



DOI:10.12404/j.issn.1671-1815.2307788

引用格式:周瑜,张梦蝶.基于长短期记忆网络的共享单车真实需求预测方法[J].科学技术与工程,2025,25(1):394-403.

Zhou Yu, Zhang Mengdie. Real demand prediction method of shared bike based on LSTM[J]. Science Technology and Engineering, 2025, 25(1): 394-403.

基于长短期记忆网络的共享单车真实需求预测方法

周瑜,张梦蝶

(内蒙古大学经济管理学院,呼和浩特 010021)

摘 要 共享单车是城市交通的重要组成部分。共享单车用户需求的随机性导致其时空需求不均衡,甚至出现“借车难”现象,无法满足高峰时段的用户需求。因此出现高频用户高频出行时段到临近站点借车的现象,这意味着存在隐性需求。对于隐性需求,首先应用租借数和归还数刻画站点状态变化,通过挖掘临近站点用户出行情况判断参考站点的临界状态,建立基于站点状态变化图和需求判定模型确定站点的隐性需求。进而依据站点真实需求,建立长短期记忆(long short-term memory, LSTM)网络预测模型,并建立基于真实需求的共享单车区域调度模型。该模型以成本最小为目标,通过遗传算法求解得到调度成本最小的路径,为基于真实需求开展平衡调度提供了参考。结果显示,在相近的运输成本下,真实需求下的调度方式能在一定程度上缓解用户借车难的问题,从而减少高频用户的流失。

关键词 共享单车;需求预测;隐性需求;长短期记忆;调度优化

中图分类号 U484;

文献标志码 A

Real Demand Prediction Method of Shared Bike Based on LSTM

ZHOU Yu, ZHANG Meng-die

(College of Economics and Management, Inner Mongolia University, Hohhot 010021, China)

[Abstract] Shared bikes represent a crucial component of urban transportation. The randomness of user demand for shared bikes with fixed piles leads to unbalanced demand in time and space, and even the difficulty in renting a bike, which cannot meet the user demand during peak hours. Therefore, high-frequency users frequently travel to nearby stations to rent a bike for serving, which means that there are implicit demands. As for the hidden demand, firstly, the state changes of the site were described by the rental number and the return number, and the critical state of the reference site was determined by mining the user travel conditions of nearby sites. The hidden demand of the site was determined based on the site state change diagram and the demand judgment model. Then, according to the real needs of the site, the long short-term memory(LSTM) network prediction model was established, and the regional scheduling model of shared bicycles based on the real needs was established. The model takes the cost minimization as the goal, and obtains the path with minimum scheduling cost through genetic algorithm, which provides a reference for balanced scheduling based on real demand. The results demonstrate that, when transportation costs are similar, the scheduling method under real demand can alleviate the problem of users' difficulty in renting a bike, thereby reducing the loss of high-frequency users.

[Keywords] shared bike; demand forecasting; implicit needs; long short-term memory; scheduling optimization

共享单车作为一种备受欢迎的出行方式,不仅具有较高的运行效率,同时也降低了对城市环境的污染,是一种经济、高效的绿色出行方式。它也有效地缓解了拥堵、碳排放等城市交通问题,加快了城市推进交通节能减排的步伐。共享单车主要是为短距离的出行需求服务,不仅可以为用户的日常通勤、休闲娱乐提供不同的交通选择,还可满足用户不同层次的出行需求^[1]。由于人们出行需求的多样性和不确定性,导致运营站点间出现不均衡的状态^[2-3]。比如在早高峰时,用户在居民区的需求高于

商业区的需求,晚高峰则反之,导致用户出现“借车难”的困境。而由此带来的则是一些用户的流失和共享单车系统收益的降低。借车难导致的用户流失情况并没有相应的数据记录,从经济学的角度来看,这部分用户的需求成为了共享单车系统的“隐性需求”^[4]。所谓隐性需求,是指用户虽然在某站点有强烈的租借共享单车的意愿,但该站点共享单车的可用水平不能满足用户需求的现象。因此,对共享单车用户的隐性需求进行研究可以提高用户的满意度,进而有助于共享单车的调度优化应用。

收稿日期:2023-10-08 修订日期:2024-07-08

基金项目:内蒙古自治区自然科学基金(2023MS07005)

第一作者:周瑜(1984—),男,汉族,内蒙古赤峰人,博士,副教授。研究方向:商业数据科学。E-mail:zhouyu@imu.edu.cn。

投稿网址:www.stae.com.cn

需求预测对共享单车系统具有重要意义。魏晋等^[5]考虑到建成环境中各影响因素的交互作用,采用梯度提升决策树模型预测共享单车出行需求。刘恒孜等^[6]针对共享单车骑行数据的周期性、非线性和随机性的特点,提出了季节性灰色 Markov 模型来预测共享单车需求量。曹旦旦等^[7]分析时间因子、气象因子等对单车需求量的影响,采用长短期记忆(long short-term memory, LSTM)神经网络模型对共享单车的短期需求量进行预测。关于隐性需求, Jin 等^[8]以纯电动的共享汽车为研究对象,建立并校准了嵌套的 Logit 模型来分析用户的模式选择行为,进一步利用实际出行数据估计电动共享汽车的隐性需求。此外,还分析了隐性需求的时空分布、纯电动共享汽车对模式分割的影响以及定价策略的影响。研究表明,隐性需求都集中发生在早晚高峰,而且大部分以居住区或综合区域为起点,以商业区或绿地为终点的两点一线模式。Kim 等^[9]也使用预测模型对客户的购买意愿进行研究,通过建立的遗传算法模型对客户的隐性需求进行预测,预测结果也起到了一定的帮助。刘林等^[10]通过所构建的映射网络(self-organizing map, SOM)对消费者进行研究,发现消费者之间也存在不同的消费类别,实现了对不同类别的消费者制定不同的策略并有效地对潜在客户进行挖掘。基于上述的文献梳理,大多数学者都是针对共享单车系统产生的数据从而对用户的需求进行研究,针对用户想借车却无车可借这部分需求的研究较少,现通过站点状态变化图和真实需求判定模型更新站点的真实需求,在合理假设的范围内建立共享单车系统的区域调度模型,并以成本最小化为目标,通过遗传算法求解得到调度成本最小的路径,进一步帮助企业制定合理的调度策略,解决用户借车困难的问题,使想借车的租车用户都有车可借。

