Computer Science and Application 计算机科学与应用, 2018, 8(5), 687-701
Published Online May 2018 in Hans. http://www.hanspub.org/journal/csa
https://doi.org/10.12677/csa.2018.85077

Hans汉斯

Research on Intrusion Detection Based on Deep Confidence Network

Lin Yu

College of Computing, Beijing University of Technology, Beijing
Email: 906973561@qq.com

Received: Apr. 29th, 2018; accepted: May 16th, 2018; published: May 23rd, 2018

Abstract

As an active security prevention technology, intrusion detection has been used in network security for a long time. But with the development and application of Internet, network attack and intrusion are constantly changing in quantity and technology level. Based on new types and concurrent attacks, traditional intrusion detection technology has been unable to meet the requirements of existing network security. As a frontier technology of machine learning and artificial intelligence, deep learning has made great achievements in speech recognition, computer vision and big data processing, and has also provided a new idea for solving the current intrusion detection problem. This paper studies the traditional intrusion detection technology based on learning method combined with the depth and the depth of the belief network, proposes an intrusion detection technology based on deep belief networks, according to the characteristics of intrusion detection data of over sampling and non [0, 1] interval of data normalization, updates the parameters in the deep belief network in the process of using variable number of the gradient descent algorithm to speed up the learning rate, the parameters of the update process, and in each batch of the training data, join the discrimination on the labels, and improve the accuracy. Experiments show that the accuracy of intrusion detection can be improved greatly by using the method proposed in this paper.

Keywords

Intrusion Detection, Deep Learning, Restricted Boltzmann Machine, Deep Confidence Network

基于深度置信网络的入侵检测研究

余　淋

北京工业大学计算机学院，北京
Email: 906973561@qq.com

收稿日期：2018年4月29日；录用日期：2018年5月16日；发布日期：2018年5月23日

文章引用：余淋. 基于深度置信网络的入侵检测研究[J]. 计算机科学与应用, 2018, 8(5): 687-701.
DOI: 10.12677/csa.2018.85077

摘　要

入侵检测作为一种积极主动的安全防范技术，在网络安全上的应用历来已久。但是随着互联网的发展和应用的不断深化，网络攻击和入侵在数量和技术水平上的不断变化，基于新类型、多并发的攻击，使传统的入侵检测技术已经无法满足现有网络安全的要求。深度学习作为目前机器学习和人工智能的前沿技术，在语音识别、计算机视觉、大数据处理等方面都取得了巨大成果，也为解决当前的入侵检测问题提供了一个新的思路。本文基于对传统入侵检测技术的研究，结合深度学习方法下的深度置信网络，提出了一种基于深度置信网络的入侵检测技术，根据入侵检测数据的特点对数据进行过取样和非[0, 1]区间的归一化，在深度置信网络的参数更新过程中，采用批梯度下降的可变学习率算法，加快了参数的更新过程，并在每批训练数据中，加入了对少类别标签的区分度，提高了准确率。实验证明，利用本文提出的方法，可以很好地提高入侵检测的准确率。

关键词

入侵检测，深度学习，限制玻尔兹曼机，深度置信网络

1. 引言

入侵检测作为一种积极主动的网络安全防御措施，在网络安全上应用已久。入侵检测通过对计算机网络或计算机系统中的若干关键节点主动地进行信息收集和分析，从中发现网络或系统中是否有违反安全策略的行为和遭到攻击的迹象，同时做出响应[1]。入侵检测从检测技术上可以分为误用检测技术和异常检测技术。误用检测通过对已知攻击行为进行特征提取，找出入侵和攻击的模式或规则库，通过对比规则库和未知行为来确定未知行为是否为攻击。异常检测则是通过建立主体正常行为模型，将攻击行为作为异常活动从大量的正常活动中区分出来，进而达到检测攻击的目的。随着大数据时代的到来，网络攻击的方式也日益多样化，新型网络攻击行为呈现出了海量化、复杂化的趋势，面对新型网络攻击的威胁，传统的基于误用和基于异常的检测技术在检测速度和准确率上表现出越来越多的瓶颈，无法满足现有网络安全的需求。近些年飞速发展的深度学习技术，在语音识别、计算机视觉、机器智能等方面都取得了巨大成果，入侵检测作为典型的分类问题，与深度学习技术的结合，为解决入侵检测难题提供了一个新的方向。本文通过对入侵检测和深度学习的分析，提出了一种利用深度置信网络来进行入侵检测的方法：首先根据入侵检测数据的特点对数据进行过取样和非[0, 1]区间的归一化，然后利用限制玻尔兹曼机栈对归一化后的数据进行无监督地特征学习，在特征学习后送入 BP 神经网络进行有监督的分类学习，利用数据集的类别标签对网络参数进行调整，最终使整个网络能较好地识别数据集中的入侵行为。

2. 相关工作

目前最常用的入侵检测算法可以归为三种，分别是误用检测算法、异常检测算法和人工智能检测算法。误用检测主要采用的算法是模式匹配、专家系统、决策树等。如燕山大学的陈传钩[2]提出的 AC-SA 串匹配算法，利用后缀自动机 SA 加快 AC 自动机的跳转，从而提高检测效率。朱俚治[3]提出的使用决

策树和决策系统的误用检测算法，通过对相似用户进行行为聚类，计算出用户行为的属性值，再由决策系统根据属性值进行决策。异常检测算法主要有支持向量机、聚类、关联规则等。如文献[4]提出的利用支持向量机和 Hadamard 矩阵，采用纠错编码解决多分类检测问题。文献[5]对聚类算法进行改进，采用多初始聚类中心和对聚类中心邻近数据进行预先搜索来减少计算量，加快了检测过程。使用人工智能的检测算法主要有基于免疫学算法的检测、基于遗传算法的检测和基于深度学习算法的检测。如文献[6]以克隆算法为基础，用 Aho-Corasick 多模式匹配机来改进免疫系统中负选择算法，取得了不错的效果。文献[7]提出了一种基于量子遗传算法的入侵检测库优化算法，利用量子旋转门更新染色体，逐代优化种群进而缩短检测时间。文献[8]提出的使用遗传算法改进神经网络，克服了神经网络学习阶段训练速度慢以及容易陷入局部最优的缺点。文献[9]提出了采用卷积神经网络对网络流中提取的特征进行学习，提取出高层特征，进而检测出僵尸云。

