



MST-GNN: Graph Neural Network with Multi-Granularity in Space and Time for Traffic Prediction

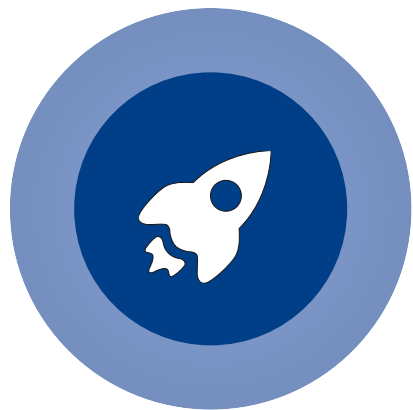
MST-GNN: 融合“空间多尺度”与“时序多粒度”信息的交通流量预测图神经网络

ICLR 2024

汇报人: 庞媛媛

2024/11/24

目 录



背景



相关工作



方法



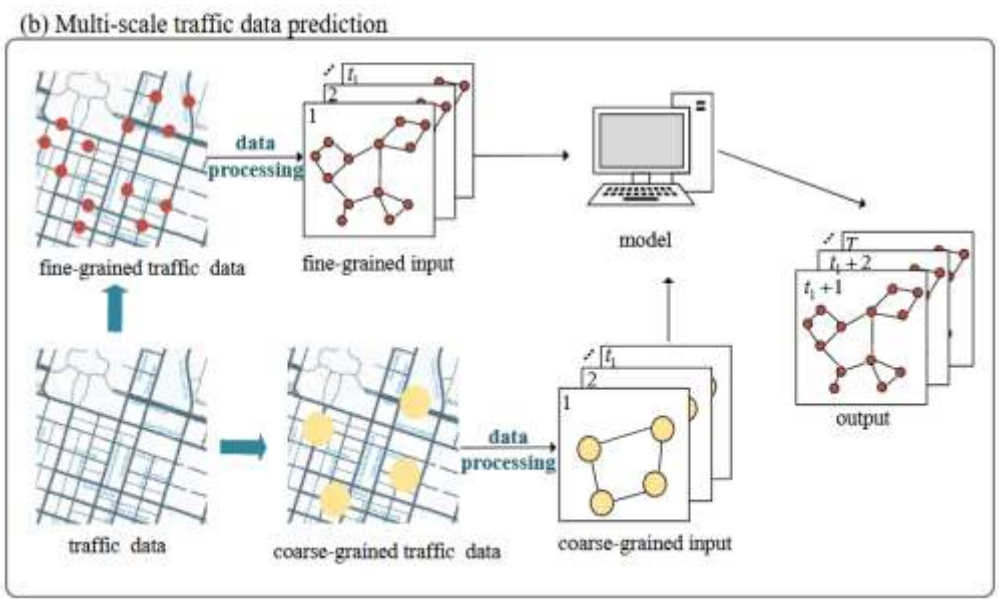
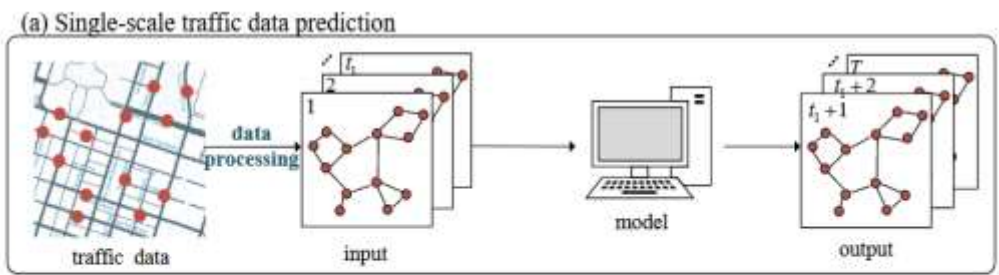
实验

2个局限性

■ 忽视跨空间尺度交通特征的整合

✓ 路段尺度（细粒度） ✗ 区域尺度（粗粒度）

必要性：帮助细粒度过程学习更宏观的交通数据，提高预测精度。

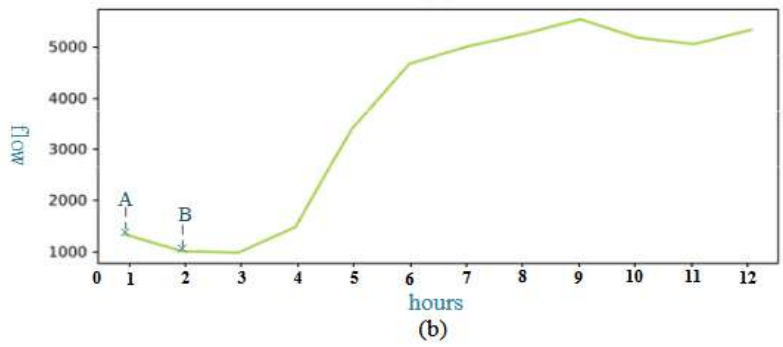
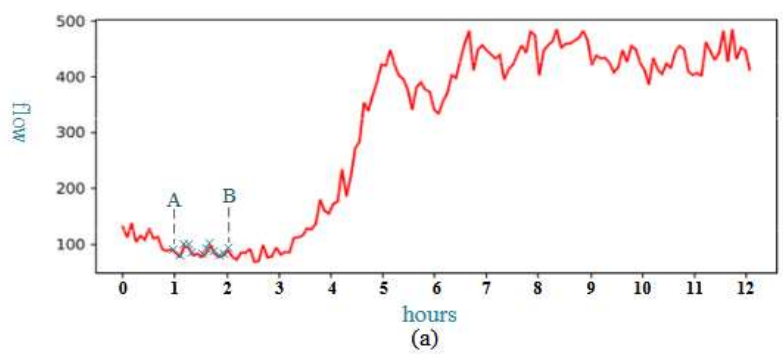


