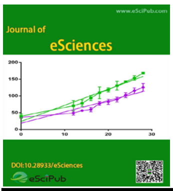
Research Article **eSciences (2020) 3:9**



Journal of eSciences
(ISSN:2637-8760)

Optimization of Task Pricing Based on Multiple Regression Analysis and Game Theory

Tianlong Wang[1], Hongyan Bao[2], Hairui Zhang[2]*

[1]College of Civil Engineering & Architecture, China Three Gorges University, Yichang, 443002, China.
[2]College of Science, China Three Gorges University, Yichang, 443002, China.

ABSTRACT

Nowadays, the Internet is developing rapidly, while in the context of the mobile Internet Making photos to make money, a kind of self-service service mode currently has the problem of unreasonable task pricing, which leads to the failure of commodity inspection. Hence, in order to solve the problem, on the basis of a set of completed project task data and the whole registered member information provided by the company, we set up Multiple regression model and Bayesian equilibrium model by means of the mechanism analysis to determine models, big data statistical to determine parameters and considering the various factors that can affect the price of task. Not only can the models established in this paper greatly optimize the task pricing scheme, but also them do improve the task completion degree as much as possible and play a positive role in the efficiency of program operation under the premise that the total cost is basically unchanged.

Keywords: K-means clustering; Multiple regression; The Bayes equilibrium; Game theory

***Correspondence to Author:**
Hairui Zhang
College of Science, China Three Gorges University, Yichang, 443002, China.

How to cite this article:
Tianlong Wang, Hongyan Bao, Hairui Zhang. Optimization of Task Pricing Based on Multiple Regression Analysis and Game Theory. Journal of eSciences, 2020; 3:9

eSciPub LLC, Houston, TX USA.
Website: https://escipub.com/

1. Introduction

In recent years, China is one of the countries with the fastest development electronic commerce and Internet technology in the world. Under the regulation of different market demands and national policies, many new business models conforming to "Internet plus" have emerged in China's mobile Internet. Among them, crowdsourcing business model which is one of the relatively new business models, has a relatively objective development prospect. Not only can the business model of crowdsourcing meet the need of fierce market competition to save operation costs, but it can also meet the needs of customers participating in product design and service planning in the network era, linking design, production and sales, and avoiding the disconnection between product design and production and customer needs to the greatest extent. In September 2015, the State Council issued the Guiding Opinions on Accelerating the Construction of Mass Entrepreneurship and Innovation Support Platform proposing to the positive development of crowdsourcing, which means that gathering the mass strives to increase employment opportunities and use the Internet or other means to complete the traditional assignment by certain enterprises and institutions. With higher efficiency and lower cost, the voluntary participants of the whole enterprises and individuals to division of labor will make the best use of public power and even meet the demand of production and life service [1-2].

Making photos to make money is a self-service pattern under the background of mobile Internet. Users register as a member of the app after downloading the app, and then receive tasks that require photos from the app (such as going to the supermarket to check the shelf status of a certain product) to earn the monetary reward designated by the app for tasks. This self-service labor crowdsourcing platform based on mobile Internet provides enterprises with a variety of business checks and information collection. Compared with the traditional market survey methods, it can greatly save the survey cost, effectively ensure the authenticity of survey data and shorten the survey cycle. Therefore, the app becomes the core of the platform operation, and task pricing in app is its core element as well. If the pricing is not reasonable, certain tasks will be ignored, leading to the failure of commodity inspection [3].

This article hopes to study the pricing rules of the *Making photos to make money* app task in order to establish a corresponding model. By formulating a reasonable pricing scheme based on the original scheme, the less cost is increasingly better on the basis that as many tasks are completed as possible. What is the most important is that it can effectively ensure the efficiency of program operation, and paly a great role in promoting the development of crowdsourcing business models [4].

This article emphases on the task data of a set of completed projects (including the location, pricing, and performance of 835 tasks) and all registered information of membership (including the location information and honor values of 1877 members). Through the study of this group of data, we need to research the company's original pricing rules and analyze the reasons why the task was not completed, according to its reason to re-formulate a new pricing plan. After that, for the task completed as much as possible on the basis of the required cost, the less cost is increasingly better.

2. Model assumptions

Before analyzing the problem and establishing the mathematical model, we need to put forward some assumptions to simplify the problem. The assumptions made in this article are as follows:

(1) Assume that all members select tasks rationally, and there is no situation where members rush to order at will;

(2) Assume that the given tasks have the same degree of difficult degree and are achievable, regardless of the impact of task descriptions and reference materials in the app on whether users accept tasks or not;

(3) Do not consider the impact of task difficult degree on task pricing;

(4) Each task is completed by one person at most, regardless of the case where multiple persons cooperate to complete a task;

(5) The position of members remains unchanged, regardless of the movement of members;

(6) Ignore the effects of weather and traffic on the problem;

