计算机研究与发展
Journal of Computer Research and Development
DOI:10.7544/issn1000-1239.2020.20190880
57(4):791-802, 2020

基于半监督学习的无线网络攻击行为检测优化方法

王 婷[1,2] 王 娜[3] 崔运鹏[1,2] 李 欢[1,2]
[1](中国农业科学院农业信息研究所 北京 100081)
[2](农业农村部农业大数据重点实验室(中国农业科学院农业信息研究所) 北京 100081)
[3](96962部队 北京 102206)
(wangting01@caas.cn)

The Optimization Method of Wireless Network Attacks Detection Based on Semi-Supervised Learning

Wang Ting[1,2], Wang Na[3], Cui Yunpeng[1,2], and Li Huan[1,2]
[1](*Agricultural Information Institute, Chinese Academy of Agricultural Sciences, Beijing* 100081)
[2](*Key Laboratory of Big Agri-Data (Agricultural Information Institute, Chinese Academy of Agricultural Sciences), Ministry of Agriculture and Rural Areas, Beijing* 100081)
[3](*Unit* 96962, *Beijing* 102206)

Abstract Aiming to optimize the attacks detection in high-dimensional and complex wireless network traffic data with deep learning technology, this paper proposed a WiFi-ADOM (WiFi network attacks detection optimization method) based on semi-supervised learning. Firstly, based on stacked sparse auto-encoder (SSAE), which is an unsupervised learning model, two types of network traffic feature representation vectors are proposed: new feature value vector and original feature weight value vector. Then, the original feature weight value vector is used to initialize the weight value of the supervised learning model deep neural network to obtain the preliminary result of the attack type, and the unsupervised learning clustering method Bi-kmeans is used to produce the corrective term for unknown attacks discrimination with the new feature value vectors. Finally, the preliminary result of the attack type and the corrective term of the unknown attacks discrimination are combined to obtain the final result of the attack type. Compared with the existing attacks detection methods with the public wireless network traffic data set AWID, the optimal performance of the method of WiFi-ADOM for network attacks detection is verified. At the same time, the importance of features in network attacks detection is explored. The results show that the method of WiFi-ADOM can effectively detect unknown attacks while ensuring detection performance.

Key words network attacks detection; network intrusion detection; semi-supervised learning; deep learning; Bi-kmeans clustering

摘 要 针对如何优化深度学习技术在海量高维复杂的无线网络流量数据中有效发现异常攻击行为的问题,提出一种基于半监督学习的无线网络攻击行为检测优化方法(WiFi network attacks detection

收稿日期:2019-12-18;**修回日期**:2020-02-24
基金项目:国家自然科学基金项目(61672101);中国农业科学院基本科研业务费院级项目(Y2020XC15)
This work was supported by the National Natural Science Foundation of China (61672101) and the Fundamental Research Funds of Chinese Academy of Agricultural Sciences (Y2020XC15).
通信作者:崔运鹏(cuiyunpeng@caas.cn)

optimization method, WiFi-ADOM).首先基于无监督学习模型栈式稀疏自编码器提出2种网络流量特征表示向量:新特征值向量和原始特征权重值向量.然后利用原始特征权重值向量初始化监督学习模型深度神经网络的权重值得到网络攻击类型的预判结果,并通过无监督学习聚类方法Bi-kmeans对网络流量的新特征值向量进行聚类以生成未知攻击类型判别纠正项.最后结合预判结果和未知攻击类型判别纠正项,得到网络攻击类型的最终判定结果.通过和已有研究方法对比,在公开无线网络攻击行为数据集AWID上验证了WiFi-ADOM方法对网络攻击行为检测的优化性能,同时探索了与网络攻击检测相关的重要特征属性的问题.实验结果表明:WiFi-ADOM方法在保证准确率等检测性能的同时能够有效检测未知攻击类型,具备优化网络攻击行为检测的能力.

关键词 网络攻击行为检测;网络入侵检测;半监督学习;深度学习;Bi-kmeans聚类

中图法分类号 TP391

无线局域网技术和移动通信设备的迅猛发展使得WiFi网络环境逐渐普及并融入人们的生活,这同时也使得WiFi成为网络攻击的目标."蹭网"、"无线钓鱼"等无线网络犯罪事件时有发生,引发了个人数据被泄露、篡改等信息安全隐患,甚至导致巨大的经济损失.腾讯手机管家发布的《2018年手机安全报告》中显示,2018年中国公共WiFi的数量近7.37亿,而风险WiFi高达46.08%.网络攻击行为的不断演化和升级使得WiFi环境下的网络安全问题愈发严峻,成为信息安全的全新困局.

网络入侵检测是目前应用最广泛也最有效的以数据驱动的网络安全主动防御方法,基于实时网络流量数据建立相应的攻击评测机制,从而实现对攻击行为的检测和预防.传统的入侵检测方法通常是通过对比已构建的网络行为模式或规则来检测当前网络连接属于正常状态还是攻击风险状态.随着互联网环境的更新换代,网络流量数据呈现出海量、高维、复杂的特点,直接用来进行攻击行为模式发现十分困难.传统的网络入侵检测方法出现检测效率低下、准确率较低、误报率和漏报率较高的问题,已不能满足网络信息安全的需求.

机器学习算法的优化和高性能计算能力的发展给网络攻击行为检测带来了新的思路和方法.深度学习作为近几年发展最快也最热门的机器学习方法,能够有效学习网络流量数据中隐藏的数据特征,从而使得进一步提高网络攻击行为检测性能成为可能[1].一个有效的网络攻击行为检测方法不仅需要能够准确识别已知网络攻击类型,还需要对未知攻击类型或新攻击类型(统称为未知攻击类型)有足够的检测能力.在学术界和工业界,已出现一些基于深度学习的网络攻击检测方法,虽然在提高检测效率和准确率方面取得一定进展,但是由于这些方法的实现都是通过已标注攻击类型的训练数据集训练完成,无法有效检测训练数据集中不包含的攻击类型,从而无法应对当前网络攻击行为千变万化和急剧增长的严峻形式[2].同时,由于网络攻击数据获取的限制,大部分研究工作都是基于已有公开数据集KDD'99或其改进版本NSL_KDD数据集[3]等面向有线通信网络的攻击行为检测进行的,针对无线局域网环境下的攻击检测研究还比较有限.

根据分析和已有问题,本文利用公开无线网络攻击行为数据集AWID,通过深度学习方法中无监督学习模型——栈式稀疏自编码器(stacked sparse auto-encoder, SSAE)构建2种网络流量数据特征表示向量,并在此基础上结合深度学习方法中监督学习模型DNN和非监督学习聚类方法Bi-kmeans,提出一种基于半监督学习的无线网络攻击行为检测优化方法(WiFi network attacks detection optimization method, WiFi-ADOM).主要贡献有2方面:

1) 基于稀疏自编码器模型提出2种网络流量数据特征表示向量:新特征值向量和原始特征权重值向量.新特征值向量作为网络流量数据的新特征表达方式,具有更强的学习能力;原始特征权重值向量表示网络流量数据不同特征属性在表现原数据特征方面的重要性.

2) 在对已有攻击行为具有优异检测性能的DNN模型的基础上,创新性地通过聚类方法Bi-kmeans引入未知攻击类型判别纠正项,解决了DNN模型不能有效检测未知网络攻击类型的问题,优化了WiFi-ADOM方法对攻击类型的最终判定.其中,原始特征权重值向量用来初始化DNN模型,对网络攻击类型预判过程起到优化作用.新特征值向量作为聚类方法Bi-kmeans的输入,解决了Bi-kmeans方法不能很好应对高维复杂数据的问题.

