

# 基于时空特征向量的城市快速路交通状态 长时段预测\*

姚博凡<sup>1,2</sup>, 邓如丰<sup>3</sup>, 熊宸<sup>1,2</sup>, 蔡铭<sup>1,2</sup>

1. 中山大学智能工程学院, 广东 深圳 518106
2. 广东省智能交通系统重点实验室, 广东 广州 510006
3. 佛山交通运行监测中心, 广东 佛山 528000

**摘要:** 针对交通状态长时段预测问题, 提出了基于时空特征向量的长时段交通状态预测模型。该模型在分析交通状态变化时空关联性的基础上, 构建了表征交通状态变化规律的时空特征向量, 包括时间轴特征、路段历史平均交通状态特征、时间常发拥堵特征和上下游路段的历史交通状态特征。单路段预测实验结果表明, 模型预测准确度较高, 全天预测准确度在 90% 左右, 高峰预测准确度在 80% 左右。此外, 与单一时间特征相比, 提出的时空特征向量能显著提高预测准确度。多路段的实验结果表明, 基于时空特征向量的 SVM 预测方法具有较好的普适性, 全天平均预测准确度达到了 95% 左右, 高峰平均预测准确度达到了 88% 左右。

**关键词:** 长时段预测; 支持向量机; 交通状态; 时空特征向量

**中图分类号:** U491 **文献标志码:** A **文章编号:** 0529-6579 (2021) 03-0115-09

## Long-term traffic status prediction of urban expressway based on spatio-temporal eigenvector

YAO Bofan<sup>1,2</sup>, DENG Rufeng<sup>3</sup>, XIONG Chen<sup>1,2</sup>, CAI Ming<sup>1,2</sup>

1. School of Intelligent Systems Engineering, Sun Yat-sen University, Shenzhen 518106, China
2. Key Laboratory of Intelligent Transportation System in Guangdong Province, Guangzhou 510006, China
3. Foshan Transportation Operations Monitor Centre, Foshan 528000, China

**Abstract:** This paper proposes a long-term traffic status prediction model based on a spatio-temporal eigenvector in order to solve the problem of long-term traffic status prediction. Based on the analysis of the spatio-temporal correlation of traffic status changes, the model constructs the spatio-temporal eigenvector that represents the law of traffic status variation, including time axis feature, historical average traffic status feature, recurrent congestion feature and historical average traffic status feature of upstream and downstream roads. The prediction results of the single roads show that the model has a good prediction performance. The accuracy of all-day prediction is about 90% and the accuracy of peak hour prediction is about 80%. Furthermore, compared with single time features, the proposed spatio-temporal eigenvector can significantly improve the prediction accuracy. The prediction results of multiple roads show that

\* 收稿日期: 2019-11-04 录用日期: 2019-12-03 网络首发日期: 2020-11-10  
基金项目: 国家重点研发计划 (2018YFB1601000); 中央高校基本科研业务费 (19lgy290)  
作者简介: 姚博凡 (1994 年生), 男; 研究方向: 交通大数据; E-mail: yaobf@mail2.sysu.edu.cn  
通信作者: 蔡铭 (1977 年生), 男; 研究方向: 交通大数据、交通环境工程和智能交通系统;  
E-mail: caiming@mail.sysu.edu.cn

the SVM prediction method based on spatio-temporal has a good universality. The average accuracy of all-day prediction is about 95% and the average accuracy of peak hour prediction is about 88%.

**Key words:** long-term prediction; support vector machine; traffic status; spatio-temporal eigenvector

近年来,城市不断发展,人口密度和路网密度不断增加,交通出行需求日益增长,特别是在一二线城市,高峰出行拥堵每日都在上演。如何准确预测交通出行,诱导出行者选择其他路径,从而分散交通出行压力,成为了解决问题的关键所在。因此,对未来交通态势进行合理的判断,对于出行者和管理者都具有重要意义。面对这一问题,有的学者从交通需求的角度出发进行预测<sup>[1-2]</sup>,有的学者则对交通流参数进行短时预测。然而,短时交通流参数预测难以满足长时段的交通出行需求。针对这一问题,本文探究了长时段交通状态预测的方法。长时段交通状态预测不仅能够为交通出行者提供最直观的交通运行信息,而且预测准确度较高,更具有实际应用价值。

