

Research on Mission Pricing of Crowdsourcing APP

Jiang Shaoyan^{1,3}, Lin Songhao², Chen Yiquan^{2,3}, Zhang Yuanbiao^{3,4,5}

¹International Business Institute, Jinan University, Zhuhai 519070, China

²Electrical and Information Engineering Institute, Jinan University, Zhuhai 519070, China

³Mathematical Modeling Innovative Practice Base, Jinan University, Zhuhai 519070, China

⁴Packaging Engineering Institute, Jinan University, Zhuhai 519070, China

⁵Key Laboratory of Product Packaging and Logistics of Guangdong Higher Education Institutes, Jinan University, Zhuhai 519070, China

Correspondence: Jiang Shaoyan, International Business Institute, Jinan University, Zhuhai 519070, China; Mathematical Modeling Innovative Practice Base, Jinan University, Zhuhai 519070, China.

Received: January 12, 2018

Accepted: February 6, 2018

Online Published: February 23, 2018

doi:10.5539/ibr.v11n4p1

URL: <https://doi.org/10.5539/ibr.v11n4p1>

Abstract

In order to have a better study of pricing of crowdsourcing APP tasks, mission data was collected about an already completed project, including the location, pricing and fulfillment of each mission. Get the law of the task pricing of the completed projects by analyzing the location distribution of the task points and the members, and then analyze the reason of the task failure. Besides, combine the information data of the members to improve the original pricing model based on the K-means clustering. Focus on the task package pricing, design a better task pricing model to improve the success rate of the task.

Keywords: crowdsourcing task pricing model, cluster analysis, package pricing

1. Introduction

In the era of big data, the collection of data presents many complicated problems. It is hard to satisfy the diversified data needs of organizations for an organization with only individual organization. The self-service crowdsourcing platform based on mobile Internet such as "making pictures and making money" adopts the shared economy mode, that is, the crowdsourcing mode, which greatly saves the investigation cost and shortens the investigation period, so that it provides a large amount of data for the organization at a lower cost.

At present, the research on crowdsourcing task mainly focuses on the developmental status of crowdsourcing and its advantages. HOWE J (2006) proposed the concept of crowdsourcing. Yan Jie et al (2017) explored the definition, patterns, composition and workflow of crowdsourcing. Zhao Jiamin (2015) obtained the conclusion that the advantages of crowdsourcing are improving efficiency and the fully elastic production scale. Huang Mincong (2016) studied the development of crowdsourcing research platform and its operating mode, and proposed the ways to operating crowdsourcing platform. Liu Ya-ru (2016) thinks crowdsourcing delivery model is more excellent than the traditional distribution methods, which achieves the integration of social idle resources. However, with the development of crowdsourcing platforms, platforms have gradually exposed problems. Pang Jianguang (2015) put forward that Internet-based crowdsourced communities are good for exploiting the wisdom of the general public, but due to the asymmetry of information, there are some risks, so the risk management mechanism of crowdsourced communities should be improved. In response to the fraudulent activities in the crowdsourcing platform, Fei Youli et al. (2015) proposed that the credit evaluation system should be improved, the fraudulent screening mechanism should be optimized, and the relevant legal system should be established and improved.

Some researchers also study the distribution of crowdsourcing tasks from the dynamic crowdsourcing task. On the premise of considering the dynamic characteristics of distribution demand, Zhao Xinglong (2016) established a dynamic model of crowdsourcing distribution path to improve the public Package distribution staff word delivery revenue and customer service experience. Qian Guang (2017) introduced personalized service into crowdsourcing logistics model to explore the impact of personalization in total logistics on customer satisfaction, namely the completion of tasks and the failure of identification. Wu Jinghong and Lv Nengfang (2016) analyzed the key players such as the outsourcing parties and outsourcing parties in the crowdsourcing mode of express delivery

companies and put forward the key tactics for the express delivery enterprises to build crowdsourcing model.

However, few articles study the crowdsourcing task based on the static distribution of crowdsourcing tasks and the pricing of tasks. Static crowdsourcing tasks are relative to the dynamic terms, and it is the task that after publishers release of tasks, the number of tasks released on the platform are fixed. In this paper, by analyzing the distribution of the tasks and membership positions of a completed project, we get the task pricing rules and tasks pricing model. Introducing the K-means clustering to a new pricing model to improve the success rate of the task, which provide a reference for the pricing of crowdsourcing APP tasks.

