



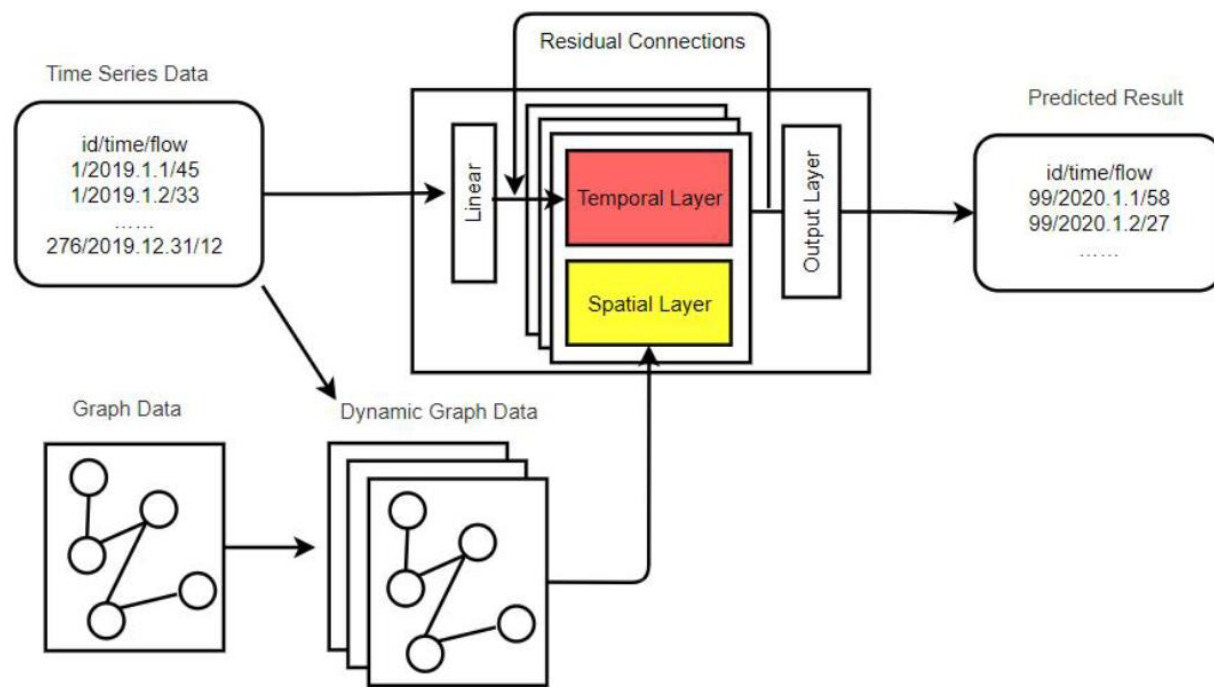
基于时空图卷积的交通预测模型

汇报人：庞媛媛

2024/05/10



通用框架、问题定义



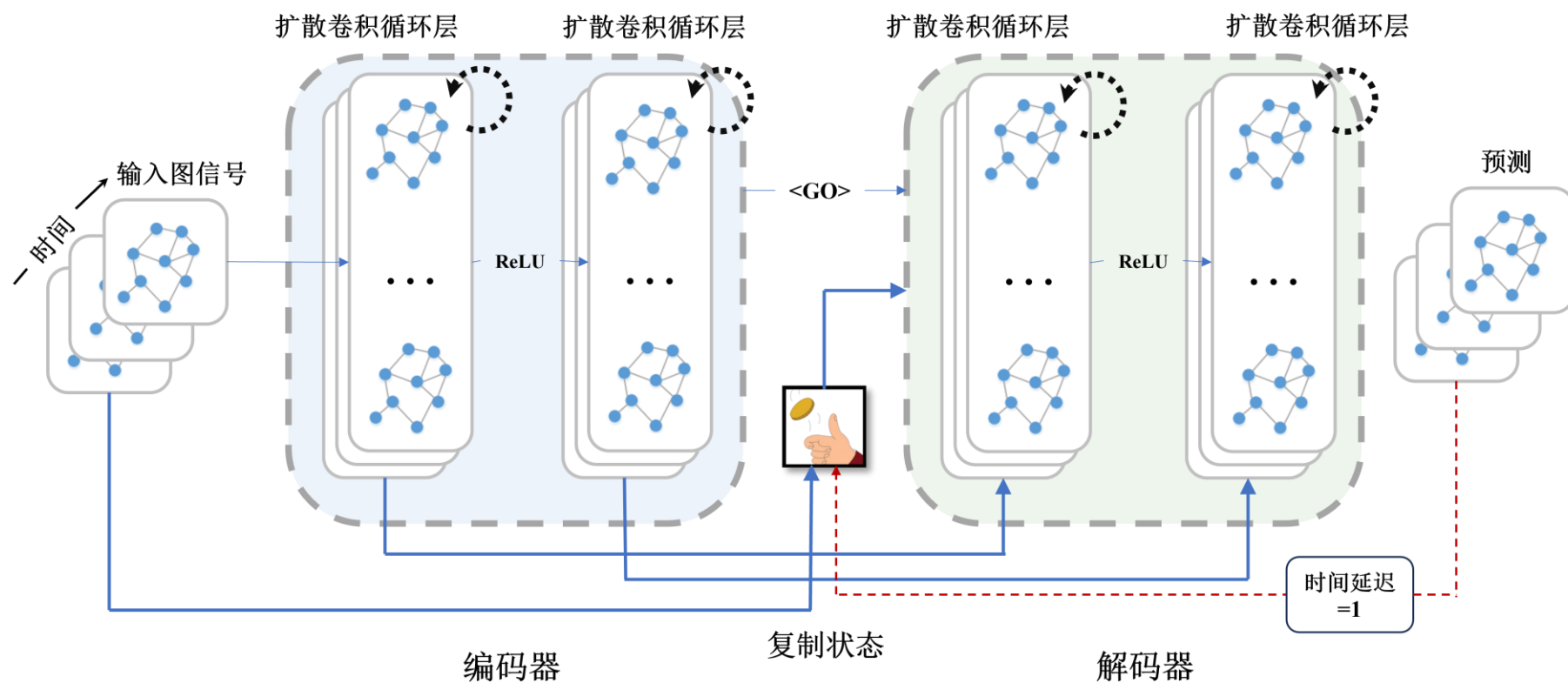
问题定义

- **目标:** 根据道路网络上 N 个传感器先观察到的交通流来预测未来的交通流
- **图定义:** 将传感器网络表示为加权有向图 $\mathcal{G} = (V, E, W)$
 - V 是节点集合, $|V| = N$
 - E 是边集合
 - $W \in R^{N \times N}$ 是权重邻接矩阵, 表示节点的接近度。
- **交通流定义:** 将 \mathcal{G} 上观察到的交通流表示为图信号 $X \in N^{N \times P}$, P 是每个节点的特征数。
- 令 $X_t \in N^{N \times P}$ 表示在时间 t 观察到的图信号 (图信号=图的动态特征矩阵 X_t)
- **本质:** 在给定图 \mathcal{G} 的情况下学习映射函数 f , 将 T' 个历史图信号映射到未来 T 个图信号。

$$[X^{(t-T'+1)}, \dots, X^{(t)}; \mathcal{G}] \xrightarrow{f} [X^{(t+1)}, \dots, X^{(t+T)}]$$

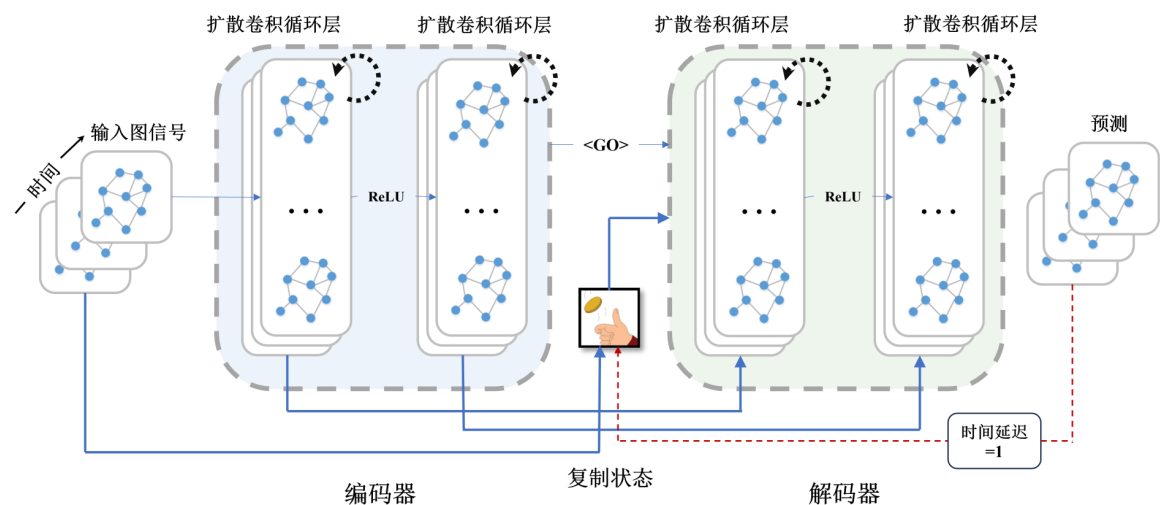
DCRNN-扩散卷积循环神经网络-2018

- 一句话概括：用由扩散过程形式化的图卷积捕捉时间依赖，用具有计划采样的编码器-解码器框架捕捉空间依赖。
