一种基于少量异常标签的 SQL 注入攻击检测方法

赵 伟[1], 周颖杰[1], 李政辉[1], 杨 松[2], 吕建成[1]
(1. 四川大学计算机学院, 成都 610065; 2. 四川大学工业互联网研究院, 成都 610065)

摘 要: SQL 注入攻击通过入侵目标数据库实现对数据的窃取或破坏,危害性极大. SQL 注入攻击检测可帮助及时发现潜在的安全威胁,从而有利于数据库安全防护. 然而在智能交通系统中,由于其内部的复杂性和 SQL 注入攻击新变种的不断涌现,可供机器学习模型训练的异常标签样本往往较少,使得现有大多数 SQL 注入攻击检测方法容易存在模型过拟合和性能退化的问题. 针对上述问题,本文综合考虑智能交通系统和 SQL 注入攻击的特点,设计了一种基于比特编码的 SQL 注入攻击检测框架. 该框架无需预训练词嵌入模型和进行语法规则解析. 基于该框架,本文提出基于注意力机制的半监督 SQL 注入攻击检测模型(ASDM). 该模型首先通过重构数据样本,学习样本特征的中心趋势和离散程度等高层次特征,表达特征后验分布和特征偏离程度;接着将该高层次特征与数据编码特征融合,突出不同类别数据间的差异;最后引入注意力机制和残差网络构造检测器输出判定结果,以使模型能够根据重要程度对特征施加不同的关注力度,同时具有较强的泛化能力. 实验结果表明:本文方法在数据标签不平衡的情况下,相较于其他 SQL 注入攻击检测方法具有更优的检测性能;并能够检测未知 SQL 注入攻击.

关键词: SQL 注入攻击; 检测; 比特编码; 注意力机制
中图分类号: TP393 **文献标识码**: A DOI: 10.19907/j.0490-6756.2022.062001

A SQL injection attack detection method based on a few abnormal labels

ZHAO Wei[1], *ZHOU Ying-Jie*[1], *LI Zheng-Hui*[1], *YANG Song*[2], *LÜ Jian-Cheng*[1]
(1. College of Computer Science, Sichuan University, Chengdu 610065, China;
2. College of Industrial Internet Research, Sichuan University, Chengdu 610065, China)

Abstract: SQL injection attacks would cause significant harm because they can steal or destroy data by intruding target database. SQL injection attack detection can find out the potential security threat in time, and it is beneficial to the database security protection. However, in intelligent transportation system, due to its internal complexity and the emergence of new varieties of SQL injection attacks, the size of abnormal samples cannot meet the requirement of machine learning model training. This would carry a significant risk of model overfitting and performance degradation. In order to solve the problem, a SQL injection attack detection framework is designed based on bit coding, considering the characteristics of intelligent transportation systems and SQL injection attacks comprehensively. In the framework, pre-training word embedding model and parsing of grammatical rules are not needed. Then, a semi-supervised SQL injection attack detection model (ASDM) is proposed based on this framework, combined

收稿日期: 2022-01-08
基金项目: 国家自然科学基金(61801315)
作者简介: 赵伟(1994—), 男, 硕士研究生, 研究方向为网络异常检测. E-mail: 2019223049289@stu.scu.edu.cn
通讯作者: 周颖杰. E-mail: yjzhou@scu.edu.cn

with the attention mechanism. In the model, the samples are reconstructed to learn the high-level features(such as the central trend and the dispersion degree of the features) and to express the feature posterior distribution and feature deviation. Then, these high-level features are fused with the data coding features to highlight the differences between different types of data. Finally, the attention mechanism and residual network are introduced to construct the detector, with the aim of exerting different attention intensity to the features according to their importance degree and guaranteeing the generalization ability of the model. The experimental results show that the proposed method has better detection performance compared with other SQL injection attack detection methods for the data with unbalanced labels, and can detect unknown SQL injection attacks.

Keywords: SQL injection; Detection; Bit encoding; Attention mechanism

1 引　言

SQL 注入攻击利用应用中存在的漏洞,欺骗其执行含恶意代码的 SQL 命令,达到窃取、破坏数据的目的,危害性极大[1,2]. 目前,智能交通系统以网站和移动应用的形式为人们服务,其安全面临日益严峻的挑战[3-5]. SQL 注入攻击是该系统中最常见且危害最大的攻击类型之一[6];它具有变种多、隐蔽性强和易突变等特点[7],且安全风险评估在 WEB 应用攻击中排名前十[8].

SQL 注入攻击检测方法能够降低应用遭受此类攻击的风险. 近年来,智能交通领域中 SQL 注入攻击方式越来越复杂,但现有的 SQL 注入攻击检测方法无法应对这种情况. 传统的 SQL 注入攻击检测方法,如静态分析、动态分析等,基于源码、语法解析,仅对部分攻击类型有效[9,10]. 基于浅层机器学习的 SQL 注入攻击检测方法,如支持向量机[11]、K 近邻[12]等,通过人工构建的 SQL 样本属性特征训练模型,检测结果依赖于特征工程,容易误报和误检.

