计算机应用研究
Application Research of Computers
ISSN 1001-3695,CN 51-1196/TP

《计算机应用研究》网络首发论文

题目：基于 ADBN 的入侵检测方法
作者：江泽涛，周谭盛子
DOI：10.19734/j.issn.1001-3695.2019.03.0149
收稿日期：2019-03-27
网络首发日期：2019-08-29
引用格式：江泽涛，周谭盛子．基于 ADBN 的入侵检测方法[J/OL]．计算机应用研究．
https://doi.org/10.19734/j.issn.1001-3695.2019.03.0149

基于 ADBN 的入侵检测方法 *

江泽涛 [a, b]，周谭盛子 [a, b]

(桂林电子科技大学 a. 广西图像图形处理智能处理高校重点实验室；b. 广西可信软件重点实验室，广西 桂林 541004)

摘 要：当下大多数入侵检测算法无法在入侵检测率和误报率之间取得较好的平衡的弊端，为了有效地避免此类问题，提出了一种基于非对称深度信念网络的入侵检测方法。该方法首先通过训练深度信念网络初始化ADBN(asymmetric deep belief network)模型中编码器部分的参数，利用正态分布初始化解码器部分的参数。然后通过计算重构误差来调优 ADBN 模型的参数，使模型能获取原始数据的最优低维表征。最后以编码器得到的数据作为分类器的输入数据并对其进行检测，采用 ADBN 模型可以提取出更有利于分类的特征且能够在模型初始化阶段节省更多的测试时间。实验结果表明，该方法可以达到更好的检测性能，对小类别样本也达到了较好的检测准确率。

关键词：入侵检测；特征提取；非对称深度信念网络；编码器；解码器
中图分类号：TP309 doi: 10.19734/j.issn.1001-3695.2019.03.0149

Intrusion detection method based on ADBN

Jiang Zetao[a, b], Zhou Tanshengzi[a, b]

(*a. The Key Laboratory of Image & Graphic Intelligent Processing of Higher Education in Guangxi*, *b. The Key Laboratory of Dependable Software of Guangxi*, *Guilin University of Electronic Technology*, *Guilin Guangxi* 541004, *China*)

Abstract: At present, most intrusion detection algorithm can't achieve a good balance between intrusion detection rate and false positive rate, in order to effectively avoid such problems, this paper proposes an intrusion detection method based on ADBN. The method first initializes the parameters of the encoder part in the ADBN(Asymmetric Deep Belief Network) model by training the deep belief network, and initializes the parameters of the decoder part by using the normal distribution. Then the parameters of the asymmetric deep belief network model are tuned by calculating the reconstruction error, so that the model can obtain the optimal low-dimensional representation of the original data. Finally, the data obtained by the encoder is used as input data of the classifier and detected. The ADBN model can extract features that are more conducive to classification and save more test time in the model initialization phase. The experimental results show that the method can achieve better detection performance and achieve better detection accuracy for small categories of samples.

Key words: intrusion detection; feature extraction; asymmetric deep belief network; encoder; decoder

0 引言

随着互联网技术和相关应用的迅速发展，人们的工作及生活等各方面越来越多的依托于互联网。在享受网络带来便利的同时，如何保障计算机网络的安全性也成为人们日益关注的重点。入侵检测系统作为信息安全综合防御系统的重要组成部分，目前已被广泛应用于各企事业单位。深度学习作为机器学习的一个分支，通过模拟人脑的机制来解释数据，在图像处理、语音识别等领域都有广泛的应用，也得到了不少入侵检测领域学者的青睐。Yin 等人[1]将循环神经网络(recurrent neural networks, RNN)应用到入侵检测中，研究了隐藏层节点与学习率对算法性能的影响，并将其与传统的机器学习方法做了对比。Sheraz 等人[2]对基于异常的深度神经网络框架进行了研究，包括卷积神经网络(convolutional neural networks, CNN)、长短期记忆网络(long short-term memory, LSTM)和自编码器(autoencoders, AE)，相较于传统的机器学习方法，如 K 近邻(k nearest neighbor, KNN)、决策树(decision tree, DT)，深度学习的方法能达到更高的入侵检测率。Sheikhan 等构建了三层 RNN[3]，该模型可以提高算法的检测精度，特别是 R2L 攻击类型，但是与传统的机器学习相比，该模型在降低误报率方面没有优势。Shone 等人[4]提出了一种非对称深度自编码(nonsymmetric deep autonecoder, NDAE)，通过无监督学习进行特征映射，最终以随机森林作为分类器，该方法在大样本数据集上表现出了很高的检测率，同时也大大缩减了训练时间。高妮等人[5]提出了一种基于自编码网络特征降维方法，该方法实现了高维空间和低维空间之间的双向映射，通过支持向量机(support vector machine, SVM)对获得的低维特征向量进行入侵识别。

神经网络的兴起有效的解决了数据的非线性问题，提高了模型的泛化能力，但传统的神经网络存在以下问题：a)随着神经元个数及隐藏层层数的增多导致模型的参数数量迅速增长，增加训练时间；b)随着神经网络层数的增多，采用梯度下降的方法容易陷入局部最优解，在反向传播的过程中也容易导致梯度弥散或梯度饱和的现象；c)随着神经网络层数的增加，模型参数也会增加，因此需要更多的标签数据用于训练模型，但现实生活中有时很难具备大量的标签数据，所以传统的神经网络对小样本数据往往达不到理想的检测效果。