在以上这些检测方法中，基于误用的检测技术误报率低，检测速度快，但是不能检测出新型入侵行为。基于异常的检测技术可以检测出异常攻击，但是也存在误报率高，检测时间较长的缺点。基于人工智能的检测技术通过在检测过程中不断学习，所以具备检测新型攻击的能力，同时误报率低，检测速度快，是目前入侵检测技术的一个新的研究方向。本文提出的方法也正是人工智能的检测技术的一种，是在深度学习基础上改进的方法。

在利用深度学习进行入侵检测的研究中，有些侧重于将深度学习与其他技术结合，来进行入侵检测。如文献[10]采用深度学习中的自编码网格模型实现对网络特征的提取，通过 softmax 分类器对特征数据进行分类，在保证较高识别率的同时，也降低了误报率。文献[11]提出的基于深度置信网络的混合入侵检测模型，模型采用 5 层结构的 DBN 对特征进行学习处理，随后采用支持向量机(SVM)进行入侵的识别和分类，实验结果与传统的 SVM 和贝叶斯网络相比有较好的识别率。文献[12]提出了一种深度学习算法和集成学习算法 MDBoot 算法相结合的 DBN-MDBoot2 方法，将多个 DBN 的分类结果进行集成，以集成后的分类结果作为最终的检测结果，取得了不错的效果。有些从深度学习网络的内部分析，寻找对深度学习网络的改进，如文献[13]提出的采用神经元映射卷积神经网络(NPCNN)作为网络结构，用 ReLU 激活器作为非线性激活函数，采用 Adam 算法进行学习，具备较少的连接和参数，具有易于训练和泛化能力强的优点。文献[14]提出的一种基于优化数据处理的深度置信网络模型的入侵检测方法，该方法将经过概率质量函数(PMF)编码和 MaxMin 归一化处理后的数据用于 DBN 模型，在不破坏网络已学习知识的基础上，变化一个参数而固定其他参数，用交叉验证的方式来选择相对最优的 DBN 结构，以提高对未知攻击的检测。本文也是从入侵检测数据和深度置信网络的特点分析，提出了一些更加合理的预处理过程和参数训练方法。

3. 深度置信网络技术分析

深度置信网络是 Geoffrey Hinton 教授在 2006 年提出的，它是一种生成模型的神经网络，通过训练网络内部神经元之间的权重和偏置，让整个网络按照最大概率生成数据，结构上深度置信网络由受限玻尔兹曼机和 BP 神经网络构成。

3.1. 受限玻尔兹曼机

受限玻尔兹曼机(RBM)是玻尔兹曼机的一种特殊拓扑结构，起源于统计物理学，是一种基于能量函数的建模方法，用来描述变量之间的高阶相互作用[15]。RBM 是一个对称连接无自反馈的两层网络模型，包含了一个可见层和一个隐含层。对称连接是指可见层和隐含层之间的节点以全连接方式连接，无自反馈是指可见层内部各节点之间以及隐含层内部各节点之间无连接。如图 1 所示。

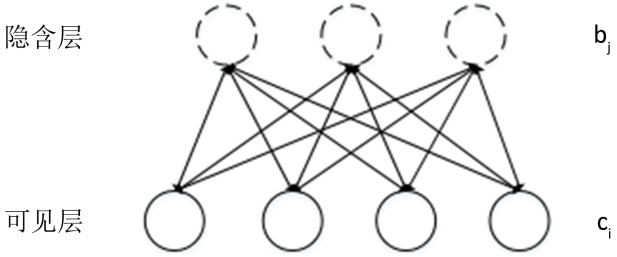

Figure 1. Structure graph of RBM
图 1. RBM 结构图

RBM 的隐含层节点的取值为 0/1 随机数，根据可见层节点取值，RBM 可以分为伯努利-伯努利受限玻尔兹曼机(Bernouli-Bernouli RBM)和高斯-伯努利受限玻尔兹曼机(Gaussian-Bernouli RBM)。前者的可见层节点的取值为 0/1 随机数，后者的可见层节点的取值为服从高斯分布的随机数。RBM 具有以下 3 个参数：

可见层和隐含层之间的权重矩阵 $w_{n,m}$；

隐含层节点偏移量 $b_j\left(j=1,2,\cdots,n\right)$；

可见层节点偏移量 $c_i\left(i=1,2,\cdots,m\right)$；

通过以上这些参数，RBM 决定了如何将一个 m 维的输入 $\left(x_1,x_2,\cdots,x_m\right)$ 编码成 n 维 $\left(y_1,y_2,\cdots,y_n\right)$ 输出，进而就实现了对原始数据的特征转换，当隐含层节点个数少于可见层节点个数时，就可以实现对原始数据的特征提取。

3.2. BP 神经网络

反向传播(BP)神经网络是一种多层的前馈神经网络，由 Rumelhart 等人于 1986 年提出[16]，BP 神经网络的网络结构由输入层、输出层以及一个或多个中间的隐含层组成，每层由若干个节点组成，网络只有在相邻层节点之间有连接。只有一层隐含层的 BP 神经网络的结构如图 2 所示。

在 BP 神经网络中，隐含层和输出层的每个节点都有一个偏置 b，输入层和隐含层之间以及隐含层和输出层之间的任意两个节点之间都有一个权重 w，用来表示节点之间的关联强度，输入层节点用来接收原始数据，对于一组输入向量，通过输入层和隐含层之间的权重和隐含层的偏置，计算出隐含层各节点的输出；同样，隐含层的输出再送给输出层，通过输出层和隐含层之间的权重以及输出层的偏置，计算出输出层各节点的输出，并由输出层节点输出和预期输出的误差，来反向调整节点之间的权重和偏置[17]；最终使输出和预期的误差达到可接受的程度。

BP 神经网络的特点是结构简单，可调整的参数多，操作性好。在深度置信网络中，BP 网络通常作为最后一层，接收上层受限玻尔兹曼机的输出结果，同时在数据标签的指导下进行有监督学习。

