

基于多模态信息融合的无监督 CAN 总线测绘框架

姚伟平, 李清宝, 陈志锋, 焦卫华

(信息工程大学, 河南 郑州 450001)

摘要: 针对黑盒环境下控制器局域网(CAN)总线拓扑测绘的挑战,提出一种基于多模态信息融合的无监督测绘模型(3Mapping)。该方法通过 JointVEM 多模态特征工程方法,同步采集电压与电磁信号,提取上升沿和下降沿的电压微分峰值、电磁峰值及其时间间隔等多维特征,有效融合发动机控制器(ECU)的硬件特征和位置信息,显著增强特征的区分度和鲁棒性。引入基于隔离森林(IF)的异常数据清洗机制,有效剔除噪声样本;结合变分自编码器(VAE)进行潜在空间特征学习,并采用高斯混合模型(GMM)实现无监督聚类。实验结果表明,JointVEM 多模态特征方法能有效提取节点硬件特征,聚类分析轮廓系数均超过 0.95, Davies-Bouldin(DB)指数低于 0.2, 多项指标上均优于对比方法,能准确完成 CAN 节点聚类和总线拓扑测绘,为后续安全检测提供可靠真实的有效标签。

关键词: CAN 总线测绘;多模态信息融合;物理层特征;聚类分析;异常检测;车载网络安全

中图分类号: TP393.08

文献标识码: A

文章编号: 1671-0673(2026)01-0088-09

3Mapping: Unsupervised CAN Bus Mapping via Multi-Modal Information Fusion

YAO Weiping, LI Qingbao, CHEN Zhifeng, JIAO Weihua

(Information Engineering University, Zhengzhou 450001, China)

Abstract: The challenge of controller area network (CAN) bus topology mapping in black-box environments is addressed in this paper through the proposal of 3Mapping, an unsupervised mapping model based on multi-modal information fusion. The proposed JointVEM multi-modal feature engineering method is used to collect voltage and electromagnetic signals synchronously and to extract a multi-dimensional feature set comprising voltage derivative peaks, electromagnetic peaks, and their temporal intervals from both rising and falling edges. Engine control unit (ECU) hardware characteristics and positional information are effectively integrated, significantly enhancing feature discriminability and robustness. An anomaly data cleaning mechanism based on isolation forest (IF) is introduced to efficiently remove noisy samples. Feature learning in a latent space is implemented using a variational autoencoder (VAE), and unsupervised clustering is performed via a Gaussian mixture model (GMM). Experimental results demonstrate that the JointVEM multi-modal feature method effectively captures node hardware characteristics, achieving silhouette scores above 0.95 and Davies-Bouldin (DB) indices below 0.2. CAN node clustering and bus topology mapping are accurately completed, providing reliable and authentic labels for subsequent security detection.

Key words: CAN bus mapping; multi-modal information fusion; physical layer features; cluster analysis; anomaly detection; in-vehicle network security

智能网联汽车(Intelligent Connected Vehicle, ICV)是指通过搭载先进传感器等装置,运用人工智能等新技术,具有自动驾驶功能,逐步成为智能移动空间和应用终端的新一代汽车^[1]。在智能网联汽车车载网络中,尽管车载以太网早已提出,但控制器局域网(Controller Area Network, CAN)总线依然是车内网络不可或缺、难以被完全替代的核心基础网络。CAN总线的固有脆弱性使其成为汽车网络安全的焦点,已有研究提出各种轻量级的加密算法与密钥管理方案加固CAN总线,如消息载荷加密^[2-3]和应用层数据认证^[4-7]。由于CAN协议本身限制及修改协议和计算代价因素,基于密码学的加固方案在实践中面临困难。

相比之下,异常检测是更加可行的选择^[8]。任何有效的异常检测,其逻辑起点是建立一个精确可靠的正常行为基准。在针对CAN总线安全的异常检测中,基准包括总线拓扑结构和正常通信模式,具体包括总线中存在哪些节点,不同ID的消息由哪个节点发送,消息周期、依赖关系和触发条件是什么。这些信息通常保存在DBC文件中^[9],但DBC文件常被汽车厂商视为机密,安全分析人员难以获得。在对汽车的安全审计、渗透测试,以及对供应链中潜在的硬件木马检测等场景中,往往处于一个黑盒环境。安全分析人员如何在不依赖厂商DBC文件的黑盒环境下,对CAN网络进行拓扑测绘,发现网络中节点、确定消息ID与发送节点的映射关系,是构建任何高级安全检测能力的基础。

近年来,很多学者提出针对CAN总线拓扑测绘的研究,和雷达辐射源个体识别类似^[10],通过提取节点硬件指纹来区分不同节点。如Schell等^[11]提出VALID方法,通过对每条报文采样1~2个显性位电平,并用轻量级递归最小二乘法累积和算法对发动机控制器(Engine Control Unit, ECU)电压特征画像,但该方法的区分能力有限。文献[12]基于时间间隔特征,提出CANvas测绘方法,但对于发送模式相似或由同一定时器触发的不同ECU,其区分能力有限,且容易受到攻击者模仿通信模式的欺骗。

在CANvas的基础上,Popa等^[13]进一步利用报文本身的特征来增强节点指纹的独特性,这种方法比单纯依赖时序更为精确,但对于信号变化不频繁或信号值范围重叠的ECU,仍然存在混淆的可能。其所依赖的特征,特别是时间间隔,容易被

