

DOI:10.19479/j.2095-719x.2403220

基于图神经网络的建筑能耗预测

杨振舰, 卢世林

(天津城建大学 计算机与信息工程学院, 天津 300384)

摘要:在城市建筑管理中,准确预测建筑能耗对实现建筑节能和构建智慧城市具有重要意义。由于能耗数据的复杂性,长期并准确地预测建筑能耗是时间序列预测中极具挑战性的难题之一。近年来,研究人员将神经网络模型应用于能耗预测任务,并取得了优秀的预测结果,然而建筑能耗会受到多维因素的影响,为了提高预测精度,提出了一种基于图神经网络的建筑能耗预测方法。该方法使用改进的图卷积网络来捕获时间序列的空间依赖关系,通过时间卷积模块来获取时间序列的时序依赖关系,并通过时空融合,更充分地挖掘多元时间序列中的时序特征,支持在端到端的框架中联合学习,在真实的能耗数据集上的实验结果证实了模型拥有更加优异的性能表现。

关键词:建筑能耗;建筑节能;图神经网络;能耗预测;空间依赖;时序特征

中图分类号:TP111.195

文献标志码:A

文章编号:2095-719X(2024)03-0220-08

Modeling of Building Energy Consumption Prediction Based on Graph Neural Networks

YANG Zhenjian, LU Shilin

(School of Computer and Information Engineering, TCU, Tianjin 300384, China)

Abstract: In urban building management, the high proportion of building energy consumption is a huge problem at present. Accurately predicting building energy consumption is of great significance to achieve building energy conservation and the building of smart cities. Due to the complexity of energy consumption data, long-term and accurate forecasting of building energy consumption is one of the most challenging problems in time series forecasting. In recent years, researchers have applied neural network models to the task of energy consumption prediction and achieved excellent prediction results. However, building energy consumption is affected by multidimensional factors. In order to improve the prediction accuracy, this paper proposes the modeling of building energy consumption prediction based on graph neural networks. The method uses a modified graph convolutional network to capture the spatial dependencies of time series, and a temporal convolution module to obtain the temporal dependencies of time series. Through the fusion of time and space, time series features that multivariate time series can be more fully mined, and joint learning in an end-to-end framework can be supported. The experimental results on the real energy consumption dataset confirm that the model has better performance.

Key words: building energy consumption; building energy efficiency; graph neural network; energy consumption forecast; spatial dependence; timing characteristics

随着全球经济的飞速发展和人类需求的快速增长,城市建筑规模及其能耗占比呈现急剧增长模式,随之带来的是复杂的气候变化和环境污染等问题。人们对能源的依赖及其供需矛盾存在很大担忧,提升能源利用率和降低能耗占比逐渐成为了行业研究热点和国家战略需求^[1]。

2021年数据显示,全球公共建筑所消耗的能源已经超过总能耗量的三分之一,所排放的二氧化碳占比超过40%^[2]。中国建筑能耗研究报告指出,到2019年,全国建筑全过程能耗总量达22.33亿tce,约占全国能源消费总量的21.2%,而建筑全过程碳排放总量高达49.97亿tco₂,约占全国碳排放总量的50.6%^[3]。根据数

收稿日期:2022-11-30;修订日期:2023-01-02

作者简介:杨振舰(1975—),男,山东茌平人,天津城建大学教授,博士。

据分析,中国将在2030年达到碳排放峰值,这将是非常不乐观的局面,国家中央统筹作出的重大战略决策明确提出了2060年前实现碳中和的战略方针^[4].因此,更有效的管理公共建筑能耗的利用率和更明确的分析公共建筑能耗数据就显得极其重要,实现建筑节能对全球节能减排也有着重要的意义.

通过预测建筑未来的能源消耗量,管理者可以提前判断建筑的能源消耗趋势,有计划地实施能源采购与能源调控策略,它是建筑节能工程里最基本的一项工作^[5].因此,建筑节能领域的研究人员对其展开了许多研究,预测建筑能耗的方法也很多.传统预测方法更多是依靠统计学方法,这类方法在研究初期广受青睐,但是面临两个问题:一是模型的复杂度随着变量的增多呈二次增长关系;二是面临着变量增多的过拟合问题.近年来,随着人工智能的发展与深入研究,出现了众多神经网络预测模型,像卷积神经网络模型、循环神经网络模型、长短期记忆模型和门控循环单元等都在精确度上有显著提升.Zhao R的预测模型结合了循环神经网络和门控循环单元,对于实际电耗数据实现了突破性的预测效果,激起了研究人员对神经网络的研究热潮^[6].Guokun Lai等设计了一种多元时间序列的通用模型,考虑实际应用中收集到的时间序列数据通常涉及长期和短期模式的混合,针对这一特性设计了全新的深度学习框架LSTNet.LSTNet使用卷积神经网络和递归神经网络来提取变量之间的短期局部依赖模式,并发现时间序列趋势的长期模式,在电力数据集上获得了非常优异的表现,从误差和准确度上都完全超越了之前的模型^[7].

