

# 一种融合时空关联与社会事件的交通流预测方法



吕明琪 洪照雄 陈铁明

浙江工业大学计算机科学与技术学院 杭州 310023

(mingqilv@zjut.edu.cn)

**摘要** 交通流预测作为智能交通系统的一个关键问题,是国内外交通领域的研究热点。交通流预测的主要挑战在于交通流数据本身具有复杂的时空关联,且易受各种社会事件的影响。针对这些挑战,提出一种用于交通流预测的深度学习框架。一方面,针对道路网络非欧氏的空间关联以及交通流时序数据的时间关联,设计了一种融合图卷积神经网络和循环神经网络的特征抽取子网络;另一方面,针对社会事件对交通流的潜在影响,设计了一种基于卷积神经网络的社会事件特征抽取子网络。最后,融合时空关联特征抽取子网络和社会事件特征抽取子网络,实现交通流预测模型。为了验证模型的有效性,文中基于真实交通流数据进行了实验。结果表明,所提模型与传统的预测模型相比具有较高的准确度,准确度提高了3%~6%。

**关键词:** 交通流预测;图卷积网络;循环神经网络;社会事件;卷积神经网络

**中图分类号** TP391

## Traffic Flow Forecasting Method Combining Spatio-Temporal Correlations and Social Events

LYU Ming-qi, HONG Zhao-xiong and CHEN Tie-ming

College of Computer Science & Technology, Zhejiang University of Technology, Hangzhou 310023, China

**Abstract** Traffic flow prediction, as a key issue in intelligent transportation system, becomes a research hotspot in the field of transportation both at home and abroad. The main challenge of traffic flow prediction is twofold. First, traffic flow has complicated spatial and temporal correlations. Second, traffic flow can be influenced by social events. Aiming at these challenges, this paper proposes a deep learning framework for traffic flow prediction. On the one hand, a sub-network by combining graph convolutional neural network and recurrent neural network is designed to extract spatio-temporal correlation features from the non-European road network space. On the other hand, a sub-network based on convolutional neural network is designed to extract social event features from textual data. Finally, the traffic flow prediction model is implemented by merging the spatio-temporal correlation feature extraction subnetwork and the social event feature extraction sub-network. In order to verify the validity of the model, experiments are conducted based on real traffic flow data. Compared with the baseline methods, the proposed method has higher accuracy, and the accuracy improves by 3% to 6%.

**Keywords** Traffic flow forecasting, Graph convolutional network, Recurrent neural network, Social events, Convolutional neural network

### 1 引言

近年来,汽车增长速度快、道路容量不足、交通管理水平低等现状使得交通拥堵问题愈演愈烈。在交通容量一定的情况下,智慧交通系统(Intelligent Transportation System, ITS)是解决交通拥堵问题的有效手段。交通流预测是ITS中的关键问题之一。准确、实时的交通流预测信息可以帮助驾驶员

进行路线规划<sup>[1]</sup>,提高交通效率和安全性<sup>[2]</sup>,进而避免或者缓解交通拥堵。

交通流预测是一项极具挑战性的任务,主要原因有:1)复杂的时空关联。交通流在空间上受道路网络拓扑结构的限制,在时间上具有邻近性、周期性等多种时序特性<sup>[3]</sup>。2)多样的影响因素。交通流受天气、假期、特殊事件等多种因素的影响<sup>[4]</sup>。

到稿日期:2020-03-16 返修日期:2020-08-28 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:国家自然科学基金联合重点项目(U1936215);浙江省自然科学基金(LY18F020033);工业互联网创新发展工程项目(TC190H3WN, TC200H01V)

This work was supported by the Joint Funds of the National Natural Science Foundation of China(U1936215), Zhejiang Provincial Natural Science Foundation (LY18F020033) and Innovative Development Project of the Industrial Internet (TC190H3WN, TC200H01V).

通信作者:陈铁明(tmchen@zjut.edu.cn)

针对第一项挑战,当前方法主要利用深度学习强大的特征学习能力从交通流数据中自动挖掘出时空关联特征。交通流中最显著的关联为时间关联性,因此长短期记忆网络(Long Short-Term Memory, LSTM)、门控循环单元(Gated Recurrent Unit, GRU)等循环神经网络(Recurrent Neural Network, RNN)是较早用于交通流预测的深度学习模型<sup>[5-7]</sup>。然而,由于RNN无法发现空间关联,因此这些方法大多面向单一道路,无法针对整个道路网络进行优化。针对此问题,少量研究将RNN与能够发现空间关联的卷积神经网络(Convolutional Neural Network, CNN)联合使用<sup>[8-12]</sup>。然而, CNN只能处理欧氏空间的网格型数据,不能适应道路网络这种非欧氏空间的数据。因此,如何针对道路网络这种非欧氏空间的数据提取时空关联特征是交通流预测首先需要解决的问题。

针对第二项挑战,当前方法大多只考虑时间、日期、天气等比较容易获取的因素<sup>[13-16]</sup>。然而,社会事件是影响交通流变化的一个重要因素。例如,交通事故、节日聚会、演唱会等会导致目标区域周围的交通流在几个小时内发生急剧变化。与时间、日期、天气等简单因素不同,社会事件的获取非常困难。此外,即使获取到社会事件,如何合理地将其融入交通流预测模型也是需要解决的一个难题。

