“拍照赚钱”的任务定价

摘要

随着“众包“模式的提出和互联网大数据的发展，”拍照赚钱“应运而生。本文通过对现有的定价模型的研究，结合城市经济物价水平以及打包发布的模式，建立新的定价模型。

针对问题一，第一，本文认为附件一中的定价模型是仅根据订单的位置建立的。首先，我们利用API标出了订单的位置并用点的大小表示价格。根据价格的等高线找出了17个最低价的中心（基本为商业中心或交通枢纽），并发现以其为圆心辐射，价格与订单离最近的圆心的远近呈近乎线性关系。但用相关性分析后发现关联度不大。我们再用MATLAB拟合函数，发现其更接近于阶梯线性关系。第二，我们用logistic回归，分析出任务未完成的原因是定价的不合理，未考虑到各市经济发展水平和物价水平的差异，城郊物价水平的差异。

1. 问题重述
   1. 背景资料与条件
   2. 需要解决的问题
2. 问题分析

2.1 问题一的分析

该问题要求研究附件一中已完成的项目定价规律，并分析任务未完成的原因。

首先，本文认为，开发者按照某种模型产生定价，这不是由于任务的完成程度或者会员位置导致的，反而是价格和位置导致任务的完成程度和会员的加入。因此，根据已有的数据，本文认为，附件一中的定价模型仅与任务的位置有关。我们在地图上标出了任务点，并用点的大小来表示定价的高低。我们发现，越靠近某些商业区和交通枢纽，价格就越低，并且以这些中心为圆心辐射出去，似乎价格与离圆心的远近呈线性关系。因此，我们根据价格做等高线，取价格最低的中心点，取每个点距离最近的中心点的距离，以某个价格为起步价，以距离为变量阶梯定价。用MATLAB拟合函数找出价格与距离的线性关系，然后用logistic回归分析出阶梯定价函数并给予证明。

2.2 问题二的分析

会员的位置和信誉度与订单的位置和定价有关

2.3 问题三的分析

2.4 问题四的分析

1. 模型假设
2. 假设接单者不会去他市完成订单
3. 假设一个市的消费水平均为该市的平均水平
4. 假设定价前没有会员
5. 假设东莞市的定价没有高于该市的物价水平

四、符号说明

* 1. 模型的建立与求解

5.1 问题一的模型建立与求解

5.1.1 附件一中项目的任务定价规律

如图1，我们利用API找出了所有点的在地图上的地理位置，并用大小表示其定价。我们发现，越靠近某些位置，点就越小。我们根据价格画等高线，再根据等高线确定17个低价的中心点。表1为这17个点的经纬度位置。