- 背景：复杂的时空依赖性+长期预测的固有困难
- 整体架构：编码器+解码器，均基于DCGRU





DCRNN-模型细节



空间依赖性建模

空间依赖性建模

扩散过程~ \mathcal{G} 上的随机游走

重启概率 $\alpha \in [0,1]$
 转移矩阵 $D_O^{-1}W$
 出度对角矩阵 $D_O = \text{diag}(W\mathbf{1})$

扩散过程的平稳分布~图上无限随机游走的加权组合

$$\mathcal{P} = \sum_{k=0}^{\infty} \alpha(1-\alpha)^k (D_O^{-1}W)^k \quad k \text{ 是扩散步数}$$

扩散卷积

$$X_{:,p} \star_{\mathcal{G}} f_{\theta} = \sum_{k=0}^{K-1} (\theta_{k,1} (D_O^{-1}W)^k + \theta_{k,2} (D_I^{-1}W^T)^k) X_{:,p} \text{ for } p \in \{1, \dots, P\}$$

$\theta \in \mathbb{R}^{K \times 2}$ 表示卷积核参数

$D_O^{-1}W, D_I^{-1}W^T$ 分别是正向扩散过程和反向扩散过程的转移矩阵

扩散卷积层 (P 维特征映射到 Q 维)

$$H_{:,q} = a \left(\sum_{p=1}^P X_{:,p} \star_{\mathcal{G}} f_{\theta_{q,p,:}} \right) \text{ for } q \in \{1, \dots, Q\}$$

