



基于混合图卷积的多通道时空交通流预测模型

张雄涛^{1,2}, 郑景玉^{1,2}, 申情^{1,2}, 孙丹枫³, 蒋云良^{1,2,4}

- (1. 湖州师范学院信息工程学院, 浙江 湖州 313000;
2. 浙江省现代农业资源智慧管理与应用研究重点实验室, 浙江 湖州 313000;
3. 杭州电子科技大学计算机学院, 浙江 杭州 310018;
4. 浙江师范大学计算机科学与技术学院, 浙江 金华 321004)

摘要: 针对交通流预测模型没有考虑道路上下文相关性和空间依赖关系动态性的问题, 提出一种基于混合图卷积的多通道时空交通流预测模型 (MHGCN)。该模型采用三明治结构 (即中间多通道空间模块, 两边时间模块) 提取时空特征, 多通道空间模块又分为静态图卷积模块和动态图卷积模块。静态图卷积模块同时从拓扑空间结构、语义空间结构及其组合中提取特定和公共的特征; 动态图卷积模块对不同的特征分配不同的权重, 从未知的图结构中提取动态的空间特征。时间模块中采用多头注意力机制提取全局时间特征, 采用时间门控机制提取局部时间特征。该模型从不同的空间结构中提取空间信息, 从不同时间间隔提取时间信息, 建立全局、全面的时空关系。实验结果表明, MHGCN 模型在 4 个公开数据集上的性能优于现有的交通流预测模型。

关键词: 智能交通; 动态图卷积; 多头注意力; 时空相关性; 多通道

中图分类号: TP393

文献标志码: A

doi: 10.11959/j.issn.1000-0801.2023173

Multi-channel spatial-temporal traffic flow prediction based on hybrid static-dynamic graph convolution

ZHANG Xiongtao^{1,2}, ZHENG Jingyu^{1,2}, SHEN Qing^{1,2}, SUN Danfeng³, JIANG Yunliang^{1,2,4}

1. School of Information Engineering, Huzhou University, Huzhou 313000, China
2. Zhejiang Province Key Laboratory of Smart Management and Application of Modern Agricultural Resources, Huzhou 313000, China
3. School of Computer, Hangzhou Dianzi University, Hangzhou 310018, China
4. School of Computer Science and Technology, Zhejiang Normal University, Jinhua 321004, China

Abstract: Aiming at the problem that the traffic flow prediction model did not consider the correlation of road context and the dynamics of spatial dependency, a multi-channel spatial-temporal traffic flow prediction based on hybrid static-dynamic graph convolution (MHGCN) was proposed. A sandwich structure (i.e. multi-channel spatial module

收稿日期: 2023-06-01; 修回日期: 2023-09-01

通信作者: 申情, sq@zjhu.edu.cn

基金项目: 国家自然科学基金资助项目 (No.U22A20102); 浙江省“尖兵”“领雁”研发攻关计划项目 (No.2023C01150)

Foundation Items: The National Natural Science Foundation of China (No.U22A20102), The “Pioneer” and “Leading Goose” Research & Development Program of Zhejiang Province (No.2023C01150)



in the middle and temporal module on both sides) was used in the model to extract spatial-temporal features, and the multi-channel spatial module was divided into static graph convolution module and dynamic graph convolution module. The static graph convolution module simultaneously extracted specific and common features from topological spatial structures, semantic spatial structures, and their combinations. The dynamic graph convolution module assigned different weights to different features and extracts dynamic spatial features from unknown graph structures. In the temporal module, the multi-head attention mechanism was used to extract the global temporal features, and the temporal gating mechanism extracted the local temporal features. The model extracted spatial information from different spatial structures and temporal information from different time intervals to establish a global and comprehensive spatial-temporal relationship. The experimental results show that the MHGCN performs better than the existing traffic flow prediction models on four real world traffic flow datasets.

Key words: intelligent transportation, dynamic graph convolution, multi-head attention, spatial-temporal correlation, multi-channel

0 引言

随着科技的发展,交通工具数量日益增加,交通拥堵和交通事故频繁发生。为解决上述问题,越来越多的国家开始重视智能交通系统(intelligent transportation system, ITS)的研究。其中,交通流预测是智能交通系统中至关重要的一部分。交通流预测是根据历史观察的道路信息(例如交通量、速度和占用率)预测未来的交通状况。交通流预测不仅可以帮助人们规划路线,方便出行,节省时间,还可以为政府在交通管理决策方面提供有效信息,以提高交通效率。

为解决交通流预测问题,早期的研究者尝试了许多方法,可以分为动态建模、数据驱动和深度学习^[1]3种。动态建模使用数学工具和物理知识对交通进行建模,这种方法计算量大、准确性低。数据驱动方法又分为统计学方法和机器学习方法。统计学方法有历史平均(historical average, HA)、差分自回归移动平均模型(auto-regressive integrated moving average model, ARIMA)以及向量自回归(vector auto-regressive, VAR)。其中,ARIMA及其变体^[2]是统计学方法中最典型的方法之一。然而,这种方法只考虑时序特征,没有考虑时空相关性和交通流的非线性,所以预测效果较

差。后来,机器学习方法也被应用在交通流预测中,相比于统计学方法,机器学习方法(如支持向量机回归(support vector regression, SVR)^[3]),可以实现更精准的预测并进行更复杂的数据建模,但是其挖掘时空关系能力有限,预测效果依然不太理想。如今,深度学习方法也被广泛应用于交通流预测任务中,研究者们将长短期记忆(long short-term memory, LSTM)网络和门控循环单元(gated recurrent unit, GRU)神经网络^[4-5]以及卷积神经网络(convolutional neural network, CNN)^[6]引入交通流预测中^[7-9]。为提取时空特征,Wu等^[10]使用CNN捕获路网交通数据的空间分布,用GRU捕获交通数据的时间分布。但由于CNN只能应用于标准网格数据,不能很好地反映路网之间的空间关系。于是,研究者引入文献[11]提出的图卷积神经网络(graph convolutional neural network, GCN)提取路网中的空间特征。GCN是一种针对图结构化数据特征提取的特殊CNN,广泛用于节点分类^[12]、图分类^[13-14]、链路预测^[15-16]和交通流预测^[17-18]中。Yu等^[19]提出的时空图卷积网络(spatio-temporal graph convolutional network: a deep learning framework for traffic forecasting, STGCN)除了利用GCN提取空间特征,还利用CNN提取时间特征。文献[20]提出的基于图卷积网络的长短期流量预测(long short-term traffic

prediction with graph convolutional network, LSGCN), 提出了一种图注意力网络, 并将图注意力网络和 GCN 集成到空间门控块中, 通过空间门控块和门控线性单元 (gated linear unit, GLU) 提取时空特征。

然而之前的方法还忽略了几个重要的方面。在时间分布上, 仅用 CNN 或 GLU 只能提取局部时间特征, 忽略了时间分布的全局时间特征。在空间分布上, 当前的交通流预测方法无法考虑不同结构的空信息。一方面, 某些节点的特征具有很强的相似性, 例如同一种功能类型的道路虽然地理位置不接近, 但交通模式是相似的。而使用 GCN 聚合邻居信息时只考虑局部拓扑空信息结构, 忽略了全局语义空信息结构。另一方面, 给定矩阵中每个节点之间的权重是固定的, 节点之间的依赖关系一般被生成静态图, 从而忽略了交通空信息依赖关系的动态特征。

为解决上述问题, 本文提出一种基于混合图卷积的多通道时空交通流预测 (multi-channel spatial-temporal traffic flow prediction based on hybrid static-dynamic graph convolution, MHGCN) 模型。在时间分布上, 为探索全局和局部的时间关系, 该模型先用多头注意力提取全局时间依赖, 再用时间门控机制提取局部时间依赖。在空间分布上, 为弥补先前模型不能提取不同结构空信息特征的缺点, 本文提出多通道空信息模块, 首先定义路网图的空间矩阵和语义矩阵, 分别表示道路的连通性和交通流特征的上下文相似性, 从拓扑空信息结构和语义空信息结构中提取空信息特征。根据输入的交通状况将空信息依赖性建模为动态加权图, 通过自适应节点嵌入, 在图结构信息未知的情况下, 捕获数据中隐藏的空信息关系。考虑拓扑空信息和语义空信息的相似性, 为提取更全面的特征, 本文提出公共图卷积模块使用参数共享的方法提取拓扑空信息和语义空信息的公共信息。最后, 通过多通道聚合模块聚合多个通道