多空间尺度数据示意图

■ 忽视不同时间尺度交通特征的整合

✓ 细粒度 or 粗粒度 ✗ 细粒度 and 粗粒度

必要性：特征相互补偿，粗粒度序列下的迭代次数少



不同时间尺度上的流量变化比较

层次理论：较小尺度的过程支撑着较大尺度的过程，较大尺度的过程通过设定边界条件来制约较小尺度的过程。

“交通信息的跨尺度整合”

贡献

- 设计多尺度时空数据挖掘模块，同时捕捉多时间尺度和多空间尺度信息；
- 提出融合多时空尺度信息的时空注意力模块；
- 骨干简单，超越现有方法，实现SOTA性能；
- 首次将多尺度时空整合应用在交通预测。

交通预测

- **STGCN (2017)** ^[1]: 图卷积与卷积卷积序列学习层结合, 对道路网络的时空依赖性进行建模;
- **T-GCN (2019)** ^[2]: 将GCN与GRU结合;
- **SFTGNN (2021)** ^[3]: 引入动态时间扭曲 (DTW) 方法生成时间图, 弥补传统模型中空间图的不足;
- **STG-NCDE (2022)** ^[4]: 基于神经控制微分方程 (NCDE) 有效实现时空特征挖掘。

上述方法仅关注交通数据的单尺度时空特征

[1] Bing Yu, Haoteng Yin, and Zhanxing Zhu. Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting. arXiv preprint arXiv:1709.04875, 2017.

[2] Ling Zhao, Yujiao Song, Chao Zhang, Yu Liu, Pu Wang, Tao Lin, Min Deng, and Haifeng Li. T-gcn: A temporal graph convolutional network for traffic prediction. IEEE transactions on intelligent transportation systems, 21(9):3848–3858, 2019.

[3] Mengzhang Li and Zhanxing Zhu. Spatial-temporal fusion graph neural networks for traffic flow forecasting. In Proceedings of the AAAI conference on artificial intelligence, volume 35, pp. 4189–4196, 2021.

[4] Jeongwhan Choi, Hwangyong Choi, Jeehyun Hwang, and Noseong Park. Graph neural controlled differential equations for traffic forecasting. In Proceedings of the AAAI Conference on Artificial Intelligence, volume 36, pp. 6367–6374, 2022.

时空注意力机制

- **ASTGCN (2019)** ^[1]: 基于注意力的时空图卷积网络，包含一个时间注意力层和一个空间注意力层，以捕捉交通数据之间的时空相关性；
- **GMAN (2020)** ^[2]: 图多注意力网络，采用了编码和解码的架构。在编码器和解码器之间加入了注意力转换模块，模拟历史时间步和未来时间步之间的关系，有助于缓解预测时间步之间的误差传播问题。

空间和时间两个角度有效学习长程依赖关系仍是挑战

[1] Shengnan Guo, Youfang Lin, Ning Feng, Chao Song, and Huaiyu Wan. Attention based spatialtemporal graph convolutional networks for traffic flow forecasting. In Proceedings of the AAAI conference on artificial intelligence, volume 33, pp. 922–929, 2019.

[2] Chuanpan Zheng, Xiaoliang Fan, Cheng Wang, and Jianzhong Qi. Gman: A graph multi-attention network for traffic prediction. In Proceedings of the AAAI conference on artificial intelligence, volume 34, pp. 1234–1241, 2020a.

社区检测

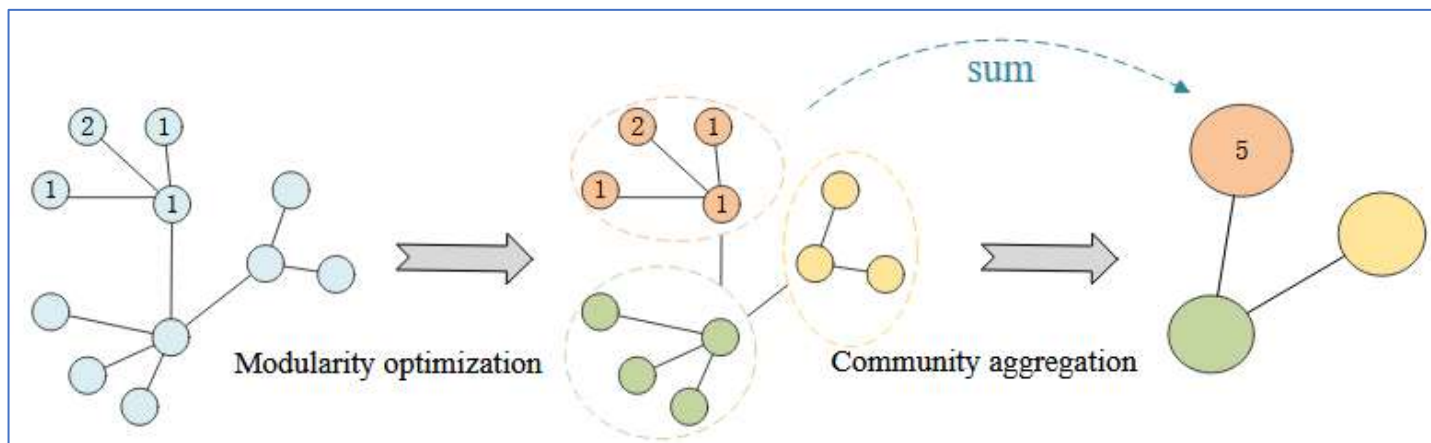
- **Louvain算法：群落检测算法**
- **两个阶段：模块优化和网络内聚**
- **评估：模块度衡量一个社区的划分是否优良**