3.3. 深度置信网络

深度置信网络(DBN)是由若干个限制玻尔兹曼机层式堆叠，再在最后一层的 RBM 后面加上一层 BP 神经网络组成，结构如图 3 所示。

每层 RBM 的作用是实现对原始输入数据的特征提取，通过 RBM 堆叠形成的 RBM 栈，最终实现高层次的、抽象特征的提取。使用深度置信网络解决分类问题的步骤如下：首先，将训练数据输入到 RBM 栈，通过对比散度(CD)算法[18]对 RBM 栈的参数进行训练，实现对数据逐层地特征提取。然后将 RBM 栈的输出结果送入到 BP 神经网络，计算 BP 神经网络的输出，BP 网络的输出值即为网络对原始数据类别的判断，再与原始数据本身的标签进行比对，两者之间的误差用来对 BP 网络的参数进行优化[19]，最后，把测试数据送入优化好的深度置信网络中，得到最终的预测分类结果。

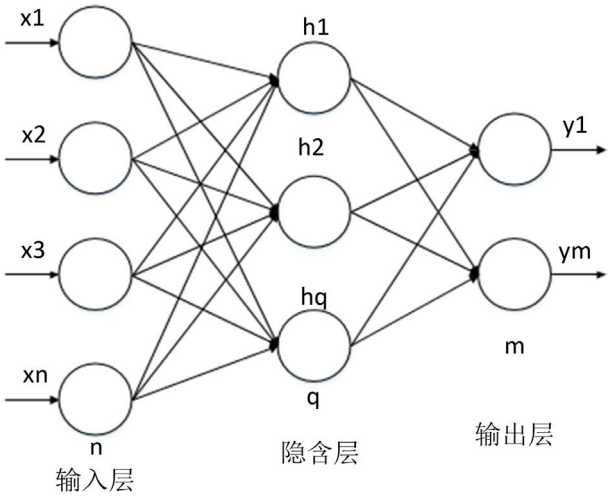

Figure 2. Structure graph of BP neural network
图 2. BP 神经网络结构图

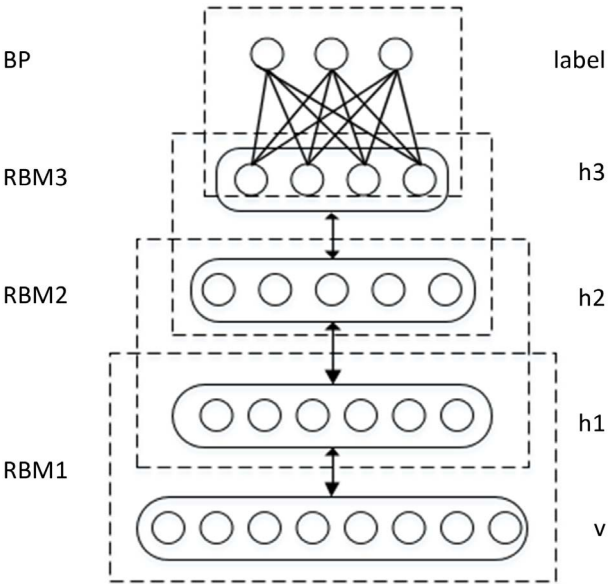

Figure 3. Structure graph of DBN
图 3. DBN 结构图

对比散度算法的过程如表 1 所示。其基本思想是：将可见层节点设置为一个训练样本的取值，并经过随机初始化后的 RBM 计算得到一个隐含层样本，再有该隐含层样本反向计算得到可见层样本，通过对比训练样本和得到的可见层样本的差异来更新 RBM 的网络参数。

4. 基于深度置信网络的入侵检测方法

4.1. 深度置信网络用于入侵检测的特点

入侵检测系统一般部署在局域网网络环境中，用于保护该网络，及时检测和发现网络中的入侵行为，这就要求入侵检测系统要实时地对网络环境中进出的大量数据包进行检测。

首先涉及到检测速度问题：传统基于误用的检测方法事先在规则库中存有大量入侵判断的规则，检测时通过抓取网络数据包，并逐一与规则库中的全部规则进行对比，来确定该数据包是否属于入侵及入

Table 1. Contrastive Divergence algorithm
表 1. 对比散度算法

Algorithm. Contrastive Divergence (CD 算法)

Input: RBM（$V_1,\cdots,V_m,H_1,\cdots,H_n$），training batch S

Output: $\Delta w_{ij},\Delta b_j,\Delta c_i$, for $i=1,\cdots,n$; $j=1,\cdots,m$

init $\Delta w_{ij}=\Delta b_j=\Delta c_i=0$ for $i=1,\cdots,n$; $j=1,\cdots,m$

for all the $v\in S$ do

$v^{(0)}\leftarrow v$ do

for $t=0,\cdots,k-1$ do

for $i=1,\cdots,n$ do sample $h_i^{(t)}\sim p\left(h_i\mid v^{(t)}\right)$

for $j=1,\cdots,m$ do sample $v_j^{(t+1)}\sim p\left(v_j\mid h^{(t)}\right)$

for $i=1,\cdots,n$, $j=1,\cdots,m$ do

$\Delta w_{ij}\leftarrow\Delta w_{ij}+p\left(H_i=1\mid v^{(0)}\right)\cdot v_j^{(0)}-p\left(H_i=1\mid v^{(k)}\right)\cdot v_j^{(k)}$

$\Delta b_j\leftarrow\Delta b_j+v_j^{(0)}-v_j^{(k)}$

$\Delta c_i\leftarrow\Delta c_i+p\left(H_i=1\mid v^{(0)}\right)-p\left(H_i=1\mid v^{(k)}\right)$

侵的类别，检测的时间复杂度按规则库中规则数量的幂基数增长。当网络流量不大时，基于误用的检测方法速度还能满足检测要求，但是当网络流量增大到一定水平时，检测速度就无法实时对所有数据包进行检测，甚至出现丢包现象。而基于深度置信网络的检测方法则不存在这种问题，因为基于深度置信网络的检测方法中没有规则库，而只有一个训练好的网络，网络的输入是预处理后的数据包，网络的输出是数据包的入侵类别，即使网络流量成倍增加，检测的时间复杂度也是同网络流量相等速率的增加，检测速度的上限较误用检测方法有了极大的提升。

其次是检测数据的属性问题，在基于误用的检测方法和基于异常的检测方法中，均要处理数据包中字符取值的属性，这就要用到字符串匹配算法，就不可避免的降低了数据包的处理速度，而基于深度置信网络的检测算法中不存在该问题，因为在将数据包送入深度置信网络计算之前，已经经过了预处理过程，将字符取值的属性转化为了合理取值范围内的数值型数据，网络只需要进行简单的数值数据的计算，这也在一定程度上提高了深度置信网络算法的检测速度。