提取到的特征, 获得全面的交通数据信息以预测未来的交通状况。

本文提出的 MHGCN 模型综合考虑了道路上下文相关性和空信息依赖动态性两个问题。与传统的交通流预测模型相比, MHGCN 引入了多通道空信息模块, 利用混合图卷积方法对交通数据进行建模。多通道空信息模块由静态图卷积和动态图卷积组成, 能够从不同的空信息结构中提取特征, 并充分融合静态和动态的空信息, 从而提高了预测精度。其次, 该模型的静态图卷积模块具有全局视野和综合性能, 能够从拓扑空信息结构、语义空信息结构以及它们的组合提取特定和公共的特征。动态图卷积模块则能够自学习节点嵌入, 捕获未知图结构下的动态空信息关系, 增加了模型的灵活性和稳健性。此外, 该模型的时间模块结合了时间门控机制和多头注意力机制, 能够捕获交通数据中的全局时间依赖性和局部时间依赖性。通过从不同时间间隔提取时间信息, 模型在长期的交通流预测任务上表现出良好的性能。最后, 在 4 个公共交通数据集上进行了广泛的实验, 实验结果表明, 该模型相比现有基准模型具有更优越的性能。本文的重要贡献体现在以下 4 个方面。

(1) 提出一种基于混合图卷积的多通道时空交通流预测模型, 该模型的多通道空信息模块由静态图卷积模块和动态图卷积模块组成, 可以从不同的空信息结构中提取空信息, 对不同部分的交通数据进行更精确的建模, 提高模型预测精度。

(2) 静态图卷积模块同时从拓扑空信息结构、语义空信息结构及其组合中提取特定和公共的特征, 建立全局、全面的空信息关系。动态图卷积使用自学习节点嵌入学习潜在的空信息关系, 在图结构信息未知的情况下捕获动态空信息关系, 具有很高的灵活性。最后, 本文通过多通道聚合模块充分融合静态和动态的空信息。

(3) 该模型的时间模块结合时间门控机制和多头注意力机制捕获时间依赖性, 从不同时间间



隔提取时间信息,使模型在长期的交通流预测任务上有较好的效果。

(4) 本文在 PEMS03、PEMS04、PEMS07 和 PEMS08 这 4 个公共交通数据集上进行了广泛的实验,实验结果表明其性能优于现有基准模型。

1 数据描述与问题定义

1.1 问题定义

在本文中,将道路网络定义为无向图 $G=(V,E,A)$,其中 V 是一组节点,即道路网络上的传感器; E 是边的集合; $A \in R^{N \times N}$ 是邻接矩阵,代表节点间的依赖关系。本文采用两种邻接矩阵,一种是空间邻接矩阵 A_{sp} ,另一种是语义邻接矩阵 A_{se} 。本文将时间 t 在 G 上观察到的交通数据表示为 $X^t \in R^{N \times P}$,其中 P 是每个节点的特征数量。假设有 T 个历史观测数据,要预测未来 M 个交通流数据。在给定道路网络图 G 的情况下,交通流预测被表述为学习函数 $f(\cdot)$:

$$[X^{t-T+1}, X^{t-T+2}, \dots, X^t; G] \xrightarrow{f(\cdot)} [X^t, X^{t+1}, \dots, X^{t+M}] \quad (1)$$

本文定义的两种邻接矩阵,形成两个子图:空间邻居子图 G_{sp} ,编码相邻道路的局部连通性;语义邻居子图 G_{se} ,该子图通过文献[21]提出的动态时间规整(dynamic time warping, DTW)算法对道路的全局上下文相似性进行编码。

1.2 邻接矩阵构造

在道路网络中,道路的连接是空间依赖性建模的重要因素。空间邻接矩阵的定义为:

$$A_{ij}^{sp} = \begin{cases} \exp\left(-\frac{d_{ij}^2}{\sigma^2}\right), & \exp\left(-\frac{d_{ij}^2}{\sigma^2}\right) \geq \varepsilon \\ 0, & \text{其他} \end{cases} \quad (2)$$

其中, d_{ij} 是节点 i 和节点 j 之间的欧几里得距离。 σ^2 和 ε 是控制矩阵 A_{sp} 稀疏性的阈值。

交通流预测不仅和道路连通性相关,还与道路的上下文相似性相关。例如,两个地理位置相

距较远的学校,其附近的交通模式有密切的相似之处,而这种相似在空间图上无法显示。因此本文使用 DTW 算法^[22]计算节点之间的交通模式的相似性。DTW 算法是计算两个时间序列数据相似度的动态规划算法,对两个长度不一样的时间序列,DTW 算法先将两个序列对齐,再计算对齐后的序列总距离。欧几里得距离和 DTW 距离之间差异的实例如图 1 所示, X 和 Y 是两个时间序列,可以看出相较于点 c ,点 a 与点 b 更具相似性。

DTW 算法定义为:

$$D(i, j) = \text{dist}(x_i, y_j) + \min(D(i-1, j), D(i, j-1), D(i-1, j-1)) \quad (3)$$

其中, $D(i, j)$ 表示子序列 $X=(x_1, x_2, \dots, x_i)$ 和 $Y=(y_1, y_2, \dots, y_j)$ 之间的最短距离, $\text{dist}(x_i, y_j)$ 表示 x_i 和 y_j 之间的绝对距离。具体来说,给定两个时间序列 $X=(x_1, x_2, \dots, x_m)$ 和 $Y=(y_1, y_2, \dots, y_n)$, X 和 Y 之间的 DTW 距离为 $\text{DTW}(X, Y) = D(m, n)$ 。

与欧几里得距离相比,DTW 距离能够更好地反映两个时间序列的相似性^[23-24],前者只能计算两个长度一致的时间序列的相似度,后者解决了两个时间序列长度不一致的问题。因此,本文通过 DTW 距离定义的语义邻接矩阵如下:

$$A_{ij}^{se} = \begin{cases} 1, & \text{DTW}(X^i, X^j) < \varepsilon \\ 0, & \text{其他} \end{cases} \quad (4)$$

其中, X^i 表示节点 i 的时间序列, ε 用来确定邻接矩阵 A_{se} 的稀疏性。

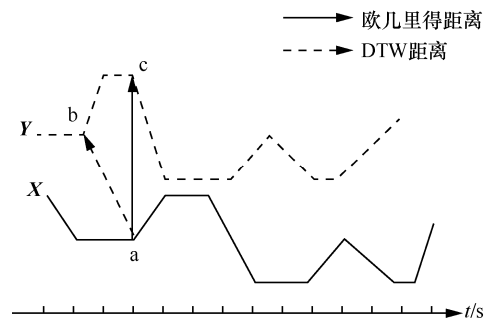


图 1 欧几里得距离和 DTW 距离之间差异的实例

2 MHGCN 模型

本文所提出的 MHGCN 模型是一种基于混合图卷积的多通道时空交通流预测模型，其框架如图 2 所示。该模型由两个时间模块、一个多通道空间模块和一个输出层组成。其中，该模型采用三明治结构（即中间多通道空间模块，两边时间模块）。模型的输入为历史交通数据 X_{in} 、空间邻居图 G_{sp} 和语义邻居子图 G_{se} ，输出为预测的交通流量。

2.1 时间模块

先前的关于时间序列特征提取的模型，如 LSTM 和 GRU 仍然存在迭代耗时、梯度不稳定和对动态变化响应缓慢的问题。因此，该模型的时间模块采用多头注意力机制提取全局时间特征，采用时间门控机制提取局部时间特征。交通数据 X_{in} 先经过线性变换映射到高维空间，增强模型拟合能力：

$$X = X_{in}W + b \quad (5)$$

其中， $X_{in} \in \mathbf{R}^{T \times N \times D}$ ， T 是历史观测数据的时间步长， N 是节点个数， D 是特征数量； $W \in \mathbf{R}^{D \times P}$ ， P 是映射后高维空间的维度； $b \in \mathbf{R}^{1 \times 1 \times P}$ 。

对于单头注意力模型，给定输入 $X^{t-T+1:t} = [X^{t-T+1}, X^{t-T+2}, \dots, X^t] \in \mathbf{R}^{T \times N \times P}$ ，将符号简化为 X ，

