



A graph-attention based spatial-temporal learning framework for
tourism demand forecasting

基于图注意的旅游需求预测时空学习框架

主要内容

Main Contents

1

摘要

2

引言

3

方法

4


模型

5

实验

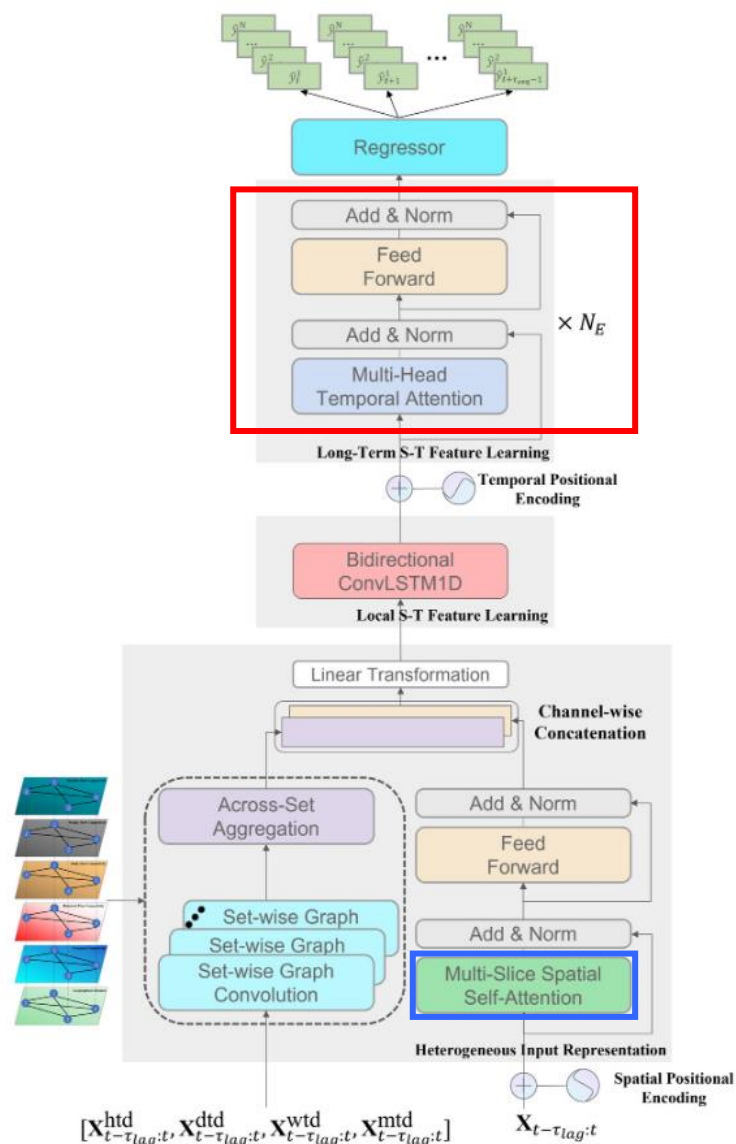
6

结论

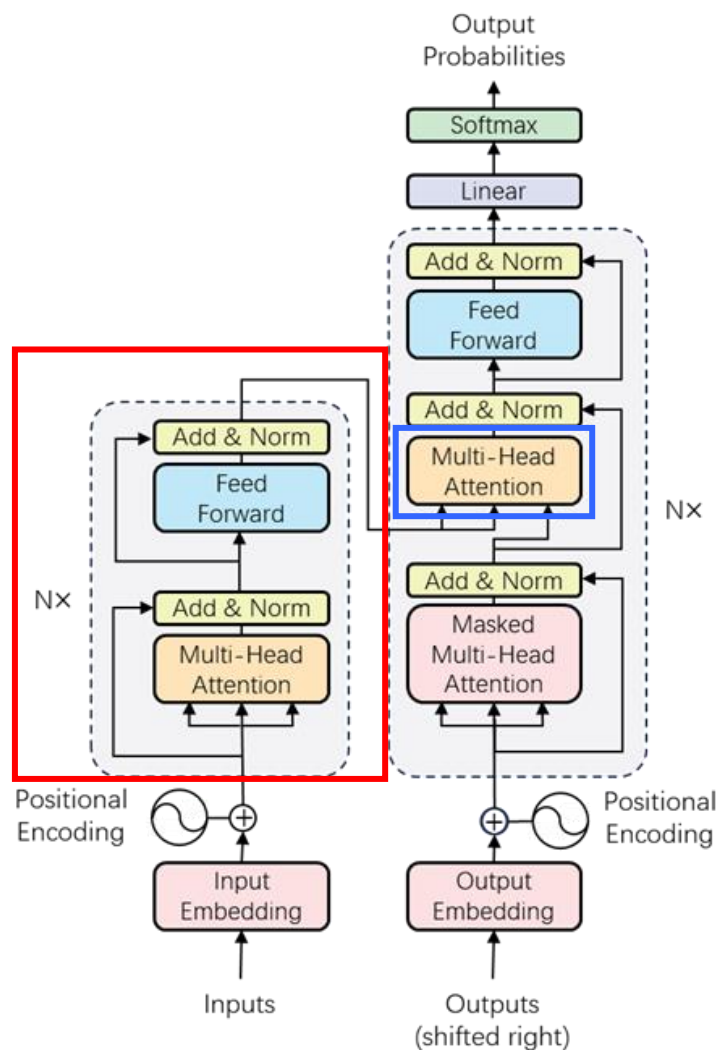


摘要






HSTGANet



Transformer

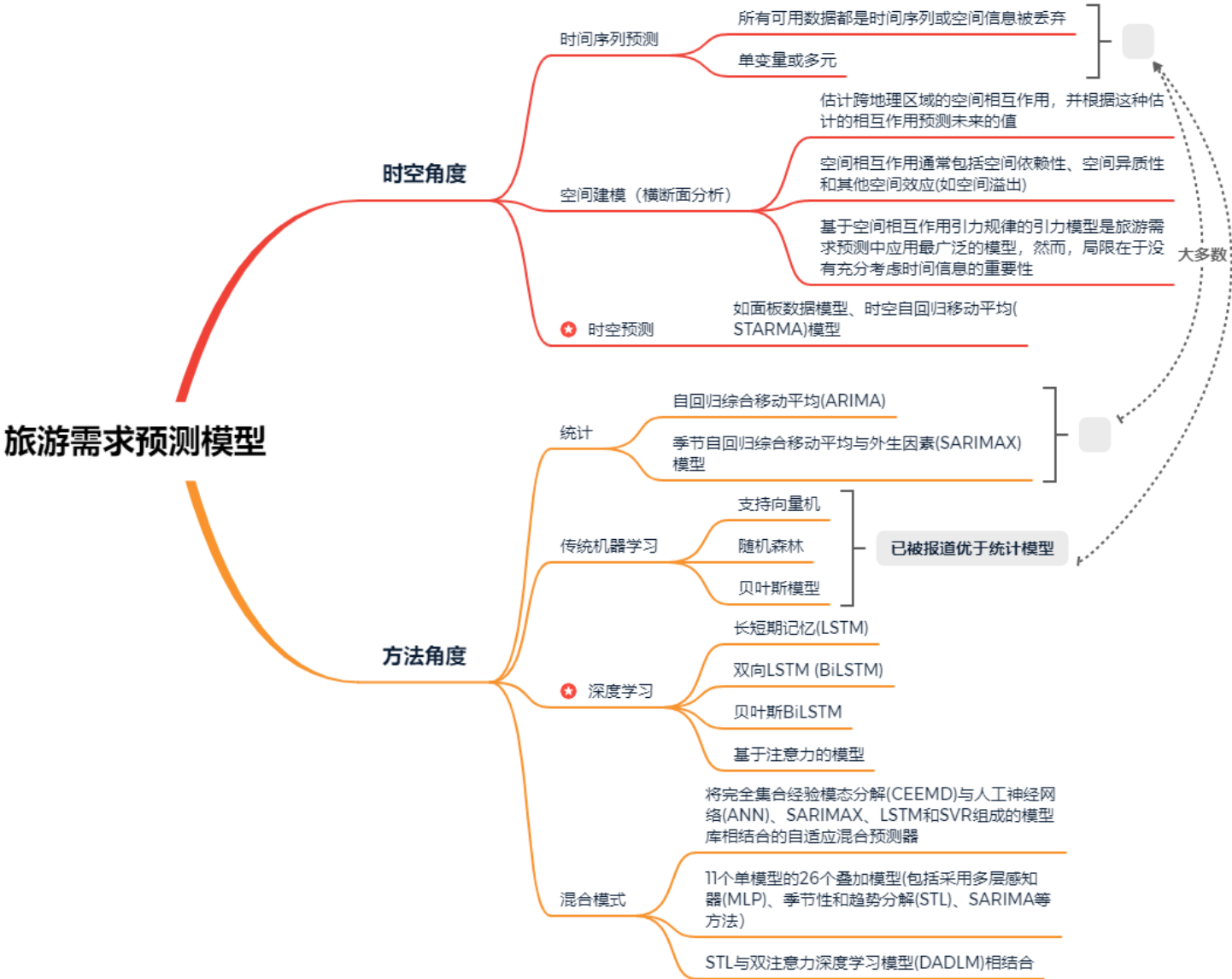
现有的时空旅游需求预测模型只探索区域之间预先指定的静态空间连接，而不考虑多个或动态的空间连接。

