

doi : 10. 3969/j. issn. 1005 - 3026. 2016. 09. 005

# 基于运营系统的出租车出行需求短时预测模型

林永杰<sup>1</sup>, 邹 难<sup>2</sup>

( 1. 美国西北大学 土木与环境工程系, 美国 埃文斯敦 60208 ;

2. 山东大学 控制科学与工程学院, 山东 济南 250100 )

**摘 要 :** 出租车系统作为城市交通运输系统的重要组成部分,其宏观规划和调度管理的合理性决定了出租车服务质量.本文主要研究出租车乘客出行需求估计及预测,为出租车规划和实时调度提供数据支持.首先,分析了出租车定位系统和计费系统,改进了传统出租车需求网格划分方法,考虑了地形、建筑群和道路网络特征,保持了网格自身出行特性的完整性.其次,根据实时收集的出租车数据,建立了易于计算的出行需求估计方法.最后,以实际数据为基础,对影响短时出行量的主要变量进行了相关性分析,提出了基于人工神经网络的短时需求预测模型.根据相关性分析确定了模型结构.以实际获取的出租车数据为例,验证了提出的需求估计和预测模型.结果证明,相比于传统自回归滑动平均模型,提出的人工神经网络模型其平均绝对误差百分比提高了32%.此外,人工神经网络模型的绝对误差百分比超过50%的概率低于10%,而自回归滑动平均模型高达23%.

**关 键 词 :** 出租车运营;乘客需求估计;网格;短时预测;人工神经网络

**中图分类号 :** U 491.5<sup>+</sup>4      **文献标志码 :** A      **文章编号 :** 1005 - 3026( 2016 )09 - 1235 - 06

## Short-Term Prediction Model of Taxi Passenger Demand Based on Operation Systems

LIN Yong-jie<sup>1</sup>, ZOU Nan<sup>2</sup>

( 1. Department of Civil and Environmental Engineering, Northwestern University, Evanston IL 60208, USA ;

2. School of Control Science and Engineering, Shandong University, Jinan 250100, China. Corresponding author : LIN Yong-jie, E-mail : yjlinsdu@gmail. com )

**Abstract :** As a critical component of urban transportation systems, the service level of taxis is significantly affected by taxi planning and dispatching. The objective of this study is to estimate and predict taxi passenger demand to support for planning and dispatching. Firstly, the data collection of the in-vehicle taxi GPS system and fare collection system are analyzed in the paper. In terms of data analysis, the traditional grid partition of taxi demand is improved by adding other factors, such as topography, buildings, and road network. The developed partition preserves the completeness of passenger demand in a grid. And then, an easy-to-use estimation method of grid-based demand is presented by the usage of real-time taxi GPS system and fare collection system. Finally, an artificial neural network ( ANN ) model is developed to predict short-term taxi demand. The structure of the ANN model is designed based on the functional characteristics of the input-output pairing correlation. Taking the field data from taxi operation system as an example, the performance of proposed estimation and prediction models is evaluated and validated. The results reveal that the proposed ANN prediction model significantly outperforms the existing autoregression-moving-average ( ARMA ) model in terms of the reduction of 32% on average absolute percentage error. Moreover, the probability of absolute percentage error greater than 50% for both ANN and ARMA models is 10% and 23%, respectively.

**Key words :** taxi operations ; passenger demand estimation model ; grid ; short-term prediction ; artificial neural network( ANN )

收稿日期 : 2015 - 04 - 28

基金项目 : 国家科技支撑计划项目( 2014BAG03B04 ) ; 中央高校基本科研业务费专项资金资助项目( 2014JC036 ) .

作者简介 : 林永杰( 1987 - ) 男, 山东潍坊人, 美国西北大学博士后研究人员 ; 邹 难( 1976 - ) 男, 北京人, 山东大学教授, 博士生导师.

作为城市交通运输系统中的重要组成部分,出租车是介于公交车与社会车辆之间的一种便捷服务工具.一方面,与公交车相比,出租车可提供舒适、快捷、个性化的服务,避免公交车拥挤、延误大和无法提供门到门服务等问题;另一方面,与私家车相比,出租车能提供同等舒适度、出行时间短、门到门等优质服务,还可避免私家车出行成本高(如车辆购置费、保养与维护费)、换乘繁琐、停车成本高(停车位售价高、停车难、停车费高)等问题.此外,随着出租车电召和合乘服务等现代化运输理念的推广与实施,出租车出行质量稳步提升、成本却能逐渐下降.从近年来出租车规模扩张和出租车运营系统大规模建设也可以看出,出租车对乘客的吸引越来越高.