通过线性变换将输入矩阵转换为 3 个矩阵，即矩阵 $Q \in \mathbf{R}^{N \times d_q}$ 、矩阵 $K \in \mathbf{R}^{N \times d_k}$ 和矩阵 $V \in \mathbf{R}^{N \times d_v}$ 。

W^Q 、 W^K 、 W^V 是 3 个可训练的参数矩阵，该过程表示为：

$$Q = XW^Q, K = XW^K, V = XW^V \quad (6)$$

接着计算注意力输出：

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (7)$$

在多头注意力机制中，不同的权值矩阵将输入特征线性映射到不同的数据子空间，并在每个子空间完成相同的注意力计算，以对时间序列进行充分学习，其中第 i 头注意力计算过程如下：

$$\text{head}_i = \text{Attention}(Q_i, K_i, V_i) \quad (8)$$

其中， $Q_i = XW_i^Q, K_i = XW_i^K, V_i = XW_i^V$ 。最后将各个 head 串联在一起并再次投影，得到多头注意力输出：

$$\text{MultiHead}(Q, K, V) = \text{Concat}(\text{head}_1, \text{head}_2, \dots, \text{head}_h)W^O \quad (9)$$

其中， $W^O \in \mathbf{R}^{hd_v \times d_{\text{model}}}$ ， h 表示多头的个数。多头注意力机制擅长捕获时间序列的全局依赖性，为探索局部时间关系，接着采用时间门控机制捕获时间特征。时间门控机制使用卷积并行提取局部时间特征，将输入序列首先经过一维卷积操作，

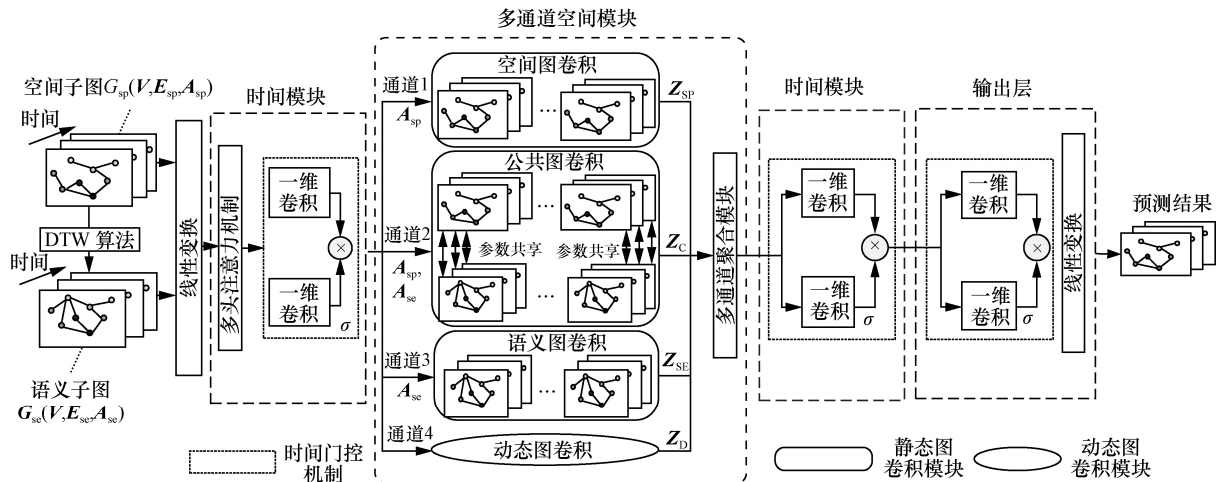


图2 基于混合图卷积的多通道时空交通流预测模型



然后分为两部分，一部分经过 Sigmoid 函数映射，另一部分不做变换，之后两部分结果相乘，得到输出。其中 W 、 V 、 b 、 c 是可训练的参数， $*$ 表示卷积运算， σ 是 Sigmoid 激活函数。计算式如下：

$$h(X) = (X * W + b) \times \sigma(X * V + c) \quad (10)$$

为避免过拟合情况的出现，第二层时间模块只使用时间门控机制提取时间特征。

2.2 多通道空间模块

交通流的空间依赖性与空间结构密切相关，为提取不同空间结构的空间特征，本文设计了多通道空间模块。该模块包括 3 个部分，分别是：静态图卷积模块、动态图卷积模块和多通道聚合模块。

2.2.1 静态图卷积模块

静态图卷积模块包含空间图卷积、语义图卷积和公共图卷积。在空间图卷积中，对空间邻居子图提取特定的特征，第 l 层输出 $Z_{sp}^{(l)}$ 可以表示为：

$$Z_{sp}^{(l)} = \text{ReLU}(\tilde{D}_{sp}^{-1/2} \tilde{A}_{sp} \tilde{D}_{sp}^{-1/2} Z_{sp}^{(l-1)} W_{sp}^{(l)}) \quad (11)$$

其中， $W_{sp}^{(l)}$ 是空间图卷积中第 l 层的权重矩阵，ReLU 是激活函数，初始 $Z_{sp}^{(0)}$ 是第一层时间模块的输出。 $\tilde{A}_{sp} = A_{sp} + I_{sp}$ ， I_{sp} 是单位矩阵， $\tilde{D}_{sp}$ 是 $\tilde{A}_{sp}$ 的对角度矩阵。本文把空间图卷积最后一层的输出表示为 Z_{sp} 。

在语义图卷积中，对语义邻居子图提取特定的特征，第 l 层输出 $Z_{se}^{(l)}$ 可以表示为：

$$Z_{se}^{(l)} = \text{ReLU}(\tilde{D}_{se}^{-1/2} \tilde{A}_{se} \tilde{D}_{se}^{-1/2} Z_{se}^{(l-1)} W_{se}^{(l)}) \quad (12)$$

其中， $W_{se}^{(l)}$ 是语义图卷积中第 l 层的权重矩阵，本文把语义图卷积最后一层的输出表示为 Z_{se} 。

事实上，空间邻居子图和语义邻居子图并非完全不相关，因此不仅需要提取这两个图的特定信息，还要提取这两个图的公共信息。在公共图卷积部分，本文采用参数共享^[25]提取这两个图的公共信息，自适应地学习空间邻居子图和语义邻居子图之间的关联信息。

利用公共图卷积从空间邻居子图中提取节点信息 $Z_{csp}^{(l)}$ ：

$$Z_{csp}^{(l)} = \text{ReLU}(\tilde{D}_{sp}^{-1/2} \tilde{A}_{sp} \tilde{D}_{sp}^{-1/2} Z_{csp}^{(l-1)} W_c^{(l)}) \quad (13)$$

其中， $W_c^{(l)}$ 是公共图卷积的第 l 层权重矩阵， $Z_{csp}^{(l-1)}$ 是第 $(l-1)$ 层节点的输出， $Z_{csp}^{(0)}$ 是第一个时间模块的输出。为提取共享信息，在利用公共图卷积从语义邻居子图中提取节点信息时，本模型在每一层公共图卷积中都共享相同的权重矩阵 $W_c^{(l)}$ ，计算式如下所示：

$$Z_{cse}^{(l)} = \text{ReLU}(\tilde{D}_{se}^{-1/2} \tilde{A}_{se} \tilde{D}_{se}^{-1/2} Z_{cse}^{(l-1)} W_c^{(l)}) \quad (14)$$

共享权重矩阵可以从两个子图中过滤出共享特征。根据不同的输入图，可以得到两个输出 Z_{csp} 和 Z_{cse} ，最后得到公共信息 Z_c 是：

$$Z_c = (Z_{csp} + Z_{cse}) / 2 \quad (15)$$

2.2.2 动态图卷积模块

道路上的车流量是变化的，导致节点间的相互影响也是动态变化的。因此除了捕获节点间的静态空间关系，还要挖掘节点间动态的空间关系。本文采用动态图卷积^[26]对空间特征进行提取，基于交通网络结构构造了一个自适应学习的图结构，可以学习隐藏的节点依赖关系。动态图卷积模块框架如图 3 所示。

动态图卷积模块的关键思想是为不同的节点动态的分配不同的权重，对于节点 i ，本文计算图中所有其他节点信息的加权和：

$$Z_i^{(l)} = \sum_{j \in N} \alpha_{i,j} \cdot Z_j^{(l-1)} \quad (16)$$

其中， $\alpha_{i,j}$ 是表示节点 j 对于节点 i 的重要性的注意力分数， $Z_i^{(l)}$ 表示第 l 层第 i 个节点的嵌入， $\sum_{j \in N} \alpha_{i,j} = 1$ 。