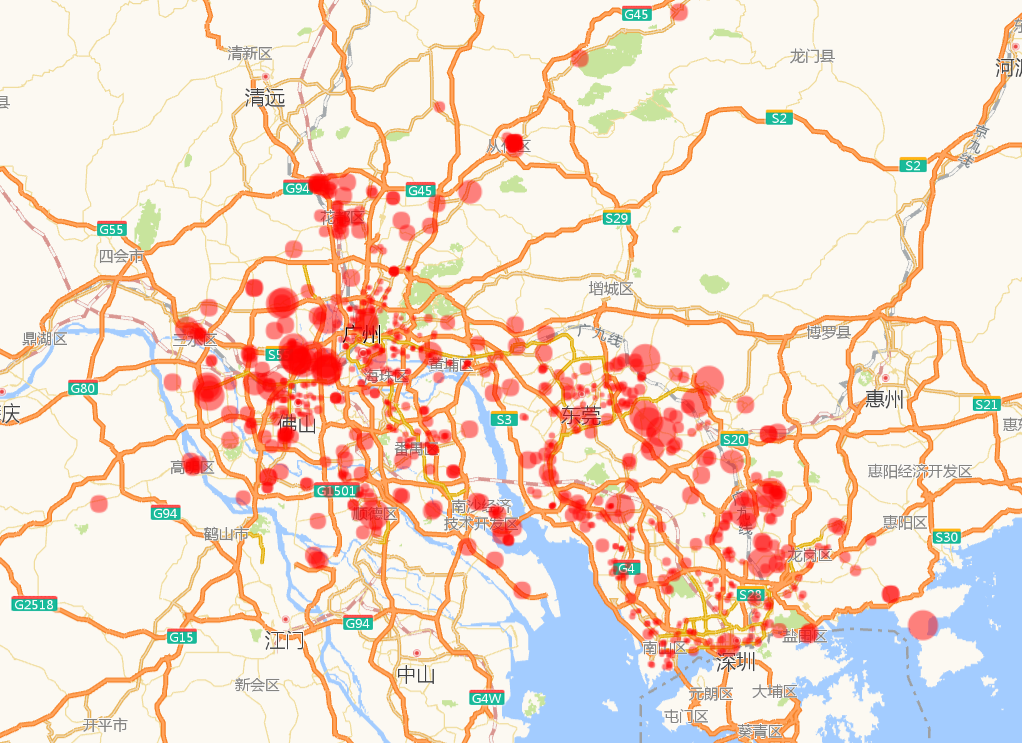


图 1

|  |  |  |
| --- | --- | --- |
| 地址 | 纬度 | 经度 |
| |  | | --- | | 深圳市民中心 | | 22.54657 | 114.0599 |
| 深圳腾讯大厦 | 22.5404 | 113.9345 |
| 深圳北站 | 22.60975 | 114.0296 |
| 虎门公园 | 22.81853 | 113.6885 |
| 东莞人民公园 | 23.03966 | 113.7592 |
| 广州公园前地铁站 | 23.12537 | 113.2641 |
| 广州天河体育中心 | 23.13746 | 113.325 |
| 广州番禺广场 | 22.9339 | 113.3843 |
| 佛山祖庙 | 23.02905 | 113.1131 |
| 顺德体育中心 | 22.83851 | 113.2497 |
| 广州塔 | 23.10647 | 113.3244 |
| 广州火车站 | 23.14896 | 113.2578 |
| 白云湖公园 | 23.22794 | 113.2372 |
| 嘉禾街道 | 23.22798 | 113.2894 |
| 塘尾 | 22.70193 | 113.8181 |
| 观澜汽车站 | 22.72077 | 114.0652 |
| 长安汽车北站 | 22.8 | 113.7942 |

