**1、文章信息**

*Dynamic Graph Convolutional Recurrent Network for Traffic Prediction: Benchmark and Solution*

面向交通预测的动态图卷积循环神经网络：基准和方法

Li, F., Feng, J., Yan, H., Jin, G., Jin, D., & Li, Y. (2021). Dynamic Graph Convolutional Recurrent Network for Traffic Prediction: Benchmark and Solution. ArXiv, abs/2104.14917.

这是清华大学发表在arxiv上的一篇文章，目前已经向计算机顶级期刊TKDE投稿，对于交通流量预测这个方面论文研究有较大学习和借鉴意义。

**2、摘要**

准确的交通预测是智慧城市建设的基础。虽然目前已有的工作已经提出了多种时空建模方法，但它们都忽略了路网中位置间相关性的动态特性。同时，大多数基于循环神经网络(RNN)的工作由于其循环操作而效率低下。此外，过往的工作缺乏不同方法之间在同一数据集上的公平比较。为了解决上述问题，本文提出了一种新的流量预测框架——动态图卷积循环网络(DGCRN)。在DGCRN中，超网络被设计为利用和提取节点属性的动态特征，而动态滤波器的参数在每个时间步生成。通过对节点嵌入进行过滤生成动态图，并与预定义的静态图集成。据我们所知，我们首先在每个时间步采用生成方法对动态图的精细拓扑进行建模。此外，为了提高效率和性能，我们采用了一种训练策略，通过限制前向和后向传播过程中译码器的迭代次数。最后，为公平比较和进一步研究提供可重现的标准化基准和全新的代表性流量的北京数据集。在三个数据集上的大量实验表明，我们的模型始终优于15个基线模型。这篇论文提出的模型和其他baseline方法的pytorch实现也已经全部开源，代码地址：https://github.com/tsinghua-fib-lab/Traffic-Benchmark（内容在逐渐完善更新中。）

**3、动机**

由于交通数据在时空维度上具有很强的动态相关性，因此建立动态的、非线性的时空相关性模型对准确的交通预测至关重要。不同位置之间的空间相关性是高度动态的，由实时交通状况和交通网络的拓扑结构决定。同时由于非线性变化和周期性的混合，时间依赖性也很复杂。由于时空关联的复杂性，我们需要应对以下挑战:

1).交通图拓扑的动态特性需要进一步建模。预先定义的邻接矩阵和自适应学习的邻接矩阵都是本质上是随时间变化的静态邻接矩阵，不足以反映现实路网拓扑的动态特性。

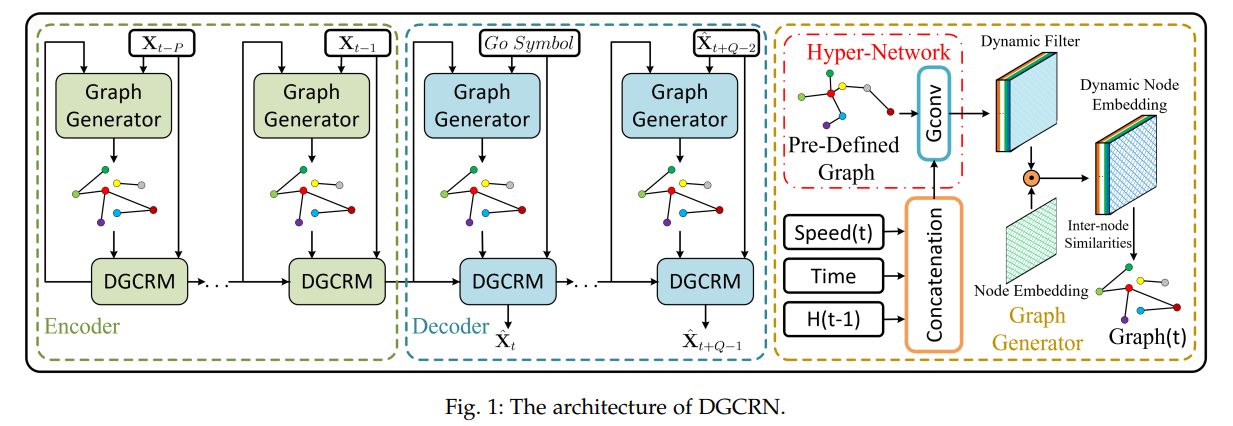
2).基于静态距离的图和基于动态属性的图，分别从不同的角度描述了交通网络的拓扑结构，从而为模型获取空间依赖性提供了更广阔的视域。但是大多数过去的工作在保持效率和避免过度平滑的同时，并没有将它们融合在一起。

