

# 共享单车流量预测与调度方法的研究及应用



## 重庆大学硕士学位论文

(学术学位)

学生姓名：何 欣

指导教师：曾 骏 副教授

学科门类：工 学

学科名称：软件工程

研究方向：流量预测

答辩委员会主席：符云清 教 授

授位时间：2020 年 6 月

# **Research and Application of Traffic Prediction and Rebalancing Method for Shared Bikes**



A Thesis Submitted to Chongqing University  
in Partial Fulfillment of the Requirement for the  
Master's Degree of Engineering

**By**

**He Xin**

**Supervised by Associate Prof. Zeng Jun**

**June, 2020**

## 摘 要

共享单车已广泛应用于诸多大城市中,为用户出行“最后一公里”提供了便捷选择。然而,一个城市的单车使用情况往往不平衡,“潮汐现象”频发。某些站点无车可用,而另一些站点却出现了车辆堆积的情况。如何有效地对各站点的单车流量进行预测,并对各站点的车辆进行合理调度,从而减少用户流失,是目前亟待解决的问题。

海量的用户骑行数据为共享单车流量预测和调度管理提供了机会,使共享单车营运公司在提升营运效率的前提下,节省资源,进而助力“智慧城市”建设。因此,针对共享单车进行流量预测和调度成为了学术界和工业界的热门研究内容。目前大多数的研究方法在建模时,仅考虑了数据本身包含的时空特性,而忽略了其他特征,如气象特征。部分研究工作将用户的历史行为对未来产生的作用视为一致,忽略了用户行为在随时间变化的过程中历史对未来行为的不同影响。此外,部分研究多以激励用户,或设定规则进行单车调度,效率较低。为了解决以上问题,本文提出了一种基于长短期记忆网络和注意力机制的共享单车流量预测模型,并基于强化学习中的 Q-Learning 算法,对共享单车调度问题进行研究。本文主要工作如下:

① 分析了共享单车流量预测和调度的背景意义、国内外研究现状、相关理论以及当前预测和调度算法中存在的不足之处。

② 提出了一种结合骑行上下文信息,并基于 LSTM 网络和注意力机制的共享单车流量预测算法。对共享单车数据集进行了详细地分析,并针对目前流量预测中大多关注时间维度特征而忽略其他影响因素的问题,融入骑行上下文信息,同时进行了特征筛选,并在真实数据集上验证了模型效果。

③ 提出了基于强化学习的共享单车调度策略。利用聚类算法将城市中的共享单车站点进行了区域的划分,在满足区域调度需求的前提下,运用所提算法辅助单车运营管理人员进行科学的调度决策。

④ 共享单车流量预测及调度管理原型系统设计与实现。基于论文中提出的算法,设计了共享单车流量预测和调度管理原型系统。该系统能为共享单车运营管理人员提供详细的用户历史骑行数据,管理人员可查询任一站点的流量预测结果,并且根据给出的建议进行车辆调度。

**关键词:** 流量预测; 单车调度; 神经网络; 注意力机制; 强化学习

## Abstract

Shared bikes have been widely used in many large cities, providing users with a convenient choice for the "last mile". However, bike usage in a city is often uneven and "tide phenomenon" occurs frequently. There are some bikes piled up in some stations, and there are no bikes available in some stations. How to effectively predict the bike traffic at each station and rebalance the bikes at each station to reduce user churn is a problem that needs to be resolved.

Massive user cycling data provides an opportunity for traffic prediction and rebalancing management of shared bikes. It enables bike-sharing companies to save resources on the premise of improving operating efficiency, thus facilitating the construction of "smart city". Therefore, traffic prediction and rebalancing for shared bikes has become a hot topic in academia and industry. Part of the research regards the effect of the user's historical behavior on the future as consistent one and ignores the different effects of the user's historical behavior on the future behavior in the process of changing with time. In addition, some studies have been conducted to incentivize users or set rules for bike rebalancing, which is not efficient. In order to solve the above problems, this thesis proposes a traffic prediction model of shared bikes, which is based on long-short term memory network and attention mechanism. We study the rebalancing problem of shared bikes based on q-learning algorithm in reinforcement learning. The main work of this thesis is as follows:

① This thesis analyzes the research background, research status, relevant theories and shortcomings of the current prediction and rebalancing algorithms for shared bikes.

② This thesis presents a shared bike traffic prediction algorithm based on LSTM network and attention mechanism, which combines cycling context information. The data set of shared bikes is analyzed in detail, and in view of the problem that most of the current traffic prediction focuses on the characteristics of time dimension and ignores other influencing factors, the cycling context information is integrated. At the same time, a method of traffic prediction based on long-short term memory network and attention mechanism is proposed. The model effect is verified on real data set.

③ This thesis proposes a rebalancing algorithm based on reinforcement learning. The clustering algorithm is used to divide the urban bike-sharing stations into regions. On the premise of meeting the regional rebalancing needs, a bike-sharing rebalancing

algorithm based on reinforcement learning is proposed to assist the bike operation manager to make scientific rebalancing decisions.

④ This thesis designs and implements the prototype system of traffic prediction and rebalancing management of shared bikes. Based on the algorithm proposed in the thesis, the traffic prediction and rebalancing management system of shared bikes is designed. The system can provide detailed overview of users' historical cycling data for bike-sharing operation managers, who can query the traffic prediction results of any station and dispatch bikes according to the given suggestions.

**Key words:** Traffic Prediction; Bike Rebalancing; Neural Network; Reinforcement Learning; Attention Mechanism

## 目 录

中文摘要.....	I
英文摘要.....	II
1 绪 论.....	1
1.1 研究背景及意义.....	1
1.2 国内外研究现状.....	2
1.2.1 流量预测方法研究.....	2
1.2.2 共享单车调度方法研究.....	3
1.2.3 研究现状相关问题分析.....	4
1.3 研究内容.....	5
1.4 论文组织结构.....	6
1.5 本章小结.....	6
2 相关理论技术分析.....	7
2.1 长短期记忆网络概述.....	7
2.2 注意力机制.....	8
2.3 Q-Learning 原理.....	10
2.4 本章小结.....	11
3 基于骑行上下文信息的共享单车流量预测模型研究.....	12
3.1 问题背景.....	12
3.2 基于骑行上下文信息的特征分析.....	13
3.2.1 时间上下文特征分析.....	13
3.2.2 气象上下文特征分析.....	16
3.3 基于骑行上下文信息的共享单车流量预测研究.....	21
3.3.1 共享单车流量预测问题定义.....	21
3.3.2 基于 LSTM 的时序特征提取.....	22
3.3.3 基于注意力机制的特征重要性计算.....	23
3.3.4 基于 LSTM 和注意力机制的共享单车流量预测模型.....	25
3.4 实验设计及分析.....	26
3.4.1 实验数据集简介.....	26
3.4.2 评价指标.....	27
3.4.3 实验结果分析.....	27
3.5 本章小结.....	30

<b>4 基于强化学习的单车调度问题研究</b> .....	31
<b>4.1 问题背景</b> .....	31
<b>4.2 基于 Q-Learning 算法的共享单车调度模型研究</b> .....	32
4.2.1 单车调度问题定义 .....	32
4.2.2 基于转移矩阵和 K-Means 的站点聚类.....	32
4.2.3 融合聚类模型的 Q-Learning 单车调度算法.....	36
<b>4.3 实验设计及分析</b> .....	38
4.3.1 实验数据集 .....	38
4.3.2 评价指标 .....	38
4.3.3 实验结果及分析 .....	38
<b>4.4 本章小结</b> .....	42
<b>5 单车流量预测与调度原型系统的设计与实现</b> .....	43
<b>5.1 应用背景</b> .....	43
<b>5.2 原型系统架构设计</b> .....	43
<b>5.3 系统数据库设计</b> .....	44
<b>5.4 原型系统功能模块设计及分析</b> .....	45
5.4.1 原型系统功能模块设计 .....	45
5.4.2 系统实现及效果分析 .....	46
<b>5.5 本章小结</b> .....	50
<b>6 结论与展望</b> .....	51
<b>6.1 主要结论</b> .....	51
<b>6.2 后续研究工作的展望</b> .....	52
<b>参考文献</b> .....	56
<b>附    录</b> .....	56
A. 作者在攻读硕士学位期间发表的论文目录.....	56
B. 作者在攻读硕士学位期间取得的科研成果目录 .....	56
C. 学位论文数据集 .....	58
<b>致    谢</b> .....	58

## 1 绪 论

### 1.1 研究背景及意义

根据《中国共享出行发展报告（2019）》显示，截止 2019 年 8 月底，共计有共享单车 1950 万辆，注册用户超过 3 亿人次，日均订单数达到 4700 万单。随着单车数量和使用人数的激增，运营商能够收集到大量的用户骑行信息。这些信息涵盖用户骑行时间、起始地点、单次骑行时长以及是否购买骑行卡等，形成了非常有价值的轨迹信息，能够真实地反映了用户的骑行需求。此外，共享单车作为公共交通的一种重要补充形式，同时也作为城市交通的“毛细血管”，能够很好地解决用户出行“最初一公里”，“最后一公里”问题，同时满足城市居民多样化出行需求。

然而，尽管共享单车系统有显著的优势，但要有效预测各站点流量和平衡站点车辆需求极具挑战性。用户的行为是动态变化的，这种变化在一个城市的每个地方不尽相同，这使得共享单车使用情况变得非常的不平衡，导致某些站点无车可用，而另一些站点出现车辆堆积的情况，严重影响了用户的体验，造成了用户的流失。虽然营运公司尝试在城市划定区域，聘请管理人员对区域内各站点的车辆进行调度，在一定程度上缓解该问题。但是由于管理人员只能凭借经验判断出站点在不同时段可能需要调度的车辆数，效率较低，没有办法彻底解决“供不应求”和“供过于求”的不平衡现象，反而影响了共享单车系统的运行效率。特别是在一些城市的边远地带，该区域复合功能低，用户的出行需求趋于一致，出行的时间相对集中，短时间内对车辆的需求量大，并且波动频繁，利用人工调度解决该地区的站点间车辆不平衡问题效率低。对于共享单车运营管理商而言，合理且高效地在各个站点之间重新分配共享单车数目至关重要。唯此，才能让共享单车系统在减少资源和人力损耗的前提下有效运行，进而提高公司自身的经济效益。

共享单车流量预测以及调度问题已经引起了工业界以及学术界的广泛关注，研究者们利用数据挖掘相关技术，从用户历史骑行数据中分析用户行为<sup>[1]</sup>，预测站点流量，辅助管理人员进行合理调度<sup>[2]</sup>，从而助力“智慧城市”建设。众多学者对此进行了研究<sup>[3][4][5][6]</sup>，但预测精度仍然有待提升。本文将针对共享单车流量预测与调度问题展开研究，而该问题目前仍面临两大挑战。首先，要进行调度，就必须确定每个站点的车辆库存水平。然而，这取决于对各站点的单车流量的准确预测，流量预测又受到时间、位置、气象等多种因素的影响。其次，共享单车调度问题是一个大规模的多容量车辆路径问题，由数百个站点组成的共享单车系统规模太大，传统的路线优化技术无法很好解决。



因此,本文详细地对用户的历史共享单车骑行数据进行了分析,探寻用户出行规律,构建并筛选了相关特征,以提高共享单车流量预测精度同时合理调度站点间车辆为目标展开研究。利用长短期记忆网络(Long Short-Term Memory, LSTM)和注意力机制对共享单车流量进行预测,并结合强化学习中的 Q-Learning 算法处理共享单车调度问题,为运营管理人员的调度提供科学决策,提前平衡各站点间的车辆数目,从而最大限度地满足用户的用车需求。利用技术解决现实难题,实现精细化管理的同时构建一个更加完善的共享单车生态系统,提升用户体验感,更好的推动智慧城市的建设与发展。

## 1.2 国内外研究现状

共享单车已经成为了城市居民不可或缺的交通出行工具。庞大的受众产生了数以亿计的行为数据,数据背后隐藏的价值吸引了国内外研究学者的广泛关注与研究<sup>[7][8][9][10]</sup>。

### 1.2.1 流量预测方法研究

#### ① 基于时序建模方法的流量预测

Lihua<sup>[11]</sup>等人根据骑行数据中的非线性以及在不同的时频空间具有不同的特点,利用差分整合移动平均自回归模型(Autoregressive Integrated Moving Average model, ARIMA)良好的线性拟合能力对数据进行处理,从而得到预测结果。但由于 ARIMA 模型仅能考虑时间层面的特征,使得预测能力相对较弱。针对该问题,Zhang<sup>[12]</sup>等人利用短期交通流预测的周期性、非线性、不确定性以及复杂性,提出了一种基于季节性差分自回归移动平均移动(Seasonal Autoregressive Integrated Moving Average model, SARIMA)模型和支持向量机(Support Vector Machine, SVM)模型的混合模型,对时间序列问题进行预测。SARIMA 模型能够发现时间序列数据的相关性,特别适用于季节性、随机时间序列的建模。SVM 对输入和输出数据具有很强的非线性映射能力。作者将两种模型混合,结合各自的优点,该模型相较于传统的 ARIMA 模型,考虑了不同季节对流量预测的影响,预测精度有了较大提升。

#### ② 基于机器学习方法的流量预测

随着机器学习的发展,研究者们逐渐弱化了对时序建模方法的使用。Ahn<sup>[13]</sup>等人提出了一种基于贝叶斯分类器和支持向量回归(Support Vector Regression, SVR)的实时流量预测方法,作者利用 3D 马尔可夫随机场对道路交通流及其关系进行时空建模,将关系较近的区域划分在一起,利用多元线性回归和 SVR 估计区域之间的依赖关系,从而进行流量预测。该方法虽然考虑了区域间的关系,但并未考虑到同一区域中不同道路之间的交通流量的关联性。徐健锋<sup>[14]</sup>等人根据城市

核心区域的交通流量数据的特点，构建了同个区域之中多条相关道路的交通流量的多维度数据模型，提出了一种基于多机器学习竞争策略的交通流量预测算法。作者首先利用时间序列聚类的方式将多维交通流量数据进行降维，然后引入多种机器学习方法进行集成学习，从而得到一个强化后的预测模型。

### ③ 基于深度学习方法的流量预测

传统的机器学习方法普遍集中于对时间和空间这两种涵盖于数据集本身的特征进行建模，在没有其他特征引入的情况下精度较为受限。相比于传统的机器学习方法，深度学习因其能够解决复杂问题的能力而备受研究人员的青睐<sup>[15]</sup>。Lv<sup>[16]</sup>等人既考虑了传统方法所涉及的交通流量预测中的时空特征，同时还利用堆叠的自编码器模型对数据降维，从而完成特征提取工作。最后将自编码器最后一层的输出作为一个回归网络的输入，进行有监督的学习，从而完成对流量的预测。该方法利用深度学习提取数据中存在的特征，相比于传统的机器学习方法，预测精度有所提升，但是与传统的方法一样，他也只考虑到了数据中涵盖的时序特征，没有做额外的特征工程。Zheng<sup>[17]</sup>等人将城市区域划分成 $N \times N$ 大小的网格，利用卷积神经网络（Convolutional Neural Network, CNN）提取网格中的时序和空间特征。此外，作者还考虑了往数据中添加其它的相关特征，如气象特征等，在考虑了时空特性的基础上引入了其他特征，结合深度学习方法更进一步提高了流量预测的精度。

### 1.2.2 共享单车调度方法研究

为了帮助共享单车营运公司提升运行效率、合理分配社会资源、合理布局城市空间等，已有研究学者提出了不同的单车调度方法。他们大多数采用整数线性规划技术<sup>[18][19]</sup>，Liu<sup>[20]</sup>等人基于大规模的历史骑行数据，采用加权 K 近邻回归方法对站点的单车使用需求进行了预测，并且在进一步分析了站点与站点之间所构成的车站网络和行车时间的基础上，采用了一个自适应能力约束的 K-centers 算法来对站点进行聚类，并提出了一种混合整数非线性规划（Mixed-Integer Nonlinear Programming, MINLP）方法解决站点内自行车数量的再平衡问题。但该方法具有较大的局限性，仅适用于规模较小的问题场景或者简化后的问题变体。对此，Petrina<sup>[21]</sup>等人基于贪心随机自适应搜索算法（Greedy Randomized Adaptive Search Procedures, GRASP）与路径重连（Path Relinking, PR）算法，提出了一种混合的调度算法。虽然 PR 的直接应用并没有在解决方案的质量上产生显著的改进，而且还在一定程度上增加了计算成本，但是这种方法能够很好的运用于中型的单车再平衡问题场景，在真实数据集上的测试也说明了这一点。随后，Lahoorpoor<sup>[22]</sup>提出了一种自底向上的基于集群的自行车系统静态再平衡模型。首先，研究了单车出行的时空模式；其次，利用层次聚类方法，定义了基于站点间形成的相似度来发

现相同类型的站点；最后，定义了单车再平衡问题的两个层次，即集群内部和集群之间。利用单一目标遗传算法，将共享单车调度问题建模为优化问题，其目标是最小化调度时间。提出的模型在真实的数据集进行了验证，分析表明，该模型能够将调度周期缩短 30%。除了从共享单车本身出发，也有研究学者从用户角度去对单车调度问题进行优化。Li<sup>[23]</sup>指出不应该对整个城市的所有自行车站点之间进行调配，这会浪费大量的人财物力。所以他们首先通过聚类算法对站点进行了分组，分组后的站点具备簇内平衡、簇间独立的特点。针对每个簇设计了一个时空强化学习模型，目标旨在最小化长时间内的客户流失。并且，为了对每个模型进行学习，作者还设计了一个深度神经网络来估计其最优长期价值函数，从而得到最优策略，最后在真实的数据集上验证了该模型的有效性。Pan<sup>[24]</sup>等人提出了一个深度强化学习框架来激励用户重新平衡共享单车系统，他们将问题建模为一个马尔可夫决策过程，并考虑了时间和空间特征，提出了层次强化定价模型（Hierarchical Reinforcement Pricing, HRP），与通常忽略空间信息并且严重依赖于精准预测方法不同的是，HRP 采用分而治之的思想，并且嵌入本地化模块捕捉空间和时间相关性。真实数据集上的结果表明，HRP 的性能在共享单车再分配和服务水平都取得了较佳的效果。

### 1.2.3 研究现状相关问题分析

国内外研究学者在单车需求预测以及共享单车调度问题上不断的进行探索，寻求新的解决方案，使得共享单车流量预测精度不断提升，共享单车调度问题逐渐优化，但仍然存在如下问题：

① 目前常用的方法所考虑的仍然是数据当中本身包含的时间与空间特性，利用机器学习或者深度学习的方法去挖掘其中的特征，或者增加额外的天气特征辅助进行流量预测。但没有对数据集进行全面细致地分析，未对特征进行筛选，存在特征冗余等问题，导致对最终的预测结果产生影响；

② 用户的出行规律、骑行喜好等会随着时间的变化而变化，机器学习方法或者神经网络模型能够捕捉到时序特征，但往往缺乏动态提取特征的能力。而注意力机制能够帮助本文更有重点的把握每个时刻所对应的特征因子。因此，在较为充分的特征工程的基础上引入注意力机制，能够构建动态的时序模型，从历史数据中捕捉到对预测结果来说更为重要的特征，而选择性的忽略不重要的特征；

③ 针对共享单车调度问题，大多数研究依据全局特征对所有站点进行调度研究，而共享单车的调度可仅在部分站点之间进行。因此，可以考虑对站点进行聚类。此外，当前的调度行为会对未来产生影响，而这种影响可以反过来限制调度行为，但目前的大多数研究并未考虑到这一点。

在共享单车流量预测以及调度问题中，随着用户的规模以及产生数据量的不断增长，对预测精度以及调度优化要求更高的同时也有着更大的研究价值，本文将针对上述问题展开进一步的研究。

### 1.3 研究内容

影响单车流量预测以及调度问题的因素复杂、多样<sup>[25][26]</sup>。因此，本文着重于对用户的骑行历史数据进行分析，找出影响共享单车流量预测的主要特征，进而提出一种共享单车流量预测算法，为共享单车调度问题提供依据和支撑。本文通过对现有流量预测和单车调度问题相关文献的研究分析，发现了此领域中还存在的问题和可以改进的方向，针对这些问题，本文提出了一种基于 LSTM 和注意力机制的流量预测模型，并展开了基于强化学习的单车调度问题研究。本文的主要研究工作如下：

① 调研。对国内外关于单车需求预测以及调度问题的研究现状进行了调查，归纳并且总结了现有的共享单车流量预测以及解决共享单车调度问题的方法，发现了目前研究中存在的不足之处，提出了可行的解决方案，为后续的科研工作做出了准备。

② 上下文信息分析。对用户的骑行历史数据进行了详尽地分析，探索了用户的骑行规律，并加入了气象数据，分析了天气对于用户骑行的影响，找出了影响用户行为的主要特征，并进行了特征筛选，去除了冗余特征。

③ 提出了一种结合骑行上下文信息，并基于 LSTM 网络和注意力机制的共享单车流量预测算法。根据前期研究，将数据集进行了划分，提取了表示周期性、趋势性、以及邻近性特征的数据，并引入上下文信息对单车的流量进行了预测，获得了最终的预测结果。

④ 提出了基于强化学习的共享单车调度策略。利用所获取的共享单车流量预测数据，结合强化学习的方法对共享单车问题进行了研究，在真实的数据集上检验了本文所提出的算法的有效性。

