软件学报 ISSN 1000-9825, CODEN RUXUEW
Journal of Software,2017,28(Suppl.(1)):1−10

E-mail: jos@iscas.ac.cn
http://www.jos.org.cn
Tel: +86-10-62562563

基于驾驶行为和速度的车内网 CAN 数据防注入攻击*

丁 男[1,2,3], 梁文斌[1,2,3], 许 力[3], 宋彩霞[1,2], 谭国真[1,2]

[1](大连理工大学 计算机科学与技术学院,辽宁 大连 116023)

[2](辽宁省物联网与协同感知工程技术研究中心,辽宁 大连 116023)

[3](软件架构国家重点实验室(东软集团股份有限公司),辽宁 沈阳 110179)

通讯作者: 丁男, E-mail: dingnan@dlut.edu.cn

摘 要: 由于车内网的开放性以及协议缺陷,其总线中数据的安全性及有效性分析是目前亟待解决的问题.利用车内 CAN 总线网络协议中车辆速度以及刹车油门等驾驶行为信息,提出了针对车内网 CAN 网络数据的防注入攻击模型.首先,基于攻击模型的分析与注入攻击特点,构建了基于驾驶行为-速度的结构模型.其次,基于该模型,利用朴素贝叶斯网络分类器,提出了面向车内网 CAN 数据防注入攻击分析模型,从而对接收到的车内网 CAN 协议中车辆行驶速度进行了有效性分析.最后通过实验仿真与验证,其结果表明,该方法能够有效地提高数据质量分析准确度.

关键词: 车内网;CAN 总线;贝叶斯网络;数据注入;攻击模型

中文引用格式: 丁男,梁文斌,许力,宋彩霞,谭国真.基于驾驶行为和速度的车内网 CAN 数据防注入攻击.软件学报,2017,28(Suppl.(1)):1−10. http://www.jos.org.cn/1000-9825/17001.htm

英文引用格式: Ding N, Liang WB, Xu L, Song CX, Tan GZ. Analysis of malicious injection attack on CAN data in in-vehicle network based on driving behavior and velocity. Ruan Jian Xue Bao/Journal of Software, 2017,28(Suppl.(1)):1−10 (in Chinese). http://www.jos.org.cn/1000-9825/17001.htm

Analysis of Malicious Injection Attack on CAN Data in In-Vehicle Network Based on Driving Behavior and Velocity

DING Nan[1,2,3], LIANG Wen-Bin[1,2,3], XU Li[3], SONG Cai-Xia[1,2], TAN Guo-Zhen[1,2]

[1](School of Computer Science and Technology, Dalian University of Technology, Dalian 116023, China)

[2](Liaoning Engineering Technology Research Center of IoT and Cooperative Sensing, Dalian University of Technology, Dalian 116023, China)

[3](State Key Laboratory of Software Architecture (Neusoft Corporation), Shenyang 110179, China)

Abstract: Because of the opening of In-vehicle network, there are several important problems to be dealt with, such as the security and validity of data. Firstly, the article builds a construct model based on driving behavior and speed. Secondly, it makes an analysis of preventing data injection by using the construct model above and the naive Bayesian network classifier, so as to take effective measures to guarantee the vehicle security. In the end, an experimental simulation is carried out to prove that the proposed method can effectively improve the accuracy of data quality analysis and lower the false rate as well.

Key words: in-vehicle network; CAN bus; Bayesian network; data injection; attack model

随着交通事故以及交通拥塞等交通问题的急迫需求,无人驾驶、智能辅助驾驶的应用被认为是解决当前交

* 基金项目: 国家自然科学基金(61471084); 软件架构国家重点实验室开放课题基金(SKLSAOP1602); 国家高技术研究发展计划(863)(2012AA111902)

Foundation item: National Natural Science Foundation of China (61471084); State Key Laboratory of Software Architecture of Open Research Fund (SKLSAOP1602); National High Technology Research and Development Program (863) (2012AA111902)

收稿时间: 2017-05-15; 采用时间: 2017-09-23

通问题的有效方法.目前,相关研究已被业内广泛关注:福特汽车的 CEO 马克·菲尔兹(Mark Fields)日前透露,福特将在 2025 年开始正式面向所有公众出售无人驾驶汽车;日本经济产业省为推动有效利用人工智能(AI)等的“第 4 次产业革命”,将重新制定截至 2030 年的自动驾驶发展目标:使每 5 辆汽车中有 1 辆成为自动驾驶汽车,将根据反向推算来制定制度.