当然基于深度置信网络的检测方法也有需要特殊处理的地方。入侵检测数据是从实际数据包中提取出的原始特征，不同于图片识别、视频处理这些场景仅涉及 0~255 之间的数值型数据，数据包中的数据特征有数值型和非数值型，离散型和连续型之分，并且不同特征的取值范围差异很大，这就导致需要先对数据进行数值化，归一化等预处理，以屏蔽量纲上的差异。其次，入侵检测数据集是从实际网络环境中抓取到的数据，是对真实网络环境的反映，所以不可避免地存在正常数据较多，而异常数据相对较少的问题，因此也需要对数据分布差异较大这一问题做特殊处理。

4.2. 少数类标签处理

在用网络数据包进行深度置信网络训练时，存在一个数据集中的类别数目差异问题，当某类标签的数据在全部数据集中的比重很大，而有些标签的数据所占比重很小时，少类别的数据往往会被预测为多类别的标签。本文采用过取样的处理方法：对少类别的样本，通过复制并加入随机噪音的方法产生出较多的新样本，将新生成的样本和原始样本一起送入网络进行训练。对多类别的样本没有采取删除部分样本的欠取样方式，是为了避免可能会丢失未知信息。在过取样之前，训练样本的类别分布如表 2 所示。

过取样后，所有数目少于 600 的类别都通过加入随机噪音的方法将数目增加到 600。

Table 2. Number distribution of training data set
表 2. 训练数据集类别数量分布

入侵类别	数量	入侵类别	数量
Normal	13416	guess_password	10
neptune	8262	ftp_wirte	1
warezclient	180	multihop	2
ipsweep	709	Rootkit	4
Portsweep	586	buffer_overflow	6
Teadrop	187	imap	5
namp	301	warezmaster	7
satan	689	Phf	2
smurf	529	land	1
Pod	38	LoadModule	1
Back	196	spy	1

4.3. 数据归一化处理

在大多数深度学习的实验中，都是将训练数据和测试数据先数值化再归一化，且一般都是归一化到[0, 1]区间，但是在实际操作中，训练数据和测试数据在一个具体的特征上，其取值范围是不同的，尤其是测试集中出现大量训练集从未出现的新记录时，这样就会导致对于一个特征，同一个数值分别在训练集和测试集归一化后的取值是不同的。如果特征原始值为字符型，经过这样就会导致同一字符在训练集和测试集中归一化后的取值也不相同，这显然不合逻辑。比如对于某一个字符型特征，训练集中可能的取值范围为(A, B, C, D, E)，测试集中的取值范围为(A, B, C, D, E, F)。对训练集数值化再归一化，可以得到对应的(0,0.25, 0.5, 0.75, 1.0)，对测试集数值化再归一化，可以得到(0, 0.2, 0.4, 0.6, 0.8, 1.0)，可以看出，对于同一个取值为 B 的特征值，它在训练阶段和测试阶段参与计算的取值分别为 0.25 和 0.2，这样的计算结果就不严谨，对于这种情况，本文采取的方法是不归一化到[0, 1]区间，而是根据训练集的取值区间来决定测试集的取值区间。步骤如下：

1) 训练集数据的非数值特征数值化；

2) 对数值化后训练集所有的特征归一化，归一化到[0, 1]区间；

3) 测试集数据的非数值特征数值化；

4) 按照训练集中所有特征的取值把测试集中的相同特征进行归一化，测试集中新出现的特征按比例归一化到[0, 1]区间外；

虽然当测试集的取值区间比训练集大时，测试集的归一化区间会超过[0, 1]，但是保证了在测试集出现新取值时，训练集和测试集归一化结果是一致的。比如上例中，测试集归一化后的结果是(0, 0.25, 0.5, 0.75, 1.0, 1.25)。这样只有测试集中新出现的 F 使用了新的数值 1.25，而其他的 A，B，C，D，E 均保留了和训练集中一样的取值。

4.4. 整体结构

基于深度置信网络的入侵检测方法采用若干层(由实验确定)RBM 堆叠构建 RBM 栈，在 RBM 栈的最后加上一个有监督的 BP 神经网络，以最后一层 RBM 的输出作为 BP 神经网络的输入，以训练集数据的标签和 BP 输出的均方差作为损失函数，训练 BP 网络，使整个 DBN 网络具备分类的能力。

整个模型的训练过程如下：

先对 RBM 堆叠栈使用贪心无监督方法逐层构建，训练集在不使用标签的情况下被用来训练 RBM 栈的参数空间 $\theta(w,b,c)$，这一过程也叫作无监督学习阶段。

使用梯度下降算法优化 BP 神经网络，训练集在附带标签的情况下用来训练 BP 神经网络的参数空间 $\theta(w,b)$。这一过程也叫作有监督学习阶段。

4.5. 参数训练方法

基于深度置信网络的入侵检测方法中存在两处参数训练的过程，分别是无监督学习阶段的参数训练和有监督学习阶段的参数训练，无监督学习阶段主要是对 RBM 栈的参数进行训练，使 RBM 栈能准确地对原始数据进行特征提取，方法是对比散度算法；有监督学习阶段训练的是 BP 网络的参数，使 BP 网络能在类别标签的指导下准确识别出特征提取后的数据包类别。在这两个参数训练过程中，均存在一定的缺陷，传统的对比散度算法每次只训练一个样本数据，迭代次数多，效率较低；BP 算法容易陷入局部最小值，收敛速度慢[20]，以及没有兼顾少类别标签，针对这些问题，本文提出了一种新的方法来训练参数。

首先在无监督学习阶段，采用批梯度下降的对比散度算法来加快参数的训练过程。将训练集随机划分为若干批次，不再针对每个数据进行参数更新，而是以批次为单位，计算每批次的权重和偏置的更新量，更新公式如下：

