)$ 为神经元激活函数, 本文选取为整流线性单元 (ReLU, rectified linear units) 激活函数。

$$f(\cdot) = \begin{cases} 0, & (x \leq 0) \\ x, & (x > 0) \end{cases} \quad (5)$$

输出层为全连接层, 利用 Softmax 函数实现特征向量到类别概率分布的转化, c 为偏置向量。

$$Y_t = W_{hy}h_t + c \quad (6)$$

$$P_{Y_t^i} = \text{softmax}(Y_t^i) = \frac{e^{Y_t^i}}{\sum_{i=1}^k e^{Y_t^i}} \quad (7)$$

Y_t^i 表示网络输出向量 Y_t 中第 i 个特征分量, $P_{Y_t^i}$ 表示相应特征分量概率输出, $0 < P_{Y_t^i} < 1$, $\sum_{i=1}^k P_{Y_t^i} = 1$ 。

隐层单元进行多层级联形成 DRNN, 引入 $W_{h^i h^{i+1}} \in \mathbf{R}^{M \times M}$ 表示层间连接权重矩阵, 层内权重共享扩展为层间权重共享, DRNN 时序结构如图 3 所示。

此时

$$h_t^i = \begin{cases} f(W_{xh}X_t + W_{h^i h^{i+1}}h_{t-1}^i + b), & i = 1 \\ f(W_{h^i h^{i+1}}h_{t-1}^{i-1} + W_{h^i h^{i+1}}h_{t-1}^i + b), & i > 1 \end{cases} \quad (8)$$

$$Y_t = W_{hy}h_t^i + c \quad (9)$$

2.3 权值调优过程

本文使用实时循环学习 (RTRL, real-time recurrent learning) 算法对 DRNN 连接权重进行调

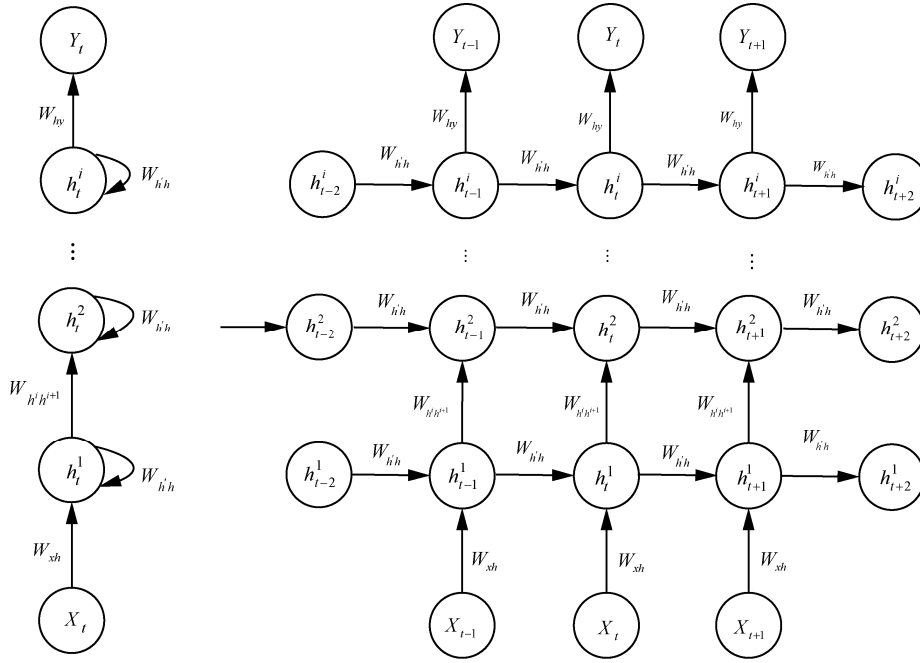


图3 DRNN 结构示意图

优。RTRL 通过梯度信息前向传播, 保证模型学习与调优过程同步进行, 同时避免了传统反向梯度传播算法在深层网络训练时面临的梯度消失和梯度爆炸问题。

由于 DRNN 每个时刻均有输出, 定义 $X_{g,t}$ 、 $Y_{g,t}$ 、 $D_{g,t}$ 、 $S_{g,t}$ 分别为节点 g 在 t 时刻输入向量, 隐层及输出层实际输出向量、隐层及输出层期望输出向量、隐层节点激活向量; W_{ij} 表示节点 i 到节点 j 连接权重; E_t 为均方误差损失函数; M, N 分别表示输入节点集合, 隐层及输出层节点集合。

