2017 高教社杯全国大学生数学建模竞赛

承 诺 书

我们仔细阅读了《全国大学生数学建模竞赛章程》和《全国大学生数学建模竞赛参赛规则》（以下简称为"竞赛章程和参赛规则"，可从全国大学生数学建模竞赛网站下载）。

我们完全明白，在竞赛开始后参赛队员不能以任何方式（包括电话、电子邮件、网上咨询等）与队外的任何人（包括指导教师）研究、讨论与赛题有关的问题。

我们知道，抄袭别人的成果是违反竞赛章程和参赛规则的，如果引用别人的成果或其他公开的资料（包括网上查到的资料），必须按照规定的参考文献的表述方式在正文引用处和参考文献中明确列出。

我们郑重承诺，严格遵守竞赛章程和参赛规则，以保证竞赛的公正、公平性。如有违反竞赛章程和参赛规则的行为，我们将受到严肃处理。

我们授权全国大学生数学建模竞赛组委会，可将我们的论文以任何形式进行公开展示（包括进行网上公示，在书籍、期刊和其他媒体进行正式或非正式发表等）。

我们参赛选择的题号是（从A/B/C/D中选择一项填写）： B

我们的参赛报名号为（如果赛区设置报名号的话）： 201719008097

所属学校（请填写完整的全名）： 广东工业大学

参赛队员（打印并签名）： 1. 林业鑫

2. 詹汉青

3. 周荧倩

指导教师或指导教师组负责人（打印并签名）： 王振友老师

（论文纸质版与电子版中的以上信息必须一致，只是电子版中无需签名。以上内容请仔细核对，提交后将不再允许做任何修改。如填写错误，论文可能被取消评奖资格。）

日期： 2017 年 9 月 17 日

赛区评阅编号（由赛区组委会评阅前进行编号）：

2017 高教社杯全国大学生数学建模竞赛

编 号 专 用 页

赛区评阅编号（由赛区组委会评阅前进行编号）：

赛区评阅记录（可供赛区评阅时使用）：

评阅人										
评分										
备注										

全国统一编号（由赛区组委会送交全国前编号）：

全国评阅编号（由全国组委会评阅前进行编号）：

“拍照赚钱”的定价模型

摘　　要

随着移动互联网的发展，新兴了一种“拍照赚钱“自助式服务模式。注册的会员通过领取相关任务，到达指定地点拍照进行商业检查和信息收集，并以此赚取标定酬金。但其中的任务定价是其核心要素。如何协调好会员接受度与预算的关系是定价的关键。

针对第一个问题，我们将附件一的数据可视化，通过excel做出3维数据地图。为了研究任务定价的规律，我们使用“控制变量法”来研究任务定价规律，发现会员的定价一方面跟任务地点的位置有关，低定价以聚集分布为主，主要集中在城市中心，而高定价主要分布在城市的边缘处或是交界处。另一方面，定价还与会员的位置，会员的信誉值有关。然后利用Matlab对未完成的位置分布进行拟合，发现未完成的任务与任务的平均收益低，会员的位置及信誉值有关。

针对第二个问题，本文首先将任务的平均预算作为基础成本，然后在通过第一问的分析后，对影响完成率的因素进行了分析，以广州市各区为例，设立周围任务数量，会员数量，会员的平均信誉值作为影响因素。我们采用多元回归模型来得出各个变量的参数，最后我们对模型进行了显著性检验，对回归模型和回归系数进行了检验。然后我们分析了它的优缺点，发现现比于初始定价方案，它能够显著提高总任务的完成率。

针对第三个问题，本文找出一个半径r,以某个任务位置作为圆心，在其半径内的任务打包发布，然后建立了一个打包定价模型。因为打包后会员实际距离缩短了，所以打包的总定价实际相比原来总价减少。本文通过计算单位距离成本，计算不同方式的打包成本。并从范围圆三个任务的定价推广到n个任务打包定价。

针对第四个问题，先根据第二问的所建立的定价模型对每个任务定好价格，然后设立打包半径为500m,.根据任务坐标的位置关系，采用贪心算法一一匹配。将综合之后的定价作为附件三的定价方案。

关键词： 可视化数据;聚类分析;多元回归分析;动态定价模型;贪心算法;控制变量法

一　问题重述

1.1 问题背景

“拍照赚钱”是移动互联网下的一种自助式服务模式。用户领取相关的任务，前往目的地拍照进行商业检查和信息搜集，赚取标定的酬金。这种众包模式大大节省企业的人力成本和调查周期。而其中的任务定价是其核心要素。定价过高，企业的成本会大幅升高。定价过低，有的任务就会无人问津。

1.2 提出问题

建立一个合理的定价模型，满足下列要求。

（1）研究附件一中的任务定价规律，分析未完成的原因。

（2）设计新的任务定价方案，与原方案进行比较。

（3）将任务联合在一起打包发布，基于这种考虑修改前面的定价模型。并考虑对最终任务完成情况的影响

（4）对于附件三的新项目给出新的任务定价方案，并评价实施结果。

二　问题分析

2.1 问题一分析

问题1要求研究附件一的任务定价规律，可以将每个任务地点的经纬度看作是一个二维特征向量，将任务标价看作是因变量。将问题的定价规律简化为求一个函数变量间的映射关系。所以我们先处理一下数据，并将任务区分为未完成和已完成，分别绘制数据地图进行可视化，通过观察找出其中规律。通过与会员的分布地图和定价关系的比较，容易发现一些任务未完成的原因。

