计算机科学 COMPUTER SCIENCE
http://www.jsjkx.com
DOI:10.11896/jsjkx.200500059

# 基于降噪自编码器和三支决策的入侵检测方法

张师鹏　李永忠
江苏科技大学计算机学院　江苏 镇江 212003
(1099682749@qq.com)

**摘　要**　入侵检测在计算机网络安全防御中起着至关重要的作用,是网络安全的关键技术之一。随着网络环境越来越复杂,网络入侵行为也逐渐表现出了多样化及智能化的特点,且越来越难以被检测到。基于上述原因,人们对已有入侵检测方法的可行性与可持续性表示担忧,具体来说就是已有的入侵检测算法很难完美地抽象出入侵行为所包含的特征,且已有的入侵检测方法在未知攻击上大都表现不佳。针对这些问题,文中提出了基于降噪自编码器和三支决策的入侵检测算法 DAE-3WD。该方法通过降噪自编码器对高维数据进行特征提取,利用多次的特征提取来构造多粒度的特征空间,然后基于三支决策理论对属于入侵或正常的行为做出立即决策,而对于疑似入侵或者疑似正常的行为则根据不同粒度的特征进行进一步的分析。深度学习具有优越的分层特征学习能力,且三支决策可以规避因信息不足而盲目分类造成的风险,该方法利用这些特性可以达到提升入侵检测表现的目的。在 NSL-KDD 数据集上进行了实验,实验结果证明,所提算法能提取到有意义的特征并能有效提升入侵检测算法的表现。

**关键词**:入侵检测;自编码器;三支决策;特征提取;网络安全
**中图法分类号**　TP309

## Intrusion Detection Method Based on Denoising Autoencoder and Three-way Decisions

ZHANG Shi-peng and LI Yong-zhong
School of Computer,Jiangsu University of Science and Technology,Zhenjiang,Jiangsu 212003,China

**Abstract**　Intrusion detection plays a vital role in computer network security. Intrusion detection is one of the key technologies of network security and needs to be kept under constant attention. As the network environment becomes more and more complex, network intrusion behaviors gradually show diversified and intelligent characteristics, and network intrusion is also becoming more difficult to detect. And the research conducted in the field of network security is also an endless study. For the above reasons, people are worried about the feasibility and sustainability of the current method, specifically, it is difficult for current intrusion detection methods to perfectly abstract the features contained in intrusion behaviors, and most of the current intrusion detection methods perform poorly on unknown attacks. In response to these problems, we propose an intrusion detection method DAE-3WD based on denoising autoencoder and three-way decisions. We hope that our method can effectively promote the research on intrusion detection. This proposed method extracts features from high-dimensional data through denoising autoencoder. Through multiple feature extractions, a multi-granular feature space can be constructed, and then an immediate decision on intrusive or normal behavior is made based on the three-way decisions, and further analysis is required for suspected intrusion or normal behavior. Deep learning has superior hierarchical feature learning ability, and three-way decisions can avoid the risk of blind classification due to insufficient information. This method uses these characteristics to achieve the purpose of improving the performance of intrusion detection. The NSL-KDD data set is used in our experiments. The experiments prove that the proposed method can extract meaningful features and effectively improve the performance of intrusion detection.

**Keywords**　Intrusion detection, Autoencoder, Three-way decisions, Feature extraction, Network security

到稿日期:2020-05-14　返修日期:2020-08-21
基金项目:国家自然科学基金(61471182);江苏省研究生科研与实践创新计划项目(KYCX20_3163);江苏省高校自然基金项目(15KJD52004)
This work was supported by the National Nature Science Foundation of China(61471182), Postgraduate Research & Practice Innovation Program of Jiangsu Province(KYCX20_3163) and Natural Science Foundation of the Jiangsu Higher Education Institutions of China(15KJD52004).
通信作者:李永忠(liyongzhong61@163.com)

## 1 引言

计算机网络的安全问题一直是一个亟待解决的棘手问题，如何识别网络攻击是一个关键问题[1]。入侵检测技术是网络安全的关键技术之一，引起了国内外学者的广泛关注[2]。入侵检测系统作为防御网络攻击的重要实体之一，通过基本手段检测网络行为，为防御者提供武器，触发针对这些行为的最佳决策计划[3-4]。入侵检测技术不仅可以保护传统的计算机网络，而且也对如云计算系统等新型的网络系统起到了一定的保护作用[5]。

将机器学习方法应用于入侵检测，能使系统具有更强的适应性、自学习性和鲁棒性，这是一个重要的研究方向[6-7]。目前，包括数据挖掘[8]、深度学习等在内的许多机器学习算法都被应用于入侵检测领域。其中，深度学习[9]可以消除浅层学习的限制，具有优秀的分层特征学习能力，能够促进对网络数据的深入分析。深度置信网络(Deep Belief Networks，DBN[10])、自编码器[11-12]等都被广泛地应用到了入侵检测领域，并且发挥着越来越重要的作用。

