

基于深度学习的时空序列预测方法综述

潘志松, 黎 维

(陆军工程大学指挥控制工程学院, 南京 210007)

摘 要: 随着数据采集技术的蓬勃发展, 各个领域的时空数据不断累积, 迫切需要探索高效的时空数据预测方法。深度学习是一种基于人工神经网络的机器学习方法, 能有效地处理大规模的复杂数据, 因而研究基于深度学习的时空序列预测方法具有十分重要的意义。在这一背景下, 针对已有的预测方法进行归纳和总结, 首先回顾了深度学习在时空序列预测中的应用背景和发展历程, 介绍了时空序列的相关定义、特点及分类; 然后按照时空序列数据的类别介绍了基于网格数据的预测方法、基于图数据的预测方法和基于轨迹数据的预测方法; 最后总结了上述预测方法, 并对当前面临的一些问题及可能的解决方案进行了探讨。

关键词: 深度学习; 数据挖掘; 时空数据; 时空序列预测

中图分类号: TP181 **文献标志码:** A

Review of Spatio-Temporal Sequence Prediction Methods Based on Deep Learning

PAN Zhisong, LI Wei

(Command and Control Engineering College, Army Engineering University of PLA, Nanjing 210007, China)

Abstract: With the vigorous development of data acquisition technology, spatio-temporal data in various fields are accumulating continuously, so it is urgent to explore efficient spatio-temporal data prediction methods. Deep learning is a machine learning method based on artificial neural networks, which can effectively process large-scale complex data. Therefore, it is of great significance to study the spatio-temporal sequence prediction methods based on deep learning. In this context, the existing prediction methods are summarized. First, the application background and development history of deep learning in spatio-temporal sequence prediction are reviewed, and the related definitions, characteristics and classification of spatio-temporal sequence are introduced. Then according to the categories of spatio-temporal sequence data, this paper introduces the prediction methods based on grid data, on graph data, and on trajectory data. Finally, the above prediction methods are summarized, and some current problems and possible solutions are discussed.

Key words: deep learning; data mining; spatio-temporal data; spatio-temporal sequence prediction

引言

近年来,随着计算机软硬件技术的快速进步,如全球定位系统(Global positioning system, GPS)^[1]和互联网^[2]的发展、各类传感器和大容量存储设备的出现,包含地理位置信息的序列数据不断累积。传统的数据挖掘方式难以应付各领域时空序列数据^[3]的爆炸性增长,因而亟需探索新的数据挖掘算法来应对规模不断增长的时空序列数据。时空序列数据通常是指具有空间相关关系的时间序列的集合^[4],图1为东北三省自2001年来18年的居民人均地区生产总值时空序列示意图。

依据已采样的历史时空数据来预测未来某些时刻被观测的时空变量的值的过程被称为时空序列预测^[4]。时空序列预测被广泛应用于多个领域,如交通规划^[5-7]、天气预测^[8-10]、经济统计分析^[11-13]和农业产量预测^[14-15]等领域。从多个角度挖掘出数据中蕴含的时空特征,精准地对时空变量进行预测能有效地为公共管理和个人规划等提供有效的决策和建议,具有重要经济价值和意义。

时空序列数据包含时间、空间和属性3个维度的信息^[16],较为复杂。进行时空序列预测时不仅需要考虑不同时刻间的关联,还需要考虑不同位置在空间上的相互影响,并且这些时空关联往往是动态变化的^[17],这给时空特征的挖掘带来了极大的挑战。现有的时空序列预测方法可大致分为两类:一是基于参数模型的传统预测方法;二是基于数据驱动和深度学习的预测方法^[18]。前者主要包含历史平均法^[19](History average model, HA)、差分自回归移动平均法^[20](Autoregressive integrated moving average model, ARIMA)、时空差分自回归移动平均法^[21](Spatio-temporal autoregressive integrated moving average model, STARIMA)、向量自回归法^[22](Vector autoregressive model, VAR)、决策树法^[23-24](Decision tree model, DT)、隐马尔科夫模型法^[25-26](Hidden Markov model, HMM)和高斯过程方法^[27-28](Gaussian process model, GP)等。后者则依据时空数据的类型可分为面向网格数据的预测方法、面向图数据的预测方法以及面向轨迹数据的预测方法。

历史平均法对时空变量在历史时刻的观测值进行了平均,将均值作为预测结果。此种方法实现简单,但缺乏对时空数据特征的深层次挖掘,在实际应用中往往精度不高;差分自回归移动平均法将时空序列看成多条相互独立的时间序列分别进行预测,然而该模型忽略了时间序列之间的空间关联;时空差分自回归移动平均法虽然考虑了时空关联,但是对于非平稳的时空序列预测效果不佳;向量自回归法是由单变量自回归模型(Autoregressive model, AR)拓展而来,将时空序列视为多维时间序列然后预测,但该模型所需参数较多,需要大量样本;决策树法利用滑动窗口分解,将时空序列预测问题分为一系列固定长度的子问题分别进行预测,该方法无法捕获时空特征的动态变化;隐马尔科夫模型法主要是在时空密度聚类算法ST-DBSCAN^[29-30]的基础上引入了时间维度,从而将时空序列预测转化状态预测问题,该方法假设条件过强,只适合特定的时空数据,泛化性能较低。高斯过程方法主要是通过计算数据的预测后验分布来得到预测值,如果观测的数据模型是非高斯模型,则还需要利用拉普拉斯近似来模拟后验分布。但高斯过程方法十分依赖于高斯核的选取,并且计算代价和空间复杂性都很高。

传统的预测方法在处理非线性、非平稳的时空序列预测问题中往往难以捕获动态的时空特征,在

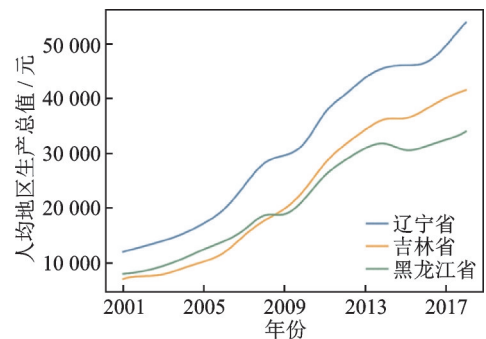


图1 东北三省近18年人均地区生产总值时空序列示意图

Fig.1 Schematic diagram of the spatio-temporal sequence of per capita GDP in the three northeastern provinces in the past 18 years