然而,神经网络在能耗预测方面一直存在着诸多问题:为了追求精度而选取一些代表性的数据,从而忽视了可能会影响预测结果的隐藏因子;进行数据预处理时,为了迎合图像领域的做法将数据处理为欧式距离上的平面数据,这会因为未同时提取到时空相关特征而导致结果无法表征实际意义.因此,在预测问题上更多的要考虑如何寻找依赖的属性关系.近几年,随着图卷积网络的提出,得益于其核心理论:图中的每个结点都因为其邻居节点和更远的节点的影响而在改变着自己的状态直到最终的平衡,关系越亲近的邻居节点影响越大.这解决了多元时间序列的一大核心问题:如何建立节点间的依赖关系.

本文以能耗数据为依据,针对卷积神经网络的不足,将图卷积神经网络与卷积神经网络融合,建立了EC-SGCN(energy consumption-signed graph convolutional network)网络模型对建筑能耗进行预测.

1 相关工作

1.1 图神经网络

图神经网络使用神经网络来学习图结构数据,提取和发掘图结构数据中的特征和模式,用以处理网络中实体间的空间依赖关系.Gori等^[8]首次提出了图神经网络,使用循环神经网络来处理图结构数据.随后Petar等^[9]提出了图注意力网络,与GCN平等对待节点的所有邻居相比,注意力机制可以为每个邻居分配不同的注意力得分,从而识别出更重要的邻居.随后Jiang等^[10]又提出了超图卷积网络DHGNN,提出并解决了图卷积网络中图不变性的问题,从而限制了节点之间的隐含关系.它们共享节点信息,本质上都是将信息从节点的邻居传递到节点本身来捕获节点的依赖关系.

然而,以上的GNN模型通常采用有向图的单向关系来处理图结构数据,往往忽视了图中可能存在的负相关关系在时空数据上的有效性,于是Tyler Derr提出了一种有符号的图神经网络模型SGCN^[11],该模型为每个实体建立关系网络,并构建正负向关系的超图,通过大量实验,该方法已被证明在人际网络中优于其他方法.但是如果将其方法应用到多元时间序列构建的关系图中,且更深入地挖掘多元实体间的依赖关系是一个新颖的方向.受此启发,本文结合了SGCN的网络思想来处理多元时间序列中的能耗预测问题.

SGCN是一个多层叠加的图卷积网络,它每一层除了考虑正向关系和负向关系,同时还考虑到正负链接的交叉链接问题^[12-13].它提出了一种平衡理论,即“朋友和敌人”的关系,如图1所示.如果需要维护两个邻接矩阵,分别是平衡矩阵 \mathbf{B} 和非平衡矩阵 \mathbf{U} ,那么第 n 层的矩阵更新为以下两种情况.

当 $n=1$ 时,此时在第一层就是起始矩阵,即

$$\mathbf{B}_i(1) = \{e_j | e_j \in N_i^+\} \quad (1)$$

$$\mathbf{U}_i(1) = \{e_j | e_j \in N_i^-\} \quad (2)$$

当 $n>1$ 时,此时需要考虑交叉链接,即

$$\mathbf{B}_i(n) = \{e_j | e_k \in \mathbf{B}_i(n-1) \cup e_j \in N_k^+\} \cup \{e_j | e_k \in \mathbf{U}_i(n-1) \cup e_j \in N_k^-\} \quad (3)$$

$$\mathbf{U}_i(n) = \{e_j | e_k \in \mathbf{U}_i(n-1) \cup e_j \in N_k^+\} \cup \{e_j | e_k \in \mathbf{B}_i(n-1) \cup e_j \in N_k^-\} \quad (4)$$

其中: $\mathbf{B}_i(n)$ 、 $\mathbf{U}_i(n)$ 表示可以从节点 i 沿着单位长度的平衡、非平衡路径到达的节点集; N_i^+ 、 N_i^- 表示节点的正、负邻居集合.

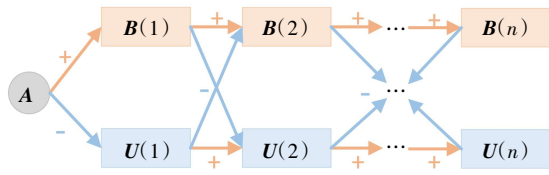


图1 交叉链接

1.2 多元时间序列

多元时间序列预测研究已经经历相当长的时间,最早其采用统计学的方法提出了动态回归模型体系,其后提出的向量自回归模型VAR是动态回归模型的多元版本^[14],用以捕获多个时间序列之间的线性依赖关系,但是模型的复杂度会随着变量成指数增长,并且面临着变量过多带来的过拟合问题.随后,结合机器学习在多元时间序列问题上的建模大大提高了预测的精度,但是仍然无法解决非线性数据问题.随着深度学习在非线形数据领域展示出极强的适应性,Shih等^[15]较早提出将深度学习应用于多元时间序列领域,并提出了预测模型. Shih等使用卷积神经网络来捕获变量之间的局部依赖关系,并使用循环神经网络来保持时间的长期依赖关系.由于卷积神经网络并不能考虑时间序列中的空间关系,因此其不能达到充分利用变量之间的潜在依赖关系的目的.