受乘客出行需求的波动性和随机性内在特征,以及出租车载客的盲目性和随意性影响,乘客与出租车服务之间总存在着较大的供需矛盾,如空载率高、候车时间长、打车难.而实际中,若司机能预先获悉短时间内乘客需求,则可实现出租车合理调度与乘客服务的双赢.尤其近年来,互联网和信息技术的普及,大部分城市出租车均已安装了车载 GPS 系统和计价系统,利用该系统采集的车辆位置和载客状态等实时数据可以估计和预测出租车乘客需求,为出租车规划和调度管理提供可靠的数据支持<sup>[1]</sup>.

近十年来,出租车运营与管理系统得到大规模普及,主要研究集中在出租车宏观规划与管理政策<sup>[2]</sup>、出租车调度与管理<sup>[3-4]</sup>、出租车出行特性分析<sup>[5-7]</sup>、出租车服务评价<sup>[8-9]</sup>等方面,而专门针对出租车出行需求预测的研究相对较少<sup>[1]</sup>.根据文献综述,已有出租车需求预测模型主要分为中短期预测和长期预测.长期预测用于出租车宏观管理、规划和政策分析等方面<sup>[10]</sup>,而中短期预测主要目的是服务于出租车实时调度<sup>[1,11-13]</sup>.本文主要研究出租车乘客中短期需求预测模型.2010 年,Chang 等提出了聚类方法估计当前单个车辆所在位置有乘客需求的概率<sup>[11]</sup>,考虑了车辆位置、星期、小时和天气因素.2011 年,Deng 等以上海市实际数据为基础,研究了上海市出租车服务的时空特点<sup>[12]</sup>.2012 年,Li 等分析了出租车行驶的目的性<sup>[13]</sup>,估计了每小时上车热点区域,并利用自回归滑动平均模型预测了每隔 20 min 乘客需求的变化.2013 年,Moreira - Matias 等分析了泊松分布和自回归滑动平均模型在短时出租车乘客需求预测方面的应用<sup>[1]</sup>,利用葡萄牙波尔图市 441 辆出租车实际采集的数据验证了模型的准确

性,以 30 min 为间隔的平均预测误差约为 24%.

根据上述研究成果,现有出租车短时需求预测模型的准确性有待进一步提高,其考虑的影响因素主要是前几个时刻的需求.然而,在实际出租车服务中,尤其是需求高峰期,乘客选择出租车的意愿受所在位置类型、时间段、星期和历史特性等多方面因素的影响.因此,本文将提出一个新型的短时出租车出行需求实时预测系统,包括出行需求估计和预测两部分.

## 1 出租车出行需求估计

出租车出行需求预测的根本在于历史出行需求,而当前国内外大部分出租车均已配备了车载 GPS 定位系统和计价收费系统,可以提供较为准确的历史出租车出行需求估计.

### 1.1 出租车运营与管理系统

出租车运营与管理系统主要以 GPS 定位系统和车载计价收费系统为主,以济南市出租车运营系统为例,两类系统的主要特点如下:

1) GPS 定位系统能够将车辆设备编号、经纬度、瞬时速度、车头方向、载客状态、时间戳,以约 15 s 固定间隔上传至控制中心.由于 GPS 数据大多采用易于管理的集中存储模式,因此中心数据存储压力很大,为了尽可能减少存储量,通常采集的信息全部以整数形式保存在数据库中,并采用人工方式定期备份.

2) 车载计价收费系统主要收集每次乘车交易信息,包括设备编号、上/下车时间、上/下车经纬度、行驶里程、等候时间、金额、空驶里程、收费类型(单程、夜间或往返)和时间戳等.由于收费系统记录的信息是以出租车驾驶员操作计价器为准,因此,数据通常在乘客下车时、载客状态由“载客”变为“空车或停运”,计价器打印乘客发票后,再上传至控制中心,数据可能存在较大延迟.如果驾驶员忘记或没有计时,则无法记录乘客出行信息.