针对上述挑战,本文提出了一种融合交通流时空关联和社会事件因素的交通流预测方法。首先,针对CNN无法从道路网络这种非欧氏空间数据中提取空间关联的问题,采用图卷积网络(Graph Convolutional Network, GCN)作为空间关联特征提取器,并结合LSTM提取时间关联特征。然后,可以从互联网上发布的相关文本数据(如位置社交网络、事件发布网站)中获取社会事件,采用TextCNN<sup>[17]</sup>从这些文本数据中抽取社会事件相关特征。最后,融合交通流时空关联特征和社会事件特征,并通过全连接层融合简单的全局特征(如日期)构建交通流预测模型。

## 2 相关工作

交通流预测指分析城市道路网络上的交通状态(如速度、流量和密度)并预测道路网络上未来交通状态的过程。交通流预测作为ITS的一个关键问题和重要分支,是国内外交通领域的研究热点。近几年来,愈来愈多的研究者投身于交通流预测领域,提出了非常多的预测方法和模型。

基于交通理论的方法是早期用于交通流预测的方法,通过利用各类物理学模型对交通流进行模拟,进而实现预测,如排队论模型<sup>[18]</sup>、元胞传输模型<sup>[19]</sup>、基本图模型<sup>[20]</sup>。然而,一方面,这些物理学模型对环境的假设过于理想,难以适应纷繁复杂的城市环境;另一方面,物理学模型中涉及大量在真实环境中难以获取和设置的参数,因此通常只用于仿真环境。

数据驱动的方法是当前用于交通流预测的主流方法,而基于统计理论的方法是较早用于交通流预测的数据驱动方法,其通过对历史交通流数据进行统计分析来实现对未来交

通流的预测,如历史平均模型<sup>[21]</sup>、卡尔曼滤波<sup>[22]</sup>、自回归移动平均模型(Autoregressive Integrated Moving Average Model, ARIMA)<sup>[23-24]</sup>。然而,基于统计理论的方法大多是基于线性假设,对非线性交通流的预测性能较差,且在挖掘交通流与各类影响因素之间的潜在关联方面能力不足。针对此问题,基于传统机器学习的交通流预测方法可学习到交通流与影响因素之间的非线性关联,大大提升了交通流预测的性能,如支持向量回归(Support Vector Regression, SVR)<sup>[25]</sup>、贝叶斯模型<sup>[26]</sup>、 $k$ 近邻模型<sup>[27]</sup>、模糊逻辑回归<sup>[28]</sup>、随机森林模型<sup>[29]</sup>。然而,传统机器学习依赖于对特征的人工定义,而人工定义的特征往往难以捕捉过于复杂的交通流时空关联,造成信息损失。

深度学习模型可以自动提取、捕捉数据相关性的特征,克服了传统机器学习需要人工定义特征的不足。深度学习模型在交通流预测领域显示出了优越的性能。RNN是处理时序数据的首选网络,被广泛应用于交通流预测<sup>[5,7,30]</sup>,但RNN无法捕捉交通流的空间关联。为了表征交通流的空间关联,研究人员利用CNN捕获交通流的空间局部特征,将RNN与能够发现空间关联的CNN联合使用<sup>[8-12]</sup>。然而,CNN最初是为欧氏域中的空间结构设计的,这些方法必须将交通网络转化为规则的网格,而交通数据是分布在道路网络上的,其结构是非欧氏的。由于数据建模的折衷,CNN不能完全捕获交通数据的空间关联。另一方面,GCN是一种基于谱图理论将卷积操作推广到非欧氏域<sup>[31-33]</sup>的模型。因此,最近的一些研究将GCN应用于道路网络的空间建模。例如,Li等<sup>[34]</sup>提出了一种用于交通流预测的深度学习框架,通过在图上使用双向随机游走来捕获空间相关性,并使用编码器-解码器体系结构捕获时间相关性。Zhang等<sup>[35]</sup>将GCN集成到一个序列框架中,以进行多步的速度预测。然而,现有的基于深度学习的交通流预测研究都聚焦在交通流时空特征上,且复杂的城市环境使得事件对交通流的影响愈发明显,因此事件是预测交通流的一个重要特征。如何有效地将事件特征融入深度学习网络实现事件敏感的交通流预测,是一个值得研究的方面。

## 3 方法

### 3.1 定义与问题描述

**定义 1(道路网络)** 道路网络用加权图  $Gr=(V, E, W_r)$  表示,  $V$  中的每个节点  $v_i$  代表一条道路, 每条边  $e_{ij}$  表示  $v_i$  和  $v_j$  之间的相关性。边  $e_{ij}$  的权重  $w_r(i, j)$  可由式(1)计算得出, 其中  $len(v_k)$  为道路  $v_k$  (或路径  $v_k$ ) 的长度。可以看出,  $w_r(i, j)$  考虑的是道路之间在道路网络中的加权距离, 距离越小权重越大。

$$w_r(i, j) = \frac{\text{median}\{len(v_k) \mid v_k \in V\}}{\text{len}(\text{the shortest path from } v_i \text{ to } v_j)} \quad (1)$$