表1

以这些中心点为圆心，设各点到最近的中心点的距离为r，定价为p，用MATLAB画出散点图，如图2。

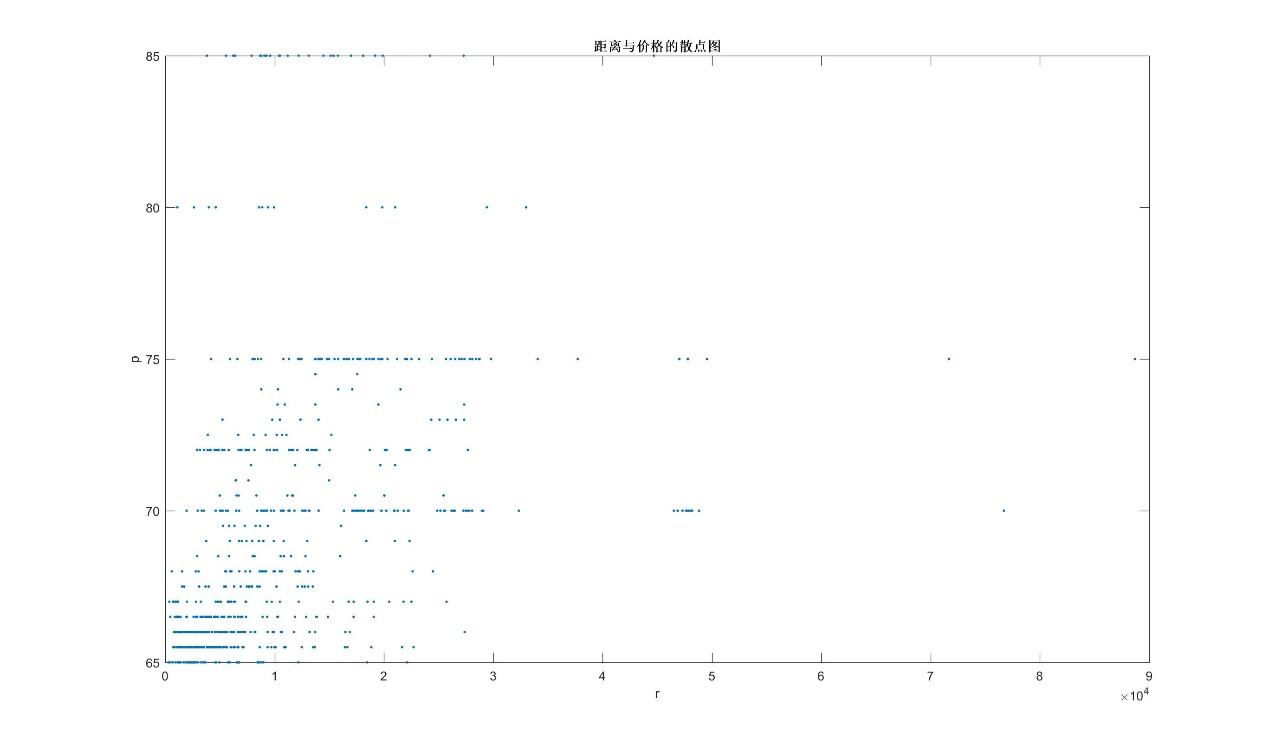


图 2

经过线性相关性分析后，发现p与r的相关系数为

10个点：SpearmanrResult(correlation=0.55025864493074839, pvalue=2.8978803985131413e-67)

