

基于时空图注意力网络的交通流量预测

邓强^{1,2*}, 李彤^{2,3#}, 苗晨^{1,2}

¹云南大学软件学院, 云南 昆明

²云南省软件工程重点实验室, 云南 昆明

³云南农业大学大数据学院, 云南 昆明

收稿日期: 2021年10月24日; 录用日期: 2021年11月22日; 发布日期: 2021年11月29日

摘要

准确的交通流预测对于提高智能交通系统的安全性、稳定性和效率至关重要。然而, 考虑到交通流之间复杂的空间和时间依赖性, 对交通网络中的时空关系进行建模是一项具有挑战性的任务。本文设计了一种新颖的时空图注意力网络(STGAL)来提取交通流动态和静态的时空相关性, 更加有效解决交通流量预测问题。具体地说, 它主要有三个模块: 1) 图注意力网络用于捕捉节点之间的空间关系, 对邻域节点进行有区别的信息聚合; 2) 长短期记忆网络用于捕获交通流的时间相关性; 3) 空间和时间注意力网络用于捕捉交通流中动态变化的时空关系。我们不仅考虑邻域节点的特征和边的权重来生成新的节点表示, 而且考虑交通流动态变化的时空模式。此外, 为了挖掘周期性数据对预测任务的影响, 我们融合了交通流最近、日周期和周周期的三个组件特征信息。在多步交通预测任务上的大量实验证明了STGAL的有效性和优越性。

关键词

智能交通, 交通流量预测, 时空相关性, 图注意力网络

Spatial-Temporal Graph Attention Networks for Traffic Flow Forecasting

Qiang Deng^{1,2*}, Tong Li^{2,3#}, Chen Miao^{1,2}

¹School of Software, Yunnan University, Kunming Yunnan

²Key Laboratory in Software Engineering of Yunnan Province, Kunming Yunnan

³School of Big Data, Yunnan Agricultural University, Kunming Yunnan

Received: Oct. 24th, 2021; accepted: Nov. 22nd, 2021; published: Nov. 29th, 2021

*第一作者。

#通讯作者。

Abstract

Accurate traffic flow forecasting is critical in improving safety, stability, and efficiency of intelligent transportation systems. However, considering the complex spatial and temporal dependence between traffic flows, modeling the spatial-temporal correlation in traffic is a challenging task. In this paper, we design a novel spatial-temporal graph attention networks (STGAL) to extract the dynamic and static spatial-temporal correlation of traffic flow simultaneously, and effectively address the problem of traffic flow forecasting. Specifically, there are three main modules: 1) Graph attention network is used to capture the spatial correlation between nodes and to aggregate the information of the neighborhood nodes differently; 2) Long short-term memory network to capture the temporal correlation of traffic flow; 3) Spatial and temporal attention networks to capture the spatial-temporal correlation of dynamic changes in traffic flow. We consider the characteristics of neighborhood nodes, weights of edges and spatial-temporal pattern of traffic flow dynamics. In addition, we integrate the recent, daily, and weekly component feature information of traffic flow to mine the impact of periodic data on prediction tasks. A large number of experiments on multi-step traffic forecasting tasks have proved the effectiveness and superiority of STGAL.

Keywords

Intelligent Transportation, Traffic Flow Forecasting, Spatial-Temporal Correlations, Graph Attention Networks

Copyright © 2021 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

及时、准确地交通流分析和预测是建设现代化智慧城市过程中必不可少的部分。交通流量预测作为智能交通系统(ITS)的重要组成部分[1], 它对于城市交通规划, 交通管理和交通控制具有非常重要的意义。通过提前准确地对车流量进行预测, 交通管理部门就能事先对车辆进行疏导, 可以提高路网的运行能力和效率, 减少交通事故的发生, 降低人们的出行时间和成本, 降低环境污染。传统的交通流量预测方法简单地采用时间序列模型, 例如历史均值(HA) [2], 向量自回归(VAR) [3]模型等机器学习方法。然而, 这些方法依赖于理想的平稳假设, 这对于高度的非线性和复杂性的流量数据往往不合乎实际情况。最近, 深度学习的方法受到了研究者的青睐, 深度递归神经网络(RNN)及其改进方法, 例如长短期记忆网络(LSTM) [4]已被广泛应用于从大量的序列信息中学习其中潜在的模式。然而, 该模型将来自不同道路的交通信息视为独立的数据流, 这就忽略了交通数据中的空间关系。因为交通数据可以表示为道路网络上的时间序列, 其中道路之间的连接性是根据道路之间的空间邻近性和模式的相似性来决定的。因此, 图卷积神经网络进入了研究者的视线, 并逐渐成为交通领域热门的研究点, 如[5] [6] [7] [8]被提出以捕捉交通流中的时空信息, 结合了循环神经网络和图卷积神经网络来建模交通流量中的空间和时间依赖关系。