本文提出了一个异构时空图注意力网络（称为HSTGANet），这是第一项将图卷积与自注意力相结合的工作，前者用于建模显式动态空间连接，后者用于建模隐式动态空间联系，用于一般的时空预测和旅游需求预测。



引言





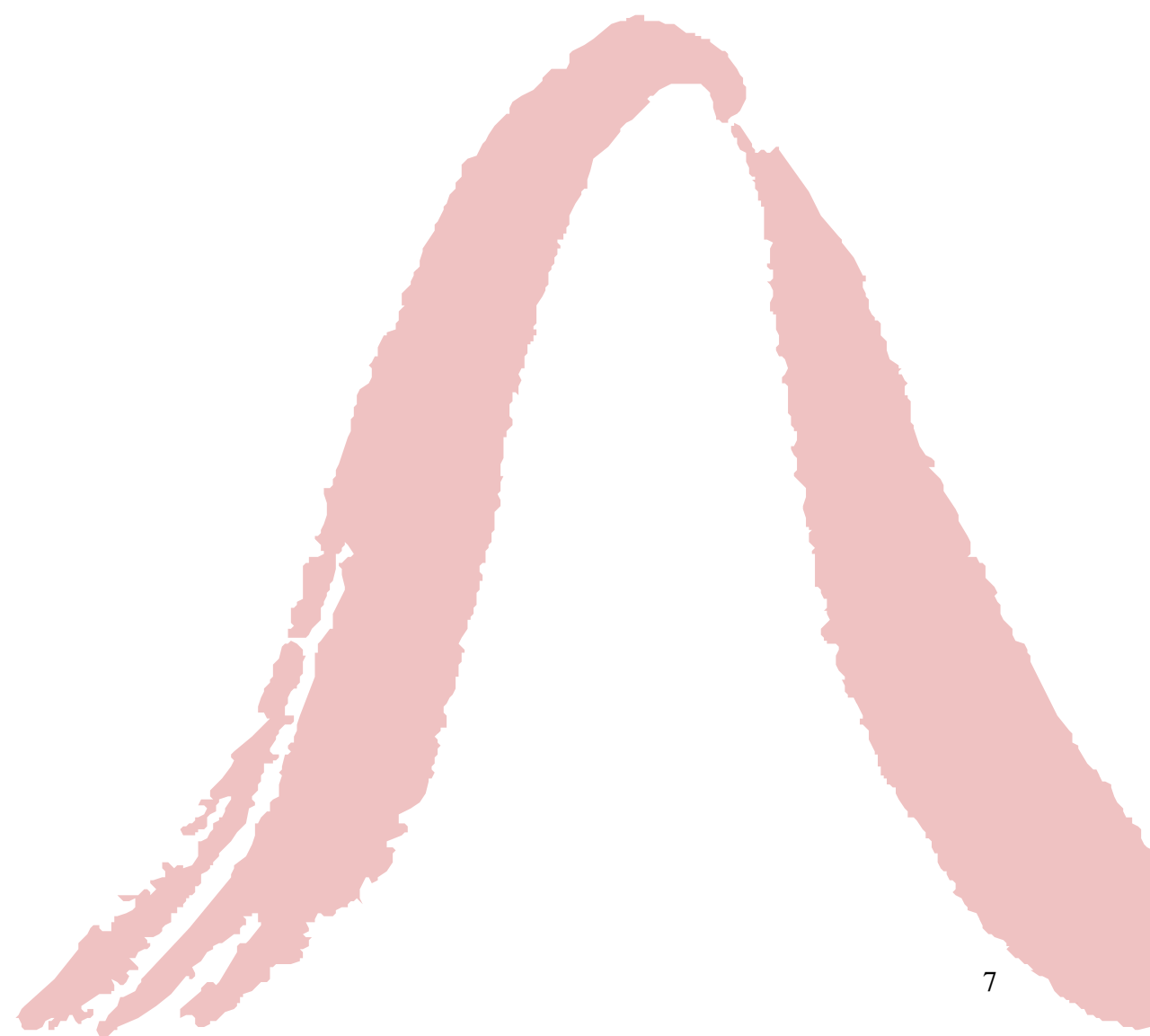
效益：帮助游客正确地安排他们的时间表，并帮助旅游从业者适当地分配旅游资源。

基于深度学习的时空预测：图神经网络通过图节点和多头注意网络之间的消息传递来捕获图的依赖关系。目前还没有研究使用深度学习模型研究时空旅游需求预测，特别是使用图神经网络和多头注意力网络。

我们将具有多个区域的景区的时空数据组织为具有多个节点和多维边的权重动态多维图。



方法



旅游文献研究的空间数据有两种:横断面数据和面板数据。横断面数据是通过在同一时间点或同一时间段观测多个区域来收集的[6]。面板数据是通过扩展具有时间维度的截面数据来观测的,其中包含了空间和时间信息。

大多数现有的时空旅游需求预测工作要么使用面板数据模型,要么使用时空自回归模型。

面板数据模型

$$\hat{\mathbf{y}}_t = \lambda \mathbf{W}_s \mathbf{y}_{t-1} + \boldsymbol{\mu} + \boldsymbol{\varepsilon}_t, \quad (1)$$

$\hat{\mathbf{y}}_t$: 所有地区在时间t的旅游需求预测向量

\mathbf{y}_{t-1} : 所有地区在时间t-1的旅游需求向量

λ : 系数

$\mathbf{W}_s \in \mathbb{R}^{N \times N}$: N个地区之间空间联系的空间加权矩阵

$\boldsymbol{\mu}$: 地区特定影响的向量

$\boldsymbol{\varepsilon}_t$: 误差项的向量

在[7]中,提出了三种不同的空间加权矩阵:基于突发事件、基于距离和基于流量的矩阵,以预先指定空间加权矩阵 \mathbf{W}_s 。

时空自回归模型通过将空间滞后纳入模型的自回归项和移动平均项来扩展传统的自回归模型，可以公式化为

$$\hat{y}_t = \sum_{i=1}^p \sum_{j=1}^m \varphi_{ij} \mathbf{W}_s^{(j)} \mathbf{y}_{t-i} + \sum_{k=1}^q \sum_{l=1}^n \theta_{kl} \mathbf{W}_s^{(l)} \boldsymbol{\varepsilon}_{t-k} + \mu + \boldsymbol{\varepsilon}_t, \quad (2)$$

