

# Analysis and Research on the Task Pricing of “Making Photos and Making Money”

Leyuan Zhao<sup>1</sup>, Bo Zhang<sup>1</sup>, Juntian Bo<sup>1</sup> and Wenfei Zhao<sup>2,\*</sup>

<sup>1</sup>Students' Brigade, Naval Aviation University, Yantai Shandong, 264001, China

<sup>2</sup>School of Basic Sciences for Aviation, Naval Aviation University, Yantai Shandong, 264001, China

\*Corresponding author

**Abstract**—We research the operation of an APP, which collects the data of related products through the members' photographs. We organize the information collected and make a regression model. We use MATLAB to make a scatter plot to understand the task in the APP in longitude - latitude - price distribution in three-dimensional coordinates, and the relationship between task pricing and membership information. The task pricing is also related to the task location. The study find that the original program has many unreasonable places. We analyze the main influencing factors and formulated the task package model, and then re-determine the task pricing scheme based on the order of magnitude. The model and scheme are properly amended by using the member's position and the position of the task. The revised pricing plan is more reasonable and more practical.

**Keywords**—linear regression; package distribution; sub-regional model

## I. INTRODUCTION

"Photo Earning Money" is a new self-service model that runs on the mobile Internet. It allows users to download APPs, register as a member, and then take a photo-taking task (such as going to the supermarket to check the availability of a product). Members earn APP for the task of the nominal reward. APP has become the core of the platform to run, and APP is the core element of the task pricing. If the pricing is not reasonable, some tasks will no one cares that lead to the failure of commodity inspection. Based on the collected data, we analyze its pricing rules and study the impact of various major factors on pricing through scatter plot. We find that the tasks in this APP can be packaged together in two-dimensional coordinates, and to re-establish a more reasonable pricing. In this study, we perform parametric analysis and testing using a binary regression equation.

## II. PACKAGING PRICING BASED ON REGRESSION EQUATION

Based on the collected data, we perform the least-squares regression analysis of the task position and the task price, and find out the regression equation. Then, we test the linearity of the regression model and the significance of regression coefficients. If the relationship between the two is not a simple linear relationship, you can change the relationship according to the scatter plot, substituting the data of the relevant data and other factors, and check whether the accuracy is improved. Then according to the improvement of the law of setting the task price to be amended and compared with the original plan.

Set up a model of task force  $W$  and task position  $(x,y)$  linear regression analysis, and we get  $n$  independent observation data for [1].  $E_n$  is the  $n$  unit matrix.

$$\begin{cases} w_i = \beta_0 + \beta_1 x_i + \beta_2 y_i + \varepsilon_i, \\ \varepsilon_i \sim N(0, \sigma^2), i = 1, \dots, n, \end{cases} \quad (1)$$

$$\begin{cases} w = X\beta + \varepsilon, \\ \varepsilon \sim N(0, \sigma^2 E_n), \end{cases} \quad (2)$$

### A. Principle of Order of Magnitude

At a certain point, the distance between all other points is calculated. We select the shortest distance from the distance, and select points that are at the same order of magnitude as the other points. We consider these points to be the closest point to the center point. We find a way to package the tasks. Pack two tasks. Here we use the order of magnitude of the most recent principle to set up a collection of points that can be packaged. We coordinate all task position coordinates on the plane of longitude and latitude, and set each point as  $P_i (i = 1, 2, \dots, n)$ . Take the  $P_1$  as an example, calculate the distance  $D_i (i = 2, 3, \dots, n)$  between all other points  $P_i (i = 2, 3, \dots, n)$  and  $P_1$ .

$$D_i = |P_i - P_1| (i = 2, 3, \dots, n)$$

$$\xi_i = \frac{D_i}{D_j} (i = 2, 3, \dots, n)$$

The shortest distance between  $P_j$  and  $P_1$  distances is  $D_j$ . The order of magnitude of  $\xi$  is calculated as  $\xi_i (i = 2, 3, \dots, n)$ . All the points  $\xi_i < 10$  are selected, and some of these points is the distance from  $P_1$  recent points. If you

find  $P_1$  to be in the same algorithm of its recent focus point nearest points, these points can be packaged with  $P_1$  together.