针对上述情况，本文从研究特征间相关性的角度出发，

收稿日期：2019-03-27；**修回日期：**2019-05-20 **基金项目：**国家自然科学基金资助项目(61572147，61762066，61876049)、广西科技计划项目(AC16380108)、广西图像图形智能处理重点实验项目(GIIP201701，GIIP201801，GIIP201802，GIIP201803)、广西研究生教育创新计划资助项目(2018YJCX46资助)、江西省自然科学基金资助项目(20171BAB212015)

作者简介：江泽涛(1961-)，男，江西南昌人，教授，博士，主要研究方向为信息安全、图像处理(zetaojiang@126.com)；周谭盛子(1993-)，女，安徽宣城人，硕士研究生，主要研究方向为信息安全.

提出了基于非对称深度信念网络(asymmetric deep belief network, ADBN)的入侵检测方法。该方法首先逐层训练堆叠的受限玻尔兹曼机(restricted boltzman machine, RBM) ，将训练后获得的参数作为编码器的参数，然后通过正态分布初始化解码器各参数并对原始数据进行还原，最后利用重构误差对网络进行调参，最终获得原始数据的高阶特征并将其作为分类器的输入数据，从而达到特征提取的目的。实验结果表明，该方法可以达到较好的在检测率和误报率之间取得平衡并且具有一定的鲁棒性。

1　相关理论介绍

1.1　自编码网络

自编码网络作为深度学习的一种，是典型的无监督学习算法。自编码网络包含输入层、隐藏层和输出层，输入层与隐藏层的连接构成编码器，隐藏层与输出层的连接构成解码器。自编码网络结构图如图 1 所示。从本质上来说，编码器就是通过将输入层的原始数据进行编码，得到能表征原始数据的低维数据，并通过隐藏层输出；解码器就是将隐藏层输出的低维数据作为输入数据并进行解码，得到原始数据的同维度表征。重构误差是自编码网络的目标函数，自编码网络的目的就是通过反向传播调节编码器与解码器的权重与偏置将误差降到最低。对于整个网络而言，最主要的是通过最小化重构误差得到隐藏层的低维数据表征。

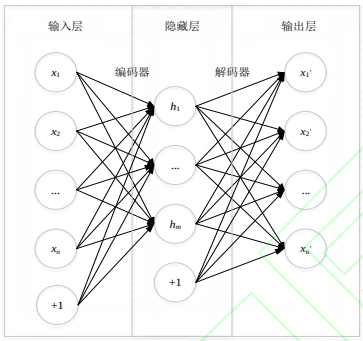

图 1　自编码网络结构图

Fig. 1　Self-encoding network structure diagram

自编码网络的重构误差定义如下：

$$L_{W,b}(x_k,x_k^{'})=\frac{1}{2}\sum_{i=1}^{p_l}(x_{ki}-x_{ki}^{'})^2 \tag{1}$$

其中：x_k 表示第 k 个样本；x_{ki} 表示第 k 个样本的第 i 维特征属性；$x_k^{'}$ 表示第 k 个样本经过解码器后在输出层的输出；$x_{ki}^{'}$ 表示第 k 个样本经过解码器后输出层第 i 个神经元；p_l 分别表示第 l 个自编码网络输入层神经元的个数。

1.2　深度信念网络

Hinton[6]于 2006 年提出深度信念网络(DBN)，自此，深度学习的研究进入了一个新的纪元。

1.2.1 受限玻尔兹曼机

DBN 由多个 RBM 堆叠而成，RBM 是深度信念网络的主要构成部分。RBM 的理论模型是一个二部图，包含可视层 v 和隐藏层 h，层内无连接，层间全连接。每个可见层节点 v_i 存在偏置 a_i，每个隐藏层节点 h_j 存在偏置 b_j，层间的连接权重用矩阵 W 表示。

假设每个节点的值均为 0 或 1，对于任意的 v_i,h_j，均有 $v_i\in\{0,1\},h_j\in\{0,1\}$。当给定可见层各节点变量时，隐藏层各节点之间是相互独立的，因此，可以得到第 j 个隐藏层节点的激活概率定义如下：

$$P(h_j=1|v,\theta)=\sigma(b_j+\sum_i W_{ij}v_i) \tag{2}$$

其中：$\theta=\{\mathrm{W,a,b}\}$ 是需要调优的学习参数。

同理可得，第 i 个可见层节点的激活概率定义如下：

$$P(v_i=1|h,\theta)=\sigma(a_i+\sum_j W_{ij}h_j) \tag{3}$$

通常 RBM 采用对比散度(contrastive divergence, CD)算法对数据进行降维，通过 k 步的 Gibbs 采样更新 θ，使得通过隐藏层重构的可见层的另一种状态与可见层节点一致，此时称隐藏层 h 为可见层 v 的另一种表达方式，达到特征提取的效果。Bengio[7]验证了当 $k=1$ 时，通过一步 Gibbs 采样就能很好的拟合输入层各节点的状态。

