

考虑网联汽车信息安全的交通流短时预测方法

王庞伟^{1,2}, 王天任^{1,2}, 李振华¹, 刘 琥³, 孙玉兰³

(1. 北方工业大学 智能交通技术交通运输行业重点实验室, 北京 100144; 2. 北方工业大学 城市道路交通智能控制技术北京市重点实验室, 北京 100144; 3. 中信科智联科技有限公司, 北京 100029)

摘要: 针对智能网联汽车因网络攻击或干扰造成的信息安全及数据缺失问题, 提出一种基于数据补全的交通流状态短时预测方法。首先, 基于边缘计算任务卸载模型, 对智能网联汽车 V2X 通信过程的异常数据动态辨识; 其次, 提出一种具有数据补全机制的图嵌入长短期神经网络模型, 实现网联汽车缺失数据补全; 再次, 通过补全后的完整数据集构建神经网络模型, 完成短时交通流状态预测; 最后, 选取北京市典型路段进行实验验证。结果表明, 该模型应用后交通流状态短时预测效果显著提高, 与其他方法相比预测误差最大降低 87.4%, 预测效果与实际交通流状态相比准确率达到 95%, 为智能网联环境下车辆信息安全与交通资源动态优化提供理论支持和技术方案。

关键词: 交通信息工程; 智能网联汽车; 信息安全; 交通流短时预测; 数据补全

中图分类号: U491.1+4

文献标志码: A

Short-term Traffic Flow Prediction Method Considering Information Security for Connected Vehicles

WANG Pangwei^{1,2}, WANG Tianren^{1,2}, LI Zhenhua¹, LIU Xiao³, SUN Yulan³

(1. Key Laboratory of Transport Industry of Intelligent Transportation Systems, North China University of Technology, Beijing 100144, China; 2. Beijing Key Lab of Urban Intelligent Traffic Control Technology, North China University of Technology, Beijing 100144, China; 3. CICT Connected and Intelligent Technologies Co. Ltd., Beijing 100029, China)

Abstract: Aiming at the information security and data missing caused by network attacks or interference for ICVs (Intelligent Connected Vehicle), this paper proposes a short-term prediction method of traffic flow state based on data imputation. Firstly, based on the edge computing task offloading model, the abnormal data of the V2X (vehicle to

everything, V2X) communication process of ICVs is identified dynamically. Secondly, a graph embedding (GE) with data imputation mechanism and long short-term memory (LSTM) neural network model is proposed to impute the missing data. Thirdly, the neural network model is established based on the complete data sets to realize the short-term traffic flow state prediction. Finally, by the proposed model applied in Beijing for field test, the final results show that the short-term prediction effect of traffic flow state is significantly improved. In comparison with other methods, the prediction error is reduced by 87.4% and the accuracy of the prediction effect is 95% according to the actual traffic flow state, which provides a novel theoretical support and technical solution for vehicular information security and dynamic optimization of traffic resources in the intelligent networked environment.

Key words: traffic information engineering; intelligent connected vehicles; information security; short-term traffic flow prediction; data imputation

随着 V2X (vehicle to everything, V2X) 通信技术不断发展, 智能网联汽车在网络安全、数据安全等方面面临众多挑战。各项研究成果表明^[1], 基于智能网联汽车 V2X 数据与多种交通传感器建立交通数据集来短时预测交通运行状态, 可以大幅提高交叉口的通行效率, 提高道路安全性, 为智能网联汽车提供信息安全保障。

目前的城市路网系统存在复杂的通信网络, 智能网联汽车通信与任务卸载过程极易受到网络干扰或攻击导致数据缺失。利用边缘计算技术来解决智能网联汽车网络传输安全与缺失数据补全问题成为发展趋势。利用 V2V (vehicle to vehicle, V2V) 与 V2I (vehicle to

收稿日期: 2022-09-08

基金项目: 北京市自然科学基金(4212034); 国家重点研发计划(2018YFB1600500); 智能交通技术交通运输行业重点实验室开放基金(F20211749)。

第一作者: 王庞伟(1982—), 男, 副教授, 工学博士, 主要研究方向为车路协同与智能驾驶。E-mail: wpw@ncut.edu.cn



infrastructure, V2I)通信技术实时获取的智能网联汽车数据判断城市路网的交通事件,并通过对智能网联汽车任务卸载数据采样加密^[2]、安全信道决策^[3]、异常数据剔除^[4]与动态传输耦合^[5]等方式提高智能网联汽车信息安全及道路通行效率。同时,在处理智能网联汽车异常数据与缺失数据的问题上,Grover等人^[6]提出叠加长短期记忆(LSTM)模型来补全智能网联汽车异常通信数据。Wu等人^[7]提出处理缺失数据的链方程多元补全(MICE)方法,利用缺失数据和观测数据之间的相关性补全原始数据。Cui等人^[8]通过在神经网络中设计插补单元来处理卸载缺失值并辅助交通预测。除上述方法外,生成对抗网络(GAN)数据插补模型^[9]、张量补全模型^[10]等方法都达到了补全缺失交通数据的效果,改善了智能交通系统运行。

在获得完整交通数据集的基础上,交通状态预测方法对交通流预测精度的影响也至关重要。传统的交通状态预测方法主要是基于交通流的模型以及统计学模型,如自适应回归综合移动平均模型^[11]、卡尔曼滤波模型^[12]及其各种变种模型等^[13-15],此类模型通常依赖于固定的假设,不能预测交通流数据的动态变化。与传统方法相比,基于变种K最邻近(K-nearest neighbors, KNN)的时空相关性预测^[16]、基于张量分解的图嵌入预测模型^[17]、综合卷积神经网络与图嵌入方法^[18-19]等混合交通预测模型的仅需一定数量的交通流原始数据与智能网联汽车数据,即可自行提取并学习其时空特征,从而对交通状态进行短时预测。Wei等人^[20]提出了一种根据相邻路段信息来预测道路交通状态的预测模型,该模型首先基于图嵌入实现路网表示,然后再根据生成式对抗神经网络实时产生交通信息预测交通状态,取得了较好的预测精度。Xie等人^[21]使用图神经网络(graph neural network, GNN)将路段连通性建模为一个图,将输入和输出表示为图序列,对城市区域路网道路中的交通流进行预测。Xu等人^[22]根据城市道路网中SCATS系统下的交通流数据时空特性,将交通路网表示为图,以交叉口为路网节点,基于图嵌入网络学习此系统下路网的时空特征,从而实时预测区域路网的整体交通状态。

综上所述,现有理论方法已取得一定研究成果,但仍存在以下问题:一方面,目前对V2X通信获取的智能网联汽车卸载缺失数据集的补全问题还有待进一步研究,另一方面,交通状态预测的研究大多采用于历史交通数据或开源数据集对交通状态进行预测,存在一定的滞后和冗余,导致预测结果实时性差、预测精度不高。

针对以上问题,通过建立基于智能网联汽车V2X

数据的交通流短时预测系统获取智能网联汽车任务卸载数据、动态辨识任务卸载异常数据和感知城市路网的交通状态。同时提出考虑智能网联汽车数据异常的交通状态短时预测方法,应用带有数据补全机制的长短期记忆神经网络模型补全智能网联汽车卸载缺失数据,最终实现对城市交叉口交通流状态的短时动态预测。

1 基于智能网联汽车V2X数据的交通流短时预测系统

针对城市交叉口场景,交通状态通常指交通环境中交通流各种车辆的运行状况,具有动态性、周期性、随机性等特性^[23]。随着V2X技术的发展,智能网联汽车V2X数据与多源传感器感知的交通状态数据可上传给交通管理系统或智能交通系统的其他子系统^[24-26],以此建立较完整的城市路交通状态数据集。

