

# 基于图卷积网络的交通预测综述

王竟成<sup>1,2</sup>, 张 勇<sup>1,2</sup>, 胡永利<sup>1,2</sup>, 尹宝才<sup>1,2</sup>

(1.北京工业大学信息学部,北京 100124; 2.北京工业大学多媒体与智能软件技术北京市重点实验室,北京 100124)

**摘要:** 交通预测是智能交通系统中的关键问题之一,精准的交通预测对于城市交通运营调整、物流运输产业提质增效以及公众出行规划等交通需求具有重要作用. 近年来,多种用于解决交通预测问题的深度学习的框架已经被提出,其中图卷积网络(graph convolutional network,GCN)及其变体在各类交通预测模型中脱颖而出,取得了可观的准确率. 因此,对基于GCN的交通流预测模型进行归纳总结,从图卷积的基本定义出发,以频域图卷积和空域图卷积为主,介绍GCN的基本原理. 随后,通过对图时空网络、图自编码器以及图注意力网络的介绍,阐明该领域模型的发展历程,分类综述不同预测模型的结构及特点. 在介绍常用交通预测数据集的基础上,以应用研究、模型研究以及多源数据融合为切入点,探讨了未来该领域的研究方向.

**关键词:** 图卷积网络; 交通预测; 深度学习; 人工智能; 图论; 智能交通系统

中图分类号: TP 389.1

文献标志码: A

文章编号: 0254-0037(2021)08-0954-17

doi: 10.11936/bjtxb2020120027

## Survey on Graph Convolutional Neural Network-based Traffic Prediction

WANG Jingcheng<sup>1,2</sup>, ZHANG Yong<sup>1,2</sup>, HU Yongli<sup>1,2</sup>, YIN Baocai<sup>1,2</sup>

(1. Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China;

2. Beijing Key Laboratory of Multimedia and Intelligent Software Technology, Beijing University of Technology, Beijing 100124, China)

**Abstract:** Traffic prediction is one of the essential issues of an intelligent transportation system. An accurate traffic prediction plays a vital role in traffic demands such as transportation operation adjustment, improvement of efficiency in logistics industry and public travel planning. In recent years, various deep learning frameworks for solving this problem have been proposed, and considerable prediction accuracy has been obtained. The traffic flow prediction model was summarized with graph convolutional network (GCN) in this paper, starting from the basic definition of graph convolution, covering frequency and spatial domain graph convolution, and introducing the basic principles of GCN. Subsequently, by introducing spatio-temporal graph neural networks, graph autoencoders, and graph attention networks, the development process of prediction models in this field was clarified, and the structure and characteristics of different prediction models were summarized. Based on the summary of commonly used datasets, the future research directions from the perspectives of application research, model research, and multi-source data fusion were discussed.

**Key words:** graph convolutional network (GCN); traffic prediction; deep learning; artificial intelligence; graph theory; intelligent transportation systems

收稿日期: 2020-12-30

基金项目: 国家自然科学基金资助项目(62072015,U1811463,61632006);国家重点研发计划资助项目(2018YFB1600903)

作者简介: 王竟成(1995—),男,博士研究生,主要从事智能交通、图卷积网络方面的研究, E-mail: wang.jc@emails.bjut.edu.cn

随着城市化进程的加快以及交通数据的爆炸式增长,智慧交通的重要性日益显著,而交通预测则是智慧交通发展的基石.精准的交通预测对于许多实际交通应用至关重要.例如:路面交通的车速及流量预测对于公众出行路径规划以及交通指挥智能调控等需求具有重要意义;轨道交通客流预测对于站点客流压力测试以及线路时刻安排同样至关重要;对于网约车的用车需求预测则可以协助网约车公司动态分配运营车辆,从而实现调度效率最大化.

交通预测问题有多种分类方式,针对不同的预测任务,可概括为流量预测、速度预测、到达时间预测以及交通需求预测等.根据交通预测的实际场景可分类为高峰期预测、平峰期预测以及异常状况下的预测等.对于交通预测问题的划分同时侧面反映了交通预测模型对于实时性、准确性以及异常状况下的可靠性要求.

不同于其他时序预测问题,交通预测的主要挑战在于包含路网结构拓扑连接关系在内的复杂耦合关联关系及其不断变化的时空特征.空间上,路网中相邻节点存在着直接相互影响的强关联关系,非相邻节点在交通流模式上也存在着基于交通出行起止点(origin-destination, OD)的隐含空间关联.时间上,交通数据呈现出极强的时变性与周期性,例如早高峰与晚高峰、周中与周末等.

传统的基于数理统计的方法以及经典机器学习方法往往针对路网中某一观测点的时序数据进行建模分析以及预测.较有代表性的方法有历史平均值(historical average, HA)<sup>[1-2]</sup>、整合移动平均自回归(autoregressive integrated moving average, ARIMA)模型<sup>[3-5]</sup>、卡尔曼滤波模型(Kalman filtering model)<sup>[6]</sup>、非参数回归模型(nonparametric regressive model)<sup>[7]</sup>以及动态模式分解(dynamic mode decomposition)<sup>[8]</sup>等.后续随着深度学习的发展,端到端的卷积神经网络(convolutional neural networks, CNN)<sup>[9]</sup>、循环神经网络(recurrent neural network, RNN)<sup>[10]</sup>及其变体长短期记忆(long short-term memory, LSTM)网络<sup>[11]</sup>以及门控循环单元(gated recurrent units, GRU)<sup>[11]</sup>等网络模型被广泛用于时序预测问题.

上述方法可以捕获时序数据的非线性特征,但针对单个节点的时序预测难以描述节点之间的相互影响.依赖路网结构的交通预测问题显然无法满足对单节点的逐一分析.虽然部分传统方法考虑了交通拓扑结构中观测点之间的连接属性,但并未进行充分利用<sup>[12]</sup>.路网作为一种离散化非规则排列且

易受距离影响的结构,图表示以及图卷积网络(graph convolutional network, GCN)<sup>[13]</sup>被自然地应用于交通预测问题且取得了优异的结果.

在交通预测领域,由于交通数据特有的时序及空间特征,诞生了许多基于图卷积的方法模型<sup>[14]</sup>,实现了对交通数据时空特征的提取.本文针对基于GCN的交通流预测模型进行归纳总结.

## 1 图卷积网络概述

### 1.1 符号定义

首先给出本文以及相关文献中常见通用的符号定义.

表1 符号定义  
Table 1 Symbol definition

符号	表示含义
$G$	图
$V$	节点集
$v$	节点
$n$	节点数
$E$	边集
$e$	边
$I_n \in \mathbb{R}^{n \times n}$	单位矩阵
$A \in \mathbb{R}^{n \times n}$	邻接矩阵
$H \in \mathbb{R}^{n \times e}$	关联矩阵
$D_e \in \mathbb{R}^{e \times e}$	边度矩阵
$D_v \in \mathbb{R}^{n \times n}$	节点度矩阵
$W \in \mathbb{R}^{n \times n}$	权重矩阵
$L \in \mathbb{R}^{n \times n}$	拉普拉斯矩阵
$\Lambda \in \mathbb{R}^{n \times n}$	特征值矩阵
$U \in \mathbb{R}^{n \times n}$	特征向量矩阵
$X \in \mathbb{R}^{t \times n \times D}$	交通数据输入矩阵

一般地, $G = (V; E; W)$ 用于表示交通预测问题中常用的无向带权图<sup>[15]</sup>. $A$ 表示图的邻接矩阵(adjacency matrix),定义节点之间的相互连通关系. $H$ 为关联矩阵(incidence matrix),用于定义图中边与节点之间的连接,其中元素的定义为

$$h_{jk} = \begin{cases} 1, & v_j \in e_k \\ 0, & \text{其他} \end{cases} \quad (1)$$

对于图中的节点,其度矩阵 $D_v$ 中元素定义为连接到该节点所有边的权重之和,即

$$d(v) = \sum_{e \in E} w(e_i) h(v, e_i), i \in I \quad (2)$$

相似地, $D_e$ 中元素定义为连接到该边的节点数

量之和,即

$$d(e) = \sum_{v \in V} h(v, e_i), i \in I \quad (3)$$

## 1.2 图卷积

传统卷积神经网络局限于对欧氏空间数据进行建模,而图卷积操作利用图表示对非欧氏空间数据进行处理,使卷积操作更加适用于交通数据结构.图卷积操作主要包括基于空域以及频域的图卷积2类.空域图卷积将图卷积操作定义为图中相邻节点之间的特征信息的聚合,基于频域的图卷积则利用图信号处理(graph signal processing),引入滤波器实现频域图卷积的推导.

### 1.2.1 空域图卷积

与深度学习中常用的例如图像等欧氏空间的卷积相比,图结构中由于节点数量及连接关系的不确定性,难以通过固定大小可学习的卷积核对特征进行提取.针对此问题,空域图卷积从节点域出发,通过特征的聚集直接在图上定义卷积操作以找到适用于图的可学习图卷积核.该过程将某中心节点与周围邻居节点的信息通过定义的聚合函数进行聚合,实现中心节点的特征更新.空域图卷积的通用聚合函数在消息传递神经网络(message passing neural network, MPNN)<sup>[16]</sup>中的定义被广泛应用.该网络提出一种空域图卷积的形式化框架,即空域卷积被分解为消息传递与状态更新2个过程,分别表示为 $M_l(\cdot)$ 与 $U_l(\cdot)$ .该框架可表示为

$$h_v^{l+1} = U_{l+1} \left( h_v, \sum_{u \in ne[v]} M_{l+1}(h_v^l, h_u^l, x_{uv}) \right) \quad (4)$$

式中: $u$ 和 $v$ 均为图中节点; $h^l$ 为图卷积在第 $l$ 层的特征,即该框架中各节点在第 $l$ 层的信息通过 $M_l(\cdot)$ 函数聚合后通过 $U_l(\cdot)$ 函数实现第 $l+1$ 层的特征更新.MPNN实现了对于整张图进行空域卷积操作,但面对庞大的图时全图卷积的方法需要占用过多计算资源.GraphSAGE<sup>[17]</sup>则针对此缺陷利用节点邻居的特征信息通过采样和聚集进行学习.该模型通过训练聚合器函数实现节点邻域的信息聚合.来自给定节点的不同跳数的或深度的信息可通过聚合器进行特征抽取并计算节点处的损失,实现节点特征的嵌入.GraphSAGE中提出了3种聚合器,分别为均值聚合器、LSTM聚合器以及池化聚合器.以LSTM聚合器为例,其表达式为

$$AGG_l^p = \max \{ \sigma(\mathbf{W}_p \mathbf{h}_{u_i}^l + \mathbf{b}), \forall u_i \in \mathcal{N}(u) \} \quad (5)$$

该聚合器将节点的前一层表示与聚合的邻居信息进行级联.Liu等<sup>[18]</sup>则通过GraphSAGE模型的均值聚合器进行了路面交通的短期车流速度预测.其

均值聚合器表示为

$$h_v^l = \sigma(\mathbf{W} \cdot \text{MEAN}(\{h_v^{l-1}\} \cup \{h_u^{l-1}, \forall u \in \mathcal{N}(v)\})) \quad (6)$$

节点 $v$ 是在节点 $u$ 的周围在固定长度上随机行走的节点.聚合器的权重矩阵由基于图的损失函数通过随机梯度下降进行调整,最终实现使用GraphSAGE模型预测交通流速度.