2.2 问题二分析

题目需要我们设一个新的任务定价方案，需要尽可能满足会员的酬金要求又要尽可能的让任务的完成率高。在这种情况，我们首先选择在原方案进行改善，评价影响定价的几个关键因素：周围任务的数量，会员数量和信誉度等。通过这些因素在原定价的基础上设计一个新的定价函数，使之尽可能提高任务的完成率又不会大幅超过预算。

2.3 问题三分析

将位置集中的任务进行打包处理，关于如何确定位置的集中，我们从任务的分布图上大概可以发现在1000m为直径的圆普遍分布着三个点，并且这个距离是平时会员所能接受的任务移动范围。所以打包发布以1000m作为位置集中的判断依据。

第二，关于打包之后的价格，当几个任务集合在一起发布时，任务的总价跟原来相比可以减低。因为任务集中在一起时，会员可以一起完成，在某种程度上减低交通成本。我们先在理想模型下建立起n个任务打包时任务定价模型。

2.4 问题四分析

这一问是结合上述建立的定价模型，给定任务位置来确定定价。所以首先采取第二问建立的模型对单个任务分别进行定价，然后通过第三问建立的模型对一些位置相近的任务进行重新打包定价。这个过程我们使用贪心算法，从上往下去匹配。

三　假设与符号

3.1 问题假设

1假设众包市场是静态的，即会员的数量和信誉值不随着时间的推移发生变化。

2忽略天气，节假日等原因对任务完成情况的影响。

3.2 符号说明

符号	意义
B	预算（元）
N	会员的密集程度（人/km^2）
P	周围相似任务的数量（个/km^2）
W	会员的平均信誉值

四　模型建立与求解

4.1 问题一

4.1.1 建立3维数据地图分析

为分析不同经纬度的任务地点的定价规律，研究任务定价的规律，我们使用“控制变量法”来研究任务定价规律，并寻找出任务未完成的原因。

从总体上看，定价大致分布在65-85之间，现将定价分为65-72 （红）、72-79（绿）、79-85（黄）三大部分进行定价与位置的可视化地图分布。大部分定价在65-72区间密集分布，而剩余的72-85区间已环状离散分布。

4.1.2 定价的分布规律

定价的位置分布　从总体上看，定价大致分布在65-85之间，现将定价分为65-72 （红）、72-79（绿）、79-85（黄）三大部分进行定价与位置的可视化地图分布。根据地图可见，定价的位置分布是以“环状”分布，定价较低在中心，定价较高在外围。圆心环部分以65-72（红）定价聚集分布为主，主要集中在城市中心，例如广州市中心地区、东莞

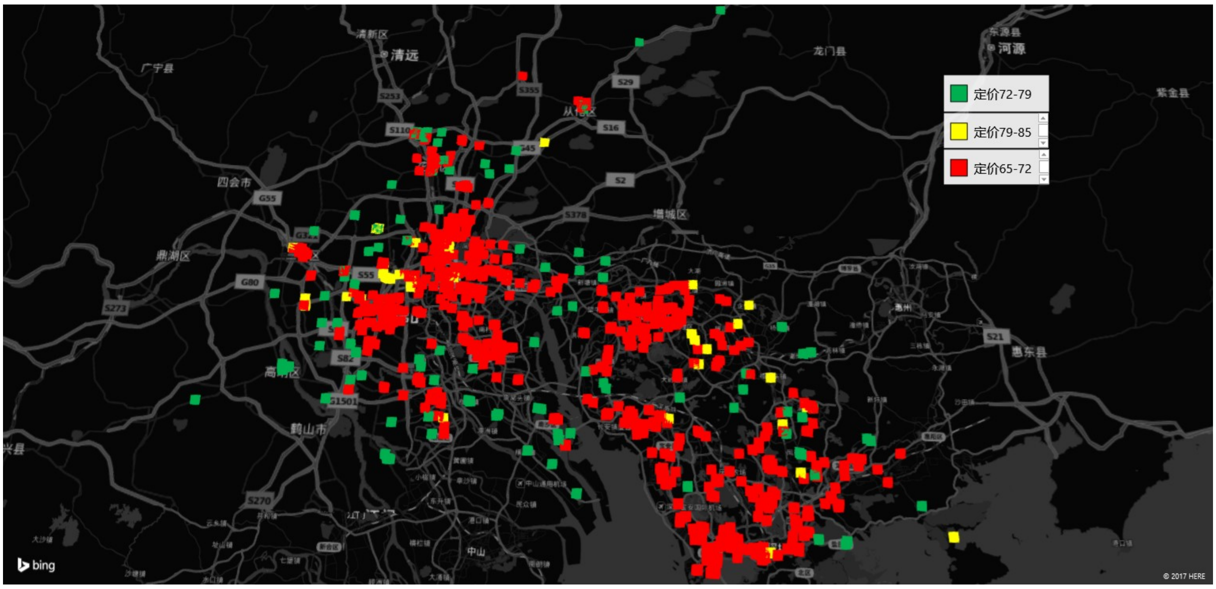

图 1: 不同定价的分布规律图示

市中心地区以及东莞市中心地区。围绕在圆心环外的第二环72-79（绿）环绕分布，主要分布在广州市中心外围预计深圳市中心外围；第三部分79-85 （黄）主要分布在第三环边缘，主要集中在城镇交界线例如广州佛山交界线，以及深圳东莞交界线。

定价与会员的人数 根据附件二的“会员信息数据”中的会员位置与定价的对比，发现会员人数越少的地区，所定的价格就越高。原因是任务完成是必要的，但是去完成的任务的人数较少，物以稀为贵，需要定价高去吸引会员去完成任务。

另外，在富裕的地区中，会员人数较多，定价就越低；反之在贫困的地区中，会员人数较少，定价就越高。

4.1.3 未完成的原因分析

我们通过matlab绘制未完成任务和会员分布图，可以发现未完成的任务聚类出现。从左往右分别出现在佛山，广州，深圳南，深圳北。利用Matlab对未完成的位置分布进行拟合。