**定义 2(交通流样本)** 时刻  $t$  道路网络的交通流数据(如车速、车流量)可表示为  $N$  维向量  $x_t$  ( $N$  是道路数量), 则交通流样本可表示为  $S_t = (\mathbf{In}_t, \mathbf{Out}_t)$ , 其中, 输入部分  $\mathbf{In}_t = [x_{t-w-1}, \dots, x_{t-1}, x_t]$  是  $W$  个历史交通流数据的序列, 输出部

分  $Out_t = x_{t+h}$  是时刻  $t+h$  的交通流数据。

交通流预测问题可以看作从大量交通流样本中学习得到函数  $f$ , 将包括当前  $t$  时刻的  $W$  个历史交通流数据映射到时刻  $t+h$  时刻的未来交通流数据, 如式(2)所示:

$$\tilde{X}_{t+h} = f(G_r; (X_{t-w-1}, \dots, X_{t-1}, X_t)) \quad (2)$$

方法的总体架构如图 1 所示, 其本质是一个深度神经网络, 主要由两部分组成: 交通流时空关联挖掘子网络和社会事件特征抽取子网络。其中, 交通流时空关联挖掘子网络基于 GCN+LSTM 的网络结构, 社会事件特征抽取子网络基于 CNN 的网络结构。两个子网络的输出与简单的全局特征融合后, 输入全连接神经网络即可实现交通流预测。

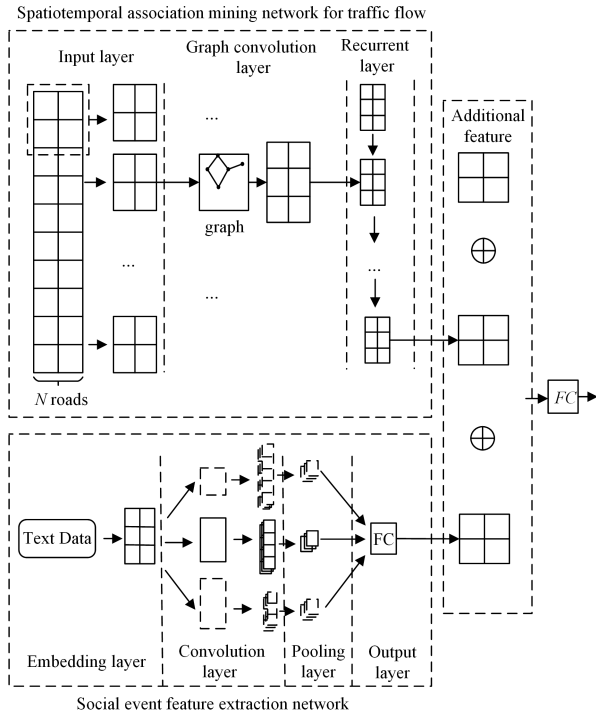


图 1 本文方法的总体架构

Fig. 1 Overall framework of proposed method

### 3.2 交通流时空关联挖掘子网络

交通流时空关联挖掘子网络基于 GCN+LSTM 的网络结构, 具体包括输入层、图卷积层和循环层。

#### 3.2.1 输入层

给定一个交通流样本  $S_t$ , 其输入部分  $In_t$  可以看作一个  $N \times W$  的矩阵, 其中  $N$  是道路数量,  $W$  是历史交通流数据的数量。使用一个大小为  $w$ 、步长为  $d$  的滑动窗口来处理  $In_t$ , 得到一个长度为  $K$  的序列(即  $S_t^1, S_t^2, \dots, S_t^K$ ), 序列中每一个元素都是一个  $N \times w$  矩阵, 为深度学习网络处理的基本单元。

使用滑动窗口处理原始交通流样本的原因如下: 首先, 时序数据具有邻近性特点, 即时间上接近的数据之间的关联较强<sup>[8]</sup>, 这些关联需要用卷积层从样本内部进行抽取, 因此每个待处理的基本单元应包含多个交通流数据(即  $w > 1$ ), 若  $w = 1$ , 则会忽略这些内部关联。其次, 将交通流样本分割成若干段输入模型(即  $w < W$ ), 这样有利于循环层捕捉样本之间的

外部关联, 若  $w = W$ , 则会忽略这些外部关联。

#### 3.2.2 图卷积层

图卷积层的意图是捕捉交通流数据中的空间关联。卷积是捕捉数据空间关联的有效手段, 但 CNN 只能处理欧氏空间中的数据(即网格型的数据), 不能适应道路网络这种非欧氏空间中的数据。由于道路网络可以表示为图的结构, 图节点之间的边可有效表示道路之间的非欧氏空间关联, 因此本文利用 GCN<sup>[32]</sup>来捕捉道路之间的空间关联。GCN 的传播规则可以由式(3)表示。其中,  $I_N$  是单位矩阵,  $X$  为道路网络图的邻接矩阵,  $\tilde{D}$  是一个对角矩阵  $\tilde{D}[i, j] = \sum_j \tilde{X}[i, j]$ ,  $W^{(l)}$  是第  $l$  层中可训练的权重矩阵。可以堆叠使用多个 GCN 层对更高阶的图节点交互进行建模。GCN 的原始输入是  $H^{(0)}$ , 它是由  $N \times w$  矩阵表示的一个序列。将交通流数据 ( $N \times w$  的矩阵)、道路网络图(即  $G_r$ )的邻接矩阵数据输入 GCN, 经过每层计算后可得到新的特征矩阵  $H^{(l)}$ 。

$$H^{(l+1)} = ReLU(\tilde{D}^{-\frac{1}{2}} \tilde{X} \tilde{D}^{-\frac{1}{2}} H^{(l)} W^{(l)}) \quad (3)$$

