

# Opportunistic Solvers Identification of a Crowdsourcing Innovation Platform

Jintong Hua<sup>1</sup> Qingliang Meng<sup>1,\*</sup>

<sup>1</sup>*School of Management & Economics, Jiangsu University of Science & Technology, Zhenjiang, Jiangsu 212000, China*  
*\*Corresponding author. Email: mengzhi007@163.com*

## ABSTRACT

Crowdsourcing innovation is characterized by loose organizational structures, free will of participants, strong innovation goals and information asymmetry, which can easily induce solvers' opportunistic behaviors, such as fraud, intellectual property theft, free-riding and so on. In order to create an attractive and supportive platform environment, identifying and avoiding opportunistic solvers is essential. In this study, a four-dimensional opportunistic solvers' identification system, including the work ability, the reputation level, the service level and the historical transaction level, is constructed. Then, the opportunistic solvers' identification model was proposed based on the PCA-BP neural network. The data is based on *www.zbj.com*, and taken into empirical research to verify the feasibility and effectiveness of the identification model.

**Keywords:** *crowdsourcing innovation, opportunistic solvers, identification system, PCA-BP neural network*

## 1. INTRODUCTION

Howe defines crowdsourcing as 'the act of a company or institution taking a function once performed by employees and outsourcing it to an undefined (and generally large) network of people in the form of an open call'. It has proved that this mode is an effective mode to acquire the wisdom of network groups[1]. *InnoCentive* brings more than 100,000 research talents from 125 countries around the world. So far more than 30,000 ideas has been submitted and 550 ideas has been implemented in *Dell's IdeaStorm*. However, as a new business model, the crowdsourcing innovation's related research is still in the exploratory stage. Therefore, the problem, such as solvers don't participate continuously[2] and low-quality of tasks[3] are more and more prominent.

The crowdsourcing is characterized by loose organizational form, solvers' free and voluntary will of participating and information asymmetry, which can easily lead to solvers' opportunistic behaviors, such as fraud, intellectual property theft. Hirth (2013) proposed two methods to monitor user fraud: majority decision-making and control group re-detection[4]. Khasraghi (2014) pointed out that solvers' participation history has an important effect on their performance, such as the participation times and the winning rate[5]. Sun (2015) proposed that increasing speculative costs can effectively avoid opportunism solvers at the same time

improve the bonus[6]. Ye (2017) summarized three threat forms of fraudulent solvers and proposed a threat defense model based on trust vector to avoid opportunistic behaviors of solvers[7]. Paswan (2017) pointed out that user fraud can be effectively restricted by applying the relational governance strategy and the role identification strategy[8].

Above researches have significance for identifying and limiting the opportunistic behaviors of solvers and ensuring the high-quality, sustainable development of the relationship between platform and solvers. While, as a new business model, crowdsourcing innovation has complex and diverse innovation tasks, extensive sources of the public, and certain information asymmetry. In order to improve the relevant theories, reduce the risk of opportunistic behaviors and ensure the healthy development of crowdsourcing, it is necessary to make research on the opportunistic solvers identification from different perspectives.

## 2. IDENTIFICATION SYSTEM

Combined with the crowdsourcing platform *www.zbj.com*, this paper analyzes and summarizes 11 variables in four dimensions to build the opportunism solver identification system in crowdsourcing. As shown in Table 1

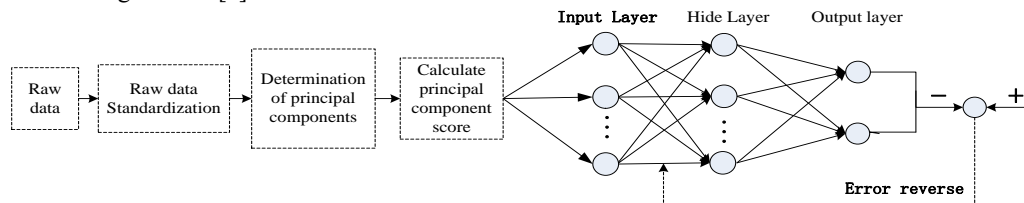
**Table 1.** Opportunistic solvers' identification system in crowdsourcing innovation

