Task Pricing Optimization Model of Crowdsourcing Platform

Li Lin[1], Xiangyue Chen[3], Yiying Lou[4], Weijian Zhang[2] & Ru Zhang[2]

[1]Business Administration Department of International Business School, Jinan University, Zhuhai, China

[2]Finance Department of International Business School, Jinan University, Zhuhai, China

[3]School of Translation, Jinan University, Zhuhai, China

[4]Internet of Things of Electrical & Information Engineering School, Jinan University, Zhuhai, China

Correspondence: Ru Zhang, Finance Department of International Business School, Jinan University, Qianshan Road 206#, Zhuhai City, Guangdong Province, Post No. 519070, China.

Abstract

In this paper, we established a task pricing optimization model by the Logistic and anti-resolve thought to work out the problem of unequal spatial distribution and overall low of the task completion rate in the crowdsourcing platforms. Combining with the actual application information, we use scatter diagram, contour map, etc. to make a qualitative study and find that the reason why some of the tasks are not accepted is because the enterprise failed to take the total task quotas around the task into consideration while pricing the task. Then, combined with the influencing factors of traditional pricing model and results of qualitative analyses, the optimization model of crowdsourcing platform is built. Next, we select an ending project in an app of "make money" in China as the example to evaluate the effectiveness of our model. We applied the method of computer simulation to solve the model, and we find that, under the new pricing plan, the task completion rate has been significantly improved, which proves the conclusion of our qualitative analysis and the validity of the optimal model.

Keywords: task quota, logistic regression, anti-resolve thought, pricing optimization model

1. Introduction

With the increasing popularity of the Internet and the rise of virtual product trading in web2.0, "Crowdsourcing" has become one of the important modes of e-commerce services (Estellés-Arolas,2012). Crowdsourcing is a business concept put forward by Howet in 2006. It refers to the practice of a company or organization to outsource the part of work performed by its employees to non-specific public in the form of tasks (Jeff,2006). And in today's world, the way for enterprise to call the public to complete the tasks is mainly by the Internet (Estellés-Arolas,2012). Compared with the former business market survey, crowdsourcing platform not only reduces the operation cost of enterprise, but also guarantees effectively the authenticity of the survey (Chen, Zheng, & Li 2017). Crowdsourcing has already penetrated into people's lives and work and more and more enterprises realize the value of crowdsourcing (Gu & Xi, 2017). However, nowadays, the task bounty for most of crowdsourcing platform is high, but the task completion rate is low and uneven distributed. Thus, it is difficult for company to obtain the comprehensive survey data. However, the success of crowdsourcing is closely related to the task bounty (Fu et al, 2018). Therefore, the issues about how to set task bounty to attract users to receive and complete the tasks and solve out the problem of low task completion rate is worth studying.

Many scholars have studied the task pricing of crowdsourcing platforms. Some research focuses into study the impact of task pricing. Yang and others studied the impact of task pricing in the crowdsourcing platform on the number of participants and made an analysis on the behavior of participants from the aspects of task duration, attributes of competitive task, and market maturity and he found that the factors influencing the participation of the members mainly include the amount of the task, the difficult degree of task, and attributes of competitive tasks within a specific period of time (Yang et al, 2009). Mason W and others, after setting different prices for different tasks in crowdsourcing platform, opined that the excessively high price dose not improve the quality of the project but just obviously increase costs and if the price of tasks is too low, it will be difficult to attract members and has a bad effect on the completion rate of the tasks (Mason W et al, 2010).

And, in recent years, many scholars have researched crowdsourcing models, tasks, and crowdsourcing web from a management or technology perspective and provided different mathematical models optimizing the task pricing to improve the utility of platform and requester's budget. Singer and Mittal (2013) put forward a constant-competitive incentive compatible mechanism that benefit to improve the maximum number of assignable tasks under the budget and found out the minimum total pricing for a given number of tasks. Pu et al. (2019) and others established multi-objective planning task configuration and pricing model based on service cost, task value and member income orientation and applied the Elitism Bee colony Algorithm to make a simulation experiment on the crowdsourcing task pricing. Cheng and Varma (2014), to retain more members to complete the tasks in the crowdsourcing platform, put forward a new pricing plan and showed that the change of the remuneration of the task has an effect on the number of members willing to complete the task. Bernstein et al. (2012)used the queuing theory to analysis the retainer model for real-time crowdsourcing to optimize and weight the cost and performance of crowdsourcing platforms. Singer and Mittal (2013) pointed out the approach of regret minimization in online learning in the pricing mechanism to improve the utility of the requester's budget. Goel and Singla (2014) designed an incentive-compatible mechanism to near-optimal utility for requester under the matching constraints of the bipartite graph of the member and task. Liu et al. (2015)considered the importance of teamwork and designed a task allocation and pricing mechanisms to reduce overall costs and increase platform efficiency.

