

基于泊松过程的寻找空驶出租车的算法

樊里略

(遵义师范学院计算机科学系 贵州遵义 563002)

摘要: 随着城市化进程的推进,城市车辆数量呈现激增的趋势。城市交通为居民的出行提供便利的同时,也带来了其它挑战,比如打车难问题。本文的主要工作分为两部分:1) 预测乘客在某个位置某个时间点的打车概率和等待空车时间;2) 在某个位置某个时间点,为乘客推荐打车位置。为了完成上述主要工作,本文首先对大规模的出租车 GPS 轨迹数据进行预处理,并生成道路特征索引。然后利用非齐次泊松过程 NPP(Nonhomogeneous Poisson Process)进行预测和推荐工作。在实验部分,我们基于真实的北京市出租车轨迹数据(由 12000 辆出租车在 30 天内产生)和路网数据,将本文方法与其它相关先进方法进行对比,实验结果表明,本文方法在准确率和运行效率方面更具优势。

关键词: GPS 轨迹,出租车,路段索引,预测,推荐

A Poisson Process-based Algorithm for Searching Vacant Taxi

FAN Li Lue

(Zunyi Normal College Department of Computer Science Gueizhou Zunyi 563002)

Abstract

With the development of urbanization, the number of cars has been proliferating fast. Urban transportation provides convenience for city resident, meanwhile, it also brings many challenges, such as causing a difficulty for taking a taxi. In this paper, main work is composed of two parts: first, we propose a method to estimate the probability and waiting time for a vacant taxi at a given time and place, second, we provide an recommendations for passengers of where to wait for a taxi. In order to complete the main work, we firstly preprocess the Large-scale Taxi GPS Traces and generate the road characteristic index. Then we use NPP (Nonhomogeneous Poisson Process) to predict and recommend. In the experimental part of the paper, based on the large-scale real-world GPS traces dataset, we compare our method and other related advanced methods and the results show that our method has more advantages in accuracy and efficiency.

Keywords: GPS traces, taxi, road segment index, prediction, recommendation

一、引言

出租车作为介于公共交通和私人交通之间的交通工具在人们的日常交通出行中起到了重要的作用[1]。然而,随着许多城市交通状况的恶化,打车难成为城市交通的一个普遍问题,一方面,城市出租车的数量与乘客之间存在供需矛盾;另一方面,乘客缺少空载出租车的位置信息。造成了部分有经验的乘客会到出租车经常停泊的地点等车,而大部分乘客只能在街道上长时间等待。因此,为乘客提供空驶的出租车信息,并告知乘客出租车需要等待的时间和打到空车的概率十分必要。

为了安全和调度的方便,许多大城市出租车都装备了 GPS 定位设备,如北京、纽约、新加坡等等。GPS 设备作为出租车移动的传感器,时刻探测着城市的交通动态[2],它们源源不断地以一定的频率向管理中心发送状态信息,如当前的经纬度、时间戳、是否载有乘客、车行速度和方向等。基于这些海量的出租车 GPS 轨迹数据,研究人员可以以独特视角研究一个城市路网动态信息[3],如人类移动模式挖掘[4-5]和出租车移动模式智能挖掘等[6-9]。

我们的研究建立在真实的北京市出租车轨迹数据集和路网信息上,该数据集是 12000 辆出租车 1 个月内的数据,数据超过 9.32 亿条。由于数据规模大,我们使用 MapReduce 来处理数据集并生成路段序列数据集索引。当给定一个乘客的具体位置和时刻时,我们使用 NPP 算法来预测等待的时间和遇到空车的概率。

- 1、基金项目:本课题得到遵义市科技局遵社科字[2010]1 号资助作者简介:樊里略(1965-),男(汉族),河南洛阳人,学士,副教授,主要研究方向计算机应用

本文的贡献主要分为下列三点:

1)、我们把每个位置映射到对应的路段上面,并提出了相应的方法建立基于路段特征的索引。面对大规模的数据集,我们利用 MapReduce 框架来完成数据预处理的工作。