⑤ 共享单车流量预测及调度管理原型系统的设计与实现。根据论文中提出的流量预测算法以及共享单车调度算法，构建了共享单车流量预测及调度管理原型系统。

本文旨在通过对共享单车流量进行预测，以及对单车调度问题研究，缓解目前实际共享单车应用场景中存在的“潮汐”问题，以提前预知、合理规划等方式为单车运营管理人员提供科学的调度决策依据，在保证共享单车系统有效运行的前提下，减少资源和人力损耗，提高公司自身的经济效益，助力“智慧城市”建设。

## 1.4 论文组织结构

本文在组织结构上包括六个章节。

第一章：绪论。主要介绍了共享单车流量预测以及调度问题的研究背景、国内外研究现状，并介绍了本文的研究内容以及组织结构。

第二章：相关理论技术分析。该章节介绍了与本文研究密切相关的技术，主要阐述了 LSTM、注意力机制、强化学习的算法原理，为本文模型提供了相应的理论基础和支撑，

第三章：共享单车流量预测模型。该章节针对用户的骑行历史数据进行了分析，引入了气象特征，构建新特征的同时也进行了特征筛选。在第三章提出一种基于 LSTM 和注意力机制的单车流量预测模型，并对算法框架、流程以及实验结果进行了详尽的分析和阐述。

第四章：共享单车调度问题研究。该章节针对单车的调度问题进行了研究，基于强化学习提出了一种共享单车调度算法，对算法的框架以及实验结果进行了详尽的阐述和分析。

第五章：共享单车流量预测及调度原型系统设计与实现。根据第三章所提出的流量预测算法，以及第四章所提出的调度策略，设计并且实现了流量预测及调度管理原型系统，为单车运营管理人员的科学决策提供依据。

第六章：总结与展望。该章节旨在对本文所研究的内容进行总结，并且对于未来可以改进的研究进行了展望。

## 1.5 本章小结

本章对所研究的共享单车流量预测以及调度问题的相关内容进行了充分调研，阐述了研究背景以及国内外研究现状，分析了目前研究中存在的不足。最后，对本文的研究内容和组织结构进行了说明。

## 2 相关理论技术分析

### 2.1 长短期记忆网络概述

本文所研究的共享单车流量预测问题，其本质是时间序列问题。循环神经网络能够挖掘序列之间的时序关系，从而很好地处理这类问题。循环神经网络源起于 Hopfield<sup>[27]</sup>网络，该网络内部有反馈连接，能够处理信号中的时间依赖。在此基础上，Jordan<sup>[28]</sup>首次在神经网络中引入了循环链接。后来，Elman<sup>[29]</sup>又在 Jordan 的研究基础上，正式提出了 RNN 模型。但是，普通的 RNN 模型在某些重要信息和输出位置之间的距离逐渐增大时，无法学习到序列的长期依赖关系。同时，普通 RNN 的循环层比较难以训练，因为在反向过程中梯度沿时间方向传递，经过权重矩阵的连乘后，得到的结果容易趋近于零或者溢出，造成“梯度爆炸”或者“梯度消失”现象。研究人员提出了长短期记忆网络<sup>[30]</sup>来解决这一问题。

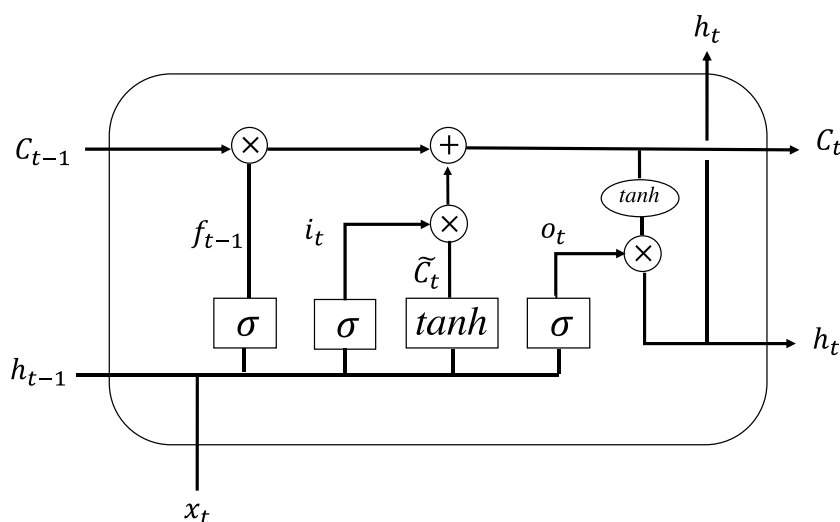


图 2.1 单个 LSTM 单元示意图

Fig.2.1 Schematic diagram of LSTM

与原始 RNN 的主要不同在于，LSTM 使用了一种结构更为复杂的循环层，单个 LSTM 单元如图 2.1 所示，每个单元被称作细胞，LSTM 的核心思想是各细胞之间的状态传递，并且通过“门”来实现信息的过滤，分别有“遗忘门”、“输入门”和“输出门”。在 LSTM 网络中，第一步由“遗忘门”决定从细胞状态中丢弃什么信息。该门会读取到输入  $h_{t-1}$  和  $x_t$ ，输出一个  $[0,1]$  之间的数值给每个细胞状态，1 表示对信息“完全保留”，“0”表示对信息“完全舍弃”，公式如 2.1 所示。

$$f_t = \sigma(W_f \cdot [h_{t-1}; x_t] + b_f) \quad (2.1)$$

其中,  $h_{t-1}$ 表示的是上一个细胞的输出,即上一个时刻对应的隐藏层状态,  $x_t$ 表示的是当前细胞的输入,  $\sigma$ 是 *sigmoid* 激活函数。

下一步是决定让多少新的信息加入到细胞状态中来, 首先利用“输入门”的 *sigmoid* 层决定哪些信息需要更新; 其次利用 *tanh* 层生成一个向量  $\tilde{C}_t$ , 它作为备选的更新内容, 将这两部分结合起来, 通过公式 2.2 和公式 2.3 对细胞状态进行更新。

$$i_t = \sigma(W_i \cdot [h_{t-1}; x_t] + b_i) \quad (2.2)$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}; x_t] + b_C) \quad (2.3)$$

最终, 利用“输出门”确定网络的输出值。最终的输出值将基于过滤后细胞的状态。利用 *sigmoid* 层确定细胞状态的输出部分, 接着将细胞状态通过 *tanh* 函数进行处理, 得到一个数值, 将其与 *sigmoid* 门的输出相乘, 得到最终的输出结果, 如公式 2.4 和 2.5 所示。

$$o_t = \sigma(W_o \cdot [h_{t-1}; x_t] + b_o) \quad (2.4)$$

$$h_t = o_t * \tanh(C_t) \quad (2.5)$$

由此可见, LSTM 网络的输出值  $h_t$ , 是受到前面历次输入值  $x_t, x_{t-1} \dots$  的影响。而本文所要解决的共享单车流量预测问题中, 用户下一个时刻的骑行行为总是受到过去诸多时刻的影响, 这种影响之间有很强的关联性, 使用 LSTM 可以对这些联系进行挖掘。输出值  $h_t$  既包含当前时间节点所蕴含的信息, 也包含了之前节点的部分重要信息。由于用户前后行为的关联性, 在预测后一个时刻的流量时, 需要前面诸多时刻的信息作为预测的特征, 而输出值  $h_t$  能够有效地保留前面时刻用户行为对后来预测结果的影响。综上, 本文选择 LSTM 作为单车流量预测算法的基础模型。

## 2.2 注意力机制

注意力 (Attention) 机制<sup>[31]</sup>最早应用于计算机视觉<sup>[32]</sup>, 随后在自然语言处理领域也开始运用, 并且获得了出众的效果。LSTM 的每一步计算结果依赖于上一步的计算结果, 因此无法并行计算。而注意力机制并不存在结果依赖, 因此能够和卷积神经网络一样进行并行处理, 保留了 LSTM 对以前信息的“记忆力”, 同时解决了其不能并行计算的问题。在注意力机制引入之前, LSTM 虽然能够进行“记忆”长距离信息, 但没有办法捕捉到与当前输出最为相关的特征信息。在自然处理领域, 对较长文本进行处理的时候, 却往往需要把握上下文之间的关系进行文段理解, 此时, 注意力机制能够很好的抓住文本或者其它信息中的重点内容, 从而获取到更多的有效信息。注意力机制的本质是根据事物之间的关系进行线性加权而求得新的表示。

注意力机制的本质如图 2.2 所示,它是给定一个和任务有关的查询,即 Query,通过计算与 Key 的注意力分布,并且附加在 Value 上,从而计算 Attention Value。这个过程体现了注意力机制在缓解神经网络模型复杂度方面的作用,由于不需要将所有信息全部输入到神经网络进行计算,只需要从信息中选取和任务相关的信息传入到神经网络进行计算即可。注意力机制的计算分为以下三个步骤:

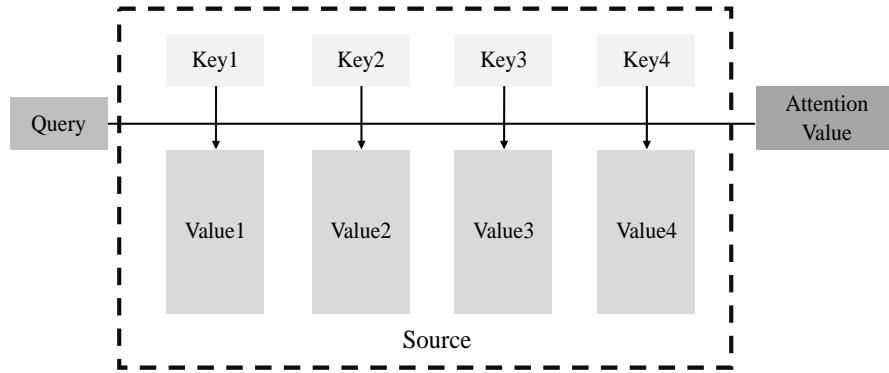


图 2.2 注意力机制实质

Fig.2.2 Essence of attentional mechanism

- ① 信息输入: 用  $X = [x_1, x_2 \dots x_n]$  表示  $n$  个输入信息
- ② 根据公式 2.6 计算注意力分布:

$$\alpha_i = \text{softmax}(s(\text{key}_i, q)) = \text{softmax}(s(x_i, q)) \quad (2.6)$$

其中  $\alpha_i$  是注意力分布,或称为概率分布,  $s(x_i, q)$  是注意力打分机制,常用的打分机制有加性模型、点积模型、微放点积模型以及双线性模型,本文在第三章采用了加性模型,公式如 2.7 所示。

$$s(x_i, q) = V^T \tanh(Wx_i + Uq) \quad (2.7)$$

- ③ 信息加权平均

注意力分布  $\alpha_i$  可以理解为在上下文查询  $q$  时,第  $i$  个信息所受到的关注程度,通过公式 2.8 加权求得最终的注意力权值。

$$\text{Attn}(q, X) = \sum_{i=1}^N \alpha_i X_i \quad (2.8)$$

在本章 2.1 小节提到, LSTM 能够有效地保留前面时刻用户行为对后来预测结果的影响。但过去时刻对未来所产生的影响在每个时刻的重要程度是不一致的,而注意力机制引入到 LSTM 可以帮助模型在保留时间序列间相互联系的情况下,更有针对性地捕捉到对预测时刻而言更为重要的那部分特征信息。



## 2.3 Q-Learning 原理

强化学习<sup>[33][34]</sup>旨在让什么都不知道计算机，通过不断地尝试，从错误中学习，最后找到规律，从而学会了达到特定目的的方法。强化学习中也包含了很多种类的算法，比如有通过行为的价值选取特定行为的方法，包括使用 Q-Learning<sup>[35]</sup>，Sarsa<sup>[36]</sup>，使用神经网络学习的 Deep Q Network<sup>[37]</sup>，还有直接输出行为的 policy gradients<sup>[38]</sup>等。本论文在后续研究中选择了 Q-Learning 方法，因为 Q-Learning 算法实现简单，并且能够通过对价值函数进行计算获得未来时刻最优的调度策略，直接可得到在何种状态 $s$ 下采取何种动作 $a$ 对当下的应用场景来说是最优的。表 2.1 给出了 Q-Learning 算法的伪代码。

表 2.1 Q-Learning 算法伪代码

Table 2.1 Q-Learning algorithm pseudo-code

输入：状态 $s$ ，动作 $a$ ，奖励 $r$
输出：动作价值函数（从动作价值函数可以获得策略和相应的动作）
1、初始化 $Q(s, a), \forall s \in S, a \in A(s)$ ，给定参数 $\alpha, \gamma$
2、Repeat:
给定状态 $s$ ，并根据 $\epsilon$ 贪婪策略在状态 $s$ 选择动作 $a$
Repeat:
(a) 根据 $\epsilon$ 贪婪策略在状态 $s_t$ 选择动作 $a_{t+1}$ ，得到回报 $r_t$ 和下一个状态 $s_{t+1}$
(b) 依据公式 2.9 更新 $Q(s, a)$
(c) $s = s', a = a'$
直到 $s$ 是终止状态，结束
直到所有的 $Q(s, a)$ 收敛，结束
3、输出最终的策略： $\pi(s) = \arg \max_a Q(s, a)$

Q-Learning 是基于值函数的强化学习方法，其本质是更新值函数，以寻找最优的动作-选择策略对。Q 就是在某一时刻的 $s$ 状态下，采取动作 $a$ 能够获得收益的期望，环境会根据 agent 反馈相应的即时奖励。在 Q-Learning 中，最重要的就是“状态”和“动作”。Q-Learning 算法的主要思想就是将状态和动作构建一张 Q 表来存储期望值，然后根据期望值来选取能够获得最大收益的动作。

具体步骤为：第一，定义一个奖励矩阵  $\mathbf{R}$ ，表示在状态  $s$  下执行动作  $a$  会得到奖励  $r$ 。第二，创建一个  $Q$  表，表示学习到的经验，与奖励矩阵  $\mathbf{R}$  同阶，它可以是一个初始值全部为 0 的矩阵。公式 2.9 为值函数更新公式。

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha(r_t + \gamma \max_a Q(s_{t+1}, a_t) - Q(s_t, a_t)) \quad (2.9)$$

$s$  表示当前的状态， $a$  表示当前的动作， $\tilde{s}$  表示下一个状态， $\gamma \in (0,1)$  为贪婪因子，若设置为 0.1，则表示更着重于考虑短期或者下一个动作的即时收益，而忽略若干个动作之后的收益；若取 0.9，则表示更着眼于长远的收益。 $\tilde{a}$  表示下一个动作， $Q$  表示在状态  $s$  和动作  $a$  下所能够获取到的期望的最大收益。 $R$  是立即得到的收益，而未来的收益则取决于下一阶段的动作。在迭代收敛之后，便可以根据  $Q$  表来选择达到目标的最优方式，也就是对于每个状态来说的最佳动作。

对于给定的策略  $\pi$ ，定义其在  $s_0$  时的价值函数为公式 2.10，价值函数能够衡量一个状态在未来的潜在价值，通俗地讲，就是能知道每一种状态的好坏。通过对价值函数的引入，对于获取最优策略的问题，就可以通过估计价值函数来间接地获取优化的策略。

$$\begin{aligned} V^\pi(s) &= E(R(s_0) + \gamma R(s_1) + \gamma^2 R(s_2) + \cdots + \gamma^n R(s_n) | \pi, s = s_0) \\ &= R(s_0) + \gamma E(R(s_1) + \gamma R(s_2) + \cdots | \pi) \end{aligned} \quad (2.10)$$

其中， $s_1, s_0 \dots$  为随机变量。由  $(R(s_1) + \gamma R(s_2) + \cdots | \pi) = \sum_{s_1} P_{s_0\pi(s_0)}(s_1) V^\pi(s_1)$  可以得到递推公式 2.11，也称之为贝尔曼方程<sup>[39]</sup>(Bellman equation)。

$$V^\pi(s) = R(s_0) + \gamma \sum_{s'} P_{s_0\pi(s_0)}(s') V^\pi(s') \quad (2.11)$$

Q-Learning 算法的选用，可以更加高效地对单车进行调度，通过计算出未来每个时刻对应的最大的价值函数，获取最优的调度方式，即当下对站点进行多少辆自行车的移动能够满足该站点未来一段时间用户对自行车的需求。本文在对城市内所有自行车站点进行调度的时候划分了区域，在区域内进行自行车调度，每个区域中的站点数目适中，降低了 Q-Learning 算法的复杂性。

## 2.4 本章小结

本章对论文中所用到的相关技术做了较为详尽的描述，主要介绍了 LSTM、注意力机制以及强化学习当中的 Q-Learning 算法，并描述了相关技术与本文所提算法之间的联系，通过分析为模型的构建提供了理论上的基础与支撑。

### 3 基于骑行上下文信息的共享单车流量预测模型研究

#### 3.1 问题背景

共享单车流量预测问题是利用用户的历史骑行数据,对未来一段时间内各个单车站点的流量进行预测。由于深度学习的成熟发展及其良好的特征提取能力和鲁棒性,研究学者们普遍使用深度学习方法进行共享单车流量预测问题研究<sup>[16][17]</sup>。部分研究学者利用堆叠的自编码器模型对数据降维,完成特征提取工作。该方法利用深度学习去找到用户骑行历史数据中存在的特征,相比于传统的机器学习方法,预测精度有所提升。但是,与传统的方法一样,它也只考虑到了数据中原有的时序特征。然而,共享单车用户的骑行行为除了与时间相关之外,与气象特征也是密切相关的。后来,有研究学者利用卷积神经网络提取网格中的时序和空间特征,同时对任何两个区域间存在的时间依赖性进行建模。并且,引入了额外的气象特征,与时空特征合并,预测每个区域的最终流量。此外,也有学者提出可将交通的关联模式抽象成图结构。利用图的顶点表示地图上的点,边则表示两个地点之间的关联。建模时空相关性对流量预测的影响,利用模型学习点和边的特性,从而完成流量预测。

但是,上述方法仍然存在如下问题:

① 在考虑特征的时候,大多数研究仅利用数据本身包含的时间特性,但对于所研究问题而言,除了单纯考虑时间特征,还应该引入与骑行数据相关的上下文特征。并且,不同的群体对出行需求是不同的,这需要根据数据集反映出的用户特性添加相应的特征。此外,特征筛选是有必要的,若存在冗余特征,会降低最终的流量预测精度;

② 用户的骑行偏好会随着时间变化。虽然机器学习方法或者神经网络模型能够捕捉到时序特征,但往往缺乏动态提取特征的能力。因此,可以考虑如何能够构建动态的时序模型,从历史数据中捕捉到对预测结果来说更为重要的特征,而选择性的忽略不重要的特征。

针对上述问题,本章从对用户骑行历史数据的详细分析入手,进行特征的构建与筛选,利用长短期记忆网络能够较好捕捉时间序列数据之间的依赖关系,结合注意力机制构建动态的流量预测模型。本章节所考虑的流量预测问题是针对有桩的共享单车,如图 3.1 所示,每辆自行车对应一个停车位,每个站点的停车位数量由运营公司预先规定。



图 3.1 单车站点示意图

Fig.3.1 Schematic diagram of station

## 3.2 基于骑行上下文信息的特征分析

### 3.2.1 时间上下文特征分析

用户的骑行历史数据其本质是时间序列数据，随着时间的变化用户的行为也会发生改变，为了探寻用户行为随着时间的变化规律。本小节从时间上下文入手，对花旗自行车（Citibike）用户历史骑行数据进行了分析。

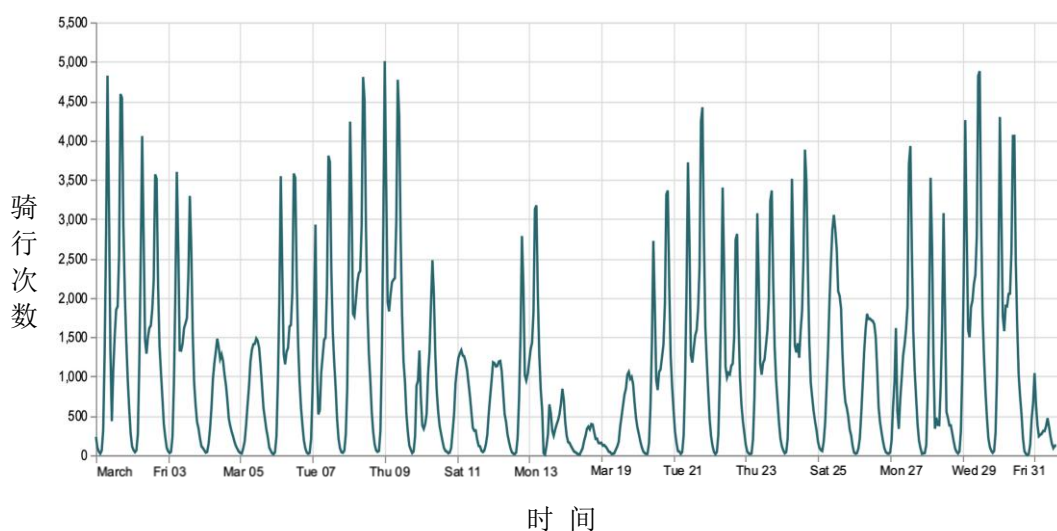


图 3.2 骑行次数随时间变化趋势图

Fig.3.2 The trend chart of cycling times with time

以 2017 年 3 月份的花旗自行车数据为例，此时间段的数据共有 15 个特征，总共包括 727665 条数据，包含 619 个单车站点，其中用户出生年份存在缺失，缺失 37982 条，缺失率 0.3%。对于缺失的年份数据，本文用所有用户的平均出生年份进行了填充。在数据集原有的 15 个特征的基础上，本文提取了新的特征，作为数据分析的依据。新特征有：骑行开始以及骑行结束所处的时间段（以 1 小时划分）。例如，2017 年 3 月 1 日 00 时 00 分 32 秒，所对应的新的时间段特征为“0”；骑行日期是工作日还是周末，工作日对应特征“0”，周末对应特征“1”。