地点	纬度	经度	拟合圆半径
佛山	113.0723	22.99083	2927.716
深圳南	114.0466	22.52331	9061.227
深圳北	113.8116	22.68519	6567.458
广州	113.2727	23.10023	5355.994

表 1: 未完成的位置分布拟合图

4.2 未完成因素猜测

1.平均收益低 根据附件一的“已结束项目任务数据”中的任务完成情况与任务标价的

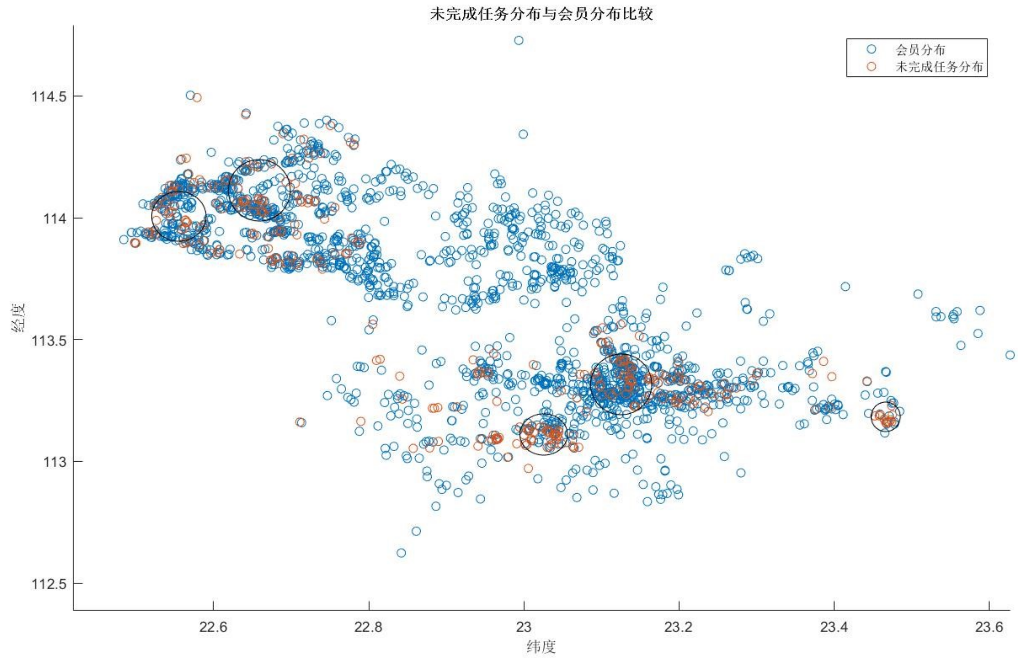

图 2: 未完成任务分布与会员分布图

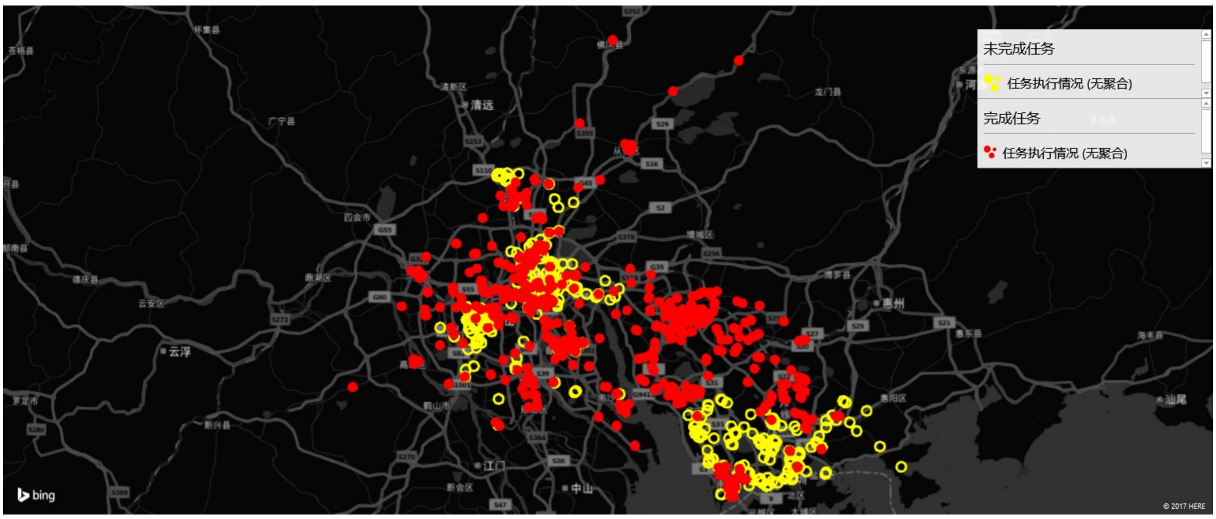

图 3: 未完成任务和完成任务的对比图

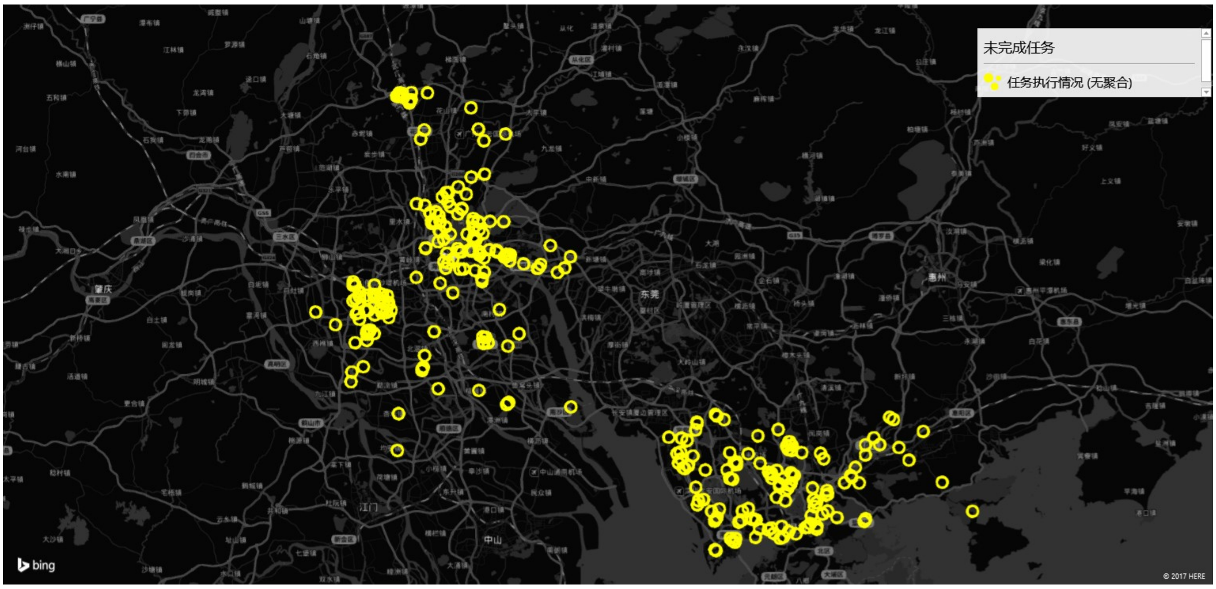

图 4: 未完成任务具体图

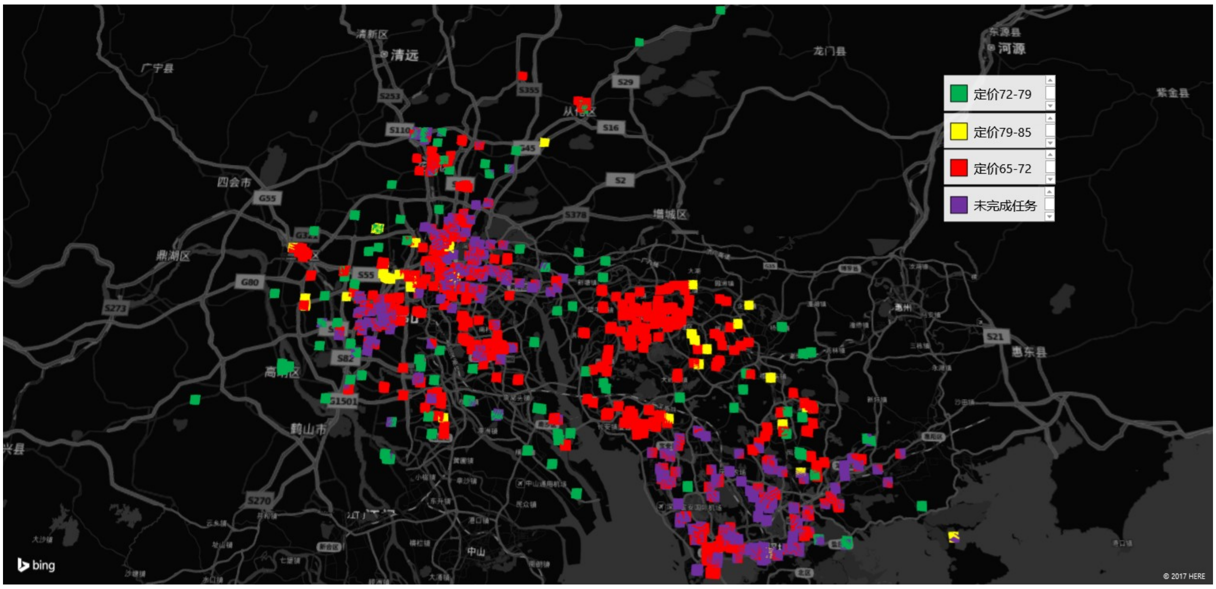

图 5: 定价与未完成任务关系图

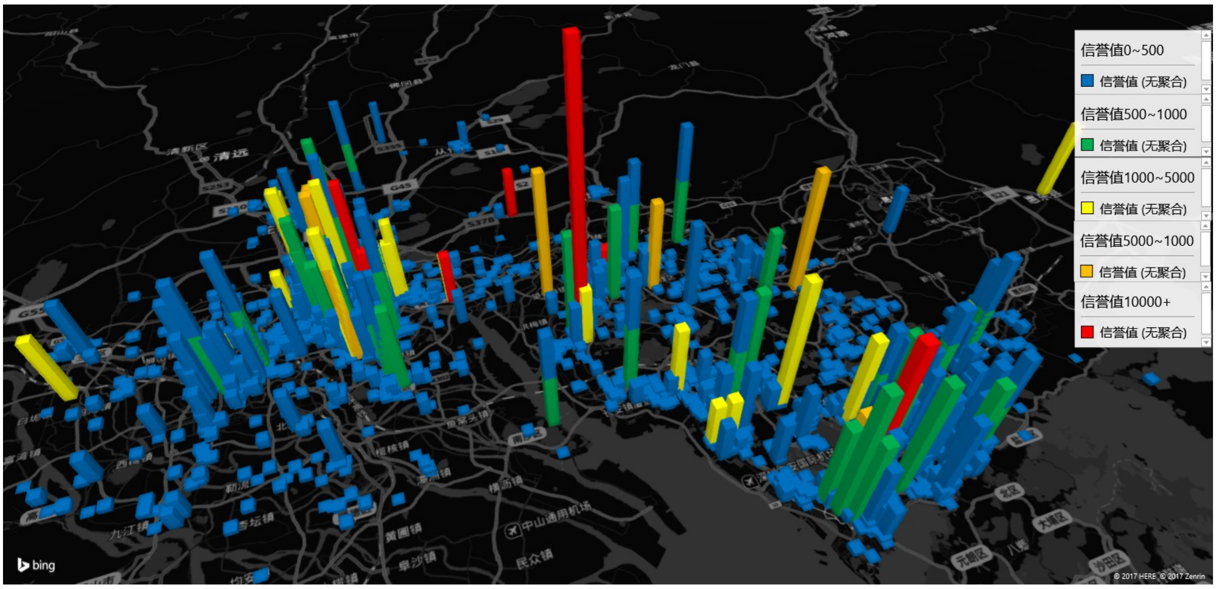

图 6: 会员位置信誉值分布情况

比较，未完成任务集中分布在定价较低的红色区域，定价为65-72。原因是定价较低，会员赚取的利润较低，平均收益较低，导致会员不愿意付出劳动力，从而导致了项目未完成。

2.会员的位置 基于定价低来说，会员与任务的距离较远，会员不愿意花过多成本去完成任务，会员获得的利润大大减少，从而导致了总体任务完成率降低。

3.会员信誉值 根据附件二的“会员信息数据”中的会员位置，我们将数据进行可视化。信誉值与任务完成情况的比较，发现会员信誉值较低，导致了会员获取的任务限额较少，使得总体任务完成率降低，恶性循环。

4.3 问题二

4.3.1 基于影响因子的动态定价模型