然而，当前在入侵检测领域的研究主要都是基于二支决策的，这种分类方法的容错能力差，且不能根据特征粒度的大小来对网络行为做出动态决策。本文结合降噪自编码器(denoising autoencoder)以及三支决策(three-way decisions)理论建立了一个入侵检测模型，通过降噪自编码器来获取粒度的不同特征，并利用三支决策理论对网络行为进行分类。基准数据集 NSL-KDD 被用于评估本文算法的性能，并在测试数据集上验证了本文算法的扩展性。

## 2 相关理论

### 2.1 降噪自编码器

降噪自编码器[13]是自编码器的一种变体，由编码器和解码器组成。通过在输入中引入随机噪声来迫使网络在学习的过程中去除随机噪声，以重构原始输入，这种学习方式能够降低网络对输入样本的敏感性，增强隐藏层特征的学习能力，提高自编码器对输入数据的泛化能力。

假设 $\tilde{X}=\{\tilde{x}_1,\tilde{x}_2,\cdots,\tilde{x}_n\}$ 为加入了噪声的输入，编码器将输入 $\tilde{X}$ 映射到一个隐藏 $X$ 的表示：

$$h=f(W\tilde{X}+b) \tag{1}$$

解码器通过式(2)把隐藏层映射到输出层。

$$\hat{X}=g(\hat{W}h+\hat{b}) \tag{2}$$

在式(1)、式(2)中，$W$，$\hat{W}$ 指权重，$b$，$\hat{b}$ 指偏置值，其中 $\hat{W}=W^{\mathrm{T}}$。

自编码器的代价函数 $J(W,b)$ 为：

$$J(W,b)=\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{2}\left\|\tilde{x}^{(i)}-\hat{x}^{(i)}\right\|^2\right)+\frac{\lambda}{2}\sum_{l=1}^{m_l-1}\sum_{i=1}^{S_l}\sum_{j=1}^{S_{l+1}}(W_{ij}^{l})^2 \tag{3}$$

其中，第一项为输入输出的均方误差，$n$ 是输入层神经元的个数，$\hat{x}$ 是 $\tilde{x}$ 的重构输出，由于训练集中的数据可能会出现过拟合的情况，因此引入第二项，即正则项(权重衰减项)，权重衰减参数 $\lambda$ 用于权衡式(3)中两个项之间的相对重要性，$W_{ij}^{l}$ 是第 $l$ 层第 $j$ 个神经元与第 $l+1$ 层第 $i$ 个神经元之间的权重，$b$ 是节点偏置，$m_l$ 是网络层数，$S_l$ 是第 $l$ 层的神经元个数，$S_{l+1}$ 是第 $l+1$ 层的神经元个数。

### 2.2 三支决策理论

三支决策[14]来源于粗糙集。当现有的信息不足以支撑做出明确的选择时，为了避免可能造成高误分率，不承诺是一种选择，因此三支决策可以规避因信息不足而盲目决策造成的风险[15]。

误分类会造成损失，不同的误分类错误会导致不同的损失。例如，在入侵检测的研究中，把一个正常的行为错误地归为异常行为可能会造成一些麻烦，但是如果把一个入侵的行为错误地分类成正常行为就有可能造成灾难性的后果[16]。

对于一个二分类问题，真实的分类标签可以表示为 P(正)和 N(负)，即接受和拒绝，可以用一个状态集 $\Omega=\{X,\neg X\}$ 来表示，即用某个数据属于 $X$ 与某个数据不属于 $X$ 来表示一个数据的归属问题。三支决策的决策集可以表示为 $D=\{D_P,D_B,D_N\}$，分别表示正向决策、边界决策以及负向决策。所有决策的代价损失函数如表 1 所列。记 $\lambda_{PP}$，$\lambda_{BP}$，$\lambda_{NP}$ 分别表示当前数据属于 $X$ 且采取行动 $D_P$，$D_B$ 以及 $D_N$ 时的损失，$\lambda_{PN}$，$\lambda_{BN}$，$\lambda_{NN}$ 分别表示当前数据不属于 $X$ 且采取行动 $D_P$，$D_B$ 以及 $D_N$ 时的损失。

表 1 三支决策的代价函数

Table 1 Cost function of three-way decisions

| | $P$ | $N$ |
|---|---|---|
| $D_P$ | $\lambda_{PP}$ | $\lambda_{PN}$ |
| $D_B$ | $\lambda_{BP}$ | $\lambda_{BN}$ |
| $D_N$ | $\lambda_{NP}$ | $\lambda_{NN}$ |