$$Q = \frac{1}{2m} \sum_c \left[\sum_{in} - \frac{(\sum_{tot})^2}{2m} \right]$$

\sum_{in} : 社区 c 内的边的权重总和, \sum_{tot} : 与社区 c 内节点相连的边的权重总和, m : 图中所有边的权重总和。

- **应用**

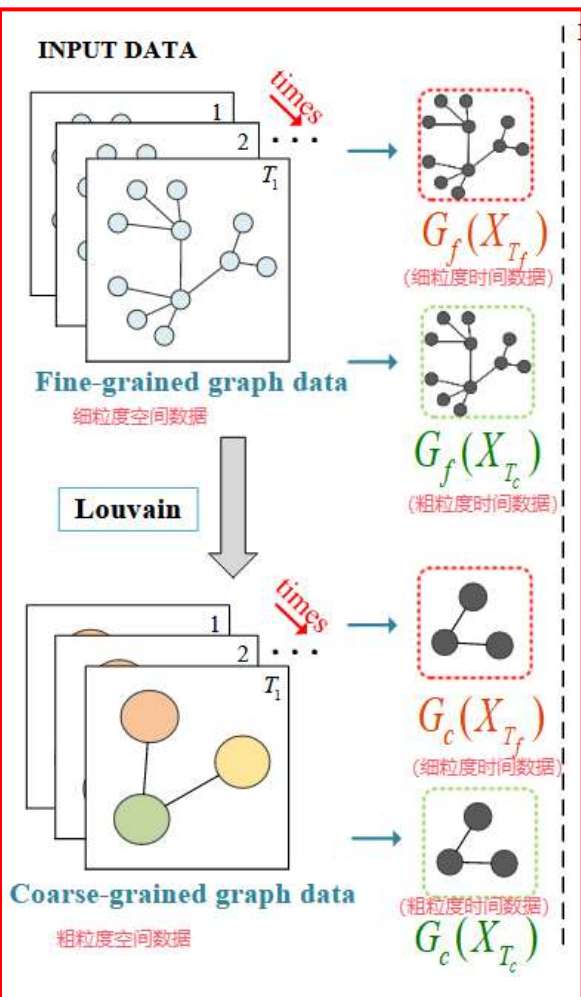
- 基于Louvain算法的新型交通区域划分方法, 根据交通特征的变化动态划分区域路网, 结果更加客观^[1];
- 基于Louvain算法对城市群落结构进行划分^[2]。



[1] Shiling Li, Jie Luo, and Jiaming Xu. Study on regional traffic sub-area division based on improved louvain algorithm and correlation degree. In 2022 China Automation Congress (CAC), pp. 35223527. IEEE, 2022.

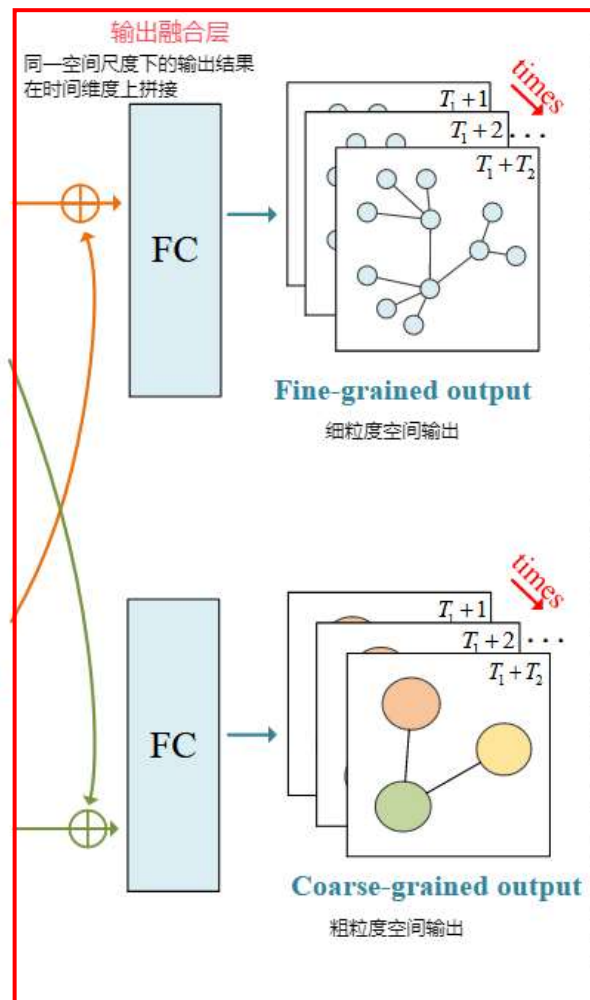
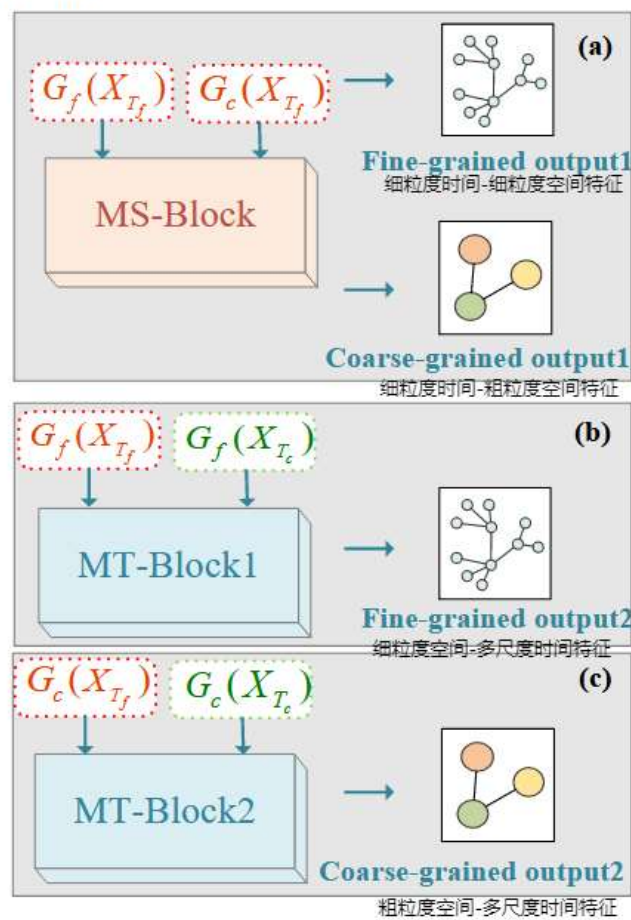
[2] Yan Zhang, Xiang Zheng, Min Chen, Yingbing Li, Yingxue Yan, and Peiying Wang. Urban finegrained spatial structure detection based on a new traffic flow interaction analysis framework. ISPRS International Journal of Geo-Information, 10(4):227, 2021.

