计算机科学 COMPUTER SCIENCE
http://www.jsjkx.com
DOI:10.11896/jsjkx.200900181

基于移动边缘计算的车载 CAN 网络入侵检测方法

于天琪[1]　胡剑凌[1]　金　炯[2]　羊箭锋[1]
1 苏州大学电子信息学院　江苏 苏州 215006
2 斯威本科技大学软件与电气工程学院　墨尔本 3122
(tqyu@suda.edu.cn)

摘　要　随着车联网技术的快速发展和广泛部署,其在为智能网联汽车提供互联网与大数据分析等智能化服务的同时,引入了网络入侵等安全与隐私问题。传统车载网络的封闭性导致现有的车载网络通信协议,特别是部署最为广泛的控制器局域网络(Controller Area Network,CAN)总线协议,在发布时缺少隐私与安全保护机制。因此,为检测网络入侵、保护智能网联汽车安全,文中提出了一种基于支持向量数据描述(Support Vector Data Description,SVDD)的车载 CAN 网络入侵检测方法。该方法提取单位时间窗内 CAN 网络报文 ID 的加权自信息量和 ID 的归一化值作为特征信息,并在移动边缘计算服务器处构建并训练 SVDD 模型,目标车辆基于训练的 SVDD 模型进行异常特征值识别,从而实现实时的车载 CAN 网络入侵检测。文中采用韩国高丽大学 HCR 实验室公开的 CAN 网络数据集,对所提方法与 3 种传统的基于信息熵的车载网络入侵检测方法在拒绝服务攻击和伪装攻击检测准确率方面进行了对比与分析。仿真实验结果表明,在少量报文入侵时,所提方法显著提高了入侵检测的准确率。

关键词:车联网;移动边缘计算;车载网络;网络入侵检测;支持向量数据描述算法
中图法分类号　TN915

Mobile Edge Computing Based In-vehicle CAN Network Intrusion Detection Method

YU Tian-qi[1],HU Jian-ling[1],JIN Jiong[2] and YANG Jian-feng[1]
1 School of Electronic and Information Engineering,Soochow University,Suzhou,Jiangsu 215006,China
2 School of Software and Electrical Engineering,Swinburne University of Technology,Melbourne 3122,Australia

Abstract　With the rapid development and pervasive deployment of the Internet of Vehicles (IoV),it provides the services of Internet and big data analytics to the intelligent and connected vehicles,while incurs the issues of security and privacy. The closure of traditional in-vehicle networks leads to the communications protocols,particularly,the most commonly applied controller area network (CAN) bus protocol,lack of security and privacy protection mechanisms. Thus,to detect the network intrusions and protect the vehicles from being attacked,a support vector data description (SVDD) based intrusion detection method is proposed in this paper. Specifically,the weighted self-information of message IDs and the normalized values of IDs are selected as features for SVDD modeling,and the SVDD models are trained at the mobile edge computing (MEC) servers. The vehicles use the trained SVDD models for identifying the abnormal values of the selected features to detect the network intrusions. Simulations are conducted based on the CAN network dataset published by the HCR Lab of Korea University,where three conventional information entropy based in-vehicle network intrusion detection methods are adopted as the benchmarks. As compared to the benchmarks,the proposed method has dramatically improved the intrusion detection accuracy,especially when the number of intruded messages is small.

Keywords　Internet of Vehicles,Mobile edge computing,In-vehicle network,Network intrusion detection,Support vector data description algorithm

1　引言

5G 移动通信网络的商用大力推动了车联网的发展[1-2]。作为车联网的主体,智能网联汽车通过车载传感系统和信息终端来实现与人、车、路等方面的信息交换,使车辆具备智能的环境感知能力,并且能够自动分析行驶状态,进而实现智能

到稿日期:2020-09-25　返修日期:2020-12-07　　本文已加入开放科学计划(OSID),请扫描上方二维码获取补充信息。
基金项目:江苏省自然科学基金(BK20200858)
This work was supported by the Natural Science Foundation of Jiangsu Province,China(BK20200858).
通信作者:羊箭锋(jfyang@suda.edu.cn)

辅助驾驶及自动驾驶[3]。

随着智能网联汽车中嵌入的电子控制单元（Electronic Control Unit，ECU）和外部通信接口数目的增长，为确保车载元件间有效的控制与通信，车载网络从简单的点对点控制总线，逐步发展成为分布式异构的多元通信与控制网络。车载网络的示意图如图 1 所示[4]。