面对规模与日俱增的时空数据时后继乏力。近年来,基于数据驱动的深度学习方法在计算机视觉^[31-32]、自然语言处理^[33-34]、语音识别^[35]和推荐系统^[36-37]等大部分领域取得了惊人的进展。在时空序列预测领域,基于深度学习的预测方法往往也强于传统的方法,在处理动态变化的场景和长期预测问题上更是具有压倒性的优势。因此,本文主要研究对象是基于深度学习的时空序列预测方法。时空序列数据大致可分为以下3种类型:网格数据、图数据和轨迹数据,因而从数据类别的角度出发可以将对应的深度学习预测算法分为对应的3大类。

本文将从时空序列预测的需求入手,从时空序列的数据特点及其应用、基于深度学习的时空序列预测方法、时空序列预测中存在的难点和未来的研究方向这3个方面出发,系统地分析该领域的发展现状和未来研究方向。

1 时空序列数据特点及实际应用

1.1 数据特点

时空序列数据是在多个区域内对一个或多个研究对象动态变化过程的采样所获得的信息,反映了多个区域内变量随时间不断变化的趋势。其主要包含时间信息(采样的时间戳)、空间信息(采样的地点区域)和属性信息(观测的值,如流量、速度、方向、次数等)。时空序列数据可以通过含定位功能的移动智能设备(手机、智能手表、智能手环等)、GPS设备、打卡点、社交媒体、监控摄像头和站点传感器等方式获取,其来源广泛,复杂多样。广义的时空序列数据中的空间不仅包括物理空间,还包括逻辑空间、网络空间等。

表1列出了6种不同类型的典型时空序列数据集。这些数据集需要经过一系列预处理^[38](筛选、去重、填充、重采样、去噪和变换等)之后才能应用于特定的预测模型之中。由表1可知时空序列具有海量性^[3]、多样性和实时性,此外时空数据还具有以下特点:

(1) 动态性。时空序列数据是一系列带有位置信息和时间戳的采样点序列,序列内部蕴含的特征会随时间不断变化,这给预测增加了许多难度。

(2) 时空相关性。时空序列数据在时间维度和空间维度并非是独立的,而是会相互影响。不同位置的时间序列会相互影响,同一位置的序列在时间维度上也有着时序关联。

(3) 数据噪声多。由于数据来源的多样化,时空序列数据往往存在很多噪声。例如不同设备采样的精度可能不一样、采样的时间维度难以完全对齐、采样的数据类别可能存在一定的不平衡等。

总而言之,时空序列数据广泛存在于现实生活的各个领域,在研究时空序列预测问题中必须充分考虑其特征并对其进行相应的处理,这样才能挖掘出其中的有用信息进而得出准确的预测结果。

表1 6种代表性的时空序列数据集

Table 1 Six representative spatio-temporal sequence datasets

数据集名称	数据类型	获取方式	采样频率	时间跨度	应用领域
BikeNYC	网格数据	站点传感器	1 h	183 d	自行车供需预测
TaxiBJ	网格数据	车载GPS设备	30 min	120 d	出租车供需预测
METR-LA	图数据	高速公路环形探测器	5 min	122 d	速度预测
PEMS-BAY	图数据	道路传感器	5 min	150 d	速度预测
UCY	轨迹数据	单目相机	0.4 s	8 min 28 s	行人检测和跟踪
ETH(Student)	轨迹数据	车载相机	0.4 s	16 min 26 s	行人检测和跟踪

1.2 实际应用

时空序列数据中蕴含大量的信息,通过对其进行建模分析和预测,准确挖掘出这些信息能够有效地辅助人们科学决策、促进产业精细化管理,便捷人们日常生活。如图2所示,时空序列预测问题已经广泛应用在天气预报、出行安排、供需调配、产量预估和股市决策等方面,深刻地影响着人们的生活方式。

(1) 天气预报。天气预报问题是一个复杂而具有挑战性的问题,准确预测天气对农业生产、工业和交通等领域有着极大的意义。在历史的天气数据中包含了丰富的信息,通过挖掘历史气候数据中潜在的规律,在结合最近的天气情况,可以对未来一段时间内的天气进行预测,为人们提供预警,有助于指导人们的生产生活。

(2) 出行安排。随着交通的快速发展,人们的出行频率越来越高,但复杂的交通路况也使得人们难以预知自己的出行时间。通过对历史出行数据的挖掘,可以获取到不同路段在不同时间以各种方式通行所花费的时间,进而为人们出行的决策(是否出行、以何种方式何时出行等)提供参考。

(3) 供需调配。出租车是人们出行的重要交通工具之一,然而很多时候人们都面临着打车困难的问题。通过对以往的出租车在不同位置的供应量和需求量数据进行挖掘分析,找出不同时刻不同位置出租车供需的变化规律。这样能够对出租车的调度进行优化,对其进行合理的分配调度,进而能有效提高用户的打车概率,减少出租车的平均空置等待时间。

(4) 产量预估。粮食安全是我国的核心战略之一,准确预估粮食产量能够有效地避免风险,进而提早作出决策,掌握粮食生产的主动权。依据历史的产量和环境等数据对粮食产量进行预测,并和实际产量对比进行分析,也能够更快找出粮食增产或减产的主要原因,对指导粮食生产具有重要意义。

(5) 股市决策。股票市场的预测是一个非常难的研究课题。将时空序列预测方法引入到股票预测之中,将不同的股票视为不同的节点进而构建预测模型,能够有效地挖掘出股票之间的联系,使得预测结果更加符合市场实际变动。