根据文献[17]的推演证明，可以得到如下两个相关阈值的计算公式：

$$\alpha=\frac{(\lambda_{PN}-\lambda_{BN})}{(\lambda_{PN}-\lambda_{BN})+(\lambda_{BP}-\lambda_{PP})} \tag{4}$$

$$\beta=\frac{(\lambda_{BN}-\lambda_{NN})}{(\lambda_{BN}-\lambda_{NN})+(\lambda_{NP}-\lambda_{PP})} \tag{5}$$

其中，$0\leqslant\beta\leqslant\alpha<1$。还可以得到如下 3 条应用到入侵检测领域的规则：

(1)如果 $P(X|[x])>\alpha$，则该网络行为被归为正类，即该网络行为是入侵行为；

(2)如果 $P(X|[x])<\beta$，则该网络行为被归为负类，即该网络行为是正常行为；

(3) 如果 $\beta\leqslant P(X|[x])\leqslant\alpha$，则表示在当前信息下，无法对该行为采取任何决策，即该行为需要被划分到边界域以等待进一步的处理。

其中，$[x]$ 表示样本在属性集下的等价类，$P(X|[x])$ 表示将等价类 $[x]$ 分为 $X$ 的概率，在入侵检测的领域则表示为一个网络行为属于入侵行为的概率。

## 3 基于降噪自编码器和三支决策的入侵检测方法

### 3.1 入侵检测算法的整体流程

入侵检测算法的总体流程如图 1 所示，主要包括 3 个模

块：数据预处理、特征提取以及利用三支决策理论进行分类。

入侵检测算法的具体步骤如算法 1 所示。

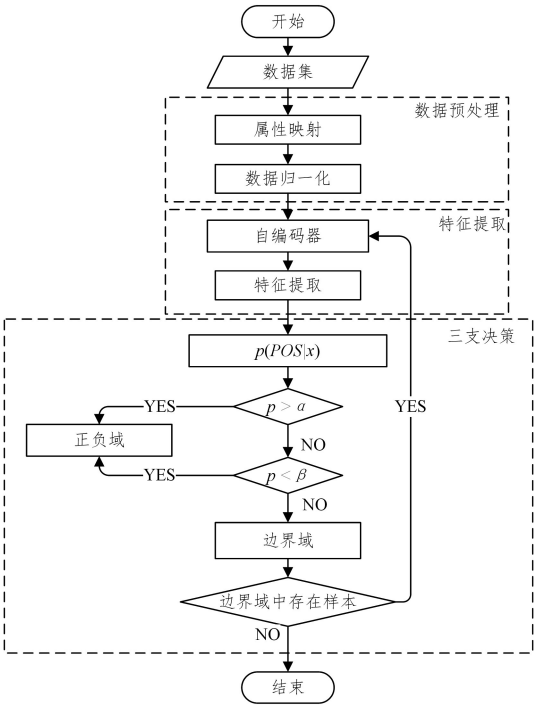


图 1 基于降噪自编码器和三支决策的入侵检测流程图

Fig. 1 Flow chart of intrusion detection based on denoising autoencoder and three-way decisions

**算法 1** 入侵检测算法

输入：网络行为数据 $X=\{x_1,x_2,\cdots,x_n\}$；阈值 $\alpha,\beta$

输出：对网络行为的最终决策

初始化：权重 W；偏置 b；代价函数 $\lambda_{PP},\lambda_{BP},\lambda_{NP},\lambda_{PN},\lambda_{BN},\lambda_{NN}$

Step1 数据预处理；

Step2 使用自编码器进行特征提取；

Step3 利用训练集训练三支决策分类器；

Step4 利用三支决策分类器得到每个网络行为的有关概率 p，即一个样本属于正类的概率；

Step5 根据贝叶斯最小风险原则，利用概率 P(X|[x])与阈值 α 及 β 的比较结果，划分当前的网络行为至正域、负域或者边界域；

Step6 如果边界域的长度大于 0，则重复 Step2－Step5。

### 3.2 数据预处理

首先，字符型的数据需要转化为数值型。其次，为了消除数据因为量纲不同或者数据相差较大而产生的误差，需要对数据进行标准化处理。本文根据式(6)将属性值归一化到同一数量级。

$$x'=\frac{x-\min_i}{\max_i-\min_i} \tag{6}$$

其中，$x$ 是第 $i$ 个属性列的一个值，$\min_i$ 是第 $i$ 个属性列的最小值，$\max_i$ 是第 $i$ 个属性列的最大值。

### 3.3 特征提取

假设输入数据 $X=\{x_1,x_2,\cdots,x_n\}$，重构出的数据为 $X'=\{x_1',x_2',\cdots,x_n'\}$，降噪自编码器的结果就是使重构结果 $X'$ 尽可能地接近原始输入数据 $X$，从而在隐藏层能够提取到代表原始数据的特征。自编码器对重构误差函数进行最小化，从而得到优化后的权重、偏置等网络参数，如式(7)所示：