对交通流状态的预测,包括流量、速度、交通状态等,国内外有许多相关的研究。在交通大数据的背景下,人工智能方法开始广泛应用,其中机器学习方法以其良好的性能而被普遍使用,比如支持向量机(SVM)<sup>[3]</sup>、k近邻算法(KNN)<sup>[4]</sup>和随机森林(RF)<sup>[5]</sup>。在流量预测方面,罗文慧等<sup>[6]</sup>利用卷积神经网络对交通流特征进行提取,将提取的特征输入到支持向量回归模型中进行流量预测,预测精度相比传统预测模型有所提升。康军等<sup>[7]</sup>为了解决用SVM进行短时交通量在线预测的计算复杂性问题,简化了Lagrange乘子的求解过程,利用滑动时间窗口来控制新样本的加入和旧样本的移除,通过线性运算完成Lagrange乘子的更新,在保证预测精度的情况下减少预测的时间。也有学者<sup>[8]</sup>提出了以SARIMA模型提取交通流数据中的周期性特征,再利用随机森林进行短时交通流量预测。Guo等<sup>[9]</sup>建立了基于KNN的混合预测框架,融合了神经网络、支持向量回归和随机森林三种预测方法,通过KNN算法来调整三种方法预测结果的权值,预测正常状态下和有交通事件发生状态的短时交通流量,与其他方法相比,结果更加精确。谢海红等<sup>[10]</sup>改进了KNN算法,用模式距离搜索方法代替欧氏距离搜索方法,利用多元统计回归模型,以下一时刻状态与对应 $k$ 个近邻下一时刻状态作为模型输入,去估计模型参数,之后借助这个多元统计回归模型,输入预测状态 $k$

个近邻的下一时刻状态,得到预测状态。在速度预测方面,史殿习等<sup>[11]</sup>认为路段的速度与路段当前的速度变化情况有关,于是选取了从当前时段开始的后4个连续时段的速度序列作为状态向量,基于速度变化趋势和密集度的变K近邻精确搜索策略,对K近邻算法的近邻搜索策略进行改进和优化,提出了一种短时交通平均速度预测模型。田瑞杰等<sup>[12]</sup>将时间序列预测和人工神经网络相结合,采用ARMA预测模型对数据建模预测,再应用人工神经网络调整预测值,得到了较好的交通流短时预测效果。孙静怡等<sup>[13]</sup>发现车流中大型车比例对车流速度具有显著影响,因此引入了大型车比例,加上前几个时间点速度数据作为支持向量机的输入,从而对高速公路交通流速度进行短时预测,结果发现大型车比例能较好提升预测精度,并且通过时间序列分析发现以前6或7个时间点的速度数据作为输入的预测方案效果较好。Wang等<sup>[14]</sup>则对SVM的核函数进行设计,构造了基于小波函数的核函数,以捕捉非平稳特征的短期交通速度数据,并利用相空间重构理论确定了模型输入空间的维度,从而建立了短时交通速度预测的混合模型,得到的预测结果优于原本的SVM模型。在交通状态预测方面,陈忠辉等<sup>[15]</sup>先是以流量、速度和占有率为指标,利用模糊C均值聚类对交通状态进行聚类,得到交通状态的标签,接着同样以流量、速度和占有率作为输入特征,以聚类得到的交通状态标签作为输出,构建了随机森林预测模型,从而预测短时交通状态。Xu等<sup>[16]</sup>采用一种基于核k近邻的时间序列道路交通状态预测算法,将当前道路交通状态数据序列与参考道路交通状态数据序列进行匹配,选择 $k$ 个最近的参考道路交通状态,预测当前道路的交通状态。Liu等<sup>[17]</sup>以天气条件、时段、道路特殊情况、道路质量和节假日为输入变量,以随机森林建立道路交通拥堵状态预测模型,结果表明该方法的预测精度较好,泛化误差较低,计算速度较快,可以有效预测。此外,也有学者利用其他方法进行交通预测,比如向量自回归(VAR)<sup>[18]</sup>、自回归整合移动平均(ARIMA)<sup>[19]</sup>、深度学习<sup>[20-21]</sup>。

但是,目前研究仍存在以下不足。第一,多

数的研究都是针对短期的交通状态预测, 在当前时刻只能预测下一个预测时刻的交通状态; 第二, 研究对象基本是一条路段, 无法说明方法的普适性。因此, 本文针对以上的不足之处, 从时间和空间两个方面的特征考虑, 构建了新的时空特征向量, 建立了一种长时段路段交通状态的预测模型; 此外, 将方法拓展应用到多个路段, 验证了方法的普适性。

## 1 长时段交通状态预测模型

### 1.1 模型框架

本文所提出的长时段交通状态预测模型框架如图1所示。框架主要包含了两部分内容: 模型构建和预测部分。在第一部分中, 特征提取是整个模型最重要的部分。对历史交通状态数据进行特征提取, 构建时空特征向量, 结合支持向量机分类算法构建交通状态预测模型; 在进行预测的时候, 先构建预测时段的时空特征向量矩阵, 输入到第一部分得到交通状态预测模型, 模型的输出即为预测时段的交通状态值。短时交通流预测指的是在当前时刻只能预测下一时刻的交通流特征参数, 时间粒度即为预测时长。而长时段交通状态预测, 指的是在当前时刻下, 对未来12 h、24 h或者更长时段的路段交通状态进行预测, 时间粒度范围可以是1~10 min, 因为本文用到的数据时间粒度是2 min, 所以本文预测的时间粒度为2 min, 但是本文的预测时长是未来24 h, 这是区别于短时预测的地方。