框架



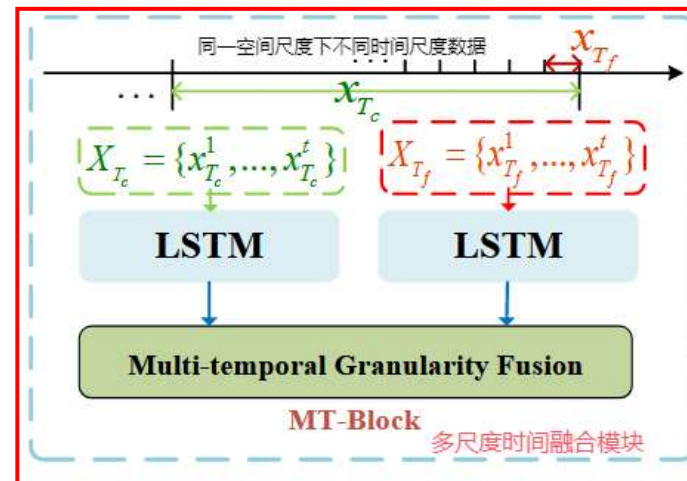
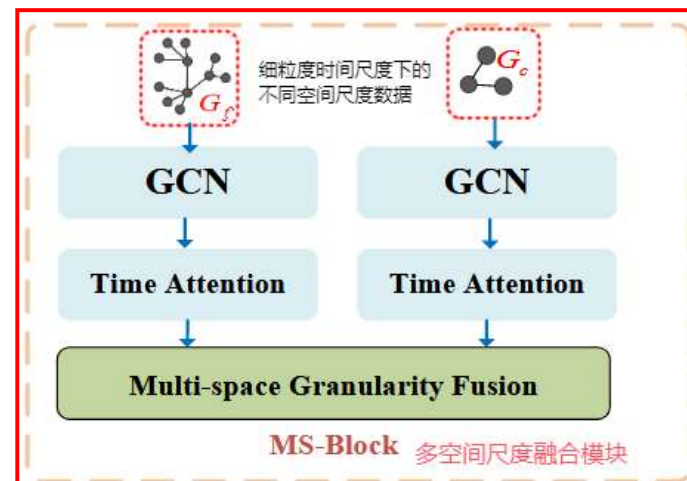
数据预处理

MST-GNN



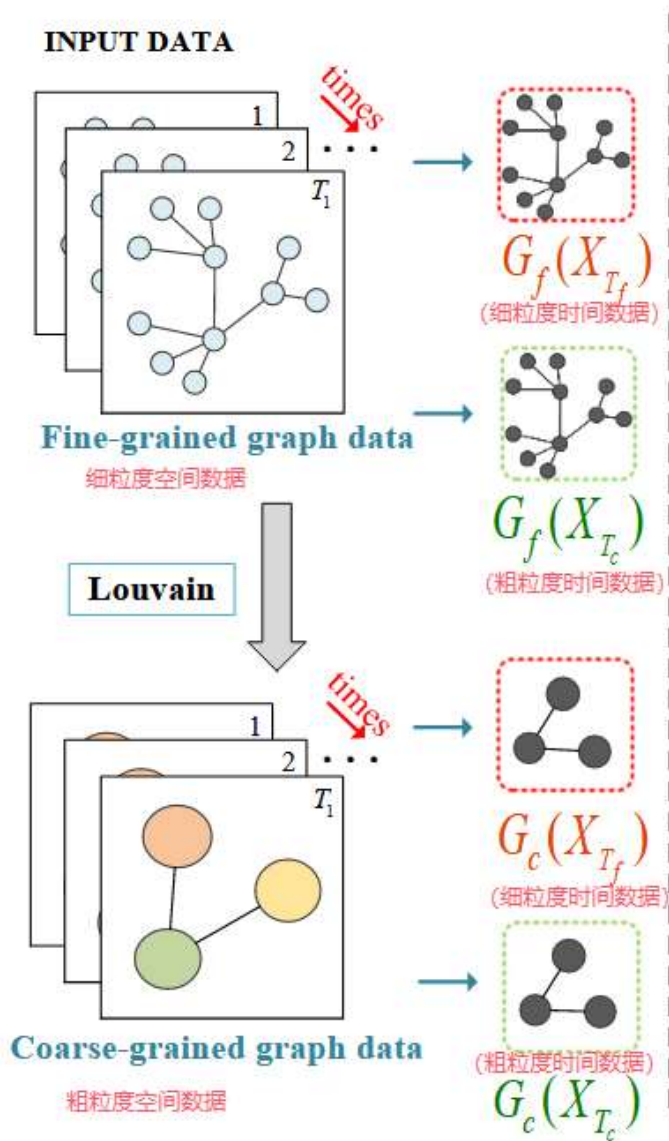
输出融合层

多尺度空间融合模块



多尺度时间融合模块

数据预处理



区域尺度的数据生成----Louvain

- ✓ 细粒度空间数据 $G_f \rightarrow$ 粗粒度空间数据 G_c
- ✓ 映射矩阵 A_{fc}
- ✓ 节点 G_c 的特征值=区域内包含的所有细粒度节点的特征值之和

