



Big Data Analysis Based on the Risk of Consumption Rebate Network Pyramid Selling

Ke Cheng¹

¹*Department of Public Security Technology, Jiangxi Police College. Nanchang city xing bay road no. 1666, China
Email: 740973316@qq.com*

Abstract

In recent years, "consumption rebate network pyramid selling" has become one of the mainstream criminal modes of pyramid selling crime. Its crime means of concealment, replication fission fast, easy to form a wild growth situation, harm the market economic order, cause mass incidents. Through big data collection and analysis, the research team screened out high-risk platforms of "consumption rebate" in e-commerce industry, mined out the scale of personnel and capital involved in multi-level network marketing, and found the law of crime and revealed market risks.

KEYWORDS: *consumer rebates; pyramid selling; network platform; Risk analysis*

1. Introduction

In the field of e-commerce, consumption rebate is a common promotion behavior, which means that consumers will get cash or cash coupons returned by merchants when they shop on e-commerce platforms, in order to promote consumers to make second purchases and create more profits. However, some criminals in the name of e-commerce, in order to "full consumption back" "consumption rebate" and other gimmicks, according to the level of the development of members, franchisees, agents, directly or indirectly to the number of development personnel as the basis for remuneration or rebate, pyramid selling illegal criminal activities.

In recent years, "consumer rebate network pyramid selling" has become one of the mainstream criminal modes of pyramid selling crime. According to incomplete statistics, since 2017, public security organs have received 566 online pyramid scheme reporting clues, including 211 consumption rebate clues, accounting for 37% of the total. Such crimes deviate from the rules of e-commerce transactions, and the capital chain is bound to break. Once the crimes are committed, the amount involved is tens of billions, involving millions of people^[1]. Against the above background, the research team carried out big data analysis. Based on the data analysis results, we provide suggestions and guidance to combat departments to jointly carry out intensive crackdown operations and resolutely curb the high incidence of crimes.

2. Research objectives

Our research objective is to discover the rules and reveal the risks. The advantages of data resources and advanced analysis technology are used to make an in-depth analysis of the risk scale, regional distribution, group status, capital links and crime techniques of consumer rebate network pyramid schemes, deeply reveal the social hazards and causes, and put forward countermeasures and suggestions for preventing and resolving risks.

3. Evaluation process

3.1. Accurate extraction of criminal characteristics

The research team comprehensively combed the crime situation of consumer rebate online pyramid schemes in recent years, sliced and analyzed major typical cases, and summarized and extracted the characteristic dimensions of illegal crimes. This kind of pyramid schemes mainly include: charging pyramid schemes entry fees, setting up complex rebate rules, forming unique hierarchical organizations, and openly deluding false propaganda^[2].

3.2. Comprehensive collection of Data resources

According to the research needs, all kinds of data resources were comprehensively collected, cleaned and

combed synchronously, and data sets were established^[3]. Data sources are as follows: 1.03 million enterprises whose business scope contains "e-commerce" were collected through the big data platform of the public security Bureau; In terms of fund data, coordinated the Anti-money Laundering Center of the People's Bank of China to transfer more than 3.43 million fund transaction data from more than 7,600 fund accounts; In terms of police data, more than 20 million participant data of more than 150 pyramid selling cases, 197,000 pyramid selling blacklist data, 192,000 pyramid selling criminal suspects data, and 2,769 pyramid selling criminal cases data were collected from the public security system. In terms of background data, nearly 300G of background data of some high-risk consumption rebate platform websites were retrieved; In terms of public data, 2.62 million pieces of public information related to Internet public opinions were retrieved through technical means.

3.3. The whole network platform inspection

3.3.1. Fully analysis of e-commerce platform data

Technical measures were taken to scan the entire network of 1.03 million enterprises whose business scope contained "e-commerce", and 109,000 e-commerce enterprises were found to have actually opened online shopping platforms.

3.3.2. Screen consumption rebate platforms

The characteristics of operation mode of e-commerce platform were extracted and mapped to 26 keywords that could be recognized by the computer as preliminary screening indicators. Out of 109,000 e-commerce platforms, 9210 platforms with consumption rebate business model were selected.