时空序列预测的应用并不仅仅局限于上述领域,随着各种定位技术和数据采集技术的发展,人们对时空序列数据挖掘的需求不断增高,其应用领域也会随之不断拓展。

2 基于深度学习的时空序列预测方法

近年来深度学习得到了蓬勃的发展,并且在多个领域取得了令人惊叹的成果。将深度学习引入时空序列预测中,能够有效地挖掘到时空大数据中隐含的信息,解决当前“数据过剩,信息匮乏”的困境^[39]。本节主要从时空序列数据的角度出发,依次介绍了面向网格数据、面向图数据和面向轨迹数据的时空序列预测算法。

2.1 相关定义

在介绍相关算法之前,首先有必要给出时空序列和时空序列预测的定义。

时空序列:一系列具有空间关联的时间序列的集合。将目标空间 χ 划分为 n 个子区域,分别记为 $\chi_1, \chi_2, \dots, \chi_n$,第 i 个子区域 χ_i 内的变量在时刻 j 的观测值为 $x_{i,j}(i \in \{1, 2, \dots, n\}, j \in \{1, 2, \dots, t\})$, t 为观测时间,区域 χ_i 对应的时间序列可表示为 $X_{i,1:t} = \{x_{i,1}, x_{i,2}, \dots, x_{i,t}\}$,则时空序列(Spatio-temporal se-

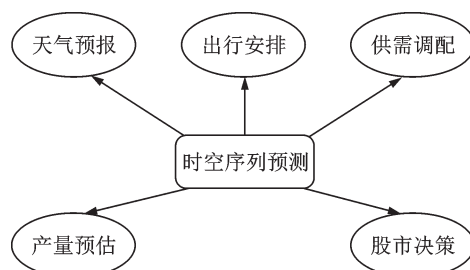


图2 时空序列预测的应用

Fig.2 Applications of spatio-temporal sequence prediction

quence, STS)可表示为

$$\text{STS} = \{X_{i,1:t} | i \in \{1, 2, \dots, n\}\} \quad (1)$$

时空序列预测:依据历史的时空序列,对未来某一段时间内时空变量的值进行估计的过程。已知历史时空序列 $\text{STS} = \{X_{i,1:t} | i \in \{1, 2, \dots, n\}\}$, 预测观测时间 t 后 L 个时刻时空变量的值, 即

$$\bar{X}_{:,t+1:t+L} = \arg \max p(X_{:,t+1:t+L} | X_{:,1:t}, E_{1:t}) \quad (2)$$

式中 $E_{1:t}$ 为时刻 1 到 t 的其他特征。

2.2 面向网格数据的预测算法

2.2.1 网格数据建模

在时空序列预测问题中,可将目标空间 χ 划分为一个包含 n 个地点 $P \times Q$ 的网格图,其中 $n = P \times Q$, 则在每一个时刻所有采样点可以视为一张图片 $P \times Q \times D$ 上的像素值,其中 P, Q, D 依次为该图片的宽度、高度和通道数(该通道数 D 对应每个采样点取值的维度数)。因此,计算机视觉领域的各种算法可以由此迁移到时空序列预测问题中来提取 n 个地点间的空间关系。

例如,在北京出租车数据集(TaxiBJ)中,将同一时刻不同节点的人口流入量建模成一张 $32 \text{像素} \times 32 \text{像素}$ 的图片(实际上应该为 $32 \text{像素} \times 32 \text{像素}$,为了简化省略了最后一个维度),则 2013 年 7 月 1 日 0:00~0:30 的流入人流量对应图片用热力图可视化如图 3 所示。其中,图右侧的长条表示热力图颜色和人群流入数量的对应关系,热力图的纵横坐标(取值范围 0~31)用于确定每个特定的节点区域。

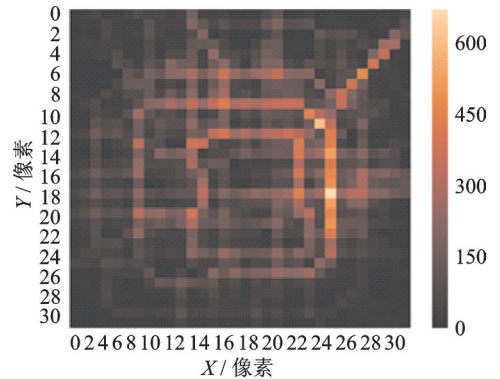


图3 北京 2017 年 7 月 1 日前半小时流入人流量热力图

Fig.3 Heatmap of inflow of people in the first half hour of July 1, 2017 in Beijing

2.2.2 面向网格时空数据的深度预测算法

时空序列数据由时间序列拓展而来,现有的时间序列预测方法如循环神经网络^[40](Recurrent neural network, RNN)、长短期记忆网络^[41](Long short-term memory, LSTM)、门控循环单元网络^[42](Gated recurrent unit, GRU)和时间卷积网络^[43](Temporal convolutional network, TCN)等可以提取出数据中的时序特征,但是在将其应用到时空序列预测中时却往往忽略了时间序列之间的空间关联。Shi 等^[44]在 2015 年提出了 ConvLSTM 模型来解决这个问题。ConvLSTM 模型是卷积神经网络^[45](Convolutional neural network, CNN)和循环神经网络 LSTM 的结合体,通过对全连接 LSTM 网络^[46](Fully connected LSTM, FC-LSTM)进行拓展,将 LSTM 中的一维张量拓展为三维张量(对应的点乘拓展为卷积操作),因而能够捕获到数据中的时空特征,该模型当时在利用雷达回波数据进行降雨量预测的问题中取得了很好的效果。ConvLSTM 模型的思想启发了将 CNN 结构引入时空序列预测中的一系列模型。

为了挖掘网格时空数据间的空间特征,可以将其建模为一系列图片序列,然后再用卷积神经网络、残差神经网络^[47](ResNet)等来提取空间特征。Zhang 等^[48]2016 年在 DeepST 模型中首次将这个思路应用到了交通流量预测这个时空序列预测领域的代表性问题中。DeepST 将每个时刻的交通流量建模成一张图片,利用不同的时间尺度(半小时、一天、一周)分别对数据采样得到 3 个序列,利用相同的 CNN 结构对这 3 个序列分别进行特征提取,之后对得到的这 3 个特征进行加权融合最后得到预测结果。2017 年, Zhang 等^[49]对 DeepST 模型提出了改进,在其卷积的基础上增加了残差模块,最后融入外部特