2)、给定某个位置某个时间点,我们利用 NPP 方法来预测乘客的打车概率和等待时间。我们还可以为乘客推荐合适的打车位置,为乘客节省等待时间。

3)、我们基于真实的数据集,进行了实地实验。通过与其它相关先进方法进行对比,我们的方法在预测等待空车时间、遇见空车的概率和乘客推荐的建议方面优于其它方法,验证我们方法的有效性和准确性。

本文结构如下:第一部分引言;第二部分相关工作;第三部分方法概述;第四部分数据处理;第五部分建立模型;第六部分对乘客的推荐;第七部分实验;第八部分总结。

二、相关工作

2.1、出租车的调度

随着智能交通系统的发展和 GPS 设备的普及,出租车调度系统的研究吸引了越来越多研究人员的关注[9]。文献[10]使用了 Naive 贝叶斯分类器和基于错误学习的方法来推测给定地点和时间的空车个数,从而增强调度系统的性能。文献[13]提出了基于模糊聚类和自适应路由的方法,安排空车去潜在乘客数期望最高的地点,从而提高调度系统的有效性;都采取集中调度和提高效率方式进行。这些工作的目标都是为了改进调度系统,与这些工作相比,本文更偏重于为乘客服务。

2.2、对出租车司机推荐系统

有许多文献已经对出租车司机基于位置推荐系统文献[11]提出了一个移动序列推荐模型,当乘客下车后,系统给司机推荐一个地点序列,当司机按照该序列行驶时,可以最大化他的收益;文献[12]研究了杭州市出租车司机的行为模式(巡航或停泊),并提出了基于 L1-范式的 SVM 的分类方法提高了分类器的性能;文献[14]从经济的角度给为租车司机提供建议。然而,这些工作都是关注与如何为出租车司机服务,本文更加关注乘客的利益。

2.3、对乘客的推荐系统

文献[15]对出租车的空驶现象进行了调查研究,利用出租车的 GPS 数据得到了不同时段出租车乘车的供给和需求的关系。通过分析不同时段的出租车空驶率,该研究旨在找出降低空驶率、实现出租车资源有效利用的手段。文献[16]提出了乘客推荐系统,该推荐系统研究了两个方面,1)基于乘客的经验建议等车的位置;2)的士司机基于载客或卸客的经验发现停泊位;其目的是使出租车司机收益最大化,本文更加注重方法的准确率和运行的效率。

三、系统概述

本部分内容首先给出与本文工作相关的定义,然后介绍系统框架。

3.1、初步定义

定义 3.1 (路段) 路段(Road Segment) r 是一条边, r 的相关信息包括两个端点(头 $r.s$ 和尾 $r.e$)、长度($r.length$)以及描述其形状的点序列($r.points$)。

定义 3.2(出租车状态)本文中出租车的三种状态 (States),它们分别是占用 (Occupied, 记为 (O)),空驶 (Cruising,记为(C)),停泊 (Parking,记为(p))空驶和停泊状态为非占用状态。它们的具体含义列在了表中。

状态	含义
占用 (O)	出租车正载有乘客的状态
空驶 (C)	出租车没有乘客且路段上行驶状态
停泊 (P)	出租车没有乘客且在某处停留等待乘客

表 (1) 出租车状态含义

3.2、系统框架

首先，我们利用 MapReduce 框架来处理大规模出租车的轨迹数据，在 MAP 阶段把轨迹数据映射到路网中的路段上，在 REDUCE 阶段以每一个路段作为 Key 值方法进行处理，建立路段序列数据集索引。然后，从数据集中提取响应信息，通过 NPP 方法计算等到空车时间、遇见空车的概率。系统框架如图 1 所示。

