

## Risk Assessment Analysis of P2P Network Loan

Xiaoling Xu, Cheng Wang\*

School of Mathematics and Economics, Hubei University of Education,  
Wuhan 430205, P. R. China

\*Corresponding Author: wangc80@163.com

**Keywords:** Risk assessment; Information Gain; P2P network loan; WOE value.

**Abstract.** P2P network loan has been widely concerned by people since it came into the market. This paper establishes an assessment analysis model to evaluate the risk level of P2P network loan. Seven important indexes are selected to constitute the evaluation index system by using the method of Information Gain. Then the analysis method of WOE value is used to evaluate the risk level of P2P network loan, and the conclusion is drawn that the larger the WOE value of complex development index, the lower the risk level of borrowers' credit.

### Introduction

Foreign scholars mainly focus on the analyses of characteristic and behavior of P2P network loan, and some scholars analyze the credit risk of borrowers from the theory of information asymmetry. The amount of domestic scholars' research literatures on the P2P network loan is relatively small, mostly based on the theory. The research focuses on the empirical analysis of credit risk measurement through Logistic regression model, and judges the validity of the model via the model's estimation results.

In the past two years, P2P network loan has been widely concerned by people from all walks of life. As a typical representative of Internet finance, P2P network loan combined the traditional lending industry with Internet technology, which is an important direction of the future development in financial sector [1-2]. Nevertheless, due to the low threshold to entry, the imperfection of credit information system and the imperfect policies and regulations, P2P online loans have some relevant risks along with the rapid development, such as credit risk, operational risk, political and legal risk, and serious cases even led to the platform run away. Therefore, it is of great significance both in theory and in practice to study the risk management of P2P network loans.

### The P2P Risk Assessment Model

**Preliminary analysis and screening of data.** This article focuses on credit risk, and classifies the risk assessment indicators [3-5], as shown in Tab 1.

Table 1. The alternative indicators of Risk assessment

Risk assessment alternative indicators	Basic Information	Age Gender Marital status Educational level
	Occupational information	Profession Length of employment
	Asset information of borrowers	Housing condition Car availability
	Income and liabilities of borrowers	With or without default more than 90 days of debt Ratio of monthly debt service to income

Based on the data from the survey conducted by the Renren-Loan in the year of 2016 [6-7], this paper organized the number of people at each level of each evaluation standard. Since the impact of each alternative index is different, the method of Information Gain [8] is used to calculate the contribution of each index compared to explanatory variables, and then delete indicators that have less effect on the risk.

**The algorithm of Information Gain.**  $S$  denotes the sample collection, and the number of samples is  $s$ .

$A$  denotes the argument,  $A = \{a_1, a_2, \dots, a_n\}$ ,  $S_{aj}$  denotes the number of samples with the feature of  $a_j$  under the variable of  $A$ .

$C$  represents the category (this article is "good customers" and "bad customers"),  $C = \{c_1, c_2\}$ ,  $S_i$  is the number of samples for the sample category of  $C_i$ .

$P_i = S_i / S$  is the probability that any sample belongs to  $C_i$ , and  $P_{ij}$  represents the probability in  $S_i$  that belongs to  $C_i$ .

$$I = -\sum_{i=1}^m \frac{S_i}{s} \log_2 \left( \frac{S_i}{s} \right)$$

$$I_j = -\sum_{j=1}^m \frac{S_{ij}}{s_j} \log_2 \left( \frac{S_{ij}}{s_j} \right)$$

$$E(A) = \sum_{j=1}^m \frac{S_{aj}}{s} I_j$$

$$Gain(A) = 1 - E(A)$$

$Gain(A)$  is Information Gain that belongs to the variable of  $A$ , and we used MATLAB to get Information Gain of the index.

Table 2. The value of Information Gain of indicators

Gain (with or without default more than 90 days of debt)	0.0965	Gain (ratio of monthly debt service to income)	0.1322
Gain (age)	0.0001	Gain (educational level)	0.0009
Gain (profession)	0.0001	Gain (gender)	0.0001
Gain (car availability)	0.0077	Gain (marital status)	0.0038
Gain (housing condition)	0.0067	Gain (length of employment)	0.0054

Data showed that the value of some indicators is too small, such as age, gender and profession, which indicates that they generate little impact on the risk, so these indicators can be excluded.

