文章编号：2095-2163(2023)01-0050-06 中图分类号：TP393 文献标志码：A

基于 FastGRNN 模型的列车通信网络入侵检测系统

方一帆，曾培峰
（东华大学 计算机科学与技术学院，上海 201620）

摘 要：本文提出了一种基于 FastGRNN(Fast Gated Recurrent Neural Network)模型，适用于列车通信网络的网络入侵检测系统。实验证明：相比其他分类模型，FastGRNN 拥有更高的分类精度，以及更低的系统占用。凭借 FastGRNN 的优良特性，该系统能在列车各车厢现存设备上部署网络入侵检测程序，在这些设备完成自己本职工作的同时，将其转化为入侵检测节点，相比于集中式的入侵检测系统，该系统拥有更高的效率和更低的部署成本。

关键词：FastGRNN 模型；循环神经网络；网络入侵检测；列车通信网络

Intrusion detection system for train communication network based on FastGRNN model

FANG Yifan, ZENG Peifeng
(School of Computer Science and Technology, Donghua University, Shanghai 201620, China)

【Abstract】 In this article, a network intrusion detection system based on FastGRNN (Fast Gated Recurrent Neural Network) model suitable for training communication networks has been proposed. Experiments show that FastGRNN has higher classification accuracy and lower system occupation than other classification models. With the excellent characteristics of the FastGRNN model, the system can deploy network intrusion detection programs on the existing equipment in each carriage of the train, and converts them into intrusion detection nodes while these types of equipment are doing their own jobs. Compared with the centralized intrusion detection system, the system has higher efficiency and lower deployment cost.

【Key words】 FastGRNN model; recurrent neural network; network intrusion detection; train communication network

0 引 言

目前，地铁及高铁等轨道交通系统正逐渐摒弃过去的 MVB、TCN 等总线技术，转而使用通信效率更高、成本更低的以太网技术[1]。在提高乘客乘坐体验的同时，以太网技术的开放性也使得轨道交通网络系统更易受到信号干扰和恶意攻击的侵害[2]。为了保护乘客的生命及财产安全，列车通信网络系统需要建立起一套完整而严格的网络安全系统。在成熟的网络安全系统中，入侵检测系统(Intrusion Detection System，IDS)往往扮演着重要角色。不同于网络防火墙，IDS 对来自系统内部和外部的流量都进行监控。当检测到异常流量时，IDS 通常会对上级系统和管理员发出警告[3]，通过一定配置，IDS 也可以直接禁止异常流量的通过，甚至对不同的异常种类采取不同的保护措施，这种 IDS 被称为入侵防护系统(Intrusion Prevention System，IPS)。

入侵检测系统主要分为基于标志的 IDS 和基于异常的 IDS[3]，前者的检测方式为识别数据包中的恶意模式，需要建立起足够庞大且可靠的模式库；后者则通过评价数据流与正常数据流的偏离程度来进行检测，这是一种典型的分类问题，因此基于异常的 IDS 通常使用机器学习方法。基于异常的 IDS 相较于基于标志 IDS 的优点是可以检测未知模式的攻击，且省去了依靠专家知识、时常需要更新的模式库；缺点是整个系统的精度很大程度上依赖训练数据集的可靠性，而这些数据集同样存在过时的风险。

近年来，以人工神经网络为主的机器学习方法正展现出其优越性。2011 年 M.A.Salama 等人[4]提出了基于深度信念网络和支持向量机(Support Vector Machine，SVM)的混合式入侵检测方法。该方法通过在 NSL-KDD 数据集上训练和测试，取得

作者简介：方一帆(1997-)，男，硕士研究生，主要研究方向：嵌入式、网络安全；曾培峰(1964-)，男，博士，教授，主要研究方向：图像处理、嵌入式、纤维的图像识别等。