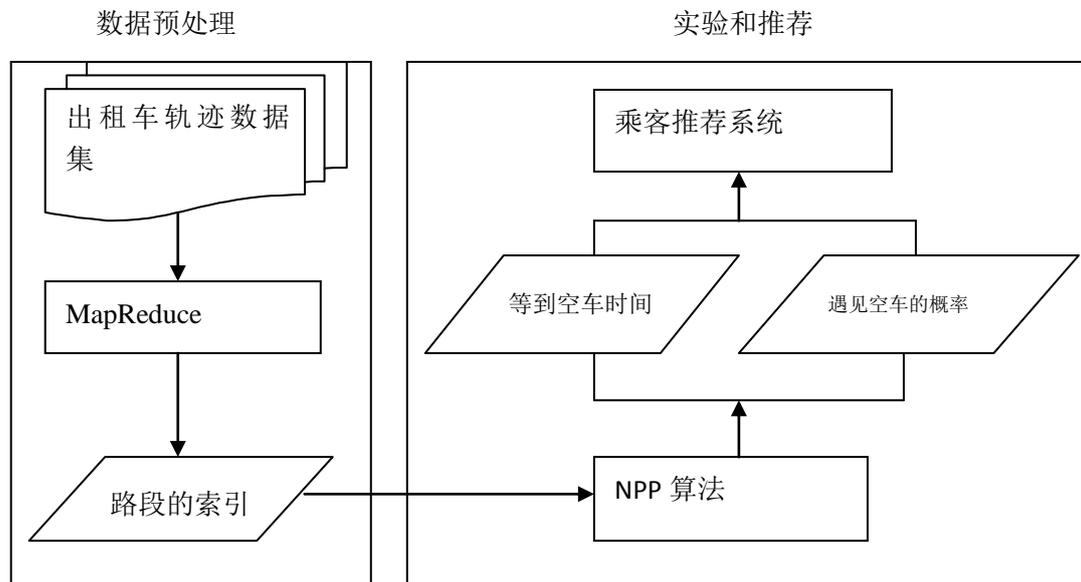


图 (1) 系统框架

四、数据的预处理

4.1、数据格式和数据的清理:

数据是北京市 2012 年 11 月 (12000 辆出租车 GPS 记录的真实的数据集)，通过安装在出租车上的 GPS 探测器，经过固定的时间间隔(50 s)采集到的包含车辆位置、行驶速度、行驶方向和时间等数据。数据格式如下：

车辆标识	触发事件	运营状态	GPS 时间	GPS 经度	GPS 纬度	GPS 速度	GPS 方位	GPS 状态
495366	4	0	20121105083627	116.3117676	39.9446449	21	276	1
063406	4	1	20121115132812	116.3302002	39.9504128	21	342	1

表 (2) 出租车轨迹数据格式

原始的轨迹数据包含大量的、不确定的、有噪声的和随机的数据，为了进行减少数据对系统的误差，对原始的轨迹数据进行处理。在出租车原始的轨迹数据集里，有些数据记录的经度和纬度的值超过北京地界经度和纬度的值；在一段时间内，有些数据的经度、纬度无变化的记录；有些出租车记录的载客状态显示错误值；对这样数据进行清理，生成出租车轨迹数据集。

4.2、路段的索引

根据定义 (3.1)，我们利用 MapReduce 框架处理出租车轨迹数据集。MapReduce 将文件拆分成 splits，并将文件按行分割形成<key,value>对；在 map 中，把<key,value>对按轨迹数据映射到路网中的路段上处理，生成<路网,轨迹>的键值对；在 Reducer 中，先对接收的数据进行排序，以每一个路段作为 Key 值方法进行处理。得到<路网-轨迹>键值对，按此输出结果。在<路网-轨迹>过程生成每个路段上的轨迹统计数据；最终形成以路网为特征数据集。把特征数据集以路段聚类，除去异常点，建立预测模型。并在模型内建立以路段对应轨迹的倒排索引，便于实时查询。

五、建立模型:

