

如何正确的预测乘客需求是智慧城市系统与交通调度系统中不可缺少的一部分,能正确的预测乘客需求有助于改善城市的交通系统,并增强用户的出行体验。本文提出一种基于SVR时空网格模型的出租车需求预测方法,建立了出租车需求序列的时间关系和空间关系,将GPS数据转换为时空网格的流量网格数据,并输入SVR时空网格模型,从而预测出城市特定区域的出租车需求。本文采用深圳市的出租车的GPS数据进行实验仿真和验证,具有很不错的效果。

随着城市人口的不断扩大,人们的出行需求也不断增加。滴滴出行和Uber等出租车需求服务也慢慢的被大众所使用。在其中,如何利用大量的出租车的轨迹数据来进行需求预测是一个很关键的问题,特别是城市内区域的出租车需求预测,精准的预测科技降低出租车的空载率从而增加司机的收入,减少汽车行驶过程中能源的消耗,根据预测结果也能进行提前的车辆调度,提高城市出租车服务的质量。因此,城市区域的出租车需求预测对整个城市的智能交通系统有十分重大的意义。

在处理交通GPS数据中比较难处理的是轨迹数据,目前的出行数据研究方法中,很多方法是将轨迹数据处理成OD对(出发点-目的地)数据,然后进行下一步的建模。本文研究通过SVR时空网格模型,能有效提高需求预测的准确性和可靠性。

## 1 需求预测框架

本文使用长期的历史GPS数据观测值,预测的是未来的出租车需求。问题可描述为,已知 $n$ 个历史观测值 $\{x_t | t = 0, 1, 2, \dots, n-1\}$ ,预测所有 $\tau$ 个时间间隔后的需求值 $x_\tau$ 。

将市区分为个网格 $M \times N$ ,每个网格为一个数据区域。因此,每个数据区域在 $\tau$ 个时间间隔将存在输入流数量和输出流数量,在第 $m$ 行和第 $n$ 列上的数据区 $(m, n)$ 的输出流和输入流的定义为:

$$\begin{cases} x_\tau^{\text{input}, m, n} = \sum_{c \in C_{\text{all}}} |P_\tau^c \in A(m, n) \cap P_{\tau-1}^c \notin A(m, n)| \\ x_\tau^{\text{output}, m, n} = \sum_{c \in C_{\text{all}}} |P_\tau^c \in A(m, n) \cap P_{\tau+1}^c \notin A(m, n)| \end{cases} \quad (1)$$

式中:  $C_{\text{all}}$ 为所有车辆集;  $P_\tau^c$ 为在 $\tau$ 时间内车辆 $c$ 所在地理位置,  $A(m, n)$ 为数据区 $(m, n)$ 的地理区域;  $| \cdot |$ 表示集合的数量。因此 $x_\tau = [x_\tau^{\text{input}}, x_\tau^{\text{output}}]$ , 其中,  $x_\tau^{\text{input}}$ 和 $x_\tau^{\text{output}}$ 分别代表在时刻内的输入和输出流。

## 2 SVR算法

支持向量机(support vector machine)是一种监督学习的线性分类器。SVR(support vector regression)则是由SVM衍生得到的,用于解决回归预测问题。SVR模型实质上是一种带约束的规划问题,能提供全局最优解,其基本原理为:

给定 $N$ 个样本向量 $\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$ ,  $x_i \in R^n, y_i \in R, i = 1, 2, 3, \dots, N$ , 其中 $x_i$ 为输入数据,  $y_i$ 为输出数据。定义一个非线性的映射函数 $\phi(x)$ , 将低维线性数据映射到高维空间, 实现非线性转化为线性的运算, 则SVR函数模型定义如下式:

$$f(x) = \omega^T \phi(x) + b \quad (2)$$

其中 $f(x)$ 决定预测值;  $x$ 为输入数据向量;  $\omega$ 为权重向

量;  $b$ 表示偏置量。

经过变换求解, 式子(2)可表示为:

$$f(x) = \sum_{i=1}^N (\beta_i - \beta_i^*) K(x_i, x) + b \quad (3)$$

式中,  $\beta_i$ 和 $\beta_i^*$ 为拉格朗日乘子,  $K(\cdot)$ 为计算高维特征空间中的两个输入向量内积的核函数。本次实验采用高斯核, 公式为:

$$K(x_i, x) = \exp\left(-\frac{\|x - x_i\|^2}{2g^2}\right) \quad (3)$$

其中,  $g$ 为高斯核密度函数。

## 3 SVR时空网格模型

轨迹数据一般都是很多的GPS数据, 描述的是每段时间车辆所在的状态。为了得到网格数据, 需要将GPS数据进行处理。处理之前需要在给定区域上进行网格的划分, 并给网格标上编号, 划分完成后就可以使用算法计算出每个数据格的输入流和输出流。之后就可以使用SVR算法进行需求的预测。基于SVR时空网格模型的预测流程可描述如下:

1)根据区域的地理信息, 将地图上的区域划分成个网格, 区域表示第行第列的数据区域。

2)处理GPS轨迹数据, 轨迹数据中最重要的是GPS坐标点, 编号等数据。依据划分的区域和选取的时间间隔, 将GPS数据处理按照公式处理成输入输出流数据, 得到区域格的输入输出流数据。