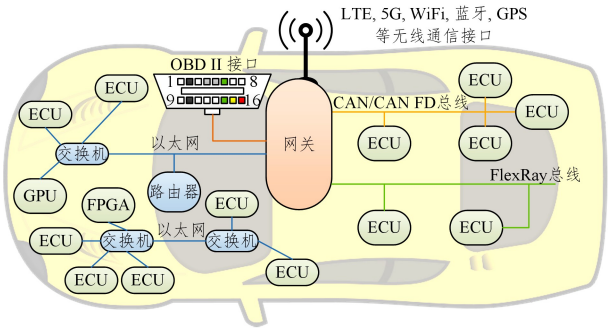

图 1　车载网络示意图

Fig. 1　Diagram of in-vehicle network

车联网通过外部通信接口在将互联网与大数据分析等智能化服务引入车辆的同时，也引入了潜在的网络入侵风险。传统车载网络的封闭性导致现有的车载网络通信协议，特别是部署最为广泛的 CAN 总线协议，在发布时缺少包括接入控制、认证、加密等在内的隐私与安全保护机制。由于缺少接入控制机制，攻击者可以通过攻破外部接口直接入侵车内网络。如果不能及时地检测恶意入侵，攻击者则可通过入侵 CAN 总线掌握车辆的绝对控制权。例如，Miller 等通过远程入侵 CAN 总线，迫使正行驶于高速公路上的 Jeep Cherokee 紧急刹车并冲向路边，为此 Chrysler 公司召回了 140 万辆汽车[5]。腾讯科恩实验室通过 Wi-Fi 接口实现了对 Tesla S 系列汽车分别在驻车和行驶状态下的远程入侵以及对 CAN 总线的绝对控制，迫使 Tesla 公司紧急更新了车载网络系统[6]。

此外，如果目标车辆不能及时地检测恶意入侵并采取措施，则不仅会导致自身的隐私泄露与安全危机，还会使攻击者通过车联网入侵来控制其他车辆，造成不可挽回的损失。然而，更新与制造成本导致以 CAN 总线为代表的车载网络协议和架构在短时间内难以被替代。与此同时，由于车载网络资源的局限性，复杂的网络安全机制难以直接应用其中。因此，亟需在现存协议与架构的基础上，展开面向车载网络入侵检测方法的相关研究，以及时准确地检测恶意入侵，进而阻止由入侵带来的进一步攻击和威胁。

针对上述问题，本文提出了一种基于支持向量数据描述的车载 CAN 网络入侵检测方法。该方法提取单位时间窗内 CAN 网络报文 ID 的加权自信息量和 ID 的归一化值作为特征信息，在移动边缘计算（Mobile Edge Computing，MEC）服务器处构建并训练 SVDD 模型。目标车辆基于下发的 SVDD 模型进行异常特征值检测，从而实现实时的 CAN 网络入侵检测。在仿真实验部分，本文采用韩国高丽大学 HCR 实验室公开的 CAN 网络数据集，对所提方法和 3 种传统的基于信息熵的检测方法在拒绝服务（Denial-of-Service，DoS）和伪装攻击的检测准确率方面进行了比较。结果表明，在单组入侵报文数量较小时，本文方法在检测准确率方面有显著提高。

2　国内外研究现状

车联网相关技术的迅猛发展和系统的广泛部署为网联汽车提供了智能化服务，但同时增加了内部车载网络遭受恶意入侵的风险。为阻止由网络入侵带来的进一步恶意攻击，及时有效的入侵检测已成为当下的研究热点。根据车载网络特征信息的类型，本文对车载网络入侵检测方法的研究现状进行了分析，包括物理指纹、网络参数以及报文信息熵等方面[7]。

物理指纹指 ECU 难以克隆的硬件特征信息。Choi 等[8]利用序列前向选择策略提取 ECU 发送的 CAN 总线电信号在时频域的关键特征，对每一个合法接入的 ECU 建模从而构建模型数据库，通过数据库以外的观测值检测 ECU 伪装攻击。文献[9]采用 ECU 的时钟偏移信息，通过递归最小二乘法构建了合法接入 ECU 的时钟偏移模型，进而利用超出阈值的观测值与模型估计值的累计差值来检测 ECU 的伪装攻击。文献[10]利用 ECU 的电压信息，对合法接入 ECU 发送的 CAN 总线电信号的高低电压出现频率进行了统计和建模，利用高低电压概率分布的偏移检测 ECU 伪装攻击。基于物理指纹的 ECU 特征建模能够精准地构建合法接入 ECU 的模型数据库，进而有效地表征非法接入 ECU 的伪装攻击，但硬件特征信息的获取难度较高且其难以描述利用合法接入 ECU 发动攻击的入侵状态。

