“拍照赚钱”任务定价讨论与研究

摘要

“拍照赚钱”是移动互联网下的一种自助式服务模式。用户下载 APP 并注册成为的会员，从 APP 上接受拍照任务，赚取任务的酬金。本文针对“拍照赚钱”任务酬金的一系列定价问题做出了研究，使用了 MATLAB、LINGO、SPSS 等软件进行了计算，得出了针对附件一至附件三中的数据在不同的情况下的定价规律与定价方案。

针对问题一，本文将附件一的数据导入地图中进行直观观察，得出了任务定价大致符合由若干个中心向四周逐渐递增的规律。据此，采用了聚类分析法计算出各中心点的经纬度，然后利用最小距离对所有任务点进行所属区域划分。最后代入附件一中的实际数据进行求解，发现符合原有观察结论：任务定价在不同区域内以该区域某一点为中心在一定范围内逐渐向外扩散递增。由于任务定价不能无限递增且中心点定价不能过低，本文考虑其定价与行政区工资水平之间的关系。经过简单的 Excel 计算可得出，任务定价符合当会员每天完成 1-2 个任务时，其任务酬金大于该行政区最低日工资水平标准小于人均日工资水平的规律。最后，通过数据导入发现所有任务点集中在广州、深圳、佛山、东莞四个城市且考虑到四个地区人均工资水平不同，进而本文利用经济学理论中的替代效应与收入效应理论，分别对四个地区对任务未完成的原因进行了分析。

针对问题二，本文运用了多目标混合整数优化模型，对任务发行方与会员方共同进行考虑，同时考虑控制任务完成情况。利用任务发行方成本、会员预期收入期望以及任务完成率建立了目标函数，然后结合实际，考虑附件二中会员的预定任务限额、预订任务开始时间等问题建立约束条件。使用 MATLAB、LINGO 软件对附件一中的数据进行带入求解，得出了针对附件一新定价方案。同时经过对任务完成率的分析对比可得出，该方案的任务完成率 70.33%大于原有的定价方案的任务完成率 62.47%。

针对问题三，本文利用了任务点与任务点之间的最小距离，不断对各个任务点之间的距离进行比较，不断将距离最小的任务点归类进行打包，建立了迭代打包算法，并利用 SPSS 实现了这一算法。接着根据问题二中的多目标混合整数优化模型，将其单个任务转换成任务打包集进行考虑，修改其目标函数与约束，最终得出其在打包情况下每个任务打包集的定价以及任务完成情况。在任务打包之后，任务完成率达到了 72.89%，说明将任务打包发布之后，有利于提高任务完成率。

针对问题四，本文从两个步骤进行分析，首先采用了问题二与问题三的模型分别对附件三中新项目数据进行求解，得出在两种情况下的任务完成比分别为 68.25%、72.91%。得出对附件三中的新任务，更适合将其打包后再发布的结论；然后考虑对问题三中针对任务打包修改后的多目标混合整数优化模型的目标函数进行线性加权，通过改变其权重，分析在不同权重下，目标函数的变化。从而得出了会员收益期望随着权重的增加而上升，任务发行方成本、任务完成率随着权重的上升而下降的规律。

关键词： 拍照赚钱 聚类分析法 多目标混合整数优化模型 迭代打包算法

一、问题重述

1.1 题目所给的信息

“拍照赚钱”是移动互联网下的一种自助式服务模式。用户下载 APP 并注册成为的会员，从 APP 上接受拍照任务，赚取任务的酬金。这种移动互联网的自助式劳务众包平台，为企业提供各类商业检查和信息搜集，与传统市场调查方式相比节省了调查成本，并有效保证了调查数据真实性，缩短了调查的周期。因此 APP 成为该平台运行的核心，其中 APP 中的任务定价是其核心要素。假如定价不合理，有的任务就会无人接受，导致商品检查的失败。

本题有三个附件，附件一是已结束项目的任务数据，包含了每个任务的位置、定价和完成情况（“1”表示完成，“0”表示未完成）；附件二是会员信息数据，包含了会员的位置、信誉值、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；附件三是一个新的检查项目任务数据，只有任务的位置信息。

1.2 所要解决的问题

问题一：研究附件一中项目的任务定价规律，分析任务未完成的原因。

问题二：为附件一中的项目设计新的任务定价方案，并和原方案进行比较。

问题三：实际情况下，多个任务可能因为位置比较集中，导致用户会争相选择，一种考虑是将这些任务联合在一起打包发布。在这种考虑下，如何修改前面的定价模型，对最终的任务完成情况又有什么影响？

问题四：对附件三中的新项目给出你的任务定价方案，并评价该方案的实施效果。

二、模型假设

假设一：由于所研究区域相对地球整体较小，则假设研究区域近似为平面；

假设二：在问题一中，假设任务定价规律仅与任务本身有关，不考虑其他相关因素；

假设三：假设文章中所查找的相关参数都符合实际情况。

三、符号说明

符号	含义
d_{ij}	任务 i 到任务 j 的距离
x_i	任务 i 的纬度
y_i	任务 i 的经度
b_j	会员 j 的预定任务限额
B_j	会员 j 的信誉度

t_j	会员 j 的预定任务开始时间

四、问题分析

对于“拍照赚钱”APP 中的定价问题，我们按照以下思路理解。

分析问题一：附件一中数据庞大，包含任务的经纬度、任务标价以及执行状况。

要寻找任务的定价规律，可先利用每个任务的经纬度，设置按价格由低到高，由浅及深的颜色导入地图中进行直观观察。通过观察后发现任务定价大致呈现由若干个中心向四周逐渐递增的规律。所以应考虑地理位置与任务定价之间的关系，运用聚类分析法，对任务进行聚类，确定中心点坐标，接着利用最小距离对所有任务进行区域划分以找出任务定价考虑。