为提高交通检测数据精度及实时性,设计了基于智能网联汽车V2X数据的交通状态感知系统。系统主要包括智能路侧传感器与边缘计算设备(mobile edge computing, MEC)、配备车载单元(on board unit, OBU)的智能网联汽车,如图1所示。系统首先对交叉口中所有车辆的状态和路侧环境状态的信息进行实时感知,然后发送信息至部署路侧的智能路侧单元(road side unit, RSU),通过V2X通信发送数据至边缘计算单元进行多源信息融合。针对复杂交通网络环境下智能网联汽车V2X数据卸载过程,系统对缺失数据或异常数据进行动态辨识,根据历史平均值和观察值插补智能网联汽车任务卸载的数据缺失值,从而得到完整的交通状态数据集。再基于构建的神经网络预测下个时间步的交通状态,最终选择以交通流量和平均速度容量作为交通状态的预测与评价的特征值,实现对交通流的综合分析。本系统中各数据的类型及来源见表1。

2 考虑智能网联汽车数据异常的交通状态短时预测方法

根据建立的基于智能网联汽车V2X数据的交通状态感知系统,提出一种带有数据补全机制的图嵌入长短期记忆神经网络(graph embedding-long short-term memory, GE-LSTM)模型用来补全智能网联汽车卸载缺失数据并短时预测交叉口的交通状态。该模型结合了复杂网络的特性,通过网络表征学习和神经网络来学习交通路网数据的时空特征,如图2所示为预测模型结构。

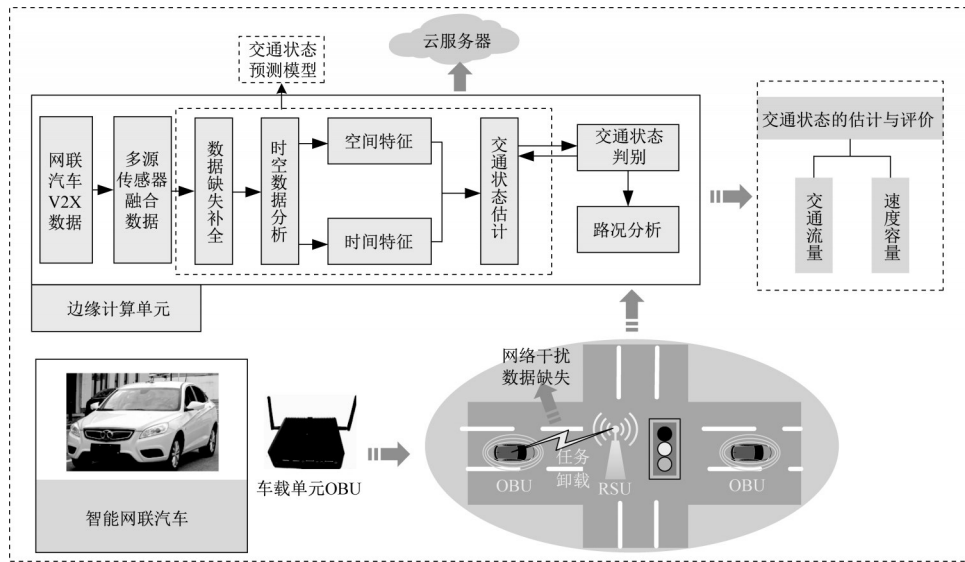


图 1 基于智能网联汽车 V2X 数据的交通状态感知系统总体框架

Fig.1 Framework of traffic state perception system based on V2X information for ICVs

表 1 交通状态数据集的类型及来源

Tab.1 Types and sources of traffic perception data sets

数据来源	数据类型
智能网联汽车 V2X 数据	速度、加速度、车牌号、车轮转速、转向角、刹车状况等
路侧传感器	车辆信息: 车牌号、经纬度、速度、水平距离、航向角 交通状态信息: 平均车速、车流量、平均排队长度、停车线位置等
交通信号机	信号周期、信号灯相位配时、信号灯色、当前相位剩余时间

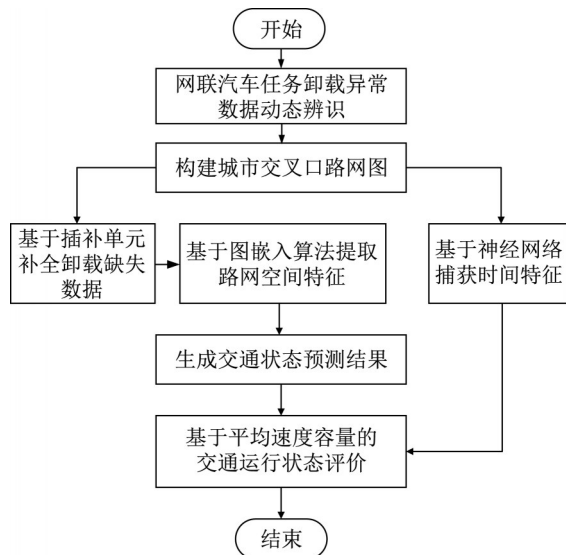


图 2 交通状态预测模型结构

Fig.2 Structure of traffic state prediction model

2.1 智能网联汽车任务卸载异常数据动态辨识

在数据任务卸载过程中,通信网络处在复杂交通网络环境下极易受到干扰,存在缺失数据或者异常数据的情况,如图 3 所示。首先采用马尔可夫链对智能网联汽车任务卸载过程建模,进一步确定任务卸载数据在通道阈值内的概率,通过 V2I 瑞利衰落信道模型得到任务卸载数据的信息安全概率,再由经验模态分解法去除噪声干扰,最后利用均值法剔除异常数据后得到缺失数据集,实现智能网联汽车任务卸载异常数据的动态辨识。

采用马尔可夫链对智能网联汽车 V2I 信道进行建模^[27],设 SNR_k 为时隙 k 中传输信道的 SNR 水平。马尔可夫卸载模型如图 3 所示。假设 V2I 信道的衰落系数为 I_{V2I} ,相应的衰落方差为 σ^2 , C_{V2I} 表示车辆和基础设施无线接入点之间的信道容量:

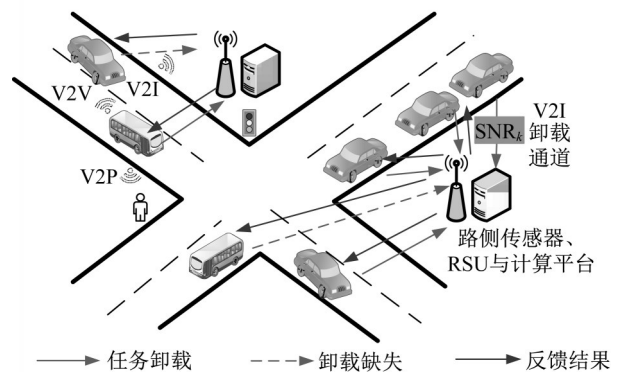


图 3 智能网联汽车任务卸载场景

Fig.3 Task offloading scenario for ICVs

$$C_{V2I} = \log_2(1 + \text{SNR}_k |I_{V2I}|^2) \quad (1)$$

同时考虑基于阈值的卸载策略,智能网联汽车选择具有最高信道传输增益的最佳信道来进行任务卸载,可以得到智能网联汽车数据在信道阈值内的概率:

$$P_{V2I}\{f_{V2I} \leq f_{\max}\} = \frac{\gamma(m_{V2I}, \frac{m_{V2I}}{\omega_{V2I}} f_{\max}^2)}{\Gamma(m_{V2I})} \quad (2)$$

式中: f_{V2I} 为智能网联汽车当前卸载频率; f_{\max} 为信道阈值; m_{V2I} , ω_{V2I} 分别表示信道的衰落参数和平均接收功率; $\Gamma(m_{V2I})$ 为伽玛函数。

采用瑞利衰落信道模型进行V2I通信,其中 $|I_{V2I}|^2$ 是一个随机变量,遵循参数 σ^{-2} 的指数分布、因此,智能网联汽车通过V2I信道在时隙 k 中卸载 α_k 位数据的信息安全概率为

$$p(s_k, \text{SNR}_k) = P\{C_{V2I} \geq \alpha_k\} = \exp\left(-\frac{2^{\alpha_k} - 1}{\sigma^2 \text{SNR}_k}\right) \quad (3)$$