目前,基于深度学习[13-15] SQL 注入攻击检测方法的研究,受到人们广泛关注,但仍存在许多问题和挑战. 在智能交通系统中,由于其内部的复杂性和 SQL 注入易突变等特点,导致含标签的 SQL 注入攻击样本少且标注困难. 现有的大多数 SQL 注入攻击检测方法在这种环境中容易存在模型过拟合,检测性能下降的问题[6]. 针对该问题,Li 等[6]提出一种数据增强方法,通过现有不同类型的 SQL 注入攻击样本的组合或融合生成新样本,增加已知类型 SQL 注入攻击样本的数量,达到缓解模型过拟合的目的. 这类方法依赖于先验知识的积累和生成的样本质量,难以解决先验知识缺乏的新型 SQL 注入攻击类型的样本标签少的问题. 同时,现有基于深度学习的 SQL 注入攻击检测方法大多需要进行 SQL 语法解析或预训练词嵌入(Word2Vector)[13]编码模型. 但这种方式不但会隐藏一些 SQL 注入攻击类型的特点,而且可能引入噪声,增加后续检测的难度.

为了解决上述问题,本文首先设计一种基于比特编码的 SQL 注入检测框架. 该框架解决了现有框架存在的上述不足. 然后基于该框架提出一种基于少量异常样本标签的半监督 SQL 注入攻击检测模型. 该模型包含数据分布学习和异常检测两个模块;其中数据分布学习模块通过重构原始样本,为异常检测模块提供高层次的样本分布特征,增加类别间样本特征的差异,从而降低检测难度;异常检测模块引入注意力机制[16,17]和残差机制[18],加强特征关注的同时保持较强的泛化能力[19],使得该模型对新型 SQL 注入攻击具有较好的检测性能.

2 相关工作

SQL 注入检测方法主要分为三种:传统检测算法、机器学习检测算法和深度学习检测算法. 传统检测算法基于源码、语法解析,仅对部分攻击类型有效. 传统检测算法中,Gould 等[9]提出一种基于静态分析的 JDBC 源码分析器,该分析器只能检测出不符合输入类型的 SQL 注入攻击. 王溢等[10]提出一种基于动态分析的方法,通过 SQL 语法解析构建语法树,进而对语法树进行检测,然而这种方式会隐藏一些 SQL 注入攻击的关键特征,致使检测性能下降.

浅层机器学习算法主要通过人工提取属性特征,训练模型进行检测,检测结果依赖于人工选取的特征. SVM[11]、KNN[12]从恶意代码中提取 n 维特征训练模型,然而这种方法的检测率并不高;Komiya 等[20]通过空白分割和令牌验证的方式提

取特征，并使用集成算法进行检测，增强了模型的泛化能力，但空白分割的特征提取方式难以应对 SQL 注入攻击的复杂情况.

深度学习因其强大的特征学习能力受到关注，越来越多的研究者尝试使用深度学习方法解决 SQL 注入攻击问题. 卷积神经网络[13,21]模型直接使用 Payload 负载的编码特征作为输入，解决了传统机器学习维度灾难的问题. Xie 等[22]提出能够处理变长序列的弹性池化卷积神经网络，解决了传统卷积神经网络难以处理不定长序列的问题. 但是这些方法无法应对智能交通系统的复杂环境，尤其是数据标签不平衡. Li 等[6]提出一种数据增强的方法，通过扩充 SQL 注入攻击样本，达到缓解模型过拟合的目的. 然而该方法无法应对新的 SQL 注入攻击类型，且依赖先验知识和生成样本的质量. 深度学习方法大都采用词嵌入对 SQL 样本编码，但这种编码方式存在以下不足：(1) 需要预先扫描已知数据构建词库，只能对词库中的词编码；(2) 需要预训练词嵌入模型，无法与检测模型端到端训练，可能引入噪声，给后续检测带来困难. 本文提出的方法解决了以上方法存在的不足，该方法基于比特编码策略，利用半监督学习端到端的对 SQL 注入攻击进行检测.

3　框架设计

为了解决传统 SQL 注入攻击检测框架采用语法解析、词嵌入编码可能隐藏 SQL 注入关键特征并引入噪声等问题，本文提出一种基于比特编码的检测框架. 该框架采用基于位的比特编码策略降低了需编码字符的数量，同时无需预训练词编码模型；包括数据采集、数据预处理、模型训练和 SQL 注入检测四个模块，如图 1 所示. 该框架首先通过数据采集和数据预处理模块构建用于模型训练和测试的高质量数据集；然后通过模型训练模块获取性能最优的模型；最后使用该模型对测试样本检测.