足够恶意攻击者模拟,因此在面对旨在规避检测的高级攻击时,鲁棒性不足。

由于不同节点时间间隔或电压特征可能重叠,基于信号传输时延差的测绘方法进入研究者的视野,这类基于信号到达时间差(Time Difference of Arrival, TDoA)的方法,能够直接区分出不同节点的物理位置,提供了更底层、更稳定的区分特征。此类方法的典型代表是TIDAL-CAN(Timing based Intrusion Detection and Localization for Controller Area Network)^[14],通过测量同一报文信号到达总线末端两个监控点的差分传播时间来识别不同发送节点。这种方法有效消除了发送时间抖动带来的影响,能够以较高的准确率区分不同位置的节点。其优势在于对替换和插入攻击具有抵抗力。

为追求更高的定位精度,CAN-SQUARE^[15]在TIDAL-CAN的基础上,聚焦于单个显性位的上升沿时间差。通过精密的信号处理和计算,该方法声称可实现分米级的定位精度。虽然这些方法最初目的是用于区分恶意节点,但能有效区分不同位置节点发送的消息,因而也具备测绘总线拓扑的能力。由于该类方法需要在总线两端或更多位置部署高精度时间同步监控器,增加了系统复杂性和成本。而且在实际黑盒环境中往往无法确定总线的端点位置,也无法准确确定信号上升沿的阈值。

针对上述问题,提出一种基于多模态融合特征的无监督CAN总线测绘框架(Unsupervised CAN Bus Mapping via Multi-Modal Fusion, 3Mapping),对未知的黑盒CAN总线拓扑进行测绘。实验结果表明,基于多模态融合特征的测绘模型能有效识别不同ECU发送的CAN消息,对消息实现准确聚类,并找到最优聚类数,实现“逻辑节点—消息ID”的有效映射,完成对网络拓扑的测绘。测绘结果能作为监督信号,为样本数据提供真实标签,有效支撑后续检测模型的训练与预测。

1 3Mapping框架设计

3Mapping系统框架图如图1所示,分为信号采集与特征提取、数据处理与聚类两个阶段,主要由多模态信号采集、特征提取与融合、异常数据清洗、潜在空间特征学习和无监督聚类5个模块组成。

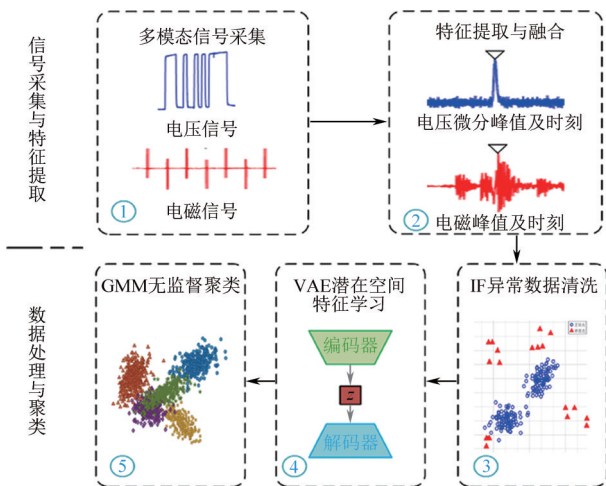


图1 系统框架图

1.1 多模态信号采集

为实现高精度的物理层特征提取,使用示波器和电压、近场电磁探头,确保获取高质量的原始信号数据,本方案适用于实验室场景下针对车辆开展深度安全分析与取证。

1.1.1 电压信号采集

采用高精度差分探头测量总线 CAN_H 与 CAN_L 差分电压,检测点可选择靠近疑似 ECU 的连接位置,以最小化总线反射对信号完整性的影响。为消除共模噪声,采集系统配备高质量的屏蔽电缆。

1.1.2 电磁信号采集

使用近场磁场探头在目标 ECU 或总线电缆附近进行非接触测量。探头定位策略基于电磁场强度扫描,通过移动探头找到信号最强的测量点。考虑到不同 ECU 的电磁辐射特性差异,设置较高的采集频带,以捕获完整的电磁特征。

1.1.3 同步触发机制

为确保电压与电磁信号的时间对齐,采用 CAN 报文 SOF 位的上升沿作为硬件触发信号。通过共享同一触发源,保证两种模态信号采集的严格同步。由于实车 CAN 总线负载一般控制在 30% 左右,使用触发采样也能最大化利用示波器采样深度。原始信号需经过位分离、帧解码和位选择等系统预处理以提取有效信息。

1.2 特征提取与融合

1.2.1 电压波形微分特征提取

CAN 总线差分电压范围为 0~2 V,随着采样速度的提高,相邻采样点之间的电压相近,波形相对平缓。通过对电压信号求一阶导数,得到电压变化率,该变化率直接反映了信号跳变的陡峭程度,间接反映了 ECU 驱动器的驱动速度,与内部晶体

管开关速度、电压源驱动能力紧密相关。

上升沿、下降沿电压微分峰值分别对应电压上升和下降最快的瞬时速率,微分峰值对应时刻标志着驱动器开关动作的瞬时效果达到最大。

1.2.2 近场电磁辐射信号特征提取

当使用近场电磁探头靠近总线测量点 P 处的电缆时,其输出信号 $V_{EM}(t)$ 正比于磁场变化率 dH/dt ,进而正比于总线电缆上的电流变化率 dI_{bus}/dt 。因此,电磁辐射峰值时刻 t_{EM} 与电流变化率 di/dt 的峰值时刻完全同步,可以通过电流模型间接推导。

由于电磁场强度正比于电流变化率,通过对电流信号求一阶导数,得到电流变化率,反映节点输出信号驱动强度和内部电感对电流的影响,其中瞬态电流峰值对应电流变化率最大的时刻。

1.2.3 多模态信号融合

在理想无损传输线中,电压与电流的变化率严格成正比,电压和电流变化趋势应该完全同步,相位差为 0,即电压微分峰值与电磁信号峰值应该同步。但在实际 CAN 总线电路中,由于传输线损耗、参数频率依赖性、阻抗失配等因素,电压与电流的变化存在显著的相位差,导致电压微分峰值与电磁信号峰值时刻常出现错位。