根据上述两类系统记录的数据可以看出:GPS 系统能够记录车辆的全程行驶轨迹、乘客出行起讫点(origin and destination, OD),而计价收费系统能够记录乘客出行 OD、出行时间、出行距离和等候时间等.而且,两类系统获取的乘客出行 OD 均以 GPS 设备定位的经纬度为准,因此可通过其中任意系统获取.但计价收费系统获取的信息会滞后于 GPS 系统.

## 1.2 出行需求估计

通过前述数据分析发现,车载 GPS 系统能够实时获取不同区域乘客上车需求.因此,本文基于 GPS 定位系统建立城市不同区域的出租车乘客出行需求估计,能为短时出行需求预测和实时调度提供依据.估计算法包括区域划分和出行需求估计两部分.

### 1.2.1 区域划分

区域划分是将一个城市的出租车出行需求进行片区划分,可以按照现有交通小区划分,也可以按照网格划分. Li 等以广州为例,将市区按照  $10\text{ m} \times 10\text{ m}$  的网格进行划分,计算出乘客上/下车点比较频繁的网格为热点<sup>[13]</sup>.但在实际中,由于部分网格受地形、区域大小等因素影响,固定  $100\text{ m}^2$  的网格划分方式可能将一个办公楼、一个路段划分成了多个不完整的子块,从而破坏了该网格乘客出行规律.当网格过小时单位小时内的出行需求也很低,降低其实用性.因此,本文提出结合城市实际的用地开发,考虑地形、建筑群和道路网络特征进行网格划分,具体步骤如下:

首先,网格初始化.以  $50\text{ m} \times 50\text{ m}$  网格将城市划分为  $N$  个子网格,第  $i$  个网格记为  $g_i(\text{lon}_{\max}, \text{lon}_{\min}, \text{lat}_{\max}, \text{lat}_{\min})$ ,其中  $\text{lon}_{\max}, \text{lon}_{\min}, \text{lat}_{\max}, \text{lat}_{\min}$  分别表示网格内点的最大经度、最小经度、最大纬度和最小纬度,即网格的边界线.

其次,网格集聚.对于任意两个空间相邻的网格  $g_1(\text{lon}_{\max1}, \text{lon}_{\min1}, \text{lat}_{\max1}, \text{lat}_{\min1})$  和  $g_2(\text{lon}_{\max2}, \text{lon}_{\min2}, \text{lat}_{\max2}, \text{lat}_{\min2})$ ,如果 2 个网格包含同一个建筑物、河流、山川、公园或路段等目标,网格间有可通行路径,且相邻边界线将目标切割成两个不完整的部分,则这两个相邻网格合并成一个新的网格  $g = \text{com}(g_1, g_2)$ ,更新边界坐标.

最后,获得物理相邻、且出行特性完整的网格.该方法能将不同的客流集聚区较为准确、完整地区分开,同时保留各自出行特征完整性.

### 1.2.2 出行需求计算

首先,定义  $x(v, k)$ ,  $y(v, k)$ ,  $f(v, k)$  分别表示第  $v$  辆出租车在  $k$  时刻的定位经度/纬度和载客状态(0—空载,1—载客,2—停运).当  $f(v, k-1) \neq 1$  且  $f(v, k) = 1$ ,证明在  $k-1$  与  $k$  时刻内有乘客上车.引入二进制变量  $p(v, x, y, q)$  代表在  $q$  时刻,坐标为  $x$  和  $y$  的位置,车辆  $v$  是否有乘客上车,其计算公式为

$$p(v, x, y, q) = \begin{cases} 1, & f(v, k-1) \neq 1, f(v, k) = 1; \\ 0, & \text{otherwise.} \end{cases} \quad (1)$$

式中:  $x = (x(k-1) + x(k))/2$ ;  $y = (y(k-1) +$

$y(k))/2$ . 根据上述定义,在  $m-1$  到  $m$  时刻内,网格  $g_i$  的出行需求总量为

$$\alpha(m) = \sum_{v=1}^V \sum_{(x,y) \in g_i} \sum_{q \in (m-1, m)} p(v, x, y, q). \quad (2)$$

式中,  $V$  表示出租车总数量.