在智能交通领域, 交通流量预测是一种典型的时空图数据挖掘问题。图 1 表示了交通流量复杂的时空关系。节点表示分布在城市道路上的传感器, 红色五角星节点表示预测结果, 节点颜色越深表示对目

标节点的空间关系影响越大，相应的虚线颜色深浅表示时间关系的影响大小。具体地，第一，在空间维度上，图 1 显示不同节点对预测目标节点的影响是不同的，而且相同节点对预测目标节点的影响也是随时间变化的。可见，不同道路节点之间的空间相关性是非线性的和高度动态的。第二，在时间维度上，交通流量具有动态时间依赖性。如图 1 所示，同一节点不同时刻的观测值呈现非线性变化，相邻时间节点之间的交通流量存在很强的相似性，然而较远时间节点的交通状态对预测时间节点的影响有时要大于较近时间点的影响。最后，交通流量在时空维度都存在很强的动态相关性，而且这种高度动态的时空相关性是同时存在的。因此，如何从这些非线性的时空数据中挖掘出这些复杂的时空相关性，是时空图数据挖掘任务的一项极大挑战。

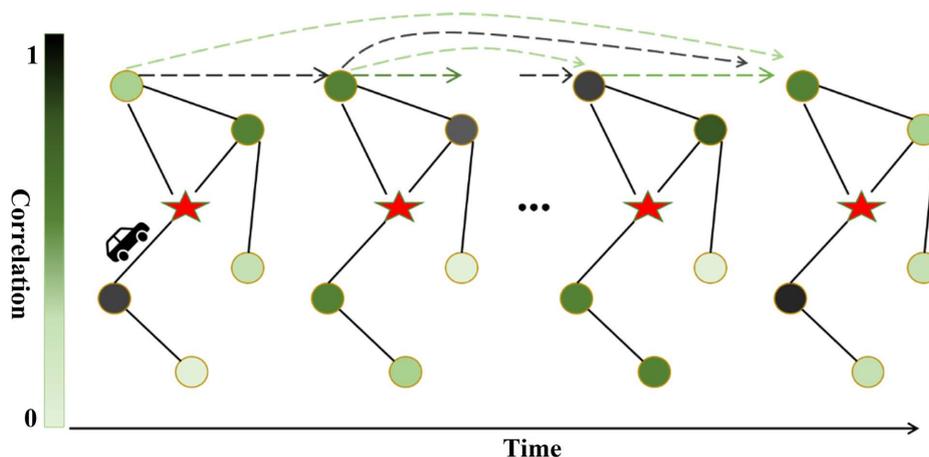


Figure 1. The complex spatial-temporal correlations of traffic flow
图 1. 交通流量复杂的时空关系

为了解决上述问题，我们提出了一种新的基于图注意力网络的深度学习模型，以有效地捕获道路交通流中复杂的时空依赖关系。本文的主要贡献：

- 我们提出了一种新颖的面向交通流量预测的时空图注意力网络(STGAL)。该模型能够更加准确提取交通流动态和静态的时空特征，有效解决交通流量预测问题。
- 我们不仅考虑邻域节点的特征和边的权重来生成新的节点表示，而且考虑交通流动态变化的时空模式。使用图注意力网络对邻域节点进行有区别的信息聚合；使用时空注意力块，用于捕捉交通流中动态变化的时空关系。

在多步交通预测任务上的大量实验，通过与各种基准方法对比分析，证明了我们的方法(STGAL)的有效性和优越性。

2. 研究方法

2.1. 问题定义

交通流量预测是根据历史交通流量数据特征来预测未来的交通流量。进一步，历史交通流量数据被表示为交通图网络上的时间序列。交通图网络定义为无向图 $G = (V, E, A)$ ，其中 V 为图上的节点集(文中是指交通道路上分布的传感器)， $|V| = N$ 是节点数目， E 为边集，表示节点 V 之间的连接关系， $A \in R^{N \times N}$ 是图 G 的邻接矩阵，表示节点 V 间的空间关系。因此，历史交通流量数据就可以表示为 $x = (x_t, x_{t-1}, \dots, x_{t-T+1}; G)$ 作为输入的长度 T 的交通序列，我们的核心目标是为了通过训练找到一个函数即 F 来预测下一个时间步长 p ，表示为 $x' = (x_{t+1}, x_{t+2}, \dots, x_{t+p})$ ，具体地定义公式如下所示，其中 θ 是可学习的参数：

$$(x_{t+1}, x_{t+2}, \dots, x_{t+p}) = F_g(x_t, x_{t-1}, \dots, x_{t-T+1}; G) \quad (1)$$

