基于自编码网络特征降维的轻量级入侵检测模型

高 妮[1,2],高 岭[1],贺毅岳[3],王 海[1]
(1. 西北大学信息科学与技术学院,陕西西安 710069;2. 西安财经学院信息学院,陕西西安 710100;
3. 西北大学经济管理学院,陕西西安 710127)

摘 要: 基于支持向量机(SVM)的入侵检测方法受时间和空间复杂度约束,在高维特征空间计算时面临"维数灾害"的问题.为此,本文提出一种基于自编码网络的支持向量机入侵检测模型(AN-SVM).首先,该模型采用多层无监督的限制玻尔兹曼机(RBM)将高维、非线性的原始数据映射至低维空间,建立高维空间和低维空间的双向映射自编码网络结构,进而运用基于反向传播网络的自编码网络权值微调算法重构低维空间数据的最优高维表示,从而获得原始数据的相应最优低维表示;最后,采用SVM分类算法对所学习到的最优低维表示进行入侵识别.实验结果表明,AN-SVM模型降低了入侵检测模型中分类的训练时间和测试时间,并且分类效果优于传统算法,是一种可行且高效的轻量级入侵检测模型.

关键词: 特征降维;自编码网络;限制玻尔兹曼机;支持向量机;入侵检测
中图分类号: TP393.08 **文献标识码:** A **文章编号:** 0372-2112 (2017)03-0730-10
电子学报URL: http://www.ejournal.org.cn **DOI:** 10.3969/j.issn.0372-2112.2017.03.033

A Lightweight Intrusion Detection Model Based on Autoencoder Network with Feature Reduction

GAO Ni[1,2], GAO Ling[1], HE Yi-yue[3], WANG Hai[1]
(1. *School of Information Science and Technology, Northwest University, Xi'an, Shaanxi* 710127, *China*;
2. *School of Information, Xi'an University of Finance and Economics, Xi'an, Shaanxi* 710100, *China*;
3. *School of Economics and Management, Northwest University, Xi'an, , Shaanxi* 710100, *China*)

Abstract: Owing to the constraints of time and space complexity, support vector machine (SVM) faced with the problem of 'curse of dimensionality' when computation happens in high-dimensional feature space. Therefore, an intrusion detection model of support vector machine based on autoencoder network (AN-SVM) is proposed. First, the multilayer unsupervised restricted boltzmann machine (RBM) in our model is employed in mapping the vector of raw dada from high-dimensional nonlinear space to low-dimensional space, and a mutual mapping autoencoder network of high-dimensional space and low-dimensional space is constructed. Then autoencoder network weights of fine-tuning algorithm based on back propagation network is employed to reconstruct the new optimal high-dimensional representation of data in low-dimensional space, and the corresponding optimal low-dimensional representation of raw data can be obtained. Furthermore, SVM classification algorithm is employed to detect intrusion from the optimal low-dimensional data. The experimental results demonstrate that AN-SVM model can effectively reduce the training time and testing time of classifier in the intrusion detection model and its classification performance outperforms those traditional methods. So, AN-SVM model is a feasible and efficient lightweight intrusion detection model.

Key words: feature reduction; autoencoder network; restricted boltzmann machine; support vector machine; intrusion detection

1 引言

随着Internet在全世界范围内的迅速发展,计算机网络安全问题已成为一个备受关注的重大问题.其中,如何识别各种网络攻击是一个不可回避的关键技术.入侵检测(Intrusion Detection,ID)作为一种主动防御技术,逐渐成为保障网络系统安全的关键技术.入侵检测系统(intrusion detection systems,IDS)的目的就是识别

收稿日期:2015-06-12;修回日期:2015-08-24;责任编辑:郭游
基金项目:国家自然科学基金(No.61373176);教育部人文社会科学研究青年项目(No.16XJC630001);陕西省自然基金(No.2015JQ7278)

不同寻常的访问或对安全内部网络的攻击. 基于机器学习的用户行为建模是 IDS 的一个重要研究课题,即通过学习网络流量和主机审计记录等观测数据来区分系统的正常行为和异常行为.

以往的研究者在 IDS 研究中引入了各种机器学习方法,如神经网络[1]、K 最近邻算法[2]、SOM[3]等方法都在入侵检测系统中取得了突破性的进展. 而支持向量机(Support Vector Machine,SVM)最早由 Vapnik[4]提出,可有效地避免经典学习方法中出现的过学习、易陷入局部极小点等问题. 尚文利[5]和 Chitrakar[6]等人将 SVM 方法应用到 IDS 中,并证明 SVM 具有很好的分类性能.

虽然 SVM 在小样本下能够取得较好效果,但实际应用于大规模的入侵检测系统中时,通常受时间和空间复杂度约束,其本质原因是由于输入特征空间具有高维、非线性的特征. 分类器所需训练样本的数目是关于样本特征空间维数的近似指数级增长函数,在高维特征空间计算时 SVM 入侵检测面临"维数灾害"的问题. 当前 IDS 面临实时处理海量数据检测速度低的问题,检测速度是衡量入侵检测系统实时性要求的一个重要评估指标. 因此,构建高效的轻量级 IDS 已成为当前研究热点.

特征维数过多是导致 IDS 检测速度低的主要原因,很多研究者通过对高维、非线性特征空间进行约简来解决此问题. 因此,对高维数据进行特征降维成为入侵检测流程中不可或缺的步骤. Kuang 等人[7,8]混合核主成分分析(KPCA)方法与遗传算法(GA)、单独的 KPCA 对高维数据进行约简,进而利用 SVM 进行入侵识别. Ahmad[9]研究了混合 PCA 与遗传算法(GA)对高维数据进行约简,进而利用 SVM 进行入侵识别. Lakhina[10]等人研究了 PCA 与 ANN 相结合做异常识别. 上述研究在入侵检测系统中对输入特征空间进行降维,有效提高了分类的整体性能. 然而,当高维特征数据呈现非线性结构时,上述方法的主要缺陷在于它们只能学习到已知数据集的低维结构,不能给出高维空间到低维空间的确定性映射,且这些方法在特征子集的检测速度上有待提高. 因此,在保证高分类精度的前提下,尽量学习到最优的低维特征空间用于提高检测速度,是本文提出特征降维算法的目的.