1.2.2 基于 DBN 的入侵检测系统

DBN 模型的提出避免了很多存在于传统神经网络中的问题。首先 DBN 是由多个 RBM 堆叠而成的，通过贪婪的方式逐层训练 RBM，每一层 RBM 的输出都是当前最优的，不需要通过训练整个 DBN 模型获得最优解。此外，DBN 包括预训练和参数微调两个过程，即分为无监督学习和有监督学习，其结构图如图 2 所示。

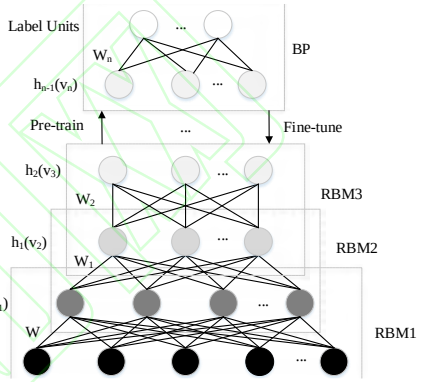

图 2　DBN 结构图

Fig. 2　DBN structure diagram

自提出 DBN 后，相关研究人员纷纷将 DBN 与入侵检测技术相结合。逯玉婧[8]提出了一种自适应深度信念网络，克服了传统的深度信念网络采用对比散度算法进行采样，最终获得的参数结果易陷入局部最优的问题。此外，还提出了面向不平衡数据的混合入侵检测模型，有效提高了小类型攻击的检测率。Gao 等人[9]研究了训练迭代次数、第一个隐藏层的节点个数、输出层节点个数等参数对 DBN 性能的影响，同时将 DBN 与自组织映射(self-organizing maps, SOM)、神经网络(neural network, NN)做了对比，突出了 DBN 在入侵准确率、检测率及误报率方面的优势。陈虹等人[10]为了解决现有的入侵检测方法对未知攻击类型检测率较低的问题，提出了一种基于优化数据处理的 DBN 模型的入侵检测方法，该方法首先对字符类型的数据进行概率质量函数(probability mass function, PMF)编码，对数值类型的数据进行归一化，然后将处理过的数据作为 DBN 的输入数据，最后得出检测结果，该方法具有良好的自适应性，提高了 DBN 的分类精度以及未知攻击类型的检测率。如汪洋等人[11]提出了一种基于深度序列加权核极限学习机的入侵检测算法(DBN-WOS-KELM)，该算法利用 DBN 做特征提取，并结合加权序列核极限学习机(kernel extreme learning machine, KELM)对低维度表示的数据进行入侵识别，实验表明该算法能有效提高小样本攻击类型的检测率。

2　基于 ADBN 的入侵检测方法

2.1　ADBN 模型

现实生活中缺乏大量的训练入侵检测算法的带标签数据，无监督学习作为机器学习中的一种方法，将会为入侵检测的后续研究贡献大部分的力量。自编码网络就是典型的无监督学习算法，通过学习样本特征属性的规律可以对高维数据进

行压缩，获得原始数据的低维表征。但是传统的自编码网络的训练时间较长，且输入层和隐藏层、隐藏层和输出层之间的连接权重以及输入层和隐藏层的偏置的初始化一般都是随机赋值，这对整个网络的训练时间及算法检测率等方面都存在一定的影响。为了克服以上问题，本文从特征提取的角度出发，提出了 ADBN 模型。ADBN 模型如图 3 所示。

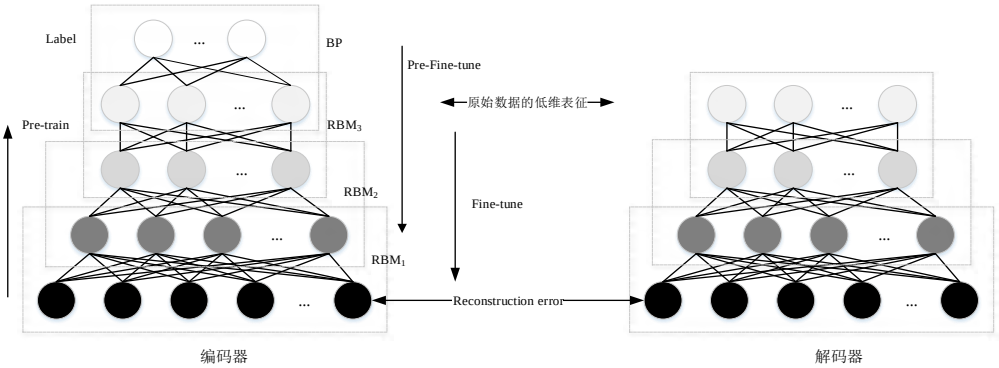

图 3　ADBN 模型图

Fig. 3　ADBN model diagram

该方法首先将堆叠的RBM作为自编码网络中的编码器，通过逐层训练 RBM 对原始数据进行降维操作，同时将训练得到的参数作为 ADBN 模型中编码器部分的参数初始值，并通过解码器还原数据，随后计算输入数据与重构后的输出数据的重构误差，并以此为目标函数利用反向传播调节 ADBN 模型的参数，最终获得原始数据的最优高阶特征，达到特征提取的目的。

