第 44 卷 第 9 期 计 算 机 工 程 2018 年 9 月
Vol. 44 No. 9 Computer Engineering September 2018

·安全技术· 文章编号：1000-3428(2018)09-0153-06 文献标志码：A 中图分类号：TP391

基于 DBN-ELM 的入侵检测研究

魏思政[1]，刘厚泉[1]，赵志凯[2]

（1. 中国矿业大学 计算机科学与技术学院，江苏 徐州 221116；
2. 中国矿业大学 物联网感知矿山研究中心，江苏 徐州 221008）

摘 要： 为了有效解决海量复杂数据的入侵检测分类问题，基于深度信念网络（DBN）和极限学习机（ELM），提出一种新的入侵检测方法。使用 DBN 对大量复杂无标签的原始数据进行特征提取，得到高度抽象的重要特征，再用 ELM 完成最终的分类工作。结合 DBN 自动提取特征的能力和 ELM 快速学习且泛化性好的优势，提高入侵检测识别率和运行效率。实验结果表明，与原始的 DBN、ELM 以及 DBN-SVM 方法相比，该方法具有更优的精确度和运行效率。

关键词： 深度学习；深度信念网络；极限学习机；混合模型；入侵检测；无监督

中文引用格式： 魏思政，刘厚泉，赵志凯. 基于 DBN-ELM 的入侵检测研究[J]. 计算机工程，2018，44(9)：153-158.
英文引用格式： WEI Sizheng，LIU Houquan，ZHAO Zhikai. Research on intrusion detection based on DBN-ELM[J]. Computer Engineering，2018，44(9)：153-158.

Research on Intrusion Detection Based on DBN-ELM

WEI Sizheng[1]，LIU Houquan[1]，ZHAO Zhikai[2]

（1. School of Computer Science and Technology，China University of Mining and Technology，Xuzhou，Jiangsu 221116，China；
2. IoT Perception Mine Research Center，China University of Mining and Technology，Xuzhou，Jiangsu 221008，China）

【Abstract】 In order to effectively solve the classification performance of massive and complex intrusion detection data，an intrusion detection method of hybrid deep learning model is proposed，which is based on Deep Belief Network（DBN）and Extreme Learning Machine（ELM）. DBN-ELM uses DBN to extract features from the massive，complex and unlabeled data to get highly abstract features at first，and then takes ELM as the classifier to finish the last classification. It totally improves the recognition rate of intrusion detection and the efficiency of the algorithm operation because it combines the ability of DBN to automatically extract features and fast learning and good generalization of ELM. Compared with DBN，ELM and DBN-ELM，experiments on KDD99 and NSL-KDD dataset show that DBN-ELM has better accuracy and efficiency of algorithm.

【Key words】 deep learning；Deep Belief Network（DBN）；Extreme Learning Machine（ELM）；hybrid model；intrusion detection；unsupervised

DOI： 10.19678/j.issn.1000-3428.0047591

0 概述

互联网带给人们方便的同时，计算机网络安全也成为了一个备受关注的重要问题。入侵检测系统（Intrusion Detection System，IDS）能够动态主动地进行安全检测，弥补了传统的安全防护方法的不足。文献[1]给出了入侵检测定义：通过监测网络数据信息，检测出入侵行为，在入侵行为造成危害前，发出警报并进行响应。

虽然经过了近 30 年的科技变革，传统的入侵检测系统已经得到了深入的发展和广泛应用，但是依然存在着很多问题，比如识别正确率低、容易发生误报、难以扩展等缺点。基于机器学习的入侵检测方法，能够利用历史入侵检测数据建立入侵检测识别模型，更有效地识别出新监测到未知攻击，同时具有高适应性和扩展性。越来越多的机器学习方法被 IDS 研究者引入到了入侵检测研究中，如文献[2]将经典的机器学习算法神经网络方法应用于入侵检测，文献[3]将基于贝叶斯算法分类的数据挖掘方法应用于入侵检测，由于极限学习机（Extreme Learning Machine，ELM）具有学习速度快且泛化性好的优势，文献[4]研究了基于 ELM 的

基金项目： 江苏省自然科学基金青年基金（BK20140216）。
作者简介： 魏思政（1991—），男，硕士研究生，主研方向为信息安全、机器学习；刘厚泉，教授、博士；赵志凯，助理研究员、博士。
收稿日期： 2017-06-14 **修回日期：** 2017-09-10 **E-mail：** 295431720@qq.com