2.2. 模型设计

图 2 给出了本文提出的 STGAL 模型的总体框架。它由三个结构相同的独立组件组成，分别对历史数据的近期相关性、日周期相关性和周周期相关性进行建模。每一个组件包含相同的时空模块并通过残差连接。具体来说，空间注意力用来提取动态的空间关系，长短期记忆网络提取时间相关性，图注意力网络用来提取空间相关性，时间注意力提取动态的时间关系。输出层采用 2-D 卷积神经网络进行多步输出。最后，我们将所有组件的卷积输出相加作为我们模型的预测，并采用均方根误差即 MSE 损失来衡量预测与它的实际情况之间的差异。接下来我们将会详细描述 ST BLOCK 中的每一个模块以及多个组件融合的方法。

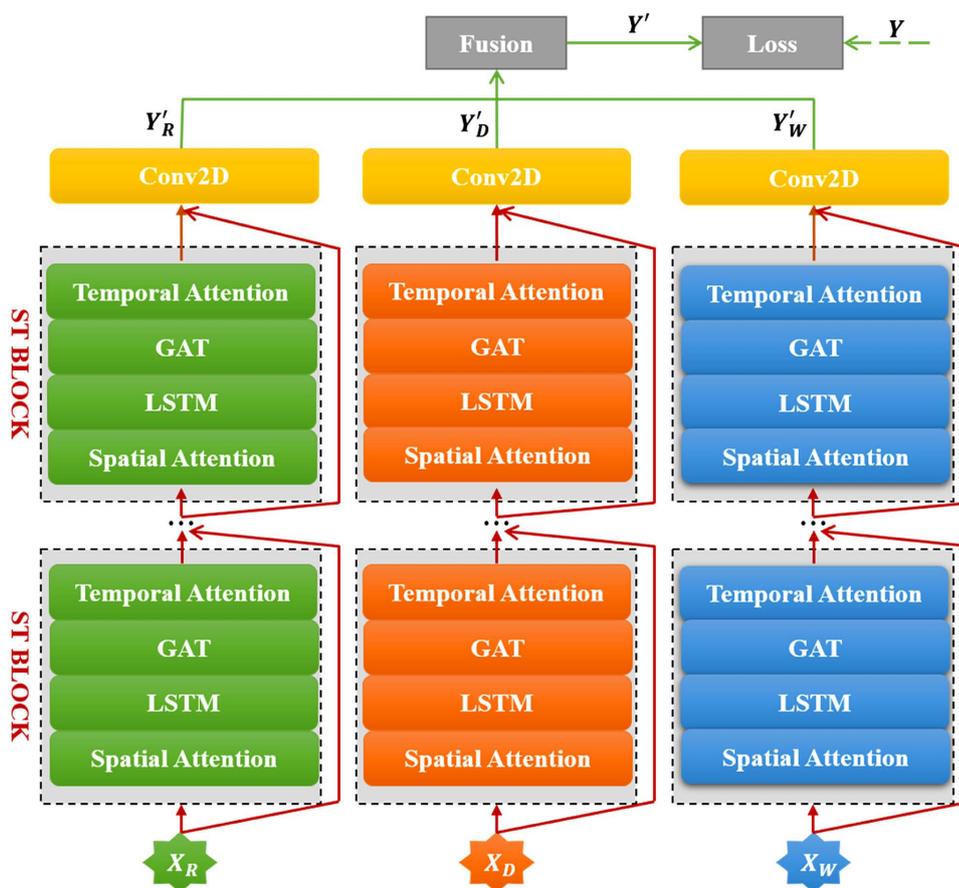


Figure 2. The framework of the spatial-temporal graph attention networks for traffic flow forecasting (STGAL)
图 2. 基于时空图注意力网络的交通流量预测模型框架(STGAL)

2.2.1. 空间注意力网络

一条道路的交通状况会受到其他道路不同程度的影响。这种影响是高度动态的，随着时间的推移而变化。为了对这些属性进行建模，我们使用一种注意力机制[9]自适应地学习空间维度中节点之间的动态相关性。其关键思想是在不同的时间步长为不同的顶点动态分配不同的权重。为了稳定学习过程，我们将空间注意机制扩展为多头注意机制。计算公式如下：

$$S = V_s \cdot \sigma((x_s W_1) W_2 (W_3 x_s^T)^T + b_s) \quad (2)$$

$$a_{i,j}^{(k)} = \frac{\exp(S_{i,j}^{(k)})}{\sum_{j=1}^N \exp(S_{i,j}^{(k)})} \quad (3)$$