收稿日期：2022-03-31

了最高 92.84%的准确率。2012 年 J.Mar 等人[5]提出了适应式模糊神经推断系统，该系统有效降低了平均检测时延，提升了拒绝服务攻击（Denial of Service，DoS）的检测精度。2017 年 J.Kim 等人[6]提出了基于深度神经网络的入侵检测方法，并在 KDD99 数据集上进行了实验，结果表明该算法最高能取得 99%以上的准确率和检测率，以及 0.01%的误报率。2018 年 M.AI-Qatf 等人[7]提出了基于稀疏自编码器和 SVM 的入侵检测算法，通过自编码器对数据集进行特征降维以提高 SVM 的分类精度和训练速度，该文献还在 J48、朴素贝叶斯、随机森林等机器学习算法上进行了广泛的实验，结果显示该算法在各性能指标上都有更好的表现。

循环神经网络（Recurrent Neural Network，RNN），因其接收上一时刻的隐层输出作为部分输入而得名。这种神经网络在保持人工神经网络非线性表达能力的同时，还能学习到序列数据间的相关性。网络流量数据通常也被认为是一种时序数据，因此 RNN 及其变种在 IDS 中的应用正逐渐成为近年的研究热点。2018 年 B.Yan 等人[8]提出的基于局部适应少数类过采样技术（Local Adaptive Synthetic Minority Over-Sampling，LA-SMOTE）和门控循环单元（Gated Recurrent Unit，GRU）的入侵检测模型；2020 年 S.Nayyar 等人[9]使用长短期记忆网络（Long Short-Term Memory，LSTM）进行入侵检测。2019 年 A.Kusupati 等人[10]提出了 FastGRNN（Fast Gated Recurrent Neural Network）和 FastRNN 算法，通过在 RNN 上直接添加门控机制，并在不同门控之间共享权重矩阵，Fast 算法能在保持 GRU 高分类精度的同时，大大减少参数数量。实验结果表明，FastGRNN 的内存占用显著低于 LSTM 和 GRU。鉴于 FastGRNN 的低资源占用表现，2021 年 P.Singh 等人[11]将该算法引入了入侵检测领域，并将系统部署到了资源严重受限的物联网设备中。实验结果表明，该算法在精度达到当时最佳水准的同时，占用内存和推断耗时都有显著下降。

对于机器学习算法而言，数据集的质量至关重要。以往入侵检测乃至网络安全领域的研究主要以 KDD99 数据集[12]为主。A.Divekar 等人[13]提出，尽管 KDD99 在领域中处于主导地位，但 KDD99 的各种缺陷正在拖累许多现代 IDS 在现实场景中的表现。这些缺陷包括：由于 KDD99 年代久远，无法囊括许多新的攻击手段；训练集中恶意样本只占总样本近 25%，模型倾向于学习多数样本，因此难以对恶意流量进行有效检测；测试集中的正常样本只有 19.4%，而仅仅 DoS 样本就占到 73.9%，与训练集相去甚远；训练集和测试集中都存在大量样本重复，训练集还存在测试集泄漏情况，使得最后训练出来的模型性能过于乐观。2015 年 N.Moustafa 等人[14]提出了 UNSW-NB15 数据集，UNSW-NB15 相比 KDD99 拥有更好的数据平衡性和稳定性，且入侵种类从 4 种提升到了 9 种，能够体现现代网络环境的复杂性。

在对近年相关文献研究的基础上，本文提出了一种适用于列车网络的入侵检测系统。该系统在各车厢放置检测节点，抓取各自车厢交换机上的网络数据包，经预处理后原地进行异常检测，再将结果汇总到位于驾驶室的总控设备中。相比由一台中央设备负责整辆列车的网络监测，该系统将运算负荷均匀地分摊到了多个设备上，因此起到了降低成本和提高系统效率的目的。此外，本文在 UNSW-NB15 数据集上搭建了基于 FastGRNN 的异常检测模型，将该模型部署到了各检测节点中，通过与其它机器学习模型在分类精度、内存占用、推断耗时上进行比较发现，借助 FastGRNN 的低系统资源占用，能够直接部署到列车现有设备上。

1　系统描述

在传统的公司网络环境中，入侵检测系统通常与防火墙一起工作。防火墙通过黑名单和白名单系统筛选来自外部网络的连接，入侵检测系统则通过监控核心交换机来保护网络安全[15]。当入侵检测系统检测到恶意攻击流量时，系统会向网络管理系统发出警报并触发进一步的安全保护措施。典型的公司网络安全体系结构如图 1 所示。