4. Scientifically calculate the scale of risks

4.1. Identify high-risk active platforms

Based on expert knowledge and case handling experience, in-depth analysis of historical cases was conducted, and 24 risk indicators from 7 characteristic dimensions were summarized and extracted as in-depth investigation indicators. A consumption rebate risk platform assessment model based on mathematical regression analysis was established to quantify the risk degree^[4]. Concrete is:

4.1.1. Ordered classification

The risk degree of consumer rebate online pyramid selling is divided into seven characteristic dimensions: freshness, activity, characteristic, suspicious, explosive degree, correlation degree and political and stability degree. Among them, the freshness mainly includes the

time point of grasping keywords; Activity mainly includes public opinion index, APP downloads, platform visits, etc. The characteristic degree mainly includes the situation of pulling head, threshold fee, establishing hierarchy relationship, etc. The degree of suspicion mainly includes commission reward, fictitious transaction, domain name characteristics, server address, program fingerprint, etc^[5]. The degree of thunder mainly includes negative public opinion, website closure, legal consultation, etc. The correlation degree mainly includes professional personnel, administrative punishment and public security information. The stability involved mainly includes the control by foreign hostile forces, foreign media reports and the gathering of participants^[6].

4.1.2. Quantization of scores

The total risk score was set at 100 points, and 20 points were assigned to feature degree and thunder degree, 15 points to suspicion degree and political stability degree, 10 points to freshness degree, activity degree and correlation degree, respectively, based on expert knowledge and case handling experience. At the same time, the 7 characteristic dimensions are refined into 24 measurable sub-dimension risk indicators.

4.1.3. Rating

Choose 10 consumer rebates kind typical cases as the training set, the list of the top 100 risk main body as the test set, repeatedly on risk assessment tool for training, testing, combined with the artificial intervention to verify, gradually improve the accuracy of risk assessment tools, eventually will score more than 60 points on model accuracy up to 83.6%, The accuracy of more than 80 points is 100%. The research team selected platforms with a model score of more than 60 points and verified by manual intervention, and finally identified 826 high-risk active platforms as the key objects of this strategic research.

The above risk platform investigation and evaluation model can be described as follows:

$$Score = q_1x_1 + q_2x_2 + q_3x_3 + q_4x_4 + q_5x_5 + q_6x_6 + q_7x_7$$

Where, x is the indicator feature; q is the weight, a positive value of 0-1, which can be adjusted according to the model establishment. Score is the total risk Score, with a maximum Score of 100.

4.2. Analysis of the scale of participants

Based on the risk degree of 826 high-risk active platforms, some platforms were randomly selected from different ranking intervals to calculate the scale of staff,

and linear regression equation was used to estimate the scale of staff of other platforms^[7]. For the current common website type membership system and APP type membership system, the mathematical model of personnel size measurement and evaluation is established respectively.

For the website type membership system. Under normal circumstances, active MLM participants will frequently log in to the platform membership system to check points, settle downlines, withdraw cash and other operations^[8]. According to the behavior characteristics of participants mentioned above, the research team coordinated with the National Internet Emergency Response Center and Tencent to monitor the visits of the platform member system servers within a certain period of time. Based on the client situation, a single occasional access is ignored and repeated access from the same client is deduplicated^[9]. The mathematical formula of the number of members of the sampling monitoring platform can be described as follows:

$$web_member_sample = \sum_{\{IP|IP \in IPset\}} (Filter_{IP}(\sum_{i=1}^T x_{IP}^{(i)}))$$

$$Filter(x) = \begin{cases} 1 & \text{if } x \geq 2 \\ 0 & \text{if } x < 2 \end{cases} x_{IP}: \text{the number of daily IP visits}$$

The number of members of the sampling monitoring platform is trained, and the mathematical formula and relevant weight parameters for predicting the number of members of the platform are obtained:

$$web_member_predict = \alpha_1 capital + \alpha_2 employee + \alpha_3 established + \alpha_4 riskscore$$

Where, α_1 is the weight of registered capital of the company; α_2 is the weight of the number of employees; α_3 is the weight of the company's operating time; α_4 is the weight of platform risk score.

According to the investigation, 252 of the 826 high-risk active platforms used the site-type membership system, and 43 of them were randomly selected. The monitoring found that the query volume of the website of the member system was 4.24 million times in a month. Using the mathematical model above, it is estimated that the number of active participants is about 15.12 million.

For App-based membership system, the mathematical formula of the number of members of sampling monitoring platform can be described as follows:

$$app_member_sample = Download/Version_Count$$

The number of members of the sampling monitoring platform is trained, and the mathematical formula and

relevant weight parameters for predicting the number of members of the platform are obtained:

$$app_member_predict = \alpha_1 capital + \alpha_2 employee + \alpha_3 established + \alpha_4 riskscore$$

Where, α_1 is the weight of registered capital of the company; α_2 is the weight of the number of employees; α_3 is the weight of the company's operating time; α_4 is the weight of platform risk score.

