

“拍照赚钱”APP 平台的任务定价的优化设计

摘要

本文主要通过对已有的某“拍照赚钱”APP 平台提供的任务项目及会员的相关信息进行研究,最终建立了具有较高实用性的任务打包定价的优化模型.

针对问题一,我们在对已知数据进行预处理的基础上,对附件一中任务按定价进行分档,结合任务的经纬度及附件二中的会员经纬度信息进行数据挖掘.通过对任务定价和人员密度的定性分析发现:该平台的任务定价主要与任务附近的人员密度呈负相关.另外,任务定价与任务复杂程度呈正相关.进一步,将任务按是否完成分为两类进行对比分析可知导致任务未完成的主要原因有以下 4 类:任务复杂程度过高、周围人员密度较小、任务周围的会员距离程度较大、任务定价过低.

针对问题二,在问题一中对平台原定价规律和未完成任务的原因分析的基础上,我们确定了影响任务定价的主要因素有:会员密度 ρ 、会员距离程度 L ,建立这两个变量与任务定价间的函数关系式,并基于穷举法对函数关系式中的参数进行优化求解,最终建立新的任务定价模型为:

$$\begin{cases} P = -3.06\rho + 0.04L + 66.32 \\ \rho = \frac{n_1}{\pi r^2} \\ L = \frac{\sum_{k=1}^{10} l_k}{10} \end{cases}$$

利用新的任务定价模型对附件一中任务进行重新定价.进一步,我们从会员及 APP 平台两个角度定义了 3 个任务定价评价指标:任务完成率 α 、任务总定价 β 、平均性价比 γ .并据此对原定价方案和新的定价方案进行综合评价,分析可知:新的定价方案下任务完成率 $\alpha = 80.23\%$ 远大于原完成率 62.51% ,任务的总定价变化不大,会员的平均性价比 $\gamma = 2.18$,说明会员的满意度较高,新的任务定价模型具有较好的实用性.

针对问题三,考虑到实际生活中存在多个任务集中分布的现象.我们基于图论与规划知识,以任务为节点,建立无向赋权网络图 $G = (V, E, D^*)$.利用就近原则对任务进行打包,对问题二中建立的任务定价模型进行改进.最后,利用附件一中的数据对该模型进行检验,我们发现:打包后的新方案的完成率 α 高达 93.37% ,劳务双方的效益大大提高;任务总定价 $\beta = 32457.9$ 元,远低于前两种方案,公司的经济成本将大大降低,收益将大大增高;会员的平均性价比 $\gamma = 2.57 > 2.18$,说明会员的满意度有了较好的提高.综上可知,我们建立的基于图论知识的任务定价的优化模型具有较高的实际价值.

针对问题四,为探究基于图论知识的任务定价的优化模型的普适性,我们利用该模型对附件三中的新的 2066 个任务进行定价,并求出任务的完成率 $\alpha = 92.35\%$ 、任务总定价 $\beta = 56062.3$ 元,平均性价比 $\gamma = 2.82$,任务的完成率非常高,任务总定价较低,平均性价比较高,则 APP 平台和会员的满意度均较好,说明该定价方案较为合理,具有较好的普适性,值得广泛应用于实际的平台运营中.

关键词: 任务定价 图论模型 完成率 参数优化

一. 问题重述

1.1 问题背景

当今，随着互联网+行业的发展，“拍照赚钱”为移动互联网下产生的一种新型自助式服务模式. 会员通过从 APP 上领取需要拍照的任务从而赚取 APP 对任务所标定的酬金. 相比传统的市场调查方式，这种基于移动互联网的自助式劳务众包模式可为企业提供各种商业检查和信息搜集，大大节省调查成本，同时有着有效地保证调查数据真实性，缩短调查的周期的优势. APP 为该平台运行的核心，但 APP 中的核心要素为任务定价. 若任务的定价不合理，可能该任务无人问津，致使任务发布失败将对任务发布企业、运营平台及会员的利益造成损失，故我们现在亟待解决任务的合理定价问题.

由某“拍照赚钱”的运行平台我们获得了相关任务和会员信息如下：

1) 已结束任务数据（附件一中）：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）；

2) 会员信息数据（附件二中）：会员的位置、信誉值、以及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（任务分配时实际上是根据预订限额所占比例进行配发）；

3) 新的检查项目任务数据（附件三中）：仅仅有任务的位置信息.

1.2 问题提出

试建立相关模型，解决下列几个方面问题：

问题一：首先由附件一给出的已结束项目的经纬度、任务定价及附件二中的会员的经纬度探究任务定价规律，再从任务位置、任务定价和会员分布等方面分析任务未完成的原因.

问题二：在问题一的基础上，分析影响定价的因素建立相关指标，从而建立新的任务定价方案；建立任务定价合理性的评价指标，通过和原方案结果进行比较对模型的合理性进行检验.

问题三：考虑到实际中会有多个任务可能因为位置比较集中，导致用户会争相选择，故从该角度将这些任务联合在一起打包发布. 对问题二中建立的定价模型进行进一步的改进，并分析所得结果.

问题四：利用附件三中所给的新项目相关数据对问题三中建立的考虑任务集合打包的任务定价方案进行检验，并讨论该方案的实施效果和实际价值.

二. 模型假设

- 1、假设任务发布时间早于会员开始预定时间；
 - 2、假设任务打包后只视作一单；
 - 3、假设附件中所给的任务数据和会员信息数据都是真实可靠的；
 - 4、假设性价比较高时，会员就一定会选择该任务，即该任务一定能被完成；
 - 5、假设 APP 上注册的会员都是真实有效的，即不存在只注册却不做任务的情况；
- （其余假设在文中另作说明）

三. 符号说明

符号	说明
ρ	会员密度
L	会员距离程度
P	任务定价
S	抢单成功率
d	任务行程
D	任务复杂程度
T_{ij}	第 j 个任务相对于第 i 位会员的会员总成本
Q_{ij}	第 j 个任务相对于第 i 位会员的任务性价比
α	任务完成率
β	任务总定价
γ	平均性价比
$G = (V, E, D)$	以任务为节点，以任务间距离赋权的无向赋权网络图

(其余符号在文中另作说明)

四. 问题一的分析与求解

4.1 问题分析

该问要求我们通过对已结束任务的经纬度及定价、APP 平台上注册会员的经纬度等数据进行分析研究任务的定价规律，并进一步探究已结束任务中部分任务未完成的原因。对此，我们从以下方面对问题进行求解：

针对任务定价规律的探究：1) 我们对附件一中给出的已结束项目的经纬度、定价等变量进行初步分析，发现任务的定价按档分布；2) 我们选取具有代表性的 3 档定价的任务，利用散点图对这 3 档任务的位置与会员的位置的关系进行分析；3) 考虑建立会员密度等指标定性分析任务的定价规律。

针对任务未完成原因的分析：1) 我们对已结束的任务项目根据是否完成进行分类；2) 从会员位置角度分析任务未完成的原因；3) 从任务定价角度分析任务未完成原因。

4.2 问题求解

4.2.1 数据预分析

(1) 误差数据的纠正

对于附件二中会员信息的数据中，我们发现部分数据存在明显错误，故先对其进行处理：会员信息表中会员编号为“B11175”的会员的位置信息中经度为：23.031824°，纬度为：113.131483°。经纬度明显反置，我们对其进行修正为经度为：113.131483°，纬度为：23.031824°。