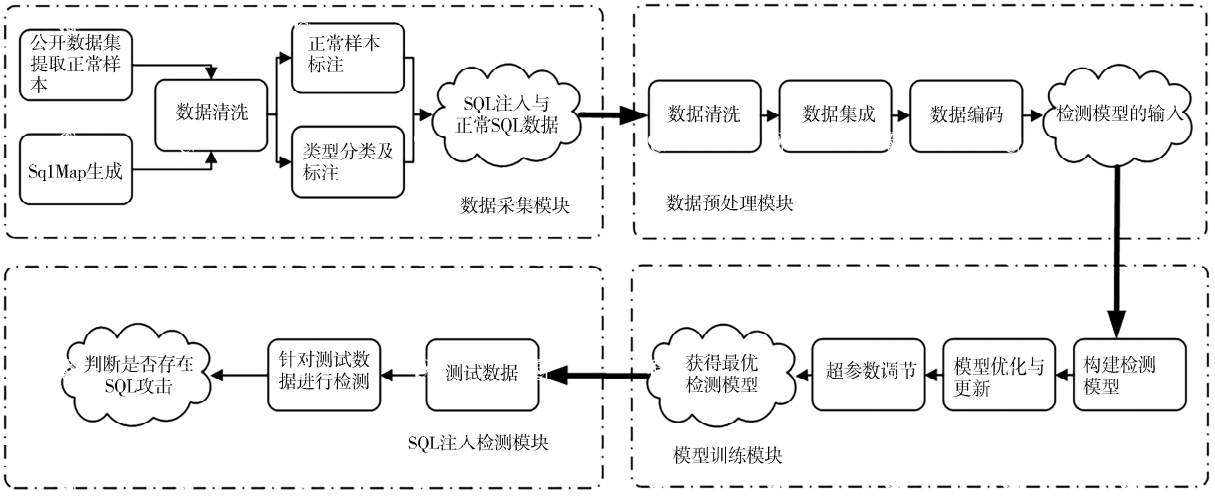

图 1　基于比特编码的 SQL 注入检测框架

Fig. 1　SQL injection detection framework based on bit coding

3.1　数据采集

数据采集模块是 SQL 注入攻击检测框架的基础，它通过不同的方式和途径采集正常、异常样本，构建用于模型训练的原始数据集. 其中正常样本采用从公开数据集中筛选的方式获取，本文采用耶鲁大学公布的 Spider[23]数据集；异常样本采用自动注入工具攻击 WEB 应用的方式生成，本文使用的自动注入工具和 WEB 应用分别为 Sqlmap[6]、DVWA[24]，并通过修改 Sqlmap 脚本 tamper 的方式丰富 SQL 注入攻击类型.

3.2　数据预处理

通过数据采集模块获得的原始样本中，可能存在一些对 SQL 注入攻击检测模型构建有干扰的因素，并不能直接用于后续步骤输入，因此需要对他们进行必要的预处理.

数据清洗包括对原始样本中可能存在冗余干扰的处理和某些特殊字段的处理两方面内容. 原始样本中可能存在冗余干扰的处理包括筛选高度相似或重复的样本、甄别和去除标注为异常的原始样本中可能包含的个别正常样本. 原始样本中特定字段的处理包括关键词大小写统一，表名、查询内容

等特定字段修正以及删除异常样本中包含的网站账号信息.

由于深度学习模型无法接收文本数据作为输入,因此需要对 SQL 样本进行编码.本文考虑到词嵌入编码存在的弊端,采用比特编码策略.该编码策略以位为单位,降低了需编码字符的数量且无需扫描已知样本构建词库用于预训练编码模型,能够客观地描述样本属性特征.具体编码方式如下:首先按 ASCII-256 将样本转换为十进制数向量,然后使用 8 位二进制数依次表示该向量中的十进制数;每条样本与文献[25]相同的方式选取蕴含了绝大部分关键信息的前 N 位字符进行编码;若样本长度不足 N 位,则使用空格填充的方式补齐,否则截断.在本工作中,N 设置为 256.

3.3　模型训练与 SQL 注入攻击检测

本框架在模型训练模块构建、优化获得基于注意力机制的半监督 SQL 注入攻击检测模型,并将该模型用于检测未知 SQL 样本是否含注入攻击.模型训练模块首先根据样本标签不平衡场景及 SQL 注入攻击特点,构建出能够获得数据分布信息和有效关注重要特征的半监督检测模型,并通过模型优化、参数调节等方式更新模型,以获得最优的检测模型.SQL 注入攻击检测模块,利用模型训练模块输出的模型对未知 SQL 样本检测,判断是否包含 SQL 注入攻击.

4　基于注意力机制的半监督检测模型

为了解决现有 SQL 注入攻击检测方法在数据标签不平衡的场景中性能下降的问题,本文提出一种基于注意力机制的半监督检测模型,如图 2 所示.该模型提出使用样本特征中心趋势和离散程度等高层次特征与样本比特编码特征融合,突出不同类别样本之间的差异;同时为了实现特征的有效关注并保持模型的泛化能力,在检测器模块中引入注意力机制[16, 17]和残差机制[18].下面介绍每个模块的具体细节.