目前,比较流行的车内网是通过控制器局域网 CAN(controller area network)将车内各执行部件组成网络,进行数据交换.然而,目前针对车载 CAN,还没有较好的网络安全机制,因而使得车载网络极易被入侵.例如 2010 年,美国的一名汽车销售员由于对公司的不满而采取恶意的报复行动:他利用公司的汽车管理账户,随意操纵已经销售出去的 100 多辆汽车,造成这些汽车的部分功能失效[1].在 2013 年 DEFCON 的黑客大会上,美国的研究人员向在场观众展示了对一些汽车的非法远程控制,他们破解了汽车的内部操纵方式,从而可以利用电脑来操纵这些汽车.2015 年 1 月,又有黑客利用宝马公司车载系统(ConnectedDrive)的安全漏洞对其进行有效的攻击.2015 年 2 月 9 日,DARPA 研究中心的工作者指出通用安吉星 OnStar 系统存在系统安全漏洞,攻击者可以利用漏洞进行远程攻击.

为了提高数据的可靠性,保证数据的质量,传统研究的主要内容是数据清洗:包括重复对象检测、异常数据检测、数据的缺失处理等.对于异常数据检测方面,主要采用数据审计的方法来解决,通过数据概化的方式找到总体的分布特征,进而以此为基础来进行数据的有效性分析[2].随着人工智能的发展,新的技术在不断地应用到入侵检测领域,数据挖掘与机器学习在网络入侵检测方面受到广大的关注[3].

尽管有许多因素能够使车辆的速度发生改变,如上下坡、路面颠簸、摩擦因子改变、天气情况等,但它们对车速的改变具有一定的连续性,不存在尖点情况.本文主要针对速度变化存在尖点的情况,提出了一种基于贝叶斯网络与驾驶行为相结合的车载网 CAN 总线数据防注入攻击分析方法.首先,本文分析了 CAN 总线数据注入的攻击特征和威胁模型;其次,本文结合速度、驾驶行为两个方面构造贝叶斯分类器,提出了数据质量分析的方法,对车载网数据包的有效性进行了分析;最后,通过实验仿真,验证了本文所提方法的有效性.

1 相关工作

Muter 等研究人员提出了一种基于传感器检测的方法:利用各种传感器来解决 CAN 总线中的数据异常行为,将多个具有不同功能的传感器放置在 CAN 总线相应的子网中,利用传感器对数据进行有效的分析与检查,从而防止攻击者的有效攻击[4].这种检测方法具有一定的合理性与实际性,但却没有给出具体的实现方案以及系统结构,同时要实现全面的异常数据检测需要设置大量的传感器,这无疑将增加成本与维护.Larson 等人[5]介绍了一种车内网中基于规格的攻击检测方法,定义了不依赖于网络协议的车辆正常行为.他们从获取的信息中来创建可应用于基于 CANopen 3.01 通信协议和对象目录部分的通信、ECU 的行为规范,还提供了一组规格例子,提出了攻击检测器的位置,并用一组攻击行为进行检测.Kammerer 等人设计了一种可用作 CAN 总线系统中所有子网的中心网关,该网关可实现对输入数据的检查与排除,其作用相当于一个防火墙,防止异常数据进入 CAN 总线中,从而保证总线的安全性[6].Kang 等人设计了一种带有深度神经网络(DNN)的入侵检测系统以提高车内网的安全性,通过抽取车内网中的数据包来构造和训练深度神经网络的结构[7].对于任意给定的一个数据包,DNN 可以给出每种类别的正常与受到攻击的可能性,从而使传感器能够识别攻击.

贝叶斯网络用一个概率的图形模型表示变量和它们之间的关系,在处理预测以及不确定问题等方面具有较高的应用价值.Mujalli 等人[8]采用 3 种不同的数据平衡技术:欠采样、饱和采样和混合采样对数据集进行平衡.对于平衡后的数据集采用多种贝叶斯分类器(AODE、WAODE 和 BNs)进行数据分类,比较各自的分类结果,从而确定应用于不同场景下的贝叶斯分类模型.他们的研究可以用来构建预测交通事故严重程度的模型.Dipti 等人[9]采用支持向量机(SVM)和朴素贝叶斯的方法来进行交通流量的有效分类.他们使用基于统计特征的流量分类以增强特征离散化,这种分类方法能够有效地提高分类性能.

2 理论模型与分析

2.1 车内网中的CAN总线结构以及威胁模型分析

汽车电子内部的各部件控制单元,主要是由 CAN 总线相互连接起来,起到传递信息和协同的作用[10].其内部 CAN 总线结构一般如图 1 所示.

