DOI:10.11918/j.issn.0367-6234.201809115

深度自编码网络在入侵检测中的应用研究

丁红卫[1,2],万 良[1,2],龙廷艳[1,2]

(1.贵州大学 计算机科学与技术学院,贵阳 550025;2.贵州大学 计算机软件与理论研究所,贵阳 550025)

摘 要: 当前网络环境下的网络数据呈现出比以往更为庞大、复杂和多维的特性,传统的机器学习方法面临复杂的高维数据需要手动提取大量特征,特征提取过程复杂且计算量大,不利于当前入侵检测实时性和准确性的要求.基于此,以降低数据维度和消除冗余信息为目的,综合利用深度自编码网络(DAN)和BP算法,提出了基于DAN-BP的入侵检测模型.首先通过叠加多个自编码网络构成深度自编码网络模型,将网络特征数据作为模型的输入,使模型能够智能的逐层抽取网络数据的分布规则,从而获得新的低维特征数据集;然后利用BP算法对学习到的低维数据进行分类识别.文中通过在自编码网络中加入正则化修正,防止训练出的自编码网络直接复制输入信息而影响训练效果;且在输入数据中添加噪声,通过学习原始数据和输出数据重构误差达到去噪的目的,从而使得学习到的新的特征数据具有更强的鲁棒性.对比了传统的降维方法和当前入侵检测方法,结果表明本文方法在分类准确率、误报率和检测速率上均具有较优的效果.

关键词: 入侵检测;深度自编码网络;BP算法;降维;自编码网络

中图分类号: TP393 **文献标志码:** A **文章编号:** 0367-6234(2019)05-0185-10

Research on the application of deep auto-encoder network in intrusion detection

DING Hongwei[1,2], WAN Liang[1,2], LONG Tingyan[1,2]

(1. College of Computer Science and Technology, Guizhou University, Guiyang 550025, China;
2. Institute of Computer Software and Theory, Guizhou University, Guiyang 550025, China)

Abstract: The network data in the current network environment is enormous, complex, and multidimensional, which is hugely different from the past. The traditional machine learning method needs to manually extract a large number of features in the face of complex high-dimensional data, and the process is complex and computationally intensive, which is not conducive to the current real-time and accuracy requirements of intrusion detection. Thus, in order to reduce the data dimension and eliminate redundant information, an intrusion detection method based on DAN-BP which combines deep auto-encoder network (DAN) and BP algorithm is proposed. First, a DAN model was constructed by overlaying several auto-encoder networks, and the network feature data was used as the input of the model, which enables the model to intelligently extract the distribution rules of the network data layer by layer, thereby obtaining a new low-dimensional feature data set. Then the low-dimensional data was classified and identified by the BP algorithm. In this research, the regularization correction was added to the auto-encoder network to prevent the trained auto-encoder network from directly copying the input information and influencing the training effect. Moreover, noise was added to the input data, and the reconstruction error of the original data and the output data was learned to achieve the purpose of denoising so that the learned new feature data is more robust. The traditional dimensionality reduction method and the proposed intrusion detection method were compared in this paper. Results show that the proposed method has better performance in classification accuracy, false alarm rate, and detection rate.

Keywords: intrusion detection; deep auto-encoder network (DAN); BP algorithm; dimensionality reduction; auto-encoder network

网络的安全性已经成为当今世界上最重要的问题之一,网络数据容易受到各种类型的攻击,从而导致网络或系统的效率降低.入侵检测系统[1-2]作为保障网络安全的重要技术之一,也越来越受业界人士的关注.入侵检测系统是一种计算机和网络的安全管理系统,核心是收集和分析计算机或网络中各个区域的信息,检查计算机或网络中的行为是否安全.入侵检测作为一种积极主动的安全防御技术,能够有效的保障网络的安全性,因此许多机器学习方法[3-7]被应用到入侵检测技术中.

准确性和实时性是当前入侵检测系统的必要要求,只有正确识别出正常数据和异常数据才不会导

收稿日期: 2018-09-17
基金项目: 贵州省科学基金黔科合LH字[2014](7634)
作者简介: 丁红卫(1992—),男,硕士研究生;
通信作者: 万 良,wanliangtr@163.com

致系统出现误报和漏报的情况,同样只有能够及时处理网络中的信息才能及时采取措施,避免带来损失[8]. 入侵检测系统处理的网络数据通常含有大量的冗余和噪声,冗余和噪声特征的存在会严重消耗计算机系统的资源,从而使得入侵检测的检测时间较长、实时性较差和准确率较低. 特征降维方法能够很好的降低数据维度和消除冗余特征,因此为了能够准确和实时的进行入侵检测,对于网络特征进行降维还是十分必要的.

相关学者针对数据降维提出了不同的策略. Lakhina 等[9]提出一种新的混合算法,即 PCA-ANN 算法,主成分分析(Principal Component Analysis, PCA)用于减少输入特征的维度,人工神经网络(Artificial Neural Network, ANN)作为分类模型,实验表明该方法能有效的减少训练时间和测试时间. 高妮等[10]结合深度信念网络(Deep Belief Networks, DBN)和支持向量机(Support Vector Machine, SVM)方法提出了 DBN-MSVM 模型,该模型使用 DBN 进行特征降维,然后使用多类支持向量机分类器进行分类,有效的提高了分类准确率. 刘珊珊[11]等结合 PCA 降维方法和粒子群优化算法(Particle Swarm Optimization, PSO)的全局寻优能力优化 BP 神经网络权值和阈值提出了 PCA-PSO-BP 的入侵检测模型,有效的提高了准确率和收敛速率. Kuang 等[12-13]设计了一种 KPCA 和 SVM 相结合的算法,核主成分分析(Kernel Principal Component Analysis, KPCA)用于降低网络的数据维度,SVM 用降维数据的识别. Sharma 等[14]提出一种基于信息增益和相关性的智能系统,利用从信息增益和相关性中获得的等级来识别有用和无用的特征,从而实现特征的减少.

当前网络数据特征呈现的是复杂的非线性关系,而上述方法对于线性相关的特征具有较好的效果,面对非线性的数据时无法进行有效的将高维数据映射到低维空间,且上述方法并不能消除网络数据中的冗余和噪声数据. 因此,提出了深度自编码网络来进行非线性网络数据的降维,并引入降噪自编码网络来提高降维后数据的鲁棒性,在保证学习到最优的低维数据的前提下,提高入侵检测的准确率和检测速率. 通过利用深度自编码网络降维降噪方法和 BP 神经网络分类算法相结合的方法,提出了基于 DAN-BP 的入侵检测模型,旨在提高入侵检测准确率和降低检测时间.