权值的修正量：

$$\Delta w = \alpha \Delta w + \eta\left(c_1 - c_2\right)/n \tag{1}$$

隐含层偏置的修正量：

$$\Delta c = \alpha \Delta c + \eta\left(h_1 - h_2\right)/n \tag{2}$$

可见层偏置的修正量：

$$\Delta b = \alpha \Delta b + \eta\left(v_1 - v_2\right)/n \tag{3}$$

其中 $c_1 = v_1 h_1$，$c_2 = v_2 h_2$，v_1 是对比散度算法中可见层第一次取值，h_1 是根据 v_1 计算出的隐含层第一次取值，v_2 是根据 h_1 计算出的可见层第二次取值，h_2 是根据 v_2 计算出的隐含层第二次取值，n 是每批次中包含的样本数目，α，η 分别为学习动量和学习率。

在有监督学习阶段，采用批梯度区分下降的可变学习率算法来训练参数，误差计算公式如下：

输出层的误差：

$$\delta_k = o_k\left(1 - o_k\right)\left(t_k - o_k\right) \tag{4}$$

隐含层的误差：

$$\delta_j = o_j\left(1 - o_j\right)\sum_k \delta_k w_{kj} \tag{5}$$

对于每批次，输出层误差为：

$$\delta_o = \sum_1^n \delta_k \rho_k \tag{6}$$

$\rho_k = n_k \cdot N/n \cdot N_k$ 隐含层误差为：

$$\delta_h = \sum_1^n \delta_j \rho_k \tag{7}$$

其中 $\rho_k = n_k \cdot N/n \cdot N_k$ 表示不同类别样本在误差度量上的区分度，n_k 表示在该批次中 k 所代表的样本类别出现的次数，n 是该批次的总样本数，N_k 表示 k 所代表的类别在整个训练样本中出现的次数，N 是整个样本

的总数目。通过加入区分度可以提高少类别样本误差在批度误差中的比重，进而加强对少类样本的拟合。

根据上面求出的训练误差，权值和偏置的修正量为：

$$\Delta w = \alpha \Delta w + \eta \delta_o o_j \tag{8}$$

$$\Delta c = \alpha \Delta c + \eta \delta_h \tag{9}$$

上面两个过程中 α，η 分别为学习动量和学习率，学习动量表示每一次权重和偏置的变化量受上一次变化量的影响程度，学习率表示权重和偏置的变化量随误差更新的步长。α 和 η 的更新规则如下：

如果训练误差在权值更新后增加了，并且超过了设定的一个阈值 ξ (根据经验 ξ 可以取 3%到 10%之间，一般取 5%)，则权值更新被取消，学习速率 η 被乘以一个因子 $\rho(0<\rho<1)$，动量系数 α 被设置为 0。

如果训练误差在权值更新后减小了，则权值更新被接受，学习速率 η 被乘以一个因子 $\rho'(\rho'>1)$，如果动量系数 α 被设置为 0，则恢复到 0 之前的值；如果 α 没有被设置为 0，则将 α 也乘以因子 ρ'。

如果训练误差在权值更新后增加了，但是增加量小于设定阈值 ξ，则此次权值更新被接受(之所以在均方误差增加的情况下仍然接受取值更新，是为了尝试跳出局部最小值点)，学习速率保持不变，如果动量系数 α 被设置为 0，则恢复到 0 之前的值；如果 α 没有被设置为 0，则将 α 乘以一个因子 $\rho(0<\rho<1)$。

5. 实验

5.1. 实验环境和数据源

本实验采用 NSL-KDD 入侵检测数据集，NSL-KDD 数据集是在 KDD CUP99 数据集的基础上再次优化得到的。和 KDD CUP99 相比，主要有以下几点改进：

1) 训练集中不存在冗余记录；

2) 数据集中每个复杂等级的记录数量和原数据集中记录所占的百分比成反比；

3) 训练集和测试集的数量设置更加合理；

在 NSL-KDD 数据集中，每行代表一条记录，每条记录包含了从一条连接中提取的 41 个特征和攻击类型，如下所示为一条记录，记录的每个特征之间用逗号隔开。

0,tcp,private,S0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,123,6,1.00,1.00,0.00,0.00,0.05,0.07,0.00,255,26,0.10,0.05,0.00,0.00,1.00,1.00,0.00,0.00,neptune,19

其中“neptune”表示攻击类型，neptune 攻击是一种拒绝服务攻击，发生该攻击时目标主机会收到大量的 SYN 报文，进而影响目标主机的正常运行。最后一列的 19 表示复杂等级，一般在检测过程中不使用。41 个特征按照内容可以分为 4 大类，分别是：

1) TCP 连接基本特征(共 9 种)

duration、protocol_type、service、flag、src_bytes、dst_bytes、land、wrong_fragment、urgentl；

2) TCP 连接的内容特征(共 13 种)

hot、num_failed_logins、logged_in、num_compromised、root_shell、su_attempted、num_root、num_file_creations、num_shells、num_access_files、num_outbound_cmds、is_hot_login、is_guest_login；

3) 基于时间的网络流量特征统计(共 9 种)

count、srv_count、serror_rate、srv_serror_rate、rerror_rate、srv_rerror_rate、same_srv_rate、diff_srv_rate、srv_diff_host_rate；

4) 基于主机的网络流量特征统计(共 10 种)

dst_host_count、dst_host_srv_count、dst_host_same_srv_rate、dst_host_diff_srv_rate、dst_host_same_src_port_rate、dst_host_srv_diff_host_rate、dst_host_serror_rate、dst_host_srv_serror_rate、

dst_host_rerror_rate、dst_host_srv_rerror_rate；

本文实验的硬件环境：windows 7 操作系统，Intel i3 2.1 GHz 双核 CPU，10 GB 内存。软件环境：Java jdk8.0。

5.2. 数据预处理

1) 数值化：数据集中存在协议类型和服务类型等字符型特征，需要先转化为数值，采用的方法为整数映射。

2) 归一化：由于不同的特征值，采用的度量单位不同，数值差异较大，为了消除这种数值差异巨大带来的不利因素，需要进行归一化处理，采用的方法为改进后的[0, 1]归一化，如上 4.3 所述。

5.3. 实验流程

实验主要分为两步：一是使用训练数据对深度置信网络训练的过程，主要是对网络中的各个参数进行更新，二是利用步骤一得到的训练好的网络，对测试数据进行预测，并把预测值和测试数据期望的真实值进行对比，统计出预测结果。实验流程图如图 4 所示。

5.4. 实验结果分析

本文实验采用国际上通用的入侵检测评价指标：包括检测准确率(简称准确率)，误报率和漏报率[21]，对实验进行有效性评价。定义 TN 为数据集中被正确识别出的正常数据数量，TP 为数据集中被正确识别出的入侵数据数量，FN 为数据集中入侵被识别为正常数据的数量，FP 为数据集中正常数据被识别为入侵的数量，则有：