入侵检测模型。但是传统的机器学习方法多属统计机器学习方法和简单的单隐层结构,表达能力有限,特征选择依赖人工,模型训练受限于人工标签的数据,面对海量复杂的真实网络应用环境产生的数据,其入侵检测识别正确率、稳定性和可靠性都得不到保证。

深度学习方法面对海量的入侵检测数据能够自动地提取特征,将低维冗余的特征经过层层映射,成为高度抽象的重要特征。文献[5-7]研究了深度信念网络对于入侵检测的高效性。文献[8]使用自编码网络提高入侵检测的识别率。文献[9]提出了一种基于支持向量机(Support Vector Machine,SVM)和深度信念网络(Deep Belief Network,DBN)的深度学习混合模型 DBN-SVM 并用于入侵检测。但已有的基于深度学习的入侵检测方法还不够完善,如面对海量的网络连接数据,DBN-SVM 运行效率低,并且 SVM 本身只支持二类分类问题。

为了进一步提高对入侵检测数据的分类识别率和算法运行效率,本文在原始的深度学习模型的基础上,提出一种新的深度学习混合模型 DBN-ELM。

1 基于 DBN-ELM 的入侵检测

基于 DBN-ELM 的入侵检测模型的总体框架如图 1 所示。首先在主机上部署网络监听系统,将监听到的历史网络访问数据集用来建立 DBN-ELM 模型,将实时数据使用建好的模型进行分类识别。对于监听到的数据需要进行预处理,主要步骤是特征映射、数字化和归一化,形成标准化数据。经过预处理的历史网络数据,需要分出小部分进行人工标签,先用大量无标签的历史数据无监督训练深度信念网络部分,再用这一小部分人工标签的数据使用训练好的 DBN 进行自动的特征提取,最后用提取到的特征训练高层的极限学习机得到完整的 DBN-ELM 模型。在训练好模型以后,就可以对监听到的新的网络访问数据进行分类识别。在这个过程中利用到了大量的无标签的历史网络访问数据,只需要人工标签一小部分数据,也不需要人工提取特征,对 DBN-ELM 来说,特征提取是通过在内部的 DBN 自动完成的。从这个结构中也可以看出,入侵检测的核心问题是如何利用网络历史数据建立模型对新的数据进行分类识别,而本文的方法不仅有能够利用海量无标签的优势,还结合了 DBN 自动提取特征的能力和 ELM 快速学习且泛化性好的优势。本节将详细介绍本文提出的 DBN-ELM 模型以及相关算法。

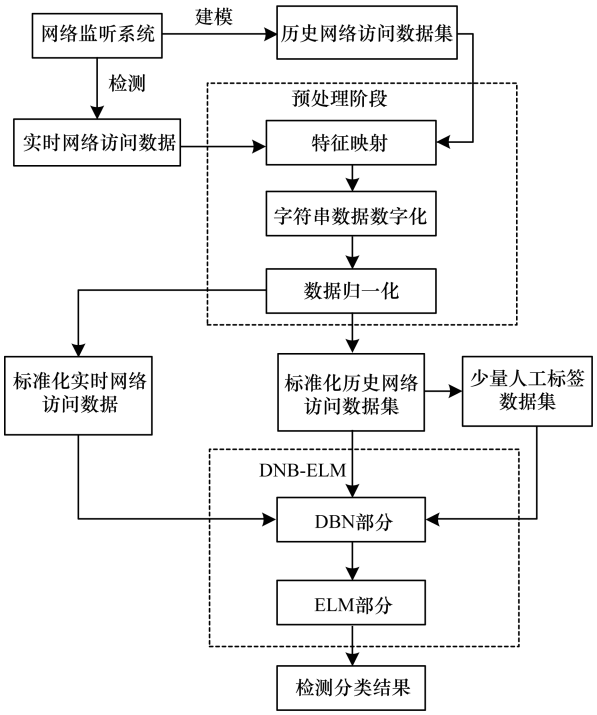

图 1 基于 DBN-ELM 的入侵检测框架

1.1 基于深度信念网络的特征提取方法

特征提取对于最终的分类是非常重要的,但是 ELM 是单隐层的神经网络,并且它的特征提取是随机映射,所以,ELM 特征提取的能力是相对匮乏的。相反的,DBN 能够从输入数据中自动的提取特征,通过多层的 RBM 特征映射,得到高度抽象的重要特征。