Li等<sup>[19]</sup>针对路况的动态变化,将交通流建模为有向图上的扩散过程,并提出了基于空域图卷积的扩散卷积神经网络(diffusion convolutional recurrent neural network, DCRNN),基于双向随机游走理论捕获空间相关性,实现了大规模路网的交通流量预测.Chen等<sup>[20]</sup>也同样在卷积网络中引入了扩散图卷积进行空域上的图卷积操作.因此,Chen等<sup>[20]</sup>为了捕获交通网络的连通性和全局性,提出了基于残差递归架构的网络模型Res-RNN,实现路面交通流预测.Song等<sup>[21]</sup>在图的邻接矩阵中增加了前后时刻的时空关联关系,并利用空域图卷积层与门控线性单元的多个组合网络构建其整体预测框架.Bruna等<sup>[22]</sup>则认为显式图结构难以表示真实的图中节点关系,并提出了基于节点嵌入学习的图神经网络架构Graph WaveNet.该网络架构中图卷积层即通过空域图卷积提取节点间的结构特征.

### 1.2.2 频域图卷积

空域图卷积借鉴了欧氏空间的卷积,频域图卷积则从信号处理的角度利用图信号的傅里叶变换实现卷积操作.频域图卷积最早由Bruna等<sup>[23]</sup>提出.对于图 $G$ ,其拉普拉斯矩阵可表示为 $\mathbf{L} = \mathbf{D}_v - \mathbf{A}$ .归一化后拉普拉斯矩阵表示为

$$\mathbf{L} = \mathbf{I}_n - \mathbf{D}_v^{-\frac{1}{2}} \mathbf{A} \mathbf{D}_v^{-\frac{1}{2}} \quad (7)$$

因拉普拉斯矩阵 $\mathbf{L}$ 为半正定的实对称矩阵,可将特征分解为 $\mathbf{L} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^T$ .与欧氏空间的卷积操作相似,对于图信号 $\mathbf{x}$ 及图卷积核 $\mathbf{g}$ 的图卷积可表示为

$$\mathbf{x} * {}_c \mathbf{g} = \mathbf{U}(\mathbf{U}^T \mathbf{x} \circ \mathbf{U}^T \mathbf{g}) = \mathbf{U}(\mathbf{U}^T \mathbf{g} \circ \mathbf{U}^T \mathbf{x}) \quad (8)$$

式中: $\circ$ 为哈达玛乘积; $*_c$ 特指图卷积操作.将 $\mathbf{U}^T \mathbf{g}$ 作为可训练的图卷积核 $\mathbf{g}_\theta$ ,图卷积操作可简化为

$$\mathbf{x} * {}_c \mathbf{g} = \mathbf{U} \mathbf{g}_\theta \mathbf{U}^T \mathbf{x} \quad (9)$$

在训练频域卷积神经网络的过程中,由于拉普拉斯矩阵的所有特征值以及特征向量的计算量过大,所以为加速特征矩阵的求解,切比雪夫网络(ChebNet.)<sup>[24]</sup>以及一阶切比雪夫网络(1stChebNet.)<sup>[25]</sup>应运而生.

切比雪夫网络首先通过切比雪夫多项式

$$\mathbf{g}_\theta(\Lambda) = \sum_{k=0}^{K-1} \theta_k T_k(\tilde{\Lambda}) \quad (10)$$

对特征值矩阵进行近似求解。式中:  $\theta \in \mathbb{R}^K$  为切比雪夫多项式系数向量;  $T_k(\tilde{\Lambda}) \in \mathbb{R}^{n \times n}$  为  $\tilde{\Lambda}$  的  $k$  阶切比雪夫多项式, 而  $\tilde{\Lambda} = 2\Lambda/\lambda_{\max} - I_n$ , 即值域归一化为  $[-1, 1]$  的特征值对角矩阵,  $\lambda_{\max}$  为特征值的最大值。对于  $k$  阶切比雪夫多项式, 有  $T_k = 2\mathbf{x}T_{k-1}(\mathbf{x}) - T_{k-2}(\mathbf{x})$ , 其中递推公式的前 2 项为  $T_0(\mathbf{x}) = 1$  以及  $T_1(\mathbf{x}) = \mathbf{x}$ 。经切比雪夫多项式近似后图上卷积操作可定义为

$$\mathbf{x} * {}_c\mathbf{g}_\theta = \mathbf{U} \left( \sum_{i=1}^K \theta_i T_k(\tilde{\Lambda}) \right) \mathbf{U}^T \mathbf{x} \approx \sum_{i=1}^K \theta_i T_k(\tilde{\mathbf{L}}) \mathbf{x} \quad (11)$$

式中  $\tilde{\mathbf{L}} = 2\mathbf{L}/\lambda_{\max} - I_n$ 。通过切比雪夫多项式近似后可将图卷积操作的复杂度从  $\mathcal{O}(n^2)$  降至  $\mathcal{O}(K|E|)$ , 其中  $K$  表示拉普拉斯算子的  $K$  阶多项式展开。借助切比雪夫网络的效率提升, Fang 等<sup>[26]</sup> 在此基础上提出了全局时空网络 (global spatial-temporal network, GSTNet), 利用切比雪夫网络的图卷积实现全局空间特征提取, 并在地铁、公交以及出租车 3 组数据集上实现交通流预测。Diao 等<sup>[27]</sup> 考虑到切比雪夫图卷积操作中固定的拉普拉斯矩阵难以捕获动态的空间变化, 因此, 设计了动态拉普拉斯矩阵估计器, 并将张量分解纳入其中, 构建了相应的动态时空图卷积神经网络 (dynamic spatial-temporal graph convolutional neural networks, DGCNN)。Cui 等<sup>[28]</sup> 则探讨了交通图卷积与频域图卷积的区别与联系。一般频域图卷积只捕捉了单跳 (one-hop) 邻居节点之间的关系, 而交通图上节点之间的关系更为复杂。Cui 等<sup>[28]</sup> 进而在此基础上研究了基于切比雪夫网络的  $k$  跳 ( $k$ -hop) 频域图卷积, 并实现了路面车速的预测。

在切比雪夫网络的基础上, 当切比雪夫多项式仅做一阶展开, 切比雪夫网络即简化为一阶切比雪夫网络, 即  $K=1$  且  $\lambda_{\max}=2$  时, 有  $\tilde{\mathbf{L}} = \mathbf{L} - I_n$ , 图卷积操作可简化为

$$\mathbf{x} * {}_c\mathbf{g}_\theta \approx \theta_0 \mathbf{x} - \theta_1 (\mathbf{L} - I_n) \mathbf{x} = \theta_0 \mathbf{x} - \theta_1 \mathbf{D}^{-\frac{1}{2}} \mathbf{A} \mathbf{D}^{-\frac{1}{2}} \mathbf{x} \quad (12)$$

式中  $\theta_0$  与  $\theta_1$  均为全图共享的可学习参数。在实际操作中, 为了进一步限制参数的数量以解决过拟合问题, 令  $\theta = \theta_0 = -\theta_1$ , 图卷积操作可进一步简化为

$$\mathbf{x} * {}_c\mathbf{g}_\theta \approx \theta (\mathbf{I}_n + \mathbf{D}^{-\frac{1}{2}} \mathbf{A} \mathbf{D}^{-\frac{1}{2}}) \mathbf{x} \quad (13)$$

为了避免深度图神经网络中梯度爆炸、消失以及数值不稳定的问题, 一阶切比雪夫网络中利用再归一化的方法, 用  $\tilde{\mathbf{D}}^{-\frac{1}{2}} \tilde{\mathbf{A}} \tilde{\mathbf{D}}^{-\frac{1}{2}}$  代替  $\mathbf{I}_n + \mathbf{D}^{-\frac{1}{2}} \mathbf{A} \mathbf{D}^{-\frac{1}{2}}$ , 其中  $\tilde{\mathbf{A}} = \mathbf{A} + \mathbf{I}_n$  且  $\tilde{\mathbf{D}}_{ii} = \sum_j \tilde{\mathbf{A}}_{ij}$ 。在此基础上, 对于  $C$  个通道的图信号  $\mathbf{X} \in \mathbb{R}^{C \times F}$ , 其一阶切比雪夫网络的图卷积可表示为

$$\mathbf{Z} = \tilde{\mathbf{D}}^{-\frac{1}{2}} \tilde{\mathbf{A}} \tilde{\mathbf{D}}^{-\frac{1}{2}} \mathbf{X} \boldsymbol{\Theta} \quad (14)$$

式中:  $\boldsymbol{\Theta} \in \mathbb{R}^{C \times F}$  为可学习权重矩阵;  $\mathbf{Z} \in \mathbb{R}^{N \times F}$  为 GCN 向前传播的输出。对于一阶切比雪夫网络, 其卷积操作复杂度为  $\mathcal{O}(FC|E|)$ , 相比于切比雪夫网络实现了进一步的近似优化。Yu 等<sup>[29]</sup> 将一阶切比雪夫图卷积操作与门控线性单元组成时空卷积模块, 并构建了时空图卷积的通用框架。在路面交通预测实验中, 切比雪夫网络表现优异且准确率仍略胜一阶切比雪夫网络一筹。Bai 等<sup>[30]</sup> 通过一阶切比雪夫图卷积构建了门控图卷积模块, 并针对不同时间范围内乘客需求的不同, 实现多步预测建模。Wang 等<sup>[31]</sup> 则为了研究城市出租车的运营模式, 将路网划分成网格, 在利用图卷积实现交通流量预测的同时, 实现交通需求的 OD 预测。

### 1.3 图卷积网络

GCN 即利用空域图卷积或频域图卷积构建的深度网络模型。Kipf 等<sup>[13]</sup> 所提图卷积神经网络模型中输入层与输出层均为图信号数据, 其中隐含层为图卷积层, 图结构在各层中共享, 最终实现节点的聚类或节点值的预测等输出。

空域图卷积网络与频域图卷积网络虽然均不是针对交通预测问题而提出的, 但由于交通数据的天然图结构属性, 令 GCN 在交通预测领域展现出优于传统方法的高效率与高准确率。无论是空域图卷积还是频域图卷积, 本质上都是对图拓扑结构的建模利用, 即空间特征的提取。例如 Lee 等<sup>[25]</sup> 利用路面传感器间的距离和位置等信息构建非欧氏关系, 并利用 GCN 提取空间特征进行交通速度预测。Geng 等<sup>[32]</sup> 则以相邻街区以及相似功能区等为节点, 构建多种基于图结构的表示, 在此基础上使用多组 GCN 进行空间相关性建模, 实现乘车需求预测。

基于 GCN 的交通预测大都建立在频域图卷积或空域图卷积的基础之上, 通过图结构约束节点间的数据关系, 利用图卷积对非欧氏空间数据进行结构化空间特征提取, 实现交通预测。

## 2 基于 GCN 的交通预测模型

目前,GCN 已成为交通预测研究的基础模型以及实验的基准方法. 由于交通数据本身为时序数据,如何挖掘时序特征且与空间特征融合预测成为用于交通预测的神经网络模型的改进重点. 此外,图注意力机制同样被用于交通预测模型且能显著提升模型性能. 同时,当前研究已不满足静态图、单图的卷积网络,动态图卷积网络、多图卷积网络以及多任务学习的框架不断被提出.