我们发现初始的定价任务完成情况并不是很好，因为在一些会员少，会员信誉值较低又或因为定价较低的因素，往往导致一些任务没有完成。充分吸收了上述的缺点，为了平衡会员和任务发布者的利益，使得任务发布者能报出的并且能够被大多数工作者接受的价格，我们设立了基础定价。

定义初始成本S1

$$S_1 = \frac{B}{n}$$

其中B是预算，n是任务个数。经过计算，初始成本为69.11 元。

其他定义因素的考虑

周围会员数量N 周围会员数量越多，任务的被接受的几率增加。所以任务的定价可以比基础价格低，而当周围会员数量较少时，只有提高任务的定价，才能增加任务的完成度。经过研究，周围会员数量对任务的定价大致可以看作是正比例函数。

$$S_2 = -k_1 * N$$

N指得是会员的密集程度。调查会员密度我们采用人口密度的形式。这里我们用广州市（经度11.2756纬度:23.1171）为例。

区镇	会员人数	面积（km2）	会员的人口密度
荔湾区	61	59.1	1.032149
越秀区	60	33.8	1.775148
海珠区	81	90.4	0.896018
天河区	148	96.3	1.536864
黄埔区	45	90.9	0.49505

表 2: 广州市各区会员密度图

周围相似任务的数量P 当周围的相似任务的数量较多时，而会员数量一定时，那些定价高的优先被选择，此时任务的定价需要提高才能提高对会员的吸引力。反过来说，周围的相似任务的数量较少时，即使定价比市场价低，此时任务也有很大几率被选择。周围相似任务的数量对定价的影响可以大致看作成一个正比例函数。

$$S_3 = k_2 * p$$

区镇	相似任务个数	面积（km2）	任务的密集度
荔湾区	53	59.1	0.896785
越秀区	33	33.8	0.976331
海珠区	30	90.4	0.331858
天河区	48	96.3	0.498442
