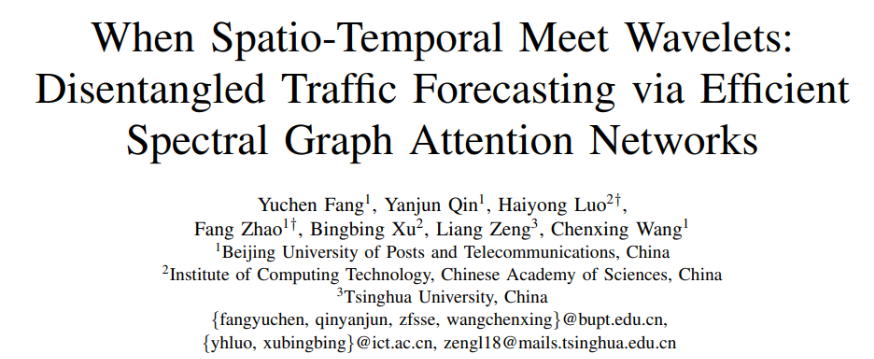
ICDE 2023 STWave：一 种新的解纠缠融合交通流量预测框架

交通流量预测对公共安全和资源优化至关重要，但由于相关数据的时间和空间动态变化的相关性，以往基于**端到端训练**的时空网络无法处理非平稳时间序列中的分布位移。本次为大家带来数据库领域顶级会议 ICDE 的文章《When Spatio-Temporal Meet Wavelets: Disentangled Traffic Forecasting via Efficient Spectral Graph Attention Networks》,针对上述问题提出了一个新的解纠缠融合框架STWave 以缓解分布移位问题.



# 一,背景

随着科技的不断发展，越来越多的传感器被应用到监测系统当中，例如测速摄像头和环路探测器等通过记录速度和流量等信息，为预测交通流量，便利道路管理提供了极大帮助。然而在时变的交通环境下，由于交通时序的复杂变化和传感器之间的空间相关性，交通预测具有很大的挑战性。

一种常见的解决方案是直接将交通数据喂给一个**时空网络模型**如STNet当中，采用基于序列和图结构的深度模型方法，以端到端的训练方式来处理时空依赖关系。但其流量预测仍然存在一些**问题**：

1. 在时间方面，由于交通时间序列与多个局部独立模块纠缠，并且局部独立模块可能发生分布移位，因此可能导致端到端STNets**过拟合**

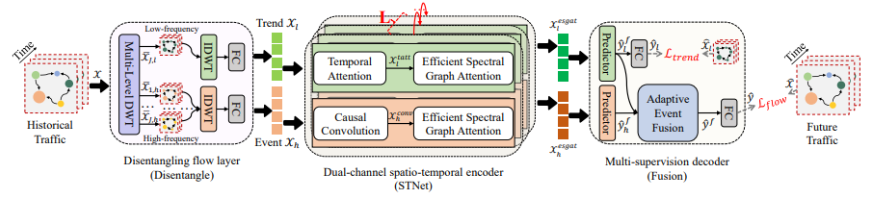
2. 在空间方面，完全基于 GAT（图注意力网络）的是学习效率成本较高，时间复杂度和空间复杂度都为平方级别，且由于只考虑值的空间信息，可能会造成**过拟合**

为了解决这些问题，该论文提出了一种新的交通预测框架 STWave.

针对上述问题, 该框架利用**离散小波变化 DWT** 将交通时间序列分解成双尺度趋势事件的形式.然后STWave设计了一个 STNet,利用相应的顺序和**基于图的方法**对不同的信息捕捉各种时间和空间变化的相关性。此外，针对GAT 方法的高复杂度和结构信息不足的问题，STWave还提出了一种新**的查询采样策略和基于图的小波位置编码方法**.最后，在STWave中使用自适应事件融合模块，将不准确的预测事件中的有用信息合并为易于预测的趋势.

# 二, 方法

下图阐述了 STWave 框架，它由以下三个重要组件组成



1. **Disentangling flow层**:给定所有传感器的历史流量数据，STWave首先利用多层离散小波变换(DWT)将所有传感器的纠缠历史时间序列分离成一个低频组件和多高频组件。为增强表征能力，STWave按时间顺序采用逆小波变换(IDWT)并在全连接层推导稳定趋势和波动事件.