主要影响因素包括以下 6 个方面:1) ECU 发射源晶体管开关速度决定信号边沿的陡峭度,进而影响峰值大小。2) ECU 发射源的输出阻抗 Z_s 和终端阻抗 R_T 失配会产生反射波,而电压与电流的反射系数存在极性差异,导致叠加后的波形峰值时刻不同。3) 实际显性位信号上升沿和下降沿并非理想阶跃,且上升沿与下降沿的非对称性会进一步加剧电压与电流的不同步。4) 不同器件和 PCB 布线存在的寄生电容和寄生电感会引入额外的延迟,且对电压和电流的影响不对称。寄生电容与传输线并联,因为电容充电会延缓电压变化,但因为分流电流会增大电流峰值但不显著改变其时序。寄生电感与传输线串联,会因为电感阻碍电流突变而延缓电流变化,同时因为分压电压而增大电压峰值,但时序变化较小。5) 传输线损耗的频率依赖性影响电压和电流的相位。在高频率下,由于集肤效应,导体电流集中于表面,单位长度导线电阻 $R_0(f)$ 随频率平方根增长,导线损耗的引入导致电压和电流的相位常数出现微小差异。例如, CAN 总线双绞线工作在 1 MHz 频率时,导体损耗可使电压相位超前电流约 $2^\circ \sim 5^\circ$,在 10 m 传输距离可导致峰值时刻错开 1~3 ns。6) 传输线分布电

感 $L_0(f)$ 和分布电容 $C_0(f)$ 对频率的依赖性,使得特性阻抗 $Z_0(f) = \sqrt{L_0(f)/C_0(f)}$ 随频率变化,破坏 $V = Z_0 I$ 的恒定关系,其中 I 为传输线上的电流,导致不同频率分量的电压与电流比例不同步。由于上升沿和下降沿包含丰富的高频成分,不同频率分量的 $Z_0(f)$ 差异会使电压和电流的波形畸变不同步,最终导致峰值时刻错位。

显性位上升沿和下降沿对应的电压微分峰值分别记为 $V_{\text{peak_rise}}$ 和 $V_{\text{peak_fall}}$,电磁峰值为 $E_{\text{peak_rise}}$ 和 $E_{\text{peak_fall}}$ 。

定义显性位上升沿和下降沿对应的电压微分峰值与电磁峰值之间的时间间隔如下:

$$\Delta t_{\text{rise}} = t_{V_{\text{peak_rise}}} - t_{E_{\text{peak_rise}}} \quad (1)$$

$$\Delta t_{\text{fall}} = t_{V_{\text{peak_fall}}} - t_{E_{\text{peak_fall}}} \quad (2)$$

通过分析可知,随着总线节点收发器的快速切换,信号在遇到节点连接点、终端电阻等阻抗不连续点处会发生反射。反射波与原始信号叠加,引发过冲、下冲和振铃,总线电压微分峰值、电流峰值及其时刻不仅与节点本身晶体管开关速度、阻抗、电感等物理特征紧密相关,也受节点位置的影响。通过采集电压和电磁两个模态的信号,分别计算显性位上升沿和下降沿电压微分峰值、电磁信号峰值和峰值之间的时间间隔,可以作为节点硬件特征和节点位置相关的指纹,用以区分不同节点。

最终,融合后的联合电压—电磁特征向量JointVEM为

$$J_{V_EM} = [V_{\text{peak_rise}}, V_{\text{peak_fall}}, E_{\text{peak_rise}}, E_{\text{peak_fall}}, \Delta t_{\text{rise}}, \Delta t_{\text{fall}}] \quad (3)$$

1.3 IF异常数据清洗

在理想模型中,表征信号特征的峰值时间间隔向量,其各个维度的数据应遵循正态分布。但在实际环境中,由于CAN总线采用双绞线,其在有效抑制差模辐射的同时,使得电缆的电磁辐射信号极为微弱。周围辐射源产生的信号可能耦合进测量系统,污染微弱的CAN总线电磁信号。特别是在实际车载环境中,由于存在点火系统、发电机等众多强电磁辐射源,使得这种现象更加严重。这种低信噪比的工况,导致使用电磁探头测量时,采集到的信号极易受环境电磁噪声与测量系统本底噪声的污染,引发峰值时刻定位的漂移与畸变。如图2(b)所示,ID 0x21消息的峰值间隔基本遵循正态分布,但图2(a)所示的ID 0x11实测数据常呈现多峰分布并夹杂离群点,严重偏离正态假设,因此简单地基于正态分布对原始数据进行异常检测的方法在实际中并不适用。

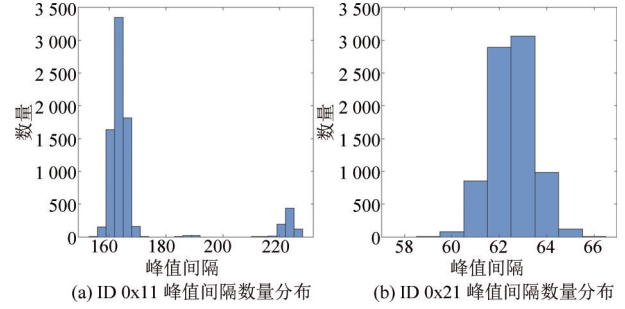


图2 原始信号分布

隔离森林(Isolation Forest, IF)是一种高效且适用于大规模数据的无监督异常检测算法,该方法不依赖于对数据分布的预先假设,转而利用异常数据本身所具有的“稀少且不同”的两个本质特征来对其进行识别。通过在高维空间中有效识别与主体数据模式显著偏离的异常样本,较好地适应本场景下数据的复杂分布特性。

