

doi: 10.7690/bgzd.2025.05.015

动静图融合和时序流注意力网络用于交通流预测

闫 敬, 王 祥, 郑 铮

(唐山师范学院计算机科学技术系, 河北 唐山 063000)

摘要: 为准确预测交通流量有利于优化交通管理、提高交通效率的问题, 提出一种新的动静态图融合和时序流注意力网络。通过图卷积网络捕捉动态和静态的空间相关性, 引入流注意力机制, 有效缓解二次复杂度问题; 设计时间相关性建模 (temporal correlation modeling, TCM) 模块替换流注意力机制的线性变换方法, 以增强模型的时序建模能力。在 4 个真实世界的交通数据集上进行了大量实验。实验结果表明: 所提出的模型具有优越的性能, 并且明显优于基线。

关键词: 交通流预测; 时空相关性; 流注意力机制; 图卷积网络; 特征融合

中图分类号: TP391.4 **文献标志码:** A

Dynamic-static Graph Fusion and Temporal Flow Attention Network for Traffic Flow Prediction

Yan Jing, Wang Xiang, Zheng Zheng

(Department of Computer Science and Technology, Tangshan Normal University, Tangshan 063000, China)

Abstract: Accurately predicting traffic flow is conducive to optimizing traffic management and improving traffic efficiency, a new dynamic-static graph fusion and temporal flow attention network is proposed. The graph convolutional network is used to capture dynamic and static spatial correlations. A flow attention mechanism is introduced to effectively alleviate the quadratic complexity problem. A temporal correlation modeling (TCM) module is designed to replace the linear transformation method of the flow attention mechanism, so as to enhance the model's temporal modeling ability. A large number of experiments are carried out on four real-world traffic datasets. The results show that the proposed model has superior performance and significantly outperforms the baselines.

Keywords: traffic flow prediction; spatio-temporal correlation; flow attention mechanism; graph convolutional network; feature fusion

0 引言

智能交通系统(intelligent traffic system, ITS)^[1]在道路交通管理中发挥着重要作用。交通预测是 ITS 的核心功能之一, 备受研究者关注。如今传感器技术成熟, 交通数据采集和存储更便利, 交通流量、速度和占有率等数据可用于交通预测研究。因交通流量是道路状况最直观的指标, 所以交通流预测成为本次研究内容。

深度学习方法常用于交通流预测研究, 但存在计算速度慢, 捕捉长周期序列相关性时易出现梯度消失或爆炸等问题, 且只能捕捉时间相关性, 无法处理空间信息, 导致预测效果不佳。

注意力机制常用于交通流预测研究, 能根据输入序列不同位置特征, 自适应计算权重, 聚焦重要位置, 但为防止注意力退化引入特定归纳偏差, 在一定程度上限制了模型适用性与适应性; 处理长序列

数据时, 注意力权重计算量呈二次增长, 会引发二次复杂度问题。

针对传统方法弊端, 笔者提出动静态图融合及时序流注意力网络。该方法融合静、动态图特征, 精准建模时空相关性, 规避传统注意力机制的归纳偏差, 缓解二次复杂度问题。因单一注意力机制能力受限, 笔者采用多头结构, 并行构建投影空间, 大幅提升模型特征提取能力。根据空间拓扑结构生成静、动态空间矩阵并融合, 经图卷积网络捕捉多尺度空间相关性; 设计时间相关性建模模块, 结合 1 维卷积 (Conv1D) 与门控循环单元 (GRU) 提取局部和长期时间特征; 设计时序流注意力机制 (temporal flow attention, TFA), 以时间相关性建模 (TCM) 模块替换传统线性变换, 增强非线性建模能力; 实验结果表明所开发模型预测准确性超越基线模型。

收稿日期: 2024-08-02; 修回日期: 2024-09-20

基金项目: 唐山师范学院科研基金项目 (2022C53)