B. Formulation of Packaging Methods and Pricing Models

The more the number of tasks packaged, the greater you deduct the remuneration. Here the problem translates to the setting of the suppression factor  $\varepsilon(n)$ . It can be seen that the coefficient of inhibition varies with the number of tasks in the package, and the problem is similar to the issue of payroll tax. The commonly used method is "sub-deduction", that is, different salary sections to pay the tax ratio is different [2].

III. APPLICATION

The number of task breakpoints  $n_1, n_2, n_3$  should be adapted to the nearest number of points at each point under the recent order of magnitude. The distribution of points obeys normal distribution  $n \sim N(\mu, \delta^2)$ , in which  $\mu=6.3, \delta=0.732$ . Therefore, we set the maximum number of tasks in the package as 10, the number of tasks breakpoint  $n_1=1, n_2=3, n_3=6, n_4=9$ , the size of the increase in proportion of  $\varepsilon_i(n) (i=1, 2, 3, 4)$  is to:

$$\varepsilon_i(n) = \begin{cases} 0, & i = 1, \\ 2.5, & i = 2, \\ 5, & i = 3, \\ 7.5, & i = 4. \end{cases}$$

By researching task price  $W$ , we let  $t$  be the gratuity growth value based on the membership credit value, and use  $t = \beta_0 + \beta_1 z + \gamma$ , obtained by MATLAB  $\hat{\beta}_0 = 3.4411, \hat{\beta}_1 = 0.1850$ , the correlation coefficient  $r = 0.0062$ . So the final packaging task to provide remuneration for

$$W = \begin{cases} y_1 + \dots + y_n + t, & n = 1, \\ y_1 + \dots + y_n - (n-1) \times 2.5 + t, & 1 < n \leq 3, \\ y_1 + \dots + y_n - 5 - (n-3) \times 5 + t, & 3 < n \leq 6, \\ y_1 + \dots + y_n - 20 - (n-6) \times 7.5 + t, & 6 < n < 10. \end{cases} \quad (3)$$

A. Sub-regional Research

We merge or remove a few tasks that are independent from dense areas and far apart from each other, and are not suitable

for packaging. In the range of longitude  $112.8^\circ \sim 114.4^\circ$  and latitude  $22.4^\circ \sim 23.8^\circ$ , it divides 20 areas with concentrated task positions and calculates the number of tasks in each area. In the same way, the member distribution is divided as shown in Figure 1 and Figure 2:

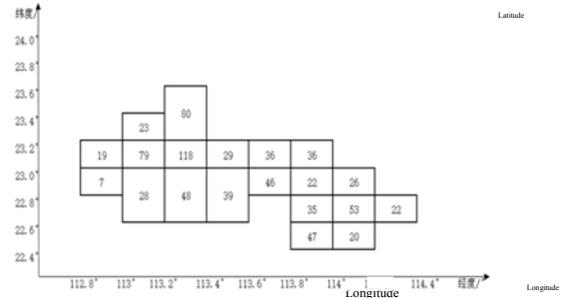


FIGURE I. THE NUMBER OF TASKS

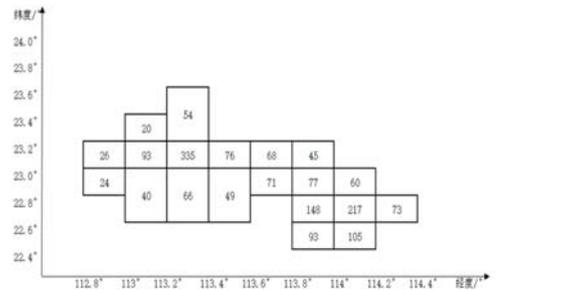


FIGURE II. THE NUMBER OF MEMBERS

The intensity of the task point and the degree of aggregation of members is basically consistent. The introduction of variables  $\alpha$ , said the number of a region can be packaged, the number of tasks set in a region is  $u$ , the number of members is  $m$ , then the number of a package can be calculated as

$$\alpha = \begin{cases} \frac{u \times m}{2 \times (u + m)}, & 500 < u, \\ \frac{u \times m}{3 \times (u + m)}, & 500 > u > 100, \\ \frac{u \times m}{4 \times (u + m)}, & u < 100. \end{cases} \quad (4)$$