实验结果表明:WiFi-ADOM 方法在保证精确率、误判率等检测性能的同时能够有效检测未知攻击类型,从而实现了无线网络环境下网络攻击行为检测的优化.

1　相关工作

1.1　网络入侵检测

入侵检测的概念是 1980 年 Anderson 首次提出的[4],之后 Heberlein 等人[5]提出基于网络流量数据检测可疑网络行为的网络入侵检测方法.根据挖掘分析的数据源,网络入侵检测可以划分为基于主机、基于网络和混合型的检测方法.基于主机的方法监测客户端级别的攻击行为,基于网络的方法监测整个网络的攻击风险,混合型的则可以同时监控整个网络和特定用户的网络安全情况.根据检测方法,网络入侵检测可以划分为 3 种方法:1)基于规则的检测方法.首先人为构建攻击行为特征规则库,与之匹配便判定为攻击行为[6].2)基于误用的检测方法.挖掘分析不同攻击类型的固有模式,网络行为数据符合其中一种模式时便被判定为攻击行为[7].3)基于异常的检测方法.通过和正常行为对比,偏离太多则判定为攻击行为[8].前 2 种方法通常具有较高的检测准确率,但是却无法有效检测未知攻击类型.而基于异常的方法虽然误判率较高,却可以有效检测未知攻击类型.

1.2　基于机器学习的网络入侵检测

基于机器学习的方法是目前常用的网络入侵检测方法,既可以作为基于误用的检测方法,又可以通过分析网络流量数据中不同特征属性的特点,进一步优化基于规则和异常的检测方法.尤其是作为目前机器学习领域热门的深度学习方法,在人工神经网络的基础上发展起来,通过构建多层次神经网络对输入数据进行深层次表达,具有强大的特征自学习能力,无需领域专家辅助就能够实现特征的有效提取和选择,非常适用于高维网络流量数据的挖掘分析[9].

已有的基于机器学习的网络入侵检测相关研究工作对网络流量数据的处理和分析过程主要可以划分为 2 种类型:

1) 首先基于原始特征数据集重新构建具有更好学习能力的新特征数据集,即特征提取,然后进行网络攻击行为分类.深度学习模型——栈式自编码器(stacked auto-encoder, SAE)是常用的特征提取方法之一.Shone 等人[10]提出的网络入侵检测模型通过构建深度非对称自编码网络提取到比传统对称自编码网络更具有学习能力的网络流量数据特征属性,然后采用支持向量机(support vector machine, SVM)方法对网络攻击行为进行分类.Thing[11]结合 SAE 模型和 softmax 分类器实现对网络攻击行为的分类检测,并通过对比不同激励函数对分类效果的影响得出激活函数 Prelu 效果最优的结论.Aminanto 等人[12]采用 SAE 模型提取网络流量数据原始特征的新特征表达,并通过对所有的原始特征和新特征进行重要级排序筛选出和网络入侵检测最相关的重要特征,然后使用 SVM 方法对网络攻击行为进行识别与分类.文献[13-16]则分别采用深度学习中其他模型,如 RNN 模型、DBN 模型、CNN 模型、神经元映射卷积神经网络模型等进行特征提取,然后在此基础上进一步实现网络攻击行为的分类.

2) 首先从原始特征集中筛选出与网络攻击行为检测相关的重要原始特征,即特征选择,然后进行网络攻击行为分类.Akashdeep 等人[17]同时以信息熵和相关性为度量,选择信息熵大以及相关性小的特征代替网络流量数据的所有特征,减少了冗余特征对网络入侵检测模型性能的影响,然后在新的特征数据集中使用人工神经网络(artificial neural network, ANN)进行攻击行为分类.Wang 等人[18]通过 ANN 模型构建的权重矩阵对初始网络流量特征进行重要性排序,然后以前 K 个原始特征属性值数据作为输入,利用 softmax 实现网络攻击行为的分类.Louvieris 等人[19]在利用 kmeans 聚类算法和朴素贝叶斯算法对网络攻击行为进行分类的基础上通过 Kruskal-Wallis 检验筛选和特定攻击类型相关的重要原始特征,并在筛选后的数据集上采用 C4.5 决策树进行分类.Zhu 等人[20]利用多目标选择算法分别从多个特征筛选度量出发形成特征的重要级排序,然后结合所有排序结果构建一个降维后的原始特征数据集.Usha 等人[21]首先通过粒子群算法优化的传统聚类方法进行特征选择,然后利用 SVM 方法实现网络攻击行为的分类检测.Kolias 等人[22]探究了网络流量数据特征是否完备对网络攻击行为分类检测的影响,以人工选择的 20 个原始特征构建精简数据集,然后同时使用完整和精简 2 种数据集作为实验数据,并分别采用随机森林、朴素贝叶斯、AdaBoost 等分类方法对比分类效果,结果表明原始特征中存在冗余,直接影响分类性能.Zhao 等人[23]通过基于特征的迁移学习方法获取到源场景和目的

场景 2 种不同网络环境下网络流量数据的特征,并在此基础上利用决策树、SVM 方法或最近邻算法对新场景下的网络攻击进行分类.

2 网络流量数据特征表示模型

本文采用深度学习方法中监督学习模型栈式稀疏自编码器(SSAE)进行网络流量数据的特征学习,并基于学习过程提出 2 种不同的网络流量数据特征表示向量:新特征值向量(feature value vector, ***FV***)和原始特征权重值向量(feature weight value vector, ***WV***).SSAE 模型的结构如图 1 所示,由 3 个稀疏自编码器(auto-encoder, AE)组成,通过对每一个稀疏 AE 的单独训练逐层提取数据的高阶特征.

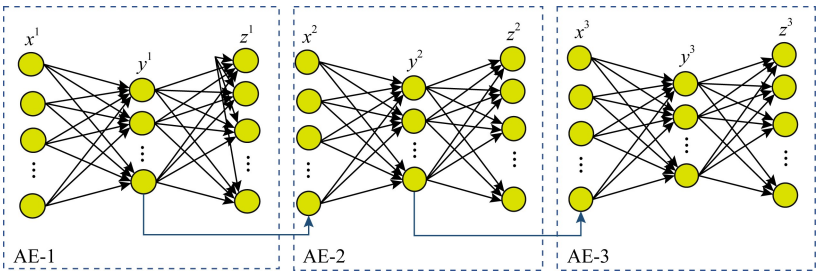

Fig. 1 Framework of network flow feature extraction model

图 1 网络流量特征表示模型结构图

2.1 SSAE 模型

AE 是一种无监督学习模型,包含输入层、隐藏层和输出层,结构分别如 AE-1,AE-2,AE-3 所示.输入层和隐藏层构成编码器,把 x 映射到 y:

$$y=g_{\mathrm{e}}(\boldsymbol{W}_{\mathrm{e}}\cdot x+\boldsymbol{b}_{\mathrm{e}}). \tag{1}$$

隐藏层和输出层构成解码器,与编码器是对称的过程,把 y 映射到 z 以实现对 x 的重建:

$$z=g_{\mathrm{d}}(\boldsymbol{W}_{\mathrm{d}}\cdot y+\boldsymbol{b}_{\mathrm{d}}), \tag{2}$$

$$g_{\mathrm{e}}=g_{\mathrm{d}}=\tanh\left(\frac{\mathrm{e}^{x_t}-\mathrm{e}^{-x_t}}{\mathrm{e}^{x_t}+\mathrm{e}^{-x_t}}\right), \tag{3}$$

其中,$\boldsymbol{W}_{\mathrm{e}}$,$\boldsymbol{W}_{\mathrm{d}}$,$\boldsymbol{b}_{\mathrm{e}}$,$\boldsymbol{b}_{\mathrm{d}}$,$g_{\mathrm{e}}$,$g_{\mathrm{d}}$ 分别表示输入数据的权重矩阵、偏置向量和激活函数;$\boldsymbol{W}_{\mathrm{e}}$,$\boldsymbol{W}_{\mathrm{d}}$,$\boldsymbol{b}_{\mathrm{e}}$,$\boldsymbol{b}_{\mathrm{d}}$ 的初始化采用 Glorot 提出的 X_{avier} 方法[24];x_t 表示 tanh 函数的输入数据.