检测准确率(Dectection Precision Rate) = (TN + TP)/(TN + TP + FN + FP) × 100%；

误报率(False Positive Rate) = FP/(TN + FP) × 100%；

漏报率(Flase Negative Rate) = FN/(TP + FN) × 100%；

为了寻找深度置信网络在入侵检测数据集上的最佳效果，分别作了以下实验：

1) 为了寻找无监督学习迭代次数对检测结果的影响，采用 3 隐含层结构，即 RBM 栈的结构数目分别为{41, 32, 23, 14, 5}，无监督学习迭代次数 N 分别取(5, 10, 30, 50, 70, 90)，有监督学习的迭代次数固定为 30 次，对比检测后的准确率如表 3 和图 5 所示。由此可以看出：随着无监督迭代次数的增加，准确率呈现出先上升后下降的趋势，当无监督学习迭代次数较少时特征提取不明显，随着迭代次数的增加特征

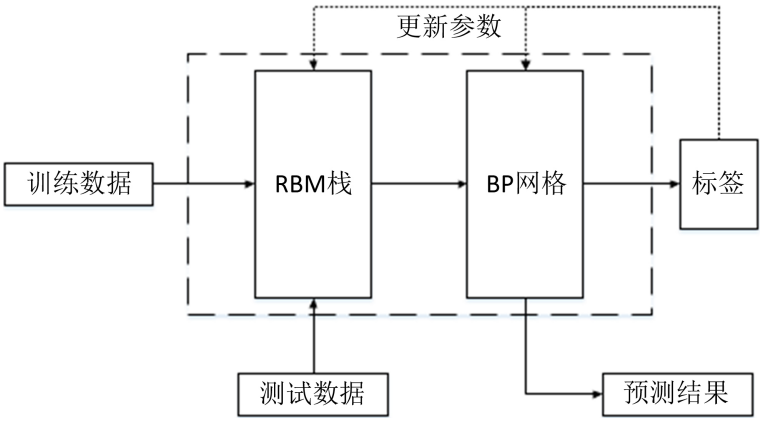

Figure 4. Intrusion detection experimental process
图 4. 入侵检测实验流程

提取效果提高，准确率上升，当迭代次数超过一定的值时，准确率反而下降，可能原因是特征提取过度而导致了丢失了某些特征信息。

2) 为了寻找有监督学习迭代次数对检测结果的影响，仍然采用 3 隐含层结构，RBM 栈的结构依然为{41, 32, 23, 14, 5}，无监督学习迭代的次数 N 固定为 50，有监督学习迭代次数 n 取值分别为(5, 10,30, 50, 70, 90)，统计检测准确率如表 4 和图 6 所示。由此可以看出，随着有监督学习迭代次数的增加，准确率呈上升趋势，且在迭代次数较低时，上升速度较快，随着迭代次数的增多，上升逐渐趋于平缓。

3) 由 1)和 2)两个实验的结果，可以看出在无监督学习迭代次数和有监督学习迭代次数分别为 50、50 时，深度置信网络模型能取得很好的效果，为了寻找 DBN 的最佳结构，将无监督和有监督迭代次数固定为(50, 50)，同时将 DBN 的结构分别置为{41, 32, 23, 14, 5}和{41, 70, 100, 80, 60, 40, 20, 5}进行实验，统计准确率如表 5 所示。

为了验证本文提出算法在入侵检测上的效果，将深度置信网络的网络结构定为{41, 32, 23, 14, 5}，无监督和有监督学习均定为 50 次，与常见的入侵检测方法 KNN，SVM，BP 神经网络算法[22]-[27]进行了对比实验。实验结果如表 6 和图 7 所示。