其中 x_s 是空间注意力机制输入的特征序列, $V_s, b_s, W_1, W_2, W_3 \in R^{N \times N}$ 是可学习的参数, $S_{i,j}^{(k)}$ 表示节点 i 和节点 j 之间的关联强度, 然后使用 Softmax 函数来确保节点的关注权重系数和为 1。通过空间注意力机制我们可以得到空间关注度矩阵 $S' \in R^{N \times N}$ 来提取交通流量动态的空间关系。

2.2.2. LSTM 网络

交通流量有非常强的时间相关性, 尤其是相近的交通流量情况往往有着相似的模式。我们使用 RNN 的强大变体 LSTM [4] 来提取交通流量的时间相关性。在该模型中, 原来的 LSTM 中的门结构和隐藏状态保持不变, 但输入用空间注意力的输出 S' 取代, 定义如下:

$$f_t = \sigma(W_f \cdot S'_t + U_f \cdot h_{t-1} + b_f) \quad (4)$$

$$i_t = \sigma(W_i \cdot S'_t + U_i \cdot h_{t-1} + b_i) \quad (5)$$

$$o_t = \sigma(W_o \cdot S'_t + U_o \cdot h_{t-1} + b_o) \quad (6)$$

$$\tilde{C}_t = \tanh(W_c \cdot S'_t + U_c \cdot h_{t-1} + b_c) \quad (7)$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (8)$$

$$h_t = o_t * \tanh(C_t) \quad (9)$$

其中 \cdot 是矩阵乘法运算符。 $W_f, W_i, W_o, W_c \in R^{N \times N}$ 是权重矩阵, 将输入 S'_t 映射到三个门和输入单元状态。 U_f, U_i, U_o 和 $U_c \in R^{N \times N}$ 是前一隐藏状态的权重矩阵。 $b_f, b_i, b_o, b_c \in R^N$ 是四个偏置向量。 f_t, i_t, o_t 分别是 LSTM 的遗忘门、输入门和输出门。 $\tilde{C}_t, C_t, C_{t-1} \in R^{N \times N}$ 为 LSTM 的状态。通过长短期记忆网络我们可以有效提取交通流的时间关系。

2.2.3. 图注意力网络

图卷积网络被广泛应用于挖掘交通道路网络中节点之间的空间关系, 但是标准的图卷积在计算过程中使用相同的卷积核参数, 使得相邻节点的信息均匀地聚集到给定的中心节点。然而, 这往往与事实情况不符合。道路节点之间的影响情况非常复杂, 交通流量往往会受早晚高峰、天气、车道数、路况、车辆密度和交通事故等因素的影响。这导致了各个邻居节点对中心节点的影响是各不相同的。图注意力网络(GAT) [10] 提出了一种利用注意力机制对相邻节点特征进行加权求和的方法。特征节点的权重完全依赖于节点的特征, 并且与图的结构无关。该方法克服了基于谱图卷积网络的瓶颈, 易于实现向不同邻域分配不同的学习权值。在这里我们采用多头注意机制来增强模型的能力, 稳定训练过程。定义如下:

$$\beta_{i,j} = \text{LeakyRelu}(\alpha [Wh_{t,i} \parallel Wh_{t,j}]) \quad (10)$$

$$e_{i,j} = \frac{\exp(\beta_{i,j})}{\sum_{v_k \in \tilde{N}(v_i)} \exp(\beta_{i,j})} \quad (11)$$

$$x'_i = \sigma \left(\frac{1}{K} \sum_{K=1}^K \sum_{v_k \in \tilde{N}(v_i)} e_{i,j}^{(k)} W^{(k)} h_{t,j} \right) \quad (12)$$

其中 W 是节点特征变换的权重参数, $\alpha(\cdot)$ 是计算两个节点相关度的函数, $h_{t,i}$ 和 $h_{t,j}$ 是 LSTM 的隐藏层输出, LeakyRelu 为激活函数, $\beta_{i,j}$ 是两个节点之间的注意力得分, $e_{i,j}^{(k)}$ 是第 k 组注意力机制计算出的权重系数, $W^{(k)}$ 是可学习的参数。我们建议对 K 个输出取平均得到最终的结果 x'_i 。