2006 年,Hinton 在《Science》上发表一篇关于"自编码网络"深度学习方法的论文[11],该方法已成为大数据和人工智能的一个热潮. 2009 年以来微软研究人员通过与 Hinton 合作,首次将限制玻尔兹曼机(Restricted Boltzmann Machine,RBM)和深度信念网络(Deep Belief Networks,DBN)[12]引入到语音识别声学模型训练中,使得语音识别的错误率相对减低 30%. 2013 年百度采用多达 9 层的 DNN 模型,更好地解决了 DNN 在线计算的技术难题[13]. 从上述案例可以看出,深度学习在复杂大规模数据的处理方面有着出色的表现,因此它是解决入侵检测速度低的一种极具前景的方法.

本文将 Hinton 提出的非线性降维的自编码网络(Autoencoder Network,AN)的深度学习方法引入到入侵检测领域中,通过具有多个隐藏层的神经网络的逐层特征变换,将样本在原空间的高维特征转换成低维特征并进一步重构样本高维特征,在该非监督学习过程中获得原始数据的低维表示,从而显著降低数据的维数. 本文的工作重点不是分类器的参数优化,而是为分类器服务的特征降维算法. 本文提出一种高效的特征降维算法来构建轻量级 IDS,旨在缩短训练时间,提高入侵分类精度.

2　深度学习模型

2.1　自编码网络

自编码网络是有 Hinton 提出的一种用于学习高效编码的人工神经网络,通过学习获得数据集的压缩编码,可以达到数据降维的目的[11]. 自编码网络能使具体的特征向量逐渐转化为抽象的特征向量. 自编码网络能很好地满足高维数据空间和低维数据空间双向映射的非线性学习,它采用自适应、多层编码(encoder)网络将高维原始数据转换成低维抽象数据,并利用类似的解码(decoder)网络从低维抽象数据重构原始数据的高维数据表示[11].

自编码网络是一种非监督学习方法,其工作原理如图 1 所示,整个结构由编码器和解码器两部分构成. 编码器用于降维,解码器用于重构,其视为编码器的逆过程. 编码器和解码器之间还存在一个交叉部分,称为"代码层". 原始高维数据输入到编码器中,被压缩后表述为代码层(code layer),code layer 是输入数据的一个低维表示.

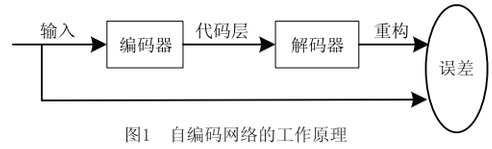

图1　自编码网络的工作原理

通过调整编码器和解码器的参数,使得重输出的重构数据和一开始的原始数据误差最小,意味着由抽象的特征向量构成的代码层数据是原输入数据的有效低维表示. 该过程中输入原始数据是无标签数据,相应地误差是通过对重构数据与原输入数据的比较计算得到.

2.2　RBM 神经网络

2.2.1　模型结构

RBM 是自编码网络的核心组件之一,是一个两层

的神经网络,其两层节点分别是可视节点和隐藏节点. RBM 的理论模型对应于一个二分图,每一层的节点之间没有连接,一层是可视层,即输入数据层($\boldsymbol{v}$),另一层是隐藏层($\boldsymbol{h}$). 可视节点用来描述输入数据特征,隐藏节点通常是机器学习自动生成. 如果假设所有的节点都是随机二值变量节点(只能取 0 或者 1 值),假设全概率分布 $p(\boldsymbol{v}|\boldsymbol{h})$ 满足 Boltzmann 分布,则称该模型为受限玻尔兹曼机.

RBM 可以看作一种无向图模型,如图 2 所示. 由于所有的 $\boldsymbol{v}$ 和 $\boldsymbol{h}$ 满足 Boltzmann 分布,因此,当输入 $\boldsymbol{v}$ 的时候,通过 $p(\boldsymbol{h}|\boldsymbol{v})$ 可以得到隐藏层 $\boldsymbol{h}$,而得到隐藏层 $\boldsymbol{h}$ 之后,通过 $p(\boldsymbol{v}|\boldsymbol{h})$ 又能得到可视层. 通过调整参数,从隐藏层获得的可视层 $\boldsymbol{v}1$ 与原来的可视层 $\boldsymbol{v}$ 如果一样,那么所获得的隐藏层就是可视层另一种表达.

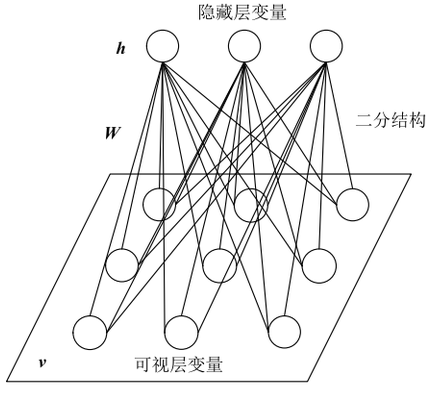

图2 RBM结构

2.2.2 模型参数分析

假设 $\boldsymbol{v}$ 是可视层单元,表示输入数据,$\boldsymbol{h}$ 是隐藏层单元,$\boldsymbol{W}$ 表示可视单元和隐藏单元之间的连接权重. 令所有单元均为二值变量,即任意 $\forall i,j,v_i \in \{0,1\},h_j \in \{0,1\}$.