（2）任务定价分档

对于附件一中的数据进行初步分析，我们发现该工作表中已结束的项目任务存在明显的分档划分现象，对此我们将任务项目按定价进行分类如下表 1：

表 1 已结束任务的定价分档表

分档	任务定价（元）	任务编号	任务数量	占比
1	65	A0022、…、A0752	65	7.78%
2	65.5	A0002、…、A0834	150	17.96%
…	…	…	…	…
21	75	A0004、…、A0757	78	9.34%
…	…	…	…	…
40	84.5	/	0	0
41	85	A0450、…、A0835	27	3.23%

（注：占比=任务数量/总任务数量×100%）

由上表我们可知：

1）任务定价为 65.5 元的占比最大为 17.96%；任务定价范围在 65 元-70 元的任务总占比达到 72.69%，说明该平台定价的大致范围集中在 65 元-70 元。

2）任务定价为 80 元及以上的任务中，定价为 80 元的仅 13 项，定价为 85 元的仅 27 项，总占比仅 4.79%，说明该类定价较少。

4.2.2 任务定价与会员位置关系的分析

为探究任务定价与会员位置的关系，我们以具有代表特征的定价三档定价任务：定价为 65 元的低价档、定价为 75 元的中价档、定价为 85 元的高价档的三档为例，利用 MATLAB 软件结合任务的经纬度信息和所有会员的经纬度信息，绘制出三档任务与会员分布的散点图如下图 2：

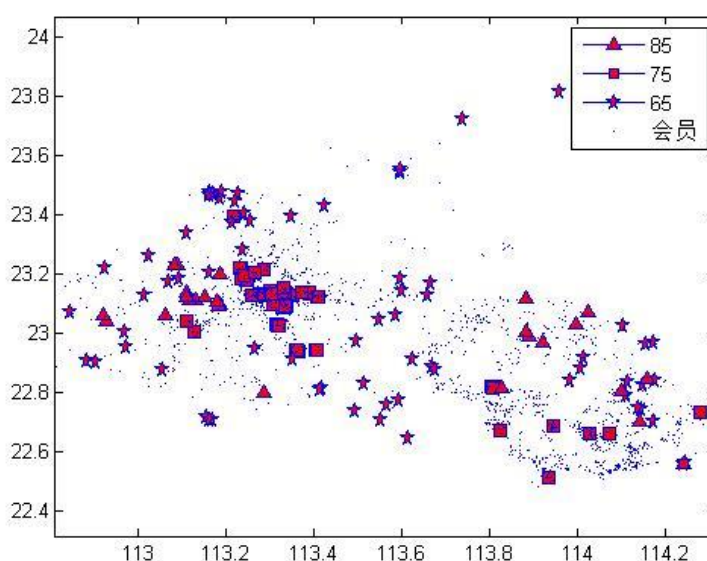


图 2 三档定价任务与会员分布散点图

由图 2, 我们不难发现:

1) 定价为 65 元低价档任务周围的会员较为密集, 定价为 75 元的中价档任务周围会员分布较为均匀, 定价为 85 元的高价档任务周围会员分布较为稀疏;

2) 任务的定价大致以由低到高向外扩散; 部分定价为 85 元的高价任务分布在人员较多的区域, 结合实际我们假定此类任务可能由于任务本身的复杂程度较高, 会员完成的成本较高.

4.2.3 任务定价和人员密度的定性分析

利用图 2 散点图我们发现任务的定价与会员的位置分布存在一定的联系, 故定义会员密度这一指标对任务定价规律进行定性分析.

(1) 会员密度指标的确定

以所研究的定价任务为圆心, 以 r 为半径确定影响区域, 会员密度 ρ 即该区域内会员数量与区域面积的比值:

$$\rho = \frac{n_1}{s_1} = \frac{n_1}{\pi r^2} \quad (1)$$

其中, n_1 表示会员数量, s_1 表示区域面积, r 为影响区域半径. 会员密度 ρ 越大, 说明单位区域内会员数量越多.

(2) 基于人员密度的人物定价的规律探究

利用式 (1), 我们取定 $r = 10km$, 求出所有已结束任务周围的会员密度, 结合每项任务的定价, 我们可知任务的定价与任务周围会员密度的关系图如下:

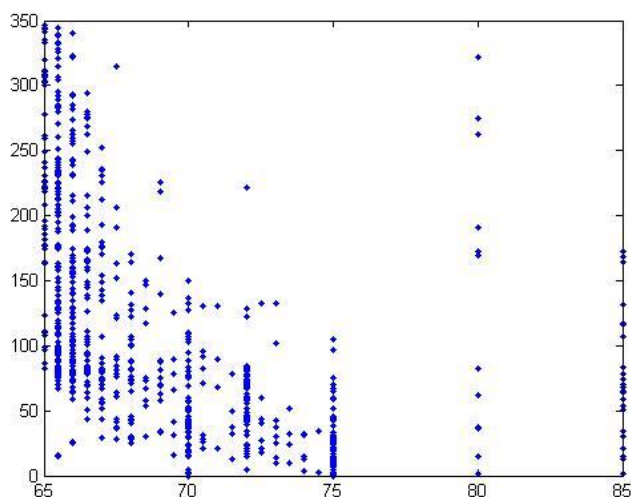


图 3 任务定价与会员密度关系图

由图 3 可知：

1) 任务的定价大致集中在 65 元-75 元的范围内。

2) 任务定价和会员密度大致呈负相关. 我们可以发现一定范围内, 会员的密度越大, 任务的定价越底; 会员的密度越小, 任务的定价越高, 说明平台在进行任务定价时, 会员密度是重要影响因素。

3) 另外, 部分高定价任务周围的人员密度也较大, 查阅相关资料, 我们可知拍照任务的复杂程度不同, 该类任务的复杂程度可能较高, 故任务定价与任务的难度亦存在一定联系, 任务难度越高, 任务定价越高。

4.2.4 任务未完成原因分析

为更好的研究任务完成与未完成的原因, 进行导致任务未完成原因的分析. 我们首先将任务按照是否完成分为两类：

表 2 任务完成情况分类表

类别	任务执行情况	任务编号	任务 GPS 经度	任务 GPS 纬度	任务标价
1	0 (未完成)	A0001	22.56614225	113.9808368	66
	
		A0831	23.0440624	113.125784	65.5
2	1 (完成)	A0003	22.57651183	113.957198	65.5
	
		A0835	23.12329431	113.1103823	85

为更直观的分析未完成部分的任务的原因, 我们将上表 2 中两类任务的位置以散点图的形式呈现如下：

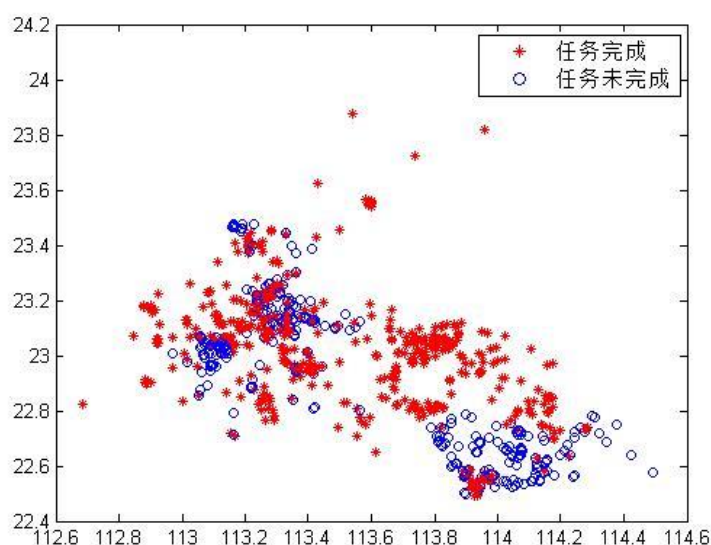


图 4 已结束任务的完成情况图

由图 4 我们可知：未完成的任务主要在整个区域的右下角、左下、中上三个子区域内非常集中；完成的任务整个区域的分布较均匀. 查阅相关资料，我们分析导致这三类子区域集中未完成任务的可能原因是由于任务的复杂程度较高、会员完成此类任务的性价比不高.