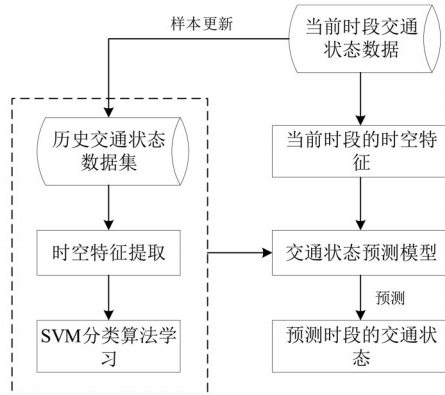


图1 交通状态预测模型框架

Fig. 1 Framework of traffic status prediction model

### 1.2 时空特征向量

现有研究考虑的特征主要有时间特征、当前时刻的预测量特征或者前几个时刻的预测量特征,

这也导致了一次预测只能预测下一时刻交通流特征参数(速度、流量等)的局限性, 如果要预测多个时刻, 势必要引入预测量, 这将会导致误差的累积。本文从考虑历史的角度出发, 引入了预测时刻的预测路段的平均历史交通状态、上下游路段的平均历史交通状态以及时间常发拥堵特征, 构建了新的时空特征向量。

**1.2.1 时间特征** 时间特征是针对预测路段本身而言的, 主要包括了三个方面: 一是时间轴特征, 也就是交通状态所在的时间, 二是路段历史平均交通状态特征, 三是时间常发拥堵特征。

1) 时间轴特征。路段交通状态本质上受人们交通出行的影响, 而交通出行是具有一定的时间规律性的, 比如工作日的上班和下班时间、非工作日的出行娱乐时间等, 因此导致了路段的交通状态变化也是具有一定的时间规律性。本文以交通状态所在的时间作为时间轴特征, 提取了时、分、秒三个特征, 由于本文所用到的实验数据时间间隔是2 min, 秒数都为0。因此, 剔除秒数特征, 保留时、分两个特征。由于交通状态变化具有一定的周期性规律, 为了描述这个周期性的变化, 加入周天特征, 周天是表示数据所在时间为星期几。将时、分、周天这三个特征分别表示为  $x_H$ 、 $x_M$  和  $x_w$ 。此外, 加入  $x_c$  代表是否工作日, 取值0和1分别表示非工作日和工作日。最终以  $x_H$ 、 $x_M$ 、 $x_w$ 、 $x_c$  四个特征作为表征交通状态的时间轴特征。

2) 历史平均交通状态特征。历史平均交通状态指的是预测路段在同个周天、同时、同分下历史数据的交通状态的平均值。假设  $W_i$  为历史数据中属于星期  $i$  的交通状态数据,  $H_j$  为历史数据中属于  $j$  时的交通状态数据,  $M_z$  为历史数据中属于  $z$  分钟的交通状态数据。

$$\bar{S}(i, j, z) = \frac{\sum_{S_k \in \Omega(i, j, z)} S_k}{n}, \quad (1)$$

其中  $\bar{S}(i, j, z)$  表示路段在星期  $i$ 、 $j$  时、 $z$  分的历史平均交通状态值,  $S_k$  为在同个周天、同时、同分下历史交通状态数据集中第  $k$  个元素交通状态值,  $n$  为元素的个数,  $\Omega(i, j, z) = W_i \cap H_j \cap M_z$ 。

3) 时间常发拥堵特征。时间常发拥堵特征是表征数据所在时刻是否在路段的常发拥堵时段内的特征。如果是, 则赋值为1, 否则, 赋值为0。时间常发拥堵特征判别算法如下:

## 算法 1 路段常发拥堵时段判别算法

---

输入: 历史数据集  $H = \{s_1, s_2, \dots, s_m\}$   
 工作日时间区间集  $T_1 = \{t_1^1, t_2^1, \dots, t_{48}^1\}$ , 区间长度 30 min  
 非工作日时间区间集  $T_2 = \{t_1^2, t_2^2, \dots, t_{48}^2\}$ , 区间长度 30 min

- 1: for  $c = 1, 2$  do
- 2:   for  $k = 1, 2, \dots, 48$  do
- 3:     从历史数据集中筛选所有在时间区间  $t_k^c$  内的记录, 记为  $H'$
- 4:     令  $H' = \{s'_1, s'_2, \dots, s'_n\}$ ;
- 5:     for  $i = 1, 2, \dots, n$  do
- 6:        令  $S = 0$
- 7:        if  $s'_i = 3$  or  $4$  or  $5$  then
- 8:           $S = S + s'_i$
- 9:        end if
- 10:      end for
- 11:      令  $p = \frac{S}{n}$
- 12:      if  $p > 0.5$  then
- 13:         $f_{t_k^c} = 1$
- 14:      else
- 15:         $f_{t_k^c} = 0$
- 16:      end if
- 17:    end for
- 18: end for

输出: 常发拥堵判定结果  $R = \{f_{t_1^1}, f_{t_2^1}, \dots, f_{t_{48}^1}, f_{t_1^2}, f_{t_2^2}, \dots, f_{t_{48}^2}\}$