时间依赖性建模

门控循环单元 GRU $\xrightarrow{\text{扩散卷积代替矩阵乘法}}$ 扩散卷积门控循环单元 DCRU

带有计划采样的编码器、解码器

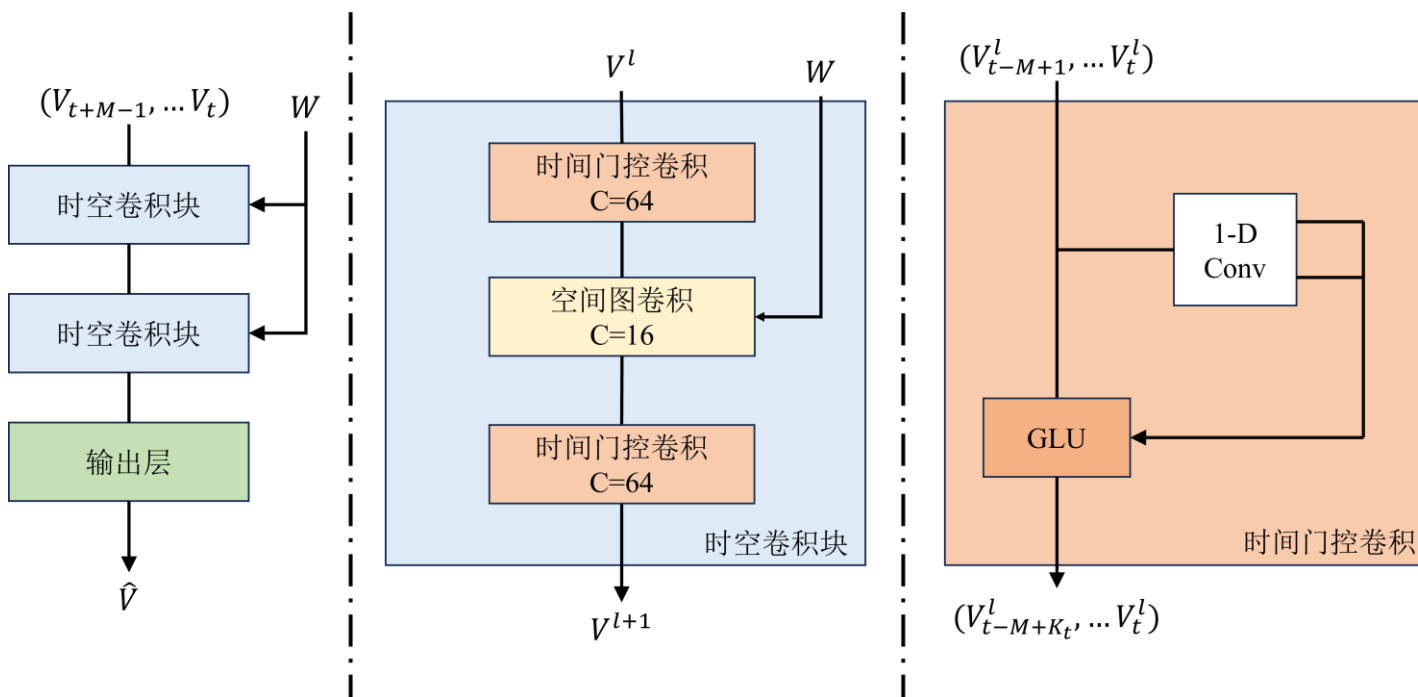
■ 编码器: 历史时间序列 \rightarrow 隐藏状态

■ 解码器: 隐藏状态 \rightarrow 预测时间序列

■ 计划采样: 弥补训练和测试的输入分布之间的差异, 训练期间向解码器提供部分真实值。

STGCN-时空图卷积网络-2018

- 一句话概括：综合图卷积和门控时间卷积捕捉时空关联
- 背景：由于交通流极高的复杂性和非线性，传统的方法不能满足中期和长期预测任务要求，且经常忽视时空依赖性。
- 整体架构：由多个时空卷积块和输出层组成，时空卷积块包含两个时间门控卷积和中间的一个空间图卷积模块。