常用的关于神经网络参数初始化的方法有全为 0、随机数初始化以及正态分布初始化。全为 0 的参数初始化方法会使隐藏层的每个节点输出相同的数，同时对损失函数求偏导也会得到相同的结果，致使每次对参数进行更新的数值也是一样的，这会导致网络失去体征提取的能力，无法达到收敛状态，所以全为 0 的参数初始化方法只适用于没有隐藏层的神经网络。使用随机数初始化网络参数时，当随机数很大时，会造成隐藏层的节点值很大或者很小，在选取 Sigmoid 函数作为激活函数时，梯度的更新趋于 0，从而出现梯度消失的现象。出于对以上因素的考虑，本文从节省神经网络训练时间的角度出发，选择用正态分布初始化解码器部分的参数，正态分布可以在一定程度上避免梯度消失的问题，在使用正态分布初始化参数时网络隐藏层节点的值不会出现很小的情况，从而也不会导致神经元学习饱和。

此外，虽然增加神经网络的层数可以降低网络误差，提高算法精度，但是网络层数的增加也会导致参数剧增、训练时间过长以及过拟合的现象，所以本文提出的 ADBN 模型只对模型中的编码器部分的参数进行训练，解码器部分的参数采用服从正态分布的随机值进行初始化，即编码器部分与解码器部分采用不同的参数初始化方法。

ADBN 模型的算法主要包括预训练、参数微调。预训练包括采用 CD 算法逐层训练 RBM 和利用标签微调参数，此处对参数的微调是有监督学习；参数微调是指通过重构误差对 ADBN 模型中的参数进行微调，此处对参数的微调是无监督学习。现将各部分算法的主要步骤做如下描述：

算法 1　采用 CD 算法逐层训练 RBM

输入：可见层各节点状态 v_0，隐藏层节点个数 n，学习率 η，RBM 层数 r。

输出：可见层的偏置 a，隐藏层的偏置 b，可见层与隐藏层之间的连接权重 W。

a) 将 a、b、W 初始化为 0。

for $j=1,2,...,n$ (对所有隐藏层节点)：

　利用式(2)计算 $P(h_j=1|v_0,\theta)$

　从条件分布 $P(h_j=1|v_0,\theta)$ 中抽取 $h_j\in\{0,1\}$

End for

for $i=1,2,...,m$ (对所有可见层节点)：

　利用式(3)计算 $P(v_i^*=1|h,\theta)$

　从条件分布 $P(v_i^*=1|h,\theta)$ 中抽取 $v_i^*\in\{0,1\}$

End for

for $j=1,2,...,n$ (对所有隐藏层节点)：

　利用式(2)计算 $P(h_j^*=1|v_i^*,\theta)$

End for

按下列各式更新各参数：

$$W\leftarrow W+\eta[P(h=1|v)v^T-P(h^*=1|v^*)v^{*T}] \tag{4}$$

$$a\leftarrow a+\eta(v-v^*) \tag{5}$$

$$b=b+\eta(P(h=1|v)-P(h^*=1|v^*)) \tag{6}$$

b) 将得到的 h^* 作为下一个 RBM 的输入层并重复步骤 a)，直到堆叠层数满足条件为止。

算法 2　利用标签微调参数

输入：由算法 1 步骤 b)获得的 $\theta=(W,b)$，样本标签 $y=y_i(i=1,2,...,n)$，目标函数阈值 λ。

输出：$\theta'(W',b')$。

a)利用下列式式子计算目标函数：

$$loss=-\frac{1}{n}[y\ln p+(1-y)\ln(1-p)] \tag{7}$$

其中：p 为检测结果。

b)若 $loss<\lambda$，则输出 $\theta'(W',b')$，否则执行步骤 c)。

c)通过对步骤 b)获得的输出层的目标函数进行式(8)的链式求导并更新权重 W：

$$\begin{aligned}W&=W+\eta\frac{\partial loss}{\partial W}=W+\\&\eta\cdot\frac{\partial loss}{\partial\sigma(z)}\frac{\partial\sigma(z)}{\partial z}\frac{\partial z}{\partial W}=W+\\&\eta\cdot\frac{\sigma(z)-y}{\sigma(z)(1-\sigma(z))}\cdot\sigma(z)(1-\sigma(z))\cdot x=\\&W+\eta\cdot x\cdot(\sigma(z)-y)\end{aligned} \tag{8}$$

通过对步骤 b)获得的输出层的目标函数进行式(9)的链式求导并更新权重 b：

$$\begin{aligned}b&=b+\eta\frac{\partial loss}{\partial b}\\&=b+\eta(\sigma(z)-y)\end{aligned} \tag{9}$$

其中: $z=W\cdot x$。

算法 3　参数微调

输入：由算法 2 获得的参数 $\theta'=(W',b')$ (用于初始化自编码网络中编码器的参数)，堆叠的 RBM 最后一层的输出 h'，服从正态分布的数值 $\theta^*=(W^*,b^*)$ (用于初始化自编码网络中编码器的参数)，训练集样本的特征属性 $S=x_i(i=1,2,...,n)$，其中 n 为样本个数。