较长时间尺度的数据生成

- ✓ 细粒度数据 $X_{T_f} = \{x_{T_f}^1, \dots, x_{T_f}^t, \dots, x_{T_f}^{T_1}\} \rightarrow$ 粗粒度数据 $X_{T_c} = \{x_{T_c}^1, \dots, x_{T_c}^t, \dots, x_{T_c}^{T'_1}\}$

计算公式:

$$x_{T_c}^t = \sum_{i=t-n*p}^t x_{T_f}^i \quad (2)$$

$$p = \frac{T_c}{T_f} \quad (3)$$

n : 单个粗粒度数据的时间步长

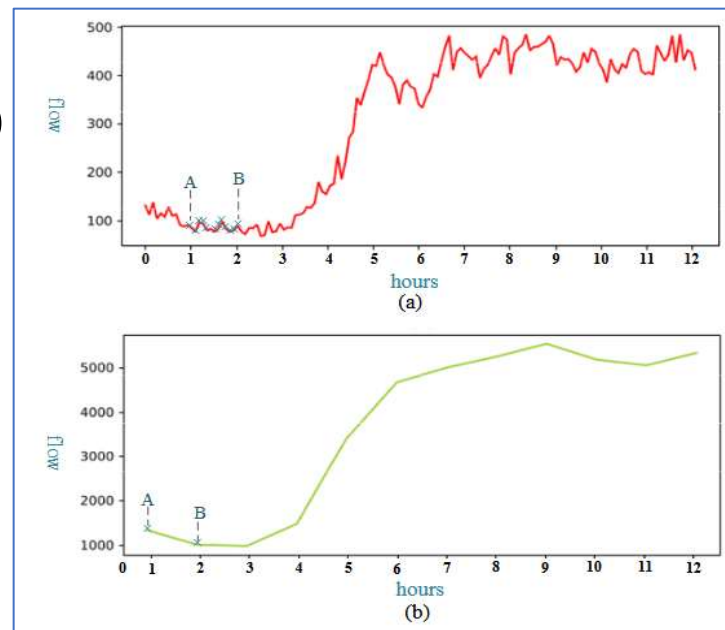
T_c : 粗粒度数据的时间尺度

T_f : 细粒度数据的时间尺度

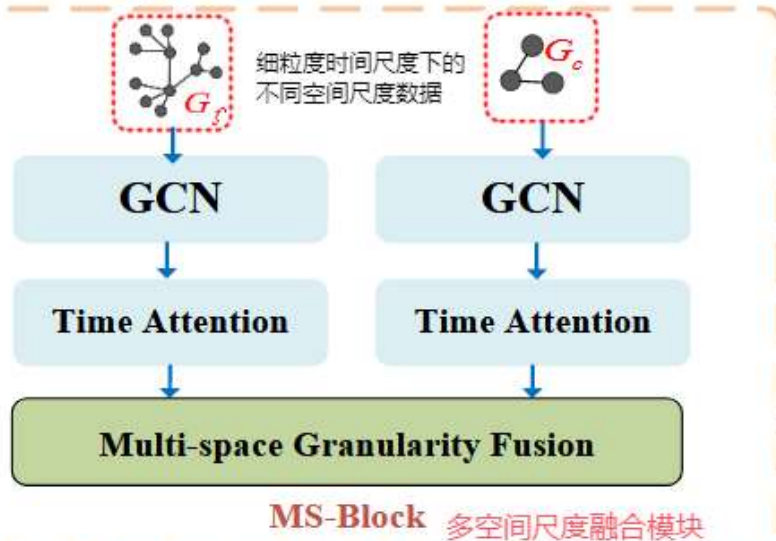
- ✓ eg: $T_f = 1min, X_{T_f} = \{x_{T_f}^1, \dots, x_{T_f}^t, \dots, x_{T_f}^{10}\}$

$$T_c = 2min, X_{T_c} = \{x_{T_c}^1, \dots, x_{T_c}^t, \dots, x_{T_c}^5\}$$

$$x_{T_c}^2 = x_{T_f}^1 + x_{T_f}^2, x_{T_c}^4 = x_{T_f}^3 + x_{T_f}^4 \dots$$



多空间尺度融合模块 MS-Block



■ GCN^[1]: 学习空间特征

图结构 $G = (V, E, A)$, 设 $H(l)$ 为所有节点的当前特征向量, 矩阵卷积运算公式:

$$H^{(l+1)} = \sigma(\tilde{D}^{-\frac{1}{2}} \tilde{A} \tilde{D}^{-\frac{1}{2}} H^{(l)} W^{(l)})$$