由于任务定价不能无限递增且中心点定价也不能太低，所以应考虑定价的上下限问题。而任务实际定价高低与行政区的工资水平相关，常用工资水平数据为工资水平最低标准与平均工资水平，所以应分析数据研究任务定价高低与其之间的关系。

分析问题二：问题一中研究了现有任务价格的定价规律，现此问题要求建立新的定价模型，使其任务的定价更为合理。考虑价格问题，首先应考虑其成本与收益问题，即任务方的成本最低与会员的收益最大，同时应该考虑定价能使任务完成率达到最大。因此可建立多目标优化模型，以成本、收益和完成率分别建立目标函数。结合实际情况，即会员任务预定限额、一个任务至多只能由一人完成等情况来建立约束条件。从而达到建立新的定价模型的目的。

分析问题三：题目中提出由于任务位置集中，导致用户会争相选择，思考将其打包联合发布。所以可根据任务与任务之间的距离进行打包，采用最小距离反复进行迭代计算将其打包分类。最后对问题二中建立的多目标优化模型进行修改，结合任务打包的因素，最后求解出模型与打包之前进行对比。

分析问题四：问题四要求对附件三中的新项目给出实际定价方案，此时应根据附件三中的数据，结合问题二问题三中的定价方法，得出其定价方案，然后根据任务完成率来判断出该方案的实施效果。

五、模型的建立与求解

5.1 问题一模型建立与求解

5.1.1 数据的预处理

根据对问题附件一中数据的分析，发现任务 A0297 与其他任务地理位置相隔较远，且结合附件二中会员地理位置来看，周围也无会员分布，所以应予以舍弃。

5.1.2 任务点分布大致观察

附件一中所包含的数据量很广，有 836 个任务的相关数据。为以最直观的形式观察其规律，先将任务相关信息直接导入地图中进行观察：

图 1 按价格分类的任务分布

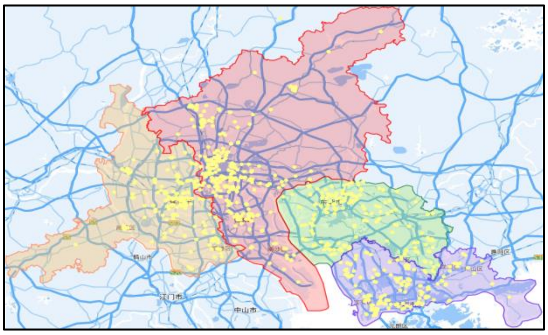


图 2 任务整体分布情况

将任务点按照其定价用不同的颜色进行分类，价格从低到高采用由浅及深的颜色进行体现，如图 1。由图 1 显示可以观察出，任务定价大致符合由若干个中心向四周逐渐递增的规律。

分析任务未完成原因时，由于不同城市人均收入水平、城市面积等相关因素的不同，需要对任务未完成的原因分城市进行探究。由图 2 可知，任务（图中黄色点）大多集中分布在广州、佛山、深圳、东莞四个城市。所以在接下来将分别对这四个城市的任务未完成原因进行讨论。

5.1.3 任务定价规律模型

5.1.3.1 基于任务地理位置的定价规律

根据以上分析可知，任务定价大致符合由若干个中心点向外扩散逐渐递增的规律。所以应首先确定每个中心点的位置，即每个区域中最低价格的任务的位置，然后根据各任务点到中心点的距离将所有任务点划分到不同的所属区域，最后对每个区域分别进行递增规律研究。

聚类分析法确定区域中心点

聚类分析法可用来描述每个任务之间的相似程度，由于定价与任务的地理位置有关，所以利用欧式距离作为聚类分析的标度，来度量样本间的相似程度，欧式距离具体计算方法如下：

$$d(x,j)=\left[\left(x_i-x_j\right)^2-\left(y_i-y_j\right)^2\right]^{\frac{1}{2}}$$

$$s.t\begin{cases} d(i,j)\ge 0, i,j\in\Omega \\ d(i,j)=0, i=j \\ d(i,j)=d(j,i), i,j\in\Omega \\ d(i,j)\le d(i,l)+d(j,l).i,j,l\in\Omega \end{cases}$$

其中i,j,l为任意三个任务，x_i,y_i分别表示任务i的经度与纬度。Ω所有任务所形成的集合。

Step 1 可建立距离矩阵：

$$D=\begin{pmatrix} d_{11} & d_{12} & \cdots & d_{1n} \\ d_{21} & d_{22} & \cdots & d_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ d_{n1} & d_{n2} & \cdots & d_{nn} \end{pmatrix}$$

Step 2 将 n 个任务首先分为 n 类，计算每类之间的平台高度

$$f=\min\left\{d(i,j)\right\}, i\in\Omega_1\ j\in\Omega_2$$

平台高度最低的为一类。

Step 3 计算新类与当前各类的距离，并不断重复步骤 Step 2，当且仅当类的个数为k时（k为理想模型分类数），则停止聚类。

通过以上步骤，结合任务定价呈现的大致规律，可以每个区域内价格最低的任务点的经纬度，计算每个区域的中心点，即：

$$r_{x_I}=\frac{\sum_{i}^{k_I}x_{i_I}}{k_I},r_{y_I}=\frac{\sum_{i}^{k_I}y_{i_I}}{k_I},$$