进一步，为讨论未任务完成原因和人员位置之间的关系，我们以编号为 A0153（已完成）、A0038（未完成）、A0025（未完成）的 3 个任务为例，取定 $r = 10km$ 进行研究.

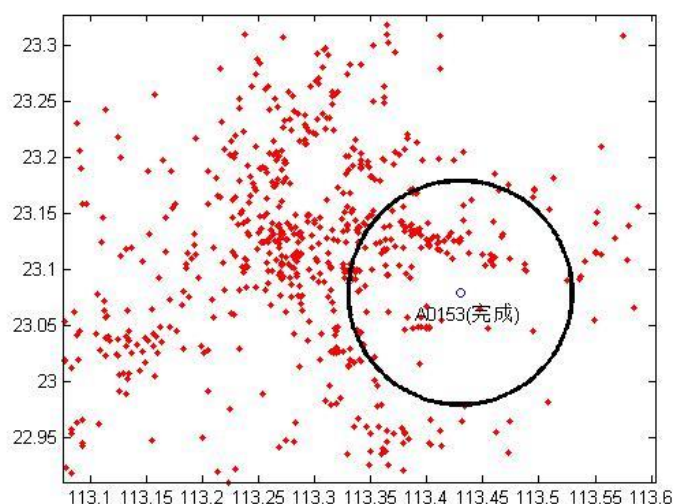


图 5 编号为 A0153（完成）的任务周围会员分布情况图

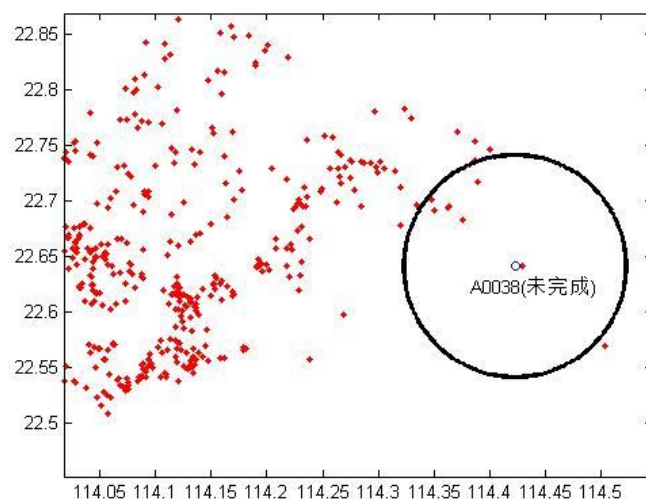


图 6 编号为 A0038（未完成）的任务周围会员分布情况图

由图 5 和图 6 对比，我们发现：

1) A0153 的标价为 70.5 元，周围的会员密度为 130；A0038 的标价为 75 元，周围的会员密度为 7. 我们发现这两个任务的标价均处于中档价位，但未完成的任务 A0038 周围的会员密度较少，故我们分析导致一部分任务无人完成是由于部分任务周围的会员密度过小.

2) 当距离任务最近的会员数 n 给定时，我们取定 $n=10$ ，可得 A0038 周围最近的 10

个会员到任务的平均距离明显小于 A0153 周围最近 10 个会员的到任务的平均距离。

对此，我们选取其余未完成和已完成的任务发现大部分未完成的任务附近最近 10 个会员距任务的平均距离均明显较大，故我们分析部分任务无法完成可能由任务周围最近的会员距任务的距离过远导致。

但我们发现有部分未完成任务周围的人员密度较大，距任务最近会员的平均距离亦较小如下图 7 所示：

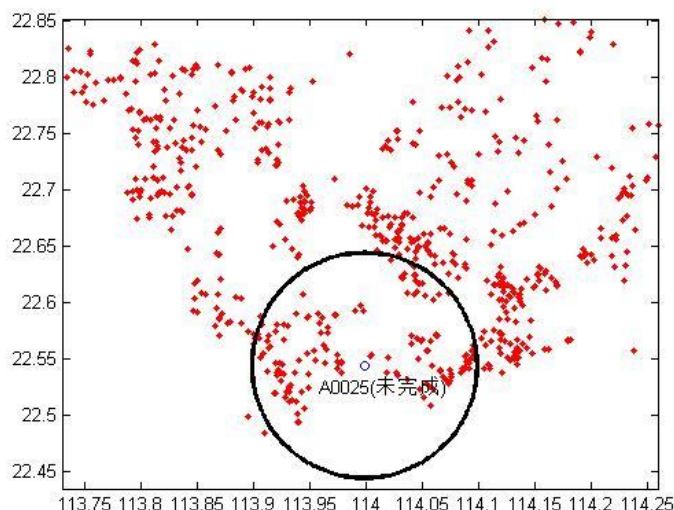


图 7 编号为 A0025（未完成）的任务周围会员分布情况图

对图 5 和图 7 进行对比，我们发现未完成任务 A0025 的周围的会员密度为 152 与已完成任务 A0153 相近，距离会员最近的 10 个会员的平均距离和 A0153 相差不大，但 A0153 的标价为 70.5 元处于中档标价，而 A0025 的标价为 67.5 元处于低档价位。故我们分析部分任务无法完成是由于任务的定价过低导致的。

五. 问题二的模型建立与求解

5.1 问题分析

在问题一中我们发现任务的定价与任务周围的会员密度、任务自身的复杂程度有关。部分任务未完成主要受任务位置与会员位置间距离、任务周围会员密度、任务的定价等因素的影响。基于此，为建立更加合理的任务定价方案及对定价方案进行评价，我们这样考虑：

针对建立合理的任务定价方案，基于问题一中对任务定价和任务未完成原因的分析，我们首先提出与定价相关几个变量：会员密度、会员距离程度，建立定价与变量间的函数关系，并采取穷举法对函数关系中的参数进行优化从而确定新的任务定价模型。

针对建立合理的定价方案评价指标，为了对原定价方案与我们新建立的定价模型进行评价，我们提出可从以下几个要求进行讨论：1）完成率提高；2）任务总定价降低，则 APP 平台所需要支出的经济成本越低，获利更高；3）会员的平均性价比变大，则会员的满意度越高，会员的接单积极性越高。显然，实际任务标价中任务的总定价的升降

必定会影响会员平均性价比的大小，故我们需要有必要协调这些要求，综合各指标对任务定价方案进行科学合理的评价。

5.2 模型建立

在问题一中，我们发现影响任务定价的主要因素为会员的密度，影响任务完成情况的主要受任务自身的复杂程度、会员密度、任务周围的会员距离程度等因素影响，同时考虑到对任务定价方案的合理评价应综合考虑劳务双方利益，故应从会员和 APP 平台两个角度进行合理评价，故我们建立该问的主要思想如下图所示：