通过经验模态分解有效降低卸载过程中产生的噪声干扰,减少数据异常。经验模态分解法变化描述如下:

$$x(t) = \sum_{i=0}^N C_i(t) + r_n(t) \quad (4)$$

式中:将智能网联汽车卸载信号 $x(t)$ 分解为信号主体分量 $C_i(t)$ 与剩余分量 $r_n(t)$ 。经滤波处理,最后对各处理后的信号主体分量和残余分量重构,得到去除噪声的卸载信号。进一步采用均值法剔除数据集中的显著异常数据后得到系统中的RSU覆盖范围内的智能网联汽车的缺失数据集 T ,表示为 $T = \{T_1, T_2, \dots, T_n\}$ 。

2.2 基于智能网联汽车V2X数据的城市区域路网建模

在城市区域路网图中,节点通常可以承载一系列交通特征,这些特征通常包括智能网联汽车的速度、加速度、经纬度位置等信息。根据第3节实验环境的设置,对道路网络进行建模,如图4所示,其中阴影区域为V2X路侧传感器所感知的范围。然后再将其表示为加权有向图,描述如下:

$$G = (V, E, X) \quad (5)$$

式中: $V = \{v_1, v_2, \dots, v_n\}$ 表示路段网络上的节点的集合, n 为节点的个数; $E = \{e_{ij}\}$ 表示各个顶点 v_i 和 v_j 之间相互连通的一组边; x 表示基于V2X路侧感知范围内的交通状态。

在时间 t 时,用图信号 $X_t \in R^{n \times c}$ 来表示图 G 的交

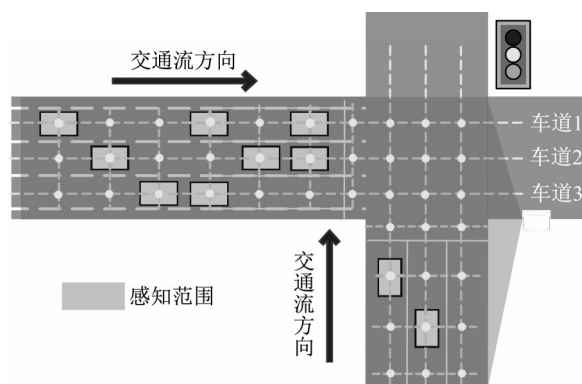


图4 城市交叉口路网图

Fig.4 Graph of urban intersection network

通状态,其中 c 表示交通状态特征的数量(例如交通流量、交通密度以及交通速度等)。在给定的 p 个历史时间步长中,图 G 中 n 个节点的交通状态观测值表示为 X , $X = (x_{t_1}, x_{t_2}, \dots, x_{t_p}) \in R^{p \times n \times c}$,对于所有节点的下 q 个时间步长的交通状态预测值表示为 Y , $Y = (Y_{t_{p+1}}, Y_{t_{p+2}}, \dots, Y_{t_{p+q}}) \in R^{q \times n \times c}$ 。将交叉口的平均速度容量 M_v 和实时交通流量 Q 作为描述交通状态的特征值,其中平均速度容量也可以用来评价交叉口的通行效率情况。对于平均速度容量的定义如下:

时间段 k 中,交叉口处的平均交通流车速为

$$\bar{V} = \frac{\sum_{i=1}^k \bar{v}_i}{k} \quad (6)$$

式中: \bar{v}_i 为 t 时刻路侧传感器感知范围内交通流的瞬时平均车速。

则时间 k 内交叉口的平均速度容量表示为

$$M_v = \frac{\bar{V}}{v_{\max}} \quad (7)$$

式中: v_{\max} 为交叉口处的最大通行车速。

2.3 基于数据补全与时空特征分析的交通流短时预测模型

在建立智能网联汽车V2X数据集与交通传感器感知数据集的基础上,交通状态预测不仅依赖于交通数据集的完整准确性,同样也会随着交叉口的空间特性和时间维度的变化而实时改变,因此对交通数据集的补全与道路时空的相关性分析可以提高预测的精度。

如图5所示为GE-LSTM模型的算法流程。模型首先基于DeepWalk从交叉口道路网络获得特征函数 $\varphi(v)$ 来提取交叉口的空间特征。然后,提取多源传感器融合数据集 E_t 与智能网联汽车缺失数据集 T ,将数据的时间特征与空间特征进行融合,构造输

入 X_t , 并将其输入至 LSTM 中的插补单元补全缺失数据并预测下一个时间步长 t 的交通状态, 最终输出的结果为 Y_t 。

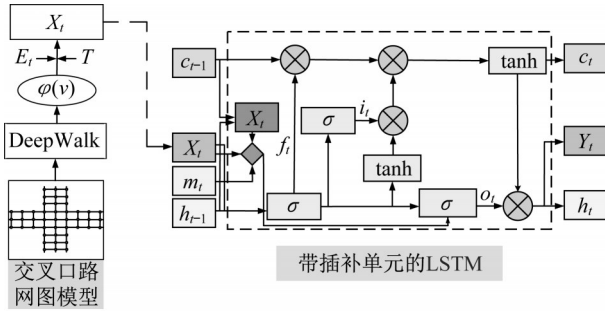


图5 预测模型的算法流程图

Fig.5 Algorithm flow chart of prediction model

2.3.1 基于图嵌入提取空间特征

在实际城市交叉口的网络空间中, 交通流具有一定的方向性, 且下游的交通流速度受到上游交通流的影响, 因此采用图嵌入 DeepWalk 算法来学习节点之间的相互作用以提取空间特征, 其原理如图 6 所示。

DeepWalk 在对学习网络的空间表示时, 首先会从中心节点开始, 通过随机游走 (Random Walk) 而生成一系列的随机游走序列, 其次基于 Skip-Gram 算法对产生的节点序列进行空间表征学习, 最后将每个节点序列嵌入到 d 维向量 ($d < n$) 表示中。

首先, DeepWalk 在路网的邻接矩阵上产生短小的随机游走进行采样, 并将其作为相似度度量 (similarity measure)。 W_i 表示从节点 v_i 开始的随机游走序列, 表示为 $\{W_i^1, W_i^2, \dots, W_i^k, \dots, W_i^m\}$, 其中 W_i^k 表示从节点 v_i 到节点 v_k ($k < m$) 的随机游走, W_i^{k+1} 表示从节点 v_k 到其邻近节点的随机游走。基于节点 v_i 产生的随机游走序列 $W_i = \{W_i^1, W_i^2, \dots, W_i^m\}$, 可以捕获中心节点下游交通流的空间相关性。

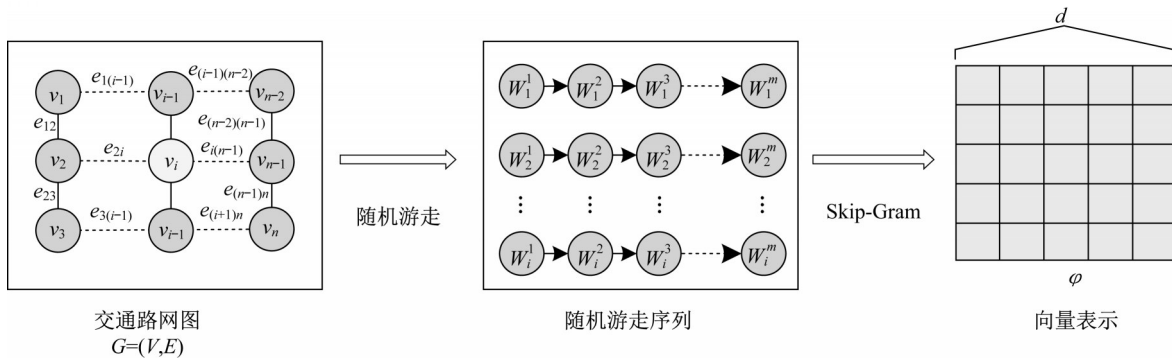


图6 DeepWalk 算法流程

Fig.6 Flow chart of DeepWalk algorithm

其次, 在完成对网络中的每个节点的随机游走后, 采用 Skip-Gram 算法来更新这些表示, 引入映射函数 $\varphi: V \rightarrow R^d$ (其中 d 是嵌入空间的维数, 且 $d \leq n$)。目的是找出与节点 v_i 相关性最大的节点, 对于随机游走序列 W , 该优化问题可以表示为

$$\begin{cases} \min_{\varphi} -\log P(W|\varphi(W_i^k)) \\ W = \{W_i^{k-l}, \dots, W_i^{k-1}, W_i^{k+1}, \dots, W_i^{k+l}\} \end{cases} \quad (8)$$