2 模型

2.1 问题定义

常见的能耗数据并不是单元的,本文中处理的能耗数据主要为多元时间序列数据.而在能耗预测任务中鲜有结合图论的方法,所以这里将多元时间序列数据中每个变量当作单个节点,通过学习每个节点之间的关系构建距离关系组成节点间的边特征,则可以将无结构的能耗数据定义为无向图 $G=(V,E,A)$,其中 V 是节点集合, E 是边特征集合, $A=\mathbb{R}^{N \times N}$ 是对应的邻接矩阵.

在给定的多元时间序列 $X=\{x_{11},x_{12},x_{13},\dots,x_{1p}\}$ 的一段时间步长为 p 的连续序列中,本文的目标是预测 $y=\{Z_{p+Q}\}$ 的 Q 步距离的能耗值数据,并结合能耗数据在时间中的周期性特征.那么,在时间步 t 时, N 个节点的能耗数据表示为 $X_t \in \mathbb{R}^{N \times D}$,其中 D 表示特征维度.因此通过构建从 X 到 Y 的映射函数 f ,能够预测接下来 t 个时刻的建筑能耗信息.

2.2 模型架构

模型是由多个单位时间步的时空融合模块组成,每个时空融合模块主要包含两大部分:空间卷积模块

和时间卷积模块,模型架构如图2所示.

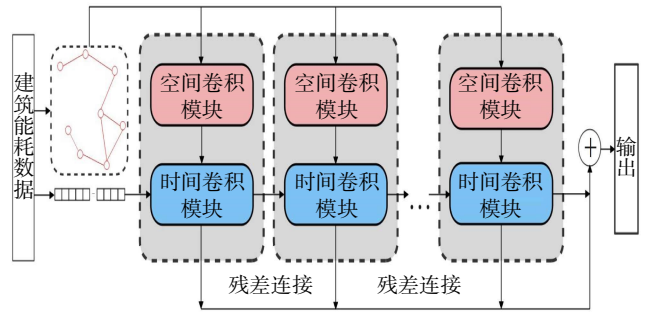


图2 模型架构

首先,将原始数据拆分成若干个单位时间步的时间序列数据作为原始输入,在每一个单位时间步中,利用图构建模块将时间序列数据构建为每个节点为原始数据的元点的邻接矩阵;再利用空间卷积模块捕获图中的拓扑结构和学习时间序列之间的隐藏关联信息;同时,将每个单位时间步的原始数据输入到时间卷积模块中,通过时间卷积模块的信息传递获取序列的动态变化,用以捕获序列的时间特征;然后,通过残差连接将空间卷积模块和时间卷积模块数据进行时空融合,为了防止梯度消失问题,本文在每个时间卷积层结束后添加残差连接;最后,得到 t 个时间步后的预测结果.

2.3 图构建模块

图构建模块的目的在于可以自适应地学习图邻接矩阵,用来捕获建筑能耗时间序列数据之间的空间关系.为了构建有向图,通常是直接利用一个点到其 K 个最近邻点的距离来确定点簇的关系,从而测量节点间相似性,一般使用KNN完成^[16].但是面临两个问题:其一是传统的KNN对训练数据依赖度特别大;其二是随着图形大小的增长计算成本呈现二次增长.

面对问题1,本文引入节点异常检测的思想,利用KNN来确定节点到它第 k 个最近邻居的距离,并生成异常分数,并将异常分数较高,即偏离太大的邻居舍弃^[17].首先,仍然使用KNN构建有向图,边特征为样本间欧式距离,如下式所示,即

$$e_{j,i} = \begin{cases} \text{dist}(X_i, X_j) & \text{if } j \in N_i \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