Dimension	Index	Variable	Variable meaning
Comprehensive work ability	$X_1$	Comprehensive score	Comprehensive evaluation score of solvers on crowdsourcing platform.
	$X_2$	Completion quality	Score given by the employer for the quality of task completion.
	$X_3$	Working speed	Score given by the employer for the speed of solver's submission.
	$X_4$	Solver level	Solver level on crowdsourcing platform
Reputation level	$X_5$	Dispute rate	Ratio of the tasks' number in dispute to total tasks completed by this solver
	$X_6$	Number of punishments	Number of times that solvers are punished by crowdsourcing platform for dishonesty or violation
	$X_7$	Amount of deposit	The amount of integrity deposit paid by solvers on the crowdsourcing platform
Service level	$X_8$	Service attitude	Score given by the employer for the solver's service attitude.
	$X_9$	Favorable rate	Ratio of praised tasks to the total number of tasks completed by solvers.
Historical transaction level	$X_{10}$	Refund rate	Ratio of the number of tasks refunded by solvers to the total number of tasks completed by solvers.
	$X_{11}$	Project completion rate	Ratio of the volume of completed orders to winning orders in the past year.

**3. METHODOLOGY**

Considering that the opportunistic solvers' identification system contains multiple dimensional indicators, there may be some correlations among the variables, and resulting in information overlaps. Moreover, too many input layers will cause low learning efficiency, slow convergence speed and local minimum state of neural network. Therefore, this paper uses principal component analysis (PCA) to process the index data, which reduces the dimension of data and retains most of the original data[9]. The structure of neural

network is simplified, and the input layers is reduced, and the convergence speed and performance are improved, so that the recognition accuracy is ensured. The principal component analysis (PCA) method simplifies the original data structure, and it uses fewer principal components to replace the complicated original variables, to make the problem clearer. The PCA-BP network model structure is shown in Fig. 1.



**Figure. 1.** The structure of PCA-BP neural network

**4. EMPIRICAL RESEARCH**

**4.1 Data collection**

www.zbj.com is the largest crowdsourcing platform in China, and its main services includes brand design, marketing promotion, website construction. This paper is based on the data of "Exposure Channel " column of integrity management center in www.zbj.com from August 2018 to April 2019. A total of 184 original data are collected, there are 34 opportunistic solvers and 150 normal solvers among the data. The main manifestations of opportunism are inducing or coercing the employer to evaluate, inability to complete tasks, manuscript plagiarism and failure to accomplish duties. The

punished solvers mainly involves in brand design, website construction, e-commerce design services. Among 184 sample data, 154 data are taken as training set and 30 data as the testing set for empirical research.

**4.2 Data analysis**

Table 2 shows that the cumulative variance contribution rate of the first three principal components is 81.873%. This shows that the first three principal components had already contains most of the information of the original variables.

**Table 2** Total variance explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	5.415	49.224	49.224	5.415	49.224	49.224	5.224	47.492	47.492
2	2.572	23.38	72.604	2.572	23.38	72.604	2.262	20.566	68.058
3	1.02	9.269	81.873	1.02	9.269	81.873	1.52	13.815	81.873
4	0.749	6.808	88.681						
5	0.587	5.339	94.021						
6	0.398	3.617	97.637						
7	0.141	1.278	98.915						
8	0.074	0.67	99.585						
9	0.024	0.215	99.801						
10	0.013	0.119	99.92						
11	0.009	0.08	100						

**Table 3** Rotated Component Matrixa

	Component		
	1	2	3
comprehensive score	0.983	-0.108	-0.056
completion quality	0.97	-0.142	-0.053
working speed	0.975	-0.116	-0.067
solver level	0.29	0.524	0.576
dispute rate	-0.153	0.897	-0.267
number of punishments	-0.18	0.756	-0.287
amount of deposit	0.317	0.433	0.68
service attitude	0.982	-0.104	-0.108
favorable rate	0.959	-0.057	-0.062
refund rate	0.569	0.492	-0.149
project completion rate	0.334	0.657	-0.155

Table 4 Component Score Coefficient Matrix

	Component		
	1	2	3
comprehensive score	0.983	-0.039	0.119
completion quality	0.973	-0.07	0.104
working speed	0.977	-0.041	0.105
solver level	0.111	0.219	0.793
dispute rate	-0.216	0.916	0.12
number of punishments	-0.22	0.798	0.036
amount of deposit	0.131	0.092	0.851
service attitude	0.99	-0.011	0.075
favorable rate	0.953	0.009	0.131
refund rate	0.519	0.526	0.206
project completion rate	0.27	0.667	0.22

As shown in table 3, in the principal component 1, the load of comprehensive score, completion quality, working speed, service attitude, favorable rate and refund rate are large, which represents the solver's comprehensive quality in the process of completing tasks. In principal component 2, dispute rate, the number of punishments and project completion rate have a large load, which could represent the opportunistic tendency of solvers in the process of communication and transaction. In principal component 3, solver level and the amount of deposit are large, which could represent the level of solver in crowdsourcing platform. It is generally true that the higher the level of solver, the more deposit they pay.