4.1　数据分布学习模块

数据分布学习模块使用基于数据分布学习的模型为检测器模块提供样本特征中心趋势和离散程度等高层次特征.本文方法的数据分布学习模块使用“Auxiliary deep generative models”中提出的辅助深度生成模型[26]实现,该模型通过引入辅助变量 a 增强了传统变分自编码器表征数据分布的能力,包含推理器和生成器两部分组成.其中推理器定义如下:

$$I(a_i|x_i)=N(a_i|u_i(x_i),\sigma_i^2(x_i)) \tag{1}$$

$$I(y_i)=\mathrm{cat}(y|\pi) \tag{2}$$

$$I(z_i|a_i,y_i,x_i)=N(z_i|u_i(a_i,x_i,y_i),\ \sigma_i^2(a_i,y_i,x_i)) \tag{3}$$

其中 $u_i,\sigma_i,x_i,a_iy_i,z_i$ 分别表示第 i 个样本的平均值、标准差、样本编码特征、辅助向量、样本标签及样本后验分布;cat(·)表示多维数据分布,本方法中对于有标签样本即为其标签的独热编码,而对于无标签样本则表示所有可能类别标签的独热编码;N 表示高斯采样;I 表示推理器.生成器定义如下.

$$G(y_i)=\mathrm{cat}(y|\pi) \tag{4}$$

$$G(a_i|z_i,y_i,x_i)=f(a_i;z_i,y_i,x_i,w) \tag{5}$$

$$G(x_i|z_i,y_i)=f(x_i;z_i,y_i,w) \tag{6}$$

其中,f 表示多层非线性变换;权重为 w;G 表示生成器.

4.2　检测器模块

本文设计了一种增强的检测器,通过关注重点特征和特征细节弥补提升效果.该检测器在传统多层感知机构建的检测器基础上引入了注意力机制和残差机制,既能对特征实施有效的关注,又保证了检测器的泛化能力.该模块融合数据分布学习模块提供的高层次特征和样本比特编码特征作为输入,通过特征提取和注意力子模块实现对特征的有效提炼和关注,进而判断样本是否包含 SQL 注入攻击.检测模块包含特征提取模块和注意力模块两部分,具体细节如下.

特征提取模块由多层非线性变换构成,可以表示为

$$f_1=\mathrm{relu}(W_1X+b_1) \tag{7}$$

$$f_i=\mathrm{relu}(W_if_{i-1}+b_i) \tag{8}$$

其中 X 为样本原始编码特征及样本特征中心趋势和离散程度等高维特征的融合;W_1、W_i 表示第 1、i 层非线性变换的权重;b_1、b_2 表示第 1、i 层非线性变换的偏置;f_1、f_{i-1}、f_i 表示第 1、$i-1$、i 层线性变换;激活函数使用 relu.特征提取模块被设计为梯形结构,实现样本特征的提取及简单去噪,尽可能保留相对重要的特征.

为了提升检测器对与任务目标联系密切特征的关注力度并保持模型泛化能力,在注意力模块引入注意力机制的同时引入残差机制.这是因为注意力机制强大的特征关注能力会使得极少量特征获得较大的权重,造成整体特征细节的丢失,使得检

测性能下降，而残差机制可以对丢失的特征细节进行补充. 注意力模块可以被定义为

$$g_i = \mathrm{SoftMax}(Wg_i + b) \tag{9}$$

$$g_k = g_i \cdot g_w \tag{10}$$

$$F = \mathrm{Concat}(g_i, g_k) \tag{11}$$

其中 g_i、g_w、g_k、F 分别代表隐层特征、特征权重、加权特征及融合特征；W 为线性变换层的权重，Concat(·)表示特征拼接.