征构建了ST-ResNet模型。通过引入残差模块,ST-ResNet确保了预测精度不会因为卷积模块层数的增加而降低。

DeepST和ST-ResNet虽然能提取到空间特征,但是对于时间特征的提取较为粗略。在此基础上,Li等^[50]于2019年提出了ST-DCCNAL模型。该模型在将同一时刻不同位置的时空变量建模为图片的基础上引入了LSTM结构。针对空间特征的提取,该模型引入了DenseNet^[51]来解决CNN在传播过程中存在信息损失的问题。ST-DCCNAL先利用DenseNet结构提取每个时刻的空间特征,再利用LSTM提取时间特征形成时空特征。考虑到不同时刻的历史数据对待预测值的影响大小不一样(时间间隔越近,影响越大),ST-DCCNAL模型还在LSTM中添加了注意力机制^[52],自动学习到了不同时刻对应的不同权重。

实际应用中,空间距离相近的区域往往具有相似的时间模式变化规律,因此Yao等^[53]提出了一个多视角的模型DMVST-Net来解决这个问题。该模型从空间视角、时间视角和语义视角来提取时空特征,其中空间和时间视角分别利用了局部CNN和LSTM来进行特征提取。语义视角则是通过构建一个加权图(节点之间边的权重代表不同区域之间功能的相似程度),然后将其映射编码到隐空间提取到语义特征,之后将语义特征和之前提取到的时空特征融合最后得到预测结果。

上述模型都有一个假设,那就是区域之间的空间关系是固定不变的。然而在实际中空间区域节点之间的空间关联往往会随着时间而变化,例如工作日和周末时商业区域和住宅区的人流量变化规律差异较大。为了捕获到这种动态的时空关联,Yao等^[54]提出了STDN模型。在提取每个时刻的空间关联时,STDN利用当前时刻的人流移动方向矩阵和人流矩阵点乘进行加权,从而刻画了每个时刻的空间关联。此外,该模型还利用了一个带有注意力机制的LSTM提取每天中蕴含的时间关联,再利用一个LSTM提取天与天之间的时间特征,从不同粒度挖掘出了数据的时空特征。

2.3 面向图数据的预测算法

尽管上述的时空序列预测算法较好地捕获了数据中动态的时空关联,但这些方法只适用于网格时空数据,即欧式空间的时空数据(如分别在一维、二维和三维欧几里得空间的文本、图片和视频等)。在实际生活中,大部分时空数据都是非欧几里得空间数据,例如知识图谱^[55]、社交网络^[56]、化合物分子结构^[57]等。这些非欧数据往往无法建模成一系列图片,强行建模成图片序列会极大地降低预测的精度。因此面向图数据的时空序列预测算法逐渐成了研究的重点和难点。

2.3.1 图数据建模

在面向图数据时空序列预测问题中,将时空序列定义在图上。对于时刻 t 在节点 i 的变量观测向量 $X_{i,t}$ 并非是独立的,而是与图中其他每个节点都有一个成对的连接。因此观测向量 $X_{i,t}$ 可以视为节点 i 定义在图 G (无向图或有向图)上的图信号。第 t 个时刻,图可表示为 $G_t=(X_{:,t}, \xi, \mathbf{W})$,其中 $X_{:,t}$ 是 t 时刻所有 n 个节点特征的集合, ξ 是边的集合,表示了节点之间的连通性, $\mathbf{W} \in \mathbf{R}^{n \times n}$ 是图 G_t 对应的加权邻接矩阵。

2.3.2 面向时空图数据的预测算法

最早将图卷积网络^[58]应用于时空序列预测问题的是北京大学的于斌,他在2018年的IJCAI会议论文中提出了STGCN^[59]算法。该算法将问题建模在图上,摒弃了RNN结构而全部用卷积结构来替代。该模型利用一个门控线性单元和一维卷积结构构建成一个时序门控卷积模块(Temporal gated-conv, TGC),然后将一个空间图卷积模块(Spatial graph-conv, SGC)夹在两个时序门控卷积模块之间形成一个基础的时空卷积单元。通过堆叠这些时空卷积单元,并在之间增加残差模块,STGCN有效地提取到

了时空图数据中的中长期时空特性,并较好地减少了模型参数,增加了训练速度。

STGCN模型尽管因为未使用RNN结构而训练较为迅速,但其对时间特征的提取较为粗略。基于此,2019年Guo等^[60]提出了ASTGCN模型,该算法在基础时空特征提取模块之中引入了注意力机制。在时间特征和空间特征提取时加入注意力机制使得该模型能够捕获动态而复杂点的时空关联。此外,ASTGCN模型还引入了类似于DeepST中的不同时间尺度(半小时、一天、一周)的子序列,分别堆叠基础时空特征提取模块后将特征融合后预测,使得模型提取的时间特征更加符合实际变化规律。

同样,针对道路网络拓扑结构的非欧性,Zhao等^[61]提出了T-GCN模型。该模型和ST-DCCNAL模型类似,都是先对每个时刻的空间特征提取,再利用RNN结构捕获空间特征序列中蕴含的时间关联。不同的是,T-GCN使用了图卷积神经网络挖掘非欧数据中的空间特性,利用了参数更加精简的GRU替代了LSTM模块。最后在两个实际数据集上验证了T-GCN模型的有效性。

STGCN在空间特征提取时只利用了一个邻接矩阵,但实际中除了物理空间距离相近的区域有类似的时间变化模式之外,如果两个区域功能相似,或者两个区域之间有交通间接连接,那么这两个区域对应的时间序列应该也有相近的变化规律。为了解决这个问题,Geng等^[62]提出了一个时空多图图卷积神经网络STMGCN。该模型从区域邻近性、功能相似性和连通性3个角度分别构建了一个图,然后在每个图中都引入了一个环境门控RNN来提取时间特征,再利用GCN提取特征并将3个图得到的时空特征进行融合得到预测结果。