本文以小时为单位对 2017 年 3 月份的骑行数据进行了统计，统计结果如图 3.2 所示，通过统计结果可以发现数据的波动呈现一定的规律。观察连续时间内折线的走势，呈现上升或者下降两种情况，在本文的论述中，本文称其为趋势性；观察每段时间区域折线的走势，可以看到某个时间区域内的走势和其他一个或者多个时间区域内的折线走势类似，在本文的论述中，本文称其为周期性，而这里的周期性本文定义为以周为单位；除此之外，以一天为单位所体现出来的周期性，本文称之为邻近性。

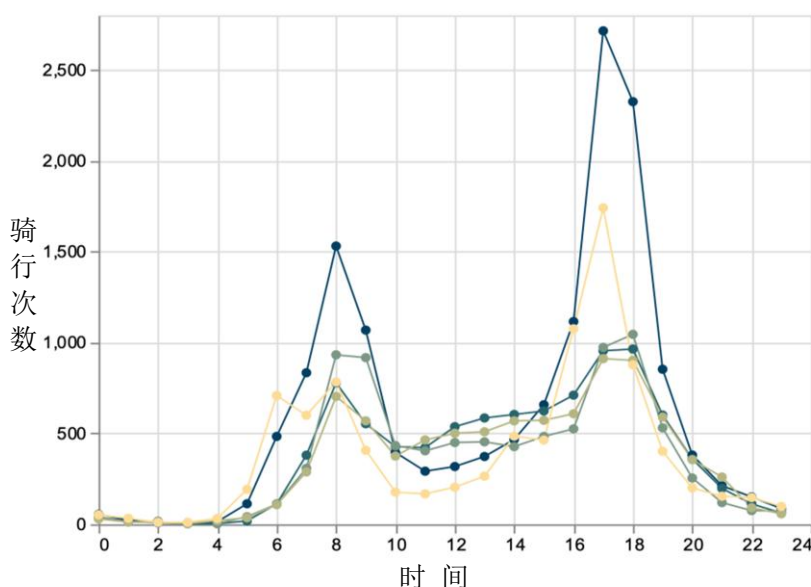


图 3.3 骑行次数前 5 的站点一天中流量统计

Fig.3.3 The traffic statistics of the top 5 stations in one day

为了更加深入的探测时间因素对流量预测的影响，本文对 2017 年 3 月份的数据以小时为单位进行了聚合，记录 3 月份用户总的骑行次数，如图 3.3 所示。本文选取了骑行热度前 5 的站点进行展示，从图 3.3 中可以看出，出现了两个明显的峰值，分别是早晨 7 点至 9 点以及傍晚 17 点至 18 点，本文也将其称之为早高峰和晚高峰，而这正好对应了大多数用户的上下班时间。所以，本文推断共享单车骑

行的主要用户可能为通勤一族。通勤一族的出行情况有着较为明显的规律，利用这一点，本文可以更好地对特征进行构建。

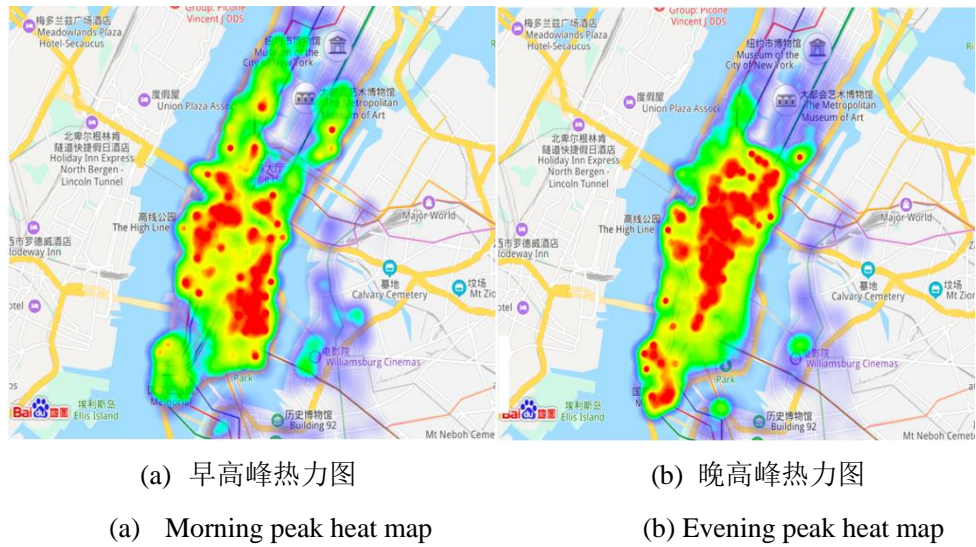


图 3.4 热力图

Fig.3.4 The heat map

为了更加直观地对此分析进行验证，本文利用百度地图 API 进行了骑行热力图分析，图 3.4(a)为早高峰时段热力图，图 3.4(b)为晚高峰时段热力图，将骑行频次转换到空间后发现通勤一族的办公地点大多在市中心，而居住地往往在城市周边。通过这些分析本文得出结论，时刻是关于单车流量预测的重要影响因素，在不同的时刻共享单车使用人群有着不同的出行规律和用车需求，但这些规律和需

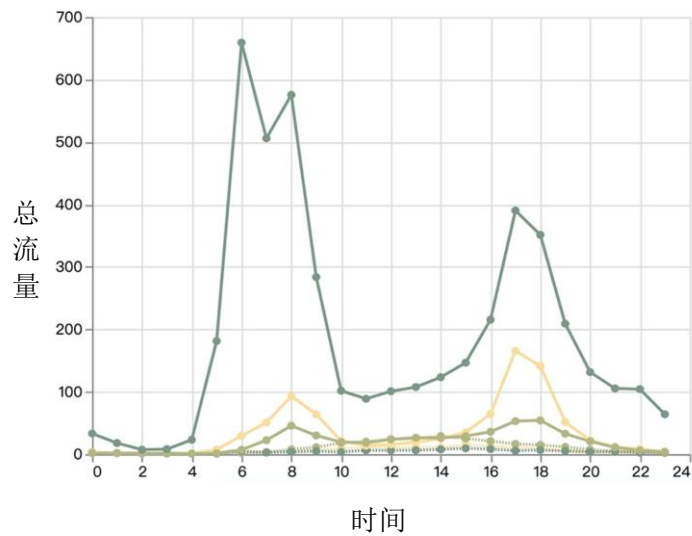


图 3.5 工作日(实线)及周末(虚线)站点流量统计

Fig.3.5 Station traffic statistics on weekdays and weekends



求都是有迹可循的，将这一点作为特征，能够让模型更好地对用户骑行习惯进行学习，从而提高流量预测精度。

本文已经从图 3.3 和图 3.4 分析出了单车的使用存在早晚高峰，也知道使用人群以上班通勤者居多，对通勤族进行分析就需要考虑该群体周末与非周末的需求差异。对此，本文基于周末与非周末对数据进行了分析，如图 3.5 所示。

本文从总流量排名前五的站点中随机选择了三个对其进行分析。从图 3.5 可以看出，各个站点周末与非周末的骑行数据存在较大差异。此数据也切合实际生活经验，通勤族总是在周一至周五上班，所以对共享单车的使用需求在周一至周五会出现早晚高峰现象，而通常在周末休息，所以周末不存在明显的峰值。通过上述分析可以发现，在后续对单车流量进行预测的时候，可以将周末以及非周末因素作为新特征进行构建，辅助模型学习用户骑行习惯，进一步提高预测精度。

### 3.2.2 气象上下文特征分析

除了时间的因素外，气象因素对共享单车流量预测也起着至关重要的作用。依据实际经验，在天气晴朗，或者多云微风的日子，用户才会更多地考虑使用共享单车出行，而遇到暴雨或者下雪等恶劣天气，没有良好的骑行条件，骑行需求会大幅度减少。接下来，本文将针对气象特征对共享单车流量的影响展开分析。

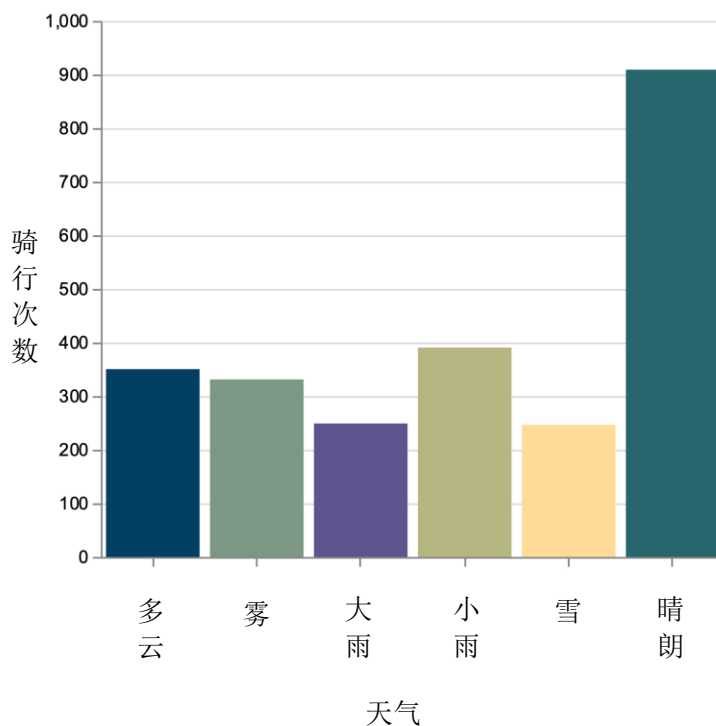


图 3.6 天气对骑行次数的影响

Fig.3.6 The influence of weather on the number of rides

首先是分析天气对单车流量的影响。在原本的数据集当中存在 16 种天气类型，本文按照表 3.1 的方式对其进行了映射。最终得到六个天气大类。天气对骑行需求的影响如图 3.6 所示，从图可以看出，不同的天气对单车流量会带来较大的影响。比如，在大到暴雨以及雪天，用车需求很小，而多云和晴朗的天气让人们更愿意骑行，分析结果也印证了本小节的猜想。

表 3.1 天气映射对应表

Table 3.1 The weather map corresponds to the table	
映射后天气类别	映射前天气类别
雾(foggy)	薄雾/雾/霾/
小雨(light rain)	小雨/中雨/光强度小雨
大雨(heavy rain)	冻雨/强降雨/雷暴雨/带小雨雷暴
雪(snow)	中雪/小雪
晴朗(sunny)	晴朗
多云(cloudy)	多云/少云/阴

此外，如图 3.7 所示，风速对骑行需求也有着不同程度的影响。从图中可以明显的看出，风速在 1 至 4 级左右时最适合骑行，此后随着风速的增加，骑行需求有着不同程度的减少。可见风速对骑行需求有着比较显著的影响，能够作为共享单车流量预测的特征。除了风速之外，如图 3.8 所示，压强对骑行需求有着不同程度的影响，并且在压强 1014kpa 和 1015kpa 时影响区别较大，即使此时压强在数值上仅相差 1kpa。

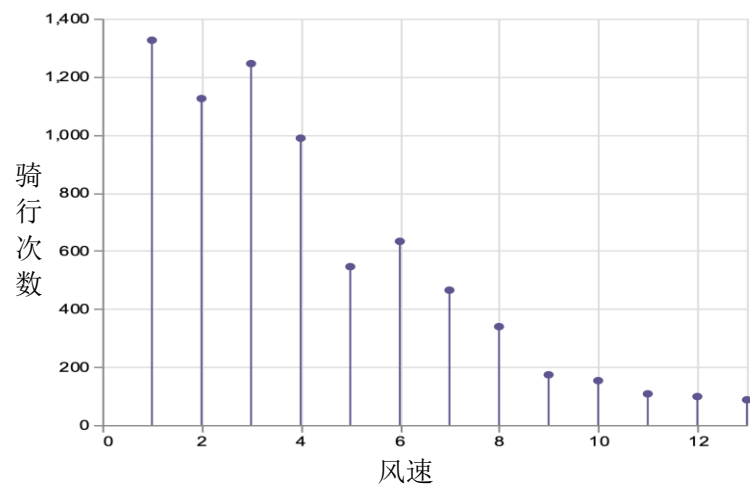


图 3.7 风速对骑行次数的影响

Fig.3.7 The influence of wind speed on the number of rides



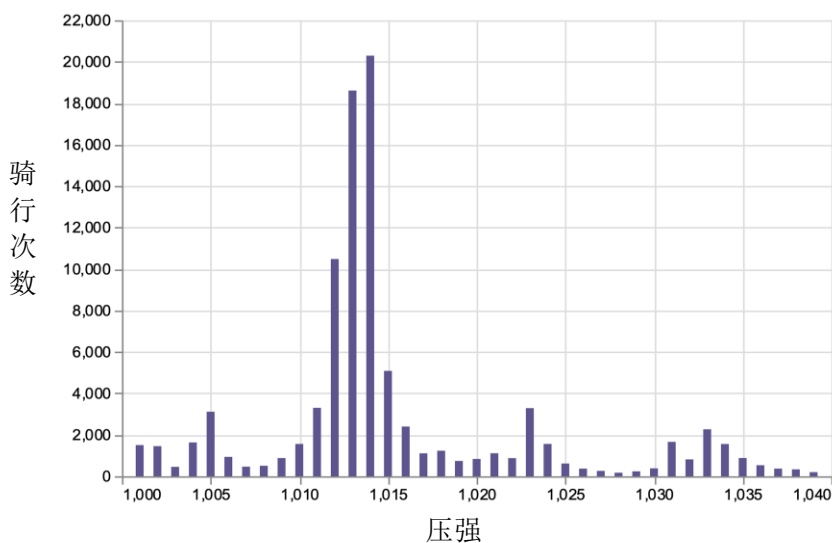


图 3.8 压强对骑行次数的影响

Fig.3.8 The influence of pressure on the number of rides

再者，温度和湿度也在一定程度上对骑行次数产生影响。但是，从图 3.9 和图 3.10 可以看出，温度和湿度对骑行次数的影响都不具备一定的规律性。根据实际生活经验，温度和湿度的确会对骑行需求产生影响。湿度过高会让人烦闷不适，影响骑行；温度过高，天气炎热，温度过低，天气寒冷，都不适于骑行。可是，图 3.9 和图 3.10 的数据分析却没有表现出过高过低的温湿度和适宜温湿度之间的差异。此时可以从特征是否出现冗余进行考虑，湿度和温度可能是受到天气的直接作用影响了骑行需求。于是本文通过两组特征对骑行数据的影响进行分析，分别是：天气与湿度，天气与温度。

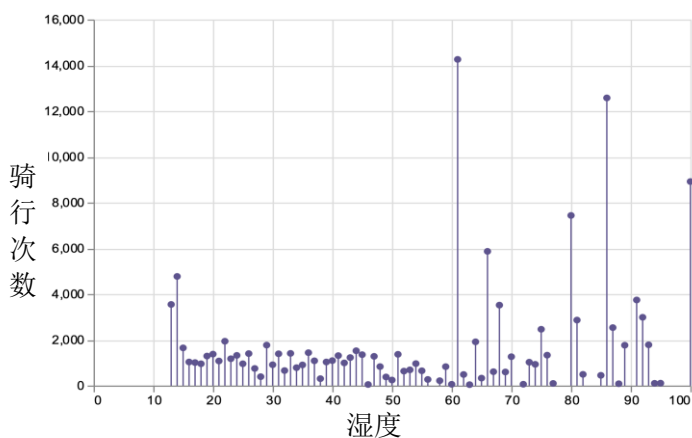


图 3.9 湿度对骑行次数的影响

Fig.3.9 The influence of humidity on the number of rides

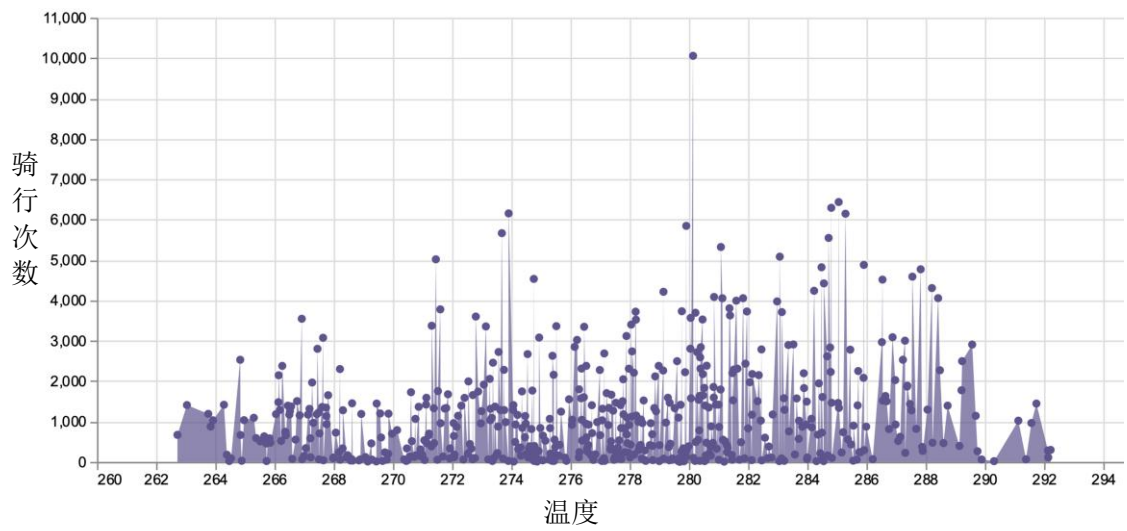


图 3.10 温度对骑行次数的影响

Fig.3.10 The influence temperature of on the number of rides

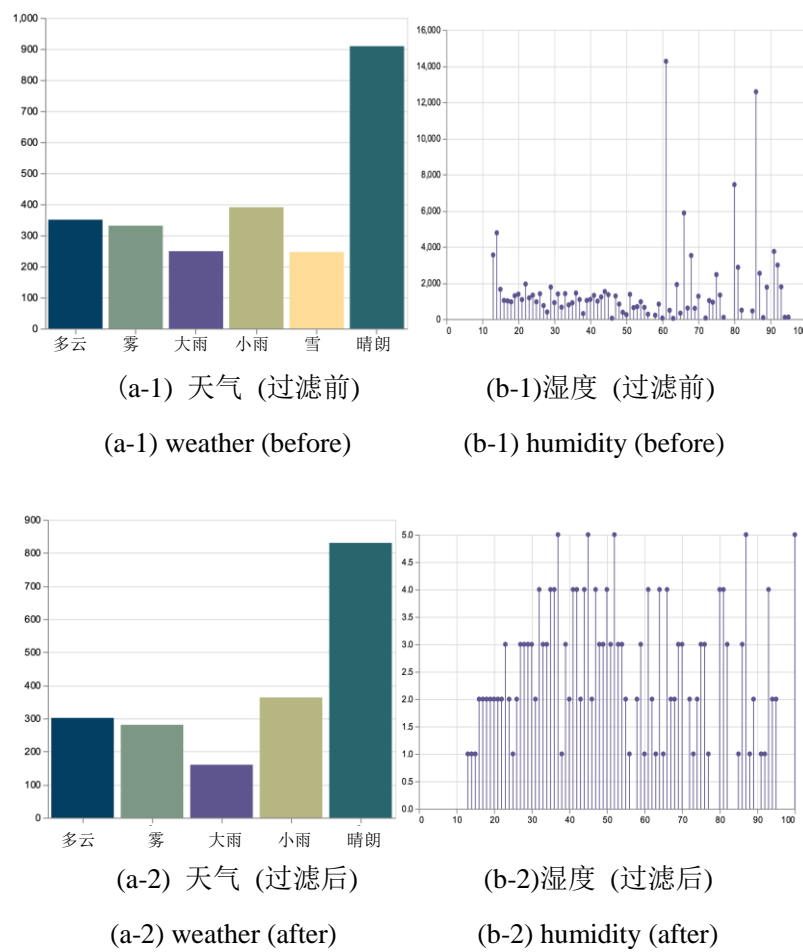


图 3.11 天气与湿度共同对骑行需求的影响

Fig.3.11 The influence of weather and humidity on the number of rides

图 3.11 是天气与湿度共同对骑行需求的影响。在获取到初步的结果后，分析得到其平均骑行次数为 3602 次，为此本文以平均值为基准对结果进行了过滤。得到了最终的结果，如图 3.11(a-1)和图 3.11(a-2)所示，也就是在过滤后天气和湿度分别对骑行需求的影响。本文发现，进行过滤之后，“雪”这一类型的天气消失了，这也侧面说明了天气对骑行需求的影响，与雪天难行的生活经验相符合。而从图 3.11(b-1)和图 3.11(b-2)可知，在过滤之后湿度并未对骑行需求产生明显有规律性的影响，所以本文判断湿度对骑行需求的影响，其实是由天气作用的。例如，暴雨会导致湿度较大，而从日常生活经验和数据分析的结果来看，显然暴雨对骑行的影响会比湿度对骑行的影响大。所以，基于上述分析，本文不将湿度作为单车流量预测的特征。