**The calculation of 'woe' value and risk assessment analysis.** Smith proposed the concept of 'woe' value in the year of 2002, 'woe' value reflect the size of the risk, and subsequently risk assessment can be carried out. Good customers mark 0, bad customers mark 1. In this paper:

$$woe = \ln \frac{\text{proportion of good customers in the category}}{\text{proportion of bad customers in the category}} = \frac{\ln(P_{y0})}{\ln(P_{y1})}$$

We used MATLAB to get the WOE value of various indicators

Table 3. The WOE value in index classification

Classification	Good customer	Bad customer	WOE value
Yes	1196	278	0.0728
No	4	22	-3.0910
(0,20%]	490	50	0.8961
(20%,50%]	604	121	0.2215
above 50%	106	129	-1.5827
postgraduate degree or above	13	1	1.1787
undergraduate degree	254	60	0.0567
below undergraduate degree	933	239	-0.0244
married	898	175	0.2491
divorced	121	28	0.0773
unmarried	181	97	-0.7625
with a house and no mortgage	234	37	0.4581
with a house and a mortgage	451	119	-0.0540
No house	515	144	-0.1119
with a car and without loan	311	50	0.4415
with a car and a loan	12	51	-0.6931
No car	787	199	-0.0114
more than five years	82	5	1.4110
(2,5]	50	11	0.1278
(0,2]	1068	284	-0.0617

Draw the scatter plot of WOE value, and then analyze the data to make a risk assessment.

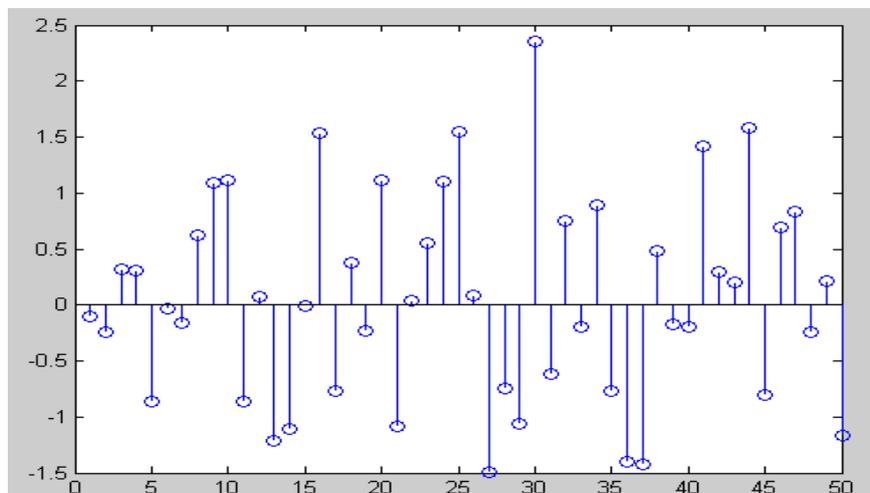


Figure 1. the scatter plot of WOE value

We analyzed the relevant data in the table and observed the distribution of points in the scatter plot. We could make relevant risk assessments:

- (1) The borrower's age, gender and its own profession have the minimal impact on the default risk.
- (2) The higher the level of education of the borrower, the lower the risk of default.
- (3) The borrower who has a normal, healthy marriage and family has the lower the risk of default.
- (4) The longer the borrower's length of employment, the lower the risk of default.
- (5) The Borrower who has a small proportion of mortgages has the lower the risk of default.
- (6) The Borrower who has a small proportion of car loans has the lower the risk of default.
- (7) The lower the ratio of monthly debt service to income borne by the borrower, the lower the risk of default occurs.

(8) The existence of default debt of the borrower does not necessarily increase the risk of default. However, if there is a default of more than 90 days in history, the default risk of the borrower is greatly increased.

## Conclusions

In this paper, we select seven important indexes to constitute the evaluation index system by using the method of Information Gain firstly, then we establish an assessment analysis model to evaluate the risk level of P2P network loan. In this model, the analysis method of WOE value is used to evaluate the risk level of P2P network loan. The conclusion can be drawn that the larger the WOE value of complex development index, the lower the risk level of borrowers' credit.

## Acknowledgments

This work is supported by the 2017 Excellent Youth Project of Hubei Provincial Department of Education (No. Q20173007).

## References

- [1] Y.X. Mo: The theory and practice of domestic and international literature review about P2P network loan, *Theory and Practice of Finance* 12 (2011) 102-103.
- [2] K.N. Fang: The analysis of credit risk under asymmetric credit information, *Economic Research*, 23 (2010) 97-107.
- [3] C.J. Rao, M. Goh, Y. Zhao and J.J. Zheng: Location selection of sustainability city logistics centers, *Transportation Research Part D: Transport and Environment* 36 (2015) 29-44.
- [4] G. Feng and S.S. Jiang: The alienation and regulation of China's P2P network loan platform, *Commercial Research* 5 (2013) 29-37.
- [5] Y. Guo: Research on the development of P2P small loans in China, *Shanghai Finance* 12 (2012) 19-22.
- [6] C.J. Rao, X.P. Xiao, M. Goh, J.J. Zheng and J.H. Wen: Compound mechanism design of supplier selection based on multi-attribute auction and risk management of supply chain, *Computers & Industrial Engineering* 105 (2017) 63-75.
- [7] S.D. Liang, Y. Li and Z.B. Fang: Comparative analysis of credit risk models, *Management Science in China* 2 (2002) 17-22.