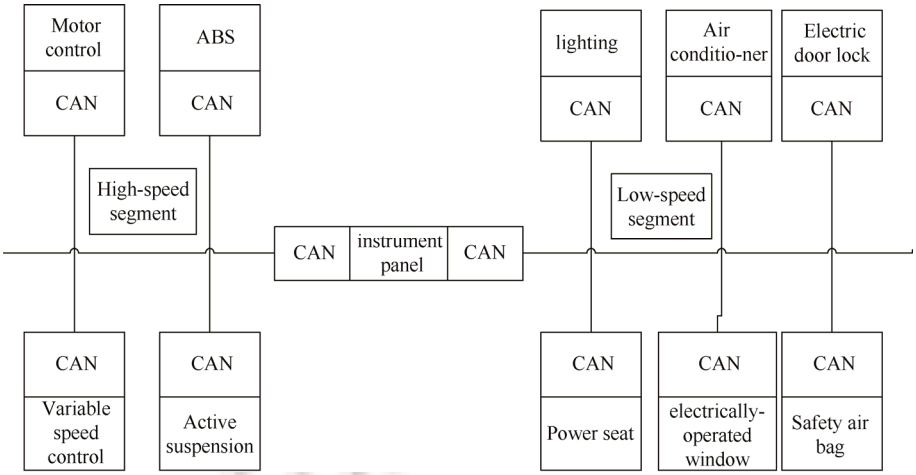


Fig.1 CAN bus architecture

图 1 CAN 总线结构

数据帧由 6 部分组成,分别为帧起始、仲裁域、控制域、数据域、检验域和帧结束.数据帧中的数据域含有本文要研究的速度数据等信息.

如图 2 所示,CAN 总线连接着各个控制操作部件,而每个部件内部都部署与之相关的 ECU,一种较为容易实现的攻击模型就是攻击者获得任何一个 ECU 的控制权,从而对其他 ECU 造成影响.可以实现的攻击类型如下.

(1) 重放攻击:在多主工作的 CAN 协议下,与总线相连的任何一个 ECU 都可以主动向总线中发送数据帧,受控的 ECU 有条件进行重放攻击;

(2) 窃听攻击:由于 CAN 协议采取广播信道进行数据传输,且不具有加密和认证过程,受控的 ECU 可以实时检测到总线上的信息帧;

(3) 伪造攻击:尽管有循环冗余校验机制对数据进行纠错,但其不具备认证信息可靠性的功能,受控的 ECU 可以向总线上发送伪造数据;

(4) 伪装攻击:CAN 协议下的任何一个 ECU 传输的数据帧中都不包含源地址和目的地址的信息,也没有认证机制,受控的 ECU 可以伪装成其他正常的 ECU 发送数据而不被察觉.

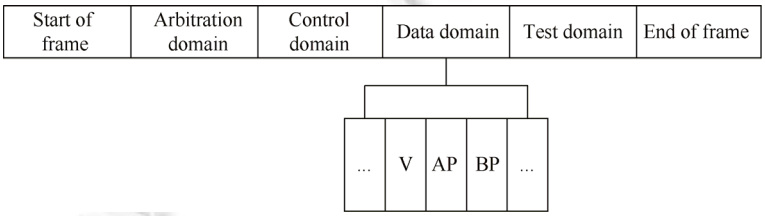


Fig.2 Data frame format

图 2 数据帧的格式

2.2 车载网络数据注入特征与分析

尽管有许多因素都能够使车辆的速度发生改变,如上下坡、路面颠簸、摩擦因子改变、天气情况等,但它们对车速的改变具有一定的连续性,不存在尖点情况.当车辆处于下坡状态时,在惯性和重力的双重作用下,车速会增加(遵循牛顿第二定律).同时,若驾驶员采取发动机制动,那么依然能够保证速度变化呈现连续性.文献[11]研究了汽车在长下坡情况下的速度控制问题,文中给出了多个不同坡度下对车辆控制的速度变化曲线.图 3 给出了在路面坡度 3%、制动初速度 60km/h、目标速 35km/h 情况下车速的变化曲线.

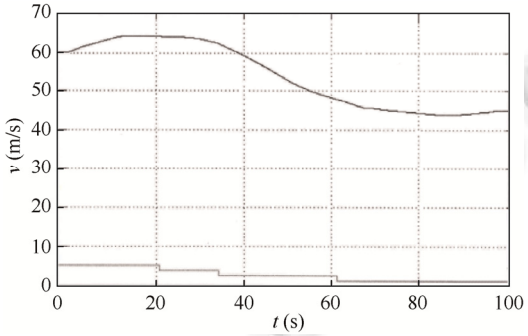


Fig.3 The change of downhill speed

图 3 下坡状态速度变化曲线

车辆行驶速度是车辆行驶状态的重要参数之一,具有交通流规律.例如,元胞自动机模型、跟驰模型等.然而,非法注入的数据无法保证其所固有的交通流规律.根据非法数据的注入形式,本文将非法数据分为两类:多重复数据和假数据.