2. **双通道时空编码器**:基于事件的稳定和波动特性, STWave利用时序注意力和因果卷积来捕捉稳定和波动的时间变化。对于时空交通数据的动态空间依赖的学习，STWave采用两种高效的频谱图注意力网络.

3. **Multi-supervision译码器**:STWave利用两个预测器来预测未来的趋势和事件，然后使用自适应事件融合模块来推导未来的流量

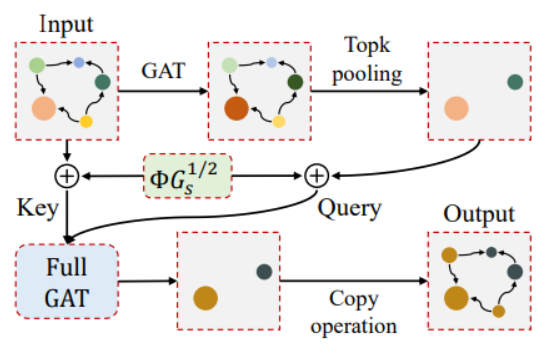
## Disentangling Flow层

DWT可以根据不同的频率使用小波滤波器从输入信号中分离出多个分量，如缓慢变化的稳定趋势对应于低频. 因此文章利用多层次DWT通过滤波器获得足够平滑的低高频和多高频分量，其中低高频分量可以代表流量时间序列中的稳定趋势和波动事件。

在 Disentangling Flow层之后，文章得到了解耦后的交通数据的趋势事件表示，它们可以在下一步进行并行处理.

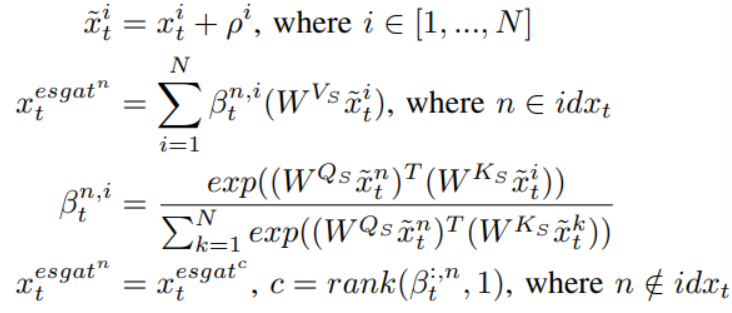
## 双通道时空编码器

双通道时空编码器通过叠加因果卷积、时序注意力和高效谱图注意网络 (ESGAT) *L*次来捕获波动时间变化、稳定时间变化和空间相关性，其体系结构如图所示

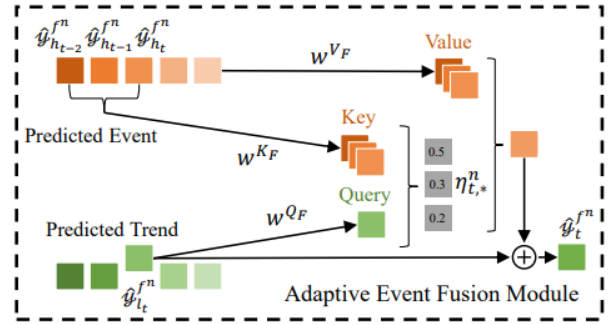


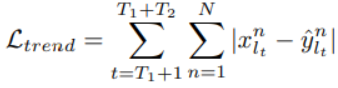
论文首先利用GAT在交通时间序列之间传递消息，然后使用 topk池来选择代表区域或社区的传感器。然后从topk池中采样 S 个传感器，通过 GAT 作为稀疏查询从其他传感器接收最大信息的数量, S 由常数采样因子 e 控制.

除此之外,为了有效地将结构信息注入到完整的GAT中，论文还提出了一种新的图位置编码方法,其有效谱图结构可以表示为

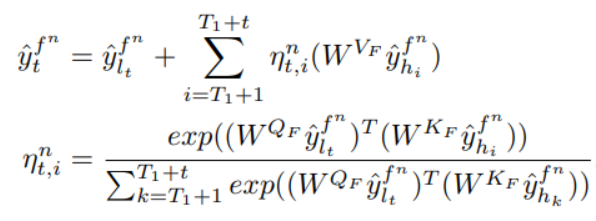
 多监督译码器

该组件的工作原理如下图所示

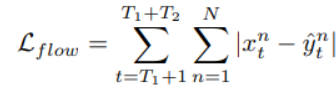
为了将双通道编码器编码的学习表征转换用于预测，论文在时间维度上使用预测器(即全连接神经网络)来推导出未来趋势和事件的表征。然后利用全连接层将趋势投影到一维值上并在模型中通过 *L1* 损失来定义监督并实现