## 2 出行需求预测模型

### 2.1 变量分析

作为交通预测模型,首要前提条件是根据实际交通问题分析与预测值相关的变量.根据文献综述和实际分析,与未来出行需求相关的变量分析如下:

1) 与前几个时刻的需求相关.因交通是一个连续系统,短时间内需求变化比较均匀,这也是自回归滑动平均预测模型的重要前提;

2) 与历史前几天同一时刻的出行需求相似,因为交通是一个具有周期性的随机系统,卡尔曼滤波模型预测旅行时间是基于该假设;

3) 与工作日、非工作日属性相关,主要受乘客出行目的的影响,是以上班为主,还是休闲娱乐为主的出行,其出发时段差异较大;

4) 与所在区域的特性相关,因现实中不同用地开发区域的出行需求和出行分布差异较大,尤其是住宅区、商业区、休闲区和购物区等,乘客出发时间和到达时间均有较大差异.

根据上述分析,以济南市收集的 2011 年 3 月以  $15\text{ min}$  为间隔的数据为基础,估计齐鲁软件园周边  $1.75\text{ km}^2$  网格内产生的出租车需求(合并了多个网格后出行需求量才比较正常,提高实用性),并进行相关性分析,如表 1 和表 2 所示,  $\alpha(w, m)$  代表第  $w$  天第  $m$  个时刻的出行需求总量.表 1 为 3 月 1 日星期二全天的数据相关性分析结果,表 2 为 3 月 7 日星期一到 3 月 13 日星期日整周的相关性分析,可以看出:

1) 在同一天内,当前时刻的出行需求与历史时间成反比,如表 1 最后一列,从整体来看,与历史最近两个时刻( $30\text{ min}$  内)的需求正相关性比较高,超过 0.7;而与更早数据相关性显著减弱,尤其与第  $75\text{ min}$  的相关性降到 0.32.

2) 对于历史天同时刻的出行需求,整体上工作日内的出行需求相关性较高,而与星期日的相关性最低,但星期六较特殊,与工作日的相关性高于星期日,这主要与该办公片区大部分企业周六上班有关.进一步分析 3 月 5 个星期二的相关性依次为 1.0, 0.74, 0.82, 0.79, 0.74,说明同一个

工作日相同时段的出行相似性也比较高.

2.2 预测模型

根据变量相关分析可知 ,出租车出行需求预测模型的输入变量不应该只局限于当天历史前几个时刻的需求<sup>[ 1 ,13 ]</sup> ,而与历史天同时刻的关系更密切 ,且要区分工作日与非工作日. 同时 ,由于交通是一个复杂的非线性时变系统 ,难以采用一般

数学模型刻画 ,而具有黑匣子特性的人工神经网络模型( artificial neural network ,ANN )为描述这一复杂关系提供了可能. 结合现有 ANN 模型在交通预测领域的优势<sup>[ 14 ]</sup> ,以及作者自身在本领域的研究成果<sup>[ 15 - 16 ]</sup> ,本文将研究利用前馈 ANN 模型进行需求预测. 选用包含输入层、隐含层和输出层的三层典型网络.

表 1 当前时刻出行需求与当天历史时刻的相关性分析

| Table 1 Correlation analysis between the current demand and previous interval ones among the same day |             |             |             |             |             |           |
|---|-------------|-------------|-------------|-------------|-------------|-----------|
| 相关系数  | $d(w, m-5)$ | $d(w, m-4)$ | $d(w, m-3)$ | $d(w, m-2)$ | $d(w, m-1)$ | $d(w, m)$ |
| $d(w, m-5)$   | 1.00        | 0.73        | 0.72        | 0.61        | 0.45        | 0.32      |
| $d(w, m-4)$   | 0.73        | 1.00        | 0.72        | 0.71        | 0.60        | 0.44      |
| $d(w, m-3)$   | 0.72        | 0.72        | 1.00        | 0.72        | 0.71        | 0.60      |
| $d(w, m-2)$   | 0.61        | 0.71        | 0.72        | 1.00        | 0.72        | 0.71      |
| $d(w, m-1)$   | 0.45        | 0.60        | 0.71        | 0.72        | 1.00        | 0.71      |
| $d(w, m)$   | 0.32        | 0.44        | 0.60        | 0.71        | 0.71        | 1.00      |