输出：ADBN 模型编码器部分隐藏层的偏置 b，可见层与隐藏层之间的连接权重 W。

a)对于训练集中的所有样本 $i(i=1,2,...,n)$，根据式(10)计算重构输出 x_i'：

$$x_k^{'} = g_{\theta'}(h_k) = Relu(W_2^l * h_k + b_2^l) = Relu(\sum_{i=1}^{p_2^l}(W_{ji}^l * h_{ki} + b_2^l)) \tag{10}$$

其中: h_k 表示第 k 个样本经过编码器后在隐藏层的输出，h_{ki} 表示第 k 个样本经过编码器后隐藏层第 i 个神经元的输出；(W_2^l, b_2^l) 表示第 l 个自编码网络解码器部分的权重与偏置；W_{ji}^l 表示第 l 个自编码器前一层第 i 个神经元和后一层第 j 个神经元的连接权重；p_2^l 分别表示第 l 个自编码网络隐藏层神经元的个数。

b)利用式(1)计算重构误差。

c)对于第 l 层的第 $j(j=1,2,3,...,m)$ 个神经元，根据式(11)计算残差值 δ_j^l：

$$\begin{aligned}\delta_j^l &= \frac{\partial}{\partial v_j^l} loss = -(x_j - h_j^l)\sigma'(v_j^l) = \\ &-(x_j - h_j^l)\sigma(x_j^l)(1-\sigma(v_j^l)) = -(x_j - h_j^l)h_j^l(1-h_j^l)\end{aligned} \tag{11}$$

其中，v_j^l 表示第 l 层的第 j 个神经元的输入值，h_j^l 表示第 l 层的第 j 个神经元的输出值。

d) 通过对步骤(3)获得的输出层的残差值进行式(12)的链式求导可得到前 $l(l=l-1,l-2,...,2)$ 层的残差值：

$$\delta_j^l = \left(\sum_{i=1}^{p^{l+1}} W_{ij}^l \delta_i^l\right) Relu(v_j^l) = \sum_{i=1}^{p^{l+1}} W_{ij}^l \delta_i^{l+1} \tag{12}$$

其中：p^{l+1} 表示第 l +1 层，W_{ij}^l 表示第 l 层的第 j 个神经元的连接权重。

e) 利用步骤(4)获得的第 l 层至第 2 层中每个神经元的残差值，通过下列式子更新网络的权重和偏置：

$$W = W + \eta \frac{\partial}{\partial W_{ij}^l} loss = W + \eta \delta_j^{l+1} \tag{13}$$

$$b = b + \eta \frac{\partial}{\partial b_j^l} loss = b + \eta \delta_j^{l+1} \tag{14}$$

2.2　基于 ADBN 的入侵检测方法

本文在 ADBN 模型的基础上提出了一种基于 ADBN 的入侵检测方法，该方法包括训练阶段和测试阶段，训练阶段主要包括基于 ADBN 的特征提取模块，测试阶段即为入侵检测模块。其流程图如图 4 所示。

a)训练阶段。将原始数据经过预处理模块处理之后作为 ADBN 模型的输入数据，通过预训练堆叠的 RBM 初始化该模块编码器部分的参数，同时使用服从正态分布的数值初始化解码器部分，此时得到 ADBN 模型编码器和解码器的所有参数，利用解码器对输入数据进行重构。利用式(7)计算目标函数，判断目标函数的值是否小于阈值，若是，则表明获取到了原始数据的最优高阶特征，此时进入侵检测模块；否则根据 2.1 节算法 2 进行参数微调，直到重构误差小于阈值为止。

b)入侵检测模块。保存基于 ADBN 的特征提取模块得到的编码器部分的参数，将测试集数据作为输入数据，通过编码器对输入数据进行特征提取，得到数据的低维表示，并采用 Softmax 分类器进行分类并得出检测结果。

3　实验与分析

3.1　实验设置

3.1.1 实验数据

KDD CUP 99 数据集是由麻省理工学院林肯实验室建立的模拟美国空军局域网的网络流量测试数据集，是目前比较权威的用于入侵检测的数据集。该数据集中的每条数据记录由 41 维特征和一个类标签组成。其中，41 维特征包括 9 个 TCP 连接基本特征、13 个 TCP 连接的内容特征、9 个基于时间的网络流量统计特征和 10 个基于主机的网络流量统计特征；该数据集主要包括四种类型的攻击：Dos(拒绝服务攻击)、Probe(端口监视或扫描)、R2L(远程主机的未授权访问)及 U2R(未授权的本地超级用户特权访问)。本文采用 10%的 KDD CUP 99 数据集作为训练集，该数据集包含了 494021 条数据记录，其攻击类型如表 1 所示。为了检验模型的泛化性能，采用 KDD CUP 99 Corrected 数据集作为测试集，该数据集包含了 311029 条数据记录，其中包括了 17 种新型攻击类型，其攻击类型如表 2 所示。