According to the survey, 574 of the 826 high-risk active platforms used web-based membership systems, 365 of which were randomly selected, and 270 million APP downloads were found. Using the mathematical model above, it is estimated that the number of active participants is about 28.12 million. The total number of participants in the two member systems is about 43.24 million.

4.3. Analysis of capital scale

Through the analysis of typical historical cases and some high-risk active platforms discovered in this strategic analysis, it is found that consumption rebate platforms often directly use company accounts for online payment and settlement in order to attract staff to join and show the legal compliance of their business activities^[10]. For this purpose, the company's account capital trading situation of high-risk active platform is queried, and the error correction is made by using the account of the company's directors, supervisors, senior personnel and counterparty account capital trading situation. Through the transaction records of the sampling platform, the total number of transactions is calculated:

$$U = \sum_{i=1}^n filter_u(x_i)$$

$filter_{u(x)}$: Indicates whether the transaction meets the count of the specified amount. If the condition is met, it is marked 1. If the condition is not met, it is marked 0

Then calculate the total capital transaction amount of the sampling platform:

$$Fund_sample = U \cdot Umoney$$

Umoney: indicates the amount corresponding to the membership fee

Train the amount of capital transaction of the sampling platform, and obtain the mathematical formula

and relevant weight parameters for predicting the scale of capital of the platform:

$$\begin{aligned} Fund_predic \\ = \alpha_1 capital \\ + \alpha_2 employee + \alpha_3 established + \alpha_4 member_total + \alpha_5 riskscore \end{aligned}$$

Where, α_1 is the weight of registered capital of the company; α_2 is the weight of the number of employees; α_3 is the weight of the company's operating time; α_4 is the weight of membership size; α_5 is the weight of platform risk score.

Through anti-money laundering inquiry of 6,804 accounts involved in 122 randomly selected high-risk platforms from 826, and using the mathematical model above, it is estimated that the total amount of funds

involved is about 1.23 trillion yuan.

5. Research results

After big data analysis of all data sets, the research team's analysis results are as follows.

5.1. The scale of industry risks is huge

Through the monitoring of the whole network, it is found that currently there are more than 109,000 e-commerce platforms, and 9,210 of them adopt the business model of consumption rebate. According to the consumption rebate risk assessment model, 5,384 platforms, accounting for 58%, have Characteristics of pyramid selling behavior. There are 826 platforms with high risk, high activity and high suspicion, accounting for 9% of the total (see figure1).

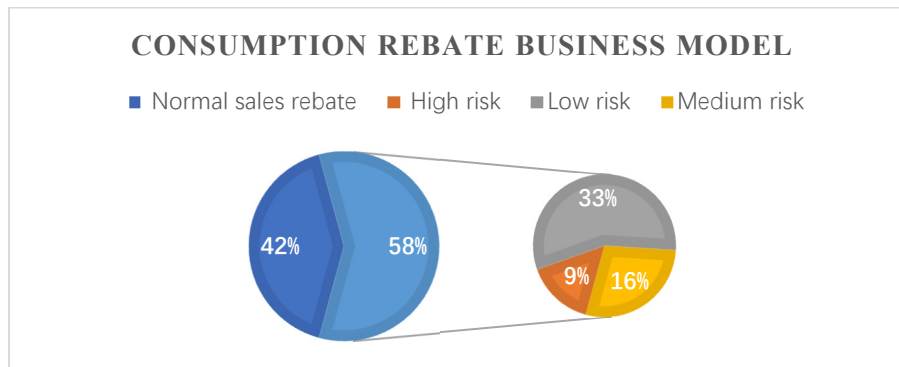


Figure 1. Figure with pyramid selling behavior characteristics of the platforms

5.2. Prominent risks in key positions

From the geographical distribution, the eastern coastal developed areas and more populous central provinces are relatively prominent risks. Guangdong, Beijing, Shanghai, Zhejiang, Sichuan, Jiangsu, Fujian, Shandong, Anhui and Henan ranked first in the total number of platforms, accounting for 86.1% of the total. From the point of view of single risk, the "head" platform has a large volume and high concentration of risk. Among the 826 high-risk platforms, 43 are more than 5 billion yuan, accounting for 6.09%, and the capital volume accounts for 81.89%.