表 2 当前时刻出行需求与历史天同时刻的相关性分析

| Table 2 Correlation analysis between the current demand and historical day ones among the same interval |           |             |             |             |             |             |             |
|---|-----------|-------------|-------------|-------------|-------------|-------------|-------------|
| 相关系数  | $d(w, m)$ | $d(w+1, m)$ | $d(w+2, m)$ | $d(w+3, m)$ | $d(w+4, m)$ | $d(w+5, m)$ | $d(w+6, m)$ |
| $d(w, m)$ 星期一   | 1.00      | 0.80        | 0.81        | 0.77        | 0.73        | 0.77        | 0.67        |
| $d(w+1, m)$ 星期二   | 0.80      | 1.00        | 0.78        | 0.86        | 0.76        | 0.76        | 0.65        |
| $d(w+2, m)$ 星期三   | 0.81      | 0.78        | 1.00        | 0.82        | 0.68        | 0.73        | 0.76        |
| $d(w+3, m)$ 星期四   | 0.77      | 0.86        | 0.82        | 1.00        | 0.73        | 0.75        | 0.62        |
| $d(w+4, m)$ 星期五   | 0.73      | 0.76        | 0.68        | 0.73        | 1.00        | 0.73        | 0.55        |
| $d(w+5, m)$ 星期六   | 0.77      | 0.76        | 0.73        | 0.75        | 0.73        | 1.00        | 0.61        |
| $d(w+6, m)$ 星期日   | 0.67      | 0.65        | 0.76        | 0.62        | 0.55        | 0.61        | 1.00        |

1) 输入层 :作为 ANN 预测模型的首要参数 ,其变量选择直接决定了模型的准确性. 本文结合前述相关性分析 ,确定输入变量包含当前日期  $w$ 、星期属性( 1 ~ 7 )、时间段  $k$ 、当天前两个时刻需求  $d(w, m-1)$ 和  $d(w, m-2)$ 、最近历史 4 个工作日同时刻需求( 如果预测非工作日则不需要该组输入变量 )、最近历史三个相同星期属性同时刻需求  $d(w-7, m)$ 、 $d(w-14, m)$ 和  $d(w-21, m)$ . 因此 ,如果预测周一至周五其输入变量为 12 个 ,预测周六或周日其输入变量为 8 个.

2) 隐含层 :是估计输入与输出变量之间的非线性或线性关系. 隐含层神经元个数对于模型的准确性具有重要影响 ,根据现有研究总结<sup>[ 17 ]</sup> ,对于单输出 ANN 模型隐含层神经元个数

$N_H \in [ 0.11N_p/N_1 + 1, 0.3N_p/N_1 + 1 ]$ . ( 3 )  
式中 : $N_1$  为输入层节点数 ; $N_p$  为训练模式的数

量.

3) 输出层 :该网络的输出层只有一个变量 ,即下一时刻的出行需求  $d(w, m)$ .

3 算例分析

根据前述出行量估计及预测方法 ,以齐鲁软件园片区的运营数据为基础估计并分析现有自回归滑动平均 ( auto-regression-moving-average ,ARMA )模型和 ANN 模型的预测效果. 为区分两种不同结构的 ANN 模型 ,分别选用星期三和星期日两种情况进行预测 ,ARMA 和 ANN 模型参数训练样本均为历史最近一周的数据 ,并采用指标 :平均绝对误差百分比 ( mean absolute percentage error ,MAPE )、最大绝对误差百分比 ( max absolute percentage error ,MAXAPE )、平均

绝对偏差( mean absolute deviation ,MAD )、累积绝对偏差( cumulative absolute deviation ,CAD )进行对比分析 ,指标定义如下 :

$$MAPE = \frac{1}{I} \sum_{i=1}^I | d_i - \bar{d}_i | / d_i ; \quad (3)$$

$$MAXAPE = \max( \{ | d_i - \bar{d}_i | / d_i , i \in [ 1 , I ] \} ) ; \quad (4)$$

$$MAD = \frac{1}{I} \sum_{i=1}^I | d_i - \bar{d}_i | ; \quad (5)$$

$$CAD = \sum_{i=1}^I | d_i - \bar{d}_i | . \quad (6)$$

式中  $I$  代表预测总次数. 为评估误差分布 ,定义绝对误差百分比小于给定值  $x$  的概率为

$$P( x ) = \text{Num}( | d_i - \bar{d}_i | / d_i \leq x ) / I . \quad (7)$$

式中 ,Num( · )表示绝对误差百分比小于给定值  $x$  发生的次数.