$$W,W',b,b'=\arg\min_{W,W',b,b'}(J(W,b)) \tag{7}$$

其中，$W,b$ 为编码器的权重和偏置，也是提取低维特征时需要的网络参数，$W',b'$ 分别表示解码器的权重和偏置。

使用梯度下降法对整个网络的权重参数进行更新，求解出目标函数的最优解。自编码器各层的激活函数选用 ELU 激活函数。基于梯度下降法对自编码器进行训练的算法步骤如算法 2 所示。

**算法 2** 自编码器训练算法

输入：数据集 X；迭代次数 T

输出：数据集 X 的低维表示 X′

Step1 $\Delta W^{(l)}=0,\Delta b^{(l)}=0$。

Step2 从 t=1 到 T：

Step2.1 假设需要优化的函数为 J(W,b)，使用 BP 算法计算出 $\nabla_{b^{(l)}}J(W,b)$ 以及 $\nabla_{W^{(l)}}J(W,b)$；

Step2.2 $\Delta b^{(l)}=\Delta b^{(l)}+\nabla_{b^{(l)}}J(W,b)$

Step2.3 $\Delta W^{(l)}=\Delta W^{(l)}+\nabla_{W^{(l)}}J(W,b)$

Step3 更新权重参数。

$$W^{(l)}=W^{(l)}-\alpha\left[\left(\frac{1}{T}\Delta W^{(l)}\right)+\lambda W^{(l)}\right]$$

$$b^{(l)}=b^{(l)}-\alpha\left(\frac{1}{T}\Delta b^{(l)}\right)$$

Step4 根据 Step1－Step3 求得的网络参数获取训练数据和测试数据的低维表示。

### 3.4 基于三支决策理论进行分类

在利用三支决策理论进行分类的过程中，根据限制条件，把整个论域分为 3 个区域(正域、负域和边界域)。在整个决策过程中，都要确定是否对当前的对象做出最终的决策，即确定该对象是属于正域还是负域，或者应该将当前的对象归为边界域。

本文将三支决策理论应用于入侵检测的领域，因此损失函数的选取就要植根于入侵检测的领域中。根据已有经验设置的损失函数如表 2 所列。

表 2 三支决策损失函数的经验值设定

Table 2 Empirical setting of the loss function

| | $P$ | $N$ |
|---|---|---|
| $D_P$ | 0 | 0.7 |
| $D_B$ | 0.3 | 0.3 |
| $D_N$ | 1 | 0 |

损失函数已经确定，则可以根据式(4)以及式(5)得出相关的阈值。

假设样本集为 $X=\{x_1,x_2,\cdots,x_n\}$，样本 $x_i$ 属于正域的概率 $p(POS|x_i)$ 需要被求解出，其中 $i=1,2,\cdots,n$。将 $p$ 值与阈值 $\alpha,\beta$ 进行比较：若 $p<\beta$，则将其分入负域；若 $p>\alpha$，则将其分入正域，否则分入边界域。

边界域中的数据在获得额外的信息后将被重新评估，对于可以分到正域或者负域的样本做出最终的决策，而对于一些样本，最终的决策则需要被再次推迟，即一部分样本需要重新被分到边界域。在所有的对象都被分到正域或负域之前，这个决策过程将一直持续下去[18]。

在自编码器网络中，提取到的特征所包含的相关的鉴别信息会随着训练时间的增加而增加，因此本文构造了一个多粒度的特征结构[13]。而不同粒度的特征将会为训练过程提供不同的信息。随着训练时间的增加，获取到的特征将会提供更多的信息来支撑边界域中样本的划分。当边界域存在时，这个过程将一直持续下去。这是 Yao[19] 所提的序贯三支决策的思想，也是本文在处理边界域时主要考虑的思想。三支决策分类器的训练流程如算法 3 所示。

**算法 3** 基于三支决策理论的分类流程

输入：训练集 Tr；测试集 Te；自编码器特征提取方式 G；阈值 α，β；初始分类器 f；正域 POS=∅，边界域 BND=∅，负域 NEG=∅

输出：正域 POS，负域 NEG

```
Do
  Tr^=G(Tr);
  Te^=G(Te);
  根据 Tr^ 训练模型分类器模型 f;
  由模型 f 得到的测试集中的每个数据属于正类的概率 P=f(Te^);
    for p,te∈P,Te
      if p>α:POS=POS∪te
      else if p<β:
          NEG=NEG∪te
      else:BND=BND∪te
      end if
    end for
      Te=BND;
Until 测试集 Te 为空
```