由于 RBM 的对称结构,RBM 层间有连接,层内无连接结构,所以在已知 $\boldsymbol{v}$ 的情况下,所有的隐藏节点之间是条件独立的,故隐藏层中第 j 个单元 h_j 的后验概率表示为:

$$p(h_j = 1 \mid \boldsymbol{v}) = \sigma\left(\sum_i w_{ij}v_i + a_i\right) \tag{1}$$

其中,运算符 $\sum_i$ 表示对 i 个可视单元 v_i 进行求和. 同理,在已知隐藏层 $\boldsymbol{h}$ 的情况下,所有的可视节点都是条件独立的,故可视层中第 i 个单元 v_i 的后验概率表示为:

$$p(v_i = 1 \mid \boldsymbol{h}) = \sigma\left(\sum_j w_{ij}h_j + b_j\right) \tag{2}$$

其中,σ 是激活函数(sigmoid):

$$\sigma(y) = \frac{1}{1+\mathrm{e}^{-y}} \tag{3}$$

在 RBM 中能量函数扮演着重要的作用,能量函数最小时的解为联合概率分布的最优解. 那么,RBM 关于某一状态 $\{\boldsymbol{v},\boldsymbol{h}\}$ 的能量函数定义为:

$$\mathrm{E}(\boldsymbol{v},\boldsymbol{h};\boldsymbol{\theta}) = -\boldsymbol{v}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{h} - \boldsymbol{b}^{\mathrm{T}}\boldsymbol{v} - \boldsymbol{a}^{\mathrm{T}}\boldsymbol{h} \tag{4}$$

其中,$\boldsymbol{\theta} = \{\boldsymbol{W},\boldsymbol{a},\boldsymbol{b}\}$ 是模型参数,$\boldsymbol{a}$ 表示隐藏单元的偏置,$\boldsymbol{b}$ 表示可视单元的偏置. 状态 $\{\boldsymbol{v},\boldsymbol{h}\}$ 的联合概率分布满足:

$$p(\boldsymbol{v},\boldsymbol{h};\boldsymbol{\theta}) = \mathrm{e}^{-E(\boldsymbol{v},\boldsymbol{h};\boldsymbol{\theta})} \tag{5}$$

2.3 对比分歧算法

为了使联合概率分布 $p(\boldsymbol{v},\boldsymbol{h};\boldsymbol{\theta})$ 最大化,对联合概率分布求解极大似然假设的梯度,表示为:

$$\frac{\partial \log p(\boldsymbol{v},\boldsymbol{h};\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \langle v_i^0 h_j^0 \rangle - \langle v_i^\infty h_j^\infty \rangle \tag{6}$$

其中,$\langle v_i^0 h_j^0 \rangle$ 为输入特征向量与其对应的隐藏层特征向量的点乘的平均值,$\langle v_i^\infty h_j^\infty \rangle$ 为马尔可夫链末端可视层特征向量与其对应的隐藏层特征向量的点乘的平均值. 由公式(6)可知联合概率分布的极大似然假设的梯度与中间状态无关,只和初始状态和最终状态有关.

利用马尔可夫链的方法求最佳联合概率 $P(v_i^\infty, h_j^\infty;\theta)$,收敛速度很难保证,难以确定步长. 对比分歧算法(contrastive divergence,CD)是由 Hinton 提出来的快速地训练限制玻尔兹曼机的方法,并在实践中取得了非常好的应用效果[14]. 两个概率分布的差异性表示为 $KL(P^0 \parallel h_\theta^\infty)$,即:

$$CD = KL(P^0 \parallel h_\theta^\infty) - KL(P_\theta^n \parallel h_\theta^\infty) \tag{7}$$

其中,P^0 为 RBM 网络初始状态的联合概率分布,P_θ^n 为经过 t 步迭代后的联合概率分布,h_θ^∞ 为马尔可夫链末端的联合概率分布.

为了达到训练目的,模型参数 $\boldsymbol{\theta} = \{\boldsymbol{W},\boldsymbol{a},\boldsymbol{b}\}$ 的更新规则应该满足公式(8):

$$\boldsymbol{\theta}^{n+1} = \boldsymbol{\theta}^n + \varepsilon(\langle v_i^0 h_j^0 \rangle - \langle v_i^n h_j^n \rangle) \tag{8}$$

其中,n 为迭代次数,ε 为学习速度. 实验证明,在 n 次迭代后,通过梯度修正参数 $\boldsymbol{\theta}$,CD 值必将趋近于 0.

3 自编码网络的入侵检测模型构建

3.1 模型设计

基于自编码网络的支持向量机(Support Vector Machine-based Autoencoder Network,AN-SVM)入侵检测模型的总体框架如图 3 所示,包含三个步骤:

步骤 1:数据预处理. 将 KDD'99 数据集[16]中符号型属性特征进行数值化,再对数据进行最小-最大规范化处理. 详见 4.1.2 节内容.

步骤 2:为预处理后的标准化数据建立高维空间和低维空间双向映射的自编码网络结构. 其中,自编码网络结构的设计详见 3.2 节内容,预训练过程详见 3.3 节内容,微调权值过程详见 3.4 节内容.

步骤 3:将步骤 2 中学习获得的低维特征数据的作为 SVM 分类器的输入,识别出来攻击类别. 将分类结果与规则库进行判断比较,对检测到的入侵做出响应处

理,如向用户发出警报、断开连接等.

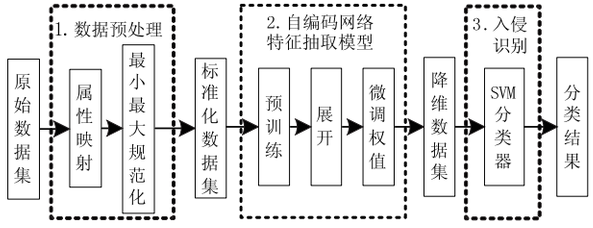

图3　基于AN-SVM模型的入侵检测系统架构

3.2　自编码网络结构设计

本文设计的自编码网络结构如图 4 所示,由数据输入层、中间 4 层 RBM 隐藏层和输出层组成,其中 RBM 隐藏层的层数决定了网络结构的非线性复杂程度. 首先将 KDD'99 数据集进行预处理后得到 122 个特征作为自编码网络的输入,采用 4 层隐藏层网络结构来表示数据之间的非线性关系,选取降维后的 5 维特征数据作为输出. 关于最优的自编码网络结构的选取将在 4.3.1 节实验部分给出.

自编码网络将具有高维和非线性的网络连接数据映射到低维空间上的实现过程包括三个步骤:

①预训练过程:采用非监督贪婪的方法对每一层 RBM 网络进行预训练,获得生成模型的权值.

②展开过程:在多次 RBM 网络预训练后,编码器和解码器使用经 RBM 训练得到的权值作为自编码网络的初始权值. 将预训练过程中得到的 RBM 网络连接起来并按照自编码网络结构将其展开.

③微调权值过程:最后按照原始训练数据与重构数据之间的误差最小化原则对自编码网络权值进行调整. 依次经过解码器和编码器利用反向传播算法(Back Propagation,BP)[15]对整个自编码网络的权值进行微调.

