

“互联网+”时代的出租车资源配置问题的研究

摘 要

随着“互联网+”时代的到来，各类公司创建打车软件平台，通过出台补贴方案，结合相关条件，合理调控出租车与乘客之间的关系。通过建立基于改进的 K-means 聚类分析评价模型结合浮动车调度模型和 RVGAM 缓解效率评分分析模型对以缓解打车问题为主导分析，研究更注重对各类指标参数特性进行定性分析与设定主要反映程度进行定量关系确定的特点，考虑的情况更为复杂，更贴近实际，因而求解得出各类反映程度值并进行分析后，更具备实际应用价值和更重要的指导意义。

问题一，主要研究在不同的时空下的出租车资源的供求匹配程度，以安徽省合肥市的出租车交通网数据库为例，立足现有条件，建立了改进的 K-means 聚类分析评价模型，结合浮动车模型，将一天的时间分为 7 个时间段，以千人拥有量、该地区出租车数量、空载率、乘客等车率为指标，将出租车资源的供求匹配程度具体量化为比值形式，并设立评分值 ε ，得出时间段 9:30~11:00 的第 71 位置点：合肥市政府广场（北纬 31.514601，东经 117.164568） $\varepsilon=0.0327$ ，供求匹配程度最好；在时间段 17:30~19:30 的合肥火车站（北纬 31.885223，东经 117.317012） $\varepsilon=0.324$ ，供求匹配程度最差。

问题二，主要研究快的打车与滴滴打车两家打车软件公司推出的打车补贴对合肥市不同时段不同地区“打车难”的缓解效果。为将打车补贴推行前后的乘客打车情况形成对比，本文提取了 2014 年 1 月 5 日，1 月 15 日，1 月 25 日三天的合肥市出租车交通网数据，并将每天的数据分为白天普通时段、交接班高峰期、夜晚普通时段三个典型时间段处理分析，并在问题一的基础上，增加指标网络订单量与乘客等车时间，建立改进的 K-means 聚类分析评价指数 R_{km} 来量化打车补贴对“打车难”的缓解效果。求解所得模型，得出打车补贴对于缓解白天普通时段普通地区的“打车难”有显著的效果，但对于缓解交接班高峰期以及夜晚普通时段的商业闹市区以及偏远郊区“打车难”没有显著的效果。

问题三，主要研究在第一，二问的基础上，考虑多种合理指标下的“供求匹配”程度以及主要打车软件推行的传统补贴方案对解决“打车难”实质问题的效果进行综合分析。以环境友好，诚信可靠，娱乐拓展，供求匹配作为方案特色发展方向；以数据库系统，“供求匹配”评分系统，补贴管理系统进行平台操作；设立虚拟货币-微币 M ，建立打车软件平台，网络公司，出租车公司，政府与司机，乘客之间的“4+2”六方合作模式体制，建立 RVGAM 缓解效率评分分析模型。以评分优越程度 ϑ 作为缓解能力评价标准，通过新补贴方案缓解效率 H_n ，传统缓解效率 H_o 与新补贴方案的微币补贴总额度 M_{alln} ，传统补贴总额度 M_{allo} 之间的定量与定性关系分析， ϑ 值越趋近于 1，越说明 RVGAM 缓解效率评分算法合理，可知当选取“供求匹配”程度 $C_{m,n}$ 的最好值 0.9673 时， ϑ 值最趋近于 1，得 ϑ 值为 0.957。所以论证得出所设计的新打车软件服务平台通过实施新补贴方案能够很好的缓解打车难问题。

关键词： K-means 聚类分析评价模型 浮动车 RVGAM 缓解效率评分模型 互联网+

一、问题重述

出租车是市民出行的重要交通工具之一，“打车难”是人们关注的一个社会热点问题。随着“互联网+”时代的到来，有多家公司依托移动互联网建立了打车软件服务平台，实现了乘客与出租车司机之间的信息互通，同时推出了多种出租车的补贴方案。

请你们搜集相关数据，建立数学模型研究如下问题：

- (1) 试建立合理的指标，并分析不同时空出租车资源的“供求匹配”程度。
- (2) 分析各公司的出租车补贴方案是否对“缓解打车难”有帮助？
- (3) 如果要创建一个新的打车软件服务平台，你们将设计什么样的补贴方案，并论证其合理性。

二、问题分析

随着“互联网+”时代的到来，打车软件的兴起使得打车服务又迈向新高度。本文在以软件打车服务为基础，收集数据，建立模型，定性评价该条件下出租车的“供求匹配”情况、“缓解打车难”问题。打车软件依靠互联网、服务器以及大数据处理在乘客与司机之间架起桥梁，有效降低了出租车空载率。分析供求匹配程度就是研究某一时间段某一地点乘客需求量与出租车数量之间的线性关系，在选取指标时，将乘客利益相关的条件优先考虑。而针对各公司的补贴方案，应当从不同角度分析对缓解打车难的帮助：优化、聚类不同时间、不同地段、不同的订单都将得到不同的方案。

问题一，主要研究不同时空条件下出租车资源配置的“供求匹配”问题，涉及 7 个时间段和合肥市 3 类 276 个地区。由于数据量庞大，对不同类型的指标的评分权重不同，所以采用改进的 K-means 聚类分析评价模型，设立供求匹配程度，对其进行量化评分。

问题二，主要研究各公司推出的出租车补贴方案对“打车难”的缓解效果。问题二是在问题一的基础上进行研究，沿用改进的 K-means 聚类分析评价模型，要将出租车补贴方案推行前后的乘客打车情况形成对比，并增加相应的评价指标，设立评价指数对缓解效果量化评分。

问题三，主要研究在问题一和问题二的基础上，充分考虑“供求匹配”程度和补贴方案，进而创建新的软件平台，制定新的补贴方案致力于缓解“打车难”程度，通过建立 RVGAM 缓解效度评分模型结合“供求匹配”程度，缓解效度和补贴额度分析出更能综合体现缓解“打车难”程度的 RVGAM 缓解效度评分优越程度 θ ，进行量化评分，并论证其合理性和优越性。

三、模型假设

- [1] 假设不考虑大数据较大变动，即认为数据库中大数据真实可靠；
- [2] 假设不考虑不可抗力的影响；
- [3] 假设忽略意外情况对统计数据的影响；
- [4] 假设“互联网+”时代已经普及；

- [5] 假设打车软件平台有效执行改进方案；
 [6] 假设城市地段人口数量变化幅度不大；
 [7] 假设某时段合肥市各地区的人口量不变；
 [8] 假设不考虑出租车燃油税的变化；
 [9] 假设模型中每辆出租车都配有打车软件司机端；
 [10] 忽略天气对模型的影响，如雨雪天气等。

四、符号说明

对本模型的符号表示说明如下：

OD	出租车起始位置至终端距离
N_C	出租车通行数量
N_D	网络订单量
L	相关载客情况
t_w	等车时间
t_r	乘客网络订单回应时间
α_μ	空载率
α_w	订单等车率
$C_{m,n}$	供求匹配程度
ε	评分值
k	簇数
X	数集
$Q(X_k)$	表示 p_i 和 X_k 的联合概率分布
$Q(p_i X_k)$	代表样本数据点 X_k 属于第 i 个聚类的条件概率
$F(n, p)$	算法目标函数
β	已下订单乘客上车率
P_n	为各目标函数值所占的权重
α_w	已下订单乘客等车率
M	虚拟货币-微币
D_k	乘客发单设定被选择距离
D_m	乘客要到达目的地距离
D_j	接单出租车与乘客之间的距离
X	补贴限制额度
δ	出租车司机诚信度
λ	出租车司机参与此类娱乐项目的程度
H	补贴方案缓解效度
\mathcal{G}	RVGAM 缓解效度评分优越程度

五、问题一：不同时空出租车资源的“供求匹配”程度的分析研究

随着社会经济的发展和人民生活水平的提高,城市的发展越来越快,随之而来机动车的数量也在不断地攀升,城市发展在给人民带来便利的同时,也造成了各种各样负面的影响。交通拥堵在大城市中成为日益严重问题,随着车辆的增加,城市生活问题也变得日益严重,交通问题影响居民的出行效率。解决城市交通拥堵,满足市民出行需求的最根本途径是大力发展公共交通,近年来城市公共交通的发展势头良好,人们对出行的多样化需求及质量要求日益提高,而出租汽车是城市公共交通的重要补充,能够满足具有一定支付能力群体日常出行需求和一般群众特殊出行需求出租汽车正好可以满足人们的这种出行需求。出租汽车能够根据客流在时间上和空间上的不同分布,提供灵活、方便、直达的运输服务,可以很好地满足人们专有乘车空间和私密要求,由此决定它是城市客运系统中不可缺少的客运方式。由于出租汽车行业相关法规规章不健全,导致该行业整体的运营组织能力和服务水平不高;受到企业经营不规范和燃油价格持续上涨等因素的影响,驾驶员和出租汽车公司双方的劳资关系紧张,出租汽车驾驶员的劳动保障问题已受到社会越来越多的关注。但是随着各类打车软件的出现,不仅方便了人们的出行,而且各类公司不同的优惠方案,也带动了市场,使得出租车公司,出租车司机师傅与乘客之间有了更高效的关系。但是针对不同时空下的不同条件,出租车的“供求匹配”问题也随之而来。能否更好通过合理的指标,改善不同时空出租车资源的“供求匹配”程度将关系着司机师傅与乘客的切身利益。基于此,本章共分如下3节:

- 不同时空出租车资源的“供求匹配”程度的分析研究;
- 基于改进的 K-means 聚类分析评价模型的建立;
- 基于改进的 K-means 聚类分析评价模型的处理求解;

问题一的研究思路如下:



5.1 不同时空出租车资源的“供求匹配”程度的分析研究

由于目前对城市出租车拥有量的控制标准中并没有上限规定,现行的《城市道路交通规划设计规范(GB 50220-95)》仅给出了出租车拥有量的下限,即大城市不少于每千人 2 辆,小城市不少于每千人 0.5 辆,中等城市可在其间取值给各城市控制出租车的发展规模与速度带来了困难。而且在不同时间段和不同的空间段,会出现人流量的高峰期和平峰期以及地域的繁华段和偏远段的区别,那么出租车的供应量和需求量都会因此存在明显的差异,所以合理分配出租车资源,高效完成出行任务便是每一辆出租车和乘客的需求。

表 1 近五年合肥每千人拥有出租车数量比例

年 份	运营车数 (辆)	标准运营 车数 (标台)	运营线路 总 长 度 (公里)	客运总量 (万人次)	出租车辆 (辆)	人口 (万人)	每千人拥有
							出租车数量
							的比例
2010	2836	3282	2020	62312	8395	745.7	1.125
2011	3522	4086	2454	67472	10546	752.1	1.402
2012	3724	4386	2695	71190	10546	757.2	1.392
2013	4315	5151	2975	72907	11078	761.1	1.455
2014	5187	6234	3122	74560	11570	769.6	1.503

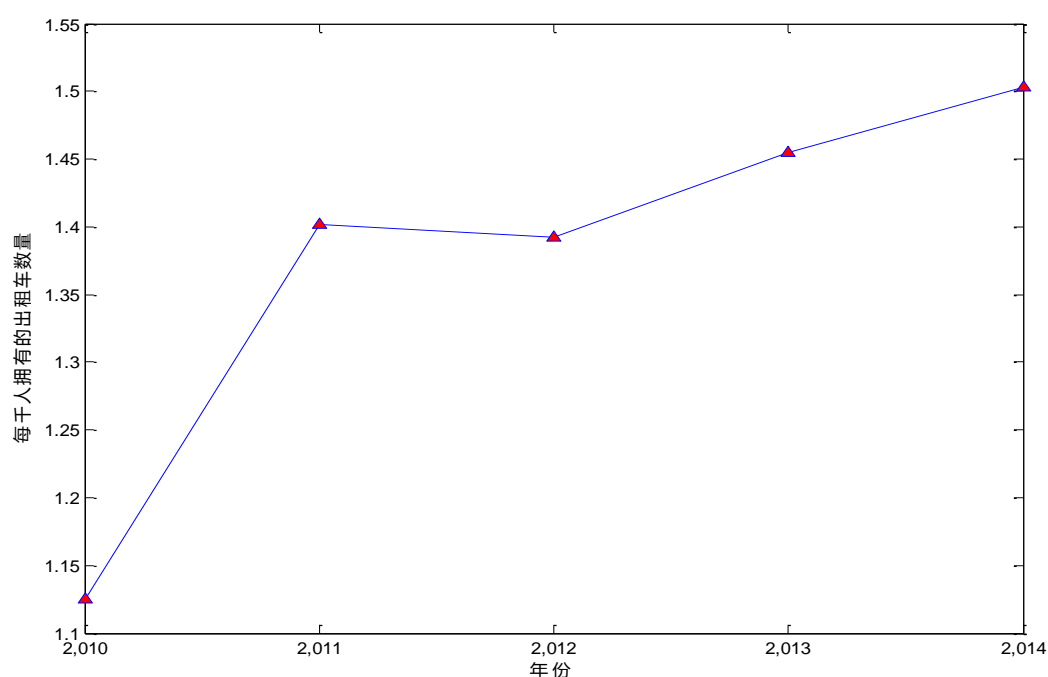


图 1 每千人拥有出租车数量年度变化

本文章以合肥为例，由表可知，近几年合肥市每千人拥有出租车数量比例虽然呈逐年上升趋势，但目前仍在在(0.5,2)之间，由规范可知合肥属于中等城市，能够较好的完成以下分析结果。