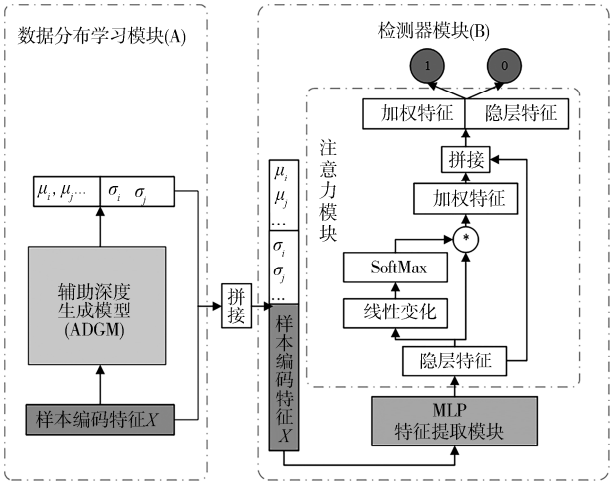

图 2　注意力半监督检测模型

Fig. 2　Attention semi-supervised detection model

5　实　验

5.1　数据集

本文基于真实 SQL 语句构建了数据集. 该数据集共含 5200 余条样本，其中正常样本 3600 余条，异常样本 1600 余条. 异常样本按文献[6]划分得到 14 种 SQL 注入攻击类型，如表 1 所示. 由于数据集的限制，本文中所有的实验均使用三折交叉验证. 为模拟智能交通系统中数据标签不平衡的场景，首先取三等份数据中的一份作为测试集，其余两份作为训练集；然后将训练集划分为有标签和无标签两部分. 具体划分如下：首先按照攻击类型随机采样分别选取 5～10 条异常样本，共计 115 条构成有标签异常样本部分；然后从正常样本中随机采样 8 倍于该数量的正常样本作为有标签正常样本部分，共计 920 条；最后剩余训练集中的样本构成无标签样本部分，如表 2 所示. 由于评估指标准确率易受标签不平衡的影响，因此对测试集中异常样本进行了上采样操作，使得测试集中正常样本和异常样本数量基本持平.

表 1　异常样本类型及样例[5]

Tab. 1　Abnormal sample type and sample[5]

类型	分类标准	样例
Blank	多空格分割	1　=1 等
CN	条件函数	Ifnull 等
CON	条件判断	=/>/</in 等
DC	常用异常字符	1=1/'/'/--等
IG	逻辑连词	&&/and/||/or
NE	不等式判断	1=2/'a'>'b'等
OC	SQL 函数名	Selectload_file()等
OP	四则计算	1+2<5 等
SP	存储过程	Declare @qnvarchar (200) select @q =等
CONN	拼接	OR+6115=6115 等
SC	SQL 动词关键字	Union all 等
TI	时延	'sleep 50'等
WAF	编码	Unicode 编码等
KS	关键字分割	A/**/ND 等

表 2　数据集

Tab. 2　Dataset

数据集	样本类型	数量
训练集	有标签正常样本	115
	有标签异常样本	920
	无标签样本	3000
测试集	正常样本	560
	异常样本	558

5.2　评价指标

在实验中采用准确率(Accuracy，ACC)、最高性能差[27](Highest Difference，HD)、检测率(Detection Rate，DR)、误报率(False Alarm Rate，FAR)作为模型检测性能评价指标. 定义如下.

$$DR = \frac{TP}{TP + FN} \cdot 100\% \tag{12}$$

$$FAR = \frac{FP}{FP + TN} \cdot 100\% \tag{13}$$

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \cdot 100\% \tag{14}$$

$$HD = DR - FAR \tag{15}$$

其中，TP 表示异常样本中被检测为异常的样本数；TN 表示正常样本中被检测为正常的样本数；FP 表示正常样本中被检测为攻击的样本数；FN 表示异常样本中被检测为正常的样本数.

表 3　SQL 注入攻击检测实验结果

Tab. 3　Experimental results of SQL injection attack detection

实验指标	LSTM/%	RNN/%	CNN/%	KNN/%	RF/%	SVM/%	LSTM(上采样)/%	LSTM(数据增强)/%	ADGM/%	本文方法/%
ACC	91.62	88.57	82.7	77.95	81.74	80.33	87.43	95.67	96.30	97.40
DR	83.49	76.93	64.44	98.83	100	100	85.66	93.05	98.32	98.00
FAR	0.50	1.08	1.08	45.51	38.80	41.77	4.30	2.00	5.5	3.08
HD	82.99	75.85	63.36	53.32	61.20	58.23	81.36	91.05	92.82	94.92

5.3　对比方法

为了评估本文方法的性能,本文选择长短时记忆人工神经网络(LSTM)[28]、循环神经网络(RNN)[25]、卷积神经网络(CNN)[13]、K 近邻算法(KNN)[12]、随机森林算法(RF)[29]、支持向量机(SVM)[11]、基于上采样的 LSTM、基于数据增强的 LSTM[6]、辅助深度生成模型(ADGM)[26]这 9 种方法作为对比方法. 其中 ADGM 可用于半监督的异常检测;基于上采样和数据增强的 LSTM 算法常用于数据标签不平衡的 SQL 注入攻击检测;而其他 6 种方法常用于数据平衡的 SQL 注入攻击检测. 对比方法与本文方法输入一致.

5.4　实验结果及分析

5.4.1　对比实验　为了验证本文方法在数据标签不平衡场景中的有效性,将该方法与其他 9 种方法进行了实验对比. 本文方法采用多层感知机搭建完成. 其中数据分布学习模块引用 ADGM[26]中的结构、参数设置;而检测器模块中特征提取子模块由两层神经元个数分别为 512 和 32 的多层感知机构建完成,注意力模块不改变输入特征的维度,最后输出判决结果. 训练过程中,采用 Adam 优化器框架对网络参数更新学习,学习率为 0.001.