1 真实需求分析

用户在高峰时期的真实需求包括成功借还需求(下文简称为显性需求)和想借车却无车可借需求(下文简称为隐性需求)两部分。共享单车的借还记录明显体现了站点中存在的显性需求,但对骑行数据进一步研究,会发现数据中不仅存在站点的显性需求,同时也可以挖掘出站点的隐性需求。因此,接下来首先将隐性需求这部分缺失的数据通过需求判定方法进行补全,进而进行租车用户真实需求的预测。

关于站点真实需求的具体研究思路如图 1 所示,得到不同站点每隔 10 min 的借还记录后,根据

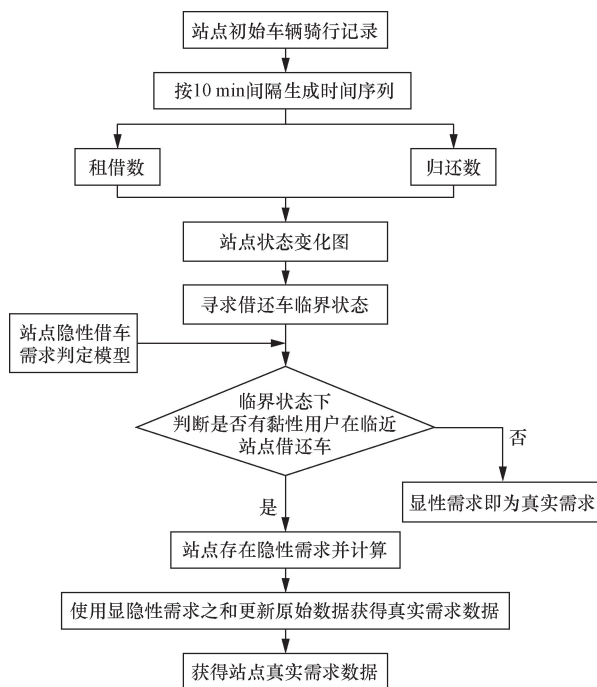


图1 真实需求分析流程

Fig. 1 Analysis process of real demand

租借数和归还数绘制出站点状态变化图,寻求站点的两种临界状态并判断在该状态下,如果有高频用户在临近站点借车,则该站点的隐性需求为该时间段下在临近站点借车的用户数量,真实需求用隐形需求和显性需求之和来更新。反之,站点不存在隐形需求,显性需求则为真实需求。

1.1 站点状态分析

采集了呼和浩特市共享单车系统 4 年的骑行数据进行用户隐形需求挖掘。为了详细分析站点的隐性需求,首先需要了解站点车辆数量随借还行为发生的变化情况,随机选取核心通勤类别中的金地商城站点并绘制其租车变化情况,由于该数据集中不包含站点初始车辆信息,因此将站点的初始状态设置为 0,结果如图 2 所示。当用户发生借车行为时,由于站点的共享单车被借走,因此其状态值变为负数;当用户在该站点进行还车时,站点的共享单

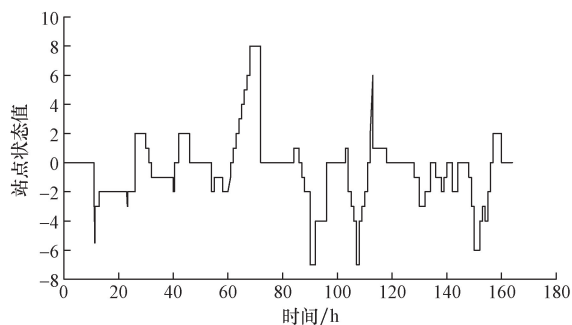


图2 站点状态变化

Fig. 2 Status change of stations

车增加,则状态值变为正。

根据该站点状态变化图可以看出,虽然波动但站点的共享单车总是处于一种动态平衡中,不难看出该站点状态变化图的两条临界线分别是 $y = 8$ 和 $y = -7$ 这两条直线,在 68 ~ 76 h,站点处于还车饱和状态,在 88 ~ 92 h,站点处于借车饱和状态,也就是说该站点在 t 时刻的车辆数 $y_2(t) = 8$ 时表示该时间段站点已经无桩可还,在 t 时刻的车辆数 $y_2(t) = -7$ 时表示已经无车可借。

对站点状态变化图继续进行研究,若站点在 $y_2(t) = -7$ 的临界状态下维持时间超过 10 min,并且在这 10 min 内有用户到达该站点想借车却无车可借,则该用户很大程度会选择附近的站点继续借车或更改出行方式;同理在 $y_2(t) = 8$ 的临界状态下,若有用户在 10 min 之内想到该站点还车却无桩可还,用户则会将共享单车还到临近站点。由于临界状态致使用户在附近站点借还车的需求称为隐性需求问题。

1.2 真实需求判定

由于站点在某一时刻的真实需求由显性需求和隐性需求两部分构成,对现有数据进行简单分析只能得到租车用户的显性需求,若想判断隐性需求还需对数据进一步探究。以金地商城站点为例,提出了判断隐形需求的方法,进而确定站点真实需求。

首先将站点中的借还车数据按照工作日和周末进行分类并将其提取出来,基于此建立站点的车辆状态变化图,对其分析无车可借和无桩可还的两种临界状态,在两种临界状态下判断是否存在该站点的高频用户到附近站点借还车并统计高频用户的数量,统计出的用户数量则为该站点该时间段下的隐性借车需求,将显性需求和隐性需求之和作为该站点该时间段下的真实需求。

具体来说,以 7 d 作为时间间隔获取站点 1 的车辆状态变化图,首先从 3 月 4 日星期四开始获取站点 1 的数据,根据 7 d 的时间间隔依次获取 3 月 11 日、3 月 18 日和 3 月 25 日的借还车数据,最后以 1 h 为时间间隔绘制车辆状态变化图,具体绘制情况如图 3 所示。