---

在该算法中, 以工作日和非工作日划分一周 (周一到周五为工作日, 周六和周日为非工作日), 将一天 24 h 划分为 48 个时间区间, 区间长度为 30 min, 工作日和非工作日共计分为 96 个时间区间。以路段历史数据为统计对象, 分别计算每个时间区间中交通状态为拥堵的频率, 在实验数据中交通状态值为 3、4 和 5 是表示拥堵的交通状态, 如表 1 所示。因此, 统计交通状态值大于 3 的频率, 如果该频率超过 0.5, 则说明该路段在该时间区间为常发拥堵。之后, 再判断数据所在时刻的时间区间是否为常发拥堵。如果是, 则赋值为 1; 否则, 赋值为 0, 即为时间常发拥堵的特征值。

**1.2.2 空间特征** 首先需要明确的一点是在没有交叉口打断的情况下, 封闭路段由多个小路段组成, 与封闭路段相接壤的路段为相邻路段。本文的研究对象是城市快速路, 是一段封闭的没有交叉口的路段, 它有与之接壤的唯一的上下游路段。由于相邻路段之间的交通状态是相互影响的, 以上下游路段的历史平均交通状态特征作为空间特征, 计算方法如式 (1) 所示, 得到在同个周天、

同时、同分下上下游路段历史数据的交通状态的平均值, 分别以  $\bar{S}^u(i, j, z)$  和  $\bar{S}^d(i, j, z)$  表示。这两个特征反映了上下游路段的历史交通状态, 表征了预测路段交通状态的空间特征。

### 1.3 支持向量分类模型

在特征提取之后, 本文构建的特征向量可以表示为

$$V = [x_H, x_M, x_w, x_c, R_i, \bar{S}(x_w, x_H, x_M), \bar{S}^u(x_w, x_H, x_M), \bar{S}^d(x_w, x_H, x_M)], \quad (2)$$

其中  $x_H$  为时特征, 取值 0~24;  $x_M$  为分特征, 取值 0~58, 间隔为 2;  $x_w$  为周天特征, 取值 1~7, 分别表示周一到周日;  $x_c$  为工作日特征, 0 表示休假, 1 表示上班;  $R_i$  为时间常发拥堵特征, 0 表示不常发拥堵, 1 表示常发拥堵;  $\bar{S}(x_w, x_H, x_M)$  为预测路段的历史平均交通状态特征;  $\bar{S}^u(x_w, x_H, x_M)$  为上游路段的历史平均交通状态特征;  $\bar{S}^d(x_w, x_H, x_M)$  为下游路段的历史平均交通状态特征。以该特征向量作为模型的输入, 模型输出是未来一段时段的交通状态值, 时间粒度为 2 min, 输出变量个数与预测时长相关, 比如预测未来 24 h 的某城市快速路段的交通状态, 则输出变量个数为 720 个。

本文的预测问题是一个非线性的分类问题, 由于支持向量机对于非线性分类具有明显优势, 于是采用了支持向量机作为学习算法, 结合构建的特征向量, 建立了交通状态的预测模型。通过对预测结果进行比较, 确定了模型的核函数为径向基 (rbf) 核函数。

## 2 实验案例

为了验证本文所提出方法的有效性和普适性, 以佛山市佛山大道北的一个路段进行单路段的交通状态长时段预测, 佛山大道北是一条城市快速路, 之后对比 SVM 与其他方法的预测结果, 以及不同核函数的预测结果。最后选取多条路段, 对应每条路段建立不同的 SVM 预测模型, 将方法应用到更多的路段。

### 2.1 数据描述及预处理

实验数据是某地图导航运营商提供的路段路况数据, 时间为 2017/12/1-31, 数据时间间隔为 2 min, 数据的有效字段包括了路段 ID、时间、道路等级、路段平均行程速度等, 参照国标《城市交通运行状况评价规范》<sup>[22]</sup>, 如表 1 所示, 通过路段平均行程速度来得到路段的交通运行状态。其

中, 交通运行状况等级包含了5种交通状态, 分别为畅通、基本畅通、轻度拥堵、中度拥堵和严重拥堵, 分别以1、2、3、4和5表示。在实验中, 以2017/12/4-24的数据为训练集。为了避免圣诞节对结果的影响, 以2017/12/26-31的数据为测试集, 并且采用最大最小归一化的方法, 对数据进行标准化处理。训练样例数据如表2所示。

表1 城市交通状态评价规范<sup>1)</sup>

Table 1 Specification for urban traffic status evaluation

运行状况等级	取值范围	交通状态值
畅通	$V_{ij} > V^f \times 70\%$	1
基本畅通	$V^f \times 50\% < V_{ij} \leq V^f \times 70\%$	2
轻度拥堵	$V^f \times 40\% < V_{ij} \leq V^f \times 50\%$	3
中度拥堵	$V^f \times 30\% < V_{ij} \leq V^f \times 40\%$	4
严重拥堵	$V_{ij} \leq V^f \times 30\%$	5