交通流量研究表明，交通状况和交通网络结构都能影响交通流的预测。因此，本文用自学习节点嵌入与隐藏状态连接，并采用缩放点积方法计算注意力：

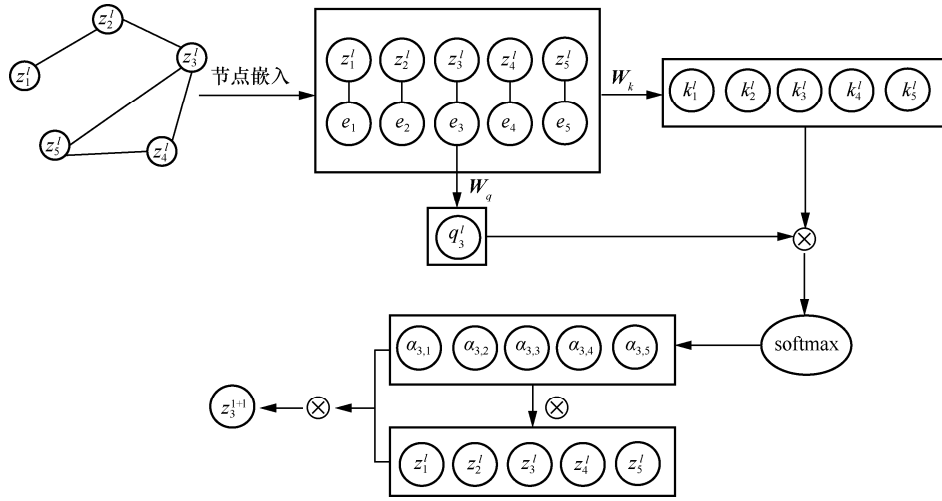


图3 动态图卷积模块框架

$$o_{i,j} = \frac{\langle W_q(Z_i^{l-1} \parallel e_i), W_k(Z_j^{l-1} \parallel e_j) \rangle}{\sqrt{d'}} \quad (17)$$

其中, $\parallel$ 表示连接操作, $\langle \cdot, \cdot \rangle$ 表示内积操作, e_i 是节点 i 的节点嵌入, 节点嵌入是随机初始化并逐渐训练的, $W_k, W_q \in R^{d' \times d_c}$ 是 key 矩阵和 query 矩阵, d_c 是 $Z_i^{l-1} \parallel e_i$ 的维数, 节点嵌入 e_i 、key 矩阵 W_k 和 query 矩阵 W_q 都是待学习的参数。接下来, 用式 (18) 计算注意力分数:

$$\alpha_{i,j} = \frac{\exp(o_{i,j})}{\sum_{k \leq N} \exp(o_{i,k})} \quad (18)$$

得到注意力分数后, 可通过式 (16) 更新节点信息。动态图卷积模块第 l 层的输出为:

$$Z_D^{(l)} = [Z_1^{(l)}, Z_2^{(l)}, \dots, Z_N^{(l)}]^T \quad (19)$$

其中, N 代表节点个数, 本文把动态图卷积模块最后的输出表示为 Z_D 。

2.2.3 多通道聚合模块

考虑预测结果可能与 4 个通道的输出之一及它们的组合相关, 于是对 4 个通道的输出进行聚合。本文提出采用注意力机制对多通道信息进行聚合。静态图卷积模块的输出分别是 Z_{SP} 、 Z_C 、 Z_{SE} , 动态图卷积模块的输出是 Z_D 。使用注意力机制 $\text{att}(Z_{SP}, Z_C, Z_{SE}, Z_D)$ 学习输出结果对应的重

要性 ($\alpha_{sp}, \alpha_c, \alpha_{se}, \alpha_d$), 计算式如下:

$$(\alpha_{sp}, \alpha_c, \alpha_{se}, \alpha_d) = \text{att}(Z_{SP}, Z_C, Z_{SE}, Z_D) \quad (20)$$

其中, $\alpha_{sp}, \alpha_c, \alpha_{se}, \alpha_d \in R^{n \times 1}$ 分别表示 Z_{SP} 、 Z_C 、 Z_{SE} 、 Z_D 的 n 个节点的注意力值。

对于节点 i , 在 Z_{SP} 中的嵌入是 $z_{sp}^i \in R^{h \times 1}$, h 是最后一层图卷积的输出通道。本文首先通过非线性函数对嵌入进行变换, 然后使用共享注意力向量 $q \in R^{h \times 1}$ 得到注意力值 ω_{sp}^i 如下:

$$\omega_{sp}^i = q^T \cdot \tanh(W \cdot (z_{sp}^i)^T + b) \quad (21)$$

其中, $W \in R^{h \times h}$ 是权重矩阵, $b \in R^{h \times 1}$ 是偏置向量。同样, 也可以分别得到嵌入矩阵 Z_C 、 Z_{SE} 、 Z_D 中节点 i 的注意力值 ω_c^i 、 ω_{se}^i 和 ω_d^i 。然后用 softmax 函数对注意力值进行归一化, 得到最终权重:

$$\alpha_{sp}^i = \frac{\exp(\omega_{sp}^i)}{\exp(\omega_{sp}^i) + \exp(\omega_c^i) + \exp(\omega_{se}^i) + \exp(\omega_d^i)} \quad (22)$$

其中, α_{sp}^i 越大意味着相应的嵌入越重要。相应地, $\alpha_c^i = \text{softmax}(\omega_c^i)$, $\alpha_{se}^i = \text{softmax}(\omega_{se}^i)$, $\alpha_d^i = \text{softmax}(\omega_d^i)$ 。对于 n 个节点, 本文学习到的权重是 $\alpha_{sp} = [\alpha_{sp}^1, \alpha_{sp}^2, \dots, \alpha_{sp}^n]^T$ 、 $\alpha_c = [\alpha_c^1, \alpha_c^2, \dots, \alpha_c^n]^T$ 、 $\alpha_{se} = [\alpha_{se}^1, \alpha_{se}^2, \dots, \alpha_{se}^n]^T$ 、 $\alpha_d = [\alpha_d^1, \alpha_d^2, \dots, \alpha_d^n]^T$, 并且定义 $\alpha_{sp} = \text{diag}(\alpha_{sp})$ 、 $\alpha_c = \text{diag}(\alpha_c)$ 、 $\alpha_{se} =$



$\text{diag}(\alpha_{se})$ 、 $\alpha_D = \text{diag}(\alpha_d)$ 。多通道聚合模块如图 4 所示, 本文结合这 4 个输出得到多通道空间模块的最终输出 Z :

$$Z = \alpha_{SP} \cdot Z_{SP} + \alpha_C \cdot Z_C + \alpha_{SE} \cdot Z_{SE} + \alpha_D \cdot Z_D \quad (23)$$

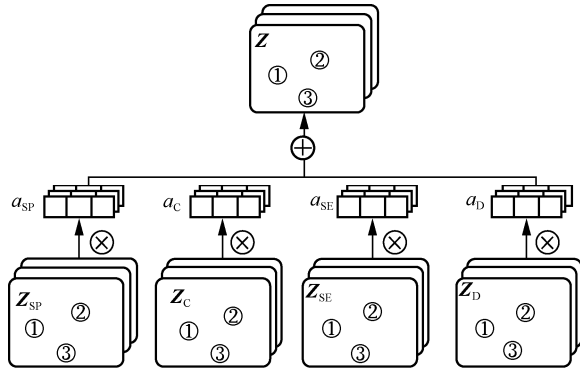


图 4 多通道聚合模块

2.3 输出层

最后, 本文附加一个时间门控机制和一个线性变换作为输出层。为避免预测不准确导致的误差叠加, 直接预测未来的 M 步交通流量。本文将线性变换层的输出通道数设为步长 M , 以获得想要的输出 $\hat{Y}$ 。使用 Huber 损失作为损失函数衡量模型的性能:

$$L(Y, \hat{Y}) = \begin{cases} \frac{1}{2}(Y - \hat{Y})^2, & |Y - \hat{Y}| \leq \delta \\ \delta|Y - \hat{Y}| - \frac{1}{2}\delta^2, & \text{其他} \end{cases} \quad (24)$$