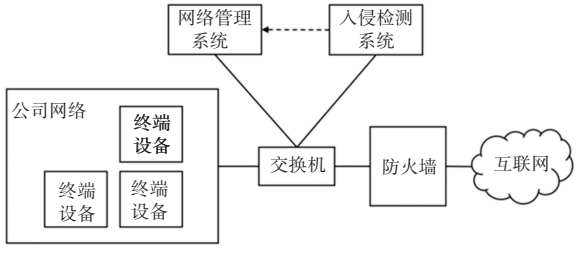

图 1　典型的公司网络安全体系结构

Fig. 1　Typical corporate network security architecture

当直接监控核心交换机时，入侵检测系统会接收进出公司网络的所有网络流量。如果网络通信繁忙，或者网络受到针对网络带宽的攻击时（如：分布式拒绝服务（Distributed Denial of Service，DDoS）攻击），入侵检测系统可能会瘫痪，这就需要入侵检测

系统采用更强大的硬件和更低延迟的入侵检测算法。同时,由于列车多车厢结构带来的网络拓扑特性,针对中心网络节点进行检测的结构也难以应用于列车通信网络。因此,在确定系统结构之前,首先要考虑列车通信网络的拓扑结构。

根据国际电工协会发布的 IEC 61375-1[16] 标准,使用以太网技术的列车通信网络,是由连接各车厢的以太网列车骨干网络(Ethernet Train Backbone Network,ETBN),以及连接车厢内部各终端设备(End Device,ED)的以太网编组网络(Ethernet Consist Network,ECN)两部分组成。列车通信网络结构如图 2 所示。

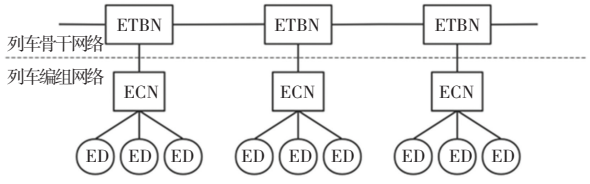

图 2 列车通信网络结构

Fig. 2 Structure of train communication network

各车厢的终端设备包括客室摄像头、LED 报站显示屏、LCD 导乘屏等,这些设备通过 ECN 进行通信,不同车厢中的设备则通过 ETBN 通信。因此,通过在各 ECN 节点放置入侵检测节点,借助分流器或端口镜像技术即可抓取流经本车厢的网络流量。结合上述讨论,本文提出的入侵检测系统如图 3 所示。

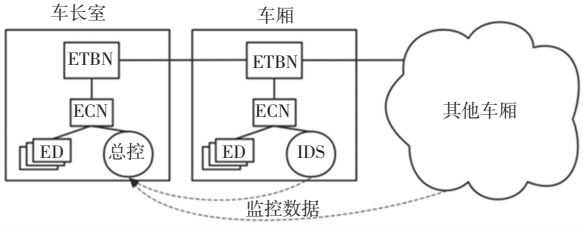

图 3 列车通信网络入侵检测系统结构

Fig. 3 Structure of intrusion detection system for train communication network

图 3 中的 IDS 为异常检测节点,当从 ECN 节点抓取到流经本车厢的网络流量后,经过特征提取和预处理后输入异常检测模型,并将检测结果告知位于司机室的总控设备。受篇幅所限,图中只列出了一个司机室和车厢,实际上所有车厢都会将检测结果发往总控设备。各异常检测节点的工作流程如图 4 所示,这些节点监控流经各车厢 ECN 节点的数据流量,对这些数据流量包进行特征提取和预处理,再输入 FastGRNN 模型,最后将模型给出的分类结果以 UDP 数据包的形式发送给司机室总控设备。