隔离森林算法的目标是通过集成学习的方式,构建 t 棵隔离树,形成一个森林。每棵隔离树的构建是一个递归的随机分割过程。

异常检测的关键在于如何量化一个样本的孤立程度。在隔离森林中,其通过计算异常分数来实现。通过标准化平均路径长度确定异常分数,计算出所有样本的异常分数 $s(x_i)$ 后,核心任务在于设定一个科学、客观的判定阈值,以明确区分正常样本与异常样本。

由于无监督学习中缺乏先验的类别标签,阈值的确定通常依赖于业务理解与数据分布的统计特性。其中,基于污染系数的判定方法因其概念直观、计算稳健而被广泛采用。该方法的核心是预设一个污染系数 α (通常 $0 < \alpha \leq 0.5$),它代表了分析者对于数据集中异常值所占比例的先验估计。此参数直接控制了模型对异常的敏感程度: α 值设置得越高,被认为异常的样本就越多,模型越敏感;反之则越保守。

基于此系数,异常判定阈值可通过以下公式确定:

$$N_{\text{threshold}} = Q_{1-\alpha}(s(X)) \quad (4)$$

式中: α 为预设的污染系数,表示数据集中预期异常值所占的比例; $Q_{1-\alpha}$ 表示所有样本异常分数集合 $s(X) = \{s(x_1), s(x_2), \dots, s(x_n)\}$ 的 $(1 - \alpha)$ 分位数。所有满足 $s(x_i) > N_{\text{threshold}}$ 的样本 x_i 被判定为异常值,在清洗过程中予以剔除。

1.4 VAE潜在空间特征学习

变分自编码器(Variational Autoencoder, VAE)^[16]

作为一种强大的深度生成模型,与主成分分析(Principal Component Analysis, PCA)等线性方法不同,在降维与特征学习方面,VAE基于概率图模型的深度生成模型,自动学习建立从隐变量 z 到观测数据 x 的生成过程,能够学习数据中复杂的非线性映射。与标准自编码器(Autoencoder, AE)相比,VAE通过引入概率编码器,强制潜在空间遵循一个先验分布,从而学习到连续、结构化且具有解释性的低维特征表示。

VAE假设每个观测数据点 x 都是由一个潜在的、不可观测的随机变量 z 通过以下过程生成模型:

$$z \sim p(z), x \sim p_\theta(x|z) \quad (5)$$

式中: $p(z)$ 是先验分布,通常选择易于采样的标准正态分布 $\mathcal{N}(0, I)$; $p_\theta(x|z)$ 是条件似然函数,由解码器(Decoder)神经网络参数化; θ 为解码器参数。

该生成过程首先从先验分布中采样隐变量 z ,然后通过生成网络 $p_\theta(x|z)$ 产生观测数据 x 。

VAE的训练过程通过重参数化技巧实现端到端的梯度反向传播,使得证据下界可以通过随机梯度下降方法有效优化。

为了在严格的资源限制下实现有效的异常检测,3Mapping框架采用经典VAE的显式分离架构,即编码器和解码器是两个由不同参数构成的独立网络,编码器专注于将输入数据映射为潜在空间分布,解码器专注于将潜在变量重建为目标输出。

1.5 GMM无监督聚类

通过VAE捕捉到输入数据在正常工况下的本质结构与潜在特征分布模式之后,可以通过对这些潜在特征进行聚类分析,来精确刻画原始数据的多元分布轮廓,通过信号聚类实现节点映射,从而为后续的异常检测建立参考基准。

1.5.1 GMM聚类模型

信号及学习到潜在空间特征可能呈现多峰分布,传统聚类算法对异常值敏感且难以处理复杂的数据分布,采用基于改进高斯混合模型(Gaussian Mixture Model, GMM)的鲁棒聚类方法,显著提升了聚类性能的稳定性和可靠性。

高斯混合模型假设观测数据由 k 个高斯分布组件混合生成,其概率密度函数可表示为

$$p(\mathbf{x}) = \sum_{k=1}^K \lambda_k \mathcal{N}(\mathbf{x} | \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) \quad (6)$$

式中: λ_k 为混合权重,满足 $\sum_{k=1}^K \lambda_k = 1$; $\boldsymbol{\mu}_k$ 和 $\boldsymbol{\Sigma}_k$ 分别表示第 k 个高斯组件的均值向量和协方差矩阵。

使用EM算法来估计参数,EM算法分为两个

步骤:期望步骤(E步骤)和最大化步骤(M步骤)。在E步骤中,计算每个数据点 x_n 属于第 k 个聚类的概率 $\gamma(z_{nk})$ 。计算方法是用第 k 个高斯分布的概率乘以它的权重,然后除以所有高斯分布的加权概率之和。简单来说,就是计算每个数据点“属于每个聚类的可能性有多大”。在M步骤中,基于E步骤计算出的概率,更新3个参数:1)更新均值,其中新的聚类中心 $\boldsymbol{\mu}_k^{\text{new}}$ 等于所有数据点的加权平均,权重就是它们属于这个聚类的概率;2)更新协方差,新的协方差矩阵 $\boldsymbol{\Sigma}_k^{\text{new}}$ 反映了数据点围绕新中心的分散程度,同样用概率作为权重;3)更新权重,新的混合权重 λ_k^{new} 等于属于这个聚类的总概率除以总数据点数,表示这个聚类在整体中的重要性。其中总数据点数 $N_k = \sum_{n=1}^N \gamma(z_{nk})$ 表示属于第 k 个聚类的有效数据点数量。