另一方面,论文对趋势中每个时间片的事件进行加权求和，权重由注意力计算，并可以通过反向传播学习，即采用数据驱动的方式自适应判断事件是否被准确预测。最终自适应事件融合可表示为:

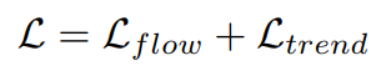


同样采用 L1 损失来定义监督, 可得到下式

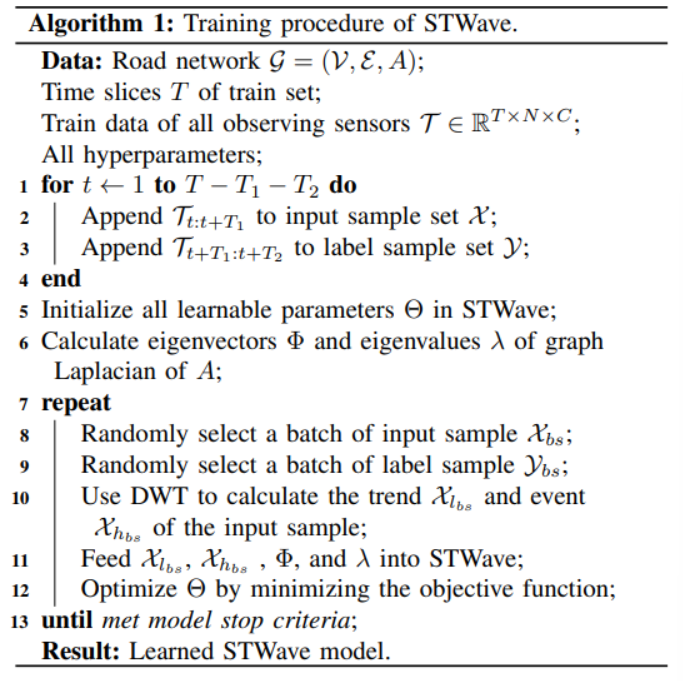


## 目标函数

综合考虑流量预测损失和趋势监控损失，STWave的目标是最小化以下目标函数



由此，可以得到 STWave 的训练过程如下所示



# 三,实验

## 实验环境

论文将提出的模型与 HA, ARIMA 等 15 个模型进行比较

**STWave相关参数设置:**

- 在 200 轮迭代下使用Adam优化器，并设置批大小为64，初始学习率为0.001. 在 20 轮训练后若损失并没有降低, 则学习率降低十分之一

- 特征个数设置为 128, 卷积核大小为 2, DWT层次为 1, 采样因子为 1, 编码器层数设为 2

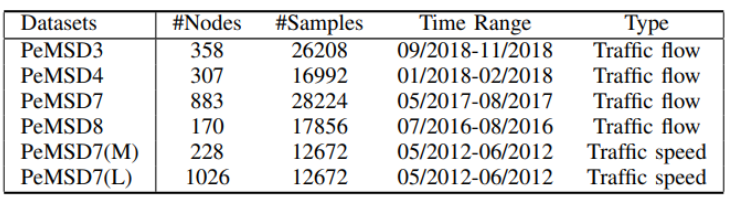
在Python 3.8.10环境下使用PyTorch 1.9.1实现

所有的实验都是在一台配置为 Intel (R) Xeon (R) hGold 6230 R CPU @ 2.10 GHz 和一张 TeslaA 100 GPU 卡上运行Ubuntu 20.04.3 LTS的机器上进行的

源码在https://github.com/LMissher/STWave

## 数据集设置和评价指标

论文采用从加州运输机构收集的六个真实数据集来评估模型.它们每5分钟实时采样一次，在以往的研究中被广泛使用.