然后,对每个节点进行信息传播,为了适应不同的聚合方式,使用可学习的信息聚合函数,对于每个节点的 k 个近邻距离构造向量,如下式所示,即

$$e^i = [e_{1,i}, \dots, e_{k,i}] \in R^k \quad (6)$$

最后将向量通过神经网络 F 映射到标量值来表示节点 i 的异常程度,更新函数结果即为聚合函数输

出的结果,异常点的距离设置为0,即

$$h_{N_i} = F(e^i, \Theta) \quad (7)$$

$$\gamma = h_{N_i} \quad (8)$$

对于问题2,提出了一种子图训练算法来处理大型图时计算量问题.在使用KNN构建图时,通常需要将所有节点存储到内存中,如果图很大就会面临内存溢出的问题. Wei-Lin Chiang等^[18]提出一种子图训练算法来解决训练时内存爆炸问题,他们应用图聚类算法将原始图划分为多个子图,并在划分的子图上训练图卷积网络.基于这种思想,模型将节点随机分成 n 组,每组按照采样节点去学习子图结构,这样每个节点都有相同的概率被分配给同一组中另一个节点,以便于可以计算和更新这两个节点之间的相似性.训练完成后,由于每个节点嵌入都经过了良好的训练,邻接矩阵也可以在预测之前预先并行计算,所以可以构建全局图用于很好的表现空间关系.

2.4 空间卷积模块

空间卷积模块目的在于融合节点的信息并聚合其邻居节点的信息来处理空间依赖关系.空间卷积模块由 n 层SGCN传播层组成,分别处理每个节点的流入和流出信息,流入信息是通过将前 $n-1$ 层传播层叠加的结果进行输出并聚合得到的.

SGCN传播层:基于一个图邻接矩阵,本文设计了SGCN传播层来处理空间依赖上的信息流.在模型的SGCN传播层中聚合和传播信息时,本文将在每一层维护两个图邻接矩阵,一个用于相应的平衡集,一个用于相应的不平衡集(不平衡集是平衡集的反转邻接矩阵).

与GCN传播规则相同,本文使用 $f_i^0 \in R^d$ 来表示初始节点特征,因此第一个聚合层(即当 $n=1$ 时)可以正常的定义传播规则,即

$$f^{B(1)} = \sigma(\mathbf{W}^{B(1)} [\sum_{j \in N_i^+} \frac{f_j^0}{N_i^+}, f_i^0]) \quad (9)$$

$$f^{B^-(1)} = \sigma(\mathbf{W}^{B^-(1)} [\sum_{j \in N_i^-} \frac{f_j^0}{N_i^-}, f_i^0]) \quad (10)$$

其中: σ 表示 \tanh 非线性激活函数,并且权重在不同的节点之间共享; $\mathbf{B}(n)$ 表示负责聚合来自节点集 $\mathbf{B}_i(n)$ 信息的聚合器; N_i^+ 、 N_i^- 表示能耗数据集节点的正、负邻居集; $f_i^{B(n)}$ ($f_i^{B^-(n)}$)为在第 n 层节点的平衡表示; $\mathbf{W}^{B(1)} \mathbf{W}^{B^-(1)} \in R^{d^{in} \times d^{in}}$ 表示第1层的线性变换矩阵, d^{in} 是内部隐藏表示的长度,类似于GCN中第 n 层中的权重矩阵.

在后续的SGCN传播层中,考虑到正负链接的交叉链接问题,聚合会显得更加复杂,正如公式3-4中,

当 $n>1$ 时, $\mathbf{B}_i(n)$ 和 $\mathbf{U}_i(n)$ 的表达式也变得极其复杂是同一个原因.那么, $n>1$ 的传播规则定义如下

$$f^{B(n)} = \sigma(\mathbf{W}^{B(n)} [\sum_{j \in N_i^+} \frac{f_j^{B(n-1)}}{|N_i^+|}, \sum_{k \in N_i^-} \frac{f_k^{B(n-1)}}{|N_i^-|}, f_i^{B(n-1)}]) \quad (11)$$

$$f^{B^-(n)} = \sigma(\mathbf{W}^{B^-(n)} [\sum_{j \in N_i^+} \frac{f_j^{B^-(n-1)}}{|N_i^+|}, \sum_{k \in N_i^-} \frac{f_k^{B^-(n-1)}}{|N_i^-|}, f_i^{B^-(n-1)}]) \quad (12)$$

与第一层类似, $\mathbf{W}^{B(1)} \mathbf{W}^{B^-(1)} \in R^{d^{in} \times 3d^{in}}$ 为层数 $n>1$ 的线性变换矩阵.通过公式11和12,可以知道在第 n 层获取节点的正相关表示(即 $f^{B(n)}$)时,它是聚合了所有正链接的邻居在 $n-1$ 层的正相关表示,同时获取了 $n-1$ 层负相关表示的平均值,即来自所有负链接邻居 $u_k \in N_i^-$ 的信息.所以对于 $n>1$ 的情况,可以理解为正相关表示实际上不仅直接从正链接的邻居处收集信息,同时从正相关的正链接的邻居和负相关的负链接的邻居处获取信息进行聚合.类似地,也可以同样的去解释 $f^{B^-(n)}$.

在多次实验中,证明两层SGCN聚合的效果表现优异,实验尝试去聚合三层及更高的层数,但是随着训练量大幅增加,叠加效果提升却不再明显,所以本文使用了两层结构.