1)  $V_{ij}$ 为路段平均行程速度,  $V^f$ 为路段自由流速度, 单位为km/h。

### 2.2 评价指标

为度量预测结果的准确性, 本文以准确度(Acc)、平均绝对误差(MAE)和均方根误差(RMSE)为评价指标。定义

$$Acc = \frac{1}{N} \sum_{i=1}^N r_i \times 100\%, \quad (3)$$

$$r_i = \begin{cases} 1, & p_i = t_i, \\ 0, & p_i \neq t_i, \end{cases} \quad (4)$$

$$MAE = \frac{1}{N} \sum_{i=1}^N (|p_i - t_i|), \quad (5)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (p_i - t_i)^2}, \quad (6)$$

其中 $p_i$ 表示预测交通状态值,  $t_i$ 表示真实交通状态值,  $N$ 表示预测交通状态值的个数。

### 2.3 实验结果及分析

2.3.1 单路段预测结果 实验采用支持向量机(SVM)对2017/12/26-31的佛山大道北一个路段的交通状态进行预测, 并将预测结果与随机森林(RF)、Adaboosting (Ada)、K-近邻(KNN)、梯度提升决策树(GBDT)和决策树(DT)的预测结果进行比较。

1) 不同方法结果比较。六种预测方法2017/12/26的预测结果如图2所示。从图中可以看出, SVM能较好地预测长时段的城市快速路交通状态。由表3中的各项评价指标可以看到, SVM对26日全天预测准确度最高, 达到了91.94%, 比其他方法高2.08%以上; MAE和RMSE是最小的, 分别为0.11和0.41。在高峰时段(7:00-9:00和17:00-19:00), SVM的预测表现同样优于其他方法。预测准确度达到了78.69%, 比其他方法高7.38%以上。预测精度比起全天要有所下降, 这是因为高峰时段的交通状态变化起伏较大、规律出现波动导致的。此外, 对27-31日的整体平均预测准确度进行计算。通过比较可知, SVM对工作日和非工作日的整体预测表现略优于其他方法。实验中, 电脑操作系统为Windows Server 2012R2, CPU为Intel Xeon E5-2630 v3, 内存为128 GB, 实验采用python编程实现。从预测时间效率来看, SVM模型的预测耗费时间稍大于其他方法, 但是时间差距只是毫秒级的。相对来说, SVM模型提升准确度所花费的时间代价比较小。

表2 训练数据原始样例

Table 2 Training data original samples

输入数据					输出数据			
时	分	周天	工作日	时间常发拥堵	历史平均交通状态	上游历史平均交通状态	下游历史平均交通状态	真实交通状态
0	0	1	1	0	1.00	1.33	1.00	1
	2	1	1	0	1.00	1.33	1.00	1
...	...	...	...	...	...	...	...	...
7	58	4	1	0	3.00	4.00	1.00	3
8	0	4	1	1	3.00	4.00	1.00	3
...	...	...	...	...	...	...	...	...
23	56	7	0	0	1.00	1.00	1.00	1
	58	7	0	0	1.00	1.00	1.00	1