对图 3 的状态变化图进一步分析,以 1 h 为单位将一天的 24 h 进行划分,并对不同时间段下共享单车的借还情况进行统计,根据站点状态变化曲线分析站点车辆数量的变化以及存在两种临界状态的时间段,对两种临界状态可以看作是该时刻该站点无法满足接下来一段时间内的车辆借/还需求,即该站点在该时间段具有隐性需求。

由图 3 可以得到,站点 1 在 3 月 4 日 17:30 之后、

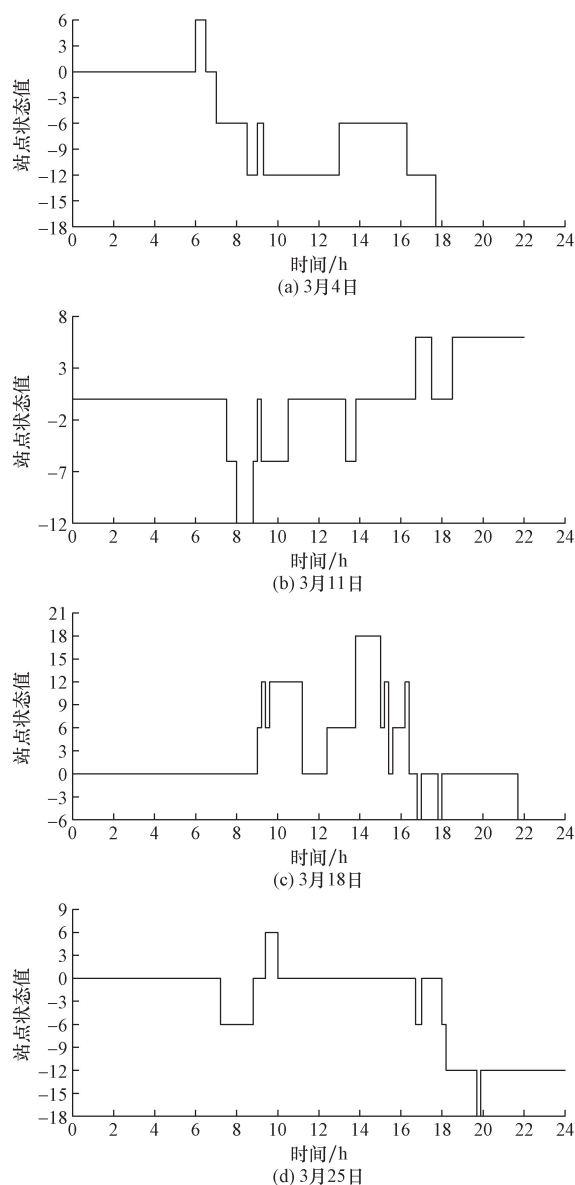


图3 金地商城站点状态连续变化

Fig. 3 Continuous change of the Golden Mall station

3 月 11 日 08:00—09:00、3 月 18 日 17:00—18:00、3 月 25 日 19:00—20:00 这几个时间段都达到最大借车需求并且持续时间超过 10 min,因此站点 1 在这几个时间段都存在隐性借车需求。以 3 月 4 日 17:30 之后的时间段为例,找出该站点频繁于该时间段借车并且 3 月 4 日 17:30 之后在附近站点借车的用户,得到该时间段下的隐性借车需求,从而补充该时间段下站点 1 的真实借车需求。图 4 给出了 3 月第一周中金地商城站点真实需求与显性需求的差异图。可以看出,真实需求和显性需求有着明显的不同,两曲线之间的部分为站点的隐性需求,存在隐性需求的时间段大多分布在早中晚 3 个高峰,也充分证实了“借车难”在高峰时刻发生的这一现状。该部分更新后的真实需求数据将作为预测的基础数据。

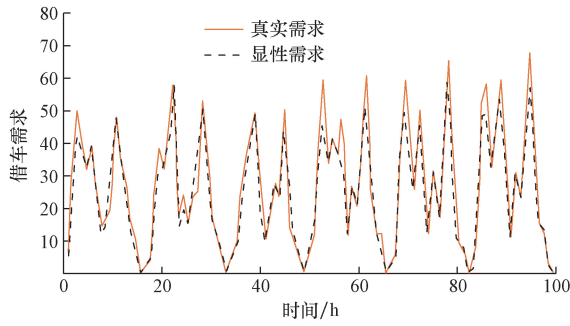


图4 真实需求与显性需求差异可视化

Fig. 4 Visualizing the difference between real and implicit demands

2 基于 LSTM 的用户真实需求预测

2.1 LSTM 方法

LSTM 模型创建了输入门、遗忘门、输出门,而对任意时刻的状态和输出进行调控,并且将短期记忆和长期记忆连接起来,解决了传统循环神经网络(recurrent neural network, RNN)由于网络层无限制地更新信息,信息变得混乱且容易消失,从而导致梯度消失问题。LSTM 结构如图 5 所示^[11-12]。

整个过程中信息的传递受到 3 个门的控制,遗忘门是通过 h_{t-1} 和 x_t 的数值来判断将哪些信息删除^[11-12]。遗忘门的表达式为

$$f_t = \sigma(W_f[h_{t-1}, x_t]) + b_f \quad (1)$$

式(1)中: $\sigma(\cdot)$ 为 sigmoid 激活函数; W_f 为遗忘门权重; b_f 为遗忘门偏置。

输入门根据 h_{t-1} 和 x_t 向记忆单元中新增信息。

$$i_t = \sigma(W_i[h_{t-1}, x_t]) + b_i \quad (2)$$

$$\hat{C}_t = \tanh(W_c[h_{t-1}, x_t]) + b_c \quad (3)$$

式中: i_t 为需要记忆的信息; $\hat{C}_t$ 为候选记忆单元,模型中被用来更新记忆单元; W_i 、 W_c 为输入门权重; b_i 、 b_c 为输入门偏置。