第一作者: 闫 敬 (1973—), 女, 河北人, 硕士。

1 方法

1.1 准备工作

1.1.1 交通网络

根据实际道路网络中各传感器之间的道路连接关系生成相应的道路网络拓扑图, 使用 $G=(V, E)$ 表示。其中, $V=\{V_1, V_2, \dots, V_N\}$ 表示整个道路网络中所有节点(道路传感器)的集合, N 表示节点数量; E 表示节点之间连通性的一组边。假如有 2 个节点之间通过道路连接, 那么 G 中就存在一条连接这 2 个节点的边。

1.1.2 时空相关性

交通流预测所使用的数据集属于时空数据, 如何从流量数据中捕捉时空相关性, 是交通流预测面临的重大挑战^[2]。时间相关性体现为不同时刻交通状况各异, 空间相关性指相互连通的不同道路间, 交通状态会相互影响, 上游道路影响下游道路, 下游道路也会因反馈效应^[3]反作用于上游道路。如图 1 所示, A、B、C 均为道路节点, 由于存在时空相关性, 时刻的节点 A 会对与其空间相连的节点 B 和 C 产生影响, 还会影响 $t+1$ 时刻的节点 A 自身。

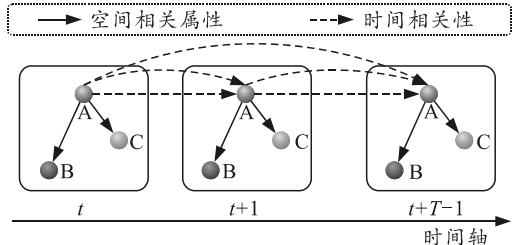


图 1 时空相关性

1.1.3 邻接矩阵

根据道路网络拓扑图 G 生成对应的邻接矩阵用 $A \in R^{N \times N}$ 表示。 a_{ij} 是 A 中的一个元素, 表示节点 V_i 与节点 V_j 之间的空间连接状态。如果存在 $V_i, V_j \in V$ 且 $(V_i, V_j) \in E$, 则 a_{ij} 的值为 1, 否则值为 0。

1.1.4 时间特征的构建

拓扑图 G 中所有节点的时间特征用 $X \in R^{N \times L \times F}$ 表示。其中, L 表示每个节点的时间序列总长度; F 表示节点的特征的类型数量, 节点的特征数据可以是交通流量、速度和道路占用率等, 笔者只考虑交通流量。交通流量预测的过程可用式(1)描述:

$$(X_{t+1}, X_{t+2}, \dots, X_{t+T'}) = \mathcal{F}[(X_{t-T+1}, X_{t-T+2}, \dots, X_t); G] \quad (1)$$

式中: T 为历史流量数据的长度; T' 为要预测的未来流量数据的长度。给定连续的时间步长 $(X_{t-T+1}, X_{t-T+2}, \dots, X_t)$ 和拓扑图 G , 通过训练一个模型 G 预测未来的 T' 个连续的时间步长 $(X_{t+1}, X_{t+2}, \dots, X_{t+T'})$ 。

1.2 动静图融合和时序流注意力网络(DSGF-TFA)

DSGF-TFA 模型结构如图 2 所示, 主要由输入层(multi-layer perceptron, MLP1)、图卷积网络(graph convolutional network, GCN)、聚合层(aggregation layer, AGG)、TFA 模块和输出层(multi-layer perceptron, MLP2)组成。