p和m: 分别是自回归项在时间和空间上的最大阶数

q和n: 分别是移动平均项在时间或空间上的最高阶数

φ 和 θ : 分别是系数

$\mathbf{W}_s^{(j)}$ 和 $\mathbf{W}_s^{(l)}$: 分别是空间加权矩阵的第j阶和第l阶

在[7]中，三个空间加权矩阵（基于突发事件的矩阵、基于距离的矩阵和基于流量的矩阵）被应用于构建时空自回归模型。在[6]中，提出了一种基于k近邻的空间加权矩阵。

尽管这些模型考虑了跨区域的空间连接并将其纳入建模中，但所提出的空间加权矩阵是预先指定的静态矩阵。我们提出了时空信息表示，同时考虑多个动态空间连接。

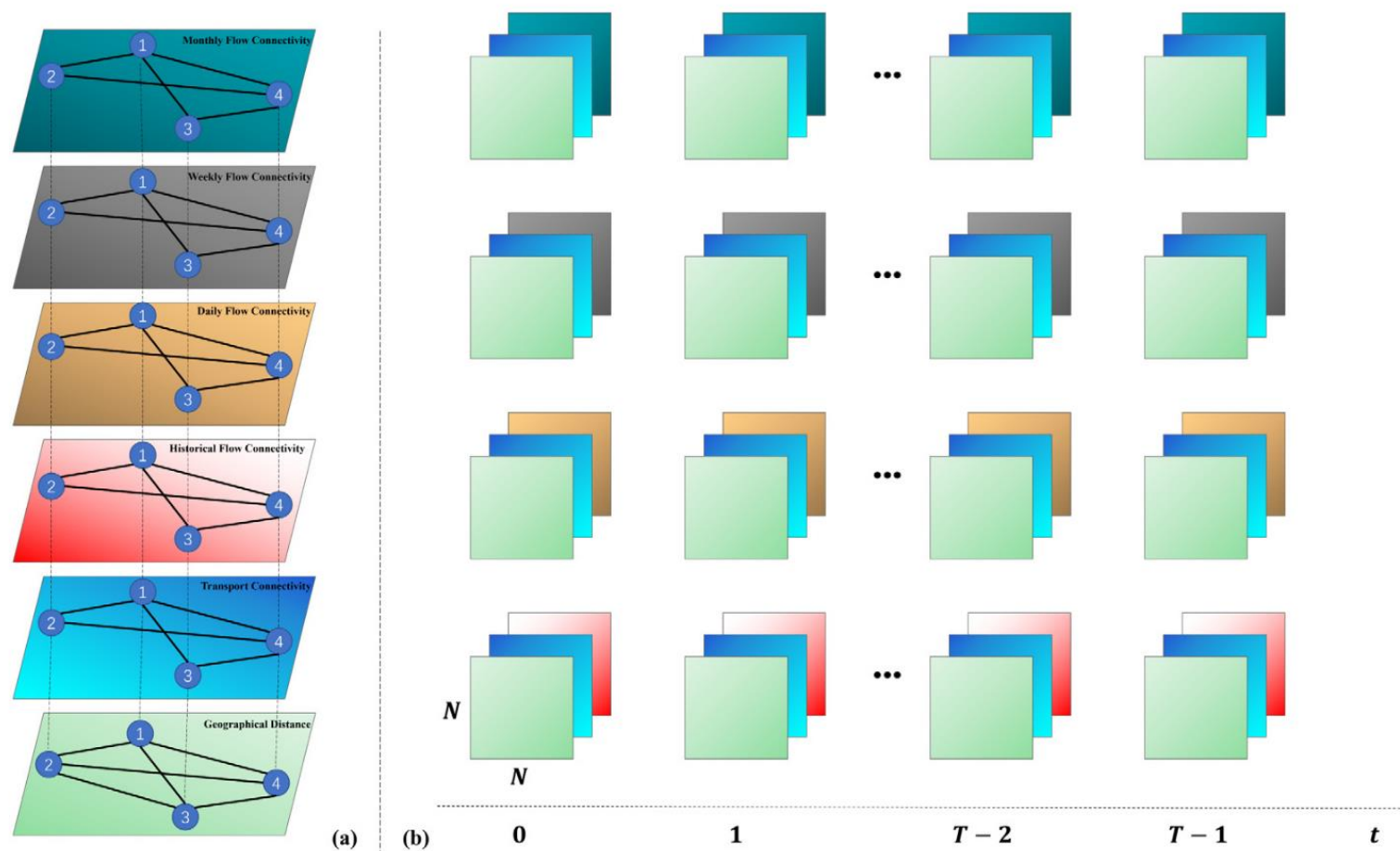


图1：提出了加权动态多维图及其相应的加权邻接矩阵：（a）图拓扑；以及（b）用于学习历史和周期性空间连接的4组加权邻接矩阵。 N 是区域的数量， T 是序列数据的时间长度。

考虑了与每个区域相关的直接相互作用（例如地理距离、旅游流量，可以明确推断）和解释变量（例如，当地旅游量、天气，只能隐含推断）。

考虑具有 N 个区域的风景区 $S=\{s_1, s_2, \dots, s_N\}$ （如图1(a)所示）。权重动态多维图表示为 $G=(V, E)$ 。 G 由表示多个区域的一个节点集 $V=\{V_1, V_2, \dots, V_N\}$ 和表示跨区域的多个显式交互的 d 个不相交子集 $E=\{E_1, E_2, \dots, E_d\}$ 组成。

空间-时间序列 X 表示为

$$\begin{aligned} X &\in \mathbb{R}^{T \times N \times c} \\ &= X^{\text{hol}} \cup X^{\text{wea}} \cup X^{\text{si}} \\ &\quad \cup X^{\text{htd}} \cup X^{\text{dtd}} \cup X^{\text{wtd}} \cup X^{\text{mtd}}, \end{aligned} \quad (3)$$

其中， c 是 X 的通道数（这里 $c=7$ ）， N 是区域数， T 是序列数据的时间长度。

holiday x^{hol} weather x^{wea} search index information x^{si} 解释变量

historical x^{htd} daily x^{dtd} weekly x^{wtd} monthly x^{mtd}

历史

周期性

为进一步的图运算，给出了G中每个维度的加权邻接矩阵。具体而言，有六种明确的相互作用：历史、每日、每周和每月的交换流量连通性、运输连通性和地理距离。

$$\mathbf{A}_{(i,j),t}^{\text{his}} = \frac{Q_{(i,j),t}^{\text{his}} + Q_{(j,i),t}^{\text{his}}}{2}, \quad (4)$$

$$\mathbf{A}_{(i,j),t}^{\text{daily}} = \frac{Q_{(i,j),t}^{\text{daily}} + Q_{(j,i),t}^{\text{daily}}}{2}, \quad (5)$$

$$\mathbf{A}_{(i,j),t}^{\text{weekly}} = \frac{Q_{(i,j),t}^{\text{weekly}} + Q_{(j,i),t}^{\text{weekly}}}{2}, \quad (6)$$

$$\mathbf{A}_{(i,j),t}^{\text{monthly}} = \frac{Q_{(i,j),t}^{\text{monthly}} + Q_{(j,i),t}^{\text{monthly}}}{2}, \quad (7)$$

其中 $Q(i, j)$ 和 $Q(j, i)$ 分别指示从 V_i 到 V_j 和从 V_j 到 V_i 的流。下标A和Q表示这些都是动态项。

$$\mathbf{A}_{i,j}^{\text{trans}} = \begin{cases} 1 & , \text{ if } v_i, v_j \text{ have accessible transport} \\ 0 & , \text{ otherwise.} \end{cases} \quad (8)$$

$$\mathbf{A}_{i,j}^{\text{dis}} = \text{dis}(L_i, L_j), \quad (9)$$

其中 L_i 和 L_j 分别是 V_i 和 V_j 的位置，而 $\text{dis}(\cdot, \cdot)$ 是一个函数，其输出是两个位置之间的测地线距离。

\mathbf{A}^{dis} and $\mathbf{A}^{\text{trans}}$ 不会随时间变化。

交通连通性和地理距离会影响区域间的空间联系。因此，我们将这两个图维与表示流连通性的四个图维中的一个图维结合起来，形成四个图维集，然后分别用于建模四个历史和周期高维时空特征。四个图维集对应的邻接矩阵如图1(b)所示。