式中: $P(W|\varphi(W_i^k))$ 表示在一个随机游走序列 W 中, 当给定一个节点的 W_i^k 时, 在长度为 l 的窗口范围内, 下一个节点出现的概率。由于交通流的方向是单向的, 因此只考虑右侧窗口内的节点, 则优化问题表示为

$$\min_{\varphi} -\log P(W_i^{k+1}, \dots, W_i^{k+l}|\varphi(W_i^k)) \quad (9)$$

然后, 根据独立假设可以对式中的条件概率进行因式分解:

$$P(\{W_i^{k+1}, \dots, W_i^{k+l}\}|\varphi(W_i^k)) = \prod_{1 \leq j \leq l} P(W_i^{k+j}|\varphi(W_i^k)) \quad (10)$$

因此, 这样就表示每个节点 v_i 将映射到其当前的特征函数 φ 中。通过 softmax 函数来近似概率分布, 以减少计算资源的消耗, 可得到概率 $P(W_i^{k+j}|\varphi(W_i^k))$ 的表示为

$$P(W_i^{k+j}|\varphi(W_i^k)) = \frac{\exp(\varphi(W_i^{k+j})\varphi(W_i^k))}{\sum_{u \in m} \exp(\varphi(W_i^u)\varphi(W_i^k))} \quad (11)$$

2.3.2 基于插补单元补全缺失数据

如果卸载到路侧端的智能网联汽车数据包含缺失值, 则将使用具有处理缺失值功能的插补单元进行处理, 目的是基于历史平均值和具有可学习衰减率的最后观察值推断智能网联汽车任务卸载的数据缺失值。此外, 从插补值补全出的值可进一步提高交通流预测的精确度。

该插补单元由前向单元输出状态 C_{t-1} 和前向输出值 h_{t-1} 作为输入,以推断后续观察值,如图5所示。由屏蔽向量 m_t 确定缺失值的位置,缺失的输入值可以通过插补单元进行插补。

推断观测 X_t' 表示为

$$X_t' = \sigma(W_i C_{t-1} + U_i h_{t-1} + b_i) \quad (12)$$

式中: W_i 和 U_i 是权重; b_i 是插补单元中的偏差; $\sigma(x)$ 表示在等式中定义的 sigmoid 函数。

然后,输入向量的每个缺失元素由推断元素更新:

$$X_t^d = m_t^d X_t^d + (1 - m_t^d) X_t'^d \quad (13)$$

式中: $X_t'^d$ 是 X_t' 的第 d 个元素。根据式(13),如果缺少 x_t^d ,则 m_t^d 为零, X_t^d 由 $X_t'^d$ 插补。

2.3.3 基于神经网络捕获时间特征

原始的 V2X 多源信息融合数据经过 DeepWalk 算法的空间特征提取,得到了一个具有更高阶特征的网络嵌入,然后再将其作为 LSTM 神经网络模型的输入,可以实现动态预测下一 m 个时间步长的交通状态 $\{y_{t+1}, y_{t+2}, \dots, y_{t+m}\}$ 。在路网图 G 中,指定的滑动窗口 l 内其交通状态表示为 $X_t = \{x_{t-l+1}, x_{t-l+2}, \dots, x_t\}$ 。令 X_t 作为 LSTM 神经网络模型的输入,其维度为 d 。

图5右侧的 LSTM 单元中,对于三个输入来说,分别设有输入门、遗忘门和输出门作为约束控制来对这些输入进行选择性地处理。

在时间 t , LSTM 单元有 3 个输入:当前交通状态 X_t ,前一隐藏层输出值 h_{t-1} 以及输出状态 c_{t-1} ;同时包含 3 个输出分别是此时隐藏层输出值 h_t 以及输出状态 c_t ,产生的预测结果为 Y_t 。3 个门的状态分别为 i_t 、 f_t 、 o_t ,为 0 到 1 之间的数值。其中,在此网络单元中, c_t 和 h_t 会传递到下一个网络,其运算过程如下:

$$i_t = \sigma(W_{xi} X_t + W_{hi} h_{t-1} + W_{ci} c_{t-1} + b_i) \quad (14)$$

$$f_t = \sigma(W_{xf} X_t + W_{hf} h_{t-1} + W_{cf} c_{t-1} + b_f) \quad (15)$$

$$o_t = \sigma(W_{xo} X_t + W_{ho} h_{t-1} + W_{co} c_{t-1} + b_o) \quad (16)$$

$$c_t = f_t c_{t-1} + \tanh(W_{xc} X_t + W_{hc} h_{t-1} + b_c) \quad (17)$$

$$h_t = o_t \tanh(c_t) \quad (18)$$

式中:用 W_{xc} 、 W_{xi} 、 W_{xf} 、 W_{xo} 分别表示交通状态的输入 X_t 的权重矩阵; W_{hc} 、 W_{hi} 、 W_{hf} 、 W_{ho} 表示隐藏层 h_t 的权重矩阵; W_{ci} 、 W_{cf} 、 W_{co} 分别表示输出状态 c_t 的权重矩阵; b_i 、 b_c 、 b_f 、 b_o 表示偏置向量;其中激活函数 \tanh 可将变量一一映射在 $[0, 1]$ 的范围中。 $\sigma(x)$ 表示在等式中定义的 sigmoid 函数,如式(19); $\tanh(x)$ 表示等式中定义的双曲正切函数,如式(20)。

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (19)$$

$$\tanh(x) = \frac{2}{1 + e^{-x}} - 1 \quad (20)$$

通过上述 LSTM 计算,得到 c_t 和 h_t ,此时网络预测输出可计算为

$$Y_t = W_y h_t + b_y \quad (21)$$

式中: W_y 表示交通预测值隐藏层 h_t 的权重矩阵; b_y 表示预测值的偏置向量。

3 实验验证与结果分析

为了验证提出的交通状态预测模型,针对典型的交叉口场景,搭建了智能路侧设备与智能网联汽车联合实验平台。实验以智能网联汽车 V2X 数据与多源交通传感器融合数据作为模型输入,对数据补全与预测模型进行验证。

3.1 实验环境及实验平台

实验选取了北京市石景山区阜石路一处典型交叉口作为实验路段。该交叉口为由东向西的单向行驶 4 车道,其中最右侧车道为专用的右转车道。如图7所示,实验自主搭建了智能路侧设备与智能网联汽车联合实验平台,用于智能网联汽车 V2X 任务卸载数据与交通状态数据的采集及实时处理,平台主要包括智能路侧设备 RSU、智能网联汽车及通信设备。配备有车载单元 OBU 的智能网联汽车通过 V2I 通信将车辆数据与任务数据卸载至智能路侧设备,智能路侧设备搭载的边缘计算平台对智能网联汽车的任务卸载数据、道路的实时交通状态与路侧多源传感器的数据融合处理,最终生成实际交叉口场景下带有缺失数据的交通状态数据集。实验平台参数如表2所示。