3.3　预训练算法

算法 1:Pre-TrainRBM ($\boldsymbol{V},\boldsymbol{H}$)

输入:可视层变量 $\boldsymbol{V}=v_1,v_2,\cdots,v_m$ 和隐藏层变量 $\boldsymbol{H}=h_1,h_2,\cdots,h_n$;

输出:模型参数 $\boldsymbol{\theta}=\{\boldsymbol{W},\boldsymbol{a},\boldsymbol{b}\}$;

(1)初始化 $W_{ij}=a_i=b_j=0(i=1,\cdots,m,j=1,\cdots,n)$,并初始化迭代次数 k;

(2)对于每一个输入特征变量$(v_1,v_2,\cdots,v_m)$:将 v_i 赋值给 $\boldsymbol{v}^0$;

(3)对网络的每个隐藏单元 j,根据公式(1)计算每个隐藏层特征向量,$h_j^{(t)}=p(h_j|\boldsymbol{v}^{(t)})$;

(4)对网络的每个可视单元 i,根据公式(2)计算每个可视层特征向量,$v_i^{(t+1)}=p(v_i|\boldsymbol{h}^{(t)})$;

(5)对于所有可视单元和隐藏单元,根据公式(6)获得初始状态和更新状态下的联合概率分布的梯度,并代入公式(8)来更新参数 $\boldsymbol{\theta}$,即:

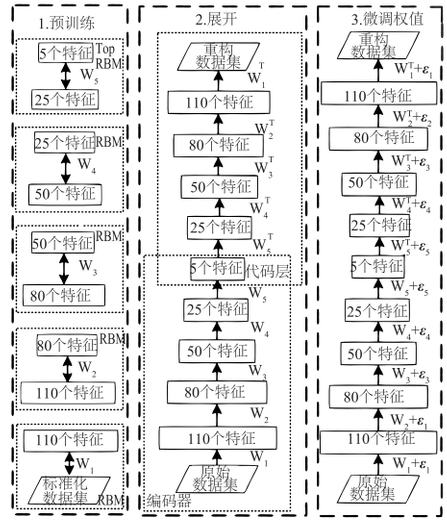

图4　自编码网络结构

$$
\begin{aligned}
W_{ij}^{(t+1)} &= W_{ij}^{(t)} + p(h_j|\boldsymbol{v}^{(0)})\cdot v_i^{(0)} - p(h_j|\boldsymbol{v}^{(k)})\cdot v_i^{(k)}\\
a_i^{(t+1)} &= a_i^{(t)} + v_i^{(0)} - v_i^{(k)}\\
b_j^{(t+1)} &= b_j^{(t)} + p(h_j|\boldsymbol{v}^{(0)}) - p(h_j|\boldsymbol{v}^{(k)})
\end{aligned}
\tag{9}
$$

(6)如果 $t=k$,保存模型参数,算法结束;

(7)如果 $t<k$,则 $t=t+1$,转向步骤(2).

在预训练过程中,首先根据公式(1)将可视向量值映射到隐藏单元;然后根据公式(2)可视单元由隐藏层单元重建;这些新可视单元再次映射到隐藏单元,这样就可以获取新的隐藏单元. 执行这种后退和前进的步骤就是我们熟悉的 Gibbs 采样,而隐藏层激活单元和可视层输入之间的相关性差别就作为权值更新的主要依据.

3.4　微调权值算法

BP 神经网络算法[15]是一种有监督的分类器,在自编码网络模型中,BP 算法起到微调整个网络中参数的作用. 基于 BP 算法的自编码网络微调权值训练过程分为两步:(1)是建立自编码网络的前向传播,获得模型参数;(2)基于给定的自编码网络结构,利用反向传播算法按照原始训练数据与重构数据计算得到的误差信息自顶向下传播至每一层 RBM,微调整个自编码网络模型的参数.

算法 2:FineTuneRBM($\boldsymbol{\theta}$,example)

输入:通过算法 1 获得模型参数 $\boldsymbol{\theta}$,训练样本为 example $=\langle v_i,t_i\rangle(i=1,2,\cdots,m)$;

输出:微调整后的模型参数 $\boldsymbol{\theta}=\{\boldsymbol{W},\boldsymbol{a},\boldsymbol{b}\}$;

(1)初始化迭代次数 k;

(2)对于每一个实例 v_i 输入自编码网络,计算 v_i 的重构表示 v'_i,使误差沿自编码网络反向传播;

(3)对每个输出单元 k,计算它的误差项 δ_k:

$$\delta_k = v'_k(1-v'_k)(v_k - v'_k) \tag{10}$$

(4)对于每个隐藏单元 $\boldsymbol{h}$,计算它的误差项 δ_h

$$\delta_h = v'_h(1-v'_h)\sum_{k\in \text{outputs}} \theta_{kh}\delta_k \tag{11}$$

(5)更新每个网络模型参数 θ_{ji}

$$\theta_{ji} = \theta_{ji} + \Delta\theta_{ji} \tag{12}$$

其中,$\theta_{ji} = \eta\delta_j x_i$,$\eta$ 为学习速率.

(6)如果 $t = k$,保存微调整后的模型参数,算法结束;

(7)如果 $t < k$,则 $t = t + 1$,转向步骤(2).

4 实验与分析

4.1 实验设置

4.1.1 实验数据

本文采用 KDD'99 数据集[16]作为 IDS 系统测试. KDD'99 数据集是由麻省理工学院林肯(Lincon)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,是目前比较权威的测试数据集.

10% 的 KDD'99 数据集包含 494021 个实例的训练数据和 311029 个测试数据,主要包括四种类型的攻击行为:DoS (Denial of service)拒绝服务攻击、R2L (Remote-to-local)远程到本地攻击、U2R (User-to-root)未经授权且试图获取超级用户和 root 权限访问以及 Probing 端口监视或扫描. KDD'99 数据集中各种类型数据的分布情况如图 5 所示.

为了验证 AN-SVM 模型的有效性,本文随机抽样形成 5 个数据集如表 1 所示.