However, in practical application, the crowdsourcing platform is still immature and there are still many problems that the theoretical knowledge cannot be solved. And the above-mentioned researches are less studied based on actual application information and less focus on the problem of uneven distribution of existing task completion rates. Thus, we make a qualitative analysis of the problem of task completion rates based on the actual application information and based on the existing problem establish a mathematical model and provide an optimal pricing strategy to solve it.

The rest of the paper is constructed in following five major sections. The next section provides "Assumptions of the questions". Then, "Qualitative analysis of the reasons for low task completion rate of crowdsourcing platform" is given to analysis the actual application information by comparison. In "Task pricing model of crowdsourcing platform", we describe the detail of task pricing model. Followed by "Case study", we apply the model to study an example to verify the validity of the model. The final section concludes the paper.

2. Hypothesis

(1) Assume that each task has the same difficulty of completion, and the content of the task has no influence on the selection and completion of members.

(2) Assume that the uncontrollable factors such as weather and traffic have little influence on the completion degree of the task and can be ignored.

(3) Assume that members are sure to complete the tasks once they receive them.

3. Qualitative Analysis on the Actual Application Information

Since most of scholars have based on the task pricing, distribution and membership of research to analyze on the task to complete. Therefore, according to the unfinished tasks, from the perspective of members, this paper researches the cause of the uneven and low special distribution of the task completion through qualitative analysis. Because general crowdsourcing platform will vest reputation value for members, the members who has the higher credibility will have the right to have a priority task and have higher amount of task allocation. The task distribution refers to the platform takes the quota each user can pick up and general assignment into consideration to distribute the task. The quota of tasks each user can pick up is given according to the user's reputation value. As a result, we believe that the reputation value and distribution of members have an impact on the completion of tasks. Due to the positive correlation between reputation, selection time and allocation amount, we attribute the influence of member reputation and time on task pricing to the influence of task allocation on task pricing. The assignments of each member are calculated as follows:

$$t_i = t_{total} \frac{u_i}{u} \tag{1}$$

In this equation, t_i represents the allocation of tasks for member i, u_i represents anticipated quota for member i, t_{total} represents the total number of tasks provided by this project.

Then, we take a finished project in the app "taking photos to make money" as an example, and mark the completion degree of tasks and the assignment amount of member tasks into the coordinate system for comparison, the result is as shown in Figure1 and 2:
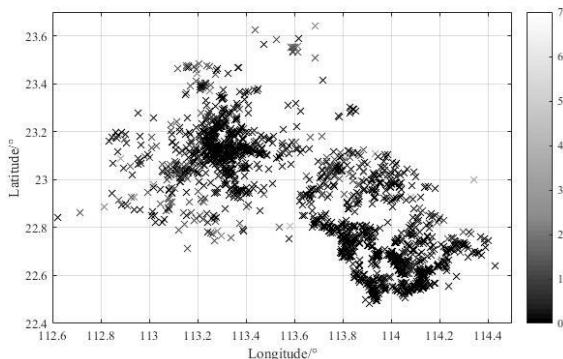
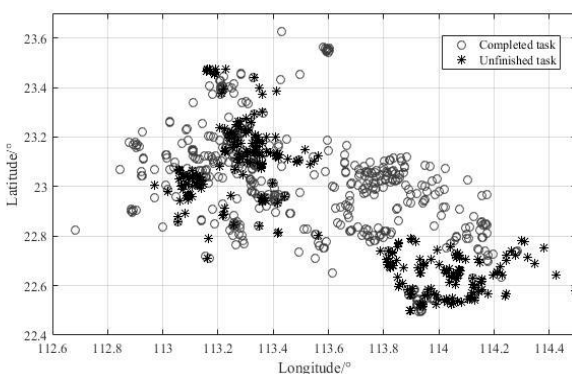Figure 1. Diagram of membership location and task assignment

Figure 2. Diagram of completion status of each task point

As the diagram shown in Fig. 1, the darker the color of the mark point is, the lower the task allocation that the member has. Contrast figure 1 with figure 2 we can know that lower area A and B of the task completion also have deeper tag color in the corresponding area in figure 1. In other words, in the lower completed section, members also have lower task distribution.