1 深度自编码网络

深度自编码网络是由多层自编码网络堆叠而成的深度神经网络模型结构,在训练深度神经网络时会面临着固有的难点,即在进行反向传播的过程中误差更新信号会逐层衰减,从而导致深度神经网络模型无法更新. 因此在训练深度自编码网络时,选择使用逐层贪婪训练方法,即对每个自编码网络进行独立的训练,将前一层的训练输出作为下一层的输入依次训练,此方法可以有效的克服随着网络层数加深而出现的“梯度消失”问题.

1.1 正则自编码网络

自编码网络[15](Autoencoder Network, AN)是一种无监督的学习算法,不需要使用数据的标签信息. AN 是由编码器和解码器两部分结构,编码器是对原始数据的降维,解码器是对降维后数据的重构. 自编码网络的学习过程就是通过训练来降低重构数据和输入数据之间的重构误差,从而学习数据的内部特征表示.

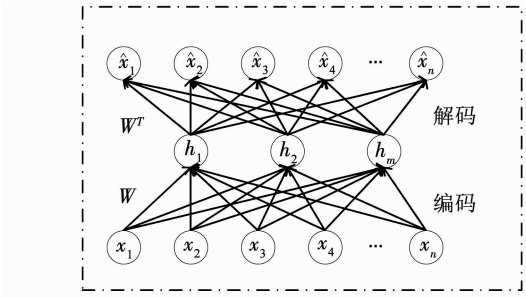


图 1 AN 结构

Fig. 1 AN structure

传统的自编码网络结构见图 1,设原始空间数据为$\boldsymbol{R}^{m\times n}$,$m$ 为原始空间中数据实例数,n 为每条实例数据的维度. $x_i \in \boldsymbol{R}^n$,$(i=1,2\cdots,m)$,编码和解码见式(1)和(2):

$$\boldsymbol{h} = S(f(x)) = S(\boldsymbol{W}x + \boldsymbol{p}). \tag{1}$$

式中:$\boldsymbol{W}$ 为输入层到隐含层之间的权值矩阵,$\boldsymbol{p}$ 为隐含层神经元的偏置量.

$$\hat{x} = S(g(\boldsymbol{h})) = S(\boldsymbol{W}^{\mathrm{T}}\boldsymbol{h} + \boldsymbol{q}). \tag{2}$$

式中:$\boldsymbol{W}^{\mathrm{T}}$为隐含层到输出层间的权值矩阵,$\boldsymbol{q}$ 为输出层神经元的偏置量.

$$G(S) = \frac{1}{1+e^{-S}}. \tag{3}$$

式(3)为 sigmoid 激活函数. 自编码网络的学习目标是最小化重构误差 L 的值,即使得输入值与输出值尽可能的接近,误差函数 L 选择均方误差损失函数见式(4):

$$L(x,\hat{x}) = \frac{1}{m}\sum_{m}^{1}(\hat{x}-x)^2. \tag{4}$$

若无任何约束,自编码网络很容易出现输出直接复制输入的情况(原本无用信息也会被加入降维

后的特征中)，由于自编码网络训练目的是降低重构误差，若出现输出直接复制输入信息，那么这样将对目标特征降维毫无意义. 因此为了防止复制输入信息，可在误差函数后加入正则化修正(稀疏性限制，稀疏性限制使得在降维过程中可以自动去除无用信息)，即可得到正则自编码网络见式(5)：

$$J_s(x,\hat{x}) = L(x,\hat{x}) + \beta\sum_j^m KL(\rho \parallel \hat{\rho}), \quad (5)$$

$$KL(\rho \parallel \hat{\rho}_j) = \rho\log\frac{\rho}{\hat{\rho}_j} + (1-p)\log\frac{1-\rho}{1-\hat{\rho}_j}. \quad (6)$$

式中：$KL(\rho \parallel \hat{\rho}_j)$作为稀疏惩罚项，$\beta$ 为稀疏惩罚项的权重，ρ 是稀疏性参数，$\hat{\rho}_j$ 为第 j 隐层的平均激活值.

深度神经网络易出现过度拟合的现象，模型发生过拟合时，拟合函数的系数通常较大，这样会导致拟合曲线的抖动非常剧烈，使得某些区间内导数的绝对值很大，正则化可以通过约束模型中的系数来减小拟合函数系数的值，使拟合曲线更加平稳，从而缓解过拟合问题. 在误差函数后加入范数惩罚的 L1 正则化方法来防止过度拟合的发生见式(7)：

$$J(x,\hat{x}) = L(x,\hat{x}) + \beta\sum_j^m KL(\rho \parallel \hat{\rho}_j) + \frac{\lambda}{n}\sum_w |w|. \quad (7)$$

式中：λ 为惩罚因子，w 为权值，n 为训练集样本数.

1.2　降噪自编码网络

Vincent 等认为如果自编码网络学习到的特征具有较强代表性，即使输入数据有所损伤也可以有效的重构出原始数据. 基于此，Vincent 等提出了降噪自编码网络，即在输入数据中加入部分损伤来训练自编码网络. 由于输入数据中带有噪声数据，则带着明确去噪目的自编码网络可以使学习到的特征数据更具有鲁棒性，因此利用这一优点来训练正则自编码网络.

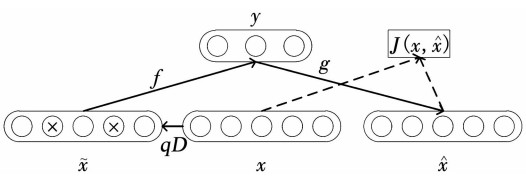


图 2　降噪自编码网络

Fig. 2　Denoising auto – encoder network

降噪自编码网络结构见图 2，x 为原始数据实例，qD 为随机映射函数. 首先使用随机映射函数将 x 转换为 $\tilde{x}$，目的是随机的在原始数据 x 中添加噪声数据；然后通过含有噪声数据的 $\tilde{x}$ 求得编码数据 $\boldsymbol{y} = S(\boldsymbol{W}\tilde{x} + \boldsymbol{p})$ 和解码数据 $\hat{\boldsymbol{x}} = S(\boldsymbol{W}^{\mathrm{T}}\boldsymbol{h} + \boldsymbol{q})$；最后通过 x 和 $\hat{x}$ 求得重构误差 J，使用梯度下降算法训练不断减小重构误差来尽可能的还原原始数据 x. 这种通过降噪自编码网络训练后的数据特征具有更强的鲁棒性.