□ 时域卷积块

- 一维卷积
- GLU激活

□ 空域卷积块

*ChebGCN*使用每个时间步的图数据在空域上高阶特征提取

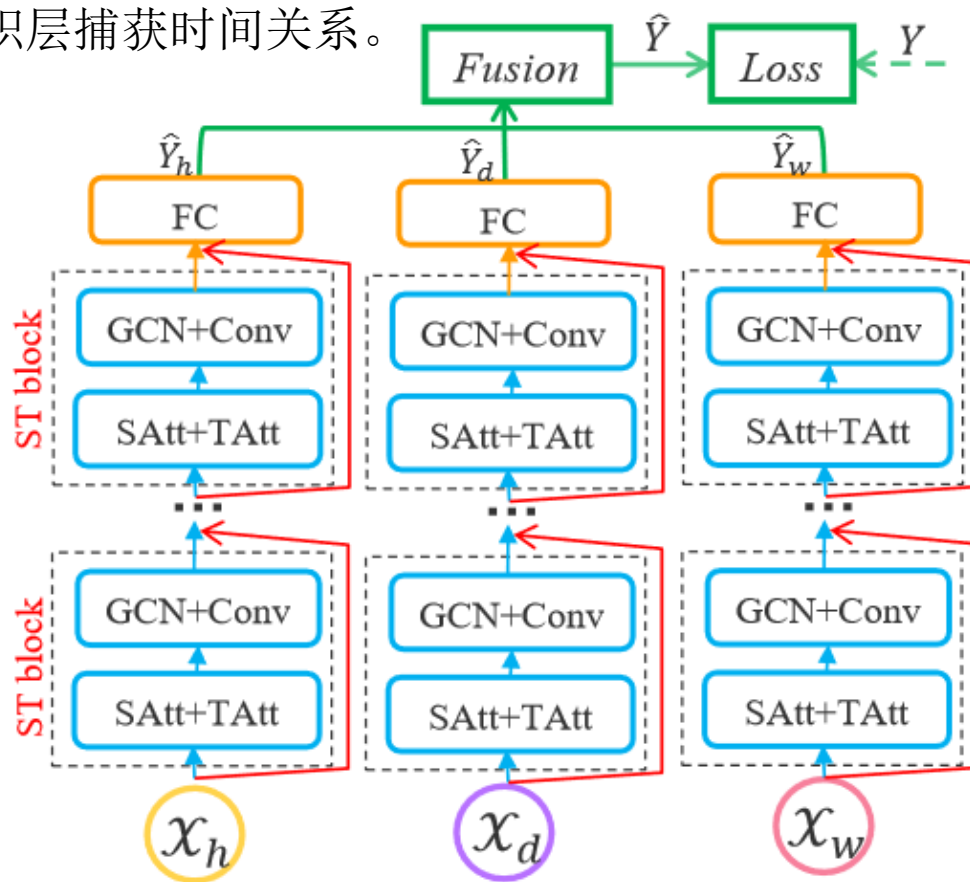
□ 输出层

- 一个时域卷积块
- 全连接层

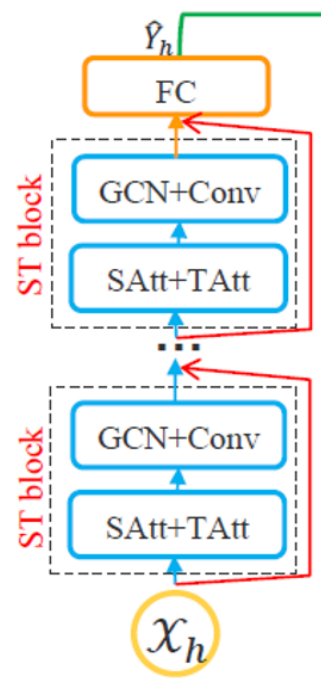


ASTGCN-基于注意力的时空图卷积网络-2019

- 一句话概括：融合时空注意力机制与时空卷积捕捉动态时空特征
- 背景：现有预测方法缺乏对交通流**动态时空关系**的建模
- 整体架构：由3个独立的组件构成，每个组件分别对交通流**3种时间特性**（recent、daily-periodic、weekly-periodic）建模，每个组件包含多个**时空注意力模块**和时卷积层，时空注意力机制捕获动态时空关系，时空卷积层通过图卷积捕获空间关系，普通卷积层捕获时间关系。



ASTGCN-模型细节



Spatial attention

第 r 个 ST block 的输入:

第 r 层输入数据的通道数

$$\mathcal{X}_h^{(r-1)} = (X_1, X_2, \dots, X_{T_{r-1}}) \in \mathbb{R}^{N \times C_{r-1} \times T_{r-1}}$$

第 r 层输入数据 时间维度的长度

注意力矩阵: 根据当前层输入动态计算

$$S = V_s \cdot \sigma((\mathcal{X}_h^{(r-1)} W_1) W_2 (W_3 \mathcal{X}_h^{(r-1)})^T + b_s) \quad (1)$$

$$S'_{i,j} = \frac{\exp(S_{i,j})}{\sum_{j=1}^N \exp(S_{i,j})} \quad (2)$$

节点 i 和节点 j 的关联强度

Temporal attention

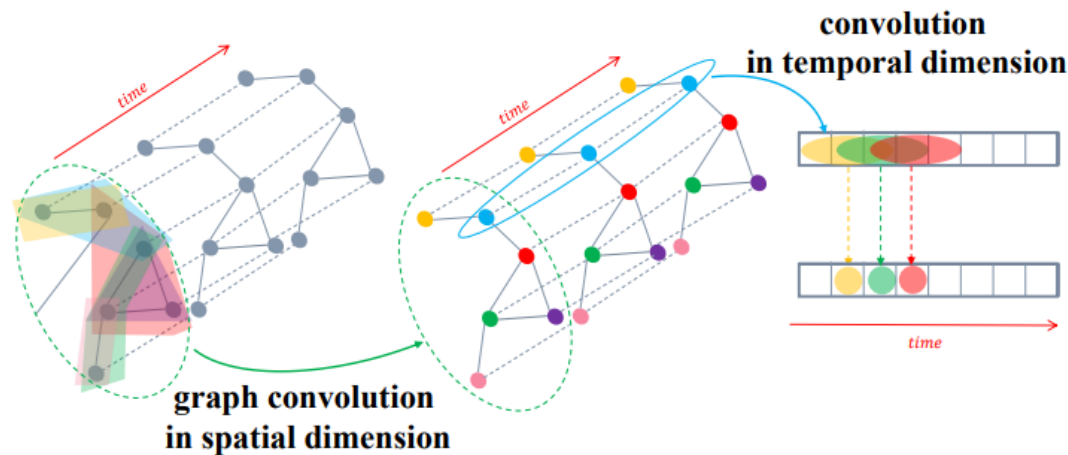
时间关联矩阵, 反映了时间 i 和时间 j 的依赖强度

$$E = V_e \cdot \sigma(((\mathcal{X}_h^{(r-1)})^T U_1) U_2 (U_3 \mathcal{X}_h^{(r-1)}) + b_e) \quad (3)$$

$$E'_{i,j} = \frac{\exp(E_{i,j})}{\sum_{j=1}^{T_{r-1}} \exp(E_{i,j})} \quad (4)$$

$$\hat{\mathcal{X}}_h^{(r-1)} = (\hat{X}_1, \hat{X}_2, \dots, \hat{X}_{T_{r-1}}) = (X_1, X_2, \dots, X_{T_{r-1}}) E'$$