2. Data Source and Processing

2.1 Data Source

Based on the original data of the 2017 National Collegiate Mathematical Contest in Higher Education (SCE), according the task data of a completed project, including the location, pricing and completion of each task("1" indicates completion and "0" indicates uncompletion), explore the project's task pricing rules and analyze the reasons for the task is not completed. According to the information of members, including the position of member, the credit value, reference to the credibility of its mission to start booking time and booking limit, design a new task pricing program to improve the success rate of the task.

2.2 Data Processing

Use the Baidu Map to mark the location of the task points, members of the seat on the map, and make a statistical processing to identify the representative of the city as an analysis example. The urban distribution of this completed mission data is shown in Table 1.

Table 1. Mission points, the number of members in each city distribution

City	Shenzhen City	Dongguan city	Foshan City	Guangzhou City	others
Number of mission points	164	175	174	321	1
The number of members	629	352	216	665	15

3. Task Pricing Rule Analysis and Establishment of Pricing Model

3.1 Task Pricing Rule Analysis

According to the available data, it is assumed that the difficulty of each task is same, so the pricing is mainly affected by the spatial location distribution. Using the triangle derivation formula to calculate the distance between the task point and each member position in the city, and then take the average value to indicate the distance between the task point and the member position. The relationship between the distance and pricing in Shenzhen are shown in Figure 1 and Figure 2.

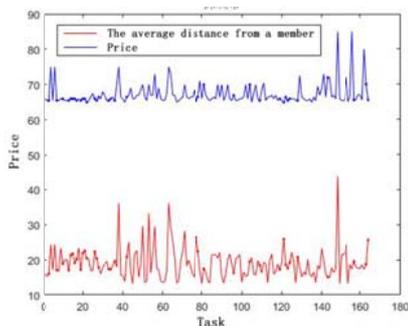


Figure 1. The curve about the relationship between the distance and pricing in Shenzhen

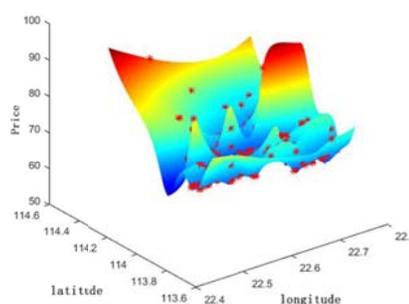


Figure 2. The three-dimensional surface map about tasks intensive distribution and pricing in Shenzhen

From Figure 1 and Figure 2, it is clear that the pricing of each task fluctuates with the distance between the task point and the member's position, especially in the vicinity of 40, 60 and 150 task points. When the task point and the member's distance at the maximum, the task point pricing will increase sharply. Contact the actual situation shows that when the task points is far away from others, their pricing will be a corresponding increase. And when the task points are close to others, their pricing is lower and more stable.

According to the pricing rules of tasks in Shenzhen and the data of other cities, it is clear that there are two rules in the pricing of tasks: the pricing of tasks is positively related to the distance of members, that is, when the

position of the task is far away from more members, the corresponding will be increased. Another is the task pricing and the distribution of the task itself has a negative correlation between the concentration, that is, when the task distribution is more intensive, the all of tasks' price will be reduced.

3.2 Establishment of Pricing Model

(1) Establishment of pricing model

In this paper, the static crowdsourcing task is considered, that is, the number of platform members does not change over time. The model includes multiple tasks and multiple platform members for one task.

Each member a_i contains two attributes: one is the cost of completing the task C_i , and another one is the number of tasks w_i that the platform members can complete, namely the task reservation quota. However, in practice, the task publishers don't know these two attributes of the members, and their purpose is to be able to receive the tasks from the platform members under the lowest cost. Therefore, this paper judges the quality of the completion of the task based on the credit value of the member, so as to maximize the task publishers' profits.

Therefore, before determining the task pricing, the paper first explore the quality of the completed work by analyzing the creditworthiness of the members. At the pricing stage, according to the cost and the number of tasks that can be accomplished by the member, we use the pricing mechanism to design a platform member. Finally, according to the first part of the quality results are sorted in descending order, priority to select high-quality members to complete the task. The specific algorithm is as follows:

Crowdsourcing platform pricing mechanism design algorithm:

Step 1: Enter the platform member attribute data $\{(C_1, w_1), (C_2, w_2), \dots, (C_n, w_n)\}$, C_j relates to the distance between members and task and member's credit value, w_j represents maximum limit for the task;

Step 2: Initialize, set $i = 1$, for each task b_i , repeat the following steps:

Step 3: Set $j = 1$, sort the member data a_j by $\{c_1 \leq c_2 \leq \dots \leq c_n\}$;

Step 4: If the current member a_j has selected the number of tasks $\bar{w}_j < w_j$, then $\bar{w}_j = \bar{w}_j + 1, p = C_i$, otherwise, repeat this step;

Step 5: End step 2 cycle;

Step 6: The algorithm is over.