其中x_{i_I},y_{i_I}分别为第I个区域内第i个价格最低的任务的经度与纬度。k_I为第I区域内价格最低的任务的个数。

运用最小距离划分任务点所属区域

根据各区域的中心点位置，计算每个任务点到每个中心点的距离：

$$h_{i_I}=sqrt\left[\left(x_i-r_{x_I}\right)^2+\left(y_i-r_{y_I}\right)^2\right]$$

其中h_{i_I}表示第i个任务到第I个区域中心点的距离。然后求出每个任务点到每个区域中心点距离的最小值：

$$h_I=\min h_{i_I}$$

该任务点到哪个区域中心点的距离最小，则属于哪个区域中心点。

将所有任务点划分到所属区域后，则可得出整体定价规律。

5.1.3.2 基于行政区日工资水平的定价规律

任务定价呈中心点向四周扩散逐渐递增，但定价水平不能无限递增（考虑发出任务的商家的成本），中心点的定价水平也不能过低（价格过低会导致无会员完成任务），所以应考虑其定价的上下限水平。

假设会员日均能完成a个任务，任务的平均定价为C，则会员依靠完成任务的日均收入为：

$$C_{day}=a\times C$$

考虑日均收入时，即最常用的数据为会员所在行政区的最低日工资水平标准C_l，会员所在行政区的人均日工资水平C_a，对数据使用 Excel 进行简单处理，观察C与C_l、C_a之间的关系。

5.1.4 模型的求解

5.1.4.1 价格数据的描述性统计

首先将附件一中的数据按照价格进行描述性统计得到结果如下：

表 1 描述性统计结果

有效个数	均值	中值	众数	极小值	极大值
835	69.11	67.00	66	65	85

得出各个价格的频数分布利用扇形图表示为：

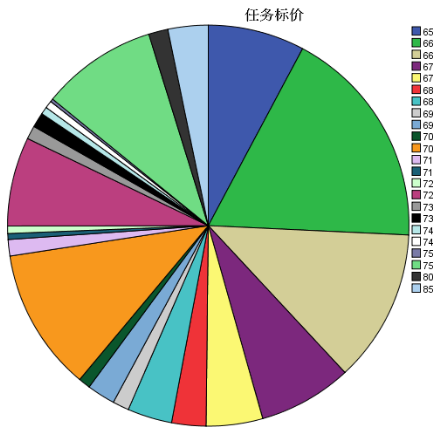


图 3 任务标价的频数

5.1.4.2 基于任务地理位置的定价规律求解

聚类分析法确定区域中心点

根据附件一中的数据，利用 MATLAB 软件画出任务点分布图，可进一步发现，任务点符合由若干个中心点向外扩散逐渐递增的规律。如下图：

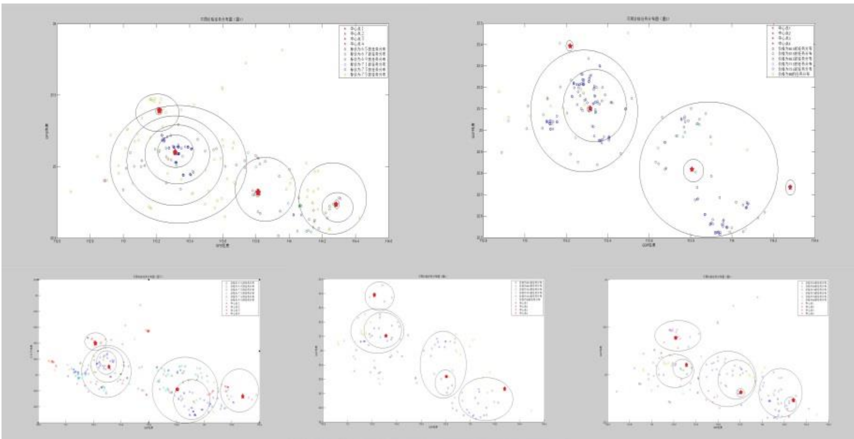


图 4 MATLAB 任务点分布图

利用 SPSS 软件聚类分析功能可求出 4 个中心点坐标为：

表 2 中心点坐标

区域	中心点经度	中心点纬度
1	113.2196	22.39286
2	113.3137	23.102885
3	114.2805	22.73565
4	113.8067	22.8185

运用最小距离划分任务点所属区域

再利用 spss 计算每个任务点到每个中心点的距离，找出最小距离，则可以将所有的任务点划分成 4 个区域，分析每个区域中任务定价和据中心点的距离的关系，令 R_i 为第 $i(i=1,2,3,4)$ 个区域中

任务位置到该区域中心点的距离，据此可以得到如下定价规律：

表 3 不同区域的定价规律

$R_1 \le$	$R_2 \le$	$R_2 \le$	$R_2 \le$	定价
0.00002	0.08653	0.18707	0.08495	65
0.13489	0.12144	0.19133	0.24749	65.5
0.14587	0.14763	0.22144	0.21089	66
0.12304	0.14638	0.24702	0.21089	66.5
—	0.14320	0.17168	0.17895	67
0.10723	0.18042	0.19092	0.20387	67.5
0.08897	0.18153	0.15305	0.22493	68
—	0.18747	0.12666	0.21825	68.5
0.15087	0.15053	0.14310	0.18597	69
—	0.22207	—	0.24065	69.5
0.22549	0.24853	0.16530	0.20074	70
0.06659	0.25064	0.10250	0.21817	70.5
—	0.30084	—	0.22125	71
0.04919	0.27497	—	0.17372	71.5
0.11289	0.26421	0.15326	0.19662	72
0.15763	0.29806	—	0.11310	72.5
—	0.32463	0.09892	—	73
—	0.31262	—	—	73.5
0.19381	0.29034	—	0.28490	74
0.14503	—	—	0.24725	74.5
0.21807	0.33535	0.19294	0.23696	75
0.34517	0.16652	0.28084	0.14943	80
0.20857	0.21764	0.19175	0.21813	85