神经元激活输出向量 $Y_{g,t}$ 在 $t+1$ 时刻表示为

$$Y_{g,(t+1)} = f_g(S_{g,t}) = f_g \left[\sum_{a \in M} W_{ga} X_{a,t} + \sum_{b \in N} W_{gb} Y_{b,t} + \varphi \right] \quad (10)$$

$f(\cdot)$ 为隐层激活函数, φ 为偏置向量, $g \in N$ 。

DRNN 损失函数为所有时刻损失函数之和。

$$E_{\text{total}}(t_0, t_1) = \sum_{\tau=t_0+1}^{t_1} E_{\tau} = \frac{1}{2} \sum_{\tau=t_0+1}^{t_1} \sum_{g \in N} (e_{g,\tau})^2 = \frac{1}{2} \sum_{\tau=t_0+1}^{t_1} \sum_{g \in N} (D_{g,\tau} - Y_{g,\tau})^2 \quad (11)$$

权重更新值如式(11)所示, 其中, α 为学习率, $\eta_{ij,t}^g$ 表示节点 g 对于权重改变量 ΔW_{ij} 的

敏感度。

$$\Delta W_{ij,t} = -\alpha \frac{\partial E_t}{\partial W_{ij}} = \alpha \sum_{g \in N} e_{g,t} \frac{\partial Y_{g,t}}{\partial W_{ij}} = \alpha \sum_{g \in N} e_{g,t} \eta_{ij,t}^g \quad (12)$$

式(9)左右两边对 W_{ij} 求微分, 可得

$$\eta_{ij,(t+1)}^g = f'_g(S_{g,t}) \left[Z_{j,t} \delta_{gi} + \sum_{b \in N} W_{gb} \eta_{ij,t}^b \right] \quad (13)$$

$$\delta_{gi} = \begin{cases} 0, & g \neq i \\ 1, & g = i \end{cases} \quad (14)$$

式(11)更新为

$$\Delta W_{ij,(t+1)} = \alpha \left[\sum_{g \in M \cup N} e_{g,t} f'_g(S_{g,t}) W_{ij,t} \right] \eta_{ij,t}^g \quad (15)$$

初始值 $\eta_{ij,t=0}^g = 0$ 。

2.4 权值调优过程

本文提出的 DRRS 入侵检测模型计算复杂度主要由 RA-SMOTE 低频样本增量算法与 RTRL 权值更新算法决定。RA-SMOTE 算法可分为样本增量与时序插入两部分: 样本增量过程的计算复杂度为 $O(n)$, n 为低频样本数量; 时序插入部分的计算复杂度为 $O(m)$, m 为新样本数量。因此, RA-SMOTE 算法总计算复杂度为 $O(nm)$ 。RTRL 算法计算复杂度为 $O(k^4)$, 主要用于更新敏感度

η_{ij}^g , 其中, k 为模型运算单元数。因此, 根据上述分析, DRRS 总计算复杂度为 $O(nm + k^4)$, 对于模型复杂度要求较低的轻量级入侵检测系统具有较好的可实现性, 在合理可控范围之内。

3 实验设置

本文实验流程如图 4 所示, 首先对实验所用入侵检测标准 NSL-KDD 数据集 (见 3.1 节) 进行数据预处理 (见 3.2 节), 包括特征数值化和归一化两项过程。同时为简化实验过程, 对原始数据及进行随机独立抽样组成多个新数据集, 并分为训练集和测试集。其次, 对训练集完成低频样本增量与时序注入过程, 形成新的训练数据集并使用新数据集对时序网络模型进行训练调优, 获取最优参数网络模型。测试集则用于验证最终实验结果有效性。实验中对每个数据集上进行多次独立重复实验, 且测试集之间交叉验证, 以保证结果无偏性, 最终实验结果取每个数据集上检测结果之和的平均值。