17个点：SpearmanrResult(correlation=0.65170700842805607, pvalue=4.1287145301858785e-102)

我们发现相关系数大约为0.65，低线性相关。用MATLAB拟合函数为y=0.0001973\*x+67.1，如图3。可以看到拟合度并不是很高。

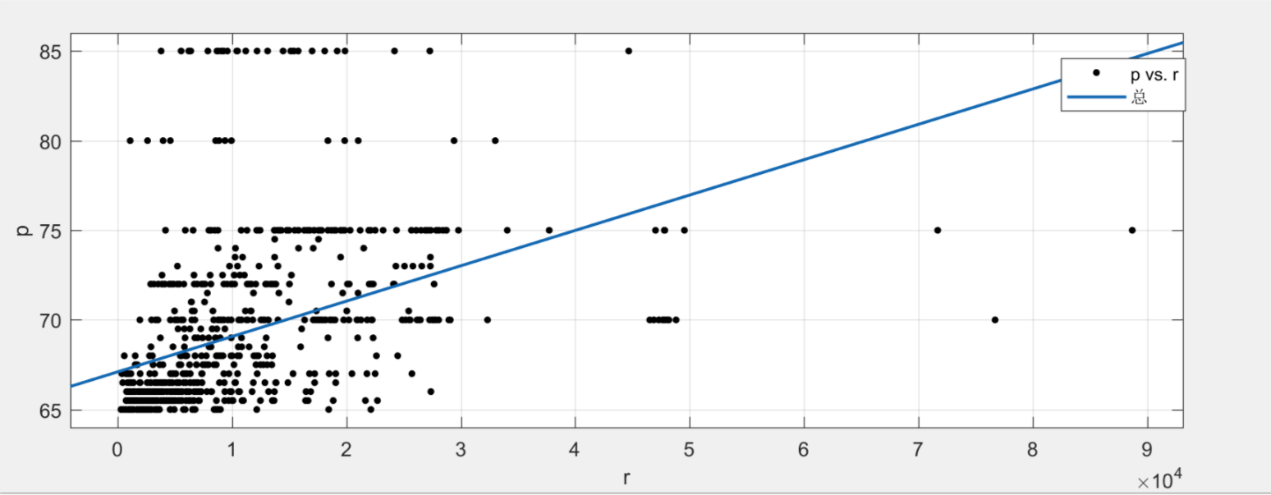
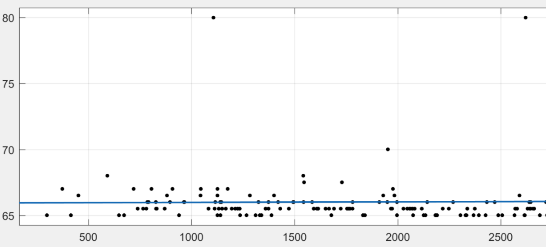
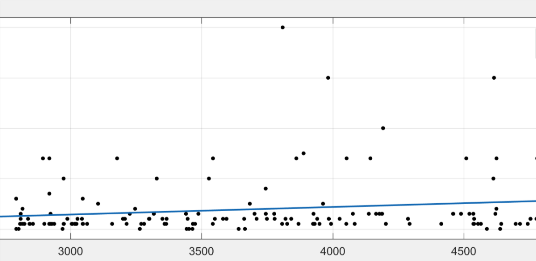