3. Analysis of problem

In order to obtain the law of task pricing, it is necessary to find out the factors that affect the task price, and find out the main influencing factors through qualitative analysis. Analyzing the two sets of known data, we first consider that the geographic location of members and task points can affect the price of tasks. First of all, the larger the number of members around a task point, it will inevitably lead to the task being selected by these members, which will drive down the price. On the contrary, if the number of members around the task point is small, the chance of the task being selected will be low, which in turn drives up the price of the task. According to this analysis, the price of the task is related to the number of members around the task. In this analysis, further in-depth consideration, if there are more tasks around a task point, that is, a better geographical location of the task point, it will also be more popular with members, so as to lower the price. On the contrary, there are fewer tasks around a task point, and the more remote the location is, the higher the price will be. In addition, the task price is also related to the ease or complexity of the task, but we do not have relevant data, so we will not discuss the difficult degree of the task here. To sum up, the price of a task is mainly related to the number of members around it and the number of tasks, and both have a negative correlation. After determining the influencing factors, we used the known big data for statistical regression analysis, determined the parameters of the model, and expressed the pricing rule of tasks quantitatively. It should be noted that before this, we need to determine the membership density and task concentration degree of each task point based on the relative distance. After confirming the task pricing law, we can analyze the reasons for the unfinished task points on the basis of the characteristics of completed task points and unfinished task points.

After finding the company's original pricing rules, we need to design a new task pricing scheme on this basis to make the pricing more reasonable and improve the completion of the task. When studying the original pricing law of the company, we only considered the geographical location of the members and did not consider the influence of the member credit value on the pricing. However, through qualitative analysis, if the member credit value near a certain task point is generally high, the task is more likely to be completed, which will cause the price to be lowered. Conversely, if the member's credit value around the mission point is low, the corresponding price will be raised. Hence, we consider the member credit value as an indicator based on the original pricing law. With the previous ideas, we selected indicators to describe the average level of member credit values around each task. On the basis of the original model 1 including a membership degree of density, concentration of tasks and tasks of the remote, we added an influencing factor which also adopts the multiple regression method to quantitatively describe the relationship between the factors affecting the pricing on this task. We hope to increase the cost of less to achieve a higher degree of task completion, and even reduce the cost. In order to improve the task completion rate, the price of the task must be consistent with the user's expected price, that is, to make the company's tasks as many as possible to be completed and the required cost to reduce as far as possible, so that the new scheme is better than the original one.

4. Sensitivity analysis

We draw the position information of the finished project in the format of longitude and latitude coordinates as follows:

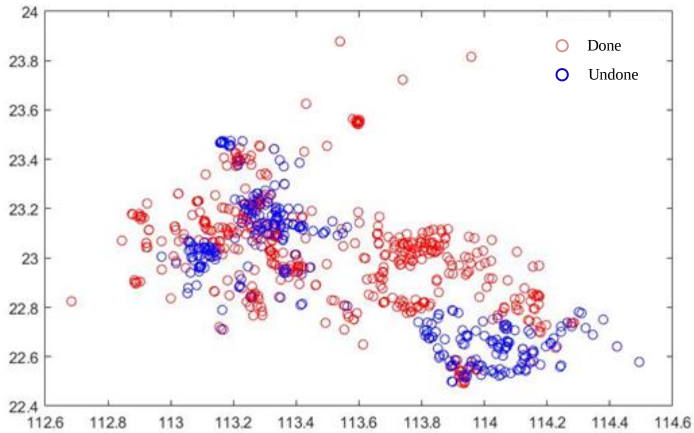


Fig. 1 Longitude and latitude coordinates of task points

In the figure above, the horizontal and vertical coordinates correspond to its latitude and longitude respectively. The hollow circle represents the coordinate position. The red circle means the completed item, while the blue circle represents the unfinished project. By analyzing the graph, we found that most of the unfinished regions are concentrated together, while the completed regions are relatively widely distributed and have special points, that is, several independent points separated from the whole large region. Combining pictures and data, we analyzed that the unfinished part in the blue area is the irrational pricing part, while the red area is the reasonably priced part. The red hollow circle in the blue area is a special case: the pricing is high.

Then we drew the latitude and longitude coordinate chart of the member's coordinate position as follows:

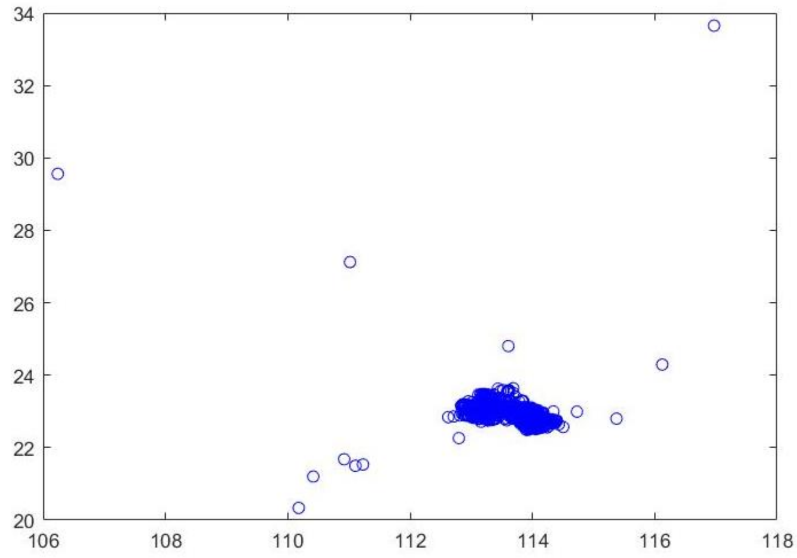


Fig. 2 Member latitude and longitude coordinates

We found that the Fig. 2 is too scattered due to the individual points, which leads to a large scope of the whole graph. We think that these points may be data errors, but it does not rule out that its location is remote. But for a more accurate analysis of these data, we excluded all remote and special points. After carefully comparing the coordinates of mission points and members, and taking into comprehensive consideration the degree of coincidence and the range of points, we only selected the points with coordinates between 22.4 ° to 23.6 ° north latitude and between 112.6 ° to 114.5 ° east longitude as the observation points. After sorting out the point diagram, we put it into the project task coordinate chart for observation, and then another longitude and latitude coordinate diagram named Fig. 3 was obtained. The diagram is as follows:

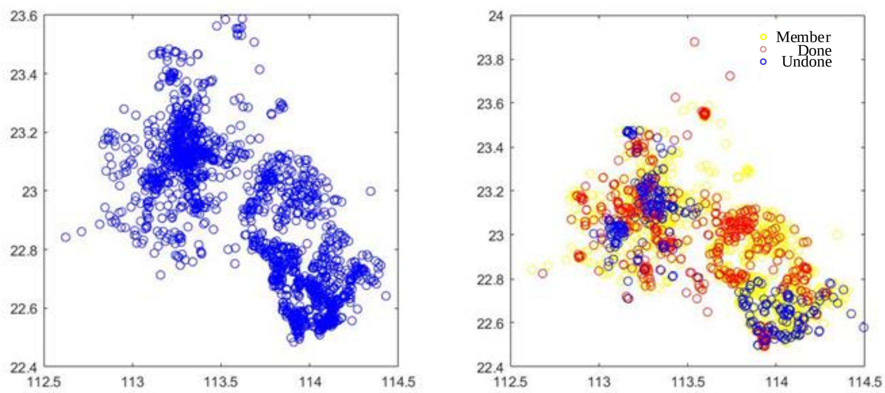


Fig. 3 Task point and member longitude and latitude coordinates

It can be clearly seen that there is an obvious correlation between the distribution of members and the distribution of task points, and the distribution area of members is wider than the distribution of tasks. However, there are also some regions with dense distribution of members and unfinished tasks. After analysis, we believe that these special regions are the regions with unreasonable prices, and we need to optimize these regions when optimizing the regions.

We divided the number of tasks completed and unfinished in each coordinate into three parts: above 69, 66 to 69, and below 66, and drew the following histogram:

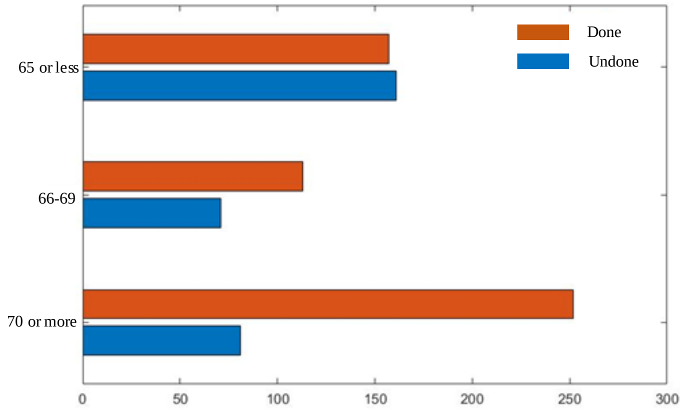


Fig. 4 Histogram of task price classification

The above histogram consists of uncompleted quantities and completed quantities, respectively, in the case of different pricing. By analyzing the Fig. 4, we can find that nearly half of the unfinished tasks are less than 66 yuan which are lower-priced tasks, and when the price increases, the number of completed tasks accounts for an increasing proportion of the total number of tasks high. It is consistent with our analysis that the higher the amount of money is, the easier it is to accept and the higher the degree of task completion will be, so we think this assumption is correct. When analyzing the unfinished part, we found that the unfinished part above 69 yuan is more than 66 to 69 yuan. We believe that the first reason is the population base, because the number of orders is higher after the price is higher. However, even if the proportion of unfinished is lower, there will be more unfinished people due to the large number of takers, which is inevitable. In addition, according to our preliminary prediction, different regions have different economic development status. In some regions, due to the good economic development status, the number of people who do not accept orders is large even though the remuneration is high.