## 4 实验分析

本节评估所提算法的性能，所有的实验均在 Windows10 PC Intel(R) Core(TM) i5-8250 CPU @1.60GHz 1.80GHz，8.00GB RAM 环境中实现。采用 Python 3.7 实现本文算法。

### 4.1 数据集

本文实验所采用的数据集是 NSL-KDD 数据集。NSL-KDD 数据集由 41 个特征属性和 1 个类属性组成。KDD 数据集包括训练集和测试集两种，总共包含 38 种攻击，其中训练集包含 22 种攻击，而测试集包含训练集中的 20 种攻击，除此之外还包含训练集中没有见过的 16 种攻击类型，因此可以使用测试集测试入侵检测方法在未知攻击上的表现。38 种攻击类型可以分为以下 4 种主要的攻击类型：拒绝服务攻击(Dos)、远程攻击(R2L)、本地用户非法提升权限的攻击(U2R)和网络刺探(Probe)。本文在实验过程中选取 20%的训练集以及全部的测试集，训练集和测试集的分布如表 3 所列。

表 3 数据分布

Table 3 Data distribution

| 类别 | Normal | DoS | Probe | R2L | U2R |
|---|---|---|---|---|---|
| 训练集 | 13449 | 9234 | 2289 | 209 | 11 |
| 测试集 | 9711 | 7458 | 2421 | 2754 | 200 |

在实验的过程中，将所有的异常样本作为一类(正类，即需要被检测出的类别)，将所有的正常样本作为一类(负类)。

### 4.2 实验评估指标

由于数据集中存在分布不均衡的现象，因此单纯凭借准确性来判断算法的优劣并不合适。在入侵检测领域，有两个评判指标比较重要，一个是误报率，另一个是漏报率，而漏报率=1−检出率；精确率反映了被预测为异常的网络行为中有多少是真正的异常行为；F1 分数综合考虑了模型查准率和查全率的计算结果，是反映算法好坏的一个重要指标。因此，本文会选择准确率(Accuracy，ACC)、误报率(False positive rate，FPR)、检出率(Detection Rate，DR)、查准率率 (Precision，PR)与 F1 分数(F1-score，F1)作为评判指标。评价指标的计算式如下：

$$ACC=\frac{TP+TN}{TP+FP+TN+FN} \tag{8}$$

$$DR=\frac{TP}{TP+FN} \tag{9}$$

$$PR=\frac{TP}{TP+FP} \tag{10}$$

$$FPR=\frac{FR}{TN+FP} \tag{11}$$

$$F1=\frac{2TP}{2TP+FP+FN} \tag{12}$$

其中，TP 和 TN 分别表示攻击记录和正常记录已正确分类；FP 表示被误认为攻击的正常记录；FN 表示错误分类为正常记录的攻击记录。

### 4.3 实验过程及结果分析

本文在处理边界域时主要考虑的是，自编码器网络提取到的信息会随着训练时间的增加而增加，即重构数据与原始数据会越来越接近，这种现象反映在相关的数字上就是重构数据与原始数据的均方误差会随着训练时间的增加而减小。图 2 给出了本文所使用的降噪自编码器网络在特征提取的过程中，重构数据与原始数据之间的均方误差。从图中曲线的走势可以看出，均方误差随着训练时间的增加而减小，这表明由编码器得到的低维的特征数据在呈现原始数据的表现上越来越好，即随着训练时间的增多，低维数据能够更好地挖掘出原始数据所包含的信息。

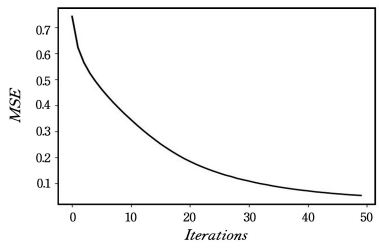


图 2 重构数据与原始数据之间的均方误差

Fig. 2 Mean square error between reconstructed data and original data

在与其他算法进行性能对比时，本文主要考虑自编码器与三支决策两个方面的性能表现，进行了以下 3 个实验：实验 1 在保证使用同样的自编码器进行特征提取的前提下，对比

基于三支决策理论进行分类的表现与基于二支决策进行分类的表现；实验 2 在保证同样使用三支决策进行分类的同时，对比自编码器与不同的特征提取方法在特征提取方面的表现；实验 3 主要对比本文算法与入侵检测领域的其他算法之间的表现。