#### 3.2.3 循环层

循环层的意图是捕捉交通流数据中的时间关联。本文使用 LSTM 实现循环层, LSTM 可以解决普通 RNN 中梯度消失或爆炸的问题, 以便更好地学习长时间相关性。

基于现有研究的经验<sup>[36]</sup>, 本文选择具有两层堆叠的 LSTM 结构。在 LSTM 层之间应用 dropout 进行正则化。如图 2 所示, 给定被分割成  $K$  段(即  $S_t^1, S_t^2, \dots, S_t^K$ )的交通流样本  $S_t$ , 将每段输入图卷积层后, 产生特征矩阵序列  $H_t^1, H_t^2, \dots, H_t^K$  ( $H_t^K$  是图卷积层对  $S_t^K$  的处理结果), 按时间顺序输入到循环层中。最后, 将 LSTM 最后一个隐状态的输出特征矩阵作为循环层的输出, 表示为  $RZ_t$  (即  $RZ_t = PM^K$ )。

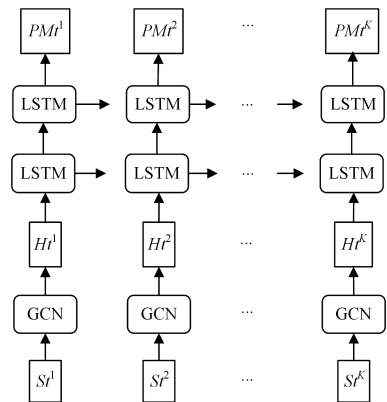


图 2 循环层结构

Fig. 2 Structure of recurrent layer

### 3.3 社会事件特征抽取子网络

社会事件可能会导致目标区域周围的交通流急剧变化。然而, 社会事件缺乏固定的模式和规律(如周期性), 因此仅观测交通流数据几乎不可能事先预测社会事件的发生<sup>[37]</sup>。针对此问题, 本文利用互联网上发布或讨论社会事件的文本资源对社会事件进行抽取。具体来说, 本文利用位置社交网络(如 Twitter)中具有位置和时间属性的短文本来抽取社会事

件,并基于深度神经网络从中抽取特征,以捕捉对未来交通流的影响。

社会事件特征抽取子网络基于 CNN 的网络结构,具体由 4 部分组成:嵌入层、卷积层、池化层和输出层。

### 3.3.1 文本信息预处理

给定一个交通流样本  $S_t$ ,根据文本发布的位置将输入部分  $In_t$  期间在所有道路附近发布的文本片段预处理后聚合成一个文档集  $D_t$ 。 $D_t(i)$  是指在道路  $v_i$  附近发布的文本片段。如果道路  $v_i$  附近无文本片段,则在  $D_t(i)$  中填入空字符串。其中,文本片段的预处理过程包括:1)文本去噪。删除网页中的标签信息,转化为纯文字。2)去除停用词。删除停用词(如英文的 of, the 等,中文的“在”“的”等)。3)中文分词。如果文本是中文,还需要进行分词操作。4)由于任何事件讨论或发布网站中都会包含大量与社会事件无关的文字信息,通常只有少量是真正关于社会事件的,因此在生成  $D_t$  之前,通常需要进行过滤操作(如采用文本分类模型)。

### 3.3.2 嵌入层

使用预训练好的词向量作为嵌入层,在模型训练过程中通过反向传播算法进行调整。词嵌入可将一个词表示为一个低维稠密的向量,使得语义相似的词在向量空间中的距离也相近。本文首先将预处理好的文本数据转化为词索引序列,词索引就是为每一个词依次分配一个整数 ID,相同的词指向相同的整数 ID。然后,使用 GloVe 算法<sup>[38]</sup>将每个词表示为一个  $q$  维向量。经过词嵌入,一个文本数据被映射成一个  $l \times q$  的矩阵  $\mathbf{A}$ ,  $l$  为文本数据的最大长度(可能会对文本矩阵进行截断或填零操作)。