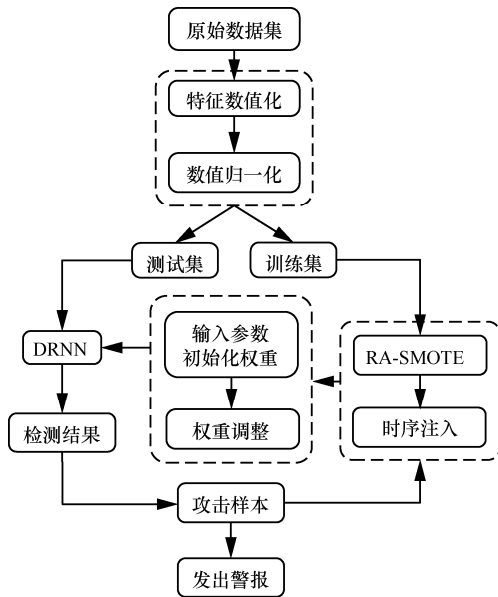


图 4 DRRS 入侵检测实验流程

3.1 实验数据集

本文实验基于入侵检测标准 NSL-KDD^[18]数据集仿真实验。NSL-KDD 数据集改进于 KDD 99 数据集, KDD 99 数据集模拟了美国空军局域网环境, 但其中包含过多冗余数据, 增加计算负担的同时会对分类结果造成误导, 因此 NSL-KDD

数据集更适用于仿真实验。

NSL-KDD 训练集和测试集分别包含 125 973 和 22 543 条连接记录, 其中主要包括 4 种类型的攻击: 拒绝服务攻击, 端口漏洞扫描攻击 (Probing)、远程控制攻击 (R2L, remote to local)、越权访问攻击 (U2R, user to root), 具体分布如表 1 所示。4 种攻击类别具体又可分为 37 种不同攻击, 训练集包含 21 种, 测试集包含 37 种, 其中有 16 种新型攻击未出现在训练集中。

表 1 NSL-KDD 数据类型分布

数据集	正常类别	攻击类别				总计
		DoS	Probing	R2L	U2R	
训练集	67 343	45 927	11 656	995	52	125 973
测试集	9 691	7 457	2 421	2 754	220	22 543

3.2 数据预处理

NSL-KDD 数据集中每一条连接记录都由 41 种分类特征组成, 其中包括 3 种符号型特征, 10 种 0-1 型特征, 15 种百分比型特征和 13 种十进制型特征。通过分析数据集, 发现特征 Num_outbound_cmds 数值全部为 0, 故将此特征移除, 因此可用特征共 40 种。数据预处理过程分为两步: 特征数值化和数值归一化, 具体过程如下所示。

1) 特征数值化

特征数值化主要针对 Protocol_type、Service 和 Flag 这 3 种符号型特征, 本文将其映射为二进制值。以 Protocol_type 特征为例, 其包括 TCP、UDP、ICMP 这 3 种协议类型, 可分别映射为 (0,0,1), (0,1,0), (1,0,0)。同理, Service 和 Flag 分别包括 70 种和 11 种符号, 因此数值化完成后数据集特征维度增加至 121 维。除此以外, 攻击类别标签同样需要数值化, 方法同理, 不再赘述。

2) 数值归一化

为简化运算和加速模型收敛速度, 同时便于实验结果分析比较, 使用 Min-Max 标准化方法将十进制型特征值统一映射到 [0,1], 如式 (16) 所示, 其中, X 表示样本特征值, $[X_{\min}, X_{\max}]$ 表示特征值取值区间, X_{norm} 表示归一化后特征值。

$$X_{\text{norm}} = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \quad (16)$$

为简化实验过程, 对训练集和测试集随机独

立抽样组成多个新数据集，如表 2 所示，并保留抽取样本在原数据集上时序排列关系。由于原始训练集中 U2R 样本数量稀少，全部保留到新数据集中不进行抽样。每个新数据集上进行多次独立重复试验，且测试集之间交叉验证，以保证结果无偏性，最终实验结果取每个数据集结果之和的平均值。