(1) 多重复数据(multi-repeat data,简称 MRD)是直接从有效数据中复制的数据,并且重复地向车辆中注入.虽然它们看起来像有效的数据,但这些数据的值是固定不变的.

(2) 假数据(fake data,简称 FD)是攻击者伪造的数据或随机产生的数据.它们是变化的,大部分不遵循交通流的规律.

根据 MRD 的特点,可以通过与其连续的数据进行检测而发现.在本文中,主要是利用对连续 3 个时刻的速度数据是否保持一致来判断 MRD.而对于 FD,检测的方法比较复杂.

2.3 基于速度-行为的朴素贝叶斯分类器

车辆速度数据的正常与否具有一定的不确定性,而贝叶斯网络是解决不确定性问题的一种重要手段,在数据挖掘、预测等方面应用较广[12,13].

2.3.1 朴素贝叶斯分类器与 Laplace 平滑

朴素贝叶斯分类器是最简单的一种情形.其假定在给定实例类标记的前提下,实例的属性值之间是相互独立的[14].本研究使用朴素贝叶斯分类器的相关说明如下.其中,表 1 列出了需要用到的状态量信息.

Table 1 State information

表 1 状态量信息

状态量	符号	备注
速度	V	m/s
加速踏板	AP	0:未踏下, 1:踏下
减速踏板	BP	0:未踏下, 1:踏下

设每个实例 x 可由属性值的合取描述,而类标记 c 从某有限集合 C 中取值,这里,测试实例为$\langle \Delta v, AP_{(t-1)}, BP_{(t-1)} \rangle$.其中,$\Delta v$ 代表当前时刻 t 的速度与前一时刻的速度差值;$AP_{(t-1)}$表示前一时刻 t–1 情况下的加速踏板状

态;$BP_{(t-1)}$表示前一时刻 t–1 情况下的减速踏板的状态.类标记的取值集合为 C={正常,异常}.

那么, $c(x)=\arg\max P(\Delta v, AP_{(t-1)}, BP_{(t-1)}\mid c)P(c), c\in C$, 而通过机器学习可以得到各特征变量的条件概率分布.为方便描述,这里采用 a_k(k=1,2,3)来代表各个属性变量.

因此,当输入一组特征变量观测值〈Δv,$AP_{(t-1)}$,$BP_{(t-1)}$〉时,可以利用贝叶斯公式计算目标变量的后验分布,实现目标变量分类[15,16].其计算公式为

$$c(x)=\arg\max P(c)\prod_{k=1}^{3}P(a_k\mid c), c\in C \tag{1}$$

$$P(c)=\frac{\sum_{j=1}^{n}\sigma(c_j,c)}{n} \tag{2}$$

$$P(a_k\mid c)=\frac{\sum_{j=1}^{n}\sigma(a_{jk},a_k)\sigma(c_j,c)}{\sum\sigma(c_j,c)} \tag{3}$$

其中,n 为训练实例的个数;c_j 为第 j 个训练实例的类标记;a_{jk} 为第 j 个训练实例的第 k 个属性值;$\delta(c_j,c)$为一个二值函数,当 c_j=c 时其值为 1,否则为 0.a_k 为 x 的第 k 个属性值.

在大多数情况下,上述计算方法是一个比较好的估计,但当接近零频率的属性值出现的时候,就有可能出现一个有偏差的过低估计.更极端的情况是:如果某个零频率的属性值进入分类器,就有可能产生公式(3)计算的结果为 0.为避免这种问题的产生,通常可以采用 Laplace 估计来平滑上述公式.重写公式(4)和公式(5).

$$P(c)=\frac{\sum_{j=1}^{n}\sigma(c_j,c)+1}{n+q} \tag{4}$$

$$P(a_k\mid c)=\frac{\sum_{j=1}^{n}\sigma(a_{jk},a_k)\sigma(c_j,c)+1}{\sum\sigma(c_j,c)+n_k} \tag{5}$$

其中,q 为类标记的个数;n_k 为训练实例第 k 个属性的取值的个数.

虽然各变量条件独立的假设在许多应用领域未必能完全满足,但这种简化的贝叶斯分类器在许多实际应用中往往得到了较好的分类精度[17,18].

