*A3T-GCN: Attention Temporal Graph Convolutional Network for Traffic Forecasting*

Zhu, Jiawei, Yujiao Song, Ling Zhao and Haifeng Li. “A3T-GCN: Attention Temporal Graph Convolutional Network for Traffic Forecasting.” ArXiv abs/2006.11583 (2020).

Github项目地址：https://github.com/lehaifeng/T-GCN/A3T

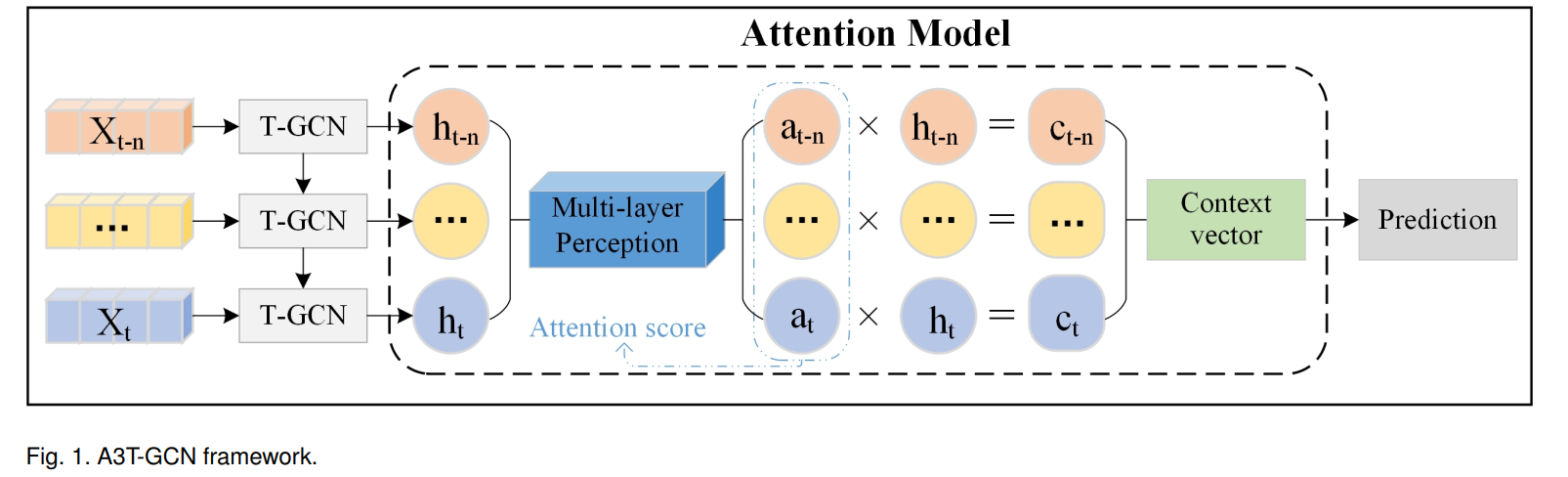
**1 论文动机**

        复杂的时空依赖性，使得交通流预测存在一些挑战：在空间维度上，由于路网的连通性，连接道路之间的交通流密切相关；在时间维度上，受到外部因素的影响，虽然通常相邻时间点之间存在趋势，但较远的过去时间节点的重要性不一定小于最近过去点的重要性。论文提出一种注意时间图卷积网络(A3T-GCN)模型，它由GCNs和GRUs组成，可同时捕获全局时间动态和空间相关性：GRU学习时间序列短时趋势，GCN学习基于道路网络拓扑的空间依赖性；同时，论文引入注意机制调整不同时间点的重要性，并集合全局时间信息，以提高预测精度。

**2 论文创新**

         A3T-GCN在T-GCN的基础上，引入注意力机制，调整不同时间点的重要性，并收集全局时间信息，以提高预测精度，获取交通状态的全局变化趋势，获得更精准的预测精度。