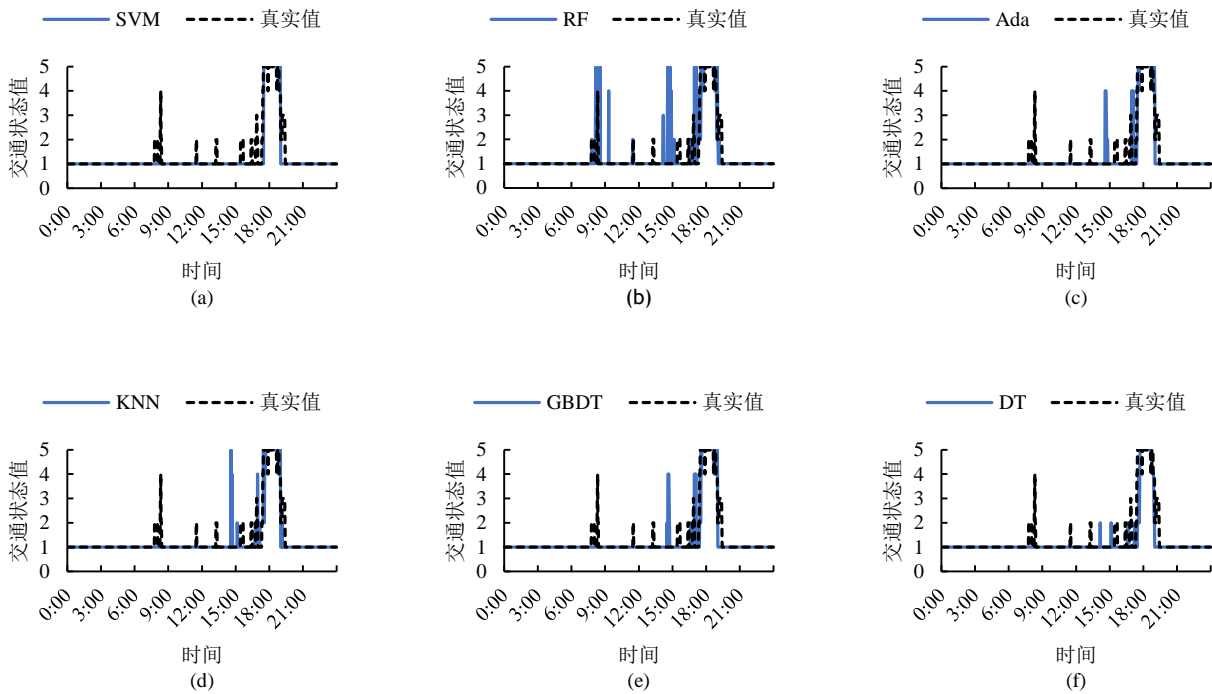


图2 不同方法的预测结果

Fig. 2 Prediction results of different methods

表3 单路段预测结果对比

Table 3 Comparison of prediction results of different methods

方法	26日全天			26日高峰			27~31日平均			预测全天平均 时间消耗/s
	Acc/%	MAE	RMSE	Acc/%	MAE	RMSE	Acc/%	MAE	RMSE	
SVM	91.94	0.11	0.41	78.69	0.29	0.73	84.94	0.32	0.95	0.09
RF	86.67	0.23	0.74	65.57	0.67	1.35	83.39	0.33	0.95	0.01
Ada	85.28	0.23	0.68	68.85	0.58	1.20	84.17	0.32	0.94	0.02
KNN	89.17	0.15	0.50	71.31	0.43	0.92	83.53	0.32	0.92	0.03
GBDT	89.31	0.16	0.54	68.03	0.48	0.99	84.17	0.32	0.94	0.01
DT	89.86	0.14	0.47	71.31	0.42	0.89	84.36	0.31	0.91	0.01

2) 不同核函数结果比较。探究了四种常用核函数对预测结果准确度的影响,如图3所示。实验结果表明,无论是在全天还是高峰时段,rbf核函数在预测准确度上都要高于linear、poly和sigmoid,这也是模型选择rbf作为核函数的原因。

3) 不同特征组合结果比较。针对本文所提出的时空特征,进一步比较了不同特征组合对预测结果准确度的影响,以验证本文提出的时间常发拥堵特征、历史平均交通状态特征能否有效提高SVM模型的预测精度。其中,历史平均交通状态特征包括了预测路段历史平均交通状态、上游路段历史平均交通状态和下游路段历史平均交通状

态。实验中,将特征按照类型的不同进行分类组合。

(a) 组合1: 时间轴特征(时、分、周天、工作日);

(b) 组合2: 时间轴特征(时、分、周天、工作日) + 时间常发拥堵特征;

(c) 组合3: 时间轴特征(时、分、周天、工作日) + 时间常发拥堵特征 + 预测路段历史平均交通状态;

(d) 组合4: 时间轴特征(时、分、周天、工作日) + 时间常发拥堵特征 + 历史平均交通状态特征(预测路段历史平均交通状态、上游路段历

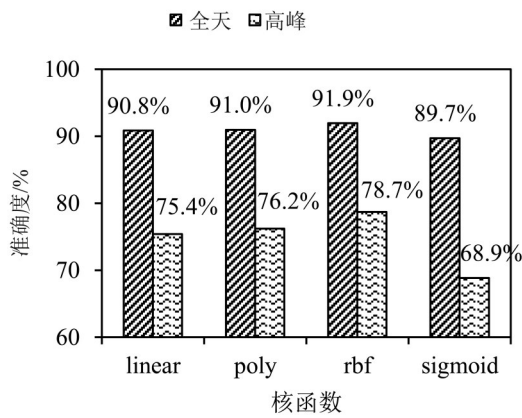


图3 不同核函数预测结果

Fig. 3 Prediction results of different kernel functions

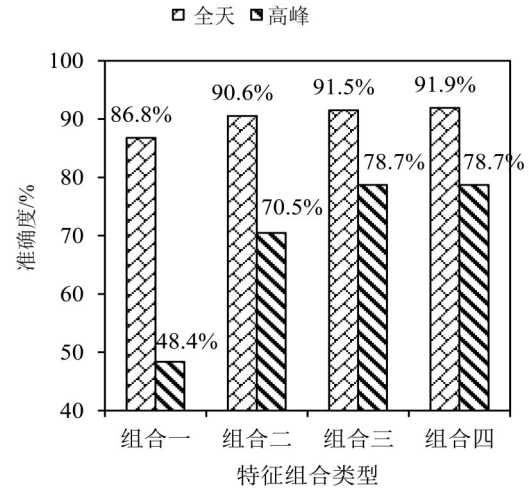


图4 不同特征组合预测结果

Fig. 4 Prediction results of different feature combinations

史平均交通状态、下游路段历史平均交通状态)。

(e) 组合4对应本文的时空特征向量。对比四种不同特征组合的预测结果准确度, 结果如图4所示。图4中, 组合4的全天预测准确度最高, 说明本文设计的时空特征向量能有效提升预测的准确度。相比组合1的预测结果, 组合2的全天和高峰预测准确度分别提高了3.8%和22.1%, 组合3的全天和高峰预测准确度分别提高了4.7%和30.3%, 组合4的全天和高峰预测准确度分别提高了5.15和30.3%, 预测精度均有大幅提升, 特别是高峰时段。从组合3和组合4的比较来看, 加入上下游路段历史平均交通状态后, 高峰预测准确度基本不变, 全天预测准确度有所提高。进一步说明了本文提出的时间常发拥堵特征和历史平均特征都能较好地反映交通状态的变化规律, 特别