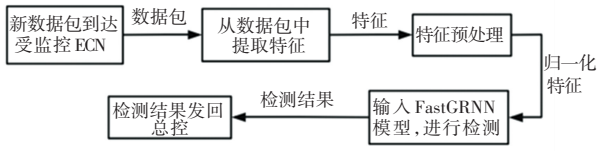

图 4 列车通信网络入侵检测系统工作流程

Fig. 4 Workflow of intrusion detection system for train communication network

2 FastGRNN 模型

FastGRNN 是 RNN 模型的一种,与 RNN 一样,FastGRNN 以外部输入和隐层上一时刻状态作为输入。FastGRNN 模型定义如下:

$$z_t = \sigma(\boldsymbol{W}\boldsymbol{x}_t + \boldsymbol{U}\boldsymbol{h}_{t-1} + \boldsymbol{b}_z) \tag{1}$$

$$\widetilde{\boldsymbol{h}}_t = \tanh(\boldsymbol{W}\boldsymbol{x}_t + \boldsymbol{U}\boldsymbol{h}_{t-1} + \boldsymbol{b}_h) \tag{2}$$

$$\boldsymbol{h}_t = (\zeta(1 - z_t) + \nu) \odot \widetilde{\boldsymbol{h}}_t + z_t \odot \boldsymbol{h}_{t-1} \tag{3}$$

式中,$\boldsymbol{x}_t$ 为 t 时刻输入向量,长度为输入特征数 D,$\boldsymbol{h}_t$ 为 t 时刻隐层输出向量,长度为隐层神经元数量 H,$\boldsymbol{W}$、$\boldsymbol{U}$ 是两种输入的权重矩阵,分别是 $H \times D$ 和 $H \times H$ 矩阵,$\boldsymbol{b}_z$、$\boldsymbol{b}_h$ 是偏置向量,长度为 H,tanh 和 σ 为非线性激活函数,分别为双曲正切和 *sigmoid* 函数,$\odot$ 为 Hadamard 积,即向量逐元素积,ζ、ν 为可训练标量,均为残差连接参数。

如果将公式(3)改写为 $\boldsymbol{h}_t = z_t \odot \widetilde{\boldsymbol{h}}_t + z_t \odot \boldsymbol{h}_{t-1}$,则得到类似于 LSTM 的网络。其中,$z \odot \boldsymbol{h}_{t-1}$ 为经过遗忘门控的记忆,而 $z_t \odot \widetilde{\boldsymbol{h}}_t$ 为新添加到记忆中的知识。相比之下,FastGRNN 在公式前半部分加入了残差连接机制 $(\zeta(1 - z_t) + \nu)$,经文献[9]证明,残差连接的加入使得 FastGRNN 的条件数(Condition Number)得到了有效控制,从而避免了传统 RNN 中严重的梯度爆炸和梯度消失问题,使得网络能够有效地学习数据集中的模式。通过让遗忘门控和输入门控共用权重矩阵,FastGRNN 只比传统 RNN 多了 $H + 2$ 个可训练参数,相比之下 LSTM 和 GRU 的可训练参数量分别为 $3H(D + H + 1)$ 和 $2H(D + H + 1)$。

本文对 UNSW-NB15 数据集进行二分类,即只判断网络流量是否为异常流量。为了方便后续计算各性能指标,该模型将进行二分类双标签预测,即输出两个值,分别代表输入正常和异常流量的概率,因此模型的输出层选用归一化指数函数(softmax)。输出层公式如下:

$$o_t = \text{softmax}(\boldsymbol{W}_o h_t) \tag{4}$$

其中,$\boldsymbol{W}_o$ 是一个 $2 \times H$ 的输出层权重矩阵。

3　实验与分析

为了确认 FastGRNN 模型是否能正确监测到列车通信网络中的异常流量，以及其被部署到资源受限的嵌入式设备中的可能性，本文在 UNSW-NB15 数据集上训练并测试了 FastGRNN 模型，记录其分类精度、模型内存占用率和运算效率，并将这些指标与其它机器学习模型进行比较。

3.1　数据集与数据预处理

UNSW-NB15 数据集由澳大利亚网络安全中心（Australian Center of Cyber Security，ACCS）提供。原始数据集共包含 254 万条记录，经筛选和过滤后，得到包含 17 万条的训练集和 8 万条数据的测试集。本文采用 UNSW-NB15 训练集作为数据集，取其中 75%作为训练集，其余 25%作为测试集，此外再在训练集中取 20%作为验证集，用于监控神经网络类模型训练时的拟合情况。UNSW-NB15 数据集有 42 个特征，其中 3 个是类型特征，其余为数值型特征。对于类型特征，本文采用独热编码（One-Hot Encoding），将可能取值数量为 n 的类型特征映射到长度为 n 的元组上；为了防止取值范围大的特征在模型中占主导地位，采用特征标准化来缩放数值特征，其公式如下：