其中, $\hat{Y}$ 是预测值, Y 是真实值, δ 是控制对异常值敏感度的超参数。

3 实验结果与分析

3.1 实验数据

为验证模型 MHGCN 的交通流预测性能, 本文在 4 个数据集上开展实验, 分别是美国加利福尼亚交通运输部门公开的高速交通数据 PEMS03、PEMS04、PEMS07 和 PEMS08, 每个数据集在不同地区采样。PEMS 项目在加利福尼亚州主要大城市的州际公路上部署了超过 39 000 个传感器站点, 每个数据集都包含交通观测的关键属性流量、速度和道路占用率。这 4 个数据集中原始的采样频率为每 30 s 一次, 交

通数据以 5 min 为时间间隔, 即每小时有 12 个流量数据。其中, 60% 的数据用于训练, 20% 用于测试, 其余 20% 用于验证。数据集具体细节如下。

- PEMS03: 从 2018 年 9 月 1 日至 11 月 30 日, 共 91 天, 在加州中北部地区共 358 个检测节点。
- PEMS04: 从 2018 年 1 月 1 日至 2 月 28 日, 共 59 天, 在旧金山湾区共 307 个检测节点。
- PEMS07: 从 2017 年 5 月 1 日至 8 月 31 日, 共 123 天, 在洛杉矶地区共 883 个检测节点。
- PEMS08: 从 2016 年 7 月 1 日至 8 月 31 日, 共 62 天, 在圣贝纳迪诺地区共 170 个检测节点。

每个数据集的空间网络都是根据实际道路网络结构构建的, 如果两个检测器在同一条路上, 则认为这两个节点在空间网络中是相连的。在实验中采用均值方差标准化 (Z-score normalization) 对特征进行标准化:

$$X' = \frac{X - \text{mean}(X)}{\text{std}(X)} \quad (25)$$

其中, X 表示交通数据, $\text{mean}(X)$ 和 $\text{std}(X)$ 分别表示交通数据的均值和标准差。

本文采用 3 个指标评估模型性能, 分别是平均绝对误差 (mean absolute error, MAE)、均方根误差 (root mean square error, RMSE)、平均绝对百分比误差 (mean absolute percentage error, MAPE)。各指标定义:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (26)$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (27)$$

$$\text{MAPE} = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (28)$$

其中, y_i 是测试集数据的真实值, $\hat{y}_i$ 表示预测值。

3.2 对比模型

为验证模型 MHGCN 的预测效果, 本文选取

以下模型与本模型进行对比。

- 全连接长短期记忆 (fully-connected long-short term memory, FC-LSTM) 网络^[27]: 一种改进后的长短期记忆网络。
- 时空图卷积网络 (STGCN)^[19]: 结合图卷积和一维卷积分别捕获空间依赖和时间依赖。
- 扩散卷积循环神经网络 (diffusion convolutional recurrent neural network: data-driven traffic forecasting, DCRNN)^[28]: 提出扩散卷积捕获时空特征, 扩散过程基于图的随机游走。
- 基于注意力机制的时空图卷积网络 (attention based spatial-temporal graph convolutional networks for traffic flow forecasting, ASTGCN)^[29]: 设计一种时空注意力机制捕获动态时空相关性。
- 用于深度时空图建模的图形波网 (graph wavenet for deep spatial-temporal graph modeling, Graph WaveNet)^[30]: 结合自适应图卷积和扩张的随机卷积捕获时空依赖关系。
- 时空同步图卷积网络 (spatial-temporal synchronous graph convolutional networks: a new framework for spatial-temporal network

data forecasting, STSGCN)^[31]: 利用多个局部时空子图模块直接同步捕获局部时空相关性。

- 用于交通流预测的时空图 ODE 网络 (spatial-temporal graph ODE networks for traffic flow forecasting, STGODE)^[32]: 通过基于张量的常微分方程捕获时空动态, 为更全面地理解网络考虑了语义矩阵。

3.3 实验设置

本文利用 12 个历史时间步 (1 h) 的交通数据去预测未来 12 个时间步的交通流量。本文选择 Adam 优化器训练模型, 初始化学率为 0.002, 批处理大小为 16, 训练轮数为 200 轮, 空间邻接矩阵的阈值 σ 和阈值 ε 分别设置为 10 和 0.5, 语义邻接矩阵的阈值 ε 设置为 0.6。静态图卷积模块中, 3 个 GCN 的隐藏层单元个数是相同的。本文的所有实验均在 NVIDIA Tesla V100-SXM2-32GB GPU 上进行。

3.4 实验结果与分析

各模型在 4 个数据集上的性能比较见表 1, 使用过去 12 个时间步预测未来 12 个时间步, 实验重复 10 次, 然后用 10 次实验的平均值表示预测值, 用误差条表征数据的分布, 其中误差条的高度是 $\pm$ 标准差。从表 1 可以看出, MHGCN 在 PEMS07 数据集上的所有指标都优于对比模型。而在

表 1 各模型在 4 个数据集上的性能比较

数据集	指标	FC-LSTM	STGCN	DCRNN	ASTGCN	Graph WaveNet	STSGCN	STGODE	MHGCN (本文模型)
PEMS03	RMSE	35.11 $\pm$ 0.50	30.12 $\pm$ 0.70	30.31 $\pm$ 0.25	29.66 $\pm$ 1.68	32.94 $\pm$ 0.18	29.21 $\pm$ 0.56	27.69 $\pm$ 0.73	27.36 $\pm$ 0.22
	MAE	21.33 $\pm$ 0.24	17.49 $\pm$ 0.46	18.18 $\pm$ 0.15	17.69 $\pm$ 1.43	19.85 $\pm$ 0.03	17.48 $\pm$ 0.15	16.51 $\pm$ 0.06	16.17 $\pm$ 0.10
	MAPE	23.33 $\pm$ 4.23%	17.15 $\pm$ 0.45%	18.91 $\pm$ 0.82%	19.40 $\pm$ 2.24%	19.93 $\pm$ 0.49%	16.78 $\pm$ 0.20%	18.38 $\pm$ 1.10%	17.52 $\pm$ 0.26%
PEMS04	RMSE	44.56 $\pm$ 0.01	35.55 $\pm$ 0.75	38.12 $\pm$ 0.26	35.22 $\pm$ 1.90	39.70 $\pm$ 0.04	33.65 $\pm$ 0.20	32.45 $\pm$ 0.27	32.23 $\pm$ 0.11
	MAE	28.70 $\pm$ 0.01	22.70 $\pm$ 0.64	24.70 $\pm$ 0.22	22.93 $\pm$ 1.29	25.45 $\pm$ 0.03	21.19 $\pm$ 0.10	20.87 $\pm$ 0.11	20.55 $\pm$ 0.07
	MAPE	19.20 $\pm$ 0.01%	14.59 $\pm$ 0.21%	17.12 $\pm$ 0.37%	16.56 $\pm$ 1.36%	17.29 $\pm$ 0.24%	13.90 $\pm$ 0.05%	14.98 $\pm$ 0.42%	14.44 $\pm$ 0.21%
PEMS07	RMSE	45.84 $\pm$ 0.57	38.78 $\pm$ 0.58	38.58 $\pm$ 0.70	42.57 $\pm$ 3.31	42.78 $\pm$ 0.07	39.03 $\pm$ 0.27	36.75 $\pm$ 0.49	34.82 $\pm$ 0.06
	MAE	29.98 $\pm$ 0.42	25.38 $\pm$ 0.49	25.30 $\pm$ 0.52	28.05 $\pm$ 2.34	26.85 $\pm$ 0.05	24.26 $\pm$ 0.14	23.95 $\pm$ 0.60	22.05 $\pm$ 0.05
	MAPE	13.20 $\pm$ 0.53%	11.08 $\pm$ 0.18%	11.66 $\pm$ 0.33%	13.92 $\pm$ 1.65%	12.12 $\pm$ 0.41%	10.21 $\pm$ 1.65%	10.98 $\pm$ 0.75%	9.82 $\pm$ 0.08%
PEMS08	RMSE	34.06 $\pm$ 0.32	27.83 $\pm$ 0.20	27.83 $\pm$ 0.05	28.16 $\pm$ 0.48	31.05 $\pm$ 0.07	26.80 $\pm$ 0.18	25.97 $\pm$ 0.19	25.89 $\pm$ 0.15
	MAE	22.20 $\pm$ 0.18	18.02 $\pm$ 0.14	17.86 $\pm$ 0.03	18.61 $\pm$ 0.40	19.13 $\pm$ 0.08	17.13 $\pm$ 0.09	16.79 $\pm$ 0.13	16.62 $\pm$ 0.07
	MAPE	14.20 $\pm$ 0.59%	11.40 $\pm$ 0.10%	11.45 $\pm$ 0.03%	13.08 $\pm$ 1.00%	12.68 $\pm$ 0.57%	10.96 $\pm$ 0.07%	11.84 $\pm$ 0.52%	11.52 $\pm$ 0.10%