由以上结果可以看出，基于深度置信网络(DBN)的入侵检测方法在准确率、误报和漏报方面都明显

Table 3. Unsupervised iteration number and accuracy
表 3. 无监督迭代次数与准确率

迭代数 N	准确率	迭代数 N	准确率
5	97.71%	50	97.94%
10	97.84%	70	97.83%
30	97.88%	90	97.81%

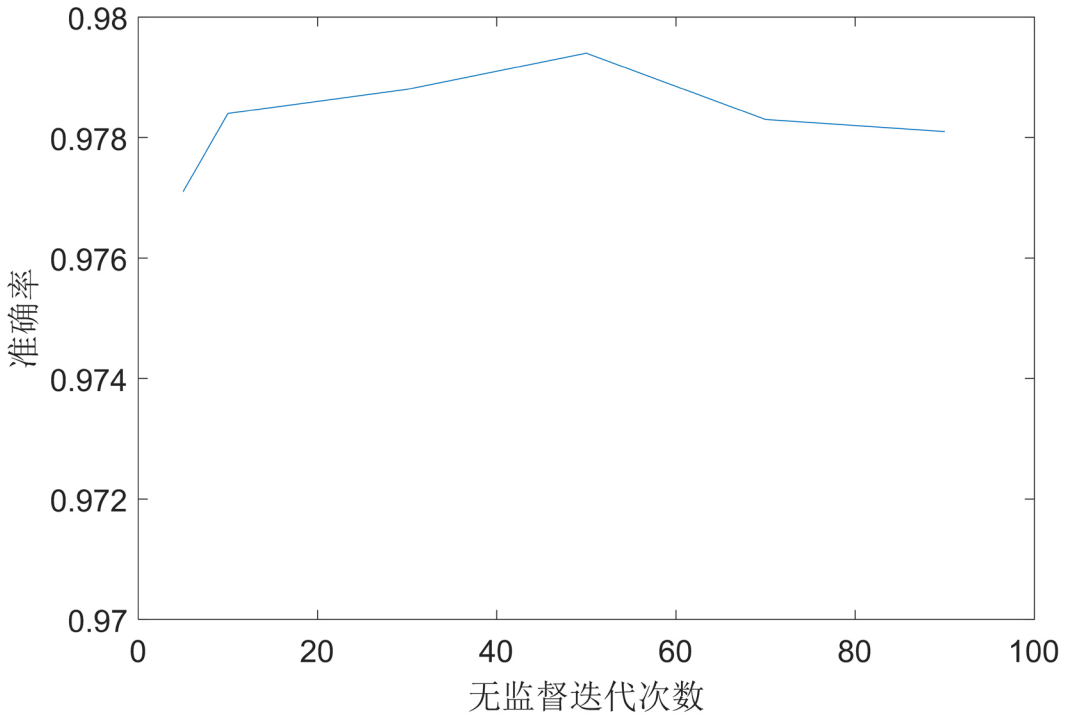

Figure 5. Intrusion detection experimental process
图 5. 入侵检测实验流程

Table 4. Supervised iteration number and accuracy
表 4. 有监督迭代次数与准确率

迭代数 n	准确率	迭代数 n	准确率
5	95.62%	50	98.19%
10	97.50%	70	98.39%
30	97.94%	90	98.44%

Table 5. Network structure and accuracy
表 5. 网络结构和准确率

网络结构	准确率
{41, 32, 23, 14, 5}	98.19%
{41, 36, 31, 26, 21, 16, 11, 5}	90.68%

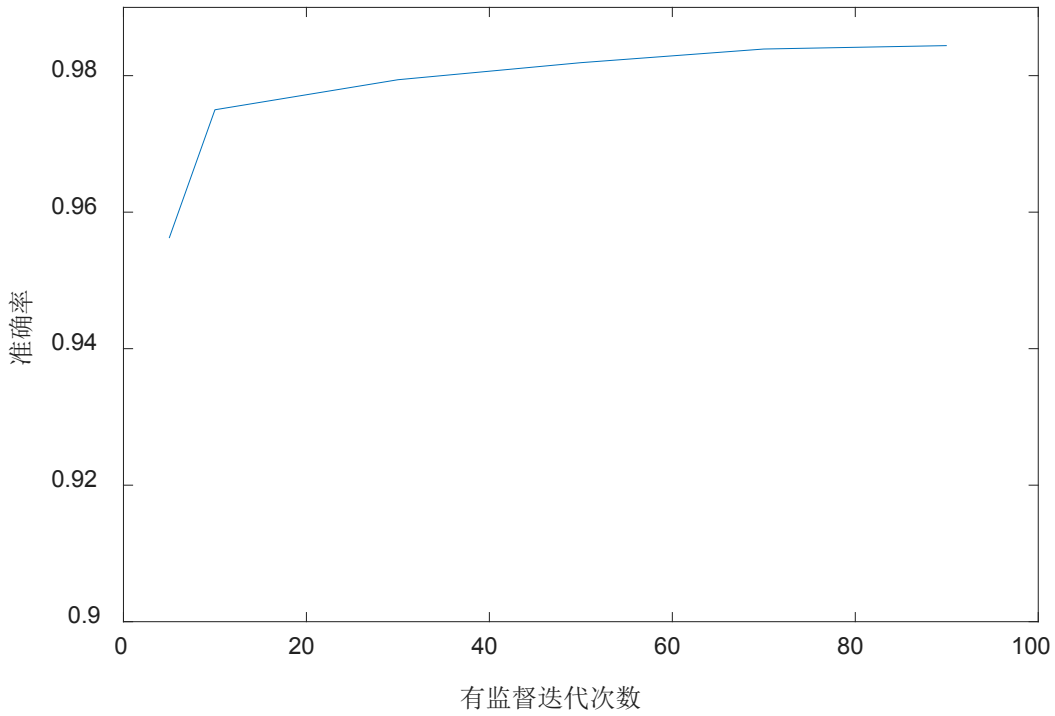

Figure 6. Intrusion detection experimental process
图 6. 入侵检测实验流程

优于 KNN 和 BP 网络，在误报率上和 SVM 相差不大，但准确率和漏报率也优于 SVM 方法。

为了验证本文提出的参数训练方法的有效性，将本文的方法同未改进的深度置信网络方法进行了对比实验，将两种深度置信网络的网络结构均设置为{41, 32, 23, 14, 5}，无监督学习和有监督学习次数均设置为(10, 10)，(20, 20)，(30, 30)，(40, 40)，(50, 50)时，对比两种方法的准确率。实验结果如表 7 和图 8 所示。

由以上结果可以看出，本文的参数训练方法在无监督和有监督学习迭代次数较低时就可以到达较好的准确率，比未改进的 DBN 参数训练方法有较大的优势。

Table 6. Comparison of the results of several detection algorithms
表 6. 几种检测算法的结果对比

模型	准确率	误报率	漏报率
KNN	76.74%	0.021	0.218
SVM	96.30%	0.004	0.071
BP	85.93%	0.015	0.202
DBN	98.19%	0.004	0.033

Table 7. Result contrast between DBN in this article and the unimproved DBN
表 7. 本文 DBN 和未改进 DBN 的结果对比

迭代次数(N,n)	本文方法准确率	未改进 DBN 方法准确率
(10, 10)	97.01%	91.76%
(20, 20)	97.43%	93.25%
(30, 30)	97.70%	93.96%
(40, 40)	97.85%	94.23%
(50, 50)	98.19%	94.77%

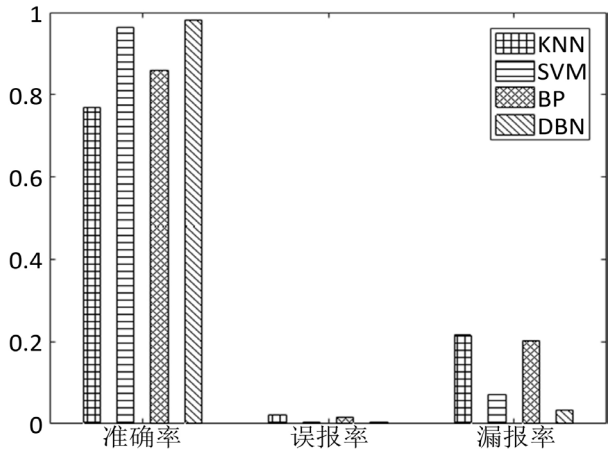

Figure 7. Comparison of the results of several detection algorithms
图 7. 几种检测算法的结果对比