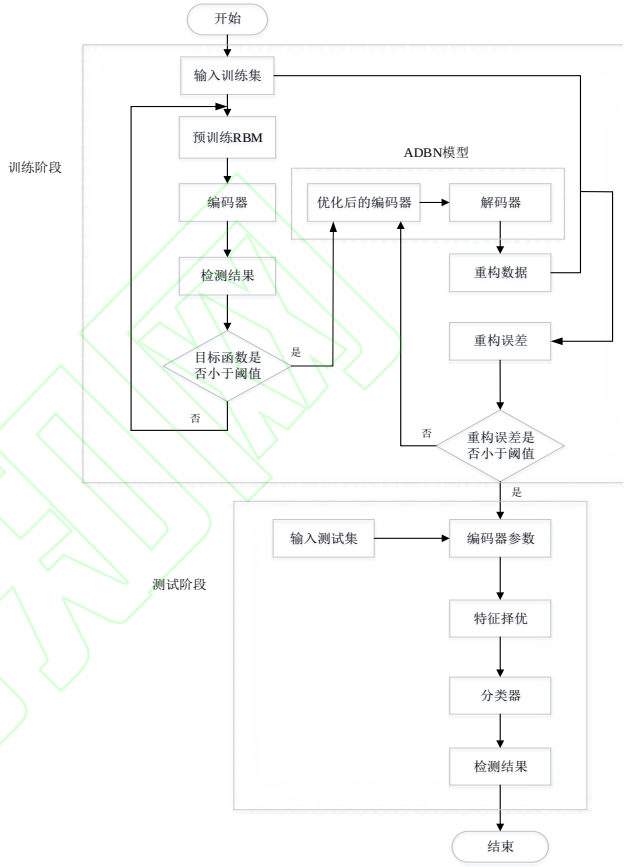

图 4　基于 ADBN 的入侵检测方法流程图

Fig. 4　Flow chart of intrusion detection method based on ADBN

表 1 10%的 KDD CUP'99 数据集攻击类型表

Tab. 1　10% KDD CUP'99 dataset attack type table

类别编号	攻击类别	攻击类型
1	Dos	back, land, neptune, pod, smurf, teardrop.
2	Probe	ipsweep, nmap, postsweep, satan.
3	R2L	ftp_write, guess_passwd, imap, multihop, phf, spy, warezclient, warezmaster.
4	U2R	buffer_overflow, loadmodule, perl, rootkit.

表 2　KDD CUP'99 Corrected 数据集攻击类型表

Tab. 2　KDD CUP'99 corrected dataset attack type table

类型编号	攻击类别	攻击类型
1	Dos	apache2, back, land, mailbomb, neptune, pod, processtable, smurf, teardrop, udpstorm.
2	Probe	ipsweep, mscan, nmap, postsweep, saint, satan.
3	R2L	ftp_write, guess_passwd, imap, multihop, named, phf, sendmail, snmpgetattack, snmpguess, warezmaster, worm, xlock, xsnoop.
4	U2R	buffer_overflow, httptunnel, loadmodule, perl, ps, rootkit, sqlattack, xterm.

3.1.2 数据预处理

KDD CUP 99 数据集中每条连接记录的 41 个特征属性有

38 个数字型特征属性和 3 个字符型特征属性构成，数据预处理首先要将字符型特征属性转换成相应的数字，然后对数据进行归一化处理，归一化公式如下所示。

$$x_{ij}^{'}=\frac{x_{ij}-x_i^{min}}{x_i^{max^{min}}} \tag{15}$$

其中：x_{ij} 为第 j 个样本第 i 维特征属性值，x_i^{min} 为第 i 维特征属性值的最小值，x_i^{max} 为第 i 维特征属性值的最大值。

为了避免原始 KDD CUP 99 数据集中的重复样本对网络参数的影响，实验采用经过去重处理的 KDD CUP 99 数据集。去重后训练集和测试集中各类别的样本数目如表 3 所示。

表 3　去重后训练集和测试集中各类别的样本数目

Tab. 3　Number of samples in each training category and in the test set after deduplication

类别编号	去重后训练集中各类别的样本数目	去重后测试集中各类别的样本数目
1	87832	47913
2	54572	23568
3	2131	2682
4	999	2913
5	52	215
总数	145586	77291

3.1.3 实验环境

本实验在 Intel CPU 3.50 GHz、8GB 内存、64 位硬件环境和 Windows 8 操作系统下，使用 Python 3.6.0 进行编码实现。

3.2　实验结果与分析

衡量一个入侵检测系统的好坏的关键在于是否具有高检测率和低误报率。因此，为了评估本文提出的入侵检测模型的性能，选取了准确率(accuracy, ACC)、精确率(precision, PRE)、检测率(detection rate, DR)、误报率(false alarm rate, FAR)四个指标。计算公式如下：

$$ACC=\frac{TP+TN}{TP+TN+FP+FN} \tag{16}$$

$$PRE=\frac{TP}{TP+FP} \tag{17}$$

$$DR=\frac{TP}{TP+FN} \tag{18}$$

$$FAR=\frac{FP}{FP+TN} \tag{19}$$

其中：TP(true positive)为正确识别的正样本数，TN(true negative)为正确识别的负样本数，FP(false positive)为错误识别的正样本数，FN(false negative)为错误识别的负样本数。

为了研究 ADBN 模型中编码器部分隐藏层个数和顶层 RBM 输出层神经元个数对算法检测率的影响，本文在如下表 4 所示的模型参数下做了对比实验并对实验结果作了分析。

表 4 ADBN 模型参数

Tab. 4　The parameters of ADBN model

模型	batch_sizes	epochs	learning_rate
RBM	500	300	0.05
DBN	500	400	0.1
ADBN	200	400	0.08