数据经过遗忘门和输出门运算过后,采用式(4)进行记忆单元更新。

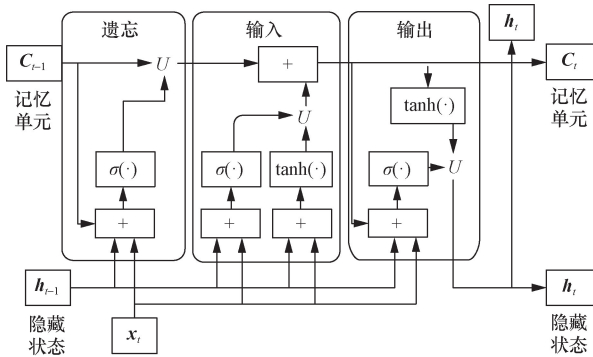


图5 LSTM 神经网络结构

Fig. 5 LSTM architecture

$$C_t = f_t \circ C_{t-1} + i_t \circ \hat{C}_t \quad (4)$$

式(4)中: $\circ$ 为哈达玛乘积。

隐藏状态 h_t 由输出们根据 h_{t-1} , x_t 和 C_t 决定。具体方式为式(5)和式(6)。

$$o_t = \sigma(W_o[h_{t-1}, x_t]) + b_o \quad (5)$$

$$h_t = o_t \circ \tanh C_t \quad (6)$$

式中: W_o 为输出门权重; b_o 为输出门偏置。

LSTM 通过“遗忘门”判断将哪些信息进行输入,“输入门”对信息进行输入,再通过“输出门”将需要的信息输出,从而解决 LSTM 模型的长期依赖问题。同时,选择 sigmoid 激活函数作为最终选择工具,用 0~1 来做权重的二次分配,并且用 tanh 函数作为变化工具。当前的单元信息是通过输入门控制之后叠加来解决梯度消失问题。

2.2 数据预处理及建模

LSTM 网络模型的建立主要由数据预处理、参数初始化、权重的确定以及训练输出 4 个部分构成。LSTM 建模流程如图 6 所示。

(1)数据预处理。由于 Sigmoidal 作为训练过程中的激活函数,因此所用的 LSTM 模型敏感度较高,对数据也有一定的要求。通过 MATLAB 中的 mapminmax 函数归一化数据,可以有效地提高模型精度,同时为后续建模奠定基础。

(2)参数初始化。通过重置输入数据的训练集形式,对优化器,损失函数以及其他基本参数进行定义。

(3)权重的确定。在模型训练过程中,最本质的是对于指定数据集中各神经元的连接关系,而权重则可以作为神经元之间联系的数据化,对网络的训练是基于量化共轭梯度(scaled conjugate gradient, SCG)算法实现的。

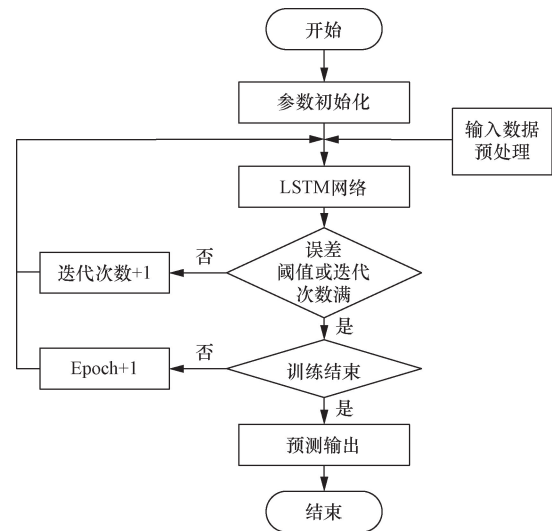


图6 LSTM 建模流程

Fig. 6 LSTM modeling process

(4) 训练输出。当训练完成后, 要对数据和网络分别进行反归一化之后便可输出结果。

针对站点真实需求的分析, 选取内蒙古医院及名都中央广场两个代表性站点进行研究, 选取 1 月 31 日—8 月 31 日期间 184 个工作日的显性需求和真实需求作为数据样本。由于共享单车在晚上 23:00—次日 06:00 作为工作人员的调度时间, 因此只针对一天中的 06:00—23:00 进行研究。站点数据的预处理时首先需要将站点真实需求和显性需求中的时间和借还数量等属性保留。以内蒙古医院为例, 整理后的站点需求数据如表 1 所示。

表 1 内蒙古医院的需求记录
Table 1 Demand records for Inner Mongolia Hospital

时间	真实需求	显性需求
2016/1/1 06:00—07:00	3	3
2016/1/1 07:00—08:00	6	6
⋮	⋮	⋮
2016/8/1 22:00—23:00	0	0
2016/8/2 07:00—08:00	7	5
⋮	⋮	⋮
2016/8/31 22:00—23:00	0	0

由于整理后的骑行数据极其分散, 不具有特征性, 不能直接训练。因此需要调用 mapminmax 函数对数据预进行处理, 按照一定的规则将原始数据映射到 $[-1, +1]$ 上, 这样可以通过传递函数优化对数据的敏感程度。

具体映射过程^[13-14]表达式为

$$y = \frac{y_{\max} - y_{\min}(x - x_{\min})}{x_{\max} - x_{\min}} + y_{\min} \quad (7)$$

式(7)中: x 和 y 分别为原始数据和处理数据; $x_{\max}$ 和 $x_{\min}$ 分别为原始数据的最大、最小值; $y_{\max}$ 和 $y_{\min}$ 分别为处理数据的最大、最小值。

2.3 建模结果与评价

对于隐藏层而言, 其本质就是获取输入数据的特征变量, 隐藏层中的层数会对模型产生不同的影响, 过多会增加模型复杂度, 同时也加长了运算时间。针对于此, 大量学者经常选择把隐藏层设置为 1~5 层作为模型参数, 层数的确定要实时的依据样本数量而变化^[11-12, 15]。

模型的拟合效果很大程度上与隐藏层中神经元的数量有关, 较少的神经元数量会导致较差的拟合结果, 而过多的神经元也会对模型产生一定的影响, 比如使模型的泛化能力降低、结果过拟合等。所以神经网络模型构建的基础是选择合适的神经元参数。所建立的预测模型对输入输出和迭代次数的设定更多地参考了大量学者的经验, 将 LSTM

的输入和输出维度均设置为 1, 隐藏层为 4, 隐藏层中神经元的个数设置为 8 进行预测, 学习速率设置为 0.1。除此之外, 为了确保模型避免过拟合的情况, 将 70% 数据作为训练样本, 其余 30% 作为测试样本进行站点真实需求的预测。