网络参数指车载网络通信过程中数据包发送的相关参数。文献[11]通过长期监测特定 ID 的发包周期和网络中数据包的时间间隔，统计出车载网络中发包间隔的区间，利用发包间隔没有落在统计区间内的异常值来检测泛洪和重播等注入式攻击。Lee 等[12]提出了一种具有特定标识符的远程帧，并通过远程帧的请求和响应过程的偏移率与时间间隔的相关系数波动以及特定节点时间间隔概率分布的峰值偏移来检测注入式攻击和伪装攻击。网络参数特征建模主要利用数据包稳定的发包周期和发包规律，因此在发包周期不规律的车载网络中，特征模型构建的难度较高。

报文信息熵指单位时间内 CAN 网络报文 ID 的信息熵。Müter 等[13]首次提出利用正常状态和入侵状态下单位时间窗内报文 ID 信息熵的波动以及报文相对熵的波动来检测泛洪攻击和特定 ID 报文重播攻击。文献[14]同样采用单位时间窗内报文信息熵的波动和报文相对距离的波动来检测泛洪攻击和报文重播攻击。Wu 等[15]在选取单位时间窗内不同报文 ID 信息熵作为网络特征的基础上，提出了一种基于固定报文数量的滑动窗策略，以避免不同波特率和非周期性 CAN 报文对信息熵的干扰，从而更加准确地表征车载网络的不同状态。基于报文信息熵的特征建模能够有效地表征洪泛攻击等有海量高频报文注入的入侵状态，但难以有效地描述少量报文注入时对信息熵影响较小的入侵状态。

本文在报文信息熵模型的基础上，构建了报文 ID 的加权自信息量模型，并与 ID 归一化值一起作为特征信息，构建并训练 SVDD 模型来检测不同 ID 报文的变化，解决了基于报文信息熵的方法在少量报文注入时检测准确率较低的问题。

3 基于移动边缘计算的车联网系统架构

基于移动边缘计算的车联网系统架构如图 2 所示[16]。

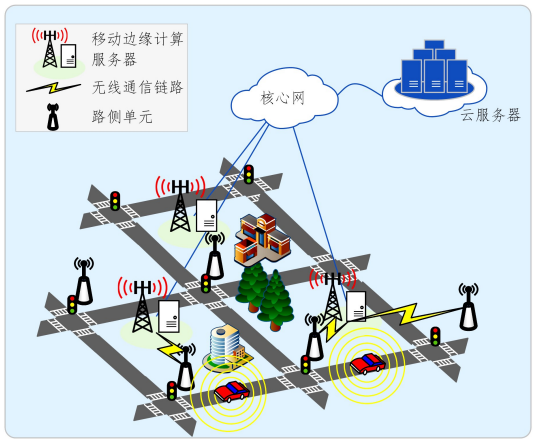

图 2 基于移动边缘计算的车联网系统架构

Fig. 2 Diagram of MEC-based IoV system architecture

(1)接入层。由智能网联汽车和路侧单元构成,后者是开放式服务接入点,车辆通过路侧单元将数据与计算任务上传至系统上层,上层反馈与控制命令通过路侧单元下发至目标车辆。

(2)移动边缘计算层。由移动通信网络(微)基站和 MEC 服务器构成,部署于车辆及其自组织网络的边缘侧,提供稳定的网络接入和本地高实时性的数据缓存与处理[17-18]。在本文方法中,MEC 服务器负责存储和处理车辆上传的网络报文数据,基于历史数据构建并训练 SVDD 模型,随后将训练模型下发至目标车辆,用于车辆端的网络入侵检测。

(3)云计算层。由云服务器构成,与移动边缘计算层通过核心网进行交互,用于系统级的设备管理、资源分配、海量数据的存储与处理。

4 基于 SVDD 的车载网络入侵检测方法

本节首先构建 CAN 报文 ID 加权自信息量模型,其次描述 SVDD 算法,最后在 CAN 报文 ID 加权自信息量模型和 SVDD 算法的基础上,提出了车载 CAN 网络入侵检测方法。

4.1 CAN 报文 ID 加权自信息量模型

CAN 网络数据帧的结构如图 3 所示。其中,ID 的大小为 11 bits,用于表明发送数据的目的地和优先级,ID 越小则优先级越高;DLC 的大小为 4 bits,用于表明数据的长度;数据域(Data Field)为所要传送的数据内容,其长度为 0～64 bits。本文在文献[13-15]的基础上,构建报文 ID 信息熵和加权自信息量模型。在单位时间窗内,报文 ID 为 id 的概率为:

$$p_{id}=NUM_{id}/NUM_{\text{total}} \tag{1}$$

其中,NUM_{id} 为该时间窗内 ID 为 id 的报文数量,NUM_{total} 为该时间窗内报文的总数量,则该报文 ID 的自信息量为:

$$I_{id}=-\log p_{id} \tag{2}$$

其加权自信息量为:

$$I_{w,id}=-p_{id}\log p_{id} \tag{3}$$

则在该时间窗内报文 ID 的信息熵为:

$$H(ID)=\sum_{id\in ID}I_{w,id}=\sum_{id\in ID}-p_{id}\log p_{id} \tag{4}$$

其中,ID 为该时间窗内报文 ID 的集合。

	Arbitration Field		Control Field			Data Field	CRC Field		ACK Field		
SOF	ID	RTR	IDE	RBO	DLC	Data	CRC	CRC Del	ACK	ACK Del	EOF
1bit	11bits	1bit	1bit	1bit	4bits	0~64bits	15bits	1bit	1bit	1bit	7bits

图 3 CAN 网络数据帧的结构

Fig. 3 Structure of CAN data frame

4.2 SVDD 算法

SVDD 是一种基于支持向量(即边界数据)的描述方法,其目标是寻求一个包含几乎所有目标样本且体积最小的超球体(域)。因此,SVDD 可作为单值分类器区分目标样本和非目标样本,常被用于异常检测和故障检测等领域。SVDD 算法的原理如下[19]。

假设有一组正类训练数据 $\boldsymbol{x}\in\mathbb{R}^{n\times d}$,其中 n 是样本个数,d 是特征维度。首先通过非线性变换函数 $\Phi:\boldsymbol{x}\rightarrow\boldsymbol{F}$ 将数据从原始空间映射到特征空间,并在特征空间中确立体积最小的超球体。为了构造该超球体,SVDD 优化问题可确立为:

$$\min_{\boldsymbol{a},R,\xi}R^2+C\sum_{i=1}^{n}\xi_i \tag{5}$$

$$\text{s.t.}\ \ \|\Phi(\boldsymbol{x}_i)-\boldsymbol{a}\|^2\leqslant R^2+\xi_i,\xi_i\geqslant 0,\forall\, i=1,2,\cdots,n \tag{6}$$

其中,R 是超球体半径,$\boldsymbol{a}$ 是超球体球心,ξ 是松弛因子,C 是一个权衡超球体体积和误分率的惩罚参数。结合拉格朗日乘子法,式(5)的优化问题可以对偶为:

$$\max_{\alpha_i}\sum_{i=1}^{n}\alpha_iK(\boldsymbol{x}_i,\boldsymbol{x}_i)-\sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\alpha_jK(\boldsymbol{x}_i,\boldsymbol{x}_j) \tag{7}$$

$$\text{s.t.}\ \ 0\leqslant\alpha_i\leqslant C \tag{8}$$

$$\sum_{i=1}^{n}\alpha_i=1 \tag{9}$$

其中,α_i 是样本 $\boldsymbol{x}_i$ 对应的拉格朗日系数。通过求解问题(7),可获取所有样本对应的拉格朗日系数。在所有训练样本中,把拉格朗日系数满足 $0<\alpha_i<C$ 的样本称为支持向量,假设训练数据集中属于支持向量的样本集合为 $\boldsymbol{SV}$,那么超球体球心和半径的计算公式分别为:

$$\boldsymbol{a}=\sum_{i=1}^{n}\alpha_i\Phi(\boldsymbol{x}_i) \tag{10}$$

$$R^2=K(\boldsymbol{x}_k,\boldsymbol{x}_k)-2\sum_{i=1}^{n}\alpha_iK(\boldsymbol{x}_k,\boldsymbol{x}_i)+\sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\alpha_jK(\boldsymbol{x}_i,\boldsymbol{x}_j) \tag{11}$$

其中,$\boldsymbol{x}_k\in\boldsymbol{SV}$;$K(\boldsymbol{x}_i,\boldsymbol{x}_j)$ 是核函数,等同于特征空间中样本的内积:

$$K(\boldsymbol{x}_i,\boldsymbol{x}_j)=\langle\Phi(\boldsymbol{x}_i),\Phi(\boldsymbol{x}_j)\rangle \tag{12}$$

测试样本 $\boldsymbol{x}_{\text{test}}$ 到超球体的距离为:

$$d=\sqrt{K(\boldsymbol{x}_{\text{test}},\boldsymbol{x}_{\text{test}})-2\sum_{i=1}^{n}\alpha_iK(\boldsymbol{x}_{\text{test}},\boldsymbol{x}_i)+\sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\alpha_jK(\boldsymbol{x}_i,\boldsymbol{x}_j)} \tag{13}$$