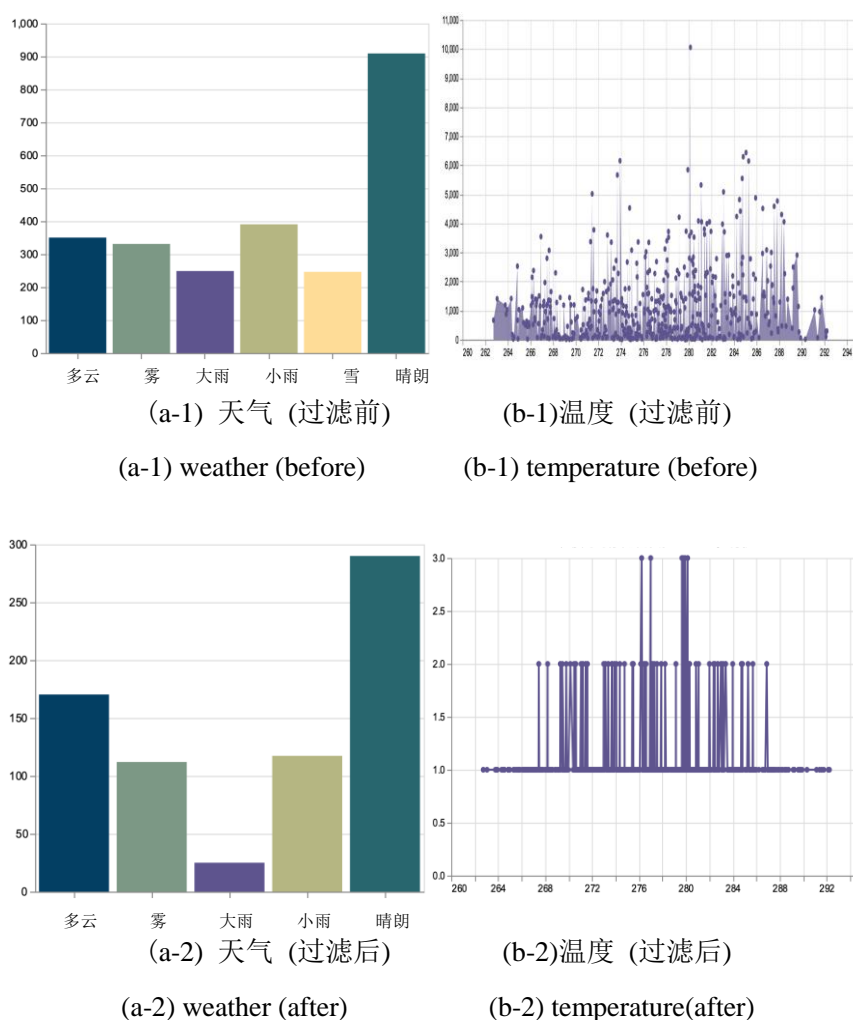


图 3.12 天气与温度共同对骑行需求的影响

Fig.3.12 The influence of weather and temperature on the number of rides

图 3.12 是天气与温度共同对骑行需求的影响。在获取到初步的结果后, 分析得到其平均骑行次数为 1173 次, 为此, 本文以平均值为基准对结果进行了过滤。得到了最终的结果, 如图 3.12(a-1)和图 3.12(a-2)所示。同样可以看到, “雪”这一类天气特征也消失了, 更进一步佐证了天气对骑行需求的显著影响。而且不管是天气单独对骑行需求的影响, 还是与温度共同作用下对骑行需求的影响, 天气对骑行需求的影响的总体趋势是大致相同的。而对于温度这一特征, 从图 3.12(b-1)和 3.12(b-2)可以看出, 无论过滤前后, 对骑行次数的影响都是无序的。再者, 温度的变化可能是由于天气引起的, 雪天导致温度低, 炎热的夏天温度会节节攀升, 从而影响骑行需求。由此, 本文认为, 温度也不适合作为单车流量预测的特征。

综上所述, 本文将时刻、周末/非周末、趋势性、周期性、邻近性、天气、风速、压强等上下文信息作为特征, 去掉温度和湿度两个冗余特征, 利用 LSTM 和注意力机制预测共享单车各站点的流量。

### 3.3 基于骑行上下文信息的共享单车流量预测研究

#### 3.3.1 共享单车流量预测问题定义

对于任意一个站点, 本文将 $x_t$ 称之为 $t$ 时刻所对应的观测值, 而流量预测问题就是根据历史观测序列 $X = \{x_1, x_2 \dots x_T\}$ , 以及目标流量序列 $Y = \{y_1, y_2 \dots y_{T-1}\}$ , 预测 $\hat{y}_T$ ,  $T$ 为时间步长大小, 即用前 $T$ 小时的观测数据去预测后 1 小时的流量值。其中, 其中 $y_1$ 表示历史观测序值 $x_1$ 所对应的站点流量。

依据 3.2.2 小节, 本文将所有的历史观测值分为 3 个部分: $S_{hour}, S_{day}, S_{week}$ , 这三部分的区别在于取值的特征窗口 $w$ 不同。当特征窗口 $w$ 为 $1hour$ 时, 时间步长 $T=3$ 时, 即以 $w$ 为单位,  $T$ 为长度进行采样, 获得历史数据作为特征; 当特征窗口 $w$ 为 $24hour$ 时, 时间步长 $T=3$ 时, 即获取要预测时段的前 3 天的历史流量数据作为特征; 当特征窗口为 $148hour$ 时, 时间步长 $T=3$ 时, 即获取要预测时段的前 3 周的历史流量数据作为特征。公式 3.1 表示如下“

$$S = [X_{t-Tw}, X_{t-(Tw-1)} \dots X_{t-1}] \quad (3.1)$$

当 $w$ 取值为 1 小时、24 小时、148 小时时,  $S$  分别记为 $S_{hour}, S_{day}, S_{week}$ 。

本文引入的上下文特征包含时刻、气象数据、周末/非周末、预测时段位于一天中的哪一个时间段以及其处于一周中的周几。上下文特征本文用  $E$  表示, 如 3.2 所示。

$$E = \{weather, windspeed, pressure, day of week, hour of day\} \quad (3.2)$$

则  $t$  时刻所对应的输入特征为 $X_{input}^t = \{S_{hour}^t, S_{day}^t, S_{week}^t, E_t\}$ 。最终共享单车流量预测问题定义为: 输入时间步长 $T$ 内的所有输入 $\{X_{input}^1, X_{input}^2 \dots X_{input}^T\}$ 和部分观测序列 $\{y_1, y_2 \dots y_{T-1}\}$ , 求 $T$ 时刻对应的流量 $\hat{y}_T$ , 公式如 3.3 所示。

$$\widetilde{y}_T = F(X_{input}^1, X_{input}^2 \dots X_{input}^T, y_1, y_2 \dots y_{T-1}) \quad (3.3)$$

其中,  $E_t$  是  $t$  时刻对应的上下文特征。 $F$  是需要学习到的一种非线性映射关系。

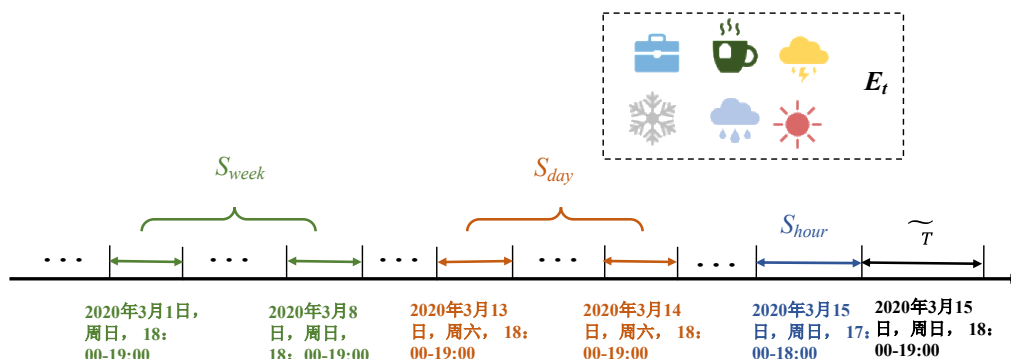


图 3.13 流量预测问题定义

Fig.3.13 Traffic prediction problem definition

### 3.3.2 基于 LSTM 的时序特征提取

对共享单车流量进行预测的本质是对 seq2seq 问题的处理, seq2seq 问题是指输入和输出都是序列, 如机器翻译<sup>[40]</sup>、问答系统、文档摘要等。编码器-解码器 (Encoder-Decoder) 模型<sup>[41]</sup>很适合处理 seq2seq 序列问题。在本文中, 编码器负责将输入的共享单车历史流量序列及相应特征转化为一个固定长度的语义向量  $C$ , 而解码器负责将编码器生成的语义向量  $C$  再转化为流量预测结果输出, 如图 3.14。

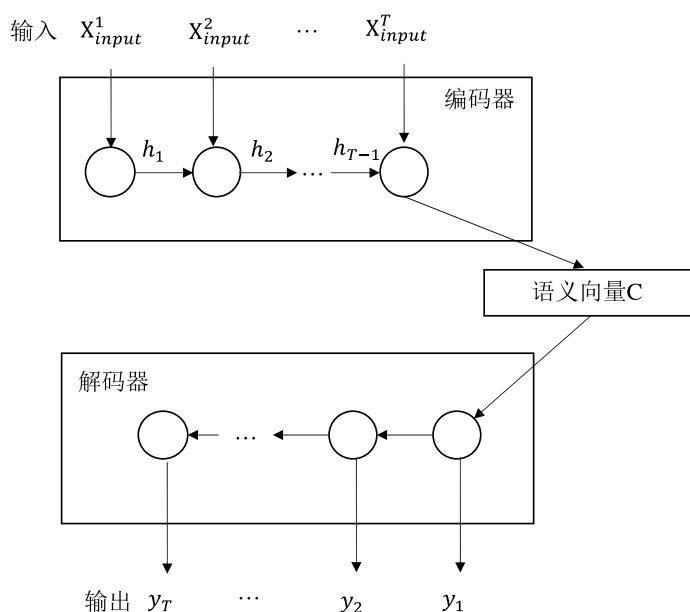


图 3.14 编码器解码器模型

Fig.3.14 Encoder-Decoder model

本文在编码器和解码器的选择上采用了 LSTM 模型, 该模型对每个神经元都采用了经典的三层门结构, 即遗忘门、输入门和输出门。LSTM 保留了经典循环神经网络模型对时间序列处理的优点, 同时门结构又能够对细胞状态进行删除或者添加信息, 克服了 RNN 无法捕捉长时间依赖关系的问题。假设编码器的输入为  $\{X_{input}^1, X_{input}^2 \dots X_{input}^T\}$ , 在 LSTM 中, 当前时间  $t$  的隐藏状态是由上一时间的状态  $h_{t-1}$  和当前时间的输入  $x_t$  共同决定的, 则根据公式 3.4 可以得到编码器中的隐藏状态。

$$h_t = f_1(h_{t-1}, x_t) \quad (3.4)$$

编码器逐个读入输入的历史流量观测数据  $\{X_{input}^1, X_{input}^2 \dots X_{input}^T\}$ , 将其编码为一个隐藏状态序列  $\{h_1, h_2 \dots h_T\}$ , 这些隐藏状态中包含了原始输入数据的特征信息, 通过某种规则  $q$  共同构成语义向量  $C$ , 供解码器使用。一种简单的语义向量  $C$  生成方法是直接取得最后的隐藏层, 如公式 3.5。但该方法忽略了前面  $T-1$  个时刻的特征信息, 最终的预测结果仅与最后一个时刻的隐藏层状态相关, 预测误差大。

$$C = q(\{h_1, h_2 \dots h_T\}) = h_T \quad (3.5)$$

解码器可以看作是编码器的逆过程。在解码器阶段, 根据给定的语义向量  $C$  和部分观测序列  $\{y_1, y_2 \dots y_{T-1}\}$ , 来共同预测下一个输出的流量值  $\hat{y}_T$ 。在解码器部分, 仍然使用 LSTM, 预测  $\hat{y}_T$  的公式为 3.6。

$$\hat{y}_T = g(s_t, y_{T-1}, C) \quad (3.6)$$

其中,  $s_t$  是 LSTM 中  $t$  时刻对应的隐藏层状态。语义向量  $C$  为编码器的输出, 它包含了编码后的输入信息。 $y_{T-1}$  是  $T-1$  时刻的输出, 反过来作为  $T$  时刻的输入,  $g$  是解码器中的 LSTM 网络。

Encoder-Decoder 模型虽然经典, 能够很好地解决 seq2seq 问题, 但它也存在一定的局限性。由于编码器和解码器之间的唯一联系是一个固定长度的语义向量  $C$ , 这要求编码器要将整个输入的历史流量序列信息压缩进一个固定长度的向量中去。但这样做存在两个弊端: 首先, 语义向量  $C$  的长度被限定, 即只有部分流量特征被编码, 语义向量  $C$  无法完整表示整个序列的信息; 其次, 时间节点靠前的流量特征信息会被时间节点靠后的流量特征稀释掉, 甚至被覆盖。但在单车流量预测问题中, 输入的序列有时候会很长, 由于编码器的限制, 语义向量  $C$  未能获得输入序列足够的有用信息。因此, 在解码的时候准确度会降低。为了解决这个问题, 本文在编码器部分添加了注意力机制, 使得在编码的时候能够保留与预测结果更为相关的特征, 而忽略掉相对不重要的特征。

### 3.3.3 基于注意力机制的特征重要性计算

注意力机制是一种能让模型对重要信息重点关注并充分学习吸收的技术。在产生输出的时候, 还会产生一个“注意力范围”表示接下来输出的时候要重点关注输

入序列的哪些部分，然后根据重点关注的部分来产生下一个输出。这样能够在产生输出的时候关注到与其最为相关的部分信息，而忽略掉重要程度较低的信息。

在共享单车流量预测中，将 $\{X_{input}^1, X_{input}^2 \dots X_{input}^T\}$ 作为输入，但这部分输入并不是全部都对 $t$ 时刻的结果产生重要意义，其中有的部分对预测结果的作用是微弱的，有的是相对重要的。传统的 LSTM 模型对所有输入的特征信息都“一视同仁”，所有特征的重要性是一致的，但对于共享单车流量预测问题，如果本文预测 18 时的流量数据，17 时的数据对于 18 时的预测来说更为重要，而 15 时流量数据的影响作用相对较弱。为了对这种重要性进行区分，使得编码器中的语义向量 $C$ 能够更有效、更完整的获取输入序列的特征信息，本文在编码器结构中引入了注意力机制。

对输入的时间序列数据 $\{X_{input}^1, X_{input}^2 \dots X_{input}^T\}$ 利用 LSTM 模型进行编码的到各输入特征对应的隐藏层状态，将隐藏向量序列 $\{h_1, h_2 \dots h_T\}$ 以加权的形式累加，如公式 3.7 所示。

$$C_t = \sum_{i=1}^T a_t^i h_i \quad (3.7)$$

此时的编码器会将输入信息编码成一个语义向量序列 $\{C_1, C_2 \dots C_T\}$ ，每一个语义向量都包含对该时刻流量预测结果更为重要的特征信息。而在解码的时候，每一步都会选择性的从语义向量序列中挑选一个子集进行进一步的处理。在解码器产生输出的时候，就能够做到充分利用输入序列所携带的特征信息，并且保证这些信息是对预测结果更为重要的部分。

由于本文在编码的时候使用了 LSTM 模型，因此可以认为 $h_i$ 包含了输入的流量特征序列中第 $i$ 个输入序列及其前面一些序列的信息。将隐藏层向量按照权重相加，表示在生成 $t$ 时刻输出的时候的注意力分配是不同的。 $a_t^i$ 的值越大，表示 $t$ 时刻对应的输出在第 $i$ 个输入序列上分配的注意力越多，也就是说第 $i$ 个输入序列所对应的特征对 $t$ 时刻来说更为重要，在生成 $t$ 时刻输出的时候受第 $i$ 个输入序列的影响也就越大。而 $a_t^i$ 是由 $t-1$ 时刻对应的输出隐藏状态 $s_{t-1}$ 和输入中各隐藏层状态共同决定的，如公式 3.9 所示。为了区分编码器与解码其中的隐藏层状态，令 $t$ 时刻编码器中的隐藏层状态为 $h_t$ ，解码器中的隐藏层状态为 $s_t$ 。

$$Score_t^i = V^T \tanh(W[h_i, s_{t-1}]) \quad (3.8)$$

$$a_t^i = \text{softmax}(Score_t^i) = \exp(Score_t^i) / \sum_{j=1}^T \exp(Score_t^j) \quad (3.9)$$

其中， $F_1$ 表示一种非线性映射关系，目的是使 $s_{t-1}$ 与每一个输入特征向量对应的隐藏层状态 $h_i$ 计算得到一个数值，然后使用 $\text{softmax}$ 得到 $t$ 时刻的流量预测输出在各输入隐藏层状态中的注意力权重，即为每个影响因子赋予了一定的权重，用以表示输入特征的重要性。在解码的时候可以根据输出对应的语义向量 $C_t$ 进行解码，

$C_t$ 包含了对该预测时候最重要的那部分信息，而忽略了不重要的特征信息，使得预测结果可以更加准确。

### 3.3.4 基于 LSTM 和注意力机制的共享单车流量预测模型

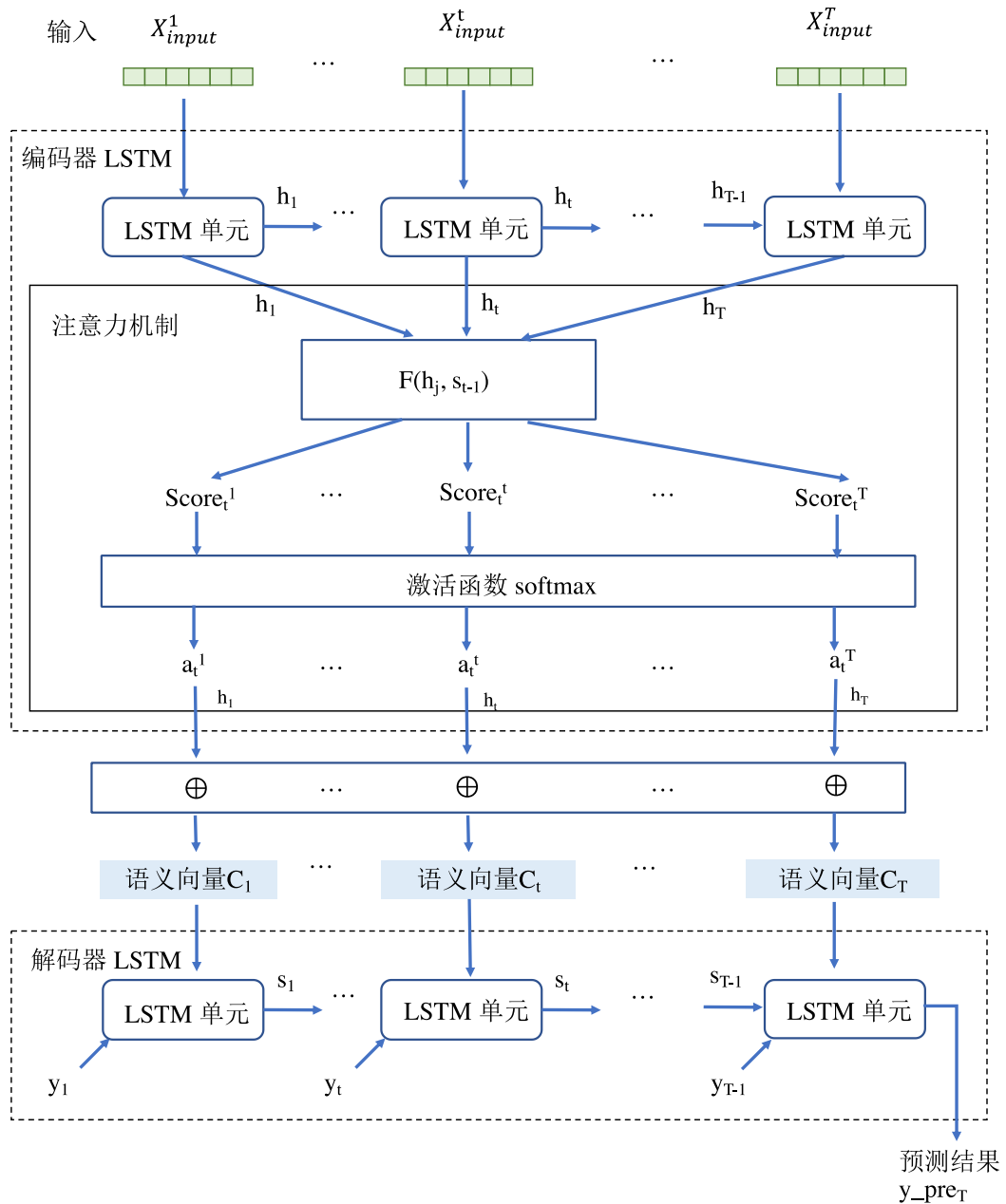


Fig.3.15 Model structure diagram

如 3.4.1 和 3.4.2 两小节所讨论的，Encoder-Decoder 模型能够很好的处理共享单车流量预测这类 seq2seq 问题，但由于编码器的限制，语义向量  $C$  未能获得输入