在本研究中,我们将任意时刻车速的异常与否作为类变量,将该时刻与前一时刻速度的变化量、前一时刻加速踏板与减速踏板的状态作为属性变量,构造结构模型进行下一步的研究.

2.3.2 极大后验假设 h_{MAP}

在利用朴素贝叶斯网络进行数据分类的时候,极大后验假设是其最后产生决策输出的理论依据.在本研究中,分类器的学习阶段需要发现在给定实例集 D=〈Δv,$AP_{(t-1)}$,$BP_{(t-1)}$〉的情况下,可能性最大的假设 c($c\in C$, C={正常,异常}).

极大后验假设如下表示:

$$h_{MAP}=\arg\max P(c\mid D), c\in C \tag{6}$$

应用贝叶斯规则得到:

$$h_{MAP}=\arg\max\frac{P(D\mid c)P(c)}{P(D)}, c\in C \tag{7}$$

因为 $P(D)$是一个不依赖于 c 的常量,所以方程可简化为

$$h_{MAP}=\arg\max P(D\mid c)P(c), c\in C \tag{8}$$

3 基于速度与行为的防注入攻击

由于可以利用 CAN 总线读取车辆运行状态(如速度、加速度以及加速踏板和减速踏板的位置等),从而为

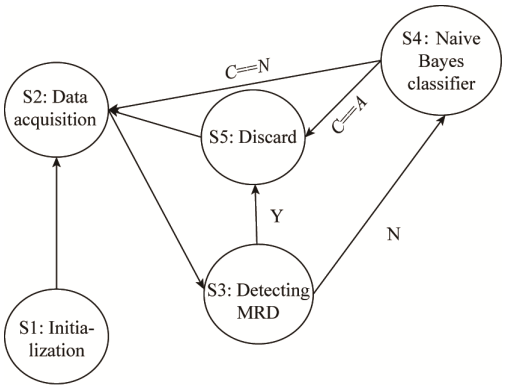


Fig.4 State machine
图 4 状态机

结合车辆的实际情况提供判断依据,准确地确定异常数据的存在.类标记的取值集合为 C={正常,异常},为方便表示,我们将正常记为 N,异常记为 A.

其状态机大致如图 4 所示.相应的执行过程如下:

S1:做好初始化工作,包括数据的接收与检测;

S2:采集来自 CAN 总线的车辆速度数据;

S3:MRD 数据检测;如果有,则进入 S5;否则,进入 S4;

S4:对原始数据采用贝叶斯网络分类器进行分类:如果分类为正常(C==N 为真),则进行下一次的数据检测(进入 S2);否则,进入 S5;

S5:认定该数据为坏数据,将其丢弃,进行下一次的数据检测(进入 S2).

4 实验及分析

4.1 实验场景的建立与相关初始工作

Sumo 仿真平台是一个微观的、空间连续的交通仿真软件.Sumo 的主要目的是给交通研究人员提供一个实现和评估自己算法的工具.在本实验中,利用 Sumo 构建了一个简单的道路文件和路径文件,通过配置文件产生仿真的实验环境,利用相关功能获取标识符为 1 的车辆在仿真期间的速度数据(记为样本 1),并对部分数据进行手动修改;采取同样的方式在不同的仿真参数下生成另外 5 个样本,用以训练数据从而构造更有效的训练集,为接下来的学习分类提供有效的支持.同时,对标识符为 2 的车辆的速度数据(记为样本 2)进行分类检测,从而判定方法的有效性和准确性.为了保证实验的稳定性,我们采取重复实验的方式进行实验仿真:按照相同的方法产生另外 11 个实验样本数据,依次进行实验,最后对各个样本的实验结果进行取平均值和方差的方式来计算总体的实验结果.这里,我们将样本分为训练样本(如样本 1)和实验样本(如样本 2)两种.

这里,我们只给出其中样本 1 和样本 2 的相关参数配置和实验过程,其余各样本实验方法与之相同(仿真参数不同).样本 1 和样本 2 的仿真参数及实验效果分别见表 2,表 2 给出了仿真环境的设置参数(包括出发时间、最大加速度、最大减速度等).