5. The establishment of Multiple regression model

In Multiple regression model, we will use the two indicators of concentration of each task point and membership density. The determination method is based on the number of other tasks within a certain distance and membership number around the task, therefore we need to use the distance between any two task points, and between the task and the member to determine. We can get the distance between two points through the mission point and the longitude and latitude of members. The specific conversion formula is deduced as follows:

The geographical position of a point is represented by the ordered number for $(\mu, \ v)$. Then μ represents the longitude and v represents the latitude. The center of the earth is o as the origin of coordinates; the equatorial plane is

xoy planes, and the plane where the circle of longitude and latitude is 0 degree is xoz planes to establish a three-dimensional cartesian coordinate system, then the following formula can be obtained:

$$\begin{cases} x = R\cos\mu\cos\nu \\ y = R\sin\mu\cos\nu \\ z = R\sin\nu \end{cases}$$

Among them, R is the radius distance of the earth which is the fixed value of $6370\ km$.

It can be proved by mathematical knowledge that the actual distance between any two points $A(\mu_A,\ \nu_A)$ and $B(\mu_B,\ \nu_B)$ is:

$$d_{AB} = R\arccos\left[\cos(\mu_A - \mu_B)\cos\nu_A\cos\nu_B + \sin\nu_A\sin\nu_B\right]$$

From the above formula, we can obtain any two task points and the distance between the task point and the member. Based on it, two distance matrices are established:

$$A = (a_{ij})_{835\times835},\ B = (b_{ik})_{835\times1877}$$

Among them, matrix A is the distance matrix of the task point, and a_{ij} is the distance between the task point i and the task point j. Matrix B is the distance matrix between task point and member, and b_{ik} is the distance between task point i and member k.

5.1 Determination of influencing factors

Before finding the law of task pricing, it is necessary to determine the factors that affect the task price. First, qualitatively determine which factors may have an impact on the task price:

a) The number of members m_i around a task point is related to the task price. From a qualitative point of view, the more members are around a task point, the greater the probability that the task will be completed, which will further reduce the price of the task. Otherwise it is raised. Therefore, membership density is one of the factors affecting the price.
b) The number of other tasks n_i around the task point is related to the price.
c) The more remote the mission point is, the farther away it is from the member gathering center, which leads to the lower the probability of being completed, thus raising the price of the mission. Therefore, the task price is related to the remoteness of the task point.

i. *Step1: Determination of membership density*

ii. Let m_i be the center of this task point p_i and define the number of members within the distance of d, where the steps to obtain m_i are as follows. Then introduce variable λ_{ik} which is either 0 or 1. And the following relationship holds.

$$\lambda_{ik} = \begin{cases} 0, b_{ik} > d \\ 1, b_{ik} \le d \end{cases}$$

In the above relation, b_{ik} is the distance between the task point i and member k. d is the artificially given reference distance, where we think d is two kilometers. Therefore, the membership density is as follows:

$$m_i = \sum_{k=1}^{b} \lambda_{ik}$$

m_i is the number of members within d, centered on task point p_i. λ_{ik} is a variable which is either 0 or 1, in order to judge whether the distance between member k and task point i is less than or equal to $2km$. b is the number of members. If the distance is no more than $2km$, the value of λ_{ik} is 1. Instead, the value is 0. And here we take b is equal to 1877.

Step2: Determination of task concentration

Let n_i be the number of other tasks centered on task point p_i and within a distance of d, where the steps to obtain n_i are as follows.

Then introduce the variable ω_{ij} whose value is 0 or 1. And the following relationship holds.

$$\omega_{ij}=\begin{cases}0, a_{ij}>d\\1, a_{ij}\le d\end{cases}$$

In the above relation, a_{ij} is the distance between the task point i and the task point j . d is the artificially given reference distance, where we think d is two kilometers. Hence, the task concentration is as follows:

$$n_i=\sum_{j=1, j\neq i}^{a}\omega_{ij}$$

Among them, n_i is the number of tasks within d , centered on task point p_i. ω_{ij} is a variable which is one of two numbers 0 and 1 for judging whether the distance between task point i and task point j is less than or equal to $2km$. a is the number of tasks. If the distance is no more than $2km$, the value of ω_{ij} is 1. Instead, the value is 0. And here we take a is equal to 835.

Step3: Determination of remoteness of task point

First and foremost, we conducted a clustering analysis on the geographical locations of members. Before clustering, we need exclude those points far away from the main region to avoid inaccurate clustering caused by these points[5]. According to the coordinate points of the member distribution in the two-dimensional coordinate system, it can be roughly divided into three regions, so we take k is equal to 3. k is the quantity to be classified. Through cluster analysis, we can divide the members into three regions k_1, k_2 and k_3. Each region has a central point, denoted as o_1, o_2 , and o_3.

Before judging the remoteness of the task, a scale is needed, that is, all task points should be divided into three regions, namely:

$$\min\{p_io_1, p_io_2, p_io_3\}$$

The region center closest to the mission point is the region of the mission. The cluster analysis diagram is as follows.