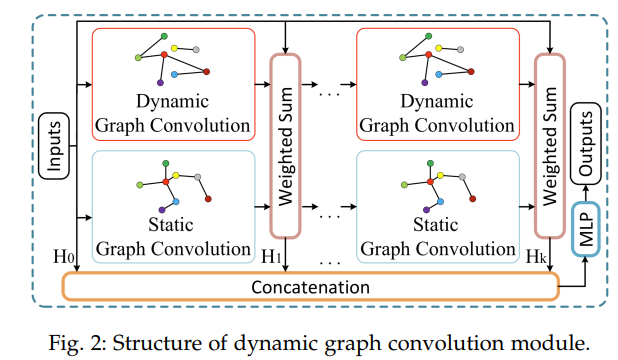
3).RNN及其变体虽然广泛应用于时间序列预测，但其训练速度受到内部递归运算的限制，阻碍了序列到序列等结构在交通预测任务中的应用。

4).随着交通预测领域的快速发展，提出的模型和交通数据集的数量也在不断增加，其中模型是在不同的数据集上进行评价，并进行相应的实验设置。这使得在模型之间进行公平比较变得困难，阻碍了该领域的发展。

**4、模型**

RNN的循环操作带来了动态知识，然而这些动态知识并没有被充分利用来捕获动态的时空关联。基于此思想，我们设计了基于序列到序列结构的DGCRN，包括编码器和解码器，如论文中的图1所示。通过DGCRN的递归运算同步生成每个时间步的动态邻接矩阵，其中两个图生成器分别用于编码器和解码器。然后，使用生成的动态图和预定义的静态图进行图卷积。接下来我们使用加权和融合每层有跳跃连接的不同图卷积的结果，如论文中的图2所示。最后，我们将传统的门控循环单元(GRU)中的所有全连接层替换为动态图卷积模块，从而得到动态图卷积循环模块(DGCRM)。





具体来说，DGCRN由图生成模块和动态循环图卷积模块两个主要部分组成。

(1)图生成模块

城市交通状况是复杂的，受不同时空关联的影响。从动态的角度来描述交通网络是很自然的。设计了一种用于动态图生成的超网络。DGCRN的核心是基于当前和历史信息在时间步长t时更新动态邻接矩阵DAt。在每个时间步长，速度Vt、时间t和隐态Ht−1串联为超网络的输入：

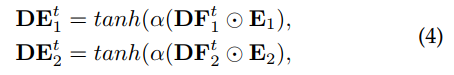


其中，It为动态节点特征，将其作为图卷积模块的输入:



这个图卷积操作中采用的是预定义的邻接矩阵A，该邻接矩阵包含了节点之间静态的基于距离的关系，可用于对动态节点状态进行消息传递过程。我们称之为超网络，该图卷积的输出是一个动态滤波器张量DFt。我们在DFt和可学习的静态随机初始化节点的嵌入E之间采用逐位相乘运算。

在实际应用中，我们使用两个超网络分别生成源节点嵌入和目标节点嵌入的两个动态滤波器，分别表示为:



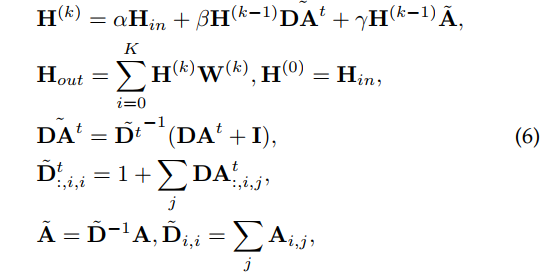
我们接下来利用节点之间的相似性来计算每一个时间步t的动态邻接矩阵：



综上所述，我们将RNN的隐藏状态加入到图生成器的输入中，将RNN的迭代与动态图的生成同时无缝地结合起来。同时，超网络内部的图卷积应用了节点间的消息传递，使图的生成更加有效，从而充分利用了动态信息。动态邻接矩阵生成模块作为图卷积的重要组成部分，可以隐式学习路网的动态表示，为静态的基于距离的邻接矩阵提供了有效的补充。

(2)      动态图卷积循环模块

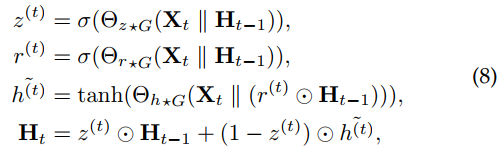
基于距离的静态图和基于节点属性的动态图从不同的角度反映了节点间的相关性。为了使模型对交通网络有更大的视界，在部署图卷积时将动态图与预定义图相结合，从而提高了交通预测的性能。具体地说，我们在每个图卷积层上采用输入图信号、预定义静态图和动态图的图卷积结果的加权求和。k-hop动态图卷积模块可表示为:



特别地，我们还采用双向图卷积来很好地利用有向图，记为:

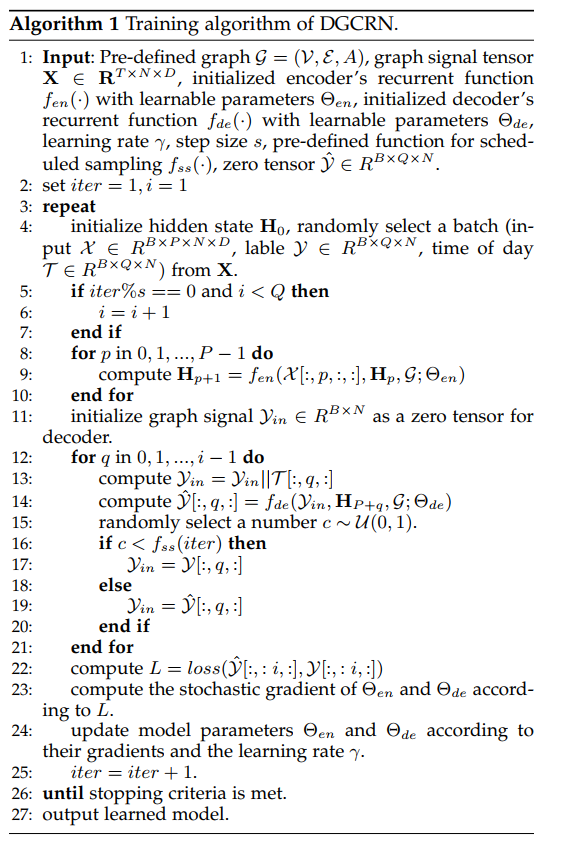


考虑到，RNN具有强大的序列相关性建模能力，GRU进一步提高了RNN在长期时间建模和避免梯度消失方面的能力。将GRU中的矩阵乘法替换为动态图卷积模块，得到DGCRM，表示为:



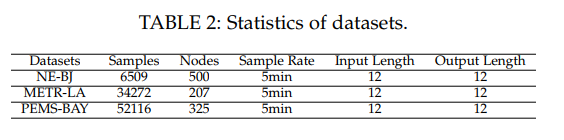
传统RNN的循环操作带来了随序列长度增加而增加的时间和内存消耗。因此，有必要使RNN执行尽可能少的重复操作。同时，如果先对基于RNN的解码器进行短期预测训练，并对早期时间步长有较强的记忆，那么经过较长的序列连续训练后，解码器可以在较长的范围内获得更好的性能。基于这些思想，我们对RNN相关模型采用了一种通用有效的训练策略，即课程学习。在反向传播之前的训练过程的前向传播中，我们不计算解码器的所有步骤，只计算第i步。随着训练过程的进行，i不断增大，直到达到最大，即预测序列的长度。这样，在训练过程中，模型无需对所有的时间步长都采用前向和后向传播，可以显著减少时间消耗，并在早期节省GPU内存。更重要的是，由于早期时间步骤的良好基础，性能也可以从该策略中受益。我们还发现RNN的调度采样可以很好地配合课程学习，进一步提高了性能。我们认为，这种训练方法可以很容易地应用于其他基于循环的解码器，从而提高时间序列预测任务的效率和性能。