表2 实验平台参数

Tab.2 Parameters of experimental platform

实验平台	参数	值
智能网联汽车	任务卸载信道带宽/MHz	20
	最大发射功率/dBm	25
	OBU 功率/W	6
智能路侧单元	激光雷达线束	32
	摄像机分辨率	1 080p
	MEC 功率/W	45
	RSU 工作带宽/MHz	20

实验数据采集的过程为 30min,获取到整个交叉口智能网联汽车以及其他交通状态的实时数据,共约 60 000 条,数据集包含的具体内容如表3、表4所示。数据集包含的多源交通传感器能获得的车流量等信息,再结合 V2I 任务卸载数据可大幅提高对整个

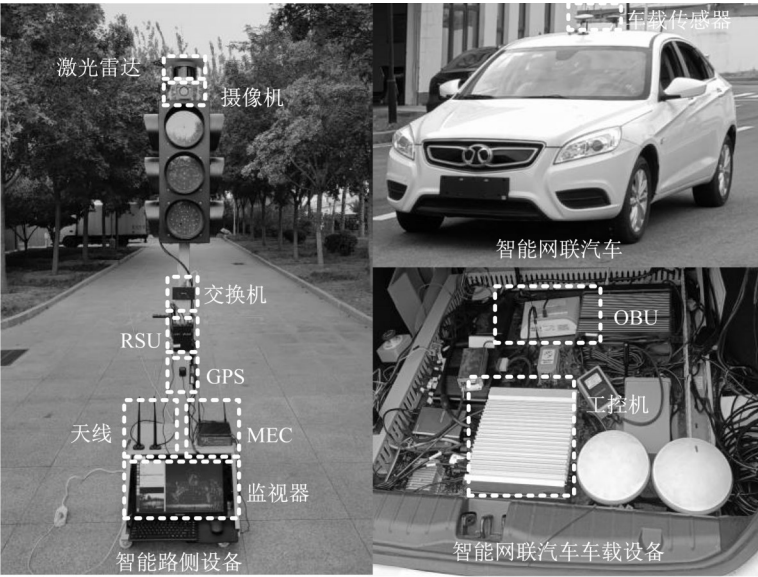


图 7 智能路侧设备与智能网联汽车联合实验平台

Fig.7 Experimental platform of intelligent roadside devices and ICVs

路段或交叉口的交通状态的实时感知能力,提高短
时预测精度。由于 V2X 通信环境中网络状态复杂,
极易因干扰或攻击导致任务卸载过程出现数据丢包
或异常的情况。因此,智能网联汽车任务卸载数据
集中包含一定数量的异常与缺失数据。

3.2 模型参数设置

实验使用了两种常用的预测评价指标均方根误
差(RMSE)和平均绝对误差(MAE)分析所提出的模
型和对比模型的性能^[28]。RMSE 和 MAE 能反映真
实观测值和预测值之间差异,其取值在 $[0, +\infty]$ 之
间。评价结果与预测效果呈负相关,当数值越大,说

明模型的预测效果越远离真实情况,反之则说明越
接近真实情况。

$$R = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}}$$
 (22)

$$M = \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{n}$$
 (23)

式中: y_i 为真实观测值; \hat{y}_i 为预测值。

实验根据时间序列把实验数据集分成了两组:
第一组将 70% 作为训练组(train data),用于模型的
训练;第二组将 30% 作为测试组(test data),用于模

表 3 智能网联汽车任务卸载数据集

Tab.3 Task offloading data sets for ICVs

V2X 通信	车牌号	经度	纬度	车轮转/(rs ⁻¹)	转向角/(°)	速度/(m·s ⁻¹)	加速度/(m·s ⁻²)	水平距离/m	航向角/(°)
是	京 N C5530	116.201 0	39.923 66	0.05	15	0.10	0.06	5.82	7.30
否	京 N 46735	116.200 9	39.923 66	—	—	1.26	—	11.71	6.52
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
是	京 A V3210	116.201 1	39.923 62	2.4	43	*	0.33	20.15	13.49
是	京 U B3957	116.201 0	39.923 62	*	*	*	*	12.43	8.03

注:*代表智能网联汽车任务卸载缺失数据。

表 4 交通状态数据集

Tab.4 Data sets of traffic state

时间戳	交通流量/ (veh·s ⁻¹)	速度/(m·s ⁻¹)	排队长度/m	信号周期/s	信号灯色 (东西向)	当前相位剩余 时间/s
1609232611	5	1.38	6.80	105	R	10
1609232612	7	0.83	14.80	105	R	9
⋮	⋮	⋮	⋮	⋮	⋮	⋮
16092344821	15	3.33	0	105	G	15
1609234482	16	5.36	0	105	G	14

注:交通流量为当前采样时间路侧平台感知范围内的所有车辆数;信号灯的灯色用 R、G、Y 分别表示红灯、绿灯和黄灯。

型的测试。此外,手动调整和设置模型的最优参数,结果如表5所示。

表5 模型参数设置
Tab.5 Model parameters

	模型参数	值
DeepWalk	嵌入维度 d	100
	滑动窗口大小 l	5
	随机游走长度	40
	随机游走次数	10
LSTM	隐藏层单元数	64
	训练批次大小	64
	训练周期	500
	学习率	0.005

3.3 异常数据动态辨识与数据补全结果分析

从实验采集的数据集中选取 10min 包含智能网联汽车任务卸载异常值与缺失值的车速数据进行异常数据动态辨识与补全效果实验。如图 8a 所示,数据集中存在显著异常值,并在相应时段内存在一定缺失数据。图 8b 为经动态辨识与数据补全后的效果,异常值被有效剔除且相应时段的车速数据得到补全。图 8c 为去除异常值与补全数据前后的对比结果,证明智能网联汽车动态辨识与数据补全效果良好。

3.4 交通状态短时预测结果分析

实验基于路侧感知平台采集的实时交通状态数据,通过 GE-LSTM 预测模型对城市交叉口的交通流量以及平均速度容量分别进行不同时间段的短时预

测,主要包含 10min、15min 和 30min 的交通预测,其中各预测对象的误差值如表 6 所示,具体的分析结果如下。

将 10min、15min、30min 时长的实验数据按 7:3 比例划分为 420s、720s、1 530s 时长的训练集数据与 180s、270s、540s 时长的测试集数据。其中图 9a、9b、9c 分别为 3 个时段交通流量测试集的真实值、数据补全后预测值与数据异常时预测值的对比结果,并得到数据补全后预测值与真实观测值之间的误差。

由图 9 可知,模型的预测结果能够准确的反映并预测出交通流量随交叉口信号灯进行周期性变化的趋势。同时,预测结果表明当智能网联汽车任务卸载数据存在缺失异常时,预测值曲线出现明显误差。依据表 6 可知,在交通流量的 10min、15min 和 30min 预测中,其数据补全后预测误差 RMSE 和 MAE 值稳定在 1 和 0.8 上下,且浮动不超过 0.1,证明数据补全后预测效果良好^[29]。

表6 预测模型的误差值

Tab.6 Error values of predicted model

预测对象	预测时长/min	RMSE	MAE
交通流量	10	1.033 2	0.843 2
	15	1.103 9	0.878 9
	30	1.045 7	0.831 2
平均速度容量	10	0.147 5	0.119 4
	15	0.128 1	0.101 8
	30	0.129 6	0.101 9