为了判断预测效果, 选择对实验结果进行量化, 选取了预测中经常使用的三类指标: 平均绝对误差 (mean absolute error, MAE)、均方根误差 (root mean square error, RMSE) 以及拟合度 R^2 评估预测的效果, 计算方式^[11-12, 16]如下。

$$y_{\text{MAE}} = \frac{1}{n} \sum_{i=1}^n |X_{\text{act}}(i) - X_{\text{pre}}(i)| \quad (8)$$

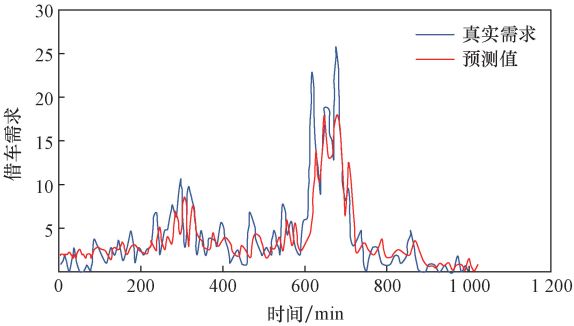
$$y_{\text{RMSE}} = \sqrt{\frac{\sum_{i=1}^n [X_{\text{act}}(i) - X_{\text{pre}}(i)]^2}{n}} \quad (9)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (X_{\text{act}} - X_{\text{pre}})^2}{\sum_{i=1}^n X_{\text{act}}^2} \quad (10)$$

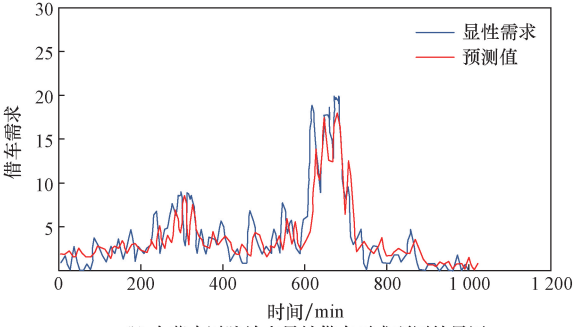
式中: n 为预测总次数; X_{act} 和 X_{pre} 分别为 i 时刻站点真实需求的真实值和预测值。与标准差的计算原理相似, RMSE 首先将真实值与预测值差值的平方和与预测个数作比再对其开平方根, MAE 可以将真实值和预测值之间的误差标记出来, 因此, MAE 和 RMSE 越小则表示模型精度越好。除此之外, 拟合度 R^2 也作为模型评估的重要指标, 其值越接近 1 表示模型预测精度越好。

对随机选取的内蒙古医院和名都中央广场两个代表性站点进行研究, 得到了 LSTM 模型对真实需求和显性需求的借车预测效果如图 7 所示, 可以看出站点在高峰时间段内显性需求和真实需求存在一定的差异, 因此站点存在隐性需求。除此之外在预测的时间段内, 预测的真实需求曲线与实际真实需求曲线大体吻合, 显性需求的预测值和真实值的偏差也比较小, 可以看出该模型对于局部值的预测也有良好的精度, 这与 LSTM 模型对于非线性的处理能力和具有序列数据特点的记忆能力是分不开的。

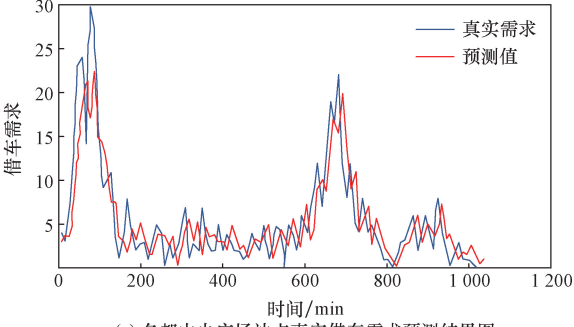
为了进一步证明 LSTM 模型的可行性, 在相同的实验条件和数据集下, 对反向传播神经网络^[17] (back propagation neural network, BPNN) 也进行了测试, 表 2 和 3 给出了在真实需求和显性需求的实验数据下 LSTM 模型和 BPNN 模型的预测效果对比, 不难看出, 在与 BPNN 的对比下, LSTM 预测模型呈现出更好的预测优势, 因此, 对于站点真实需求的预测是行之有效的。



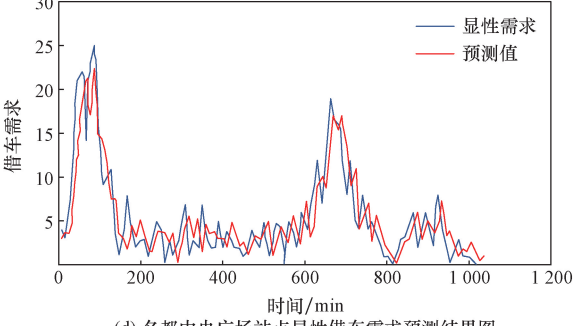
(a) 内蒙古医院站点真实借车需求预测结果图



(b) 内蒙古医院站点显性借车需求预测结果图



(c) 名都中央广场站点真实借车需求预测结果图



(d) 名都中央广场站点显性借车需求预测结果图

图7 LSTM 预测结果

Fig.7 LSTM prediction results

表2 内蒙古医院站点需求预测误差比较

Table 2 Comparison of demand prediction errors for Inner Mongolia Hospital station

模型 需求类型	LSTM		BPNN	
	真实 需求	显性 需求	真实 需求	显性 需求
拟合度 R^2	0.668 3	0.731 5	0.438 0	0.521 8
平均绝对误差(MAE)	2.034 9	1.719 4	2.374 1	2.081 2
均方根误差(RMSE)	3.105 2	2.830 6	3.377 6	2.916 3

表3 名都中央广场站点需求预测误差比较

Table 3 Comparison of demand prediction errors for Mingdu Central Plaza station

模型 需求类型	LSTM		BPNN	
	真实 需求	显性 需求	真实 需求	显性 需求
拟合度 R^2	0.684 6	0.613 7	0.490 2	0.430 6
平均绝对误差(MAE)	2.128 9	2.306 4	2.517 3	2.712 9
均方根误差(RMSE)	3.081 7	3.159 2	3.353 2	3.492 1