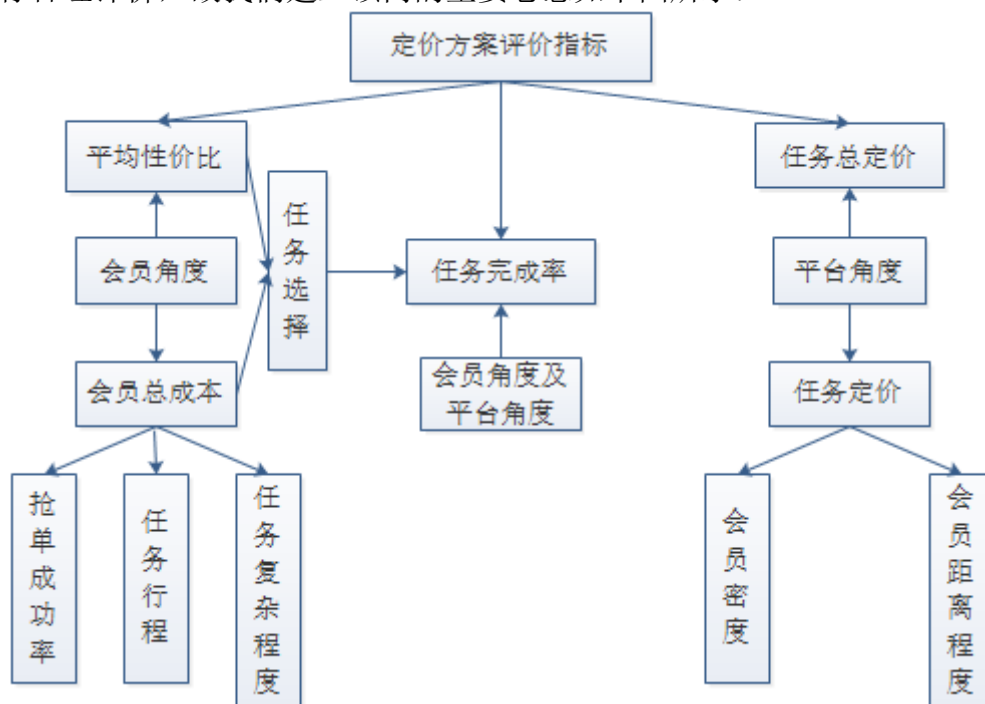


图 8 任务定价及方案评价主要影响因素结构图

5.2.1 影响定价的变量的确定

(1) 会员密度 ρ

会员密度即为单位区域内会员数量与区域面积的比值. 其中我们以所需定价的任务为圆心，以 r 为半径确定影响区域。

$$\rho = \frac{n_1}{s_1} = \frac{n_1}{\pi r^2}$$

其中， n_1 表示会员数量， s_1 表示区域面积. 会员密度越大，则说明单位区域内会员数量越多，接任务的人越多，因此任务的定价应当越低。

(2) 会员距离程度 L

以所需定价的任务为中心，分别求取与该中心最近的 10 个会员的距离，用 10 个距离的平均值来表示会员距离程度. 会员平均距离越大说明会员与该任务的靠近程度越大，会员所需要的成本也越低，会员接任务的积极性越高，任务定价应越低。

$$L = \frac{\sum_{k=1}^{10} l_k}{10}, \quad k = 1, 2, \dots, 10$$

(2)

其中, l_k 表示与任务第 k 个近的会员和任务之间的距离. 其中, 考虑到我们所获得为会员及任务的经纬度信息, 而我们所处的地球为球状体, 故我们利用球面坐标内的距离公式^[1]求解.

$$l_k = 2R \cdot \arcsin(\sqrt{\sin^2(\frac{\varphi_0 - \varphi_k}{2}) + \cos\varphi_i \cos\varphi_0 \sin^2(\frac{\lambda_0 - \lambda_k}{2})}) \quad (3)$$

其中, φ_k, φ_0 分别为第 k 个近的会员和任务的纬度, λ_k, λ_0 分别为第 k 个近的会员和任务的经度, R 表示地球半径, 故本文中取 $R = 6371km$.

5.2.2 任务定价模型的建立

在问题一的基础上, 我们已经发现任务定价与会员密度 ρ 成负相关性, 即会员密度越高, 任务定价越低. 通过对任务未完成原因进行分析, 发现会员距离程度 L 也与任务定价有关, 且会员距离程度越高, 任务定价越高. 因此我们建立新的任务定价模型:

$$\begin{cases} P = -k_1\rho + k_2L + C_1 \\ \rho = \frac{n_1}{\pi r^2} \\ L = \frac{\sum_{k=1}^{10} l_k}{10} \end{cases} \quad (4)$$

其中, C_1, k_1, k_2 为常系数, P 为任务定价, ρ 为会员密度, L 为会员距离程度, n_1 表示会员数量, r 表示区域面积, $l_k, k = 1, 2, \dots, 10$ 表示与任务第 k 个近的会员和任务之间的距离.

5.2.3 影响会员任务选择的因素的确定

由图 8 可知, 会员对任务选择将影响任务的完成率, 同时任务选择也会影响会员的获利, 故通过确定影响会员选择的因素, 我们可以更好的对定价模型进行调整, 使得会员和 APP 平台劳务双方达到双赢效果.

(1) 抢单成功率 S

在进行任务分配时实际上是根据预订限额所占比例进行配发, 我们定义抢单成功率为某参与某次抢单的会员的实际额数与抢夺该单任务的所有会员的总实际额数的比值. 抢单成功率越大说明该会员得到任务越容易, 会员所要花费的成本越低.

$$S = \frac{m_1}{m_{\text{总}}} \quad (5)$$

其中, m_1 表示实际额数; $m_{\text{总}}$ 表示抢夺该单任务的会员. 其中, 实际额数为今日限额与今日已抢到的任务数量的差.

$$m_1 = m_x - m_q \quad (6)$$

其中, m_x 表示该抢单会员的限定额数; m_q 表示今日已抢到任务数量. 已抢任务数

量越大，则实际额数就越小，抢单成功率较低。

(2) 任务行程 d

我们定义任务行程即为以会员为起点，会员与其准备接单的任务之间的距离。任务行程越大，说明会员到任务之间的距离越大，会员所需花费的成本越高。

$$d(u_i, v_j) = 2R \cdot \arcsin(\sqrt{\sin^2(\frac{\varphi_j - \varphi_i}{2}) + \cos\varphi_i \cos\varphi_j \sin^2(\frac{\lambda_j - \lambda_i}{2})}) \quad (7)$$

其中， φ_i, φ_j 是会员点 u_i 和任务点 v_j 的纬度， λ_i, λ_j 是会员点 u_i 和任务点 v_j 的经度，本文中 R 取地球半径为 $R = 6371km$ 。

(3) 任务复杂程度 D

任务复杂程度，是任务难易度的反映。不同的任务难度会影响会员的选择，原则上，任务难度越大会员选择的概率越小，任务难度越小会员的接任务积极性越高。任务因此我们引入任务复杂程度，记作 D 。

5.2.4 会员任务选择的判定指标的确定

对于会员而言，任务价格并不是选择任务的唯一评价标准。对于一些偏远地区或者任务难度较大的任务，即使任务定价很高，大部分会员还是不会选择这些任务，为了建立一个统一标准，需要综合考虑会员的获利和成本情况，因此我们先定义了总成本 T_{ij} 和性价比 Q_{ij} 。

(1) 会员总成本 T_{ij}

会员从收到任务消息后，在抢单过程中、去往任务位置的过程中、在任务点完成任务等过程中均需要消耗会员自身的成本，故我们综合考虑抢单成功率 S 、任务行程 d 、任务复杂程度 D 这3个变量确定会员的总成本。会员所需的总成本越低，其获利越少，则会员对任务的积极性越低、对任务选择的可能性越小。

$$T_{ij} = k_3 d - k_4 S + k_5 D + C_2 \quad (8)$$

$$\begin{cases} d(u_i, v_j) = 2R \cdot \arcsin(\sqrt{\sin^2(\frac{\varphi_j - \varphi_i}{2}) + \cos\varphi_i \cos\varphi_j \sin^2(\frac{\lambda_j - \lambda_i}{2})}) \\ S = \frac{m_1}{m_{\text{总}}} \\ m_1 = m_x - m_q \end{cases}$$