(1)实验 1

主要探究将三支决策理论引入入侵检测领域所产生的影响，在同样使用自编码器进行特征提取的情况下，对比三支决策与 K 近邻(K-Nearest Neighbor，KNN)[20]、支持向量机(Support Vector Machine，SVM)[21]、随机森林(Random Forest，RF)[22]等传统的基于二支决策进行分类的方法在入侵检测领域的表现。表 4 列出了在 NSL-KDD 测试集上不同分类方法的准确率(ACC)、检出率(DR)、误报率(FPR)、查准率(PR)以及 F1 分数(F1)。

表 4 在 NSL-KDD 测试集上不同分类方法的评估指标值

Table 4 Performance of different classification methods on NSL-KDD test set

| | *ACC* | *DR* | *FPR* | *PR* | *F1* |
|---|---|---|---|---|---|
| DAE-3WD | **0.926** | **0.905** | 0.046 | **0.962** | **0.933** |
| DAE-KNN | 0.804 | 0.689 | **0.044** | 0.954 | 0.800 |
| DAE-RF | 0.764 | 0.623 | 0.049 | 0.944 | 0.750 |
| DAE-SVM | 0.841 | 0.761 | 0.054 | 0.949 | 0.845 |

由表 4 的结果可以看出，本文提出的基于自编码器和三支决策的入侵检测方法 DAE-3WD 在准确率(ACC)、检出率(DR)、F1 分数(F1)、精确率(PR)4 个指标上均优于其他入侵检测方法，甚至在其中的 3 个评价指标上远优于其他分类方法。以上结果表明，把三支决策理论应用于入侵检测产生了积极的影响，基于三支决策理论的入侵检测方法优于传统的基于二支决策的方法。

几种方法的 ROC 曲线图如图 3 所示。从 ROC 曲线图中可以看出，本文提出的 DAE-3WD 方法的曲线更接近(0，1)这个点，且 DAE-3WD 的曲线完全包住了其他方法所产生的 ROC 曲线，佐证了 DAE-3WD 优于其他方法。

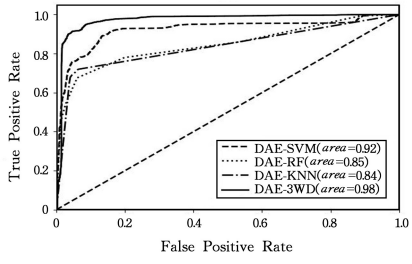


图 3 不同入侵检测方法的 ROC 曲线图

Fig. 3 ROC curves of different classification methods

(2)实验 2

在保证不同方法同样使用三支决策进行分类的同时，探究自编码器与不同的特征提取方法在特征提取方面的表现，对比的特征提取方法选用主成分分析(Principal Component Analysis，PCA)[23]、独立成分分析(Independent Component Correlation Algorithm，ICA)[24]以及奇异值分解(Singular Value Decomposition，SVD)[25]3 种特征提取方法。表 5 列出了使用不同的特征提取方法，但使用同样的基于三支决策理论的分类方法在 NSL-KDD 测试集上的评估指标值。

表 5 不同的特征提取方法在 NSL-KDD 测试集上的评估指标值

Table 5 Performance of intrusion detection methods on NSL-KDD test set

| | *ACC* | *DR* | *FPR* | *PR* | *F1* |
|---|---|---|---|---|---|
| DAE-3WD | 0.926 | **0.905** | 0.046 | **0.962** | **0.933** |
| PCA-3WD | **0.938** | 0.850 | **0.019** | 0.957 | 0.900 |
| ICA-3WD | 0.927 | 0.865 | 0.042 | 0.910 | 0.887 |
| SVD-3WD | 0.928 | 0.842 | 0.029 | 0.935 | 0.886 |

从表 5 中可以看出，基于降噪自编码器的特征提取方法在检出率(DR)、精度(PR)、F1 分数(F1)3 个指标上略优于其他特征提取方法得到的结果，这表明通过降噪自编码器提取到的特征在某些方面略优于基于传统的特征提取方法得到的特征，结果表明将降噪自编码器用于提取网络行为的特征是可取的。

具有不同特征提取方式的入侵方法的 ROC 曲线图如图 4 所示。

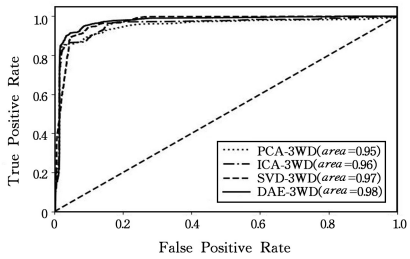


图 4 特征提取方式不同的入侵方法的 ROC 曲线图

Fig. 4 ROC curves of intrusion methods with different feature extraction methods

从图 4 可以看出，使用了自编码器进行特征提取的 DAE-3WD 算法与其他方法相比，在分类的表现上都极其接近，但是 DAE-3WD 还是略优于其他方法。以上结果证明，使用自编码器确实提取到了相对较优的特征。