3 基于真实需求预测的调度优化

由于用户在不同时间段下的骑行需求不同,导致共享单车无车可借现象频频发生,使共享单车在出行方式中的分担率直线下降。因此,以真实需求的分析结果为依据,以成本最小作为目标函数建立调度模型,从而对共享单车系统的再平衡进行深入研究。与以往的车辆路径问题相比,由于共享单车具有高灵活性,因此对其调度系统的效率也有更高的要求。在一定的约束条件下,对共享单车系统进行合理的调度可以增加共享单车的利用率,进而提高用户满意度。

在对共享单车系统的调度进行研究时,需要将共享单车、站点、调度中心、目标函数等作为构成要素纳入考虑范围之内。假设所有共享单车具有相同质量和运输条件,除此之外,调度车具有相同的行驶速度和容量,不考虑其内在因素对调度产生的影响。

3.1 调度模型建立

作为一个复杂系统问题,共享单车系统的调度模型会受到很多参数以及条件的影响,为了使调度模型得到更优的调度路径,首先对模型进行了以下假设。

(1)调度区域:由站点和调度中心组成。在调度过程中,由于各个站点距离较近且需求量较小,因此假定一辆调度车即可完成一个区域的调度工作,其次,不考虑外界及其他因素对调度车产生的影响,对调度车的容量及行驶里程也会事先设定。假定站点之间的调度最多跨越2个站点;比如调度路线为1-2-3;站点1需求量为10,站点2需求量为-5,站点3需求量为-2(负数代表调入,正数代表调出),此时站点1可向站点2,站点3分别调入各子需求量。

(2)调度中心及站点:调度区域中有多个站点,由于站点间需求不同,需在调度区域中建立一个调度中心,调度车由此出发完成调度,最后回到调度中心。

(3)共享单车:共享单车具有相同的质量、型号

且总量固定,除此之外,共享单车只有在桩和被骑行两种状态。

(4)时间间隔:调度过程以 1 h 为周期,对一个周期内的借还需求进行统计和预测,假设站点的调度需求周期内不发生变化,每个站点也都有可以接受的时间窗。

对建模过程中所需要的变量和参数如表 4 和表 5 所示。

根据对提出的基于真实需求的调度优化问题,以及参数和变量的定义,参考已有研究^[6,18-19],构建了以成本最小为目标的调度模型,如式(11)~式(22)所示。

以调度总成本最小作为目标函数,其中总成本包括调度车的固定成本、行驶成本以及惩罚成本。

表 4 参数及含义

Table 4 Symbols and explanations

参数	含义
t	$t = 1, 2, 3, \dots$ 表示需要调度的时间段
v	$v = 1, 2, 3, \dots$ 表示调度车辆
B_i^0	调度区域初始共享单车数量
U_i^t	站点 i 在 t 时间段内的借还车数量
A_i^t	站点 i 在 t 时间段内对车辆的需求
Q	调度车所能承载的最大共享单车数量
C_v	调度车的固定成本
E_v	调度车的行驶成本
ω	惩罚成本
$\bar{t}$	调度过程中,每辆共享单车的调度时间
d_{ij}	两个站点之间的距离
t_{ijv}	调度车在两个站点间的行驶时间

表 5 变量符号及含义

Table 5 Variables and explanations

变量	含义
y_{ijv}^t	某时间段内调度车对 i, j 站点的任务完成情况, 完成为 1, 反之为 0
x_{ijv}^t	某时间段内站点 i 向 j 的调入量
z_{1ij}^t	站点 i 对 j 的调入量 (≥ 0)
z_{2ij}^t	站点 i 对 j 的调出量 (≥ 0)
ts_{iv}	调度车的服务时间
TU_{ijv}^t	调度车的超时时长

其中约束条件(12)表示调度区域当前车辆总数等于上一阶段末调度后的车辆数,式(13)表示区域内共享单车的总量要超过用户的总需求,约束条件(14)对调度车的容量进行了限定,式(15)~式(17)对调度时间进行设定,式(18)表示调入站点 j 的共享单车数量等于调出站点 i 的车辆数,式(19)对调度车的数量进行判断,式(20)规定调度车的行驶距离不能超过其最大行驶距离,式(21)、式(22)表示对变量的约束。

3.2 数值算例

根据对租车用户骑行偏好的研究可知,共享单车在早中晚高峰时期下的借还频次最高,其余时间段用户对共享单车的需求相对较少。基于此可以认为在非高峰时期,通过其本身的运营,共享单车系统可以自己达到平衡状态,但在早中晚的高峰时刻,由于站点借还频率较高难以自主平衡,因此在高峰期需要人工调度进行再平衡。

$$\min Z = C_v \sum_{t \in T} \sum_{i=1}^N \sum_{j=1}^N \sum_{v=1}^V y_{ijv}^t + E_v \sum_{t \in T} \sum_{i=1}^N \sum_{j=1}^N \sum_{v=1}^V d_{ij} y_{ijv}^t + w \sum_{t \in T} \sum_{i=1}^N \sum_{j=1}^N \sum_{v=1}^V TU_{ijv}^t \quad (11)$$

$$B_i^t = B_i^{t-1} + \sum_j z_{1ij}^{t-1} - \sum_j z_{2ij}^{t-1} - u_i^t, \quad \forall i \in I, t \in 2, 3, \dots, T \quad (12)$$

$$B_i^{t+1} \geq A_i^t, \quad \forall i \in I, t \in T \quad (13)$$

$$x_{ijv}^t \leq Q, \quad \forall i, j \in I, t \in T, v \in V \quad (14)$$

$$x TU_{ijv}^t \geq t_{ijv} + ts_{iv} + ts_{jv} - 1, \quad \forall v \in V \quad (15)$$

$$ts_{iv} \geq \sum_i z_{ij}^t \bar{t}, \quad \forall i \in I, v \in V, t \in T \quad (16)$$

$$\text{s. t. } \begin{cases} ts_{iv} \geq \sum_j z_{2ij}^t \bar{t}, & \forall i \in I, v \in V, t \in T \end{cases} \quad (17)$$

$$z_{1ij}^t = z_{2ij}^t, \quad \forall i, j \in I, t \in T \quad (18)$$

$$y_{i,j,v}^t \geq y_{i,j,v+1}^t, \quad \forall i, j \in I, t \in T, v \in V \quad (19)$$

$$\sum_{i=1}^N \sum_{j=1}^N d_{ij} x_{ijv}^t \leq D_v, \quad \forall t \in T, v \in T \quad (20)$$

$$y_{ijv}^t \in \{0, 1\}, \quad \forall i \in I, v \in V, t \in T \quad (21)$$

$$x_{ijv}^t, z_{ij}^t, B_i^t, ts_{iv}, TU_{ijv}^t \geq 0, \quad \forall i, j \in I, t \in T, v \in V \quad (22)$$