2.5 时间卷积模块

考虑到建筑能耗数据的线性结构,时间卷积模块主要用来捕获时间序列数据之间的时间依赖性关系.时间卷积模块采用混合扩张层和门控循环单元来提取复杂时间特征.对于需要同时处理不同大小范围的多种时间模式和长时间序列段的序列关系,本文提出了混合扩张层.它使用多种尺寸的滤波器处理单元,再结合混合扩张卷积,同时为了防止梯度弥散和梯度爆炸引入了门控循环单元.

混合扩张层:由两个扩张层组成^[19],其一是由一个扩张层连接一个 \tanh 激活函数构成;另一个是由一个初始扩张层连接一个sigmoid激活函数构成.其作用是作为一个门控单元来控制过滤器.每个扩张层由多个扩张卷积构成^[20].

首先,使用扩张卷积需要选择合适的卷积内核,如果内核大小设定过大,卷积层可能因为内核太大而无法发现短期时序特征,设置太小可能无法发现长期时序特征,在图像处理中,广泛采用策略称为inception^[21].它在二维的输出具有三种不同内核大小的卷积滤波器,分别为 1×1 , 3×3 和 5×5 ,但是运用在时间序列存在三个待解决的问题:①如何从二维图像迁移到一维的时间序列中;②叠加的扩张卷积产生的栅格效应,因为扩张卷积的内核并不连续,也就是并不是所有的

序列段都用来计算了,因此这里将信息看作 checker-board 的方式将会损失信息的连续性,即膨胀卷积不能覆盖所有的时序特征;③扩张卷积的设计更适合用于获取长期时序特征,这样会丢失较小的时序特征,因此,如何兼顾设计扩张卷积内核是关键问题。

首先,对于问题 1,考虑到图像中使用的滤波器集合,可以设置为 1×1 、 1×3 和 1×5 的滤波器集合。但是考虑到建筑能耗数据具有的周期性,例如 7, 12, 24, 60 等,原始滤波器集合不能很好地包含这些周期,于是提出了 1×2 、 1×3 、 1×6 和 1×7 的滤波器集合^[2]。针对问题 2 和问题 3:为了防止栅格效应,应该将膨胀率设置为不存在大于 1 的公约数,同时膨胀率需要满足下式,即

$$M_i = \max[M_{i+1} - 2r_i, M_{i+1} - 2(M_{i+1} - r_i), r_i] \quad (13)$$

其中: r_i 是第 i 层的膨胀率; M_i 是指在 i 层的最大膨胀率,假设总共有 n 层,那么 $M_n = r_n$ 。最后确定为 $(1 \times 2$ 、 1×3 、 1×5 、 $1 \times 7) \times 2$ 的循环滤波器集合。

于是,在给定一维时间序列输入 $z \in \mathbb{R}^t$ 时,使用混合扩张层,过滤器的输出根据最大的过滤器宽度被截断为相同的长度,并在整个卷积中连接通道维度,所以扩张卷积层定义为

$$z = \text{concat}[2 \times (\sum_{s=0}^{k-1} f_{1 \times 2}(s), \sum_{s=0}^{k-1} f_{1 \times 3}(s), \sum_{s=0}^{k-1} f_{1 \times 5}(s), \sum_{s=0}^{k-1} f_{1 \times 7}(s)) \times (t - d - s)] \quad (14)$$

其中: d 是膨胀因子。

其次,循环神经网络在处理时间序列数据方面一直存在优势,所以本文将混合扩张层处理过的一维数据通过 GRU 网络来防止梯度弥散和梯度爆炸^[23]。GRU 接受两个输入,其一是上一个状态中聚合了空间信息的数据嵌入表示 H ,另一个是当前状态的时间信息状态表示 S 。GRU 通过门控机制来控制当前时刻输入和上一状态的写入程度,更新门的值越大,重置门的值越小,则表示上一状态的信息被写入的越少。以这种方式,时间卷积模块能够捕获能耗数据长期的时间相关性,其详细操作如下式所示,即

$$r_i = \sigma(W_r[H_{i-1}, S_i] + b_r) \quad (15)$$

$$z_i = \sigma(W_z[H_{i-1}, S_i] + b_z) \quad (16)$$

$$\hat{h}_i = \lambda(W_h[z_i, r_i \times H_{i-1}] + b_h) \quad (17)$$

$$H_i = (1 - z_i) \times H_{i-1} + r_i \times \hat{h}_i \quad (18)$$

式中: z 为 GRU 的更新门大小; r 为 GRU 的重置门大小; σ 为 sigmoid 非线性激活函数; λ 为 tanh 非线性激

活函数; W 和 b 是可学习参数。

2.6 残差连接

由于时间卷积模块的接收域大小取决于网络深度 n 、滤波器大小 k 和膨胀因子 d ,因此如何使时间卷积模块变深变大是获取接收域的核心^[24]。接收域可以叠加大量的层并且选择较细尺度的过滤器尺寸使网络变深变窄。在残差网络中,整个网络采用跳跃连接,可以加快训练进程,并且可以避免模型的梯度消失^[25]。