乘客和街道上行驶的出租车是随机出现，所组成乘客序列和出租车序列相互独立，且与

当前状态有关，乘客序列或出租车序列中的每个随机变量仅依赖于它的前一个随机变量，而独立于其他前面的所有随机变量，所以乘客序列和行驶出租车序列符合泊松过程。算法思想：利用泊松过程，依靠于不同路段不同时间段的空车到达率的历史记录来预测出当前情况下（即用户的所在位置和所处时刻）一辆空驶出租车到达的需要等待的时间。在模型中以非齐次泊松过程为条件，解决两个问题：1、预测等待时间；2、遇到空车的概率。

用 T_{taxi} 表示出租车自身的到来时间，则在时间点 t_0 和位置点 p_0 ，等车时间 $T_w^{t_0,p_0}$ 可以表示为：

$$T_w^{t_0,p_0} = f_{road}(T_{taxi}) \quad (1)$$

其中 f_{road} 为路段函数。

下面讨论怎样估计 T_{taxi} 的值,并给出 $T_w^{t_0,p_0}$ 的表达形式。

5.1、预测等待空车的时间：

在某一路段上 R,出租车接单行序列而入，在 t_0 时刻达到 R 路段 p_0 位置。有两种情况：1、 t_0 时刻到达 p_0 位置有车等候 $T=0$ ；2、 t_0 时刻到达 p_0 位置无车等候 $t \in \{0 < t < T\}$ 。

根据定义 (3.2)，在 R 路段上定义出租车的三种状态：其中空车状态有两种，一种在停泊位停留的车 P，另一种是出租车在空驶情况 C。当到达街边有停留的车， $t=0$ ；C 状态符合非齐次的泊松过程。因此，有

$$T_{taxi} = \Pr(P|t_0, p_0) * 0 + \Pr(C|t_0, p_0) * T_{next} \quad (2)$$

进一步地，

$$\Pr(C|t_0, p_0) = \frac{\sum_{k=\lfloor(t-\Delta t)/\tau\rfloor}^{\lfloor(t+\Delta t)/\tau\rfloor} \sum_{j \in D(p_0, \delta)} \#_{k,j}(C)}{\sum_{k=\lfloor(t-\Delta t)/\tau\rfloor}^{\lfloor(t+\Delta t)/\tau\rfloor} \sum_{j \in D(p_0, \delta)} (\#_{k,j}(C) + \#_{k,j}(P) + \#_{k,j}(O))} \quad (3)$$

在 R 路段上，出租车状态#C 表示为路段上空车行驶的个数；#P 表示在某个停泊位停了车的个数；#O 表示为行驶在路段上有乘客使用状态的出租车的个数。

接下来，只要预测 T_{next} 的值即可。

用 $N(t)$ 表示出租车到达情况的随机变量，用 $\Lambda(t)$ 表示出租车到达情况的累积强度，根据非齐次泊松过程，我们得到如下性质：

性质 1

$$\Lambda(t) = \int_0^t \lambda(\tau) d\tau \quad (4)$$

性质 2

$$\Pr\{N(t) = k\} = \Pr\{N(t+s) - N(s) = k\} = e^{-\Lambda(t)} \frac{\Lambda(t)^k}{k!}, \quad N(t) = \{0, 1, 2, \dots, n\}. \quad (5)$$

且 $N_{(0)} = 0$

所以，在 t 时刻内，出租车到达的累积分布函数（CDF）为：

$$\begin{aligned} \Pr\{T_{next} < t\} &= 1 - \Pr\{T_{next} > t\} \\ &= 1 - \Pr\{N(t) = 0\} \\ &= 1 - e^{-\Lambda(t)} \end{aligned} \quad (6)$$

因此， T_{next} 的密度函数（PDF）为累积分布函数（CDF）的导数。即为：

$$P(T_{next}) = \frac{d \Pr\{T_{next} < t\}}{dt} = \lambda(t) e^{-\Lambda(t)} \quad (7)$$