表 2 实验数据集

编号	训练集		测试集	
	正常样本	攻击样本	正常样本	攻击样本
No.1	8 647	1 237	3 783	1 373
No.2	9 268	1 364	4 178	1 438
No.3	7 899	1 083	4 903	1 312
No.4	10 045	1 544	3 966	1 194
No.5	8 902	1 395	3 834	1 011
No.6	9 570	1 528	2 045	939

3.3 实验参数

本文实验使用计算机硬件配置为 Inter Core i7-7700 四核处理器，8 GB 内存，256 GB 固态硬盘。操作系统为 64 位 Windows 10 系统。其中，RA-SMOTE 算法基于 R 语言编程实现，DRNN 基于 Google 开源深度学习框架 TensorFlow 实现，编程语言为 python。

DRRS 模型主要的参数变量包括 RA-SMOTE 算法中最近邻半径和样本合成过采样率，DRNN 结构层数和隐层单元数，RTRL 算法中权重更新学习率等。变量参数值如表 3 所示，具体参数选择过程见 4.1 节。

表 3 实验变量参数值

变量名称	变量值
最近邻半径	55
过采样率	600%
DRNN 层数	5
输入层神经元个数	121
第一隐层神经元个数	100
第二隐层神经元个数	65
第三隐层神经元个数	35
输出层神经元个数	5
权重学习率	0.01
最大迭代次数	50

3.4 评价指标

实验所用评价指标基于混淆矩阵表示，包括正确率 (ACC, accuracy rate)，精确率 (PRE, precision)，虚警率 (FAR, false alarm rate)，具体定义如式 (16)~式 (18) 所示。TP (true positive) 表示分类正确的正类样本数，FP (false positive) 表示分类错误的正类样本数，TN (true negative) 表示分类正确的负类样本数，FN (false negative) 表示分类错误的负类样本数。

$$ACC = \frac{TP + TN}{TP + FP + TN + FN} \quad (17)$$

$$PRE = \frac{TP}{TP + FP} \quad (18)$$

$$FAR = \frac{FP}{FP + TN} \quad (19)$$

除此以外，训练时间 (T_{train}) 和测试时间 (T_{test}) 用来评价 DRRS 模型建立及检测速度。

4 仿真实验及结果分析

实验仿真验证过程分为 2 个部分：4.1 节采用控制变量法对 DRRS 模型算法最优参数进行测定，探究 RA-SMOTE 算法最近邻半径、采样率、DRNN 层数及隐层神经元数对模型最终检测结果的影响，并选出最优参数，构建最优参数模型；4.2 节则采用 4.1 节测定的最优参数构建 DRRS 模型，并在整体正确率、虚警率、训练及测试时间等方面与已有入侵检测法进行对比实验，验证本文模型有效性。

4.1 最优参数选择分析

1) RA-SMOTE 算法最近邻半径的影响

RA-SMOTE 算法样本增量过程需要对每个低频攻击样本划定最近邻半径，根据半径内高频样本数量差异，选择不同增量方式。因此最近邻半径的大小将影响样本区域划分，进而改变样本增量方式。半径过大，高频样本比例升高，可能导致原本属于 SPR 区域集合内的低频样本被划分为 DPR 区域集合，从而合成不必要的新样本，造成虚警率上升；反之，半径过小将导致 DPR 区域集合内低频样本点划分为 SPR 区域集合，无法有效合成新样本。