据此，可得出任务定价规律为：在不同区域内以该区域中心点为中心在一定范围内逐渐向外扩散递增。

5.1.4.3 基于行政区日工资水平的定价规律求解

根据数据查找，广东省的人均最低日工资水平标准为 57.875 元[2]，人均日工资水平为 145.08 元[1]。根据附件一，任务定价范围在 65 元-85 元左右，通过 Excel 简单处理发现，当 $a=1,2$ 时，即会员每日完成 1-2 个任务时，满足：

$$C_l \le C_{day} \le C_a$$

即当会员每天完成 1-2 个任务时，任务酬金不小于日工资最低水平标准，不高于日均工资水平的定价规律。

5.1.5 任务未完成原因分析

微观经济学中有以下两个概念：

替代效应：一种商品的名义价格发生变化后，将同时对商品的需求量发生两种影响：一种是因

该种商品名义价格变化，而导致的消费者所购买的商品组合中，该商品与其他商品之间的替代。

收入效应：另一种是在名义收入不变的条件下，因一种商品名义价格变化，而导致消费者实际收入变化，而导致的消费者所购商品总量的变化。

表 4 不同城市情况表[1]

城市	人均收入水平（日/元）	城市面积（平方公里）
广州	160	7434
深圳	173.3	1996.85
佛山	128.1	3875
东莞	118.93	2465

由于不同城市人均收入水平、城市面积等相关因素的不同，现分别对广州、深圳、佛山、东莞这几个任务集中的行政区的任务未完成原因进行讨论。

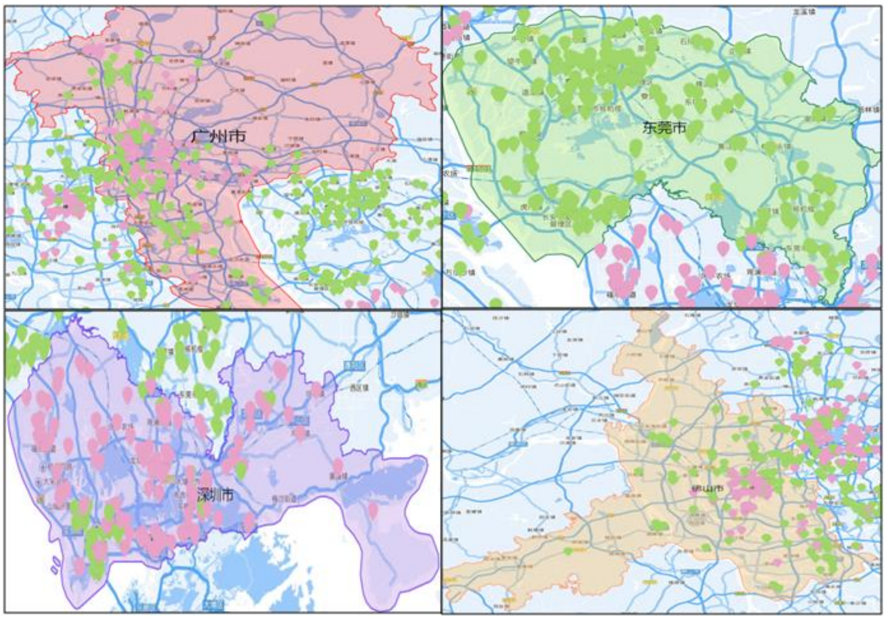


图 5 四个城市任务完成情况

以上图中粉红色点代表未完成任务，绿色点代表已完成任务。接下来对未完成任务的原因进行分析：

广州市：广州完成比例约为 60.8%。广州的日平均工资较高，任务酬金只能至多达到会员的日平均工资的一半左右，对于会员来讲，此时任务酬金对于会员来说的替代效应略大于收入效应，使得任务完成率略高于替代效应与收入效应相等时的完成率。

佛山市：佛山市的任务完成比例为 66.09%，佛山市未完成的任务较少，且集中在佛山市中心，较外围的任务均已完成。佛山市的收入水平位于四个城市最低的位置，此时任务的酬金对于会员来讲替代效应大于收入效应，所以任务完成情况较好。根据中心点来看，佛山市中心附近处于价格中心点地带，即酬金最低的位置，所以会员更愿意完成较远的任务以此来得到更多的酬金。

深圳市：深圳任务完成比例为 20.9%，已完成任务较少。深圳的人均工资水平较高，位于四个

城市之首。此时任务酬金给会员带来的替代效应远小于收入效应，所以人们更愿意选择闲暇而放弃任务的酬金。

东莞市：东莞市的任务完成比例为 100%，原因在于东莞市的收入水平处于较低的状态，城市面积也相对最小，任务酬金给会员带来的替代效应远大于收入效应，且到达任务点距离范围较短。所以任务完成情况最好。

5.2 问题二模型建立与求解

问题一对附件一进行了分析，得出了依据现有价格的任务定价规律。现结合附件二中的会员信息，针对任务成本、会员收益、以及任务完成比三个角度，来建立多目标优化模型，为任务制定新的定价方案。

5.2.1 数据的预处理

对问题附件二中数据的进行分析，发现一些会员离其最近的任务点相隔较远，需要驾车前往，对有车的会员来讲，任务酬金带来的收入效应远大于替代效应，对于相隔较远的任务点该类会员并不会驾车前往完成。所以应对间隔最近的任务点距离大于 100 公里的数据予以舍弃，即 B0136、B1708、B0039、B0472、B0048、B1822、B0082、B0022、B1712。