表 4 中各参数均由经验值获得，整个 ADBN 模型训练包括以下三部分：CD 算法训练单层的 RBM 模型；贪婪的逐层训练堆叠的 RBM，即 DBN 模型；利用重构误差微调 ADBN 模型中的参数。表 4 中的 batch_sizes 表示每次取的数据样本个数，epochs 表示整个数据集样本被轮的次数，learning_rate 表示各模型的学习率。

实验 1　隐藏层个数对算法检测率的影响如图 5 所示。

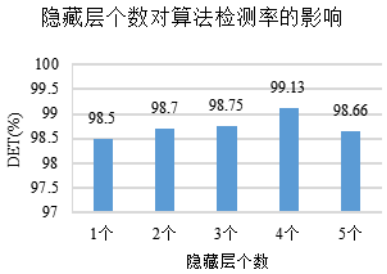

图 5　隐藏层个数对算法检测率的影响

Fig. 5　The effect of the number of hidden layers on the detection rate of the algorithm

图 5 中，1 层表示 ADBN 模型中编码器部分包含 1 个隐藏层，即整个编码器部分的神经元依次为：41-23-5；2 层表示 ADBN 模型中编码器部分包含 2 个隐藏层，即整个编码器部分的神经元依次为：41-28-14-5；3 层表示 ADBN 模型中编码器部分包含 3 个隐藏层，即整个编码器部分的神经元依次为：41-30-20-10-5；4 层表示 ADBN 模型中编码器部分包含 4 个隐藏层，即整个编码器部分的神经元依次为：41-35-26-19-12-5。从实验结果可以看出，模型在包含 4 个隐藏层时算法的检测率达到最高。当隐藏层个数小于 4 时，随着隐藏层个数的增加，算法的检测率也会提高，这是由于模型在训练阶段可以通过更多的隐藏层提取能够表征原始数据的特征。当隐藏层个数大于 4 时，会导致模型出现拟合过度现象，从而降低算法的检测率。

实验 2　该实验采用 4 个隐藏层，分析顶层 RBM 输出层神经元个数对算法检测率的影响，结果如图 6 所示。

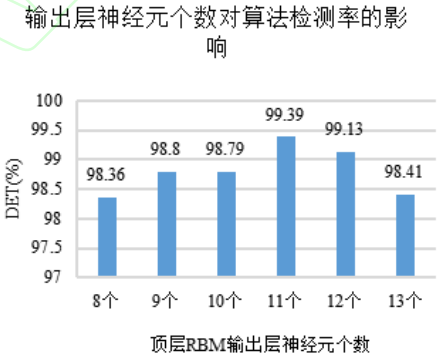

图 6　输出层神经元个数对算法检测率的影响

Fig. 6　Influence of the number of neurons in the output layer on the detection rate of the algorithm

改变顶层 RBM 输出层神经元的个数，图 6 中，8 个表示 ADBN 模型中整个编码器部分的神经元依次为：41-35-26-19-8-5；9 个表示 ADBN 模型中整个编码器部分的神经元依次为：41-35-26-19-9-5；10 个表示 ADBN 模型中整个编码器部分的神经元依次为：41-35-26-19-10-5；11 个表示 ADBN 模型中整个编码器部分的神经元依次为：41-35-26-19-11-5；12 个表示 ADBN 模型中整个编码器部分的神经元依次为：41-35-26-19-12-5；13 个表示 ADBN 模型中整个编码器部分的神经元依次为：41-35-26-19-13-5。从实验结果可以看出，当顶层 RBM 输出层神经元个数为 11 时，算法的检测率最好。当神经元个数小于 11 时，模型不能保留更多的对分类有贡献的特征。当神经元个数大于 11 时，模型的参数会增多，造成结果过拟合的现象，同时对算法的检测率有明显的负面影响。

由实验 1、2 的实验结果可以看出，本章提出的 ADBN 模型具有更好的鲁棒性，算法的检测率均能保持在 98%以上。

实验 3　将本章提出的入侵检测方法与其他方法在准确率和精确率方面进行比较，实验结果如表 5 所示。

表 5　算法准确率与精确率对比表

Tab. 5　Algorithm accuracy and accuracy comparison table

分类模型	性能指标	分类器性能(%)				
		normal	Dos	Probe	R2L	U2R
DBN	ACC	**99.49**	99.65	14.19	89.25	7.14
	PRE	94.51	98.74	86.66	**100**	**38.46**
文献[4]	ACC	99.49	**99.79**	**98.74**	9.31	7.14
	PRE	**100**	**100**	**100**	**100**	0
文献[9]	ACC	96.77	96.41	93.85	92.17	86.54
	PRE	90.26	98.35	85.18	90.43	36.75
本章	ACC	92.58	97.26	97.58	**96.22**	**99.59**
	PRE	89.53	97.31	85.5	14.29	5.22