$$x' = \frac{x - \bar{x}}{\sigma} \tag{5}$$

其中，x、x' 分别为标准化前后的特征；$\bar{x}$ 为特征均值；σ 为特征标准差。标准化后的数字代表其距离均值相差多少个标准差。

经预处理后数据集的特征数量增加到了 190 个。UNSW-NB15 共有 10 种标签，其中包括 1 个正常标签以及 9 个异常标签。由于本文提出的系统对网络流量进行二分类，因此将异常标签合并成 1 个标签。预处理后的数据集数据分布见表 1。

表 1　数据集数据分布

Tab. 1　Distribution of the dataset

标签	训练集	百分比/%	测试集	百分比/%
正常	27 749	44.94	9 251	44.94
异常	34 000	55.06	11 332	55.06
合计	61 749	100	20 583	100

3.2　性能指标

本文使用准确率（Accuracy）、精确率（Precision）、召回率（Recall）和 F1 分数来评判模型的分类精度。之所以不单独使用准确率作为评判标准是因为其缺乏可参考性：对于高度不平衡的数据集，例如 99%的标签都是正常的情况下，模型不用做任何判断便可得到 99%的准确率。上述 4 种指标的计算方式如下：

$$A = \frac{\text{正确分类的样本数}}{\text{样本总数}} \tag{6}$$

$$P = \frac{TP}{TP + FP} \tag{7}$$

$$R = \frac{TP}{TP + FN} \tag{8}$$

$$F_1 = 2 \times \frac{P \times R}{P + R} \tag{9}$$

其中，A、P、R、$F1$ 分别为准确率、精确率、召回率和 $F1$ 分数，TP、FP、FN 分别为将正类分类为正类、负类分类为正类、正类分类为负类的数量。精确率的含义是当模型认为一个样本是正类时，其确实是正类的比率；召回率的含义是所有正类被模型正确分类的比率，即使在数据集极度不平衡的情况下，只要正类是少数类，精确率和召回率也可以体现出模型的性能。例如：对于上述不做任何判断的模型，其精确率是 0，召回率是无效数值，因为 TP 和 FN 都是 0；$F1$ 分数是精确率和召回率的调和平均数，其优势是可以综合考虑模型的精确率和召回率，且相比算数平均数，使用调和平均数的 $F1$ 分数可以更有效地惩罚精确率或召回率接近 0 的情况。

3.3　实验环境及参数设置

本文实验的硬件环境为 Intel i7-2600、NVIDIA GTX1660；软件环境为 Python3.7；FastGRNN 基于开源库 tensorflow 中的 RNNCell 基类和公式（1~3）；对比实验中的神经网络模型来自开源库 Keras，其它模型来自开源库 scikit-learn。对于输出为概率的分类模型，分类阈值设为 0.8。

FastGRNN 模型参数见表 2。

表 2　FastGRNN 模型参数

Tab. 2　Model parameters of FastGRNN

模型参数名称	参数值
输入层维度	38
隐层维度	40
输出层维度	2
输出层激活函数	softmax
优化器	自适应矩估计（Adam）
学习率	0.01
批大小	100
迭代批次	125
损失函数	多类交叉熵

3.4 实验结果

除 FastGRNN 模型外,本文还使用相同的实验流程,在 UNSW－NB15 数据集上对支持向量机(Support Vector Machine,SVM)、K 近邻算法(K－Nearest Neighbors)、随机森林、多层感知机(Multilayer perceptron,MLP)、长短期记忆网络(Long Short－Term Memory,LSTM)、门控循环单元(Gated Recurrent Unit,GRU)等机器学习模型进行了测试。各模型的性能指标测试结果见表 3。

从实验结果可以看出,FastGRNN 模型拥有最优的分类精度(准确率、精确率、召回率、*F*1 分数),并且内存占用显著低于其它神经网络模型。虽然非神经网络模型(SVM、KNN、随机森林)只占用很少的内存,但相比 FastGRNN 模型,这些模型都在分类精度或计算时间上存在短板。