原始输入 * 注意力矩阵, 以动态调整输入



□ 时空注意力

- 空间维度捕获不同区域交通状况相互影响的动态性
- 时间维度捕获不同时间片段交通流存在的相关性
- 动态调整图信号

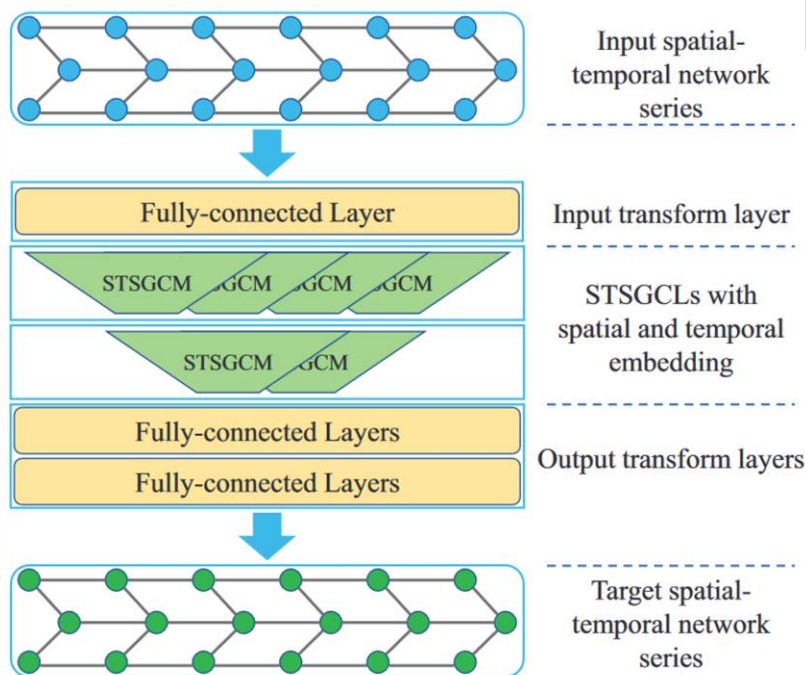
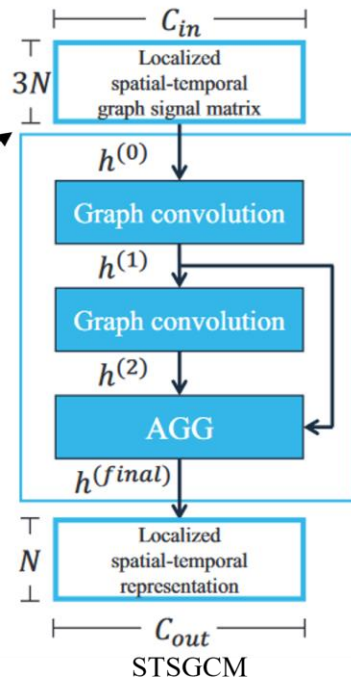
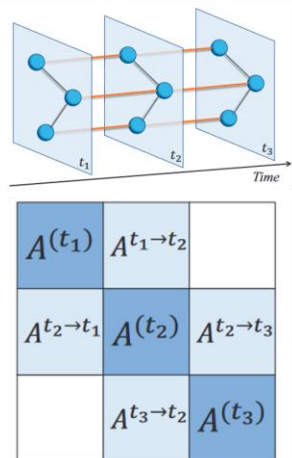
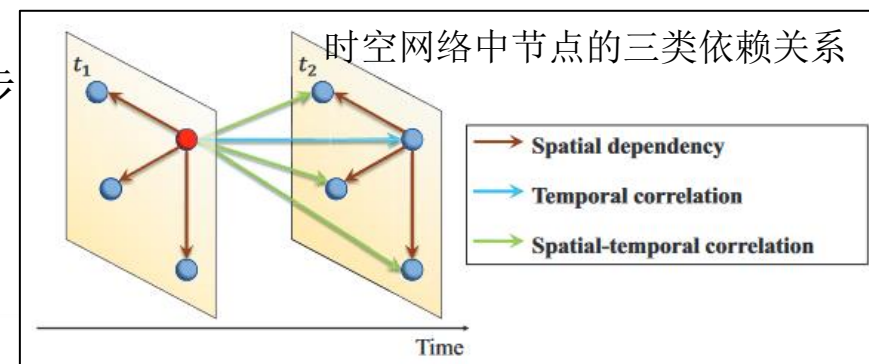
□ 时空图卷积层

- 图卷积捕获来自近邻节点的空间依赖性
- 标准卷积合并相邻时间步信息

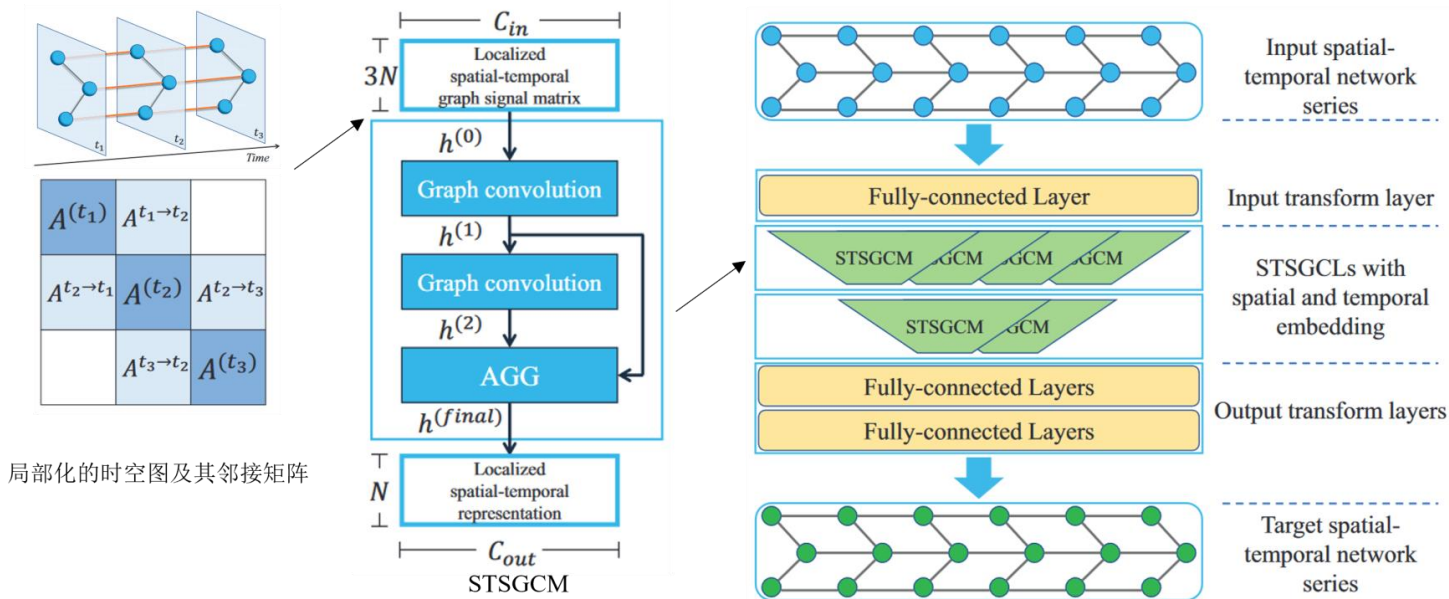


STSGCN-时空同步图卷积网络-2020

- 一句话概括：综合图卷积和门控时间卷积捕捉时空关联
- 背景：多数方法仅捕获前两种关系，然后将空间表征喂入时间维度建模得到第三种关系，**不是同时捕获时空网络中的依赖关系**；没有很好的处理**时空网络的异质性**。
- 整体架构：由一组时空同步图卷积层、全连接层和输出层组成，时空同步图卷积层包含多个时空同步卷积模块，**STSGCM捕捉时空相关性**，**STSGCL捕捉整个网络序列的长程时空相关性即异质性**。