表 1 5 个数据样本集

序号	训练集			测试集		
	正常	异常	总数	正常	异常	总数
DS1	9442	1408	10850	9530	1606	11136
DS2	6943	4107	11050	6984	4166	11150
DS3	5980	1520	7500	8022	1228	9250
DS4	6924	1126	8050	8191	1189	9380
DS5	7705	2345	10050	9570	1980	11550

4.1.2 数据预处理

KDD'99 数据集的每个连接记录由 41 个属性特征组成,其中包含 38 个数字型属性特征和 3 个符号型属性特征. 数据预处理主要包括两个步骤:①符号型属性特征的数值化;②最小-最大规范化.

① 符号型属性特征的数值化

本文采用属性映射方法,将符号型特征转变成二进制数值特征. 其中,属性特征 protocol_type 有 3 种不同的取值:'tcp'、'udp'和'icmp',将其扩展到 3 维特征向量. 'tcp'表示为[1,0,0],'udp'表示为[0,1,0],'icmp'表示为[0,0,1]. 同理,属性特征'service'的 70 种符号取值和'flag'的 11 种符号取值都可以建立符号值与相应数值的映射关系. 以此类推,41 维特征经过变换之后有 122 维特征.

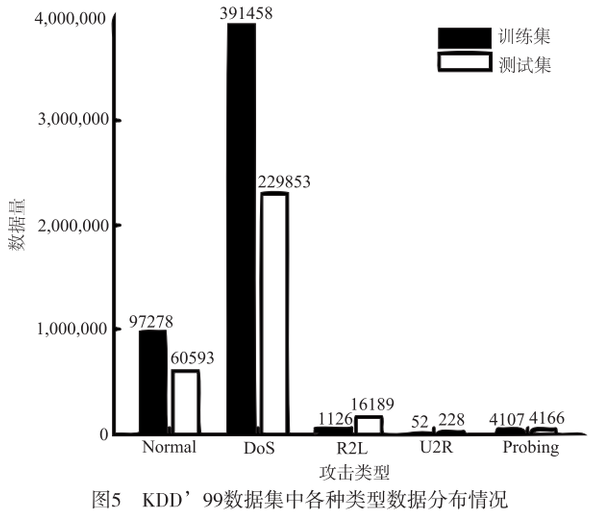

图5 KDD' 99数据集中各种类型数据分布情况

②最小-最大规范化

对原始属性值进行最小-最大规范化处理后,各属性处于同一数量级,适合进行综合对比评价. 根据公式(13)将数字型数据线性映射到[0,1] 范围.

$$x' = \frac{x - \text{MIN}}{\text{MAX-MIN}} \tag{13}$$

其中,x 是属性值,MIN 是该属性的最小值,MAX 是该属性的最大值.

4.1.3 模型参数设置

本实验在 Intel CPU 3.00 GHz、2GB 内存、64b 硬件环境和 Windows 7 操作系统下,使用 MATLAB 2010a 进行编码实现. SVM 分类器实现采用了开源工具 LIBSVM[17]. 其中,AN-SVM 模型表示为:AN^i-SVM(i 表示 AN-SVM 模型包含的 RBM 的层数). AN-SVM 模型参数如表 2 所示:

本文利用自编码网络对特征向量进行降维处理,将获得的低维特征数据的作为 SVM 分类器的输入. KDD'99 数据集经数据预处理后为 122 维特征,故自编码网络的输入层节点数为 122. 第 1 隐藏层节点数选取接近输入层节点数的维数 110,之后的高层节点数依次设置为 80,50,25 和 5,进而根据 3.1 节所述自编码网络实现过程建立高维空间和低维空间双向映射的自编码网络结构. 在建立好的网络结构上对每一层 RBM 网络进行预训练,采用 1 次吉布斯采样对模型参数进行更新,文献[18]论证了 $k = 1$ 时,RBM 模型趋于平稳分布,能得到了最好的实验结果. 在每一层 RBM 训练时,迭代次数为 50 次,基于 BP 算法的自编码网络微调权值算法的迭代次数为 200 次.

表 2　AN^5-SVM 模型参数列表

AN^5-SVM parameter	Value
RBM pre-training learning rate	0.01
BP fine-tuning learning rate	0.1
k step for CD－k algorithm	1
Number of nodes in input layer	122
Number of nodes in 1st hidden layer	110
Number of nodes in 2nd hidden layer	80
Number of nodes in 3rd hidden layer	50
Number of nodes in 4th hidden layer	25
Number of nodes in output layer	5
RBM pre-training max epoch	50
BP fine-tuning max epoch	200
SVM kernel type	N-RBF[8]
SVM gamma value in kernel function	0.00001
C value	1000
n for n-fold cross validation mode	10

为了进一步缩短入侵分类的训练和测试时间,本文采用文献[8]提出的 N-RBF 核函数作为 SVM 的核函数,该核函数通过嵌入特征属性的均值和均方差对属性进行归一化,以避免基于 RBF 的分类训练中因属性取值范围差异过大而导致产生过多支持向量的问题. 设置惩罚因子 $C=1000$,核函数参数 gamma = 0.00001 时,通过 10 折交叉验证获得最高的交叉验证准确率.

4.2　评估标准

在入侵检测性能评估对比实验中,采用训练时间、测试时间、准确率、检测率等作为评价指标来衡量 AN-SVM 模型的性能. 其中准确率(Accuracy)、检测率(Detection Rate)和误报率(False Alarm Rate)定义如下:

$$AC=(TP+TN)/(TP+TN+FP+FN) \quad (14)$$

$$DR=TP/(TP+FP) \quad (15)$$

$$FAR=FP/(TN+FP) \quad (16)$$

其中,TP (True Positive)是正确识别的正常记录数,TN (True Negative)是正确识别的攻击记录数,FP (False Positive)是错误识别的正常记录数,FN (False Negative)是错误识别的攻击记录数. 训练时间用 Tr 表示,测试时间用 Te 表示.

4.3　实验分析

为了验证基于自编码网络入侵检测模型的有效性,本文设计了 2 组实验:

实验 1:分析 AN-SVM 模型中各参数对入侵检测效果的影响.

实验 2:对 AN-SVM 模型与其他方法的分类精度和检测速度进行对比分析.

在试验中,选用数据集 DS1 对 AN-SVM 模型参数选择的效率进行评估. 通过实验发现,当 AN-SVM 模型采用 5 层 122-110-80-50-25-5 的自编码网络结构,总体分类性能最高. 下面讨论不同网络深度、隐藏层节点数和输出层节点数等关键因子对入侵识别性能的影响.