AE 以最小化代价函数为优化目标,基于反向传播法则不断调整模型参数的取值以获取最优值,训练完成时隐藏层的输出 y 即作为 x 的新特征表达.每一个 AE 隐藏层的输出 y 都作为下一层 AE 的输入,最后一个 AE 隐藏层的输出作为网络流量数据的最终特征表达.由于隐藏层中神经单元的数量小于输入层,所以对高维复杂的网络流量数据起到了降维的作用.SAE[25-27]则是在 AE 训练过程中加入一个稀疏正则化项 J_{sparsity} 来约束隐藏层的神经单元数量,能够减少模型参数以降低训练难度,同时能够有效解决训练结果局部最小化和过拟合的问题.为了更好地防止模型过拟合的问题,本文同时增加了 L_2 正则化项 J_{weights},SSAE 的代价函数为

$$E_{\mathrm{S}}=J_{\mathrm{S}}+\lambda\times J_{\mathrm{weights}}+\beta\times J_{\mathrm{sparsity}}, \tag{4}$$

$$J_{\mathrm{S}}=\frac{1}{N}\sum_{n=1}^{N}\sum_{k=1}^{K}L(x_{nk},z_{nk}), \tag{5}$$

$$J_{\mathrm{weights}}=\frac{1}{2N}\sum_{k=1}^{K}\sum_{m=1}^{M}(w_{km})^2, \tag{6}$$

$$J_{\mathrm{sparsity}}=\sum_{m=1}^{M}KL(\rho\,\|\,\rho'_m)=$$
$$\sum_{m=1}^{M}\left[\rho\log\left(\frac{\rho}{\rho'_m}\right)+(1-\rho)\log\left(\frac{1-\rho}{1-\rho'_m}\right)\right], \tag{7}$$

$$\rho'_m=\frac{1}{N}\sum_{n=1}^{N}g_{\mathrm{e}}(\boldsymbol{W}_m x_n+\boldsymbol{b}_m), \tag{8}$$

其中,$L(x_{nk},z_{nk})$表示 x_{nk} 和 z_{nk} 之间的交叉熵损失函数,x_{nk} 和 z_{nk} 分别表示输入数据中特征属性的期望值和观测值,N,K 和 M 分别表示训练样本数量、特征数量和神经单元数量;x_n 表示第 n 个训练样本,$\boldsymbol{W}_m$ 和 $\boldsymbol{b}_m$ 分别表示第 m 个神经单元对应的权重向量和偏置值,ρ 表示稀疏性常数(本文取值为 0.05),ρ'_m 表示第 m 个神经单元的平均激活度;λ 和 β 则属于需要在训练过程中优化的模型超级参数.

2.2 网络流量数据特征表示向量

本文基于深度学习模型对网络流量数据的 2 种处理过程,提出 2 种网络流量数据特征表示向量:新特征值向量和原始特征权重值向量.其中,前者形成的是新的特征表达方式,属于特征提取;后者表示的是原始特征属性在表现原数据特征方面的重要性,属于特征选择,具体为:

1) 网络流量数据特征表示向量 1——新特征值

向量 $\boldsymbol{FV}$.由 SSAE 模型最后一层隐藏层中所有神经单元构建的新特征值组成,$\boldsymbol{FV}=(f_1,f_2,\cdots,f_n)$,$n$ 表示神经单元的数量.

2) 网络流量数据特征表示向量 2——原始特征权重值向量 $\boldsymbol{WV}$.由 SSAE 模型中所有神经单元表示的流量数据原始特征权重值累计组成,$\boldsymbol{WV}=(w^1,w^2,\cdots,w^i,\cdots,w^m)$,$m$ 表示原始特征值的数量,w^i 表示原始特征对应的最终权重值.所有权重值形成一个权重网络,每个节点所表示的权重值由上层每个节点权重值和边权重值的乘积求和构成.以原始特征 i 为例,构建其权重网络,其结构如图 2 所示:

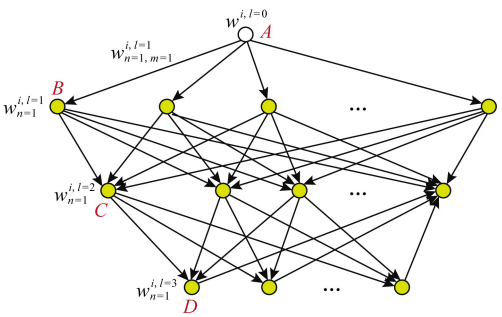

Fig. 2　Framework of original weigh network

图 2　原始特征权重值网络结构图

图 2 中,l 表示权重网络的层数,n 表示当前网络层 l 层节点所对应的顺序标识,m 表示网络层 $l-1$ 层节点所对应的顺序标识;$w^{i,l=0}$ 表示流量数据特征 i 的初始权重值,所有初始权重值都为 1;$w_n^{i,l}$ 表示 l 层第 n 个节点所对应的权重值,比如 1 层第 1 个节点 B 的权重值为 $w_{n=1}^{i,l=1}$;$w_{n,m}^{i,l}$ 表示 l 层第 n 个节点和 $l-1$ 层第 m 个节点之间的边所对应的权重值,比如 1 层第 1 个节点 B 和 0 层第 1 个节点 A 之间边的权重值为 $w_{n=1,m=1}^{i,l=1}$.$\boldsymbol{WV}$ 向量的具体计算过程为:

① 第 1 层节点的权重值是当下节点和初始节点之间的边所对应的权重值,比如节点 B 的权重值为 $w_{n=1}^{i,l=1}=w_{n=1,m=1}^{i,l=1}$;

② 第 2 层节点 C 的权重值是每个上层节点权重值和与本节点之间边权重值乘积的和,比如节点 C 的权重值为 $w_{n=1}^{i,l=2}=\sum_{m=1}^{M}w_{n=1,m}^{i,l=2}$;

③ 第 3 层依次类推,特征 i 在第 n 个神经单元的最终权重值 $w_n^{i,l=3}=\sum_{m=1}^{M}w_{n,m}^{i,l=3}$,第 3 层所有神经单元的最终权重值构成 $\boldsymbol{WV}$ 向量.