其中， k_3, k_4, k_5, C_2 均为常数， T_{ij} 为会员 i 选择任务 j 花费的总成本， d 为行走路程， S 为抢单成功率， D 任务复杂程度， φ_i, φ_j 是会员点 u_i 和任务点 v_j 的纬度， λ_i, λ_j 是会员点 u_i 和

任务点 v_j 的经度， m_1 表示实际额数， $m_{\text{总}}$ 表示抢夺该单任务的会员总实际额数， m_x 表示限定额数， m_q 表示今日已抢到任务数量。

(2) 任务性价比 Q_{ij}

我们定义任务性价比 Q_{ij} 为目标任务的定价与会员完成该任务的总成本的比值：

$$Q_{ij} = \frac{P_j}{T_{ij}} \quad (9)$$

其中， Q_{ij} 为会员 i 选择任务 j 的性价比； P_j 为任务 j 定价（会员获利）； T_{ij} 为会员 i 选择任务 j 花费的总成本。

一般而言，会员都会选择性价比较高的任务，基于此建立基于性价比的任务选择的0-1函数：

$$f(Q_{ij}) = \begin{cases} 1, & Q_{ij} \geq \delta \\ 0, & Q_{ij} < \delta \end{cases} \quad (10)$$

其中， δ 表示阈值，1表示会员 i 会选择任务 j ，0则代表不会选择该任务。

综上，我们建立会员的任务选择判定函数为：

$$\begin{cases} f(Q_{ij}) = \begin{cases} 1, & Q_{ij} \geq \delta \\ 0, & Q_{ij} < \delta \end{cases} \\ Q_{ij} = \frac{P_j}{T_{ij}} \\ T_{ij} = k_3d - k_4S + k_5D + C_2 \end{cases} \quad (11)$$

5.2.5 任务定价方案评价指标的确立

对于任务定价方案的优劣，我们主要从APP平台、会员两方面提出以下3个角度进行考虑：1）完成率大小；2）任务总定价高低，任务的总定价越低则APP平台所需要支出的经济成本越低，获利更高；3）会员的平均性价比的高低，平均性价比越高则会员的满意度越高，会员的接单积极性越高。我们需要综合各指标对任务定价方案进行合理的评价。

(1) 任务完成率 α

任务的完成率是指某地区内任务完成量与总任务数量的比值。任务的完成率越大反映任务完成情况越好，APP平台和会员的获利均增大，从而反映定价越合理。任务完成率 α 计算公式：

$$\alpha = \frac{x_1}{x_{\text{总}}} \quad (12)$$

其中, x_1 表示任务完成量; $x_{\text{总}}$ 表示总任务数量.

(2) 任务总定价 β

任务总定价为某地区内所有任务定价的总和. 任务总定价 β 计算公式:

$$\beta = \sum_j^n P_j \quad (13)$$

其中, P_j 指第 j 个任务的定价. 任务总定价越大反映了该 APP 平台发布任务的经济成本越大.

(3) 平均性价比 γ

平均性价比为某区域内所有会员所做任务性价比的平均值. 会员的平均性价比反映了会员的满意度情况, 平均性价比越高, 说明会员满意度越大.

$$\gamma = \frac{1}{m} \sum_{j=1}^n \sum_{i=1}^m f(Q_{ij}) \cdot Q_{ij} \quad (14)$$

其中, m 为会员总数, n 为任务总数.

5.3 模型求解

我们基于穷举法^[4]对任务定价模型和会员总成本函数中参数优化求解. 算法设计如下:

Step1: 分别给定参数 $k_1, k_2, C_1, k_3, k_4, k_5, C_2$ 的初始区间 $[a_i, b_i], i = 1, 2, \dots, 6$, 并确定步长 k , 分别依步长取点;

Step2: 根据所给参数, 计算任务完成率;

Step3: 取所有元素中任务完成率最高的点, 以其为中心, $4k$ 长度为边长, 建立新的初始区间;

Step4: 仍以 k 步长, 依次取点, 计算任务总定价;

Step5: 取所有元素中任务总定价最低的点, 以其为中心, $2k$ 长度为边长, 建立新的初始区间;

Step6: 仍以 k 步长, 依次取点, 计算平均性价比;

Step7: 取所有元素中平均性价比最高的点, 即为较优参数.

利用上述算法, 我们取定初始步长 $k = 0.01$, 利用 MATLAB 软件^[3]计算可得结果如下:

(1) 任务定价模型为:

$$\begin{cases} P = -3.06\rho + 0.04L + 66.32 \\ \rho = \frac{n_1}{\pi r^2} \\ L = \frac{\sum_{k=1}^{10} l_k}{10} \end{cases} \quad (15)$$

其中, P 为任务定价, ρ 为会员密度, L 为会员距离程度, n_1 表示会员数量, r 表示区域面积, $l_k, k = 1, 2, \dots, 10$ 表示与任务第 k 个近的会员和任务之间的距离.

(2) 会员总成本为:

$$T_{ij} = 0.82d - 0.13S + 1.05D + 5.47 \quad (16)$$

其中, T_{ij} 为会员 i 选择任务 j 花费的总成本, d 为行走路程, S 为抢单成功率, D 任务复杂程度.

(3) 会员的任务选择判定函数为:

$$\begin{cases} f(Q_{ij}) = \begin{cases} 1, Q_{ij} \geq \delta \\ 0, Q_{ij} < \delta \end{cases} \\ Q_{ij} = \frac{P_j}{T_{ij}} \\ T_{ij} = 0.82d - 0.13S + 1.05D + 5.47 \end{cases} \quad (17)$$

其中, Q_{ij} 为会员 i 选择任务 j 的性价比; P_j 为任务 j 定价(会员获利); T_{ij} 为会员 i 选择任务 j 花费的总成本.

5.4 模型应用

利用我们新的任务定价模型我们对附件一中的任务重新进行定价所得结果如下:

表 3 任务的新定价情况表

任务号码	会员距离程度	会员密度	新方案定价
A0001	109.60	0.40	74.4
A0002	35.52	0.33	66.7
A0003	69.52	0.37	67.9
A0004	349.31	0.08	79.8
A0005	80.41	0.30	68.6
...
A0407	55.08	0.07	68.3
A0408	13.94	0.45	65.5
A0409	249.24	0.04	76.0
A0411	151.39	0.12	71.9

...
A0831	41.36	0.23	67.2
A0832	145.36	0.06	71.9
A0833	88.01	0.30	75.9
A0834	52.19	0.20	67.8
A0835	123.52	0.18	74.6

我们取性价比的阈值为 $\delta = 1.6$.利用式 (16)、(17)，我们计算可得到每一位会员完成每项任务的总成本，计算相应的性价比. 通过对新方案下的每项任务的执行情况进行计算，得到如图 9 所示的任务完成情况图：