表 3 展示了本文方法与其他方法的实验结果,本文方法相较于 ADGM、基于数据增强的 LSTM 整体性能 HD 分别提升 2.1%、3.87%,显著优于其他方法. ADGM 方法与本文方法检测率基本相同,但误报率更高(2.42%). LSTM 和基于数据增强的 LSTM 相较于本文方法,拥有更低的误报(0.5%和 2%);但 LSTM 方法的检测能力较差(83.49%),基于数据增强的 LSTM 方法检测率相较于本文方法低 4.95%. 其他 5 种数据标签平衡时常用的检测方法,在数据不平衡时,效果较差(最优整体性能 HD 仅为 75.85%). 综合所有方法的实验结果分析,可以得出本文方法明显优于其他方法.

5.4.2　消融实验　本文通过引入样本特征中心趋势、离散程度等高层次信息以及注意力子模块两方面提升模型的检测效果,为验证上述两部分因素分别对模型性能的影响,分别进行了消融实验研究.

为评估引入样本特征中心趋势、离散程度信息对模型性能的影响,将本文方法(ASDM)与未引入样本特征后验分布信息的方法(ASDM-Z)对模型性能的影响进行了比较. 实验结果如图 3 所示. 使用样本特征中心趋势和离散程度等高层次特征能够使模型性能提升. ADSM 模型相较于 ASDM-Z 最高性能差(HD)、准确率分别提升 1.02%、0.62%,误报率降低了 1.34%. 本文方法针对无标签样本既考虑其为正常的可能,也考虑其为异常的可能,造成模型学习获得的样本特征中心趋势、离散程度所表征的分布信息与实际的分布信息有些差异;但是影响是比较小的,从实验结果中也可看出,检测率仅降低 0.32%,几近持平.

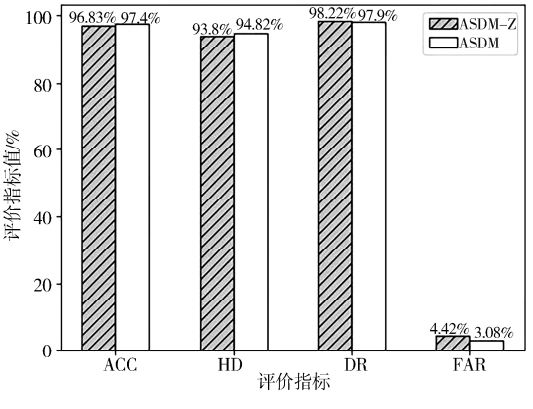

图 3　第一组消融实验结果

Fig. 3　Results of ablation experiment 1

为评估注意力子模块对模型性能影响,将本文方法(ASDM)与不使用注意力子模块的方法(ASDM No AM)对比. 实验结果如图 4 所示. 注意力模型通过注意力机制加强模型对于重要特征的关注,降低次要甚至负面特征的影响;同时是使用残差机制弥补必要的特征细节,提升模型的泛化能力;使得模型性能明显提升. ADSM 模型相较于

ASDM-No AM 最高性能差(HD)、准确率分别提升 0.98%、0.35%，误报率降低 0.92%，检测率几近相同.

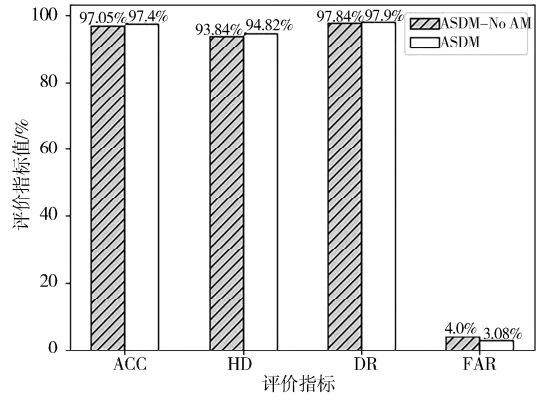

图 4　第二组消融实验结果

Fig. 4　Results of ablation experiment 2

通过以上消融研究，验证了本文方法的有效性. 这是因为样本特征中心趋势、离散程度信息不但能够表达特征的后验分布，还能表达特征偏离程度等信息，而注意力模块加强了模型对于特征的关注能力. 因此本文提出的方法能够有效检测 SQL 注入攻击.

5.4.3　*数据效率实验*　由于新型 SQL 注入攻击难以标记，因此数据效率是评估半监督方法的重要因素. 由于异常样本包含多种类型，若使用简单随机下采样的方式获取有标签异常样本可能造成部分类型有标签样本较多，部分类型有标签较少，使得模型对不同类型的异常样本检测存在偏差. 因此我们对每种类型异常样本等量随机下采样，每组实验选取 3、5、10、15 条样本，对本文方法的数据效率进行实验验证. 为避免实验结果是由于随机采样得到的偶然结果，我们对每组实验多次等量随机下采样，计算每次实验结果的平均作为最终结果.

本文方法性能的总体趋势随着异常样本标记数量的增加而增加，但并未发生剧变，性能稳定. 实验结果如图 5 所示. 随着可标注的异常样本的增加，本文方法整体性能(HD、ACC)由 94.05%、96.96%提升至 95.72%、97.8%，误报率由 4.7%降至 2.5%. 检测率的结果与总体趋势相反，这是因为不同的异常之间分布不同，还可能存在相互冲突的信息，随着标记异常样本数量的增加，这种趋势会被增大，因此检测性能降低.