STSGCN-模型细节



□ 局部时空图的构建

- 直接捕捉当前和相邻时间步的每个节点对其邻居的影响
- 将所有节点与相邻时间步的节点连接起来

- 局部时空图的邻接矩阵 $A'_{i,j} = \begin{cases} 1, & \text{if } v_i \text{ connects to } v_j \\ 0, & \text{otherwise} \end{cases}$

□ 时空嵌入

- 解决局部时空图掩盖时间属性的问题
- 可学习的时间嵌入矩阵和空间嵌入矩阵帮助模型捕捉时空关联建模能力

□ 时空同步图卷积模块STSGCM

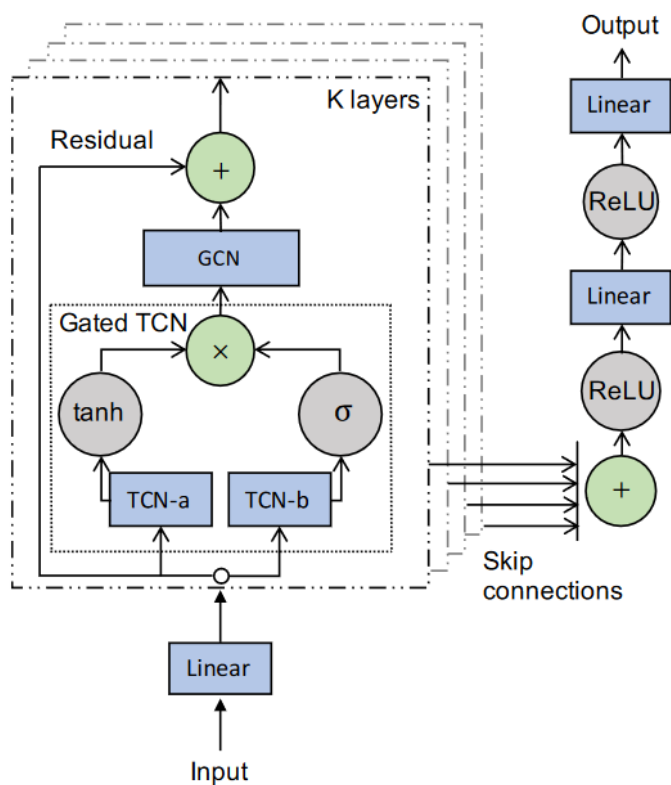
- 输入局部时空图邻接矩阵
- 两个图卷积聚合每个节点与邻居节点的特征
- 带有激活函数的全连接层将节点特征转换到新空间

□ 时空同步图卷积层STSGCL

- 捕捉长程时空相关性
- 使用滑动窗口切出不同时期并使用多个STSGCM对
不同时期建模， 提取长程空间-时间特征。

GWNET-Graph WaveNet-2019

- **一句话概括：** 采用自适应依赖性矩阵并通过节点嵌入学习来精确捕获数据中隐藏的空间依赖性，整合扩张因果卷积和扩散卷积捕获时空依赖性，同时能够处理长序列。
- **背景：** 固定图结构的假设导致空间依赖性的捕捉受限+RNN或CNN无法捕捉远程时间序列
- **整体架构：** 由堆叠的时空层和输出层组成，时空层由图卷积层（GCN）和门控时间卷积层（Gated TCN）组成，通过堆叠多个时空层处理不同时间级别的信息，底层接收短期时间信息，顶层接收长期时间信息。



模型细节

- **图卷积网络：** 采用GCN捕捉空间依赖关系
- **自适应图建模：** 通过节点嵌入学习自适应邻接矩阵来探索隐藏的空间依赖关系
- **时间卷积网络：** 采用扩张因果卷积捕获时间依赖关系
- **门控机制：** 时间卷积网络的基础上加入门控机制保留部分有效信息