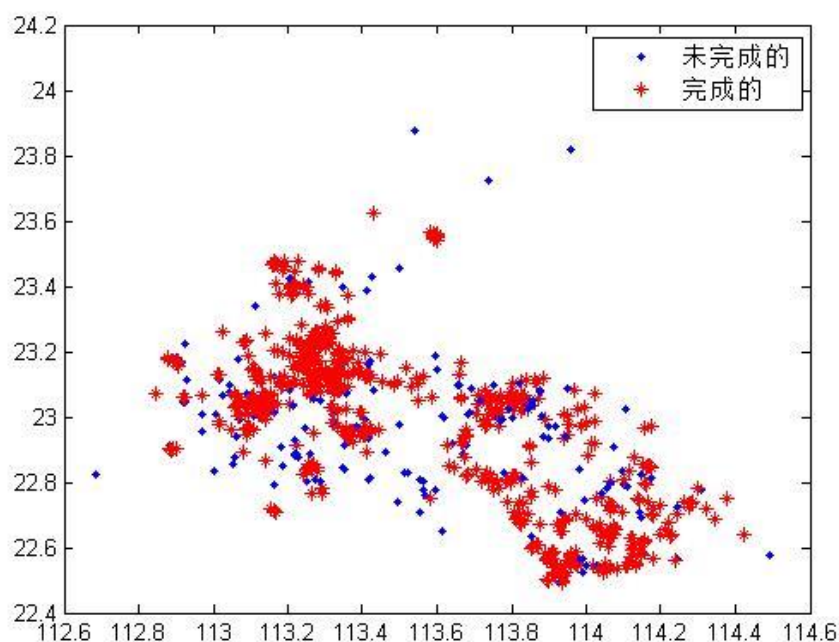


图 9 新定价方案下的任务完成情况图

对比图 4 与图 9，我们不难发现：新定价方案下任务的完成量明显高于原定价方案下的任务完成量，且完成任务的分布更加均匀。

接着，我们原方案与新方案下的任务定价与任务执行情况进行了统计分析，所得结果如下表 4 所示：

表 4 原方案与新方案的定价、执行情况对比表

任务号码	原方案任务 定价	新方案任务 定价	原方案下任务执 行情况	新方案下任务执 行情况
A0001	66	74.4	0	0
A0002	65.5	68.7	0	1
A0003	65.5	67.9	1	1
A0004	75	79.8	0	0
A0005	65.5	68.6	0	1
...
A0409	75	76.0	1	0
A0410	73.5	78.7	0	0
A0411	75	71.9	1	1
...
A0831	65.5	67.2	0	1
A0832	72	71.9	1	1
A0833	85	75.9	1	0
A0834	65.5	67.8	1	1
A0835	85	74.6	1	1

由表 4 我们发现：新定价方案下，未完成部分的任务标价相较原任务定价普遍有所增加；原定价高且完成的任务的标价普遍有所减少，而原完成且定价不高的任务的标价增减幅度不大。

进一步，我们利用式（12）、（13）、（14）中的任务定价评价指标对原定价方案和新定价方案更加精确的比较如下：

表 5 原方案与新方案的定价评价指标情况表

	完成率	任务总定价	平均性价比
原方案	62.51%	57707.5	/
新方案	80.23%	59189.9	2.18

从表 5 中可以分析得出：

- 1) 新定价方案下的任务的完成率远高于原方案，有更多的任务被会员们完成，这说明新方案下 APP 平台和会员的获利都将增加；
- 2) 新方案中任务总定价与原方案相近，公司的总经济成本有所变化不大；
- 3) 新方案的平均性价比较高，说明新定价方案下会员们满意度较高。

总的来说，新的任务定价方案可使 APP 平台和会员双方的获利均增大，同时任务的完成率增大，更好的利用网络平台的信息资源；会员们的满意度较高，说明该新的定价方案较合理科学，同时现实应用价值较高。

六. 问题三的模型建立与求解

6.1 问题分析

考虑到实际生活中会有多个任务所在位置较为集中,为了防止因会员抢单导致部分任务无人接单,同时可以用多个任务的打包发布减少会员完成任务时的总成本,从而提高会员的满意度;增加任务的完成率,降低任务的总定价,使 APP 平台获益更多. 我们可以将每个任务的位置看作节点,以两节点间的距离作为边建立无赋权网络图,先将任务进行合理打包后看作一个新的任务,再利用问题二中所建立的任务定价模型进行定价,建立基于图论知识的任务打包优化定价模型.

6.2 模型建立

考虑到现实中,会有多个任务的位置较为集中,若单个发布资源利用率过低且消耗成本过大,APP 平台和会员的获利降低等,我们提出在问题二所建立的定价模型的基础上,先利用图论知识对任务进行打包成一个新任务,在利用已有定价模型进行定价,从而对问题二中提出的定价模型进行优化.

(1) 基于任务位置的网络图^[2]的建立

我们将任务的经纬度信息反映到图像上,以任务位置看作节点,假设任意节点间均有边相连(即任意两任务位置间均有路),以两节点间的距离作为边的权值,则可将任务位置关系抽象成一无向赋权网络图 G , 记为:

$$G = (V, E, D^*) \quad (18)$$

其中, $V = \{v_1, v_2, \dots, v_{835}\}$ 为任务位置的节点集, $E = \{e_1, e_2, \dots, e_{C_{835}^2}\}$ 为连接 G 各点间的边集. 再记边权集为:

$$D^* = \{d(e_i) | i = 1, 2, \dots, 140\} \quad (19)$$

其中, $d(e_i)$ 表示边 e_i 的权值. 而对于图中任意两点 v_i 与 v_j , 用 $d(v_i, v_j)$ 代表两点间的球面距离, 具体计算公式如下:

$$d(v_i, v_j) = 2R \cdot \arcsin\left(\sqrt{\sin^2\left(\frac{\varphi_j - \varphi_i}{2}\right) + \cos\varphi_i \cos\varphi_j \sin^2\left(\frac{\lambda_j - \lambda_i}{2}\right)}\right)$$

其中, φ_i, φ_j 是节点 v_i 和节点 v_j 的纬度, λ_i, λ_j 是节点 v_i 和节点 v_j 的经度, R 表示地球半径, 在此取 $R = 6371km$.

(2) 特殊节点间距离改进

考虑到任务的完成情况不同, 为了尽量使可能完成的节点与无法完成的节点打包在一起, 我们对这类特殊节点间距离做如下处理:

$$d(v_i, v_k) = 1.6 R \cdot \arcsin(\sqrt{\sin^2(\frac{\varphi_j - \varphi_i}{2}) + \cos\varphi_i \cos\varphi_j \sin^2(\frac{\lambda_j - \lambda_i}{2})}) \quad (20)$$

其中, φ_i, φ_j 是预计可完成任务的节点 v_i 和预计无法完成任务的节点 v_j 的纬度, λ_i, λ_j 是预计可完成任务的节点 v_i 和预计无法完成任务的节点 v_j 的经度. 取定 R 为地球半径, $R = 6371km$.

通过特殊节点间距离的处理后, 预计无法完成任务的节点与可能完成任务的节点间的理论距离减小, 按距离靠近的打包原则, 将大大增加了预计无法完成任务和预计可完成任务打包分配的可能性.

基于距离改进后的各任务节点的邻接矩阵部分结果如下:

$$D^* = \begin{pmatrix} 0 & 31.6 & 58.4 & & 2680.5 & 2896.3 & 3049.4 \\ 31.6 & 0 & 47.1 & \cdots & 4387.3 & 3144.2 & 3573.9 \\ 58.4 & 47.1 & 0 & & 3163.7 & 3867.5 & 5837.4 \\ & \vdots & & \ddots & & \vdots & \\ 2680.5 & 4387.3 & 3163.7 & & 0 & 52.6 & 269.3 \\ 2896.3 & 3144.2 & 3867.5 & \cdots & 52.6 & 0 & 89.6 \\ 3049.4 & 3573.9 & 5837.4 & & 269.3 & 89.6 & 0 \end{pmatrix} \quad (21)$$

(3) 任务打包分配的原则确立

为了合理对任务进行打包, 我们以最短距离为原则, 同时考虑可打包成一起的任务间最大距离, 将 835 个任务进行打包分配, 即将问题转化为对顶点集 $V = \{v_1, v_2, \dots, v_{835}\}$ 的划分问题, 使得各子集满足以下要求:

1) 所有子集的并集为全集, 任意两个子集的交集为空集, 即:

$$v_i \in V_i, \bigcup_{i=1}^n V_i = V \text{ 且 } \forall i, j \text{ 有 } V_i \cap V_j = \emptyset \quad (23)$$

其中, n 为子集最终划分个数.