由于传统GMM对异常值敏感,采取基于后验概率的异常值识别方法进一步排除异常,有:

$$\text{outlier_score}(\mathbf{x}_i) = 1 - \max_k \gamma(z_{ik}) \quad (7)$$

该分数反映了数据点 $\mathbf{x}_i$ 属于任意聚类的确定性程度。分数越高,表明该点与所有聚类中心的匹配度越低,越可能是异常值。

1.5.2 最优聚类数选择

在聚类分析中,聚类数量 k 的选择是决定聚类结果合理性与实用价值的核心前提。若 k 取值过小,会导致不同类别数据被强行合并,簇内数据异质性过高,无法反映数据的真实分布结构;若 k 取值过大,将同一类别的数据拆分为多个冗余簇,引入虚假的类别差异。

传统最优 k 选择方法普遍存在局限性:一方面,单一评估指标的片面性显著。例如,组内平方和仅衡量簇内数据的紧凑性,随 k 增大持续下降,难以确定肘部位置的客观阈值;轮廓系数虽能平衡簇内紧凑性与簇间分离度,但对密度不均的数据鲁棒性差。另一方面,聚类稳定性常被忽略。例如, k -means聚类算法高度依赖初始聚类中心的随机选择,特别是在数据类别边界模糊时,随机初始化带来的稳定性问题会进一步放大。

为解决上述问题,实现对最优聚类 k 的全面、可靠评估,使用一种多指标加权评分机制。首先在合理的 k 取值范围内,对每个候选 k 多次运行 k -means以评估聚类稳定性;其次,同步计算覆盖“簇内紧凑性”(WCSS)、“簇间分离度—簇内紧凑性平衡”(轮廓系数)、“簇间—簇内差异比”(DB指数、

CH 指数)的多维度指标;最后,通过加权融合各指标得分,具体权重可根据任务对“紧凑性”“分离度”“稳定性”的侧重灵活调整,以综合得分最高的 k 值作为最优聚类数,从而兼顾聚类结果的有效性、合理性与稳定性。

2 实验评估

为了全面评估所提出的无监督特征学习与分类框架的有效性,本节先介绍实验设置,并从特征工程有效性、潜在空间特征学习能力评估以及整体方法对比实验 3 个方面展开详细评估。

2.1 实验设置

只有使用真实节点标签作为参照,才能有效验证所提方法的有效性。为此构建了一套包含 6 个节点的 CAN 总线原型系统,核心目标是模拟真实 CAN 总线环境,人为控制节点数量和消息 ID、数据内容、发送周期等规则,生成可追溯的真实节点标签,用于评估算法性能。

原型系统每个节点均由 Arduino UNO R4 Minima 开发板^[17]与 CAN 收发器组成。其中,开发板搭载 Renesas RA4M1 处理器^[18],集成 CAN 控制器,可直接支持 CAN 总线通信功能

为了丰富总线电压信号的特征多样性,以模拟实际车载环境中不同硬件配置带来的信号差异,选用了 NXP TJA1050T^[19]和 TI SN65HVD230^[20]两种常见的不同型号 CAN 收发器。为了引入不同总线距离带来的信号传输差异,节点之间通过 5 段不同长度的 AWG26 规格双绞线连接。各节点的消息 ID、发送周期等具体参数如表 1 所示。系统通信速率设置参考了典型车载 CAN 总线标准,设置为 500 kbps。

在总线上选定 5 个不同的测量点,对电压信号和电磁辐射信号进行联合采样。受限于示波器存储深度,通过电平触发方式在每个测量点各采集 1 600 个 CAN 消息,通过数据预处理和位选择,筛选出其他非冲突位显性位信号具体数据集,情况如表 2 所示。

表 1 CAN 原型节点参数设置			
节点	收发器型号	ID	消息周期/ms
1	SN65HVD230	0x11	20
		0x12	25
2	TJA1050T	0x21	20
		0x22	50
		0x23	40
		0x31	125
3	TJA1050T	0x32	50
		0x33	40
		0x34	20
		0x41	20
4	TJA1050T	0x42	80
		0x51	20
5	SN65HVD230	0x52	25
6	SN65HVD230	侦听节点	

表 2 电压—电磁特征向量数据集			
数据集	有效帧数量	显性位数量	非冲突位样本数量
d_1	1 571	62 330	56 197
d_2	1 555	61 715	55 629
d_3	1 593	63 189	56 971
d_4	1 591	63 123	56 910
d_5	1 589	63 033	56 830

2.2 多模态物理层特征工程有效性评估

为验证 JointVEM 多模态物理层特征工程的有效性,从已有基于电压模态特征工程方法中,选取 VALID、IdentifierIDS^[21]、BTMonitor^[22]3 种特征工程方法作为基线进行对比验证实验。为确保实验的公平性和可比性,所有分类器均在相同条件下进行测试,确保观察到的性能差异主要源于特征工程方法本身,而非分类器参数或其他外部因素。从 d_5 数据集中选取 6 500 个样本作为测试数据,并根据不同特征工程方法分别提取特征。

2.2.1 分类性能评估

表 3 对比了采用 4 种不同特征工程方法提取数据集特征后,支持向量机(Support Vector Machine, SVM)、随机森林(Random Forest, RF)和 k 近邻(k -Nearest Neighbors, k -NN)3 种具有代表性分类器上的性能结果。

表 3 不同特征工程方法的分类性能比较						
特征工程	SVM 准确率	SVM F1	RF 准确率	RF F1	k -NN 准确率	k -NN F1
VALID	0.788±0.004	0.811±0.004	0.750±0.014	0.775±0.011	0.781±0.008	0.803±0.010
IdentifierIDS	0.980±0.003	0.980±0.003	0.997±0.001	0.997±0.001	1.000±0.000	1.000±0.000
BTMonitor	0.579±0.042	0.538±0.039	0.704±0.042	0.673±0.051	0.633±0.038	0.597±0.057
JointVEM	0.984±0.001	0.984±0.001	1.000±0.000	1.000±0.000	1.000±0.000	1.000±0.000