AGCRN –自适应图卷积循环网络-2020

□ **一句话概括：**通过自适应模块增强传统图卷积，并将其结合到 RNN 中，以捕捉细粒度的时空相关性。

□ **背景：**

■ 现有基于GCN的交通预测模型对不同数据序列的**特定细粒度模式处理不准确**，偏向捕捉共享模式；

■ 基于节点在路网中的实际距离得到的邻接矩阵，但**先验路网可能不完整**。

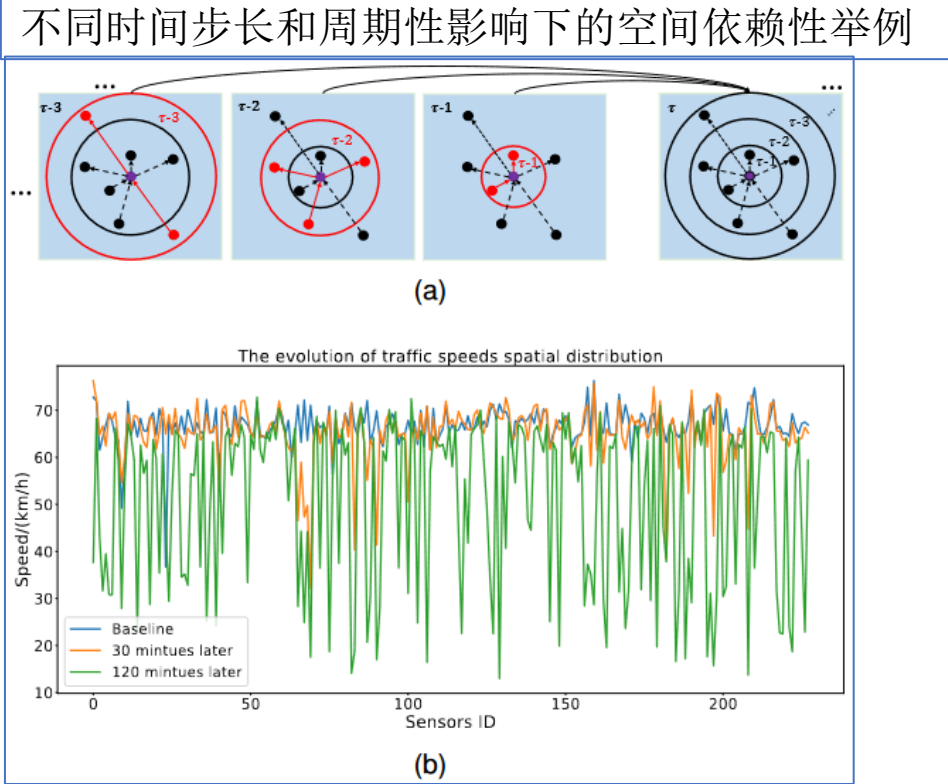
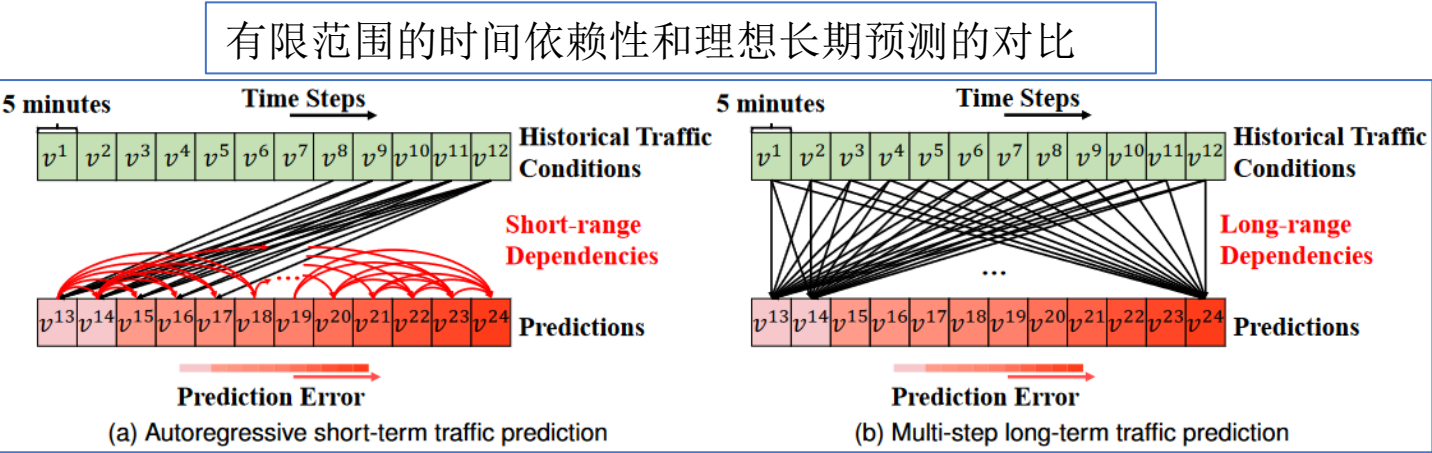
□ **核心：**提出两种机制增强目前的GCN方法

■ **自适应参数学习模块：**进行 GCN 特征变换时，对每个节点使用单独的权重和偏置以便学习每个节点特有的交通模式。

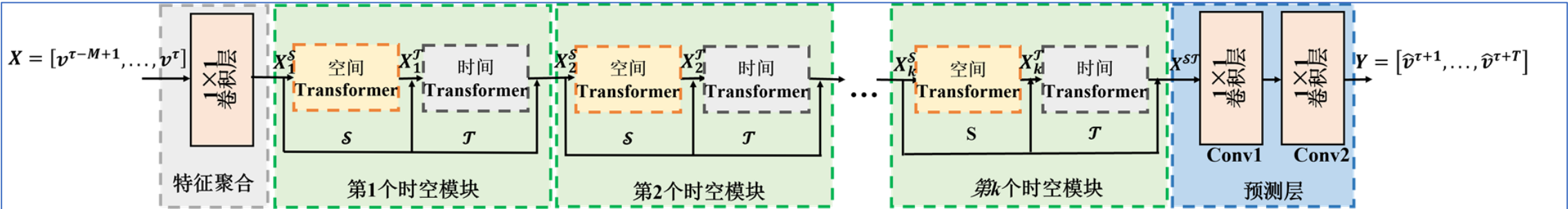
■ **数据自适应图生成模块：**使用节点嵌入向量自适应的通过数据驱动的方式学习邻接矩阵

STTN-时空Transformer网络-2020

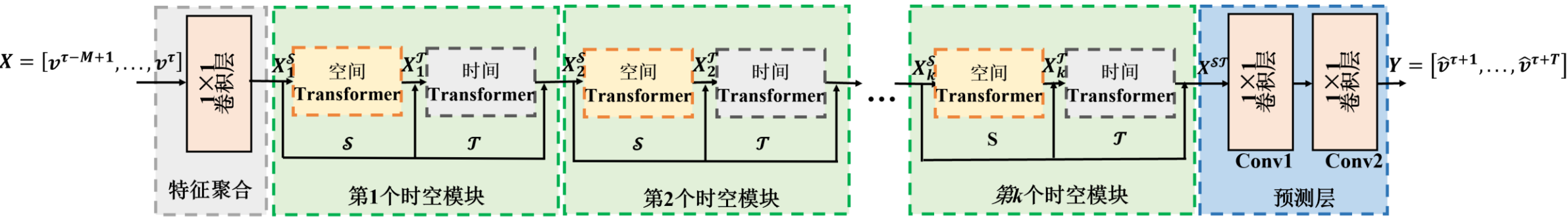
- 一句话概括：综合图卷积和门控时间卷积捕捉时空关联
- 背景：固定的空间依赖性+有限范围的时间依赖性
→准确预测困难，长期预测挑战