3　无线网络攻击行为检测优化方法

本文在 2.2 节所构建的 2 种网络流量数据特征表示向量的基础上,结合深度学习方法中监督学习模型 DNN 和非监督学习聚类方法 Bi-kmeans 提出一种基于半监督学习的网络入侵检测方法 WiFi-ADOM.WiFi-ADOM 方法的总框架如图 3 所示,主要包括 5 个部分:数据预处理、数据特征表示(SSAE 模型)、攻击类型初步判别(DNN 模型)、未知攻击类型判别纠正项生成(Bi-kmeans 聚类方法)和攻击类型最终判别.

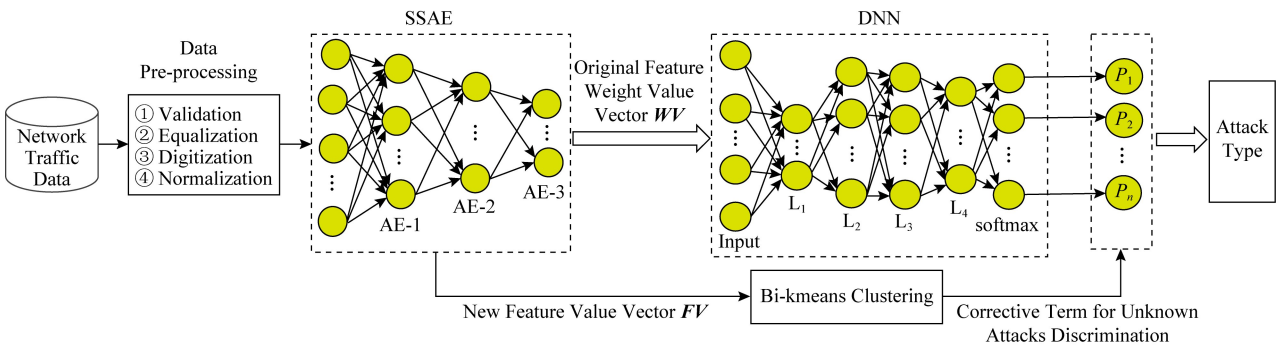

Fig. 3　Architecture of proposed method of WiFi-ADOM

图 3　基于半监督学习的网络入侵检测方法结构图

3.1　WiFi-ADOM 方法

1) 攻击类型初步判别

攻击类型初步判别采用深度学习模型 DNN,最后一层隐藏层的激活函数使用 softmax,将上一层的输入向量映射到另一个同维度向量,新向量元素的取值范围为[0,1],且所有元素的和为 1,如式(9)所示.模型的优化函数中增加了 L_2 正则项[28-29],如式(10)所示:

$$\sigma(z_{\text{softmax}})=\frac{e^{z_{\text{softmax}}}}{\sum_{m=1}^{M}e^{z^{m}_{\text{softmax}}}},\tag{9}$$

$$E_{\text{D}}=J_{\text{D}}+J_{\text{weights}},\tag{10}$$

$$J_{\text{D}}=\frac{1}{N}\sum_{n=1}^{N}\sum_{k=1}^{K}L(y_{nk},z_{nk}),\tag{11}$$

其中,M 表示神经单元的数量,z^{i}_{softmax} 表示 softmax 层第 i 个神经单元的输入数据;$L(y_{nk},z_{nk})$ 表示 y_{nk} 和 z_{nk} 之间的交叉熵损失函数,y_{nk} 和 z_{nk} 分别表示样本攻击类型的期望值和观测值,N 和 K 分别表示训练样本数量和特征数量.模型参数包括 2 种类型:基本参数,即权重值和偏置值;超级参数,包括批量训练样本数 *bath-size*、训练迭代次数 *epoch*、最大步长 *step*、学习率初始值 *learning rate*.为了加强模型全局寻优能力和加快模型收敛速度,本文针对训练过程中基本参数的初始化和更新采用 3 个优化措施:

① 使用 SSAE 生成的 ***WV*** 向量初始化 DNN 模型的权重矩阵,以避免权重参数失效问题.w 的初始化直接影响整个模型的训练效果和速度.如果初始值过小则输入的特征信息会在隐藏层之间前向传递时逐渐衰减而导致丢失.如果初始值过大则会严重影响模型训练过程中的梯度下降速度和学习速度,甚至导致结果无法收敛.另外,如果初始值设定为 0,则会导致所有神经单元的作用相同从而失去训练的意义.

② 采用 Adam 动量更新优化方法来调整模型基本参数:

$$w=w-\partial\frac{\dfrac{\beta_1 v_{\text{d}w}+(1-\beta_1)\text{d}w}{1-\beta_1}}{\sqrt{\dfrac{\beta_2 v_{\text{d}w}+(1-\beta_2)\text{d}w}{1-\beta_2}}+\varepsilon},\tag{12}$$

$$b=b-\partial\frac{\dfrac{\beta_1 v_{\text{d}b}+(1-\beta_1)\text{d}b}{1-\beta_1}}{\sqrt{\dfrac{\beta_2 v_{\text{d}b}+(1-\beta_2)\text{d}b}{1-\beta_2}}+\varepsilon},\tag{13}$$

其中,w 和 b 分别表示权重值和偏置值;β_1 和 β_2 分别表示第 1 个动量参数和第 2 个动量参数,本文中取值为 0.9 和 0.999;设置非零项 ε 以防止分母为零,本文取值为 10^{-8}.

③ 采用自适应学习步长来调整学习率 α,在训练迭代前期学习率较大,随着模型参数趋于稳定,学习率逐渐降低,α 可计算为

$$\alpha=\frac{\alpha_{\max}-\alpha_{\min}}{\exp(\lambda n/N)}+\alpha_{\min},\tag{14}$$

其中,$\alpha_{\max}$ 和 $\alpha_{\min}$ 分别表示学习率的最大值和最小值;λ 为学习率调节参数,本文取值为 40;n 为当前迭代次数,N 为总迭代次数.

2) 未知攻击类型判别纠正项

本文以 SSAE 模型生成的 ***FV*** 向量数据集作为输入数据,采取能够克服 kmeans 聚类存在的局部收敛问题的变种方法 Bi-kmeans 的聚类结果来生成未知攻击类型判别项.未知攻击行为判定纠正项为

$$P_{\text{u}}=ACC_{\text{kmeans}}\times P_{\text{kmeans}},\tag{15}$$

其中,ACC_{kmeans} 表示在 WiFi-ADOM 方法训练阶段 Bi-kmeans 方法对训练集分类所得到的准确率;P_{kmeans} 表示在方法有效性验证阶段 Bi-kmeans 方法对测试集中攻击行为的判别结果,判定为未知攻击类型时值为 1,否则为 0.

未知攻击行为判别项的生成过程包括 2 个部分:以训练集数据作为输入进行 Bi-kmeans 聚类以获取 ACC_{kmeans} 和判定过程需要用到的簇中心点信息;测试集数据中样本作为输入对攻击行为进行判别,即获取 P_{kmeans}.其中,Bi-kmeans 聚类的具体过程为:

① 将训练集中所有数据构建的所有点作为一个簇分成 2 部分;

② 选择其中一个可以最大强度降低误差平方和残差平方和的簇,使其继续划分为 2 部分;

③ 重复步骤②直至满足设定簇数 k,其中 k 值的选择采取斯坦福大学 Robert 教授提出的 Gap Statistic 方法,W_k 跌落最快的点便是最佳 k 值;

④ 记录各个簇中心点并计算准确率 ACC_{kmeans}.