若 $d \leqslant R$，表明 $\boldsymbol{x}_{\text{test}}$ 在超球体内或超球面上，并判定为正常样本；反之，则判定为异常样本。

4.3 基于 SVDD 的车载网络入侵检测方法

在 CAN 网络报文 ID 加权自信息量建模和 SVDD 算法的基础上，本节提出了基于 SVDD 的车载 CAN 网络入侵检测方法，具体如下。

在初始化阶段，智能网联汽车将车载 CAN 网络的报文数据，通过路侧单元上传至 MEC 服务器构建历史数据集，用于 SVDD 模型的训练。

在 SVDD 模型训练阶段，首先提取用于训练的 CAN 网络报文 ID 的时间序列为 $\boldsymbol{id}_{\text{Train}}=\{id^{(1)},id^{(2)},\cdots,id^{(t)},\cdots\}$。假设时间窗的窗长为 W，则时间窗 i 内的报文 ID 序列为 $\boldsymbol{id}_W^{(i)}$，将其代入式(1)－式(3)得到加权自信息量 $\boldsymbol{I}_w^{(i)}$。选择归一化的 $id(\tilde{id}=id/0\text{x}7\text{ff})$和加权自信息量作为 SVDD 模型的特征信息，则训练数据及其对应的标签标记分别为 $\boldsymbol{x}^{(i)}=\{(0.328,\boldsymbol{I}_{w,0x2a0}^{(i)}),\cdots,(\tilde{id},\boldsymbol{I}_{w,id}^{(i)}),\cdots\}$，$\boldsymbol{y}^{(i)}=\{1,\cdots,1,\cdots\}$。利用 $\boldsymbol{x}=\{\boldsymbol{x}^{(1)},\cdots,\boldsymbol{x}^{(i)},\cdots\}$和 $\boldsymbol{y}=\{\boldsymbol{y}^{(1)},\cdots,\boldsymbol{y}^{(i)},\cdots\}$ 进行模型训练，并将得到的 svdd 模型下发至目标车辆用于网络入侵检测。该训练过程如算法 1 所示。

算法 1 MEC 服务器端 SVDD 模型的训练

输入：$(\mathbf{id}_{\text{Train}},\text{W})$

输出：svdd 模型

1. for i←1 to ♯ $\mathbf{id}_{\text{Train}}$/W do
2. $\quad\mathbf{id}_{\text{W}}^{(\text{i})}\leftarrow\{\text{id}^{((\text{i}-1)*\text{W}+1)},\cdots,\text{id}^{(\text{i}*\text{W})}\}$
3. $\quad\mathbf{I}_{\text{w}}^{(\text{i})}\leftarrow$substitute $\mathbf{id}_{\text{W}}^{(\text{i})}$ into (1)－(3)
4. $\quad\mathbf{x}^{(\text{i})}\leftarrow\{\tilde{\text{id}},\mathbf{I}_{\text{w}}^{(\text{i})}\},\mathbf{y}^{(\text{i})}=\{1,1,\cdots,1\}$
5. end
6. svdd←SVDD($\mathbf{x}$,$\mathbf{y}$)/＊ SVDD 模型训练 ＊/

在网络入侵检测阶段，智能网联汽车将时间窗 W 内的实时 CAN 网络通信报文 $\boldsymbol{id}_{W\_Test}^{(t)}$ 代入式(1)－式(3)中，得到加权自信息量 $\boldsymbol{I}_{w\_Test}^{(t)}$，并构建测试数据集 $\boldsymbol{x}_{\text{Test}}^{(t)}$。利用已训练的 svdd 模型，得到相应的 $\boldsymbol{y}_{\text{Test}}^{(t)}$，若标签为－1，则该 id 的报文检测为入侵报文；若标签为 1，则为正常网络报文。该检测过程如算法 2 所示。

算法 2 基于 svdd 模型的网络入侵检测

输入：svdd 模型

输出：入侵报文 ID

1. while t do
2. $\quad\mathbf{id}_{\text{W\_Test}}^{(\text{t})}\leftarrow\{\text{id}^{((\text{t}-1)*\text{W}+1)},\cdots,\text{id}^{(\text{t}*\text{W})}\}$
3. $\quad\mathbf{I}_{\text{w_Test}}^{(\text{t})}\leftarrow$substitute $\mathbf{id}_{\text{W\_Test}}^{(\text{t})}$ into (1)－(3)
4. $\quad\mathbf{x}_{\text{Test}}^{(\text{t})}\leftarrow\{\tilde{\text{id}},\mathbf{I}_{\text{w\_Test}}^{(\text{t})}\}$
5. $\quad\mathbf{y}_{\text{Test}}^{(\text{t})}\leftarrow\text{svdd}(\mathbf{x}_{\text{Test}}^{(\text{t})})$
6. $\quad$if $-1\in\mathbf{y}_{\text{Test}}^{(\text{t})}$
7. $\quad\quad$intrusion detected id←$\mathbf{y}_{\text{Test}}^{(\text{t})}(\text{id})=-1$
8. $\quad$end
9. end