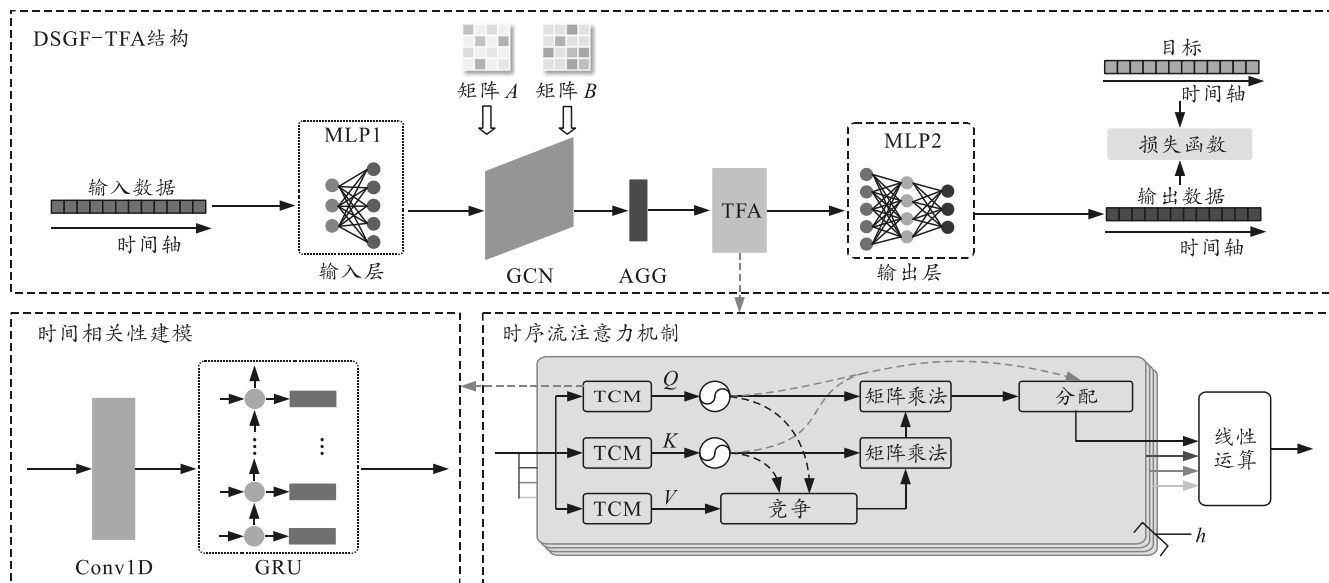


图 2 DSGF-TFA 详细结构

2 实验

2.1 数据集

在 PeMS03、PeMS04、PeMS07 和 PeMS08 4 个公共交通数据集^[4]上评估 DSGF-TFA 模型性能。数据集由 Caltrans 性能测量系统在加州高速公路采集, 每 30 s 实时采集一次, 编译后聚合为 5 min 间隔数据, 每台传感器一天数据有 288 个时间步长。

PeMS 系列原始数据含交通流量、速度和占有率 3 种指标, 笔者仅用交通流量数据做交通预测研究。

2.2 基线方法

将 DSGF-TFA 与 LSTM^[5]、DCRNN^[6]、STGCN^[7]、ASTGCN(r)^[8]、STG2Seq^[2]、Graph WaveNet^[9]、STSGCN^[4]、STGODE^[10]、STDSGNN^[11]、STGPCN(Kronecker)^[12]、LEISN-ED^[13]、STSGRU^[14]等基线方法进行比较。

2.3 实验设置

公平起见, 与其他基线保持相同的数据划分方法, 将所有数据集以 6:2:2 的比率划分为训练集、验证集和测试集。使用 1 h 的历史数据来预测未来 1 h 的交通流量, 即使用过去 12 个连续时间步

长预测未来 12 个连续时间步长。

实验均在具有 NVIDIA GeForce RTX3090 GPU 服务器上运行。切比雪夫多项式的项数 $K=3$, 注意头数为 4, 节点嵌入维度为 10, GRU 隐藏单元数量为 64。使用 Adam 优化器训练模型, 学习率为 0.001, 批大小为 16。

使用平均绝对误差(MAE)、平均绝对百分比误差(MAPE)和均方根误差(RMSE)评估模型的性能, 如式(2)~(4)所示:

$$\text{MAE} = \frac{\sum_{t=1}^T \sum_{n=1}^N |Y_t^n - \hat{Y}_t^n|}{(N \times T)}; \quad (2)$$

$$\text{MAPE} = \frac{\sum_{t=1}^T \sum_{n=1}^N \frac{|Y_t^n - \hat{Y}_t^n|}{Y_t^n}}{(N \times T)}; \quad (3)$$

$$\text{RMSE} = \sqrt{\frac{\sum_{t=1}^T \sum_{n=1}^N (Y_t^n - \hat{Y}_t^n)^2}{(N \times T)}}. \quad (4)$$