1.3　深度自编码网构建

深度自编码网络是由多个正则自编码网络和降噪自编码网络相结合的变形网络堆叠而成的深度神经网络. 深度自编码网络的降维原理是逐层的减少隐藏层神经元的个数，以较少的深层特征来表达原始的浅层特征. DAN 的训练过程分为两个阶段，即预训练和微调，共 3 个步骤进行.

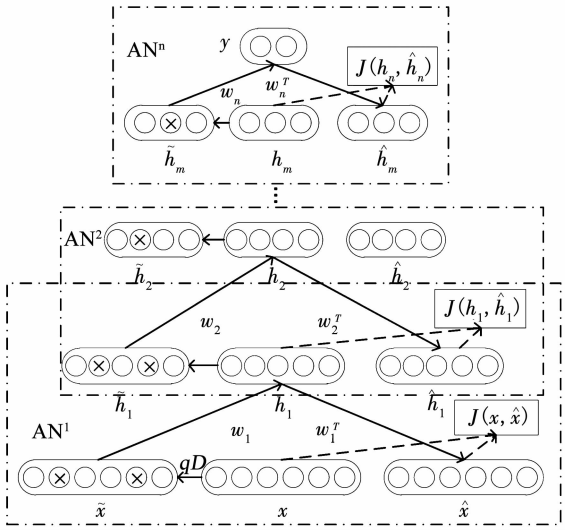


图 3　预训练

Fig. 3　Pre-training

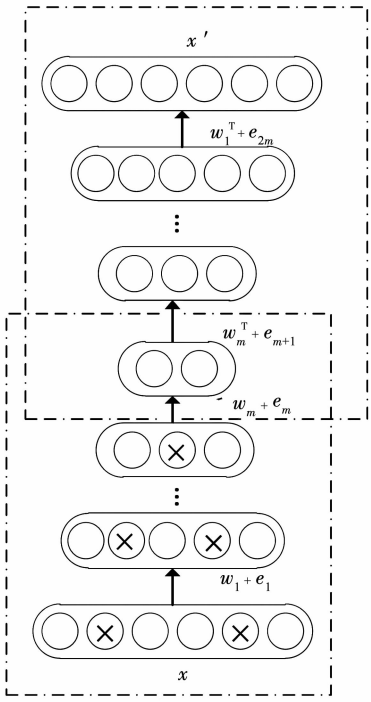


图 4　权值微调

Fig. 4　Fine-tuning

1.3.1 预训练

预训练过程见图 3. 预训练过程即为逐层的训练深度自编码网络的过程. 由于深度自编码网络为深度神经网络模型,传统的方法会导致发生“梯度消失”的问题,而逐层训练每个 AN 可以有效的避免此问题的发生. 预训练过程如下:

步骤 1:初始化参数 $\theta = \{w, b\}$ 和 t,即权值、偏置和最大迭代次数.

步骤 2:将每条实例特征赋值给输入神经元 x.

步骤 3:添加噪声,通过 qD 随机映射函数在输入特征 x 中添加噪声数据($x \sim qD(\tilde{x}/x)$).

步骤 4:计算隐含层神经元输出 $\boldsymbol{h}_1$,每个隐含层的神经元的输入可由式(1)得出.

步骤 5:重构输入特征,由隐含层计算得到的特征值通过式(2)得到重构特征 $\hat{x}$.

步骤 6:计算重构误差,原始特征 x 和重构特征 $\hat{x}$ 根据式(7)计算出重构误差.

步骤 7:更新 θ,依据梯度下降算法更新模型参数 θ 的值,更新见式(8)和(9).

$$w = w - \alpha \frac{\partial}{\partial w} J(x, \hat{x}), \tag{8}$$

$$b = b - \alpha \frac{\partial}{\partial b} J(x, \hat{x}). \tag{9}$$

步骤 8:若迭代次数 $k < t$ 则转向步骤 2,否则进行下一步训练.

步骤 9:将隐含层的输出作为下个 AN 的输入,根据步骤 3 ~ 8 的过程逐层贪婪的训练整个深度自编码网络.

1.3.2 构建深度自编码网络

通过预训练可得到训练完成的堆叠自编码网络,将堆叠的自编码网络展开后,得到的深度自编码网络见图 4. 其中模型的第一层至第 m 层为编码器,编码器中的每个自编码网络的权值由预训练得出 $\boldsymbol{W} = \{\boldsymbol{w}_1, \boldsymbol{w}_2, \cdots \boldsymbol{w}_m\}$;模型的第 m + 1 层至 2m 层为解码器,解码器中的每个自编码网络的权值设置为与编码器权值对应的转置,即 $\boldsymbol{W}^T = \{\boldsymbol{w}_1^T, \boldsymbol{w}_2^T, \cdots \boldsymbol{w}_m^T\}$.

1.3.3 微调

使用 BP 算法作为深度自编码网络的微调算法,微调的过程见图 4. 首先将 x 作为深度自编码网络的输入,从而求得输出 x';然后通过梯度下降算法最小化输入和输出误差,对网络模型参数进行更新. 微调算法具体过程如下.

步骤 1:正向传播过程. 将 x 作为输入数据,求得重构表示 x'.

步骤 2:根据误差反向传播,使用梯度下降算法进行权值和偏置的反向更新.

1)计算输出层误差,输出层误差见式(10):

$$Err_k = x_{k'}(1 - x_{k'})(x_k - x_{k'}). \tag{10}$$

2)计算隐含层误差,隐含层误差见(11):

$$Err_h = x_h'(1 - x_h') \sum_k Err_k w_{hk}. \tag{11}$$

式中:$x_{h'}$为隐含层输出,Err_k 为输出层误差,w_{hk}为隐含层和输出层的连接权值.

步骤 3:更新网络模型参数 w 和 b. 权值和偏置更新见式(12)和(13):

$$w_{hk} = w_{hk} + (L) Err_k x_{h'}, \tag{12}$$

$$b_h = b_h + (L) Err_h. \tag{13}$$

式中 L 为学习率.

步骤 4:达到最大迭代结束算法,否则转向步骤 1.

2 DAN-BP 模型构建

2.1 模型设计

基于 DAN-BP 入侵检测模型架构见图 5,该入侵检测框架具体过程如下:

步骤 1:原始数据预处理过程如下

1)属性映射,将字符型网络数据特征转换为数值型数据.

2)数据归一化,由于同种属性特征的数据之间相差较大,从而影响训练效果,因此要将数据归一化到[0,1]区间内.

步骤 2:DAN 特征降维

1)预训练,即逐层贪婪的训练每个自编码网络,前一个 AN 的输出是下一个的输入.

2)权值微调,由于每个自编码网络是单独训练的,因此只能保证每个自编码网络的权值最优,而无法保障整体权值最优,所以要进行整体权值的微调.