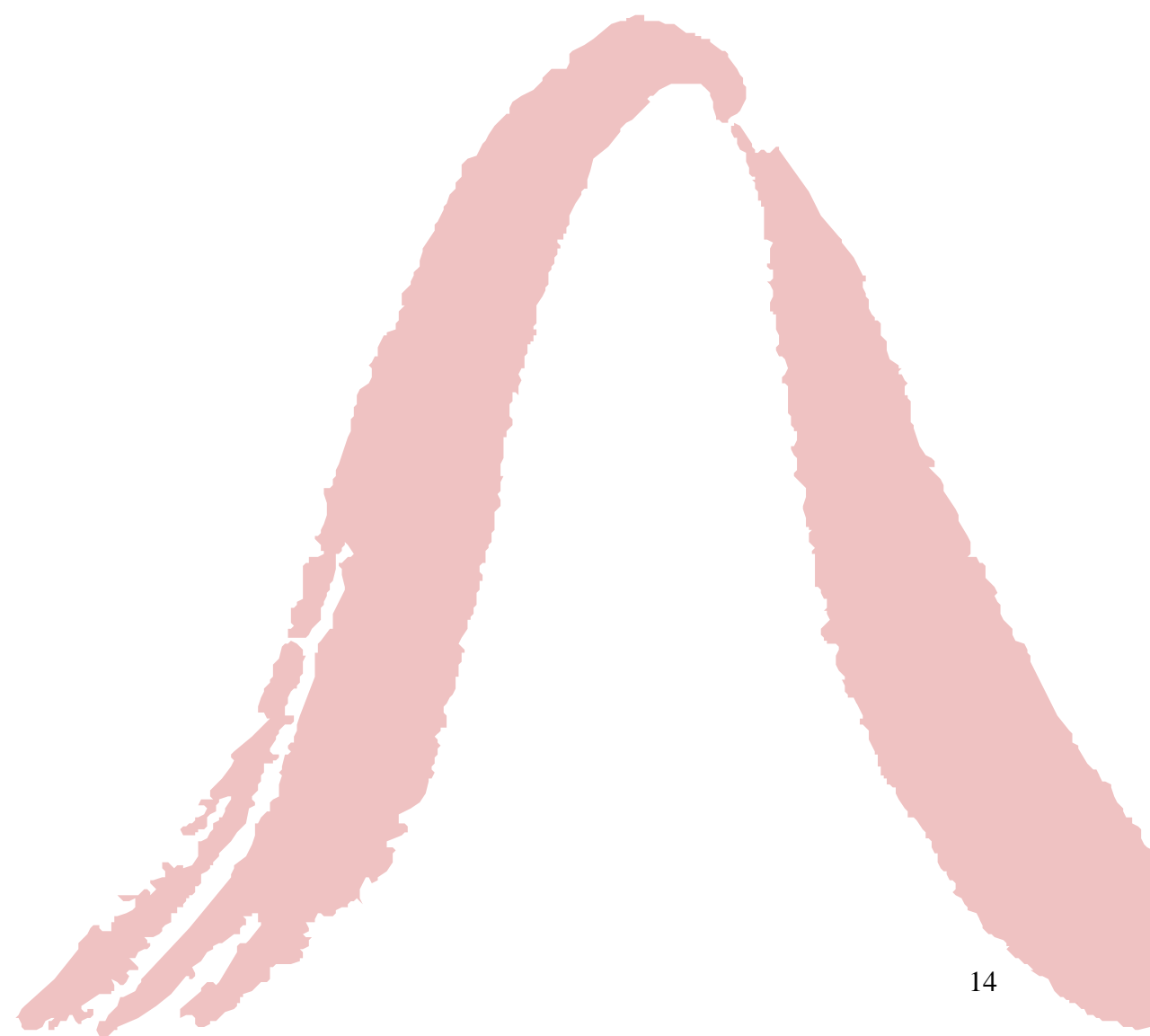
具有 τ_{lag} -step 时空输入的 τ_{seq} -step 旅游需求预测可以公式化为

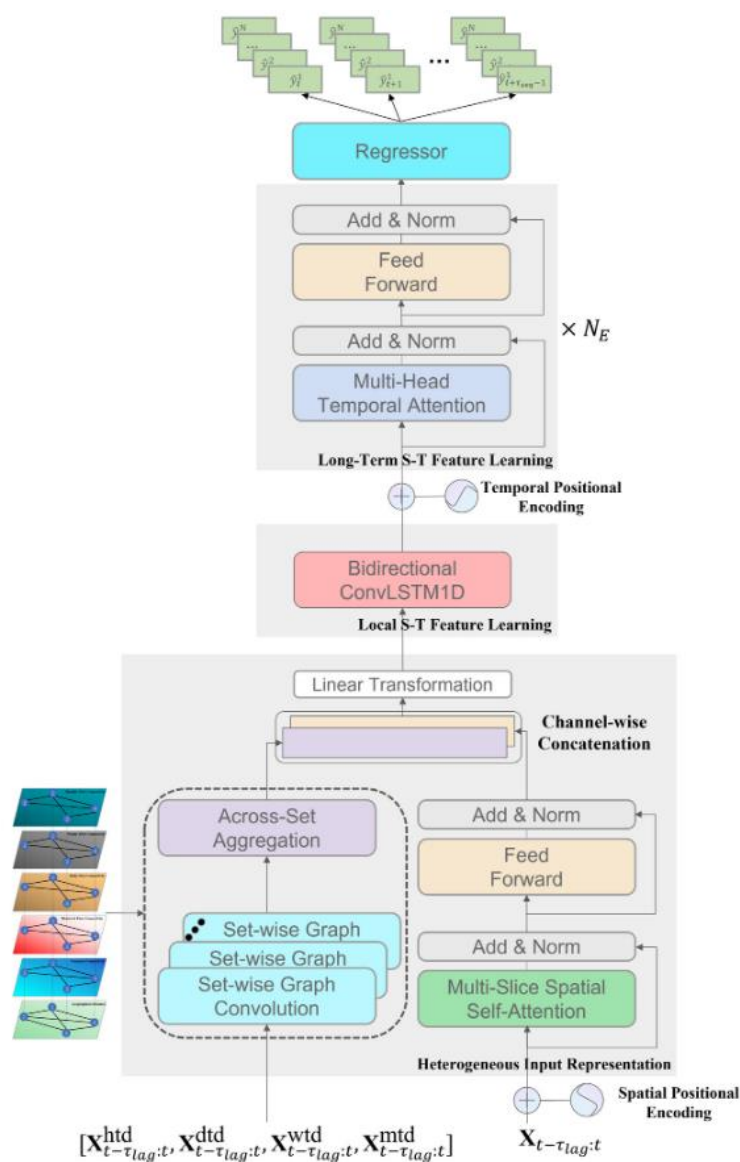
$$\{\hat{\mathbf{y}}_{t:t+\tau_{seq}}^i | i = 1, 2, \dots, N\} = F_{\theta}(\mathbf{X}_{t-\tau_{lag}:t}, \mathcal{G}), \quad (10)$$

$\hat{\mathbf{y}}$ 是旅游需求预测， F 表示所提出的HSTGANet， θ 表示可学习参数。



模型





HSTGANet由异构输入表示层、局部时空特征学习层、NE时空特征学习块和最终预测层组成。

利用异质输入表征层有效学习反映动态空间联系的隐藏表征，利用局部时空特征学习层和长期时空特征学习块分别学习有助于预测时空旅游需求的局部和长期时空特征。

异构输入表示层

时空信息由时空序列 \mathbf{X} 和多维图 \mathbf{G} 表示。跨区域的空间连接可以由图结构中的多个图维度集和相应的通道显式推断，也可以由时空序列隐式推断。

为了从输入的两个部分中提取嵌入复杂时空特征的隐藏表示，我们提出了一种多维图嵌入层和多层空间自注意力层来形成异构输入表示层来提取时空特征的两个部分。

多维图嵌入层

多维图嵌入层提取集合表示 $\mathbf{h}_{t-\tau_{lag}:t}^{\text{his}}$, $\mathbf{h}_{t-\tau_{lag}:t}^{\text{daily}}$, $\mathbf{h}_{t-\tau_{lag}:t}^{\text{weekly}}$, and $\mathbf{h}_{t-\tau_{lag}:t}^{\text{monthly}}$, 分别表示历史、每日、每周和每月的空间连接模式。

然后从所有集合中聚集这些表示形成最终的图嵌入(即从显式动态空间连接中学习到的最终隐藏表示 $\mathbf{h}_{t-\tau_{lag}:t}^{\text{exp}}$)。