- 整体架构：由堆叠的空间-时间块和预测层组成，空间-时间块由空间Transformer和时间Transformer组成，动态依赖的背景下共同提取时空特征。时空模块进一步堆叠形成深度时空特征的深度模型。预测层利用两个 1×1 卷积层聚合时空特征进行交通预测。



STTN-模型细节



- 时空模块

 - 整合空间和时间Transformer，共同模拟交通网络内的空间和时间依赖关系，以实现精确预测。
 - 第*l*时空模块的输入是第*l-1*个时空模块提取的*N*个节点在时间步长 $\tau - M + 1, \dots, \tau$ 的 d_g 维特征的三维张量 $X_l^S \in \mathbb{R}^{M \times N \times d_g}$
 - 空间Transformer和时间Transformer被叠加产生3-D输出张量，残差连接用于稳定训练。

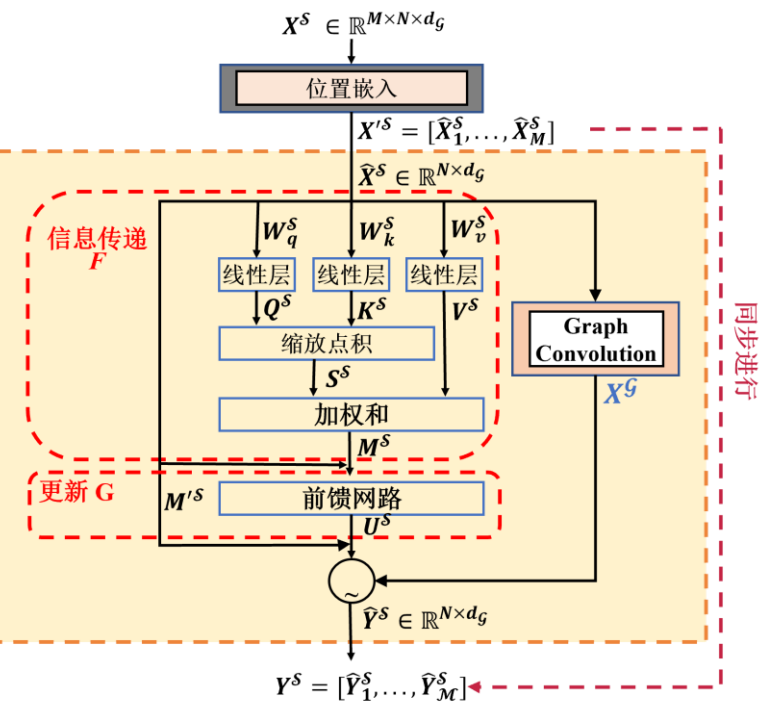
Step1 空间Transformer从输入节点特征 X_l^S 和图邻接矩阵 W 中提取空间特征 Y_l^S 即 $Y_l^S = \mathcal{S}(X_l^S, A)$

Step2 Y_l^S 与 X_l^S 结合生成输入 X_l^T ，进入时间Transformer得到时空特征 Y_l^T 即 $Y_l^T = \mathcal{T}(X_l^T)$

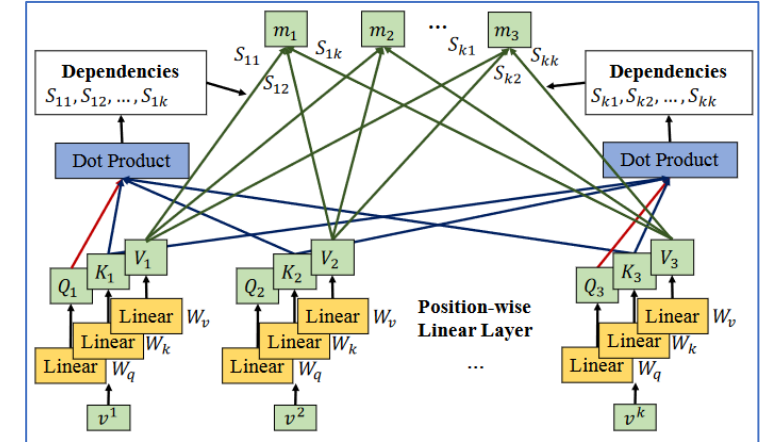
Step3 计算输出张量 $X_{l+1}^S = Y_l^T + X_l^T$ 送入下一时空模块

Step4 根据最后一个时间步的时空特征进行预测，输入是由时间长 τ 的*N*个节点的 d^{st} 维时空特征组成的二维张量 $X^{ST} \in \mathbb{R}^{N \times d^{ST}}$ ，预测 $Y = \text{Conv}(\text{Conv}(X^{ST}))$

STTN-模型细节



(a) 空间Transformer S



- 空间-时间位置嵌入层
 - 两个节点之间的空间依赖关系由它们的距离和时间步决定，不用先验信息
 - 采用可学习的空间和时间位置嵌入层学习空间-时间嵌入到每个节点特征中
 - 空间和时间可学习嵌入矩阵 \mathcal{D}^S 、 \mathcal{D}^T 分别由图邻接矩阵、单次时间编码初始化以考虑节点之间的连通性和距离、时间步关联
 - 被送入固定图卷积层和动态图卷积层用于空间特征学习
- 固定图卷积层
 - 捕捉道路拓扑结构中的静态空间依赖
 - 采用基于切比雪夫多项式近似的图卷积（ChebNet）学习节点特征
- 动态图卷积层
 - 从实时交通状况和突发变化中捕捉动态空间依赖
 - 采用基于注意力机制的动态图卷积学习节点特征
 - 通过学习多个线性映射来模拟在不同潜在子空间中受各种因素影响的动态定向空间依赖关系
 - 三个潜在子空间：Q、K、V，QK提取动态空间依赖关系，并与V更新节点特性
- 信息融合的门机制
 - 融合从动态和静态图卷积学到的空间特征

谢 谢！