### 3.3.3 卷积层

句子中相邻的词的关联性较高,而经过预处理和嵌入层之后的文本数据是一个矩阵的形式,因此本文采用 CNN 对这种关联进行捕捉。本文采用大小为  $q \times h$  的卷积核对  $l \times q$  的文本矩阵进行卷积操作。该卷积核的思路与  $h$ -Gram 算法类似,即对  $h$  个词的关联进行捕捉。卷积操作可由式(4)表示:

$$o_i = \mathbf{w} \cdot \mathbf{A}[i:i+h-1], i=1, 2, \dots, l-h+1 \quad (4)$$

其中,  $\mathbf{w}$  表示参数矩阵。卷积之后通常使用激活函数激活,得到所需要的特征,公式如下:

$$c_i = f(o_i + b) \quad (5)$$

本文使用 3 个不同高度的卷积核( $h=3, 4, 5$ )来得到更丰富的特征表示,每个高度的卷积核个数为 100。经过卷积后,得到 300 个一维向量。

### 3.3.4 池化层

本文使用了 max-over-time-pooling,在保持主要特征的情况下,大大降低了参数的数量。max-over-time-pooling 指对卷积层输出的每一个 feature map 取最大值作为保留特征。如图 3 所示,本文对卷积后得到的 300 个一维向量分别取最大值,然后拼接在一起,则池化层输出长度为 300 的一维向量。

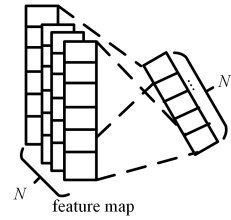


图 3 max-over-time-pooling 操作

Fig. 3 max-over-time-pooling operation

输出层的作用是把前面提取到的特征综合起来,并通过设置神经元个数来控制输出特征向量的维数。

### 3.4 数据融合

数据融合主要是指融合交通流时空关联挖掘子网络和事件特征抽取子网络,以及其他的外部因素,本文主要考虑两种外部因素,即日期特征和周期特征。

**日期特征。**给定交通状态样本  $S_t$ ,日期特征指  $S_t$  所处时刻是星期几、是否是节假日。本文提取  $t+h$  时刻的日期特征,将样本  $S_t$  的日期特征矩阵表示为  $\mathbf{TZ}_t$ 。

**周期特征。**城市道路上的交通流具有很强的周期性(例如,在连续的工作日,高峰时段的交通状态通常是相似的)<sup>[39]</sup>。如果要直接从原始时序数据中学习大尺度的周期特征,输入模型的数据必须跨度很长(例如,要学习以周为尺度的周期特征,则输入模型的数据跨度至少为一周),但过长的数据会因为梯度消失问题而导致 RNN 模型难以训练,因此  $W$ (即交通流样本输入部分的长度)不能设置太长,这使得循环层只能捕捉短期周期特征。针对此问题,在给定交通流样本  $S_t$  的情况下,将  $t+h$  时刻的一天前  $t-n^d$  时刻和一周前  $t-n^w$  时刻的交通流数据视为长期周期特征,其中  $n^d$  和  $n^w$  分别为一天和一周的时间长度。将  $S_t$  的周期特征矩阵表示为  $\mathbf{PZ}_t$ 。

在得到所有特征矩阵之后,将  $\mathbf{RZ}_t, \mathbf{EZ}_t, \mathbf{TZ}_t, \mathbf{PZ}_t$  融合在一起。具体地,将  $\mathbf{RZ}_t, \mathbf{EZ}_t, \mathbf{TZ}_t, \mathbf{PZ}_t$  连成一个长的特征矩阵  $\mathbf{Z}_t = \mathbf{RZ}_t \oplus \mathbf{EZ}_t \oplus \mathbf{TZ}_t \oplus \mathbf{PZ}_t$ 。之后,在  $\mathbf{Z}_t$  上堆叠两个全连接层。最后,使用 sigmoid 作为激活函数,预测  $t+h$  时刻的交通流数据。对模型进行训练,使得预测的交通流数据和真实交通流数据之间的均方根误差(Root Mean Square Error, RMSE)最小。

## 4 实验

### 4.1 实验设置

1)数据集。本文基于真实数据集 PEMS10 进行实验评测。PEMS10 由加利福尼亚交通运输部采集。加利福尼亚州所有主要都会区的高速公路系统上总共部署了 39 000 个传感器。本文在加利福尼亚州第 10 区中随机选择 608 个传感器作为数据源(如图 4 所示)。PEMS10 的时间段为 2018 年 1 月 1 日至 2018 年 3 月 31 日。数据集以 15 min 为间隔进行汇总,并以平均车速作为要预测的交通流数据。PEMS10 包含每个传感器的 8 640 条记录。需要注意的是:首先,

PEMSD10 中的传感器和道路并不是一一对应的(同一条道路可能部署多个传感器),因此,使用阈值高斯核方法<sup>[33]</sup>构建道路网络图,即传感器  $v_i$  和  $v_j$  之间的权重  $w_r(i, j)$  由式(6)计算。其中  $dist(v_i, v_j)$  表示传感器  $v_i$  和  $v_j$  之间的物理距离,  $\sigma$  是距离的标准差,  $\kappa$  是距离阈值。其次,PEMSD10 还从交通事故信息页面收集事故报告。事件报告包含有关事件的日期、时间、位置和文本描述的信息,作为事件挖掘的数据源。

$$w_r(i, j) = \begin{cases} e^{-\frac{dist(v_i, v_j)^2}{\sigma^2}}, & \text{if } dist(v_i, v_j) \leq \kappa \\ 0, & \text{otherwise} \end{cases} \quad (6)$$