为了方便起见，我们将 $\lambda(t)$ 简化为分段线性函数，并确定 5 分钟为分段的间隔

$$\text{因此得到} \quad \hat{T}_{next} = \int_0^{+\infty} t \lambda e^{-\lambda t} dt = \frac{1}{\lambda} \quad (8)$$

接下来使用最大似然估计（MLE）对 λ 进行估计，记 $N_{k,j}^i(C)$ 为第 i 天，在时间点 k，位置 j 的空车数目，则得到似然函数

$$L(\lambda) = \prod_{i=1}^N \Pr\{N(\Delta t) = N_{k,j}^i(C)\} \quad (9)$$

$$\text{令 } \frac{\partial \ln L(\lambda)}{\partial \lambda} = 0, \text{ 解得 } \lambda = \frac{\sum_{i=1}^n N_{k,j}^i(C)}{N \Delta t}$$

最终，我们得到

$$T_{\text{taxi}} = \frac{\sum_{k=\lfloor (t-\Delta t)/\tau \rfloor}^{\lfloor (t+\Delta t)/\tau \rfloor} \sum_{j \in D(p_0, \delta)} \#_{k,j}(C) N \Delta t}{\sum_{k=\lfloor (t-\Delta t)/\tau \rfloor}^{\lfloor (t+\Delta t)/\tau \rfloor} \sum_{j \in D(p_0, \delta)} (\#_{k,j}(C) + \#_{k,j}(P) + \#_{k,j}(O)) \sum_{i=1}^n N_{k,j}^i(C)} \quad (10)$$

$T_w^{t_0, p_0}$ 形式为:

$$T_w^{t_0, p_0} = f_{\text{road}} \left(\frac{\sum_{k=\lfloor (t-\Delta t)/\tau \rfloor}^{\lfloor (t+\Delta t)/\tau \rfloor} \sum_{j \in D(p_0, \delta)} \#_{k,j}(C) N \Delta t}{\sum_{k=\lfloor (t-\Delta t)/\tau \rfloor}^{\lfloor (t+\Delta t)/\tau \rfloor} \sum_{j \in D(p_0, \delta)} (\#_{k,j}(C) + \#_{k,j}(P) + \#_{k,j}(O)) \sum_{i=1}^n N_{k,j}^i(C)} \right) \quad (11)$$

5.2、预测遇见空车的概率:

我们定义了 τ 时刻内一个空置的出租车到达概率为 P_{taxi} 。根据公式 (6), P_{taxi} 为

$$P_{\text{taxi}} = \Pr\{T_{\text{next}} < t\} = 1 - e^{-\Lambda(t)} \quad (12)$$

六、基于位置的乘客推荐系统

乘客打车时,有时在某个位置长时间等不到车,部分乘客会换地点打车。哪里能打到车?给乘客一个推荐。当一个空置出租车从某位置 p_0 和某时刻 t_0 行驶过来,使用 NPP 算法预测等车时间和遇见出租车的概率。乘客打车时,通常乘客的不耐烦性是乘客特性比较重要的一个方面,它表示了乘客对于等车的忍耐程度。推荐系统考虑这个因素。假设乘客等车忍耐极限时间为 $T_{\text{endurance}}$,能提供的另一个等车位置离原位置不远,花费时间不能超过 $T_{\text{endurance}}$ 。目标是协调等车时间和步行距离。为乘客推荐的候选的位置 $P_{\text{candidata}}$ 。

$$P_{\text{candidata}} = \{p: \text{distance}(p_0, p) < S T_{\text{endurance}}\} \quad (13)$$

S 为步行的速度, Distance 为 p_0 位置到 P 位置的距离。定义步行时间为 T_{walk} ,等于从 p_0 到 P 位置距离与速度的比值 $\frac{\text{distance}(p_0, P)}{S}$ 。然后,定义从 p_0 到 P 位置花费的总时间为 T_{total} 。

$$T_{\text{total}} = T_{\text{walk}} + T_w^{t_0 + T_{\text{walk}}, p} \quad (14)$$