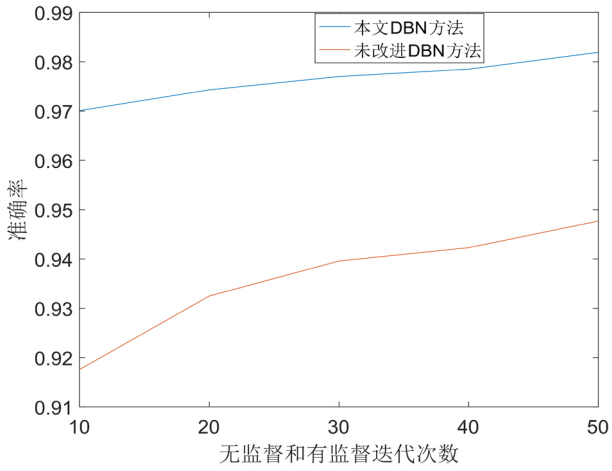

Figure 8. Intrusion detection experimental process
图 8. 入侵检测实验流程

6. 总结

本文通过对入侵检测和深度置信网络的研究，利用深度学习在处理分类问题上的优势，将深度置信网络运用在入侵检测问题上，设计出了基于深度置信网络的入侵检测方法，结合入侵检测数据的特点对数据进行过取样和非[0, 1]区间的归一化，在对深度置信网络的参数进行更新过程时，采用批梯度下降的可变学习率算法，加快了参数的更新过程，并在每批训练数据中，加入了对类别的区分度，提高检测的准确率。通过实验先确定了深度置信网络的参数最佳选择，包括无监督学习次数、有监督学习次数以及网络结构，然后同 KNN，SVM，BP 神经网络等方法进行了对比，在准确率，误报率和漏报率等几个常用的入侵检测指标上都有较好的表现，证明了本文所提方法的有效性。但是本方法也存在一些缺点，比如在网络结构参数的选择上，不同的参数会存在一定的影响，同时参数选择也缺乏理论的支持，这些是需要改进的地方。

参考文献

[1] 刘文涛. Linux 网络入侵检测系统[M]. 北京: 电子工业出版社, 2004: 2-19.

[2] 陈传钧. 基于模式匹配的入侵检测研究[D]: [硕士学位论文]. 秦皇岛: 燕山大学, 2006.

[3] 朱俚治. 一种基于决策系统和决策树的误用检测算法[J]. 计算机与数字工程, 2016, 44(12): 2353-2355 + 2391.

[4] 赵伟. 基于 SVM 的入侵检测研究[D]: [硕士学位论文]. 北京: 北京交通大学, 2007.

[5] 杜强. 基于改进聚类分析算法的 IDS 模型构建[D]: [硕士学位论文]. 太原: 山西大学, 2011.

[6] 许铭. 基于免疫机理的入侵检测系统的研究[D]: [硕士学位论文]. 淮南: 安徽理工大学, 2010.

[7] 张宗飞. 量子遗传算法在网络误用检测中的应用[J]. 计算机工程与设计, 2010, 31(12): 2933-2935 + 2939.

[8] 屈洪春, 王帅. 一种基于进化神经网络的混合入侵检测模型[J]. 计算机科学, 2016, 43(S1): 335-338.

[9] 寇广, 汤光明, 王硕, 宋海涛, 边媛. 深度学习在僵尸云检测中的应用研究[J]. 通信学报, 2016, 37(11): 114-128.

[10] 李春林, 黄月江, 王宏, 牛长喜. 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.

[11] 杨昆朋. 基于深度学习的入侵检测[D]: [硕士学位论文]. 北京: 北京交通大学, 2015.

[12] 蔡之鑫. DBN 和 MDBoost2 在入侵检测中的应用[D]: [硕士学位论文]. 广州: 广东工业大学, 2016.

[13] 钱铁云, 王毅, 张明明, 刘俊恺. 基于深度神经网络的入侵检测方法[J]. 华中科技大学学报(自然科学版), 2018, 46(1): 6-10.

[14] 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(6): 1636-1643 + 1656.

[15] Nicolas, L.R. and Yoshua, B. (2008) Representational Power of Restricted Boltzmann Machines and Deep Belief Networks. *Neural Computation*, 20, No. 6.

[16] Rumelhart, D.E. (1986) Learning Representation by BP Errors. *Nature*, **7**, 64-70.

[17] van der Smagt, P.P. (1994) Minimisation Method for Training Feed forward Neural Network. *Neural Networks*, **7**, 1-11.

[18] Hinton, G.E. (2007) Learning Multiple Layers of Representation. *Trends in Cognitive Sciences*, **11**, 428-434. https://doi.org/10.1016/j.tics.2007.09.004

[19] Behera, L., Kumar, S. and Patnaik, A. (2006) On Adaptive Learning Rate That Guarantees Convergence in Feed forward Networks. *IEEE Transactions on Neural Networks*, **17**, 1116-1125.

[20] Bengio, Y., Lamblin, P., Popovici, D., *et al*. (2007) Greedy Layer-Wise Training of Deep Networks. *Advances in Neural Information Processing Systems*, **19**, 153.

[21] Gafney, J.E. and Ulvila, J.W. (2001) Evaluation of Intrusion Detectors: A Decision Theory Approach. *Proceedings IEEE Symposium on Security and Privacy*, Oakland, 14-16 May 2000.

[22] Ghosh, P., Shakti, S. and Phadikar, S. (2016) A Cloud Intrusion Detection System Using Novel PRFCM Clustering and KNN Based Dempster-Shafer Rule. *International Journal of Cloud Applications and Computing*, **6**, 18-35.

[23] Yu, Q., Wang, S., Wang, J.L. and Zhang, B.H. (2011) Research for SVM with Self-Reacting Feature Weighted in IDS. *Advanced Materials Research*, **204-210**, 604-607. https://doi.org/10.4028/www.scientific.net/AMR.204-210.604

[24] Song, J.H., Zhao, G. and Song, J.Y. (2013) Research on Property and Model Optimization of Multiclass SVM for NIDS. *Applied Mechanics and Materials*, **347**, 616-619.

[25] Xu, J. (2013) IDS Method Based on Improved SVM Algorithm under Unbalanced Data Sets. Springer, New York.

[26] Wei, M., Su, J., Jin, J. and Wang, L. (2014) Research on Intrusion Detection System Based on BP Neural Network. Springer, Berlin Heidelberg.

[27] Yuan, J.S. and Wang, Y. (2013) The Development of Intrusion Detection System Based on Improved BP Neural Network. *Advanced Materials Researc*h, **718-720**, 1973-1979.