步骤 3:使用 BP 分类算法进行降维后数据的分类预测.

1)初始化参数,对 BP 分类模型进行参数初始化.

2)模型训练,使用训练数据对 BP 分类模型进行训练.

3)参数调优,根据每次的训练结果调整模型参数,直到模型达到最优.

步骤 4:将预测数据输入到训练完成的 BP 模型中,从而得到每条预测数据的预测结果.

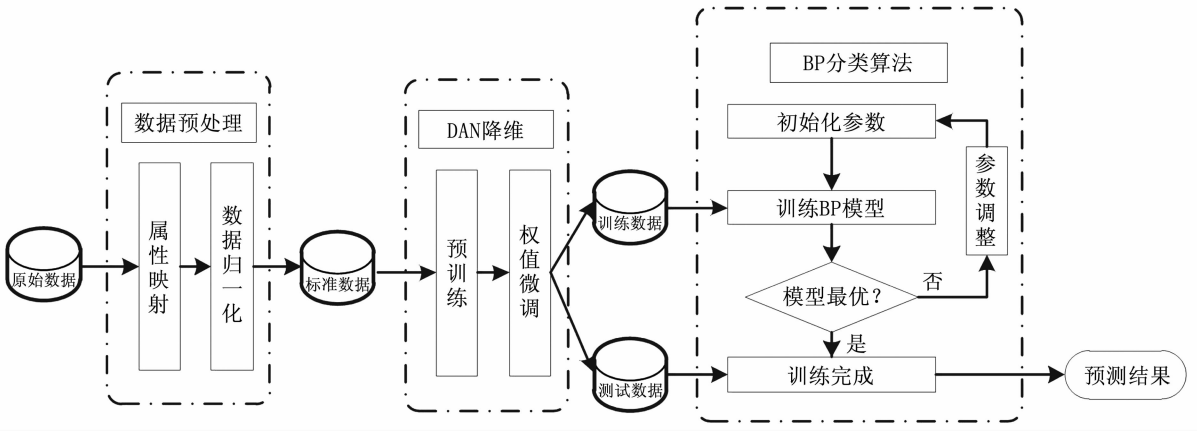


图 5　入侵检测总体架构图

Fig. 5　Overall architecture diagram of intrusion detection

2.2　数据预处理

使用 KDD CUP99[16] 数据集作为入侵检测实验数据，KDD CUP99 数据集因包含字符型数值，因此要进行预处理操作. 预处理过程如下：

2.2.1　数值化

原始的 KDD CUP99 数据集中包含 41 个特征属性，其中有 3 个属性为字符型的特征属性，分别为 protocol_type（协议类型）、service（目标主机的网络服务类型）、flag（连接正常或错误的状态）. protocol_type 包含 3 种协议类型，service 包含 70 种服务类型，flag 包含 11 种状态. 分别对这 3 种字符型数值进行数值化编码处理，具体编码过程见图 6：

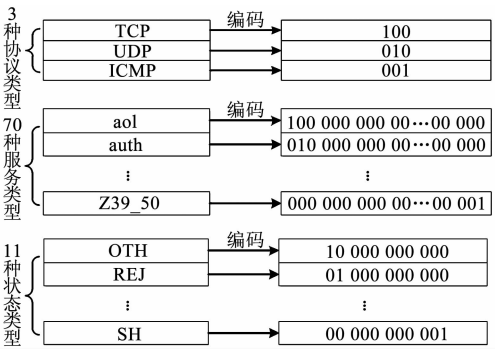


图 6　数值化

Fig. 6　Numeralization

2.2.2　归一化

数值化处理之后，数据集中的数据转换为数值型数据，但数值型数据中数值差异较大，如特征属性 duration（连接持续时间），取值范围是 0 ~ 58 329. 数值差异较大容易引起网络收敛较慢和神经元输出饱和等问题，因此要对原始数据进行归一化处理. 使用最大-最小归一化方法将数据集中的数据归一化到 [0,1] 区间之内见式(14).

$$x^* = \frac{x - x_{\min}}{x_{\max} - x_{\min}}. \tag{14}$$

式中：x^* 为归一化后的数据，x 为当前原始数据，$x_{\min}$ 为当前属性中的最小的数据值，$x_{\max}$ 为当前属性中的最大数据值.

2.3　DAN-BP 算法过程

1）初始化 AN 的权值 W 和偏置 b；

2）将预处理后的训练集 x 作为训练数据输入；

3）FOR i in epoch1

4）使用预训练方法训练 AN；

5）RETURN W,b；

6）构建深度自编码网络；

7）FOR i in epoch2

8）输入训练数据 x；

9）使用微调方法进行 DAN 参数微调；

10）END

11）向训练完成的深度自编码网输入训练数据和测试数据；

12）获得训练集和测试集的编码器降维结果 x_train、x_test；

13）构建 BP 分类模型；

14）FOR i in epoch3

15）将训练集的降维结果作为训练数据训练 BP 分类模型；

16）根据训练结果调整参数，直至模型效果最优；

16）END

17）将测试集输入到训练好的 BP 分类器中，得出预测结果.

3　实验

实验在 Intel Core i5 CPU 2.8GHz、4G RAM 环境和 Windows10 操作系统下实现，使用 Python3.5 进行仿真实验，BP 算法采用 Keras 框架实现.

3.1　数据集选择

KDD CUP99 数据集是当前入侵检测实验通用标

准数据集. KDD CUP 99 数据集源自林肯实验室的一项入侵检测评估项目,林肯实验室模拟空军局域网的一个网络环境,仿真各种不同用户和类型的网络攻击,使它就像一个真实的网络环境. 它是一个 LAN 上 9 周的模拟原始 TCPdump(*)转储数据的集合,训练数据从 7 周的网络流量获得,有大约 500 万个连接记录,最后两周产生测试数据约 200 万连接记录. 在数据中共有 4 大类攻击类型,其中又分为 39 种小类,训练数据共有 22 种攻击类型,新的 17 种攻击是测试数据集中存在的附加攻击, 而不存在于训练数据集中. 数据集中每条实例数据包含 41 个特征属性和一个标签属性,其中标签属性分为 5 大类,即 Normal、DOS、Probe、R2L 和 U2R 5 种类型.

3.2 评价指标

对比实验中,采用准确率(Accuracy,AC),误报率(False Alarm Rate,FA)和召回率(Recall,RE)等作为本实验效果优劣的评判标准,见式(15) ~(17):

$$\mathrm{AC}=\frac{T_{\mathrm{P}}+T_{\mathrm{N}}}{P+N}\ , \tag{15}$$

$$\mathrm{FA}=\frac{F_{\mathrm{N}}}{P}, \tag{16}$$

$$\mathrm{RE}=\frac{T_{\mathrm{P}}}{T_P+F_{\mathrm{N}}}. \tag{17}$$

具体参数含义见表 1.