5.1.1 合理指标的建立

由于浮动车数据记录了精确的时间信息和车辆位置坐标，因此，对应某运营行为的始末时刻，能够确定出租车出行的 OD，可由相应的匹配后经纬度表示。并且可以记录某一空间或某一时间段下的出租车通行数量 N_c 。结合该空间和时间段下查询打车软件应用商的数据库可了解打车软件的网络订单量 N_D ，相关载客情况 L ，等车时间 t_w ，乘客网络订单回应时间 t_r 以及推算出的空载率 $\alpha_\mu = \frac{N_c - L}{N_c} \times 100\%$ ，订单等车数

$D_s = N_D \alpha_w$ ，建立起相关指标。

5.1.2 供求匹配程度评分值算法的建立

根据上文建立的相关指标，定义供求匹配程度： $C_{m,n} = \frac{\alpha_w N_D}{\alpha_\mu N_C}$ ，由供应量与需求量的关系得， $C_{m,n}$ 越接近 1，该时段该地区的出租车供应量与需求量就越合理，设供求匹配程度评分值为 ε ，建立简单算法，即 $\varepsilon = |1 - C_{m,n}|$ 。由于供应量与需求量之间紧密联系，通过比值能够直观体现二者之间的匹配程度，所以建立评分值后，根据 ε 值的大小可以有效判断供求匹配程度。

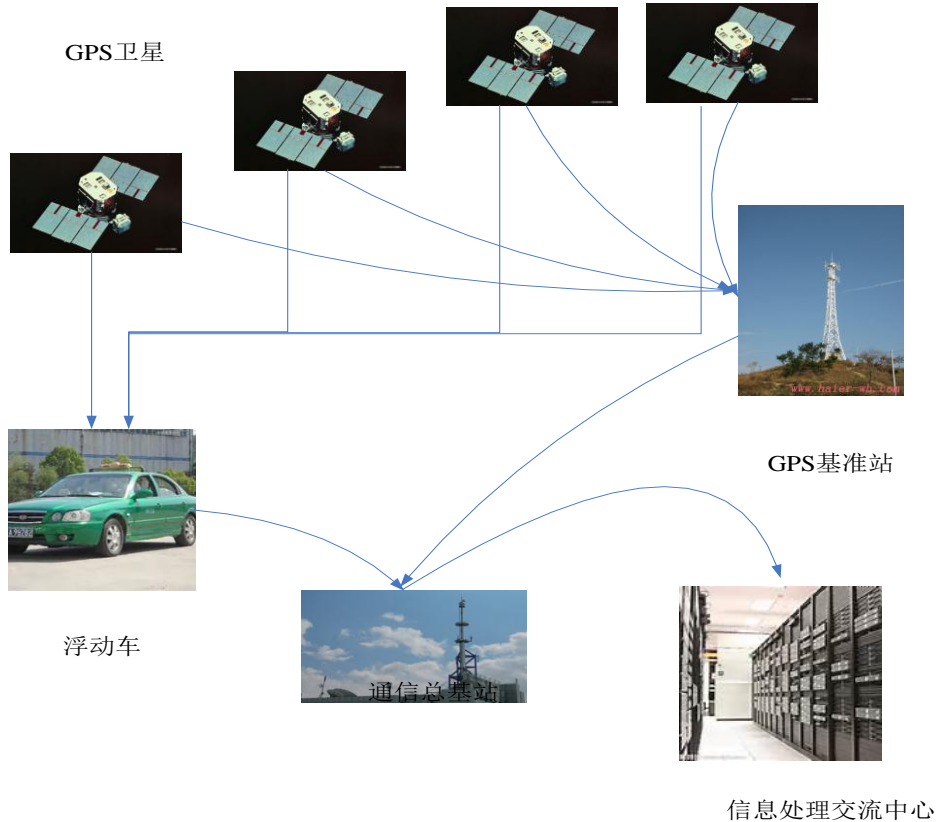


图 2 浮动车 GPS 数据交流

5.2 基于改进的 K-means 聚类分析评价模型的建立

K-means 改进算法是一种混合算法，它将出租车调度及浮动车数据采集匹配地图算法与 K-means 算法进行互补结合，融合二者长处。前者先将出租车作为能够传回 GPS 坐标数据的浮动车，依靠高效的地图匹配，以出租车作为浮动车的 GPS 数据中，包含了记录出租车运营过程的事件触发信息和状态信息，每个运营过程对应于时间上连续的一段记录集。依据相关指标并且依次列成相关的表格。最后利用 K-means 算法对不同条件下的表格数据进行四种不同程度的聚类分析和精确化计算，最终得到精确度高的聚类结果。由于文章分析出租车资源的“供求匹配”程度，则可以通过匹配相关的评分数值大小来描述程度高低。

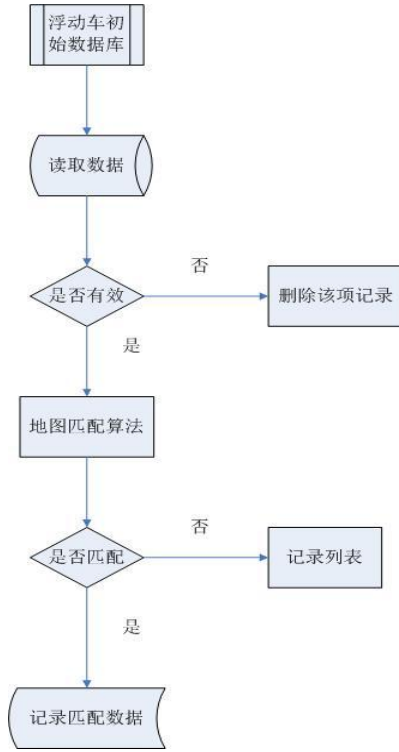


图3 浮动车地图匹配算法程序

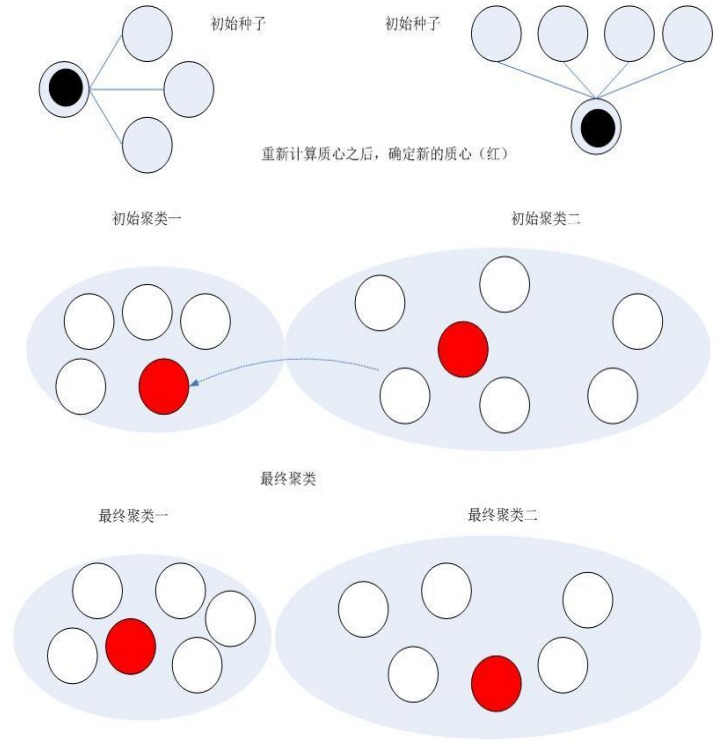


图4 K-means 改进算法示意

5.2.1 改进的 K-means 聚类分析评价模型的数学描述

(1) K-means 改进算法的目标函数

给定一个数据集 $X = \{x_1, x_2, \dots, x_n\}$ ，假设其中 k 个簇的簇类中心为 $P = \{P_1, P_2, \dots, P_n\}$ ，那么有概率聚类的目标函数形式如下：

$$F(n, p) = \sum_{k=1}^k Q(X_k) \sum_{i=1}^k Q(p_i | X_k) \|X_i - p_i\|^2 \quad (5-1)$$

其中， $Q(X_k)$ 表示 p_i 和 X_k 的联合概率分布， $Q(p_i | X_k)$ 代表样本数据点 X_k 属于第 i 个聚类的条件概率。此公式即为 K-means 改进算法的目标函数。

(2) 改进的 K-means 聚类分析评价模型的参数及步骤

设改进的 K-means 聚类分析评价模型的算法参数是 $F(n, p)$ ，表示每个聚类的相关信息，其中 n 代表不同的簇中的数据的数据点数， p 代表所有簇中的数据点数 n 的平均值，即 K-means 算法中的初始簇心的个数。

该混合算法，即改进的 K-means 聚类分析评价模型的算法的实施具体可以分为两个部分，第一步对浮动车数据记录了精确的时间信息和车辆位置坐标进行拓展处理，然后利用改进的 K-means 算法根据 K 值对聚类得到的结果进行处理，最终得到最优的处理结果。混合算法实施的具体步骤介绍如下。

- Step1：确定算法参数 K，以确定聚类过程及最终的簇数；
- Step2：随机抽取 K 个对象，并将它们看作初始聚类的初始簇心；
- Step3：将数据集中剩下的数据分配到初始的 K 个簇中，分配原则是距离最短原则，即将剩余数据分配到距离最近的簇心所在的簇；

- *Step4*: 将每个初始聚类簇中的所有对象重新计算,以此获得其平均值,并根据此平均值选择新的簇心;
- *Step5*: 不断重复步骤(4),直到聚类分析算法的目标函数 $J(n, p)$ 不再发生变化为止。

改进的 K-means 算法根据 K 值处理阶段: 首先检查集合 D 中的每个数据,找出每一个不曾被聚类的数据点,计算该点到所有簇心的距离,将其归类到距离最近的簇中;将所有数据点聚类之后重新计算每个簇的簇心及各个簇的数据点的数目,即 $F(n, p)$;重复上述两个步骤,直到函数 $F(n, p)$ 不再发生改变为止。

5.3 基于改进的 K-means 聚类分析评价模型的处理求解

基于以上内容的介绍,根据改进后的 K-means 算法思想及算法设计步骤,建立基于不同目的的数据处理模型,根据需要分别建立相应的数据模型,并依据该算法思想,最终实现大数据在改进型模型和数据平台上的分析与处理。下面将介绍浮动车 GPS 数据处理模型的建立过程。

5.3.1 浮动车 GPS 数据预处理

文中研究所用的 GPS 数据在数据获取的过程中有可能受到环境、天气及操作误差等方面的影响,导致数据误差较大。所以为了避免产生非常大的误差,就必须对这些大数据进行使用前的预先处理。通过对浮动车的浮动状态以及 GPS 数据的数据特点进行分析,在数据使用之前可以预先对数据进行处理,这样可以使最终的数据处理结果更加理想。

本章研究的方法采用空间和时间限制法,即将 GPS 数据的经纬度确定在某一个空间或时间范围内,对该范围内的数据进行处理研究,范围外的数据根据情况进行剔除。由于限制选择对算法及数据处理模型的正确性及有效性的影响很小,可以忽略不计,所以研究结果可以代表整个空间和时间数据处理的研究结果。文章中划定范围为:空间: 经度坐标范围从东经 116.8409 至 117.1949,纬度坐标范围从北纬 38.3846 至北纬 33.7035; 时间为 7:30~9:30, 9:30~11:00, 11: 00~13:00, 13:00~17:30, 17:30~19:30, 19:30~0:00, 0:00~7:30 这 7 个时间段。划定范围之后,需要对没有意义的数据进行删选。在数据处理过程中,没有意义的干扰数据主要包括以下几个方面:

第一,如果处理器收到的出租车的坐标数据长时间不发生变化或者在极小的范围内变动时,表示车辆静止,速度为零。这样的数据属于没有意义数据,将其进行排除。

第二,如果在某一时间处理器收到的坐标数据在短时间发生跳跃性变化超出了最高运行速度限制级别,则认为这样的数据是无意义的,不能真实反映车辆运行轨迹,所以不作为参考数据使用。

5.3.2 依据改进算法设置限制条件

分析显示合肥市工作日早高峰为 7:30~9:30, 下班高峰为 11: 00~13:00, 晚高峰为 17:30~19:30, 运用改进的 K-means 聚类分析的方法将三个高峰时间段作为时间带分段,把全天的时间分为 7:30~9:30, 9:30~11:00, 11: 00~13:00, 13:00~17:30, 17:30~19:30, 19:30~0:00, 0:00~7:30 这 7 个时间段把每个时间段中的不同地区(共 276 个地区)分为

三个类别，分别为供求匹配程度较好，供求匹配程度一般，供求匹配程度较差。

m 为时间段编号区间(1,7)取整， n 为地区编号区间(1,276)取整， $C_{m,n}$ 为某时间段某地区的供求匹配程度， $F(x)$ 为改进 K-means 聚类分析后所得到的各指标的目标函数值， P_n 为各目标函数值所占的权重。

- 已下订单乘客上车率： β ；
- 已下订单乘客等车率： $\alpha_w = 1 - \beta$ ；
- 空车率： α_μ ；
- 供求匹配程度： $C_{m,n} = \frac{\alpha_w N_D}{\alpha_\mu N_c}$ ，由供应量与需求量的关系得， $C_{m,n}$ 越接近 1，该时段

该地区的出租车供应量与需求量就越合理，已知评分值 ε ，则 ε 值越接近 0，供求匹配程度越好。

5.3.3 模型结果的表示与分析

本文章以合肥为例，在忽略节假日影响的情况下，利用以上算法进行数据处理。首先分析合肥一天中各地区的出租车需求量分布情况。合肥市一天中各地区的出租车需求量与实际出租车数量分布情况图如下，可看出二者变化规律为正相关，且相比较之下差异性不大。

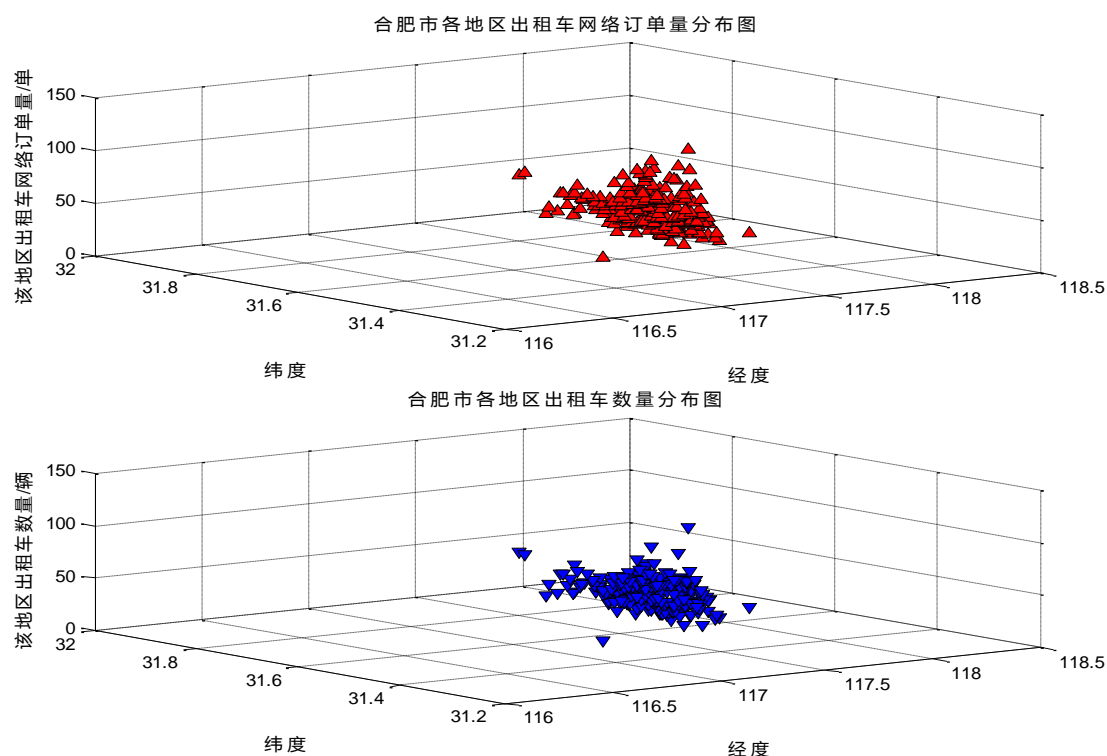


图 5 合肥市单日各地区出租车数量及网络订单量

由图可知合肥不同地区对应出租车需求程度存在明显差异。对图中进行数据处理。利用地图匹配算法，调用浮动车 GPS 数据库数据，大致得到合肥市出租车供应和需求的主要地段，即火车站地段，商业街地段，景点地段以及新市政府地段，如图所示：



图6 火车站地段



图7 商业街地段



图8 景点地段



图9 新市政府地段

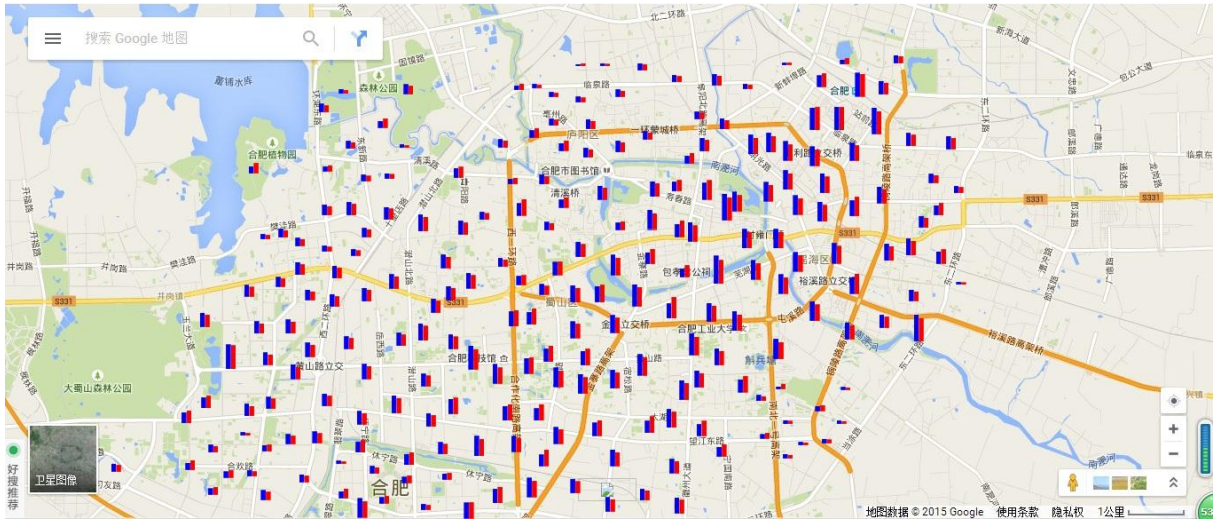


图10 合肥市出租车数量及网络订单量地图匹配

根据图中所示，可知不同地段的出租车供应量和需求量是不同的，往往在火车站，商业圈，景点，市政府等地段的出租车供应量和需求量明显高于其它地段。

综合考虑时空因素，分析显示列表；设各时间段各地区段的供求匹配程度：

表2 各时间段各地区段的供求匹配程度

时间	供求匹配程度
7:30~9:30	$C_{1,1} \sim C_{1,276}$
9:30~11:00	$C_{2,1} \sim C_{2,276}$
11: 00~13:00	$C_{3,1} \sim C_{3,276}$
13:00~17:30	$C_{4,1} \sim C_{4,276}$
17:30~19:30	$C_{5,1} \sim C_{5,276}$

19:30~0:00	$C_{6.1} \sim C_{6.276}$
0:00~7:30	$C_{7.1} \sim C_{7.276}$

利用 K-means 改进算法，运用 SPSS 软件对附录中各类指标在不同时间段以及不同空间条件下进行聚类分析，得到附录中相关重要数据。

因为 F 检验应仅用于描述性目的，因为选中的聚类将被用来最大化不同聚类中的案例间的差别。观测到的显著性水平并未据此进行更正，因此无法将其解释为是对聚类均值相等这一假设的检验。

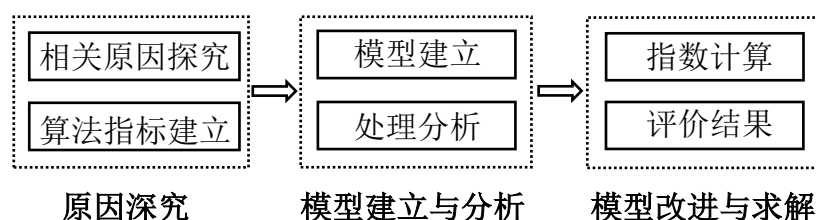
通过编程计算比较得出，得出在时间段 9:30~11:00 的第 71 合肥市政府广场（北纬 31.514601, 东经 117.164568） $\varepsilon = 0.0327$ ，供求比配程度最好；在时间段 17:30~19:30 的合肥火车站（北纬 31.885223，东经 117.317012） $\varepsilon = 0.324$ ，供求比配程度最差。

六、问题二：分析各公司的出租车补贴方案是否对“缓解打车难”有帮助

本章主要研究在上一问的基础上，利用已建立的供应匹配评分制度，将“供应匹配”程度作为是否缓解打车难的评判标准。由于出租车是市民出行的重要交通工具之一，然而近年来“打车难”却是人们关注的一个社会热点问题。随着“互联网+”时代的到来，有多家网络技术公司依托移动互联网建立起打车软件服务平台，实现了出租车司机与乘客之间的信息互通，建立起乘客，出租车司机，出租车公司，政府四方之间的紧密联系。为了拓宽市场，高效权衡供求关系，各家打车软件公司同时推出了多种出租车的补贴方案。然而这些方案是否切实可行，能否有效缓解打车难问题需要进一步探讨，故建立基于改进的 K-means 聚类分析评价算法模型，通过对比城市有无实行补贴方案在不同指标和不同条件下的“供求匹配”程度来量化反映是否缓解打车难问题。基于此，本章共分如下 3 节：

- 打车难的原因探究；
- 基于改进的 K-means 聚类分析评价模型建立与分析；
- 基于改进的 K-means 聚类分析评价模型模型的求解；

问题二的研究思路如下：



6.1 打车难的原因探究

6.1.1 客观原因分析

在移动网络以及智能手机功能的迅速提升下，近年来中国的网民数目呈现飞速增长。早在 13 年的一项调查中，中国网民已达 6.18 亿，其中手机用户规模达到 5 亿，占

整体网民的 80.9%。

伴随着社会和科技的迅速发展，各主流城市均面临着人口过多、工作压力大、生活节奏快等社会问题。在这种背景下，人们对交通运输的要求也越来越苛刻。公交车、地铁作为上班族最依赖的交通运输方式，每逢工作日高峰期，都将会面临一场“上车难”大战。公交车挤破门，地铁站爆满。而出租车作为城市客运交通系统中的润滑剂，理应能够弥补运输资源短暂性不足或为特殊需求人群提供便利，尤其是当夜间公交车收车、地铁停运时。但事实却是每逢上下班高峰、节假日、雨雪天气时半天打不到车。出租车作为城市的移动名片，“打车难”却成了社会的弊病，出租车公司政策、制度、规定、理念严重落后以及在车辆配置方面的不合理使得这种弊病向着越来越严重的趋势发展。

