

文章编号: 1674—8247(2024)05—0067—05
DOI:10.12098/j.issn.1674-8247.2024.05.011

高速铁路信号系统网络入侵检测技术研究

曹 峰¹ 林瑜筠²

(1. 南京铁道职业技术学院轨道交通工程实践中心, 南京 210031; 2. 南京铁道职业技术学院, 南京 210031)

摘 要:入侵检测作为一种网络主动防御技术,能够有效阻止来自黑客的多种手段攻击。随着机器学习的发展,相关技术也开始应用到入侵检测中。本文采用 sklearn 库中 preprocessing 模块的函数对 KDD CUP 99 数据集进行预处理,基于朴素贝叶斯和逻辑回归算法,建立了网络入侵检测模型,并利用信息增益算法对入侵相关特征进行选择,然后进行训练与预测。实验结果表明,选择特征子集进行训练和预测能够保证预测准确率并大幅提高检测效率。研究成果可为高速铁路信号系统网络入侵检测模型的设计和建立提供参考。

关键词:信号系统;入侵检测;机器学习;KDD CUP 99 数据集;朴素贝叶斯;逻辑回归

中图分类号:U284 **文献标志码:**A

Study on Network Intrusion Detection Techniques for High-speed Railway Signal Systems

CAO Feng¹ LIN Yujun²

(1. Rail Transit Engineering Practice Center, Nanjing Vocational Institute of Railway
Technology, Nanjing 210031, China;

2. Nanjing Vocational Institute of Railway Technology, Nanjing 210031, China)

Abstract: Intrusion detection, as an active defense mechanism in networking, effectively thwarts diverse forms of attacks by hackers. With the advancements in machine learning, related technologies are increasingly being employed in intrusion detection systems. This study utilized preprocessing functions from the sklearn library's preprocessor module to preprocess the KDD CUP 99 dataset. Based on Naive Bayes and logistic regression algorithms, a network intrusion detection model was constructed, followed by feature selection using the information gain algorithm prior to training and prediction. Experimental results demonstrate that training and predicting with a subset of selected features ensures prediction accuracy while significantly boosting detection efficiency. The findings provide valuable reference for the design and establishment of network intrusion detection models in high-speed railway signal systems.

Key words: signal systems; intrusion detection; machine learning; KDD CUP 99 dataset; Naive Bayes; logistic regression

作为社会基础设施建设的重要组成部分,网络安全
安全已上升为国家安全战略。高速铁路信号系统在提升

其信息化、数字化、网络化和智能化水平的同时也面临
诸多网络安全威胁,对高速铁路运营安全造成严重

收稿日期:2020-03-15

作者简介:曹峰,(1987-),男,工程师。

基金项目:教育部高铁安全协同创新中心、江苏省高铁安全工程技术研究开发中心科研项目(GTAQ202204)

引文格式:曹峰,林瑜筠.高速铁路信号系统网络入侵检测技术研究[J].高速铁路技术,2024,15(5):67-71.

CAO Feng, LIN Yujun. Study on Network Intrusion Detection Techniques for High-speed Railway Signal Systems[J]. High Speed Railway Technology, 2024, 15(5):67-71.

影响^[1]。

入侵检测是当今网络环境下实现信息保障的纵深防御框架中一项重要技术。随着网络流量剧增及复杂性提升,数据量大幅增长且维度不断扩展,入侵检测面临着高效准确处理海量数据的巨大挑战。因此,在网络入侵检测中,从大量的网络流量中选择最重要的特征,去除无关数据、降低维数,并利用它们快速有效地检测入侵,提高检测精度,具有十分重要的意义^[2]。

近几年来,随着机器学习逐渐兴起,国内外对入侵检测的研究不断深入。相关学者采用决策树、贝叶斯网络模型、支持向量机等算法构建了入侵检测模型^[3-4]。

然而,目前的入侵检测系统存在许多问题,如低检测率,不同类型攻击的检测率差别较大、误检率较高。此外,由于存在不相关和冗余属性,检测过程变得更加复杂,同时也会降低检测率,协议识别、病毒过滤等一些额外属性也可能增加运算时间,影响检测率。为解决以上问题,对训练集在训练之前进行特征选择,从而降低数据维数,提高分类效率^[5-6]。