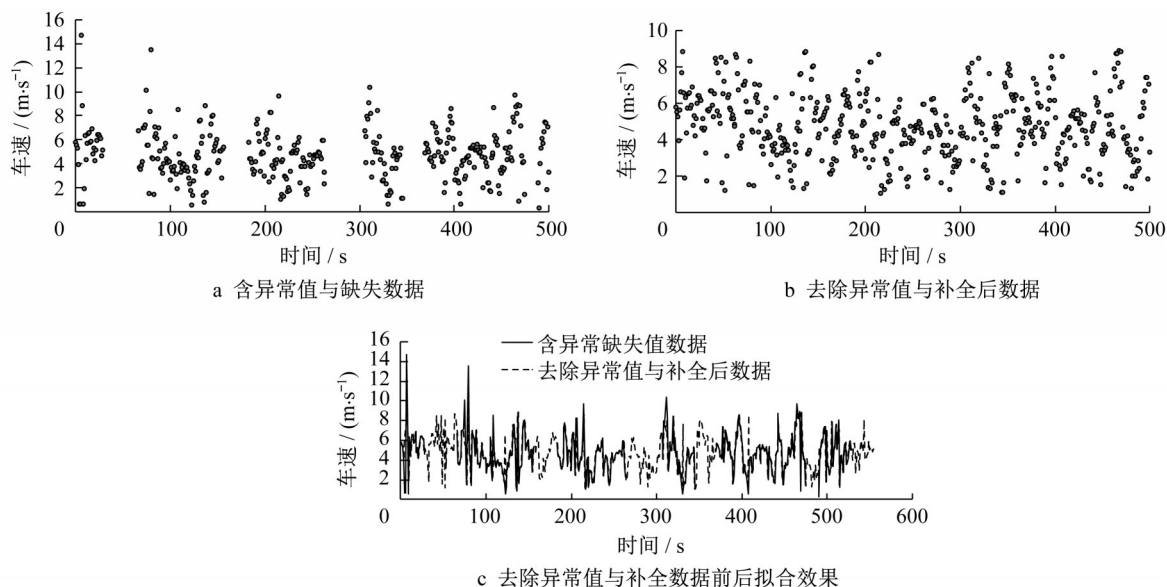


图8 异常数据动态辨识与数据补全结果

Fig.8 Results of abnormal data dynamic identification and data imputation

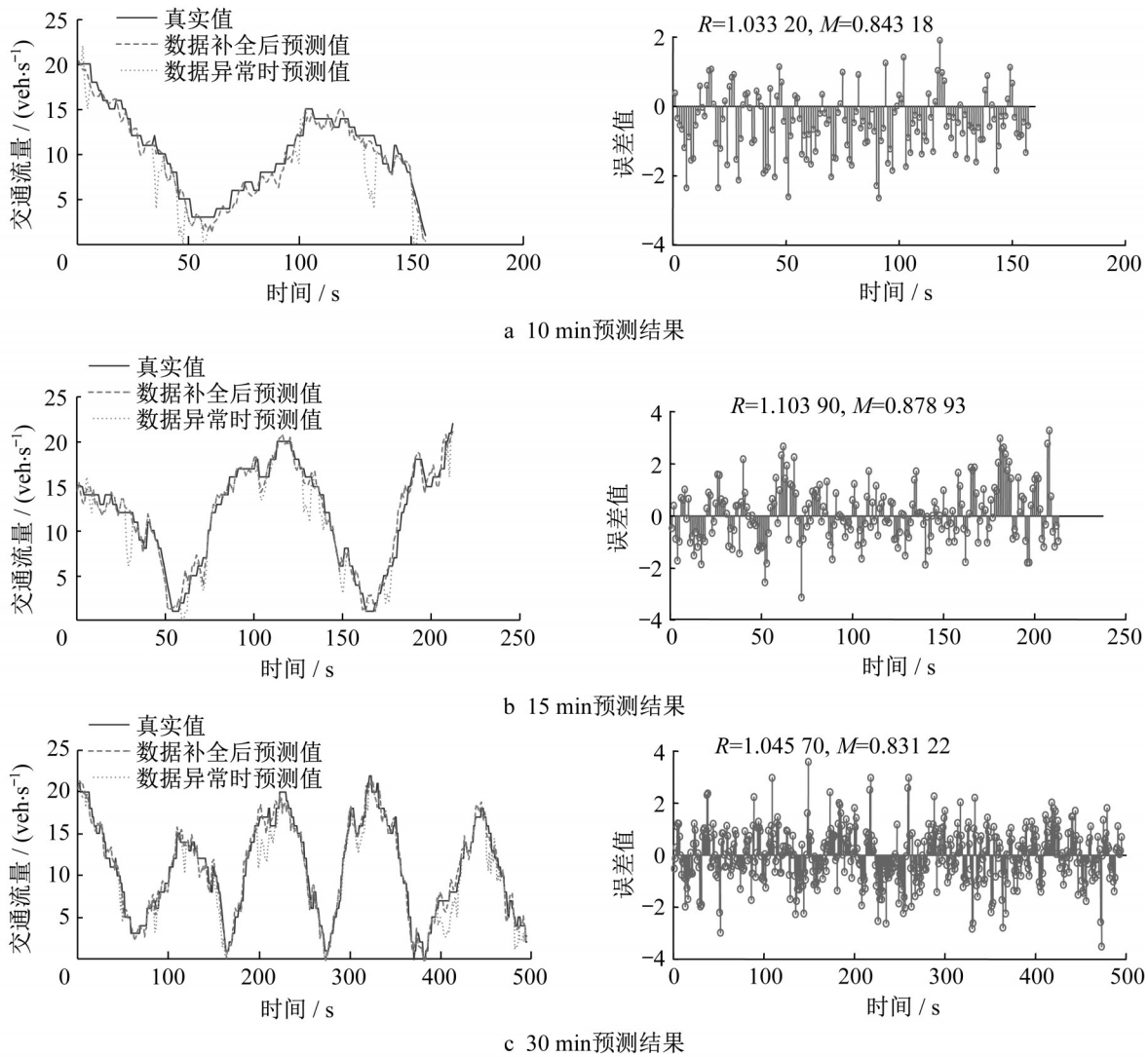


图9 不同时段下交通流量预测结果

Fig.9 Prediction results of average traffic flow in different time

如图10所示,图10a、10b、10c分别为预测3个时段平均速度容量测试集的真实值、数据补全后预测值与数据异常时预测值的对比结果,并得到数据补全后预测值与真实观测值之间的误差。

平均速度容量(M_v)通常被用来评价和分析道路交通状态的表现情况, M_v 越大,则说明当前交叉口的表现良好,反之则说明当前交叉口的交通状态表现越差,发生交通事故的概率会大幅增加。实验根据通过交叉口的最大行程时间^[30]将 M_v 的阈值设置为0.6,即 M_v 大于0.6则说明交通状态表现良好。如图10所示,在平均速度容量未来时间步长平均速度容量超过阈值0.6所对应的时间点,由此可实时感知并判断交叉口的交通运行状况,也为交通管理者对交叉口进行交通实时诱导及疏解交叉口排队压力提供了数据支持。

3.5 对比实验结果分析

为了直观体现所提预测模型(GE-LSTM)的准确性,在实验结果中引入其他神经网络预测模型进行对比^[20-22],主要包括SVM、LSTM、KNN、CNN和RNN模型,参数设置方面,上述模型与所提预测模型具有相同的隐藏单元的数量、训练批次大小、训练周期以及学习率,如表5所示。

在对比实验中,将上述模型分别对交通流量和平均速度容量进行分时段预测,最终各模型对比结果如图11、12所示。

在10min、15min和30min的3个时段的交通流量和平均速度容量预测中,各模型预测误差RMSE和MAE值随预测时间的增大而减小。相比其他5类模型,GE-LSTM模型RMSE和MAE值最小,其中交通流量的预测误差RMSE和MAE值分别减少了