6.1.2 主观原因分析

出租车公司调度配置不合理主要体现在以下几个方面：

- 我们发现，打车难只是发生在某一特定路段、特定时间域内的一种间歇性现象。如商业区晚 6 点至 9 点，平均等车时间较长。而在另一路段的同一时间域内，却存在着大量出租车空驶现象。如郊区的晚 6 点至 9 点，出租车空载率较高。针对这种情况，简单的增加出租车数量是不会有有效的解决问题的，反而会造成更高的空载率以及更会乱的出租车市场。
- 出租车的收费标准存在不合理性。目前各省市的计价器，除了起步价上略有不同，整体计价方式都是按照“每公里价钱*里程数+低速行驶（等候）*时间+起步价”为最终钱数。这种计算方式是针对多年前的路况设计的，对于日新月异的现代社会，显然存在某些弊端。出租车所面向的服务群体是大众，是对公交、地铁在交通网中的补充。出租车以其方便高效，灵活迅速的特点，旨在解决人群短距离出行的困难，而不是长途运输的窘境。在如今的交通路况下，应当适当提高低速行驶和长距离运输的费用。
- 出租车行业主管部门监管监督制度不到位。现出租车公司基本实行 24 小时换岗制，两名司机白班夜班倒替。这样做一定程度上提高了出租车的运营率，但问题也紧随之而来：换班时间大都几种在下班高峰阶段 17:00~19:00，这样就为百姓打车带来不便，距离较远或交通拥堵地段很多司机借换班为借口，直接拒载。下班高峰是一天中出租车需求量最大的时间段，过多的司机换班无法保证充足的出租车在位，自然就造成了“打车难”。

打车软件应运而生，乘客客户端通过软件向终端发送自己的位置以及要去的地方，终端同时收集出租车所在位置与载客状况，服务器处理后向乘客位置周围的出租车司机推送用车需求即发送订单；出租车司机进行抢单，抢到订单后与乘客取得联系，完成订单处理。整个模式如图所示：

应用打车软件优化了出租车资源配置，有效将乘客用车和出租车所提供的服务进行了对接，使得社会上供与应之间可以相互弥补，最大限度的降低出租车空载率。以目前最重要的两款打车软件的功能实际上是把乘客的用车需求与出租车司机的服务供给进行对接，解决了乘客与出租车司机信息不对称的问题，优化了出租车资源的配置。

6.1.3 各打车软件公司的补贴政策

经过相关资料的查询，占国内打车软件主要份额的快的打车与滴滴打车在 2014 年初各推出了各自的打车补贴：

表 3 各打车软件补贴

快的打车			滴滴打车		
时间	乘客返现	司机返现	时间	乘客返现	司机返现
1 月 10 日	10 元	10 元	1 月 20 日	10 元	10 元
2 月 17 日	11 元	5-11 元	2 月 17 日	10-15 元	首单 50 元
2 月 18 日	13 元	5-11 元	2 月 18 日	12 至 20 元	
3 月 4 日	10 元/单	5-11 元	3 月 7 日	6-15 元	
3 月 5 日	5 元		3 月 23 日	3-5 元	
3 月 22 日	3—5 元				

经过观察，我们取快的打车推出补贴方案之前的 1 月 5 日，快的打车推出补贴方案后，滴滴打车推出补贴方案之前的 1 月 15 日，滴滴打车推出补贴方案后的 1 月 25 号，这三天内合肥市 276 个地区在中午交接班高峰时段（11:00~13:00），白天普通时段（9:30~11:00，13:00~17:30），夜晚普通时段（19:30~0:00）三个典型时间段内的出租车相关数据。

6.1.4 算法指标建立

目前，本文在国内外参考文献的基础上，对传统的体系指标做了改进，早先是采用功效系数法，对出租车千人拥有量、空载率和乘客出行方式三个指标来评价某地区“打车难”的程度，对每个指标赋予一个功效系数，之后，分为“优”“中”“差”三个等级，用这三个等级分别对每一个指标进行评价，得出功效系数，最后，把所有的功效系数进行加权平均，得出最终的评价值，此评价方法选取的指标不完善，评价过程也有所欠缺，在此基础上，本论文对这种方法进行了改进，将空载率更换为更直观的供求匹配程度，乘客的出行方式假定为打车出行，在“互联网+”时代的到来与打车软件的影响条件下，增加了网络订单数、该地区出租车数量、乘客等车时间三项指标，用共个五指标对运力规模进行评价，指标体系图如图 11 所示。

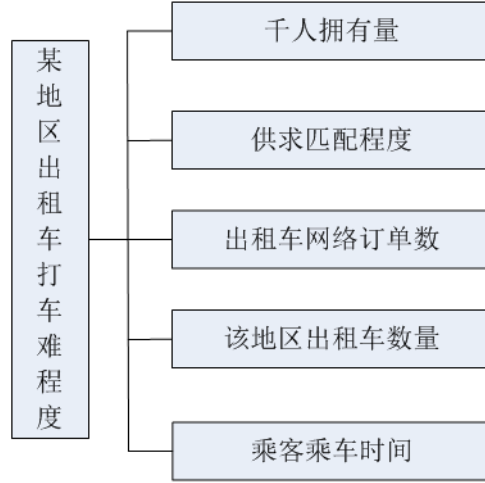


图 11 指标体系

- (1) 千人拥有量:即每千人拥有出租车的数量，反映了出租车的客观需求，实践研究表明，若城市的千人拥有量偏高，则造成出租车空驶率过高需求，可能造成行业内部与行业之间的恶性竞争，不利于社会稳定，供给大于若城市千人拥有量偏低，则会给乘客出行带来不便，造成等车困难，服务水平下降的情况。

- (2) 供求匹配程度: $C_{m,n} = \frac{\alpha_w N_D}{\alpha_\mu N_C}$ ，由供应量与需求量的关系得， $C_{m,n}$ 越接近 1，该时段该地区的出租车供应量与需求量就越合理，已知评分值 ε ，则 ε 值越接近 0，供求匹配程度越好。

- (3) 网络订单数: 即某一时间段内某地区的打车软件下单量，可以直观具体地反应打车软件的补贴政策对出租车市场需求量的影响，计算方法为：

$$\text{出租车网络订单数} = \frac{\text{该时段出租车网络订单总数}}{\text{时间段长度} \cdot \text{地区数量}} \quad (6-1)$$

- (4) 该地区出租车数量: 即某一时段某地区内的出租车数量，可以直观具体地反应打车软件的补贴政策对出租车市场供应量的影响，计算方法如下：

$$\text{该地区出租车数量} = \text{上一时段出租车数量} + \text{流入量} - \text{流出量} \quad (6-2)$$

- (5) 乘客等车时间: 在一定程度上反映了出租车的服务水平，指的是乘客在选择出租车这种出行方式时，在出租车到来时候的等待时间，计算方法为：

$$\text{乘客平均等车时间} = \frac{\sum \text{等车时间}}{\text{总候车次数}} \quad (6-3)$$

6.2 基于改进的 K-means 聚类分析评价模型的建立与分析

6.2.1 基于改进的 K-means 聚类分析评价模型的建立

设需要进行评价的众对象是一个集合，定义集合为 $Q = (p_1, p_2, \dots, p_n)$ ，则令 A_i 为 $X_i (i=1, 2, 3 \dots 6)$ 的权重系数， $X_{ij} (i=1, 2, 3 \dots 6; j=1, 2, \dots, n)$ 为第 j 个评估对应于指标 X_i 值的功效系数值。可以得出，第 k 个评价的对象所对应的运力规模评价指数的代数式是：

$$R_{kn} = \sum_{i=1}^5 A_i X_{ik} \quad (6-4)$$

通过改进的 K-means 聚类分析评价指数，即 R_{kn} 越大，打车补贴方案对缓解“打车难”的效果就越显著。

由于 R_{kn} 取不同的区间值，就会对应不同的评估结果，那么不同的区间值对应的评价结果如下所示：

$$R_{kn} \in \begin{cases} (0,0.6] & \text{差} \\ (0.6,0.8] & \text{中} \\ (0.8,1.0] & \text{优} \end{cases} \quad (6-5)$$

6.2.2 模型指标权重的建立

文章所提五项指标根据可重要程度，分别赋予与之相应的权重系数。可知等待时间的多少可以反应出租车服务水平的高低，而出租车需要最大化的满足乘客的乘车需求，因此，等待时间的这项指标显得最为重要；当然出租车空载率可以反映出租车公司的经营效益，而经营效益与出租车的拥有量也有重大影响；里程利用率在一定程度上可以反映出租车的运行效率，并且运行效率的高低决定了出租车的利用效率和服务质量；简要分析可知公交车出行结构比重、万人拥有量和道路网密度是交通服务的基本指标，重要程度一般。根据公共交通的服务性质，即需要达到社会效益最大化，那么经过上述论证，再根据五项指标的重要程度，分别赋予权重值 $A_1: A_2: A_3: A_4: A_5=3:4:3:4:6$ 。

6.2.3 模型指标评价功效系数的确定

通过改进的 K-means 聚类分析评价模型算法得出该五个指标在不同的区间内的概率聚类的条件概率 $Q(p_i|X_k)$ ，根据指标评价系数规定的区间 (0,1]，通过式 (5.1) 计算得出概率聚类的目标函数 $F(n,p)$ ，即指标评价功效系数。

具体权重、评价功效系数见表 4。

表 4 评价功效系数

指标	权重	指标值	功效系数	指标	权重	指标值	功效系数
千人拥有量/辆/千人	0.15	0~0.5	0.5	网络订单数/单	0.15	20~35	0.3
		0.5~1.0	0.7			35~50	0.5
		1.0~1.5	1.0			50~65	0.7
						65~80	1.0
供求匹配程度	0.2	0.50~0.70	0.5	出租车数量/辆	0.20	20~35	0.3
						35~50	0.5
		0.70~0.90	0.7			50~65	0.7
						65~80	1.0

		0.90~1.10	1.0	乘客等车时间/分钟	0.3	2~3.5	1.0
		1.10~1.30	0.7			3.5~5	0.7
						5~6.5	0.5
						6.5~8	0.3

通过 SPSS 数学软件运用改进的 K-means 聚类分析评价模型进行评价,可以全方位、多角度的从各项影响指标的重要程度出发,对打车补贴缓解合肥市出租车“打车难”问题的效果程度进行评价,有效的避免了由于评价指标选择不当或者选择较少而造成对评价结果的影响。由于评估结果准确率高,可参考性较强,而且人为影响因素较小,所以是一种科学的评价方法。

6.2.4 基于改进的 K-means 聚类分析评价模型的处理与分析

经过对相应数据进行分类处理,可以得出以下结论,做出以下图示:

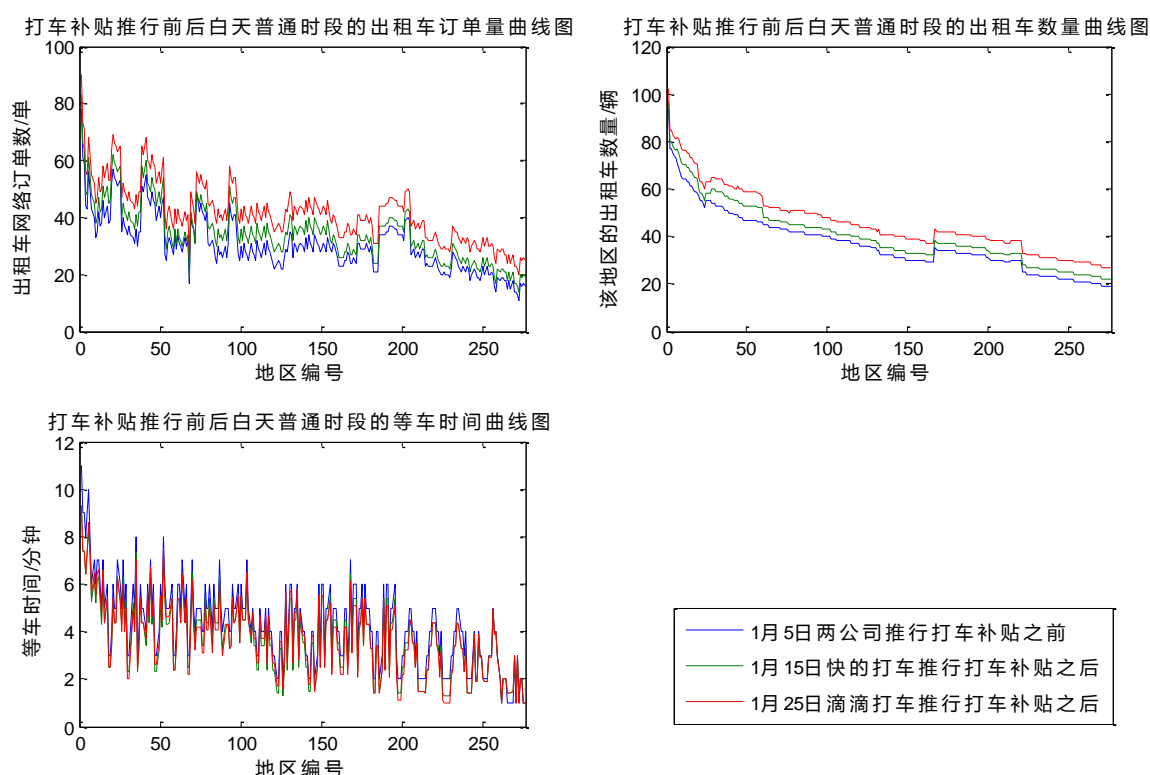


图 12 打车补贴推行前后白天普通时段等车相关指标变化

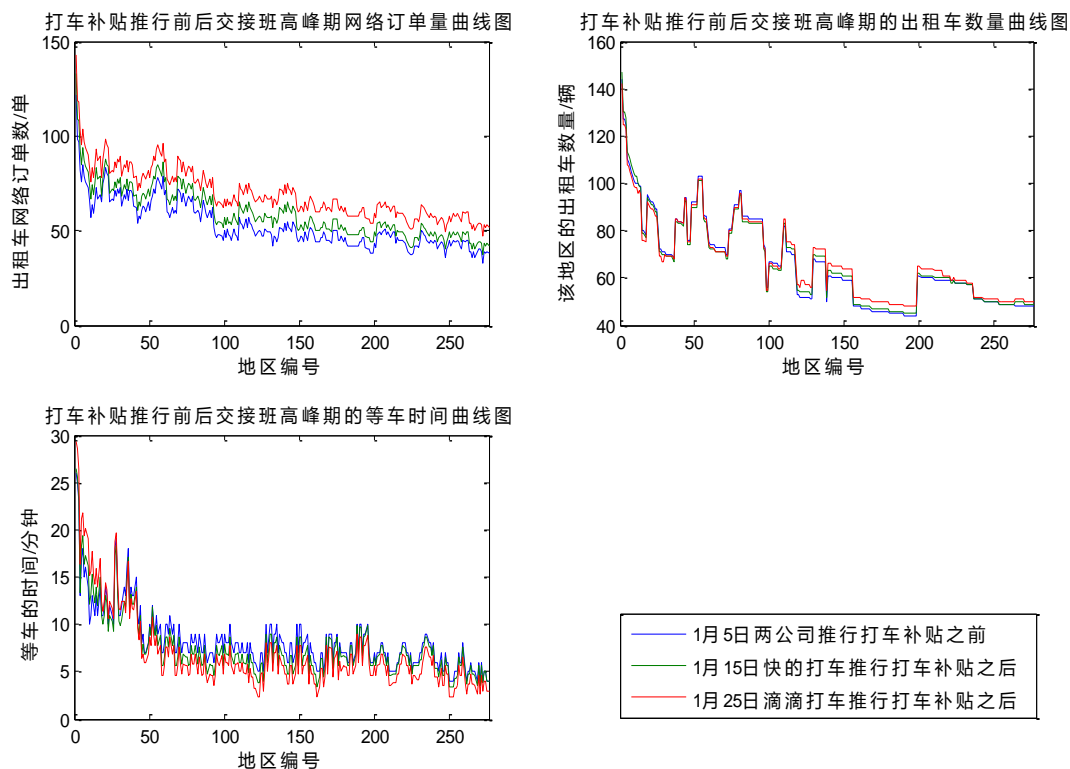


图 13 打车补贴推行前后交接班高峰期等车相关指标变化

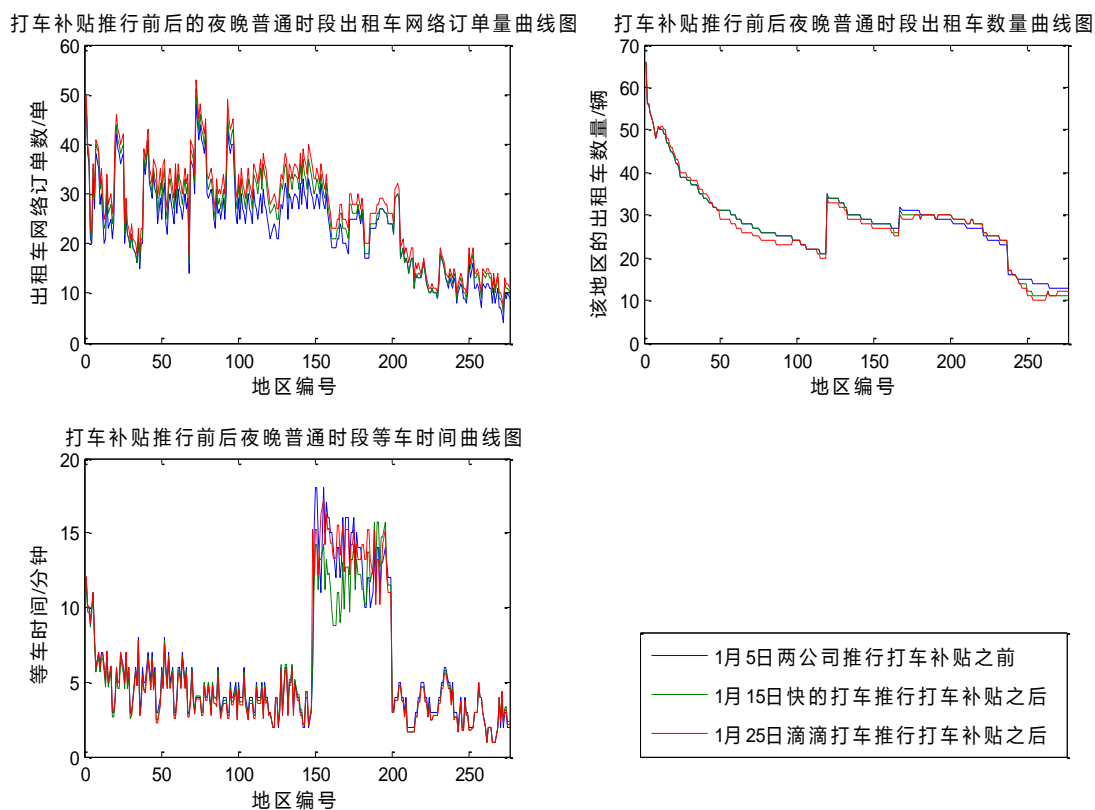


图 14 打车补贴推行前后夜晚普通时段等车相关指标变化

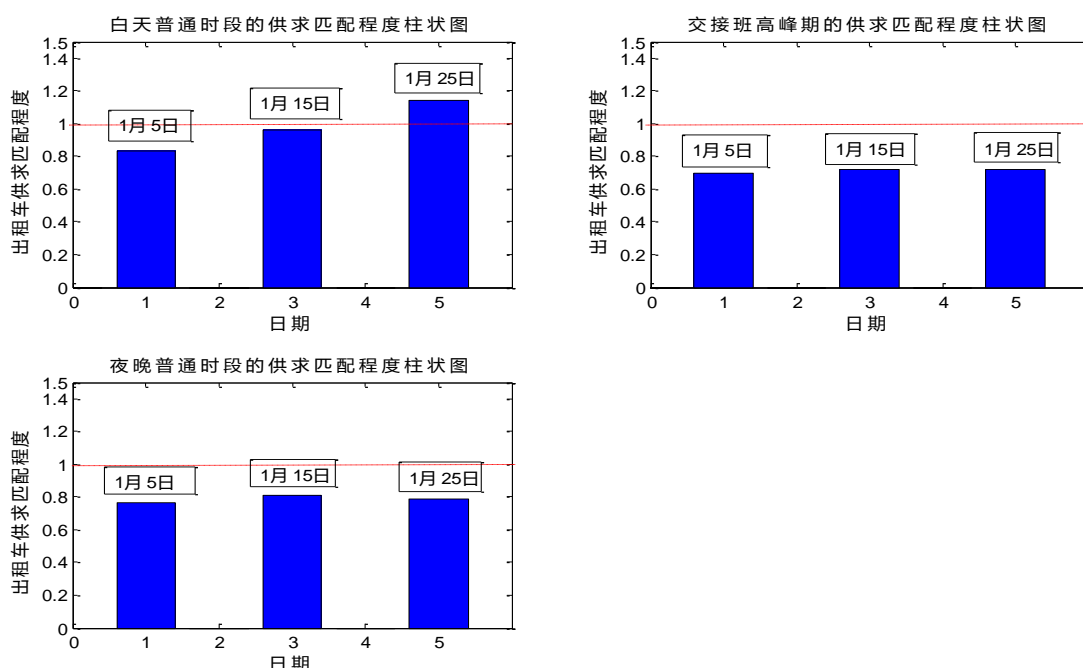


图 15 不同主要时段出租车供求匹配程度

随着两家打车软件公司先后推出各自的打车补贴方案，从图 12 可看出，在白天普通时段合肥市各地区的出租车的网络订单数在不断地增加，出租车的实际数量也随网络订单数增加，乘客的等车时间相对减少，说明打车补贴方案的推行对缓解白天普通时段合肥市各地区的“打车难”程度有显著地帮助。

从图 13 可看出，在交接班高峰期，合肥市各地区的出租车的网络订单数在大幅度地增加，出租车的实际数量也随网络订单数增加，只有商业闹市区（地区 1~30）和偏远郊区（地区 150~200）的出租车数量相对不变，甚至减少；而且乘客的等车时间相对减少，只有商业闹市区的等车时间有反弹增加的现象，说明打车补贴的推行对缓解交接班高峰期，合肥市非闹市商业地区的“打车难”程度有显著地帮助，而对高峰闹市区起不到明显的作用。

从图 14 可看出，在夜晚普通时段，合肥市各地区的出租车的网络订单数会出现不同程度的增加，其中偏远郊区的网络订单增加比较明显，但是出租车的实际数量却没有明显幅度的增加甚至有小幅度的减少，偏远郊区的出租车数量出现明显的减少，在商业闹市区和普通路段地区乘客的等车时间相对减少，但在偏远郊区夜晚的等车时间并没有任何改观，说明打车补贴的推行对缓解夜晚普通时段合肥市非闹市商业地区的打车难有显著地帮助，而对夜晚偏远郊区“打车难”程度起不到明显的作用。

从图 15 可看出，在两家打车软件公司推出打车补贴先后的三个典型时间段合肥市出租车供求匹配程度，其中白天普通时段的 C_n 在打车补贴推行后接近 1 的程度更大，说明打车软件的补贴方案对缓解白天普通时段的“打车难”有一定的作用，而交接班高峰期和夜晚普通时段的 C_n 在打车补贴推行后并没有更接近 1 的趋势，尤其在 1 月 20 日滴滴打车推出打车补贴后，交接班高峰期和夜晚普通时段的 C_n 均有更偏离 1 的趋势，说明打车软件的补贴方案对缓解交接班夜晚普通时段的“打车难”程度没有明显的作用。

6.3 改进的 K-means 聚类分析评价模型的评价求解

6.3.1 改进的 K-means 聚类分析评价模型的评价指数计算

(1) 千人拥有量：根据上述调查，2014 年合肥市共拥有出租汽车 11570 台，城市人口达 769.6 万人，计算得合肥市千人拥有量为 1.503 台。

(2) 其他指标：

功效值与权重的计算选取见表 5。

表 5 功效值与权重计算值

日期	时段	供求匹配程度	网络订单量/单	出租车数量/辆	乘客等车时间/分钟
1 月 5 日	白天普通时段	0.831	30.48	36.98	4.25
	交接班高峰时期	0.697	53.07	65.90	8.26
	夜晚普通时段	0.765	23.48	28.23	5.98
1 月 15 日	白天普通时段	0.962	34.68	40.47	3.70
	交接班高峰时期	0.720	59.31	66.10	9.13
	夜晚普通时段	0.805	25.46	28.29	5.43
1 月 25 日	白天普通时段	1.144	41.64	45.73	3.49
	交接班高峰时期	0.719	69.25	67.35	8.35
	夜晚普通时段	0.791	27.28	27.74	6.13

(3) 根据公式 (2-1) 计算得到打车补贴方案推行前后三个日期中三段典型时间段的改进的 K-means 聚类分析评价指数，见表 6。

表 6 不同时间评价值

日期	时段	评价指数	日期	时段	评价指数	日期	时段	评价指数
1 月 5 日	白天普通时段	0.61	1 月 15 日	白天普通时段	0.73	1 月 25 日	白天普通时段	0.79
	交接班高峰时期	0.52		交接班高峰时期	0.54		交接班高峰时期	0.54
	夜晚普通时段	0.56		夜晚普通时段	0.60		夜晚普通时段	0.59