是时间常发拥堵特征对于高峰期的预测准确度提升效果显著。

**2.3.2 多路段预测结果** 为了验证基于时空特征向量的SVM预测方法的普适性, 实验随机选取了佛山市500条城市快速路路段进行测试, 与单条路段的实验类似, 针对每条路段建立不同的SVM预测模型。对每一条路段的评价指标进行平均, 结果如表4所示。实验表明, 该预测方法同样适用于其他路段, 在随机选取的500条路段上的平均预测准确度较好, 26号全天预测平均准确度达到了95.32%, 高峰时段预测平均准确度达到了88.63%, 27号到31号的预测平均准确度达到了94.94%, 在六种方法中表现最优, 充分表明了本文所提出的预测方法的普适性。

表4 多路段预测结果对比

Table 4 Comparison of prediction results of multiple roads

方法	26日全天			26日高峰			27~31日平均		
	Acc /%	MAE	RMSE	Acc /%	MAE	RMSE	Acc /%	MAE	RMSE
SVM	95.32	0.07	0.27	88.63	0.19	0.43	94.94	0.09	0.35
RF	94.60	0.08	0.30	86.95	0.21	0.46	94.02	0.10	0.39
Ada	82.70	0.26	0.51	75.60	0.37	0.63	82.57	0.28	0.59
KNN	94.91	0.08	0.28	87.48	0.20	0.44	94.39	0.09	0.36
GBDT	94.96	0.08	0.28	87.61	0.20	0.44	94.40	0.09	0.36
DT	94.97	0.08	0.28	87.49	0.20	0.45	94.55	0.09	0.35

### 3 结论

本文通过挖掘路段交通状态的时空关联性, 对特征向量进行设计, 提出了以时间特征和空间

特征为基础的特征向量来表征交通状态的规律性变化, 并结合支持向量机分类算法, 预测长时段的路段交通状态。以单路段和多路段实验, 验证

了模型的有效性和普适性,表明了本文提出的时空特征向量能较好地描述交通状态的变化规律,预测精度在全天和高峰时间的表现都较好。本文的创新在于弥补了短时交通状态预测的不足,构建了新的时空特征向量,提出了一种长时段交通状态预测的有效方法,并且引入的新特征能有效提高长时段交通状态预测准确度。此外,将方法应用于多个路段,针对每条路段建立不同的SVM预测模型,验证了方法的普适性,对于大规模路

网的应用具有一定的可行性。

同时,研究也存在可以进一步完善的地方。比如:由于缺乏节假日、气象等数据,本文在构建特征向量的时候没有考虑节假日和气象的影响。在后续的研究中,可以考虑将节假日和气象纳入特征向量当中,以提高该模型在其他特殊交通需求模式下的预测能力。此外,可以挖掘更多的特征,进一步提高预测精度。