### 2.1 图时空网络

交通数据是标准的时间序列数据,如图 1 所示,即同一数据集内所有数据内容依时间按照统一时间口径进行排列组织. 其中:橙色虚线表示某节点延边对邻居节点产生影响;蓝色虚直线表示某节点交通数据该时刻对下一时刻的影响;蓝色虚曲线表示跨时刻的长远影响. 由于交通数据存在明显的时序特征,许多时序预测模型便被应用于交通预测问题,例如最常用的传统交通预测方法 ARIMA<sup>[33]</sup>. 该方法基于单个交通节点的时序数据构建统计模型,进而分析数据实现预测. 传统机器学习方法,例如支持向量回归(support vector regression, SVR)<sup>[34]</sup>,相较于基于统计的方法能更好地捕获复杂的线性关系. 上述方法大多用于单一时序数据的识别与预测问题,难以对全局交通数据进行端到端的统一整体化分析及预测.

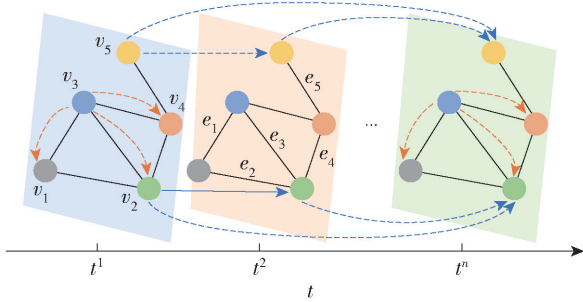


图 1 时序交通数据

Fig. 1 Time series traffic data

图神经网络将空间特征提取蕴于图卷积操作之中,但当提取空间特征与时序特征相结合的方法被用于交通预测问题时,人工智能在该领域的潜力才得以真正释放. 时序特征的提取方法主要有 CNN 以及 RNN.

#### 2.1.1 基于 CNN 的时序特征提取

在 Gehring 等<sup>[35]</sup> 所提出的卷积 Sequence-to-Sequence 模型(convolutional sequence to sequence,

ConvS2S)的基础上, Yu 等<sup>[29]</sup> 代表性地利用其门控线性单元(gated liner units, GLU) 组建时序卷积层并构建时空图卷积网络,实现时序特征的提取.

STGCN 的模型结构见图 2, 2 层时序卷积层与 1 层图卷积层组成了时空卷积模块. 其中时序卷积层将一维因果卷积作用于时间轴,该操作可表示为

$$\Gamma *_{\tau} Y = P \circ \sigma(Q) \in \mathbb{R}^{(M-K_i+1) \times C_o} \quad (15)$$

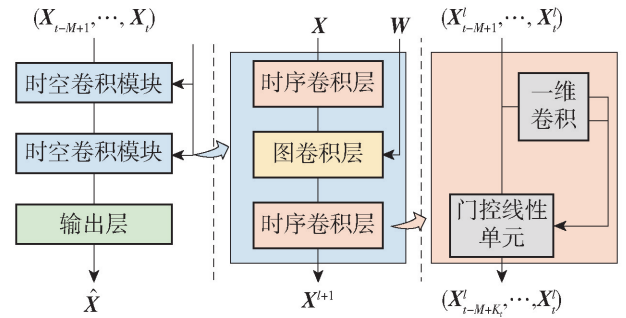


图 2 时空图卷积网络<sup>[29]</sup>

Fig. 2 Spatio-temporal graph convolutional networks<sup>[29]</sup>

式中:  $\Gamma \in \mathbb{R}^{K_i \times C_i \times 2C_o}$  为因果卷积核,  $C_i$  与  $C_o$  分别为输入与输出的通道宽度;  $Y \in \mathbb{R}^{M \times C_i}$  为一维因果卷积的输入序列;  $M$  表示序列的长度;  $P$  和  $Q$  均表示门控线性单元的输入,其长度各为因果卷积输出长度的一半且通道宽度不变. 在此基础上,通过 Sigmoid 作为门限机制的激活函数  $\sigma(Q)$  来控制时序卷积层的输出,实现时序特征的提取. Diao 等<sup>[27]</sup> 同样在其所提框架 DGCNN 的时序卷积层中采取了相似的门控卷积层(gated CNN) 作为时序特征的提取方法.

在利用因果卷积提取交通数据的时序特征的基础之上, Graph WaveNet<sup>[22]</sup> 采用了扩张因果卷积(dilated casual convolution), 亦称作膨胀因果卷积或者空洞因果卷积. 如图 3 所示,扩张因果卷积通过跳过一定的步长在输入序列上沿时间轴进行滑动,从而增加时序卷积操作的感受野. 随着扩张因果卷积层数量的增加,其捕获的时序感受野的范围呈指

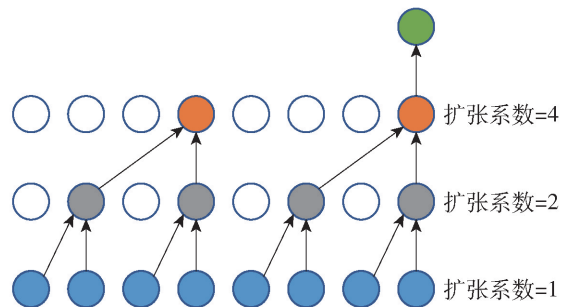


图 3 扩张因果卷积<sup>[20]</sup>

Fig. 3 Dilated casual convolution<sup>[20]</sup>

数级增加. 该卷积操作可表示为

$$\mathbf{x} *_{\tau} \mathbf{f}(t) = \sum_{s=0}^{K-1} \mathbf{f}(s) \mathbf{x}(t-d \times s) \quad (16)$$

式中:  $\mathbf{x} \in \mathbb{R}^T$  为输入序列;  $\mathbf{f} \in \mathbb{R}^K$  表示扩张卷积核;  $d$  表示控制跳过步长的扩张系数. 通过将扩张因果卷积层按照扩张系数递增的顺序进行堆叠, 使网络以更浅的深度捕获更长的序列特征. 但该模型中的门控机制不同于时空图卷积网络, 该模型中不仅用 Sigmoid 作为门控激活函数, 还将门控单元的另一半输入通过双曲正切函数 (hyperbolic tangent, tanh) 进行激活.

Fang 等<sup>[26]</sup> 同样在其所提框架 GSTNet 的时空卷积模块中利用扩张卷积构建了多分辨率时序模块, 即多层因果卷积层的堆叠. Guo 等<sup>[36]</sup> 不仅用卷积提取时空特征, 并在其所提深度时空 3D 卷积神经网络 (deep spatial-temporal 3D convolutional neural networks, ST-3DNet) 中引入了 3D 卷积, 从而实现端到端的交通拥堵状态预测以及人群流量预测. Yao 等<sup>[37]</sup> 针对不平衡空间分布的交通数据, 提出使用少量训练数据实现模型训练, 并通过迁移学习实现不同城市的交通流预测问题. 在其提出的 MetaST 模型中将 CNN 与 LSTM 相结合, 构成其元学习网络模型.

### 2.1.2 基于 RNN 的时序特征提取

同样为了在深度图卷积网络中提取时序特征并用于交通预测, Zhao 等<sup>[38]</sup> 较早地提出了时序图卷积网络 T-GCN. 他们认为 RNN 作为处理序列数据的最广泛使用的模型, 其变体 LSTM 网络与 GRU 能克服训练过程中梯度爆炸与消失的缺陷. LSTM 模型与 GRU 模型均使用门控机制来“记忆”尽可能多的长期信息.

T-GCN 的网络架构如图 4 所示, 其以历史交通数据作为模型的输入, 通过 GCN 以及门控循环单元后获得预测结果. 为了获取交通数据的时序特征, T-GCN 利用了参数较少并且训练更快的门控循环单元而非 LSTM 网络. 其计算过程表示为

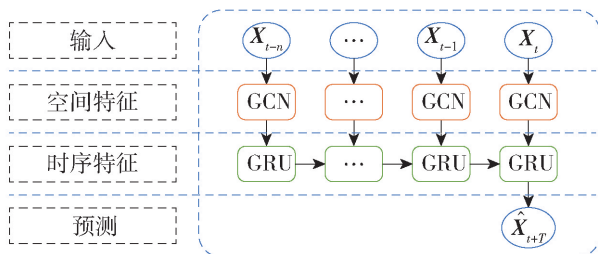


图 4 时序图卷积网络<sup>[38]</sup>

Fig. 4 Temporal graph convolutional networks<sup>[38]</sup>

$$\begin{aligned} \mathbf{u}_t &= \sigma(\mathbf{W}_u[f(\mathbf{A}, \mathbf{X}_t), \mathbf{h}_{t-1}] + \mathbf{b}_u) \\ \mathbf{r}_t &= \sigma(\mathbf{W}_r[f(\mathbf{A}, \mathbf{X}_t), \mathbf{h}_{t-1}] + \mathbf{b}_r) \\ \mathbf{c}_t &= \tanh(\mathbf{W}_c[f(\mathbf{A}, \mathbf{X}_t), (\mathbf{r}_t \circ \mathbf{h}_{t-1})] + \mathbf{b}_c) \\ \mathbf{h}_t &= \mathbf{u}_t \circ \mathbf{h}_{t-1} + (1 - \mathbf{u}_t) \circ \mathbf{c}_t \end{aligned} \quad (17)$$

式中:  $\mathbf{u}_t$  和  $\mathbf{r}_t$  分别表示  $t$  时刻的更新门和复位门;  $\mathbf{h}_t$  和  $\mathbf{h}_{t-1}$  分别表示  $t$  时刻和  $t-1$  时刻的输出;  $\sigma(\cdot)$  表示 Sigmoid 激活函数;  $f(\mathbf{A}, \mathbf{X}_t)$  表示图卷积操作;  $\mathbf{W}$  和  $\mathbf{b}$  则分别表示训练过程中的权重以及偏置项. Chen 等<sup>[20]</sup> 在门控循环单元存储长期依赖关系的基础上, 利用跳跃链接 (hop-links) 的方案捕捉周期性的时间相关性. 同时将残差加入循环图网络中, 从而改善深层网络里长期反向传播 (long-term back-propagation) 的梯度爆炸和消失的问题. Li 等<sup>[39-40]</sup> 则分别在其共享单车需求预测的模型以及出租车需求预测模型中使用了 LSTM 进行时序建模. 该建模过程可描述为

$$\begin{aligned} \mathbf{i}_t^i &= \sigma(\mathbf{W}_i \mathbf{g}_t^i + \mathbf{U}_i \mathbf{h}_{t-1}^i + \mathbf{b}_i) \\ \mathbf{f}_t^i &= \sigma(\mathbf{W}_f \mathbf{g}_t^i + \mathbf{U}_f \mathbf{h}_{t-1}^i + \mathbf{b}_f) \\ \mathbf{o}_t^i &= \sigma(\mathbf{W}_o \mathbf{g}_t^i + \mathbf{U}_o \mathbf{h}_{t-1}^i + \mathbf{b}_o) \\ \boldsymbol{\theta}_t^i &= \tanh(\mathbf{W}_g \mathbf{g}_t^i + \mathbf{U}_g \mathbf{h}_{t-1}^i + \mathbf{b}_g) \\ \mathbf{c}_t^i &= \mathbf{f}_t^i \circ \mathbf{c}_{t-1}^i + \mathbf{i}_t^i \circ \boldsymbol{\theta}_t^i \\ \mathbf{h}_t^i &= \mathbf{o}_t^i \circ \tanh \mathbf{c}_t^i \end{aligned} \quad (18)$$