6.3.2 改进的 K-means 聚类分析评价模型的评价结果

对合肥市两打车软件公司推行的打车补贴前后的改进的 K-means 聚类分析评价指标计算得，在 1 月 5 日，1 月 15 日，1 月 25 日这三天，白天普通时段的评价指数明显递增，交接班高峰时期的评价指数平稳，虽有所递增，但不明显，夜晚普通时段的评价指数呈现先增后减的形式。综上所述，打车补贴对于缓解白天普通时段普通地区的“打车难”程度有显著的作用，但对于缓解交接班高峰期以及夜晚普通时段商业闹市区以及偏远郊区的“打车难”程度没有显著的作用。

七、问题三：创建新打车软件服务平台

本章主要研究在第一问，第二问的基础上，考虑不同时间，不同空间下“供求匹配”程度需要，综合分析已有的不同公司对缓解打车难而提出的打车补贴方案，结果认为有一定缓解效果，但总体效果不佳，所以需要创建一个新的打车软件服务平台，重新设计相关的补贴方案，并论证其合理性，在利用问题一的方法和借鉴问题二的方案建立 RVGAM 缓解效度评分模型进行分析求解。基于此，本章共分如下 3 节：

- 创建新打车软件平台；
- RVGAM 缓解效度评分模型的建立；
- RVGAM 缓解效度评分模型的求解；

问题三的研究思路如下：



7.1 创建新打车软件平台

本着创新要求，吸纳滴滴打车和快的打车等打车软件平台的优秀补贴方案思想，以环境友好型，诚信可嘉型，娱乐拓展型，供求匹配型四个主要特色目标作为补贴方案发展方向，基于此创建新打车软件平台，旨在减少环境污染，缓解交通压力，节约各自时间，提高双方效率的同时，通过软件平台合理调整补贴，宏观调控出租车与乘客之间的供需关系，最终缓解打车难。

7.1.1 预先设定服务平台规则

新的打车软件平台先与大众网络娱乐公司建立合作关系，设立双方可直接或间接交易的虚拟货币-微币 M 。平台需建立出租车乘客 GPS 定位数据库系统，实时掌握位置时间关系；还需建立“供求匹配”评分系统，根据评分宏观调控供需关系；最后建立补贴管理系统，通过出租车司机和乘客之间下单前后进行补贴，将补贴转化为虚拟货币，打入司机和乘客在打车软件平台设立的各自的账户中，司机和乘客可以将补贴返现，也可以参与网上指定大众平台买卖商品。

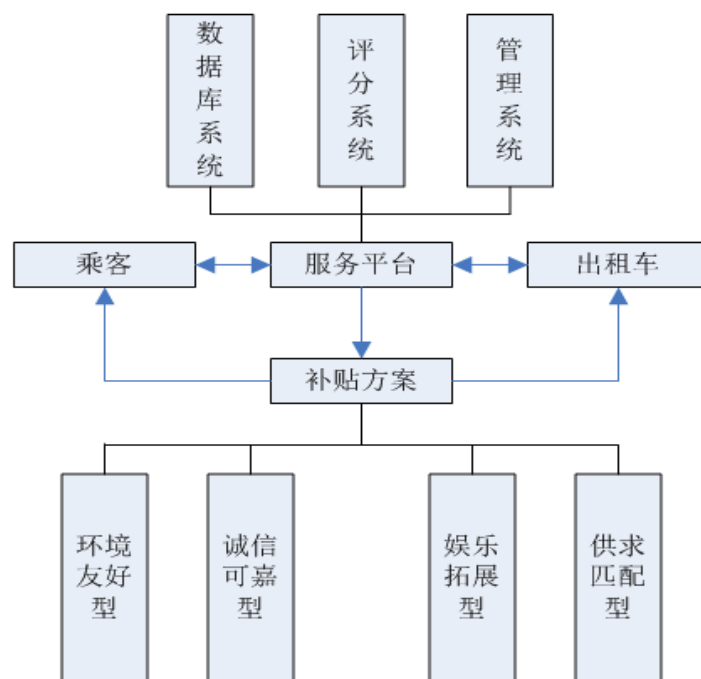


图 16 新服务平台工作模式

通过与出租车公司调控，政府监督条件下，建立起打车软件平台，大众网络娱乐公司，出租车公司，政府与司机，乘客之间的“4+2”六方合作模式体制，在方案推动中，通过各类补贴有效缓解打车难问题，又推动虚拟货币发展，刺激消费，使得司机和乘客为缓解打车难情况下获得补贴，这样就能通过服务平台规则使得以上情况形成良性循环，最终缓解打车难问题。

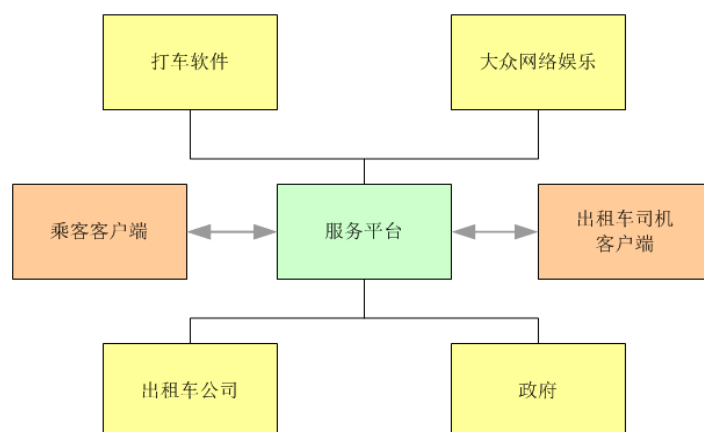


图 17 “4+2”模式示意

7.1.2 乘客补贴方案分析

- 1.环境友好型条件下：乘客发单可设定被选择距离 D_k 和目的地距离 D_m ，在被选择距离内可预约选择意向出租车，该出租车会收到特殊信号。如果接单出租车与乘客之间的距离 D_j 加上目的地距离 D_m 之后值越小，说明对环境影响程度越小。设定补贴限制额度为 X_{hc} ，则该条件下微币补贴：

$$M_{hc} = \left(1 - \frac{D_j}{D_k}\right) \cdot \frac{D_j}{D_m + D_j} \cdot X_{hc} \quad (7-1)$$

- 2.诚信可嘉型条件下：乘客发单和取消订单具有较高的自由度，但爽约情况下必定带来一定的经济损失和诚信缺失。设定乘客发单数目 S_{sc} ，发单后爽约数目为 S_{bc} ，当发单数目远高于爽约数目时，体现乘客的诚信度较高。设定补贴限制额度为 X_{xc} ，则该条件下平均微币补贴：

$$M_{xc} = \left(1 - \frac{S_{bc}}{S_{sc}}\right) \cdot \frac{1}{S_{sc} - S_{bc}} \cdot X_{xc} \quad (7-2)$$

- 3.娱乐拓展型条件下：乘客多在互联网中参与各种大众平台，而且关联度较高的主推游戏和网上购物。打车软件则与游戏和网上购物公司建立合作关系，使得补贴能够在之间流通，拓展关联度，刺激更多消费，巩固利益链，促进有利循环。根据三方之间数据交流，可设乘客参与此类娱乐项目的程度为 γ ，设定补贴限制额度为 X_{tc} ，则该条件下微币补贴：

$$M_{tc} = \gamma \cdot X_{tc} \quad (7-3)$$

- 4.“供求匹配”型条件下：根据第一问和第二问可知，不同时间和不同地段条件下，“供求匹配”程度 $C_{m,n}$ 存才一定的差异，由供应量与需求量的关系得， $C_{m,n}$ 越接近 1，该时段该地区的出租车供应量与需求量就越合理。往往“供求匹配”程度能够反映出时间里平峰期和高峰期，空间里繁华段和偏远段的出租车供需关系。往往高峰期，繁华段乘客需求量大，但出租车越是满足乘客需求，越会导致高峰期延长，繁华段拥堵；往往平峰期，偏远段需求量小，出租车不愿意前往，造成不接单，放空车。为避免此类事件，可对乘客进行相关补贴，合理指导乘客出行。设定补贴限制额度为 X_{cc} ，则该条件下微币补贴：

$$M_{cc} = C_{m,n} \cdot X_{cc} \quad (7-4)$$

7.1.3 出租车司机补贴方案分析

- 1.环境友好型条件下：出租车司机可根据乘客设定被选择距离 D_k 和目的地距离 D_m 以及在被选择距离内是否收到被预约选择的特殊信号作为意向出租车的条件下进行自主接单。如果接单时所在位置与乘客之间的距离 D_j 越小或者乘客与目的地距离 D_m 越小，说明能够高效到达乘客所在位置或者对环境的影响程度越小。设定补贴限制额度为 X_{hs} ，则该条件下微币补贴：

$$M_{hs} = \left(1 - \frac{D_j}{D_k}\right) \cdot \frac{D_j}{D_m + D_j} \cdot X_{hs} \quad (7-5)$$

- 2.诚信可嘉型条件下：对出租车司机建立诚信平台系统，定义诚信度 δ ，主要通过诚信度与补贴挂钩，通过诚信度高低，涨跌变化适当调整补贴额度，并且诚信度高的出租车司机可优先抢单。

(1) 顺风车：在已经载客情况下，若沿途有顺路乘客，经过车上乘客同意后可继续接单搭载，则每一笔单可提高诚信度 $a\%$ 。

(2) 道德信用：若出现拒载或者道德问题被乘客举报，则每一笔单可降低诚信度 $b\%$ 。

(3) 合理接单：当有乘客发单后迟迟未有接单的情况定义为“差单”，当乘客发单后，不论时间，空间或者乘客增加费用的情况下，使得接单比较迅速的情况定义为“好单”。当每接一笔“差单”，可提高诚信度 $c\%$ ；当每接一笔“好单”，可降低诚信度 $d\%$ 。设定补贴限制额度为 X_{xs} ，则该条件下微币补贴：

$$M_{xs} = \delta \cdot (a\% + c\% - b\% - d\%) \cdot X_{xs} \quad (7-6)$$

- 3.娱乐拓展型条件下：出租车司机也多在互联网中参与各种大众平台，而且关联度较高的首推游戏和网上购物。出租车司机所属于出租车公司和所使用打车软件公司则与游戏和网上购物公司建立合作关系，使得补贴能够在之间流通，拓展关联度，刺激更多消费，巩固利益链，促进有利循环。根据三方之间数据交流，可设出租车司机参与此类娱乐项目的程度为 λ ，设定补贴限制额度为 X_{ts} ，则该条件下微币补贴：

$$M_{ts} = \lambda \cdot X_{ts} \quad (7-7)$$

- 4.“供求匹配”型条件下：根据第一问和第二问可知，不同时间和不同地段条件下，“供求匹配”程度 $C_{m,n}$ 存才一定的差异，由供应量与需求量的关系得， $C_{m,n}$ 越接近 1，该时段该地区的出租车供应量与需求量就越合理。往往“供求匹配”程度能够反映出时间里平峰期和高峰期，空间里繁华段和偏远段的出租车供需关系。往往高峰期，繁华段，出租车越满足乘客需求，越会导致高峰期延长，繁华段拥堵；往往平峰期，偏远段需求小，出租车不愿意前往，造成不接单，放空车。为避免此类事件，可对出租车司机进行相关补贴，鼓励或限制车辆出行，有效调控交通。设定补贴限制额度为 X_{cs} ，则该条件下微币补贴：

$$M_{cs} = C_{m,n} \cdot X_{cs} \quad (7-8)$$

7.2RVGAM 缓解效度评分模型建立

所谓缓解效度评分（RELIEVE VALIDITY GRADE ANALYSE）是指将打车软件平台实施的补贴方案对缓解打车难问题的程度进行量化，建立缓解效度评分模型，对传统补贴方案的缓解效度与新设想方案的缓解效度进行对比分析，建立分析算法，研究其稳定程度和效度走向趋势，进而得到新方案的缓解效果程度。

7.2.1RVGAM 缓解效度评分分析算法缓解效度 H 的确定

(1) 缓解效度 H 的相关关系

首先定义缓解效度 H ，由于不论是传统补贴方案还是新设想方案都是基于缓解打车难的情况下建立推广实施的，那么最终的补贴额度的高低往往影响缓解程度。根据经验，当补贴较少时，出租车司机和乘客往往倾向于传统招手打车方式，就会增加乘客的打车难度和司机的接单难度；而当补贴较高时，司机和乘客为积累补贴得到的虚拟货币-微币 M ，往往依赖打车软件不同补贴方案，就会使得乘客愿意下单打车，司机愿意接单接车，那么打车将更加容易。所以微币补贴总额度 M_{all} 的多