2.2.4. 时间注意力网络

除了捕捉动态的空间关系, 在时间维度上, 不同时间段的交通状况之间存在相关性, 并且在不同的情况下这种相关性也是不同的。因此我们还需要时间注意力机制[11]来关注交通流量长期的动态时间关系。定义如下:

$$E = V_e \cdot \sigma\left(\left((x'_{i,t})^T U_1\right) U_2 (x'_{i,t} U_3) + b_e\right) \quad (13)$$

$$T_{i,j}^{(k)} = \frac{\exp(E_{i,j}^{(k)})}{\sum_{j=1}^N \exp(E_{i,j}^{(k)})} \quad (14)$$

其中 V_e, b_e, W_1, W_2, W_3 都是可学习的参数。 $E_{i,j}^{(k)}$ 是第 k 组注意力机制计算出的时间 i 和时间 j 之间的关联强度。同样地使用 Softmax 函数来确保时间节点的关注权重系数和为 1。通过时间注意力机制我们可以提取到交通流的动态时间相关性。

2.2.5. 多组件融合方法

图 3 给出了对原始数据周期性特征的提取。输入数据的特征与最近、每日和每周观测的交通数据密切相关。如图所示, 我们沿着时间轴截取长度为 T_r 、 T_d 、和 T_w (T 为采样的时间长度, n 为正整数) 的三个时间序列段, 分别作为最近周期数据 X_R , 日周期数据 X_D , 即过去几天同一预测时刻的采样数据, 周周期数据 X_W , 即过去几周同一预测时刻的采样数据。因此, 我们采用了一种基于参数矩阵的方法来融合 STGAL 中的三个组件(最近, 每日和每周交通流模式), 如下所示:

$$Y' = \sigma(W_R \otimes Y'_R + W_D \otimes Y'_D + W_W \otimes Y'_W) \quad (15)$$

其中, W_R, W_D, W_W 是学习参数, 表示其对预测目标的影响程度; \otimes 是矩阵对应元素相乘的哈达马乘积; Y'_R, Y'_D, Y'_W 分别是近期, 日周期, 周周期 3 个组件的输出。

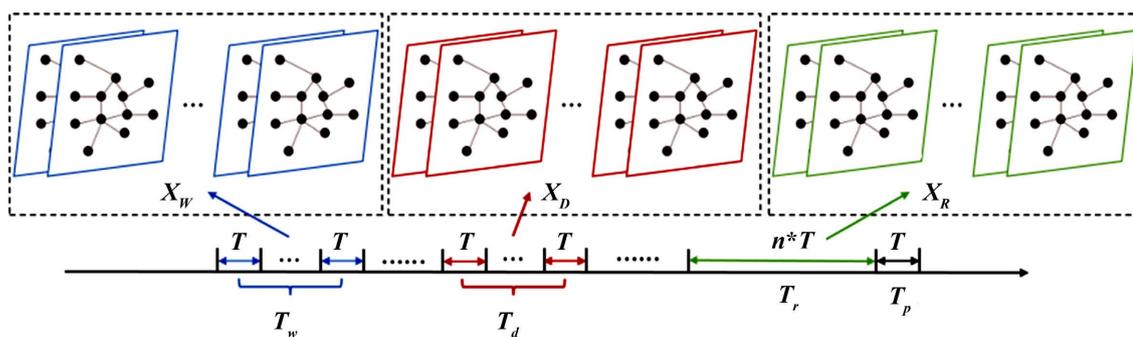


Figure 3. Example of the input time series construction

图 3. 输入时间序列片段构建示例

3. 实验

3.1. 实验数据集

我们在两个真实交通数据集[8] PeMSD4 和 PeMSD8 上进行实验。这些数据来自 Caltrans Performance

Measure System。每 30 秒实时收集一次，原始交通流量数据被聚合成 5 分钟间隔，这意味着每小时的流量数据中有 12 个点。我们使用过去一小时的交通流量数据来预测下一小时的流量情况。数据集中的缺失数值通过线性插值来填充。通过标准归一化方法将数据集标准化，使训练过程更加稳定。关于这两个数据集的详细统计信息如表 1 所示。

Table 1. Description of experimental datasets

表 1. 实验数据集描述

数据集	传感器数量	时间范围	数据量
PeMSD4	307	1/1/2018~2/28/2018	16,992
PeMSD8	170	7/1/2016~8/31/2016	17,856