深度信念网络是最早的并成功应用于多个领域的深度学习方法之一[10],它由多层受限制的玻尔兹曼机(Restricted Boltzmann Machine,RBM)和一层反向传播(Back Propagation,BP)网络组成,如图 2 所示。

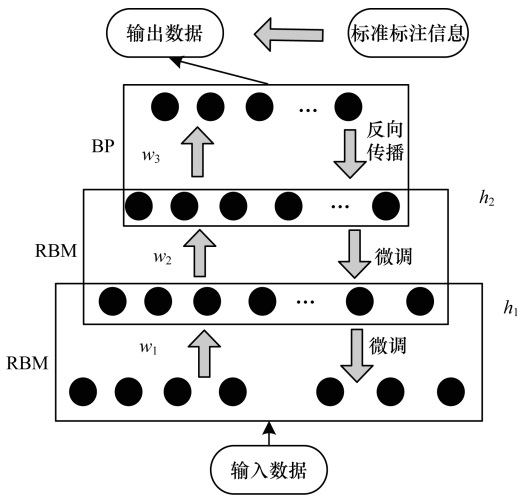

图 2 DBN 结构

为了使用 DBN 来进行自动的特征提取,先利用大量的无标签的入侵检测数据对 DBN 进行无监督的训练,从而初始化整个 DBN 的网络参数,主要是下面 2 个参数:在两层 RBM 之间的链接权重和各层神经元

的偏置。训练时采用逐层训练的方法,训练好当前层后,这一层 RBM 的输出作为下一层 RBM 的输入。下面将以 1 层 RBM 来介绍,结构如图 3 所示。

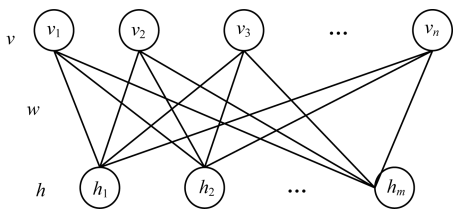

图 3　RBM 结构

RBM 是一种二分图,它包含一个可视层 v 和一个隐含层 h,每层之间进行全连接,但是同一层之间是没有连接的。若一个 RBM 包含 n 个可视单元和 m 个隐含层单元,则这个 RBM 的能量函数计算方法如下:

$$E(v,h|\theta) = -\sum_{i=1}^{n} a_i v_i - \sum_{j=1}^{m} b_j h_j - \sum_{i=1}^{n}\sum_{j=1}^{m} v_i W_{ij} h_j \quad (1)$$

其中,v_i 代表的是第 i 个可视单元的状态,h_j 则表示的是第 j 个隐含层单元的状态。RBM 的参数 $\theta = \{W_{ij}, a_i, b_j\}$,$W_{ij}$ 是可视单元 i 和隐含层单元 j 之间的权重,a_i 是可视单元 i 的偏置,b_j 是隐含层单元 j 的偏置。由此得分布函数:

$$P(v,h|\theta) = \frac{e^{-E(v,h|\theta)}}{Z(\theta)} \quad (2)$$

其中,$Z(\theta)$ 是配分函数,用来做归一化。联合概率分布 $P(v,h|\theta)$ 的似然函数为:

$$P(v|\theta) = \frac{1}{Z(\theta)}\sum_{h} e^{-E(v,h|\theta)} \quad (3)$$

RBM 的学习过程就是求出 θ,然后拟合样本。假设训练样本的数量是 T,那么根据极大似然法,也就是通过对数似然函数的最大化来求 θ。而其中最重要的一步是求 $\ln P(v^{(t)}|\theta)$ 关于 θ 的偏导数:

$$\frac{\partial \ln P(v|\theta)}{\partial \theta} = \sum_{t=1}^{T}\left(\left\langle\frac{\partial(-E(v^{(t)},h|\theta)}{\partial\theta}\right\rangle_{P(h|v^{(t)},\theta)} - \left\langle\frac{\partial(-E(v,h|\theta)}{\partial\theta}\right\rangle_{P(v,h|\theta)}\right) \quad (4)$$