论文使用前12个时间片的观测流量来预测接下来的12个时间片，并按时间顺序将它们分为训练集(60%)、验证集(20%)和测试集(20%)

论文采用三个广泛用于所有任务的指标，即平均绝对误差(MAE),平均绝对百分比误差(MAPE)和归一化均方误差(RMSE)

## 实验与结果

接下来，论文就以下 5 个问题设计并进行了实验

- STWave 是否比其它方法优秀？

- STWave中的不同组件(例如ESGAT)如何影响模型性能？

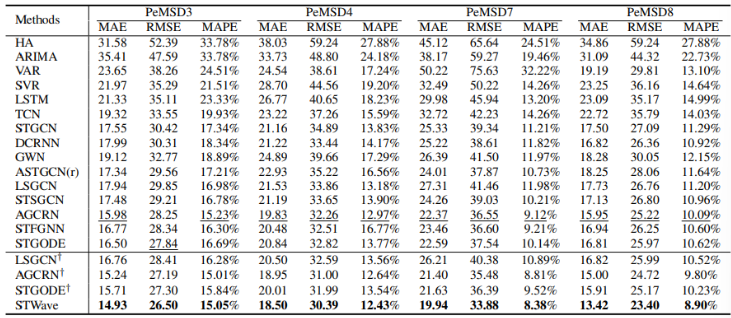
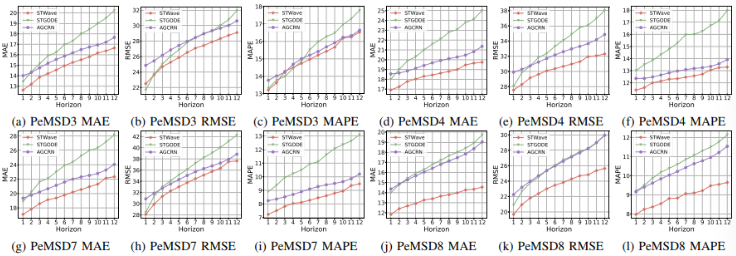
- 框架中的参数如何影响 STWave 的效果？

- 论文提出的 ESGAT 是否有效？

- STWave 的预测结果是否可靠合理？

**\*\*问题一\*\***

论文进行在 6 个数据集上对 15 种模型进行了对比实验，结果如下图所示

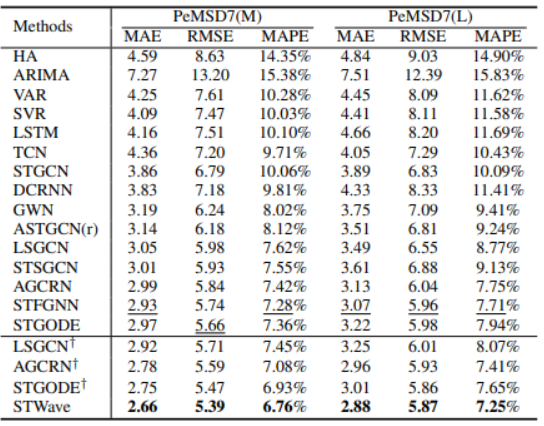
 

可以发现论文提出的方法在所有数据集下都具有最好的表现

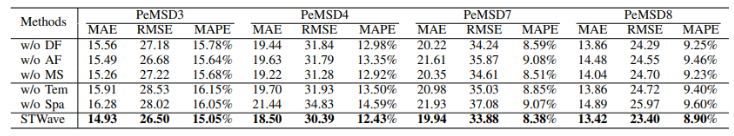
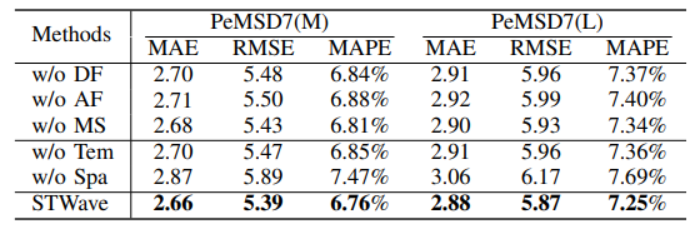
**\*\*问题二\*\***