少与缓解效率程度 H 呈正相关。

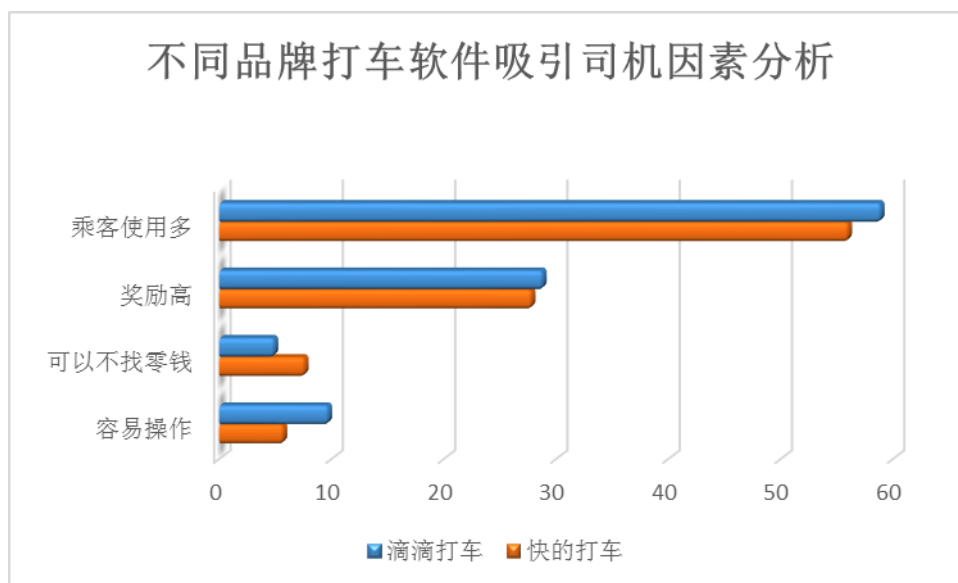


图 18 不同品牌打车软件吸引司机因素分析

由第一问知“供求匹配”程度 $C_{m,n}$ 越趋近于 1，说明该时间该空间条件下的供需的平衡度较好，体现打车较方便。在这里以 $|1-C_{m,n}|$ 作为一种度量，当 $|1-C_{m,n}|$ 值越小，说明打车越容易，即越能体现缓解打车难，即 $|1-C_{m,n}|$ 与缓解效率程度 H 呈负相关。

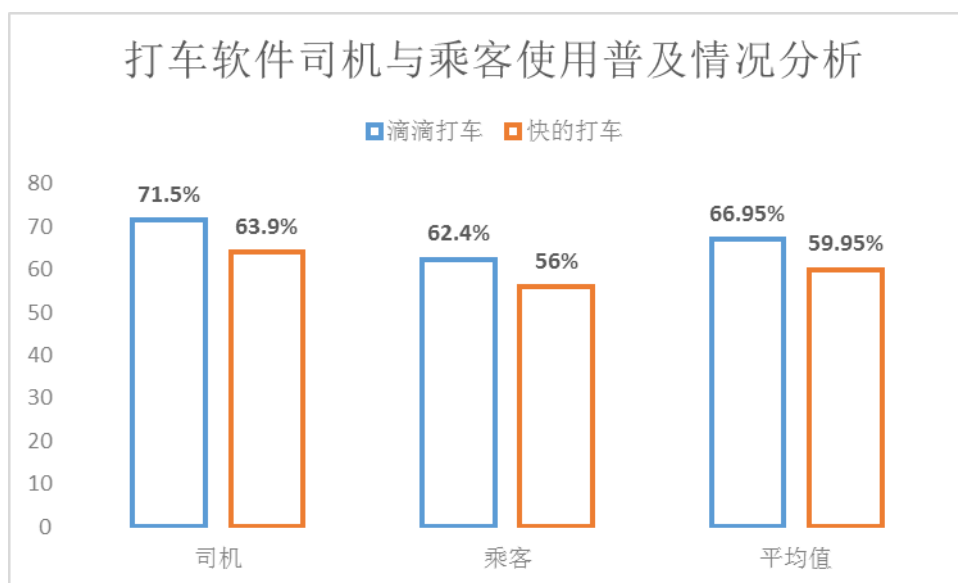


图 19 打车软件司机与乘客使用普及情况分析

(2) 缓解效率 H 的定义表达

综合上文所述，为使得微币补贴总额度 M_{all} 的多少与度量 $|1-C_{m,n}|$ 都成正相关，可以定义缓解效率 H 的表达式：

$$H = \frac{M_{all}}{|1-C_{m,n}|} \quad (7-9)$$

由于市场上打车软件 APP 主要为滴滴出行和快的打车，在这里将二者所提供的补贴方案共同视为传统补贴方案。实际上，两个公司一直在变化补贴方案，在这里将二者结合进行平均，得到传统补贴方案的微币补贴总额度 M_{allo} ，进而得到传统补贴方案缓解效率 H 的表达式：

$$H_o = \frac{M_{allo}}{1 - C_{m,n}} \quad (7-10)$$

本文设计出新的打车软件服务平台，根据上文的分析可知，每一单的补贴额度都包括乘客和司机的各类情况，所以得到新补贴方案的微币补贴总额度 M_{alln} ，表达式为：

$$M_{alln} = M_{hc} + M_{xc} + M_{tc} + M_{cc} + M_{hs} + M_{xs} + M_{ts} + M_{cs} \quad (7-11)$$

所以进而得到新补贴方案缓解效率 H 的表达式：

$$H_n = \frac{M_{alln}}{1 - C_{m,n}} \quad (7-12)$$

7.3 RVGAM 缓解效率评分分析算法的模型求解

通过计算传统补贴方案的微币补贴总额度 M_{allo} 和新补贴方案的微币补贴总额度 M_{alln} ，进而通过图表可对比出传统补贴方案缓解效率 H_o 和新补贴方案缓解效率 H_n 的稳定程度和走向趋势，可以解出二者的优越程度。

7.3.1 RVGAM 缓解效率评分优越程度的确定

根据文章要求，如果新补贴方案缓解效率 H_n 高于传统补贴方案缓解效率 H_o ，则说明新的打车软件服务平台能够更好地缓解打车难问题，为了更好地反映两个效率之间的对比程度，更好地论证 RVGAM 缓解效率评分算法的合理性和所设计的新打车软件服务平台高效的缓解能力，所以在这里定义 RVGAM 缓解效率评分优越程度：

$$g = \frac{H_n - H_o}{H_o} + 0.4 \quad (7-13)$$

根据市场规则可知，新补贴方案缓解效率 H_n 不可能远远高于传统补贴方案缓解效率 H_o ，一般高幅不会超过 40%，如式中所示加常量 0.4，那么根据式（7-13）中可知，当 RVGAM 缓解效率评分优越程度 g 越趋近于 1，越说明 RVGAM 缓解效率评分算法的合理性越好，所设计的新打车软件服务平台具有高效的缓解能力，能够切实缓解打车难问题。

7.3.2 RVGAM 缓解效率评分分析算法结果表示

通过近期传统打车软件公司实施的补贴方案，如图 20 所示，可计算出传统补贴方案的微币补贴总额度 $M_{allo}=8.22$ ，以合肥市为例，由第一问得知选取“供求匹配”程度 $C_{m,n}$ 最好值 0.9673，进而求得传统补贴方案缓解效率 $H_o=251.376$ 。

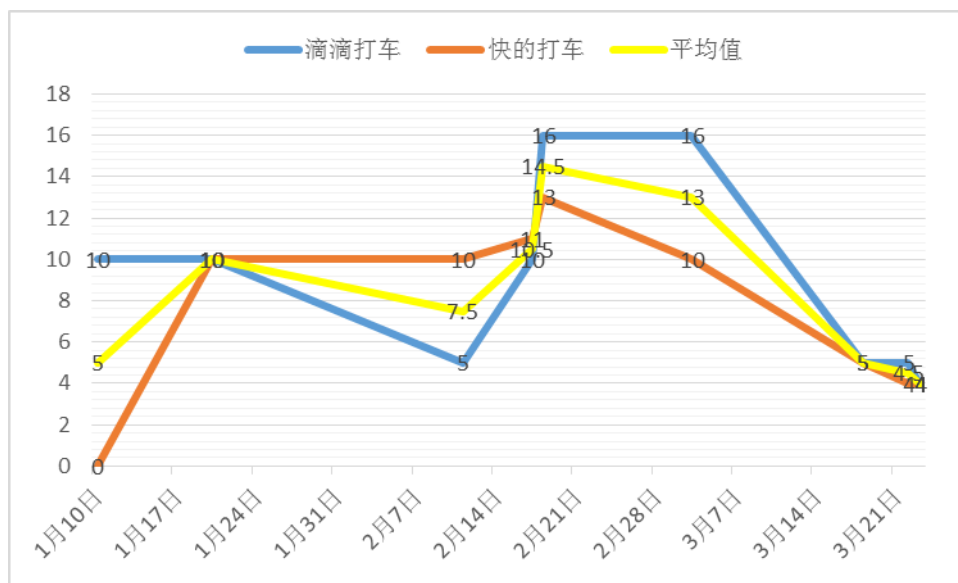


图 20 主要公司补贴方案走势

通过所设计的新打车软件服务平台实施的补贴方案，对所有条件下的参数赋值，设定补贴限制额度 $C_{m,n} = M_{hc} = M_{xc} = M_{tc} = M_{cc} = M_{hs} = M_{xs} = M_{ts} = M_{cs} = 4$ ，根据经验可知总的变量影响不会超过 20%，所以取最大影响为 20% 情况下，可计算得出新补贴方案的微币补贴总额度 $M_{alln} = 12.8$ ，以合肥市为例，由第一问得知选取“供求匹配”程度 $C_{m,n}$ 最好值 0.9673，进而求得新补贴方案缓解效率 $H_n = 391.437$ 。

当“供求匹配”程度 $C_{m,n}$ 取一般值时，即从最好值 0.9673 到最差值 1.3240 取值代入，分别将传统补贴方案缓解效率 H_o 和新补贴方案缓解效率 H_n 值做在图中。

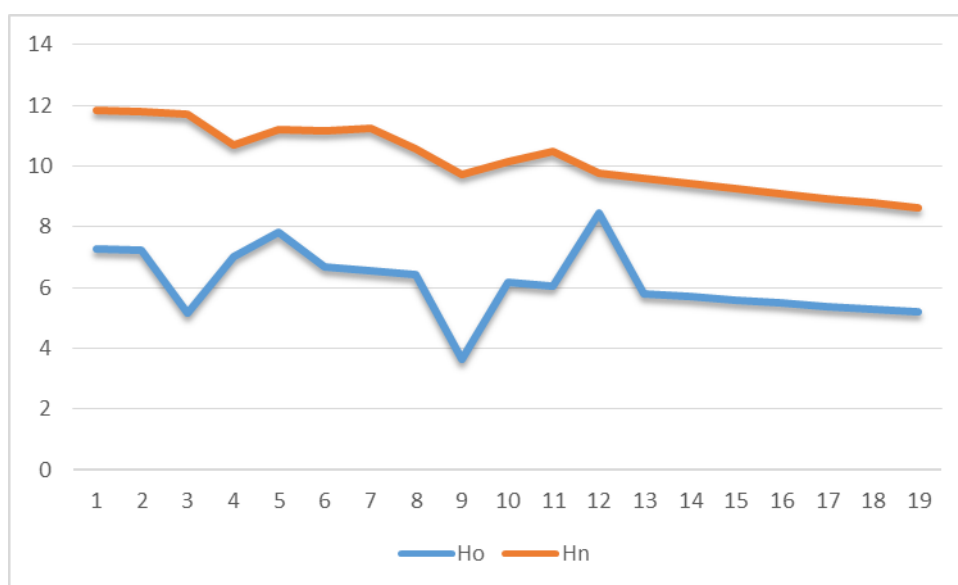


图 21 传统补贴方案缓解效率 H_o 和新补贴方案缓解效率 H_n 值变化趋势

根据图中可知，传统补贴方案缓解效率 H_o 浮动较大，呈下降趋势，新补贴方案缓解效率 H_n 比较稳定，呈平稳趋势。则可以说明新方案稳定性较好，缓解能力有较好态势。将二者代入式 (7-12) 中求出 ϑ ，可知当选取“供求匹配”程度 $C_{m,n}$ 的最好值 0.9673 时， ϑ 值最趋近于 1，得 ϑ 值为 0.957。所以论证得出所设计的新打车软件服务平台通过