残差连接层本质上是 $1 \times L_i$ 标准卷积,其中 L_i 是第 i 个残差连接层的输入序列长度。它将序列标准化为相同长度的序列并跳转到输出模块。输出模块由两个 1×1 标准的卷积层组成,其作用将输入的维度转换为所需的输出维度。

2.7 预测

经过 T 个时间步的时空融合后,ECSGCN 采用了残差连接层得到了最后的预测结果。因为能耗预测本质是回归任务,所以采用相对平方误差(RSE)和相对绝对误差(RAE)作为损失函数,其计算公式为

$$RSE = \frac{\sum_{i=1}^n (p_i - a_i)^2}{\sum_{i=1}^n (\bar{a} - a_i)^2} \quad (19)$$

$$RAE = \frac{\sum_{i=1}^n |p_i - a_i|}{\sum_{i=1}^n |\bar{a} - a_i|} \quad (20)$$

3 实验

3.1 实验设置

实验中,采用 Python 语言和 Pytorch 深度学习框架对模型进行实现,并采用 Adam 优化器设置梯度减切为 5 进行训练。训练参数中,学习率设置为 0.001,为了防止模型过拟合,在每个小模块结束后设置 Dropout 为 0.3,空间卷积层的深度设置为 2,节点嵌入维度设置为 40。所有实验均在具有 8G RAM 的 NVIDIA 2060super GPU 上进行。本文提出的模型与以下 5 种基线方法进行比较。

ARIMA:自回归整合移动平均模型是统计学里经典的自回归模型,其中 VARIMA 可对多元时间序列数据进行预测。

GP^[26]:该模型是使用核方法的高斯模型。核函数的采用保证了高斯过程具有很强的处理非线性问题的能力,能够对时间序列数据进行有效的估计。

RNN-GRU^[27]:该模型是具有全连接的门控循环单元的递归神经网络。

LSTNet-skip:该模型结合了卷积神经网络和循环神经网络.利用一维普通卷积捕捉短期局部信息,利用LSTM捕捉长期宏观信息,对并对输入数据维度整理,捕捉更长期的信息并充分利用序列的周期特性。

TPA-LSTM:该模型是基于注意力的循环神经网络模型,时序注意力机制通过使用CNN提取输入信息

中的定长时序模式,使用评分函数确定各时序模式的权值,再利用LSTM捕捉长期信息。

为了验证ECSGCN模型的有效性,采用RSE、RAE和Correlation评价指标来对比各模型性能。

3.2 实验结果分析

ECSGCN和其他基线模型在相同数据集上进行了周期为3,6,12,24的能耗预测的比较,结果如表1所示.图3-4显示了模型在不同周期下的损失函数.结合表1和图3-4可以获得以下分析结果。

表1 ECSGCN与其他基准模型的性能比较

模型	周期							
	3		6		12		24	
	RSE	CORR	RSE	CORR	RSE	CORR	RSE	CORR
ARIMA	0.099 5	0.884 5	0.103 5	0.863 2	0.105 0	0.859 1	0.105 4	0.859 5
GP	0.150 0	0.867 0	0.190 7	0.833 4	0.163 2	0.839 4	0.127 3	0.881 8
RNN-GPU	0.110 2	0.859 7	0.114 4	0.862 3	0.118 3	0.847 2	0.129 5	0.865 1
LSTNet-skip	0.086 4	0.928 3	0.093 1	0.913 5	0.100 7	0.907 7	0.100 7	0.911 9
TPA-LSTM	0.082 3	0.943 9	0.091 6	0.933 7	0.096 4	0.925 0	0.100 6	0.913 3
ECSGCN	0.075 6	0.947 8	0.084 5	0.936 9	0.094 5	0.929 0	0.097 1	0.925 3