图 3

因此，我们尝试用阶梯型线性函数将其拟合。如图4

P=

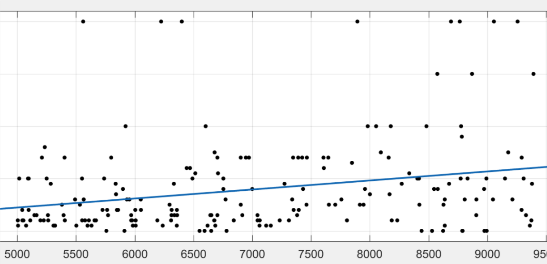
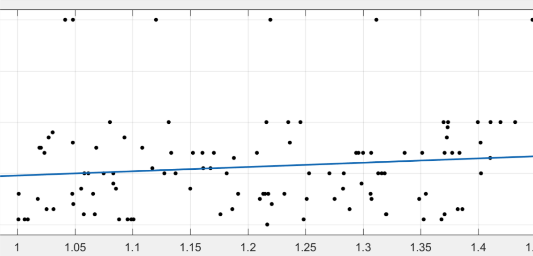


图 4

5.1.2 分析任务未完成的原因

首先我们利用地图API做出完成和未完成的任务的位置图，如图5

紫色的点为未完成的任务位置，黑色的点为完成的任务位置

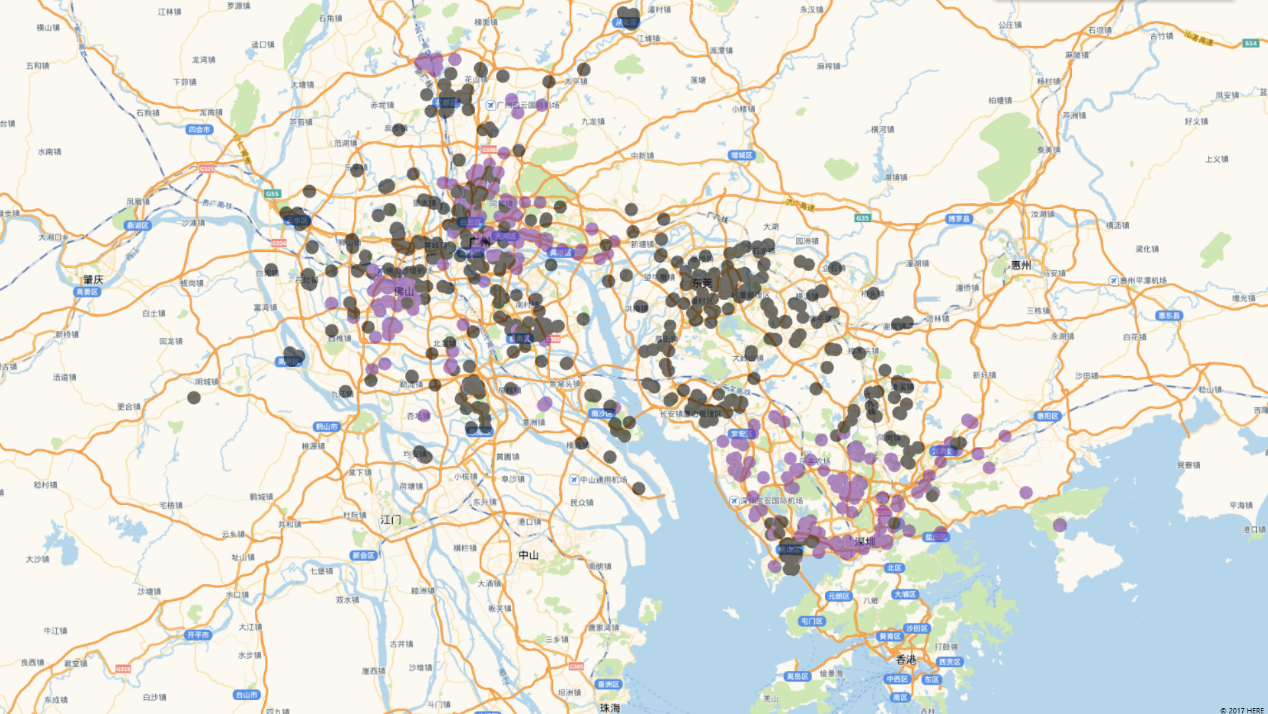


图 5

其次，利用logistic回归来检验上述定价方案是否与任务完成情况有关。设距离为x1，价格为x2。当价格与任务完成度线性相关时，y=1；反之，y=0。p为y=1的概率。

（数据？）

我们发现，定价方案与任务完成情况的相关度不大，即定价方案的不合理导致任务完成率低。

5.1.2.1 未考虑城市间经济物价水平的差异

从图5中我们可以看到，东莞基本上所有的任务都完成了，而深圳、广州、佛山的任务完成度非常低。（各市任务完成度）

根据广东省、广州市、深圳市2016年统计年鉴，我们得到了东莞市、广州市、深圳市的居民人均日均消费支出，如表2。不难发现，东莞市的人均日均消费水平最低，附件一中的定价与东莞市的物价水平的匹配度最高，因此任务完成度最好。然后我们根据东莞市的物价水平与其他城市物价水平的比例建立价格系数，以东莞市为价格模型求出其他市的价格模型。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | 广州市 | 深圳市 | 佛山市 | 东莞市 |
| 人均日均  消费支出 | 97.95 | 88.66 | 85.76 | 77.41 |
| 价格系数 | 1.265 | 1.145 | 1.108 | 1 |
| 起步价 | 82 | 74.5 | 72.01 | 65 |

5.1.2.2 未考虑市区与郊区消费水平的差异

城郊的任务完成度在大体上比市中心完成度高。根据附件一的定价模型，城郊的定价普遍比市区高，但市区的消费水平普遍高于郊区。因此，该定价模型没有考虑到市区和郊区消费水平的差异，导致市区的任务完成度过低。

5.1.2.4 定价模型未考虑打包处理

任务分布密集处没有打包处理，这一方面导致用户争相选择任务密集点而忽视其他任务，另一方面导致任务没有批量化处理，与已完成任务距离十分接近的未完成任务较多。（离已完成较近的未完成数据）

5.2 问题二的模型建立与求解

5.2.1 新的定价方案

* 1. 模型的评价
  2. 模型的改进和推广

7.1 模型的改进

可以通过调查，得到距离和价格意愿的相关性。这样就能根据会员与任务点的距离，自动匹配价格。

可以根据会员的动态位置，以会员到周围订单点的距离做出动态定价

7.2 模型的推广