实验中对最近邻半径取值在 [0,100] 范围内的模型检测正确率及虚警率进行统计，结果如图 5 所示，可知当最近邻半径取值范围在 [50,60] 内时，

检测正确率达到平稳,此后不再显著增加,且最近邻半径大于60时,虚警率逐渐增加,因此综合考虑,取最近邻半径55为最优值。

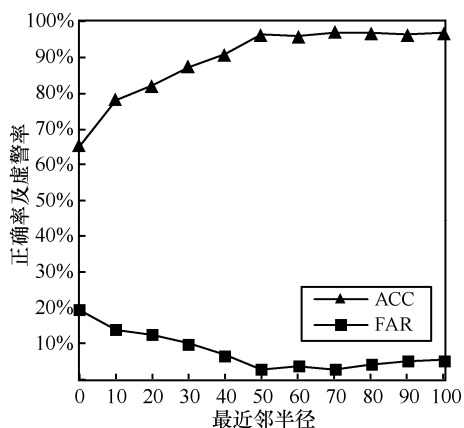


图5 最近邻半径对模型性能指标的影响

2) RA-SMOTE 算法采样率的影响

模型检测结果同样受到 RA-SMOTE 算法采样率的影响:采样率过低将导致新生成样本数量不足,影响后续 DRNN 对样本特征的学习;采样率过高则造成样本数量上的浪费。图6为检测率百分比在[100,1000]范围内取100整数倍时的模型检测正确率及虚警率。由实验结果可知,当采样率为600%时,正确率及虚警率处于综合最佳状态。百分比继续升高,正确率与虚警率无明显改善,表明此时新生成的样本属于过量样本。

3) DRNN 层数及隐层单元数的影响

DRNN 模型层数及隐层单元数对模型检测正确率等指标具有重要影响。已有研究表明,深层网络结构模型层数及隐层单元数越多,模型对于

数据的非线性拟合能力越强。同时深层结构更有利于实现原始高维分类特征向其低维表示的抽象转化,加速后续分类过程。但模型层数及隐层单元数过多,将导致训练时间急剧增加,网络连接权重调优过程复杂度升高,模型难以收敛。因此如何在模型结构与复杂度之间找到平衡点至关重要。

实验通过对已有文献中深层网络结构进行统计,选择其中5种不同模型结构进行实验,获取最优结构。模型结构分别为 RNN2 (121-5, 二层), RNN3 (121-60-5, 三层), RNN4 (121-90-45-5, 四层), RNN5 (121-100-65-35-5, 五层), RNN6 (121-110-80-55-30-5, 六层)。实验结果如图7和图8所示,由图可知模型层数到达五层时检测正确率已达99%以上,且继续增加层数正确率无明显提升,而此时模型训练时间和测试时间随着层数增加,增长率不断升高,呈非线性增长趋势,因此 [121-100-65-35-5] 五层结构适用于本文 DRRS 入侵检测模型。