本文基于 KDD CUP 99 数据集,利用信息增益进行特征筛选,通过朴素贝叶斯、逻辑回归算法建立模型完成高速铁路信号系统网络入侵检测。

1 入侵检测原理

入侵检测是一种网络主动防御技术,它从计算机网络系统中的若干关键点收集信息,并进行分析是否存在违反安全策略的行为和遭到攻击的迹象,在不影响网络性能的情况下对网络进行监测,从而提供对内部攻击、外部攻击和误操作的实时保护^[7-9]。

入侵检测的意义在于能够将数据流中正常与异常数据区分,实现对攻击行为的报警。通常将数据集中一部分训练集用于训练分类器,然后用另一部分测试集预测分类器性能,分类器可以是基于贝叶斯、决策树、神经网络等算法构建的机器学习模型^[10-11],基本流程如图 1 所示。

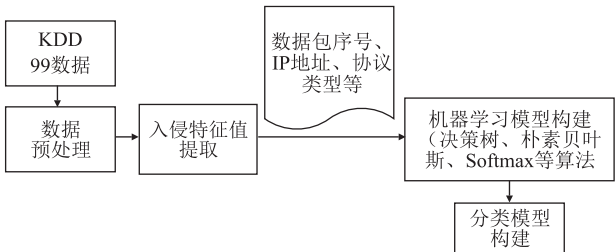


图 1 基于机器学习的网络入侵检测图

2 KDD CUP 99 入侵检测数据

KDD CUP 99 数据集源自相关机构参与资助和研

发的入侵检测项目。项目中使用海量的网络通信数据作为研究对象,主要用来建立更高效的入侵检测模型和算法,评估入侵检测系统的检测性能。由于机要部门网络敏感性,采集了数量众多、种类齐全的实际攻击行为的数据记录^[12]。采集长达 3 个多月,最终收集将近 700 多万条以标识类型格式保存的网络连接数据。其中,每条信息用 41 个特征来描述,标签有 5 种,分别为 Normal、DOS、Probing、R2L、U2R。正常的网络连接标记为 Normal,其余 4 种为异常网络攻击,每一种网络攻击又包含多种攻击类型。在训练数据集中共有 22 种攻击类型,在测试数据集中还包含另外 14 种攻击类型,如表 1 所示。

表 1 网络连接标识分类表

标识类型	含义	具体分类标识
Normal	正常记录	Normal
DOS	拒绝服务攻击	Back、land、neptune、pod、smurf、teardrop
Probing	监视和其他探测活动	ipsweep、nmap、portsweep、satan
R2L	来自远程机的非法访问	ftp-write、guess-passwd、imap、multihop、phf、spy、warezclitent、warezmaster
U2F	普通用户对本地超级用户特权的非法访问	buffer-overflow、loadmodule、perl、rootkit

3 算法

3.1 信息增益法

信息增益法是一种简单高效的特征选择方法,其基本思路是将各个特征与结果的类别关系进行量化排序,选择保留关联性强的特征。在信息增益中,重要的衡量标准是特征能为最后的分类带来多少信息量,信息量越多,此特征越重要。本文用熵的定义来对信息量进行评估,而信息增益可由熵和条件熵进行计算。

3.1.1 信息熵及相关定义

一个事件或系统,准确的说是一个随机变量,它有着一定的不确定性。随机变量的不确定性越高,要消除其不确定性则需引入更多的信息,这些信息的度量称作信息熵。对随机事件的信息量求期望,得熵的定义:

$$H(X) = - \sum_{x \in X} p(x) \ln p(x) \tag{1}$$

式中: x ——随机变量 X 的一个特定取值;

X ——随机变量 X 所有可能取值的集合;

$H(X)$ ——随机变量 X 的熵,衡量了其不确定性或信息量;

$p(x)$ ——随机变量 X 取特定值 x 的概率。

熵是随机变量不确定性的度量,不确定性越大,熵

值就越大。两个随机变量的联合分布可形成联合熵,用 $H(X,Y)$ 表示。 (X,Y) 发生所包含的熵减去 X 单独发生所包含的熵表示在 X 发生的前提下, Y 发生“新”带来的熵^[13]。根据条件熵的定义可得:

$$H(X,Y) - H(X) = - \sum_{x \in X, y \in Y} p(x,y) \log p(y|x) = \sum_{x \in X} p(x) H(Y|X=x) \quad (2)$$