$\tilde{A} = A + I_N$, I_N : 单位矩阵

\tilde{D} : 阶数矩阵, $\tilde{D} = \sum \tilde{A}_{ij}$

■ 时间注意力层: 输入数据 X^l , 学习时间嵌入, 捕捉时间特征。

$$Q^l = X^l \cdot w_q$$

$$K^l = X^l \cdot w_k$$

$$V^l = X^l \cdot w_v$$

$$O^l = \text{softmax}(Q^l \cdot (K^l)^T \cdot V^l)$$

■ 多空间尺度特征融合层: 融合两个空间尺度的数据特征

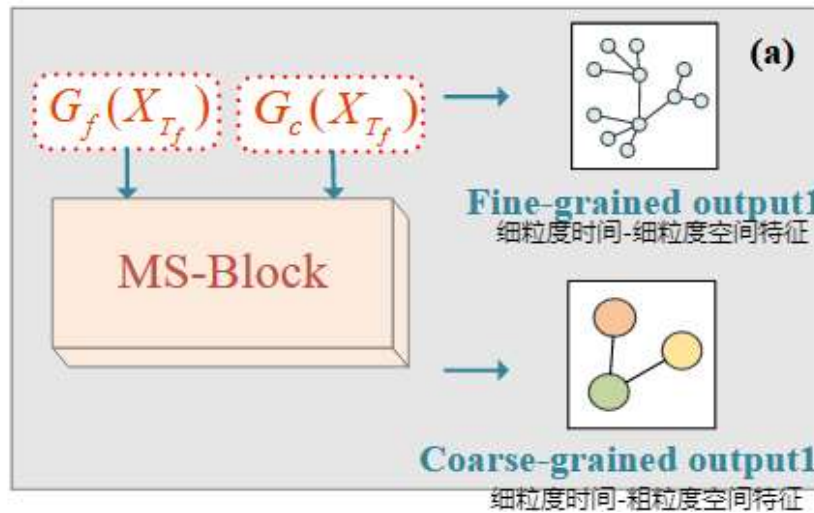
➢ 输入 (空间) 粗粒度数据和 (空间) 细粒度数据 X_c^l 和 X_f^l

➢ 不同尺度数据的空间特征通过映射矩阵相互影响

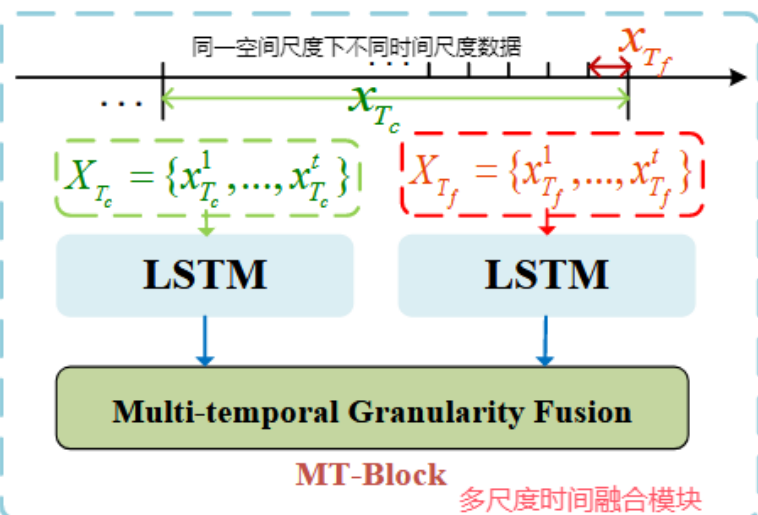
$$X_c^{l+1} = X_c^l + \eta_1 \cdot \text{softmax}(A_{fc} X_f^l W_c)$$

$$X_f^{l+1} = X_f^l + \eta_2 \cdot \text{softmax}(A_{fc}^T X_c^l W_f)$$

(A_{fc} 是数据预处理阶段的映射矩阵)



多时间尺度融合模块 MT-Block



■ LSTM层：提取不同时间尺度交通数据的时间特征

➢ 输入时间序列 $X^l \in \mathbb{R}^{s \times t \times f}$, LSTM 模型提取嵌入 $O \in \mathbb{R}^{s \times t \times f}$

➢ 全连接层线性变换, 得到LSTM 层的输出 $X^{l+1} \in \mathbb{R}^{s \times n \times f}$

(s : 空间维度, t : 输入的时间步长, n : 输出的时间步长, f : 特征维度)

■ 基于空间注意力的多尺度融合层:

➢ 输入经过LSTM处理得到的 (时间) 粗粒度和 (时间) 细粒度两组输出

➢ 将 (时间) 粗粒度特征融入细粒度特征

① 计算细粒度数据与粗粒度数据的相似度, 确定权重值

② 对粗粒度数据加权求和得到 “粗粒度数据相对于细粒度数据的融合值”

③ 将融合值与上一层细粒度数据特征相加

➢ 输出 (时间) 细粒度数据特征 (粗粒度空间、细粒度空间分别输出)