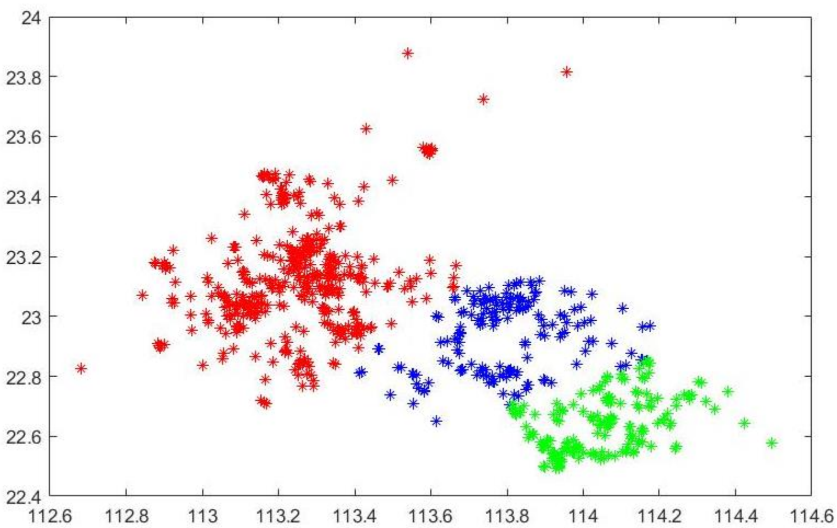


Fig. 5 Longitude and latitude coordinates of task points after cluster analysis

Step4: Data normalization processing

To eliminate the influence of dimension on the results, we normalized the data of membership density, task concentration and task remoteness. The specific processing method is as follows:

①Processing of membership density

$$x_1=\frac{m_{\max}-m_i}{m_{\max}}$$

By the equation, x_1 is a matrix with 835 rows and 1 column, $m_{\max}$ is the maximum value in m. And m_i is the number of members within

$2\mathrm{km}$ distance centered on the task point p_i.

②Processing of task concentration

$$x_2 = \frac{n_{\max} - n_i}{n_{\max}}$$

Among the equation, x_2 is a matrix with 835 rows and 1 column, $n_{\max}$ is the maximum value in n. And n_i is the number of other task points within $2km$ distance centered on the task point p_i.

③Dealing with the remoteness of the task

$$x_3 = \frac{|p_i o_j|}{|p_i o_j|_{\max}}$$

In the above relation, x_3 is a matrix with 835 rows and 1 column. $|p_i o_j|$ represents the distance between the task point p_i and the center point of its region. $|p_i o_j|_{\max}$ is the maximum distance in the region.

5.2 Multiple regression analysis

Therefore, we can establish the relationship between task pricing and the three influencing factors[6-7], namely:

$$y_i = x_0 + \alpha x_1 + \beta x_2 + \gamma x_3 + \delta$$

Among them, $x_0, \alpha, \beta, \gamma$ and δ are an unknown parameters independent of x_1, x_2 and x_3, where $x_0, \alpha, \beta, \gamma$ and δ are called the regression coefficient.

Now we get 835 observation data, namely the data $[b_i, a_{i1}, a_{i2}, a_{i3}]$ of 835 task points. In these data, b_i is the observed value of y, while a_{i1}, a_{i2} and a_{i3} are the observed value of x_1, x_2 and x_3 respectively. From the above equation, we can obtain:

$$\begin{cases} b_i = x_0 + \alpha a_{i1} + \beta a_{i2} + \gamma a_{i3} + \delta_i \\ \delta_i \sim N\left(0, \sigma^2\right), \qquad i = 1,2,3 \end{cases}$$

Remember to:

$$X = \begin{bmatrix} 1 & a_{11} & a_{12} & a_{13} \\ 1 & a_{21} & a_{22} & a_{23} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & a_{n1} & a_{n2} & a_{n3} \end{bmatrix} \quad , \quad Y = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

$$\delta = [\delta_1, \delta_2, \cdots, \delta_n]^T \quad , \quad B = [x_0, \alpha, \beta, \gamma]^T$$

At the same time, we take n is equal to 835. The parameter x_0, α, β and γ in the above multiple regression model are estimated by the least square method, that is, the estimated value $\hat{B}$ should be selected to make the sum of the squares of errors minimize. So the sum of the squares of the errors is listed below.

$$Q = \sum_{i=1}^{n} \delta_i^2 = \sum_{i=1}^{n} \left(b_i - \hat{b}_i\right)^2$$

Same as above, we take n is equal to 835. Obviously, it's easy to prove this mathematically.

$$X^T X B = X^T Y$$

The solution to the above equation is:

$$\hat{B} = \left(X^T X\right)^{-1} X^T Y$$

Substitute $\hat{B}$ back to the original model to get the estimated value of y, and then we made a significance test for the regression equation, namely:

$$R^2 = \frac{\sum_{i=1}^{n} \left(\hat{y}_i - \bar{y}\right)^2}{\sum_{i=1}^{n} \left(y_i - \bar{y}\right)^2}$$

In the above relation, $\hat{y}_i$ is the forecast value, and the observed value is y_i. And $\bar{y}$ is the mean value of the observed value.