3.2. 评价指标

为了评估交通预测模型的预测性能，我们主要测量预测值与真实值之间的误差。误差越小，预测值越接近真实值。由于不同区域的交通情况可能不同，使用绝对误差可能表明模型对相对简单的样本进行过拟合，而平方误差对困难和不可预测的样本惩罚更多，可以更好地表现复杂情况下的性能。因此，我们采用了三个广泛使用的度量：平均绝对误差(MAE)、均方根误差(RMSE)和平均绝对百分比误差(MAPE)来衡量预测模型的性能，它们数值越小，预测效果越好。

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |Y - Y'| \quad (16)$$

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (Y - Y')^2} \quad (17)$$

$$\text{MAPE} = \frac{100\%}{N} \sum_{i=1}^N \left| \frac{Y - Y'}{Y} \right| \quad (18)$$

3.3. 实验设置

为了评估我们模型的整体性能，我们将 STGAL 与广泛使用的基线和最先进的模型进行了比较，包括 1) HA [2]: 历史平均法。我们使用最后 12 个时间片的平均值来预测下一个值。2) VAR [3]: 向量自回归是一种先进的时间序列模型，它能够捕捉时间序列之间的成对关系。3) LSTM [4]: 用于时间序列预测的长期短期记忆网络，是循环神经网络改进的变体。4) DCRNN [5]: 扩散卷积递归神经网络利用扩散图卷积网络和序列到序列结构分别对空间信息和时间信息进行编码。5) STGCN [6]: 完全使用卷积结构来捕获空间和时间特征，其时空卷积模块结构类似三明治，上下是时间卷积，中间是空间卷积。6) Graph Wave-Net [7]: 它引入了自适应的图建模方法，充分考虑了现实世界中的复杂关系及不确定性，结合了图形卷积和扩张的随意卷积来捕获时空依赖关系。7) ASTGCN [8]: 基于注意力的时空图卷积网络使用三个组件对时间序列上的近期、日周期、周周期三个模式进行建模。在 K 阶切比雪夫多项式展开的图卷积内融入了注意力机制来捕获空间动态性，在时间维度上使用注意力机制捕获动态时间关系。

所有基于深度学习的模型，包括我们的 STGAL，都是用 Python3.8 和 Pytorch 1.6.0 实现的。并在带有一个 NVIDIA GTX 2080Ti 的 GPU 服务器上执行。我们使用 ADAM 优化器对所有模型进行优化，对于超参数，我们将所有模块的隐藏单位设置为 64，批量大小设置为 16。我们选择均方误差 MSE 作为模型的损失函数，最初的学习速率设置为 0.0005，学习速率的衰减效率设置为 0.92。所有深度学习模型的最佳参数都是通过对验证集进行仔细的参数调整来选择的。

3.4. 实验结果及分析

将我们的模型与 PeMSD4 和 PeMSD8 上的上述七种有代表性的基线方法进行了比较。表 2 显示了未来一小时即 12 个预测步长的交通流量预测性能的平均结果。我们可以观察到：1) 基于 GCN 的方法优于 HA、VAR 传统的机器学习方法和 LSTM 深度学习方法，说明了显式建模空间相关性的重要性和 GCN 在交通流预测任务中的有效性；2) 我们的方法进一步改进了基于 GCN 的方法，并取得了显著的效果。STGAL 对 PeMSD4 和 PeMSD8 数据集的 MAE 和 MAPE 的现有最好结果进行了 5% 以上的相对提升，以及对 PeMSD4 和 PeMSD8 数据集的 RMSE 的现有最好结果进行了 3% 以上的相对提升，特别是对 PeMSD4 数据集的 MAE 的现有最好结果有着高达 8.76% 的相对提升。3) 实验结果表明，STGAL 能够更加准确捕捉交通流量序列中的时间和空间相关性，并取得最好的预测效果。

Table 2. Performance of STGAL and baseline models for PeMSD4 and PeMSD8

表 2. STGAL 以及基准方法在数据集 PeMSD4 和 PeMSD8 上的性能表现

Model	PeMSD4			PeMSD8		
	MAE	RMSE	MAPE	MAE	RMSE	MAPE
HA	38.03	59.24	27.88%	34.86	52.04	24.07%
VAR	23.75	36.66	18.09%	23.46	36.33	15.42%
LSTM	26.94	41.38	17.80%	22.02	33.74	13.61%
DCRNN	24.48	37.86	16.75%	17.83	27.78	12.42%
STGCN	22.06	34.80	16.20%	17.88	27.63	12.30%
Graph Wave-Net	22.13	33.94	16.05%	17.49	26.15	12.21%
ASTGCN	21.80	32.82	15.80%	16.63	25.27	12.18%
STGAL (ours)	19.89	31.71	14.84%	15.77	24.42	11.54%
Improvements	+8.76%	+3.38%	+6.07%	+5.17%	+3.36%	+5.25%