5.2.2 多目标混合整数优化任务定价模型

5.2.2.1 模型目标函数的建立

任务的价格与任务发行方和会员相关，即与成本和收益相关。发行方希望将其成本控制尽量小，会员则希望收益尽量大。同时，在本题中，任务完成率也是一个需要关注的问题，任务价格也需要尽可能的使任务完成率达到最大。

任务发行方成本目标函数

考虑定价的合理性，任务发行方成本自然是不可忽略的重要因素，在发行任务时，需要考虑将其成本在一定条件下降到最低。建立任务成本函数：

$$\min g_1 = \sum_{i=1}^{m} w_i$$

其中 $w_i \geq \left(i = 1, 2 \cdots N\right)$ 为任务 i 的定价。

任务完成率目标函数

在达到以上两个目标的同时，还需要考虑其任务完成情况，定价过高或过低都会影响其任务完成情况。所以应考虑将任务完成率达到最大。

$$\max g_3 = \frac{\sum_{i=1}^{n}\sum_{j=1}^{m} y_{ij}}{n}$$

其中 $y_{ij} = 1$ 时表示任务 i 会被会员 j 完成，否则则为 0。，n 为任务总数。

会员收益期望目标函数

前两个目标函数都是从任务方角度出发，现从会员方来看，建立会员方的目标函数。会员在执行任务时会先考虑其任务的价格合理性，根据 5.1.5 可知，当任务酬金给会员带来替代效应大于收入效益时，会员才会执行其任务。所以会员期待任务酬金较高。建立会员收益期望目标函数：

设 P_{ij} 为任务 i 被会员 j 执行的概率。则 P_{ij} 与匹配半径 r_0 的会员数、任务 i 周围任务的密度、信誉度、任务预订时间相关。

P_{ij} 与匹配半径 r_0 内的会员数成反比关系，即距离越大，概率越小。即：

$$p_{ij} = k_1 e^{-d_{ij}}$$

$$s.t. p_{ij} = \begin{cases} k_1 e^{-d_{ij}}, d_{ij} \le r_0 \\ 0, otherwise \end{cases}$$

P_{ij} 与任务 i 的周围任务密度成正比关系，即：

$$p_{i1} = k_2 N_i$$

其中 N_i 为任务 i 周围的任务个数， N_i 反应了任务 i 与周围任务的密度程度。

同理， P_{ij} 也与会员的信誉度成正比关系，和开始预订任务的时间成正比关系，即：

$$p_{1j} = k_3 B_j$$

$$p_{2j} = k_4 e^{-T_j}$$

其中 B_j 为每个会员 j 的信誉度。

$$T_j = \left(t_j - t_{\max}\right) \times 60$$

其中 t_j 为会员 j 开始预订的时间， $t_{\max}$ 为会员最晚开始预订时间。

利用概率之和为 1，确定其参数 $k_i \left(i = 1, 2, 3, 4\right)$：

$$\sum_{i=1}^{n} \sum_{j=1}^{m} p_{ij} = 1$$

$$\sum_{i=1}^{n} p_{i1} = 1$$

$$\sum_{i=1}^{m} p_{1j} = 1$$

$$\sum_{i=1}^{m} p_{2j} = 1$$

从而得出为：

$$P_{ij} = \omega_1 p_{ij} + \omega_2 p^{'}$$

ω_{11}, ω_{12} 为信誉度与开始预订时间的线性权重， ω_1, ω_2 为 p_{ij} 与 $p^{'}$ 之间的线性权重。其中：

$$p^{'} = p_{i1}{}^{T}\left(\omega_{11}p_{1j} + \omega_{12}p_{2j}\right)$$

$$s.t. p^{'} = \begin{cases} p^{'}, d_{ij} \le r_0 \\ 0, otherwise \end{cases}$$

则会员收益期望目标函数为：

$$\max g_2 = \sum_{i=1}^{n}\sum_{j=1}^{m} w_i y_{ij} P_{ij}$$

5.2.2.2 模型约束条件的建立

考虑会员预定任务限额

根据附件二，可知每个会员有规定的任务限额。在会员预定任务时，会员预定的任务总额应不超过其预定限额。即：

$$\sum_{i=1}^{n} y_{ij} \le b_j, j = 1, 2 \cdots m$$

其中b_j为第 j 个会员的预定限额

一个任务至多一个会员完成

一个任务不可能由多个会员同时完成，即：

$$\sum_{j=1}^{m} y_{ij} \le 1, i = 1 \cdots n$$

任务价格与当地收入水平相关

考虑任务定价应符合当地收入水平的行情，高于当地的收入水平，即：

$$w_i \ge \zeta_i, i = 1, \cdots, n$$

其中ζ_i是第i个任务的当地收入水平。

y_{ij}应满足条件

$$y_{ij} = \begin{cases} 1 \\ 0 \end{cases}$$

当任务i被会员 j 执行时，$y_{ij} = 1$；否则$y_{ij} = 0$。

5.2.3 模型的求解

结合附件二会员信息的数据，利用前面所提到的多目标优化任务定价模型，在考虑到会员收益期望目标函数时，利用 matalb 求出参数k_1, k_2, k_3, k_4的值分别如下：