When the unit price of task is $p = C_i$, the number of tasks assigned to platform members a_i can be expressed as \bar{w}_i . There are three main reasons for designing this pricing strategy: First, the cost is sorted according to the distance between platform member and task location, which can avoid pricing too high or too low; Secondly, under the condition of maximizing the profit of the task publisher, select the tasks are priced low enough to control the cost of the task; Thirdly, for the platform members, the task price is in line with the expected return of most members, so it will have a high acceptance, and will not make the task unattended, leading to the failure of the task.

(2) Taking Shenzhen as an Example to Solve Task Pricing Model

Taking Shenzhen as an example, substitute the task information data and membership information of Shenzhen into the pricing model based on mechanism design theory to implement this pricing scheme for different tasks in this project, and the crowdsourcing task are priced. The part of the results shown in Table 2:

Table 2. Pricing of Tasks in Shenzhen Based on Mechanism Design Theory

Task number	New pricing	Task number	New pricing	Task number	New pricing
A0001	65.37	A0011	66.36	A0021	67.02
A0002	67.47	A0012	66.44	A0022	66.50
A0003	65.17	A0013	65.70	A0023	66.46
A0004	70.22	A0014	66.74	A0024	66.28
A0005	65.51	A0015	66.84	A0025	65.96
A0006	70.18	A0016	66.16	A0026	66.99
A0007	65.65	A0017	66.43	A0027	66.49
A0008	65.38	A0018	67.21	A0028	65.77
A0009	67.21	A0019	65.72	A0029	65.41
A0010	66.10	A0020	66.88

(3) Comparative Analysis of New Pricing and Original Programs

The Shenzhen pricing data shows that the mean price of the task is 68.00 yuan based on design of the mechanism, while the average of the original pricing scheme is 67.33 yuan, so the price increase is not large. However, based on the design of the mechanism, task pricing fully accounts of the membership credit value of the impact, improve the task completion rate.

4. Pricing Model of Task Based on K-Means Clustering

4.1 K-Means Clustering Tasks to Package

The K-means clustering algorithm ^[8] has high efficiency for processing large data sets, especially when the sample distribution is in-class agglomeration, which can achieve good clustering effect. The basic idea is as follows: the objects are divided into k clusters, making the clusters have a high degree of similarity, while the similarity between clusters is lower. The general steps are as follows:

Step1: Randomly select k objects from n data object as the initial cluster center;

Step2: Calculate the Euclidean Distance between each object in the sample set and these center objects according to the mean value (center object) of all the objects in each cluster, and then divide the corresponding objects according to the minimum distance again;

Step3: Recalculate the mean (center object) of each (changing) cluster;

Step4: Loop Step2 to Step3 until each cluster no longer changes, and output k cluster satisfying the minimum variance standard.

Considering that the completion of some tasks in the project has not been successfully completed, which still has failure. In order to improve the completion rate of the task, starting from the cost of task publisher and the reward of platform members, the weighted European distance K-means clustering algorithm is used to divide the positions of the task points. For reference, Zhou Shibing proposed a new algorithm to determine the optimal clustering number of K-means clustering algorithm based on the design of a new clustering effectiveness index, identified the clustering numbers k when the task location was divided, so as to further realize the task packing in similar distance.

Using MATLAB software to perform K-means clustering on the task positions of completed projects in Shenzhen, the packing situation of similar tasks is obtained, with a total of 40 clusters. The relative positions of all the task points after clustering are shown in Figure 3.