在实际应用中,道路网络时空数据往往是非平稳的,并且缺乏环境信息,为了增强模型预测性能,Dai等^[63]在2020年提出了HSTGCN。HSTGCN旨在预测两个地点之间的通行时间,为导航应用提供后台算法支撑。为了解决环境信息匮乏的问题,该模型除了利用历史的通行时间时空序列之外,还引入了和通行时间强相关的交通流量数据。利用一个Transformer结构对交通流量数据进行特征提取,然后与通行时间数据分别送入一个时序门控卷积模块中提取其时间特征并融合。之后再将得到的特征送入GCN网络中之后再利用时序门控卷积模块提取时间特征最后得到预测结果。HSTGCN可以看作是只有一个基础时空卷积单元的STGCN模型,不同的是HSTGCN增加了交通流量数据来辅助进行预测,此外还对原有的邻接矩阵进行了重定义使得交通固有的邻近性得到了更好的表示。

尽管和面向网格时空数据的深度预测算法处理的数据类型不同,这两类算法在某些方面有共同的特点。这两者的一些特征提取模块可以相互借鉴迁移,一些相应的技巧也是通用的。

2.4 面向轨迹数据的预测算法

行人轨迹预测广泛应用于自动驾驶^[64]和机器人导航^[28]等领域。准确地预测行人的轨迹使得自动驾驶的车辆、机器人等能及时地避开障碍,对于它们的安全行驶具有重大意义。在传统的轨迹预测方法中,有的利用贝叶斯推断来研究周边环境进而预测行人移动意图^[65];有的从人的角度出发,利用社会力模型构建Social Force^[66]模型,通过研究人的社会属性来预测行人的轨迹。本节主要介绍基于深度学习的时空轨迹预测算法。

2.4.1 轨迹数据建模

轨迹预测是依据预测对象历史的轨迹位置预测未来一段时间内位置的过程。对于场景中的每一个研究对象 $i(i \in \{1, 2, \dots, n\})$,其在时刻 t 所处的位置可用坐标 $p_{i,t} = (x_{i,t}, y_{i,t})$ 表示。给定研究对象历史的位置序列 $p_{i,t}(t \in \{1, 2, \dots, T\})$,预测未来 L 个时刻他们的位置 $p_{i,T+1:T+L}$,未来第 k 个时刻的位置可表示为 $\bar{p}_{i,T+k} = (\bar{x}_{i,T+k}, \bar{y}_{i,T+k})$ 。

2.4.2 面向时空轨迹数据的深度预测算法

轨迹预测可以视为一个序列生成任务,经典的时间序列处理算法如RNN、LSTM等可以被用在轨迹预测之中。将每个行人的轨迹视为一条独立的时间序列,虽然可以提取到各自的时间特征,但彼此之间的空间关联却被忽略了。为了解决这个问题,Alahi等^[67]于2016年提出了Social-LSTM模型。该模型将每个行人视为一条时间序列,用LSTM结构来提取各自的时间特征。同时Social-LSTM引入了一个“Social”池层共享LSTM的隐藏状态来模拟行人之间的交互。通过将特定半径内所有LSTM的隐藏状态汇集在一起作为下一时间步的输入,该模型能自动学习在时间上一致的轨迹之间发生的交互。

尽管Social-LSTM能够依据邻近性来刻画行人间的交互,但这并不完全符合实际情况。实际上不同距离不同方向的行人对自动驾驶的汽车(或机器人)的影响大小也不一样。因而Vemula等^[68]提出了Social Attention模型,该模型利用注意力机制来捕获不同人的相对重要程度,根据神经网络的隐藏状态所代表的运动模式来分配不同的权重。与基于池化机制的模型相比,基于注意力的模型能获得更丰富的交互信息,实现更准确的预测。

行人轨迹不仅受行人本身影响,还会和周围的目标进行交互。之前的方法都是利用各种集结方法将学习到的不同行人状态融合在一起,Social-STGCNN^[69]则通过构建图的方式替代了这个集结方法。通过GCN提取到一个时空的隐状态,然后送入到一个时间外推卷积神经网络中得到预测结果。实验表明,Social-STGCNN模型不仅提高了预测精度,还有效地提高了数据利用率,仅用五分之一的数据就超越了之前的最好结果。

为了更充分地利用行人交互信息,周于涛等^[70]提出了一个基于LSTM和GCN的模型,该模型利用图卷积网络中的节点表示行人的运动信息,节点之间的边表示行人之间的交互并且利用了自注意力机制来计算行人间的交互程度。此外该模型还捕获了行人间的伴随行为特征,并在损失函数中体现了这一点,实验表明该模型相比于之前方法预测性能有了较大的提升。

为了进一步对各种时空序列预测算法的特征提取进行对比,表2列出了各类预测算法中一些代表性算法的特征提取比较。

本文提及的基于深度学习的各种预测方法总结如表3所示。

3 下一步研究方向

基于深度学习的时空序列预测方法虽然在不断发展,但目前仍然存在许多问题。针对一个应用领域的时空序列预测方法如何迁移到别的应用领域中,目前也缺乏一个统一的框架来对不同领域的各种预测算法进行性能评估比较;建模时往往默认为数据的分布不会随着时间变化,但实际上时空变量受

表2 部分时空序列预测算法特征提取对比

Table 2 Feature extraction and comparison of partial spatio-temporal sequence prediction algorithms

模型名称	参数模型	时间特征	空间特征	时空关联	外部特征
HA ^[19]	✓	✓			
ARIMA ^[20]	✓	✓			
STARIMA ^[21]	✓	✓			
DeepST ^[48]			✓		✓
DMVST ^[53]		✓	✓	✓	✓
STGCN ^[59]		✓	✓	✓	
STMGCN ^[62]		✓	✓	✓	
Social-STGCNN ^[69]		✓	✓	✓	✓

表3 基于深度学习的时空序列预测方法总结

Table 3 Summary of spatio-temporal sequence prediction methods based on deep learning