表 5 参数表

参数	k_1	k_2	k_3	k_4
数值	1.059036882	7.265857734	0.00000192	0.021171

参数已知后则可以得到任务i被会员 j 完成的概率 p_{ij}，则可以建立多目标优化模型，利用 lingo

软件求解可得每个任务的价格，节选部分价格如下（完整定价文件见支撑材料）：

表 6 部分任务定价表

任务编号	gps 经度	gps 纬度	任务执行情况	原始定价	重新定价
A0001	113.9808368	22.56614225	0	66	79.42688
A0002	113.9405252	22.68620526	0	65.5	81.39167
A0003	113.957198	22.57651183	1	65.5	74.25246
A0004	114.2445711	22.56484081	0	75	66.24778
A0005	113.9507227	22.55888775	0	65.5	68.54218
A0006	114.2413174	22.55899906	0	75	65.72086
A0007	113.9722597	22.54900371	1	65.5	69.60256
A0008	113.9565735	22.56277351	0	65.5	81.49829
A0009	113.8956606	22.50001192	0	66	84.70732
A0010	113.9239778	22.5437861	1	66	66.18471
A0011	113.9308596	22.52486369	0	65.5	70.94052
A0012	113.9358436	22.519087	0	65.5	69.83677
A0013	113.977909	22.54797243	1	65.5	69.24743
A0014	113.9314284	22.50616871	1	66	67.73667

还可以得出整个地区的任务完成比约为 70.33%，比之前的 62.44%增大约 8 个百分比点，说明该模型在原来的基础上有一定的优化作用。

5.3 问题三模型建立与求解

5.3.1 任务打包模型

利用距离矩阵：

$$D=\begin{pmatrix} d_{11} & d_{12} & \cdots & d_{1n} \\ d_{21} & d_{22} & \cdots & d_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ d_{n1} & d_{n2} & \cdots & d_{nn} \end{pmatrix}$$

Step 1 寻找距离矩阵中的最小距离：$\min\left(d_{ik}\right), i \neq k$

Step 2 寻找除任务 k 外距离任务 i 的最小距离 $\min\left(d_{ia}\right), i \neq a$，同时寻找除任务 i 外距离任务 k 的最小距离 $\min\left(d_{bk}\right), b \neq k$。比较 $\min\left(d_{ia}\right)$ 与 $\min\left(d_{bk}\right)$ 的最小值，即：

$$\min\left[\min\left(d_{ia}\right), \min\left(d_{bk}\right)\right]$$

则找出距离最小的任务加入任务 i,k 的打包。以此步骤进行迭代，直到 $\max d_I \leq d_0$，其中 d_0 为所能容忍的打包任务之间的最大距离，d_I 为打包任务集中的各个任务之间距离的集合。并计算打包集合中任务之和 α_i。

Step 3 剔除被打包的任务任务后，以 Step 1、Step 2 循环执行此步骤。直到将所有任务全部打包结束。

5.3.2 基于任务打包模型的定价模型修改

5.3.2.1 目标函数的修改

引入目标变量打包任务集i的定价$w_i^{'}\left(i=1,2\cdots n^{'}\right)$与：

$$y_{ij}^{'}=\begin{cases}1\\0\end{cases}$$

当打包集i被会员j执行时，$y_{ij}^{'}=1$；否则$y_{ij}^{'}=0$。

根据问题二的多目标优化定价模型，打包之后。目标函数修改为：

任务发行方成本目标函数

由于我们将任务打包为任务集，则任务成本应为打包任务集价格之和，即：

$$\min g_i^{'}=\sum_{i=1}^{n^{'}}w_i^{'}$$

任务完成率目标函数

$y_{ij}^{'}$表示打包任务集i是否被会员j执行，则所有被完成的任务应为：

$$\sum_{i=1}^{n^{'}}\sum_{j=1}^{m}y_{ij}^{'}\alpha$$

则任务完成率应为：

$$\max g_2^{'}=\frac{\sum_{i=1}^{n^{'}}\sum_{j=1}^{m}y_{ij}^{'}\alpha_s}{n}$$

会员收益期望目标函数

由于任务被打包，则打包任务集i被会员j执行的概率为$P_{ij}^{'}$ 应为其任务集内各个任被会员j执行的概率之和：

$$P_{ij}^{'}=\sum_{i\in I_s}^{\alpha_s}p_{ij}^{'}\left(s=1.2\cdots Q\right)$$

其中I_s表示为第s个打包集合，Q为所有打包集合数，α_s为第s个打包任务集合中任务总数。则目标函数应为：

$$\max g_2^{'}=\sum_{i=1}^{n^{'}}\sum_{j=1}^{m}w_i^{'}y_{ij}^{'}P_{ij}^{'}$$

5.3.2.2 约束条件的修改

考虑会员预定任务限额

被预定的任务个数应小于会员预定限额，而预定任务个数应为选取的各打包集中的任务个数之和与单个预定任务数求和。即：

$$\sum_{i\in I} y_{ij}^{'}\alpha_s + y_{ij} \le b_j, j=1,2,3\cdots m$$

一个任务至多一个会员完成

每个打包任务集至多只能被一个会员完成，即：

$$\sum_{j=1}^{m} y_{ij}^{'} \le 1, i=1,2\cdots n^{'}$$

任务价格与当地收入水平相关

打包任务集中每个任务平均分摊的酬金应大于当地的收入水平。同时，每个独立的任务的定价也应大于当地收入水平。即：

$$w_1^{'} \ge \alpha_s a_i \left(i=1\cdots n^{'}, i\in I\right)$$
$$x_i \ge a_i \left(i=1\cdots n\right)$$