其中,$\langle\cdot\rangle_P$ 代表分布 P 的数学期望。$P(h|v^{(t)},\theta)$ 代表可视单元为样本 $v^{(t)}$ 时隐含层单元的分布,而 $P(v,h|\theta)$ 为可视单元和隐含单元的联合分布。

通过以上推导的公式就可以使用大量无标签的入侵检测数据来逐层无监督的训练好每一层 RBM,从而初始化整个网络。但是多层的 RBM 组成的深度网络只是用来特征提取,还需要顶层的分类器完成最终的分类,而 DBN-ELM 就是结合了 ELM 作为分类器。

1.2　基于极限学习机的分类方法

ELM 是一种单隐层的前馈神经网络,由输入层、隐含层和输出层组成。传统的神经网络,需要用迭代算法来求得大量的网络节点参数,这个迭代过程是非常消耗时间的,因此,文献[11]提出了 ELM 算法。

ELM 不需要更新隐含层神经元的偏置和输入层到隐含层的权重,只需要设置隐含层神经元的个数,并且能够通过一次学习就得到唯一的最优值。入侵检测数据具有量大、变化快、未知的攻击多等特点,因此,使用极限学习机作为入侵检测的分类器,能够有效发挥其快速学习的优势和很好的泛化能力。

假设 ELM 输入神经元、隐含层神经元和输出神经元的个数分别为 n、l 和 m。输入层与隐含层之间的连接权重和隐含层和与输出层之前的连接权重分别是 w 和 β,并且假设有 N 个样例用来学习:(x,t),那么这个神经网络可以表示为:

$$\sum_{i=1}^{l}\beta_i g(w_i,b_i,x_j) = y_j, j = 1,2,\cdots,N \quad (5)$$

其中,函数 $g(\cdot)$ 是隐含层的激活函数,β 是隐含层神经元和输出神经元之间的连接权重,b 是隐含层神经元的偏置。

单隐层前馈神经网络(SLFN)为了逼近 N 个输入样本,并且零错误,因此:

$$\sum_{i=1}^{l}\beta_i g(w_i,b_i,x_j) = t_j, j = 1,2,\cdots,N \quad (6)$$

用 $\boldsymbol{H}$ 表示输出矩阵,$\boldsymbol{O}$ 为期望输出矩阵,因此:$\boldsymbol{H\beta = O}$。

ELM 能够在训练之前随机的生成 w 和 b[11-12],因此,只要明确 ELM 隐含层的神经元个数和隐含层的激活函数,就能够计算出 β。下面是 ELM 训练的过程:首先,确定好隐含层神经元的数量,随机的设置好隐含层神经元的偏置 b 和输入层和隐含层的权重 w。接着,确定激活函数并算出输出矩阵 $\boldsymbol{H}$。最后计算出隐含层和输出层之间的权重 $\beta = \boldsymbol{H}^{\mathrm{T}}\boldsymbol{H}$。

1.3　DBE-ELM 模型

1.3.1　DBN-ELM 网络深度的确定方法

只有先确定 DBN-ELM 的网络深度,才能进行整体的网络结构设计。本文通过对 DBN-ELM 隐含层输出样本之间的互相关系数的变化来确定网络的深度。因为 ELM 是单隐层的结构,所以主要是确定 DBN 部分的深度。

DBN 是由多层的 RBM 组成的,前一层的隐含层作为后一层的可视层,设 v_i、h_j 分别为可视层和隐含层的神经元状态,a_i、b_j 分别是它们对应的偏置,两层之间的权重用 w_{ij} 表示。则隐含层神经元的输出为:

$$h_j = \sum_i v_i w_{ij} + b_j \quad (7)$$

为了测试样本之间的相关性将样本分为两大类:正常与攻击。设 $h_j^{(1)}$、$h_j^{(2)}$ 分表示这两类样本在 j 层的

输出，i 层的样本输入用 $v_i^{(1)}$、$v_i^{(2)}$ 表示，则可以得到：

$$h_1 = h_j^{(1)} = \sum_i^N v_i^{(1)} w_{ij} + b_j \tag{8}$$

$$h_2 = h_j^{(2)} = \sum_i^N v_i^{(2)} w_{ij} + b_j \tag{9}$$

互相关系数的表达式则为：

$$\rho = \frac{\sum h_1 h_2}{\sqrt{h_1 h_2}} \tag{10}$$

在完全识别的情况下，不同类别的互相关系数会随着层数的增长而减小，$\rho_k < \rho_{k+1}$。这也解释了为什么 DBN 采用的深度结构，会因为层数的增加而降低错误的概率[13]。