2) 同一打包任务内距离约束, 即:

$$\forall v_i, v_j \in V_i, \max d(v_i, v_j) \leq \tau \quad (24)$$

其中, τ 为阈值限制.

3) 同一打包任务内个数约束, 即:

$$|V_i| \leq 5 \quad (25)$$

其中, $|V_i|$ 为集合 V_i 内所含元素个数.

4) 以距离较短为原则进行打包, 在满足条件 2)、3) 下, 将节点分配到距离最近的集合. 即:

$$\forall v_i \in V_k, l_{ik} \leq l_{ij}$$

$$l_{ij} = \max_{v_k \in V_j, i \neq k} \text{mind}(v_i, v_k), \forall v_k \in V_j \text{ 且 } i \neq k$$
(26)

其中, l_{ij} 为节点 v_i 到集合 V_j 的距离; V_j 为加入 v_i 后不违背条件 1)、2) 中的集合.

6.3 模型求解

基于任务打包分配的原则, 我们利用如下算法对问题进行求解:

Step1: 任取一点 v_k , 计算 v_k 到其他集合的距离 d (单独一点看作单点集), 记最大值为 d_{max} , $i = 1$;

Step2: 取所有距离中第 i 小的集合, 记为 V_j , 距离记为 d_i ;

Step3: 判断该集合 V_j 中元素个数 $|V_j|$ 是否小于等于 4; 若是, 则进入下一步; 反之, 直接进入 Step6;

Step4: 计算 v_i 到 V_j 中所有点的距离, 记其中的最大值为 l_{max} ;

Step5: 判断 l_{max} 是否小于阈值, 若是, 则将 v_i 归于 V_j 中, 返回 Step1; 否则, 进入下一步;

Step6: 判断 l_i 是否小于 d_{max} , 若是, 则取 $i = i + 1$, 重复 Step2; 反之, v_i 单独成为一个点集, 返回 Step1.

该算法的具体流程如下图所示:

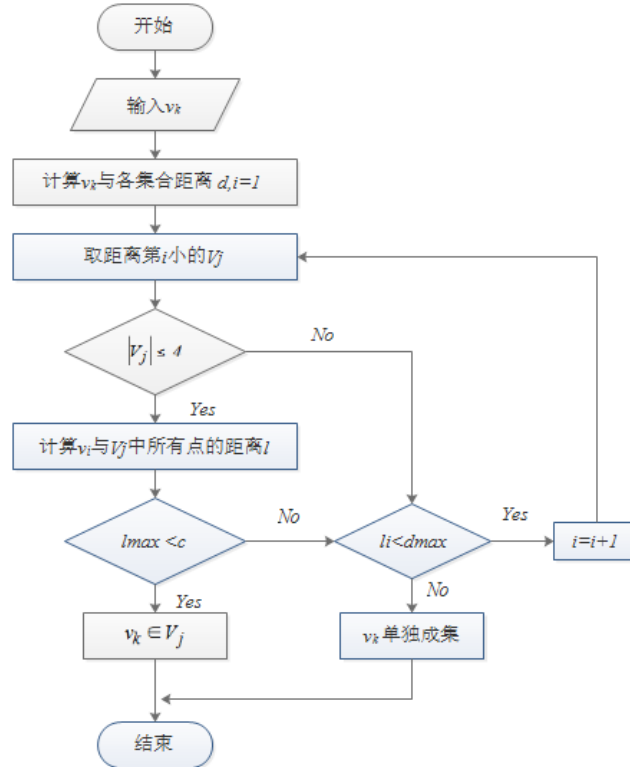


图 10 基于任务打包原则的模型求解算法示意图

为更加清晰的分析任务打包的定价的效果，我们选取了如图 11 所示的任务分布情况为例，以各任务为节点，任务间的距离给各节点间的边赋权，建立无向赋权图，根据特殊节点间距离改进式(20)，得到图中各点之间的距离. 再利用上述算法，我们用 MATLAB 软件编程可将图中的点进行划分最终得到如图 11 所示的打包结果：

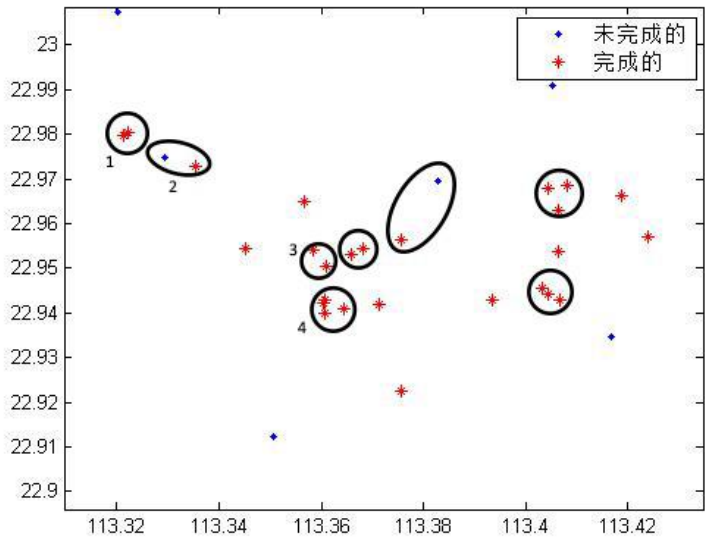


图 11 基于图论知识的部分任务打包示意图

由图 11 我们可知，我们将图中原有的任务重新分配成了 8 类打包任务和部分未打包任务. 其中，以第 1、2、3、4 类打包为例，任务的位置情况如下表 6 所示：

表 6 图 2 中 4 类打包任务位置情况表

类别	所包含任务的位置
1	(113. 3222, 22. 9803), (113. 3212, 22. 9798)
2	(113. 3294, 22. 9748) (未完成), (113. 3354, 22. 9728)
3	(113. 3584, 22. 9541), (113. 3610, 22. 9506)
4	(113. 3644, 22. 9410), (113. 3607, 22. 9429), (113. 3607, 22. 9400), (113. 3604, 22. 9423)

我们对打包发布的任务运用改进过的定价模型进行求解，得到第 1、2、3、4 类打包发布的任务定价情况如下：

表 7 图 11 中 4 类打包任务的定价情况表

类别	所包含任务的原定价	会员距离程度	会员密度	打包后的新定价
1	66, 66. 5	86. 73	0. 33	68. 7
2	75, 66. 5	86. 03	0. 35	69. 6
3	65. 5, 65. 5	58. 16	0. 29	67. 7
4	65, 65, 65, 65,	58. 71	0. 39	68. 4

由表 7 可知：打包后的任务定价远远小于该类任务的原定价和。

类似上述任务的打包方式，我们将所有任务进行相应的打包处理，得到打包后的新方案，将该方案的定价评价指标与之前的两种方案进行对比，得到如下表所示的定价评价情况表：

表 8 三种定价方案的定价评价指标情况表

	完成率	任务总定价	平均性价比
原方案	62.51%	57707.5	
新方案	80.23%	59189.9	2.18
打包后的新方案	93.37%	32457.9	2.57