表 1 参数定义

Tab. 1 Definitions of parameters

P	N	T_{P}	F_{N}	T_{N}	F_{P}
正样本数	负样本数	正样本被判对	正样本被判错	负样本被判对	负样本被判错

3.3 DAN 结构分析

3.3.1 参数设置

从训练集和测试集中分别选出 40 000 条训练数据和 10 000 条测试数据进行实验,各类数据的具体情况见表 2.

表 2 各类数据详情

Tab. 2 Particulars of various data

数据集	Normal	Dos	Porbe	U2R	R2L	总数/条
训练集	10 000	24 719	4 107	52	1 122	40 000
测试集	2 700	6 004	1 010	35	251	10 000

在进行最终实验测试时,首先要进行参数调优. 目前对于参数调优并没有自动化的方法,只能在前人提供的参考值的基础上进行大量的实验,对比分析参数的优劣,通过反复试验调优,最终确定的参数见表 3.

表 3 参数设置

Tab. 3 Parameter settings

DAN 参数设置		BP 参数设置	
参数	数值	参数	数值
网络结构	122-60-30-15-5	网络结构	5-28-10-5
AN 学习率	0.01	学习率	0.001
L1 正则化系数	0.005	Dropout	0.2
稀疏惩罚因子 β	0.15	训练轮数	30
稀疏参数 ρ	0.05		
预训练轮数	100		
微调轮数	50		

3.3.2 降维效果分析

将 DAN 方法与 PCA(主成分分析)方法和传统 AN 方法的降维结果做了对比实验. 3 种降维方法同时对测试数据集进行降维处理,降维后的二维分布见图 7.

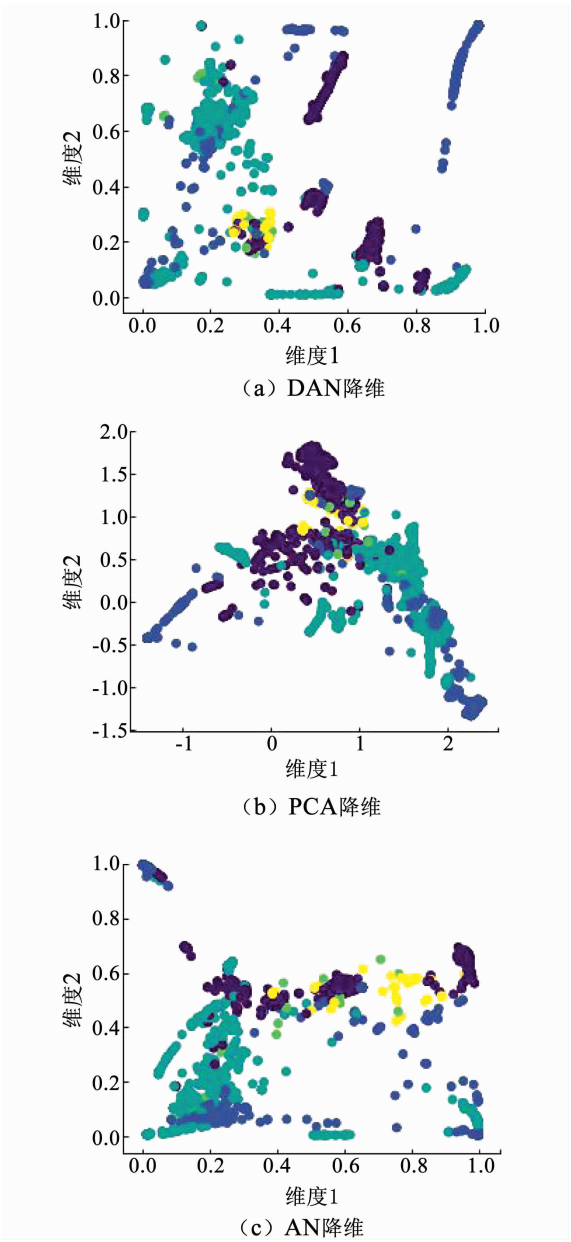


图 7 降维效果对比

Fig. 7 Comparison of dimensionality reduction effects

由图 7 中的二维分布结果可知，DAN 降维方法要优于 PCA 和 AN 方法. 因为 DAN 方法为非线性降维方法，随着 DAN 深度的增加，其非线性映射的次数也会越多，可以提取更为抽象的特征，因此也可保留更多有用的原始信息；PCA 方法为线性降维方法，将高维数据映射至较低的维度会丢失较多的原始信息，PCA 降维后的累积贡献率为 76.3%，以此也可看出 PCA 降维信息损失较为严重；传统的 AN 方法为浅层的网络，由于网络层次较浅，而不能抽取原始数据中的有效信息，从而导致降维效果不理想.

3.3.3　结构分析

1）网络层数分析

深度自编码网络的结构对于降维的结果有较大的影响，通常随着网络层次的加深，编码后提取特征就会更加抽象，对于原始数据的表示层次也会更高. 因此为了比较得出最优的深度模型结构，设置了不同深度的自编码网络结构进行对比实验. 通过测试 5 种不同结构的 DAN 模型来比较各自的性能，5 种 DAN 降维后的训练误差见表 4，将 5 种降维数据输入至 BP 模型进行训练和预测的结果见图 8.

表 4　不同结构的 DAN 训练误差

Tab. 4　DAN training errors for different structures

网络结构	层数	误差
122-5	2	0.004 60 ± 0.000 02
122-60-5	3	0.001 23 ± 0.000 06
122-60-25-5	4	0.000 99 ± 0.000 02
122-60-30-15-5	5	0.000 70 ± 0.000 03
122-100-60-30-15-5	6	0.000 94 ± 0.000 09

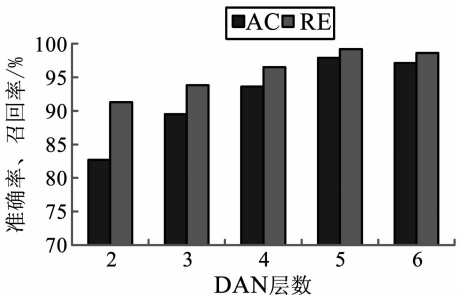


图 8　不同层数性能比较

Fig. 8　Performance comparison of different layers

由结果可知，具有 5 层网络结构的 DAN 具有更好的训练误差和测试效果，较浅或较深都会使误差增大. 通常随着网络深度的增加，DAN 的概括抽象能力也会增强. 浅层的 DAN 网络由于网络层次较浅，导致概括抽象能力较差，容易损失较多的原始信息，从而使得训练误差较大；太过于深层的 DAN 网络，随着深度的增加会导致抽取的信息过于抽象，使得原始数据的细节信息丢失，也会使训练误差增加. 因而选用具有 5 层结构的 DAN 进行数据降维.