数据类型	方法名称	简要介绍
网格数据	ConvLSTM ^[44]	将LSTM中的一维张量拓展到三维,对应的点乘换为卷积操作
	DeepST ^[48]	将时空序列建模成图片序列,并从不同时间尺度进行采样后融合
	ST-ResNet ^[49]	将残差模块引入时空数据对应的图片序列卷积操作当中
	ST-DCCNAL ^[50]	利用DenseNet提取空间特征,后再利用带注意力机制的LSTM提取时间特征
	DMVST-Net ^[53]	从时间、空间、语义3个视角对时空序列进行建模
	STDN ^[54]	利用带门控机制的CNN挖掘动态的空间关联,并堆叠两个LSTM从不同层次挖掘时间特征
图数据	STGCN ^[59]	首次将GCN引入时空序列预测问题,利用时序门控卷积和空间图卷积构建基础的时空卷积单元,之后再堆叠和残差连接
	ASTGCN ^[60]	在STGCN的基础上引入注意力机制,同时对数据以不同采样间隔进行重采样
	T-GCN ^[61]	先利用GCN捕获空间特性,再送入GRU中提取时间关联
	STMGCN ^[62]	从区域邻近性、功能相似性和连通性3个视角分别构建GCN
	HSTGCN ^[63]	引入和预测变量强相关的时空变量辅助进行预测,并设计了新的邻接矩阵
轨迹数据	Social-LSTM ^[67]	利用不同的LSTM捕获行人的时序特征,并将LSTM的隐状态进行共享来刻画行人间的交互
	Social attention ^[68]	利用注意力机制来获取不同目标的重要程度,捕获更丰富的交互信息
	Social-STGCNN ^[69]	构建图来替代行人状态的信息融合,较好地提高了数据利用率

到环境的影响,当环境变化时时空序列数据的分布也有可能发生漂移,因此需要建立一个自适应的时空序列处理框架来应对变化更加复杂的时空数据。此外,基于深度学习的模型严重依赖于大量数据去训练,对于某些关键数据匮乏的区域则力有未逮,因而需要探索新的方式来解决数据匮乏问题;在实际应用中,某些领域对于模型的规模和实时性要求较高,如何简化模型,提高预测的时间效率也是研究的重点。综上所述,基于深度学习的时空序列预测算法还可以沿下面的方向进一步研究:

(1) 建立统一框架。可以从数据特征出发,建立一个统一的框架,例如将模型分为时间特征提取、空间特征提取、环境特征提取3大部分,具体提取的方式又可以往下细分。这样不仅可以及时空序列预测模型迁移到不同的应用领域,也方便下一步对各个模型进行评估比较。此外,一个统一的预测框架还能够指导创建新的模型,对时空序列预测领域有着极大的意义和价值。

(2) 解决数据分布漂移问题。从数据分布的视角出发,已有的预测方法都是建立在数据独立同分布的假设之上,这也是当前阻碍模型预测精度提高的瓶颈之一。针对时空序列数据的分布随时间变化的实际情况,可以考虑利用自适应学习的方式来解决数据分布漂移问题。

(3) 解决数据匮乏问题。针对关键数据匮乏问题,一是考虑利用迁移学习的方式,将数据充足且有相似点的区域学到的模型迁移到数据匮乏的区域;二是进行数据增广,将研究数据进行重采样等变换,也可将与研究对象相关的其他数据引入进来辅助进行预测。

(4) 模型压缩和实时性。为了满足模型能在轻量化的设备上使用的要求,需要对模型进行剪枝压缩,在保证一定精度的情况下对模型做出简化。对于某些对实时性要求较高的应用场景,可以考虑将在线学习融入到时空序列预测之中,进而使模型能够及时更新和响应预测。

4 结束语

本文首先阐述了时空序列的背景,简要介绍了传统的时空序列预测方法,随后对时空序列的数据特点和实际应用做了详尽的描述,然后介绍了在网格数据、图数据和轨迹数据这3类主要的时空数据中如何利用深度学习来进行预测。时空序列预测在现实生活的各领域都起着重要作用,如何利用深度学习进一步挖掘出时空数据中的隐含特征,提高预测的精度是研究者们需要努力的方向。

参考文献:

- [1] NEREM R S L, LARSON K M. Global positioning system, theory and practice, 5th edition[J]. Eos, Transactions American Geophysical Union, 2001, 82(33): 365-365.
- [2] 中国互联网络信息中心. 第32次中国互联网络发展状况统计报告[J]. 互联网天地, 2013(10): 74-91.
China Internet Network Information Center. The 32nd statistical report on Internet development in China [J]. Internet World, 2013(10): 74-91.
- [3] 王佳璆,邓敏,程涛,等. 时空序列数据分析和建模[M]. 北京: 科学出版社, 2012.
WANG Jiaqiu, DENG min, CHENG Tao, et al. Spatiotemporal data analysis and modeling [M]. Beijing: Science Press, 2012
- [4] 黎维,陶蔚,周星宇,等. 时空序列预测方法综述[J]. 计算机应用研究, 2020, 37(10): 2881-2888.
LI Wei, TAO Wei, ZHOU Xingyu, et al. Review of spatiotemporal series prediction methods [J]. Computer Application Research, 2020, 37 (10): 2881-2888.
- [5] LV Y, DUAN Y, KANG W, et al. Traffic flow prediction with big data: A deep learning approach[J]. IEEE Transactions on Intelligent Transportation Systems, 2015, 16(2): 865-873.
- [6] 王艳辉,贾利民,王卓,等. 基于时空序列的铁路客运量数据挖掘预测[J]. 中国铁道科学, 2005, 26(4): 130-135.
WANG Yanhui, JIA Limin, WANG Zhuo, et al. Data mining and forecasting of railway passenger volume based on temporal and spatial series [J]. China Railway Science, 2005, 26 (4): 130-135.
- [7] POLSON N G, SOKOLOV V O. Deep learning for short-term traffic flow prediction[J]. Transportation Research Part C Emerging Technologies, 2017, 79: 1-17.
- [8] OLAIYA F, ADEYEMO A B. Application of data mining techniques in weather prediction and climate change studies[J]. International Journal of Information Engineering & Electronic Business, 2012, 4(1): 51-59.
- [9] RADZUAN N, OTHMAN Z, BAKAR A A. Uncertain time series in weather prediction [J]. Procedia Technology, 2013, 11: 557-564.
- [10] 谢考宪. 山东省降水量时空分布的诊断分析与未来趋势预测[J]. 山东气象, 1992(1): 31-34.
XIE Kaoxian. Diagnostic analysis and future trend prediction of precipitation in Shandong Province [J]. Shandong Meteorology, 1992(1): 31-34.
- [11] AUGUSTIN N H, BORCHERS D, CLARKE E D, et al. Spatiotemporal modelling for the annual egg production method of stock assessment using generalized additive models[J]. Canadian Journal of Fisheries & Aquatic Sciences, 1998, 55(12): 2608-2621.
- [12] PU G, WENG Y. Value-at-risk forecasts by a spatiotemporal model in Chinese stock market[J]. Physica A: Statistical Mechanics and Its Applications, 2016, 441: 173-191.
- [13] KIM K J. Financial time series forecasting using support vector machines[J]. Neurocomputing, 2003, 55(1/2): 307-319.
- [14] LETSON D, LACIANA C, BERT F E, et al. Value of perfect ENSO phase predictions for agriculture: Evaluating the impact of land tenure and decision objectives[J]. Climatic Change, 2009, 97(1/2): 145-170.
- [15] CAO L, SAN X, ZHAO Y, et al. The application of the spatio-temporal data mining algorithm in maize yield prediction[J]. Mathematical & Computer Modelling, 2013, 58(3/4): 507-513.
- [16] 张山山. 地理信息系统时空数据建模研究及应用[D]. 成都: 西南交通大学, 2001.
ZHANG Shanshan. Research and application of GIS spatiotemporal data modeling [D]. Chengdu: Southwest Jiaotong