表 6 各站点的需求情况
Table 6 Demand of stations

序号	站点名称	坐标 $X/(^{\circ})$	坐标 $Y/(^{\circ})$	真实需求量	显性需求量	调入说明
1	市政府	111.755 426	40.848 119	-10	-4	调出
2	内蒙古教育厅	111.739 961	40.846 132	12	5	调入
3	内蒙古政协	111.740 818	40.844 128	13	8	调入
4	交通设计院	111.704 874	40.835 108	-4	-1	调出
5	鼓楼广场	111.695 257	40.830 810	-3	-4	调出
6	内蒙古医院	111.690 564	40.841 826	5	3	调入
7	呼铁佳园	111.717 505	40.847 953	-6	-4	调出
8	名都中央广场	111.695 671	40.858 265	10	7	调入
9	呼呼和佳地	111.680 633	40.857 369	12	9	调入
10	海天花园	111.699 330	40.847 253	-2	-2	调出
11	金地商城	111.647 486	40.859 091	3	1	调入
12	内蒙古广播电视台	111.686 330	40.833 154	1	1	调入
13	艺术厅北街	111.721 371	40.845 575	3	0	调入
14	芳汀花园	111.690 902	40.829 282	11	7	调入
15	电子信息质检院	111.697 413	40.833 613	-2	-1	调出

随机选择该类别下的 15 个站点作为一个调度区域进行模型的求解。并以 8 月 31 日为例,根据 LSTM 模型求得各个站点的真实需求和显性需求及预测结果计算出高峰时期下的调度需求量,如表 6 所示。模型求解中,将调度区域的中心作为调度中心,调度车从调度中心出发,完成对各个站点的调度后再回到调度中心。采用遗传算法对该模型进行求解,目标函数和遗传算法的参数取值如表 7 和表 8 所示。

最终得到真实需求下共享单车的调度方案为:

0→1→11→10→8→9→7→12→3→13→5→15→2→6→4→14→0。

表 7 目标函数参数取值

Table 7 Parameter of the objective function

参数	取值	说明
Q	100	调度车对共享单车的最大承载数量
C_v	100	调度车的固定成本
E_v	4	每公里调度车的行驶成本
ω	80	调度车在周期内未完成任务的惩罚
$\bar{t}$	60	1 h 内能调度的最大单车数量
D^v	80	调度过程中调度车的最远行驶距离

表 8 遗传算法参数取值

Table 8 Parameter of the genetic algorithm

参数	数值	参数	数值
种群大小	100	变异概率	0.05
交叉概率	0.90	最大迭代次数	600

得到显性需求下共享单车的调度方案为:

0→3→5→14→12→10→2→9→1→6→8→7→15→11→4→13→0。

表 9 给出了真实需求和显性需求在每个过程中依次的调度量和在运量。在各调度阶段的调度量中,负值表示在当前阶段调度车要从站点调出共享单车,正值则表示调入,可以看出真实需求和显性需求的调度量存在差异,原因在于真实需求下的调度方式考虑了隐性需求。对于在运量而言,由于各站点的需求不同,显性需求下的在运量一直减少,最大在运量则为初始调度车的载运量,而真实需求下的在运量呈现出了先增加后减少的趋势。

表 10 给出了调度过程中的初始装载量和调度总量,对于显性需求而言,调度车的初始装载量为 30 辆,整个调度过程中调度总量为 57 辆车,而在考虑隐性需求后,真实需求调度过程中的初始装载

表 9 各阶段调度数量

Table 9 Scheduling results

调度阶段		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
显性需求	调度量	0	8	-4	7	1	-2	5	9	-4	3	7	-4	-1	1	-1	0
	在运量	30	22	26	19	18	20	15	6	10	7	0	4	5	4	5	5
真实需求	调度量	0	-10	3	-2	10	12	-6	1	13	3	-3	-2	12	5	-4	11
	在运量	43	53	50	52	42	30	36	35	22	19	22	24	12	7	11	0

表 10 具体调度情况对比

Table 10 Comparison of scheduling results

需求类型	初始装载量/辆	调度总量/辆
显性需求	30	57
真实需求	43	82

量为 43 辆,调度总量为 82 辆。由于调度车对共享单车的最大承载量为 100 辆,因此两种需求下都只需要一辆调度车就可以完成所有调度,两种需求下的成本差距微乎其微。可以看出在相近的运输成本下,真实需求下的调度方式能在一定程度上缓解用户借车难的问题,从而减少高频用户的流失。

4 结论

作为绿色交通的重要组成部分,共享单车不仅提供了绿色、经济的出行方式,对居民而言也解决了其最后一公里的问题。针对借还车困难引起的用户流失及站点真实需求问题,通过站点状态变化图判断站点是否存在隐性需求,并通过需求判定模型将其真实需求进行校正。通过建立的 LSTM 预测模型实现对站点真实需求的预测。在与 BPNN 模型的对比之下,所使用的预测模型表现出较好的预测精度。最后,对共享单车系统的预测结果进行应用研究。基于以上预测结果进行共享单车调度模型的建立,通过遗传算法求解得到成本最小的调度路线,在提升用户满意度的同时也为共享单车系统的运营减轻了成本。研究结果为共享单车的精准管理提供了参考。