2.4 实验结果和分析

2.4.1 在 PeMS 系列数据集上的结果

表 1 展示 DSGF-TFA 与基线模型的交通流预测结果, 粗体为最佳指标。DSGF-TFA 在 4 个真实数据集上预测效果显著优于基线模型。LSTM 仅考虑流量数据时间维度依赖, 其他基线模型与 DSGF-TFA 兼顾时间与空间相关性, 故效果优于 LSTM。

表 1 DSGF-TFA 与基线模型之间交通流量预测性能的比较

%

Model	PeMS03			PeMS04			PeMS07			PeMS08		
	MAE	MAPE	RMSE	MAE	MAPE	RMSE	MAE	MAPE	RMSE	MAE	MAPE	RMSE
LSTM	21.33	23.33	35.11	27.14	18.20	41.59	29.98	13.20	45.84	22.20	14.20	34.06
DCRNN	18.18	18.91	30.31	24.70	17.12	38.12	25.30	11.66	38.58	17.86	11.45	27.83
STGCN	17.49	17.15	30.12	22.70	14.59	35.55	25.38	11.08	38.78	18.02	11.40	27.83
ASTGCN(r)	17.69	19.40	29.66	22.93	16.56	35.22	28.05	13.92	42.57	18.61	13.08	28.16
STG2Seq	19.03	21.55	29.73	25.20	18.77	38.48	32.77	20.16	47.16	20.17	17.32	30.71
Graph WaveNet	19.85	19.31	32.94	25.45	17.29	39.70	26.85	12.12	42.78	19.13	12.68	31.05
STSGCN	17.48	16.78	29.21	21.19	13.90	33.65	24.26	10.21	39.03	17.13	10.96	26.80
STGODE	16.50	16.69	27.84	20.84	13.77	32.82	22.99	10.14	37.54	16.81	10.62	25.97
STDSGNN	16.12	16.15	25.59	20.67	13.83	32.40	22.91	10.06	34.95	16.73	10.84	25.59
STGPCN(Kronecker)	17.11	16.48	28.99	20.96	13.78	33.35	24.02	10.08	38.77	16.41	10.43	25.60
LEISN-ED	15.83	14.66	26.05	-	-	-	-	-	-	15.94	10.18	24.96
STSGRU	15.45	15.85	24.13	20.11	13.86	31.80	21.50	9.08	34.40	15.68	10.67	25.12
DSGF-TFA(our)	15.30	13.29	22.69	19.50	12.46	29.29	21.14	8.23	33.34	16.00	9.16	23.94

STSGCN 借局部时空子图挖掘时空数据相互影响, 效果优于旧基线方法。STGODE 结合张量常微分方程, 在全局依赖上比 STSGCN 更具优势。最新 4 个基线方法中, STSGRU 通过共享权重与周期性建模, 预测精度最高; DSGF-TFA 在各数据集表现更优, 如在 PeMS03 上, MAPE 和 RMSE 分别提升 16.15% 和 5.97%; 在 PeMS07 和 PeMS08 上, MAPE 分别提升 9.36% 和 14.15%。因其能自适应生成动态空间矩阵, 还引入流注意力机制及新时间相关性建模模块, 增强了时空建模与预测能力。

2.4.2 可视化

如图 3 所示, 绘制 DSGF-TFA 和 STGODE 在 4 个数据集上的散点图, 横坐标和纵坐标分别表示预测值和真实值。可以看出: 所提出的 DSGF-TFA 模型的散点聚集程度明显比 STGODE 更高, 说明 DSGF-TFA 的预测性能优于 STGODE。

图 4 展示 DSGF-TFA 模型在 PeMS03 数据集上预测值与真实值的绝对误差热力图。因数据集规模大, 仅选取 12 条道路的前 60 个时间步长, 并绘制 5 和 60 min 预测层位的热力图。交通状况复杂多变,