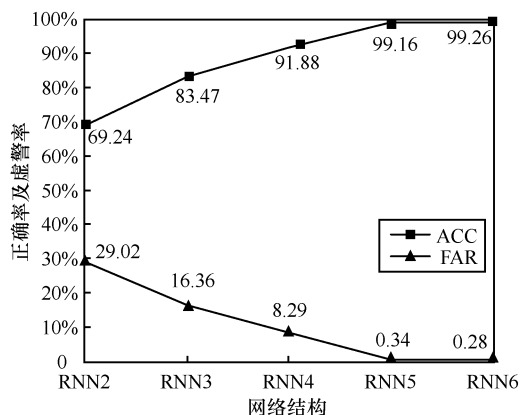


图7 RNN 结构对模型性能指标的影响

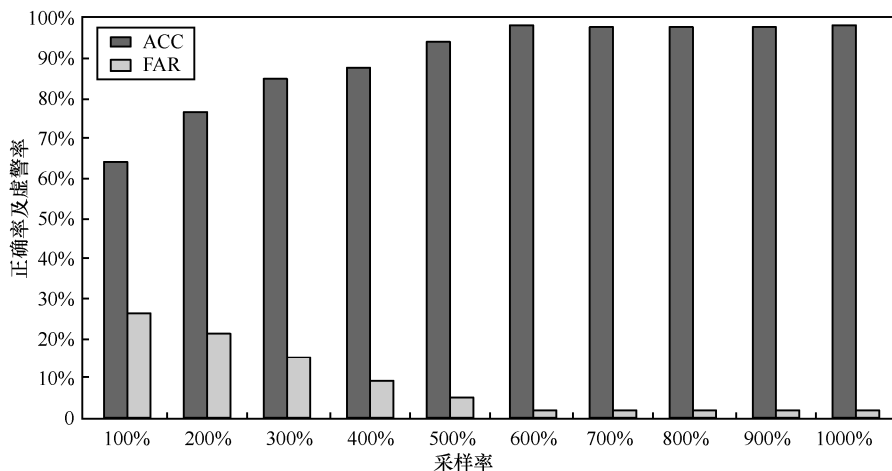


图6 采样率对模型性能指标的影响

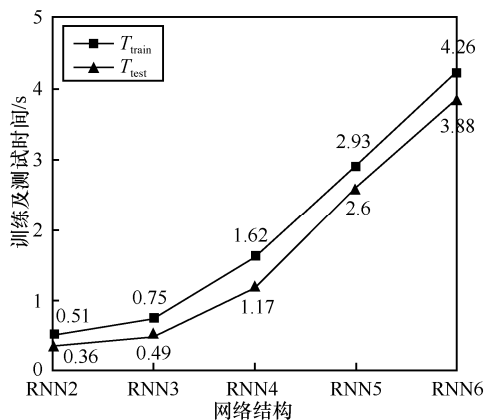


图8 RNN 结构对模型训练测试时间的影响

4.2 对比实验结果分析

在 4.1 节最优参数选择实验的基础上, 采用选择出的最优参数建立检测模型并与已有入侵检测算法在 PRE、FAR 等参数方面进行实验对比, 验证本文所提出的 DRRS 模型的有效性。

1) 表 4 为 RA-SMOTE 算法有效性的对比验证实验。分别在对原始数据集不采用过采样算法、采用 SMOTE 算法和采用 RA-SMOTE 算法的基础上训练 DRNN 模型并进行验证。从表中可以看出, 对于 Normal、DoS 和 Probing 这 3 类高频样本集合, 3 种 DRNN 模型在检测精确率与虚警率指标上并无过大差异。然而对于 R2L 和 U2L 低频样本集合, 过采样算法的使用对检测精确率的提升和虚警率的抑制都起到了明显的效果。虽然相比于单独 DRNN 模型, SMOTE 算法和

表 4 过采样算法实验性能对比

样本类别	评价指标	DRNN	DRNN+SMOTE	DRNN+RA-SMOTE
Normal	PRE	98.52%	98.66%	99.01%
	FAR	1.47%	1.28%	1.29%
DoS	PRE	97.92%	98.34%	98.76%
	FAR	2.27%	2.04%	1.97%
Probing	PRE	97.88%	98.56%	98.42%
	FAR	3.04%	1.54%	1.23%
R2L	PRE	69.26%	94.63%	95.39%
	FAR	32.78%	6.67%	3.99%
U2L	PRE	24.32%	92.54%	94.02%
	FAR	77.50%	7.27%	5.60%
模型训练时间 T_{train}/s		2.68	3.24	2.93

RA-SMOTE 算法时间消耗分别增长了 20.90% 和 9.33%, 但由此换来的性能改善是值得的。更进一步地, RA-SMOTE 算法比 SMOTE 算法对模型训练时间的影响缩短了 9.29%, 表明 RA-SMOTE 算法以更高效的方式合成了训练所需的新样本。

2) 本文选择已有文献中提出的 6 种入侵检测模型与 DRRS 进行对比实验[3,5,7-10], 检测指标包括整体正确率与虚警率, 结果如图 9 和图 10 所示。由结果可知, 在整体检测正确率方面, DRRS 模型实验结果优于其他检测模型, 说明 DRRS 模型充分挖掘出了样本集数据之间存在的时序关系, 并对其进行了有效的学习训练, 从而提高了模型整体检测正确率。在整体虚警率指标上, DRRS 模型同样获得了不错的效果, 仅略高于文献[3]中提出的 CHI-SVM 模型。