$\alpha$  is rounded to the number of packages for the end of a regional plan for. [3] Calculated by MATLAB,  $\hat{\beta}_0 = -19.437497, \hat{\beta}_1 = 0.278544, \hat{\beta}_2 = 2.476796$ . However, the  $R^2$  is too small, so the two are not simple linear relationships.

Therefore, we observe the relationship between the task price and the distribution of mission locations through a scatter graph, as shown in Figure 3.

The picture shows that there is a different price stage for the same location, and it is calculated that the two are not simply linear. From this we can conclude that when a business sets a price for a task, and it does not mainly consider the distribution of locations. Obviously, the task position is to affect the task pricing. [4]

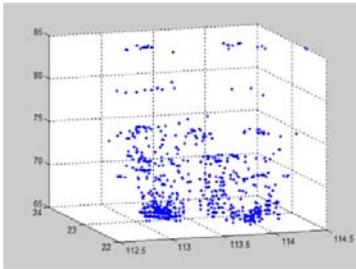


FIGURE III. TASK LIST PRICE AND TASK LOCATION DIAGRAM

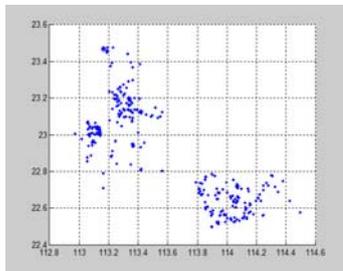


FIGURE IV. FAILED TASK DISTRIBUTIO

Figure 4 shows that the failed tasks are mainly distributed in two relatively concentrated areas, longitude  $113^{\circ} \sim 113.6^{\circ}$  latitude  $22.8^{\circ} \sim 23.2^{\circ}$  and longitude  $113.8^{\circ} \sim 114.3^{\circ}$  latitude  $22.5^{\circ} \sim 22.8^{\circ}$ .

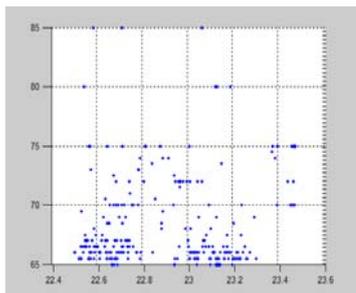


FIGURE V. FAILURE TASK LATITUDE-PRICE CHART

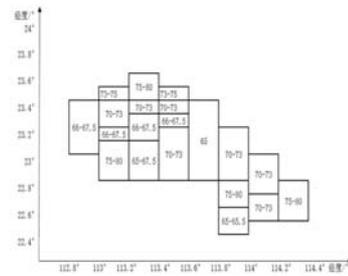


FIGURE VI. SUB-REGIONAL PLAN PRICING RANGE CHART

As can be seen from Figure 6, the prices of failed tasks are generally low except for a few. From Figure 4 and Figure 5, we know that the prices of failed tasks are low and concentrated in the longitude  $112.8^{\circ} \sim 114.6^{\circ}$  and latitude  $22.4^{\circ} \sim 23.6^{\circ}$  regions. We take a uniform distribution of 15 points. In the success of the task we randomly found with each point and the price difference is less than 3 Yuan each of the five points. We let the 75 points longitude, latitude, task value into the MATLAB to calculate the value of stats and substituted several times. When the resulting stats  $>= 0.8$  (stats = 0.861 in a trial), we think such arrangement meeting the needs, which will greatly improve the success of the task rate.