参 考 文 献

- [1] Luo H. Comparative life cycle assessment of station-based and dock-less bike sharing systems[J]. Resources, Conservation, and Recycling, 2019, 146(3): 180-189.
- [2] Wang Y, Szeto W Y. An enhanced artificial bee colony algorithm for the green bike repositioning problem with broken bikes[J]. Transportation Research Part C: Emerging Technologies, 2021, 125: 102895.
- [3] Zhang B, Li X, Saldanha-da-Gama F. Free-floating bike-sharing systems: new repositioning rules, optimization models and solution algorithms[J]. Information Sciences, 2022, 600: 239-262.
- [4] 徐雅婷. 乡镇企业主导下乡村振兴的发展途径——基于消费者隐性需求挖掘视角[J]. 经济研究导刊, 2022(13): 7-9.
Xu Yating. The development way of township enterprises' industrial revitalization[J]. Economic Research Guide, 2022(13): 7-9.
- [5] 魏晋, 安实, 张炎棠. 考虑建成环境交互影响的共享单车需求预测[J]. 科学技术与工程, 2023, 23(26): 11424-11430.
Wei Jin, An Shi, Zhang Yantang. Prediction of shared bicycle demand based on environmental factor interactions[J]. Science Technology and Engineering, 2023, 23(26): 11424-11430.
- [6] 刘恒孜, 贺玉龙, 宋太龙, 等. 共享单车需求预测及调度优化

- [J]. 科学技术与工程, 2021, 21(35): 15247-15254.
- Liu Hengzi, He Yulong, Song Tailong, et al. Demand forecast and scheduling optimization of shared bicycles[J]. Science Technology and Engineering, 2021, 21(35): 15247-15254.
- [7] 曹旦旦, 范书瑞, 张艳, 等. 基于长短期记忆神经网络模型的共享单车短时需求量预测[J]. 科学技术与工程, 2020, 20(20): 8344-8349.
Cao Dandan, Fan Surui, Zhang Yan, et al. Short-term demand forecasting of shared bicycles based on long short-term memory neural network model[J]. Science Technology and Engineering, 2020, 20(20): 8344-8349.
- [8] Jin F, Yao E, An K. Analysis of the potential demand for battery electric vehicle sharing: mode share and spatiotemporal distribution[J]. Journal of Transport Geography, 2020, 82: 102630.
- [9] Kim E, Kim W, Lee Y. Combination of multiple classifiers for the customer's purchase behavior prediction[J]. Decision Support Systems, 2003, 34(2): 167-175.
- [10] 刘林, 喻国平. 基于自组织特征映射(SOM)网络对潜在客户的挖掘[J]. 南昌大学学报: 理科版, 2006, 30(5): 507-510.
Liu Lin, Yu Guoping. Application of discovering the potential customers based on self-organizing feature map(SOM) neural network[J]. Journal of Nanchang University (Natural Science), 2006, 30(5): 507-510.
- [11] 王永志, 刘博, 李钰. 一种基于 LSTM 神经网络的电力负荷预测方法[J]. 实验室研究与探索, 2020, 39(5): 41-45.
Wang Yongzhi, Liu Bo, Li Yu. A power load data prediction method based on LSTM neural network model[J]. Research and Exploration in Laboratory, 2020, 39(5): 41-45.
- [12] 李娜, 刘强, 张森, 等. 基于长短期记忆的煤炭含水率智能预测与实验研究[J]. 实验室研究与探索, 2022, 41(4): 35-39, 56.
Li Na, Liu Qiang, Zhang Miao, et al. Intelligent prediction and experimental research of coal moisture content based on LSTM[J]. Research and Exploration in Laboratory, 2022, 41(4): 35-39, 56.
- [13] 李光旭, 彭怡, 寇纲. 不确定幂加权几何平均算子的动态多目标决策[J]. 系统工程理论与实践, 2015, 35(7): 1855-1862.
Li Guangxu, Peng Yi, Kou Gang. Dynamic multiple criteria decision making method with uncertain power geometric weighted average operators[J]. Systems Engineering-Theory & Practice, 2015, 35(7): 1855-1862.
- [14] 李艳虹, 寇纲, 彭怡, 等. 基于矩阵相似性的异构群体偏好信息集结方法[J]. 系统科学与数学, 2019, 39(10): 1672-1683.
Li Yanhong, Kou Gang, Peng Yi, et al. Heterogeneous group preference information aggregation method based on matrix similarity[J]. Journal of Systems Science and Mathematical Sciences, 2019, 39(10): 1672-1683.
- [15] 杨忠富, 常俊, 许妍, 等. 联合胶囊和双向 LSTM 网络的 VPN 加密流量识别[J]. 计算机工程与应用, 2023, 59(23): 246-253.
Yang Zhongfu, Chang Jun, Xu Yan, et al. VPN encrypted traffic identification for joint capsule and bidirectional LSTM networks[J]. Computer Engineering and Applications, 2023, 59(23): 246-253.
- [16] 周瑜, 寇纲, 白杨, 等. 基于变点-算术失效强度减的设备维

- 修质量评估[J]. 系统工程学报, 2021, 36(2): 247-255, 263.
- Zhou Yu, Kou Gang, Bai Yang, et al. Assessing equipment maintenance quality based on change point-arithmetic reduction of failure intensity model[J]. Journal of Systems Engineering, 2021, 36(2): 247-255, 263.
- [17] Xu X, Ye Z, Li J, et al. Understanding the usage patterns of bicycle-sharing systems to predict users' demand: a case study in Wenzhou, China [J]. Computational Intelligence and Neuroscience, 2018, 2018: 9892134.
- [18] 陈植元, 林泽慧, 金嘉栋, 等. 基于时空聚类预测的共享单车调度优化研究[J]. 管理工程学报, 2022, 36(1): 146-158.
- Chen Zhiyuan, Lin Zehui, Jin Jiadong, et al. Optimization of shared bicycle scheduling based on spatio-temporal clustering prediction[J]. Journal of Industrial Engineering and Engineering Management, 2022, 36(1): 146-158.
- [19] 李珍萍, 刘璐, 王晨. 共享单车调度优化问题研究[J]. 数学的实践与认识, 2021, 51(6): 30-40.
- Li Zhenping, Liu Lu, Wang Chen. Research on scheduling optimization problem of shared bikes[J]. Journal of Mathematics in Practice and Theory, 2021, 51(6): 30-40.