

2022科大讯飞A.I.开发者大赛
考虑时空依赖及全局要素的城市道路
交通流量预测挑战赛
决赛答辩

团队：BUAA

CONTENTS

- 一 团队介绍
- 二 算法方案解析
- 三 下阶段优化思路

算法方案解析

- 交通预测：带有时空附加属性的**时间序列预测**

- X_t : t 时刻的交通时空数据, \hat{Y}_t : 预测 t 时刻交通系统的某一状态
- T : 输入因素序列的长度, T' : 预测序列的长度



历史交通时空数据 $X_{t-T+1}, X_{t-T+2}, \dots, X_t$



未来交通状态 $\hat{Y}_{t+1}, \hat{Y}_{t+2}, \dots, \hat{Y}_{t+T'}$

算法方案解析

- **时空神经网络 (Spatio-Temporal Graph Neural Networks)**

- **时间相关性的建模**

- 时间序列的邻近性、趋势性和周期性



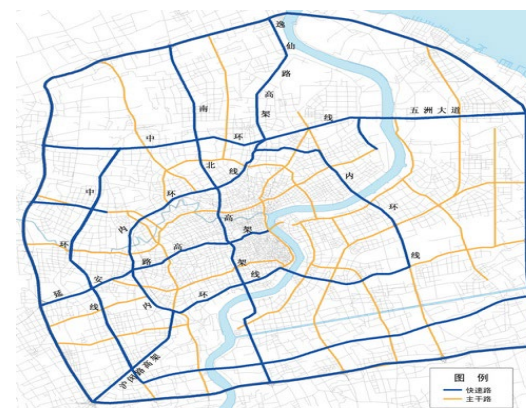
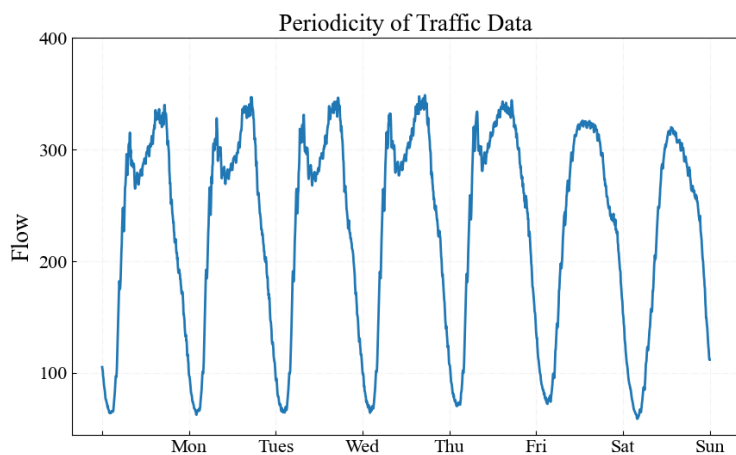
循环神经网络GRU

- **空间相关性的建模**

- 道路节点在路网上的邻接关系 (如：上下游的影响)