为了简化以后的描述, 集合表示统一表示为 $\mathbf{h}_{t-\tau_{lag}:t}^{\text{set}}$

学习每个集合表示的输入, 用 $\mathbf{X}_{t-\tau_{lag}:t}^{\text{set}}$ 表示

$$\mathbf{X}_{t-\tau_{lag}:t}^{\text{set}} = \begin{cases} \mathbf{X}_{t-\tau_{lag}:t}^{\text{htd}} & , \text{ if set = his} \\ \mathbf{X}_{t-\tau_{lag}:t}^{\text{dtd}} & , \text{ if set = daily} \\ \mathbf{X}_{t-\tau_{lag}:t}^{\text{wtd}} & , \text{ if set = weekly} \\ \mathbf{X}_{t-\tau_{lag}:t}^{\text{mtd}} & , \text{ if set = monthly.} \end{cases} \quad (11)$$

每个集合表示都是通过图卷积操作获得的维表示的集合。在t处的聚合表示：

$$\mathbf{h}_t^{\text{set}} = (\mathbf{h}_t^{\text{set,dis}}, \mathbf{h}_t^{\text{set,trans}}, \mathbf{h}_t^{\text{set,set}}) \mathbf{W}^{\text{set}}, \quad (12)$$

$$\mathbf{h}_t^{\text{set,dis}} = \tilde{\mathbf{A}}^{\text{dis}} \mathbf{X}_t^{\text{set}} \mathbf{W}^{\text{set,dis}}, \quad (13)$$

$$\mathbf{h}_t^{\text{set,trans}} = \tilde{\mathbf{A}}^{\text{trans}} \mathbf{X}_t^{\text{set}} \mathbf{W}^{\text{set,trans}}, \quad (14)$$

$$\mathbf{h}_t^{\text{set,set}} = \tilde{\mathbf{A}}_t^{\text{set}} \mathbf{X}_t^{\text{set}} \mathbf{W}^{\text{set,set}}, \quad (15)$$

$\mathbf{W}^{\text{set}} \in \mathbb{R}^{(3 \times d_{\text{exp}}) \times d_{\text{exp}}}$, $\mathbf{W}^{\text{set,dis}} \in \mathbb{R}^{1 \times d_{\text{exp}}}$, $\mathbf{W}^{\text{set,trans}} \in \mathbb{R}^{1 \times d_{\text{exp}}}$, $\mathbf{W}^{\text{set,set}} \in \mathbb{R}^{1 \times d_{\text{exp}}}$ 是可学习的权重。

$$\tilde{\mathbf{A}}_t = \mathbf{I} + (\mathbf{D}_t)^{-\frac{1}{2}} (\mathbf{A}_t) (\mathbf{D}_t)^{-\frac{1}{2}}, \quad (16)$$

其中 \mathbf{D}_t 和 \mathbf{A}_t 是 G 中相应维度的 t 处的度矩阵和邻接矩阵， \mathbf{I} 是单位矩阵。

$$\mathbf{h}_t^{\text{exp}} = \text{ReLU}((\mathbf{h}_t^{\text{his}}, \mathbf{h}_t^{\text{daily}}, \mathbf{h}_t^{\text{weekly}}, \mathbf{h}_t^{\text{monthly}}) \mathbf{W}^{\text{exp}} + \mathbf{b}^{\text{exp}}), \quad (17)$$

\mathbf{W}^{exp} 和 \mathbf{b}^{exp} 分别是可学习的权重和偏差。

多层面空间自注意

多头注意力可以计算为

$$\text{MultiHead}(\mathbf{Q}^H, \mathbf{K}^H, \mathbf{V}^H) = \text{Concat}(\mathbf{o}_1^{\text{Head}}, \dots, \mathbf{o}_H^{\text{Head}}) \mathbf{W}^O,$$

$$\mathbf{o}_i^{\text{Head}} = \text{Attention}\left(\mathbf{Q}^H \mathbf{W}_i^Q, \mathbf{K}^H \mathbf{W}_i^K, \mathbf{V}^H \mathbf{W}_i^V\right), \quad (18)$$

$\text{Attention}(\cdot)$ 表示缩放的点积注意力，定义为

$$\text{Attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{softmax}\left(\frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{d_k}}\right) \mathbf{V}, \quad (19)$$

$\mathbf{Q}^H \in \mathbb{R}^{\tau \times d_{\text{model}}}$, $\mathbf{K}^H \in \mathbb{R}^{\tau \times d_{\text{model}}}$, and $\mathbf{V}^H \in \mathbb{R}^{\tau \times d_{\text{model}}}$ 分别是查询矩阵、键矩阵和值矩阵

$\mathbf{Q} \in \mathbb{R}^{\tau \times d_k}$, $\mathbf{K} \in \mathbb{R}^{\tau \times d_k}$, $\mathbf{V} \in \mathbb{R}^{\tau \times d_v}$ 分别是用于计算标度点积注意力的变换查询矩阵、键矩阵和值矩阵

τ 是时间片的输入维度； d_k 是查询和键的输入维度； d_v 是值的输入维度

$\mathbf{W}_i^Q \in \mathbb{R}^{d_{\text{model}} \times d_k}$, $\mathbf{W}_i^K \in \mathbb{R}^{d_{\text{model}} \times d_k}$, $\mathbf{W}_i^V \in \mathbb{R}^{d_{\text{model}} \times d_v}$ $\mathbf{W}^O \in \mathbb{R}^{H d_v \times d_{\text{model}}}$ 是可学习的权重，它们从原始表示空间投影到不同的子空间

H 是注意力头的数量， d_{model} 是变换矩阵的输入维度， $d_k = d_v = d_{\text{model}}/H$

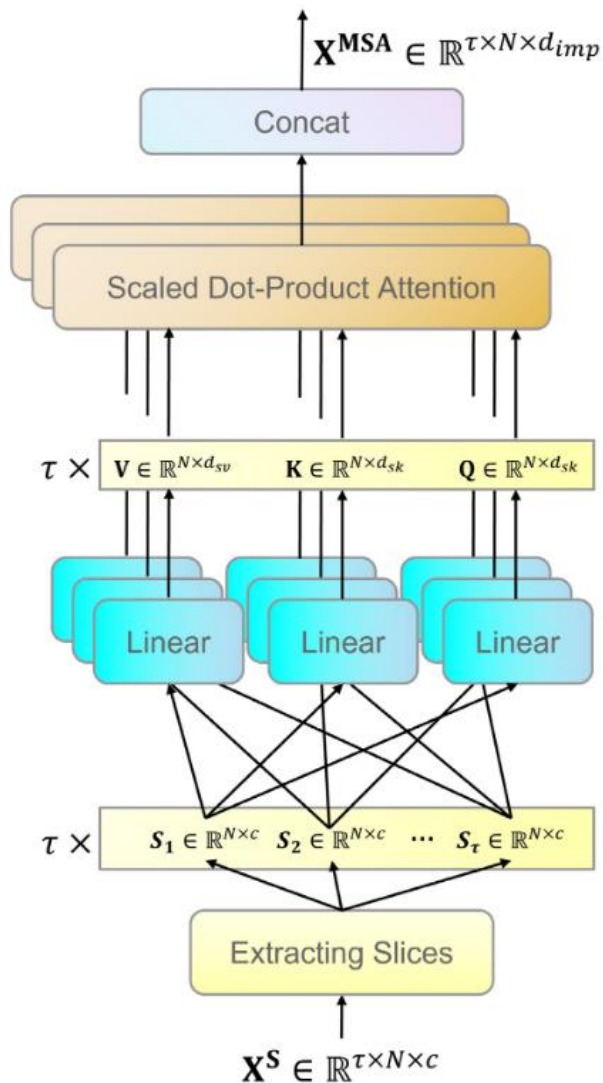


Fig. 3. Proposed multi-slice spatial self-attention.