2）降维维度分析

为研究降维维度对预测结果的影响，找到最小最优的降维维度，选择降维至 1 到 8 维的维度进行了分析. 根据网络层数分析的结果，选用具有 5 层结构的 DAN 进行分析，实验结果如图见图 9. 通过实验结果可知，降维至 5 维的 DAN 具有更好的入侵检测准确率和召回率.

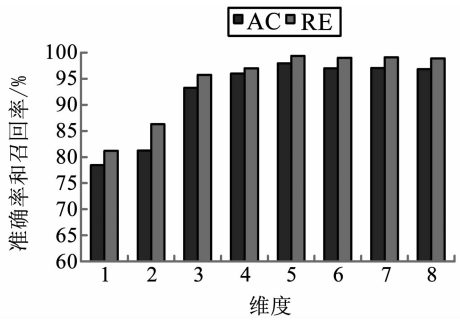


图 9　降维维度分析

Fig. 9　Dimensionality reduction analysis

将降维后数据作为 BP 神经网络的输入进行训练和预测，结果见图 10 和图 11. 由两组图像可知，检测曲线呈现稳定上升的趋势且训练准确率较高，代价函数值呈现稳定下降的趋势且训练误差较低，表明神经网络具有良好的训练效果. 因此，选择 122-60-30-15-5 结构的 DAN-BP 模型作为实验模型.

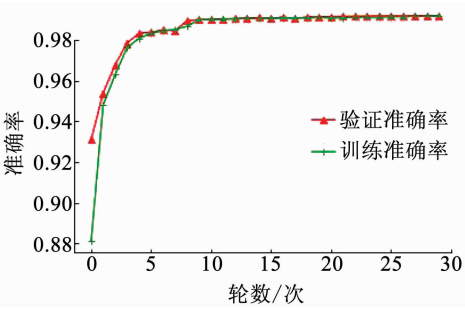


图 10　准确率曲线

Fig. 10　Accuracy curves

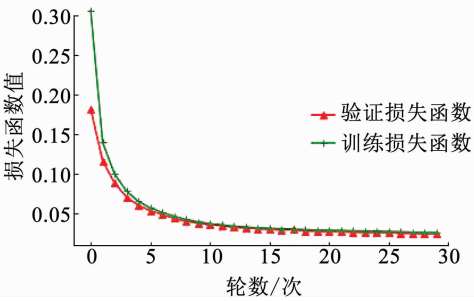


图 11　损失函数曲线

Fig. 11　Loss function curves

3.4 攻击实验分析

3.4.1 单独攻击实验

从数据集中选出 6 种最常见的入侵类型作为单独入侵检测的实验对象，每种类型数据选出 2 000 条数据进行测试（其中正常类型的数据为 1 900 条，入侵类型的数据为 100 条）. 实验结果见表 5.

表 5 单独攻击实验结果

Tab. 5 Results of individual attack experiment

入侵类型	正常数据/条	入侵数据/条	准确率/%	误报率/%
back	1 900	100	95.60	1.12
guess_passwd	1 900	100	95.71	1.16
ipsweep	1 900	100	98.50	1.15
neptune	1 900	100	99.42	1.05
smurf	1 900	100	95.57	1.10
portsweep	1 900	100	99.61	1.00

由表 5 可看出 DAN-BP 方法对于常见的攻击类型具有较好的识别效果，其实验准确率和误报率都要优于传统的入侵检测方法[17]. 因此，DAN-BP 方法对于常见的单独攻击可以进行有效的识别.

3.4.2 混合攻击实验

由于网络数据中的攻击类型通常较为复杂，因此设置了不同的攻击组合类型来测试 DAN-BP 算法对于复杂网络攻击的有效性. 分别对数据集中的 4 类攻击类型数据和正常类型数据进行了入侵检测的实验，不同攻击组合类型的检测准确率见图 12. 图中的横轴代表不同攻击组合的数据，纵轴代表准确率. 攻击组合中所包含的攻击类型见表 6.

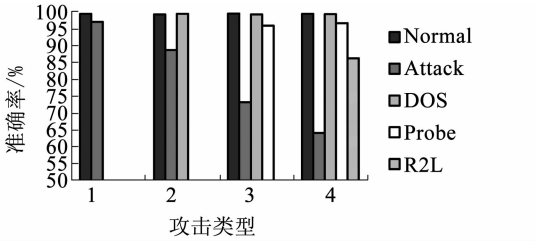


图 12 各类攻击准确率

Fig. 12 Accuracies of various types of attacks

表 6 攻击组合说明

Tab. 6 Description of attack combinations

类别	Attack			
	DOS	Probe	R2L	U2R
1	T	T	T	T
2	F	T	T	T
3	F	F	T	T
4	F	F	F	T

表 6 中给出了实验中攻击组合中所包含的攻击类型的信息. 其中“T”表示攻击组合中包含了该类的攻击类型，“F”表示攻击组合中没有包含该类攻击，但该类攻击被作为一个单独的攻击类型进行实验.

由实验结果可知，当只包含正常和异常两类数据时，正常和异常的检测率都较高. 当包含 3 类数据时，对 Normal 类型和 DOS 攻击类型具有很好检测效果. 当包含 4 类数据时，对 Normal 类型、DOS 类型和 Probe 类型的数据都具有较好的检测效果. 由于 U2R 和 R2L 类型的攻击的训练数据很少，所以导致训练不足，因此得到的检测结果在 5 种数据类型中的准确率稍低. 总体来说这些实验结果都要优于传统的入侵检测算法[18-19].

3.5 与现有方法对比分析

为了验证算法的有效性，从数据集中随机抽取了 4 组数据进行实验验证，并将实验结果和传统的降维算法及现有的深度学习方法进行了对比. 4 组数据见表 7.

表 7 4 组数据抽样结果

Tab. 7 Sampling results of the 4 groups of data

数据集	训练集			测试集		
	正常	异常	总量/条	正常	异常	总量/条
D1	10 280	9 619	19 899	6 139	2 536	8 675
D2	11 223	12 410	23 633	6 856	2 709	9 565
D3	9 302	11 708	21 010	9 352	2 439	11 791
D4	10 626	13 410	24 036	8 436	2 138	10 574

3.5.1 准确率、误报率和时间分析

通过对比 4 组数据的准确率（AC/%）、误报率（FA/%）、训练时间（Tr/s）和测试时间（Te/s）来测试算法的有效性，对比结果见图 13，DAN-BP 模型在准确率、误报率和检测时间（训练时间和测试时间）方面都要优于其它常用的入侵检测算法，具体对比结果见表 8.