Table 2 Simulation parameters
表 2 仿真参数

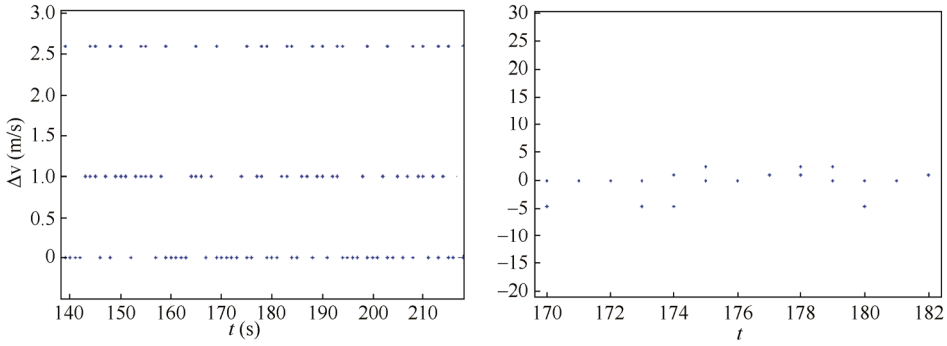
ID	Depart (出发时间:s)	Accel (最大加速度:m/s^2)	Decel (最大减速度:m/s^2)	Sigma (跟驰模型参数)	Length (车长:m)	MaxSpeed (最大速度:m/s)	备注
1	0.00	2.6	4.5	0.5	5	70	车辆 1
2	2.00	2.6	4.5	0.5	5	70	车辆 2

4.2 车辆行驶速度数据变化量离散化

按照上文提出的基于速度-行为的朴素贝叶斯分类器,测试实例为$\langle \Delta v, AP(t-1), BP(t-1)\rangle$.其中,$\Delta v$ 代表当前时刻 t 的速度与前一时刻的速度差值;$AP(t-1)$表示前一时刻 $t-1$ 情况下的加速踏板状态;$BP(t-1)$表示前一时刻 $t-1$ 情况下的减速踏板的状态.针对数据样本 1,利用 Matlab 将实际获取的Δv 进行离散化处理,以便于训练集的构造与学习分类.离散化处理的规则如下:

$$\Delta v'=\begin{cases}\text{NAN}, & \Delta v>2.7\\ 2.7, & 0.5\prec\Delta v\leqslant 2.7\\ 0.5, & 0\prec\Delta v\leqslant 0.5\\ 0, & -0.5\leqslant\Delta v\leqslant 0\\ -4.7, & -4.7\leqslant\Delta v\prec 0.5\\ \text{NAN}, & \Delta v\succ -4.7\end{cases} \tag{9}$$

其中,被离散到 1 000 和−1 000 的Δv 被认为是攻击的数据.图 5(a)所示为样本 1 的速度数据的变化量,图 5(b)所示为相应的离散化图像.



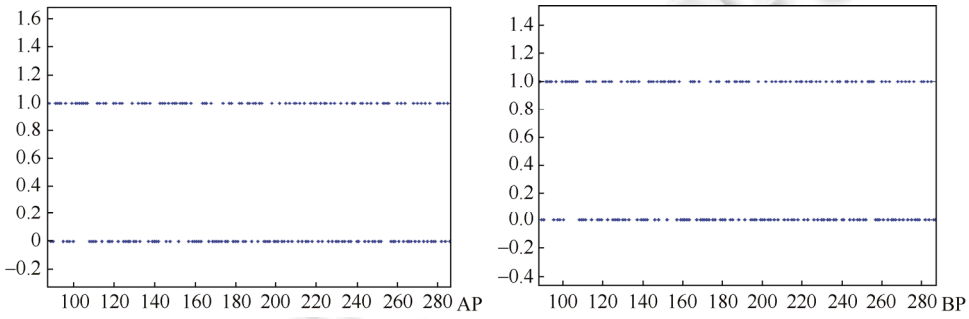
(a) 速度数据的变化量 (b) 离散化

Fig.5 Sample 1 and discrete

图 5 样本 1 与离散

4.3 训练集的构造

由于贝叶斯分类器的分类效果与性能在很大程度上取决于训练集构造的好坏,因此,训练集的构造是实验的重中之重[19].为了保证训练集的构造达到高检出率的性能,在样本 1 的 500 个数据集中(包含车辆速度、加速踏板状态和减速踏板状态)进行适当合理的修改(掺入一定的错误数据,从而可以有效地提高检出率).修改包括以下几项:(1) 一定的车辆速度;(2) 一定的车辆速度和对应的前一时刻加速踏板或减速踏板的状态;(3) 单独修改部分加速踏板状态;(4) 单独修改部分减速踏板状态.加速踏板和减速踏板都有两种状态:即 AP=1 表示加速踏板踏下,AP=0 表示加速踏板未踏下;BP=1 表示减速踏板踏下,BP=0 表示减速踏板未踏下.图 6 所示为训练集的加速踏板与减速踏板的状态,用 0、1 两个数值表征不同的状态.