$y_{ij}^{'}$ 应满足条件

$$y_{ij}^{'} = \begin{cases} 1 \\ 0 \end{cases}$$

当打包集 i 被会员 j 执行时，$y_{ij}^{'}=1$；否则 $y_{ij}^{'}=0$。

5.3.3 问题三模型求解

将附件一中的数据按照距离运用 SPSS 进行打包，共分成了 89 个打包任务集，每个集有 1 到 43 个任务点不等，如图（为使图清晰直观，只包含了部分集）：

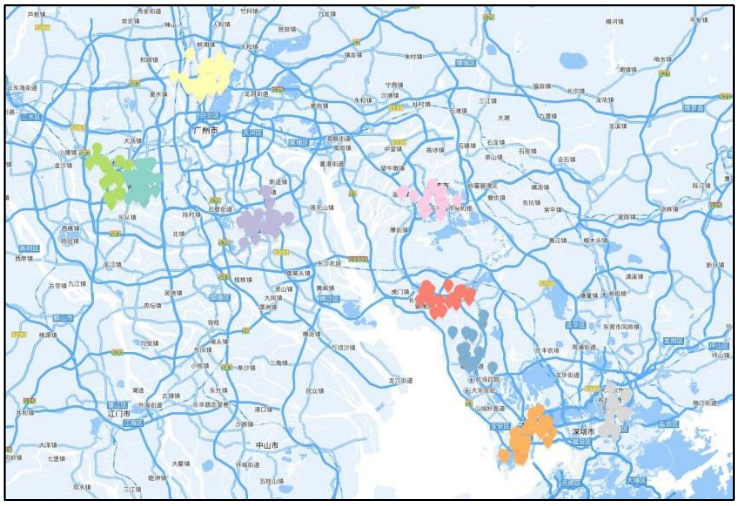


图 6 部分打包分区情况

建立多目标优化模型，用 Lingo 进行求解，得到每个集的定价如下：

表 7 部分集定价表

分组情况	任务标价_max	任务标价_min	个数	重新标价
1	72	65	31	2070
2	75	65	18	1196
3	75	65	43	2844
4	80	65	43	3251
5	67	65	29	1903
6	68	65	12	787
7	66.5	65	13	849
8	75	65	36	2407
9	80	65	19	1245
10	85	65	21	1445
11	72.5	65	20	1375
12	69.5	65	36	2356

在此种定价模型下的任务完成比为 72.89%，对比未打包的情况任务完成比有所提高。

5.4 问题四：附件三新项目的任务定价

5.4.1 采用问题二、问题三中的模型求解对比

根据问题二以及问题三所建立的模型，结合附件三中的数据情况，可考虑使用问题二中多目标混合整数优化模型以及问题三中打包修改过后的多目标混合整数优化模型对其进行打包定价求解。将附件 3 的数据运用模型二来进行求解，得到重新标价，如下图（下图为部分数据）：

表 8 采用问题二模型求解的部分新项目任务集定价表

任务号码	任务 GPS 纬度	任务 GPS 经度	重新标价
C0001	22.73004117	114.2408795	67
C0002	22.72704287	114.2996199	81
C0003	22.70131065	114.2336007	69
C0004	22.73235925	114.2866672	73
C0005	22.71839144	114.2575495	79
C0006	22.75392493	114.3819253	65
C0007	22.72404221	114.2721836	76
C0008	22.71937803	114.2732478	85
C0009	22.73028254	114.2304955	84
C0010	22.7187968	114.267027	74
C0011	22.65746229	114.3476957	74
C0012	22.71611614	114.2474716	72
C0013	22.72986374	114.2939012	77
C0014	22.746174	114.285869	69
C0015	22.7333086	114.2943071	71
C0016	22.74551648	114.2744633	69
C0017	22.72282036	114.2654813	77
C0018	22.72492675	114.2731306	79

计算出任务的完成率为 68.25%。

现将附件三中新的任务项目进行打包，共分成了 517 个任务打包集，采用问题三的模型，用 lingo 软件进行求解，得到每个任务打包集的定价，部分数据如下图：

表 9 采用问题三模型求解的部分新项目任务集定价表

分集	个数	重新标价
1	8	517
2	3	202
3	15	1190
4	5	359
5	3	218
6	1	75
7	11	840
8	4	243
9	7	458
10	6	365
11	21	1280
12	4	274
13	2	156

得出在此种定价模型下的任务完成比为 72.91%，说明该方案实施效果较好。

通过对比其完成比度可以发现，采用问题三中将任务打包后再利用模型定价的方式的任务完成率比采用问题二中直接利用模型定价的方式的任务完成率略高，说明对附件三中的新项目可先将任务打包后再发布。

5.4.2 针对目标函数变化规律的加权分析

由 5.4.1 可得出，附件三中的新项目适合将其打包后再发布。即采用问题三中的模型更为合理。现对问题三中根据任务打包修改后的多目标混合整数优化模型的三个目标函数进行线性加权，通过任意改变其权重可以得出：

表 10 权重变化表

权重比例的变化	ω_1	ω_2	ω_3
1	0.3	0.3	0.4
2	0.3	0.5	0.2
3	0.5	0.3	0.2
4	0.5	0.1	0.4

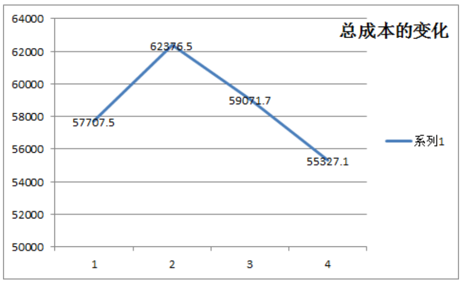
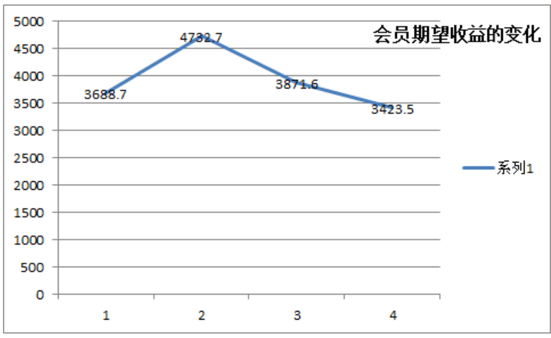