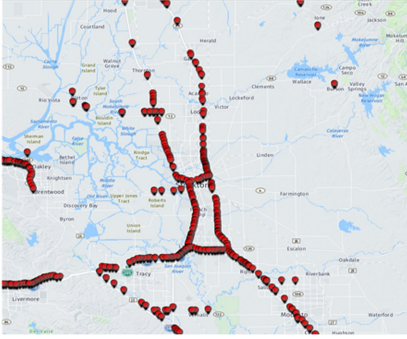


图4 PEMS10 的传感器分布

Fig. 4 Sensor distribution of PEMS10

2) 预处理。首先,采用线性插值的方法来填充平均车速数据的缺失值。其次,在输入模型之前,使用 Min-Max 归一化方法将平均车速值映射到  $[0, 1]$  的范围内。再次,由于日期特征都是类别型变量,因此使用 one-hot 编码来为其形成数字型向量。

3) 参数设置。PEMSD10 的道路数量为  $N=608$ 。设置交通流样本输入部分的长度  $W=96$ (即 24 小时),  $\tau=8$ , 距离阈值  $\kappa=2$  km, 词向量的维数  $q=100$ , 社会事件提取网络输出的特征向量的维数为 50。在实验中,采用了 5 折交叉验证策略。使用 Adam 优化器训练深层神经网络。将学习率设置为 0.001, batchsize 设置为 32, training epoch 设置为 2000。

4) 评估指标。用 RMSE 和平均绝对百分比误差 (Mean Absolute Percentage Error, MAPE) 评估预测方法的性能,其公式如下:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (X_i - \tilde{X}_i)^2} \quad (7)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{X_i - \tilde{X}_i}{X_i} \right| \times 100\% \quad (8)$$

其中,  $\tilde{X}_i$  和  $X_i$  分别是预测值和真实值,  $n$  是所有预测值的数量。

## 4.2 比较实验

为了评测所提方法 (Event Temporal Graph Convolutional Network, Event-TGCN) 的性能,将其与以下基线方法进行比较。

HA: 历史平均模型<sup>[37]</sup>, 使用前几个周期的平均值作为预测。本文将周期设置为一周,因此预测值为前若干周同一时间的平均车速。

ARIMA: 标准的 ARIMA 模型<sup>[40]</sup>。

SVR: 基于 SVR 算法<sup>[41]</sup> 训练交通流预测模型, 实验中使用线性核。

FNN: 基于 FNN 训练交通流预测模型<sup>[42]</sup>。本文使用两个隐藏层, 第一隐藏层包含 64 个单元, 第二隐藏层包含 32 个单元。

LSTM: 使用 LSTM 模型进行交通流预测<sup>[30]</sup>。本文堆叠两个 LSTM 层, 这两个层都包含 32 个单元。

CLTFP: 结合 CNN 和 LSTM 进行交通流预测<sup>[43]</sup>。使用一维 CNN 捕获空间相关性, 然后使用两层 LSTM 挖掘时间相关性。

T-GCN: 即 Event-TGCN 忽略社会事件特征。

这些方法中, HA, ARIMA 和 T-GCN 没有考虑事件文本信息和全局特征, 而 SVR, FNN, LSTM 和 CLTFP 考虑了文本信息和全局特征。所有的基线方法都经过优化, 以输出最佳性能。

具体实验结果如表 1 所列。

表 1 不同方法在平均车速预测上的对比

Table 1 Comparison of different methods for average speed prediction

	Prediction horizon									
	1 hour		2 hours		3 hours		4 hours		5 hours	
	RMSE	MAPE/%	RMSE	MAPE/%	RMSE	MAPE/%	RMSE	MAPE/%	RMSE	MAPE/%
HA	4.18	4.09	4.18	4.09	4.18	4.09	4.18	4.09	4.18	4.09
ARIMA	3.88	3.72	5.01	4.34	5.84	4.74	6.72	5.12	7.66	5.48
SVR	3.49	3.51	3.62	3.69	3.70	3.78	3.75	3.84	3.78	3.89
FNN	3.03	3.23	3.17	3.35	2.77	2.92	2.79	2.92	2.80	2.98
LSTM	2.69	2.85	2.89	3.15	3.09	3.49	3.31	3.67	3.35	3.74
CLTFP	2.83	3.01	2.92	3.13	3.02	3.25	3.02	3.28	3.05	3.31
T-GCN	2.92	3.06	2.94	3.06	2.97	3.15	2.98	3.13	2.99	3.15
Event-TGCN	2.9	3.04	2.92	3.07	2.95	3.07	2.96	3.08	2.97	3.10

从结果可以看出: 1) 基于深度学习的方法(包括 FNN, LSTM, CLTFP, T-GCN 和 Event-TGCN) 与传统方法(包括 HA, ARIMA 和 SVR) 相比, 具有更好的性能, 这表明基于深度学习的方法能够较好地捕捉非线性时空相关性。2) 基于循

环深度学习的方法(包括 LSTM, CLTFP, T-GCN 和 Event-TGCN) 优于 FNN, 这表明交通流具有很强的时间相关性。3) 基于时空深度学习的方法(包括 CLTFP, T-GCN 和 Event-TGCN) 优于 LSTM, 证明了空间相关性在交通流预测中的有