4.3.1　模型参数选择实验分析

(1)网络深度的影响

自编码网络的深度对入侵识别效果起到非常重要的作用. Hinton 的相关研究也指出 3 层 RBM 网络已能提取有效的特征用于分类任务[18]. 随着 RBM 层数的增加,深度学习模型的建模能力也随之增强,高层特征的表示能力更抽象,更能提高分类性能. 然而,文献[18]说明了随着 RBM 层数的增加,神经网络节点增加,训练时间也大幅增加,过多的层数容易导致过拟合现象.

本实验设置了 6 种不同的 AN-SVM 网络结构,性能对比分析结果如表 3 所示. 设置 AN^1-SVM 为浅层 122-5 的 RBM 网络结构,同理,AN^2-SVM、AN^3-SVM、AN^4-SVM、AN^5-SVM 和 AN^6-SVM 分别表示为 122-60-5、122-80-40-5、122-100-70-40-5、122-110-80-50-25-5 和 122-110-80-60-40-20-5 的 RBM 网络结构. 由表 3 可知,实验结果表明,AN-SVM 采用 5 层的 RBM 网络结构的准确率、检测率最大. AN-SVM 通过多层映射单元提取出主要的结构信息,其性能优于浅层结构.

表 3　不同的 AN-SVM 结构的性能比较

模型结构	AC(%)	DR(%)
AN^1-SVM	71.18	70.35
AN^2-SVM	74.22	75.60
AN^3-SVM	82.65	81.30
AN^4-SVM	95.07	94.69
AN^5-SVM	98.89	97.33
AN^6-SVM	96.49	96.33

(2)第 1 隐藏层节点数的影响

第 1 隐藏层表达了原始数据的首次约简特征,该层的表达能力决定了上层的特征抽取能力,故第 1 隐藏层节点数的选取非常关键. 文献[18]指出,神经元节点数的增多能提高网络的逼近能力,但也面临过学习问题,会降低网络的泛化能力. 因此自编码网络结构逐层降维幅度不宜过大,以最小信息损失获得原始数据的低维表示,防止结果剧烈震荡.

本实验中,设置深层 122-110-80-50-25-5 的 AN^5-SVM 模型,其他参数不变,改变第 1 隐藏层节点数,获得如图 6 所示的实验结果,可知第 1 隐藏层节点数为 110 的分类效果最好,因为 110 接近输入特征向量的维数.

(3)输出层节点数的影响

本文另外一个重要的研究是选择最小的特征向量维数来提高入侵识别准确率. 为了测试输出层节点数是如何影响入侵检测性能,本文采用深层 122-110-80-50-25-5 的 AN^5-SVM 模型,其他参数不变,输出层节点数从 1 变

化至10,结果如图7所示.由图7可知,AN^5-SVM模型设置5个输出层节点的准确率和检测率最大.

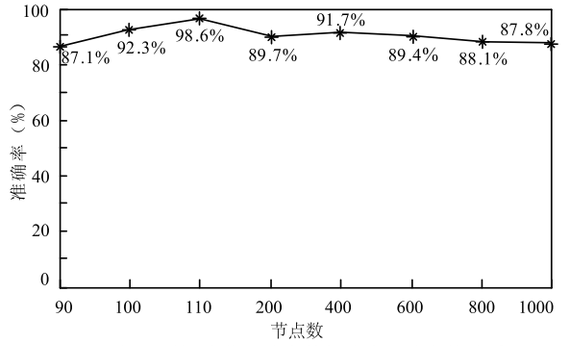

图6 第1隐藏层节点数与分类正确率的关系

4.3.2 与其他方法的对比分析

AN^5-SVM模型与入侵检测相关工作做了实验对比.表4给出了AN^5-SVM模型与其他方法的检测率(DR)、误报率(FAR)、训练时间(Tr)和测试时间(Te)等方面的性能对比.

(1)从图8和图9可以看出,在分类检测率和误报率方面,设置5个输出层节点的AN^5-SVM模型提高了SVM在网络入侵识别的性能,在一定程度上优于已有的一些方法.因此,该模型在保证最大化检测率的同时,也具有将特征维数最小化到5的能力,这是本文提出特征降维算法的目的.