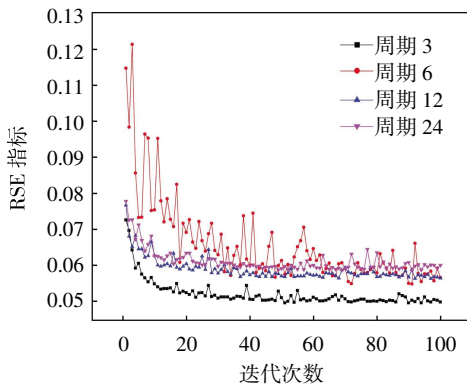


图3 RSE指标对比

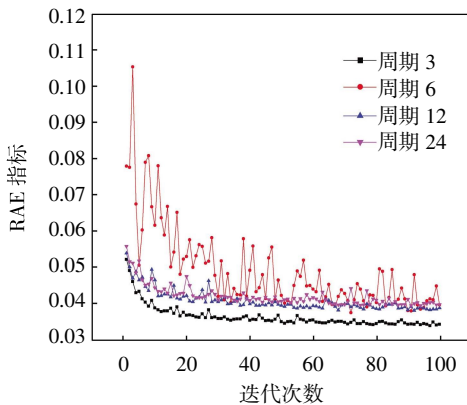


图4 RAE指标对比

基于图神经网络的方法在各项评估指标中均比传统的时间序列分析方法和机器学习方法具有更低的预测误差且表现了更强的相关性.导致这样的结果

主要原因在于ARIMA和GP模型在处理非线性数据时能力有限,并对数据的平稳性要求过高.同时,ECSGCN与深度学习方法相比均有不同程度的提升,ECSGCN不仅擅长对非线性数据进行建模,还可以更好地发掘时间序列数据的复杂关系.因此,基于图神经网络的模型表现出更好的性能。

ECSGCN在长期和短期预测上都实现了比其他深度学习模型更好的预测精度.在周期为3的短期预测中,ECSGCN在RSE上比传统的RNN网络提升了3.5%,比TPA-LSTM网络提升了0.67%;类似的提升在周期为24的长期预测中也同样明显,在RSE上比传统RNN网络提升了3.3%,比TPA-LSTM网络提升了0.35%.这主要因为采用了可学习的方式构建了能耗数据的图结构,能够挖掘时序数据中节点序列之间的关联信息,使得节点的嵌入表达更加丰富,提高了预测精度.同时,模型采用了循环聚合的方法实时融合了节点嵌入,更有效地提高了学习细粒度。

但是随着时间周期的增长,ECSGCN模型的预测效果逐渐降低,如图5所示.这是因为ECSGCN在面对长时间任务时,模型中时间卷积模块的扩张卷积感受野大小会随着内核大小呈线性增长,要处理长时间的序列,则需要更深的网络和更大的内核,在周期阈值内预测精度呈现增长趋势,但是超过阈值则呈现下降趋势.在周期为6时,RSE和RAE表现波动非常大,并且效果并不理想,但是随着epochs提升到100以上

时,仍然可以趋于稳定,同时线性相关性处于预期范围内。

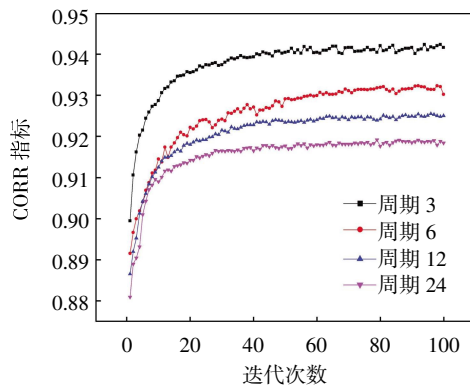


图5 CORR指标对比

4 结论

本文以提高建筑能耗预测精度为研究目标,提出了一种端到端的图神经网络模型。该模型利用可学习的自适应矩阵从时间序列数据中获取节点间的相关性,并引入了图神经网络提取建筑能耗数据的空间依赖关系,实现了能耗数据时空特征融合,并在多项指标上优于现有模型,同时解决了不同周期下的预测问题。考虑到实际的建筑能耗预测问题仍然受到许多隐藏因素的影响,例如天气、温度、季节等,因此,在未来的工作中,将引入更多的额外因素作为研究特征来提高建筑能耗预测精度。

参考文献:

- [1] 王伟兴,康庆华,董帆,等. 国内外能源利用现状分析[J]. 云南化工, 2019, 46(6): 48-49.
- [2] PARVIN K, LIPU M, HANNAN M A, et al. Intelligent controllers and optimization algorithm for building energy management towards achieving sustainable development: challenges and prospects[J]. IEEE Access, 2021, 9(99): 41577-41602.
- [3] MENG D, WU L F. Forecasting the building energy consumption in China using grey model[J]. Environmental Processes: An International Journal, 2020, 7(3): 1009-1022.
- [4] 马丁, 陈文颖. 中国2030年碳排放峰值水平及达峰路径研究[J]. 中国人口资源与环境, 2016, 26(1): 1-4.
- [5] 周海珠, 王雯翥, 魏慧娇, 等. 我国绿色建筑高品质发展需求分析与展望[J]. 建筑科学, 2018, 34(9): 148-153.
- [6] RUI Z, WANG D, YAN R, et al. Machine health monitoring using local feature-based gated recurrent unit network[J]. IEEE Transactions on Industrial Electronics, 2017, 6(99): 1-10.
- [7] LAI G, CHANG W C, YANG Y, et al. Modeling long and short-term temporal patterns with deep neural networks[J]. ACM, 2018, 8(41): 95-104.
- [8] GORI M, MONFARDINI G, SCARSELLI F. A new model for learning in graph domains[J]. IEEE ICDM, 2005, 2(97): 729-734.
- [9] VELOCKOVIC P, CUCURULL G, et al. Graph attention network[J]. Machine Learning, 2017, 1050(20): 10-15.
- [10] WANG J, ZHANG Y, WEI Y, et al. Metro passenger flow prediction via dynamic hypergraph convolution networks[J]. IEEE Transactions on Intelligent Transportation Systems, 2021, 22(12): 7891-7903.
- [11] DERR T, MA Y, TANG J. Signed graph convolutional networks[J]. IEEE International Conference Data Mining, 2018, 41(5): 929-934.
- [12] HEIDER F. Attitudes and cognitive organization[J]. The Journal of Psychology, 1946, 2(1): 107-120.
- [13] CARTWRIGHT D, HARARY F. Structural balance: a generalization of Heider's theory[J]. Psychological Review, 2019, 63(5): 277-282.
- [14] BARTHOLOMEW D J, ZHANG F. Time series analysis forecasting and control[J]. Journal of the Operational Research Society, 2017, 22(2): 199-201.
- [15] SHIH S Y, SUN F K, LEE H Y. Temporal pattern attention for multivariate time series forecasting[J]. Machine Learning, 2019, 108(8): 1421-1441.
- [16] LI L J, WANG H. A supervised similarity measure for link prediction based on KNN[J]. International Journal of Modern Physics C, 2021, 32(9): 1-14.
- [17] GOODGE A, HOOI B, NG S K, et al. Lunar: unifying local outlier detection methods via graph neural networks[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2022, 36(6): 6737-6745.
- [18] CHIANG W L, LIU X, SI S, et al. Cluster-GCN: an efficient algorithm for training deep and large graph convolutional networks[J]. CORR, 2019, 19(10): 257-266.
- [19] CHEN L C, PAPANDEOU G, SIS, MURPHY K, et al. DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs[J]. IEEE Transactions on Pattern Analysis Machine Intelligence, 2018, 40(4): 834-848.
- [20] SZEGEDY C, LIU W, JIA Y, et al. Going deeper with convolutions[J]. IEEE CVPR, 2015, 7(12): 1-9.
- [21] ALOM M Z, HASAN M, YAKOPCIC C, et al. Inception recurrent convolutional neural network for object recognition[J]. Machine Vision and Applications, 2021, 32(1): 1432-1449.
- [22] WU Z, PAN S, LONG G, et al. Connecting the dots: multivariate time series forecasting with graph neural networks[J]. ACM, 2020, 20(11): 753-763.

- [23] 王 杨,郑 津,刘 影.基于自适应门控图神经网络的交通流预测[J].计算机应用研究,2022,3(28):1-6.
- [24] HE Y D,ZHAO J B. Temporal convolutional networks for anomaly detection in time series [J]. Journal of Physics:Conference Series,2019,1213(4):50-59.
- [25] JIE H,LI S,GANG S.Squeeze and excitation networks[J]. IEEE Transactions Pattern Analysis Machine Intelligence,2019,42(8):2011-2023.
- [26] FRIGOLA R.Bayesian time series learning with gaussian processes[D].Cambridge:University of Cambridge,2015.
- [27] XIA M,SHAO H,MA X.A Stacked GRU-RNN based approach predicting renewable energy and electricity load for smart grid operation[J]. IEEE Transaction on Industrial Informatics,2021,17(10):7050-7059.

(上接第 219 页)

- [13] ERDŐS P,RÉNYI A. On the evolution of random graphs[J]. Mathematical Institute of the Hungarian Academy of Sciences, 1960,5(1): 17-60.
- [14] ALBERT R,BARABÁSI A L. Statistical mechanics of complex networks[J]. Reviews of Modern Physics,2002,74(1): 47-48.
- [15] GUIMERÀ R. Networks: an introduction[M]. Oxford: Oxford University Press,2012.
- [16] BARABÁSI A L,ALBERT R J S. Emergence of scaling in random networks[J]. Science,1999,286(5439): 509-512.
- [17] MENDONCA N C,BOX C,MANOLACHE C,et al. The monolith strikes back: why istio migrated from microservices to a monolithic architecture[J]. IEEE Software,2021,38(5): 17-22.
- [18] 杨 平. 服务网格(Service Mesh)简介[J]. 现代电视技术, 2020,13(5): 5-6.
- [19] CALHEIROS R N,RANJAN R,BELOGLAZOV A,et al. CloudSim: a toolkit for modeling and simulation of cloud computing environments and evaluation of resource provisioning algorithms[J]. Software: Practice and Experience,2011,41(1): 23-50.
- [20] CHUN B,CULLER D,ROSCOE T,et al. Planetlab: an overlay testbed for broad-coverage services[J]. ACM Sigcomm Computer Communication Review,2003,33(3): 3-12.