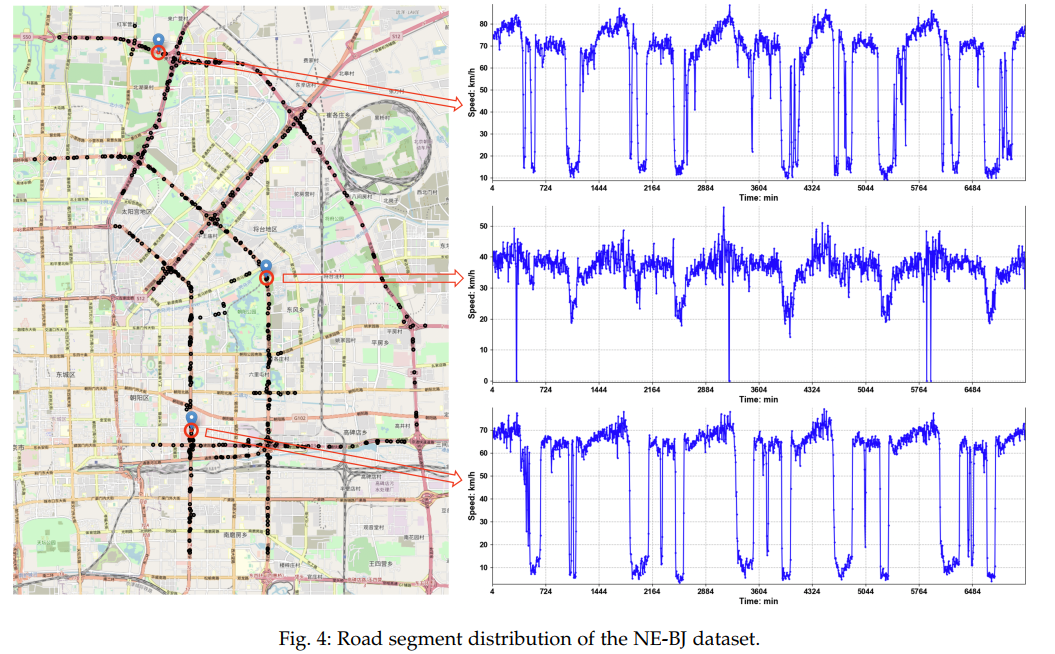
(3)      DGCRN算法描述

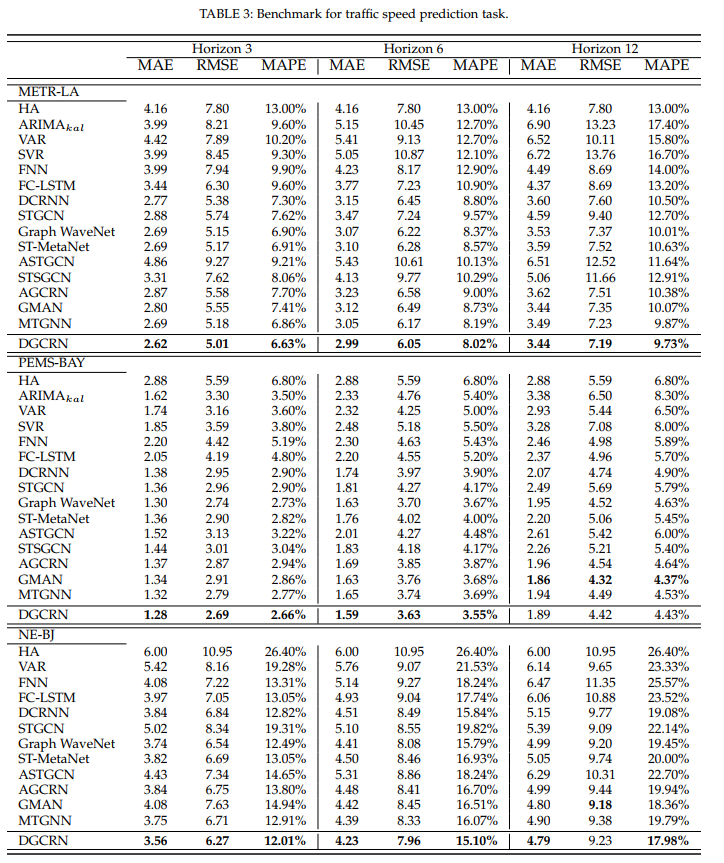


**5、实验结果**

论文中使用METR-LA, PEMS-BAY和NE-BJ数据集将DGCRN与15种baseline方法进行了比较。论文中使用包括绝对百分比误差（MAPE），平均平均误差（MAE）和均方根误差（RMSE）的三个指标来评估所有方法。其中， 以往提出的交通数据集，例如METR-LA, PEMS-BAY，要么是在交通速度相对较高且较为简单的高速公路上根据车速建立的，要么不能充分反映复杂的城市交通状况。为了克服这些不足，使交通预测任务更具挑战性，我们在2020年7月的工作日发布了由腾讯地图导航数据中收集和提取的NE-BJ数据集。速度的单位是km/h。该数据集选取了北京东北角地区交通拥堵发生率较高的500条主干道路段。该数据集能够充分反映北京市中心城区的交通状况，具有进一步的研究价值。在三个数据集上的对比实验结果如下：