The scores of three principal components are calculated by formula (1), then there will be the input variables of BP neural network.

$$\begin{cases}
 F_1 = 0.983Z_{x_1} + 0.973Z_{x_2} + 0.977Z_{x_3} + 0.111Z_{x_4} - 0.216Z_{x_5} - 0.22Z_{x_6} \\
 \quad + 0.131Z_{x_7} + 0.99Z_{x_8} + 0.953Z_{x_9} + 0.519Z_{x_{10}} - 0.27Z_{x_{11}} \\
 F_2 = -0.039Z_{x_1} - 0.07Z_{x_2} - 0.041Z_{x_3} + 0.219Z_{x_4} + 0.916Z_{x_5} + 0.798Z_{x_6} \\
 \quad + 0.092Z_{x_7} - 0.011Z_{x_8} + 0.009Z_{x_9} + 0.526Z_{x_{10}} + 0.667Z_{x_{11}} \\
 F_3 = 0.119Z_{x_1} + 0.104Z_{x_2} + 0.105Z_{x_3} + 0.793Z_{x_4} + 0.12Z_{x_5} + 0.036Z_{x_6} \\
 \quad + 0.851Z_{x_7} + 0.075Z_{x_8} + 0.131Z_{x_9} + 0.206Z_{x_{10}} + 0.22Z_{x_{11}} \quad \dots(1)
 \end{cases}$$

### 4.3 Results contrast

In this paper, the original variables are synthesized into three principal components by PCA and the scores of each principal component are obtained. 184 data are divided into training data and testing data according to the ratio of nearly 10:2, which are used for training and testing of BP neural network. There are 154 training data, including 23 opportunistic solvers and 131 normal solvers; 30 testing data, including 10 opportunistic solvers and 20 normal solvers. The BP neural network model is trained by training data, and the setting error is achieved at step 4552. The error curve is shown in Fig.2.

PCA-BP neural network model is used for recognition, with 27 correct and 3 wrong identifications, and the recognition

accuracy is 90%; The recognition results of traditional BP neural network are 23 correct and 7 wrong, the recognition accuracy is 76.7%. Therefore, the BP neural network model improved by PCA method have higher recognition accuracy.

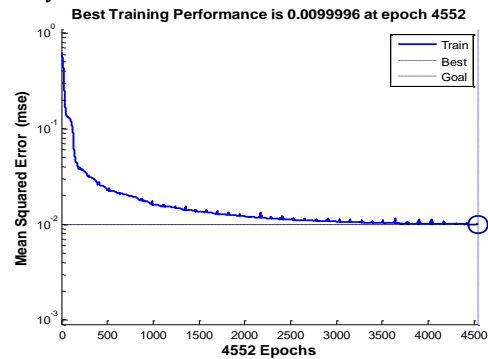


Figure. 2. The error curve graph

## 5. DISCUSSION

Based on the above theoretical analysis and empirical results, this paper proposes three strategies to avoid the opportunistic behavior of solvers.

(1) Setting up a scientific opportunistic solvers' identification mechanism. The identification methods of opportunistic solvers in crowdsourcing innovation are mostly post-test methods, that is, to evaluate the quality of submitted tasks or to identify the accused solvers. Therefore, it is necessary to set up a reasonable opportunistic solvers' identification mechanism to transform the post-event control into process identification. According to the solver's past historical data and participation information, it is effective to identify the solvers with higher opportunistic tendency, and improve

the crowdsourcing innovation efficiency.

(2) Improving the punishment mechanism of solvers' opportunistic behaviors. The crowdsourcing platform will punish the opportunistic solvers after the employer safeguards their legitimate rights and interests. The punishment methods are mostly to reduce the integrity score or to give poor evaluation. If the solver is punished again within the specified time, the credit score will be deducted multiple times, and the degradation punishment will be carried out. At the same time, the third-party dispute settlement mechanism is introduced. For solvers who cause losses to the employer and the platform, the deposit paid by them will be deducted according to the severity of the event, and the opportunistic behavior of solvers will be avoided by increasing the penalty cost.