To avoid visual error, we randomly select uncompleted task points in the area of lower completion rate and count total quota of member tasks within the scope of 2 km, so that we can research the relationship between uncompleted task points and member quotas, some results as Table 1 shown:

Table 1. The relationship between total task allocation and task completion

Coordinate	Total assignment of tasks	Completion of execution
(23.20014,113.37491)	0	No
(23.10125,113.54843)	0	No
(23.12849,113.35062)	3	No
(23.15934,113.33871)	2	No

The Table 1 shows that most of the total task allocation near the unfinished task point is low, and there is a situation that all the members have not been assigned to a task quota in some individual task points within the scope of 2 km.

This indicates that there is a serious oversupply in this area. In consequence, this task point should increase its pricing in order to attract a small number of members with quotas within the scope, or to attract more members with sufficient quotas outside the scope to receive and fulfil the task. We summarize and compare the task pricing of all task points and present the task pricing of each task point in the form of topographic map. The results are as follows:
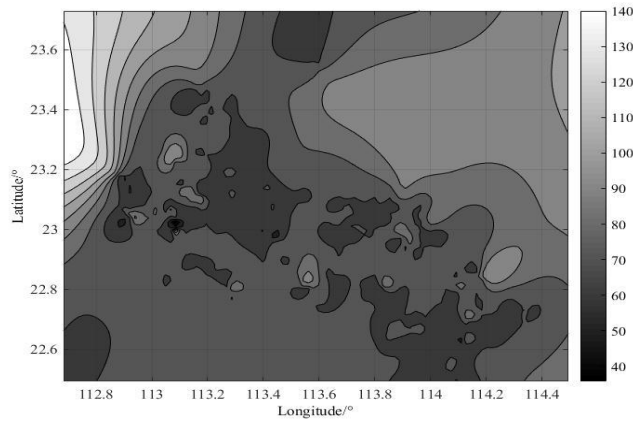
Figure 3. Pricing distribution of each task point

It can be seen from Fig. 3 that the pricing of the unfinished task is relatively low, contrary to the above statement, the task pricing has certain irrationality. As a consequence, we can conclude that the traditional task pricing of crowdsourcing platform only considers the factors such as tasks and geographical and the overall numbers of members, it doesn't make a comprehensive analysis to combine with the dynamic reputation value of the member itself, so that there is an uneven and low special distribution of task completion rate. Thus, we propose the task pricing optimization model of crowdsourcing platform in the following part. On the basis of considering the factors of the traditional pricing model, we will consider more about task quotas members possess based on their reputation value.

4. Task Pricing Model of Crowdsourcing Platform

4.1 Influential Factors of the Traditional Pricing Model of Crowdsourcing Platform

In the traditional task pricing model of crowdsourcing platform, task pricing is related to geographical distribution of the task, number of members and quality of the task (Jin, 2018). As above-mentioned, we have been assumed equal quality of tasks, we first consider geographical distribution of the task and the number of members in the task pricing model of crowdsourcing platform. And geographical distribution is regarded as the distance between each task point and the member. To avoid regional differences, this paper only researches the task released in the same city. On account of different sizes of each city and different numbers of members, we replace the number of members with the density.

4.1.1 Membership Density

As for the same city, the total membership density has similar influence on each pricing of the task, and for the city's area is too big, it is meaningless to research the pricing of each task point and it is hard to regard it as the influential factor of pricing difference. Therefore, the membership density is expressed as the membership density within a certain range of each task.

$$\rho_k = \frac{q_k}{s_k}, k = 1, 2, 3, \cdots, m \tag{2}$$

Among which, ρ_i represents the membership density of the k task, q_i represents the total number of members in the k task point area, s_i represents the area of the k task, and m represents the total number of tasks issued by a city.