效性。4) 基于图的深度学习方法(包括 T-GCN 和 Event-TGCN)优于 CLTFP,这表明与 CNN 相比,图能够更好地模拟路网中的空间相关性。5) Event-TGCN 优于 T-GCN,这说明社会事件对交通流具有一定的影响。6) 与最佳基线方法相比,Event-TGCN 将预测误差降低了约 3%~6%。这表明,在 Event-TGCN 中,将事件文本信息、全局特征(时间、周期)和交通流数据结合起来,可以获得更好的性能。

**结束语** 本文提出了一种基于深度学习的道路网络交通流预测模型。该模型同时考虑交通流数据的时空关联和外部影响因素,一方面利用图卷积网络提取道路网络空间关联特征,并通过循环神经网络提取时间关联特征,另一方面从互联网上的文本数据中挖掘社会事件特征。最后,通过融合时空关联特征和社会事件特征实现预测模型。本文通过一系列基于真实交通数据集的实验,证明了所提模型的有效性。

本文提出的方法尚存在以下不足,拟在未来工作中进行深入研究。首先,本文实验中采用的事件文本是已经整理好的归属于某条道路的交通相关事件,而实际中无法达到如此理想的条件。因此,未来将研究从互联网数据中挖掘时空交通事件的方法。其次,除道路网络的拓扑结构外,城市环境的复杂性导致存在更多深层的时空关联。因此,未来将研究能够融合更多语义化关联的深度学习模型。

## 参 考 文 献

- [1] YUAN J, ZHENG Y, XIE X, et al. Driving with knowledge from the physical world[C]//Proceeding of KDD, 2011:316-324.
- [2] BELLETTI F, HAZIZA D, GOMES G, et al. Expert level control of ramp metering based on multi-task deep reinforcement learning [J]. IEEE Transactions on Intelligent Transportation Systems, 2017, 99:1-10.
- [3] ZHAO L, SONG Y J, ZHANG C, et al. Temporal graph convolutional network for urban traffic flow prediction method [J]. arXiv:1811.05320v2, 2018.
- [4] XIE P, LI T R, LIU J, et al. Urban flow prediction from spatio-temporal data using machine learning: A survey [J]. Information Fusion, 2020, 59:1-12.
- [5] ZHAO Z, CHEN W, WU X, et al. LSTM network: A deep learning approach for short-term traffic forecast [J]. IET Intelligent Transport Systems, 2017, 11(2):68-75.
- [6] KANG D, LYU Y, CHEN Y. Short-term traffic flow prediction with LSTM recurrent neural network [C] // Proceeding of ITSC, 2017:1-6.
- [7] DUAN Y, LYU Y, WANG F Y. Travel time prediction with LSTM neural network[C]//Proceeding of ITSC, 2016.
- [8] ZHANG J, ZHENG Y, QI D, et al. Predicting citywide crowd flows using deep spatio-temporal residual networks [J]. Artificial Intelligence, 2018, 259(6):147-166.
- [9] YU H Y, WU Z H, WANG S Q, et al. Spatiotemporal Recurrent Convolutional Networks for Traffic Prediction in Transportation Networks [J]. Sensors, 2017, 17(7):1501.
- [10] YAO H X, WU F, KE J T, et al. Deep multi-view spatial-temporal network for taxi demand prediction [J]. arXiv: 1802.08714v2, 2018.
- [11] WANG D J, YANG Y, NING S M. DeepSTCL: A Deep Spatio-temporal ConvLSTM for Travel Demand Prediction[C]// Proceeding of IJCNN, 2018.
- [12] ZHANG J B, ZHENG Y, QI D K. Deep spatio-temporal residual networks for citywide crowd flows prediction[C]// Proceeding of AAAI, 2016.
- [13] HAN F, LI C C, ZHAO N, et al. Modeling the impacts of weather, spatial and temporal factors on traffic operation[C]// Proceeding of Advanced Forum on Transportation of China, 2012.
- [14] KEAY K, SIMMONDS I. The association of rainfall and other weather variables with road traffic volume in Melbourne, Australia [J]. Accident Analysis & Prevention, 2005, 37(1):109-124.
- [15] ZHAO Y, SADEK A W, FUGLEWICZ D. Modeling the impact of inclement weather on freeway traffic speed at macroscopic and microscopic levels [J]. Transportation Research Record: Journal of the Transportation Research Board, 2012, 227(2):173-180.
- [16] DUNNE S, GHOSH B. Weather adaptive traffic prediction using neurowavelet models [J]. IEEE Transactions on Intelligent Transportation Systems, 2013, 14(1):370-379.
- [17] KIM Y. Convolutional Neural Networks for Sentence Classification [C] // Proceeding of the 2014 Conference on Empirical Methods on Natural Language Processing, 2014:1746-1751.
- [18] XU X Y, LIU J, LI H Y, et al. Analysis of subway station capacity with the use of queueing theory [J]. Transportation Research Part C: Emerging Technologies, 2014, 38:28-43.
- [19] WEI P, CAO Y, SUN D. Total unimodularity and decomposition method for large-scale air traffic cell transmission model [J]. Transportation Research Part B: Methodological, 2013, 53(7):1-16.
- [20] XU F F, HE Z C, SHA Z R. Impacts of traffic management measures on urban network microscopic fundamental diagram [J]. Journal of Transportation Systems Engineering & Information Technology, 2013, 13(2):185-190.
- [21] GUO M, XIAO X, LAN J H. Summary of Short-term Forecast Methods for Road Traffic Flow [J]. Automation Technology and Application, 2009, 28(6):8-9, 16.
- [22] GUO J, HUANG W, WILLIAMS B M. Adaptive Kalman filter approach for stochastic short-term traffic flow rate prediction and uncertainty quantification [J]. Transportation Research Part C: Emerging Technologies, 2014, 43(1):50-64.
- [23] KUMAR S V, VANAJAKSHI L. Short-term traffic flow prediction using seasonal ARIMA model with limited input data [J]. European Transport Research Review, 2015, 7.
- [24] WILLIAMS B M, HOEL L A. Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: Theoretical basis and empirical results [J]. Journal of Transportation Engineering, 2003, 129(6):664-672.
- [25] HU W, YAN L, LIU K, et al. A short-term traffic flow forecast-