式中, $H(X,Y)$ 表示随机变量 X 和 Y 的联合熵; $H(Y|X=x)$ 表示在给定随机变量 $X=x$ 的条件下随机变量 Y 的条件熵。

两个随机变量 X,Y 的互信息,定义为 X,Y 的联合分布和独立分布乘积的相对熵:

$$I(X,Y) = \sum_{x \in X, y \in Y} p(x,y) \log \frac{p(x,y)}{p(y)p(x)} \quad (3)$$

式中: $I(X,Y)$ ——随机变量 X 和 Y 之间的互信息。

因此,条件熵也可用 $H(X) - I(X,Y)$ 来计算。

3.1.2 信息增益定义及计算

特征 A 对训练数据集 D 的信息增益 $G(D,A)$, 定义为集合 D 的经验熵 $H(D)$ 与特征 A 给定条件下 D 的经验条件熵 $H(D,A)$ 之差,即:

$$G(D,A) = H(D) - H(D,A) \quad (4)$$

显然,这就是训练数据集 D 和特征 A 的互信息。

在实际进行特征选择过程中,先计算整个数据集 D 的经验熵。再遍历所有特征,对于特征 A , 计算特征 A 对数据集 D 的经验条件熵 $H(D,A)$, 接着计算 A 的信息增益 $G(D,A)$ 。最后,将信息增益按从大到小排序,根据多次实验结果得出选择信息增益大的前 10 个特征作为特征选择效果最佳,并将其作为入侵检测分类的依据。

3.2 机器学习算法

3.2.1 朴素贝叶斯

贝叶斯算法是以贝叶斯原理为基础,使用概率统计的知识对样本数据集进行分类,其特点是结合先验概率和后验概率,避免了主观偏见和过拟合现象。朴素贝叶斯算法是在贝叶斯算法基础上进行了简化。朴素贝叶斯分类通过给定的训练集,假定其属性之间条件独立,学习从输入到输出的联合概率分布,再基于数据模型,求出后验概率最大的输出^[14]。

贝叶斯分类的工作原理为:

$$P(y_i | X) = - \frac{P(X | y_i) P(y_i)}{P(X)} \quad (5)$$

式中, $X = (x_1, x_2, \dots, x_n)$ 是特征向量; y_i 是类别。

因此,贝叶斯分类就是寻找 $P(y_i | X)$ 对应的最大值,最大值对应的类别即为预测类别。然而,在实际情

况中 $P(y_i | X)$ 往往无法直接求得,所以根据贝叶斯公式将求解 $P(y_i | X)$ 转化为求解:

$$\hat{y} = \arg \max_y P(X | y_i) P(y_i) P(X) \quad (6)$$

朴素贝叶斯中的“朴素(naive)”是指假设特征向量 $X = (x_1, x_2, \dots, x_n)$ 中的各个特征之间相互独立,即:

$$P(X | y_i) P(y_i) = P(x_1, x_2, \dots, x_n | y_i) P(y_i) = P(y_i) \prod_{j=1}^n P(x_j | y_i) \quad (7)$$

通过最大后验概率来进行估计:

$$\hat{y} = \arg \max_y \prod_{i=1}^n P(x_i | y) \quad (8)$$

式中: \hat{y} ——预测的类别标签。

朴素贝叶斯分类器的差异大部分来自于处理分布时的所做的假设不同。对于高斯朴素贝叶斯,其公式为:

$$P(x_i | y) = \frac{1}{\sqrt{2\pi\sigma_y^2}} e^{-\frac{(x_i - \mu_y)^2}{2\pi\sigma_y^2}} \quad (9)$$

式中: x_i ——给定样本的第 i 个特征;

μ_y ——给定类别 y 下特征 x_i 的均值;

σ_y^2 ——给定类别 y 下特征 x_i 的方差。

3.2.2 逻辑回归

逻辑回归是一种广义的线性回归分析模型,它将数据拟合到一个 logit 函数中,从而完成对事件发生概率的预测^[14]。对于现实生活中常见的分类问题,线性回归很难完成一个鲁棒性很好的分类器。逻辑回归在线性回归的基础上进行改进,使得输出结果映射为上的概率值,帮助判断结果。