4.1.2 Distance from Members

As can be seen from Fig.1 to Fig.3, the closer the member is to the task point, the more the member is inclined to complete the task at that point, which will affect the reduction of the pricing at that point. So, the distance between tasks and members is represented by the distance between the members closest to the task point and the task point.

$$D_k = min\{c_{k1}, c_{k2}, c_{k3}, \cdots, c_{kl}\} \tag{3}$$

Among which, D_i represents the closest distance between the k task and the member, and c_{il} represents the distance between the k task and the member l.

4.2 Task Pricing Model Based on Logistic Regression

The traditional task pricing scheme of crowdsourcing platform does not take into account the task quota, which causes uneven regional distribution of task completion rate of crowdsourcing platform. As a result, we set the task quota as variables in the new task pricing scheme. In this paper, we mainly use inverse solution of logistic model. First of all, we conclude the conditional function of completing task based on the original data. Then we establish the task pricing optimization model of crowdsourcing platform according to the factors affecting task pricing. And we obtain the optimal pricing under the completed task through the inverse solution. In the end, we constrain the price based on the actual situation.

4.2.1 Logistic Regression Model of Conditions of Task Completion

In order to improve the task completion rate, it is necessary to study the basis and conditions of members' completion of tasks. Then, logistic regression is used to conduct simulation training for members at all points. The logistic regression model is as follow:

$$h_0(x) = g(\theta^T x) = \frac{1}{1+e^{-\theta^T x}} \tag{4}$$

$$\theta^T x = \theta_0 + \theta_1 x_1 + \cdots + \theta_n x_m = \sum_{i=1}^{n} \theta_i x_i \tag{5}$$

Among which, $h_0(x)$ is expressed as the prediction function; x_i is the input component of the model; θ_i represents learning parameters. Due to the member's choice of tasks and the relationship between the task and the task center distance only lies in the price, we do not repeat input. In consequence, $x_i(i=1,2,3,4)$ are expressed as membership density ρ_k of a certain range, task quota of members, the shortest distance D_k from members and the original pricing p of the task.

4.2.2 Task Pricing Mode

When the task is completed, the output is "1", and when the task is uncompleted, the output is "0". In order to study the required task pricing at a high task completion rate, the result of the prediction function constructed by logistic regression was "1".

$$h_0(x) = g(\theta^T x) = \frac{1}{1+e^{-\theta^T x}} = 1 \tag{6}$$

At the same time, the price in the input component of the model x_i is set as unknown quantity and the learning parameters θ remain unchanged. Then the new price for each task are obtained when the output of the prediction function is all "1".

Meanwhile, out of consideration of enterprise cost, the tasks released by crowdsourcing platform cannot continuously increase the price in order to improve the completion rate. According to the price fluctuation theory, the task price should fluctuate up and down around its value, and its fluctuation range is restricted by the task value and has a certain range of fluctuation. In addition, because this paper seeks the optimal price through the inverse solution, it is unnecessary to consider the minimum price of users to accept, that is, it is unnecessary to restrict the lower limit of pricing. As a consequence, we still need to constrain and adjust of prices after operating the logistic regression algorithm, namely for high price, price shall be adjusted to the maximum within the range of fluctuation. The price upper limit formula $P_{f\max}$ is as follows:

$$P_{f-max} = \frac{\sum_{1}^{Q} P_f}{Q} \times (1+\gamma) \tag{7}$$

In this equation, P_{f-max} and P_{f-min} represents the maximum and minimum of task price; P_f represents each task pricing before the processing of pricing constraints; Q represents the total number of members in the region; γ denotes the range of fluctuation.

5. Case Study

We take a finished project in the app "taking photos to make money" as an example, and selects the task released in Foshan city, Guangdong Province, China as the research scope. The input components of the model are the shortest distance between members and tasks, the quota of member tasks, the membership density and the original price of tasks within a certain range.

Considering that this app collects information mainly based on people's convenience, and the range of people's daily activity is usually within a kilometer of the residential area. Therefore, we transform the member density of each city into the membership density within 1km of each task point in the determination of membership density. Combined with SPSS and MATLAB, the task pricing optimization formula is:

$$p = -1 \div 0.113 \left(log\left(\frac{1}{h-1}\right) - 6.35 - 0.549 x_5 - 0.177 x_3 - 0.302 x_1 \right);$$