JointVEM多模态特征工程方法在所有分类器评估指标上均表现出最优或接近最优的性能。在SVM分类器上达到 0.984 ± 0.001 的准确率和 $F1$ 分数,在RF和 k -NN分类器上均实现了完美的100%的分类性能。这一卓越表现验证了JointVEM特征工程的有效性,通过综合多种特征提取技术的优势,实现了对节点硬件特征的精准识别。该方法极低的标准差进一步证明了其出色的稳定性和可靠性。

2.2.2 聚类性能评估

为保障实验的严谨性与结果的可靠性,聚类分析过程采用 k -means算法执行。具体步骤为:首先,结合数据真实类别分布特征与实验目标需求,将聚类数量设定为真实节点标签数量($k = 5$);其次,通过10次重复实验降低随机误差干扰;最后,针对10次重复实验得到的调整兰德指数(Adjusted Rand Index, ARI)、标准化互信息(Normalized Mutual Information, NMI)、轮廓系数、同质性及完整性等核心聚类质量指标,分别计算算术平均值。最终形成不同特征工程方法的聚类质量比较结果,如表4所示,以此客观量化不同特征工程方法对聚类效果提升作用。

表4 不同特征工程方法的聚类质量比较

方法	ARI	NMI	轮廓系数	同质性	完整性
VALID	0.363	0.712	0.786	0.754	0.667
IdentifierIDS	-0.031	0.449	0.520	0.535	0.352
BTMonitor	-0.603	0.015	0.666	0.194	-0.202
JointVEM	0.598	0.767	0.535	0.768	0.766

基于表4的多维度评估结果,4类特征工程方法的聚类表征能力呈现显著差异。不同特征工程方法的聚类表现出不同质量和特点,其中JointVEM在捕捉数据真实类别结构、提升聚类质量方面表现出最优有效性,主要表现为如下两点。

1) 聚类与真实标签的一致性最优,ARI、NMI是反映“聚类—真实标签匹配度”的核心指标,JointVEM均达峰值,其中ARI为0.598,不仅是唯一超过0.5的方法,更是远高于VALID(0.363)、IdentifierIDS(-0.031)、BTMonitor(-0.603)。结合ARI的取值范围 $[-1, 1]$,表明JointVEM的聚类结果与真实标签的一致性显著优于随机水平。

2) NMI为基于信息论衡量聚类与真实标签的信息共享程度,JointVEM该值为0.767,是4类方法中最高值,较次优的VALID(0.712)提升约7.7%,说明JointVEM提取的特征能高效传递“真实类别→聚类结果”的信息,避免类别信息在特征转化中丢失。

2.2.3 特征空间可视化分析

为直观呈现高维特征中“同类样本聚集、异类

样本分离”的空间分布特性,通过高维特征降维可视化与类间可分性量化对比相结合的方式,构建特征空间可视化分析体系,借助t-SNE(t-distributed Stochastic Neighbor Embedding)算法将高维特征映射至二维平面,在保留样本局部邻域关系的前提下直观呈现簇分布结构,具体结果如图3所示。由于VALID算法只有一维特征,在此不做对比,采用JointVEM特征提取方法得到的特征空间中,不同类别的样本形成了更清晰的簇结构,类间重叠区域明显减少。

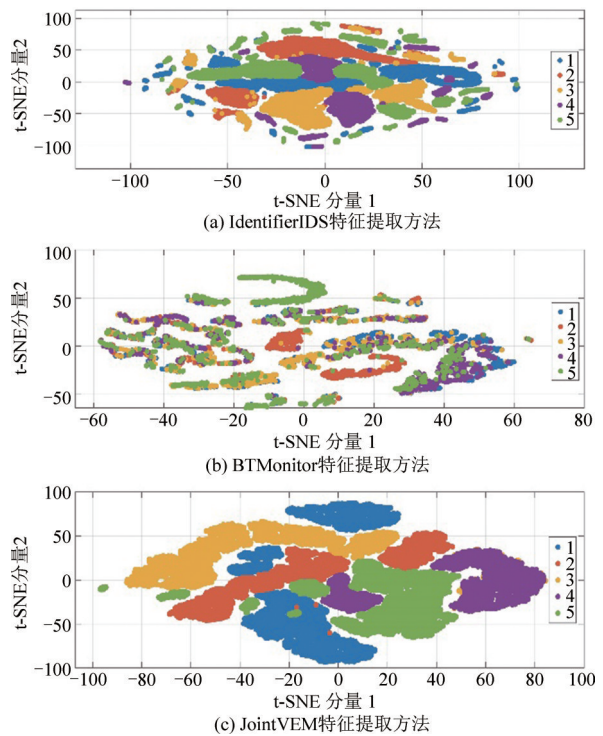


图3 特征空间 t-SNE 可视化

2.3 潜在空间特征学习能力评估

根据真实节点标签,对比分析标准AE和VAE两种模型隐空间特征的类内距离标准差。通过计算可得,标准AE的类内距离平均标准差为0.591,分离式VAE的类内距离平均标准差为0.317,分离式VAE隐空间特征类内距离更加紧凑,如图4所示。可以观察到VAE结构学得特征均能形成比标准AE更清晰、分离度更高的簇群。