黄埔区	18	90.9	0.19802

表 3: 广州市各区任务密度图

会员的平均信誉值w 可以发现会员的信誉值会影响任务的完成情况。高信誉值所以在一些会员数量较少的地区，需要适当地提高酬金，才能让任务完成度最好。即N小于某个阈值时，此时该地区的平均会员信誉值会影响任务的完成情况。

$$S_4 = -k_3 \; * \; w$$

一种比较简单的计算平均信誉会员信誉值是

$$\overline{w}=\frac{\sum_{q=1}^{n}W_q}{n}$$

但容易发现这种在存在较高或较低的信誉值，会显著拉高或拉低信誉会员信誉值，这对实验结果的误差很大。经过分析，我们采取阈值比例的算法。即

$$\overline{w}=\frac{\overline{n_1}}{\overline{n_1}+\overline{n_2}}$$

我们设计阈值为30，大于这个数为高信誉值，低于这个数为低.$\overline{w_1}$信誉值代表用第一种算法算得高信誉值会员平均信誉值人数，$\overline{w_2}$指的是总信誉值会员人数.

区镇	高信誉值会员人数	低信誉值会员人数	面积	平均信誉会员信誉值
荔湾区	23	38	59.1	0.377049
越秀区	24	36	33.8	0.4
海珠区	29	52	90.4	0.358025
天河区	28	120	96.3	0.189189
黄埔区	9	36	90.9	0.2

表 4: 广州市各区平均信誉会员信誉值图

任务难易繁琐程度　总所周知，任务越难，相对应的酬金要越高。当然，任务的难度的度量历来存在争议。另一方面，任务越复杂繁琐，完成的时间成本越高，会影响接其他任务。所以任务的繁琐程度也影响着定价。但是我们在这个问题中，由于每个任务都是到达指定地点拍照，所以我们认为这些任务的难易繁琐程度是相同的。这个因素在这个问题中不予考虑。