最后,从 $P_{\text{candidata}}$ 集内挑选一个最短路段 P_{best}

$$P_{\text{best}} = \{p \in P_{\text{candidata}} : \text{argmin} T_{\text{total}}\} \quad (15)$$

七、实验:

本节进行全面的实验来评估我们的方法。实验利用北京市不同区域的路段来验证等待空车时间和遇见空车的概率。在第一个实验中,实验验证空置出租车等车的时间和遇见车的概率,并与文献[16]方法比较运行的有效性和准确性。在第二个实验,实验评估乘客推荐模型,并与文献[17]比较运行的效率。数据是北京市 2012 年 11 月份(12000 辆出租车 GPS 记录的数据集)数据记录超过 9.32 亿条。把 GPS 数据集分成两部分。30 日数据作为训练集。

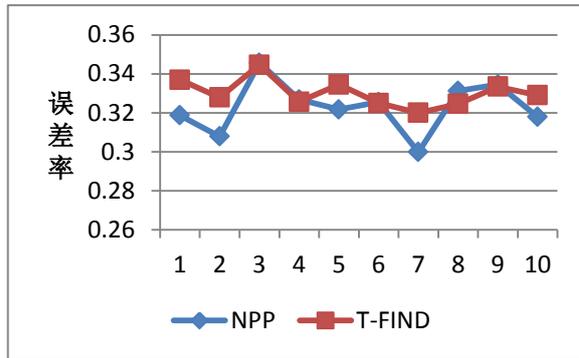
路网数据:该实验用到北京市的路网数据,它包含 106,579 个路网结点和 141,380 条路段。

7.1、预测等待空车的时间和遇见空置出租车的概率

利用北京市不同区域三条路段来测试我们方法等待空车时间。北京市内选取中关村北大街 (p_1, p_2, p_3, p_4)、知春路 (p_5, p_6, p_7, p_8) 和中关村东路 ($p_9, p_{10}, p_{11}, p_{12}$)。每个路段都测试 4 个不同的位置。在几个给定时间段里,每一个位置预测等到第一辆空置出租车到来的时间。在给定的位置和时间段里,每个位置记录第一辆空置出租车等待时间真实数据。在这些路段和给定时间段中,我们的方法预测第一辆空置出租车到来某位置需要时间,并与文献[16]方法比较运行的准确性。准确性可以用失误率低来描述。利用等待空车时间的相对误差率表示这两种方法的差别。等待空车时间误差率表示为:

$$\text{误差率} = \frac{|\text{真实值} - \text{预测值}|}{\text{真实值}} \times 100\%$$

(16)



图(2) 等车时间的误差率

图(2)显示的是 NPP 方法和 T-Finder 方法预测等待空车时间的相对误差率。NPP 方法平均误差率为 32.29%；T-Finder 方法平均误差率为 33.01%。结果显示是合理的，NPP 方法比 T-Finder 方法要更精确。

为了预测获得空置出租车的概率，在中关村北大街、知春路和中关村东路，每条路段都测试 4 个不同的位置。对于每个位置，实验记录了给定的时间段通过该路段的空车数。见表(3) 这里是基于打到空车的概率排序的结果。#记为在给定的时间内通过该位置的真实空车数。Rank 记为该路段空车数的排序结果。在 Rank_{T-Finder} 和 Rank_{NPP} 方法中，每个位置排序结果与 Rank 排序结果不合者用粗体字表示。结果说明：NPP 方法好于 T-Finder。