### 参考文献:

- [1] 葛梦雪,宋国华,王芃森,等.基于手机应用软件测算OD数据的区域交通需求预测[J].公路交通科技,2019,36(5):152-158.  
GE M X, SONG G H, WANG P S, et al. Regional traffic demand forecasting based on OD data calculated by mobile phone application software [J]. Journal of Highway and Transportation Research and Development, 2019, 36(5): 152-158.
- [2] ZHAO Y, ZHANG H, AN L, et al. Improving the approached of traffic demand forecasting in the big data era [J]. Cities, 2018, 82: 19-26.
- [3] 李林超,张健,杨帆,等.基于核函数切换和支持向量回归的交通量短时预测模型[J].东南大学学报(自然科学版),2017,47(5):1032-1036.  
LI L C, ZHANG J, YANG F, et al. Traffic volume prediction based on support vector regression with switch kernel functions [J]. Journal of Southeast University (Natural Science Edition), 2017, 47(5): 1032-1036.
- [4] HABTEMICHAEL F, CETIN M. Short-term traffic flow rate forecasting based on identifying similar traffic patterns [J]. Transportation Research Part C: Emerging Technologies, 2016, 66: 61-78.
- [5] GUO F, KRISHNAN R, POLAK J. The influence of alternative data smoothing prediction techniques on the performance of a two-stage short-term urban travel time prediction framework [J]. Journal of Intelligent Transportation Systems, 2017, 21(3): 214-226.
- [6] 罗文慧,董宝田,王泽胜.基于CNN-SVR混合深度学习模型的短时交通流预测[J].交通运输系统工程与信息,2017,17(5):68-74.  
LUO W H, DONG B T, WANG Z S. Short-term traffic flow prediction based on CNN-SVR hybrid deep learning model [J]. Journal of Transportation Systems Engineering and Information Technology, 2017, 17(5): 68-74.
- [7] 康军,段宗涛,唐蕾,等.一种LS-SVM在线式短时交通流预测方法[J].计算机应用研究,2018,35(10):2965-2968.  
KANG J, DUAN Z T, TANG L, et al. LS-SVM online short term traffic flow prediction method [J]. Application Research of Computers, 2018, 35(10): 2965-2968.
- [8] 熊亨,戚湧,张伟斌,等.基于时空相关性的短时交通流预测模型[J].计算机工程与设计,2019,40(2):501-507.  
XIONG T, QI Y, ZHANG W B, et al. Short term traffic flow forecasting model based on temporal-spatial correlation [J]. Computer Engineering and Design, 2019, 40(2): 501-507.
- [9] GUO F, POLAK J, KRISHNAN R. Predictor fusion for short-term traffic forecasting [J]. Transportation Research Part C: Emerging Technologies, 2018, 92: 90-100.
- [10] 谢海红,戴许昊,齐远.短时交通流预测的改进K近邻算法[J].交通运输工程学报,2014,14(3):87-94.  
XIE H H, DAI X H, QI Y. Improved K-nearest neighbor algorithm for short-term traffic flow forecasting [J]. Journal of Traffic Transportation Engineering, 2014, 14(3): 87-94.
- [11] 史殿习,丁涛杰,丁博,等.一种基于非参数回归的交通速度预测方法[J].计算机科学,2016,43(2):224-229.  
SHI D X, DING T J, DING B, et al. Traffic speed forecasting method based on nonparametric regression [J]. Computer Science, 2016, 43(2): 224-229.
- [12] 田瑞杰,张维石,翟华伟.基于时间序列与BP-ANN的短时交通流速度预测模型研究[J].计算机应用研究,2019(11):1-8.  
TIAN R J, ZHANG W S, ZHAI H W. Short-term traffic flow velocity prediction model based on time series and BP-ANN [J]. Application Research of Computers,

- 2019(11):1-8.
- [13] 孙静怡, 牟若瑾, 刘拥华. 考虑大型车因素的支持向量机短时交通状态预测模型研究[J]. 公路交通科技, 2018, 35(10): 126-132.
- SUN J Y, MOU R J, LIU Y H. Study on short-term traffic state forecasting model of SVM considering proportion of large vehicles [J]. Journal of Highway and Transportation Research and Development, 2018, 35(10): 126-132.
- [14] WANG J, SHI Q. Short-term traffic speed forecasting hybrid model based on Chaos-Wavelet Analysis-Support vector machine theory [J]. Transportation Research Part C: Emerging Technologies, 2013, 27: 219-232.
- [15] 陈忠辉, 凌献尧, 冯心欣, 等. 基于模糊C均值聚类 and 随机森林的短时交通状态预测方法[J]. 电子与信息学报, 2018, 40(8): 106-113.
- CHEN Z H, LING X Y, FENG X X, et al. Short-term traffic state prediction approach based on FCM and random forest [J]. Journal of Electronics & Information Technology, 2018, 40(8): 106-113.
- [16] XU D W, WANG Y D, PENG P, et al. Real-time road traffic state prediction based on Kernel-KNN[J]. Transportmetrica A: Transport Science, 2020, 16(1): 104-108.
- [17] LIU Y, WU H. Prediction of road traffic congestion based on random forest [C]//2017 10th International Symposium on Computational Intelligence and Design (ISCID). IEEE, 2017.
- [18] 刘欢, 邵社刚, 王赵明. 基于向量自回归模型的城市快速路交通状态短时预测研究[J]. 公路交通科技(应用技术版), 2019, 15(2): 314-319.
- LIU H, SHAO S G, WANG Z M. Study on short-term traffic status prediction of urban expressway based on vector autoregression model[J]. Journal of Highway and Transportation Research and Development (Applied Technology Edition), 2019, 15(2): 314-319.
- [19] XU D, WANG Y, JIA L, et al. Real-time road traffic state prediction based on ARIMA and Kalman filter [J]. Frontiers of Information Technology & Electronic Engineering, 2017, 18(2): 287-302.
- [20] POLSON N, SOKOLOV V. Deep learning for short-term traffic flow prediction [J]. Transportation Research Part C: Emerging Technologies, 2017, 79: 1-17.
- [21] WANG J, CHEN R, HE Z. Traffic speed prediction for urban transportation network: A path based deep learning approach[J]. Transportation Research Part C: Emerging Technologies, 2019, 100: 372-385.
- [22] 国家质量监督检验检疫总局. 城市交通运行状况评价规范:GB/T 33171-2016[S]. 北京: 国家标准化管理委员会, 2016.
- General Administration of Quality Supervision, Inspection and Quarantine. Specification for urban traffic performance evaluation: GB/T 33171-2016 [S]. Beijing: Standardization Administration, 2016.

(责任编辑 王海蓉)