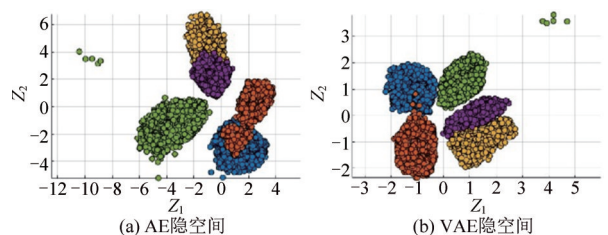


图4 AE和VAE结构隐空间对比

2.4 整体方法评测

对提出的3Mapping测绘框架进行全面测试,测试数据集为实验设置中描述的5个具有真实节点标签的数据集,通过测试评估聚类准确性,计算轮廓系数、DB指数,并和真实标签比较聚类结果准确性。

首先对 d_1 数据集进行GMM无监督聚类分析,该数据集包含59197个样本,对应5个ECU节点发送的13个不同ID的消息。测试结果如图5所示。使用簇内平方和(Within-Cluster Sum of Squares, WCSS)评估聚类效果,并使用肘部法则确定聚类数为5,使用轮廓系数和DB指数也得到同样的结果。进一步检查不同类对应的ID列表,发现不同ECU发送的ID消息被正确聚类为同一类,与CAN总线原型设置完全一致。

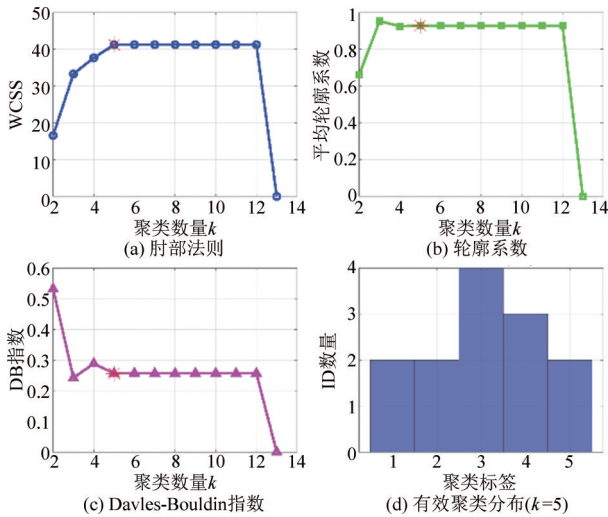


图5 d_1 数据集聚类结果

对其他4个数据集的测试结果如表5所示,均能准确聚类到5类,且不同ID对应的聚类类别与实验设置完全一致,超过0.95的轮廓系数和低于0.2的DB指数意味着样本聚类非常有效。

实验结果表明,在总线不同点测量电压—电磁多模态信号,均能对消息实现准确聚类,找到最优聚类数等于CAN总线中主动发送消息的节点数,实现“逻辑节点—消息ID”的有效映射,并完成对网络拓扑的测绘。

表5 整体方法聚类测试结果

数据集	分类结果	轮廓系数	DB指数
d_1	5	0.961	0.198
d_2	5	0.992	0.092
d_3	5	0.981	0.149
d_4	5	0.977	0.174
d_5	5	0.972	0.187

3 结束语

针对现有方法存在的不足,提出一种基于多模态信息融合的无监督CAN总线测绘框架。通过综合总线电压微分瞬态特性与近场电磁辐射脉冲多模态特征,构建了高维联合特征空间,准确表征硬件指纹。引入基于IF无监督异常数据清洗流程,有效剔除了由外部电磁干扰与测量误差引入的噪声样本,净化了特征分布。最终,结合VAE和高斯混合模型的聚类分析,该方法能够在无任何先验知识的黑盒环境下,稳定、准确地将不同的消息ID按其物理发送源进行归类,从而实现了车载CAN总线逻辑拓扑结构的准确测绘,为后续安全检测提供可靠真实标签。但本方法有效性的前提是基于目标节点主动发送信号驱动总线,因此还无法探测处于静默模式下的隐藏节点,如何探测总线上的隐藏节点还需进一步研究。

参考文献:

- [1] 国家发展改革委,中央网信办,科技部,等.关于印发《智能汽车创新发展战略》的通知[J].科学中国人, 2020(9):62-65.
- [2] SODERI S, COLELLI R, TURRIN F, et al. SENECAN: secure KEy Distribution OvEr CAN through watermarking and jamming[J]. IEEE Transactions on Dependable and Secure Computing, 2023, 20(3):2274-2288.
- [3] CUI J, CHEN Y N, ZHONG H, et al. Lightweight encryption and authentication for controller area network of autonomous vehicles[J]. IEEE Transactions on Vehicular Technology, 2023, 72(11):14756-14770.
- [4] ZHANG G Q, SHEN J, LI J T, et al. Payload processor: message authentication for in-vehicle CAN bus using data compression and tag filling[J]. Computer Networks, 2025, 259:No.111061.
- [5] UMAR M, WANG J D, LI F, et al. Physical layer authentication in the Internet of vehicles through multiple vehicle-based physical attributes prediction[J]. Ad Hoc Networks, 2024, 152:No.103303.
- [6] SUN H, SU H B, WENG J, et al. MTDCAP: moving target defense-based CAN authentication protocol[J]. IEEE Transactions on Intelligent Transportation Systems, 2024, 25(9):12800-12817.
- [7] BELLA G, BIONDI P, COSTANTINO G, et al. TOUCAN: a protocol to secure controller area network[C]//Proceedings of the ACM Workshop on Automotive Cybersecurity. New York, USA: ACM, 2019:3-8.
- [8] HAYWARD J, TOMLINSON A, BRYANS J. Adding cy-