PEMS03、PEMS04 和 PEMS08 数据集上, MHGCN 除了指标 MAPE 略高于 STSGCN, 其余指标均优于对比模型。从实验结果可以观察到, 深度学习方法优于传统的机器学习方法, 因为传统的机器学习方法, 例如 FC-LSTM 只考虑了时间相关性, 忽略了时空结合在交通流预测方面的重要性。而除了 FC-LSTM 之外的方法, 都利用了空间信息, 因此比仅用时间序列预测的方法具有更好的性能。STGCN 使用一维卷积提取时间特征, DCRNN 使用循环神经网络提取时间特征, 这两种模型都使用图卷积提取空间特征, 所以比只考虑时间特征的模型效果好。尽管 Graph WaveNet 是基于图结构的模型, 但是由于该模型结构复杂, 模型参数较多, 所以预测效果较差。ASTGCN 考虑了预测任务的周期性, 因此效果优于其他基于图结构的模型。STSGCN 考虑了空间信息和时间信息之间的相互作用, 因此预测效果不错, 但其只考虑局部信息的特点导致其预测性能不高。STGODE 构建了两类型矩阵, 在空间特征提取时既考虑空间相关性又考虑上下文相关性, 比只考虑空间相关性的模型效果好, 但其忽略了交通空间的动态特征, 对空间特征提取不够充分。本文提出的模型不仅考虑空间相关性和上下文相关性, 同时使用自学习节点嵌入学习潜在的动态空间关系, 因此本文的模型比这些方法具有更好的性能。

与对比模型相比, 本文提出的模型在 4 个数据集上的性能大部分优于对比模型, 这也表明本模型的泛化能力强。本文提出的模型结构简单, 没有纯粹地堆叠网络层数, 在静态图卷积部分采用共享参数的策略减少模型参数, 这也是本模型效果好于对比模型的原因。各模型在 4 个数据集上预测未来 1 h 的平均性能如图 5 所示, 展示了本文模型和对比模型在 4 个数据集上预测未来 1 h 的平均性能比较, 利用折线图更直观地反映不同模型之间性能的差异。

3.5 参数敏感性实验

图卷积中的隐藏神经元个数表示空间信息汇聚

维度的大小, 对模型的预测精度影响较大, 所以本文通过对比实验确定静态图卷积中的两个隐藏神经元的个数 n_1 和 n_2 。设置 n_1 和 n_2 的网格搜索范围为 $\{64, 128, 256\}$, 选择最优参数组合。MHGCN 模型不同参数组合的性能比较见表 2, 可以看到在这 9 组组合中, $n_1 = 256$, $n_2 = 128$ 的组合在 PEMS03 数据集上, MAPE 指标略高于组合 $n_1 = 128$, $n_2 = 256$, 而 RMSE 和 MAE 指标均优于其他组合。在 PEMS04、PEMS07 和 PEMS08 这 3 个数据集上, 该组合的所有指标均优于其他组合。因此本文选择效果最好的一组参数 $n_1 = 256$, $n_2 = 128$ 。

3.6 消融实验

本节通过消融实验验证 MHGCN 模型中各个模块的有效性, MHGCN 变体的预测性能对比见表 3。为分析各个模块的有效性, 本文从以下 4 个方面在 4 个数据集上进行消融实验:

- MHGCN-two 包括两层时空模块;
- MHGCN-S 删除多通道空间模块中的动态图卷积模块, 仅保留静态图卷积模块;
- MHGCN-T 删除第一层时间模块中的多头注意力机制;
- MHGCN-D 删除多通道空间模块中的静态图卷积模块, 仅保留动态图卷积模块。

从表 3 可以看出, MHGCN 的性能在数据集 PEMS07 和 PEMS08 上均优于其他模型, 在数据集 PEMS03 和 PEMS04 上除了 MAPE 指标略高于 MHGCN-T, MAE 指标和 RMSE 指标都优于其他模型。MHGCN-two 模型堆叠了两层时空模块, 表明该模型一层时空模块比两层时空模块性能更好。MHGCN-S 模型表明动态图卷积模块的重要性, 仅用静态图卷积提取空间信息导致节点之间的权重固定, 不能灵活地提取邻居信息。MHGCN-T 模型表明时间模块中多头注意力模块的重要性, 在时间特征提取上使用多头注意力和时间门控机制可以有效地提取全局时间信息和局部时间信息。MHGCN-D 模型表明静态