为了学习空间自我注意，首先将 $\mathbf{X}^S \in \mathbb{R}^{\tau \times N \times c}$ 分成 τ 个片 $\{\mathbf{S}_i \in \mathbb{R}^{N \times c} | i = 1, 2, \dots, \tau\}$ 然后投影到 $\mathbf{Q} \in \mathbb{R}^{N \times d_{sk}}$, $\mathbf{K} \in \mathbb{R}^{N \times d_{sk}}$ and $\mathbf{V} \in \mathbb{R}^{N \times d_{sv}}$ 用于计算标度点积注意力。

多片空间注意力可以用以下方程表示:

$$\text{MultiSlice}(\mathbf{X}^S) = \text{Concat}(\mathbf{O}_1^{\text{Slice}}, \dots, \mathbf{O}_\tau^{\text{Slice}}),$$

$$\mathbf{O}_i^{\text{Slice}} = \text{Attention}(\mathbf{S}_i \mathbf{W}_i^{SQ}, \mathbf{S}_i \mathbf{W}_i^{SK}, \mathbf{S}_i \mathbf{W}_i^{SV}), \quad (20)$$

$\mathbf{W}_i^{SQ} \in \mathbb{R}^{c \times d_{sk}}$, $\mathbf{W}_i^{SK} \in \mathbb{R}^{c \times d_{sk}}$, and $\mathbf{W}_i^{SV} \in \mathbb{R}^{c \times d_{sv}}$ 是可学习的权重，从原始表示空间投射到基于切片的时间子空间; $d_{sk} = d_{sv} = d_{imp}$ ，其中 d_{imp} 为多片空间自注意层的输出维数。

类似于Transformer中的多头注意力模块。

前馈层和层规范化

前馈层由两个线性变换组成，线性变换组成中间是ReLU激活。多切片空间自注意的输出（即，多切片（ \mathbf{X}^S ））表示为 \mathbf{h}^{at} ，第一层归一化的输出表示为 \mathbf{h}^{ln} 。学习到的隐含空间连接 \mathbf{h}^{imp} 可以表示为

$$\mathbf{h}^{\text{ln}} = \text{LayerNorm}(\mathbf{h}^{\text{at}}) + \mathbf{h}^{\text{at}}, \quad (21)$$

$$\mathbf{h}^{\text{imp}} = \text{LayerNorm}((\text{ReLU}(\mathbf{h}^{\text{ln}}\mathbf{W}_1 + \mathbf{b}_1)\mathbf{W}_2 + \mathbf{b}_2)) + \mathbf{h}^{\text{ln}}, \quad (22)$$

其中 \mathbf{W}_1 和 \mathbf{W}_2 是可学习的权重， \mathbf{b}_1 和 \mathbf{b}_2 分别是第一和第二线性变换的可学习的偏差。

通道级连接和完全连接层

将多维图嵌入层学习到的隐藏表示与多片空间自注意通过通道级联进行连接，然后进行线性变换，可以表示为

$$\mathbf{h}^{\text{emb}} = \text{Concat}(\mathbf{h}^{\text{exp}}, \mathbf{h}^{\text{imp}})\mathbf{W}_3 + \mathbf{b}_3, \quad (23)$$

$\mathbf{W}_3 \in \mathbb{R}^{(d_{\text{exp}}+d_{\text{imp}})\times d_{\text{emb}}}$ 和 \mathbf{b}_3 是可学习的权重和偏差， d_{emb} 是异构输入表示层的输出维度。

局部时空特征学习层

为了进一步从异构输入表示层学习的隐藏表示中学习空间-时间信息，我们采用了一维卷积LSTM（ConvLSTM1D）层。

ConvLSTM1D层中的卷积和递归运算有助于学习局部时空特征（ $\mathbf{h}^{\text{local}}$ 表示）。为了充分利用时间信息，我们应用双向ConvLSTM1D层来同时提取前向和后向时间信息。

长期时空特征学习块

利用堆叠的长时特征学习块更好地学习长时特征(用 \mathbf{h}^{long} 表示)。每个长期时空特征学习块由多头时间注意层和前馈层组成。

多头时间注意首先将输入 $\mathbf{h}^{\text{local}} \in \mathbb{R}^{\tau \times N \times d_{\text{local}}}$ 重塑为 $\tilde{\mathbf{h}}^{\text{local}} \in \mathbb{R}^{\tau \times (N \times d_{\text{local}})}$ ，然后将其变为 $\mathbf{Q}^H \in \mathbb{R}^{\tau \times d_{\text{model}}}$, $\mathbf{K}^H \in \mathbb{R}^{\tau \times d_{\text{model}}}$, and $\mathbf{V}^H \in \mathbb{R}^{\tau \times d_{\text{model}}}$ ，最后进行Eq.(18)中定义的多头自注意。

位置编码

由于尺度点积注意等地对待输入矩阵的元素(行)，因此它们的空间或时间位置(也可能包含非常有用的信息)自然被注意机制忽略。因此，我们将空间位置编码（SPE）添加到多层空间自注意层之前的输入 $\mathbf{X} \in \mathbb{R}^{\tau \times N \times c}$ ，并将时间位置编码（TPE）添加到多头时间自注意层前面的 $\tilde{\mathbf{h}}^{\text{local}} \in \mathbb{R}^{\tau \times (N \times d_{\text{local}})}$ ，以分别指示输入元素的空间和时间位置。

我们研究了两种被广泛采用的位置编码:正弦和可学习位置编码。

正弦位置编码可以指定为
$$S(T)PE(pos, 2i) = \sin\left(\frac{pos}{10000^{2i/d_{pe}}}\right), \quad (24)$$

$$S(T)PE(pos, 2i + 1) = \cos\left(\frac{pos}{10000^{2i/d_{pe}}}\right), \quad (25)$$

其中pos为输入元素的空间或时间位置，i为位置编码的维度索引，dpe为位置编码的维度。对于SPE, dpe = c, 对于TPE, dpe = N × dlocal。

和Transformer中的位置编码一样。

可学习位置编码简单地采用一组权值来存储从训练中学习到的位置信息。基于位置编码的消融研究，如表8所示，我们选择正弦位置编码来形成空间和时间位置编码。

Table 8
Ablation studies for the effectiveness of positional encodings.

Model	Region	MAE	RMSE
HSTGANet	Overall	20.06	38.46
HSTGANet w/o PE	Overall	22.88	43.74
HSTGANet w learnable SPE	Overall	23.82	44.46
HSTGANet w learnable TPE	Overall	21.97	41.92
HSTGANet w learnable PE	Overall	20.44	39.09

输出回归量

输出回归器由两个线性变换层组成，并对给定的学习到的时空特征进行最终回归。具体地说， τ_{seq} -step旅游需求预测可以表示为

$$\hat{\mathbf{y}} = (\mathbf{h}^{\text{long}} \mathbf{W}_4 + \mathbf{b}_4) \mathbf{W}_5 + \mathbf{b}_5, \quad (26)$$

$\mathbf{W}_4 \in \mathbb{R}^{d_{\text{model}} \times d_{\text{model}}}$ and $\mathbf{W}_5 \in \mathbb{R}^{d_{\text{model}} \times \tau_{seq}}$ 是可学习的权重， \mathbf{b}_4 和 \mathbf{b}_5 是偏差。