为了解决这个问题，论文设计了消融实验

为验证解耦层的效果，论文分别用 LSGCN, AGCRN 和 STGODE 替换了 STNet，效果如下图所示



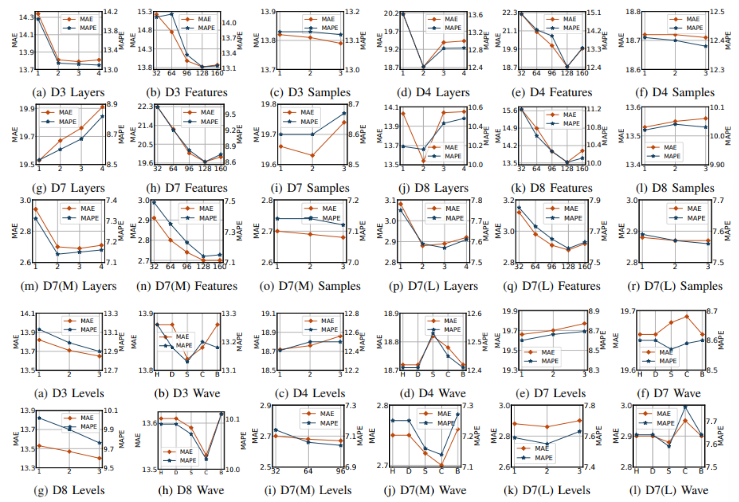
可以发现性能均比 STNet 效果差。除此之外，论文还别对框架中的 DF, AF, MS, Tem, Spa等步骤或组件进行了消融测试，结果如下图所示

可以发现完整的 STWave 具有最优的表现，且可以发现“w/o Spa”的结果远差于“w/o Tem”，表明在多变量交通预测任务中，空间维度比时间维度起着更重要的作用

**\*\*问题三\*\***

对双通道编码器的层数，STWave 中的特征数和 ESGAT 的采样因子的参数进行测试，分别从[1,2,3,4]，[32,64,96,128,160]和[1,2,3]的空间中进行搜索，结果如下图所示



可以发现:

- 模型的性能随着双通道编码器层数的增加而提高，并且当有2层时趋于稳定

- 其次，当特征个数为128时,模型可以达到最佳性能.显然，增加神经网络的规模可以提高表征能力，但过多的特征可能会在学习到的表征中引入噪声

**\*\*问题四\*\***

为了解决这个问题，论文同样针对性地提出了如下多个变种并与 STGODE, LSGCN, STFGNN等模型进行测试

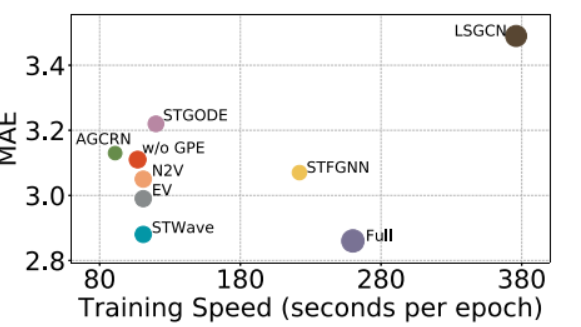
- "Full":STWave中没有查询采样策略来计算所有空间相关性

- “w/o GPE”:不再使用图形位置编码 GPE

- EV”:利用图拉普拉斯特征向量作为GPE

- N2V”:使用Node2vec来学习本地感知的图位置编码

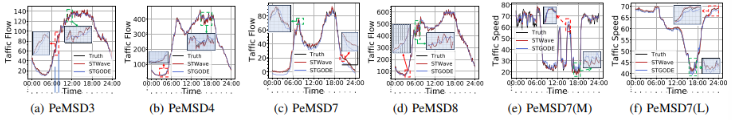
最终结果如下图所示



可以发现 STWave 在训练速度与性能上取得了最好的表现

**\*\*问题五\*\***

论文在所有数据集上设计了案例研究，论文将交通时间序列的一些预测曲线可视化，如下图所示

由于STWave将流量分解为不同的分量，且易预测的趋势不受波动事件的干扰，论文模型比STGODE的预测曲线更精确

# 总结

论文提出了一种新的解纠缠融合的交通预测框架STWave，不遵循复杂的交通端到端建模范式.具体而言,STWave首先通过DWT将交通时间序列分解为趋势和事件，从而使两个独立的分量互不干扰。为此论文提出了一个双通道时空编码器，通过因果卷积、时间注意力和论文的ESGAT来捕获不同时间环境下的稳定时间变化、波动时间变化和空间相关性。此外，利用ESGAT, STWave可以高效提取全局空间下的动态相关性。最后，STWave利用多监督解码器与自适应事件融合模块同时预测稳定趋势和流量时间序列，在六个交通数据集上表现均比现有方法优秀。

本文作者

贾轲

重庆大学计算机科学与技术（卓越）专业本科在读学生，重庆大学Start Lab成员。

主要研究方向：时序数据压缩，数据挖掘