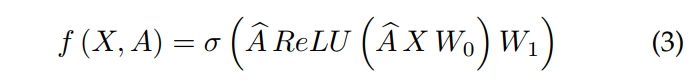
**3 论文模型**



        A3T-GCN模型，由GCN模块、GRU模块和Attention 模块组成。

1）GCN模块

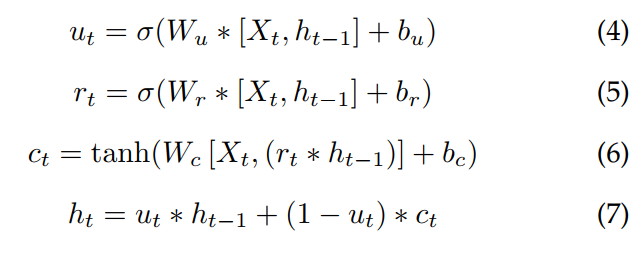
        捕获路网的拓扑特征以获得空间依赖性。两层GCN卷积输出：



         通过确定中心路段与周围路段的拓扑关系，GCNs可以同时编码路网的拓扑结构和路段的属性。

2）GRU模块

         捕获节点属性的动态变化，以获得交通状态的局部时间趋势。（7）代表GRU输出。



       ht-1是t-1的隐藏状态

       Xt是当前时刻的交通速度

       rt是控制前一刻忽略状态信息程度的重置门

       ut是更新门，用于控制进入当前状态的前一时刻的状态信息量。

       ct是存储在当前时刻的内存内容

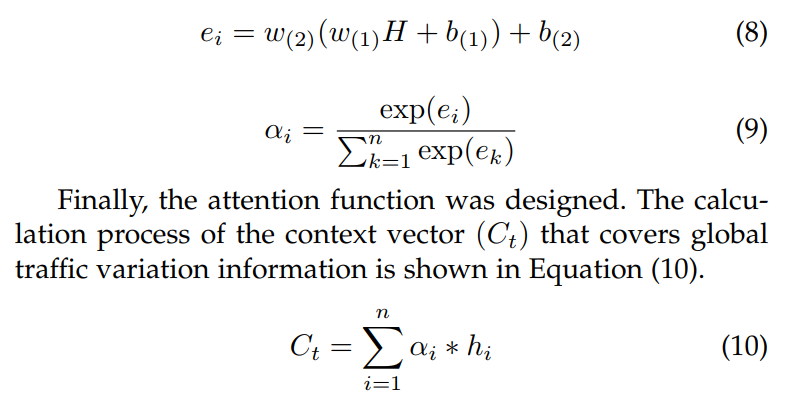
       ht是当前时刻的输出状态。

       GRU通过使用前一时刻的隐藏状态和当前时刻的流量信息作为输入来确定当前时刻的交通状态（速度）。由于门控机制，GRU在捕获当前交通信息时保留了历史交通信息的变化趋势。因此，可以从交通数据中捕捉动态时间变化特征，学习交通状态的时间变化趋势。

3）Attention模块

       用于捕捉交通状态的全局变化趋势，有利于实现准确交通预测。

       时间序列xi(i=1，2，...，n)，其中n是时间序列长度。 软注意力模型的设计过程如下：1）用CNNs（及其变体）或RNNs（及其变体）计算不同时刻的隐态hi(I=1，2，...，n)，并将其表示为H={h1，h2，...，hn}。2）设计了一个评分函数来计算每个隐藏状态的得分/权重。3）设计一个注意函数来计算能够描述全局流量变化信息的上下文向量（Ct）。4）利用上下文向量得到最终的输出结果。



        w(1)、b(1)分别是第一层的权重和偏置

        w(2)、b(2)分别是第二层的权重和偏置

        Ct是上下文向量

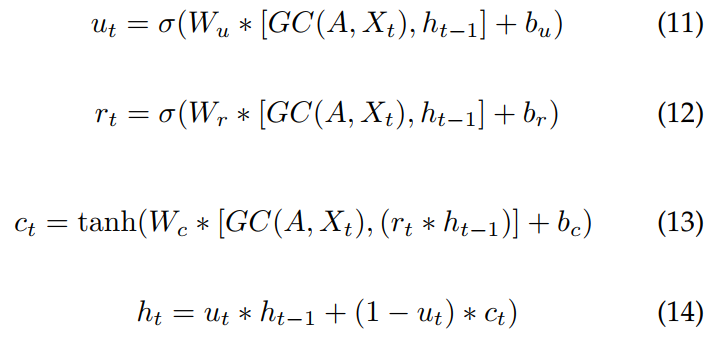
        et是注意力分数。

全过程：

结合GCN和GRU建立T-GCN模型。由式（11）计算得到n个隐态(h)，代表时空特征的：{ht-n，...，ht-1，ht}。

将隐态输入到注意模型中，得到覆盖全局交通变化信息的上下文向量Ct。

预测结果采用全连接层输出。



        ht-1是t-1的输出

        GC代表图形卷积过程

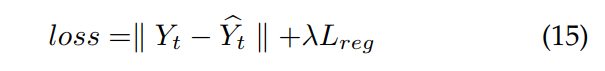
        ut和rt是t时刻的更新门和重置门，

        Ct是当前存储的内容

        ht是t时刻的输出状态

        W和b是训练过程中的权重和偏置，

损失函数设计（和TGCN相同）：



**4 论文实验**

4.1 dataset

1）SZ-taxi。由深圳2015年1月1日-1月31日的出租车轨迹数据组成，选取罗湖区156条主要道路作为研究区域。实验数据包括两部分：一是156\*156的邻接矩阵，它描述了道路间空间关系，每行表示一条道路，矩阵中的值表示道路之间的连接性；二是特征矩阵，它描述了每条道路上的速度随时间的变化，每行代表一条路，每列是不同时段交通速度。每15分钟算一次每条路上的车速。