任务期限　一个任务的期限同时也影响着任务的酬金。期限越短，任务越紧急，相应的酬金越高才能使会员立刻领取。但由于题目没有给出任务期限这个限制条件，所以我们一律将任务期限定为3 天，并不把任务期限当作定价因素。

所以，总的定价函数如下

$$y=c_0+c_1x_1+c_2x_2+c_3x_3$$

假设y,x1,x2,x3的观察值分别为b_i,a_i1,a_i2,a_i3 且 用最小二乘法算出c0,c1,c2,c3的估计值，即应选取估计值, 使得c_j=$\widehat{c_j}$,j=0,1,2,3时，误差平方和

$$Q=\sum_{i=1}^{18}{\varepsilon_i}^2=\sum_{i=1}^{18}\left(b_i-\widehat{b_i}\right)^2=\sum_{i=1}^{18}\left(b_i-c_0-c_1a_{i1}-c_2a_{i2}-c_3a_{i3}\right)^2$$

达到最小。为此，令

$$\frac{\partial Q}{\partial c_j}=0,j=0,1,2,3$$

序号	定价	周围会员密度	周围任务密度	会员平均信誉值
1	66	1.032149	0.896785	0.377049
2	66.5	1.032149	0.896785	0.377049
3	65.5	1.032149	0.896785	0.377049
4	75	1.775148	0.976331	0.4
5	65.5	1.775148	0.976331	0.4
6	65	1.775148	0.976331	0.4
7	65.5	0.896018	0.331858	0.358025
8	65.5	0.896018	0.331858	0.358025
9	65.5	0.896018	0.331858	0.358025
10	65	1.536864	0.498442	0.189189
11	66.5	1.536864	0.498442	0.189189
12	65.5	1.536864	0.498442	0.189189
13	67	0.49505	0.19802	0.2
14	75	0.49505	0.19802	0.2
15	75	0.49505	0.19802	0.2
16	72	0.115	0.19090	0.35714
17	70	0.115	0.19090	0.35714
18	66	0.115	0.19090	0.35714

图 7: 选取的18个点相应的指标图示

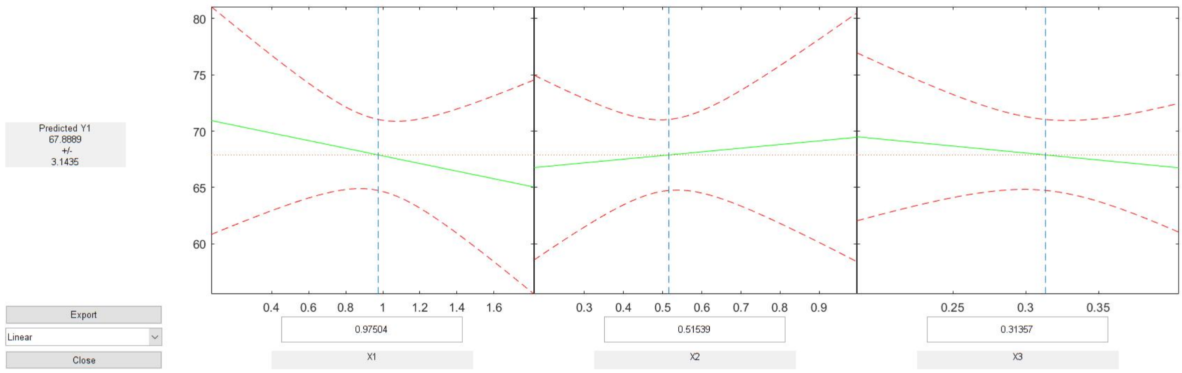

图 8: 多元线性回归方程各系数关系图

得到正规方程组，求解正规方程组得c0,c1,c2,c3的估计值

$$[\widehat{c}_0, \widehat{c}_1, \widehat{c}_2, \widehat{c}_3] = (X^T X)^{-1} X^T Y$$

利用Matlab程序，求得各参数c_0=1.3918 c_1=0.6363 c_2=2.7161 c_3=15.2336

4.4 基于动态分配的双向定价模型

为了使任务发布者和会员满意，我们的基础定价方式采用双向定价的方式。即每个会员先提交一个可接受的定价，然后这个模型将自动输出每个任务的定价，使得每个会员即能有满意的报酬，又让任务发布者可以尽可能的降低成本并且能够完成任务。

假设众包工人p1⋯..pj提交的数据分别为c1⋯⋯cj; 将提交的数据对按照$c_1 \leq c_2 \leq \dots\dots\dots \leq c_n$排序,当

$$ci \leq \frac{B}{\sum\limits_{j<i} \overline{w_j}}$$

时, 任务标价为$p = c_i$,预算为$B - p\sum_{j=1}^{i-1} \overline{w}_j$, 会员被分配的任务数量可表示为

$$\overline{w}_i = \min\{w_i, [\frac{B}{p}] - \sum_{j<i} \overline{w}_j\}$$

也就是说，在会员自己能分配的任务与定价为ci剩余预算能完成的任务数量之间选择最小值作为每个会员分配的实际任务数量。

这样的分配能够让仅可能多的人分配到合适的任务，又能让任务发布者节约成本。然后参考其信誉给出的任务开始预订时间和预订限额，原则上会员信誉越高，越优先开始挑选任务，其配额也就越大。