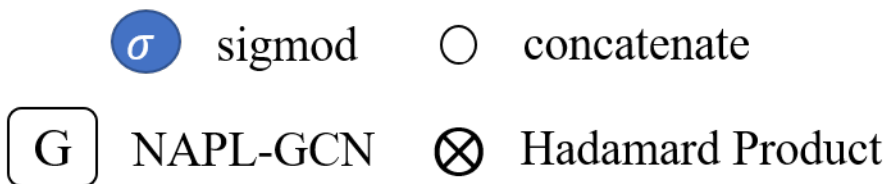
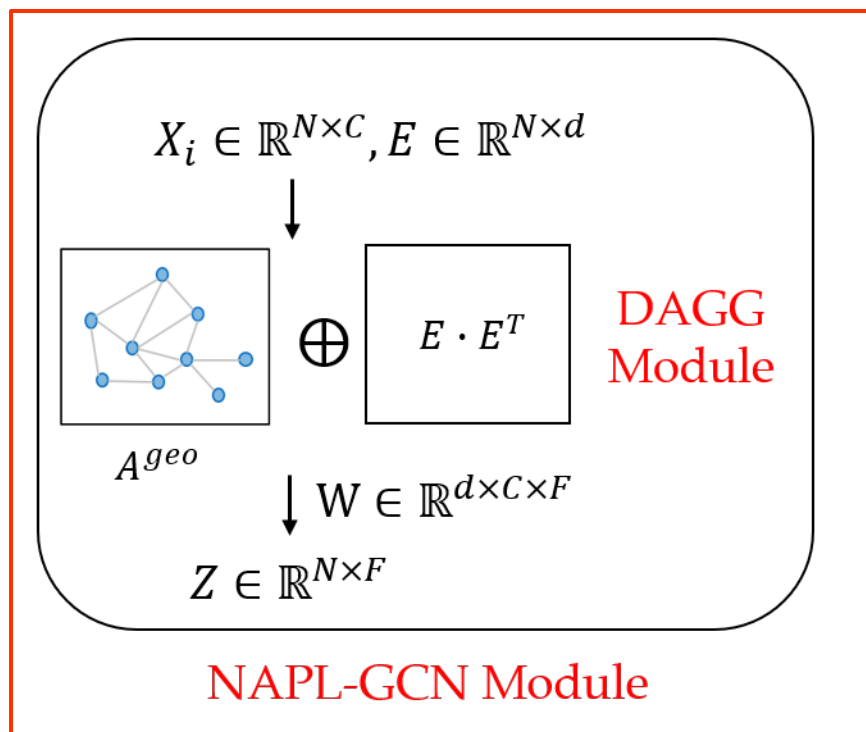
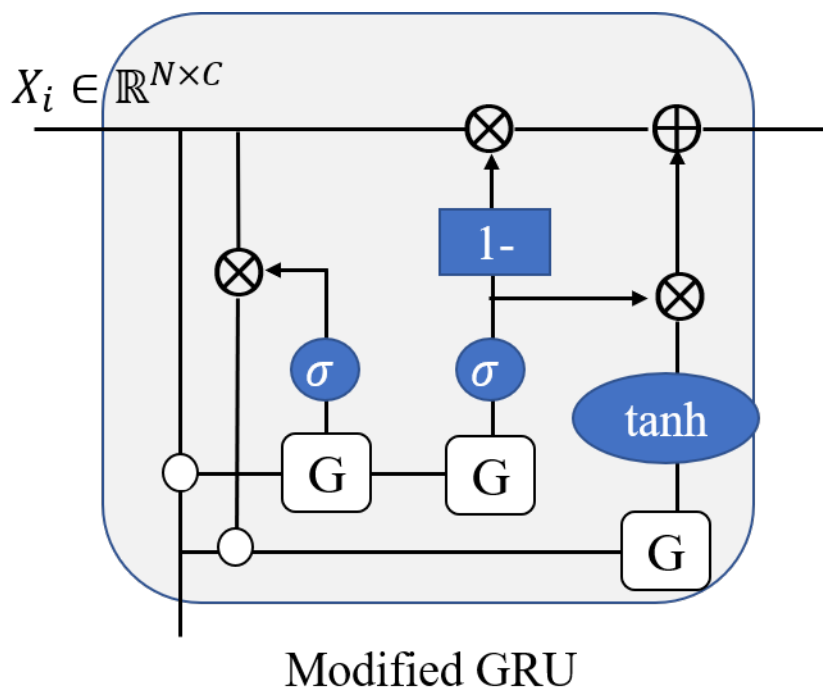
图卷积神经网络GCN