5 实验及结果分析

5.1 实验数据库

本文采用韩国高丽大学 HCR 实验室的 CAN 网络数据集[12]，针对所提出的网络入侵检测方法进行性能分析。数据集的相关信息如表 1 所列。

表 1 HCR 实验室 CAN 网络数据库

Table 1 CAN dataset from HCR lab

	Count	ID Range
Attack-free	30000	0x001～0x7ff
DoS Attack	31000～40000	0x000～0x7ff
Spoofing Attack	31000～40000	0x001～0x7ff

如表 1 所列，采用汽车在正常行驶状态下采集的网络数据 30000 条，报文 ID 的范围是 0x001～0x7ff。为检测 DoS 攻击和伪装攻击，随机插入 200 组攻击，每组入侵报文数量为 5～50。DoS 攻击采用 ID 优先级最高的 0x000 进行泛洪式攻击，阻止 CAN 网络的一切通信和服务；伪装攻击通过伪装发送合法的报文 ID，驱使车辆进行非驾驶员控制的特定操作，如紧急刹车、持续提速等。实验中，正常行驶状态下的数据用于模型训练，入侵状态下的数据用于测试。

5.2 评估参数

评估过程中采用的参数有真阳率(true positive rate，TPR)、假阳率(false positive rate，FPR)和准确率(accuracy)。

$$TPR=TP/(TP+FN) \tag{14}$$

$$FPR=FP/(FP+TN) \tag{15}$$

$$Accuracy=\frac{TP+TN}{TP+FP+TN+FN} \tag{16}$$

其中，TP，FP，TN，FN 分别为异常数据检测为异常、正常数据检测为异常、正常数据检测为正常、异常数据检测为正常的数量。

5.3 实验结果分析

在性能评估部分，本文提出的基于 SVDD 的入侵检测方法的参数设置如下：信息熵时间窗 W 的大小为 50 条报文，SVDD 模型训练数据集的大小为 500，即 $\boldsymbol{x}_{\text{Train}}\in\mathbb{R}^{500\times 2}$，所采用的核函数为拉普拉斯核(Laplacian Kernel)，SVDD 模型训练采用 Python SVDD 工具箱[20]。模型训练数据和训练所得到的支持向量如图 4 所示。

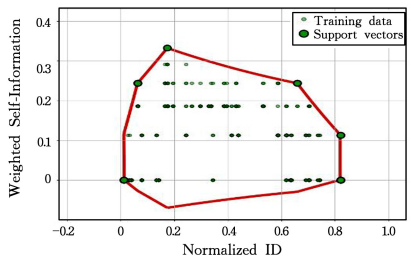

图 4 SVDD 模型的训练数据和支持向量

Fig. 4 Training data and support vectors of SVDD model

为验证本文方法在检测准确率方面的提升效果，将其与文献[13]和文献[14]中基于单位时间窗内信息熵(Traditional Entropy)和相对熵(Relative Entropy)的方法，以及文献[15]中基于固定报文数的滑动窗内信息熵(Sliding-window Entropy)的方法，对 DoS 和伪装两种攻击下的检测准确率进行了对比与分析。

基于单位时间窗内相对熵的方法不适用于 DoS 攻击的

检测，因此图 5 仅针对本文方法与基于单位时间窗内信息熵和基于固定报文数的滑动窗内信息熵的方法，在 DoS 攻击不同单组入侵报文数量的条件下进行了对比。如图 5 所示，随着单组入侵报文数量从 5 增长至 50，基于信息熵和基于滑动窗信息熵的方法的检测准确率分别从 79.02%和 78.59%增长至 94.22%和 95.33%。本文提出的基于 SVDD 的方法在不同入侵报文数量的条件下，检测准确率稳定在 97.70%以上。由此可见，在 DoS 攻击下，单组入侵报文数量低于 20 时，本文方法相比基于信息熵和基于滑动窗信息熵的方法在检测准确率方面有显著提高；在单组入侵报文数量分别为 5，10，20 时，检测准确率相比前者分别提高了 25.55%，22.43%，20.71%，相比后者分别提高了 26.24%，21.24%，15.30%。