3.5. 消融研究

为了进一步研究 STGAL 不同模块的影响，我们进行了全面的消融研究，我们设计了 STGAL 模型的四个变体。具体来说 1) N-SA：去掉 STGAL 中的空间注意力模块；2) N-TA：去掉 STGAL 中的时间注意力模块；3) N-LSTM：去掉 STGAL 中的长短期记忆网络；4) N-GAT：这里是用标准的图卷积网络来代替我们 STGAL 中的多头图注意力网络。

我们将这四个变体和 STGAL 在 PeMSD4 和 PeMSD8 上进行实验，如图 4、图 5 所示。我们可以观察到：1) N-TA 的各个评价指标都处于最大值，即时间注意力模块对两个数据集的影响最大，说明了用时间注意力来提取交通流长期的动态的时间相关性是非常有必要的。2) 去掉空间注意力模块，STGAL 的实验效果明显降低，我们推测其原因是空间注意力提取了交通流的动态空间相关性，符合了现实中交通流动态变化的事实。3) 用标准的图卷积网络代替多头图注意力网络，模型的性能在两个数据集上都表现很差，是因为标准的图卷积网络在卷积计算的过程中共享卷积核参数，使得邻居节点的信息被均匀地聚合到中心节点上，而 GAT 网络可以对邻域节点进行有区别的信息聚合，它不仅考虑邻域节点的特征和边的权重来生成新的节点表示，而且考虑交通流动态变化的时空模式。4) 与四个变体相比较，STGAL 具有最佳的性能。一方面表明了 STGAL 中各个模块的重要性，另一方面说明了 STGAL 能够更加准确提取交通流量序列中的时空相关性。