算法方案解析

• 模型总体结构

- 将扩展后的自适应图卷积模块嵌入到循环神经网络的计算中，共同捕获时空相关性。



E : Node Embedding
 X_i : Node Features
 W : Node Specific Parameters

算法方案解析

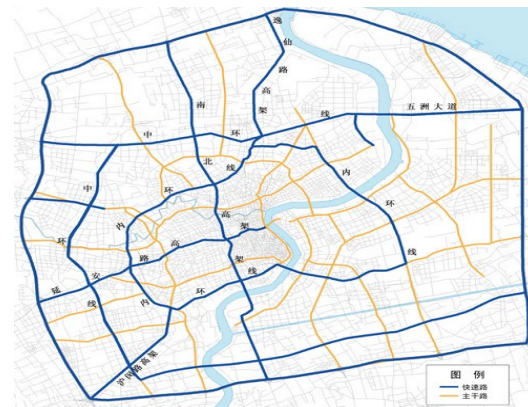
- 图卷积部分

- 地理距离图

- 路网节点之间的相对地理位置关系，如上游与下游的关系，对交通节点的流量有很大影响。
 - 根据数据中的节点的相对距离，构造了一个地理距离邻接矩阵。

$$A_{i,j}^{geo} = \begin{cases} 1, & \exp(-\frac{(D_{i,j}^{geo})^2}{\sigma^2}) \geq \epsilon \\ 0, & \exp(-\frac{(D_{i,j}^{geo})^2}{\sigma^2}) < \epsilon, \end{cases}$$

- 其中 $D_{i,j}^{geo}$ 表示节点 i 和节点 j 之间的距离， σ 是所有距离的标准差， $\epsilon = 0.1$ 是控制图的稀疏程度的参数。



• 图卷积部分

– 标准图卷积
$$Z = (I_N + D^{-\frac{1}{2}} A^{geo} D^{-\frac{1}{2}}) X_i \Theta + b$$

- A^{geo} 是图的邻接矩阵, D 是度矩阵, $X_i \in R^{N \times C}$ 和 $Z \in R^{N \times F}$ 是图卷积的输入和输出。
- $\Theta \in R^{C \times F}$ 和 $b \in R^F$ 分别为可学习的权重和偏置。
- 存在的问题: Θ 和 b 是所有图节点共享的!
- 为了学习特定于节点的模式, 即 $\Theta \in R^{N \times C \times F}$ 和 $b \in R^{N \times F}$, 考虑到参数太多, 通过矩阵分解的方法, 引入一个节点嵌入矩阵 $E \in R^{N \times d}$, 参数矩阵 $\tilde{W} \in R^{d \times C \times F}$, 参数矩阵 $\tilde{b} \in R^{d \times F}$, 而 $d \ll N$ 。这样即得到 $\Theta = E\tilde{W}$, $b = E\tilde{b}$ 。减少了参数数量。

– 节点自适应参数学习GCN:
$$Z = (I_N + D^{-\frac{1}{2}} A^{geo} D^{-\frac{1}{2}}) X_i E\tilde{W} + E\tilde{b}.$$