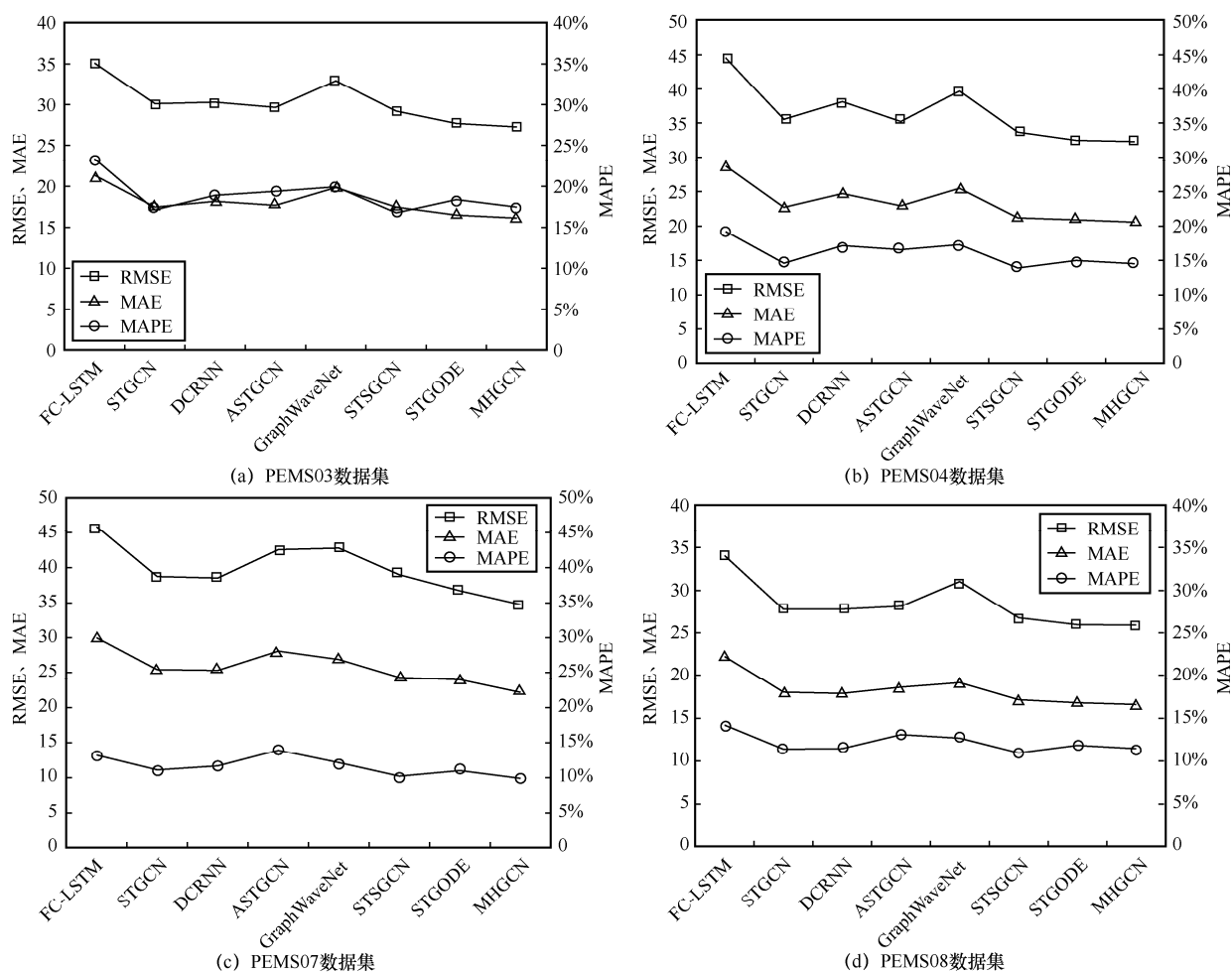


图5 各模型在4个数据集上预测未来1h的平均性能

表2 MHGCN模型不同参数组合的性能比较

数据集	指标	$n_1 = 64$			$n_1 = 128$			$n_1 = 256$		
		$n_2 = 64$	$n_2 = 128$	$n_2 = 256$	$n_2 = 64$	$n_2 = 128$	$n_2 = 256$	$n_2 = 64$	$n_2 = 128$	$n_2 = 256$
PEMS03	RMSE	27.30	27.47	27.90	27.73	27.98	27.97	28.51	27.30	27.60
	MAE	16.17	16.43	16.52	16.27	16.37	16.53	16.73	16.14	16.24
	MAPE	17.33%	18.15%	18.26%	18.35%	17.39%	17.20%	21.21%	17.33%	17.77%
PEMS04	RMSE	32.32	32.61	32.65	32.34	32.51	33.32	32.40	32.23	32.67
	MAE	20.72	20.77	20.69	20.71	20.67	21.04	20.79	20.55	20.65
	MAPE	14.50%	14.55%	14.58%	14.59%	14.57%	14.42%	14.58%	14.34%	14.36%
PEMS07	RMSE	35.11	35.08	34.94	35.00	34.73	34.77	35.00	34.71	34.94
	MAE	22.61	22.46	22.44	22.43	22.15	22.18	22.45	22.13	22.38
	MAPE	10.59%	10.08%	10.11%	10.10%	9.92%	10.05%	10.11%	9.84%	10.13%
PEMS08	RMSE	26.25	25.97	26.48	26.24	26.14	26.16	26.13	25.89	26.25
	MAE	16.93	16.83	17.14	16.83	16.84	16.81	16.71	16.67	16.86
	MAPE	11.64%	11.52%	12.00%	11.55%	11.70%	11.75%	11.90%	11.51%	11.46%

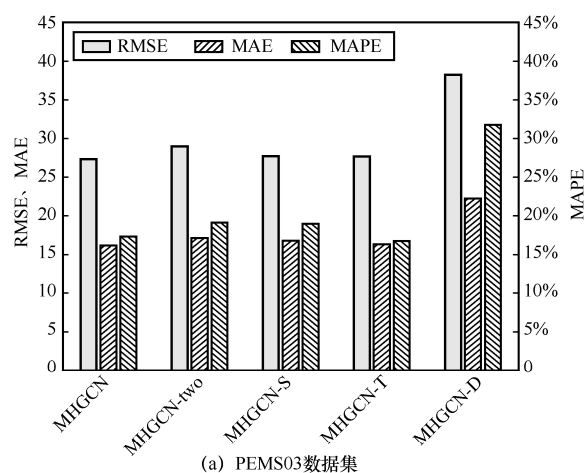


表3 MHGCN 变体的预测性能对比

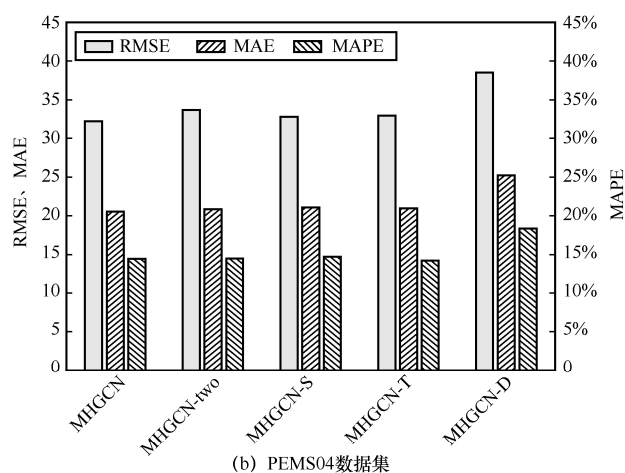
数据集	指标	MHGCN-two	MHGCN-S	MHGCN-T	MHGCN-D	MHGCN
PEMS03	RMSE	28.99	27.71	27.68	38.25	27.30
	MAE	17.12	16.80	16.33	22.23	16.14
	MAPE	19.10%	18.98%	16.76%	31.77%	17.33%
PEMS04	RMSE	33.68	32.81	32.98	38.54	32.23
	MAE	20.85	21.08	20.97	25.22	20.55
	MAPE	14.46%	14.69%	14.23%	18.33%	14.34%
PEMS07	RMSE	36.51	35.58	35.57	56.86	34.71
	MAE	22.70	22.88	22.90	33.90	22.13
	MAPE	10.12%	10.39%	10.45%	15.75%	9.84%
PEMS08	RMSE	27.65	26.74	26.70	32.29	25.89
	MAE	17.64	17.12	17.16	21.15	16.67
	MAPE	11.43%	12.50%	11.37%	14.40%	11.51%

图卷积模块的重要性。其中静态图卷积模块对模型的影响最大,预测精度最高。可以看出,本文提出的模型 3 个模块都能提高模型的预测性能。

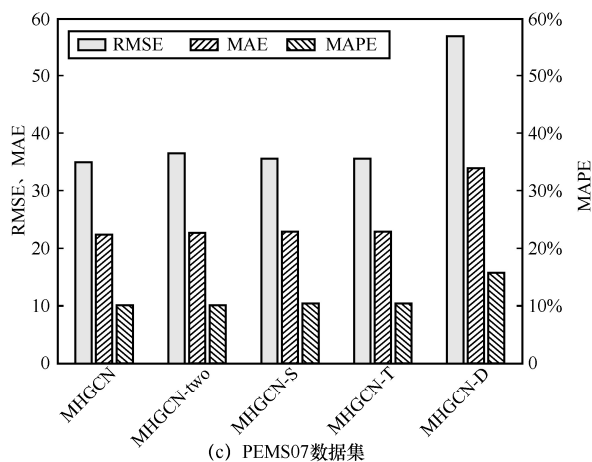
MHGCN 各变体模型预测性能对比如图 6 所示,其使用柱状图更直观地反映了 MHGCN 各变体的预测性能。



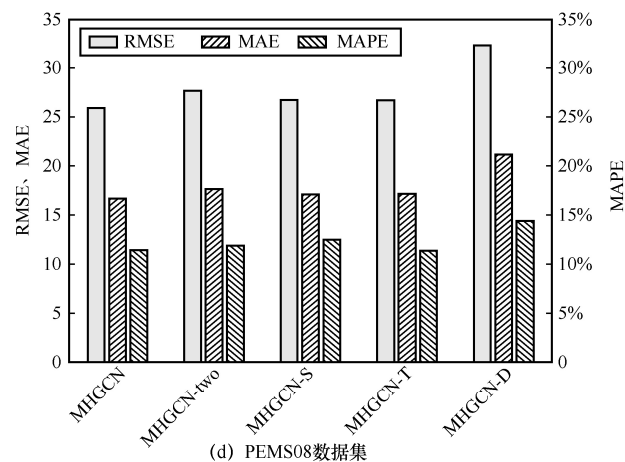
(a) PEMS03数据集



(b) PEMS04数据集



(c) PEMS07数据集



(d) PEMS08数据集

图6 MHGCN 各变体模型预测性能对比

4 结束语

本文提出一种基于混合图卷积的多通道时空交通流预测模型, 针对交通流数据的动态性、时空性与非线性等特点, 采用多通道空间模块, 既在拓扑空间结构和语义空间结构上提取静态空间信息, 又引入动态图卷积在未知图结构中提取动态空间信息; 在静态图卷积模块中, 对空间子图、语义子图及其组合中提取特定和公共的特征; 在动态图卷积模块中, 使用自学习节点嵌入学习动态信息; 在时间模块, 采用多头注意力捕获全局时间依赖性。本文在 4 个公开数据集上进行实验, 实验结果表明 MHGCN 模型在这 4 个数据集上的表现优于所对比的基准模型。这表明 MHGCN 模型可以很好地捕获不同结构的特征, 同时, 消融实验的结果表明该模型的每个模块都有效地提高了模型的预测精度。