(3)实验 3

主要对比本文算法与研究人员在入侵检测领域所进行的其他算法研究。文献[26]提出了一种基于独立成分分析 ICA 与深度神经网络 DNN 的入侵检测模型 ICA-DNN；文献[27]提出了一种基于前馈神经网络的模型 SFID；文献[28]提出了一种基于层叠非对称深度自编码器的入侵检测方法 SNADE。表 6 列出了本文提出的方法 DAE-3WD 与其他方法在几个评价指标上的比较结果。

表 6 不同入侵检测方法的比较

Table 6 Comparison of different intrusion detection methods

| | *ACC* | *DR* | *FPR* | *PR* | *F1* |
|---|---|---|---|---|---|
| DAE-3WD | **0.926** | **0.905** | 0.046 | 0.962 | **0.933** |
| ICA-DNN | 0.922 | 0.862 | 0.048 | 0.898 | 0.880 |
| SFID | 0.923 | 0.857 | **0.030** | 0.934 | 0.894 |
| SNADE | 0.916 | 0.885 | 0.043 | **0.964** | 0.923 |

从实验结果可以看出，本文算法在准确率(ACC)、检出率

(DR)以及 F1 分数(F1)上优于其他方法,而在误报率(FPR)以及精度上略逊于其他方法。整体来看,DAE-3WD 是优于其他方法的。

本文提出的 DAE-3WD 方法与对比文献中的几种方法得到的 ROC 曲线图如图 5 所示。

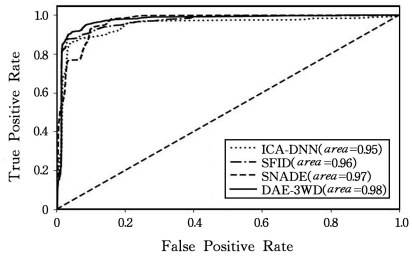


图 5 本文方法与已有方法的 ROC 曲线图

Fig. 5 ROC curves of the method proposed in this article and the existing method

从图 5 可以看出,DAE-3WD 方法得到的曲线相比其他方法得到的曲线更接近左上角,且 DAE-3WD 得到的 AUC 面积大于其他算法得到的 AUC 面积。上述结果表明,本文方法的表现略优于本文中的其他对比方法。

**结束语** 本文在已有研究的基础上提出了一种基于降噪自编码器与三支决策理论的入侵检测算法,通过无监督的降噪自编码器来获取能够表达出原始数据的抽象特征,之后利用三支决策理论进行分类决策。经过仿真实验表明,本文算法相对其他部分算法的表现更优。

本文算法在利用三支决策理论进行分类决策的过程中,对于边界域的处理没有考虑时间成本,例如,当边界域中的样本非常少时,仍然需要重新采用基于三支决策的分类方法对边界域进行处理,这样可能会导致较小的边界域耗费大量时间的情况出现,在未来的研究中可以考虑将时间成本纳入对边界域的处置中。

## 参 考 文 献

[1] GAO N, GAO L, GAO Q, et al. An Intrusion Detection Model Based on Deep Belief Networks[C]// 2014 Second International Conference on Advanced Cloud and Big Data. IEEE, 2014: 247-252.

[2] QIAN Y Y, LI Y Z, YU X Y. Intrusion Detection Method Based on Multi-label and Semi-Supervised Learning [J]. Computer Science, 2015, 42(2): 134-136.

[3] NESPOLI P, PAPAMARTZIVANOS D, MÁRMOL F G, et al. Optimal Countermeasures Selection Against Cyber Attacks: A Comprehensive Survey on Reaction Frameworks[J]. IEEE Communications Surveys & Tutorials, 2017, 20(2): 1361-1396.

[4] DÍAZ-LÓPEZ D, DÓLERA-TORMO G, GÓMEZ-MÁRMOL F, et al. Dynamic Counter-Measures for Risk-Based Access Control Systems: An Evolutive Approach[J]. Future Generation Computer Systems, 2016, 55: 321-335.

[5] LU Y. Research on a New Hybrid Intrusion Detection Algorithm for Cloud Computing[J]. Journal of Chongqing University of Technology (Natural Science), 2020, 34(10): 153-159.

[6] LI Y Z, ZHANG J. Intrusion Detection Algorithm Based on Cluster and Cloud Model[J]. Computer Science, 2015(2): 33.

[7] GAO L Y, TIAN Z S, LI L X, et al. A SVDD-Based Method for WLAN Indoor Passive Intrusion[J]. Journal of Chongqing University of Posts and Telecommunications (Natural Science Edition), 2020, 32(2): 200-209.