攻击行为判定过程为:

① 本文采用欧氏距离作为聚类过程中的距离度量.计算测试集中样本数据所构建的点到每个已有簇内中心点及其邻近点的距离,取平均值作为测试样本点的簇距离.如果最小簇距离大于所对应簇内 2 点之间最大距离,且存在一定数量的类似节点,则判定为未知攻击类型,基于这些节点构建新簇.

② 如果未存在新簇,则重复步骤①.如果存在新簇,则计算测试样本点到每个已有簇内中心点及其邻近点的距离,取平均值作为测试样本点的簇距离.当测试样本点的最小距离簇为新簇时,判定为未知攻击类型,否则为已知攻击类型.

3.2 网络攻击行为检测流程

本文以网络流量数据作为输入,通过 WiFi-ADOM 方法实现对网络攻击行为的类型判别,具体实现过程为:

1) 对网络流量数据进行预处理,包括有效化、

均衡化、数字化和归一化,获得的数据作为 SSAE 模型的输入,分别构建训练数据集和测试数据集;

2) 通过训练 SSAE 模型获取网络流量数据新特征值向量 $\boldsymbol{FV}$ 和原始特征权重值向量 $\boldsymbol{WV}$;

3) 通过所有特征的 $\boldsymbol{WV}$ 向量构建权重矩阵,基于此初始化 DNN 模型的权重值,训练模型得到网络攻击行为的初始判别概率向量 $\boldsymbol{P}_{\mathrm{DNN}}=(P_{\mathrm{n}},P_{\mathrm{im}},P_{\mathrm{f}},P_{\mathrm{in}})$和模型准确率 ACC_{DNN}.其中 $P_{\mathrm{n}},P_{\mathrm{im}},P_{\mathrm{f}},P_{\mathrm{in}}$ 分别代表正常(normal)行为、伪装(impersonation)攻击行为、洪泛(flooding)攻击行为和注入(injection)攻击行为发生的概率;

4) $\boldsymbol{FV}$ 向量作为 Bi-kmeans 聚类方法的输入,得到未知攻击类型判别纠正项;

5) 基于初始判别概率添加未知攻击类型判别纠正项,得到网络行为的最终判别概率向量

$$\boldsymbol{P}=\left(\frac{ACC_{\mathrm{DNN}}}{ACC_{\mathrm{DNN}}+ACC_{\mathrm{kmeans}}}\times\boldsymbol{P}_{\mathrm{DNN}},\ \frac{ACC_{\mathrm{kmeans}}}{ACC_{\mathrm{DNN}}+ACC_{\mathrm{kmeans}}}\times P_{\mathrm{u}}\right),$$

所有元素的和为 1,值最大的元素所对应的攻击行为类别即为最终判别结果.当 Bi-kmeans 方法对测试样本的判别结果为未知攻击类型时,未知攻击类型判别纠正项为 1,Bi-kmeans 方法的判别结果和 DNN 模型的判别结果同时决定测试样本的攻击类型,ACC_{kmeans}越大则未知攻击类型的可能性越大;反之,未知攻击类型判别纠正项为 0,DNN 模型对测试样本攻击类型的判别起主导作用,ACC_{kmeans}取值的大小不影响判别结果.

WiFi-ADOM 算法伪代码为:

① function WiFi-ADOM
② 　function 数据预处理(训练数据集)
③ 　　for 样本数据 do
④ 　　数据有效化;
⑤ 　　数据均衡化;
⑥ 　　数据数值化;
⑦ 　　数据归一化(式(16));
⑧ 　　end for
⑨ 　　return 数据集 R;/ * 95 维 * /
⑩ 　function 网络流量数据特征表示向量生成(数据集 R)
⑪ 　　for $i=1$ to h do
　　/ * $h=3$, SSAE 隐藏层数量 * /
⑫ 　　for 训练样本 do
⑬ 　　计算 y(式(1));
⑭ 　　计算 z(式(2));
⑮ 　　最小化 E_{S}(式(4));
⑯ 　　更新 $\theta_i=\{w,b\}$;
⑰ 　　end for
⑱ 　　$y\leftarrow L_i$; / * SSAE 中 AE 的隐藏层,维度为 70,50,30 * /
⑲ 　end for
⑳ 　新特征值向量 $\boldsymbol{FV}\leftarrow$30 维向量;
　　/ * 详见 2.2 节 1) * /
㉑ 　原始特征权重值向量 $\boldsymbol{WV}\leftarrow$95 维向量;
　　/ * 详见 2.2 节 1) * /
㉒ 　return $\boldsymbol{FV}$,$\boldsymbol{WV}$;
㉓ function DNN 模型训练(训练数据集,$\boldsymbol{WV}$)
㉔ 　基于 $\boldsymbol{WV}$ 初始化 DNN;
㉕ 　训练 DNN;
㉖ 　最小化目标函数 E_{D}(式(10));
㉗ 　return $\theta=\{w,b\}$和 DNN 准确率 ACC_{DNN};
㉘ function Bi-kmeans 聚类($\boldsymbol{FV}$)
㉙ 　Bi-kmeans 聚类; / * 详见 3.1 节 2) * /
㉚ 　return 准确率 ACC_{kmeans},聚类各簇中心点;
㉛ function 测试样本攻击类型判定(测试数集、θ、ACC_{DNN}、ACC_{kmeans}、簇中心点)
㉜ 　数据预处理(测试数据集);
㉝ 　基于 $\theta=\{w,b\}$生成攻击类型初判概率向量;
㉞ 　生成未知攻击类型判别纠正项;
　　/ * 详见 3.1 节 2) * /
㉟ 　生成最终判别概率向量 $\boldsymbol{P}$;/ * 详见 3.2 * /
㊱ 　攻击类型←向量 $\boldsymbol{P}$ 中元素的最大值;
㊲ 　return 攻击类型.

4　实验与结果

本节在无线网络攻击数据集 AWID 上分别测试了 WiFi-ADOM 方法对已知攻击类型和未知攻击类型的检测性能.本文的实验环境为 Windows Server 2010 操作系统、Intel Xeon 2.0 GHz CPU、512 GB 内存,WiFi-ADOM 方法和对比方法基于 python3.6 和 TensorFlow1.3 实现.

4.1　数据集

AWID 数据集来源于 Kolias,是数量最大也是

最全面的真实 WiFi 网络环境下采集的网络攻击数据集[22].按照攻击类型级别,数据集被划分为 2 种数据子集:4 种大攻击类型的 CLS 数据集和 16 种子攻击类型的 ATK 数据集.后者的 16 种子攻击类型包含在前者的 4 种大攻击类型中,比如:ATK 数据集中的 Caffe-Latte, Hirte, Honeypot 和 EvilTwin 攻击类型属于 CLS 数据集中的伪装攻击类型.同时 AWID 数据集包含完整数据集和精简数据集 2 个版本.本文使用精简版本的 CLS 数据集,其分布情况如表 1 所示:

Table 1 The Distribution of AWID

表 1 AWID 数据集的分布情况

Type	Training Set	Testing Set
Normal	1 633 190	530 785
Flooding	48 484	8 097
Impersonation	48 522	20 079
Injection	65 379	16 682
Total	1 795 575	575 643

数据集的预处理过程包括数据有效化、均衡化、数值化和归一化 4 个部分.

1) 数据有效化

网络流量数据中有些特征值属于缺失状态,为了保证输入数据的有效性,删除属性值中只有 20% 属于正常状态的特征和所有值都相同的特征,其余所有缺失值都用 0 来填充.最后 154 维特征减少为 95 维.