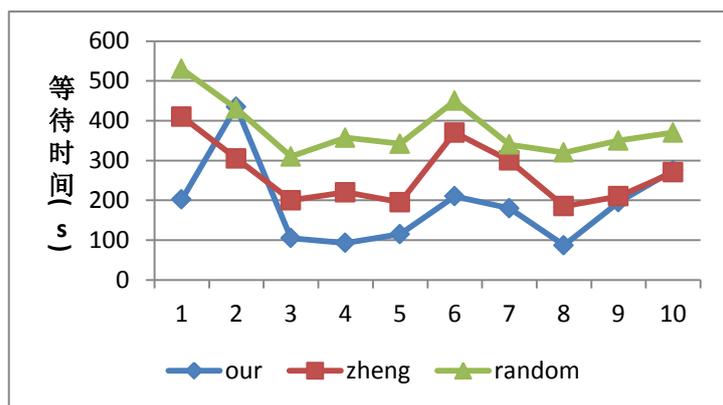
路段	中关村北大街				知春路				中关村东路			
时间	8: 30--8: 35				9: 30—9: 35				10: 05—10: 10			
位置	p ₁	p ₂	p ₃	p ₄	p ₅	p ₆	p ₇	p ₈	p ₉	p ₁₀	p ₁₁	p ₁₂
#	5.1	3.2	3.0	1.2	1.8	2.2	3.2	8.3	6.2	2.4	3.3	3.2
Rank	1	2	3	4	4	2	3	1	1	4	2	3
Rank _{T-Finder}	1	3	2	4	2	3	4	1	1	2	3	3
Rank _{NPP}	1	2	3	4	4	3	2	1	1	4	3	2

路段	中关村北大街				知春路				中关村东路			
时间	13: 30--13: 35				14: 30—14: 35				15: 05—15: 10			
位置	p ₁	p ₂	p ₃	p ₄	p ₅	p ₆	p ₇	p ₈	p ₉	p ₁₀	p ₁₁	p ₁₂
#	4.2	2.8	3.2	1.3	2.5	3.2	2.2	5.3	4.2	5.7	1.5	2.4
Rank	1	3	2	4	3	2	4	1	2	1	4	3
Rank _{T-Finder}	2	1	3	4	2	3	4	1	1	2	3	3
Rank _{NPP}	1	2	3	4	3	2	4	1	3	1	4	2

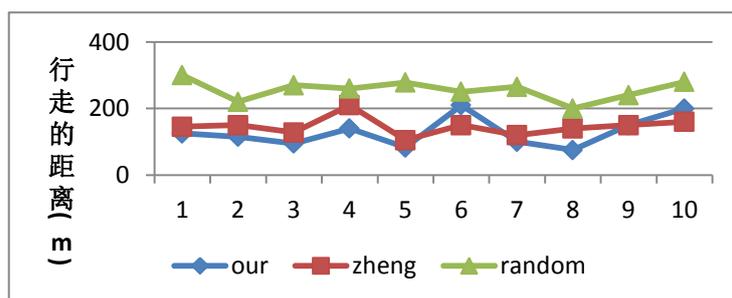
表(3) 预测空置出租车的概率

7.2、评估乘客推荐系统

进一步评估乘客推荐模型，从 p₁, p₂...p₁₁, p₁₂ 位置中随机选取 10 个位置，对每个位置，进行实验比对，验证每个位置平均等待时间。用我们的乘客推荐模型与文献[17]方法和乘客自由走动的方法进行比较。我们的推荐方法平均等待时间低于其它两种方法。图(3)表明不同位置平均等待时间的差异。结果表明，我们的推荐方法优于文献[17]方法和乘客自由走动的方法。图(4)显示步行距离的差异，我们的方法行走距离比文献[17]和乘客自由走动方法行走距离要少。



图(3)十个不同位置平均等待的时间



图(4)十个不同位置行走的距离

八、结论

文章通过对大规模出租车 GPS 轨迹数据的进行处理, 建立路段特征索引, 利用 NPP 算法预测了等待空车时间和遇见空车的概率, 通过一系列的实验, 并与其它方法进行比较, 效果高于其它方法。进一步提出了乘客推荐系统, 有目的指导乘客到就近的地方寻觅空置的出租车。由于出租车数量(不到北京市出租车总数的1/5)和行驶范围覆盖率有限, 某些路段存在有误差, 这将是今后需要进一步解决的问题。