- ting method based on the hybrid PSO-SVR [J]. *Neural processing letters*, 2016, 43(1):155-172.
- [26] CASTILLO E, JOSÉ M, SANTOS S C. Predicting traffic flow using Bayesian networks [J]. *Transportation Research Part B: Methodological*, 2008, 42(5):482-509.
- [27] CAI P L, WANG Y P, LU G Q, et al. A spatiotemporal correlative k-nearest neighbor model for short-term traffic multistep forecasting [J]. *Transportation Research Part C: Emerging Technologies*, 2016, 62:21-34.
- [28] KANO H, FURUKAWA T, TSUKAHARA S, et al. Short-term traffic prediction using fuzzy C-means and cellular automata in a wide-area road network[C]//*Proceeding of IEEE Intelligent Transportation Systems*. 2005.
- [29] YANG S, QIAN S. Understanding and predicting roadway travel time with spatio-temporal features of network traffic flow, weather conditions and incidents[C]//*Proceeding of Transportation Research Board*. 2018.
- [30] MA X, TAO Z, WANG Y, et al. Long short-term memory neural network for traffic speed prediction using remote microwave sensor data [J]. *Transportation Research Part C: Emerging Technologies*, 2015, 54:187-197.
- [31] BRUNA J, ZAREMBA W, SZLAM A, et al. Spectral Networks and Locally Connected Networks on Graphs [J]. arXiv:1312.6203v3, 2014.
- [32] DEFFERRARD M, BRESSON X, VANDERGHEYNST P. Convolutional neural networks on graphs with fast localized spectral filtering [J]. *Advances in Neural Information Processing Systems*, 2016:3844-3852.
- [33] KIPF T N, WELING M. Semi-supervised classification with graph convolutional networks[C]//*Proceeding of ICLR*. 2017.
- [34] LI Y G, YU R, CYRUS S, et al. Diffusion convolutional recurrent neural network: Data-driven traffic forecasting [J]. Arxiv. org/abs/1707.01926.
- [35] ZHANG Z, LI M, LIN X, et al. Multistep speed prediction on traffic networks: A deep learning approach considering spatio-temporal dependencies [J]. *Transportation Research Part C: Emerging Technologies*, 2019, 105(8):297-322.
- [36] YAO S C, HU S H, ZHAO Y R, et al. DeepSense: A unified deep learning framework for time-series mobile sensing data processing [J]. arXiv:1611.01942v2, 2017.
- [37] GUO M, XIAO X, LAN J H. A summary of the short-time traffic flow forecasting methods [J]. *Techniques of Automation & Applications*, 2009, 28(6):8-9, 16.
- [38] PENNINGTON, JEFFREY, SOCHER, et al. Glove: Global vectors for word representation[C]//*Proceeding of the 2014 Conference on Empirical Methods in Natural Language Processing*. 2014.
- [39] KE J T, YANG H, ZHENG H Y, et al. Hexagon-based convolutional neural network for supply-demand forecasting of ride-sourcing services [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2019, 20(11):4160-4173.
- [40] HAMED M M, AL-MASAEID H R, SAID Z M B. Short-term prediction of traffic volume in urban arterials [J]. *Journal of Transportation Engineering*, 1995, 121(3):249-254.
- [41] WU C H, WEI C C, SU D C, et al. Travel time prediction with support vector regression[C]//*Proceeding of ITSC*. 2003.
- [42] PARK D, RILETT L R. Forecasting freeway link travel times with a multilayer feedforward neural network [J]. *Computer Aided Civil & Infrastructure Engineering*, 2002, 14(5):357-367.
- [43] WU Y, TAN H. Short-term traffic flow forecasting with spatial-temporal correlation in a hybrid deep learning framework [J]. arXiv:1612.01022v1, 2016.



**LYU Ming-qi**, born in 1982, Ph.D, associated professor, is a member of China Computer Federation. His main research interests include data mining and ubiquitous computing.



**CHEN Tie-ming**, born in 1978, Ph.D, professor, Ph.D supervisor, is a member of China Computer Federation. His main research interests include data mining and information security.