• 图卷积部分

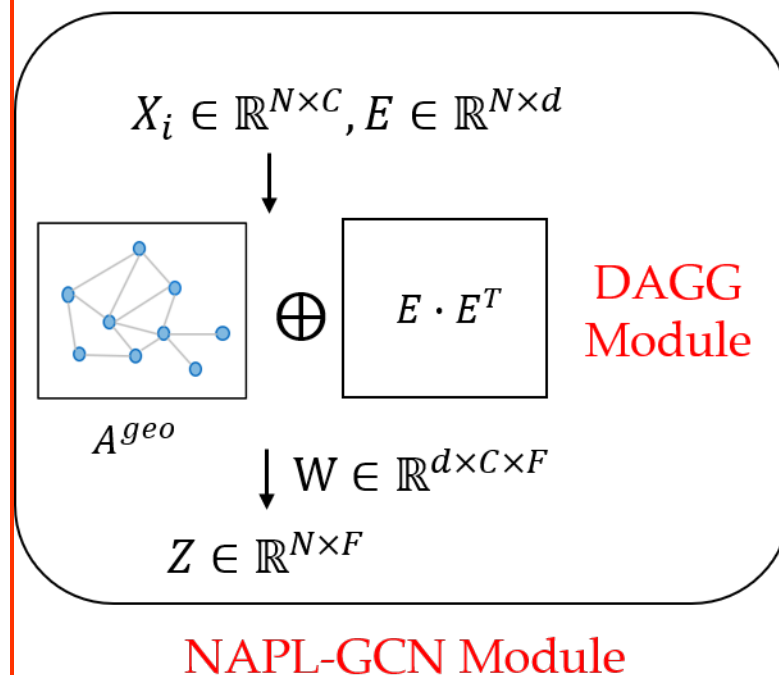
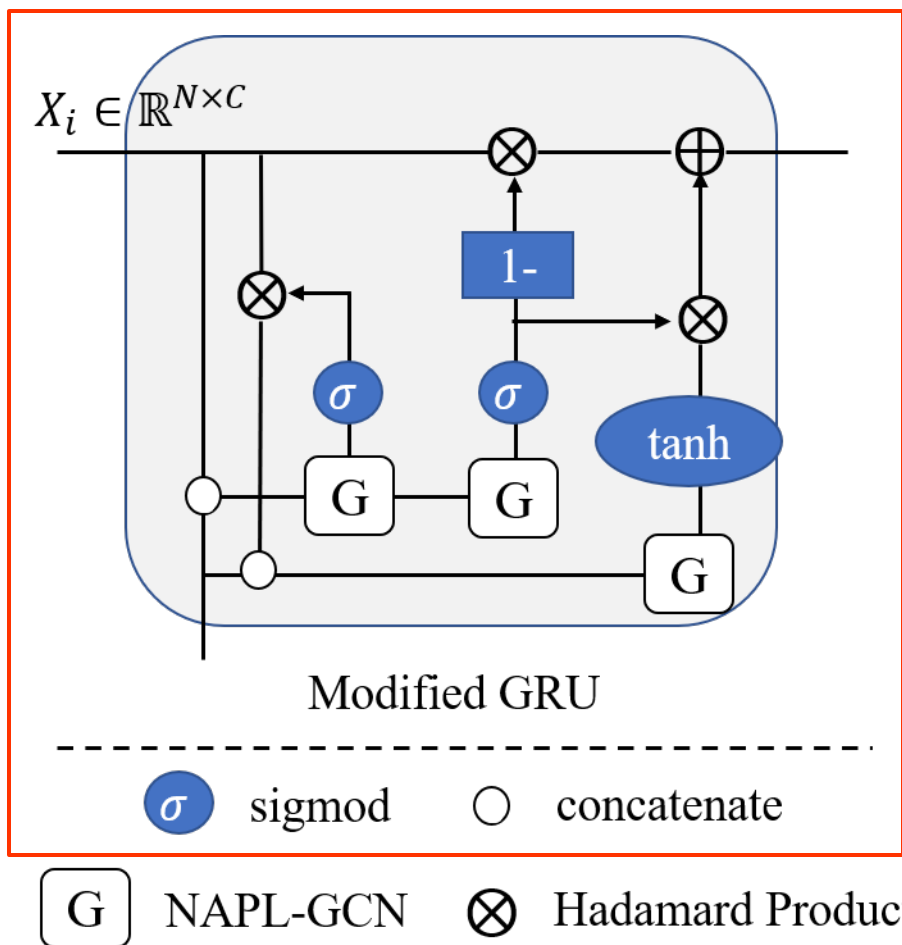
- 节点自适应参数学习GCN: $Z = (I_N + D^{-\frac{1}{2}} A^{geo} D^{-\frac{1}{2}}) X_i E \tilde{W} + E \tilde{b}.$
 - 存在的问题：空间关系的学习依赖于静态的预定义的空间邻接矩阵 A^{geo} 。
 - 为了自适应的学习节点对之间的空间依赖关系，使用节点嵌入矩阵 $E \in R^{N*d}$ 与转置矩阵 $E^T \in R^{N*d}$ 的乘积 $EE^T \in R^{N*N}$ 来模拟节点间的关系，这样一来这个空间关系矩阵是可以在模型训练过程中自我学习的。
- 节点自适应参数学习+数据自适应图生成GCN:

$$Z = (I_N + \boxed{\text{softmax}(\text{ReLU}(E \cdot E^T))} + D^{-\frac{1}{2}} A^{geo} D^{-\frac{1}{2}}) X_i E \tilde{W} + E \tilde{b}$$