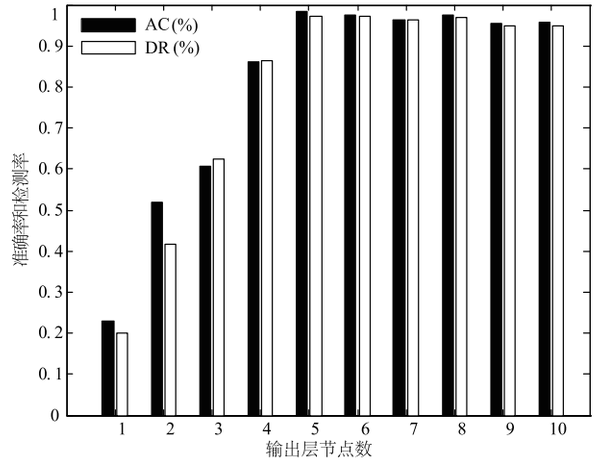

图7 不同输出层节点数的SVM性能比较

表4 不同分类器的检测性能对比

分类器 \ 数据集		DS1	DS2	DS3	DS4	DS5
AN^5-SVM-5	DR(%)	98.23	94.73	97.48	98.31	96.82
	FAR(%)	1.225	4.48	3.67	1.71	3.11
	Tr(s)	1.57	2.05	1.26	1.35	1.833
	Te(s)	1.19	1.21	0.921	0.894	1.37
AN^5-SVM-10	DR(%)	96.789	94.57	95.5	96.61	95.76
	FAR(%)	4.61	5.68	4.68	3.97	4.78
	Tr(s)	3.76	3.36	2.977	2.435	3.06
	Te(s)	2.3	1.97	1.632	1.47	2.39
KPCA-GA-SVM[8]	DR(%)	98.2	94.753	97.47	98.07	96.737
	FAR(%)	2.26	4.69	3.2	1.95	3.32
	Tr(s)	2.81	2.21	1.91	1.333	2.661
	Te(s)	1.821	1.925	1.5	1.64	2.03
PCA-GA-SVM[9]	DR(%)	97.011	95.54	95.16	97.51	96.16
	FAR(%)	3.8	5.6	4.92	3.13	4.35
	Tr(s)	9.724	12.552	8.854	7.356	13.112
	Te(s)	7.67	11.91	7.18	5.577	9.615
KPCA-SVM[7]	DR(%)	95.16	93.11	94.32	95.48	94.27
	FAR(%)	8.6	7.34	7.04	4.45	6.14
	Tr(s)	16.28	18.02	12.42	13.34	15.992
	Te(s)	14.14	15.95	11.553	12.511	13.085
SVM-122	DR(%)	95.2	92.77	94.49	95.2	93.57
	FAR(%)	8.31	7.67	6.78	4.89	6.77
	Tr(s)	37.639	38.7	24.66	25.17	35.135
	Te(s)	25.24	26.3	15.94	18.837	18.85
PCA-ANN[10]	DR(%)	91.063	89.65	88.94	90.654	87.236
	FAR(%)	9.08	9.74	9.57	8.53	8.94
	Tr(s)	25.663	24.23	19.547	21.678	27.42
	Te(s)	21.254	19.65	13.47	19.524	17.85

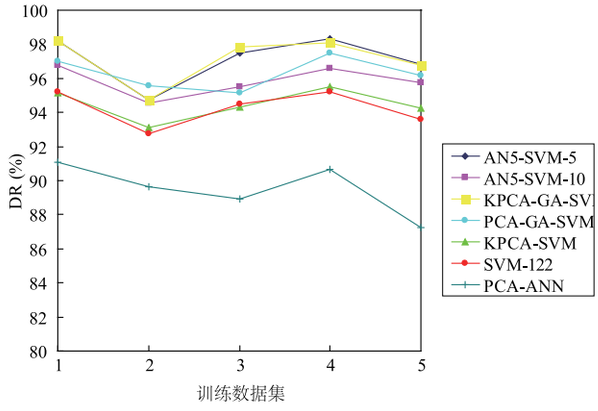

图8　不同分类器的检测率对比

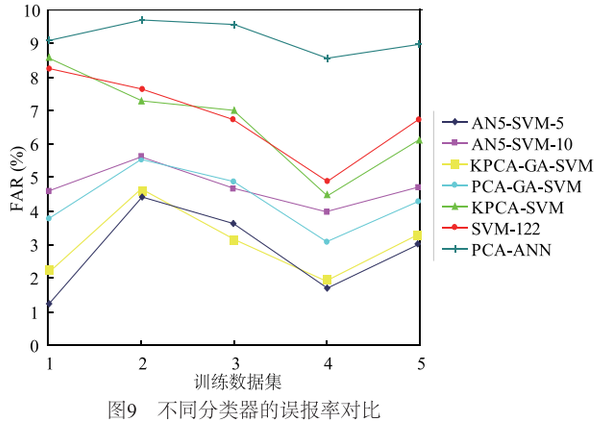

图9　不同分类器的误报率对比

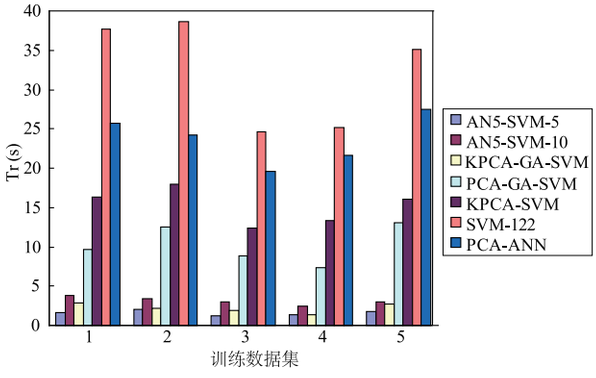

图10　不同分类器的训练时间对比

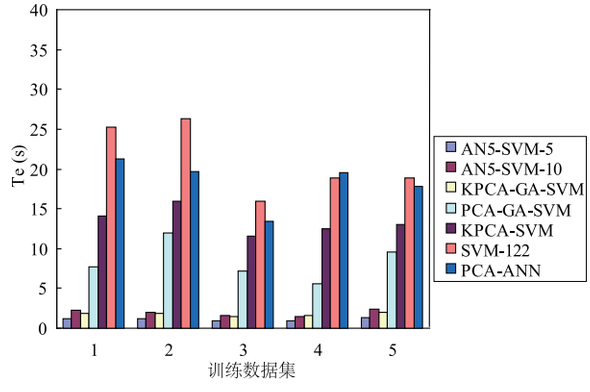

图11　不同分类器的测试时间对比

表 5　不同攻击类别的检测性能对比

分类器 \ 攻击类别		Normal	DoS	R2L	U2R	Probing
AN^5-SVM-5	DR (%)	99.68	99.89	87.58	18.31	98.83
	FAR (%)	0.022	0.027	15.8	79.64	3.08
AN^5-SVM-10	DR (%)	98.88	98.5	85.9	16.71	95.96
	FAR (%)	1.22	1.18	18.4	80.63	4.77
KPCA-GA-SVM[8]	DR (%)	99.56	98.85	87.66	18.07	96.737
	FAR (%)	0.126	1.06	15.8	85.6	3.32
PCA-GA-SVM[9]	DR (%)	99.5	99.2	75.16	17.51	96.16
	FAR (%)	0.38	0.65	24.85	83.6	4.35
KPCA-SVM[7]	DR (%)	99.42	98.99	74.32	15.48	94.27
	FAR (%)	0.61	0.84	28.43	87.22	6.14
SVM-122	DR (%)	98.2	97.74	64.49	13.2	93.57
	FAR (%)	0.31	1.62	33.78	88.5	6.77
PCA-ANN[10]	DR (%)	96.38	90.49	62.687	18.513	92.962
	FAR (%)	2.986	9.32	34.46	78.683	6.94

(2)本文对基于不同特征降维算法的 SVM 入侵检测模型进行了两个计算性能指标——训练时间和测试时间的对比实验. 文献[19]强调基于特征降维算法的入侵检测系统更多关注测试时间,而不是检测率,虽然

不同的特征降维算法对入侵检测系统检测率的影响不大,但是训练时间与测试时间有显著的差异. 文献[8]指出使用约简后的数据集对分类器有明显的速度提高.