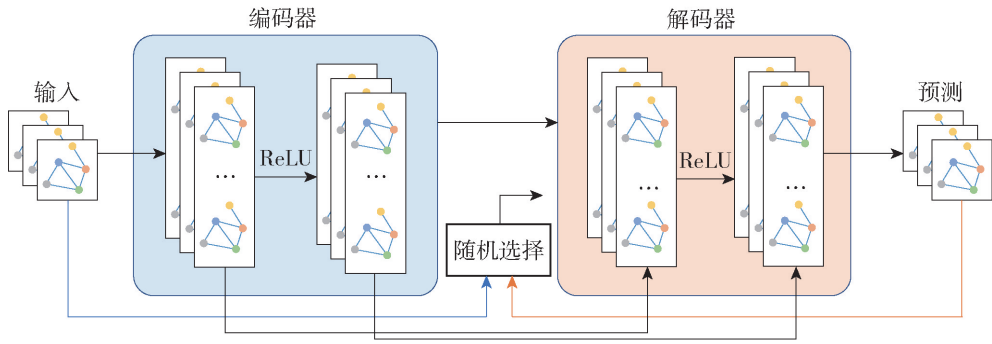
式中  $\mathbf{W}_a$ 、 $\mathbf{U}_a$ 、 $\mathbf{b}_a$  ( $a \in \{i, f, o, g\}$ ) 均为可学习变量. Wang 等<sup>[41]</sup> 针对交通时间预测, 提出了 2 层 LSTM 堆叠的模型, 并考虑道路交叉口以及交通信号灯对预测结果的影响以提升模型性能.

GRU 与 LSTM 作为时序依赖的循环网络建模方法, 在车速预测<sup>[42-43]</sup>、道路占有率预测<sup>[44]</sup>、人群密度预测<sup>[45]</sup>、网约车需求预测<sup>[46]</sup>、出租车需求预测<sup>[47-48]</sup>、出租车 OD 预测<sup>[49]</sup>、公交到站时间预测<sup>[50-51]</sup>以及空气质量预测<sup>[52]</sup>等实际需求问题上均取得了优异的预测效果.

## 2.2 图自编码器

以交通数据等标准的序列化数据作为输入, 自编码器 (auto-encoder, AE) 架构在预测问题上的效果更加稳定且泛化效果更好<sup>[53]</sup>. Li 等<sup>[19]</sup> 所提 DCRNN 模型中, 在使用门控循环单元提取时序特征的基础上利用了自编码器的架构. 其中编码器和解码器都是基于 DCRNN 的递归神经网络. 其模型的基本框架如图 5 所示.

在训练过程中, 历史交通数据序列首先被输入编码器, 并使用其最终状态初始化编码器. 在测试过程中, 预测值被用于替换真实值. 基于图的自编

图5 扩散卷积递归神经网络<sup>[19]</sup>Fig.5 Diffusion convolutional recurrent neural network<sup>[19]</sup>

码器较早由 Kipf 等<sup>[54]</sup>提出,包括变分图自编码器 (variational graph auto-encoders, VGAE) 以及图自编码器 (graph auto-encoders, GAE). VGAE 将变分自编码器迁移到了图领域,以图卷积作为编码操作,用已知图通过编码学习节点向量化表示的分布对其采样,从而得到节点的向量表示,通过解码操作重新构图以实现链路预测等任务. GAE 则包括 2 层 GCN,解码器计算连接 2 个节点边的存在概率进而重构图. Lv 等<sup>[55]</sup>较早地利用自编码器模块化地构建深层网络模型并用于交通流预测. 其提出的堆叠式自编码器 (stacked auto-encoder, SAE) 模型的输入与输出可分别表示为

$$\begin{aligned} y(\mathbf{x}) &= f(\mathbf{W}_1 \mathbf{x} + \mathbf{b}_1) \\ z(\mathbf{x}) &= g(\mathbf{W}_2 y(\mathbf{x}) + \mathbf{b}_2) \end{aligned} \quad (19)$$

式中:  $\mathbf{x}$  为训练样本;  $y(\mathbf{x})$  为编码器的隐藏表示,并于解码器中得到  $z(\mathbf{x})$ ;  $\mathbf{W}_1$  和  $\mathbf{W}_2$  以及  $\mathbf{b}_1$  和  $\mathbf{b}_2$  分别为编码器和解码器的权重矩阵以及偏置向量. 在此基础上, Lv 等<sup>[55]</sup>将多个自编码器堆叠,在获得第 1 个隐藏层后,将第  $k$  个隐藏层的输出作为第  $k+1$  个隐藏层的输入,最后通过输出层实现回归预测. 在此基础上, Zhao 等<sup>[56]</sup>设计并提出了一种分层递归自编码器,使用 3 层堆叠式自编码器架构来获取时序依赖并使用递归神经网络进行预测. 在 Bai 等<sup>[30]</sup>的工作中将捕捉长期时序特征与捕捉短期时序特征的编解码器分别同步进行训练用以缓解基于 RNN 的解码器模型固有的误差积累问题. Lin 等<sup>[57]</sup>则利用基于自编码器架构的图滤波器模型提取共享单车与出租车用车数据的时空特征,并预测共享单车的需求量. 自编码器的架构灵活多变,可基于 CNN<sup>[58]</sup>,亦可基于 LSTM<sup>[59-61]</sup>. 编码器与解码器在功能与架构上相辅相成,在 GEML<sup>[31]</sup>、Forecaster<sup>[62]</sup>、ST-GRAT<sup>[63]</sup>及 ARU<sup>[64]</sup>等模型中均展现出优异的效果.

### 2.3 图注意力网络

注意力机制首先被提出用于自然语言处理<sup>[54]</sup>,如今已被广泛应用到许多序列相关的任务之中. 例如: Liang 等<sup>[65]</sup>利用多级注意力的循环神经网络自适应地调整多个地理传感器采集到的时序数据间的相关性并用于空气质量等预测. 其优点在于可以放大序列中重要部分的影响,将注意力机制引入图神经网络,其可同样受益. GCN 由于依靠拉普拉斯矩阵的特征值,使得卷积操作难以抽离于整体的静态图结构. 在图注意力网络 (graph attention networks, GAT)<sup>[66]</sup>中,不同节点被分配以不同权重,训练过程凭借成对的相邻节点,而非具体图结构,即不依赖于整个图的全部信息. 在时空图卷积网络的基础上,注意力机制被引入并广泛用于时空特征的提取.

在处理空间特征时,以路面交通为例,某道路的交通状态在一定程度上必然受到直接相邻或间接相邻道路的影响. 此种影响的范围以及程度随位置、距离以及时间等因素不断变化. 在 Guo 等<sup>[67]</sup>提出基于注意力的时空图卷积网络 (attention based spatial-temporal graph convolutional networks, ASTGCN) 中,空间注意力矩阵可定义为

$$\mathbf{S} = \mathbf{V}_s \sigma((\mathbf{X}_h^{r-1} \mathbf{W}_1) \mathbf{W}_2 (\mathbf{W}_3 \mathbf{X}_h^{r-1})^T + \mathbf{b}_s) \quad (20)$$

式中:  $\mathbf{X}_h^{r-1}$  为第  $r$  个时空图卷积模块的输入;  $\mathbf{V}_s$ 、 $\mathbf{b}_s$  以及  $\mathbf{W}_n$  ( $n \in \{1, 2, 3\}$ ) 均为可学习变量. 注意力矩阵  $\mathbf{S}$  中元素  $S_{i,j}$  的值表示节点之间的相关性,并使用 Softmax 函数确保注意力权重和为 1.

$$S'_{i,j} = \frac{\exp(S_{i,j})}{\sum_{j=1}^N \exp(S_{i,j})} \quad (21)$$

在进行图卷积操作的过程中,邻接矩阵  $\mathbf{A}$  与注意力矩阵  $\mathbf{S}'$  将同时动态地调整节点之间的权重关系以捕获空间维度上节点之间的动态相关性. Pan 等<sup>[68]</sup>同样认为,两节点之间的空间相关性与其地理

信息相关. 在其所提模型中,提取空间特征的注意力机制与元学习(meta learning)相结合,使其注意力模块的权重从元知识中学习而来. 在用于共享单车需求预测的 STG2Vec<sup>[39]</sup>模型中同样利用注意力机制对空间特征进行提取. Chen等<sup>[69]</sup>则为双分量图卷积(bicomponent GCN)提出了多范围注意力机制(multi-range attention mechanism). 其模型首先分别构建节点图与边缘图,在通过双分量图卷积实现边与节点交互的基础上,利用多范围注意力聚合邻域的信息,从而动态地了解不同聚合范围的重要性. Zhang等<sup>[70]</sup>则在所提门控注意力网络(gated attention networks, GaAN)中利用多头注意力机制(multi-head attention mechanism)对节点及其邻居的特征进行聚合. 与传统多头注意力机制不同, GaAN使用了卷积子网(convolutional sub-network)来控制每个注意力头的重要性,并用以解决交通速度预测问题.

与空间特征方法类似,在处理时序特征时,注意力机制被作用于时间维度以对长期序列数据进行建模. 例如: Yao等<sup>[71]</sup>为解决动态的空间依赖以及周期性的时间依赖,设计了一种周期性转移的注意力机制来处理长期周期性时序偏移. 在同一模型中将注意力机制同时用于空间特征以及时序特征的提取已被证明卓有成效<sup>[67,72-75]</sup>.

注意力机制可被用于辅助时空特征的提取,也可以摒弃卷积操作,借鉴 Transformer<sup>[76]</sup>架构仅利用注意力机制完成交通预测任务. 在图变换网络(graph transformer networks, GTN)<sup>[77]</sup>相关工作中提出,传统 GCN 的结构固定,即其中邻接矩阵为图的固有属性不随模型训练过程而改变. GTN 则能够挑选有价值的多跳连接的元路径(meta-path),从而实现基于可学习图结构的预测.

GCN 无论是引入注意力机制还是与 Transformer 架构的结合<sup>[78]</sup>,目的均是更有效地提取空间特征与时序特征. Li等<sup>[62]</sup>利用该框架提出了 Forecaster 模型. 从学习图结构开始,将学习到的结构表示不同位置数据之间的空间依赖,再基于图拓扑对 Transformer 进行稀疏处理以提升模型强度,从而实现精准的出租车需求预测.

### 3 公开数据集

鉴于大范围跨时空交通数据采集的难度,高质量的数据集对于交通预测问题的研究至关重要. 交通数据类别丰富,包括但不限于交通运输数

据、交通管理数据以及用于辅助的气象数据和事件数据等.

其中交通运输数据主要采集于路面交通、轨道交通以及航运交通的实际运营过程,例如路面车流量、路面车速、路面车道占有率、出租车需求量、网约车需求量、共享单车轨迹、公交车客流量、公交车到站时间、地铁进出站客流量以及地铁换乘客流量等. 交通管理数据则包括信号灯调度数据、潮汐车道数据以及私家车限号数据等. 在此基础上,气象数据与交通数据密切相关,主要源自气象部门的信息化系统,包括气象卫星以及地面气象监测站等获取的数据,其中重点关注雨、雪、雾、霾等各种恶劣天气对交通情况的影响. 事件数据则包括大型活动数据以及交通事故数据等. 本文将重点介绍数个目前应用较多的全球公开交通数据集.