2）Los-loop。由环形探测器在洛杉矶高速公路上实时采集。选择207个传感器和从2012年3月1日-3月7日的交通速度。每5分钟算一次交通速度。数据由邻接矩阵和特征矩阵组成。利用交通网络中传感器之间的距离计算邻接矩阵。缺失数据，用线性插值方法填充。

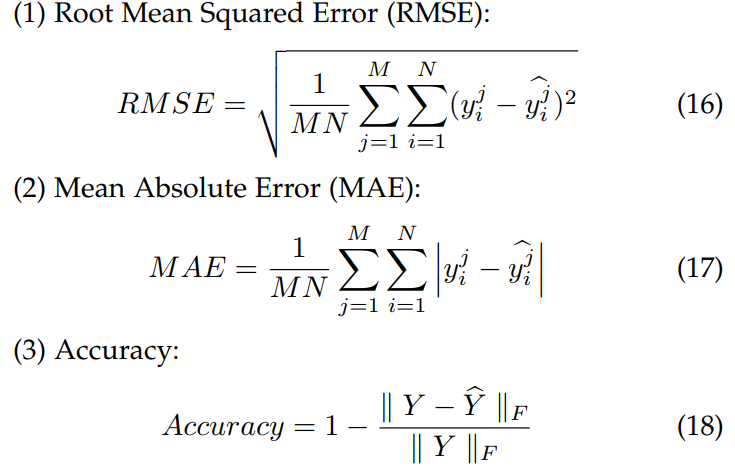
        实验中，输入数据归一化为区间[0,1]。

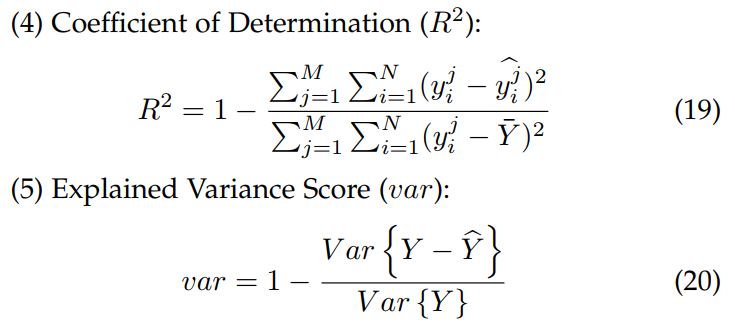
        按时间顺序，前80%的数据作为训练集，后20%作为测试集。预测未来15分钟，30分钟，45分钟和60分钟的交通速度。