通过以上的理论推导，只需要通过计算隐含层的不同类别之间的互相关系数，求出能够使互相系数稳定的最小隐含层数作为最终的网络深度，避免了浪费，提高了训练的效率。详细的层数确定数据，将在第 2 节实验部分详细讲解。

1.3.2 DBN-ELM 的网络结构

DBN-ELM 由负责特征提取的 DBN 和负责分类的 ELM 两部分组成，结合了 DBN 自动提取特征的能力和 ELM 快速学习且泛化性好的优势，提高了分类识别正确率和算法效率。通过 1.3.1 节的方法确定了模型结构隐含层的层数为 n，则把输入层到第 $n-1$ 个隐含层作为 DBN 部分，第 n 个隐含层作为 ELM 的隐含层，与作为 ELM 输入层的第 $n-1$ 个隐含层和输出层组成一个完整的 ELM，如图 4 所示。

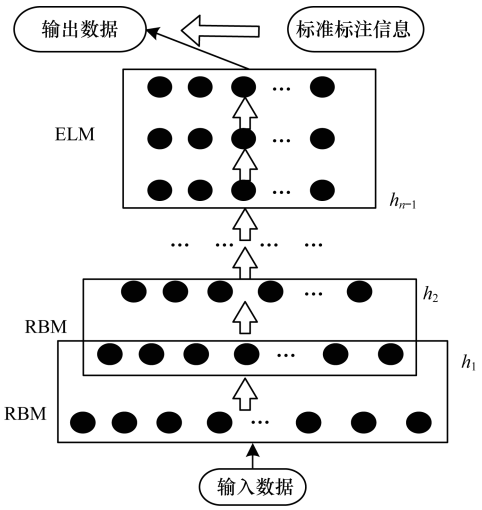

图 4 DBN-ELM 结构

确定网络的层数以后先用无监督的训练方法训练好 DBN 部分，也就是输入层到第 $n-1$ 个隐含层，再用 ELM 算法确定剩下的第 $n-1$ 层到输出层的权重和偏置[14]。

设 N 为第 n 个隐含层的神经元的数量，M 为第 $n-1$ 个隐含层神经元的数量，则此网络可表示为：

$$\sum_{i=1}^N \beta_i g(w_i h_{n-1} + b_i) = o_j, j = 1, 2, \cdots, M \tag{11}$$

w_i 和 b_i 分别为第 $n-1$ 个隐含层到第 n 个隐含层的权重和偏置，第 n 个隐含层到输出层的权重为 ρ_i，最小化输出误差如式(12)所示。

$$\begin{cases} \sum_j^N \| o_j - t_j \| = 0 \\ t_j = \sum_{i=1}^N \beta_i h_{nj}, j = 1, 2, \cdots, M \end{cases} \tag{12}$$

最终可以将问题转化为 $h_n\beta = \boldsymbol{O}$，$h_n$ 表示第 $n-1$ 个隐含层到第 n 个隐含层的输出：

$$h(w_1, w_2, \cdots, w_N, b_1, b_2, \cdots, b_N, h_{(n-1,1)}, h_{(n-1,2)}, \cdots, h_{(n-1,m)}) = \begin{bmatrix} g(w_1 \cdot h_{(n-1,1)} + b_1) \cdots g(w_N \cdot h_{(n-1,1)} + b_N) \\ g(w_1 \cdot h_{(n-1,m)} + b_1) \cdots g(w_N \cdot h_{(n-1,m)} + b_N) \end{bmatrix} \tag{13}$$

可以将问题化解为求 $\hat{w}_i, \hat{b}_i, \hat{\beta}$ 使得：

$$\| \boldsymbol{h}_n(\hat{w}_i, \hat{b}_i)\hat{\beta} - \boldsymbol{O} \| = \min_{w,b,\beta} \| \boldsymbol{h}_n(w_i, b_i)\beta - \boldsymbol{O} \| \tag{14}$$