(3) Creating a free and positive crowdsourcing innovation platform environment. The healthy development of crowdsourcing platform is inseparable from the sound opportunistic solvers' identification mechanism and punishment mechanism. Meanwhile, it is also necessary to create a free and positive environment for crowdsourcing innovation platform. Combined with the typical characteristics and operation process of crowdsourcing innovation, an effective multi-party interaction strategy and trust mechanism are established to maximize the performance of crowdsourcing innovation.

## 6. CONCLUSION

To identify and avoid the opportunistic solvers of crowdsourcing innovation, it is urgent to guide and encourage solvers to participate continuously, reduce the risk of opportunistic behaviors, and create a good platform environment. This paper constructs a four-dimensional opportunistic solvers' identification system, which includes the comprehensive work ability, the reputation level, the service level and the historical transaction level. Then, it proposes the opportunistic solvers' identification model based on the PCA-BP neural network and carries out the empirical research with *www.zbj.com*. The results show that the principal component analysis (PCA) method simplifies the original data structure, and improves the recognition accuracy. However, there are still some limitations in the paper. The empirical data mainly comes from the domestic crowdsourcing platform *www.zbj.com*, which has relatively simple innovation tasks, and shorter task completion time. Subsequent research can be combined with typical foreign crowdsourcing platforms, such as *Innocentive*, *IdeaStorm*, to carry out relevant empirical research, which will be more representative.

## ACKNOWLEDGMENT

This work was supported by Foundation of Humanities and Social Sciences of Ministry of Education of China No. 19YJA630055; Postgraduate Research & Practice Innovation Program of Jiangsu Province No. KYCX19-1652.

## REFERENCES

- [1] Lang M, Bharadwaj N, How crowdsourcing improves prediction of market-oriented outcomes, *Business Research*.69 (10) (2016) 4168-4176. DOI: <http://doi.org/10.1016/j.jbusres.2016.03.020>
- [2] Wexler, M, Reconfiguring the sociology of the crowd: exploring crowdsourcing, *International Journal of Sociology and Social Policy*.31(1-2) (2011) 6-20. DOI: <https://doi.org/10.1108/01443331111104779>
- [3] Boudreau K J, Lakhani K R, Using the Crowd as An Innovation Partner, *Harvard Business Review*.91(4) (2013)60-9, 140.D OI: <http://doi.org/10.1080/08985626.2012.721008>
- [4] Hirth M, Hoßfeld T, Analyzing Cost Sand Accuracy of Validation Mechanisms for Crowdsourcing Platforms, *Mathematical & Computer Modelling*.57(11-12) (2013) 2918-2932. DOI: <http://doi.org/10.1016/j.mcm.2012.01.006>
- [5] Khasraghi H J, Aghaie A, Crowdsourcing Contests: Understanding the Effect of Competitors' Participation History on Their Performance, *Behaviour & Information Technology*. 33(10-12) (2014)1383-1395. DOI: <http://doi.org/10.1080/0144929x.2014.883551>
- [6] Sun Y, Wang N, Yin C, et al, Understanding the Relationships Between Motivators and Effort in Crowdsourcing Marketplaces: A Nonlinear Analysis, *International Journal of Information Management*. 35(3) (2015) 267-276. DOI: <http://doi.org/10.1016/j.ijinfomgt.2015.01.009>
- [7] B. Ye, Y. Wang and L. Liu, Crowd Defense: A Trust Vector-Based Threat Defense Model in Crowdsourcing Environments, in: *IEEE International Conference on Web Services (ICWS)*, Honolulu, HI, 2017, pp. 245-252. DOI: <http://doi.org/10.1109/ICWS.2017.39>
- [8] Paswan A K, Hirunyawipada Tanawat, Iyer P, Opportunism, Governance Structure and Relational Norms: An Interactive Perspective, *Business Research*.77(2017) 131-139. DOI: <http://doi.org/10.1016/j.jbusres.2017.04.012>
- [9] Esbensen K H, Geladi P, 2.13 – Principal Component Analysis: Concept, Geometrical Interpretation, Mathematical Background, Algorithms, History, Practice, Comprehensive Chemometrics. (2009) 211-226. DOI:<http://doi.org/10.1016/B978-04452701-1.00043-0>