我们设计一个优先度自然序列$t_1 \cdots t_x$,t=0，1，2⋯⋯.,各会员优先度大小取决于信誉值的。然后用贪婪算法优先分配完高优先度的会员。

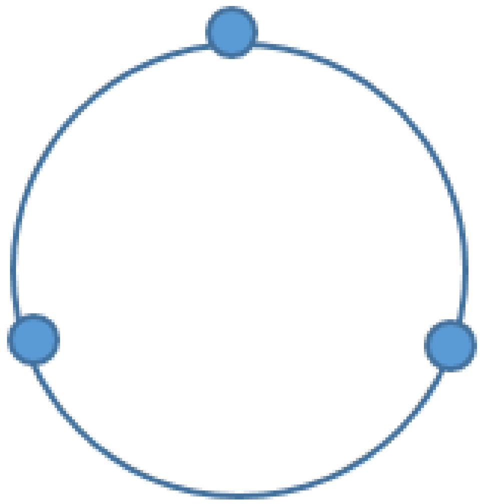

图 9: 3个任务分布图示

4.5 问题3

4.5.1 打包任务的定价模型

先从3个任务打包的情况分析 设这个圆的半径为r，这三个任务的地点x_1,x_2,x_3为均匀分布在圆上,假设这些任务的初始定价为s_1,s_2,s_3; 假设会员初始处于圆心的位置。

现在计算未打包的情况下，完成每个任务后要回到出发点，完成所有任务的路程和y_1

$$y_1 = 6r$$

由于打包任务，会员可以一下子执行三个任务，此时完成所有任务的路程和y_2

$$y_2 = 2 * \frac{\frac{3}{2}r}{\sqrt{3}} + 2r$$

用原来的总定价与总路程的比例作为每单位的成本p

$$p_3 = \frac{\sum_{i=1}^{n} x_i}{y_1}$$

所以打包后的定价s为

$$s_3 = p * \frac{y_2}{y_1}$$

然后我们把这种情况推广到n个任务上

$$y_1 = 2nr$$

$$\theta = \frac{2\pi}{n}$$

其中y1是未打包时完成所有任务的总路程，θ 指得是圆心到相邻两点的射线的夹角。利用余弦定理可以算出n边形的边长。

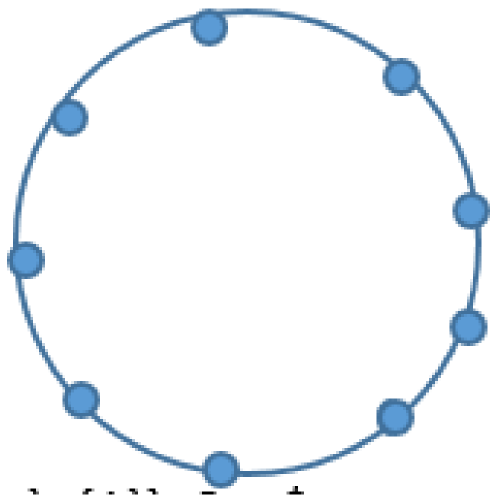

图 10: n个任务分布图示

所以打包后总路程为

$$y_2 = (n-1) * \sqrt{2r^2 - 2r^2 \cos\theta} + 2r$$

打包后的总成本为

$$s_n = p_n * \frac{y_2}{y_1} = \frac{y_2}{{y_1}^2} \sum_{i=1}^{n} x_i$$

4.5.2 模型评价

该模型是假定每个任务点都是同等距离，而且位在一个单位圆的边上，然而实际上更大的可能是落在单位圆里面，所以有一定的误差。

4.6 问题4

这一问是结合上述建立的定价模型，给定任务位置来确定定价。所以首先采取第二问建立的模型对单个任务分别进行定价，然后通过第三问建立的模型对一些位置相近的任务进行重新打包定价。这个过程我们使用贪心算法，从上往下去匹配。

这里仍以天河区作为例子，计算其的最终定价。

根据上面的模型，由于缺少预算，我们仍以69.11元作为初始成本。位于天河区大概有475个任务。通过计算会员人口密度为1.536864人/km，任务数量密度4.9325个/km，会员信誉平均值为0.189189。由公式

$$y = c_0 + c_1x_1 + c_2x_2 + c_3x_3$$

可以得出任务的定价。

将天河区的任务按照纬度进行排序。然后利用贪心算法从上往下匹配需要打包的任务。我们把半径设定为500m，按照这样的方式，将任务的定价按需要一起打包来分

图 11: 天河区任务位置分布图示

类，最后用打包定价模型对原定价进行修正。

$$s_n = p_n * \frac{y_2}{y_1} = \frac{y_2}{{y_1}^2} \sum_{i=1}^{n} x_i$$

Sn即为最后每个任务的最终定价。

其他地点的任务定价依照此流程，将其汇总这样我们就可以最终的定价方案了，然后将其实时显示在app上，优先展示给高信誉值的会员。

五　检验回归模型和回归系数

因变量y与自变量x_1,x_2,x_3之间是否存在线性关系是需要检验，显然，如果所有的都很小，y和x_1,x_2,x_3的线性关系就不明显，所以可令原假设为

$$H0 : c_j = 0\ ,\ j\ =\ 1\ ,\ 2\ ,\ 3.$$

记m =3, n=25,$Q = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (b_i - \widehat{b}_i)^2$,$U = \sum_{i=1}^{n} (b_i - \overline{b}_i)^2$ 这里$\widehat{b_i} = \widehat{c}_0 + \widehat{c}_1 a_{i1} + ... + \widehat{c}_m a_{im} (i = 1, ..., n)$ ，$overlineb = \frac{1}{n} \sum_{i=1}^{n} b_i$。当成立时统计量

$$F = \frac{U/m}{Q/(n-m-1)} \tilde{} F(m, n-m-1)$$

在显著性水平下，若

$$F_{1-\alpha/2}(m, n-m-1) < F < F_{\alpha/2}(m, n-m-1)$$

则接受H_0，否则拒绝式。利用Matlab程序求得统计量F=35.6612，查表得上$\alpha/$ 2 分位数=3.8188，因而拒绝式（1）的原假设，模型整体上通过了检验。当式（1）的H_0被拒绝时,β_j不全为0，但是不排除其中若干个等于0.所以应进一步作如下m+1 个检验：

$$H_0^{(j)}: c_j = 0j \ = \ 0\ ,1\ ,\ 2\ ,\ldots..,\ m\ ,$$

当$H_0^{(j)}$成立时，有

$$t_j = \frac{\widehat{B_j}/\sqrt{c_{jj}}}{\sqrt{Q/(n-m-1)}}\tilde{}t(n-m-1)$$

式中：c_{jj}为$\left(X^T X\right)^{-1}$中的第（j，j）元素，对给定的α，若$|\mathrm{t_j}| < t_{\frac{\alpha}{2}}(n-m-1)$，则接受$H_0^{(j)}$，否则拒绝。