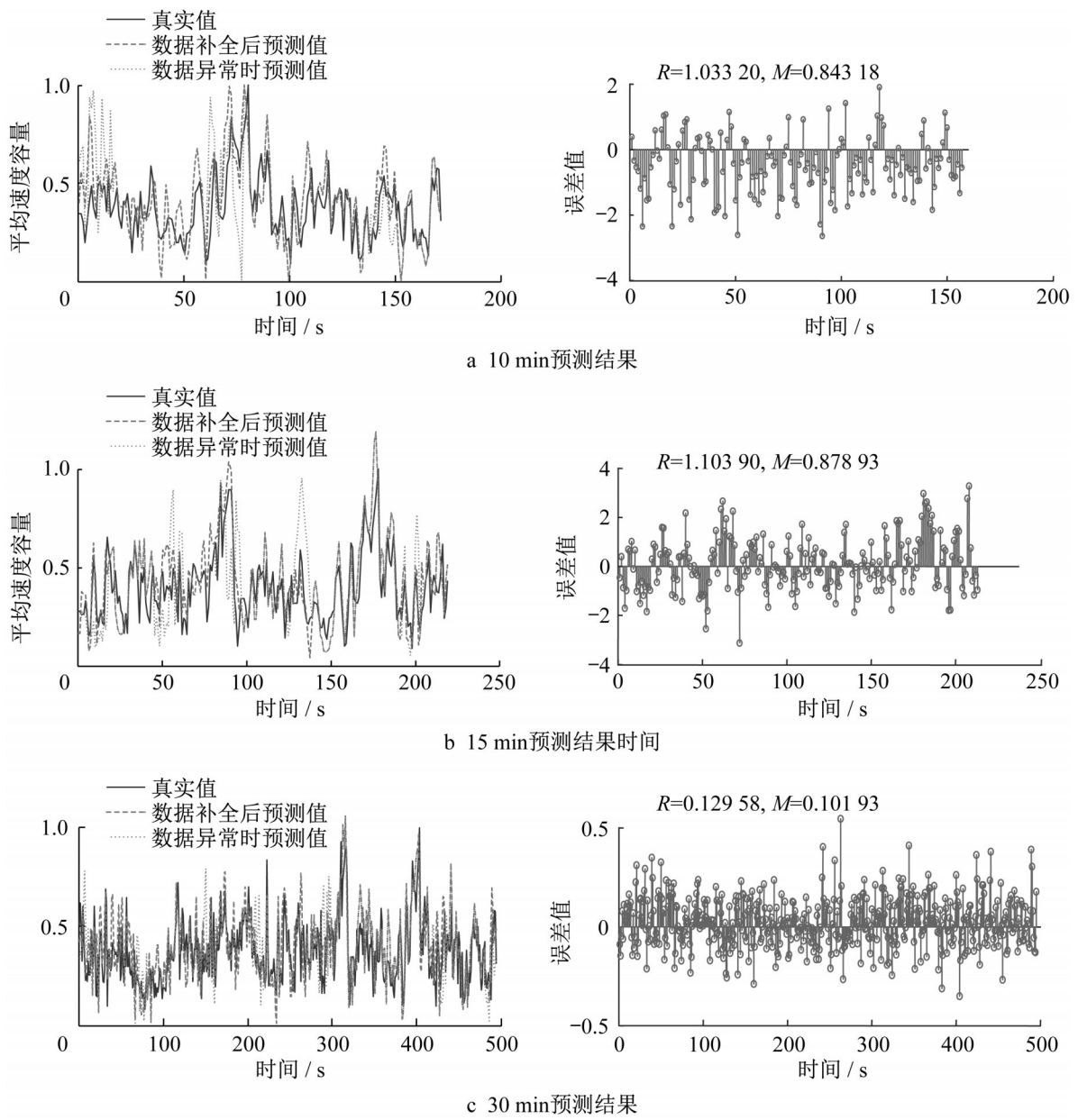


图 10 不同时段下平均速度容量预测结果

Fig.10 Prediction results of mean speed capacity in different time

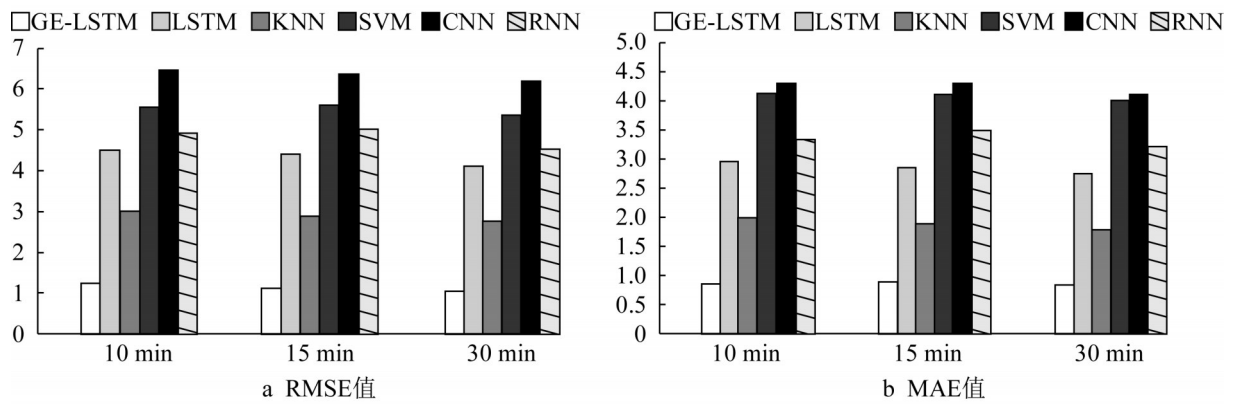


图 11 交通流量的对比结果

Fig.11 Comparison results of average traffic flow

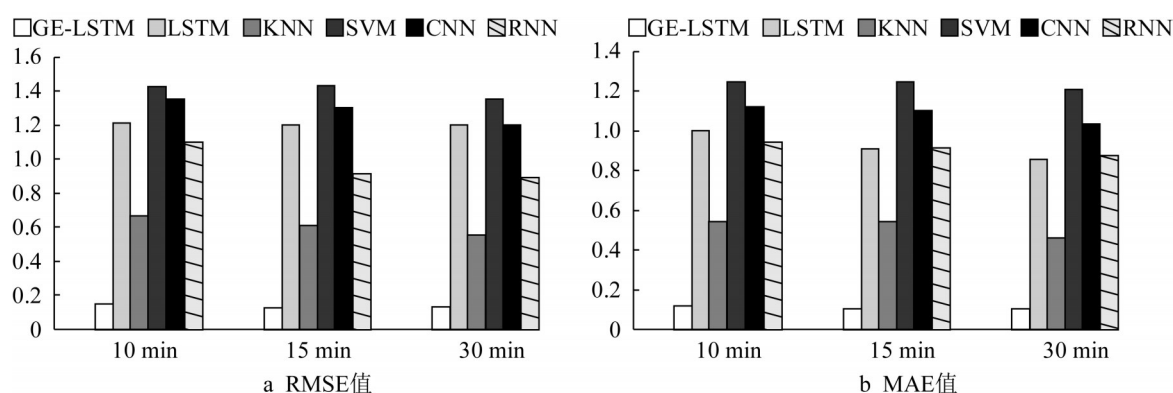


图 12 平均速度容量的对比结果

Fig.12 Comparison results of mean speed capacity

74.6%、71.7%，平均速度容量分别减少了86.5%、87.4%。此外，根据GE-LSTM与LSTM模型对比可知，GE-LSTM模型的预测误差值最小，证明在交通状态预测的过程中，通过引入图嵌入(GE)模型学习道路网络中的空间特征，对于基于LSTM进行时间上的动态预测有促进作用，进一步证明了GE-LSTM模型预测的准确性。

4 结论

(1)将智能网联汽车V2X数据与路侧交通传感器实时感知数据引入城市道路交通流状态短时预测方法。通过动态辨识智能网联汽车异常缺失数据与分析交通状态数据的时空特征，在传统长短期记忆神经网络上加入插补单元，实现交通状态的实时感知、补全与预测，并通过实验对比给出数据补全后的预测效果。

(2)通过对智能网联汽车数据异常缺失时与补全后的交通流预测结果对比分析，显示所提补全与预测方法的有效性，并引入其他神经网络预测模型作对比实验，实验结果表明所提模型的预测误差值最小，证明了模型预测的准确性。