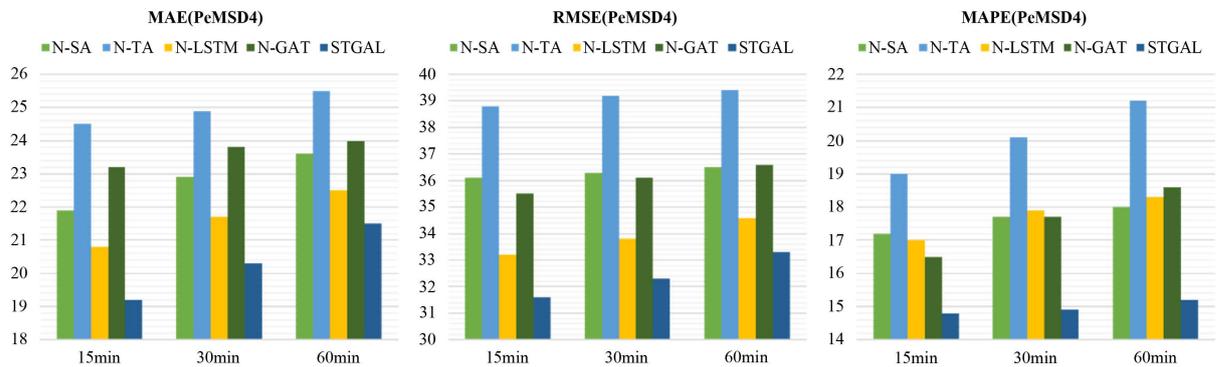


Figure 4. Ablation study on PeMSD4

图 4. 在 PeMSD4 上的消融研究

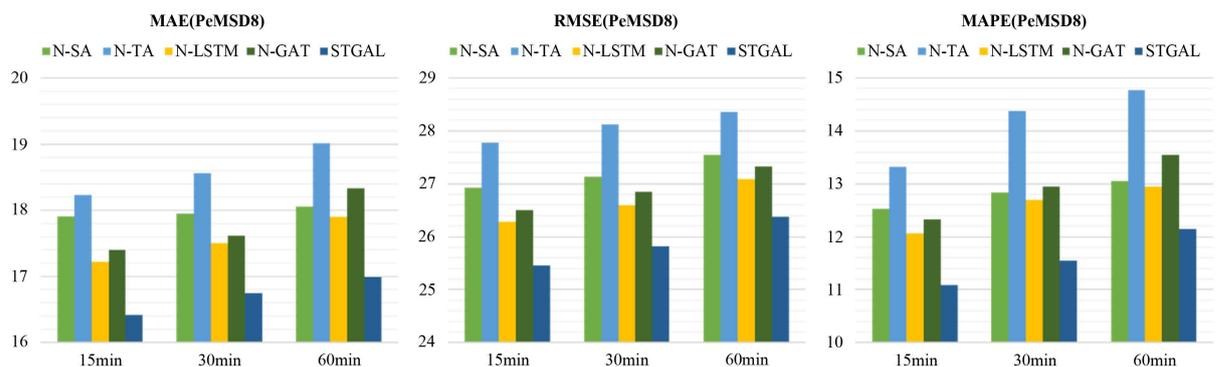


Figure 5. Ablation study on PeMSD8

图 5. 在 PeMSD8 上的消融研究

4. 结语

本文提出了一种新颖的面向交通流量预测的时空图注意力网络 STGAL。我们使用图注意力网络和短期记忆网络来提取城市道路中交通流量的静态的时空关系，而且我们使用空间和时间注意力网络来捕捉交通路网节点中动态变化的时空关系。在多步交通预测任务上的大量实验结果证明了 STGAL 以及它的各个模块的有效性和优越性。在下一步的工作中，我们将会研究所提出的模型是否可以应用于解决其他的时空数据挖掘问题，例如时空车辆轨迹数据的挖掘任务。

基金项目

云南省软件工程重点实验室开放基金资助项目(2020SE404)。

参考文献

- [1] 赵娜, 袁家斌, 徐晗. 智能交通系统综述[J]. 计算机科学, 2014, 41(11): 7-11+45.
- [2] 杨春霞, 秦家鹏, 王庆, 李欣栩. 基于多车道加权融合的短时交通流预测研究[J]. 公路交通科技, 2021, 38(1): 121-127.
- [3] Lippi, M., Bertini, M. and Frasconi, P. (2013) Short-Term Traffic Flow Forecasting: An Experimental Comparison of Time-Series Analysis and Supervised Learning. *IEEE Transactions on Intelligent Transportation Systems*, **14**, 871-882. <https://doi.org/10.1109/TITS.2013.2247040>
- [4] 李磊, 张青苗, 赵军辉, 聂逸文. 基于改进 CNN-LSTM 组合模型的分时段短时交通流预测[J]. 应用科学学报, 2021, 39(2): 185-198.
- [5] Li, Y., Yu, R., Shahabi, C. and Liu, Y. (2018) Diffusion Convolutional Recurrent Neural Network: Data-Driven Traffic

-
- Forecasting. *International Conference on Learning Representations*, Vancouver, 30 April-3 May 2018, 1-10. <https://openreview.net/forum?id=SJiHXGWAZ>
- [6] Yu, B., Yin, H. and Zhu, Z. (2018) Spatial-Temporal Graph Convolutional Networks: A Deep Learning Framework for Traffic Forecasting. *Proceedings of the 27th International Joint Conference on Artificial Intelligence*, Stockholm, 13-19 July 2018, 3634-3640. <https://doi.org/10.24963/ijcai.2018/505>
- [7] Wu, Z., Pan, S., Long, G., Jiang, J. and Zhang, C. (2019) Graph Wave Net for Deep Spatial-Temporal Graph Modeling. *Proceedings of the 28th International Joint Conference on Artificial Intelligence*, Macao, 10-16 August 2019, 1907-1913. <https://doi.org/10.24963/ijcai.2019/264>
- [8] Guo, S., Lin, Y., Feng, N., Song, C. and Wan, H. (2019) Attention Based Spatial-Temporal Graph Convolutional Networks for Traffic Flow Forecasting. *Proceedings of the AAAI Conference on Artificial Intelligence*, **33**, 922-929. <https://doi.org/10.1609/aaai.v33i01.3301922>
- [9] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L. and Polosukhin, I. (2017) Attention Is All You Need. *Proceedings of the 31st International Conference on Neural Information Processing Systems*, Long Beach, 4-9 December 2017, 5998-6008.
- [10] Veličković, P., Cucurull, G., Casanova, A., Romero, A., Lio, P. and Bengio, Y. (2018) Graph Attention Networks. *International Conference on Learning Representations*, Vancouver, 30 April-3 May 2018, 1-12. <https://openreview.net/forum?id=rJXMpikCZ>
- [11] Feng, X., Guo, J., Qin, B., Liu, T. and Liu, Y. (2017) Effective Deep Memory Networks for Distant Supervised Relation Extraction. *Proceedings of the 26th International Joint Conference on Artificial Intelligence*, Melbourne, 19-25 August 2017, 4002-4008. <https://doi.org/10.24963/ijcai.2017/559>