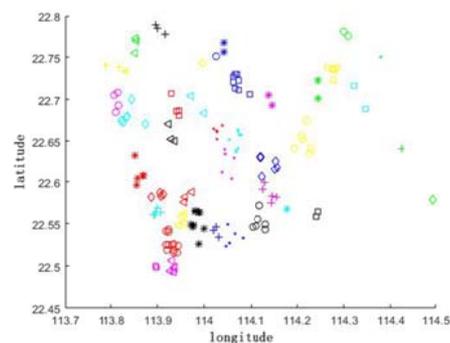


Figure 3. The relative positions of all the task points after clustering in Shenzhen

Observing the figure above, we find that the locations of some of the mission sites are relatively concentrated. Therefore, these tasks can be packaged and released jointly, which not only avoids publisher's trouble with the follow-up communication members about completing the "one-on-one" task, but also reduces the pricing of the release of these tasks. At the same time, for the platform members, the choice of multiple tasks within a small range not only be able to get a considerable price benefits, but also because of its mission from multiple tasks within a small range control the cost of the member to complete the task. In the end, it can reach a win-win situation, while meeting the task of the publisher and platform members expected goals.

4.2 Solved and Analyzed Pricing Result of Packaged Task

Based on the improvement of task pricing model established in 3.2, firstly, package the previous location of the task to price and distribute a task pricing and distribution. In determining the pricing of a task package, the task using each package the maximum value of the task package plus the package task pricing averages. Secondly, multiply by the number of tasks contained in the package and then multiply by a certain percentage of the price

of the task package. Among them select the maximum value of the task can offset the members alone to complete each task required in the package, and the average value of the task pricing is representative in the package. On the other hand, in order to meet the maximization of the revenue of the task publisher and the platform member, the ratio can be 0.618, the golden section, so that the task publisher has a minimum cost, and membership can receive the most effective task.

Task package pricing expression is as follows:

$$P = \max p_i + \bar{p}_i \times 0.618 \times (n-1) \{i=1, 2, \dots, n\}$$

Among them, P represents the cost of a task package, n represents the number of tasks packaged in a task package, $\max p_i$ represents the maximum value of the pricing of tasks p_i received by a member within a task package, \bar{p} represents the mean value of the pricing of tasks received by members within a task package.

Based on the mechanism design theory of the task pricing model, using MATLAB software to calculate the task data in Shenzhen, get the tasks' numbers and price after being packaged as shown in Table 3.

Table 3. Shenzhen the tasks' numbers and price after being package

Package the number of tasks	Pricing						
3	152.14	13	551.19	2	117.14	7	321.61
5	240.80	1	73.18	1	69.80	1	72.24
4	192.13	6	285.86	2	109.66	4	193.26
5	248.39	5	237.70	14	599.33	10	435.15
2	108.74	5	233.12	3	150.19	2	114.64
5	233.32	4	196.07	6	278.13	5	232.38
1	74.82	2	118.83	3	155.89	2	111.97
1	68.74	5	230.39	2	113.61	1	72.03
8	366.08	5	233.53	2	108.00	4	193.24
3	156.11	3	155.04	4	196.68	3	151.20

Comparing with 3.2 results, found that the packaged task in Shenzhen City, the total cost was 7952.33 yuan, and the total cost of packaging before the task was 11,152.33 yuan. We can see that a more concentrated package of tasks can reduce the cost of the task publisher of publishing the task of pricing after being integrated to a great extent.

On the other hand, members with higher creditworthiness can choose the task package closest to them within the scope of their pre-defined task quota, and accomplish multiple tasks within the task package at one time, resulting in higher economic benefits. Compared with the selection of members after each task is released separately, the distribution of focused tasks after being packaged meets the target of reducing the member's paying cost and obtaining a considerable profit.

At the same time, by observing the formula of the task package pricing, the more number k of K-means clustering, the smaller of cost of the task publisher, but each member will be bound by the task quota when receiving the task. Therefore, when the number of clustering is too small, the number of tasks in each task package will be too much, which will result in many members can't receive the task package, resulting in the increase of the failure rate of the task. Therefore, it is necessary to determine the reasonable K-means clustering number k .

5. Conclusion

5.1 Practical Significance

Through the research of this paper, it is found that the task attribute and the membership attribute lead to different competitive conditions in the crowdsourcing task market, which causes the task publisher to adopt different bidding strategies.

Specific with the actual situation, the task of pricing may also have the following conditions: the higher price of crowdsourcing platform similar tasks recently, the higher the task pricing; the shorter the task requirements of the completion of the task, the higher the pricing, etc. All of these results are determined by the incentive mechanism theory. On the crowdsourcing task platform, the task of publishers and members are often expected to be able to get their own income maximization. Platform members will choose some of their more favorable tasks to complete. Therefore, high pricing of tasks is a big attraction for members. And task publishers balance other factors based on pricing, such as ensuring completion of tasks, quality of accomplishment, and so on.