5.3 The result analysis of Multiple regression model

According to the above model, we used $MATLAB$ software to solve the problem. Firstly, we used $K-means$ algorithm to cluster the member locations, and obtained the longitude and latitude coordinates of the center points of the three regions as follows:

Table 1: Latitude and longitude coordinate table of the center points of the three regions

	o_1	o_2	o_3
Longi-tude	113.2757	114.0689	113.8294
Latitude	23.1235	22.6414	22.9159

The regression coefficients and correlation coefficients R^2 of multiple regression equations are shown in the following table:

Table 2: Regression coefficient and correlation coefficient R^2 data table

x_0	α	β	γ	R^2
59.837	7.956	3.245	1.018	0.741

That is, the pricing law of tasks is as follows:

$$y = 59.837 + 7.956x_1 + 3.245x_2 + 1.018x_3$$

Among the equation, y is the pricing of the task and x_1 is the membership density of the task point. The concentration degree of the task is x_2, and the remoteness degree of the task is x_3.

Combining the results with the information on the map, we got the reasons for the unfinished work as follows:

(1) Mission point is too remote, but the price is not high, which is not up to the expected value of members.

(2) Even though the geographical location is in the concentrated area of the members, the price of the task itself is still low. It does not reach the expected value of the members either.

To sum up, the reason why most tasks are not completed is that their prices are lower than they should be.

We made a pie chart for the weights of the three variables as follows:

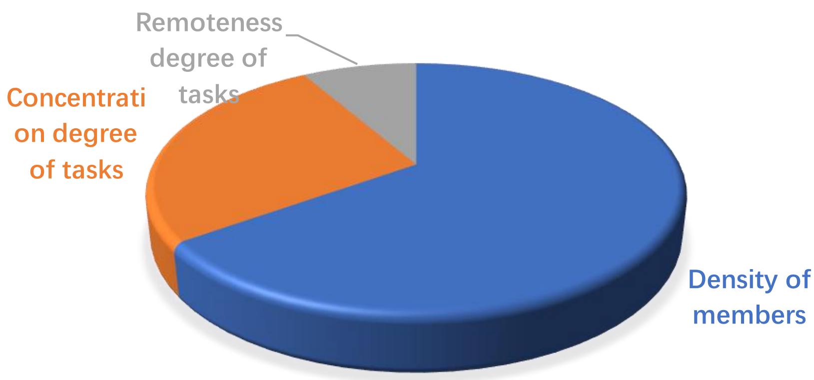


Fig. 6 Variable weight pie chart

According to life experience and data analysis, the higher the membership density of a place is, the more likely the tasks of that place are to be accepted and completed, while the importance of remoteness and concentration of tasks is not very high. Since the remoteness of the task is not directly related to the location of the member, that is, the member may also appear in the remote place, and the member will preferentially choose the task which is closer to him and easier to complete. Therefore, we think the results are more reasonable and consistent with the actual situation.

6. The establishment of Improved multiple regression model and Bayesian equilibrium model

6.1 Determination of influencing factors

First of all, we conducted qualitative analysis on the influencing factors, and then added the factor of the credit value of members on the basis of the above three influencing factors.

Step1: Divide the credit value of members into ranges

Due to the large difference of members' credit values and the extreme disunity of dimensions, we adopted the method of grading by segments, and replaced the honor values of members with the numbers in the interval from 0 to 1. The specific dividing intervals and corresponding scores are shown in the following table.

Table 3: Dividing intervals and corresponding score table

>20000	5000-20000	800-5000	100-800	10-100	<10
1	0.8	0.6	0.4	0.2	0

Then each member corresponds to his own credit score g_i.

Step2: The credibility of the mission point

It is necessary to describe the credibility degree of members around each mission point. We used the average of the member's credit score within $2km$ to represent the credibility of the task point, that is:

$$x_4 = \frac{1}{m_i}\sum_{i=1}^{m_i} g_i$$

Among them, x_4 is a matrix of 835 rows and 1 column, representing the credibility of each task point. m_i represents the membership density of task point p_i. And g_i represents a member's credit score.

Step3: Multiple regression analysis

Define a relationship as follows.

$$y = x_0 + \alpha x_1 + \beta x_2 + \gamma x_3 + \theta x_4$$

Based on the relation, y is the observed value of the completed task. $x_0, \alpha, \beta, \gamma$ and θ are the regression coefficient of the multiple regression equation. x_1, x_2, x_3 and x_4 are the membership density, concentration, remoteness and reputation of the mission. If it is necessary to prove the significance of the regression equation, we will select the observed values of the completed tasks for fitting.

6.2 Establishment of Bayesian Equilibrium Model

The above modeling process has solved the optimization of the task pricing scheme. Below we calculated the task completion degree under the new scheme. However, whether a task will be completed by a member depends on whether the price of the task has reached the member's expected price. Of course, members hope that the price is becoming higher, which is much better for them, while the company wants to lower the price. At the same time, the company hopes that tasks are completed as much as possible, so there is a game in this process. The premise is that members don't know the company's highest level, the company don't know the member of the lowest price expectations either. Therefore, we establish a Bayesian equilibrium model[8][9] to calculate the task completion degree under the scheme, so as to evaluate the effect of the new scheme.