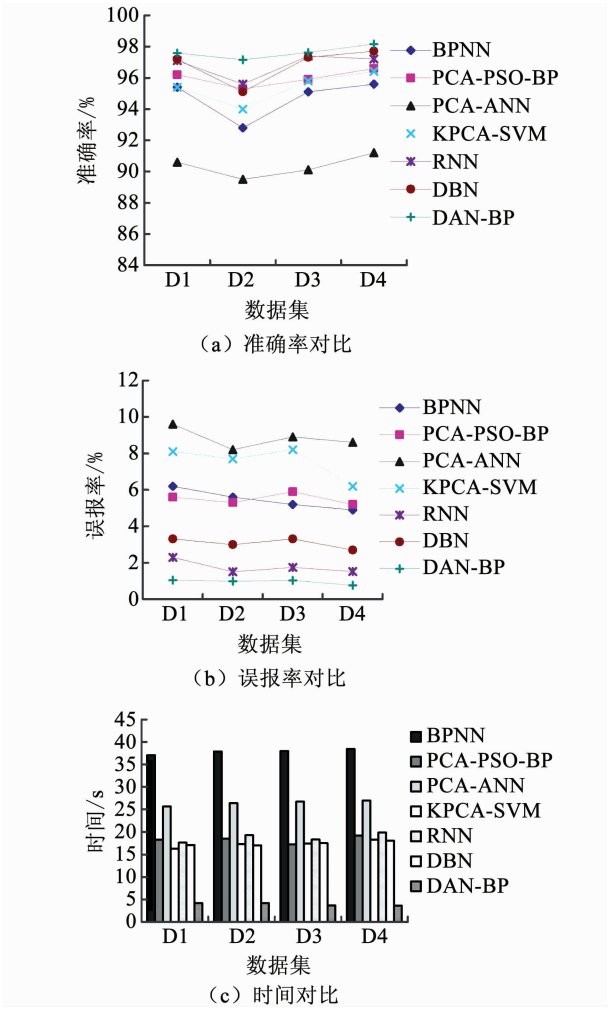
表 8 实验结果对比

Tab. 8 Comparison of experimental results

分类器	评价指标	D1	D2	D3	D4
BPNN	AC/%	94.5	92.8	95.1	95.6
	FA/%	6.2	5.6	5.2	4.9
	Tr/s	27.43	28.11	27.79	28.56
	Te/s	9.65	9.76	10.21	9.89
PCA-PSO-BP[11]	AC/%	96.2	95.3	95.9	96.6
	FA/%	5.6	5.3	5.9	5.2
	Tr/s	12.53	13.16	12.57	13.41
	Te/s	5.71	5.32	4.73	5.75
PCA-ANN[9]	AC/%	90.6	89.5	90.1	91.2
	FA/%	9.6	8.2	8.9	8.6
	Tr/s	16.73	17.06	16.89	17.21
	Te/s	8.91	9.32	9.83	9.75

续表

分类器	评价指标	D1	D2	D3	D4
KPCA-SVM[12]	AC/%	95.4	94.0	95.8	96.4
	FA/%	8.1	7.7	8.2	6.2
	Tr/s	11.25	12.02	11.47	12.39
	Te/s	5.06	5.32	5.93	5.89
RNN[20]	AC/%	97.1	95.6	97.4	97.2
	FA/%	2.3	1.5	1.74	1.52
	Tr/s	12.53	13.79	12.36	13.87
	Te/s	5.11	5.47	5.96	5.98
DBN[21]	AC/%	97.19	95.10	97.31	97.71
	FA/%	3.3	3.0	3.3	2.7
	Tr/s	11.69	11.56	11.79	12.37
	Te/s	5.42	5.49	5.71	5.63
DAN-BP	AC/%	97.58	97.15	97.63	98.16
	FA/%	1.04	0.99	1.02	0.76
	Tr/s	3.05	3.21	2.41	2.87
	Te/s	1.12	0.96	1.25	0.76



（a）准确率对比

（b）误报率对比

（c）时间对比

图 13　与现有方法对比

Fig. 13　Comparison with existing methods

3.5.2　各类型数据准确率对比

为进一步验证算法的有效性，分别对 Normal、Dos、Probe、U2R 和 R2L 5 种不同类型的数据分别进行了比较分析，实验结果见表 9. DAN-BP 模型对于 Normal 和 Dos 类型数据具有较高的检测率. 由于 U2R 和 R2L 类型的攻击数据的训练数据较少且未知攻击较多，因此检测准确率稍低，但相比于其它的分类器模型的检测效果有了较大的提升.

表 9　各数据类型准确率对比

Tab. 9　Comparison of accuracy rate of each data type

分类器	准确率/%	Normal	Dos	Probe	U2R	R2L
KPCA-SVM	AC	98.87	98.59	93.27	26.48	72.12
BPNN	AC	96.23	97.52	91.56	23.21	64.11
PCA-PSO-BP	AC	97.28	98.49	92.11	26.52	65.46
RNN	AC	98.78	98.98	93.21	28.53	73.55
DBN	AC	99.01	98.95	94.16	32.51	75.86
DAN-BP	AC	99.51	99.26	96.35	61.11	82.73

4　结　论

入侵检测数据中原始数据的维度较高，且冗余较多，直接用于入侵检测中会增加计算复杂度和增大计算机资源消耗，因此提出了一种基于深度自编码网络的降维方法. 深度自编码网络是一种深层神经网络结构，通过隐藏层逐层抽取有效的信息，可以很好的进行特征降维和去除冗余特征，文中提出的 DAN 综合使用了正则自编码网络和降噪自编码网络，有效的提高了网络的泛化能力和降维特征的鲁棒性. 基于此，综合利用 DAN 高效的非线性降维能力和 BP 算法对于低维数据优秀的分类能力，提出了基于 DAN-BP 的入侵检测模型，该模型首先使用 DAN 对高维数据进行高维至低维的非线性映射，然后使用最优的低维数据进行 BP 神经网络训练和预测. 通过使用 KDD CUP99 数据集验证表明，本文方法非常适用于高维数据的信息抽取，从而有效的降低了训练时间和测试时间，非常符合对当前网络入侵检测的实时性要求，并且实验结果的预测准确率和误报率相比于当前常用的入侵检测方法有了更好的提升. 因此 DAN-BP 入侵检测模型不仅提高了入侵检测预测准确率且加快了入侵检测速度，是一种适用于当前高维、复杂网络数据的方法，可以为当前的网络入侵检测研究提供一种新的思路.