根据 ELM 算法，第 $n-1$ 个隐含层到第 n 个隐含层的权重 w_i 和偏置 b_i，可以随机的初始化并得到唯一的输出矩阵 $\boldsymbol{h}_n$，并根据期望输出 $\boldsymbol{O}$ 将 DBN-ELM 的训练问题最终转化成了求解线性系统 $\boldsymbol{h}_n\beta = \boldsymbol{O}$，则：

$$\hat{\beta} = \boldsymbol{h}_n^+ \boldsymbol{O} \tag{15}$$

2 实验结果与分析

2.1 数据集及其预处理方法

KDD99 数据集由麻省理工学院林肯实验室模拟美国空军局域网环境而建立的网络流量测试数据集，多年来一直被国内外众多学者用来评价入侵检测算法的性能，尽管这个数据是 1999 年发布的数据集，但是现有的网络协议、操作系统等并没有发生巨大的变化，这个数据集依然能够有效评价入侵检测算法的性能，目前还是热门的入侵检测算法验证数据集。但是 KDD99 数据量庞大还有一些固有问题，而 NSL-KDD 数据集是一种包含了一部分 KDD99 数据集的新数据集[15]。NSL-KDD 数据集剔除了 KDD99 的大量冗余，并调整了测试集和训练集比例，使得更适合用于入侵检测实验。在本文实验中，KDD99 数据集只作为无标签的数据，NSL-KDD 数据集作为可靠的标签的数据。NSL-KDD 调整后的训练集和测试集分别有 25 192 个和 11 850 个。但是攻击行为仍然为 DoS、U2R、R2L、Probe。

NSL-KDD 有 41 维特征，其中包括数值型、字符型，需要进行预处理，如下：

1)将字符型数据映射为数值型。例如属性特征“protocal_type”有 3 种取值：tcp，udp 和 icmp，则可分别用二进制向量[1,0,0]、[0,1,0]和[0,0,1]表示。根据这个方法，最终将 41 维的原始特征变换成了

122 维的输入特征。

2)对数据进行归一化处理,将数据的大小范围缩小到 0 到 1 之间,使得各个属性特征处于同一个量级,方法如下:

$$y = \frac{y - MIN}{MAX - MIN} \tag{16}$$

3)用向量对 5 种标签类型进行编向量的 1 位 ~ 5 位分别表示为正常、DoS、U2R、R2L、Probe,输出时也用对应的映射编码表示,分别用向量表示为[1,0,0,0,0]、[0,1,0,0,0]、[0,0,1,0,0]、[0,0,0,1,0]和[0,0,0,0,1]。

2.2　网络深度分析

根据理论推导,需要观察隐含层样本之间的互相关系数,来确定最终的网络深度。先从训练集中随机的选取 20 000 个正常数据,再分层随机选取 20 000 个攻击数据,用来训练预设的 DBN-ELM 模型。

初始设置模型的隐含层数为 8,各隐含层神经元个数依次为 100-90-80-70-60-50-40-30,输出各个层2 类样本之间的互相关系数如图 5 所示。可以看出,当隐含层层数达到第 4 层时,互相关系数就近似于 -1,之后就开始保持不变,说明此时达到了最好的分类状态,所以,本文通过此实验设定隐含层的数量为 4。

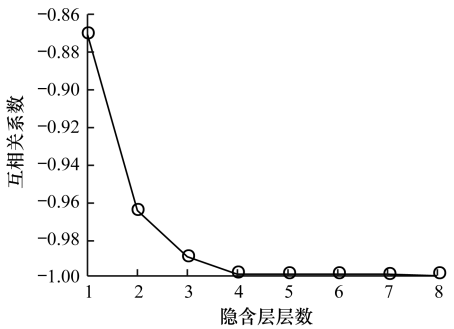

图 5　互相关系数趋势

2.3　DBN-ELM 的网络参数

DBN-ELM 的结构参数设置情况如表 1 所示,根据降维递减法 DBN 部分从输入到第 3 个隐含层的神经元设置为 122-100-70-30。第 4 个隐含层是 ELM 的隐含层,根据 Kolmgorov 定理中的定义,ELM 的输入层和隐含层的神经元的数目的关系为$k = 2m + 1$,因此,设定第 4 个隐含层的神经元的数目为 $2 \times 30 + 1 = 61$,输出层神经元个数为 5,因为需要分为 5 类。