序列足够的有用信息。而注意力机制能够让模型对重要信息重点关注并充分学习吸收,不再局限于将所有的输入信息 $\{X_{input}^1, X_{input}^2 \dots X_{input}^T\}$ 都编码为一个固定长度的语义向量 $C$ ,而是将输入信息编码成语义向量序列 $\{C_1, C_2 \dots C_T\}$ ,每一个语义向量都包含对该时刻流量预测结果更为重要的特征信息,弥补了 Encoder-Decoder 模型的不足。因此,本章以 Encoder-Decoder 模型为框架,在 Encoder 部分加入了注意力机制,提出了一种基于 LSTM 和注意力机制的共享单车流量预测模型,模型结构如图 3.15 所示。输入信息 $\{X_{input}^1, X_{input}^2 \dots X_{input}^T\}$ 通过 LSTM 网络计算得到隐藏层向量 $\{h_1, h_2 \dots h_T\}$ (如公式 3.4 所示),每一个隐藏层向量与 $s_{t-1}$ 计算得到注意力得分(如公式 3.8 所示),对所有注意力得分使用激活函数 $softmax$ 得到注意力权重(如公式 3.9 所示),所有注意力权重的总和为 1。此时这些注意力权重表示在计算某时刻流量输出时,对应的输入序列各自的重要性。通过该方式计算得到语义向量序列 $\{C_1, C_2 \dots C_T\}$ ,通过对 LSTM 作为解码得到的语义向量序列 $\{C_1, C_2 \dots C_T\}$ 进行解码,得到最终的预测结果。

$$s_t = f_2(s_{t-1}, U[y_{t-1}; C_{t-1}] + b) \quad (3.10)$$

其中, $y_{t-1}$ 是 $t-1$ 时刻对应的观测值,即 $t-1$ 时刻对应的共享单车流量, $C_{t-1}$ 是 $t-1$ 时刻对应的语义向量,包含与 $t-1$ 时刻流量预测值最为相关的输入特征信息。 $[y_{t-1}; C_{t-1}]$ 表示将二者进行拼接,同 $t-1$ 时刻对应的隐藏层状态一起作为 LSTM 网络的输入, $f_2$ 即表示通过 LSTM 模型进行计算。 $U$ 和 $b$ 是网络中需要学习到的参数。

$$\widehat{y}_T = V(W[s_T; C_T] + b_w) + b_v \quad (3.11)$$

利用公式 3.11 可以得到最终的预测结果。 $\widehat{y}_T$ 是 $T$ 时刻所对应的流量预测值, $V$ 、 $W$ 、 $b_w$ 和 $b_v$ 都是网络中需要学习到的参数。 $s_T$ 对应 $T$ 时刻解码器的隐藏层状态,公式如 3.10 所示, $C_T$ 为 $T$ 时刻对应的编码得到的语义向量。

### 3.4 实验设计及分析

为了验证本文所提模型的有效性,围绕如下几个方面展开了实验:

- ① 以站点预测结果为据,验证本章所提算法有效性。
- ② 与其它共享单车流量预测算法比较,验证本章所提算法有效性。
- ③ 验证本章在 3.2.3 节对气象数据特征的取舍的合理性。

#### 3.4.1 实验数据集简介

实验所用数据集来自花旗自行车数据<sup>[42]</sup>,它收集了自 2013 年以来的用户骑行历史数据。实验中采用 2017 年 3 月至 10 月份的用户历史骑行数据,共计 11942180 条。本文添加了同时段的气象数据,最终实验数据集详情(包括上下文特征数据)

如表 3.2 所示。实验时按照时间选择全部数据的 70%作为训练集，余下的 30%作为测试集。

表 3.2 数据集详情

Table 3.2 Detail of the datasets

数据集	花旗单车数据
地点	纽约
时间跨度	2017/03/01-2017/10/30
数据字段	15 种
缺失值	年龄（缺失率 0.5%，平均数填充）
数据条数	11942180
总站点个数	621
天气类型	6 类（如：小雨、雪等）
温度	[262.7, 292.3]华氏度
风速	[0, 12]
压强	[995, 1040]
湿度	[12, 95]

3.4.2 评价指标

本文采用了平均绝对误差（Mean Absolute Error, MAE）和均方根误差（Root Mean Squared Error, RMSE）作为评价指标。平均绝对误差能够反映出模型的拟合效果，公式如 3.12 所示，MAE 累积了每一个预测值与真实值之间的误差，MAE 越大说明预测误差越大，故一个预测能力较好的模型应该保证 MAE 越小越好。均方根误差能够对模型预测误差较大的值进行放大，能够较好地对比不同模型之间的鲁棒性，其公式如 3.13 所示，一个较好地模型也应该保证 RMSE 越小越好。

$$MAE = 1/n \sum_{i=1}^n |y_i - \hat{y}_i| \tag{3.12}$$

$$RMSE = \sqrt{1/n \sum_{i=1}^n (y_i - \hat{y}_i)^2} \tag{3.13}$$

其中， $n$ 为测试样本个数， $y_i$ 为真实值， $\hat{y}_i$ 为预测值。

3.4.3 实验结果分析

① 流量预测方法有效性验证

以站点 Avenue D & E 3 (ID=302)为例，按照小时为预测单位，图 3.16 和图 3.17 展示了该站点在未来一天以及一周当中的流量预测结果。对比真实值和预测值可以发现，本文所提的模型能够较好地拟合出序列的变化趋势。特别地，如图 3.16

所示, 7-8 时许和 17-19 时许出现了波峰和波谷, 这也体现出了 3.2.2 小节所分析到的早高峰和晚高峰情况, 说明本文在时间维度上所提取的特征是有效果的, 模型较好地对高峰情况进行了学习。

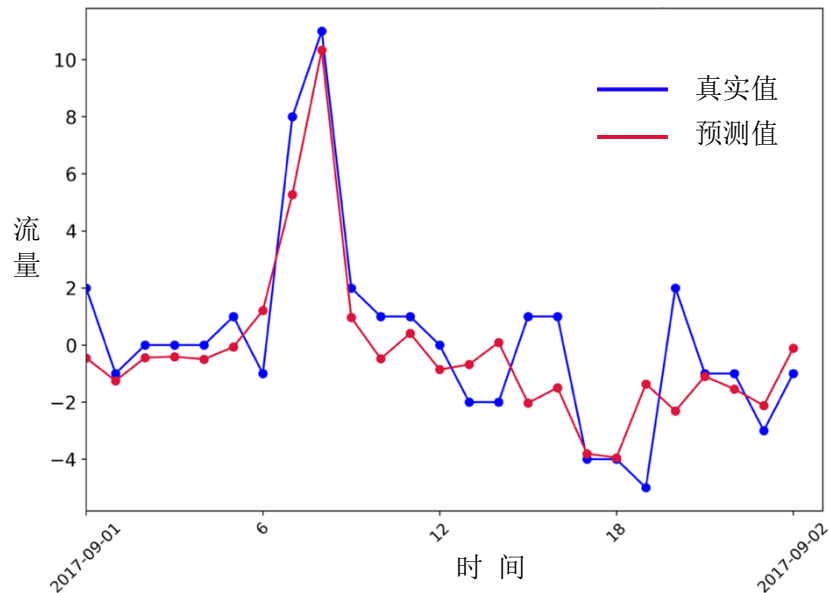


图 3.16 站点 Avenue D & E 3 (ID=302)未来一天内的流量预测结果

Fig.3.16 Traffic forecast for the next day for site Avenue D & E 3 (ID=302)

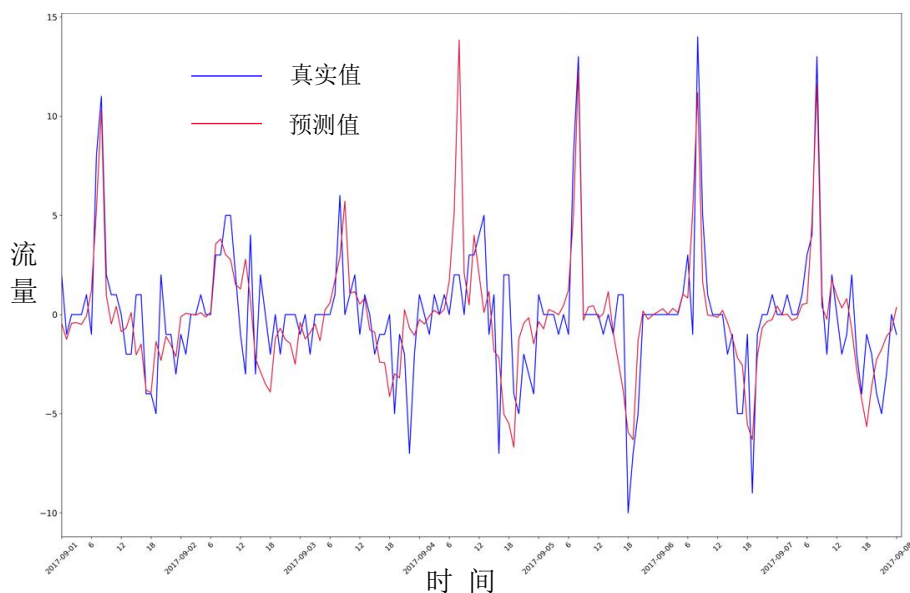


图 3.17 站点 Avenue D & E 3 (ID=302)未来一周的流量预测结果

Fig.3.17 Traffic forecast for the next week for site Avenue D & E 3 (ID=302)

图 3.17 是该站点一周的流量预测结果。2017 年 9 月 1 日为星期五, 图 3.17 从周五开始, 展示了未来一周的共享单车流量预测结果。从图中可以看出, 9 月 2 日

与 3 日的数据与其它 5 天的数据有所不同, 差异在于波峰和波谷并未有其余几天明显, 这也恰好对应了本文前面的分析, 周末和非周末用户的出行规律和用车需求是不同的, 同时说明了本文所提的模型较好地捕捉到了周末与非周末这一时间层面的特性。

## ② 和其它流量预测方法进行对比

本章提出的共享单车流量预测算法以编码器-解码器为框架, 利用 LSTM 和注意力机制对流量进行预测。为了更好验证模型的有效性, 本文将本章节所提算法与下列算法做了对比:

HA: 利用历史所有观测值的平均值来预测未来值。

ARIMA<sup>[43]</sup>: 通过发现时序数据之间的相关性来预测未来值。

RNN<sup>[44]</sup>: 利用循环神经网络捕捉时序数据之间的相关性来预测未来值。

ST-RNN<sup>[45]</sup>: 将时空特性融合到带有时间转移矩阵和距离转移矩阵的 RNN 模型中预测未来值。

表 3.3 模型误差分析表-1

Table 3.3 The analysis table-1 of model error		
方法	MAE	RMSE
HA	0.697	1.054
ARIMA	0.576	0.801
RNN	0.328	0.445
ST-RNN	0.287	0.325
<b>LSTM+Attention</b>	<b>0.239</b>	<b>0.294</b>

对比结果如表 3.3 所示。从表中可以看出本章节所提的算法 LSTM+Attention 相对于其他算法, 在评价指标 MAE 和 RMSE 上均有所提升。本章所提算法 LSTM+Attention 在两种指标上均远超 HA 算法, 这说明采用历史平均值对流量进行预测误差很大, 需要从历史流量数据中提取出时间相关的特征, 也说明了本文从特征入手进行考虑的合理性与有效性。本章所提算法相比于 ARIMA 和 RNN 在预测能力上都有明显提升。这说明本章所提模型引入 RNN 能够较好地对时间序列模型进行处理。由于 ARIMA 仅考虑了目标序列 $\{y_1, y_2 \dots y_t\}$ , 而忽略了驱动序列 $\{x_1, x_2 \dots x_t\}$ , 所以预测效果不佳, 也说明了本章考虑输入序列特征的有效性。对于 ST-RNN, 虽然其对输入的历史流量数据进行了时间上的建模, 但相比本章提出的模型, 未能捕获到对于预测结果来说更为重要的特征信息, 说明了本章对引入气象数据, 以及利用注意力机制捕获对预测结果来说更为重要的特征的有效性。

虽然本章所对比的多个方法流量预测误差 MAE 已经在 1 以下,更加准确的预测结果能够帮助单车营运公司进行诸如调度、换电等其它业务,在数值上提升的幅度或许较小,但转换到实际场景中,可能为公司节约了几十万甚至几百万的成本。

### ③ 气象数据取舍的合理性验证

表 3.4 模型误差分析表-2

Table 3.4 The analysis table-2 of model error		
方法	MAE	RMSE
LSTM+Attention-HT	0.271	0.315
LSTM+Attention	0.239	0.294

本章在 3.3.2 小节针对气象特征对用户历史骑行数据进行了分析,为了验证其合理性,本文通过实验对包含湿度和温度的 LSTM+Attention-HT 模型和 LSTM+Attention 模型进行了对比,结果如表 3.4 所示。从表中可以看出, LSTM+Attention 模型在两个评价指标上都要优于 LSTM+Attention-HT 模型,这说明 3.3.2 小节对于特征的取舍是合理的。也说明了本文猜想的合理性,即温度、湿度与天气三者之间确实存在特征冗余,而特征冗余会影响模型的预测能力,删除冗余的特征可以帮助提升模型性能。该结论的得出从侧面说明了本章节所做特征分析与筛选工作的重要性。

综上所述,本章所提的 LSTM+Attention 算法能够构建动态的时序模型,解决了无法区分输入特征重要性的问题,同时考虑到了大部分研究工作所忽略的气象特征。实验表明本章所提算法能够提高单车流量预测的效果。

## 3.5 本章小结

本章针对用户历史骑行数据集进行分析,筛选和构建了相关特征,并提出了一种基于 LSTM 和注意力机制的共享单车流量预测方法。本章首先描述了目前共享单车流量预测中存在的问题。接着对本章所提算法的思想进行了详尽地阐述。最后利用花旗自行车数据集对本章所提算法进行了验证,实验结果表明本章节所提出的算法能够提高预测能力,在 MAE 和 RMSE 两种评价指标上表现良好,说明了所提算法的有效性,并且通过实验验证了特征筛选的合理性。

## 4 基于强化学习的单车调度问题研究

### 4.1 问题背景

共享单车各个站点之间的自行车流量是不均衡的,会导致一些热点地区的站点在某时段的自行车数目严重不足,部分用户无车可用,而一些自行车站点的车辆长时间无人使用。这会造成用户用车不便,导致用户流失,降低用户粘性。因此,如何有效地解决这种不平衡问题,是目前面临的主要挑战。部分研究学者将共享单车调度问题建模为一个马尔可夫决策过程,从改变定价的角度去激励用户将车辆停放至目的地附近的其它站点,以缓解部分站点的停放压力。也有研究学者考虑定制“红包车”,通过为用户提供选择来归还自行车至指定站点,以获取金钱奖励。这两种方法的目的都是让用户从指定的车辆数目较多的站点骑走车辆,或者归还车辆到车辆数目较少的站点。

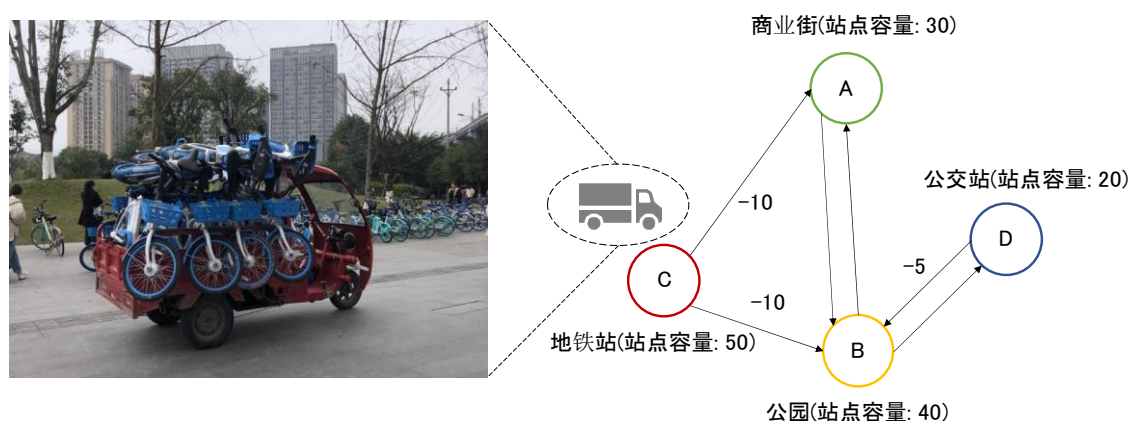


Fig.4.1 Bike rebalancing diagram

但是,上述方法涉及的主要调度者是用户,具有不确定性,无法保证用户一定能够按照推荐的方式去选择车辆。为了避免由用户带来的不确定性,同时提高调度效率,有研究学者利用调度运输工具,解决在高峰时段维护共享单车系统再平衡问题,并且在夜间重新平衡以准备系统在高峰时段的使用。但是,该研究主要集中在发现全局特征上,假设站点内的自行车需求对全局特征的反应相等,这在城市面积大且高度多样化的情况下会带来较大的误差。因为,在区域最南边的车辆与最北边的车辆在实际情况中不存在调度关系。

以用户为主的研究方法，调度效率较低，同时会存在资源浪费的情况。而一些研究仅考虑了全局特征对调度情况的影响，却忽略了调度问题是可以在特定区域内进行的。为了解决上述问题，本章首先利用基于转移矩阵的 K-Means 聚类算法将研究的自行车站点划分为不同的区域，然后利用强化学习中的 Q-Learning 方法对每个区域中站点间的车辆进行调度，在考虑到局部特征的情况下最大效率的进行调度优化。利用图 4.1 所示的红色运输车辆，将各个站点的自行车进行挪动，使得站点之间的自行车数目相对平衡，最大限度的满足用户的用车需求。

## 4.2 基于 Q-Learning 算法的共享单车调度模型研究

### 4.2.1 单车调度问题定义

**站点：**一个自行车站点有  $N$  个停车位，能够容纳  $N$  个车辆。所有的自行车站点可用有向图  $G = (S, E)$  表示，以每个站点  $s \in S$  为节点，边  $E$  是每个自行车站点中的所有连边，有  $e_{ij} = (s_i, s_j) \in E$ ，其中  $i, j$  是第  $i, j$  个站点。

**调度车：**共享单车营运公司会派驻调度管理人员，利用调度车辆往返各站之间，将即将拥有过多车辆的站点中的车辆挪走，紧接着去到车辆使用需求大的站点查看是否需要增添车辆，如图 4.1 所示。

**调度信息：**调度信息包含四个部分，站点库存能力  $s_i^c$ ，即这个站点最多能够容纳多少辆自行车停放。站点目前库存车辆  $s_i^n$  以及  $t$  时刻站点的流量  $s_i^{flow(t)}$ ，站点调度车辆数目  $s_i^{target}$ ，即需要调走或者放下的车辆数目。

**共享单车调度问题：**通过给定的站点库存能力  $s_i^c$ ，站点目前库存车辆  $s_i^n$  以及  $t$  时刻站点的流量  $s_i^{flow(t)}$ ，在调度时间最少的情况下，求出每个站点的需要调度车辆数目  $s_i^{target}$ 。

调度问题的实质就是通过调度车辆，将每个站点的自行车数量恢复到目标值。本文的调度目标是使得某个区域内各站点之间的耗时  $T$  最小。因此，调度车辆数目  $s_i^{target}$  由式 4.1 决定。其中， $OP_{i,w}$  表示站点  $s_i$  在时间窗口  $w$  内的最优调度状态。

$$OP_{i,w} = \operatorname{argmin}_{s_i^{target}} \{T \mid 0 \leq s_i^n + s_i^{target} + \sum_{t=0}^w s_i^{flow(t)} \leq s_i^c\} \quad (4.1)$$

### 4.2.2 基于转移矩阵和 K-Means 的站点聚类

图 4.2 展示了本文所选取的纽约市自行车数据集中每个站点的分布情况。从图中可以看出，在单车调度的时候，并不需要对整个区域进行调度，即不必从最北边往最南边进行调度，对站点进行聚类后再对区域内进行单车调度，能够节省资源，提升效率。

聚类的目的要使得簇内的站点之间联系紧密，而簇间交互较少。传统的基于网格的聚类算法直接以经纬度作为聚类依据，无法合理的考虑站点之间的交互情况。聚类算法应选取原理简单、实现方便、收敛速度快的算法。同时本章所使用

的数据集较大，采用的聚类算法需要在大规模也能保持较好地可伸缩性和高效性。依据这个标准，本文选择了 K-Means 作为基础聚类算法。

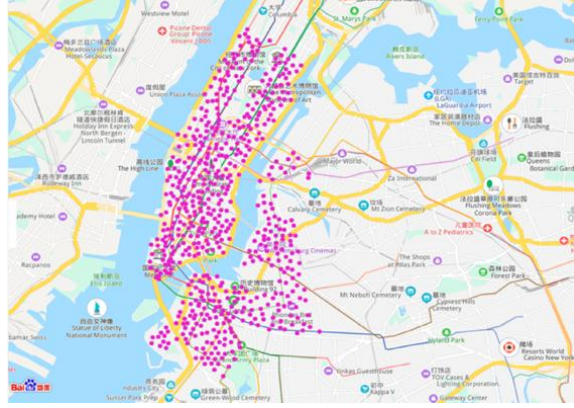


图 4.2 站点分布

Fig.4.2 The station distribution

首先，随机选择  $k_m$  个聚类中心， $k_m$  属于质心集合  $K_1$ ，站点集合  $S$  中的所有数据点  $s$  基于公式 4.2 被分配到多个聚类中。

$$\operatorname{argmin}_{k_m \in K_1} \operatorname{dist}(k_m, S)^2 \quad (4.2)$$

其中， $\operatorname{dist}(\cdot)$  是标准  $L_2$  欧式距离，将第一次聚类中指向第  $m$  个聚类质心的站点集合设定为  $Clu_{1,m}$ 。