(3)所提方法与实验平台能够达到有效解决智能网联汽车任务卸载信息安全问题的目的，后续将所研究的交叉口场景扩展至城市区域路网内的交通流状态预测，为智能网联汽车信息安全与交通资源合理调度提供理论支持和技术方案。

作者贡献声明：

王庞伟：提出论文选题，设计论文架构，指导实验开展，论文审阅与修改；

王天任：理论方法实现及实验方案实施，数据整理与处

理，论文撰写与修改；

李振华：技术指导与论文修改；

刘斌：文献整理，实验结果验证，论文撰写与绘图；

孙玉兰：技术指导与论文修改。

参考文献：

- [1] 鹿应荣,许晓彤,丁川,等. 车联网环境下信号交叉口车速控制策略[J]. 交通运输系统工程与信息, 2018, 18(1): 50.
LU Yingrong, XU Xiaotong, DING Chuan, *et al.* A speed control strategy at signalized intersection under connected vehicle environment [J]. Journal of Transportation Systems Engineering and Information Technology, 2018, 18(1): 50.
- [2] CHEN S, ZHU X, ZHANG H, *et al.* Efficient privacy preserving data collection and computation offloading for fog-assisted Iot [J]. IEEE Transactions on Sustainable Computing, 2020, 5(4): 526.
- [3] PEI X, YU H, WANG X, *et al.* Noma-based pervasive edge computing: Secure power allocation for IoV [J]. IEEE Transactions on Industrial Informatics, 2020, 17(7): 5021.
- [4] USMAN M, JOLFAEI A, JAN M A. Rasec: an intelligent framework for reliable and secure multilevel edge computing in industrial environments [J]. IEEE Transactions on Industry Applications, 2020, 56(4): 4543.
- [5] HAN X, TIAN D, SHENG Z, *et al.* Reliability-aware joint optimization for cooperative vehicular communication and computing [J]. IEEE Transactions on Intelligent Transportation Systems, 2020, 22(8): 5437.
- [6] GROVER H, ALLADI T, CHAMOLA V, *et al.* Edge computing and deep learning enabled secure multitier network for internet of vehicles [J]. IEEE Internet of Things Journal, 2021, 8(9): 14787.
- [7] WU R, HAMSHAW S D, YANG L, *et al.* Data imputation for multivariate time series sensor data with large gaps of missing data [J]. IEEE Sensors Journal, 2022, 22(11): 10671.
- [8] CUI Z, KE R, PU Z, *et al.* Stacked bidirectional and unidirectional LSTM recurrent neural network for forecasting

- network-wide traffic state with missing values [J]. Transportation Research Part C: Emerging Technologies, 2020, 118(1):102.
- [10] LI Q, TAN H C, WU Y K, *et al.* Traffic flow prediction with missing data imputed by tensor completion methods [J]. IEEE Access, 2020, 8(2):63188.
- [11] BIERNACKI A. Improving quality of adaptive video by traffic prediction with (f) arima models [J]. Journal of Communications and Networks, 2017, 19(5): 521.
- [12] GUO J H, WILLIAMS B. Real-time short-term traffic speed level forecasting and uncertainty quantification using layered kalman filters [J]. Transportation Research Record Journal of the Transportation Research Board, 2010, 2175(1):28.
- [13] 许岩岩, 翟希, 孔庆杰, 等. 高速路交通流短时预测方法[J]. 交通运输工程学报, 2013, 13(2): 114.
- XU Yanyan, ZHAI Xi, KONG Qingjie, *et al.* Short-term prediction method of freeway traffic flow [J]. Journal of Traffic and Transportation Engineering, 2013, 13(2): 114.
- [14] GUO X C, LI Y, YANG J. Proactive traffic responsive control based on state-space neural network and extended kalman filter [J]. Journal of Southeast University, 2010, 26(3): 466.
- [15] 张矢宇, 杨宇昊, 陈尘, 等. 基于回归卡尔曼滤波组合模型的航道整治区域船舶交通流时空预测[J]. 大连海事大学学报, 2021, 47(1): 37.
- ZHANG Shiyu, YANG Yuhao, CHEN Chen, *et al.* Spatio-temporal prediction of ship traffic flow in waterway regulation area based on regression kalman filter combination model [J]. Journal of Dalian Maritime University, 2021, 47(1): 37.
- [16] TAK S, WOO S, YEO H. Data-driven imputation method for traffic data in sectional units of road links [J]. IEEE Transactions on Intelligent Transportation Systems, 2016, 17(6): 1762.
- [17] WU Y K, TAN H C, LI Y, *et al.* A fused cp factorization method for incomplete tensors [J]. IEEE Transactions on Neural Networks and Learning Systems, 2019, 30(3): 751.
- [18] BRONSTEIN M M, BRUNA J, *et al.* Geometric deep learning: going beyond euclidean data [J]. IEEE Signal Processing Magazine, 2017, 34(4): 18.
- [19] GROVER A, LESKOBEC J. Node2vec: scalable feature learning for networks [C]//Proceedings of the 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining. San Francisco California USA: Association for Computing Machinery, 2016, 16(8):855-864.
- [20] WEI C C, XU D W, PENG P, *et al.* GE-GAN: a novel deep learning framework for road traffic state prediction [J]. Transportation Research Part C: Emerging Technologies, 2020, 117(8): 102635.
- [9] LEE S Y, CONNERTON T P, LEE Y W, *et al.* Semi-GAN: an improved GAN-based missing data imputation method for the semiconductor industry [J]. IEEE Access, 2022, 10(1):72328.
- [21] XIE Z P, LV W F, *et al.* Sequential graph neural network for urban road traffic speed prediction [J]. IEEE Access, 2020, 8(5): 63349.
- [22] XU D W, DAI H W, *et al.* Road traffic state prediction based on a graph embedding recurrent neural network under the scats [J]. Chaos, 2019, 29(10): 103125.
- [23] 朱琳. 城市快速路交通态势评估理论与方法研究[D]. 北京: 北京交通大学, 2013.
- ZHU Lin. Theory and method studies on traffic situation assessment for urban expressways [D]. Beijing: Beijing Jiaotong University, 2013.
- [24] 朱敏. 视频车辆检测器产品的设计与开发[D]. 杭州: 浙江工业大学, 2019.
- ZHU Min. Design and development of video vehicle detector product [D]. Hangzhou: Zhejiang University of Technology, 2019.
- [25] WANG P, DENG H, ZHANG J, *et al.* Model predictive control for connected vehicle platoon under switching communication topology [J]. IEEE Transactions on Intelligent Transportation Systems, 2021, 23(7):7817.
- [26] 王庞伟, 于洪斌, 张为, 等. 城市车路协同系统下实时交通状态评价方法[J]. 中国公路学报, 2019, 32(6): 176.
- WANG Pangwei, YU Hongbin, ZHANG Wei, *et al.* Real-time traffic status evaluation method for urban cooperative vehicle infrastructure system [J]. China Journal of Highway and Transport, 2019, 32(6):176.
- [27] ZHOU J, TIAN D, WANG Y, *et al.* Reliability-oriented optimization of computation offloading for cooperative vehicle-infrastructure systems [J]. IEEE Signal Processing Letters, 2018, 26(1): 104.
- [28] 赵阳阳, 夏亮, 江欣国. 基于经验模态分解与长短时记忆神经网络的短时地铁客流预测模型[J]. 交通运输工程学报, 2020, 20(4): 194.
- ZHAO Yangyang, XIA Liang, JIANG Xinguo, *et al.* Short-term metro passenger flow prediction based on emd-lstm [J]. Journal of Traffic and Transportation Engineering, 2020, 20(4):194.
- [29] LU W, YI Z, WU R, *et al.* Traffic speed forecasting for urban roads: a deep ensemble neural network model [J]. Physica A: Statistical Mechanics and Its Applications, 2022, 593(5):126.
- [30] MINH Q T, KAMIOKA E. Traffic state estimation with mobile phones based on the "3r" philosophy [J]. Ieice Trans Communication, 2011, 94(12):3447.