利用Matlab程序，求得统计量

t_0= 16.4595 t_1 = -1.206 t_2= 0.5220 t_3=-0.8415

查表得上α/2分位数$t_{0.025}$（21）=2.1754 对于式（1）得检验，在显著水平α = 0.05时，接受x$H0: Cj = 0j = 01$.拒绝$H0: Cj = 0j = 23$，即变量x_1 对模型得影响是不显著的。建立线性模型时，可以不使用x_0。

六 模型的优缺点

6.1 优点：

1.与原的定价机制进行比较，可以发现我们的模型是基于初始的定价，然后综合考虑了周围会员密度，会员的平均信誉值，周围任务的数量，任务的难度复杂度，任务的期限等等，是对初始的定价方式的完善。因为在尽量不超出预算的情况下，又使得总的任务的完成率尽可能高。

2.本模型采用多个数据来求得k值，所求的k值较准确。并且对模型的变量进行了显著性检验，分析变量之间的关系。

6.2 缺点：

1.基础成本的计算过于简单，在各个影响因子的作用下有可能超出原来的预算

2. 由于材料的不全，一些意外因素没有考虑，例如会员的活跃度及行动路径等等。这些会对总任务的完成率产生一定影响，有可能使结果出现偏离。

参考文献

[1] 宋天舒,童咏昕,王立斌,许可. 空间众包环境下的3类对象在线任务分配[J/OL]. 软件学报,2017,28(03):611-630. (2016-11-29)[2017-09-16]. http://kns.cnki.net/kcms/detail/11.2560.TP.20161129.1335.018.htmlDOI：10.13328/j.cnki.jos.005166

[2] 孙信昕. 众包环境下的任务分配技术研究[D].扬州大学,2016.

[3] 梁帅童. 服装网络零售定制的定价研究[D].上海工程技术大学,2014.

[4] 刘晓钢. 众包中任务发布者出价行为的影响因素研究[D].重庆大学,2012.

附录

代码一 多元回归模型 **Matlab** 代码

```
y=[66
66.5
65.5
75
65.5
65
65.5
65.5
65.5
65
66.5
65.5
67
75
75
72
70
66
];
x1=[1.032149
1.032149
1.032149
1.775148
1.775148
1.775148
0.896018
0.896018
0.896018
1.536864
1.536864
1.536864
0.49505
0.49505
0.49505
0.115
0.115
0.115
];
x2=[0.896785
0.896785
```

```
0.896785
0.976331
0.976331
0.976331
0.331858
0.331858
0.331858
0.498442
0.498442
0.498442
0.19802
0.19802
0.19802
0.1909
0.1909
0.1909
];
x3=[0.377049
0.377049
0.377049
0.4
0.4
0.4
0.358025
0.358025
0.358025
0.189189
0.189189
0.189189
0.2
0.2
0.2
0.35714
0.35714
0.35714
];
x123=[x1,x2,x3];
X=[ones(18,1),x123];
[beta,betaint,r,rint,st]=regress(y,X);
q=sum(r.^2);
ybar=mean(y);
yhat=X*beta;
u=sum((yhat-ybar).^2);
m=3;
```

```
n=length(y);
F=u/m/(q/(n-m-1));
fw1=finv(0.025,m,n-m-1);
fw2=finv(0.975,m,n-m-1);
c=diag(inv(X'*X));
t=beta./sqrt(c)/sqrt(q/(n-m-1));
tfw=tinv(0.975,n-m-1);
save xydata y x123
```

代码二 双向定价模型代码

```
#include <iostream>
using namespace std;

int main()
{
	int B;
	cin>>B;                          //输入预算
	int q[1000];
	int i=0;
	while(cin>>q[i])
		i++;                     //根据信誉值得出每个工作者的质量
	int sumq=i;

	int c[1000];
	int j=0;
	while(cin>>c[j])
		j++;                         //输入工作者提交的数据
	sort(c[],j);
	for(int i=0;i<j;i++)
	{
		for(int q=0;q<j;q++)
			int sum;
			sum+=w[q];
		if(c[p]<=(B/(sum+1)){
			s=c[i];
			w[i]=min(w[i],abs(B/p)-sum);
			}                        //利用定价机制设计得出任务单价 p 以及每个工作者能够
被分配的任务数量 w[i]
	}
	sort(q[],sumq);                  //将工作者进行排序
	if(total<B)                          //如果不超过预算时
	{
	for(int j=0;j<sumq;j++)      /按信誉值高的人从高到低去分配任务
		if(w[i]!=0){
```

```
                W[i--];
                B=B-p*w[j];
                j++;}
        }
        else
            for(int z=j+1;z<sumq;z++)
                if(w[i]!=0){
                    W[i--];
                    B=B-p*w[z];
                    z++;}          //否则，删除质量最高的王作者，把质量次高的工作者排在第
一位 ，
        }
```