After forecasting the completed project, and the price results of each task point are as follows:
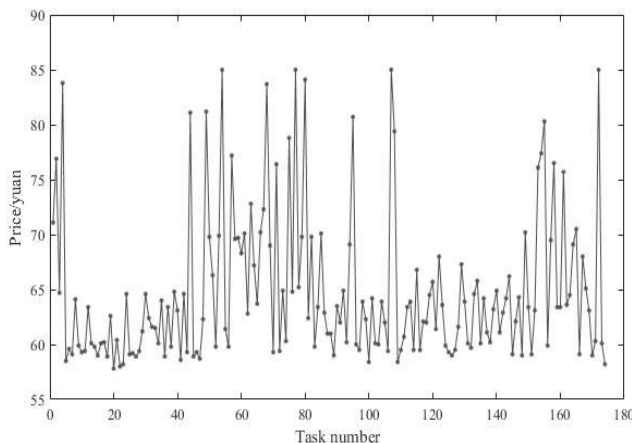

Figure 4. New price of the "take photos to make money" app task

According to figure 4, some tasks in the predicted result were overpriced, so the task price was adjusted in combination with price constraints. Among them, $\gamma=0.15$, $Q=177$. After operating Matlab2016 b calculation, the results are as follows:

Table 2. Constraints of task pricing

Task value	Upper price
71.4425	82.11

By comparing the data before and after the constraint, it can be found that the price of 6 task points is restrained by the upper limit price. Then, combined with the information and completion conditions of the member in the crowdsourcing platform, the completion situation of the new task pricing after pricing constraint processing and Logistic regression are in the following Table 3:

Table 3. Comparative results of new task pricing completion rate and original completion rate

Original average price of the task	New average price of the task	Original completion rate	New completion rate
71.44	64.65	66%	95.98%

As can be seen from Table 3 above, the average price of new task pricing on crowdsourcing platform has been decreased after adjustment, but the completion degree of task has been increased instead. However, according to the analysis of the fourth quarter in this paper, the original scheme does not take low reputation value of regional members into account, which leads to relative low task quota, so that part of the task price is too low and cannot be completed. In consequence, the optimized pricing should be higher than the original plan pricing, but here the results are contrary to the conclusion. Therefore, we further analyze the sampling data of the task point, and the sampling results are shown in Table 4 below:

Table 4. Sampling analysis of task points

No.	Task Location	Original Price	Completion Status	New Price	Completion Status
1	(23.00570，113.12436)	65.5	0	71.1	1
3	(23.02630，113.14035)	66	0	76.4	1
4	(22.85760，113.05312)	70.5	1	62.4	1
5	(23.00974，113.08552)	65.5	0	63.5	1

It can be seen from Table 4 that the original task pricing of unfinished task points has been increased in the original pricing scheme, which is in line with the expected conclusion. At the same time, the pricing of completed task points

has been decreased. This shows that the new task pricing model also optimizes the pricing of the completed task points, reduces the enterprise cost, has strong practical significance and added value. In addition, it can be observed that although the No.5 task was originally an unfinished task, its price still fell after adjustment, but the falling range was small, and its completion status was still "1". The existence of the special point could be influenced by the real complexity of the task, local transportation, or the change of peripheral task price, but the change is not large, and it still can keep the completion of the task. Therefore, the scheme has high feasibility and practical significance, which can obviously improve completion rate of the crowdsourcing platform.

6. Conclusion

In this paper, the task pricing optimization model is established for the problem of uneven distribution and overall low level of the task completion rate of the crowdsourcing platform. We innovatively use scatter plots and contour maps for comparative analysis and make a qualitative analysis with the actual application information of the low and uneven distribution task completion rate. Then, we find that the reason on why the rate is unevenly distributed and overall low is that the tradition task pricing model does not take into account the impact of membership reputation. Thus, based on these result, the task pricing optimization model is obtained by using Logistic regression and anti-solution thinking and combining the data in the platform. Comparing with the traditional pricing scheme, we find that the task completion rate under the new task pricing is higher than before, and the task average pricing is lower than the traditional task average pricing, which means that the model has certain feasibility and added value. Besides, it also shows that blindly increasing the price of the task does not enhance the task completion rate, but only increase the cost of the company, and Mason and Watts (2010) also made the same conclusion.

Differences in the complexity of different tasks, climate and the possibility of membership repentance are not considered in this paper and the profit of crowdsourcing platform is not taken into account as well. Thus, in the further study, some objective factors should be considered, and the model should combine the enterprise and members so as to achieve maximum profit while achieving a higher task completion rate.