The lowest price acceptable to members is ω_s, and the highest price acceptable to the company is ω_b. The member's quotation is q_s, and the company's quotation is q_b.

For any given ω_s, the member's quote q_s should maximize the expected profit. Because on the basis of the assumption of Bayesian equilibrium model, this task can be completed only when q_b is greater than or equal to q_s. After the transaction, members' profits are:

$$(q_s+q_b)/2-\omega_s$$

When the task is completed, the profit is 0, so q_s should meet the condition:

$$\max_{q_s}\left\{\frac{q_s+E\left[q_b \mid q_b \ge q_s\right]}{2}-w_s\right\}P\left\{q_b \ge q_s\right\} \quad (1)$$

(Where $E[\ ]$ represents the conditional expectation of q_b under conditions $q_b \ge q_s$, and $P\{\ \}$ represents the probability of events)

Similarly, q_b should meet the condition:

$$\max_{q_b}\left\{\frac{q_b+E\left[q_s \mid q_b \ge q_s\right]}{2}-w_s\right\}P\left\{q_b \ge q_s\right\} \quad (2)$$

If combination (q_s,q_b) satisfies both equations (1) and (2), it is an equilibrium of both sides. There are many equilibria for this game problem, and here we chose linear price equilibrium. Assume that the quotations of the member and the company are linear functions of the value of the task to both, expressed as:

$$\begin{cases} g_s = a_s + c_s\omega_s \\ g_b = a_b + c_b\omega_b \end{cases}$$

We determine the coefficient a_s,c_s,a_b and c_b to try our best to equalize equations (1) and (2).

It can be proved mathematically that g_s obeys the uniform distribution on $\left[a_s,a_s+c_s\right]$.At this point, for a given ω_b, the optimal response of the company is to find g_b that satisfies equation (2). When $g_b \in \left[a_s,a_s+c_s\right]$, then $p\{g_b \ge g_s\}=(g_s+g_b)/2$, therefore formula (2) is:

$$\max_{q_b}\left\{\omega_b-\frac{g_b+(a_s+g_b)/2}{2}\right\}\cdot\frac{g_b-a_s}{c_s}$$

This is the optimization of a quadratic function, and the optimal solution is:

$$g_b=\frac{2}{3}\omega_b+\frac{1}{3}a_s$$

Similarly, formula (1) can obtain the optimal response of members:

$$g_s=\frac{2}{3}a_s+\frac{1}{3}(a_b+c_b)$$

It can be proved that under Bayesian linear price equilibrium, its task completion degree is:

$$f=\frac{\int_l^h\int_0^{\omega_b-l}(\omega_b-\omega_s)\,d\omega_s d\omega_b}{\int_0^h\int_0^{\omega_b}(\omega_b-\omega_s)\,d\omega_s d\omega_b}$$

According to the equation above, ω_b is the highest price the company can pay. ω_s is the lowest price acceptable to members. h is the maximum value in the estimated price. l is one fourth of the maximum value, that is,

$l = \frac{1}{4}h$.

6.3 The result analysis of Improved multiple regression model and Bayesian equilibrium model

Firstly, the regression coefficients $x_0, \alpha, \beta, \gamma$ and θ of the above multiple regression equation were calculated by using software $MATLAB$. The specific values are shown in the following table:

Table 4: Regression coefficient table

x_0	α	β	γ	θ
58.9001596	10.8338773	4.08061201	1.73053334	-5.85930728

That is the new task pricing scheme is designed as follows:

$$y = 58.9 + 10.83x_1 + 4.08x_2 + 1.73x_3 - 5.86x_4$$

Based on the new task pricing scheme, the maximum estimated task price is 73.9 . For the sake of calculation, let's take $h = \hat{y}_{\max} = 76$, that is, $l = \frac{1}{4}h = 19$. Then it was substituted into the above task completion formula and calculated the task completion $f = 0.8437$ under the new pricing scheme, which means 84.37% of the tasks would be completed. However, the task completion degree in the original scheme was only 0.6251 . Therefore, it can be considered that the task completion degree in the new scheme is higher than that in the original scheme.

We plotted the completion degree before and after optimization into a pie chart as follows:

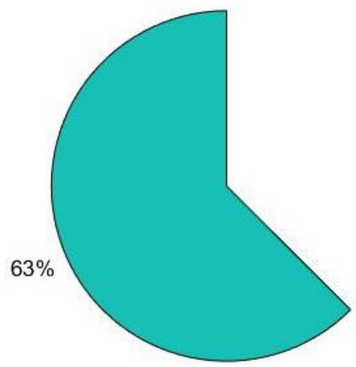


Fig. 7 Original scheme completion pie chart

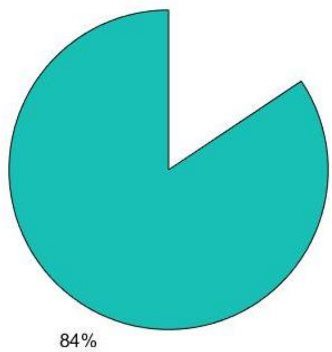


Fig. 8 new scheme completion pie chart

Comprehensive analysis of two charts, we found that the new design scheme greatly surpassed the previous plan in terms of completion. We think that 84% of the completed degree is quite high, and it is 21% higher than the original 63%, leading to a great promotion of task completion. So we believe that the new task pricing scheme is more scientific and effective than the company's original scheme.