#### 1) PeMS

PeMS 为加利福尼亚运输局的性能测量系统(performance measurement system)的缩写. 该数据集的内容由 39 000 余个独立传感器以 5 min 为时间间隔实时收集,其范围覆盖了加利福尼亚州所有主要城市区域的高速公路系统. 该数据集可由其官方网站 <http://pems.dot.ca.gov/> 获得. 由于 PeMS 全部数据体量较大,目前衍生出数个常用子数据集,包括 PeMS-03、PeMS-04、PeMS-07、PeMS-08、PeMS-SF 以及 PeMS-BAY 等. 子数据集覆盖了不同大小的区域,包含不同数量传感器采集的不同时间跨度以及不同时间粒度的交通流信息.

#### 2) METR-LA

该数据集采集自洛杉矶高速路的 207 个传感器,每组数据的时间间隔为 5 min,该数据由其官网 <https://www.metro.net/> 获得,并用于 DCRNN<sup>[19]</sup>模型的实验测试中首次使用.

#### 3) LOOP<sup>[28]</sup>

该数据集由部署在华盛顿州西雅图地区的 4 条相连的高速公路上的 323 个地感线圈收集而来. 该数据集包含 2015 年全年的速度数据,时间间隔为 5 min. 数据地址: <https://github.com/zhilyongc/Seattle-Loop-Data>.

#### 4) NYC Taxi

该数据集由纽约出租车与轿车委员会(Taxi and Limousine Commission)提供,收集了纽约出租车的行程记录,包括上下车的日期与时间、上下车的位置、行程距离、分项票价、费率类型、付款类型以及乘客人数等信息. 数据集所含数据的时间跨度为



2009—2020年且仍在不断更新之中. 该数据集可通过网址 <https://www1.nyc.gov/site/tlc/about/tlc-trip-record-data.page> 获得.

## 4 未来研究方向

随着智慧交通以及深度学习的发展,近年交通预测领域已取得长足的进步,但目前仍存在许多富有研究价值与意义的问题和挑战. 本文将从应用、模型以及多源数据的角度分类讨论未来研究的方向,以及目前富有价值与意义的尝试.

### 4.1 应用研究

近年来,交通预测领域内对图卷积网络的应用已日渐丰富,包括但不限于道路车辆流量预测<sup>[58,79-83]</sup>、道路车辆平均速度预测<sup>[84-85]</sup>、道路车辆占有率预测<sup>[44]</sup>、道路交通状态预测<sup>[28,72,86-87]</sup>、行人轨迹预测<sup>[88-89]</sup>、车辆轨迹预测<sup>[90]</sup>、通勤时间预测<sup>[91-92]</sup>、轨道交通客流预测<sup>[93-94]</sup>、物流时间预测<sup>[95]</sup>、出租车需求预测<sup>[31-32]</sup>以及共享单车需求预测<sup>[60,96]</sup>等.

然而,目前的交通预测成果大多集中于短时常态预测,对于长时以及非常态下的研究相对较少.

#### 1) 长时交通预测

对于长时交通预测而言,由于更复杂的时空依赖性和更多不确定性因素,长时预测比中短时预测更加困难. Wang等<sup>[97]</sup>即利用LSTM网络以及自编码器架构对数据的长时特征进行学习,并在其自编码器中提出了一种基于学习相似模式的硬注意力机制以增强神经元的记忆并减少误差在传播过程中的累积. 但上述模型均未以图结构对数据进行建模表示且模型中利用CNN及RNN对单时序数据的时空特征进行提取,忽略了时序数据之间的相互影响. 因此,未来可针对此问题在此基础上通过图结构对数据进行表示,并引入GCN以实现更加精准的长时交通预测.

#### 2) 非常态交通预测

交通异常状况即发生于交通营运过程之中,或发生于交通体系之外,但影响到交通正常运行的情况及事件,例如突发交通事故,人群大规模聚集,恶劣天气导致的路面结冰、积水等非常态情形. 虽然非常态数据体量较小,但可借鉴小样本学习以及迁移学习的思想. 异常交通事件对于公共出行安全具有极大威胁,因此,该研究方向具有重要社会价值. 交通异常事件往往具有突发性强以及连锁影响范围广等特点. Guo等<sup>[98]</sup>测试了KNN、RNN以及TDNN算法在异常状况下交通预测的准确率,并表示KNN

效果较好. Li等<sup>[99]</sup>则发现当同时使用来自多个传感器的交通信息时,可以在短期趋势上提高预测准确性. 其研究表明,对于异常状况下的交通流预测,交通流序列之间的空间关系变得更加重要. 因此,对于该问题,将图卷积网络引入其中可作为未来重点研究方向之一.

### 4.2 模型研究

GCN由于对图结构数据进行卷积操作,所以,图表示与图卷积操作均具有深入研究价值. 对于构建图,目前研究热点包括但不局限于动态图卷积网络、多图卷积网络以及深层图卷积网络.

#### 1) 动态图建模

由于许多交通网络本质上均为图拓扑结构,为充分利用该拓扑信息,交通数据均依靠图建模进行表示. 在构建图的过程中,邻接矩阵承担了空间信息的载体. 大多数用于交通预测的GCN均采用了静态邻接矩阵,即假定节点之间的关联关系不随时移而改变. 该邻接矩阵往往由研究人员根据实际交通情况手工设计而来,最常见的即表示路面节点间连接关系的0-1矩阵以及表示节点之间实际地理距离的距离矩阵.

但目前有研究者认为预定义的邻接矩阵并不能很好地反映节点之间真正的相关性及其时空依赖<sup>[89]</sup>,反而从数据中学得的动态邻接矩阵在相关交通预测问题上表现出更加准确的效果. 例如:Guo等<sup>[100]</sup>在训练阶段通过数据驱动的方式学习优化图,从而从交通数据中揭示了路段之间的潜在关系. 在图卷积的过程中,拉普拉斯矩阵由邻接矩阵计算而来,Diao等<sup>[27]</sup>设计了动态拉普拉斯矩阵估计器,给定一个代表交通网络静态结构的全局拉普拉斯矩阵,同时利用短期交通流变化估计实时拉普拉斯矩阵,从而实现动态建模.

GCN无疑依赖于良好定义的图结构信息<sup>[101]</sup>. 当基于空间位置的邻接矩阵不明确时,对于交通数据这种多元时间序列可分析多元变量间的因果关系,并动态地构建邻接矩阵.

此外,超图<sup>[102]</sup>作为一种图的表示方法,由于可以表示个体之间的高阶关系,目前被广泛应用于社交网络或通信网络等领域之中. 与普通图相比,超边可同时连接多个节点,从而突破两两关系的限制. 例如:Wang等<sup>[103]</sup>在轨道交通预测中利用静态的超边表示轨道交通的线路,在此基础上对交通流OD进行分析,挖掘不同时间跨度下的车流或客流的出行规律. 以北京地铁为例,天通苑站及周围站点与

中关村及其周围站点在轨道网络中并非直接相连,但由 OD 分析可发掘出两者之间存在客流强关联且该关联随时移呈现出不同模式.因此,可据此构建动态超边,与静态超边同时纳入超图之中,并通过超图时空卷积网络实现交通预测.

## 2) 多图卷积网络

随着 GCN 的发展,单图所承载的信息越来越难以满足交通预测的需求.研究者即通过多图卷积网络,尤其是异质图卷积网络,对交通预测问题进行了更加精细化建模.

Lv 等<sup>[104]</sup>认为现有工作多数局限于利用时空模型进行交通预测,而道路之间的语义相关性 (semantic correlation) 同样重要.在其所提时序多图卷积网络 (temporal multi-graph convolutional network, T-MGCN) 中利用道路之间的空间相关性以及语义相关性构建多图,对多图分别进行图卷积后将结果融合,再通过 GRU 提取时序特征,实现交通流预测. Chai 等<sup>[105]</sup>则通过共享自行车站点之间的距离以及行驶记录相关性等关联关系构造多图.与 Lv 等<sup>[104]</sup>的方法不同的是,Chai 等<sup>[105]</sup>在处理多图卷积时,先进行图融合操作,再对融合图做图卷积.类似地,Ke 等<sup>[106]</sup>也通过空间距离以及语义相关性为出租车的 OD 预测构造多图.针对轨道交通,进出站客流在不同时间跨度下的出行模式也可以通过多个超图进行建模<sup>[103]</sup>.针对路面交通, Song 等<sup>[21]</sup>针对时序特征,设计了不同时间段的多个模块以有效地捕获时空图中的异质性. Zhang 等<sup>[107]</sup>则通过路网节点之间的距离、方向以及位置关系构建多图神经网络.在复杂多层次交通数据的建模上,多图卷积网络取得了良好的效果.如何挖掘更深层次的多图表示,如何实现更高效的多图卷积操作,仍是目前及未来的研究热点.

## 3) 多任务图网络

与多图卷积网络不同,多任务图网络借鉴了多任务学习 (multi-task learning) 的架构.多任务学习的初衷在于仅关注单个任务可能会忽略相关任务中潜在但有益的信息.通过在一定程度上共享任务之间的参数或特征,可提升主要任务的表现.通常,多任务可以分为同构任务和异构任务.同构任务可直接共享模型的参数,从而降低训练中过拟合的风险.异构任务的模型不同,但通过在训练中共享中间层特征,从而相互提供额外信息<sup>[108]</sup>.

在交通预测领域, Huang 等<sup>[109]</sup>提出了一种基于多任务深度网络架构,该架构由底部的深度信念网

络 (deep belief network) 以及顶部的多任务回归层组成. Zhang 等<sup>[110]</sup>将地理位置划分为交通单元,利用多任务学习对每个交通单元通过时间卷积实现乘客需求预测. Gao 等<sup>[111]</sup>则利用多任务学习融合了 GPS 轨迹、智能手机数据和道路网络结构以估计出租车乘客的出行时间.多任务学习为深度网络模型提供了优化并利用了多源数据的潜力,在交通预测领域基于 GCN 的多任务学习框架仍具有很高的研究价值.

## 4) 图卷积优化

GCN 可以使图信号更加平滑,这是图卷积的固有优势.但是,在 GCN 的层数不断加深时,其训练结果极易出现过平滑的情况<sup>[14]</sup>.由于图卷积是特殊形式的拉普拉斯平滑,在图卷积聚合邻居节点特征的过程中,平滑操作使信号在特征层面更加一致,从而使信号失去其多样性,导致相关预测任务的性能急剧下降,该现象在小数据集上更加明显.因此,GCN 不能像一般卷积模型那样不断地堆叠,但浅层神经网络又存在感受野和特征提取能力有限的问题.

针对该问题,一种解决思路是基于随机游走 (random walk) 的协同训练 (co-training)<sup>[112]</sup>方法.因为随机游走可以探索全局图结构,弥补了图卷积操作局部平滑的缺陷.另一种增大图卷积感受野的思路是自训练 (self-training) 方法.该方法首先训练带给定标签的 GCN,随后为每个类别选择最可靠的预测并将其添加到标签集中.在此基础上,使用预先训练的 GCN 继续使用扩展标签后的数据集训练 GCN.上述 2 种方法也可以同时进行,其目的都是扩展训练集.