由表 5 可以看出，传统的 DBN 网络模型对 Probe 类攻击和 U2L 类攻击的检测准确率都较低，分类精度也不高。文献[4]提出的 NDAE 相较于传统的深度信念网络大大的缩短了网络模型训练时间，且对于 Probe 类攻击检测性能有明显的优势，对除了 U2R 类攻击外的类型都达到了百分百的检测精度，但是该算法对小样本攻击类型的检测准确率较大多数传统的机器学习方法还有一定差距。文献[9]通过训练 RBM 得到前向传播的网络参数，再反向传播调参的过程中，也是用该参数进行数据的重构，这在一定程度上影响了算法的检测精度。本章提出的入侵检测方法解决了文献[9]存在的问题，通过非对称的方式对 ADBN 模型中的参数进行初始化，使模型能够得到更好的低维表征。实验结果表明，该方法对各类攻击均能保持较高的检测准确率，特别是针对小类型样本，但是由于算法对小类型样本的误报率较高，所以检测精度有待提高。

实验 4　将本章提出的入侵检测方法与其他方法在检测率和误报率进行比较，实验结果如表 6 所示。

表 6　算法检测率与误报率对比表

Tab. 6　Algorithm detection rate and false alarm rate comparison table

分类模型	数据集	DET/%	FAR/%
DBN	KDD	97.91	2.10
文献[3]	KDD	94.10	0.38
文献[4]	KDD	97.85	2.15
本章	KDD	99.39	0.31

由表 6 可以看出，传统的 DBN 网络模型和文献[4]提出的 NDAE 都取得了较高的检测率，但也存在误报率高的情况。文献[3]利用基于误用检测的方法，通过连接 RNN 模型约简模型中神经元的个数，降低了算法的误报率，虽然该模型可以提高 R2L 攻击类型的分类精度，但是相比于其他分类器而言，其总体检测率还有待提高。本章提出的入侵检测方法能在检测率和误报率之间取得更好的平衡。

实验 3、4 均采用 41-35-26-19-11-5ADBN 模型，各参数如表 5 所示。

4　结束语

本章提出了一种基于 ADBN 的特征择优方法，为了克服自编码网络的训练时间长及网络参数随机初始化的问题，结合 AE 和 DBN 的思想，通过逐层训练堆叠的 RBM 初始化 ADBN 模型中编码器部分的参数。同时为了避免网络层数的增加而导致的网络参数过于复杂，出现网络模型过拟合的现象，本章对 ADBN 模型中的解码器部分的参数通过服从正态分布的数值进行初始化，这也为模型的训练节省了一半的时间。实验结果表明，该算法能在入侵检测的综合性能上取得更好的平衡，针对小样本类别有更高的检测准确率，同时算法随 ADBN 模型隐藏层及顶层 RBM 输出层神经元个数影响不明显，所以具备一定的鲁棒性。

参考文献：

[1] Yi Chuanlong, Zhu Yuefei, Fei Jinlong, He Xinzheng. A Deep Learning Approach for Intrusion Detection Using Recurrent Neural Networks [J]. IEEE Access, 2017, PP (99): 21954-21961.

[2] Sheraz N, Yasir S, Shehzad K, *et al.* Enhanced Network Anomaly Detection Based on Deep Neural Networks [J]. IEEE Access, 2018, 6: 48231-48246.

[3] Sheikhan M, Jadidi Z, Farrokhi A. Intrusion detection using reduced-size RNN based on feature grouping [J]. Neural Computing & Applications, 2012, 21 (6): 1185-1190.

[4] 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.

[5] 高妮, 高岭, 贺毅岳, 王海. 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 (03): 730-739. (Gao Ni, Gao Ling, He Yiyue, Wang Hai. A Lightweight Intrusion detection Model Based on Autoencoder Network with Feature Reduction [J]. Acta Electronica Sinica, 2017, 45 (03): 730-739.)

[6] Hinton G E, Salakhutdinov R R. Reducing the dimensionality of data with neural networks. [J]. Science, 2006, 313 (5786) .

[7] Bengio Y. Learning Deep Architectures for AI [J]. Foundations & Trends® in Machine Learning, 2009, 2 (1): 1-127.

[8] 逯玉婧. 基于深度信念网络的入侵检测算法研究 [D]. 河北师范大学, 2016. (Lu Yujing. Intrusion Detection Algorithm Based on Deep Belief Networks [D]. Hebei Normal University, 2016.)

[9] Gao Ni, Gao Ling, He Yiyue, *et al.* Intrusion detection model based on deep belief nets [J]. Journal of Southeast University, 2015: 247-252.

[10] 陈虹, 万广雪, 肖振久. 基于优化数据处理的深度信念网络模型的入侵检测方法 [J]. 计算机应用, 2017, 37 (06): 1636-1643+1656. (Chen Hong, Wan Guangxue Xiao Zhenjiu. Intrusion detection method of deep belief network model based on optimization of data processing [J]. Journal of Computer Applications, 2017, 37 (06): 1636-1643,1656.)

[11] 汪洋, 伍忠东, 朱婧. 基于深度序列加权核极限学习的入侵检测算法 [J/OL]. 计算机应用研究,2020,37(3). (2019-01-29) [2019-03-27]. http://www.arocmag.com/article/02-2020-03-045.html. (Wang Yang, Wu Zhongdong, Zhu Jing. Intrusion detection algorithm based on depth sequence weighted kernel extreme learning [J/OL]. Application Research of Computers, ,2020,37(3). (2019-01-29) [2019-03-27]. http://www. arocmag. com /article/02-2020-03-045.html.)