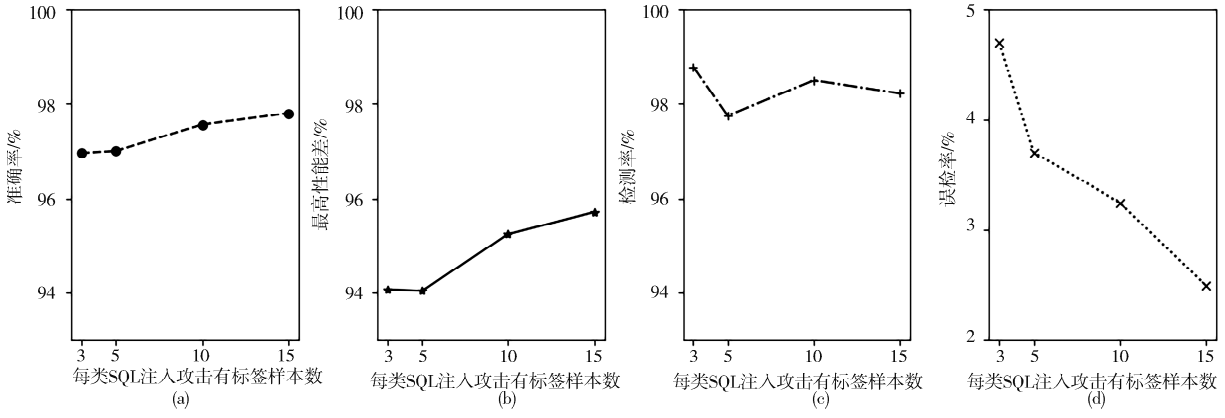

图 5　异常样本数量对模型性能的影响分析

Fig. 5　Analysis of the influence of the number of abnormal samples on the performance of the model

5.4.4　*新型 SQL 注入攻击检测*　为了验证本文方法对新型 SQL 注入攻击的检测能力，并保障验证结果并非偶然原因造成，进行了三组新型 SQL 注入攻击检测实验. 本实验使用的数据设置与对比试验中的数据设置不同之处在于：本实验训练集中有标签异常样本由对比实验训练集的有标签异常样本随机选取 7～8 种攻击类型的样本组成(大约 45 至 80 条).

实验结果如表 4 所示，可以看出当有未被标注出的新类型 SQL 注入样本出现时，本文方法依然具有良好的检测能力：检测率达到 93%以上，误报率在 5%以下. 本文方法之所以能够对新型 SQL 注入攻击依然有较优的检测效果，原因有以下两点：一方面是因为引入样本特征中心趋势和离散程度等高层次特征，增加了异常与正常样本之间的差异；另一方面是因为注意力模块中引入残差机制，增强了模型的泛化能力.

表 4　新攻击类型检测验证实验结果

Tab. 4　Experimental results of new attack type detection and verification

实验指标	新攻击类型验证实验 1/%	新攻击类型验证实验 2/%	新攻击类型验证实验 3/%
ACC	96.05	95.66	95.13
DR	96.94	95.37	93.06
FAR	4.75	4.08	3
HD	92.19	91.29	90.06

6　结　论

本文通过对智能交通系统和 SQL 注入攻击特点的分析,从编码策略和数据标签不平衡两方面考虑,使用比特编码策略解决传统编码存在的不足,同时提出一种基于注意力的半监督检测方法. 在真实 SQL 语句构建的数据集上进行了一系列的实验,证实了本文方法在数据标签不平衡时具有优异的检测效果,而且对于新型 SQL 注入攻击也有一定的检测能力.

参考文献:

[1] 黄琨茗, 张磊, 赵奎, 等. 基于最长频繁序列挖掘的恶意代 码检测[J]. 四川大学学报: 自然科学版, 2020, 57: 681.

[2] 刘繁艳, 姜瑜. SQL 注入攻击研究[J]. 中国科技信息, 2005, 17A: 1.

[3] Sui P, Li X, Bai Y. A study of enhancing privacy for intelligent transportation systems: k-correlation privacy model against moving preference attacks for location trajectory data [J]. IEEE Access, 2017, 5: 24555.

[4] Alwan Z S, Younis M F. Detection and prevention of SQL injection attack: a survey [J]. Int J Comput Sci Mob Comput, 2017, 6: 5.

[5] 陈兴蜀, 江天宇, 曾雪梅. 基于多维时间序列分析的网络异常检测[J]. 工程科学与技术, 2017, 49: 7.

[6] Li Q, Wang F, Wang J. LSTM-based SQL injection detection method for intelligent transportation system [J]. IEEE T Veh Technol, 2019, 68: 4182.