(a) 加速踏板的状态 (b) 减速踏板的状态

Fig.6 The status of the pedal (0: Not down, 1: Step down)

图 6 踏板的状态(0:未踏下,1:踏下)

4.4 执行检测算法及性能分析

对于 MRD 数据,检测的方法为:通过比较连续 3 个时刻下,车辆速度的变化情况来判断是否存在多重复数据.在处理多重复数据之后,对每一时刻的车速数据,按照所设计的分类器的属性变量的要求,进行相应的计算以进行分类检测.

为了验证本研究方法的有效性,本文参考文献[20]中采用的方法,对相同的数据集(实验样本)进行分析比较.文献[20]中采用的是基于时间序列数据有效阈值比较法对数据的有效性进行分析.相对于本文的实验环境设置:实验仿真车辆的最大加速度和最大减速度分别为 2.6m/s^2 和–4.5m/s^2,设置相应的阈值来判断数据的正确性.我们将实验样本收集到的数据作为测量数据,计算相邻时间速度的变化量.采取与上述相同的数据集作为检测数据,实验结果如下:在修改的 100 组数据中,被检测出来有误的为 60 个.该方法只是针对数据的变化情况来进行判断,没有实际考虑到驾驶员的驾驶行为(急加速与急减速)对车辆速度变化的影响.同时,由于只是考虑速度的变化是否处于合理的范围之内,也就无法识别多重复数据的存在.而采用本文的方法,可以很好地检测出所有修改的 MRD 数据,同时,基于贝叶斯网络分类器的速度-行为决策模型可以有效地检测出不符合实际情况的数据,提高检测效率.

单个实验样本情况下,两种方法的检测结果见表 3.

Table 3 Result of experiment sample No.2

表 3 实验样本 2 的实验结果

方法	FD	MRD	检测为修改的数据
本文	80	10	90
文献[20]	60	0	60

对实验样本 2 的实验分析:对比样本 2 的检测结果,设计的分类器具有良好的分类性能,检出率达到 90%.在对应实验中,车辆 2 的最大加速度和最大减速度是 2.6m/s^2 和–4.5m/s^2.通过查看检测到的速度数据对应的时刻,在检测到错误的速度数据中,这些被修改的速度值与前一时刻速度值的变化量要么大于 2.6,要么小于–4.5,因而可以被分类器很好地检测出来.在 10 个未被检测到的时刻当中,这些对应修改的速度数据值的变化量均在[–4.5,2.6].同时,与前一时刻的踏板状态成正相关关系,很难被检测到.

所有实验样本情况下,两种方法的检测结果计算后见表 4.

Table 4 Result summary of all experiment samples

表 4 所有实验样本结果汇总

方法	实验样本组数	平均 FD 数量	FD 方差	平均 MRD 数量	MRD 方差
本文	12	78	5.67	7	0
文献[20]	12	56	11.23	0	–

由实验汇总结果可以看出:在实验样本数据相同的情况下,对于 MRD 数据,采用本文的方法能够完全将其检测出来;而文献[20]中 MRD 数据无法检测到;对于 FD 数据,本文方法的平均数量为 78,方差相对文献[20]较小.当实验数据的变化量处于各个实验设置的最小加速度与最大加速度之间时,很难被检测到.

5 总结与展望

采用贝叶斯网络的方法,为实现判断数据异常与否提供了可能.通过适当地修改原始数据集来构造训练集,以便于训练集能够有效地检测出异常.引入 Laplace 平滑估计之后,为了减少错误提供了有力依据.综合考虑速度变化量、加速踏板以及减速踏板的状态,将它们作为属性变量进行分类器的评估,从而更加有效地提高了检测水平.通过实验可以看出,本实验的设计方法能够达到比较好的检测效率.根据本实验的数据检测模型和检测算法,可以更加准确地识别异常数据,提高可靠性,为判断攻击行为提供一种有效的检测方法,为之后的应对措施提供判定条件,如可对该数据包的标识符进行锁定和屏蔽,以防止其进一步的攻击.同时,在考虑到驾驶员驾

驶行为的基础上,可以进一步提高数据的质量.但整个实验还有待改进:比如现在考虑的只是单个车辆的情况,需要综合考虑周围的车辆才更有意义.

References:

[1] Yan C, Xu WY. Thought of the security of vehicle intelligentization. China Computer Federation, 2016,12(1): 20−26 (in Chinese with English abstract).

[2] Buczak AL, Guven E. A survey of data mining and machine learning methods for cyber security intrusion detection. IEEE Communications Surveys & Tutorials, 2016,18(2):1153−1176.