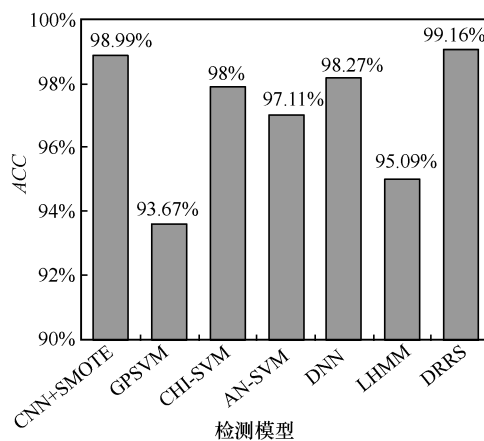


图9 不同检测模型整体正确率对比

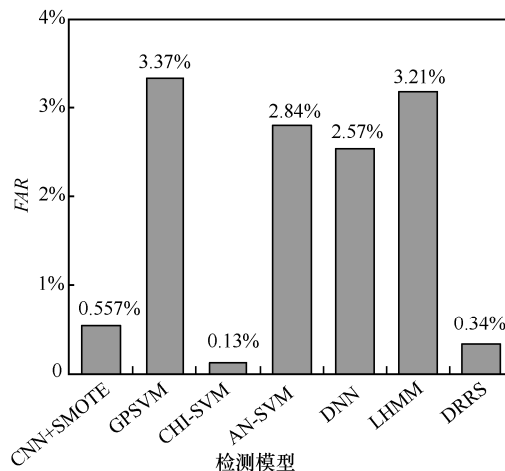


图10 不同检测模型整体虚警率对比

表5 不同检测模型单项检测精确度对比

样本类别	PRE						
	CCNN+SMOTE ^[7]	GPSVM ^[8]	CHI-SVM ^[3]	AN-SVM ^[5]	DNN ^[9]	LHMM ^[10]	DRRS
Normal	—	—	96.10	99.68	99.52	—	99.01
DoS	—	91.32	98.87	99.89	97.00	98.43	98.76
Probing	—	85.50	95.80	98.83	98.56	98.61	98.42
R2L	92.97	90.72	76.92	87.58	71.00	31.32	95.39
U2L	55.91	89.28	96.37	18.31	71.00	86.66	94.02

3) 不同入侵检测模型在 5 种样本类型上的单项检测精确度对比结果如表 5 所示。对于 Normal、DoS 和 Probing 这 3 类高频样本, DRRS 模型检测精确度与已有模型基本持平, 表明本文模型对于高频样本检测精确度已达到平均标准, 满足预期要求。对于 R2L 和 U2L 两类低频攻击样本, DRRS 模型检测精确度优势明显, 除 U2L 检测精确度略低于 CHI-SVM 模型外, 均高于其余模型精确度。

综合上述各项实验结果, 说明本文提出的 DRRS 模型不仅可以有效处理低频攻击样本检测率低的问题, 且 RA-SMOTE 算法以更小的时间代价实现了样本增量过程。同时 DRRS 模型在整体检测正确率及虚警率等指标上均取得不错的效果, 对于测试数据集中存在的未知攻击样本同样具有检出能力。

5 结束语

传统入侵检测模型无法有效处理低频攻击样本, 且模型构建过程未考虑样本间时序关系。本文提出了一种新的基于深度循环神经网络和区域自适应合成过采样算法的组合式入侵检测模型 DRRS。首先, DRRS 利用 RA-SMOTE 算法自适应的对低频样本实现样本增量及时序注入过程, 以更少的时间消耗改善了原始数据集中数据非平衡分布状况。然后, 利用 DRNN 模型特有的内部循环结构, 完成对新样本集的时序积累学习与检测。在 NSL-KDD 数据集上, 实验结果表明本文提出的 DRRS 模型有效解决了上述问题, 且各项指标均优于已有入侵检测模型, 对于模型复杂度需求较低的轻量级入侵检测模型具有很好的实用性, 适用于实际网络环境, 为入侵检测问题提出了新的研究思路。

参考文献:

- [1] 国家互联网应急中心. 2017 年 7 月我国互联网安全威胁报告[R]. 2017.
National Internet Emergency Center. The report of China's Internet security threat in July[R]. 2017.
- [2] LEI Y, LIU J, YIN H. Intrusion detection techniques based on improved intuitionist fuzzy neural networks[J]. Applied Mechanics & Materials, 2014, 713-715(1): 2507-2510.
- [3] THASEEN I S, KUMAR C A. Intrusion detection model using fusion of chi-square feature selection and multi class SVM[J]. Journal of King Saud University-Computer and Information Sciences, 2016, 29(4):462-472.
- [4] DASTANPOUR A, IBRAHIM S, MASHINCHI R. Comparison of genetic algorithm optimization on artificial neural network and support vector machine in intrusion detection system[C]//IEEE International Conference on Open Systems, 2014: 72-77.
- [5] ABDLHAMED M, KIFAYAT K, SHI Q. Intrusion prediction systems[J]. Information Fusion for Cyber-Security Analytics, 2017, 69(1): 155-174.
- [6] PARSAEI M, ROSTAMI S, JAVIDAN R. A hybrid data mining approach for intrusion detection on imbalanced NSL-KDD dataset[J]. International Journal of Advanced Computer Science and Applications, 2016, 7(6): 20-25.
- [7] POZI M, SULAIMAN M, MUSTAPHA N. Improving anomalous rare attack detection rate for intrusion detection system using support vector machine and genetic programming[J]. Neural Processing Letters, 2016, 44(2): 279-290.
- [8] 高妮, 高岭, 贺毅岳. 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. A lightweight intrusion detection model based on autoencoder network with feature reduction[J]. Acta Electronica Sinica, 2017, 45(3): 730-739.
- [9] DIRO A, CHILAMKURTI N. Distributed attack detection scheme using deep learning approach for Internet of Things[J]. Future Generation Computer Systems, 2018, 82(1):761-768.
- [10] CHINCHORE R, SAMBARE S. Intrusion detection system by layered approach and hidden Markov model[J]. International Journal of Computer Application, 2015, 5(2):7-14.

- [11] CHAWLA NV, BOWYER KW, HALL LO, et al. SMOTE: synthetic minority over-sampling technique[J]. Journal of Artificial Intelligence Research, 2002, 16(1):321-357.
- [12] HINTON G, SALAKHUTDINOV R. Reducing the dimensionality of data with neural networks[J]. Science, 2006, 313(28): 504-507.
- [13] 徐彬, 陈渤, 刘宏伟. 基于注意循环神经网络模型的雷达高分辨率距离像目标识别[J]. 电子与信息学报, 2016, 38(12): 2988-2995.
- XU B, CHEN B, LIU H W. Attention-based recurrent neural network model for radar high-resolution range profile target recognition[J]. Journal of Electronics & Information Technology, 2016, 38(12): 2988-2995.
- [14] THANDA A, VENKATESAN S M. Audio visual speech recognition using deep recurrent neural networks[C]. IAPR Workshop on Multimodal Pattern Recognition of Social Signals in Human-Computer Interaction, 2016: 98-109.
- [15] GUAMAN F, JOTY S, MARQUEZ L, et al. Machine translation evaluation with neural networks[J]. Computer Speech & Language, 2017, 45(1): 180-200.
- [16] JORDAN MI. Attractor dynamics and parallelism in connectionist sequential machine[C]. Eighth Conference of the Cognitive Science Society, 1986:531-546.
- [17] SONG J, TAKAKURA H, OKABE Y. Statistical analysis of honeypot data and building of Kyoto 2006+ dataset for NIDS evaluation[C]. The Workshop on Building Analysis Datasets & Gathering Experience Returns for Security. 2011:29-36.
- [18] TAVALLAEE M, BAGHERI E, LU W. A detailed analysis of the KDD CUP 99 data set[C]. IEEE International Conference on Computational Intelligence for Security and Defense Applications. 2009:53-58.

[作者简介]



燕曷昊 (1994-), 男, 山西吕梁人, 国家数字交换系统工程技术研究中心硕士生, 主要研究方向为机器学习、入侵检测、网络安全。



韩国栋 (1964-), 男, 山东莱西人, 博士, 国家数字交换系统工程技术研究中心副教授、硕士生导师, 主要研究方向为宽带信息处理、信息安全、芯片设计技术。