其次，利用公式 4.3 进行质心的更新，得到最终的聚类结果。

$$k_m = 1/|Clu_{1,m}| \sum_{s_i \in Clu_{1,m}} s_i \quad (4.3)$$

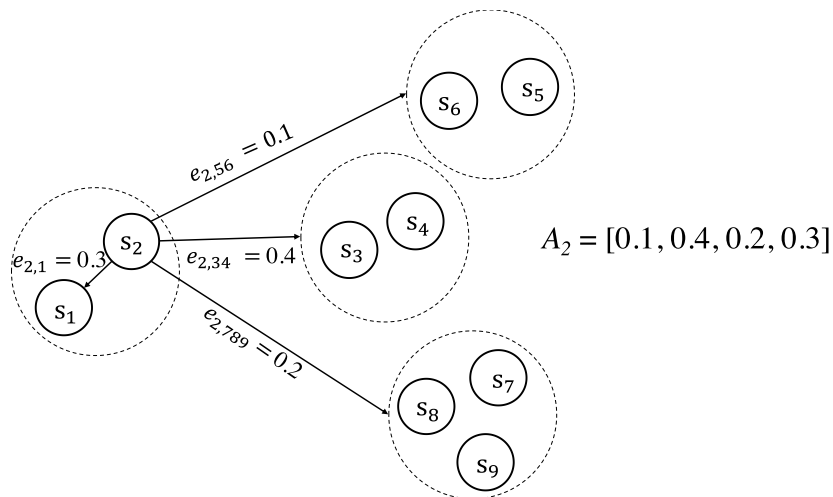


图 4.3 转移矩阵示意图

Fig.4.3 Transition matrix diagram



但是仅用 K-Means 算法在实际应用场景中仍有局限性, 相邻两个联系紧密的站点被划分在了两个类中, 这两个站点分属不同调度人员管理。但其实, 两个交互频繁的站点可以同属于一个调度管理人员, 更大可能节省资源。为了解决这个问题, 本文在 K-Means 算法中引入了转移矩阵。

转移矩阵由数值组成, 这些数值是 24 小时内, 第  $i$  个站点到所有不包含自身站点的转移概率, 即从  $i$  到其它站点的概率。转移矩阵示意图如图 4.3 所示。其中,  $e_{2,56}$  指的是从  $s_2$  到  $s_5$  及  $s_6$  的概率总和,  $\{s_5, s_6\} \in Clu_{1,m}$ , 此概率也就表示了站点到  $s_2$  到  $s_5, s_6$  所组成的簇中的概率, 如此迭代, 计算第  $i$  个站点的转移矩阵  $A_i$ 。概率的计算能够帮助后面二次聚类的时候筛选类中相关性不大或类间相关性较大的站点, 增加聚类的合理性。

然后, 随机选择  $k_n$  个聚类中心,  $k_n$  属于质心集合  $K_2$ , 转移矩阵  $A_i$  中的所有数据点  $a$  基于公式 4.4 被分配到一个聚类中。

$$\operatorname{argmin}_{k_n \in K_2} \operatorname{dist}(k_n, A_i)^2 \quad (4.4)$$

将第二次聚类指向第  $n$  个聚类质心的站点集合设定为  $Clu_{2,n}$ 。接着, 利用公式 4.5 进行质心的更新, 得到最终的聚类结果。

$$k_n = 1/|Clu_{2,n}| \sum_{a_i \in Clu_{2,n}} a_i \quad (4.5)$$

最后, 对于所有  $n$ ,  $n \in (1, K_2)$ , 将  $Clu_{2,n}$  中的所有站点分类至簇  $[(Clu_{2,n,num} * K_1/N)]$  中, 以此更新  $Clu_{1,1}, Clu_{1,2}, \dots, Clu_{1,K_1}$ 。其中  $Clu_{2,n,num}$  表示第  $Clu_{2,n}$  个聚类中的所有站点的个数,  $N$  表示所有站点的个数。

本文通过手肘法则对两次 K-Means 聚类算法所使用的  $K$  值进行了选取。

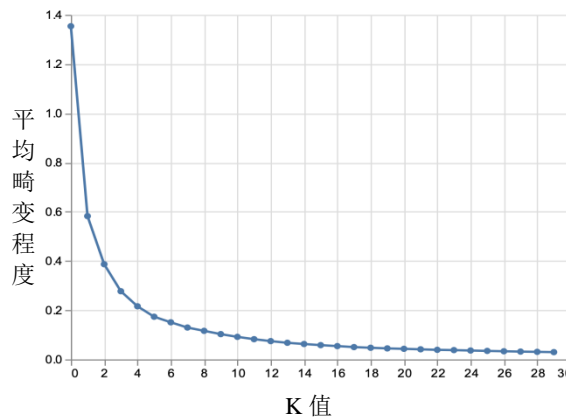


图 4.4 平均畸变程度随 K 值的变化-K1

Fig.4.4 The mean distortion degree varies with K value -K1

如图 4.4, 横坐标表示  $K$  的取值, 纵坐标表示平均畸变程度, 其公式如 4.6 所示。

$$SSE = \sum_{i=1}^k \sum_{p \in Clu_{1,m}} |p - mean_i|^2 \quad (4.6)$$

其中,  $Clu_{1,m}$  是第一次聚类第  $m$  个簇,  $p$  是  $C_i$  所有的样本点,  $mean_i$  是  $C_i$  中所有样本的均值, 误差平方和 (Sum of The Squared Error, SSE) 表示的是所有样本的聚类误差, 衡量聚类效果的好坏。根据图 4.4, 并且考虑实际情况, 设定一个簇由一个网格员进行车辆调度, 本文将 K-Means 的簇数目定为 20。最终本文所使用到的数据集中共有 621 个站点, 则一个网格员人均负责的站点数约为 30。

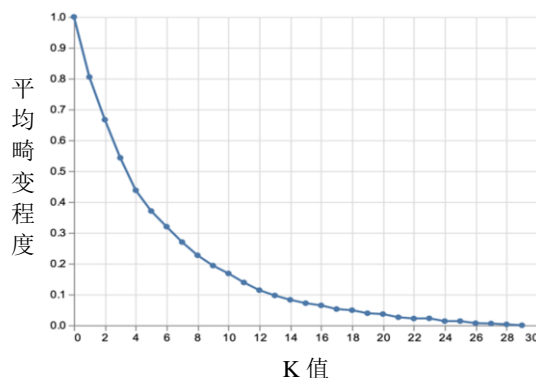


图 4.5 平均畸变程度随 K 值的变化-K2

Fig.4.5 The mean distortion degree varies with K value -K2

根据图 4.5, 本文设定  $K_2=10$ 。经过聚类之后得到聚类结果, 如图 4.6 所示。

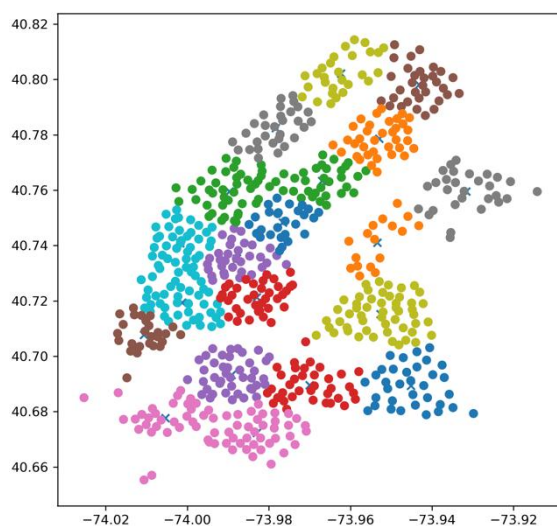


图 4.6 聚类结果

Fig.4.6 Clustering results

图 4.6 是对所有站点进行聚类后的结果。调度管理人员在区域之内进行调度即可, 不用考虑簇间调度。这是由于在实际的操作过程中, 单车营运公司会根据调

度管理人员的调度情况对其进行工资分配，若进行了簇间调度，会引发调度管理人员之间的矛盾以及调度工作量认定出现分歧，不利于公司管理。

### 4.2.3 融合聚类模型的 Q-Learning 单车调度算法

本节融合聚类模型，提出一种基于 Q-Learning 算法的单车调度策略，算法整体框架如图 4.7 所示。原始站点在城市区域的各个地方分布，通过 4.2.2 小节所描述的基于转移矩阵和 K-Means 聚类算法对原始站点进行聚类，在后续的处理中，只针对每个聚类簇中的站点进行调度研究。因为在实际的营运过程中，共享单车管理公司会雇佣相应的调度管理人员，每个人员负责特定的区域，区域间的调度被忽略，此举避免了调度管理人员之间的分歧，最大程度节省资源的同时公司能够更好地对人员进行管理。对于聚类后的每一个区域，都会单独运用 Q-Learning 算法进行调度处理，在每次的迭代中，都将随机选择到一个起始状态，如果智能体判断本次调度没有进入目标状态，即本次调度没有获取到最大的奖励价值函数或者未更新完区域内所有站点。则从当前状态下所有可能的动作中选择一个给出最大奖励的一对动作和状态，直到获取到目标状态或者将区域内所有的站点更新完，则更新结束，输出调度结果。

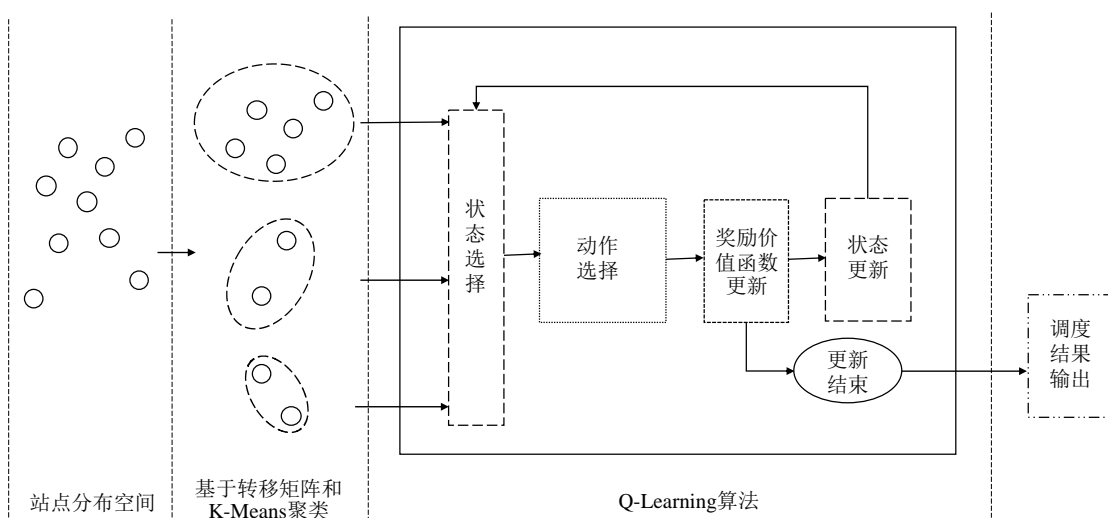


图 4.7 算法整体框架图

Fig.4.7 Overall framework diagram of algorithm

为了更好地阐述本章所提算法，首先就 Q-Learning 中的几个基本概念结合共享单车调度场景进行简单介绍。

**状态(state):** 状态是一个有限的集合，是人为标定的特征向量。在共享单车调度问题中状态表示的是当前站点中的自行车数量。某个特定的状态用  $s$  表示， $s \in S$ 。

动作(action): 动作也是一个有限的集合, 是指调度车辆应该在某一站点调走或者放置的自行车数目。某个特定的动作用 $a$ 表示,  $a \in A$ 。

奖励(reward): 奖励是智能体采取某个动作后得到的即时奖励, 它为强化学习提供了目标。用 $R(s, a)$ 表示在状态 $s$ 执行了 $a$ 动作后所得到的即时奖励。

策略(policy): 策略是指根据当前的状态 $s$ 如何产生下一个动作 $a$ 。某个特定的策略用 $\pi$ 表示。

表 4.1 动作与奖励设置

Table 4.1 Action and reward settings

取值	
动作	$[-10, -5, -3, 0, 3, 5, 10]$
奖励	移动后 $Min_{num} \leq Num_{now} \leq Max_{num}$ : $r = -0.5 * \sqrt{a}$
	移动后 $Max_{num} \leq Num_{now}$ : $r = -30$
	移动 $Num_{now} \leq Min_{num}$ : $r = -30$

本章所提算法中对动作和奖励的设置如表 4.1 所示, 其中 $Min_{num}$ 指该站点最少车辆数, 本文设置为 0, 因为一个站点即使没有车辆也不会存在车辆为负数的情况,  $Num_{now}$ 是调度后站点车辆的库存情况,  $Max_{num}$ 是站点最大的车辆容纳情况, 这个是由单车运营调度公司设定的, 在本文中取值为 60。

本文的目标是使模型获得最大化期望累计奖励, 即得到 $r_1 + r_2 + \dots + r_t + r_{t+1} + r_{t+2} \dots$ 的最大值,  $r_t$ 即 $t$ 时刻之前的奖励, 是已经获取到的, 而 $t$ 时刻之后如何选取动作, 即调度车辆数目, 从而使得单车调度方法最优, 这时候通过定义 $t$ 时刻之后回报 $G_t$ , 如公式 4.7 所示, 并结合价值函数  $Q$  的更新求解。

$$G_t = r_{t+1} + \gamma r_{t+2} + \gamma^2 r_{t+3} \dots + \gamma^{t_{end}-t} r_{t_{end}} \quad (4.7)$$

其中,  $t_{end}$ 整个强化学习过程结束的时刻, 在共享单车调度场景中, 本文以一天为调度周期,  $t_{end}$ 也就指一天结束的最后一个小时。 $\gamma \in [0, 1]$ 是折扣因子, 表示对未来回报的关注度,  $\gamma$ 的引入将一个无限长度的问题转化为了一个拥有最大值上限的问题。当 $\gamma = 0$ 时, 表示只关注与眼前的利益, 也就是说在车辆调度的过程中近似于贪心地选取下一个将带来回报最大的动作, 当 $\gamma$ 趋近于 1 时, 则表示对未来的回报和当前的回报一样注重, 即认为对晚高峰时刻的车辆调度数目和正常时段的调度数目的影响程度一样。

接下来, 定义最优动作价值函数, 如公式 4.8 所示。最优动作价值函数描述的是在 $t$ 时刻根据当前站点中的自行车数量选择一个动作 $a$ 进行车辆移动后, 能够获取的最大未来回报的期望。

$$Q(s_t, a_t) = \max_{\pi} E_{\pi}[G_t | S_t = s, A_t = a, \pi] \quad (4.8)$$

在获取到了每种状态下执行每个动作的最优动作价值函数后，即每个站点目前的库存自行车数量和一状态下，未来发生每种调度动作所能获取到的最优动作价值函数后，可以通过公式 4.9 获取当前状态下最优的动作集合。

$$a_t^* = \operatorname{argmax}_{a_t} Q(s_t, a_t) \quad (4.9)$$

Q-Learning 的最终目的是要构建一个状态和动作的二维表，当要采取动作时从这个表中选择使得当前状态的 Q 值最大的那个动作即可，对于最优价值函数的估计，通常使用 Bellman 方程迭代求解，如公式 4.10 所示。

$$Q(s, a) = (1 - \alpha)Q(s, a) + \alpha[R(s, a) + \gamma \max_a Q(s', a)] \quad (4.10)$$

其中， $\alpha$  是学习速率，学习率越大，保留之前训练的效果就越少，学习率越小，保留之前的训练效果就越多， $\alpha \in [0, 1]$ 。

### 4.3 实验设计及分析

为了验证本章节模型的有效性，围绕如下几个方面展开了实验：

- ① 与其它单车调度算法比较，验证本章所提算法有效性。
- ② 验证模型有效性：将未进行调度与最后一次实验结果（最终调度结果）进行比较，验证所提模型能够有效解决共享单车调度问题。
- ③ 分析“动作”与“奖励”变化：“动作”能够反映运输车辆在站点移动的车辆数目，“奖励”可反映整个迭代过程中学习变化情况，分析“动作”与“奖励”的变化能更完整地验证模型的有效性。
- ④ 分析调度“成功率”：除了验证模型能够合理地对共享单车进行调度外，由于有桩共享单车站点的可容纳车辆数目是固定的，利用“成功率”可反映在结束调度后站点内车辆在  $[0, 60]$  范围内的百分比，0 指站点内车辆数目不会小于 0，60 是各站点容量上限的平均数。

#### 4.3.1 实验数据集

本节使用的实验数据集来自花旗自行车（Citibike），与第三章所用数据集一致。

#### 4.3.2 评价指标

本章节采用调度完成所耗费的时间去衡量模型的效率，一次好的调度应该在尽可能短的时间内完成，这样才能让更多用户想要使用单车时就能有车可用，如公式 4.11 所示。站点之间的行驶距离由经纬度计算得到，站点之间行驶所耗费的时间由距离除以运输车辆的速度得到。

$$total\_time = \min \sum_i^{N-1} T(S_i, S_{i+1}) \quad (4.11)$$

4.3.3 实验结果及分析

本章所提的方法能够结合第三章的流量预测结果，对单车进行调度。本文在真实的数据集上进行了实验，验证了方法的有效性。

① 和其它调度方法进行对比

本章提出的共享单车调度策略在对所有站点聚类的基础上，利用 Q-learning 算法进行调度规划。为了更好地验证模型的有效性，本小节将所提算法与下列算法做了对比：

**Random：**采用随机选择站点的方法模拟无经验的管理人员进行调度。

**遗传算法（Genetic Algorithm，GA）：**遗传算法是研究人员在受到生物系统的计算机模拟技术的启发后创造的智能算法，该算法可以通过模拟生物的进化和遗传行为自适应概率优化，非常适合解决类似共享单车调度这样的复杂系统问题。

**Q-learning：**本章节所提出的共享单车调度策略。

表 4.2 调度时间对比

Table 4.2 Rebalancing time comparison	
方法	耗费时间(秒)
Random	3943
GA	2340
Q-learning	<b>1680</b>

传统的共享单车调度方法通常按照地理位置简单划分得到多个调度区域，每个调度区域由一个管理人员完成调度任务。**Random** 调度方法，模拟了无经验的管理人员，从任意站点出发，路径的选择也是凭借经验由调度人员进行选择。对于三种对比算法，本文均在相同的聚类中完成调度。从表 4.2 可以看出，**Random** 调度方法所耗费的时间最多，而且该调度方法以调度管理人员的经验为依据，不确定因素较多，会使得单车运营公司管理不便。本章节所提出的 **Q-learning** 调度策略完成调度所耗费的时间最短，相较于 **Random** 方法减少了 57%的时间花费，较遗传算法减少了 28%的时间花费，说明了本文所提策略能够有效地减少调度时间。

② 模型有效性验证

首先，本文为验证所提调度方法的有效性，将没有使用任何调度方法与使用本文所提调度方法所得到的调度结果进行了对比，结果如图 4.8 所示。两条红色的线分别表示站点容量的上限和下限。站点至少能够容纳 0 辆，因为即使该站点没有车辆停放，容量也不会为负；至多容纳的车辆数本文在此设置为 60 辆，取每个站点容量上限的平均值。站点容量上限及下限的设置代表在调度的时候需要考虑

实际情况，而非随意进行调度。如果调度的车辆太多，使得该站点车辆总数超过了 60 辆，也属于无效调度，因为没有多的停车桩可供使用，反而浪费资源降低效率。如果不进行调度，那么站点在未来的 24 小时内站点内车辆的数目走势，可以看出，站点的初始车辆数目为 20，5 小时内的波动值很小，但早晨 6 点之后，车辆数目陡然增加，这依据本文在第三章的分析可知，早高峰即将到来，而且该站点可能位于公司附近，所以会有大量车辆驶入，导致站点无法容纳。

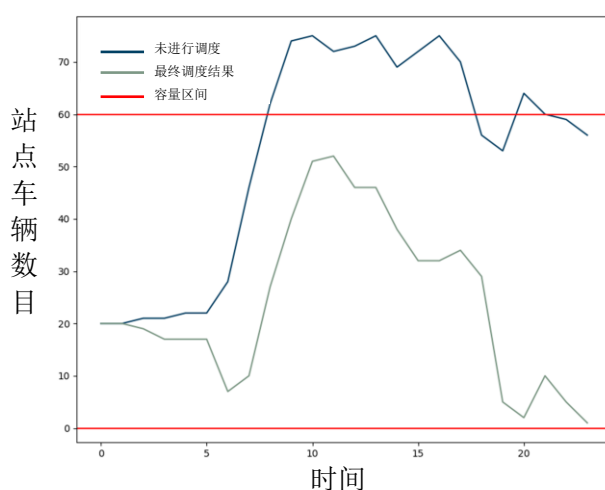


图 4.8 调度结果图

Fig.4.8 Rebalancing result graph

随着时间的增加，如果不加以调度，最终在 8 时左右，站点会没有停车位置提供，而这些无车位可停的用户只能选择将车辆停到旁边的站点，或者提供调度费让管理人员帮忙停放，但这势必影响用户的体验感，从而降低用户粘性。在后续的时间段内，车辆数目的走势有所波动，但总体波动幅度不大，这是因为在这段时间，只有少量的用户有用车需求，也仅有少量的用户到达该站点，而到了 17 时左右，由于晚高峰使站内车辆数目有较为明显的下降，但仍然在上限附近波动。因此，若不做调度，不但影响用户体验感，也不利于运营公司的管理与维护。利用 Q-Learning 方法对区域内的各站点车辆进行调度后，明显缓解了没有位置停放车辆的现象。本文进行了 400 次迭代实验，最终调度结果即训练 400 次之后的结果。可以看到，一天中所有时间段的站内自行车数目均在设定的范围内。这也证明了本文所提方法的有效性。