本文设置了 6 种不同的分类器:AN-SVM、KPCA-GA-SVM、PCA-GA-SVM、KPCA-SVM、PCA-ANN 和单独的 SVM. 通过每一种特征降维算法获得约简的特征子集来构建入侵检测模型,并记录相应分类器的训练时间与测试时间,实验结果如图 10 和图 11 所示. 实验结果表明 AN^5-SVM-5 具有最短的训练时间与测试时间. 可以说,AN^5-SVM 模型是从降低 SVM 分类训练时间与测试时间这一角度出发,利用自编码网络高效降维能力构建轻量级入侵检测模型.

(3)本文进一步在数据集 DS1 上对不同攻击类别的检测率(DR)和误报率(FAR)进行对比分析,实验结果如表 5 所示. AN^5-SVM 模型在 DoS 和 Probe 攻击行为预测的检测率是最好的. 6 种分类器在 R2L 和 U2R 攻击行为预测的检测性能都不太理想. 但使用特征降维算法约简的特征子集比基于所有特征建立的 SVM 分类器的整体检测性能要好.

5 结论

自编码网络的深度学习方法对于入侵检测是一种全新的机器学习方法,该方法不仅能提供从高维空间到低维空间的双向映射关系,而且可以给出相反的逆映射. 本文提出了一种基于高效的自编码网络的支持向量机入侵检测模型,首先该模型利用自编码网络方法对高维、非线性的原始数据进行特征降维,然后对学习过程中获得最优的低维特征向量采用 SVM 进行攻击识别. 本文在 KDD'99 数据集上验证 AN-SVM 模型的有效性,实验结果表明本文提出的 AN-SVM 模型非常适用于高维空间的信息抽取任务,降低了入侵检测模型中分类的训练时间和测试时间,能很好的满足入侵检测的实时性要求,而且检测性能优于传统算法. AN-SVM 模型提高了入侵检测率且加快了入侵检测速度,是一种可行的、高效的轻量级入侵检测模型,为入侵检测提供了一种新的研究思路.

参考文献

[1] Wang G, Hao J, Ma J, et al. A new approach to intrusion detection using artificial neural networks and fuzzy clustering[J]. Expert Systems with Applications, 2010, 37(9): 6225 – 6232.

[2] Lin W C, Ke S W, Tsai C F. CANN: An intrusion detection system based on combining cluster centers and nearest neighbors [J]. Knowledge-Based Systems, 2015, 78: 13 – 21.

[3] Hoz E D L, Hoz E D L, Ortiz A, et al. Feature selection by multi-objective optimisation: Application to network anomaly detection by hierarchical self-organising maps [J]. Knowledge-Based Systems, 2014, 71: 322 – 338.

[4] Vapnik V N. The nature of statistical learning theory[J]. Neural Networks IEEE Transactions on, 1995, 10(5): 988 – 999.

[5] 尚文利,张盛山,万明,等. 基于 PSO-SVM 的 Modbus TCP 通讯的异常检测方法[J]. 电子学报,2014,11(42): 2314 – 2320.
Shang Wenli, Zhang Shengshan, Wang Ming, et al. Modbus/TCP communication anomaly detection algorithm based on PSO-SVM[J]. Acta Electronica Sinica, 2014, 11(42): 2314 – 2320. (in Chinese)

[6] Chitrakar R, Huang C. Selection of candidate support vectors in incremental SVM for network intrusion detection [J]. Computers & Security, 2014, 45(3): 231 – 241.

[7] Kuang F, Xu W, Zhang S, Wang Y. A novel approach of KPCA and SVM for intrusion detection [J]. Journal of Computational Information Systems, 2012, 8(8): 3237 – 3244.

[8] Kuang F, Xu W, Zhang S. A novel hybrid KPCA and SVM with GA model for intrusion detection [J]. Applied Soft Computing, 2014, 18(4): 178 – 184.

[9] Ahmad I, Abdullah A, Alghamdi A, et al. Optimized intrusion detection mechanism using soft computing techniques [J]. Telecommunication Systems, 2013, 52(4): 2187 – 2195.

[10] 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): 1790 – 1799.

[11] Hinton G E, Salakhutdinov R R. Reducing the dimensionality of data with neural networks[J]. Science, 2006, 313(28): 504 – 507.

[12] Hinton G E, Osindero S. A fast learning algorithm for deep belief nets[J]. Neural Computation, 2006, 18(7): 1527 – 1554.

[13] 余凯,贾磊,陈雨强,等. 深度学习的昨天、今天和明天[J]. 计算机研究与发展,2013,50(9):1799 – 1804.
Yu Kai, Jia Lei, Chen Yuqiang, et al. Deep learning: yesterday, today, and tomorrow[J]. Journal of Computer Research and Development, 2013, 50(9): 1799 – 1804. (in Chinese)

[14] Hinton G E. Training products of experts by minimizing contrastive divergence[J]. Neural computation, 2002, 14(8): 1771 – 1800.

[15] Rumelhart D E, Hinton G E, Williams R J. Learning representations by back-propagating errors[J]. Nature, 1986, 323(6088):533 – 536.

[16] Stolfo S J, Fan W, Lee W K, et al. Cost-Based Modeling for Fraud and Intrusion Detection: Results from the JAM Project [EB/OL]. http://kdd. ics. uci. edu/databases/kddcup99/kddcup99. html, 2011 – 06 – 27.

[17] Chang C C. LIBSVM: a library for support vector machines[J]. ACM Transactions on Intelligent Systems & Technology, 2011, 2(3):389 – 396.

[18] Larochelle H, Bengio Y, Louradour J, et al. Exploring strategies for training deep neural networks[J]. Journal of Machine Learning Research, 2009, 10(6):1 – 40.

[19] Williams N, Sebastian Z, Armitage G. A preliminary performance comparison of five machine learning algorithms for practical IP traffic flow classification[J]. ACM SIGCOMM Computer Communication Review, 2006, 36(5): 5 – 16.

作者简介

高　妮　女, 1982 年生于陕西省咸阳市. 博士, 讲师. 主要研究方向为网络安全、机器学习等.

E-mail: gaoni@ nwu. edu. cn