图 7 任务方发行成本随权重变化情况



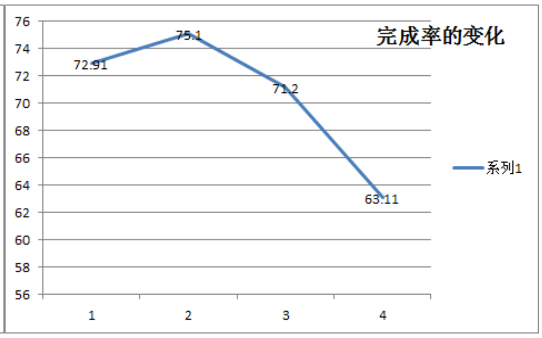


图 8 会员收益期望随权重变化情况　　　　**图 9 任务完成率随权重变化情况**

结合表中数据与折线图可知，会员收益期望随着权重的增加而上升，任务发行方成本、任务完成率随着权重的上升而下降。

六、模型评价

6.1 模型的优点

1、本文数据与图形相结合，与实际情况基本吻合，模型具有可行性；

2、问题二从任务发布者和会员两方面同时考虑，并将定价与任务完成率同时考虑，使定价模型较为严谨；

3、问题三定义了一个任务打包算法，以任务距离进行打包对任务打包考虑完善。

6.2 模型的缺点

1、由于问题二数据较多，考虑较为复杂，计算有所不便。

七、参考文献

[1]广东省统计信息网 http://www.gdstats.gov.cn/
[2] http://www.liuxue86.com/a/3114409.html

附录

一、计算任务 i 被会员 j 完成的概率 pij（操作环境：MATLAB）

将任务经纬度、会员经纬度、会员荣誉值、会员开始时间的标准化数据分别导入到 matlab 中形成 A、B、C、D 四个矩阵，再进行如下操作：

1. 考虑到任务 i 与会员 j 之间的距离 d，算出 pij1

```
function d=d1(A,B)
d=zeros(835,1877);
for i=1:835
    for j=1:1877
       d(i,j)=sqrt((A(i,1)-B(j,1))^2+(A(i,2)-B(j,2))^2);
    end
end
k1=1/sum(sum(exp(-d)))
pij1=zeros(835,1877);
for i=1:835
```

```
    for j=1:1877
        if d(i,j)<1.2
            pij1(i,j)=k1*exp(-d(i,j));
        else
            pij1(i,j)=0;
        end
    end
end
```

2. 考虑任务 i 周围的密集程度，计算概率 pi

```
function c=c1(A)
c=zeros(835,835);
for i=1:835
    for j=1:835
      c(i,j)=sqrt((A(i,1)-A(j,1))^2+(A(i,2)-A(j,2))^2);
    end
end
n=zeros(835,1);
n(1,1)=0;
for i=1:835
    for j=1:835
        if c(i,j)<=0.05
            n(i,1)=n(i,1)+1;
        else continue
        end
    end
end
k2=1/sum(n)
pi=k2.*n
```

3. 考虑到会员 j 的荣誉度，计算概率 pj1

```
k3=1/sum(C)
pj1=k3.*C
```

4. 考虑到会员的开始时间，计算概率 pj2

```
k4=1/sum(exp(-D))
pj2=k4.*exp(-D)
```

5. 通过加权算出总概率 pij

```
pj=0.5*pj1+0.5*pj2
pij2=pi*pj'
pij=0.5*pij1+0.5*pij2
```

二、求解多元优化混合整数规划模型（操作环境：lingo）

1.利用附件一中的数据参照规划模型进行求解

```
model:
sets:
task/1..835/:x;
```

```
person/1..1877/:b;
link(task,person):y,p;
endsets
data:
p=@ole('C:\Users\JSJSYS\Desktop\jz.xlsx','p');
b=@ole('C:\Users\JSJSYS\Desktop\限额.xlsx','b');
a=67.66;
enddata
max=0.3*@sum(task(i):(@sum(person(j):x(i)*y(i,j)*p(i,j))));
min=0.3*@sum(task(i):x(i);
max=0.4*(@sum(link:y(i,j)))/835;
@for(person(j):
   @sum(task(i):y(i,j))<=b(j));
@for(task(i):
   @sum(person(j):y(i,j))<=1);
@for( task(o):x(o)<=100.66);
@for(task(i):x(i)>=67.66);
@for(link:@bin(y);

End
```

2.将附件一中的任务进行打包后参照规划模型求解

```
model:
sets:
task/1..89/:x;
person/1..1877/:b;
link(task,person):y,p;
endsets
data:
p=@ole('C:\Users\JSJSYS\Desktop\jz.xlsx','p');
b=@ole('C:\Users\JSJSYS\Desktop\限额.xlsx','b');
a=67.66;
enddata
max=0.3*@sum(task(i):(@sum(person(j):x(i)*y(i,j)*p(i,j))));
min=0.3*@sum(task(i):x(i);
max=0.4*(@sum(link:y(i,j)))/89;
@for(person(j):
   @sum(task(i):y(i,j))<=b(j));
@for(task(i):
   @sum(person(j):y(i,j))<=1);
@for( task(o):x(o)<=100.66);
@for(task(i):x(i)>=67.66);
@for(link:@bin(y));
End
```