However, platform members may not be able to accomplish the task efficiently and effective after receiving the task. To ensure the quality of the task completion, the crowdsourcing platform should improve the business risk management. The platform not only needs to verify the real name of the platform member "package", but also

improve the credit rating system. Tu Shuli (2015) mentioned that in the delivery of crowdsourcing delivery model, in order to avoid the loss and damage of items in the distribution process, the platform will take full insurance insurers, cargo loss compensation protection and other measures. Therefore, in order to ensure the successful completion of the task, the platform can imitate the delivery mode of crowdsourcing management, to collect margin members, the failure of the task of compensation protection and other measures. In the meantime, some certain task recipients will get quick answers but not right, which can optimize their time efficiency and earn more money. To avoid this phenomenon, crowdsourcing platforms should also take steps to identify counter-frauds.

5.2 Essay Innovation

In the process of solving the problem, this paper integrates the mechanism design theory and task pricing - attribute information theory, and applies these theories in the new field of mobile Internet - the task pricing of self-service labor crowdsourcing platform. The paper explores the way for the development of crowdsourcing model in new field, and points out the influencing factors of task bidding, which is conducive to the adoption of shared economy model crowdsourcing website platform to better serve the task of publishers, but also providing some reference for the tasks of the task of pricing publisher.

References

- Calculate distance based on two latitude and longitude.360 Personal Library. Retrieved September 17, 2017, from http://www.360doc.com/content/13/0410/17/11500595_277406769.shtml
- Eickhoff, C., & Vries, A. P. D. (2013). Increasing cheat robustness of crowdsourcing tasks. *Information Retrieval*, 16(2), 121-137. <https://doi.org/10.1007/s10791-011-9181-9>
- Fei, Y. L., Tian, J., & Deng, J. (2015). Research on Causes and Countermeasures of Fraud Behaviors in Crowdsourcing Contest. *Journal of Jiangsu University of Science and Technology (Social Science Edition)*, 15(4), 82-86.
- Howe & Jeff. (2006). The rise of crowdsourcing. *06 Jenkins H Convergence Culture Where Old & New Media Collide*, 14(14), 1-5.
- Huang, M. C. (2016). Development status and evolution trend of crowdsourcing research platform. *Guangdong Science and Technology*, 25(8), 17-20.
- Liu, Y. R. (2016). Study on Model of Crowdsourcing Express and Its Development Tendency. *Logistics Engineering and Management*, 38(4), 32-33.
- Pangjian, G. (2015). The Risk Management Mechanism of Crowdsourcing Community Innovation. *China Soft Science*, 2, 183-192.
- Qian, G. (2017). Reseach on Service Recommendation for Crowdsourcing Distribution. (Doctoral dissertation, Zhejiang Sci - Tech University).
- Tu, S. L. (2015). Discussion on Problems and Countermeasures for Crowdsourcing of Tourism Logistics Network of Jiangxi. *Logistics Technology*, 1, 45-48.
- Wang, Y., & Yi, S. K. (2007). An analysis of incentive issues in mechanism design theory. *Indian Ocean Economic Research*, 22(5), 51-52.
- Wu, J. H., & Luneng, F. (2016). Research on the Crowdsourcing Model for End Distribution in Express Delivery Enterprises. *Journal of Xichang University: Natural Science*, 30(2), 69-73.
- Yan, J., Liu, R. J., & Liu, H. (2017). A Literature Review of Domestic and Foreign Crowdsourcing Research. *China Science and Technology Forum*, 8, 59-68.
- Zhang, Z. L., Cao, Z. Y., & Li, Y. T. (2010). Research Based on Euclid Distance with Weightsof K_means Algorithm. *Journal of Zhengzhou University (Engineering Science)*, 31(1), 89-92.
- Zhao, J. M. (2015). Crowdsourcing can be used cheaply to get the job done? *Family Services*, 9.
- Zhao, X. L. (2016). Network Optimization of Crowdsourcing Distribution Based on K-means-Genetic Algorithm. (Doctoral dissertation, Beijing Jiaotong University).
- Zhou, S., Xu, Z. Y., & Tang, X. Q. (2010). Method for determining optimal number of clusters in K-means clustering algorithm. *Computer Applications*, 30(8), 1995-1998. <https://doi.org/10.3724/SPJ.1087.2010.01995>

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).