[7] 代威, 黄金杰, 刘畅. 典型内部网络 SQL 注入攻击与防范 [J]. 网络安全技术与应用, 2020, 2: 19.

[8] Wichers D, Williams J. Owasp top-10 2017[EB/OL]. [2021-12-21]. https://owasp. org/www-project-top-ten/owasp_Top_Ten_2017/.

[9] Gould C, Su Z, Devanbu P. JDBC checker: a static analysis tool for SQL/JDBC applications [C]//Proceeding of the International Conference on Software Engineering. [S. l.]: IEEE, 2004.

[10] 王溢, 李舟军, 郭涛. 防御代码注入式攻击的字面值污 染方法[J]. 计算机研究与发展, 2012, 49: 2414.

[11] Choi J, Kim H, Choi C. Efficient malicious code detection using n-gram analysis and SVM [C]//Proceeding of the 2011 14th International Conference on Network-Based Information Systems. [S. l.]: IEEE, 2011.

[12] Hashem I, Islam M, Haque S M. A proposed technique for simultaneously detecting DDoS and SQL injection attacks [J]. Int J Comput Appl, 2021, 183: 50.

[13] Falor A, Hirani M, Vedant H. A deep learning approach for detection of SQL injection attacks using convolutional neural networks [M]//Proceedings of the Data Analytics and Management. [S. l.]: IEEE, 2022.

[14] Feng F, Liu X, Yong B. Anomaly detection in ad-hoc networks based on deep learning model: a plug and play device [J]. Ad Hoc Netw, 2019, 84: 82.

[15] Wang J, Huang M, Nie Y, *et al*. Static analysis of source code vulnerability using machine learning techniques: a survey [C]//Proceeding of the 2021 4th International Conference on Artificial Intelligence and Big Data. [S. l.]: IEEE, 2021.

[16] Vaswani A, Shazeer N, Parmar N. Attention is all you need [C]//Proceeding of the Advances in neural information processing systems. [S. l.]: IEEE, 2017.

[17] 晋儒龙, 卿粼波, 文虹茜. 基于注意力机制多尺度网络的自然场景情绪识别[J]. 四川大学学报: 自然科学版, 2022, 59: 012003.

[18] Yu T, Zhang R, Yang K. Spider: a large-scale human-labeled dataset for complex and cross-domain semantic parsing and text-to-SQL task [C]//Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing. [S. l.]: Association for Computational Linguistics, 2018.

[19] 邓描, 刘强, 陈洪刚, 等. 一种基于特征正则约束的异常检测方法[J]. 四川大学学报: 自然科学版, 2020, 57: 1077.

[20] Komiya R, Paik I, Hisada M. Classification of malicious web code by machine learning [C]//Proceeding of the 2011 3rdInternational Conference on

Awareness Science and Technology. [S. l.]: IEEE, 2011.

[21] Falor A, Hirani M, Vedant H. A deep learning approach for detection of SQL injection attacks using convolutional neural networks [M]//Proceedings of the Data Analytics and Management. [S. l.]: IEEE, 2022.

[22] Xie X, Ren C, Fu Y. SQL injection detection for web applications based on elastic-pooling cnn [J]. IEEE Access, 2019, 7: 151475.

[23] Yu T, Zhang R, Yang K. Spider: a large-scale human-labeled dataset for complex and cross-domain semantic parsing and text-to-SQL task [C]//Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing. [S. l.]: IEEE, 2018.

[24] Durai K N, Subha R, Haldorai A. A novel method to detect and prevent SQLIA using ontology to cloud web security [J]. Wireless Pers Commun, 2021, 117: 2995.

[25] Liu H, Lang B, Liu M, *et al*. CNN and RNN based payload classification methods for attack detection [J]. Knowl-Based Syst, 2019, 163: 332.

[26] Maaløe L, Sønderby C K, Sønderby S K. Auxiliary deep generative models [C]//Proceedings of the International conference on machine learning. [S. l.]: IEEE, 2016.

[27] Jokar P, Arianpoo N, Leung V. Electricity theft detection in AMI using customers' consumption patterns [J]. IEEE T Smart Grid, 2017, 7: 216.

[28] Feng F, Liu X, Yong B. Anomaly detection in ad-hoc networks based on deep learning model: a plug and play device [J]. Ad Hoc Netw, 2019, 84: 82.

[29] Aggarwal P, Kumar A, Michael K. Random decision forest approach for mitigating SQL injection attacks [C]// Proceedings of the 2021 IEEE International Conference on Electronics, Computing and Communication Technologies. [S. l.]: IEEE, 2021.

引用本文格式：

中　文：赵伟，周颖杰，李政辉，等. 一种基于少量异常标签的 SQL 注入攻击检测方法[J]. 四川大学学报：自然科学版，2022，59：062001.

英　文：Zhao W, Zhou Y J, Li Z H, *et al*. A SQL injection attack detection method based on a few abnormal labels [J]. J Sichuan Univ: Nat Sci Ed, 2022, 59: 062001.