模型训练

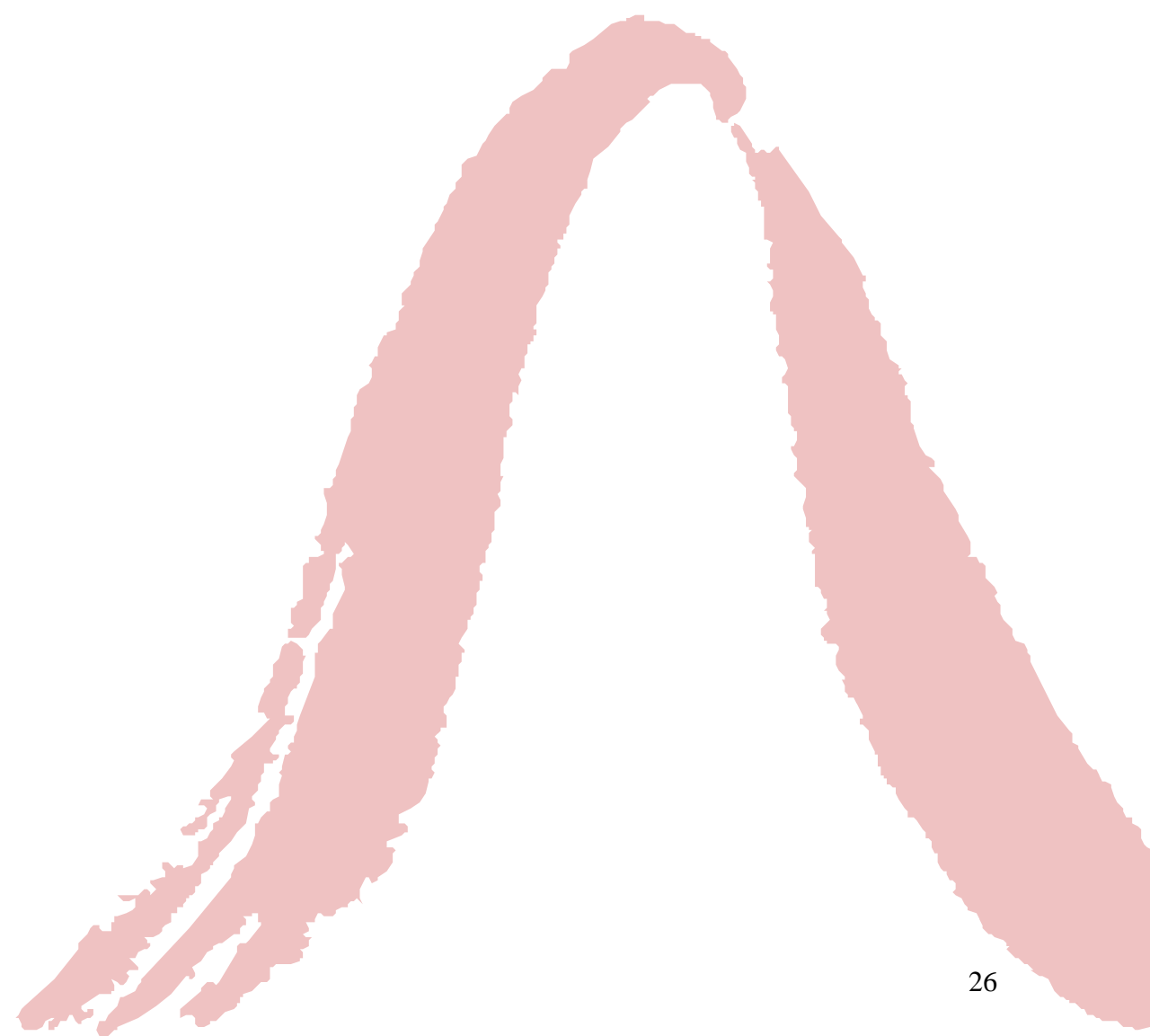
通过最小化旅游需求预测和所有地区和预测范围内的真实旅游需求之间的平均绝对误差(MAE)来优化时空旅游需求预测模型:

$$\underset{\theta}{\operatorname{argmin}} \mathcal{L}_{MAE} = \frac{1}{N * \tau_{seq}} \sum_{i=1}^N \sum_{t=1}^{\tau_{seq}} |y_t^i - \hat{y}_t^i|, \quad (27)$$

其中 θ 表示模型的可学习参数， y 和 \hat{y} 分别是真实和预测旅游需求。



实验



实验数据集

由于公共轮渡是游客唯一可用的交通工具，因此可以从轮渡票务数据中获得历史旅游需求和换乘流量（包括进出流量）。收集的数据范围为2017年1月1日至2019年12月31日。每天共有11次观测。

表1收集的万山群岛六个地区的旅游需求统计数据。

Table 1
Statistics of collected tourism demand for the six regions of Wanshan Archipelago.

Region	mean	std	min	25%	50%	75%	max
Xiangzhou	141.83	165.68	0	0	94	213	1091
Guishan	35.90	54.75	0	0	4	63	445
Wailingding	62.12	101.49	0	0	0	97	709
Hengqin	7.42	24.87	0	0	0	0	437
Dongao	36.66	65.83	0	0	0	54	500
Wanshan	12.03	24.17	0	0	0	10	323

mean: Mean of tourism demand;
std: Standard deviation of tourism demand;
min: Minimum tourism demand;
max: Maximum tourism demand;
p%: p percentile.

我们进一步收集了所有六个地区的地理位置、中国公众假期信息、百度搜索引擎的搜索指数数据和中国气象数据服务中心的天气数据，形成了输入序列。在本案例研究中，所有六个地区共享相同的假期信息。

实验设置

Table 2

Search space and best hyper-parameter settings for hyper-parameters of the proposed HSTGANet.

Hyper-parameter	Search space	Best setting
Hidden dimensions ($d_{\text{exp}} = d_{\text{imp}} = d_{\text{emb}} = d_{\text{local}}$)	{32, 64, 128}	64
Kernel size of ConvLSTM1D	{3, 5}	3
Number of ST learning blocks (N_E)	{1, 2, 3}	3
Number of attention heads (H)	{4, 8}	4
Multi-head attention dimension (d_{model})	{64, 128, 256}	128
Feed-forward dimension (d_{ff})	{64, 128, 256}	128
Dropout rate	{0.3, 0.5, 0.7}	0.5

两个指标用于评估所提出的HSTGANet和基线的性能，包括平均绝对误差（MAE）和均方根误差（RMSE）。

基线模型和实施

空间模型很长一段时间没有应用于旅游需求预测。我们使用时间序列预测和时空预测基线模型。

时间序列预测基线模型：

选择ARIMA和SARIMAX模型作为具有代表性的统计模型。在拟合SARIMAX模型时，我们使用假期、天气和搜索索引信息作为解释变量。

选择支持向量回归（SVR）和随机森林回归（RFR）作为代表性的传统机器学习模型。作为具有代表性的深度学习模型，我们选择了两个基于RNN的模型，LSTM和BiLSTM，以及一个基于注意力的模型，即多头注意力CNN（MHACNN）。

时空预测基线模型：

面板数据模型和时空自回归模型。我们在计算方程时使用了基于流量的空间加权矩阵（1）和（2）。在这两个模型中没有考虑解释变量。

此外，还有深度卷积神经网络（DCNN）、时空图卷积网络（STGCN）和时空变换器（STTN）（长期交通流预测的SOTA之一）。还考虑了两种先进的基于GNN的时空模型，即图多注意力网络（GMAN）和基于注意力的时空图神经网络（ASTGNN）。

表3和表4分别显示了所提出的HSTGANet与万山群岛六个地区77步旅游需求预测的时间序列和时空预测基线模型在MAE和RMSE方面的性能比较。

Table 3
Performance comparison of 77-step tourism demand forecasting for the six regions in Wanshan Archipelago for HSTGANet and the time series forecasting baseline models.

Region	Metric	ARIMA	SARIMAX	SVR	RFR	LSTM	BiLSTM	MHACNN	HSTGANet
Xiangzhou	MAE	119.97	118.06	74.00	59.75	77.44	75.09	69.77	50.81
	RMSE	151.02	153.61	120.61	93.22	102.20	100.93	93.78	79.65
Guishan	MAE	42.42	38.34	26.46	24.77	30.75	28.64	28.12	20.87
	RMSE	60.43	60.18	52.62	43.19	46.81	44.48	41.29	42.04
Wailingding	MAE	57.34	55.79	32.61	25.91	34.70	38.39	30.89	21.06
	RMSE	73.57	73.61	59.18	49.87	51.66	54.73	48.42	42.93
Hengqin	MAE	9.06	16.14	7.88	11.49	9.67	11.91	16.28	6.97
	RMSE	26.67	27.04	21.26	24.27	19.76	20.70	22.70	22.59
Dongao	MAE	35.63	38.41	25.22	16.72	26.45	23.11	20.87	13.34
	RMSE	45.87	50.78	48.11	32.03	37.74	33.34	29.27	27.1
Wanshan	MAE	13.28	14.21	8.41	5.86	7.56	6.37	7.65	7.28
	RMSE	21.31	20.89	20.77	13.12	14.03	12.52	13.28	16.46
Overall	MAE	46.28	46.83	29.10	24.08	31.10	30.59	28.93	20.06
	RMSE	63.15	64.35	53.76	42.62	45.37	44.45	41.46	38.46

机器学习模型与统计模型相比、最好的模型

Table 4
Performance comparison of 77-step tourism demand forecasting for the six regions in Wanshan Archipelago for HSTGANet and the spatial-temporal forecasting baseline models.