References

Bayus, B. L. (2013). Crowdsourcing new product ideas over time: an analysis of the dell ideastorm community. *Social Science Electronic Publishing, 59*(1), 226-244. http://dx.doi.org/10.2139/ssrn.1979557

Bernstein, M. S., Karger, D. R., Miller, R. C., & Brandt, J. (2012). Analytic methods for optimizing realtime crowdsourcing. Computer Science.

Chen, R., Zhou, Z., & Zhu, L. (2017). Task pricing rule model of crowdsourcing platform. Economic and trade practice (21).

Cheng, M., & Varma, V. (2014). Scaling-up the Crowd: Micro-Task Pricing Schemes for Worker Retention and Latency Improvement. AAAI Conference on Human Computation and Crowdsourcing (Vol.15, pp.193-208). http://dx.doi.org/10.13140/2.1.3692.8649

Dong, C. (2013). Multi-label classification model based on case and logistic regression [D]. Nanjing university.

Estellés, A. E. (2012). Towards an integrated crowdsourcing definition. *Journal of Information Science, 38*(2), 189-200. https://doi.org/10.1177/0165551512437638

Fu, X., & Feng, R. (2018). Research on task pricing scheme in crowdsourcing business model. *Modern management*, *8*(2), 162-171. https://doi.org/10.12677/mm.2018.82020

Goel, G., Nikzad, A., & Singla, A. (2014). Allocating tasks to workers with matching constraints: truthful mechanisms for crowdsourcing markets. International Conference on World Wide Web (pp.279-280). ACM. http://dx.doi.org/10.1145/2567948.2577311

Gregory Mankiw, Principles of Economics, United States: Dryden, 1998

Gu, S., & Chen, X. (2017). Research summary of crowdsourcing platform and establishment of performance mechanism of crowdsourcing platform. *Scientific and technical progress and countermeasures, 34*(22), 153-160. https://doi.org/10.6049/kjjbydc.2017040665

Jeff, H. (2006) The Rise of Crowd Sourcing. *Wired, 14,* 1-4.

Karger, D. R., Oh, S., & Shah, D. (2011). Budget-optimal task allocation for reliable crowdsourcing systems. Operations Research, 62(1), págs. 1-24. http://dx.doi.org/10.1287/opre.2013.1235

Mason, W., & Watts, D. J. (2010). Financial Incentives and the Performance of Crowds. *ACM SIGKDD Explorations Newsletter, 11,* 100-108. https://doi.org/10.1145/1809400.1809422

Pu, D., Fan, Z., Yuan, G., & Yang, Y. (2019). Elite swarm optimization algorithm and stakeholder crowdsourcing pricing model. *Computer application research,* (5). https://doi.org/10.3969/j.issn.1001-3695.2017.11.0748

Qing, L., Tie, L., Ruiming, T., & Stephane, B. (2015). An efficient and truthful pricing mechanism for team formation in crowdsourcing markets. IEEE International Conference on Communications (Vol.15, pp.567-572). IEEE. http://dx.doi.org/10.1109/ICC.2015.7248382

Singer, Y., & Mittal, M. (2013). Pricing mechanisms for crowdsourcing markets. International Conference on the World Wide Web (pp.1157-1166). ACM. https://doi.org/ 10.1145/2488388.2488489

Singla, A., & Krause, A. (2013). Truthful incentives in crowdsourcing tasks using regret minimization mechanisms. International Conference on World Wide Web (pp.1167-1178). ACM. http://dx.doi.org/10.1145/2488388.2488490

Vuculescu, O., & Bergenholtz, C. (2014). How to solve problems with crowds: a computer-based simulation model. *Creativity & Innovation Management, 23*(2), 121-136. https://dx.doi.org/10.1111/caim.12059

Wang, L., & Xu, A. (2015). The research of the transfer phenomena about the employment of undergraduate to go to second-tier cities - take Chongqing as an example [J]. Journal of the Chinese market, (26), 113-121. https://doi.org/10.13939 / j.carol carroll nki ZGSC. 2015.26.113

Yang, Y., Chen, P. Y., & Pavlou, P. A. (2009). Open Innovation: An Empirical Study of Online Contests. International Conference on Information Systems, Icis 2009, Phoenix, Arizona, Usa, December (Vol.6, pp.13). DBLP. https://dx.doi.org/10.1145/2567948.2577311