由表 8 可以发现：

1) 打包后的新方案的完成率较前两种定价方案有显著的提高，说明对任务的打包处理可以使较多原来未完成任务，可以随理论上执行完成的任务一起得以完成；

2) 打包后的任务总定价比比前两种方案都低，说明对任务的进行打包处理也可以降低任务的定价，一方面可降低公司的经济成本，另一方面可提高公司的利润收入；同时，平均性价比也有明显的提高，说明会员的满意度有了较好的提高。

综上，我们可以发现基于图论知识的任务打包定价优化模型的实际价值较高，能给运行平台和会员均带来较高的效益和满意度，值得推广运用于实际运营中。

七. 问题四的分析与求解

7.1 问题分析

为对我们建立的基于图论知识的任务打包定价优化模型进行实际生活中的应用价值的探究，对模型的实用性和合理性进行检验，我们利用附件三中所给的任务经纬度信息，利用我们建立的基于图论知识的任务打包定价模型进行定价，分析任务的定价效果，预测任务的执行情况和定价方案的效果。

7.2 问题求解

根据附件三所给任务经纬度信息，我们通过地理位置的查找发现附件三中的大部分经纬度位置仍然在广东省的广州市与深圳市，为了探究问题三中所建模型在该地区是否具有普遍适用性，我们假定该 APP 平台的会员信息不变，将问题三中所建立的基于图论知识的任务打包定价优化模型应用于附近三的新任务项目，得到以下结果，其中前 15 类打包为例，其定价等相关情况如下：

表 9 新项目的任务标价与执行情况表

打包编号	任务号码	会员距离程度	会员密度	任务定价	执行情况
1	C0001、C0070、C0072	125.39	0.34	71.2	1
2	C0002、C0822、C1984	64.53	0.41	67.6	1
3	C0003、C0782	47.80	0.44	66.8	1
4	C0004、C1004、C1240	55.29	0.37	68.2	1
5	C0005、C0348、C1490	78.75	0.70	69.0	1
6	C0006	225.85	0.06	75.0	0
7	C0007、C1897	63.36	0.41	69.5	1
8	C0008、C1032	72.64	0.24	70.2	0
9	C0009、C0090、C1930	143.43	0.30	72.5	1
10	C0010、C0131、C1214、 C1215	70.84	0.54	68.7	1
11	C0011	283.91	0.06	77.3	0
12	C0012、C0092	90.68	0.22	69.4	1
13	C0013、C0132、C0201、 C0239	57.63	0.72	67.8	1
14	C0014	94.03	0.13	69.6	1
15	C0015、C1841	55.54	0.40	68.3	1

同时，该项目的定价评价指标如下表 10 所示：

表 10 新项目定价评价指标情况表

	完成率	任务总定价	平均性价比
新项目方案	92.35%	56062.3	2.82

根据任务的打包情况以及定价评价指标的情况，我们发现问题三中所建立的基于图论知识的任务打包定价优化模型的实际适用性较好且较为科学合理，应用于附件三的新项目后，该定价方案的完成率高达 92.35%，近 1900 个任务将被会员们完成，任务总定价适中，平均性价比较高说明会员的满意度较高，值得在现实运营中进行推广。

八. 模型评价

8.1 模型优点

(1) 问题二中，我们不仅站在公司角度考虑了会员密度和会员距离程度，进一步从会员角度定义了性价比，最后从完成率、任务总定价和平均性价比三个角度来评价定价方案，使得模型较为完整；

(2) 问题三的任务打包中，我们对预计可完成任务和预计无法完成任务间的距离进行特殊处理，大大增加了将可完成任务和无法完成任务的捆绑销售的概率，最后打包效果较好；

(3) 问题三中, 我们建立图论模型, 将打包问题转化为子集划分问题, 并考虑实际情况加以条件约束, 在实际应用中有广阔的前景;

(4) 问题四中, 我们用该定价模型对其他地区进行应用, 得到效果较好, 说明我们的定价模型实施性较强, 符合普遍情况.

8.2 模型缺点

(1) 由于数据不足, 我们在定价时没有考虑任务本身属性, 如任务难度, 任务周期等.

(2) 利用穷举法对第二问中系数进行优化时, 由于系数较多, 计算量较大.

九. 参考文献

[1] 石立兴, 胡访宇. 二维空间上基于呼叫详细记录数据的用户移动性分析[J]. 计算机应用, 2015, 35(09): 2453-2456+2469.

[2] 卜月华, 王维凡, 吕新忠. 图论及其应用(第2版). 南京: 东南大学出版社, 2015, 5.

[3] 姜启源, 谢金星, 叶俊. 数学模型(第四版). 北京: 高等教育出版社, 2011, 1.

[4] 刘卫国主编. MATLAB 程序设计教程. 北京: 中国水利水电出版社, 2005.

附录：Matlab 程序：

问题二中：

(1)

% 到任务点最近的 10 个距离的平均值

d=[];

for i=1:835

 k=0;

 for j=1:1877

 if (wupin(i,1)-huiyuan(j,1))^2+(wupin(i,2)-huiyuan(j,2))^2<0.01

 k=k+1;

 end

 end

 d=[d;k];

end

x2=[];

for i=1:835

 x=[];

 for j=1:1877

x1=6371*acos(cos(wupin(i,2))*cos(huiyuan(j,2))*cos(wupin(i,1)-huiyuan(j,1))+sin(wupin(i,2))
*sin(huiyuan(j,2)));

 x=[x x1];

 end

 xx=sort(x,'ascend');

 s=0;

 for k=1:10

 s=s+xx(k)/10;

 end

 x2=[x2;s];

end

(2)

%判断任务是否完成

X_J_B1=[];F_1=[];

for i=1:835

 dingjia=C1-K1*wupin(i,7)+K2*wupin(i,6)+0.5*wupin(i,8);

 for j=1:1877

Y_H_lucheng=6371*acos(cos(wupin(i,2))*cos(huiyuan(j,2))*cos(wupin(i,1)-huiyuan(j,1))+sin(
wupin(i,2))*sin(huiyuan(j,2)));

 chengben=K3*Y_H_lucheng-K4*huiyuan(j,3)+K5*wupin(i,8)+C2;

 X_J_B=dingjia/chengben;

```

        if X_J_B<1.6
            F_1(i,j)=0;
        else F_1(i,j)=1;
        end
    end
end
end

```

```

k1=[];
for i=1:835
    if sum(F_1(i,:))==0
        k1=[k1 i];
    end
end
end

```

(3)

% 65,75,85 以及会员的分布图

```

for i=1:835
    if wupin(i,3)==85
        plot(wupin(i,1),wupin(i,2),'marker','^','MarkerFacecolor','r','MarkerSize',6);hold on;
    end
    if wupin(i,3)==75
        plot(wupin(i,1),wupin(i,2),'marker','p','MarkerFacecolor','r','MarkerSize',8);hold on;
    end
    if wupin(i,3)==65
        plot(wupin(i,1),wupin(i,2),'marker','s','MarkerFacecolor','r','MarkerSize',6);hold on;
    end
end
end
plot(huiyuan(:,1),huiyuan(:,2),'b.','MarkerSize',4.5);
h1=plot(wupin(6,1),wupin(6,2),'marker','^','MarkerFacecolor','r','MarkerSize',6);
h2=plot(wupin(6,1),wupin(6,2),'marker','s','MarkerFacecolor','r','MarkerSize',5);
h3=plot(wupin(450,1),wupin(450,2),'marker','p','MarkerFacecolor','r','MarkerSize',8);
h4=plot(huiyuan(1,1),huiyuan(1,2),'b.','MarkerSize',4.5);
legend([h1 h2 h3 h4],'85','75','65','会员')

```