在二分类问题中,通常选择函数作为基础函数,即:

$$\text{sigmoid}(z^{(i)}) = \frac{1}{1 + e^{-z^{(i)}}} \quad (10)$$

可表示为:

$$g(z) = \frac{1}{1 + e^{-z}} \quad (11)$$

式中: $g(z)$ ——Sigmoid函数的输出,它代表一个概率值,该值在 0 到 1 的范围内。

函数 $y = g(z)$ 在 $z=0$ 时为 1/2,随着 z 逐渐变小,函数值趋于 0, z 逐渐变大时函数值趋于 1。如果定义线性回归的预测函数为:

$$z^{(i)} = \omega^T x^{(i)} + b \quad (12)$$

式中: $z^{(i)}$ ——第 i 个样本的目标变量的预测值;

ω^T ——权重向量,包含了除偏置项外的所有模型参数;

$x^{(i)}$ ——第 i 个样本的特征向量;

b ——偏置项,也称为截距。

那么逻辑回归的预测函数为:

$$y^{(i)} = \text{sigmoid}(z^{(i)}) \quad (13)$$

式中: $y^{(i)}$ ——对于第 i 个样本的预测概率,即给定输入特征下目标变量 y 为 1 的概率。

得到预测值后,计算对应的损失函数:

$$L(a^{(i)}, y^{(i)}) = -y^{(i)} \log(a^{(i)}) \quad (14)$$

式中: $a^{(i)}$ ——第 i 个样本的实际标签,通常是 0 或者 1;

L ——对于第 i 个样本,预测值 $y^{(i)}$ 和实际标签 $a^{(i)}$ 之间的损失。

得到预测值后,计算对应的损失函数:

$$J = \frac{1}{m} \sum_{i=1}^m L(a^{(i)}, y^{(i)}) \quad (15)$$

式中: J ——成本函数,通常是所有单个样本损失的平均值;

m ——样本数量。

计算出代价函数后,选择优化方法(如梯度下降法)来最小化代价函数,最终得到合适的参数 ω 和 b 。

对于多分类问题中,通常选择 softmax 函数作为基础逻辑函数。

当给定一个特征向量 (x_1, x_2, \dots, x_n) 时,要计算出其属于类别 y_k 的概率,则需要将 k 个线性回归函数值压缩至区间 $[0, 1]$, 此时采取 softmax 将输入 $h_\theta(x)$ 表示为:

$$h_\theta(x) = [p(y = 1 | x, \theta), \dots, p(y = K | x, \theta)]^T = \frac{[\exp(\theta_1^T(x)), \dots, \exp(\theta_K^T(x))]^T}{\sum_{j=1}^K \exp(\theta_j^T(x))} \quad (16)$$

式中, $h_\theta(x)$ 表示给定输入 x 和模型参数 θ 时,样本属于每个类别 K 的概率, $\theta_j \in R_n + 1$ 表示每一个线性回归函数的参数,分母用来对概率分布进行归一化,使得概率之和等于 1。

4 入侵检测具体步骤及分析

4.1 数据选取

本文使用 KDD CUP 99 中网络入侵数据包 corrected 作为训练集, kddcup_data_10percent 作为测试集。kddcup_data_10percent 数据包是对 kddcup_data 数据包 10% 的抽样。

4.2 数据预处理

4.2.1 数据标准化

在进行数据分析之前,首先要对数据进行标准化处理,再利用处理后的数据进行分析。数据标准化将离散字符型数据预处理为数值型数据。本文选用了 0, 1, 2, 3, 4 分别代表 Normal、DOS、Probing、R2L、U2R

五个大类,将所有分类标识均用这 5 个数字代替。作为 Python 中强大的第三方库, sklearn 库(全称 scikit-learn)提供了很多用于数据标准化处理的模块函数。

4.2.2 数据归一化

数据归一化是机器学习的预处理阶段,由于训练集和测试集特征范围相差较大,取值范围不同,无法取得相同的离散区间,因此在进行离散化前需将部分数据进行归一化。本文利用过滤算法中的 unsupervised-instance-normalize 算法将标准化的每个数值归一化到 $[0, 1]$ 区间,公式为:

$$x^* = \frac{x - \min}{\max - \min} \quad (17)$$

式中: \max ——样本数据的最大值;