表 1　DBN-ELM 结构参数设置

DBN-ELM 的参数	参数值
输入层神经元数	122
隐含层 1 神经元数	100
隐含层 2 神经元数	70
隐含层 3 神经元数	30
隐含层 4/ELM 隐含层神经元数	61
输出层神经元数	5

2.4　实验的评价标准

本文主要采用准确率和误报率作为入侵检测的评价指标,如下:

$$AC = \frac{TN}{N}$$

$$FA = \frac{FM}{M} \tag{17}$$

其中,AC 表示准确率,TN 表示正确分类的样本数目,N 表示样本总数,FA 表示误报率,FM 表示被误报为入侵的正常样本数,M 为正常样本总数。

2.5　结果分析

为了证明 DBN-ELM 比原始的 DBN 和 ELM 性能好,本文实验首先对需要无监督训练的部分统一采用 20% 的 KDD99 数据作为无标签的数据训练,然后重点讨论各模型在不同比例的 NSL-KDD 数据集下有监督训练后的效果。DBN、ELM 和 DBN-ELM 分别用 20%、30% 和 40% 的 NSL-KDD 训练数据集完整最终的训练,然后在测试集上进行了测试并对比,其中,DBN 的设计采用和 DBN-ELM 相同的层数以及其他参数。

如图 6 所示,DBN-ELM 达到了 97.5% 的准确率,这是在 40% 的训练集上训练后的结果,还可以看出,从 30% 训练集到 40% 训练集准确度提升相对平缓了,说明 DBN-ELM 发挥了 DBN 强大的抽象表示能力,只需要很少的数据集就可以训练出表达能力强的模型,同时也体现了 ELM 泛化性好的优势,在未知的测试集上准确度表现好。

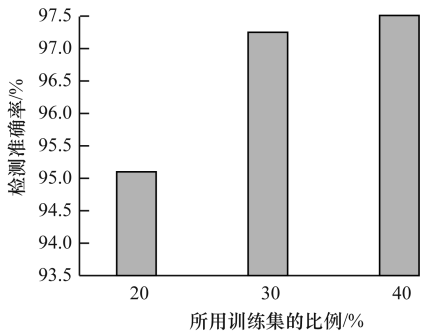

图 6　DBN-ELM 的检测精度

表 2 为 DBN、ELM 和 DBN-ELM 分别在 20%、30% 和 40% 训练集上训练后的准确度,其中,ELM 达到了 91.55% 的精确度,DBN 达到了95.32% 的精确度,并且 DBN-ELM 达到了 97.5% 的精确度,通过此表对比可以看出,DBN-ELM 的准确度明显优于 DBN 和 ELM,显著提高了入侵检测的正确识别能力。

表 2　ELM、DBN 和 DBN-ELM 检测准确度　%

训练数据集比例	ELM	DBN	DBN-ELM
20	88.11	93.73	95.10
30	89.82	95.01	97.25
40	91.55	95.32	97.51

为了进一步探索本文方法的性能,本文还使用了另一个深度学习混合模型 DBN-SVM 和本文相关算法进行对比。如图 7 所示,在 20% 样本训练的情况下,DBN-SVM 准确率比 DBN-ELM 稍高了0.2%,这个不可否认 SVM 在小样本情况下确实有优势,但是和本文方法差距极小,并且从整体看 DBN-ELM 模型的准确率优于其他 2 种方法,而从图 8 也可以看出,DBN-ELM 的误报率情况也是整体最好的。