[3] Chandola V, Banerjee A, Kumar V. Anomaly detection: A survey. ACM Computing Surveys, 2009,41(3):75−79.

[4] Muter M, Groll A, Freiling FC. A structured approach to anomaly detection for invehicle networks. In: Proc. of the 6th Int'l Conf. on Information Assurance and Security (IAS). IEEE, 2010. 92−98.

[5] Larson UE, Nilsson DK, Jonsson E. An approach to specification-based attack detection for invehicle networks. In: Proc. of the Intelligent Vehicles Symp. IEEE, 2008. 220−228.

[6] Kammerer R, Fromel B, Wasicek A. Enhancing security in CAN systems using a star coupling router. In: Proc. of the 7th IEEE Int'l Symp. on Industrial Embedded Systems (SIES). IEEE, 2012. 237−246.

[7] Kang MJ, Kang JW. Intrusion detection system using deep neural network for invehicle network security. PLoS ONE, 2016, 11(6):e0155781. [doi: 10.1371/journal.pone.0155781]

[8] Mujalli RO, López G, Garach L. Bayes classifiers for imbalanced traffic accidents datasets. Accident Analysis & Prevention, 2016,88:37−51.

[9] Tiwari D, Mallick B. SVM and naïve Bayes network traffic classification using correlation information. Int'l Journal of Computer Applications, 2016,147(3):1−5.

[10] Gmbh RB. CAN Specification, Version2.0-1991. Stuttgart: Bosch. 1991. 1−73.

[11] Lu CJ. The study on vehicle speed control in the long downhill [Ph.D. Thesis]. Chongqing: Chongqing Jiaotong University, 2012 (in Chinese with English abstract).

[12] Li W, Li QX. Using naïve Bayes with AdaBoost to enhance network anomaly intrusion detection. In: Proc. of the 3rd Int'l Conf. on Intelligent Networks and Intelligent Systems (ICINIS). IEEE, 2010. 486−489.

[13] Ahirwar DK, Saxena SK, Sisodia MS. Anomaly detection by naïve Bayes & RBF network. Int'l Journal of Advanced Research in Computer Science & Electronics Engineering, 2012,1(1):22−27.

[14] Ge J, Xia Y, Wang J. A naïve Bayesian classifier in categorical uncertain data streams. In: Proc. of the Int'l Conf. on Data Science and Advanced Analytics. IEEE, 2015. 392−398.

[15] Park SH, Fürnkranz J. Efficient implementation of class-based decomposition schemes for naïve Bayes. Machine Learning, 2014,96(3):1−15.

[16] Li WJ, Xiong XF, Mao YM. Classification method for interval uncertain data based on improved naïve Bayes. Journal of Computer Applications, 2014,34(11):3268−3272 (in Chinese with English abstract).

[17] Ren J, Lee SD, Chen X, *et al*. Naïve Bayes classification of uncertain data. In: Proc. of the IEEE Int'l Conf. on Data Mining. IEEE, 2009. 944−949.

[18] Mukherjee S, Sharma N. Intrusion detection using naïve Bayes classifier with feature reduction. Procedia Technology, 2012,4(11): 119−128.

[19] Barot V, Chauhan SS, Patel B. Feature selection for modeling intrusion detection. Int'l Journal of Computer Network & Information Security, 2014,6(7):56−62.

[20] Ki Y, Baik D. Model for accurate speed measurement using double-loop detectors. IEEE Trans. on Vehiclar Technology, 2006,55:1094−1101.

附中文参考文献:

[1] 闫琛,徐文渊.汽车智能化的安全思考.中国计算机学会通讯,2016,12(1):20−26.

[11] 卢从娟.汽车长下坡速度控制研究[博士学位论文].重庆:重庆交通大学,2012.

[16] 李文进,熊小峰,毛伊敏.基于改进朴素贝叶斯的区间不确定性数据分类方法.计算机应用,2014,34(11):3268−3272.

丁男(1978－),男,辽宁沈阳人,博士,副教授, CCF 专业会员,主要研究领域为大数据分析及其应用,车联网,智能交通.

梁文斌(1993－),男,硕士,主要研究领域为车联网,数据质量分析.

许力(1982－),男,博士,高级工程师,CCF 高级会员,主要研究领域为概率图模型理论,大数据分析,云计算.

宋彩霞(1977－),女,博士,讲师,主要研究领域为车联网数据安全,交通流模型.

谭国真(1960－),男,博士,教授,博士生导师,CCF 杰出会员,主要研究领域为车联网,智能交通,人工智能.