\min ——样本数据的最小值;

x ——标准化后的数据。

4.3 特征选择

利用信息熵增益构建数据集中的重要特征,通过比较提取特征前后检测的准确率,将特征对最终预测结果的影响程度排序,选取 10 种对结果影响最大的特征进行入侵检测。

```
% matplotlib inline
```

```
import matplotlib.pyplot as plt
```

```
kdd.hist(bins=50,figsize=(20,15))
```

plt.show() 用 hist 函数对特征值进行处理,可看出特征值的影响程度。

在划分训练集和测试集前,计算初始信息熵和基于不同特征的条件增益,两者相减得到基于不同特征的信息增益,将全部 41 个信息增益进行排序,选取信息增益大小位于前 10 的特征用于后续模型的训练和预测。

4.4 实验结果

为比较不同分类器及特征选择对预测结果的影响,验证特征选择后能够提高程序运行效率。本文在上述数据预处理及特征选择后,分别利用贝叶斯以及逻辑回归分类算法对提取所有 41 个特征和提取 10 个特征两种情况进行分类仿真。定义的 data_processingO 函数里有 1 个参数 all_features,默认为 True。表示提取所有 41 个特征进行训练;如果选择 all_features 为 False,则表示提取 10 个特征进行训练。采用 sklearn 库 model_selection 模块中 sklearn.naive-bayes. GaussianNB 与 sklearn.linearmodel.LogisticRegression 函数将数据集划分为两部分,一方面用于模型训练,另一方面用于模型预测,以验证模型的准确性。其中,训练集占 60%,测试集占 40%。

首先,对 5 种标签的准确率进行预测,结果如表 2

所示。

表2 分类准确率表(%)

算法	全部41个特征	提取10个特征
朴素贝叶斯	92.38	84.77
逻辑回归	99.82	99.26

由表2可知,无论是否进行特征选择,逻辑回归分类器的准确率都要高于朴素贝叶斯。在进行特征选择后,朴素贝叶斯分类的准确率有显著下降,而逻辑回归分类的准确率下降很少。

其次,对4种攻击的识别率进行了实验,仿真结果如表3所示。

表3 攻击种类识别率表(%)

算法	全部41个特征	提取10个特征
朴素贝叶斯	99.16	94.58
逻辑回归	99.83	98.85

由表3可知,朴素贝叶斯和逻辑回归分类器针对4种攻击的识别率都比标签高。尤其是朴素贝叶斯分类识别攻击类型的准确率较识别标签高出不少。在特征选择后,两个分类器对攻击类型的识别率有所下降,这与识别标签结果相同。

最后,特征选择的主要目的是为了在分析大样本、多特征数据时减少程序运行时间从而提高检测效率。因此,有必要对训练以及预测所用时间进行分析。在实验中,各个模型训练与预测时间情况,如表4所示。

表4 程序运行时间表(s)

算法	全部41个特征		提取10个特征	
	训练时间	预测时间	训练时间	预测时间
朴素贝叶斯	0.58	16.72	0.34	4.45
逻辑回归	228.75	0.70	21.49	0.44

由表4可知,在进行特征选择前,朴素贝叶斯模型运行时间主要用于预测,而逻辑回归模型运行时间主要用于训练。在进行特征选择选出10个特征后,朴素贝叶斯模型预测时间和逻辑回归模型训练时间均显著降低。仿真结果表明特征选择对提高模型的检测效率有显著作用。

5 结束语

本文基于朴素贝叶斯和逻辑回归算法,建立了网络入侵检测模型,并对入侵相关特征进行训练与预测。得到主要结论如下:逻辑回归分类的准确率和攻击识别率要比朴素贝叶斯模型更高;进行特征选择后,朴素贝叶斯分类的准确率显著下降,攻击识别率有所下降,逻辑回归分类的准确率和攻击识别率略有下降;进行特征选择后,两个模型的程序运行时间明显缩短。总

体研究表明,进行特征选择能够极大提高网络入侵检测效率,对于保障高速铁路信号系统网络安全具有重要意义。

参考文献:

[1] 余超,雷雳. 铁路移动终端安全管控方案探讨[J]. 高速铁路技术, 2022, 13(5): 10-13, 30.
YU Chao, LEI Li. Discussion on Safety Control Solution of Railway Mobile Terminals [J]. High Speed Railway Technology, 2022, 13(5): 10-13, 30.