在图卷积操作过程中,由于中心节点的特征易于向邻居节点传播,而边缘节点的特征则很难辐射至其余边缘节点.为克服此缺点,可引入一种通用的自动编码器网络架构,将数据内容信息补充到 GCN,实现类似于残差网络的效果,或使用跳跃图卷积网络实现特征的有效传递.为了更好地捕获数据样本之间的高阶关系,同时可以利用注意力融合机制对可判别信息进行突出.基于上述思路,在深度堆叠的 GCN 中,每一层都用于捕获数据的不同潜在特征.

## 4.3 多源数据融合

近年来交通预测方法不断由模型驱动转变为数据驱动,即由数学模型推导演变为挖掘并学习数据本身的特征.目前,基于单一数据源的学习已逐渐

完备,但实际中交通数据多源异构的特点决定了该领域研究必将从单源走向多源。同时空内,跨媒体交通数据类型多样并结构复杂,但同时相互影响且互为补充。针对多源交通数据的挑战,本文将小样本学习、跨媒体数据融合以及交通知识图谱嵌入作为未来研究方向的参考。

### 1) 小样本学习

目前,大多数交通预测的解决方案均通过密集的历史交通数据进行训练,但由于交通数据采集等问题,城市之间交通数据量级不同,部分城市交通数据量较小或质量较差。同时,用以辅助交通预测的极端天气、交通事故、交通管制以及其他异常事件等发生频率远低于交通数据采集频率,使得学习过程异常困难。因此,交通预测领域的小样本学习问题同样为未来研究方向之一。

元学习在监督学习领域为小样本学习提供了解决思路,Pan 等<sup>[68]</sup>即提出了元-图注意力模块与元-循环神经网络模块。其中注意力模块的权重是通过地理属性中提取的节点和边的元知识生成的,因此,理论上可以对各种空间相关性进行建模,而循环神经网络根据每个节点的元知识生成门控循环单元的所有权重,从而实现元学习的 GCN 模型。

解决此问题的另一种思路是通过迁移学习来执行跨城市的交通预测任务。该思路旨在将交通知识从数据源丰富的城市转移到数据稀缺的目标城市。但目前该研究方向尚未进行彻底的探究,例如如何设计一个高质量泛用性强的用于交通预测的迁移学习框架,或者如何实现小样本外部数据的嵌入学习等。

### 2) 跨媒体数据融合

目前,交通领域已经汇聚了监控设备采集的视频、图像,线圈卡口采集的流量、速度、占有率以及车载设备采集的 GPS 等海量多源异构的高维交通数据,同时恶劣天气、交通事故以及交通管制等辅助数据同样对于交通预测具有不可忽视的价值。由于难以从单一数据源全面且精准地捕获交通特征,并且针对体量大、时空跨度广以及结构复杂等特点,对多源交通数据进行融合、协同分析已成为该领域未来研究方向之一。对于多源交通数据融合,本文认为需研究跨媒体数据的有效表达方法及协同分析模型。

无论是图时空网络或图注意力网络,其关注点均主要在于时空特征的提取。在此基础上,若引入交通流数据之外有价值的信息,势必对交通预测问

题有所提升。Ni 等<sup>[113]</sup>通过研究社交媒体的动态进而感知交通事件的发生,并通过分析表明轨道交通的客流量与社交媒体的发帖率之间存在正相关关系。基于该发现,Ni 等<sup>[113]</sup>提出了基于标签的事件监测算法并与 SARIMA 模型相结合,用于轨道交通的客流预测。同样针对轨道客流预测问题,Chen 等<sup>[114]</sup>则利用刷卡数据对异常事件进行捕获,并结合 ARIMA 模型实现预测。社交媒体数据<sup>[115]</sup>、刷卡数据<sup>[116]</sup>、遥感数据<sup>[117]</sup>及天气数据<sup>[118]</sup>均被用作提升预测精度的外源数据。Zhang 等<sup>[118]</sup>在其所提框架中将雨雪天气等外部因素通过门控机制融合至预测模型之中。但目前上述方法均未考虑交通数据的图表示及 GCN,因此,在 GCN 的基础上,通过多任务学习等方法融合多源异构交通数据具有广阔的研究前景。

### 3) 交通知识图谱嵌入

知识图谱是对某领域的知识整合以及利用的重要工具,由大量知识实体及其之间的关系交织而成。交通运输领域的知识存在于海量多源异构的交通数据之中。对于交通知识图谱,未来主要研究图谱构建以及知识嵌入两方面内容。

知识图谱构建方面,早期主要为泛领域化的语义网络,强调广义概念之间的关联关系。近年来面向特定行业应用领域的知识图谱逐步涌现,例如语义网络知识图谱<sup>[119]</sup>以及医疗知识图谱<sup>[120]</sup>等。对于交通领域,则需要明确知识实体以及实体之间的多层次关联关系,挖掘跨域全时交通数据之间的深度关联与知识聚合,实现交通知识图谱构建。在构建交通知识图谱的基础上,可研究知识的嵌入模型。目前,主流方法有基于 TransE<sup>[121]</sup>及其扩展和变体的平移距离模型以及基于 Rescal<sup>[122]</sup>的扩展和变体语义匹配模型等。交通知识的嵌入对于交通预测模型而言可提供无法由交通流数据中学得的额外信息。因此,如何对交通知识进行抽取并构建相关知识图谱,如何将交通知识嵌入交通预测模型之中,如何将富有价值意义的交通知识应用于实际交通预测问题之中是未来交通预测领域的重点研究方向之一。

## 5 结论

1) 本文首先梳理了交通预测问题以及图卷积神经网络的基本定义。在此基础上对不同类型的基于 GCN 的交通预测模型进行分类介绍,包括图时空网络、图自编码器以及图注意力网络等。

2) 针对交通预测实验,本文介绍了目前全球范围内广泛用于科学研究的部分公开交通数据集,包括其数据来源、数据内容以及数据获取地址等信息。

3) 讨论了未来的几点研究方向,包括应用研究方向、模型研究方向以及多源数据研究方向。本文适合用于快速了解 GCN 在交通预测领域的模型及应用。

#### 参考文献:

- [1] PAN B, DEMIRYUREK U, SHAHABI C. Utilizing real-world transportation data for accurate traffic prediction[C]// IEEE 12th International Conference on Data Mining. Piscataway: IEEE, 2012: 595-604.
- [2] SUN Y, ZHANG G, YIN H. Passenger flow prediction of subway transfer stations based on nonparametric regression model [J]. *Discrete Dynamics in Nature and Society*, 2014, 2014: 397154.
- [3] WILLIAMS B M, DURVASULA P K, BROWN D E. Urban freeway traffic flow prediction; application of seasonal autoregressive integrated moving average and exponential smoothing models [J]. *Transportation Research Record*, 1998, 1644(1): 132-141.
- [4] LEE S, FAMBRO D B. Application of subset autoregressive integrated moving average model for short-term freeway traffic volume forecasting [J]. *Transportation Research Record*, 1999, 1678(1): 179-188.
- [5] WILLIAMS B M. Multivariate vehicular traffic flow prediction; evaluation of arimax modeling [J]. *Transportation Research Record*, 2001, 1776(1): 194-200.
- [6] XIE Y, ZHANG Y, YE Z. Short-term traffic volume forecasting using Kalman filter with discrete wavelet decomposition [J]. *Computer-Aided Civil and Infrastructure Engineering*, 2007, 22(5): 326-334.
- [7] CLARK S. Traffic prediction using multivariate nonparametric regression [J]. *Journal of Transportation Engineering*, 2003, 129(2): 161-168.
- [8] YU Y, ZHANG Y, QIAN S, et al. A low rank dynamic mode decomposition model for short-term traffic flow prediction [J/OL]. *IEEE Transactions on Intelligent Transportation Systems*, 2020: 1-14 [2020-12-28]. <http://dx.doi.org/10.1109/TITS.2020.2994910>.
- [9] MA X, DAI Z, HE Z, et al. Learning traffic as images: a deep convolutional neural network for large-scale transportation network speed prediction [J]. *Sensors*, 2017, 17(4): 818.
- [10] LV Z, XU J, ZHENG K, et al. LC-RNN: a deep learning model for traffic speed prediction [C] // *Proceedings of the 27th International Joint Conference on Artificial Intelligence*. San Francisco, CA: Morgan Kaufmann, 2018: 3470-3476.
- [11] FU R, ZHANG Z, LI L. Using LSTM and GRU neural network methods for traffic flow prediction [C] // 31st Youth Academic Annual Conference of Chinese Association of Automation. Piscataway: IEEE, 2016: 324-328.
- [12] WU Z, PAN S, CHEN F, et al. A comprehensive survey on graph neural networks [J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2021, 32(1): 4-24.
- [13] KIPF T N, WELING M. Semi-supervised classification with graph convolutional networks [EB/OL]. *CoRR*, 2012 [2020-12-28]. <https://arxiv.org/abs/1609.02907>.
- [14] NAGY A M, SIMON V. Survey on traffic prediction in smart cities [J]. *Pervasive and Mobile Computing*, 2018, 50: 148-163.
- [15] WEST D B. *Introduction to graph theory* [M]. Upper Saddle River: Prentice Hall, 2001: 1-8.
- [16] GILMER J, SCHOENHOLZ S S, RILEY P F, et al. Neural message passing for quantum chemistry [C] // *International Conference on Machine Learning*. San Diego: PMLR, 2017: 1263-1272.
- [17] HAMILTON W, YING Z, LESKOVEC J. Inductive representation learning on large graphs [EB/OL]. *CoRR*, 2017 [2020-12-28]. <https://arxiv.org/abs/1706.02216>.
- [18] LIU J, ONG G P, CHEN X. GraphSAGE-based traffic speed forecasting for segment network with sparse data [J/OL]. *IEEE Transactions on Intelligent Transportation Systems*, 2020: 1-12 [2020-12-28]. <http://dx.doi.org/10.1109/TITS.2020.3026025>.
- [19] LI Y, YU R, SHAHABI C, et al. Diffusion convolutional recurrent neural network: data-driven traffic forecasting [EB/OL]. *CoRR*, 2017 [2020-12-28]. <https://arxiv.org/abs/1707.01926>.
- [20] CHEN C, LI K, TEO S G, et al. Gated residual recurrent graph neural networks for traffic prediction [C] // *Proceedings of the AAAI Conference on Artificial Intelligence*. Palo Alto, CA: AAAI Press, 2017(33): 485-492.
- [21] SONG C, LIN Y, GUO S, et al. Spatial-temporal synchronous graph convolutional networks: a new framework for spatial-temporal network data forecasting [C] // *Proceedings of the AAAI Conference on Artificial*