4.2 对比基线

        历史平均模型(HA)、ARIMA、SVR、GCN和GRU

4.3 评价指标（5个）





4.4 实验结果

1）实验1

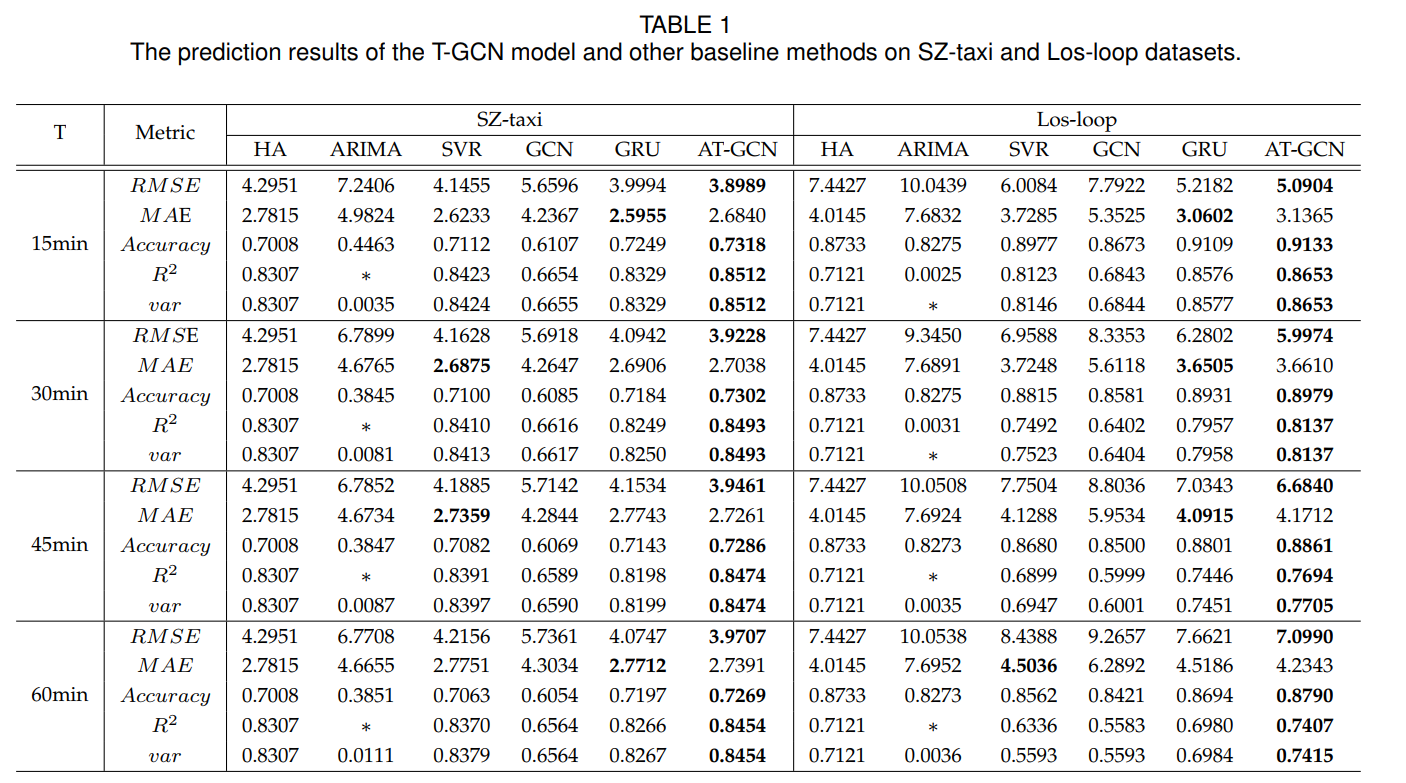
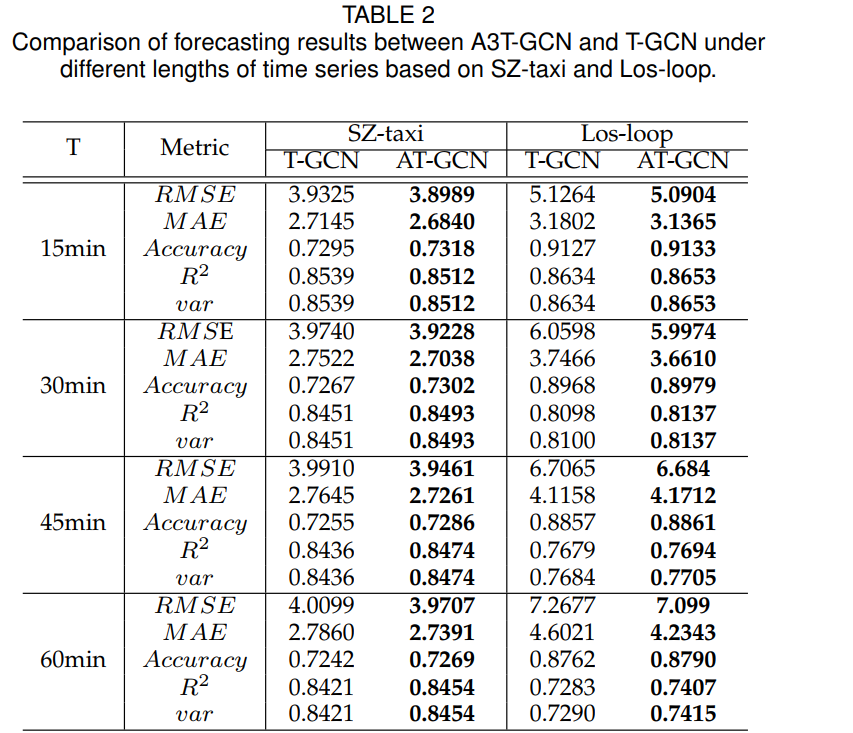


        表1显示不同模型和两个数据集在各种交通速度的预测精度方面的比较。在两个数据集中，A3T-GCN模型可以获得所有指标的最优预测性能，从而证明了A3T-GCN模型在时空交通预测任务中的有效性和优越性。

2）实验2



        为了验证引入注意力机制以捕捉全局变化的有效性，表2列出了A3T-GCN和T-GCN在两个数据集上的实验结果。由表可知，相比T-GCN模型，A3T-GCN模型预测误差更低，且精度更高，对全局特征捕获能力更强。