一般预测层位增大, 模型预测性能会下降。但从热力图4可见, DSGF-TFA 模型在 5 min 短时与 60 min

长时预测效果均佳, 表明其在捕捉长期序列时空相关性上优势明显。

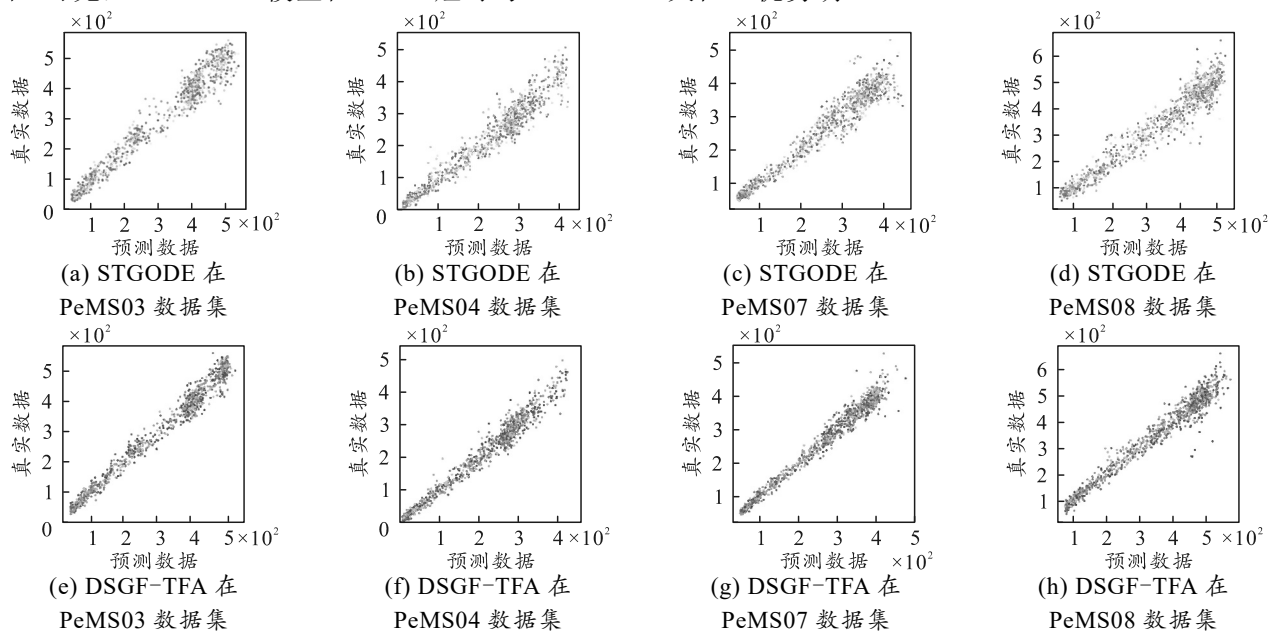


图3 STGODE 和 DSGF-TFA 散点图

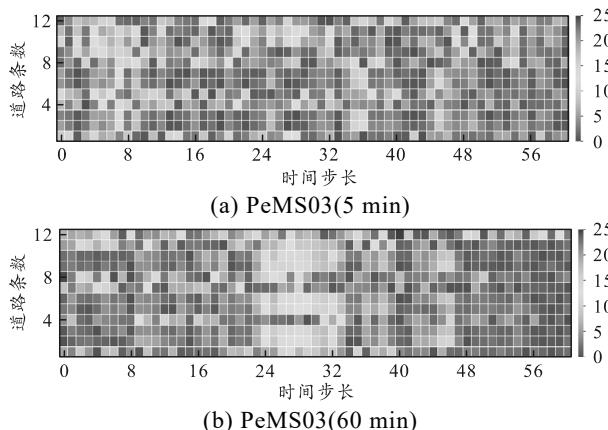
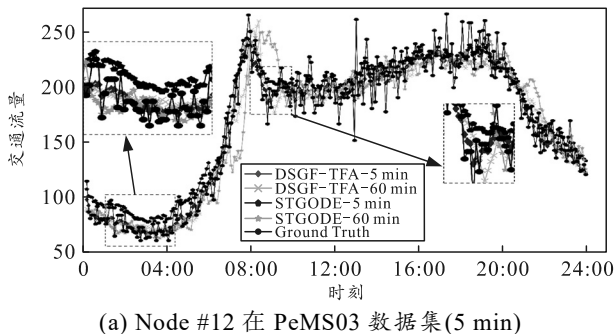
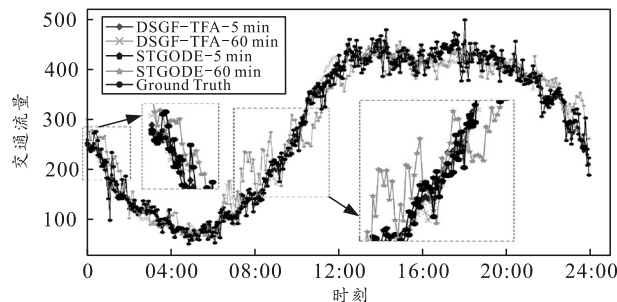


图4 热力图

如图5所示, 在 PeMS03 和 PeMS07 测试集截取一天时段, 绘制 DSGF-TFA 和 STGODE 在 5 min、60 min 的预测曲线, 并与真实流量对比。流量升降时, DSGF-TFA 比 STGODE 对变化预测更敏感、迅速, 数值更准。DSGF-TFA 在 5 min 短时及 60 min 长时预测中, 效果均优于 STGODE。



(a) Node #12 在 PeMS03 数据集(5 min)



(b) Node #112 在 PeMS07 数据集(5 min)

图5 DSGF-TFA 和 STGODE 预测

2.5 消融实验