University, 2001.

- [17] ZHU Q, CHEN J, ZHU L, et al. Wind speed prediction with spatio-temporal correlation: A deep learning approach[J]. *Energies*, 2018, 11(4): 705.
- [18] SHI X, YEUNG D Y. Machine learning for spatiotemporal sequence forecasting: A survey [J]. arXiv preprint arXiv: 180806865, 2018.
- [19] HOBEIKA A G, KIM C K. Traffic-flow-prediction systems based on upstream traffic[C]//Proceedings of VNIS'94-1994 Vehicle Navigation and Information Systems Conference. [S.l.]: IEEE, 1994: 345-350.
- [20] BROCKWELL P J, DAVIS R A, BERGER J O, et al. Time series: Theory and methods[M]. [S.l.]: Springer, 2015.
- [21] MARTIN R L, OEPPEN J. The identification of regional forecasting models using space: Time correlation functions[J]. *Transactions of the Institute of British Geographers*, 1975(1): 95-118.
- [22] CHRISTIANO L J, CHRISTOPHER A. Sims and vector autoregressions[J]. *The Scandinavian Journal of Economics*, 2012, 114(4): 1082-1104.
- [23] QUINLAN J R. Induction of decision trees[J]. *Machine Learning*, 1986, 1(1): 81-106.
- [24] ROKACH L, MAIMON O. Decision trees[M]. *Data Mining and Knowledge Discovery Handbook*. [S.l.]: Springer, 2005: 165-192.
- [25] BAUM L E, PETRIE T. Statistical inference for probabilistic functions of finite state Markov chains[J]. *The Annals of Mathematical Statistics*, 1966, 37(6): 1554-1563.
- [26] BAUM L E, PETRIE T, SOULES G, et al. A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains[J]. *The Annals of Mathematical Statistics*, 1970, 41(1): 164-171.
- [27] ROBERT C. Machine learning: A probabilistic perspective [M]. USA:MIT Press, 2014.
- [28] TRAUTMAN P, KRAUSE A. Unfreezing the robot: Navigation in dense, interacting crowds[C]// Proceedings of the 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems.[S.l.]: IEEE, 2010: 797-803.
- [29] BIRANT D, KUT A. ST-DBSCAN: An algorithm for clustering spatial-temporal data [J]. *Data & Knowledge Engineering*, 2007, 60(1): 208-221.
- [30] ESTER M, KRIEGL H P, SANDER J, et al. A density-based algorithm for discovering clusters in large spatial databases with noise[C]// Proceedings of 2nd International Conference on Knowledge Discovery and Data Mining. [S.l.]:[s.n.], 1996: 226-231.
- [31] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2017, 39(6): 1137-1149.
- [32] HE K, ZHANG X, REN S, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 37(9): 1904-1916.
- [33] DEVLIN J, CHANG M W, LEE K, et al. Bert: Pre-training of deep bidirectional transformers for language understanding[J]. arXiv preprint arXiv:181004805, 2018.
- [34] COLLOBERT R, WESTON J. A unified architecture for natural language processing: Deep neural networks with multitask learning[C]//Proceedings of the 25th International Conference on Machine Learning.[S.l.]:[s.n.], 2008: 160-167.
- [35] GRAVES A, MOHAMED A-R, HINTON G. Speech recognition with deep recurrent neural networks[C]// Proceedings of the 2013 IEEE International Conference on Acoustics, Speech and Signal Processing.[S.l.]:IEEE, 2013: 6645-6649.
- [36] SARWAR B, KARYPIS G, KONSTAN J, et al. Item-based collaborative filtering recommendation algorithms[C]// Proceedings of the 10th International Conference on World Wide Web.[S.l.]:[s.n.], 2001: 285-295.
- [37] RICHARDS S, AZIZ N, BALE S, et al. Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American college of medical genetics and genomics and the association for molecular pathology[J]. *Genetics in Medicine*, 2015, 17(5): 405-423.
- [38] 刘明吉, 王秀峰, 黄亚楼. 数据挖掘中的数据预处理[J]. *计算机科学*, 2000(4): 56-59.
- LIU Mingji, WANG Xiufeng, HUANG Yalou. Data preprocessing in data mining[J]. *Computer Science*, 2000(4): 56-59.