[8] ZHANG Y S, JIANG S Y. Research on Network Intrusion Detection Based on Rick Data Mining Tracking Technology[J]. Journal of Chongqing University of Technology (Natural Science), 2019, 33(10): 127-135.

[9] HINTON G E, OSINDERO S, TEH Y W. A Fast Learning Algorithm For Deep Belief Nets[J]. Neural Computation, 2006, 18(7): 1527-1554.

[10] WEI P, LI Y, ZHANG Z, et al. An Optimization Method for Intrusion Detection Classification Model Based on Deep Belief Network[J]. IEEE Access, 2019, 7: 87593-87605.

[11] YANG Y Q, ZHENG K F, WU B, et al. Network Intrusion Detection Based on Supervised Adversarial Variational Auto-Encoder with Regularization [J]. IEEE Access, 2020, 8: 42169-42184.

[12] LI Y Z, ZHANG S P, LI Y, et al. Research on Intrusion Detection Algorithm Based on Deep Learning and Semi-Supervised Clustering[J]. International Journal of Cyber Research and Education (IJCRE), 2020, 2(2): 38-60.

[13] VINCENT P, LAROCHELLE H, BENGIO Y, et al. Extracting and composing robust features with denoising autoencoders [C]// Proceedings of the 25th International Conference on Machine Learning. ACM, 2008: 1096-1103.

[14] YAO Y Y. Three-way decision: an interpretation of rules in rough set theory[C]// International Conference on Rough Sets and Knowledge Technology. Berlin, Heidelberg: Springer, 2009: 642-649.

[15] ZHANG Y B, MIAO D Q, ZHANG Z F. Multi-granularity text sentiment classification model based on three-way decisions[J]. Computer Science, 2017, 44(12): 188-193.

[16] ZHANG L B, LI H X, ZHOU X Z, et al. Sequential three-way decision based on multi-granular autoencoder features[J]. Information Sciences, 2020, 507: 630-643.

[17] LIU D, LIANG D C. Generalized three-way decisions and special three-way decisions[J]. Journal of Frontiers of Computer Science and Technology, 2017, 11(3): 502-510.

[18] MALDONADO S, PETERS G, WEBER R. Credit scoring using three-way decisions with probabilistic rough sets[J]. Information Sciences, 2020, 507: 700-714.

[19] YAO Y Y. Granular computing and sequential three-way decisions[C]// International Conference on Rough Sets and Knowledge Technology. Berlin, Heidelberg: Springer, 2013: 16-27.

[20] SENTHILNAYAKI B, VENKATALAKSHMI K, KANNAN A. Intrusion Detection System using Fuzzy Rough Set Feature Selection and Modified KNN Classifier[J]. International Arab Journal of Information Technology, 2019, 16(4): 746-753.

[21] AL-QATF M,LASHENG Y,AL-HABIB M,et al. Deep learning approach combining sparse autoencoder with SVM for network intrusion detection[J]. IEEE Access, 2018, 6: 52843-52856.

[22] REN J D,LIU X Q,WANG Q,et al. A Multi-Level Intrusion Detection Method Based on KNN Outlier Detection and Random Forests[J]. Journal of Computer Research and Development, 2019,56(3):566-575.

[23] DING Y,LI Y Z. Research on Intrusion Detection Algorithm Based on PCA and Semi-Supervised Clustering[J]. Journal of Shandong University (Engineering Science), 2012,42(5): 41-46.

[24] DU Y,ZHANG Y D,LI M H,et al. Improved Fast ICA algorithm for data optimization processing in intrusion detection[J]. Journal on Communications,2016,37(1):42-48.

[25] GHASEMI J,ESMAILY J. Intrusion Detection Systems Using a Hybrid SVD-Based Feature Extraction Method[J]. International Journal of Security and Networks,2017,12(4):230-240.

[26] LIU J H,MAO S P,FU X M. Intrusion Detection Model Based on ICA Algorithm and Deep Neural Network[J]. Netinfo Security,2019,19(3):1-10.

[27] FENG W Y,GUO X B,HE Y Y,et al. Intrusion Detection Model Based on Feedforward Neural Network[J]. Netinfo Security,2019,19(9):101-105.

[28] SHONE N,NGOC T N,PHAI V D,et al. A Deep Learning Approach to Network Intrusion Detection[J]. IEEE Transactions on Emerging Topics in Computational Intelligence,2018,2(1): 41-50.

**ZHANG Shi-peng**, born in 1994, postgraduate. His main research intrests include computer network security and machine learning.

**LI Yong-zhong**, born in 1961, M.S, professor, M.S supervisor. His main research interests include computer network security and information security, intelligent information processing, and application of embedded system.