- Intelligence. Palo Alto, CA: AAAI Press, 2020 (34): 914-921.
- [22] BRUNA J, ZAREMBA W, SZLAM A, et al. Graph waveNet for deep spatial-temporal graph modeling [EB/OL]. CoRR, 2019 [2020-12-28]. <https://arxiv.org/abs/1906.00121>.
- [23] BRUNA J, ZAREMBA W, SZLAM A, et al. Spectral networks and locally connected networks on graphs [EB/OL]. CoRR, 2013 [2020-12-28]. <https://arxiv.org/abs/1312.6203>.
- [24] DEFFERRARD M, BRESSON X, VANDERGHEYNST P. Convolutional neural networks on graphs with fast localized spectral filtering [C]//Proceedings of the 30th International Conference on Neural Information Processing Systems. Red Hook, NY: Curran Associates, 2016: 3844-3852.
- [25] LEE K, RHEE W. DDP-GCN: multi-graph convolutional network for spatiotemporal traffic forecasting [EB/OL]. CoRR, 2019 [2020-12-28]. <https://arxiv.org/abs/1905.12256>.
- [26] FANG S, ZHANG Q, MENG G, et al. GSTNet: global spatial-temporal network for traffic flow prediction [C]//Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence. San Francisco, CA: Margan Kaufmann, 2019: 2286-2293.
- [27] DIAO Z, WANG X, ZHANG D, et al. Dynamic spatial-temporal graph convolutional neural networks for traffic forecasting [C]//Proceedings of the AAAI Conference on Artificial Intelligence. Palo Alto, CA: AAAI Press, 2019 (33): 890-897.
- [28] CUI Z, HENRICKSON K, KE R, et al. Traffic graph convolutional recurrent neural network: a deep learning framework for network-scale traffic learning and forecasting [J]. IEEE Transactions on Intelligent Transportation Systems, 2019, 21(11): 4883-4894.
- [29] YU B, YIN H, ZHU Z. Spatio-temporal graph convolutional networks: a deep learning framework for traffic forecasting [EB/OL]. CoRR, 2017 [2020-12-28]. <https://arxiv.org/abs/1709.04875>.
- [30] BAI L, YAO L, KANHERE S, et al. Stg2seq: spatial-temporal graph to sequence model for multi-step passenger demand forecasting [EB/OL]. CoRR, 2019 [2020-12-28]. <https://arxiv.org/abs/1905.10069>.
- [31] WANG Y, YIN H, CHEN H, et al. Origin-destination matrix prediction via graph convolution: a new perspective of passenger demand modeling [C]//Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. New York: Association for Computing Machinery, 2019: 1227-1235.
- [32] GENG X, LI Y, WANG L, et al. Spatiotemporal multi-graph convolution network for ride-hailing demand forecasting [C]//Proceedings of the AAAI Conference on Artificial Intelligence. Palo Alto, CA: AAAI Press, 2019 (33): 3656-3663.
- [33] WILLIAM S, BILLY M, LESTER A H. 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.
- [34] WU C H, HO J M, LEE D T. Travel-time prediction with support vector regression [J]. IEEE Transactions on Intelligent Transportation Systems, 2004, 5(4): 276-281.
- [35] GEHRING J, AULI M, GRANGIER D, et al. Convolutional sequence to sequence learning [EB/OL]. CoRR, 2017 [2020-12-28]. <https://arxiv.org/abs/1705.03122>.
- [36] GUO S, LIN Y, LI S, et al. Deep spatial-temporal 3D convolutional neural networks for traffic data forecasting [J]. IEEE Transactions on Intelligent Transportation Systems, 2019, 20(10): 3913-3926.
- [37] YAO H, LIU Y, WEI Y, et al. Learning from multiple cities: a meta-learning approach for spatial-temporal prediction [C]//Proceedings of The World Wide Web Conference. New York: Association for Computing Machinery, 2019: 2181-2191.
- [38] ZHAO L, SONG Y, ZHANG C, et al. T-gen: a temporal graph convolutional network for traffic prediction [J]. IEEE Transactions on Intelligent Transportation Systems, 2019, 21(9): 3848-3858.
- [39] LI Y, ZHU Z, KONG D, et al. Learning heterogeneous spatial-temporal representation for bike-sharing demand prediction [C]//Proceedings of the AAAI Conference on Artificial Intelligence. Palo Alto, CA: AAAI Press, 2019 (33): 1004-1011.
- [40] LI Y, ZHU Z, KONG D, et al. Deep multi-view spatial-temporal network for taxi demand prediction [EB/OL]. CoRR, 2018 [2020-12-28]. <https://arxiv.org/abs/1802.08714>.
- [41] WANG D, ZHANG J, CAO W, et al. When will you arrive? estimating travel time based on deep neural networks [C]//Proceedings of the AAAI Conference on Artificial Intelligence. Palo Alto, CA: AAAI Press, 2018, 32(1): 2500-2507.
- [43] 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.
- [44] LAI G, CHANG W C, YANG Y, et al. Modeling long- and short-term temporal patterns with deep neural networks [ C ] // *Proceedings of the 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*. New York: Association for Computing Machinery, 2018: 95-104.
- [45] ZONOOZI A, KIM J, LI X L, et al. Periodic-CRN: a convolutional recurrent model for crowd density prediction with recurring periodic patterns [ C ] // *Proceedings of the 27th International Joint Conference on Artificial Intelligence*. San Francisco, CA: Margan Kaufmann, 2018: 3732-3738.
- [46] KE J, ZHENG H, YANG H, et al. Short-term forecasting of passenger demand under on-demand ride services: a spatio-temporal deep learning approach [ J ]. *Transportation Research Part C: Emerging Technologies*, 2017, 85: 591-608.
- [47] DAVIS N, RAINA G, JAGANNATHAN K. Grids versus graphs: partitioning space for improved taxi demand-supply forecasts [ J/OL ]. *IEEE Transactions on Intelligent Transportation Systems*, 2020: 1-10 [2020-12-28]. <http://dx.doi.org/10.1109/TITS.2020.2993798>.
- [48] XU J, RAHMATIZADEH R, BÖLÖNI L, et al. Real-time prediction of taxi demand using recurrent neural networks [ J ]. *IEEE Transactions on Intelligent Transportation Systems*, 2017, 19(8): 2572-2581.
- [49] LIU L, QIU Z, LI G, et al. Contextualized spatial-temporal network for taxi origin-destination demand prediction [ J ]. *IEEE Transactions on Intelligent Transportation Systems*, 2019, 20(10): 3875-3887.
- [50] PANG J, HUANG J, DU Y, et al. Learning to predict bus arrival time from heterogeneous measurements via recurrent neural network [ J ]. *IEEE Transactions on Intelligent Transportation Systems*, 2018, 20(9): 3283-3293.
- [51] HE P, JIANG G, LAM S K, et al. Travel-time prediction of bus journey with multiple bus trips [ J ]. *IEEE Transactions on Intelligent Transportation Systems*, 2018, 20(11): 4192-4205.
- [52] TANG X, YAO H, SUN Y, et al. Joint modeling of local and global temporal dynamics for multivariate time series forecasting with missing values [ C ] // *Proceedings of the AAAI Conference on Artificial Intelligence*. Palo Alto, CA: AAAI Press, 2020, 34(4): 5956-5963.
- [53] DAI A M, LE Q V. Semi-supervised sequence learning [ J ]. *Advances in Neural Information Processing Systems*, 2015, 28: 3079-3087.
- [54] KIPF T N, WELING M. Variational graph auto-encoders [ EB/OL ]. *CoRR*, 2016 [2020-12-28]. <https://arxiv.org/abs/1611.07308>.
- [55] LV Y, DUAN Y, KANG W, et al. Traffic flow prediction with big data: a deep learning approach [ J ]. *IEEE Transactions on Intelligent Transportation Systems*, 2014, 16(2): 865-873.
- [56] ZHAO J, ZHU T, ZHAO R, et al. Layerwise recurrent autoencoder for real-world traffic flow forecasting [ C ] // *International Conference on Intelligent Science and Big Data Engineering*. Berlin: Springer, 2019: 78-88.
- [57] LIN L, HE Z, PEETA S. Predicting station-level hourly demand in a large-scale bike-sharing network: a graph convolutional neural network approach [ J ]. *Transportation Research Part C: Emerging Technologies*, 2018, 97: 258-276.
- [58] JIANG R, SONG X, HUANG D, et al. DeepUrban-Event: a system for predicting citywide crowd dynamics at big events [ C ] // *Proceedings of 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. New York: Association for Computing Machinery, 2019: 2114-2122.
- [59] LIAO B, ZHANG J, WU C, et al. Deep sequence learning with auxiliary information for traffic prediction [ C ] // *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. New York: Association for Computing Machinery, 2018: 537-546.
- [60] ZHU L, LAPTEV N. Deep and confident prediction for time series at uber [ C ] // *Proceedings of 2017 IEEE International Conference on Data Mining Workshops*. Piscataway: IEEE, 2017: 103-110.
- [61] YE J, SUN L, DU B, et al. Co-prediction of multiple transportation demands based on deep spatio-temporal neural network [ C ] // *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. New York: Association for Computing Machinery, 2019: 305-313.
- [62] LI Y, MOURA JM. Forecaster: a graph transformer for forecasting spatial and time-dependent data [ EB/OL ]. *CoRR*, 2019 [2020-12-28]. <https://arxiv.org/abs/1909.04019>.
- [63] PARK C, LEE C, BAHNG H, et al. ST-GRAT: a novel spatio-temporal graph attention network for accurately forecasting dynamically changing road speed [ EB/OL ]. *CoRR*, 2019 [2020-12-28]. <https://arxiv.org/abs/>