### ③ “动作”与“奖励”变化分析

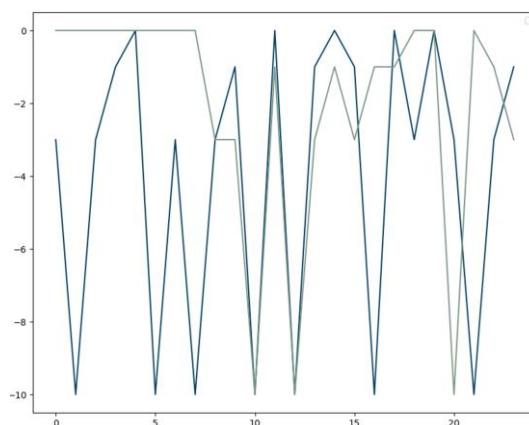


图 4.9 “动作”变化

Fig.4.9 The "action" changes

在实验中，本文总共进行了 4 轮训练，每轮迭代 100 次。图 4.10 中所示的是最后一轮训练 24 小时内的“动作”结果。在本文的模型里，动作也就是每次需要调度的车辆数目。若移动数目为“0”，即保持现有站内车辆数目不变；为“+3”，即表示此站点需要调入 3 辆自行车；为“-3”，即表示此站点需要调走 3 辆自行车。

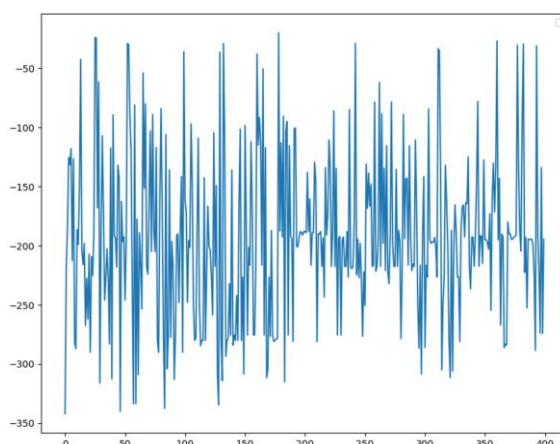


图 4.10 每轮迭代的“奖励”变化

Fig.4.10 The "reward" changes for each iteration

对于奖励，本文的设置：若任意时刻调度后站内车辆数目大于了上限 60 或者低于了下限 0，得到的奖励为“-30”；如果在一天结束，车站内的数目小于等于 60 或者大于等于 0，则给予“+20”奖励，否则给予“-20”奖励；若动作不为“0”，奖励设置为  $-0.5 * \sqrt{action}$ ， $action$  表示动作，即移动的车辆数目。而奖励的波动说明了所提算法在训练的时候会根据动作和状态的变化主动向者奖励较大的方向学习，从而得到最终的调度结果。从图 4.10 也可以发现，前 200 轮训练得到的奖励



波动范围比后 200 轮的波动范围总体要大一些, 在开始的时候, 强化学习并不知道何种状态下选择何种动作是最优的, 但是通过不断的试错与尝试, 所获得的奖励逐渐向优化的方向发展, 最终学习到最优的调度结果。

#### ④ 调度“成功率”分析

除了调度结果的有效性验证外, 调度“成功率”能够反映出自行车展站点的车辆库存情况是否在合理范围内, 即设定的容量上限与下限之间。因为有桩共享单车每个站点可容纳车辆数是固定的, 超过了容纳数量用户没有车位停放自行车, 与此同时, 无论如何调度, 站点中自行车的数量都不会小于 0。图 4.11 展示的是成功率, 成功率表示的是在一天的所有调度行为中, 调度后使得站内车辆处于 $[0, 60]$ 之间的占比。可以看出, 在第一轮的实验中, 调度成功率为 82%, 第二次调度是有所上升, 第三次大约 1% 的下降, 但是到最后一轮训练, 调度成功率已经达到了 91.5%。

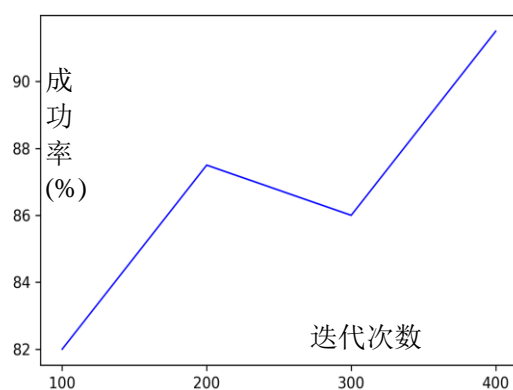


图 4.11 每轮迭代的调度成功率

Fig.4.11 Rebalancing success rate of each iteration

从上述的实验结果可以发现, 本文的方法可以有效的对站点内的车辆进行有效的调度, 并且使得站内的车辆数目稳定在一个区间内, 这样可以让用户在想要使用车辆时就有车可用, 想要停放车辆时, 也不会因站内车辆过多而要另找站点停放, 提升了用户的体验感, 从而能够在一定程度上提高了用户粘性。

## 4.4 本章小结

本章首先对所有的站点依据 K-Means 以及站点之间的交互情况进行了聚类, 随后考虑到现实中的管理运营情况, 将车辆的调度限制于每个簇中, 利用 Q-Learning 方法对站点间的车辆进行调度。在真实的数据集上进行了实验, 验证了所提方法的有效性。

## 5 单车流量预测与调度原型系统的设计与实现

### 5.1 应用背景

单车惠及众多用户,有效缓解了用户“最初一公里”和“最后一公里”出行难的问题。但同时也存在一些局限性,由于地区间用户行为的不一致性,使得部分站点“供大于求”,站内车辆长期闲置;而部分地区“供小于求”,用户有时无车可用,影响了用户的体验感,减少了用户粘性,不利于共享单车运营管理公司的发展。目前,国内的共享单车运营管理公司多采用的是根据站点分布划分区域,聘请单车调度管理人员对区域内车辆进行调度管理,而这些管理常常是依据经验来的。依据经验的调度往往比较低效,不利于公司进行高效的运营和管理。

针对上述问题,结合本文第三章提出的流量预测算法以及第四章提出的单车调度算法,本章设计了一个单车流量预测与调度管理原型系统,使用对象为共享单车公司的运营管理人员。管理人员通过使用该系统,能够通过查看用户的历史骑行记录获取用户信息(若涉及隐私则脱敏处理)、获取各个站点的单车使用情况、未来时刻各站点共享单车流量预测情况、以及相应的调度方案,辅助运营管理人员做出科学决策。

### 5.2 原型系统架构设计

本章所提出的共享单车流量预测及调度管理原型系统采用了三层架构,如图 5.1 所示。层次划分旨在实现“高内聚,低耦合”,采用“分而治之”的思想,将问题划分为若干个小问题,分开解决,易于控制、延展、以及资源分配。

表示层(UI层),即与用户进行交互的主要界面,用于接收输入并展示结果,是系统界面的展示,让单车运营管理人员明白系统所有具备的功能,能够进行的操作。在系统设计上,表示层是原型系统与共享单车运营管理人员交互的接口,应注重实用性与简洁性相结合,让共享单车运营管理人员操作方便。所有拥有权限的单车运营管理人员可以通过登录进入到该系统。

业务逻辑层(BLL层),它是数据访问层与表示层之间的桥梁,主要是对业务逻辑的实现,并且依据第三章所提的算法对共享单车流量进行预测并返回预测结果、根据第四章所提算法对共享单车进行调度并且返回调度建议等。

数据访问层(DAL层),是与数据库进行交互,将存储在数据库中的数据提交给业务层。在本文中,数据访问层主要存储共享单车流量预测所需要的用户骑行历史数据、站点信息数据、用户信息数据、历史气象数据等。

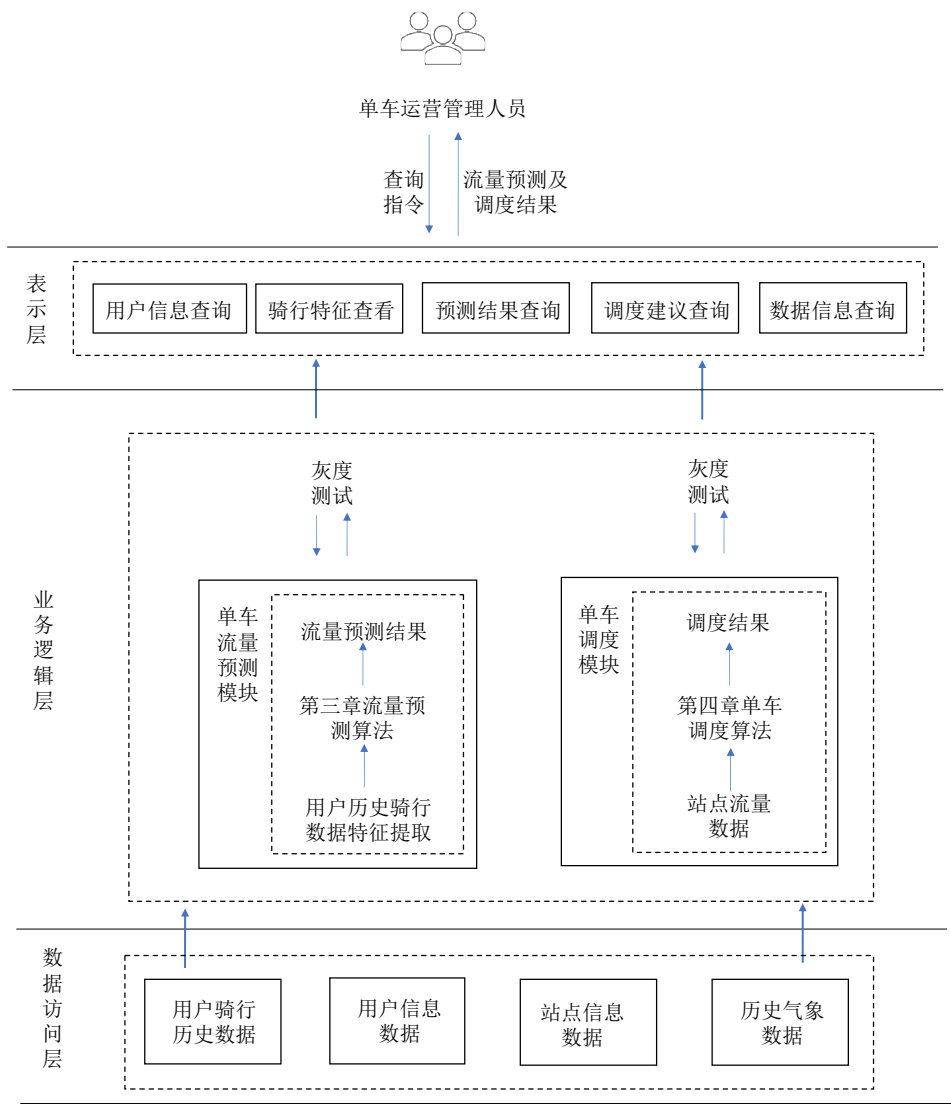


图 5.1 系统架构图

Fig.5.1 The system architecture diagram

5.3 系统数据库设计

本文所设计的系统选用了 Mysql 8.0.11 数据库。在单车流量预测及调度管理原型系统中，共设计了四个数据表：用户骑行信息表、历史气象数据表、站点信息表以及单车运营管理人员信息表，ER 图如 5.2 所示。单车运营管理人员信息表主要存储拥有系统使用权的单车运营公司管理人员的基本信息，包括管理者姓名、管理者工号、管理者职级、管理者密码等。用户骑行信息表主要涉及使用单车使用用户的基本信息，包括用户出生年份、是否购买骑行会员卡、用户性别、骑行起始时间、骑行时长、骑行起始站点。历史气象数据表主要涵盖该地区的观测气象数据，包括观测时间（具体日期及时间段）、温度、天气类别（晴朗、小雨、

暴雨、降雪、微风)、海平面压强、风速、湿度等。站点信息表为各个自行车站点的基本信息,包括站点的ID、站点经度、站点纬度。

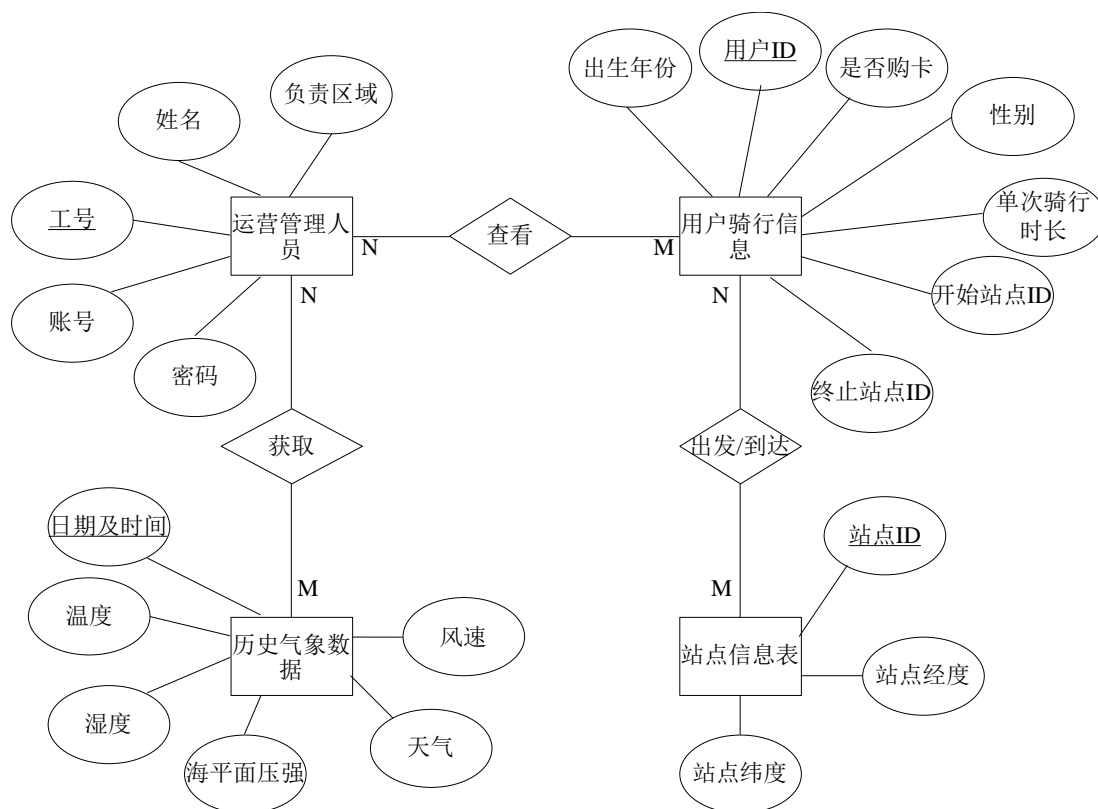


图 5.2 系统 E-R 图

Fig.5.2 System E-R diagram

## 5.4 原型系统功能模块设计及分析

### 5.4.1 原型系统功能模块设计

本系统主要包括四个模块,分别是系统登录模块、数据查询模块、单车流量预测模块和单车调度模块。系统功能模块结构图如图 5.3 所示。

各个模块的具体功能如下:

① 系统登录模块。包含单车运维管理人员登录,以及个人信息的管理。此模块为单车流量预测及调度系统的入口,管理人员可由此入口登录系统。

② 数据查询模块。登录后的管理人员可对阶段性的历史数据进行查询,会生成数据报表,主要包含内容有:数据集总体概况、数据字段名、各数据分别有多少条、有无缺失值以及每个字段下数据的具体情况。同时,拥有高级别权限的管理人员也能够对生成的数据报表进行下载。此外,该模块以用户的历史骑行数据为依据,生成用户画像,反映用户群体在骑行自行车的一系列行为反应出的特征,如周期性、趋势性、邻近性等,以及可根据用户的骑行轨迹,对功能区进行判断,

辅助单车运营管理人员对后续站点是否增加或者减少做出科学的判断与决策。

③ 单车流量预测与调度模块。该模块位于系统的业务逻辑层，应用了论文第三章以及第四章所提出的算法，根据用户的历史骑行数据，对各站点的单车流量进行预测，并根据预测值合理调整各站点内停车桩的数目，进行科学规划。此外，给出调度建议，告知调度人员共享单车站点需要调度的车辆数目，为运营管理人员提供科学的决策依据，更有效合理的对站点内的车辆进行管理。

④ 单车报修模块。运营管理人员除了负责共享单车的调度工作，在实际的操作过程中，从其它站点调度的车辆可能是有损坏的，这也会影响用户体验。所以，调度人员也需要对调度过程中发现的有损车辆进行报修。

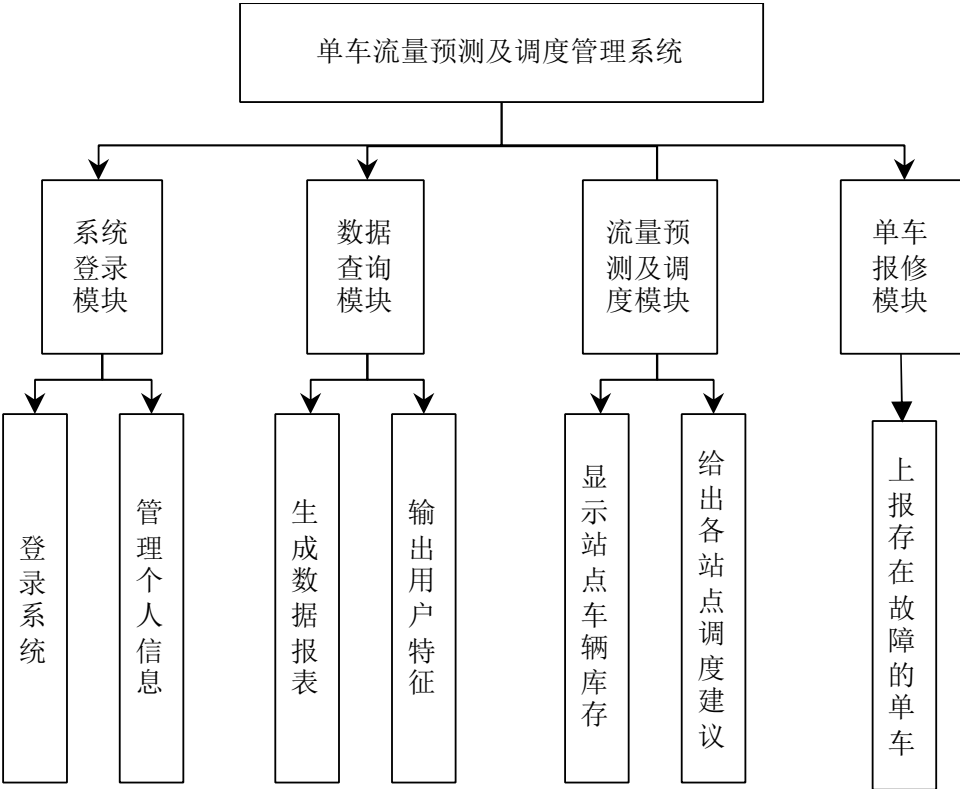


图 5.3 系统功能模块结构图

Fig.5.3 The structural diagram of system modules

5.4.2 系统实现及效果分析

在详细的设计之后，本文对原型系统进行了实现，该系统提供给共享单车营运公司指定的调度管理人员。接下来将通过对共享单车流量调度管理原型系统的展示来进一步说明本文算法思想。



图 5.4 系统登录界面

Fig.5.4 The interface of system login

图 5.4 展示的是共享单车流量预测以及调度管理原型系统，共享单车调度管理人员可以通过输入用户名和密码登录到该系统。本系统只能通过键入用户名和密码，登录成功后使用。因为每个共享单车调度管理人员所负责的区域是不同，只拥有自己负责区域的调度管理权限。若密码遗失，可通过忘记密码进行密码的找回。登录成功后，即可进入到本系统的主页面，系统主页介绍了共享单车流量预测以及调度的应用背景及其作用。管理人员可以查询模块查看到历史用户骑行数据的数据，但这些并不是原始数据，是经由公司筛选剔除掉关键个人信息后的数据。同时也能了解各个站点一天即一周内的流量预测情况，流量预测模块是经由第三章所提基于 LSTM 和注意力机制的预测算法计算得到的。此外，也能够查看单车调度模块，该模块会根据第四章所提出的基于 Q-Learning 算法的单车调度算法计算得到的。通过流量预测模块，管理人员能够了解到每个站点的流量走势，知晓该站点的车辆使用情况，单车调度模块能够帮助管理人员决策所需调度的车辆以及区域内的调度路线。