为验证 DSGF-TFA 各模块有效性, 设计其 4 个变体模型, 并在最佳超参数下进行实验。将这 4 个变体与 DSGF-TFA 在 PeMS04 和 PeMS08 数据集上的预测结果作对比分析。

1) REPL-Att: 将 TFA 模块替换为传统的注意力机制, 从而验证 TFA 模块的有效性。

2) REPL-FAAtt: 将 TFA 模块中的 TCM 模块替换为流注意力机制中传统的线性变换方法, 以证明 TCM 模块建模时间相关性的能力。

3) Only-Dyn: 去除静态空间图, 仅保留自适应动态图, 从而证明完全忽略现实中的静态空间连接关系是否合理。

4) Only-Sta: 去除动态图, 仅留静态图, 验证融合静、动态图空间特征信息的必要性。消融实验显示, DSGF-TFA 各关键模块对交通流预测有积极

作用。REPL-Att 因传统注意力机制的特定归纳偏差效果最差；REPL-FAtt 受传统线性变换局限表现不佳；Only-Dyn 和 Only-Sta 片面，融合动静空间相关性能有效提升模型捕捉时空相关性的能力。

2.6 超参数的影响

在同一模型中设置不同超参数，会影响其预测效果。从图 6 和 7 在 PeMS04 数据集上的实验可知：当单独改变注意力头数、节点嵌入维度数等超参数（其他超参数设为最佳）时，超参数设置过小，模型建模时空相关性能力不足；适当增大超参数能提升预测效果，但增加过多会因过拟合致预测性能下降。只有将超参数控制在合适范围，模型性能才达最优。