- berattacks to an industry-leading CAN simulator [C]//Proceedings of the 2019 IEEE 19th International Conference on Software Quality, Reliability and Security Companion. Piscataway USA: IEEE, 2019:9-16.
- [9] VECTOR GMBH. DBC File Format [EB/OL]. (2023-10-21) [2025-10-17]. https://nusolar.github.io/training-f23/pdfs/DBC_File_Format_Documentation.pdf.
- [10] 庞启龙, 王建涛, 郑远浩. 基于 VAE-DANN 的变调制雷达辐射源个体识别算法[J]. 信息工程大学学报, 2025, 26(4):401-407.
- [11] SCHELL O, KNEIB M. VALID: voltage-based lightweight intrusion detection for the controller area network [C]//Proceedings of the 2020 IEEE 19th International Conference on Trust, Security and Privacy in Computing and Communications. Piscataway, USA: IEEE, 2020:225-232.
- [12] KULANDAIVEL S, GOYAL T, AGRAWAL A K, et al. CANvas: fast and inexpensive automotive network mapping [C]//Proceedings of the 28th USENIX Conference on Security Symposium. Berkeley, USA: USENIX Association, 2019:389-405.
- [13] POPA L, GROZA B, JICHICI C, et al. ECUPrint: physical fingerprinting electronic control units on CAN buses inside cars and SAE J1939 compliant vehicles [J]. IEEE Transactions on Information Forensics and Security, 2022, 17:1185-1200.
- [14] MURVAY P S, GROZA B. TIDAL-CAN: differential timing based intrusion detection and localization for controller area network [J]. IEEE Access, 2020, 8:68895-68912.
- [15] GROZA B, MURVAY P S, POPA L, et al. CAN-SQUARE-decimeter level localization of electronic control units on CAN buses [M]//BERTINO E, SHULMAN H, WAIDNER M, et al. Computer Security-ESORICS 2021. Cham, Switzerland: Springer, 2021:668-690.
- [16] KINGMA D P, WELING M. Auto-encoding variational bayes [C/OL]//International Conference on Learning Representations, 2014. <https://arXiv.org/abs/1312.6114>.
- [17] Arduino Team. UNO R4 Minima [EB/OL]. (2025-01-01) [2025-10-17]. <https://docs.arduino.cc/hardware/uno-r4-minima/>.
- [18] Renesas Electronics Corporation. RA4M1 Datasheet Group [EB/OL]. (2023-09-29) [2025-10-17]. <https://www.renesas.com/en/document/dst/ra4m1-group-datasheet?r=1054146>.
- [19] Philips. TJA1050 high speed CAN transceiver datasheet [EB/OL]. (2003-10-22) [2025-10-17]. <https://www.nxp.com/docs/en/data-sheet/TJA1050.pdf>.
- [20] INSTRUMENTS T. SN65HVD23x 3.3-V CAN bus transceivers datasheet (Rev. O) [EB/OL]. (2018-04-01) [2025-10-17]. <https://www.ti.com/lit/ds/symlink/sn65hvd230.pdf>.
- [21] DENG Z Y, LIU J J, XUN Y J, et al. IdentifierIDS: a practical voltage-based intrusion detection system for real in-vehicle networks [J]. IEEE Transactions on Information Forensics and Security, 2023, 19:661-676.
- [22] ZHOU J, JOSHI P, ZENG H B, et al. BTMonitor: bit-time-based intrusion detection and attacker identification in controller area network [J]. ACM Transactions on Embedded Computing Systems, 2019, 18(6):1-23.
- (编辑: 邵 华)
-
- (上接第 87 页)
- [13] GUO D Y, REN S, LU S, et al. GraphCodeBERT: pre-training code representations with data flow [DB/OL]. (2020-09-17) [2025-12-10]. <https://doi.org/10.48550/arXiv.2009.08366>.
- [14] DETTMERS T, PAGNONI A, HOLTZMAN A, et al. QLoRA: efficient finetuning of quantized LLMs [C]//Proceedings of Advances in Neural Information Processing Systems. New York, USA: Curran Associates, Inc., 2026, 36:10088--10115.
- [15] HU E J, SHEN Y L, WALLIS P, et al. LoRA: low-rank adaptation of large language models [DB/OL]. (2021-06-17) [2025-07-07]. <https://arXiv.org/abs/2106.09685>.
- [16] LIU B C, CHEN C Y, GONG Z, et al. MFTCoder: boosting code LLMs with multitask fine-tuning [DB/OL]. (2024-08-24) [2025-12-08]. <https://dl.acm.org/doi/abs/10.1145/3637528.3671609>.
- [17] FAN J H, LI Y, WANG S H, et al. A C/C++ code vulnerability dataset with code changes and CVE summaries [C]//Proceedings of the 17th International Conference on Mining Software Repositories. New York, USA: ACM, 2020. DOI:10.1145/3379597.3387501.
- [18] ZHOU Y Q, LIU S Q, SIOW J K, et al. Devign: effective vulnerability identification by learning comprehensive program semantics via graph neural networks [C]//Proceedings of the 33rd International Conference on Neural Information Processing Systems. Red Hook, USA: Curran Associates Inc., 2019:No.915.
- [19] KIM Y. Convolutional neural networks for sentence classification [C]//Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing. Stroudsburg, USA: ACL, 2014. DOI:10.3115/v1/D14-1181.
- [20] LIU Z X, TANG Z J, ZHANG J W, et al. Pre-training by predicting program dependencies for vulnerability analysis tasks [C]//Proceedings of the IEEE/ACM 46th International Conference on Software Engineering. New York, USA: ACM, 2024:1863-1875.
- [21] WANG Y, WANG W S, JOTY S, et al. CodeT5: identifier-aware unified pre-trained encoder-decoder models for code understanding and generation [C]//Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing. Stroudsburg, USA: ACL, 2021:8696-8708.
- [22] ROZIÈRE B, GEHRING J, GLOECKLE F, et al. Code Llama: open foundation models for code [DB/OL]. (2023-08-24) [2025-12-08]. <https://arXiv.org/abs/2308.12950>.
- (编辑: 邵 华)