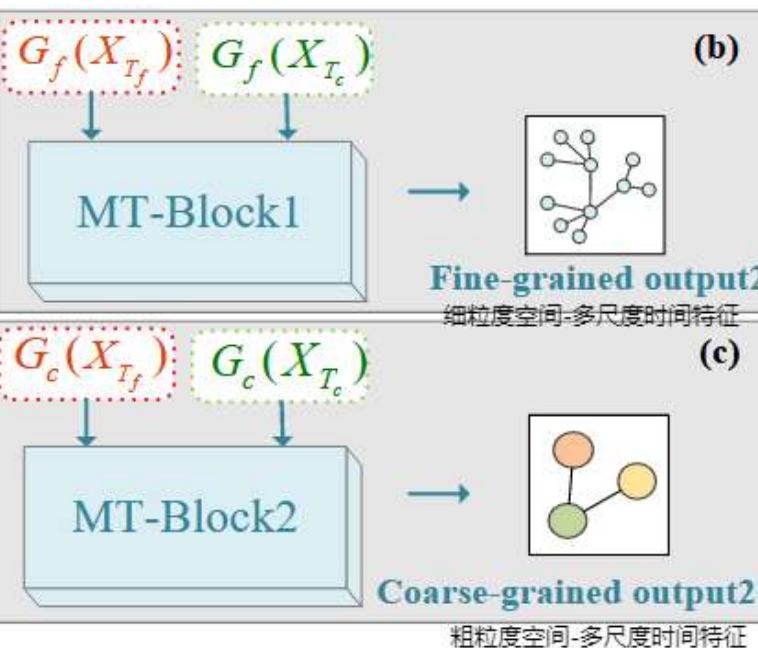
$$Q^l = X_f^l \cdot w'_q$$

$$K^l = X_c^l \cdot w'_k$$

$$V^l = X_c^l \cdot w'_v$$

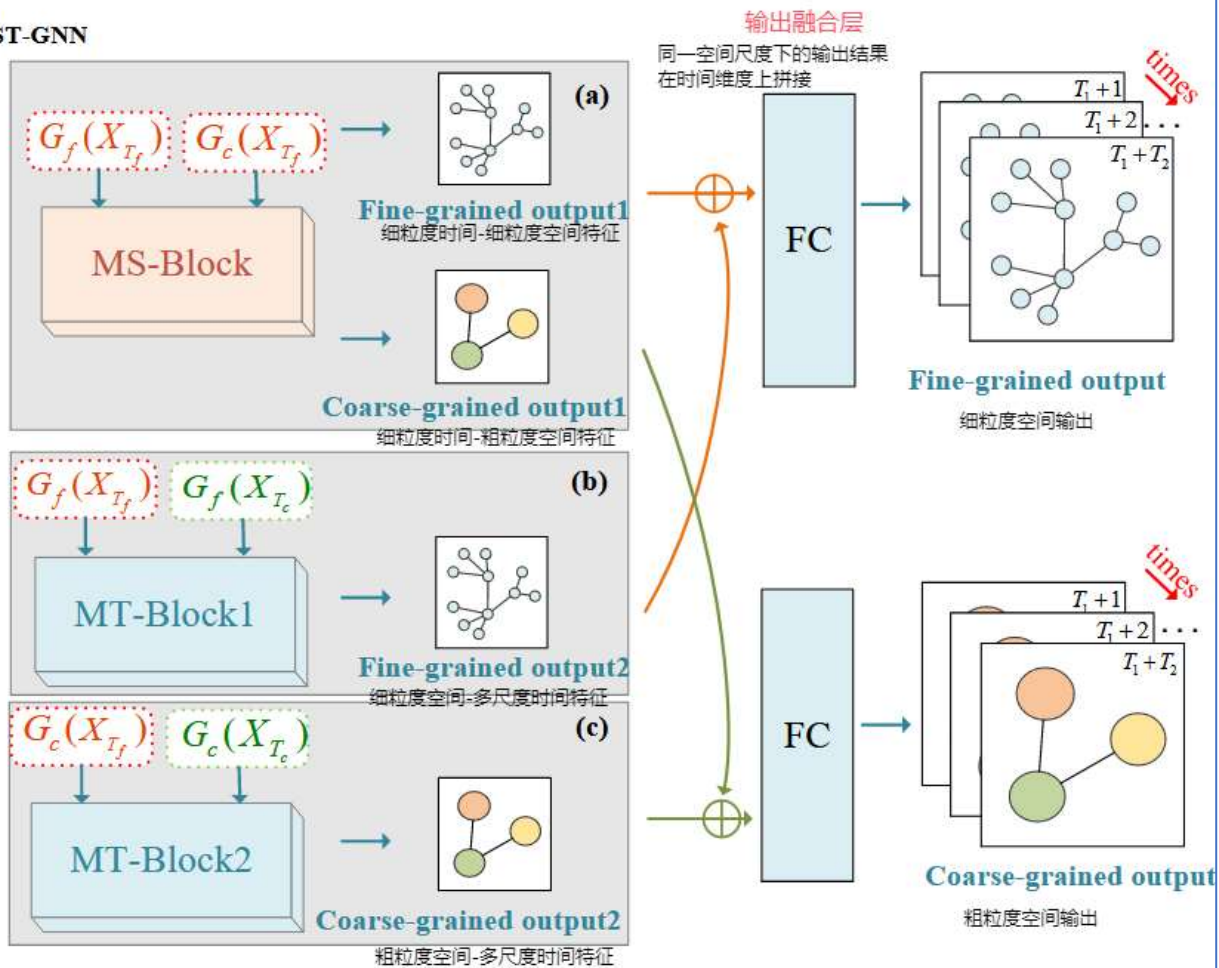
$$O^l = \text{softmax}(Q^l \cdot (K^l)^T \cdot V^l)$$

$$X_f^{l+1} = O^l + X_f^l$$



输出融合层

MST-GNN



上述两个模块的输出

- **MS-Block:** 细粒度空间数据 X_f^s 、粗粒度空间数据 X_c^s
- **2个MT-Block:** 细粒度空间数据 X_f^t 、粗粒度空间数据 X_c^t
- ※ 所有输出都具有细粒度时间尺度

输出融合层

- 同一空间尺度下的输出结果在时间维度上拼接
- 全连接层线性变换

$$X_{f_out} = FC([X_f^s, X_f^t])$$

$$X_{c_out} = FC([X_c^s, X_c^t])$$

- 输出细粒度空间预测值、粗粒度空间预测值

损失函数

$$\mathcal{L} = \eta_1 \|X_{f_target} - X_{f_out}\|^2 + \eta_2 \|X_{c_target} - X_{c_out}\|^2$$

其中, $\eta_1 = \eta_2 = 0.5$.