2) 数据均衡化

数据集中正常行为记录和攻击行为记录的比例高达 10:1,这种数据不平衡问题将会严重影响模型的效果,所以本文对数据进行了均衡化.本文从原始数据集中随机抽取 10% 的正常行为记录与原有的攻击行为记录组成新的数据集.

3) 数据数值化

把数据集中十六进制表示的特征值统一转化为十进制,由于 mac 地址既非数值其总量也非常庞大,所以本文把属性值转换为 mac 地址在整个数据集中出现的次数.标签列即攻击行为类型列有 4 种属性值,分别映射为 one-hot 形式的四维向量,比如:对于泛洪攻击行为映射为 0001.

4) 数据归一化

数据集中不同特征的值域大不相同,为了消除这种差别给模型带来的影响,对数值型数据进行归一化.本文采用最值归一化法把属性值映射到区间[0,1]之间:

$$y_i=\frac{x_i-\min(x)}{\max(x)-\min(x)}, \tag{16}$$

其中,y_i 表示第 i 个特征值归一化之后的值,x_i 表示第 i 个特征值,$\min(x)$和 $\max(x)$分别代表特征属性值取值范围内的最小值和最大值.

4.2 聚类性能度量标准

本文采用准确率(accuracy, ACC)、召回率($Recall$)、误判率(false alarm rate, FAR)、F_1 作为网络攻击行为检测方法的性能评价指标.

1) 准确率(ACC),是衡量检测方法的整体有效性,可计算为

$$ACC=\frac{TP+TN}{TP+TN+FP+FN}. \tag{17}$$

2) 误判率(FAR)表示被错误判定为攻击行为的数量和正常行为总量的比值,可计算为

$$FAR=\frac{FP}{TN+FP}. \tag{18}$$

3) 召回率($Recall$)表示准确判定为攻击行为的数量占攻击行为总量的比值,可计算为

$$Recall=\frac{TP}{TP+FN}. \tag{19}$$

4) F_1 指数综合考量了准确率和召回率的评估内容,其中准确率表示准确检测到的攻击行为占检测到的攻击行为总量的比例,可计算为

$$F_1=\frac{2TP}{2TP+FP+FN}, \tag{20}$$

其中,TN(true negative)表示把正常行为正确判别为正常行为的数量,TP(true positive)表示把攻击行为正确判别为攻击行为的数量,FN(false negative)表示把攻击行为错误判别为正常行为的数量,FP(false postive)表示把正常行为错误判别为攻击行为的数量.

4.3 实验和结果

本文分别从已知攻击类型检测和未知攻击类型检测 2 个方面来评估 WiFi-ADOM 方法的检测性能.在未知攻击行为检测中,通过在原始数据集中删除特定类别的攻击类型,构建包含未知攻击类型的数据集.为了保证检测性能评估的全面性,除了正常行为类型外,其他 3 种大攻击类型:泛洪攻击、伪装攻击、注入攻击都分别设定为未知攻击类型,且每次实验都重复 10 次,取其平均值作为最后的结果.通过对比不同网络层数和神经单元数量设定下模型的重构误差和性能指标表现,SSAE 模型的网络结构为 95:70:50:30,DNN 模型的网络结构为 95:30:60:40:20:4.

1) 面向攻击行为检测的特征重要性分析

根据 2.2 节中原始特征权重网络的构建方法,构建 WiFi-ADOM 方法 DNN 模型中原始特征的权重网络,把获取到的每一个原始特征的最终权重值作为特征重要性度量,值越大表示对攻击检测的作用越大,值越小则表示作用越小.根据最终权重值,选择前 15 个特征属性作为与攻击检测相关的重要特征.本文分别从 2 个方面分析了攻击行为检测中特征属性的重要性:DNN 模型不同隐藏层对特征属性重要性的影响;通过对比 WiFi-ADOM 和 C4.5,D-FES-SVM[12],DNN,SVM 不同方法的重要特征选择结果探究网络攻击检测中特征属性的重要性.

结果如图 4 所示,行表示不同检测方法,列表示不同特征,浅红色为底色,深红色区域表示列所对应的特征是行所对应检测方法的重要特征属性.由图 4 可以看出,重要性热度最高的是特征 82,107,154,作为重要特征的比例是 5/5;其次是特征 38,71,108,122,作为重要特征的比例是 4/5.由于 DNN 模型中第 4 层隐藏层的选择结果即最终选择,所以用 WiFi-ADOM 标识,不同隐藏层选择的相同重要特征有 10 个,其余 5 个各有不同;相对于第 2 层隐藏层,第 3 层隐藏层和最终选择结果相似度更高.

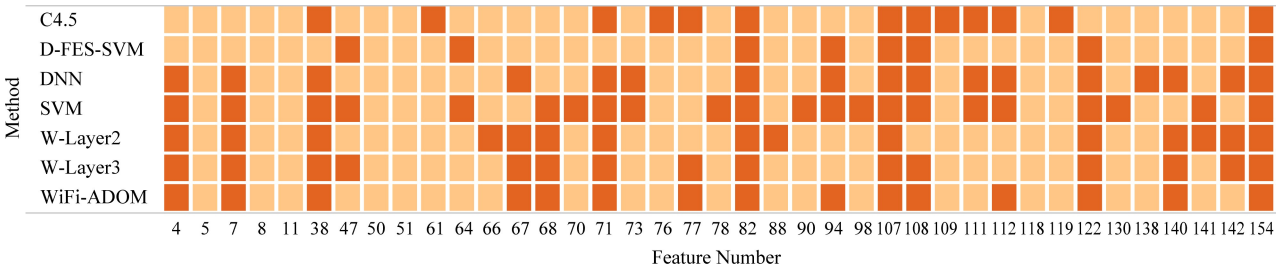

Fig. 4　The importance heat map of characteristic in attracts detection

图 4　面向攻击检测的原始特征重要性热点图

图 5 分析了 WiFi-ADOM 方法中 DNN 模型不同隐藏层 Top20 原始特征属性的权重值排序的变化情况.其中,点代表不同特征的权重值排序,浅蓝色和蓝色的点分别表示第 2 层和第 3 层隐藏层选择的特征属性;深蓝色的点表示第 4 层隐藏层的选择结果即最终选择,用 WiFi-ADOM 标识;红色常量线表示排序为 5,红线以上的设定为重要特征.由图 5 可得,不同隐藏层选择的重要特征的重要程度并不是一致的,有的逐渐上升,如特征 142;有的逐渐下降,如特征 82;有的则差别比较大,如特征 112.从整体来看,相比第 2 层,第 3 层隐藏层的选择结果和第 4 层的更相近.结合图 4 中第 3 层比第 2 层更接近最终结果的分析,可知深度学习模型 DNN 的隐藏层数对网络流量特征属性重要性的分析结果有直接影响,在模型已完成调优的特定层数范围内,隐藏层越深其分析结果越准确.