预测结果各指标如表 3 所示 ,可以发现 :

表 3 ARMA 和 ANN 预测模型的评价结果  
Table 3 Evaluated results of ARMA and ANN prediction models

| 场景                               | 15 min 平均<br>出行量 | 方法   | 评价指标   |         |      |     |            |                |
|----------------------------------|------------------|------|--------|---------|------|-----|------------|----------------|
|                                  |                  |      | MAPE   | MAXAPE  | MAD  | CAD | $P( 0.3 )$ | $1 - P( 0.5 )$ |
| 2011/3/30 星期三<br>07 :45 到 13 :45 | 12 人次            | ANN  | 24. 1% | 85. 7%  | 3. 6 | 81  | 66. 7%     | 4. 2%          |
|                                  |                  | ARMA | 36. 7% | 119. 7% | 4. 8 | 116 | 58. 3%     | 25. 0%         |
| 2011/3/30 星期三<br>17 :00 到 22 :45 | 15 人次            | ANN  | 35. 0% | 140. 0% | 4. 9 | 112 | 56. 5%     | 17. 4%         |
|                                  |                  | ARMA | 61. 9% | 286. 6% | 7. 5 | 172 | 39. 1%     | 34. 8%         |
| 2011/3/27 星期日<br>11 :15 到 21 :15 | 7. 4 人次          | ANN  | 26. 4% | 100. 0% | 2. 0 | 79  | 60. 0%     | 10. 0%         |
|                                  |                  | ARMA | 32. 6% | 133. 3% | 2. 4 | 96  | 52. 5%     | 15. 0%         |

4 结 论

本研究结合出租车运营系统实时采集的车辆位置和载客状态信息 ,提出了基于网格的出租车乘客需求实时估计模型和短时预测模型. 在需求估计模型中改进了传统网格划分方法 ,加入了地形、建筑群和道路网络特征等因素 ,保持了网格出行需求的完整性. 在预测阶段 ,分析了主要输入输出变量间的相关性 ,提出了基于人工神经网络的预测模型 ,并根据相关性分析确定了模型结构. 根据实际数据结果证明 :提出的模型预测性能显著优于现有自回归滑动平均模型 ,其平均绝对误差百分比提高了 32% . 本文所研究的出租车乘客需求估计和预测模型 ,较为准确地预计了短时出租车需求量 ,为出租车实时调度和出租车规划提供了可靠的信息支撑.

1 ) 该片区 3 月份出行总量约为 2. 4 万人次 ,最大和最小日出行量分别为 983 人次和 428 人次. 全天各时段需求差异性也较大 ,最大和最小需求分别为每小时 87 人次和 1 人次 ,需求低峰在 0 ~7 点 ,高峰集中在 17 :00 到 18 :00 下班时间. 这符合该片区以商业办公为主的土地利用特性.

2 ) ANN 准确性均优于现有的 ARMA 模型 ,尤其是星期日 ,ARMA 模型没有考虑历史天出行需求的相似性 ,难以捕捉其变化趋势.

3 ) 本文提出的 ANN 模型预测的可靠性也比较高 ,在三种情况下 ,预测绝对误差百分比超过 50% 的概率低于 10% ,而 ARMA 模型高达 23% . 然而 ,ANN 模型在星期三晚上预测的平均绝对误差百分比和最大绝对误差百分比约为 35% 和 140% ,这与该时间段内需求波动剧烈有关( 其方差高于其余两种情况方差的两倍 ).

参考文献 :

[ 1 ] Moreira-Matias L ,Gama J ,Ferreira M ,et al. Predicting taxi-passenger demand using streaming data [ J ]. *IEEE Transactions on Intelligent Transportation Systems* ,2013 ,14 ( 3 ) :1393 - 1402.

[ 2 ] Arnott R. Taxi travel should be subsidized[ J ]. *Journal of Urban Economics* ,1996 ,40 ( 3 ) :316 - 333.

[ 3 ] Liao Z. Taxi dispatching via global positioning systems[ J ]. *IEEE Transactions on Engineering Management* ,2001 ,48 ( 3 ) :342 - 347.