We take into account the location of members and the location of characters and price relationship, according to the region to develop price plans, price should be within the scope of the plan. Plan as shown in Figure 6. In the completed tasks, the few independent of the dense outer area should be merged or eliminated. Count different regions and price of the task, the success rate reached 72.9%. The new mission location point is plotted in MATLAB, as shown in Figure 7:

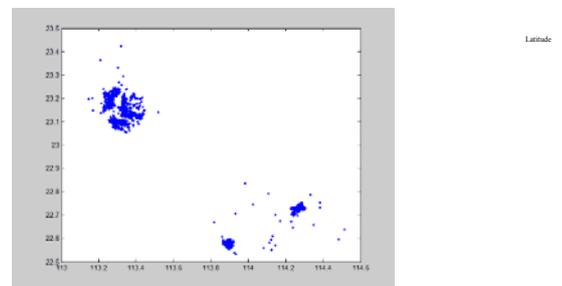


FIGURE VII. NEW MISSION DISTRIBUTION

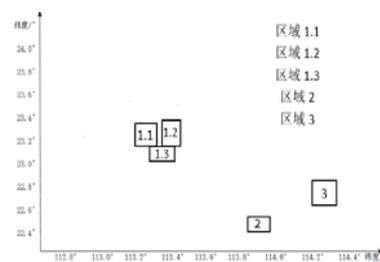


FIGURE VIII. NEW TASK AREA DIVISION DIAGRAM

Seen from the figure, mainly concentrated in three regions:

- Longitude  $113.1^{\circ} \sim 113.5^{\circ}$  and latitude

23° ~ 23.3° lines surround the area.

- Longitude 113.8° ~ 114.0° and latitude 22.5° ~ 22.6° lines surround the area.

- Longitude 114.2° ~ 114.4° and latitude 22.7° ~ 22.8° lines surround the area.

#### A. Develop A Package Plan

According to the recent order of magnitude, the task point in the area 1,2and3 are relatively far from each other. So we do not consider cross-regional packaging. The distribution of task points in area 1 can be further divided into 3 areas, and we have to arrange packing schemes separately in these 5 areas. Regional distribution shown in Figure 8:

Without affecting the regularity of statistics, the number of tasks in each region was calculated after merging the task points close to each region and removing the task points without considering packing. The number of tasks is 1.1: 509 in the area, 1.2: 712 in the area, 1.3: 208, Zone 2: 253, Zone 3: 292.

In the same way, the number of members in each region is 1.1: 195 in the region, 1.2: 194 in the region, 1.3: 168 in the region, 2:93 in the region, and 3:73 in the region.

Using formula (4), the planned number of planned packages in each region is 1.1: 74, 1.2: 76, 1.3: 27, 2: 23, and 3: 20.

#### B. Develop A Unit Price

Taking into account the membership and the intensity of the task, set the number of tasks in a region, the number of members, to develop the unit price of E:

$$E = 60 + \frac{5 \times u}{m}$$

Set the unit price for the region 1.1: 72 Yuan, the region 1.2: 73 Yuan, the region 1.3: 65 Yuan, the region 2:74 Yuan, the region 3:80 Yuan.

#### C. Set A Package Price

Amending regression equation coefficients to construct a model of single factor linear regression of packaging task price  $W$  and membership credit value

$$W = \beta_0 + \beta_1 z + \varepsilon$$

$\beta_0$ ,  $\beta_1$  is regression coefficients, calculated by *MATLAB*  $\beta_0=5.741$ ,  $\beta_1=0.835$ .

Each region finally formulates the specific price according to the formula (3) combining the member's specific credit value and the specific task.

#### REFERENCES

- [1] Sijin An and Jian Zhai. Analysis and prediction model of the influence factors of software crowdsourcing engagement [N]. Computer system application, volume 24, volume 10, 2015.
- [2] Shoukui Si and Xijing Sun. Mathematical modeling algorithm and application [M]. Defense industry press, 2011.
- [3] Huiwen Wang. The least squares regression method and its application [M]. Beijing: national defense industry press, 2000.
- [4] Zhuolin Li. The graphical method of statistical data and its application [D]. Chongqing: Chongqing University, 2015.