5.3. Overlapping risks and hidden dangers

Some platforms have equity links. Through equity penetration analysis, 32 of the 826 high-risk platforms, accounting for 3.8%, have equity connection; There is money lending between some platforms, once "thunder", easy to trigger a chain reaction; Some platforms have state-owned assets or have been promoted by local government leaders, which can easily form "moral hazard" and increase the difficulty of disposal. Some of the start-up funds of some platforms come from financial

institutions, and risks may penetrate into financial institutions.

5.4. Criminal techniques are constantly updated

In terms of publicity methods, it has gradually shifted from simple "consumption rebate" to self-created concepts such as "sharing economy", "new retail" and "social e-commerce". In terms of promotion channels, in addition to offline propaganda, online shopping malls and traditional media, most platforms have specially developed App, and some platforms also rely on Wechat and small programs for shopping malls. Some platforms even rely entirely on Jingdong, Taobao and other formal shopping malls for promotion. From the perspective of "commodity" category, in addition to food, clothing, health care products, daily necessities and other kinds of physical commodities, but also gradually expanded to futures, options, consumer coupons, virtual currency and other virtual commodities, and eventually evolved into pure capital plate. From the perspective of rebate mode, it has changed from simple "multi-level rebate" to "single-layer rebate and team payment", and the form of rebate has also changed from simple capital to "points", "virtual currency" and "options", etc., with the boundary of illegal crimes becoming increasingly blurred.

5.5. Obvious professionalization characteristics

According to the data analysis of 16.26 million people who participated in network pyramid selling cases, 1.02 million people participated in two or more cases, mainly concentrated in Guangdong, Shandong, Heilongjiang, Hunan, Henan, Sichuan and other provinces, and some people even participated in 14 cases at most. Among them, 9 closely related core gangs were involved in three or more cases at the same time, showing obvious professionalization characteristics.

5.6. Risks related to stability are more prominent

According to the network monitoring, there were 103,000 negative public opinions on the above-mentioned 826 high-risk platforms, among which 1,393 were the most. Some groups frequently incite cross-provincial gatherings to "protect their rights", and some even go abroad to make complaints.

6. Conclusions

Our research model finds that consumer rebate platforms involve a large number of people and a huge amount of capital. Once the capital chain breaks and becomes a pyramid scheme criminal case, it will breed systemic financial risks and affect social stability. In the frontier field of economic crime investigation, our research provides criminal investigation departments with strategies to prevent pyramid schemes and points out the direction of combating crimes. In the future, there will still be new development of consumption rebate platform, and our research model may need to add new variables for optimization.

Acknowledgments

Jiangxi Police College Scientific Research Project: Public Security Big Data Application research (No: 202019B); Illegal Fund Transfer Payment Approach research (No. 2020YB001)

References

- [1] Lee Cheng. 2019. *Empirical study on new characteristics and investigation countermeasures of network pyramid selling crime*. *Journal of Henan Police College*. vol28: p 120-123
- [2] Ma Zhenpeng. 2020. *Investigation countermeasures of consumption rebate network pyramid selling crime*. *Journal of People's Public Security University of China*, vol 12:p91-93
- [3] Sun Daoning, Zhao Guangye. 2020. *Re-examination of network pyramid selling case investigation train of thought*. *Journal of Beijing Police College*. Vol 05: p 83-85
- [4] Cheng Xiaobai, Deng Changzhi. 2019. *Technology of illegal funds analysis and investigation*. *China People's Public Security University Press*, Beijing
- [5] Xu Sijia. 2017. *Investigation countermeasures of consumer rebate network pyramid selling crime cases*. *Journal of Qiqihar University*, vol 12:p56-58
- [6] Du Xiuping, Yu Long. 2021. *Research on the governance countermeasures of consumer Rebate type network pyramid selling crime*. *China Anti-counterfeiting Report*, vol 01:p73-75
- [7] Qian Ying, Fang Xiunan. 2022. *Multiple linear regression model and its application*. *China Science and Technology Information*. Vol 04: p 73-74
- [8] Cheng Xiaobai, Tu Mi. 2021. *Public Security Big data application tutorial*. *China People's Public Security University Press*, Beijing
- [9] Xie Zerun, Liu Deli. 2019. *Investigation countermeasures of network pyramid selling*. *Journal of Wuhan Public Security Executive Leadership Academy*. Vol 03: p 26-27
- [10] Tian Xinnan, Wang Congcong. 2019. *The application of big data in the investigation of network pyramid selling criminal cases*. *Law and society*. Vol 01: p 230-231

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