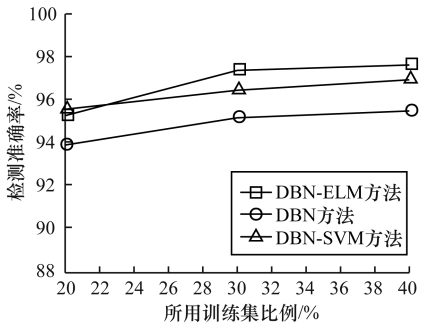

图 7 DBN-ELM、DBN、DBN-SVM 的准确率对比

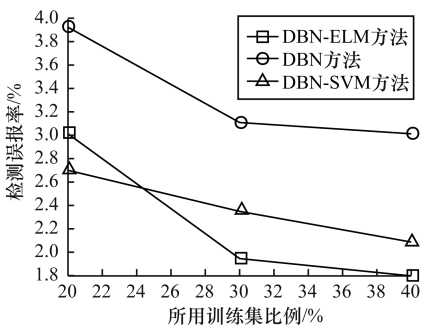

图 8 DBN-ELM、DBN、DBN-SVM 的误报率对比

表 3 为 ELM、DBN、DBN-SVM 和 DBN-ELM 分别在 20% 、30% 、40% 训练集上训练所消耗的时间,可以看出,DBN-ELM 在训练运行时间方面远远优于文献[6]提出的另一个深度学习混合模型 DBN-SVM,也低于原始的 DBN。所以,从总体的精确度和高效性来看,DBN-ELM 大幅提高了入侵检测的检测能力。

表 3 ELM、DBN、DBN-SVM 和 DBN-ELM 训练时间

训练数据集比例/%	ELM/s	DBN/s	DBN-SVM/s	DBN-ELM/s
20	0.021	0.27	2.54	0.12
30	0.032	0.61	3.96	0.21
40	0.043	1.53	5.07	0.35

3 结束语

本文提出一种新的用于入侵检测的深度学习混合模型 DBN-ELM。该模型先使用海量无标签的入侵检测数据进行无监督训练,再用 DBN 进行特征提取,之后用提取好的特征训练顶层的 ELM 部分,最后用 ELM 完成最终的分类。实验结果表明,DBN-ELM 具有比传统方法更好的性能。但是本文目前只是在 KDD99 和 NSL-KDD 数据集上进行了实验,下一步考虑搭建实时入侵检测环境,分析和解决该模型在实际应用中的问题。

参考文献

[1] DENNING D E. An intrusion-detection model[J]. IEEE Transactions on Software Engineering, 1987(2): 222-232.

[2] RYAN J, LIN M J, MIIKKULAINEN R. Intrusion detection with neural networks[J]. Advances in Neural Information Processing Systems, 1997, 28(10): 915.

[3] LEE W, STOLFO S J. A framework for constructing features and models for intrusion detection systems[J]. ACM Transactions on Information & System Security, 2000, 3(4): 227-261.

[4] 关亚文,刘 涛,黄 干. 无线传感器网络中基于 ELM 的混合入侵检测方案[J]. 计算机工程,2015,41(3): 136-141.

[5] 杨昆朋. 基于深度学习的入侵检测[D]. 北京:北京交通大学,2015.

[6] 杨昆朋. 基于深度信念网络的入侵检测模型[J]. 现代计算机(专业版),2015(2):10-14.

[7] 逯玉婧. 基于深度信念网络的入侵检测算法研究[D]. 石家庄:河北师范大学,2016.

[8] 李春林,黄月江,王 宏,等. 一种基于深度学习的网络入侵检测方法[J]. 信息安全与通信保密,2014(10): 68-71.

[9] AMBUSAIDI M, HE X, NANDA P, et al. Building an intrusion detection system using a filter-based feature selection algorithm[J]. IEEE Transactions on Computers, 2016, 65(10): 2986-2998.

[10] HINTON G, OSINDERO S, TEH Y W. A fast learning algorithm for deep beliefnets[J]. Neural Computation, 2006, 18(7): 1527-1554.

[11] HUANG Guangbin, ZHU Qinyu, SIEW C K. Extreme learning machine: a new learning scheme of feedforward neural networks[J]. IEEE International Joint Conference on Neural Networks, 2004, 2: 985-990.

[12] FENG Guorui, HUANG Guangbin, LIN Qingping, et al. Error minimized extreme learning machine with growth of hidden nodes and incremental learning[J]. IEEE Transactions on Neural Networks, 2009, 20(8): 1352.

[13] 高 强,马艳梅. 深度信念网络(DBN)网络层次数量的研究及应用[J]. 科学技术与程,2016(23): 234-238,262.

[14] 赵志勇,李元香,喻 飞,等. 基于极限学习的深度学习算法[J]. 计算机工程与设计,2015,36(4):1022-1026.

[15] TAVALLAEE M, BAGHERI E, LU Wei, et al. A detailed analysis of the KDD CUP 99 data set[C]// Proceedings of IEEE International Conference on Computational Intelligence for Security & Defense Applications. Washington D. C., USA: IEEE Press, 2009: 1-6.

编辑 刘 冰