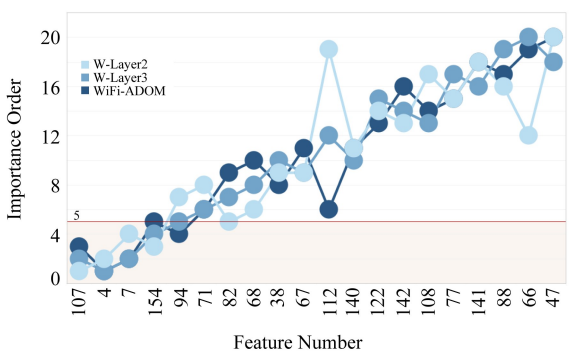

Fig. 5　Important order of original features detection

图 5　面向攻击检测的原始特征重要性排序

2) 已知攻击类型检测性能分析

表 2 分析了 SSAE 模型中不同 AE 隐藏层对 WiFi-ADOM 方法性能的影响.AE-1,AE-2,AE-3 中隐藏层形成的新特征值作为 Bi-kmeans 聚类方法的输入数据,根据检测情况分别计算不同性能指标.表 2 中,性能指标列表现最优的结果用粗体格式加以标识.由表 2 可得,*ACC* 和 F_1 均随着隐藏层的增加而不断增大;*FAR* 的值依次降低;*Recall* 的值则先稍稍降低后又增加.栈式稀疏自编码器模型基于特征提取对网络流量特征的整体学习能力优于单个稀疏自编码器模型.

Table 2　Impacts of Different AEs for SSAE Model

表 2　SSAE 模型中不同 AE 隐藏层对检测性能的影响 %

Layer	*ACC*	*FAR*	*Recall*	F_1
AE-1	95.14	0.76	95.15	94.53
AE-2	97.20	0.13	95.07	95.11
AE-3(WiFi-ADOM)	**98.56**	**0.05**	**97.21**	**97.99**

Note: The best values are in bold.

表 3 对比了不同 AE 隐藏层对不同攻击类型检测准确率的影响.其中,F 表示泛洪攻击,In 表示注入攻击,Im 表示伪装攻击.从表 3 可以看出,对正常行为检测的准确率几乎没有改变,都高于 99.80%;对伪装攻击和注入攻击检测的 *ACC* 都是逐渐增加,且都维持在较高的水平.对泛洪攻击检测的 *ACC* 增加最多,但是 *ACC* 值却低于 80%,有待进一步改进.在对不同网络攻击类型的检测中,栈式稀疏自编码器模型基于特征提取对网络流量特征的学习能力优于单个稀疏自编码器模型,但是在数据不充分的攻击类型检测中具有局限性.

Table 3 Impacts of Hidden Layer in Different AEs on *ACC* for SSAE Model

表 3 SSAE 模型中不同 AE 隐藏层对检测准确率的影响 %

Layer	Normal	F	In	Im	Total
AE-1	99.81	69.93	99.03	95.07	98.02
AE-2	99.85	70.95	99.20	96.97	98.31
AE-3(WiFi-ADOM)	**99.87**	**72.31**	**99.21**	**96.99**	**98.56**

Note: The best values are in bold.

表 4 对比了 WiFi-ADOM 方法和 SAE+DNN,DNN,SVM 方法在不同检测性能指标上的表现.由表 4 可知:WiFi-ADOM 在 *ACC*,*Recall* 和 F_1 分别取值最高,即 98.56%,97.21%,97.99%,在 *FAR* 取值最小,即 0.05%,总体性能最优;然后依次是 SAE+DNN 和 DNN,最后是 SVM.DNN 和 SAE 模型结合的检测性能优于 DNN 模型.本文根据 SSAE 模型生成的原始特征权重向量初始化 DNN,在 DNN 进行特征选择和分类的基础上,不仅接收了 SSAE 能够提取特征的能力,也保留了完整的原始特征信息,所以进一步优化了 SAE+DNN 方法的检测性能.

Table 4 Performance Comparation of Different Methods

表 4 不同网络攻击行为检测方法的检测性能对比 %

Methods	*ACC*	*FAR*	*Recall*	F_1
WiFi-ADOM	**98.56**	**0.05**	**97.21**	**97.99**
SAE+DNN	97.85	0.06	99.01	95.03
DNN	97.51	1.24	98.95	95.22
SVM	95.79	0.07	97.02	5.5

Note: The best values are in bold.

表 5 对比了 WiFi-ADOM 方法和 SAE+DNN,DNN,SVM 方法对不同攻击类型检测的性能指标 *ACC*.从表 5 可以看出,WiFi-ADOM 方法对 4 种攻击类型检测的 *ACC* 均取得最高值,在准确率上表现性能最优.4 种方法对正常行为类型和注入攻击类型的检测指标 *ACC* 均大于 95%,处于较优性能,但是对泛洪攻击的检测准确率均小于 80%.对于伪装攻击的检测准确率,SVM 方法的 *ACC* 只有 15.5%;其他 3 种方法的 *ACC* 都高于 95%,处于较优性能.由于篇幅有限,本文只展示了不同方法在检测准确率指标上的性能表现,其他性能指标对比结果类似,WiFi-ADOM 方法的检测性能优于其他方法.

Table 5 Performance Comparison of Different Methods on *ACC*

表 5 不同检测方法对不同攻击类型的检测准确率对比 %

Methods	Normal	F	In	Im	Total
WiFi-ADOM	**99.87**	**72.31**	**99.21**	**96.99**	**98.56**
SAE+DNN	97.85	70.35	99.01	95.03	96.32
DNN	97.51	69.71	98.95	95.22	95.97
SVM	95.79	69.52	97.02	15.5	91.76

Note: The best values are in bold.

3) 未知攻击类型检测性能分析

表 6 对比了 3 种攻击类型:洪泛攻击 F、入侵攻击 In、伪装攻击 Im 分别作为未知攻击类型的情况下,WiFi-ADOM 方法在性能指标 *ACC* 上的表现.由表 6 可知,3 种情况下对正常行为的检测准确率几乎没有变化,均处于较优性能;因为在检测过程中未知攻击行为形成新簇需要一定时间,在此期间并不能有效检测到未知攻击行为,对整体准确率有所影响,所以未知攻击行为的 *ACC* 均比其他已知攻击行为的 *ACC* 下降得要稍微多一些,但是在可接受的范围内;在泛洪攻击为未知攻击行为的情况下,泛洪攻击检测 *ACC* 下降得最多.综上所述,WiFi-ADOM 方法在保证检测性能的同时可以有效检测未知攻击行为,实现了对网络攻击行为检测的优化.

Table 6 Performance of WiFi-ADOM on Unknown Attacks

表 6 WiFi-ADOM 方法对未知攻击类型的检测结果 %

Type	Normal	F	In	Im	Total
F	99.41	59.01	99.02	95.76	96.47
In	99.35	69.23	97.65	96.21	98.31
Im	99.34	70.15	99.17	95.21	98.02
None	99.87	72.31	99.21	96.99	98.56

5 总 结

在无线网络攻击行为日渐猖狂的当下,有效的

网络攻击行为检测成为当务之急,既要保证对已知攻击类型的检测性能:高准确率、低误判率、低漏判率等,又要具备检测未知攻击类型的能力.虽然相对于传统检测方法,目前利用机器学习方法进行的相关研究在提高准确率、降低误报率方面已取得一定的进展,但是由于网络流量数据不断指数级增长和无线网络攻击行为持续演化升级,依然存在无法有效检测未知攻击类型等多种问题.因此,本文结合属于非监督学习的 SSED 模型、Bi-kmeans 聚类方法和属于监督学习的 DNN 模型,提出一种基于半监督学习的网络攻击行为检测优化方法 WiFi-ADOM,并探究了在不同攻击类型检测中特征属性的重要性.实验结果表明,本文提出的 WiFi-ADOM 方法在攻击检测性能方面达到了很好的优化效果,尤其是对未知攻击类型的检测,但是对于泛洪攻击作为未知攻击类型的情况,WiFi-ADOM 方法的优化效果还比较有限.未来的研究会从 2 个方面展开:1)探究网络流量数据中特征属性的具体特性及其在网络攻击行为检测中的作用,并在 WiFi-ADOM 方法的基础上针对泛洪攻击作为未知攻击类型的情况进一步优化;2)评估子攻击类型(比如 Caffe-Latte 攻击)作为未知攻击类型的情况下 WiFi-ADOM 方法的检测效果.