参考文献

- [1] LU, J., WANG, W.: Confirming method of urban taxi quantity. Journal of Traffcand Transportation Engineering 4(1) (2004) 92-95
- [2]Zheng, Y., Liu, Y., Yuan, J., Xie, X.: Urban computing with taxicabs. In: Proceedings of the 13thinternational conference on Ubiquitous computing, ACM (2011)89-98
- [3]Castro, P.S., Zhang, D., Li, S.: Urban tra_c modelling and prediction using largescale taxi gps traces. In: Pervasive Computing. Springer (2012) 57-72
- [4]Gonzalez, M.C., Hidalgo, C.A., Barabasi, A.L.: Understanding individual humanmobility patterns. Nature 453(7196) (2008) 779-782
- [5]Li, X., Pan, G., Wu, Z., Qi, G., Li, S., Zhang, D., Zhang, W., Wang, Z.: Predictionof urban human mobility using large-scale taxi traces and its applications. Frontiersof Computer Science 6(1) (2012) 111-121

- [6]Chang, H.w., Tai, Y.c., Hsu, J.Y.j.: Context-aware taxi demand hotspots prediction. *International Journal of Business Intelligence and Data Mining* 5(1) (2010)3-18
- [7]Reades, J., Calabrese, F., Sevtsuk, A., Ratti, C.: Cellular census: Explorations in urban data collection. *Pervasive Computing, IEEE* 6(3) (2007) 30-38
- [8]Wong, K., Wong, S., Bell, M., Yang, H.: Modeling the bilateral micro-searching behavior for urban taxi services using the absorbing markov chain approach. *Journal of advanced transportation* 39(1) (2005) 81-104
- [9]桂智明,向宇,李玉鉴.基于出租车轨迹的并行城市热点区域发现. *华中科技大学学报*, 2012, 40(12): 187--190
- [10] Lee D, Wang H, Cheu R, Teo S Taxi dispatch system based on current demands and real-time traffic conditions *Transportation Research Record Journal of the Transportation Research Board*, (2004),1882(1),193-200.
- [11]Phithakkitnukoon, S., Veloso, M., Bento, C., Biderman, A., Ratti, C.: Taxi-awaremap: identifying and predicting vacant taxis in the city. In: *Ambient Intelligence*. Springer (2010) 86-95
- [12] Ge Y, Xiong H, Tuzhilin A, Xiao K, Gruteser M, Pazzani M An energy-efficient mobile recommender system In *Proc. KDD 2010* 899-908
- [13] Li B, Zhang D, Sun L, Chen C, Li S, Qi G, Yang Q Hunting or waiting? Discovering passenger-finding strategies from a large-scale real-world taxi dataset In *Pervasive Computing and Communications Workshops (PERCOM Workshops)*, 2011 IEEE International Conference on, 2011 63 -68
- [14]Yamamoto K, Uesugi K, Watanabe T Adaptive Routing of Cruising Taxis by Mutual Exchange of Pathways In: *Knowledge-Based Intelligent Information and Engineering Systems Springer*, 2010. 559-566.
- [15] 傅常伦. 基于车辆定位系统的数据融合及其在城市交通中的应用[D]. 同济大学硕士学位论文, 2008: 10-12.
- [16] Yuan, N., Zheng, Y., Zhang, L., Xie, X.: T_finder: A recommender system for finding passengers and vacant taxis. (2012)
- [17] Zheng, X., Liang, X., Xu, K.: Where to wait for a taxi? In: *Proceedings of the ACM SIGKDD International Workshop on Urban Computing*, ACM (2012) 149-156
- [18]康留旺. 基于 FCD 的城市出租车空驶率及拥有量的计算[J]. *数字技术与研究*, 2010 (07): 81-82.
- [19]胡立, 陈健. 基于用户轨迹聚类分析的推荐算法研究[J]. *计算机研究与发展*, 2012 (49): 250-256.
- [20]张朋. 基于浮动车数据的城市居民出行规律研究[D]. 中南大学硕士学位论文 2012.