3)将数据按照一定比例分为训练集和测试集, 训练集用于训练SVR模型, 测试集用于得到计算模型精度。

4)在SVR模型中, 以历史的 $h$ 个时间段的输入输出流量作为输入, 下个时间段的输入输出流量, 构建样本。



图1 网格选取图

5) 将样本集放入SVR模型中进行计算,并计算出不同的评价指标结果。

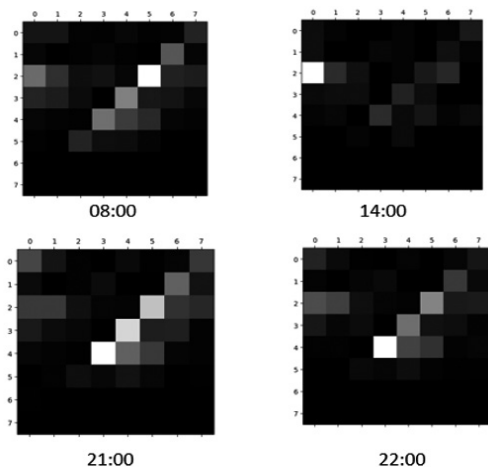


图2 不同时间点输

#### 4 实验分析

本次实验所使用的出租车数据集来源于深圳市2016年9月的出租车数据,并从中选取了一天的车数量,区域内的车量有12667辆。

其中,划取深圳市流量较大的区域,并划分8×8为的网格区,如图1所示。

为了保证数据的准确性,本实验的时间间隔设置为1分钟,并在数据处理中选取6:00到23:00的数据进行计算。图2选取了3个的数据,为了更

好的显示,将网格数据绘制成灰度图像,如图2所示。从图二中可以看出,随着时间的变化,各区域的交通流数据呈现不同的数值。本次实验选取h为5。使用SVR时空网格模型的方法,能使用前5个数据有效的考虑时序的因素,对整个预测有很好的效果。

对于回归的预测指标,通常有RMSE(均方根误差)、NMAE(标准平均绝对误差)、MAE(平均绝对误差)这几种。SVR时空模型完成后,本次实验将测试集输入模型,计算了这几个回归预测指标,数据如表格所示表1示。

表1 实验结果

性能指标	数值
RMSE	14.3589
NMAE	0.5292
MAE	9.0801

#### 5 总结

本次实验使用了SVR时空网络模型来进行需求预测,使用了交通输入输出流的方法而不是传统的通过OD对的方法,取得了良好的效果。所预测出的需求也能用于给整个智能交通系统做下一步的车辆调度,对智能交通系统有重要意义。

作者简介:王成安(1996—),男,汉族,广东梅州人,硕士,现就读于广东工业大学,研究方向为智能交通、大数据。

DOI:10.19353/j.cnki.dzsj.2020.03.026

在信息化时代,人们获得信息的方式发生了翻天覆地的变化,给传统的广电传统媒体造成巨大的冲击。在新时期为了能够积极顺应时代发展的趋势,推进广电传统媒体与新兴媒体的有机融合,必须积极构建完整的媒体融合技术体系,明确广电传统媒体和新兴媒体产业融合的主要特点以及融合动因,推动产业融合的水平全面提升。

在2016年国家新闻出版广电总局颁布了关于推动广电传统媒体和新兴媒体融合发展的指导意见,对当前新形势下媒体融合工作开展的现状做出了重要部署,广电传统媒体与新兴媒体融合发展取得了区域性的进展,也已经逐渐成为当前媒体深度融合的重要组成。在新时期要想促进广电与互联网的融合发展,最主要的就是保证传统广电传统媒体顺应互联网发展趋势,积极做到社交化、移动化、视频化的转变,通过利用先进的网络信息技术以及创新思维促进传统广电传统媒体实现从覆盖优势转为受众优势,从数字化成果转为传播力优势,一体化优势转变为平台优势,从内容安全管理转变为全媒体融合管理,确保广电传统媒体能够符合新形势下党和国家的要求。

#### 1 广电传统媒体与新兴媒体融合的主要特点

在新时期广电传统媒体与新媒体产业融合的过程中,最主要的就是注重融合发展布局以及打造媒体融合宣传的优势,通过顶层设计培养全媒体的专家型人才,对广电传统媒

体原有的采编播等流程进行改进与优化,通过完善融合项目方案推动项目深度发展,最后要积极探索以视频为中心,以新闻为首要的重点,通过以用户数据为核心,形成“三微一端”的融媒体平台,充分运用大数据云计算等新型技术,保证融媒体中心的全面发展,努力推动传统广电传统媒体向智慧型媒体转型升级。

#### 2 影响广电传统媒体与新媒体产业融合的重要因素

##### 2.1 市场因素

随着新媒体时代的到来,为了提高市场的集中度和产品差异化发展及时促进。广电传统媒体的广告收入网络收入满足理想状态,必须要依托广电传统媒体自身的综合实力以及资源优势进行差异化发展。

##### 2.2 市场行为

市场行为主要包括服务竞争、行为内容、竞争行为和价格竞争行为,其主要的竞争对手,包括行业内的方方面

广电传统媒体与新兴媒体融合技术体系探究  
蛟河市融媒体中心 郑强