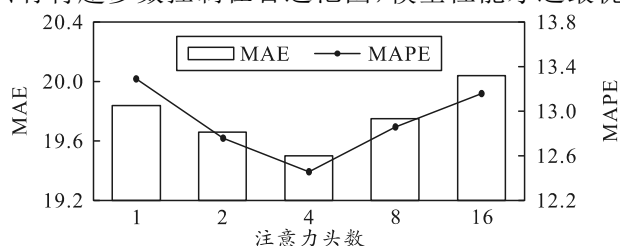


图 6 不同注意力头数的影响

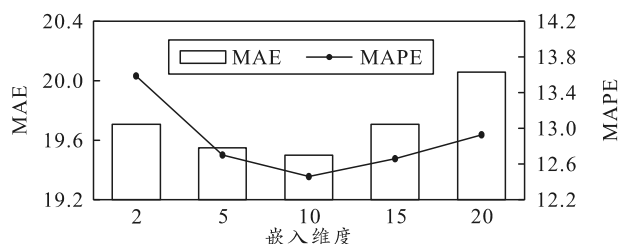


图 7 不同节点嵌入维度数的影响

3 结论

引入流注意力机制，使用源竞争机制和汇分配机制替换传统的注意力权重计算模块，在捕捉关键特征信息的同时降低了模型的复杂度，缓解了传统注意力机制面临的二次复杂度问题。设计时间相关性建模模块取代全连接层，实现非线性特征映射。兼顾静态和动态空间特征信息，融合多尺度特征提升建模能力。经 4 个数据集实验，证明了 DSGF-TFA 的预测性能和泛化能力。模型虽有效果，但存在不足，如因道路差异，共享参数难捕捉特性。未来拟为节点分配独立参数捕捉多尺度交通模式，融合天气、温度等外部数据，提升建模能力和预测精度。

参考文献：

[1] ZHANG J, WANG F Y, WANG K, et al. Data-driven

intelligent transportation systems: A survey[J]. IEEE Transactions on Intelligent Transportation Systems, 2011, 12(4): 1624-1639.

- [2] BAI L, YAO L, KANHERE S S, et al. STG2Seq: Spatial-temporal graph to sequence model for multi-step passenger demand forecasting[C]. IJCAI: Proceedings of the 28th International Joint Conference on Artificial Intelligence, 2019: 1981-1987.
- [3] DONG C J, SHAO C F, ZHUGE C X, et al. Spatial and temporal characteristics for congested traffic on urban expressway[J]. Bijing Gongye Daxue Xuebao, 2012, 38(8): 1242-1246, 1268.
- [4] 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[C]. In Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34: 914-921.
- [5] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural computation, 1997, 9(8): 1735-1780.
- [6] LI Y G, YU R, SHAHABI C, et al. Diffusion convolutional recurrent neural network: data-driven traffic forecasting[C/OL]. arXiv: Learning (2017). <https://doi.org/10.48550/arXiv.1707.01926>.
- [7] LI C L, CUI Z, ZHENG W M, et al. Spatio-Temporal Graph Convolution for Skeleton Based Action Recognition[C]. In AAAI Conference on Artificial Intelligence, 2018: 3482-3489.
- [8] GUO S N, LIN Y F, FENG N, et al. Attention based spatial-temporal graph convolutional networks for traffic flow forecasting[C]. In Proceedings of the AAAI Conference on Artificial Intelligence, 2019: 922-929.
- [9] WU Z H, PAN S R, LONG G D, et al. Graph wavenet for deep spatial-temporal graph modeling[C]. IJCAI: Proceedings of the 28th International Joint Conference on Artificial Intelligence, 2019: 1907-1913.
- [10] FANG Z, LONG Q Q, SONG G J, et al. Spatial-temporal graph ODE networks for traffic flow forecasting[C]. the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining, 2021: 364-373.
- [11] ZHANG R, XIE F, SUN R, et al. Spatial-temporal dynamic semantic graph neural network[J]. Neural Computing and Applications, 2022, 34: 16655-16668.
- [12] TAN Z, ZHU Y F, LIU B. Learning spatial-temporal feature with graph product[J]. Signal Processing, 2023, 210: 109062.
- [13] LAI Q, CHEN P. LEISN: A long explicit-implicit spatio-temporal network for traffic flow forecasting[J]. Expert Systems with Applications, 2024, 245: 123139.
- [14] SUN X Y, CHEN F H, WANG Y C, et al. Short-term traffic flow prediction model based on a shared weight gate recurrent unit neural network[J]. Physica A: Statistical Mechanics and its Applications, 2023, 618: 128650.