- [39] 康军, 黄山, 段宗涛, 等. 时空轨迹序列模式挖掘方法综述 [J]. 计算机应用, 2021(4): 1-9.
KANG Jun, HUANG Shan, DUAN Zongtao, et al. Overview of methods for mining spatiotemporal trajectory sequence patterns [J]. Journal of Computer Applications, 2021(4): 1-9.
- [40] MIKOLOV T, KARAFIÁT M, BURGET L, et al. Recurrent neural network based language model[C]//Proceedings of the Eleventh Annual Conference of the International Speech Communication Association.[S.l.]:[s.n.], 2010.
- [41] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural Computation, 1997, 9(8): 1735-1780.
- [42] CHUNG J, GULCEHRE C, CHO K, et al. Empirical evaluation of gated recurrent neural networks on sequence modeling[J]. arXiv preprint arXiv:14123555, 2014.
- [43] BAI S, KOLTER J Z, KOLTUN V. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling [J]. arXiv preprint arXiv:180301271, 2018.
- [44] SHI X, CHEN Z, WANG H, et al. Convolutional LSTM network: A machine learning approach for precipitation nowcasting [C]//Proceedings of the 28th International Conference on Neural Information Processing Systems-Volume 1. [S.l.]: ACM, 2015: 802-810.
- [45] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.
- [46] GRAVES A. Generating sequences with recurrent neural networks[J]. Computer Science, 2013(1): 1-43.
- [47] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.[S.l.]:IEEE, 2016: 770-778.
- [48] ZHANG J, ZHENG Y, QI D, et al. DNN-based prediction model for spatio-temporal data[C]//Proceedings of the 24th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems.[S.l.]:ACM, 2016: 1-4.
- [49] ZHANG J, ZHENG Y, QI D. Deep spatio-temporal residual networks for citywide crowd flows prediction[C]//Proceedings of the AAAI Conference on Artificial Intelligence.[S.l.]:AAAI, 2017.
- [50] LI W, TAO W, QIU J, et al. Densely connected convolutional networks with attention LSTM for crowd flows prediction[J]. IEEE Access, 2019, 7: 140488-140498.
- [51] HUANG G, LIU Z, VAN DER MAATEN L, et al. Densely connected convolutional networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.[S.l.]:IEEE, 2017: 4700-4708.
- [52] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]//Proceedings of the 31st International Conference on Neural Information Processing Systems.[S.l.]:[s.n.], 2017: 6000-6010.
- [53] YAO H, WU F, KE J, et al. Deep multi-view spatial-temporal network for taxi demand prediction[C]//Proceedings of the AAAI Conference on Artificial Intelligence.[S.l.]:AAAI, 2018.
- [54] YAO H, TANG X, WEI H, et al. Revisiting spatial-temporal similarity: A deep learning framework for traffic prediction[C]// Proceedings of the AAAI Conference on Artificial Intelligence.[S.l.]:AAAI, 2019: 5668-5675.
- [55] WANG Z, ZHANG J, FENG J, et al. Knowledge graph embedding by translating on hyperplanes[C]//Proceedings of the AAAI Conference on Artificial Intelligence.[S.l.]:AAAI, 2014.
- [56] SCOTT J. Social network analysis[J]. Sociology, 1988, 22(1): 109-127.
- [57] HERZBERG G. Molecular spectra and molecular structure-Vol I [M]. [S.l.]: Read Books Ltd, 2013.
- [58] BRUNA J, ZAREMBA W, SZLAM A, et al. Spectral networks and deep locally connected networks on graphs[C]// Proceedings of 2nd International Conference on Learning Representations.[S.l.]: ICLR, 2014.
- [59] YU B, YIN H, ZHU Z. Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting[C]// Proceedings of the 27th International Joint Conference on Artificial Intelligence.[S.l.]:[s.n.], 2018: 3634-3640.
- [60] GUO S, LIN Y, FENG N, et al. Attention based spatial-temporal graph convolutional networks for traffic flow forecasting[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2019, 33: 922-929.
- [61] ZHAO L, SONG Y, ZHANG C, et al. T-GCN: A temporal graph convolutional network for traffic prediction[J]. IEEE Transactions on Intelligent Transportation Systems, 2019, 21(9): 3848-3858.

- [62] GENG X, LI Y, WANG L, et al. Spatiotemporal multi-graph convolution network for ride-hailing demand forecasting[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2019, 33: 3656-3663.
- [63] DAI R, XU S, GU Q, et al. Hybrid spatio-temporal graph convolutional network: Improving traffic prediction with navigation data[C]//Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. [S.l.]: ACM, 2020: 3074-3082.
- [64] YURTSEVER E, LAMBERT J, CARBALLO A, et al. A survey of autonomous driving: Common practices and emerging technologies[J]. IEEE Access, 2020, 8: 58443-58469.
- [65] BEST G, FITCH R. Bayesian intention inference for trajectory prediction with an unknown goal destination[C]// Proceedings of the 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). [S.l.]:IEEE, 2015: 5817-5823.
- [66] HELBING D, MOLNAR P. Social force model for pedestrian dynamics [J]. Physreve, 1995, 51(5): 4282.
- [67] ALAHI A, GOEL K, RAMANATHAN V, et al. Social LSTM: Human trajectory prediction in crowded spaces[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.[S.l.]:IEEE, 2016: 961-971.
- [68] VEMULA A, MUELLING K, OH J. Social attention: Modeling attention in human crowds[C]// Proceedings of the 2018 IEEE international Conference on Robotics and Automation (ICRA).[S.l.]: IEEE, 2018: 4601-4607.
- [69] MOHAMED A, QIAN K, ELHOSEINY M, et al. Social-STGCNN: A social spatio-temporal graph convolutional neural network for human trajectory prediction[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. [S.l.]:IEEE, 2020.: 14424-14432.
- [70] 周于涛,吴华意,成洪权,等. 结合自注意力机制和结伴行为特征的行人轨迹预测模型[J]. 武汉大学学报(信息科学版), 2020, 45(12): 164-171.
ZHOU Yutao, WU Huayi, CHENG Hongquan, et al. Pedestrian trajectory prediction model combining self-attention mechanism and companion behavior characteristics[J]. Journal of Wuhan University (Information Science Edition), 2020, 45 (12): 164-171.

作者简介:



潘志松(1973-),通信作者,男,教授,研究方向:模式识别与机器学习,E-mail: hotpzs@hotmail.com。



黎维(1995-),男,博士研究生,研究方向:时空序列预测、深度学习,E-mail: 1280358009@qq.com。

(编辑:夏道家)