实验

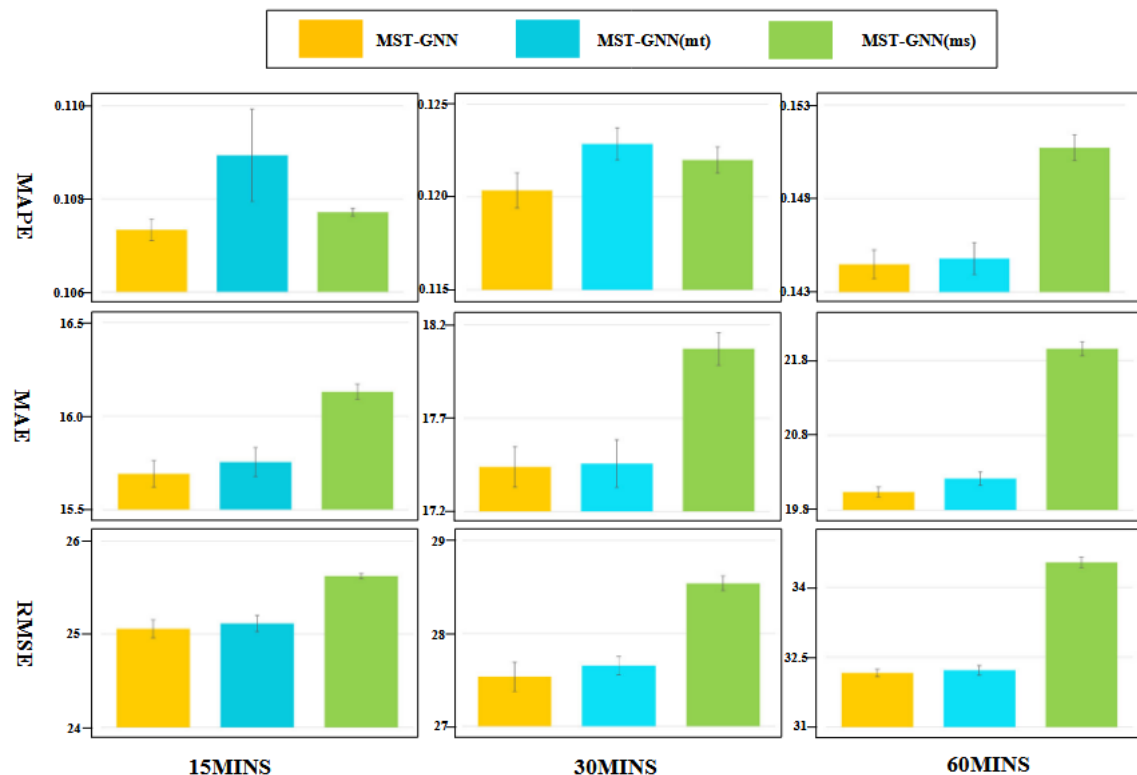
- 数据集：出租车流量数据集PeMSD4、PEMSD8（Louvain算法处理收，分别包含31、15个社区）
- Baseline：
 - 深度学习模型：GCN、LSTM
 - 最新交通流预测图神经网络：ASTGCN
 - 交通预测经典图神经网络模型：T-GCN、
 - 最新多尺度信息图神经网络：HGCN
- 实验结果 STGCN
- 短期交通流预测模型：STG-NCDE

Table 1: Performance comparison of MST-GNN and baseline models on PeMSD4 dataset.									
Time	5mins			15mins			30mins		
Metric	MAPE	MAE	RMSE	MAPE	MAE	RMSE	MAPE	MAE	RMSE
GCN	0.213	28.18	39.68	0.219	29.29	41.13	0.227	30.60	42.89
LSTM	0.138	21.15	31.82	0.140	21.65	31.39	0.154	23.55	34.13
TGCN	0.192	23.59	35.01	0.194	23.79	35.37	0.221	25.83	37.90
STGCN	0.149	19.79	30.80	0.147	20.84	32.63	0.173	22.67	35.13
ASTGCN	0.201	30.97	51.13	0.181	26.47	42.17	0.184	27.17	43.29
HGCN	0.125	17.73	28.10	0.140	19.68	29.41	0.154	20.67	32.27
STG-NCDE	0.120	18.50	29.28	0.123	18.86	30.03	0.133	19.49	30.99
MST-GNN	0.098	14.51	23.22	0.107	15.69	25.06	0.120	17.44	27.73

Table 2: Performance comparison of MST-GNN and baseline models on PeMSD8 dataset.									
Time	5mins			15mins			30mins		
Metric	MAPE	MAE	RMSE	MAPE	MAE	RMSE	MAPE	MAE	RMSE
GCN	0.210	34.02	46.09	0.208	34.76	46.98	0.212	35.49	48.00
LSTM	0.099	18.00	26.61	0.121	21.69	29.96	0.116	20.96	29.88
TGCN	0.244	22.80	31.82	0.230	22.46	31.52	0.250	23.42	33.08
STGCN	0.125	16.24	24.93	0.134	17.45	26.70	0.143	19.53	29.60
ASTGCN	0.143	21.35	34.88	0.138	20.06	32.70	0.142	20.65	33.72
HGCN	0.092	13.32	20.44	0.102	14.66	22.62	0.109	15.75	24.66
STG-NCDE	0.096	14.27	21.92	0.090	13.73	21.55	0.107	15.98	24.81
MST-GNN	0.063	9.98	15.26	0.067	10.86	16.89	0.075	11.89	18.57

Table 3: The computation time on PeMSD4 dataset.		
Method	Training	Inference
	Computation Time (s/epoch)	Computation Time (s)
HGCN	37.3	4.12
STG-NCDE	426.5	42.7
MST-GNN	4.1	0.45

消融实验



2种变体

- 仅多空间标模型 (MST-GNN(ms))
- 仅多时标模型 (MST-GNN(mt))

结果表明:

- MSTGNN 在 RMSE 指标上的改进幅度最大;
- 随着预测时间的增加, MST-GNN 的精度优势更加明显;
- 验证了 MST-GNN 模型中多时空尺度模块有效性。

总结

结论

- ✓ 提出用于交通流预测的MST-GNN模型。针对现有模型忽视交通数据跨时间尺度和空间尺度特征的问题，研究提出了简单有效的尺度融合解决方案。
- ✓ 模型包含一个多时间尺度特征融合模块和一个多空间尺度特征融合模块。
- ✓ 在两个公共数据集上进行对比实验，都表现出 SOTA 性能。
- ✓ 测试模型的计算时间，模型运行效率高。

未来方向

- 整合更多尺度的信息

谢 谢！