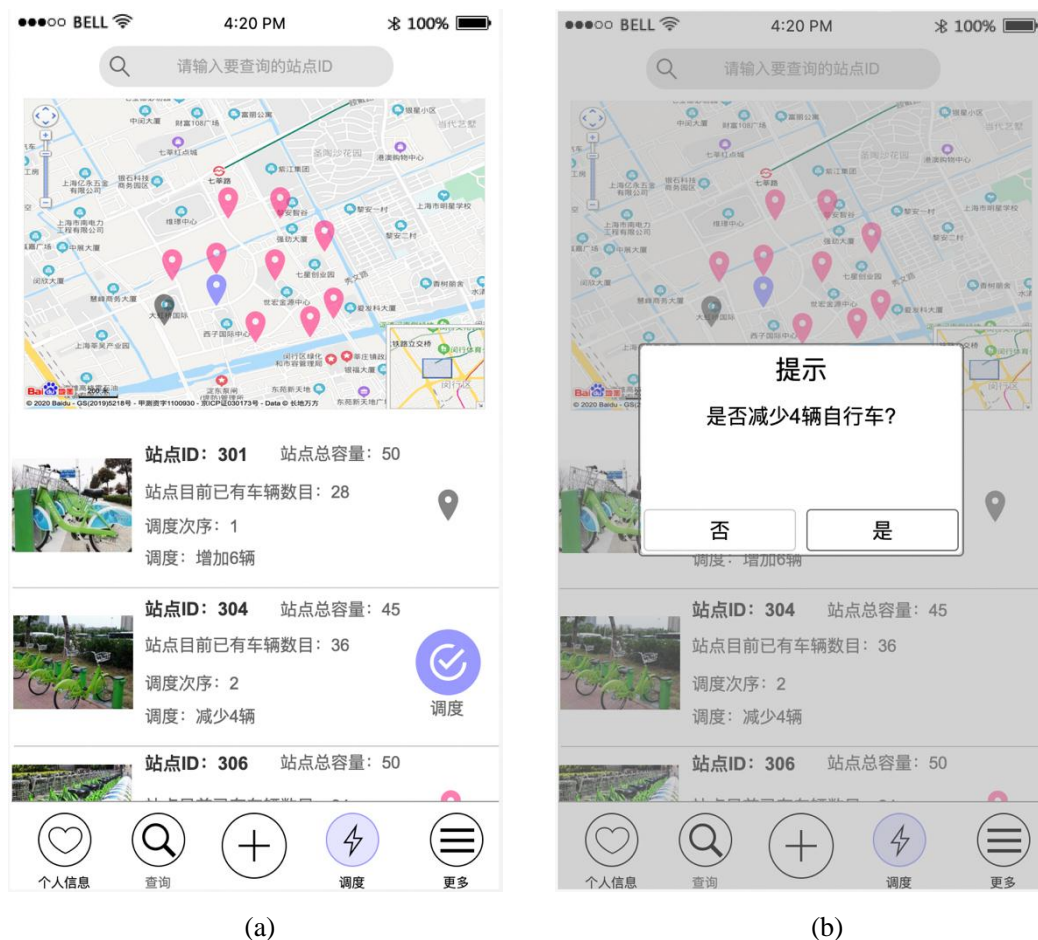


图 5.5 共享单车调度界面

Fig.5.5 The interface of shared bike rebalancing

共享单车调度界面如图 5.5 所示, 共享单车调度管理人员可以在本模块接收到系统返回的调度指令。在图 5.5(a)模块的顶端, 设置了一个查询窗口, 共享单车调度管理人员可在此处查询任意站点的站点总容量、目前拥有共享单车数量、调度次序以及系统所给出的调度指令。在该查询窗口的下方, 给出了当前登录本系统的共享单车管理人员的管理范围, 并且将每个共享单车站点在地图上标示了出来。图中灰色的图标表示共享单车调度管理人员已经完成了对该站点的车辆站点调度, 紫色的图标表示当前共享单车调度管理人员正处于的站点, 红色的图标表示该管理人员负责区域内的其他未调度的站点。地图是调用的百度地图, 管理人员可以通过放大缩小以及获取目前所处位置的坐标获取到更加精确的地理位置信息。地图的下方是各个站点的情况, 包含每个站点的图示、当前站点的 ID、站点总共可以容纳的共享单车数量、站点目前已经停放的共享单车数量、该站点在本时段调度活动中所处的调度次序以及该站点需要调度的车辆。当共享单车调度管理人员到达当前调度站点后, 可根据系统建议进行车辆的放置, 在完成车辆的放置或者



调离之后，选择图 5.5(a)中站点 304 右侧的紫色调度图标，点击该图标，跳转到图 5.5(b)所示界面，确定放置或者调离的车辆数目，完成本次调度工作。车辆调度信息的反馈到后台实时的更新系统中去，重新计算当前站点的库存容量，通过第三章的共享单车流量预测算法进行下一个时刻的单车流量预测以及下一个时刻的调度结果计算。



图 5.6 共享单车报修界面

Fig.5.6 Repair interface for Shared bikes

在完成本时段的调度工作之后，共享单车调度管理人员还需要检查共享单车车辆是否完好，若有损坏的车辆需要及时上报，将故障信息上报给运营公司相应管理系统，公司会安排该区域的维修管理人员前来对损坏的自行车进行维修。通过本界面最上端的扫描功能对共享单车二维码进行扫描，或者手动输入车辆编号，接着在下方选择车辆具体的损坏部位，若有未涉及到的其他损坏部位，可通过拍照以及留言的方式告知。在完成共享单车的报修之后，共享单车运营管理人员须返回到调度界面，在调度确认处重新填写该站点的调度信息，若系统建议增加 6



辆共享单车，实际所调度的 6 辆自行车中有一辆损坏。此时可以选择多调度一辆车，实际增加 7 辆共享单车，信息填写时填入“实际增加 7 辆，有损 1 辆”。

通过整个共享单车流量预测以及调度管理原型系统，可以从用户的历史骑行数据中获取特征，结合单车调度管理人员返回的建议，以及站点周围实际情况对后续算法模型进行升级调整。此系统能够让单车运营管理人员在调度车辆的时候有所科学依据，而非仅凭经验进行站点间的车辆调度，从而使各个自行车站点的单车数量趋于合理，避免“无车可用”以及“无位可停”的情况。

## 5.5 本章小结

本章节将论文中第三章与第四章所提及的算法融入到了共享单车流量预测与调度管理原型系统中，并进行了展示。

## 6 结论与展望

### 6.1 主要结论

作为公共交通的重要补充形式,单车在人们的日常出行中占据着重要的位置。在使用过程中,积累了大量的用户骑行数据,这些数据具有重要的价值。通过对数据进行分析挖掘,能够了解用户的出行行为模式,进而辅助管理者做出科学的决策,助力“智慧城市”的建设与发展。但由于各个站点使用情况的差异较大,导致了用车需求的不均衡,存在某些站点“供不应求”,有些站点“供大于求”的现象。而在实际过程中,对于单车车辆的移动往往是靠区域单车运营管理人员自身经验,这种方式存在很大的局限性。

因此,本文调研了国内外各领域中有共享单车流量预测以及调度问题,了解到目前领域中存在的问题和进步空间,针对这些不足,本文从对花旗自行车的用户历史骑行数据的详尽分析入手,结合已有的气象数据,对用户骑行行为进行了分析,根据分析结果构造新特征。据此,提出了一种基于 LSTM 和注意力机制的单车流量预测算法,提高了模型预测精度,并且利用强化学习中的 Q-learning 算法实现了单车调度。最后以提出的两类算法为基础,设计并实现了一个共享单车流量预测以及调度管理原型系统。论文主要完成的工作如下:

① 调查分析。对本课题相关的研究背景和意义进行了调研,根据国内外对单车流量预测的研究现状,归纳了现有的预测方法,总结了其中仍存在的不足之处,并提出本文的主要研究内容。

② 提出了一种结合骑行上下文信息,并基于 LSTM 网络和注意力机制的共享单车流量预测算法。该算法利用 LSTM 挖掘流量数据之间的时序特征,同时结合注意力机制对与输出时刻最为相关的输入特征进行提取,从而提高了共享单车流量预测的精度。此外,在第三章也对该算法整体框架做了描述,对算法核心内容做了详细解释,通过实验验证了本章所提算法的有效性以及特征筛选的合理性。

③ 提出了基于强化学习的共享单车调度策略。利用 Q-learning 算法计算出最优的价值函数,从而找到在当前站点库存车辆数目下需要调度的车辆数。并且,在真实的数据集上进行了验证,实验证明所提算法能够合理有效地对各站点之间的共享单车进行调度,平衡区域内各站点之间的使用需求。

④ 共享单车流量预测及调度管理原型系统设计与实现。基于第三章和第四章所提的算法,设计了共享单车流量预测及调度管理原型系统。

## 6.2 后续研究工作的展望

本文主要在单车流量预测和单车调度领域做了研究，通过 LSTM 和注意力机制对单车流量进行预测，通过 Q-learning 算法对单车进行调度。但是，本文所提算法仍然存在一些值得改进的地方，论文接下来的工作将重点包括以下几个方面。

① 考虑图卷积方法。本文详尽的分析了用户历史骑行数据，结合气象数据构建了上下文特征，结合时序特征对单车流量进行了预测。但是，站点分散在各个不同的地方，空间特性明显，因此，可以考虑利用图卷积模型对空间特征进行有针对性的处理，这样结合三部分特征，进一步提高预测精度。

② 在构建共享单车调度策略时，可以更多考虑到其它因素，如路况、调度过程中额外耗费的成本等。

③ 完善系统。本文所设计的单车流量预测及调度管理原型系统较为简单，相对于实际可投入使用的预测和调度系统而言，还有很多可完善的地方。在后续工作中，本文将遵循软件工程开发流程，完善本系统功能，提高系统的易用性和可靠性。

## 参考文献

- [1] O'Mahony E , Shmoys D B. Data Analysis and Optimization for (Citi)Bike Sharing[C]. Twenty-ninth Aai Conference on Artificial Intelligence, AAAI, 2015, 687-694.
- [2] Chemla D, Meunier F, Calvo R W, et al. Bike sharing systems: Solving the static rebalancing problem[J]. Discrete Optimization, 2013, 10(2), 120-146.
- [3] Yang Z, Hu J, Shu Y, et al. Mobility Modeling and Prediction in Bike-Sharing Systems[C]. international conference on mobile systems, applications, and services. 2016, 165-178.
- [4] Wang B, Kim I. Short-term prediction for bike-sharing service using machine learning[J]. Transportation research procedia, 2018, 171-178.
- [5] Tang Y, Pan H, Fei Y, et al. Research on Users' Frequency of Ride in Shanghai Minhang Bike-sharing System[J]. Transportation research procedia, 2017, 4979-4987.
- [6] Singhvi D, Singhvi S, Frazier P I, et al. Predicting Bike Usage for New York City's Bike Sharing System[C]. national conference on artificial intelligence, 2015, 110-114.
- [7] A. Bargar, A. Gupta, S. Gupta, and D. Ma, Interactive Visual Analytics for Multi-City Bikeshare Data Analysis[C]. 3rd Int. Work. Urban Comput, 2014, 1-5.
- [8] Dellamico M, Iori M, Novellani S, et al. The Bike sharing Rebalancing Problem with Stochastic Demands[J]. Transportation Research Part B-methodological, 2018, 362-380.
- [9] Vogel P, Greiser T, Mattfeld D C, et al. Understanding Bike-Sharing Systems using Data Mining: Exploring Activity Patterns[J]. Procedia - Social and Behavioral Sciences, 2011, 514-523.
- [10] Yan Y, Tao Y, Xu J, et al. Visual analytics of bike-sharing data based on tensor factorization[J]. Journal of Visualization, 2018, 21(3), 495-509.
- [11] Lihua N, Xiaorong C, Qian H, et al. ARIMA model for traffic flow prediction based on wavelet analysis[C]. international conference on information science and engineering, 2010, 1028-1031.
- [12] Zhang N, Zhang Y, Lu H, et al. Seasonal Autoregressive Integrated Moving Average and Support Vector Machine Models: Prediction of Short-Term Traffic Flow on Freeways[J]. Transportation Research Record, 2011, 22(15), 85-92.
- [13] Ahn J Y, Ko E, Kim E Y, et al. Predicting Spatiotemporal Traffic Flow Based on Support Vector Regression and Bayesian Classifier[C]. international conference on big data and cloud computing, 2015, 125-130.
- [14] 徐健锋, 汤涛, 严军峰, 基于多机器学习竞争策略的短时交通流预测[J]. 交通运输系统工程与信息, 2016, 32-46.

- [15] 辛壮, 邓烜堃, 万良, 丁红卫, 基于深度学习的交通流量预测研究[J]. 计算机工程与应用, 2019, 55(02), 234-241
- [16] 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.
- [17] Zhang J, Zheng Y, Qi D, et al. Deep Spatio-Temporal Residual Networks for Citywide Crowd Flows Prediction[C]. national conference on artificial intelligence, 2016, 1655-1661.
- [18] Benchimol M, Benchimol P, Chappert B, et al. Balacing the stations of a self service bike hire system [J]. Rairo-operations Research, 2011, 45(1), 37-61.
- [19] Wang Y, Yin H, Chen H, et al. Origin-Destination Matrix Prediction via Graph Convolution: a New Perspective of Passenger Demand Modeling[C]. knowledge discovery and data mining, 2019, 1227-1235.
- [20] Liu J, Sun L, Chen W, et al. Rebalancing Bike Sharing Systems: A Multi-source Data Smart Optimization[C]. knowledge discovery and data mining, 2016, 1005-1014.
- [21] Papazek P, Kloimullner C, Hu B, et al. Balancing Bicycle Sharing Systems: An Analysis of Path Relinking and Recombination within a GRASP Hybrid[C]. parallel problem solving from nature, 2014, 792-801.
- [22] Lahoorpoor B, Farooqi H, Sadeghiniaraki A, et al. Spatial Cluster-Based Model for Static Rebalancing Bike Sharing Problem[J]. Sustainability, 2019, 11(11), 3205-3226.
- [23] Li Y, Zheng Y, Yang Q, et al. Dynamic Bike Reposition: A Spatio-Temporal Reinforcement Learning Approach[C]. knowledge discovery and data mining, 2018, 1724-1733.
- [24] Pan L, Cai Q, Fang Z, et al. A Deep Reinforcement Learning Framework for Rebalancing Dockless Bike Sharing Systems[C]. national conference on artificial intelligence, 2019, 33(01), 1393-1400.
- [25] Wang S, He T, Zhang D, et al. Towards Efficient Sharing: A Usage Balancing Mechanism for Bike Sharing Systems[C]. the web conference, 2019, 2011-2021.
- [26] Faghihimani A, Hampshire R C, Marla L, et al. An empirical analysis of bike sharing usage and rebalancing: Evidence from Barcelona and Seville[J]. Transportation Research Part A-policy and Practice, 2017, 177-191.
- [27] Hopfield J J. Neural networks and physical systems with emergent collective computational abilities[J]. Proceedings of the National Academy of Sciences of the United States of America, 1982, 79(8), 2554-2558.
- [28] Jordan M I. Serial Order: A Parallel Distributed Processing Approach[J]. Advances in psychology, 1997, 471-495.
- [29] Elman J L. Finding Structure in Time[J]. Cognitive Science, 1990, 14(2), 179-211.

- [30] Hochreiter S, Schmidhuber J. Long short-term memory[J]. *Neural Computation*, 1997, 9(8), 1735-1780.
- [31] Bahdanau D, Cho K, Bengio Y, et al. Neural Machine Translation by Jointly Learning to Align and Translate[J]. *Computation and Language*, 2014.
- [32] Itti L, Koch C. Computational modelling of visual attention[J]. *Nature Reviews Neuroscience*, 2001, 2(3), 194-203.
- [33] Sutton R S, Barto A G. Reinforcement Learning: An Introduction[C]. *neural information processing systems*, 1999, 342-360.
- [34] M. E. Harmon, W. Aacf, and S. S. Harmon, Reinforcement Learning: A Tutorial Scope of Tutorial[J], *Synth. Lect. Artif. Intell. Mach. Learn*, 2010, 81-89
- [35] Watkins C, Dayan P. Technical Note: Q-Learning[J]. *Machine Learning*, 1992, 8(3), 279-292.
- [36] Sprague N, Ballard D H. Multiple-goal reinforcement learning with modular Sarsa(O)[C]. *international joint conference on artificial intelligence*, 2003, 1445-1447.
- [37] Li H, Wei T, Ren A, et al. Deep reinforcement learning: Framework, applications, and embedded implementations: Invited paper[C]. *international conference on computer aided design*, 2017, 847-854.
- [38] Schulman J, Abbeel P, Chen X, et al. Equivalence Between Policy Gradients and Soft Q-Learning[J]. *Learning*, 2017.
- [39] Bertsekas D P, White C C. Dynamic Programming and Stochastic Control[J]. *IEEE Transactions on Systems, Man, and Cybernetics*, 1977, 7(10), 758-759.
- [40] Cho K, Van Merriënboer B, Gulcehre C, et al. Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation[J]. *Computation and Language*, 2014, 1724-1734.
- [41] Cho K, Van Merriënboer B, Bahdanau D, et al. On the Properties of Neural Machine Translation: Encoder-Decoder Approaches[J]. *Computation and Language*, 2014, 103-111.
- [42] Xie P, Li T, Liu J, et al. Urban flow prediction from spatiotemporal data using machine learning: A survey[J]. *Information Fusion*, 2020, 1-12.
- [43] 窦慧丽, 刘好德, 吴志周, et al. 基于小波分析和 ARIMA 模型的交通流预测方法[J]. *同济大学学报 (自然科学版)*, 2009, 37(4), 486-489.
- [44] Fu R, Zhang Z, Li L, et al. Using LSTM and GRU neural network methods for traffic flow prediction[C]. *youth academic annual conference of chinese association of automation*, 2016, 324-328.
- [45] Liu Q, Wu S, Wang L, et al. Predicting the Next Location: A Recurrent Model with Spatial and Temporal Contexts[C]. *AAAI*, 2016, 194-200.

## 附 录

### A. 作者在攻读硕士学位期间发表的论文目录

- [1] J. Zeng, **X. He**, H. Tang, and J. Wen, Predicting the next location: A self-attention and recurrent neural network model with temporal context[J]. Trans. Emerg. Telecommun. Technol, 2020. (SCI)
- [2] J. Zeng, **X. He**, H. Tang, and J. Wen, A Next Location Predicting Approach Based on a Recurrent Neural Network and Self-attention[C]. Collaborative Computing: Networking, Applications and Worksharing, 2019, 309-322. (CCF-C 类)
- [3] J Zeng, **X. He**, Y. Li, J. Wen, and W. Zhou, A point of interest recommendation method using user similarity[J]. Web Intell, 2018, 16(02), 105-112. (CCF-C 类)
- [4] J. Zeng, **X. He**, Y. Wu, and S. Hirokawa, User Behavior Analysis of Location-Based Social Network[C]. IIAI-AAI, 2018, 21-25. (EI)
- [5] J. Zeng, F. Li, **X. He**, and J. Wen, Fused Collaborative Filtering With User Preference, Geographical and Social Influence for Point of Interest Recommendation[J]. Int. J. Web Serv. Res, 2019, 16(04), 40-52. (SCI)
- [6] J. Zeng, H. Tang, Y. Li, and **X. He**, A deep learning model based on sparse matrix for point-of-interest recommendation[C]. Proceedings of the International Conference on Software Engineering and Knowledge Engineering, 2019, 379-492. (CCF-C 类)

### B. 作者在攻读硕士学位期间取得的科研成果目录

- [1] 发明专利：曾骏，**何欣**，唐浩然，文俊浩，基于循环神经网络和注意力机制的下一个地点预测方法（专利申请号：201910872506.4）
- [2] 发明专利：曾骏、李烽、**何欣**、文俊浩、柳玲，基于用户偏好、社交信誉度和地理位置的兴趣点推荐方法（专利号：ZL 2017 1 0899405.7）
- [3] 发明专利：曾骏、李英华、唐浩然、**何欣**，一种基于用户签到稀疏矩阵的深度学习兴趣点推荐方法（专利申请号：2019103088661.1）

## C. 学位论文数据集

关键词		密级		中图分类号	
流量预测；单车调度；LSTM； 注意力机制；强化学习		公开		TP	
学位授予单位名称	学位授予单位代码	学位类别		学位级别	
重庆大学	10611	学术学位		硕士	
论文题名		并列题名		论文语种	
共享单车流量预测与调度方 法的研究及应用		无		中文	
作者姓名	何欣	学号		201724021006	
培养单位名称		培养单位代码			
重庆大学		10611			
学科专业	研究方向	学制		学位授予年	
软件工程	流量预测	3 年制		2020 年	
论文提交日期	2020 年 6 月	论文总页数		65 页	
导师姓名	曾骏	职称		副教授	
答辩委员会主席		符云清 教授			
电子版论文提交格式					
文本（√）    图像（）    视频（）    音频（）    多媒体（）    其他（）					



## 致 谢

时间总是匆匆，三年一晃而过。

在这三年的时光里，我也收获了很多。在这里我要对帮助和支持我的老师们、同学们、家人以及朋友们表达我最真挚的谢意。

首先，我要真诚地感谢我的导师曾骏老师。曾老师在学术研究和论文写作方面给予了我很多的帮助，从进入实验室后的第一篇论文到毕业论文，曾老师都十分耐心，一字一句地帮我进行修改，悉心指导。此外，曾老师还教会了我很多为人处事的道理，让我受益匪浅。

其次，我要感谢我们团队的各位老师和同学们。正是有各位老师的带领，我们才能够在讨论班中“头脑风暴”，交流和传递学术知识，分享科研成果。非常感谢各位老师的辛苦付出。此外，我还要感谢李烽师兄、李英华师兄，无论是科研、学习还是到实习、找工作，两位师兄都给了我非常细致的帮助；还有与我同级的陈开润同学、李颖巧同学和申汝平同学，平时我们会一起交流讨论学术，这也为我研究生的科研工作提供了很大的帮助。

再次，我要衷心感谢重庆大学大数据与软件学院为我提供了优秀的科研平台和宝贵的学习资源，愿学院越来越好。

此外，我还要感谢我的父母和我的朋友们，他们无论在何时何地都给予了我最无私的理解与包容，科研受阻的时候，更能给我以鼓励。

乘风破浪会有时，直挂云帆济沧海。也感谢生活和学习中遇到的困难和挫折，锻造了更好的自己。

最后，我也要感谢对论文进行评阅以及答辩委员会的专家们。

何 欣

二〇一〇年四月 于重庆