[2] 赖英旭,刘增辉,蔡晓田,等. 工业控制系统入侵检测研究综述[J]. 通信学报, 2017, 38(2): 143-156.
LAI Yingxu, LIU Zenghui, CAI Xiaotian, et al. Research on Intrusion Detection of Industrial Control System [J]. Journal on Communications, 2017, 38(2): 143-156.

[3] 张蕾,崔勇,刘静,等. 机器学习在网络空间安全研究中的应用[J]. 计算机学报, 2018, 41(9): 1943-1975.
ZHANG Lei, CUI Yong, LIU Jing, et al. Application of Machine Learning in Cyberspace Security Research [J]. Chinese Journal of Computers, 2018, 41(9): 1943-1975.

[4] 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.
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.

[5] 杨印根,王忠洋. 基于深度神经网络的入侵检测技术[J]. 网络安全技术与应用, 2019(4): 37-41.
YANG Yingen, WANG Zhongyang. Intrusion Detection Technology Based on Deep Neural Network [J]. Network Security Technology & Application, 2019(4): 37-41.

[6] 解滨,李清扬,董新玉. 面向网络入侵检测数据的对抗样本生成方法[J]. 山东大学学报(理学版), 2021, 56(3): 28-36.
XIE Bin, LI Qingyang, DONG Xinyu. Adversarial Examples Generation Method for Network Intrusion Detection Data [J]. Journal of Shandong University (Natural Science), 2021, 56(3): 28-36.

[7] 王晓程,刘恩德,谢小权. 攻击分类研究与分布式网络入侵检测系统[J]. 计算机研究与发展, 2001, 38(6): 727-734.
WANG Xiaocheng, LIU Ende, XIE Xiaoquan. Attack Classification Research and a Distributed Network Intrusion Detection System [J]. Journal of Computer Research and Development, 2001, 38(6): 727-734.

[8] 丁龙斌,伍忠东,苏佳丽. 基于集成深度森林的入侵检测方法[J]. 计算机工程, 2020, 46(3): 144-150.
DING Longbin, WU Zhongdong, SU Jiali. Intrusion Detection Method Based on Ensemble Deep Forests [J]. Computer Engineering, 2020, 46(3): 144-150.

[9] 李勇,张波. 一种基于深度CNN的入侵检测算法[J]. 计算机应用与软件, 2020, 37(4): 324-328.
LI Yong, ZHANG Bo. An Intrusion Detection Algorithm Based on Deep Cnn [J]. Computer Applications and Software, 2020, 37(4): 324-328.

(4)导风屏障在1/4跨位置以及跨中均能够显著地减小来流风速的大小,可满足强风时列车运行风速的限值,高速列车可安全通过桥梁。

(5)在数值模拟和风洞试验的基础上,通过实桥测试的形式评估、检验导风屏障产品的挡风性能,实现该技术产品多手段评估的闭合。研究成果可为此类数值模拟及风洞试验结果的验证提供参考。

参考文献:

- [1] 郭薇薇,夏禾,张田. 桥梁风屏障的气动效应及其对高速列车运行安全的影响分析[J]. 工程力学, 2015, 32(8): 112-119, 128.
GUO Weiwei, XIA He, ZHANG Tian. Analysis on Aerodynamic Effects of Bridge Wind Barrier and Its Influence on Running Safety of a High-speed Train[J]. Engineering Mechanics, 2015, 32(8): 112-119, 128.
- [2] 沈广旭,金阿芳,闻腾腾. 高速铁路防风栅的挡风效果数值模拟研究[J]. 佳木斯大学学报(自然科学版), 2019, 37(4): 569-572.
SHEN Guangxu, JIN Afang, WEN Tengting. Numerical Simulation of Windscreens Effect of High Speed Railway Indiscreet[J]. Journal of Jiamusi University (Natural Science Edition), 2019, 37(4): 569-572.
- [3] 张田. 强风场中高速铁路桥梁列车运行安全分析及防风措施研究[D]. 北京: 北京交通大学, 2013.
ZHANG Tian. Study on Running Safety of Trains and Windproof Measures for High-speed Railway Bridges in Strong Wind Field[D]. Beijing: Beijing Jiaotong University, 2013.
- [4] 许自强,何德华,于卫东. 大风工况动车组运行速度限值研究[J]. 铁道机车车辆, 2016, 36(1): 39-43.
XU Ziqiang, HE Dehua, YU Weidong. Research on Limit Operational Speed of CRH Train under High-speed Wind Condition[J]. Railway Locomotive & Car, 2016, 36(1): 39-43.
- [5] 黄双林. 兰新高铁防风标准研究[J]. 铁道工程学报, 2019, 36(6): 14-17, 73.
HUANG Shuanglin. Research on the Wind Break Standard of Lanzhou-Urumqi High-speed Railway[J]. Journal of Railway Engineering Society, 2019, 36(6): 14-17, 73.
- [6] 潘新民,马秀清,徐洁. 兰新高铁挡风墙防风效果分析评估[J]. 干旱气象, 2019, 37(3): 496-499.
PAN Xinmin, MA Xiuqing, XU Jie. Analysis and Evaluation about Anti-wind Efficiency of Windbreak Experimental Section in Lan-Xin High Railway[J]. Journal of Arid Meteorology, 2019, 37(3): 496-499.
- [7] 闫宏凯,潘新民,叶文军. 南疆线风区铁路风特性与行车控制关键风速测点研究[J]. 机电传动, 2019(3): 64-68.
YAN Hongkai, PAN Xinmin, YE Wenjun. Study of Wind Characteristics and Key Monitoring Points for Traffic Control in Nanjiang Railway[J]. Electric Drive for Locomotives, 2019(3): 64-68.
- [8] 吕娜,刘伟,谢海清,等. 叶片式导风屏障挡风性能优化研究[J]. 高速铁路技术, 2022, 13(1): 78-82, 88.
LV Na, LIU Wei, XIE Haiqing, et al. Study on Optimization of Wind-proof Performance of Vane-type Wind Deflector[J]. High Speed Railway Technology, 2022, 13(1): 78-82, 88.
- [9] 代晓巍,李振兴,赵丽莉. 飞行器外测数据连续型野值的抽取剔除方法[J]. 电子设计工程, 2015, 23(12): 68-70.
DAI Xiaowei, LI Zhenxing, ZHAO Lili. Study on Continuous Outliers Eliminating Based on Picking up Method for Trajectory Measurement Data of Aircraft[J]. Electronic Design Engineering, 2015, 23(12): 68-70.
- [10] 陈红岩,胡非,曾庆存. 处理时间序列提高计算湍流通量的精度[J]. 气候与环境研究, 2000, 5(3): 304-311.
CHEN Hongyan, HU Fei, ZENG Qingcun. Dealing with Imperfect Data to Improve Estimation Precision of Turbulence Flux[J]. Climatic and Environmental Research, 2000, 5(3): 304-311.

(上接第71页)

- [10] 曹峰. 计算机联锁系统安全评估分析与研究[J]. 高速铁路技术, 2015, 6(4): 1-3.
CAO Feng. Analysis and Research on Safety Assessment of Computer Interlocking System[J]. High Speed Railway Technology, 2015, 6(4): 1-3.
- [11] 张玲,张建伟,桑永宣,等. 基于随机森林与人工免疫的入侵检测算法[J]. 计算机工程, 2020, 46(8): 146-152.
ZHANG Ling, ZHANG Jianwei, SANG Yongxuan, et al. Intrusion Detection Algorithm Based on Random Forest and Artificial Immunity[J]. Computer Engineering, 2020, 46(8): 146-152.
- [12] KDD Cup 1999 Data. Irvine, CA(USA), Information and Computer Science University of California, Irvine[EB/OL]. <http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html>, 2005-6-8.
- [13] 张全龙. 基于深度学习模型的网络入侵检测研究[D]. 天津: 天津理工大学, 2021.
ZHANG Quanlong. Research on Network Intrusion Detection Based on Deep Learning Model[D]. Tianjin: Tianjin University of Technology, 2021.
- [14] WANG Wei, HE Yongzhong, LIU Jiqiang, et al. Constructing Important Features from Massive Network Traffic for Lightweight Intrusion Detection[J]. IET Information Security, 2015, 9(6): 374-379.