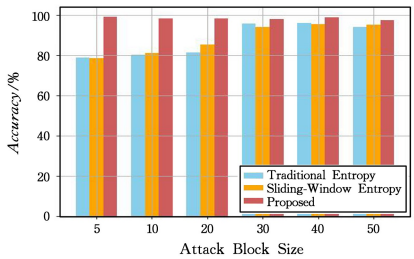

图 5 不同入侵报文数量下 DoS 攻击检测准确率的对比

Fig. 5 Comparison on detection accuracy of DoS attack under different attack block sizes

此外，在 DoS 攻击单组入侵报文数量为 5 时，3 种方法的 ROC 曲线如图 6 所示。ROC 曲线的横轴为 FPR，纵轴为 TPR。ROC 曲线下的面积(Area Under Curve，AUC)为 0～1，越接近于 1 表明检测效果越好。图 6 同样表明，在单组入侵报文数量为 5 时，本文方法优于基于信息熵和基于滑动窗信息熵的方法。本文方法在少量报文攻击时，对检测准确率有显著提高的原因在于，区别于总体统计单位时间窗内报文的信息熵，本文方法分别计算了不同报文 ID 的加权自信息量，对报文入侵更加敏感。

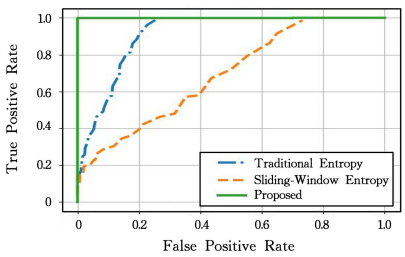

图 6 单组入侵报文数量为 5 时 DoS 攻击检测的 ROC 曲线

Fig. 6 ROC curve of DoS attack detection when attack block size is 5

图 7 给出了本文方法与基于单位时间窗内信息熵、基于固定报文数的滑动窗内信息熵、基于单位时间窗内相对熵这 3 种方法，在伪装攻击不同入侵报文数量下的对比情况。图 8 给出了伪装攻击单组入侵报文数为 5 时，4 种方法的 ROC 曲线。图 7、图 8 表明，在伪装攻击单组入侵报文数量低于 20 时，本文方法优于基于信息熵和基于滑动窗信息熵的方法。在单组入侵报文数为 5，10，20 时，准确率相较于前者分别提高了 24.54%，24.29%，21.22%，相较于后者分别提高了 32.29%，23.69%，17.19%。

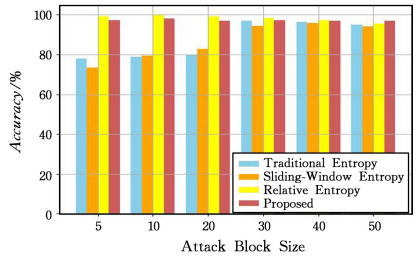

图 7 不同入侵报文数量下伪装攻击检测准确率的对比

Fig. 7 Comparison on detection accuracy of spoofing attack under different attack block sizes

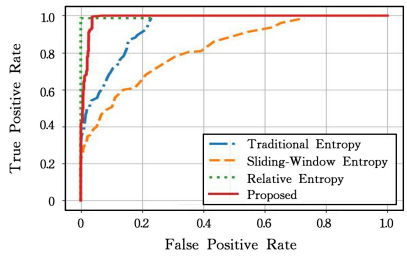

图 8 单组入侵报文数量为 5 时伪装攻击检测的 ROC 曲线

Fig. 8 ROC curve of spoofing attack detection when attack block size is 5

图 7 还表明，本文方法在不同入侵报文数量的条件下，检测准确率在 96.98%以上；而基于相对熵的方法，检测准确率也在 95.72%以上。如图 8 所示，在单组入侵报文数为 5 时，基于相对熵的方法的 AUC 略高于本文方法。其原因在于基于相对熵的方法与本文方法相似，分别计算了不同报文 ID 单位时间窗内的概率变化，对报文入侵敏感，对合法报文 ID 的伪装攻击能够取得较高的检测准确率。然而，相对熵方法由于自身定义的局限性，无法用于 DoS 攻击和未知报文 ID 入侵攻击的检测，具有一定的应用局限性。本文定义了报文 ID 加权自信息量并使用了 SVDD 算法，在保留了分别计算不同报文 ID 对入侵检测敏感性的同时，突破了由相对熵定义所带来的应用局限性，对 DoS 攻击和伪装攻击都有具有较高的检测准确率。