FastGRNN 的高分类精度使得基于其搭建的 IDS 拥有更高的可靠性。市面上许多嵌入式芯片的 L1 缓存便足以容纳下该模型(如 STM32MP1 系列)。这意味着将该模型部署到硬件平台后,能在拥有高检测精度的同时保持低功耗。

表 3 实验结果

Tab. 3 Experiment result

模型名称	准确率/%	精确率/%	召回率/%	*F*1 分数	内存占用/kb	每样本计算时间/ms
FastGRNN	96.54	96.56	96.54	0.97	14.06	29.41
SVM	92.45	92.47	92.46	0.92	4.22	2 184.83
KNN	92.98	93.16	92.98	0.93	1.95	1 597.95
随机森林	90.00	90.00	90.00	0.90	3.42	2.83
MLP	89.78	91.58	89.78	0.90	37.70	14.63
LSTM	93.69	94.37	93.69	0.94	1 223.44	36.34
GRU	93.47	94.20	93.47	0.93	917.98	37.23

4 结束语

本文提出了一种适用于列车通信网络的入侵检测系统,通过将检测节点分散到各车厢,原地完成检测工作并将结果汇总到总控设备,该系统将任务负载分摊到多个节点上,从而起到了提高效率和降低成本的作用。此外,本文基于 UNSW－NB15 数据集测试了 FastGRNN 模型。根据实验结果,该模型能达到 0.97 的 *F*1 分数、14.06 KB 的内存占用和29.41 ms的平均单样本检测时间,能够部署到列车现有设备上,只需对列车现有硬件结构做有限修改即可引入网络安全保护。

参考文献

[1] 王中尧. 以太网技术在列车通信网络中的应用探究[J]. 中国管理信息化, 2018, 21(16): 141－142.

[2] 常振臣, 牛得田, 王立德, 等. 列车通信网络研究现状及展望[J]. 电力机车与城轨车辆, 2005(3): 5－7,60.

[3] LIAO H J, RICHARD LIN C H, LIN Y C, et al. Intrusion detection system: A comprehensive review[J]. Journal of Network and Computer Applications, 2013, 36(1): 16－24.

[4] SALAMA M, EID H, RAMADAN R, et al. Hybrid Intelligent Intrusion Detection Scheme[J]. Advances in Intelligent and Soft Computing, 2011, 96:295－302.

[5] MAR J, HSIAO I F, YEH Y C, et al. Intelligent intrusion detection and robust null defense for wireless networks[J]. International Journal of Innovative Computing, Information and Control, 2012, 8: 3341－3359.

[6] KIM J, SHIN N, JO S, et al. Method of intrusion detection using deep neural network[EB/OL]. 2017: 313－316. https://doi.org/10.1109/BIGCOMP.2017.7881684.

[7] AL－QATF M, LASHENG Y, AL－HABIB M, et al. Deep Learning Approach Combining Sparse Autoencoder With SVM for Network Intrusion Detection[J]. IEEE Access, 2018, 6: 52843－52856.

[8] YAN B, HAN G. LA－GRU: Building Combined Intrusion Detection Model Based on Imbalanced Learning and Gated Recurrent Unit Neural Network[J]. Security and Communication Networks, 2018, 2018: e6026878.

[9] NAYYAR S, ARORA S, SINGH M. Recurrent Neural Network Based Intrusion Detection System [C]//2020 International Conference on Communication and Signal Processing (ICCSP). 2020: 136－140.

[10] KUSUPATI A, SINGH M, BHATIA K, et al. Fastgrnn: A fast, accurate, stable and tiny kilobyte sized gated recurrent neural network[J]. Advances in neural information processing systems, 2018, 31.

[11] SINGH P, PANKAJ A, MITRA R. Edge－detect: Edge－centric network intrusion detection using deep neural network[C]//2021 IEEE 18th Annual Consumer Communications & Networking Conference (CCNC). IEEE, 2021: 1－6.

[12] KDD Cup 1999 Data[EB/OL]. [1999－10－28]. http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html.

[13] DIVEKAR A, PAREKH M, SAVLA V, et al. Benchmarking datasets for Anomaly－based Network Intrusion Detection: KDD CUP 99 alternatives[C]//2018 IEEE 3rd International Conference on Computing, Communication and Security (ICCCS), 2018: 1－8.

(下转第 59 页)