算法方案解析

• 模型总体结构

- 将扩展后的自适应图卷积模块嵌入到循环神经网络的计算中，共同捕获时空相关性。



E : Node Embedding

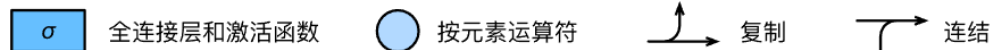
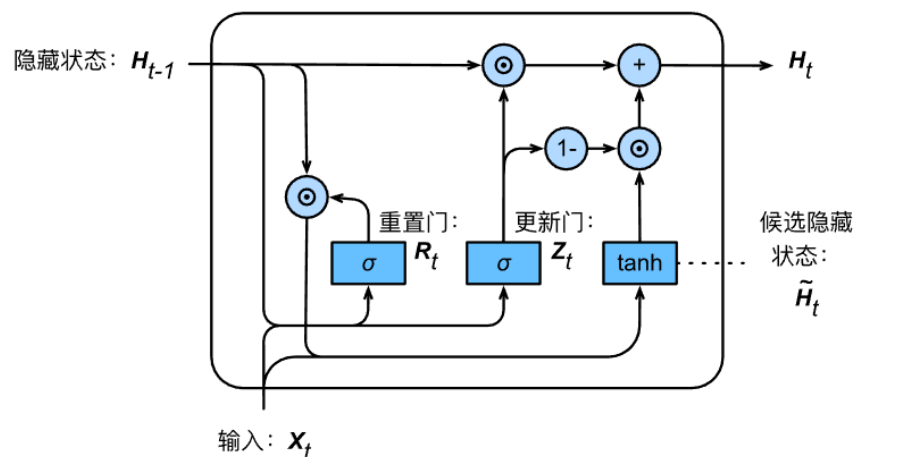
X_i : Node Features

W : Node Specific Parameters

算法方案解析

• 循环神经网络部分

– 选用门控循环单元GRU网络



当前时间步输入 x_t , 上一时间步隐藏状态 H_{t-1}

$$r^{(t)} = \sigma(\Theta_r \star_G [h^{(t)}, H^{(t-1)}] + b_r),$$

$$u^{(t)} = \sigma(\Theta_u \star_G [h^{(t)}, H^{(t-1)}] + b_u),$$

$$c^{(t)} = \tanh(\Theta_c \star_G [h^{(t)}, (r^{(t)} \odot H^{(t-1)})] + b_c),$$

$$H^{(t)} = u^{(t)} \odot H^{(t-1)} + (1 - u^{(t)}) \odot c^{(t)},$$

• 时空模块融合

- 将上述图卷积模块嵌入到循环神经网络的计算中, 替换其中的若干全连接层。

• 模型训练与融合

- 预测模块：直接进行12步的预测，获取12步结果
- 训练Loss: MSE
- 十折交叉训练：数据划分成10份，每次选择一份作为验证集，其他作为训练集，进行训练
- 模型融合：根据模型在验证集上的loss的倒数做权重，对不同模型的预测结果进行融合（loss越大，融合权重越小），得到最终预测结果

下阶段优化思路

- **神经网络**

- 自适应图学习部分学习到的图在不同时间步长是相同的，可以改成随着时间而动态变化的，更加符合交通预测这个多步时间序列预测的任务。

- **梯度提升树**

- 尝试树模型在交通预测上的应用
- 尝试模型融合

参考文献

- Bai L, Yao L, Li C, et al. Adaptive graph convolutional recurrent network for traffic forecasting[J]. Advances in neural information processing systems, 2020, 33: 17804-17815.
- Wang J, Jiang J, Jiang W, et al. Libcity: An open library for traffic prediction[C]//Proceedings of the 29th International Conference on Advances in Geographic Information Systems. 2021: 145-148.

交流与讨论