7. Advantages of the model

(1) In Multiple regression model, we used the method of mechanism analysis to determine the model and big data statistical regression to determine parameters to find the pricing task law, which makes the model solution convenient and convincing.

(2) When making regression statistics, the data are normalized to eliminate the influence of

dimension, that is, the influence of each factor on the task price can be judged by its regression coefficient.

(3) The method of cluster analysis is cleverly used to express the remoteness of the task conveniently.

(4) In Bayesian equilibrium model, with the clever use of Bayesian equilibrium theory, it is more convenient to obtain reasonable pricing of each task point and members' expected price of each task, so as to determine the completion of the task.

8. Shortcomings of the model

(1) In Multiple regression model, only the linear relation is considered, and the nonlinear relation of multiple times is not discussed, resulting in a low degree of fitting.

(2) Due to the lack of data, the impact of the difficulty of the task on the completion of the task was not taken into account, and the model processing was not rigorous enough.

9. Promotion of the model

First of all, in modern statistics, the determination of influencing factors often relies on the methods of natural science and the use of mathematical tools. Then determining models and parameters through mechanism analysis and big data analysis, so as to reflect the weight of influencing factors and simplify the complex statistical regression problem. In addition, multiple regression analysis is not affected by data distribution, sample size and the number of indicators. It is more flexible, convenient and highly reliable, and can be used in industrial production, economic returns and other industries. Therefore, the model is suitable for most of the multi-factor and multi-variable problems and problems related to game theory. Furthermore, the application background of the model in this paper is the labor crowdsourcing platform based on smart phones and mobile Internet. Under similar background, there are many applications, such as food delivery application, Didi taxi, express delivery service platform and so on, which all involve the relationship between commodity pricing, geographical location and member positivism.

10. Conclusion

In this paper, aimed at the task pricing of *Making photos to make money,* a multivariate regression model and a Bayesian equilibrium model are established. We obtained the four main factors affecting the task pricing, that is, member density, concentration of tasks, degree of tasks remoteness, degree of credibility of members around the task points and their respective weights by means of determining the model and parameters by the mechanism analysis and large data statistics. We also found the rule of company of the original task pricing and analyzed the reason why most of the tasks are not completed is that the price is too low, failing to live up to the expectations of the member price. On a solid basis of the original model, this paper reformulates the pricing rules, provides or reduces the price of some tasks appropriately, and improves the task completion rate by 21% on the basis that the total cost is basically unchanged. In general, this article provides the company with a reasonable task pricing scheme. On the basis of the company's original pricing law, we added other indicators affecting the price of task, which not only retained the company's original pricing characteristics, but also made the price of the task more reasonable.

Furthermore, compared to the original scheme of the company, the new scheme can greatly improve the task completion with the almost constant cost, leading to the efficiency of the program operation and the promotion of the development of crowdsourcing business model to a certain extent.

References

1. Xiaoyun Chen. Research on the Business Model of Retail Enterprises in the Era of "Mobile Internet +" Sharing Economy——Retail Crowdsourcing Shopping Model [J] .Journal of Baoji University of Arts and Sciences (Social Science Edition), 2016,36(06):76-81.
2. Shu Wang. Research on Collaborative Innovation of Crowdsourcing Model of Online Business Platform [D] .Zhejiang University, 2012.
3. Yixin Cui, Ziwei Shen, Xu Zhang. Research on the new business model of "taking pictures to make money" [J] .Modern Business, 2019(36):1 3-16.
4. Yajie Zhu. Research on Crowdsourcing Business Model Element Model and Operation Mechanism [D] .Shandong University, 2011.
5. Xing Wang. Analysis of electricity consumption behavior of residents in Jingzhou City based on K-means clustering algorithm [J] .Communications World, 2019(08):276-277.
6. Jie Fang, Zhonglin Wen, Dongmei Liang, Nini Li. Analysis of Moderating Effects Based on Multiple Regression [J]. Psychological Science, 2015,38(03):715-720.
7. Qiyuan Jiang, Jinxing Ye. Mathematical Model (Fifth Edition), Beijing: Higher Education Press, 1986.6.
8. Jianqiang Fan, Haicheng Xu. Bidding Model Based on Bayesian Game Equilibrium at Reasonable Low Price [J] .Statistics and Decision, 2008(18):65-67.
9. Frank R. Giordano, William P. Fox, Steven B. Horton. Mathematical Modeling, Beijing: Mechanical Industry Press, 2014.6.