在未来的工作中, 将在以下方向进行拓展。首先, 考虑交通系统中条件传播存在延迟性, 计划对空间信息传播中的时延进行建模。其次, 计划研究多属性图学习方法, 交通流预测存在许多外部干扰因素, 如天气、兴趣点、事故等。未来如何充分挖掘这些数据, 建立定量的评价指标, 对于后续模型的改进非常重要。对多属性变量进行相关性建模反映变量之间的相互作用, 从而更准确地预测, 是一个值得研究的方向。

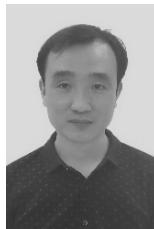
参考文献:

- [1] JIANG Y L, FAN J B, LIU Y, et al. Deep graph Gaussian processes for short-term traffic flow forecasting from spatiotemporal data[J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(11): 20177-20186.
- [2] LUO X L, NIU L Y, ZHANG S R. An algorithm for traffic flow prediction based on improved SARIMA and GA[J]. KSCE Journal of Civil Engineering, 2018, 22(10): 4107-4115.
- [3] LUO X L, LI D Y, ZHANG S R. Traffic flow prediction during the holidays based on DFT and SVR[J]. Journal of Sensors, 2019: 1-10.
- [4] KNOL D, LEEUW F D, MEIRINK J F, et al. Deep learning for solar irradiance nowcasting: a comparison of a recurrent neural network and two traditional methods[C]//International Conference on Computational Science. Heidelberg: Springer-Verlag, 2021: 309-322.
- [5] ZHAO Z, CHEN W H, WU X M, et al. LSTM network: a deep learning approach for short-term traffic forecast[J]. IET Intelligent Transport Systems, 2017, 11(2): 68-75.
- [6] GUO S N, LIN Y F, LI S J, 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.
- [7] ZHANG J B, ZHENG Y, QI D K. Deep spatio-temporal residual networks for citywide crowd flows prediction[C]//Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence. New York: ACM Press, 2017: 1655-1661.
- [8] YAO H X, TANG X F, WEI H, et al. Revisiting spatial-temporal similarity: a deep learning framework for traffic prediction[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2019, 33(1): 5668-5675.
- [9] 廖熙雯, 冷甦鹏, 明显君, 等. 基于数字孪生的城市交通流智能预测与导引策略[J]. 电信科学, 2023, 39(3): 70-79.
- [10] LIAO X W, LENG S P, MING Y J, et al. Digital twin based intelligent urban traffic forecasting and guidance strategy[J]. Telecommunications Science, 2023, 39(3): 70-79.
- [11] WU Y K, TAN H C, QIN L Q, et al. A hybrid deep learning based traffic flow prediction method and its understanding[J]. Transportation Research Part C: Emerging Technologies, 2018(90): 166-180.
- [12] BRUNA J, ZAREMBA W, SZLAM A, et al. Spectral networks and locally connected networks on graphs[EB]. 2013: arXiv: 1312.6203.
- [13] ZHAO T X, ZHANG X, WANG S H. GraphSMOTE: imbalanced node classification on graphs with graph neural networks[C]//Proceedings of the 14th ACM International Conference on Web Search and Data Mining. New York: ACM Press, 2021: 833-841.
- [14] LUO X, JU W, QU M, et al. DualGraph: improving semi-supervised graph classification via dual contrastive learning[C]//Proceedings of 2022 IEEE 38th International Conference on Data Engineering (ICDE). Piscataway: IEEE Press, 2022: 699-712.
- [15] ZENG J Q, XIE P T. Contrastive self-supervised learning for graph classification[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2021, 35(12): 10824-10832.
- [16] ROSSI A, BARBOSA D, FIRMANI D, et al. Knowledge graph embedding for link prediction: a comparative analysis[J]. ACM Transactions on Knowledge Discovery from Data, 2021, 15(2): 1-49.
- [17] DING K Z, XU Z, TONG H H, et al. Data augmentation for deep graph learning[J]. ACM SIGKDD Explorations Newsletter, 2022, 24(2): 61-77.



- [17] ZHANG J L, CHEN F, GUO Y N, 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.
- [18] ZHENG C P, FAN X L, WANG C, et al. GMAN: a graph multi-attention network for traffic prediction[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(1): 1234-1241.
- [19] YU B, YIN H T, ZHU Z X. Spatio-temporal graph convolutional networks: a deep learning framework for traffic forecasting[C]//Proceedings of the 27th International Joint Conference on Artificial Intelligence. Menlo Park: AAAI Press, 2018: 3634-3640.
- [20] HUANG R Z, HUANG C Y, LIU Y B, et al. LSGCN: long short-term traffic prediction with graph convolutional networks[C]//Proceedings of the Twenty-Ninth International Conference on International Joint Conferences on Artificial Intelligence. 2021: 2355-2361.
- [21] BERNDT D J, CLIFFORD J. Using dynamic time warping to find patterns in time series[C]//Proceedings of the 3rd International Conference on Knowledge Discovery and Data Mining. Menlo Park: AAAI Press, 1994: 359-370.
- [22] WANG L, KONIUSZ P. Uncertainty-DTW for time series and sequences[C]//Proceedings of European Conference on Computer Vision. Heidelberg: Springer-Verlag, 2022: 176-195.
- [23] SUN H, SUN M S, WENG J, et al. Analysis of ID sequences similarity using dtw in intrusion detection for CAN bus[J]. IEEE Transactions on Vehicular Technology, 2022, 71(10): 10426-10441.
- [24] 杜辉, 郑长亮, 苗春雨, 等. 一种动态特征匹配的部分重叠点云配准方法[J]. 电信科学, 2021, 37(4): 97-107.
- DU H, ZHENG C L, MIAO C Y, et al. A partial overlapping point cloud registration method based on dynamic feature matching[J]. Telecommunications Science, 2021, 37(4): 97-107.
- [25] WANG X, ZHU M Q, BO D Y, et al. AM-GCN: adaptive multi-channel graph convolutional networks[C]//Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. New York: ACM Press, 2020: 1243-1253.
- [26] TIAN C Y, CHAN W K, VICTOR C. Spatial - temporal attention wavenet: a deep learning framework for traffic prediction considering spatial - temporal dependencies[J]. IET Intelligent Transport Systems, 2021, 15(4): 549-561.
- [27] SUTSKEVER I, VINYALS O, LE Q V. Sequence to sequence learning with neural networks[C]//Proceedings of the 27th International Conference on Neural Information Processing Systems. Cambridge: MIT Press, 2014: 3104-3112.
- [28] LI Y G, YU R, SHAHABI C, et al. Diffusion convolutional recurrent neural network: data-driven traffic forecasting[C]//Proceedings of International Conference on Learning Representations. [S.l.:s.n.], 2018: 1-15.
- [29] GUO S N, LIN Y F, FENG N, et al. Attention based spatial-temporal graph convolutional networks for traffic flow forecasting[C]//Proceedings of the AAAI conference on artificial intelligence. Menlo Park: AAAI Press, 2019: 922-929.
- [30] WU Z H, PAN S R, LONG G D, et al. Graph WaveNet for deep spatial-temporal graph modeling[C]//Proceedings of the 28th International Joint Conference on Artificial Intelligence. Menlo Park: AAAI Press, 2019: 1907-1913.
- [31] SONG C, LIN Y F, GUO S N, et al. Spatial-temporal synchronous graph convolutional networks: A new framework for spatial-temporal network data forecasting[J]//Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(1): 914-921.
- [32] FANG Z, LONG Q Q, SONG G J, et al. Spatial-temporal graph ODE networks for traffic flow forecasting[C]//Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining. New York: ACM Press, 2021: 364-373.

[作者简介]



张雄涛 (1984-), 男, 博士, 湖州师范学院副教授、硕士生导师, 主要研究方向为机器学习、深度学习、模糊系统等。



郑景玉 (1997-), 女, 湖州师范学院硕士生, 主要研究方向为深度学习、智慧交通。



申情 (1982-), 女, 湖州师范学院教授、硕士生导师, 主要研究方向为智能决策、智能信息处理。

孙丹枫 (1988-), 男, 博士, 杭州电子科技大学副教授、硕士生导师, 主要研究方向为模式识别与深度学习。

蒋云良 (1967-), 男, 博士, 浙江师范大学教授、博士生导师, 主要研究方向为智能信息处理、GIS 等。