- 1911.13181.
- [64] DESHPANDE P, SARAWAGI S. Streaming adaptation of deep forecasting models using adaptive recurrent units [C] // Proceedings of 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. New York: Association for Computing Machinery, 2019: 1560-1568.
- [65] LIANG Y, KE S, ZHANG J, et al. Geoman: multi-level attention networks for geo-sensory time series prediction [C] // Proceedings of the 27th International Joint Conference on Artificial Intelligence. San Francisco: Margan Kaufmann, 2018: 3428-3434.
- [66] VELIČKOVIĆ P, CUCURULL G, CASANOVA A, et al. Graph attention networks [EB/OL]. CoRR, 2017 [2020-12-28]. <https://arxiv.org/abs/1710.10903>.
- [67] GUO S, LIN Y, FENG N, et al. Attention based spatial-temporal graph convolutional networks for traffic flow forecasting [C] // Proceedings of AAAI Conference on Artificial Intelligence. Palo Alto, CA: AAAI Press, 2019, 33(1): 922-929.
- [68] PAN Z, LIANG Y, WANG W, et al. Urban traffic prediction from spatio-temporal data using deep meta learning [C] // Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. New York: Association for Computing Machinery, 2019: 1720-1730.
- [69] CHEN W, CHEN L, XIE Y, et al. Multi-range attentive bicomponent graph convolutional network for traffic forecasting [C] // Proceedings of the AAAI Conference on Artificial Intelligence. Palo Alto, CA: AAAI Press, 2020, 34(4): 3529-3536.
- [70] ZHANG J, SHI X, XIE J, et al. Gaan: gated attention networks for learning on large and spatiotemporal graphs [EB/OL]. CoRR, 2018 [2020-12-28]. <https://arxiv.org/abs/1803.07294>.
- [71] YAO H, TANG X, WEI H, et al. Revisiting spatial-temporal similarity: a deep learning framework for traffic prediction [C] // Proceedings of the AAAI Conference on Artificial Intelligence. Palo Alto, CA: AAAI Press, 2019, 33(1): 5668-5675.
- [72] ZHENG C, FAN X, WANG C, et al. Gman: a graph multi-attention network for traffic prediction [C] // Proceedings of the AAAI Conference on Artificial Intelligence. Palo Alto, CA: AAAI Press, 2020, 34(1): 1234-1241.
- [73] PARK C, LEE C, BAHNG H, et al. ST-GRAT: a novel spatio-temporal graph attention network for accurately forecasting dynamically changing road speed [EB/OL]. CoRR, 2019 [2020-12-28]. <https://arxiv.org/abs/1911.13181>.
- [74] LIANG Y, KE S, ZHANG J, et al. Geoman: multi-level attention networks for geo-sensory time series prediction [C] // Proceedings of the 27th International Joint Conference on Artificial Intelligence. San Francisco, CA: Margan Kaufmann, 2018: 3428-3434.
- [75] YI X, DUAN Z, LI T, et al. Citytraffic: modeling citywide traffic via neural memorization and generalization approach [C] // Proceedings of the 28th ACM International Conference on Information and Knowledge Management. New York: Association for Computing Machinery, 2019: 2665-2671.
- [78] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [C] // Advances in Neural Information Processing Systems. Red Hook, NY: Curran Associates, 2017: 5998-6008.
- [79] YUN S, JEONG M, KIM R, et al. Graph transformer networks [C] // Advances in Neural Information Processing Systems. Red Hook, NY: Curran Associates, 2019: 11983-11993.
- [80] CAI L, JANOWICZ K, MAI G, et al. Traffic transformer: capturing the continuity and periodicity of time series for traffic forecasting [J]. Transactions in GIS, 2020, 24(3): 736-755.
- [81] YIN H, WONG S, XU J, et al. Urban traffic flow prediction using a fuzzy-neural approach [J]. Transportation Research Part C: Emerging Technologies, 2002, 10(2): 85-98.
- [82] KANG Z, XU H, HU J, et al. Learning dynamic graph embedding for traffic flow forecasting: a graph self-attentive method [C] // 2019 IEEE Intelligent Transportation Systems Conference. Piscataway: IEEE, 2019: 2570-2576.
- [83] SUN J, ZHANG J, LI Q, et al. Predicting citywide crowd flows in irregular regions using multi-view graph convolutional networks [J/OL]. IEEE Transactions on Intelligent Transportation Systems, 2020: 1-12 [2020-12-28]. <http://dx.doi.org/10.1109/TKDE.2020.3008774>.
- [84] ASIF M T, DAUWELS J, GOH C Y, et al. Spatiotemporal patterns in large-scale traffic speed prediction [J]. IEEE Transactions on Intelligent Transportation Systems, 2013, 15(2): 794-804.
- [85] 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: 297-322.
- [86] PAN T L, SUMALEE A, ZHONG R X, et al. Short-term traffic state prediction based on temporal-spatial correlation [ J ]. *IEEE Transactions on Intelligent Transportation Systems*, 2013, 14(3): 1242-1254.
- [87] CHEN W, CHEN L, XIE Y, et al. Multi-range attentive bicomponent graph convolutional network for traffic forecasting [ C ] // *Proceedings of the AAAI Conference on Artificial Intelligence*. Palo Alto, CA: AAAI Press, 2020, 34(4): 3529-3536.
- [88] SUN Y, HE T, HU J, et al. Socially-aware graph convolutional network for human trajectory prediction [ C ] // *2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference*. Piscataway: IEEE, 2019: 325-333.
- [89] MOHAMED A, QIAN K, ELHOSEINY M, et al. Social-STGCNN: a social spatio-temporal graph convolutional neural network for human trajectory prediction [ C ] // *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Piscataway: IEEE, 2020: 14424-14432.
- [90] KOSARAJU V, SADEGHIAN A, MARTÍN-MARTÍN R, et al. Social-bigat: multimodal trajectory forecasting using bicycle-gan and graph attention networks [ C ] // *Advances in Neural Information Processing Systems*. Red Hook, NY: Curran Associates, 2019, 32: 137-146.
- [91] HU J, GUO C, YANG B, et al. Stochastic weight completion for road networks using graph convolutional networks [ C ] // *2019 IEEE 35th International Conference on Data Engineering*. Piscataway: IEEE, 2019: 1274-1285.
- [92] ZHANG X, RICE J A. Short-term travel time prediction [ J ]. *Transportation Research Part C: Emerging Technologies*, 2003, 11(3/4): 187-210.
- [93] LIU Y, LIU Z, JIA R. DeepPF: a deep learning based architecture for metro passenger flow prediction [ J ]. *Transportation Research Part C: Emerging Technologies*, 2019, 101: 18-34.
- [94] WEI Y, CHEN M C. Forecasting the short-term metro passenger flow with empirical mode decomposition and neural networks [ J ]. *Transportation Research Part C: Emerging Technologies*, 2012, 21(1): 148-162.
- [95] SIMROTH A, ZÄHLE H. Travel time prediction using floating car data applied to logistics planning [ J ]. *IEEE Transactions on Intelligent Transportation Systems*, 2010, 2(1): 243-253.
- [96] CHAI D, WANG L, YANG Q. Bike flow prediction with multi-graph convolutional networks [ C ] // *Proceedings of the 26th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*. New York: Association for Computing Machinery, 2018: 397-400.
- [97] WANG Z, SU X, DING Z. Long-term traffic prediction based on LSTM encoder-decoder architecture [ J/OL ]. *IEEE Transactions on Intelligent Transportation Systems*, 2020: 1-11 [ 2020-12-28 ]. <http://dx.doi.org/10.1109/TITS.2020.2995546>.
- [98] GUO F, POLAK J W, KRISHNAN R. Comparison of modelling approaches for short term traffic prediction under normal and abnormal conditions [ C ] // *13th International IEEE Conference on Intelligent Transportation Systems*. Piscataway: IEEE, 2010: 1209-1214.
- [99] LI L, SU X, ZHANG Y, et al. Traffic prediction, data compression, abnormal data detection and missing data imputation: an integrated study based on the decomposition of traffic time series [ C ] // *17th International IEEE Conference on Intelligent Transportation Systems*. Piscataway: IEEE, 2014: 282-289.
- [100] GUO K, HU Y, QIAN Z, et al. Optimized graph convolution recurrent neural network for traffic prediction [ J ]. *IEEE Transactions on Intelligent Transportation Systems*, 2021, 22(2): 1138-1149.
- [101] WU Z, PAN S, LONG G, et al. Connecting the dots: multivariate time series forecasting with graph neural networks [ C ] // *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. New York: Association for Computing Machinery, 2020: 753-763.
- [102] FENG Y, YOU H, ZHANG Z, et al. Hypergraph neural networks [ C ] // *Proceedings of the AAAI Conference on Artificial Intelligence*. Palo Alto, CA: AAAI Press, 2019, 33: 3558-3565.
- [103] WANG J, ZHANG Y, WEI Y, et al. Metro passenger flow prediction via dynamic hypergraph convolution networks [ J/OL ]. *IEEE Transactions on Intelligent Transportation Systems*, 2021: 1-13 [ 2020-12-28 ]. <http://dx.doi.org/10.1109/TITS.2021.3072743>.
- [104] LV M, HONG Z, CHEN L, et al. Temporal multi-graph convolutional network for traffic flow prediction [ J/OL ]. *IEEE Transactions on Intelligent Transportation Systems*, 2020: 1-12 [ 2020-12-28 ]. <http://dx.doi.org/10.1109/TITS.2020.2983763>.
- [105] CHAI D, WANG L, YANG Q. Bike flow prediction with multi-graph convolutional networks [ C ] // *Proceedings of the 26th ACM SIGSPATIAL International*



- Conference on Advances in Geographic Information Systems. New York: Association for Computing Machinery, 2018: 397-400.
- [106] KE J, QIN X, YANG H, et al. Predicting origin-destination ride-sourcing demand with a spatio-temporal encoder-decoder residual multi-graph convolutional network[J]//Transportation Research Part C: Emerging Technologies, 2021, 122: 102858.
- [107] ZHANG J, CHEN F, GUO Y, et al. Multi-graph convolutional network for short-term passenger flow forecasting in urban rail transit [J]. IET Intelligent Transport Systems, 2020, 14(10): 1210-1217.
- [108] RUDER S. An overview of multi-task learning in deep neural networks [EB/OL]. CoRR, 2017 [2020-12-28]. <https://arxiv.org/abs/1706.05098>.
- [109] HUANG W, SONG G, HONG H, et al. Deep architecture for traffic flow prediction: deep belief networks with multitask learning[J]. IEEE Transactions on Intelligent Transportation Systems, 2014, 15(5): 2191-2201.
- [110] ZHANG K, LIU Z, ZHENG L. Short-term prediction of passenger demand in multi-zone level: temporal convolutional neural network with multi-task learning [J]. IEEE Transactions on Intelligent Transportation Systems, 2019, 21(4): 1480-1490.
- [111] GAO R, GUO X, SUN F, et al. Aggressive driving saves more time? multi-task learning for customized travel time estimation[C]//Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence. San Francisco, CA: Margan Kaufmann, 2019: 1689-1696.
- [112] LI Q, HAN Z, WU X M. Deeper insights into graph convolutional networks for semi-supervised learning[C]//Proceedings of the AAAI Conference on Artificial Intelligence. Palo Alto, CA: AAAI Press, 2018, 32(1): 3538-3545.
- [113] NI M, HE Q, GAO J. Forecasting the subway passenger flow under event occurrences with social media [J]. IEEE Transactions on Intelligent Transportation Systems, 2016, 18(6): 1623-1632.
- [114] CHEN E, YE Z, WANG C, et al. Subway passenger flow prediction for special events using smart card data [J]. IEEE Transactions on Intelligent Transportation Systems, 2019, 21(3): 1109-1120.
- [115] ESSIEN A, PETROUNIAS I, SAMPAIO P, et al. A deep-learning model for urban traffic flow prediction with traffic events mined from twitter [C]//Proceedings of The World Wide Web Conference. New York: Association for Computing Machinery, 2020, 14: 1-24.
- [116] LI Y, WANG X, SUN S, et al. Forecasting short-term subway passenger flow under special events scenarios using multiscale radial basis function networks [J]. Transportation Research Part C: Emerging Technologies, 2017, 77: 306-328.
- [117] ZHANG Y, DONG X, SHANG L, et al. A multi-modalgraph neural network approach to traffic risk forecasting in smart urban sensing [C]//2020 17th Annual IEEE International Conference on Sensing, Communication, and Networking. Piscataway: IEEE, 2020: 1-9.
- [118] ZHANG J, ZHENG Y, SUN J, et al. Flow prediction in spatio-temporal networks based on multitask deep learning [J]. IEEE Transactions on Knowledge and Data Engineering, 2019, 32(3): 468-478.
- [119] FRANCO-SALVADOR M, GUPTA P, ROSSO P, et al. Cross-language plagiarism detection over continuous-space-and knowledge graph-based representations of language [J]. Knowledge-Based Systems, 2016, 111: 87-99.
- [120] ROTMENSCH M, HALPERN Y, TLIMAT A, et al. Learning a health knowledge graph from electronic medical records [J]. Scientific Reports, 2017, 7(1): 1-11.
- [121] BORDES A, USUNIER N, GARCIA-DURAN A, et al. Translating embeddings for modeling multi-relational data [C]//Advances in Neural Information Processing Systems. Red Hook, NY: Curran Associates, 2013, 26: 2787-2795.
- [122] NICKEL M, ROSASCO L, POGGIO T. Holographic embeddings of knowledge graphs [EB/OL]. CoRR, 2015 [2020-12-28]. <https://arxiv.org/abs/1510.04935>.

(责任编辑 梁洁)