Region	Metric	Panel	STARMA	DCNN	STGCN	STTN	GMAN	ASTGNN	HSTGANet
Xiangzhou	MAE	152.95	151.43	73.68	76.55	60.11	60.25	63.08	50.81
	RMSE	215.8	212.9	96.57	102.71	86.77	95.09	99.91	79.65
Guishan	MAE	49.14	48.25	33.36	26.55	20.33	23.82	21.1	20.87
	RMSE	84.54	82.57	46.38	42.63	36.6	49.8	44.48	42.04
Wailingding	MAE	56.72	56.57	43.99	41.56	34.4	29.61	24.82	21.06
	RMSE	97.57	95.99	58.11	59.4	52.18	55.25	49.58	42.93
Hengqin	MAE	8.89	8.85	18.03	12.25	14.31	6.82	6.69	6.97
	RMSE	28.09	27.12	25.14	20.2	22.78	21.63	19.7	22.59
Dongao	MAE	34.45	34.55	33.13	26.89	18.98	17.68	14.24	13.34
	RMSE	62.2	60.44	42.76	36.68	29.7	34.33	28.87	27.1
Wanshan	MAE	14.91	15.52	12.09	8.54	6.66	7.03	5.64	7.28
	RMSE	30.68	29.68	16.62	15.33	13.38	16.74	13.26	16.46
Overall	MAE	52.84	52.53	35.71	32.06	25.80	24.20	22.60	20.06
	RMSE	86.48	84.78	47.60	46.16	40.24	45.47	42.63	38.46

基于深度学习的模型与统计模型相比
基于图注意与基于卷积和图卷积的相比
与表3的最好模型相比

训练时间序列预测模型以获得每个区域的最佳预测性能，而基于整体预测性能，对所有六个区域训练一个单一的时空预测模型。因此，优化的时空预测模型在某些区域可能不会优于特定于区域的时间序列预测模型，特别是对于旅游需求规模较小的模型（例如，BiLSTM 在预测万山岛旅游需求的 RMSE 方面实现了最佳预测性能）。

不同时空预测模型捕获的时空特征可能不同，因此对六个区域的关注程度也可能不同，导致跨区域的预测性能不一致(即，这些模型可能在不同的区域实现最佳的区域特定性能)。

消融研究

从HSTGANet中去除多维图嵌入(MDGE)层或多层空间自我注意(MSSSA)层，分别验证显式动态空间连接和隐式动态空间连接的有效性，以及两层从两个输入中提取高维时空特征的有效性。

表 5 所提出的多维图嵌入层和多层空间自注意力层有效性的消融研究。

Table 5
Ablation studies for the effectiveness of the proposed multi-dimensional graph embedding layer and multi-slice spatial self-attention layer.

Model	Region	MAE	RMSE
HSTGANet	Overall	20.06	38.46
HSTGANet w/o MDGE	Overall	27.14	48.37
HSTGANet w/o MSSSA	Overall	44.23	72.92

多维图的有效性

表 6 中显示的实验结果证明了采用的六个显式空间交互的有效性，由多维图建模。

Table 6
Ablation studies for the effectiveness of the six spatial interactions in the multi-dimensional graph.

Model	Region	MAE	RMSE
HSTGANet	Overall	20.06	38.46
HSTGANet w/o A^{dis}	Overall	21.73	40.22
HSTGANet w/o A^{trans}	Overall	23.71	43.89
HSTGANet w/o A^{his}	Overall	21.83	41.33
HSTGANet w/o A^{daily}	Overall	25.02	45.01
HSTGANet w/o A^{weekly}	Overall	28.61	52.25
HSTGANet w/o $A^{monthly}$	Overall	21.93	40.59

周期性观测的有效性

表 7 显示，删除任何周期性观察对长期旅游需求预测的预测性能有负面影响。

Table 7
Ablation studies for the effectiveness of periodic observations.

Model	Region	MAE	RMSE
HSTGANet	Overall	20.06	38.46
HSTGANet w/o X^{dtd}	Overall	21.34	38.84
HSTGANet w/o X^{wtd}	Overall	24.26	44.71
HSTGANet w/o X^{mtd}	Overall	21.80	40.00

位置编码的有效性

- (1) HSTGANet，正弦位置编码;
- (2) HSTGANet，去除位置编码(不含PE的HSTGANet);
- (3)正弦SPE替换为可学习的SPE (HSTGANet with 可学习的SPE);
- (4)正弦TPE替换为可学习的TPE (HSTGANet with可学习的TPE);
- (5)正弦位置编码和TPE都被可学习的位置编码(具有可学习PE的HSTGANet)所取代。

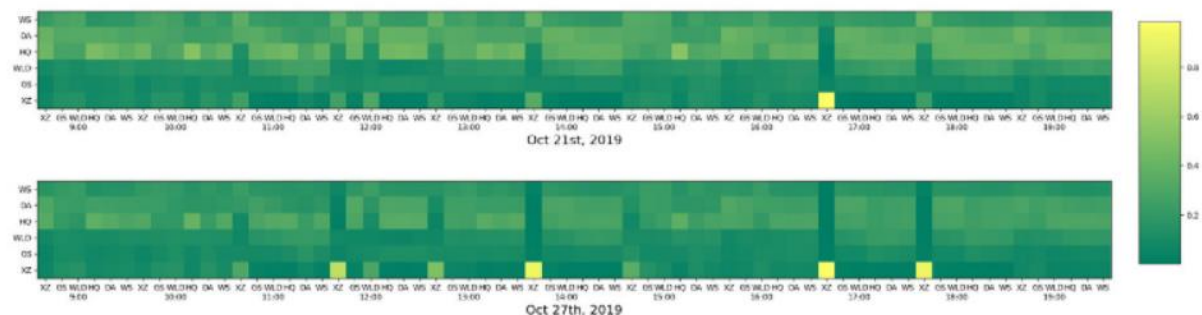
Table 8
Ablation studies for the effectiveness of positional encodings.

Model	Region	MAE	RMSE
HSTGANet	Overall	20.06	38.46
HSTGANet w/o PE	Overall	22.88	43.74
HSTGANet w learnable SPE	Overall	23.82	44.46
HSTGANet w learnable TPE	Overall	21.97	41.92
HSTGANet w learnable PE	Overall	20.44	39.09

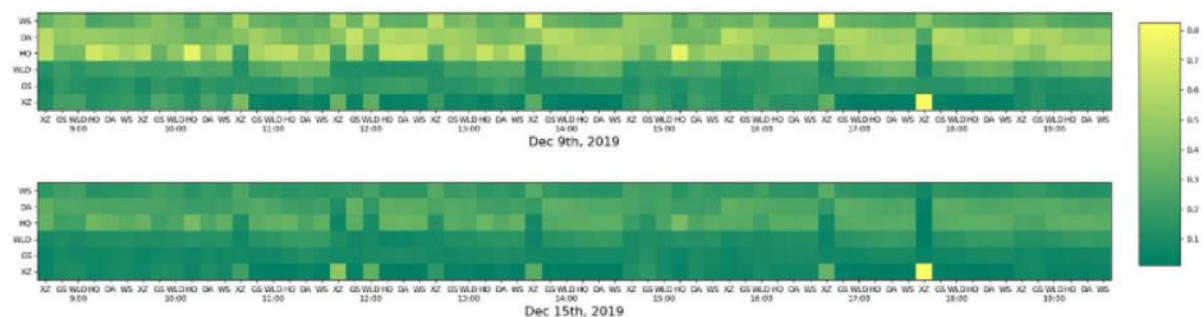
表 8 清楚地表明位置信息对于注意力层很重要。具体来说，正弦位置编码比可学习的位置编码更有效。实验结果表明，使用混合位置编码可能会导致更差的旅游需求预测性能（（3）、（4）的情况）。

学习隐式动态空间连接的可视化


利用所提出的多层空间自注意，对于滑动输入，得到一个表示学习到的隐式动态空间连接的注意加权矩阵Wimps



(a) Implicit dynamic spatial connections on Oct. 21st, 2019, and Oct. 27th, 2019.



(b) Implicit dynamic spatial connections on Dec. 9th, 2019, and Dec. 15th, 2019.



结论

本文提出了一种基于图注意力的旅游需求预测时空学习框架。

通过多维图嵌入和多片空间自关注，实现了高维动态空间连接的全面嵌入。通过利用局部和长期时空特征学习层，本文提出的HSTGANet在学习时空特征和预测旅游需求方面表现出显著的能力。



谢谢