参　考　文　献

[1] Zhang Yuqing, Dong Ying, Liu Caiyun, et al. Situation, trends and prospects of deep learning applied to cyberspace security [J]. Journal of Computer Research and Development, 2018, 55(6): 1117-1142 (in Chinese)
(张玉清, 董颖, 柳彩云, 等. 深度学习应用于网络空间安全的现状、趋势与展望[J]. 计算机研究与发展, 2018, 55(6): 1117-1142)

[2] Sommer R, Paxson V. Outside the closed world: On using machine learning for network intrusion detection [C] //Proc of 2010 IEEE Symp on Security and Privacy. Piscataway, NJ: IEEE, 2010: 305-316

[3] Shirazi H M, Kalaji Y. An intelligent intrusion detection system using genetic algorithms and features selection [J]. Majlesi Journal of Electrical Engineering, 2010, 4(1): 33-43

[4] Anderson J P. Computer security threat monitoring and surveillance [R]. Fort Washington, CA: James P. Anderson Co, 1980

[5] Heberlein L T, Dias G V, Levitt K N, et al. A network security monitor [C] //Proc of 1990 IEEE Computer Society Symp on Research in Security and Privacy. Piscataway, NJ: IEEE, 1990: 296-304

[6] Kelly J P, Cook S F, Kaufman D W, et al. Prevalence and characteristics of opioid use in the US adult population [J]. Pain, 2008, 138(3): 507-513

[7] Farooqi A H, Khan F A. Intrusion detection systems for wireless sensor networks: A survey [C] //Proc of Int Conf on Future Generation Communication and Networking. Berlin: Springer, 2009: 234-241

[8] Scarfone K, Mell P. Guide to intrusion detection and prevention systems (IDPS) [R]. Washington: US Department of Commerce, 2012

[9] Potluri S, Diedrich C. Accelerated deep neural networks for enhanced intrusion detection system [C] //Proc of 2016 IEEE 21st Int Conf on Emerging Technologies and Factory Automation (ETFA). Piscataway, NJ: IEEE, 2016: 1-8

[10] 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

[11] Thing V L L. IEEE 802.11 network anomaly detection and attack classification: A deep learning approach [C] //Proc of 2017 IEEE Wireless Communications and Networking Conf (WCNC). Piscataway, NJ: IEEE, 2017: 1-6

[12] Aminanto M E, Choi R, Tanuwidjaja H C, et al. Deep abstraction and weighted feature selection for WiFi impersonation detection [J]. IEEE Transactions on Information Forensics and Security, 2018, 13(3): 621-636

[13] You Lina, Li Yujun, Wang Yue, et al. A deep learning-based RNNs model for automatic security audit of short messages [C] //Proc of 2016 16th Int Symp on Communications and Information Technologies (ISCIT). Piscataway, NJ: IEEE, 2016: 225-229

[14] Kang M J, Kang J. Intrusion detection system using deep neural network for in-vehicle network security [J]. Plos One, 2016, 11(6): e0156530

[15] Liang Jie, Chen Jiahao, Zhang Xueqin, et al. One-hot encoding and convolutional neural network based anomaly detection [J]. Journal of Tsinghua University: Science and Technology, 2019, 59(7): 523-529 (in chinese)
(梁杰, 陈嘉豪, 张雪芹, 等. 基于独热编码和卷积神经网络的异常检测[J]. 清华大学学报: 自然科学版, 2019, 59(7): 523-529)

[16] Zhang Sicong, Xie Xiaoyao, Xu Yang. Intrusion detection mehtod based on a deep convolutional neural network [J]. Journal of Tsinghua University: Science and Technology, 2019, 59(1): 46-54 (in Chinese)
(张思聪, 谢晓尧, 徐洋. 基于 dCNN 的入侵检测方法[J]. 清华大学学报: 自然科学版, 2019, 59(1): 46-54)

[17] Akashdeep, Manzoor I, Kumar N. A feature reduced intrusion detection system using ANN classifier [J]. Expert Systems with Applications, 2017, 88: 249-257

[18] Wang Wei, Zhu Ming, Wang Jinlin, et al. End-to-end encrypted traffic classification with one-dimensional convolution neural networks [C] //Proc of 2017 IEEE Int Conf on Intelligence and Security Informatics (ISI). Piscataway, NJ: IEEE, 2017: 43-48

[19] Louvieris P, Clewley N, Liu Xiaohui. Effects-based feature identification for network intrusion detection [J]. Neurocomputing, 2013, 121: 265-273

[20] Zhu Yingying, Liang Junwei, Chen Jianyong, et al. An improved NSGA-III algorithm for feature selection used in intrusion detection [J]. Knowledge-Based Systems, 2017, 116: 74-85

[21] Usha M, Kavitha P. Anomaly based intrusion detection for 802.11 networks with optimal features using SVM classifier [J]. Wireless Networks, 2017, 23: 2431-2446

[22] Kolias C, Kambourakis G, Stavrou A, et al. Intrusion detection in 802.11 networks: Empirical evaluation of threats and a public dataset [J]. IEEE Communications Surveys & Tutorials, 2016, 18(1): 184-208

[23] Zhao Juan, Shetty S, Pan Janwei, et al. Transfer learning for detecting unknown network attacks [J]. EURASIP Journal on Information Security, 2019, 1: 63-71

[24] Glorot X, Bengio Y. Understanding the difficulty of training deep feedforward neural networks [C] //Proc of the 13th Int Conf on Artificial Intelligence and Statistics. Palo Alto, CA: The Association for the Advancement of Artificial Intelligence, 2010: 249-256

[25] Olshausen B A, Field D J. Sparse coding with an overcomplete basis set: A strategy employed by V1? [J]. Vision Research, 1997, 37(23): 3311-3325

[26] Eskin E, Arnold A, Prerau M, et al. A geometric framework for unsupervised anomaly detection [J]. Applications of Data Mining in Computer Security, 2002, 4: 77-101

[27] Almusallam N Y, Tari Z, Bertok P, et al. Dimensionality reduction for intrusion detection systems in multi-data streams—A review and proposal of unsupervised feature selection scheme [J]. Emergent Computation, 2017, 24: 467-487

[28] Guerra L, McGarry L M, Robles V, et al. Comparison between supervised and unsupervised classifications of neuronal cell types: A case study [J]. Developmental Neurobiology, 2011, 71(1): 71-82

[29] Møller M F. A scaled conjugate gradient algorithm for fast supervised learning [J]. Neural Networks, 1993, 6(4): 525-533

Wang Ting, born in 1987. PhD, assistant researcher. Her main research interests include information management and data mining, big data analytics.

Wang Na, born in 1985. Bachelor, engineer. Her main research interests include information management and network maintenance.

Cui Yunpeng, born in 1972. PhD. Professor. His main research interests include the application of machine learning, natural language processing, data analysis and visualization in agriculture.

Li Huan, born in 1992. MA, assistant researcher. Her main research interests include natural language processing and information retrieval.