结束语 本文提出了一种基于 SVDD 的车载 CAN 网络入侵检测方法，提取单位时间窗内 CAN 网络报文 ID 的加权自信息量和 ID 的归一化值作为特征信息，在 MEC 服务器处构建并训练 SVDD 模型。目标车辆基于下发的 SVDD 模型进行异常特征值检测，从而实现实时的车载 CAN 网络入侵检测。在仿真实验部分，本文采用韩国高丽大学 HCR 实验室公开的 CAN 网络数据集，对本文方法和 3 种传统的基于信息熵的方法在 DoS 和伪装两种攻击的检测准确率方面进行了比较，结果表明，在单组入侵报文数量较小时，本文方法在检测准确率方面有显著提高。本文工作仍存在一定的局限性，即目前只针对 DoS 和伪装两种攻击进行了分析，在未来工作中将针对更多类型的网络入侵攻击进行分析与检测。

参 考 文 献

[1] LIU Z,ZHANG T. Research on automatic lane change method based on vehicle network information[J]. Journal of Chongqing University of Technology (Natural Science),2020,34(4):11-17.

[2] CHEN L,ZHANG D,LIANG J. The Driving active service selection method based on QoS for Internet of Vehicle environment[J]. Journal of Chongqing University of Technology (Natural Science),2019,33(12):8-17.

[3] LI Y,LUO Q,LIU J,et al. TSP security in intelligent and connected vehicles: challenges and solutions [J]. IEEE Wireless Communications,2019,26(3):125-131.

[4] WU W,LI R,XIE G,et al. A survey of intrusion detection for in-vehicle networks [J]. IEEE Transactions on Intelligent Transportation System,2020,13(3):919-933.

[5] MILLER C,VALASEK C. Remote exploitation of an unaltered passenger vehicle [R]. BlackHat USA,2015.

[6] KEEN SECURITY LAB. Car hacking research: Remote attack Tesla motors [EB/OL]. [2020-09-24]. https://www.blackhat.com/docs/us-17/thursday/us-17-Nie-Free-Fall-Hacking-Tesla-From-Wireless-To-CAN-Bus.pdf.

[7] YOUNG C,ZAMBRENO J,OLUFOWOBI H,et al. Survey of automotive controller area network intrusion detection systems [J]. IEEE Design & Test,2019,36(6):48-55.

[8] CHOI W,JOO K,JO H J,et al. VoltageIDS: Low-level communication characteristics for automotive intrusion detection system[J]. IEEE Transactions on Information Forensics Security,2018,13(8):2114-2129.

[9] CHO K T,SHIN K G. Fingerprinting electronic control units for vehicle intrusion detection[C]// 25th USENIX Conference on Security Symposium. 2016:911-927.

[10] SHIN K G,CHO K T. Viden: Attacker identification on in-vehicle networks[C]// ACM SIGSAC Conference on Computer Communication Security. 2017:1109-1123.

[11] SONG H M,KIM H R,KIM H K. Intrusion detection system based on the analysis of time intervals of CAN messages for in-vehicle network[C]// International Conference on Information Networks (ICOIN). 2016:63-68.

[12] LEE H,JEONG S H,KIM H K. OTIDS: A novel intrusion detection system for in-vehicle network by using remote frame [C]// 15th IEEE PST. 2017:5709-5757.

[13] MÜTER M,ASAJ N. Entropy-based anomaly detection for in-vehicle networks[C]// IEEE Intelligent Vehicles Symposium. 2011:1110-1115.

[14] YU H,QIN G H,SUN M H,et al. Cyber security and anomaly detection method for in-vehicle CAN[J]. Journal of Jilin University (Engineering Edition),2016,46(4):1246-1253.

[15] WU W,HUANG Y,KURACHI R,et al. Sliding window optimized information entropy analysis method for intrusion detection on in-vehicle networks[J]. IEEE Access,2018,6:45233-45245.

[16] YU C,LIN B,GUI P,et al. Deployment and dimensioning of fog computing-based Internet of Vehicle infrastructure for autonomous driving[J]. IEEE Internet of Things Journal,2019,6(1):149-160.

[17] YU X,LIU Y,SHI X,et al. Mobile edge computing offloading strategy under Internet of Vehicles scenario [J]. Computer Engineering,2020,46(11):29-34,41.

[18] LING F,DUAN J,LI C,et al. Research on dynamic load balancing algorithm for C-V2X edge server[J]. Computer Engineering,2020,46(12):201-206,221.

[19] TAX M J D,DUIN P W R. Support Vector Data Description [J]. Machine Learning,2004,54:45-66.

[20] Support Vector Data Description (SVDD) Toolkit [EB/OL]. [2020-09-24]. https://github.com/iqiukp/SVDD.

YU Tian-qi, born in 1991, Ph.D, lecturer. Her main research interests include Internet of Things, edge computing and sensor data analytics.

YANG Jian-feng, born in 1978, Ph.D, senior experimentalist. His main research interests include signal processing and electronic countermeasure.