实施新补贴方案能够很好的缓解打车难问题。

八、模型评价

8.1 模型优点

1、研究了不同时空出租车资源的匹配程度问题，具有相当的难度与一定的创新性。建立的改进的 K-means 聚类分析评价模型，着重关注了出租车市场供应量与需求量的匹配指标，考虑的情况也更为复杂与贴近客观实际，因而更具备实际应用价值。

2、问题二考虑互联网+时代智能手机与基于智能平台的打车软件的普及，增加功效系数与评价指标，对原模型进行了改进，提高了原模型的适应性。

3、本文建立的模型均以真实的合肥市出租车交通网数据库为基础，数据内容详实，类型多样，并用 SPSS 与 MATLAB 联合编程求解，能够更直观准确的处理大数据，更加接近问题的实质。

4、第三问中建立的 RVGAM 缓解效度评分分析模型，可以通过利用评分优越程度 θ 更好地反映新补贴方案缓解效度 H_n 高于传统补贴方案缓解效度 H_o 两个效度之间的对比程度，简单地通过定性和定量分析得，优越程度 θ 越趋近于 1，越说明 RVGAM 缓解效度评分算法的合理性越好，所设计的新打车软件服务平台具有高效的缓解能力，能够切实缓解打车难问题。

8.2 模型缺点

1、考虑的实际因素有所抽象、简化，可能得到的结果不能很好地适应复杂环境。实际上天气条件是变化的；合肥市各地区人口数是不断变化的；出租车在不同时间段的运营成本也是变化的；这些都假设不考虑，模型离准确如实地反映现实可能还存在较大的距离。

2、搭载不同的打车公司软件司机端的出租车的行驶范围、载客路径会不一样，本文建立的模型没有考虑这一点。在不同时间、不同地区，针对乘客订单上的不同目的地，司机判断是否接单的结果也可能是不同的。

3、第三问中建立的 RVGAM 缓解效度评分分析模型和新打车软件服务平台没有更细致给出额度影响的参数量，而参数量影响相应的微币额度，就会导致新补贴方案缓解效度 H_n 和传统补贴方案缓解效度 H_o 可能存在问题，致使优越程度 θ 值产生误差，进而使缓解能力的分析不准确。

参考文献

- [1] 李冬新, 栾洁. 滴滴打车的营销策略与发展对策研究[J]. 青岛科技大学学报(社会科学版), 2015, 31(1): 105-108;
- [2] 孙聪毅. 基于 SAMR 的多数据融合城市出租车载客引导方法研究[D]. 山东大学硕士学位论文, 2015, TP399: 19-24;
- [3] 增宪培, 陈鹏. 从打车软件想到的出租车管制分析[J]. 交通企业管理, 2014 年第 10 期 总第 314 期, 31(1): 1-5;
- [4] 百度文献, 苍穹智能出行平台, 2015 年 9 月 11 日;
- [5] 司守圭, 孙玺菁. 数学建模算法与应用[M]. 北京: 国防工业出版社, 2011;
- [6] Mark M.Meerschaert 著. 刘来福, 杨淳, 黄海洋译. 数学建模方法与分析[M]. 北京: 机械工业出版社, 2005.6.;
- [7] 苏金明, 阮沈勇. MATLAB 6.0 实用指南 [M]. 北京: 电子工业出版社, 2002.

附录

附录 1 SPSS 软件对附录中各类指标在不同时间段以及不同空间条件下的聚类分析

表 1 最终聚类中心间的距离

聚类	1	2	3
1		37.362	62.383
2	37.362		25.269
3	62.383	25.269	

表 2 最终聚类中心

	聚类		
	1	2	3
订单量	76	54	38
出租车数量	92	64	44
空载率	.206	.220	.192
等车时间	14	8	6
呼叫回应时间	30	21	19

ANOVA

	聚类		误差		F	Sig.
	均方	df	均方	df		
网络订单量	12073.19	2	24.730	273	488.20	.000
	9				2	
出租车数量	18614.82	2	38.746	273	480.42	.000
	4				8	
载客情况	9906.884	2	27.135	273	365.09	.000
					0	
空载率	121.307	2	40.381	273	3.004	.051
订单上车率	117.365	2	57.347	273	2.047	.131
网络订单回应时间	636.663	2	7.055	273	90.248	.000
/秒						

F 检验应仅用于描述性目的，因为选中的聚类将被用来最大化不同聚类中的案例间的差别。观测到的显著性水平并未据此进行更正，因此无法将其解释为是对聚类均值相等这一假设的检验。

附录 2 程序

1、供求匹配程度评分值算法程序

<pre> %改程序为某时间段某地区的出租车供求 匹配程度评分值计算算法 Nd=[]; %导入某时间段该地区的网络出租车订单 量 Nd=Nd'; Nc=[]; %导入某时间段该地区的实际出租车数量 Nc=Nc'; au=[]; %导入某时间段该地区的空车率 au=au'; aw=[]; %导入某时间段该地区的等车率 aw=aw'; C=[]; %建立某时间段某地区的出租车供求匹配 程度数组 e=[]; %建立某时间段某地区的出租车供求匹配 </pre>	<pre> 程度评分数组 for i=1:276 C(1,i)= Nd(1,i)*aw(1,i)/(Nc(1,i)*au(1,i)); e(1,i)=abs(1-C(1,i)); end %计算评分值 min=e(1,1); for i=1:276 if e(1,i)<min min=e(1,i); j=i; end end %迭代求评分值最小项 disp('该时间段中供求匹配程度度最好的 是地区')j disp('其供求匹配程度评分值为: ')min %输出结果 </pre>
--	--

附录 3 第一问中运用 SPSS 处理数据部分结果

聚类成员															
案例 号	等车 时间/ 分钟	聚类	距离	案例 号	等车 时间/ 分钟	聚类	距离	案例 号	等车 时间/ 分钟	聚类	距离	案例 号	等车 时间/ 分钟	聚类	距离
1	21	1	47.288	139	7	3	7.654	70	9	2	12.919	208	6	3	9.107
2	19	1	18.648	140	6	3	21.313	71	7	2	14.785	209	7	3	12.710
3	16	1	14.145	141	6	3	9.610	72	6	2	9.955	210	5	3	12.161
4	10	1	15.324	142	5	3	26.642	73	7	2	21.449	211	5	3	6.312
5	14	1	16.925	143	5	3	17.026	74	7	2	12.160	212	5	3	15.696
6	15	1	8.240	144	7	3	14.290	75	7	2	18.451	213	5	3	15.622
7	12	1	10.431	145	7	3	8.013	76	6	2	14.106	214	5	3	9.829
8	10	1	16.817	146	5	3	26.593	77	6	2	12.482	215	6	3	6.289
9	11	1	21.183	147	6	3	17.093	78	8	2	10.882	216	6	3	5.203
10	7	2	23.568	148	9	3	13.906	79	7	2	14.961	217	7	3	5.446
11	8	2	25.769	149	6	3	9.508	80	7	2	12.333	218	7	3	5.783
12	10	2	21.386	150	9	3	12.872	81	8	2	11.229	219	8	3	7.042

13	9	2	21.143	151	9	3	19.140	82	6	2	14.075	220	8	3	7.201
14	8	2	20.473	152	6	3	15.892	83	8	2	11.101	221	7	3	12.707
15	10	2	20.712	153	8	3	9.643	84	7	2	12.825	222	7	3	9.952
16	8	2	21.722	154	8	3	25.713	85	7	2	14.414	223	6	3	13.289
17	9	2	17.970	155	9	3	12.660	86	7	2	18.956	224	7	3	15.156
18	6	2	17.379	156	6	3	9.524	87	9	2	14.408	225	5	3	15.362
19	6	2	14.843	157	8	3	19.208	88	7	2	15.457	226	5	3	12.884
20	7	2	15.397	158	7	3	15.892	89	6	2	13.653	227	5	3	15.630
21	9	2	13.610	159	7	3	10.075	90	7	2	18.956	228	5	3	14.937
22	8	2	12.387	160	6	3	4.120	91	7	2	13.621	229	5	3	18.636
23	8	2	11.490	161	6	3	6.331	92	7	2	21.189	230	6	3	17.772
24	10	2	10.680	162	5	3	6.533	93	6	2	15.949	231	6	3	9.588
25	9	2	9.125	163	5	3	6.355	94	6	2	14.244	232	7	3	9.363
26	8	2	12.021	164	7	3	2.950	95	8	3	16.233	233	7	3	7.534
27	10	2	10.848	165	7	3	3.757	96	7	2	14.688	234	8	3	7.607
28	8	2	8.201	166	5	3	3.402	97	7	3	14.838	235	8	3	8.647
29	9	2	9.307	167	6	3	3.179	98	8	3	14.017	236	7	3	12.631
30	6	2	11.806	168	9	3	9.846	99	6	3	14.316	237	7	3	8.514
31	6	2	9.273	169	6	3	9.729	100	8	2	15.443	238	6	3	9.863
32	7	2	5.462	170	9	3	10.269	101	7	3	14.018	239	7	3	13.095
33	9	2	6.417	171	9	3	16.534	102	7	2	18.448	240	5	3	8.596
34	8	2	10.328	172	6	3	7.905	103	7	3	14.366	241	5	3	9.949
35	11	2	7.595	173	8	3	8.067	104	9	3	14.554	242	5	3	22.381
36	6	2	8.140	174	8	3	7.924	105	7	3	14.579	243	4	3	19.841
37	7	2	6.038	175	9	3	7.905	106	6	2	20.694	244	6	3	12.973
38	9	2	4.061	176	6	3	8.579	107	7	3	14.228	245	4	3	16.448
39	8	2	3.578	177	8	3	8.012	108	7	3	12.774	246	4	3	13.979
40	8	2	7.208	178	7	3	14.697	109	7	3	11.173	247	6	3	13.990
41	10	2	10.217	179	7	3	3.338	110	6	2	22.442	248	5	3	17.345
42	9	2	5.605	180	6	3	12.412	111	6	3	16.627	249	5	3	11.161
43	8	2	7.600	181	6	3	4.779	112	8	3	15.197	250	4	3	10.171
44	10	2	11.895	182	5	3	9.939	113	7	3	11.473	251	4	3	19.841
45	8	2	5.112	183	5	3	9.939	114	7	2	20.784	252	4	3	10.419
46	9	2	6.808	184	7	3	9.868	115	8	3	13.571	253	5	3	11.936
47	6	2	8.424	185	7	3	9.853	116	6	2	25.283	254	5	3	11.948
48	6	2	19.645	186	5	3	9.939	117	8	3	17.043	255	5	3	10.058

49	7	2	8.405	187	6	3	10.401	118	7	3	14.598	256	7	3	10.751
50	9	2	17.379	188	9	3	9.939	119	7	3	10.907	257	6	3	11.437
51	8	2	7.942	189	6	3	9.853	120	6	3	8.411	258	6	3	14.083
52	11	2	8.053	190	9	3	10.170	121	6	3	10.723	259	8	3	9.928
53	9	2	14.221	191	9	3	10.156	122	5	3	9.048	260	4	3	10.857
54	8	2	20.487	192	6	3	5.485	123	5	3	9.124	261	3	3	11.424
55	10	2	9.019	193	8	3	3.602	124	7	3	8.411	262	4	3	10.964
56	8	2	15.323	194	8	3	3.602	125	7	3	10.318	263	4	3	12.221
57	9	2	15.356	195	9	3	6.685	126	5	3	10.206	264	5	3	12.505
58	6	2	19.286	196	6	3	5.743	127	6	3	11.231	265	4	3	20.867
59	6	2	13.147	197	5	3	12.618	128	9	3	9.293	266	4	3	13.378
60	7	2	4.998	198	5	3	12.985	129	6	3	7.469	267	4	3	20.867
61	9	2	14.608	199	5	3	12.618	130	9	3	19.544	268	3	3	11.726
62	9	2	8.578	200	6	3	12.653	131	9	3	18.311	269	5	3	16.477
63	8	2	5.333	201	6	3	17.943	132	6	3	6.137	270	3	3	16.477
64	10	2	6.505	202	7	3	9.232	133	8	3	15.587	271	5	3	20.020
65	8	2	13.586	203	7	3	12.710	134	8	3	12.313	272	3	3	41.962
66	9	2	7.168	204	8	3	9.107	135	9	3	7.292	273	4	3	19.447
67	6	2	6.857	205	8	3	12.507	136	6	3	12.418	274	4	3	22.476
68	6	2	7.276	206	7	3	16.223	137	8	3	15.236	275	3	3	17.240
69	7	2	14.897	207	7	3	10.005	138	7	3	12.313	276	3	3	20.526

附录 4 合肥市出租车交通网数据库数据见压缩文件包