[ 4 ] Glaschenko A ,Ivaschenko A ,Rzevski G ,et al. Multi-agent real time scheduling system for taxi companies[ C ]//8th International Conference on Autonomous Agents and Multiagent Systems. Budapest ,2009 :29 - 36.

[ 5 ] Wong K I ,Wong S C ,Bell M G H ,et al. Modeling the bilateral micro-searching behavior for urban taxi services using the absorbing Markov chain approach[ J ]. *Journal of Advanced Transportation* ,2005 ,39 ( 1 ) :81 - 104.

[ 6 ] 李艳红 ,袁振洲 ,谢海红 ,等. 基于出租车 OD 数据的出租车出行特征分析[ J ]. *交通运输系统工程与信息* ,2007 ,7

( 5 ) 85 – 89.

( Li Yan-hong , Yuan Zhen-zhou , Xie Hai-hong , et al. Analysis on trips characteristics of taxi in Suzhou based on OD data[ J ]. *Journal of Transportation Systems Engineering and Information Technology* 2007 ,7( 5 ) 85 – 89. )

[ 7 ] 边扬 ,王伟 ,陆建. 城市出租车运营网络平衡模型[ J ]. *交通运输工程学报* 2007 ,7( 1 ) 93 – 98.

( Bian Yang ,Wang Wei ,Lu Jian. Equilibrium model of urban taxi service network [ J ]. *Journal of Traffic and Transportation Engineering* 2007 ,7( 1 ) 93 – 98. )

[ 8 ] Yang H ,Lau Y W ,Wong S C ,et al. A macroscopic taxi model for passenger demand ,taxi utilization and level of services[ J ]. *Transportation* 2000 ,27( 3 ) 317 – 340.

[ 9 ] Wong K I ,Wong S C ,Yang H. Modeling urban taxi services in congested road networks with elastic demand[ J ]. *Transportation Research Part B :Methodological* ,2001 ,35 ( 9 ) 819 – 842.

[ 10 ] 王昊 ,王伟 ,陈峻 ,等. 城市出租车交通分布预测模型[ J ]. *公路交通科技* 2006 ,23( 6 ) :145 – 156.

( Wang Hao ,Wang Wei ,Chen Jun ,et al. A forecasting model of trip distribution for urban taxis[ J ]. *Journal of Highway and Transportation Research and Development* ,2006 ,23 ( 6 ) :145 – 156. )

[ 11 ] Chang H ,Tai Y ,Hsu J Y. Context-aware taxi demand hotspots prediction[ J ]. *International Journal of Business Intelligence and Data Mining* 2010 ,5( 1 ) 3 – 18.

[ 12 ] Deng Z ,Ji M. Spatiotemporal structure of taxi services in Shanghai using exploratory spatial data analysis[ C ]//19th International Conference on Geoinformatics IEEE. Shanghai , 2011 :1 – 5.

[ 13 ] Li X ,Pan G ,Wu Z ,et al. Prediction of urban human mobility using large-scale taxi traces and its applications[ J ]. *Frontiers of Computer Science* 2012 ,6( 1 ) :111 – 121.

[ 14 ] Dougherty M. A review of neural networks applied to transport [ J ]. *Transportation Research Part C :Emerging Technologies* ,1995 ,3( 4 ) 247 – 260.

[ 15 ] Lin Y J ,Yang X F ,Zou N ,et al. Real time bus arrival prediction :a case study of Jinan China[ J ]. *Journal of Transportation Engineering ASCE* ,2013 ,139( 11 ) :1134 – 1140.

[ 16 ] 李大铭 ,赵新良 ,林永杰 ,等. 基于模糊神经网络的短时公交到站时间预测[ J ]. *东北大学学报( 自然科学版 )* 2011 ,32( 3 ) :443 – 446.

( Li Da-ming ,Zhao Xin-liang ,Lin Yong-jie ,et al. Short-term bus arrival time prediction using a fuzzy neural network[ J ]. *Journal of Northeastern University ( Natural Science )* 2011 ,32( 3 ) :443 – 446. )

[ 17 ] Basheer I A , Hajmeer M. Artificial neural networks : fundamentals ,computing design and application[ J ]. *Journal of Microbiological Methods* 2000 ,43( 1 ) 3 – 31.