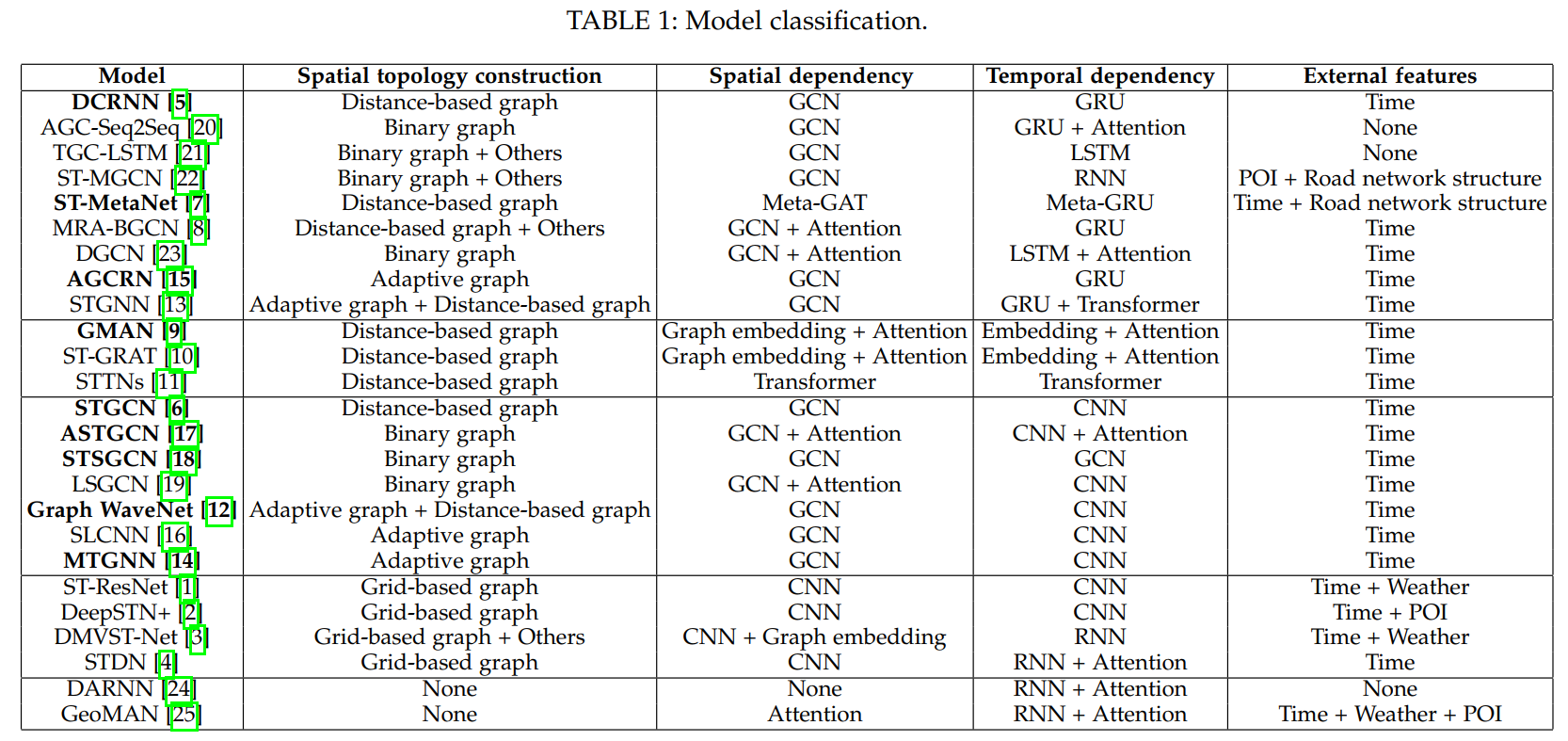






**6、创新点**

本文提出了一种自适应逐级生成动态邻接矩阵的超网络，其中动态节点特征的消息传递提高了图的生成效率。然后将静态图和动态图集成到图卷积模块中，显著提高了性能。对于更具挑战性的预测任务，本文中还发布了一个具有代表性的与拥堵相关的流量数据集。最后提出了一个基于三个公共数据集和15个基线的开源基准，以进行公平的比较和进一步的研究。这是一篇介于研究论文和综述之间的文章，即提出了更优的模型也对在交通流量预测领域过往的工作做了一个很好的总结，例如文中的表格对各种知名的模型做了详细的分类：



这篇论文承接了过往的工作，也开拓了未来再交通流量预测这个研究方向继续研究的思路，还为后人的研究提供了便利，是交通预测领域的一篇良心之作，具有很高的学习和研讨价值。