参考文献

[1] ASHFAQ R A R, WANG Xizhao, HUANG Zhexue, et al. Fuzziness based semi-supervised learning approach for intrusion detection system[J]. Information Sciences, 2017, 378(C): 484. DOI:10.1016/j.ins.2016.04.019

[2] 卿斯汉，蒋建春，马恒太，等. 入侵检测技术研究综述[J]. 通信学报，2004，25(7)：19. DOI:10.3321/j.issn:1000-436X.2004.07.003

QING Sihan, JIANG Jianchun, MA Hengtai, et al. Research on

intrusion detection techniques: a survey [J]. Journal of China Institute of Communications, 2004, 25(7): 19. DOI:10.3321/j.issn:1000-436X.2004.07.003

[3]范自柱, 徐勇, 徐保根, 等. 一种快速 KMSE 算法及其在异常入侵检测中的应用[J]. 哈尔滨工业大学学报, 2011, 43(3): 90. DOI:10.11918/j.issn.0367-6234.2011.03.019
FAN Zizhu, XU Yong, XU Baogen, et al. A fast KMSE algorithm and its application on anomaly intrusion detection[J]. Journal of Harbin Institute of Technology, 2011, 43(3): 90. DOI:10.11918/j.issn.0367-6234.2011.03.019

[4]BASU A, ROY S S, ABRAHAM A. A novel diagnostic approach based on Support Vector Machine with linear kernel for classifying the erythemato-squamous disease[C]//2015 International Conference on Computing Communication Control and Automation. NY: IEEE, 2015: 343

[5]ROY S S, VISWANATHAM V M. Classifying spam emails using artificial intelligent techniques [J]. International Journal of Engineering Research in Africa, 2016, 22: 152. DOI:10.4028/www.scientific.net/JERA.22.152

[6]TAN Bin, TAN Yang, LI Yuanxu. Research on intrusion detection system based on improved PSO-SVM algorithm [J]. Chemical Engineering Transaction, 2016, 51: 583. DOI: 10.3303/CET1651098

[7]MITTAL D, GAURAV D, ROY S S. An effective hybridized classifier for breast cancer diagnosis [C]//IEEE International Conference on Advanced Intelligent Mechatronics. NY: IEEE, 2015: 1026

[8]陈友, 程学旗, 李洋, 等. 基于特征选择的轻量级入侵检测系统[J]. 软件学报, 2007, 18(7): 1639. DOI:10.1360/jos181639
CHEN You, CHENG Xueqi, LI Yang, et al. Light weight intrusion detection system based on feature selection[J]. Journal of Software, 2007, 18(7): 1639. DOI:10.1360/jos181639

[9]LAKHINA S, JOSEPH S, VERMA B. Feature reduction using principal component analysis for effective anomaly-based intrusion detection on NSL-KDD [J]. International Journal of Engineering Science & Technology, 2010, 2(6): 3175

[10]高妮, 贺毅岳, 高岭. 海量数据环境下用于入侵检测的深度学习方法[J]. 计算机应用研究, 2018, 35(4): 1197. DOI:10.3969/j.issn.1001-3695.2018.04.050
GAO Ni, HE Yiyue, GAO Ling. Deep learning method for intrusion detection in massive data[J]. Application Research of Computers, 2018, 35(4): 1197. DOI:10.3969/j.issn.1001-3695.2018.04.050

[11]刘珊珊, 谢晓尧, 景凤宣, 等. 基于 PCA 的 PSO-BP 入侵检测研究[J]. 计算机应用研究, 2016, 33(9): 2795. DOI:10.3969/j.issn.1001-3695.2016.09.054
LIU Shanshan, XIE Xiaoyao, JING Fengxuan, et al. Research on network intrusion detection based on PCA PSO-BP[J]. Application Research of Computers, 2016, 33(9): 2795. DOI:10.3969/j.issn.1001-3695.2016.09.054

[12]KUANG Fangjun, XU Weihong, ZHANG Siyang, et al. A novel approach of KPCA and SVM for intrusion detection[J]. Journal of Computational Information Systems, 2012, 8(8): 3237

[13]KUANG Fangjun, XU Weihong, ZHANG Siyang. A novel hybrid KPCA and SVM with GA model for intrusion detection[J]. Applied Soft Computing Journal, 2014, 18: 178. DOI:10.1016/j.asoc.2014.01.028

[14]SHARMA A, MANZOOR I, KUMAR N. A feature reduced intrusion detection system using ANN classifier[J]. Expert Systems with Applications, 2017, 88: 249. DOI:10.1016/j.eswa.2017.07.005

[15]BENGIO Y, LAMBLIN P, DAN P, et al. Greedy layer-wise training of deep networks [J]. Advances in Neural Information Processing Systems, 2007, 19: 153. DOI:citeulike-article-id:4640046

[16]TAVALLAEE M, BAGHERI E, LU W, et al. A detailed analysis of the KDD CUP 99 data set[C]//IEEE International Conference on Computational Intelligence for Security & Defense Applications. NY: IEEE, 2009:1

[17]刘胜会. 聚类分析在入侵检测中的应用研究[D]. 重庆: 重庆大学, 2014
LIU Shenghui. Research and application of clustering analysis in intrusion detection[D]. Chongqing: Chongqing University, 2014

[18]POTLURI S, DIEDRICH C. Accelerated deep neural networks for enhanced Intrusion Detection System [C]//IEEE International Conference on Emerging Technologies and Factory Automation. NY: IEEE, 2016: 1

[19]SADEK R A, SOLIMAN M S, ELSAYED H S. Effective anomaly intrusion detection system based on neural network with indicator variable and rough set reduction [J]. International Journal of Computer Science Issues, 2013, 10(6): 227

[20]YIN Chuanlong, ZHU Yuefei, FEI Jinlong, et al. A deep learning approach for intrusion detection using recurrent neural networks [J]. IEEE Access, 2017, 5(99): 21954. DOI: 10.1109/ACCESS.2017.2762418.

[21]高妮, 高岭, 贺毅岳. 面向入侵检测系统的 Deep Belief Nets 模型[J]. 系统工程与电子技术, 2016, 38(9): 2201. DOI:10.3969/j.issn.1001-506X.2016.09.33
GAO Ni, GAO Ling, HE Yiyue. Deep Belief Nets model oriented to intrusion detection system [J]. Systems Engineering and Electronics, 2016, 38(9): 2201. DOI:10.3969/j.issn.1001-506X.2016.09.33

(编辑 苗秀芝)