Advances in Social Science, Education and Humanities Research, volume 598
Proceedings of the 7th International Conference on Social Science and Higher Education (ICSSHE 2021)

Regional Atlas Analysis of the Extent of Concern About Telecom Fraud

Bin Liu[1,*] and Pengrong Zhao[1]

[1] Department of Public Security Administration, Jiangsu Police Institute, Nanjing, Jiangsu 210031, China
*Corresponding author. Email: liubin000@163.com

ABSTRACT
Public awareness of anti-telecommunications fraud is the key to an effective reduction of the occurrence of fraud. Thus, this study was conducted to look into the extent of public concern about telecom fraud. It firstly analyzed the population characteristics and spatial characteristics of the extent of concern of telecom fraud from 2011 to 2020 by Baidu Index. It then discussed the factors influencing user concern by multiple linear regression, based on the 2016 Analytical Report on Telecommunications Fraud in China, China Statistical Yearbook 2017 and Baidu Index. The results indicate that the degree of telecom fraud attacks, public detection ability, education level, and local economic development could impact the extent of public concern over telecom fraud.
Keywords: *telecommunications fraud; regional differences; multiple linear regression; Baidu Index*

1. INTRODUCTION

The Internet, with its popularization and application, has brought great convenience to people. Yet, it has also become an important carrier of non-contact crimes, among which telecommunications fraud (hereinafter referred to as telecom fraud) is the most prominent. One of the main reasons why criminals of telecommunications fraud have repeatedly succeeded is the public ignorance or lack of knowledge. The lack of professional knowledge easily leads people into financial fraud, the misjudgment of emotions makes them fall into the trap of "pig-butchering scam" (also called romance scam), and the ignorance of technology leads to the leakage of personal information. The Internet, rather than an accomplice of crime, should be an important information resource, providing the public with more information on network security. Therefore, the public should be actively concerned about telecom fraud through the Internet and acquire skills and knowledge on anti-telecom fraud, thus curbing the spread of telecom fraud from the very onset. The extent of network concern about telecom fraud reflects the netizens' concern about telecom fraud and is also related to the effect of telecom fraud information release.

2. ANALYSIS OF POPULATION CHARACTERISTICS ON THE EXTENT OF USER CONCERN

According to the Statistical Report on China's Internet Development, as of December 2020, the number of search engine users in China had reached 770 million. Market data showed that search engines, social platforms and e-commerce platforms were the top three channels for netizens interviewed to obtain information, accounting for 69.6%, 49.0% and 35.3% respectively. [1] Baidu Index, an information platform based on massive netizens' behavior data, can accumulate user concern data through netizens' concern. [2] This study adopted Baidu Index to obtain netizens' concern about telecom fraud. Since there had been no unified name for online fraud, this study took the two most widely used keywords, "online fraud" and "telecom fraud", to search for the extent of concern from Internet users and the media.

2.1. The extent of User Concern [3]

Researchers collected data from January 1, 2011 to March 16, 2020. The average daily Baidu Index value for "online fraud" was 706, and the value for "telecom fraud", 840, indicating that the public was more aware of "telecom fraud". This term had attracted a high degree of public attention and was showing a slow upward trend.
In addition, there had been several peaks in the public concern about "telecom fraud". The first peak was from January 13-19, 2014, when Baidu Index suddenly rose to 5849. During this time period, a famous actress called Wei Tang fell into the trap of telecom fraud, which caused a heated discussion in the society due to celebrity effect; the second peak was from April 18-24, 2016, with Baidu Index rising to 2411; the third peak was from August 29 to September 4, 2016, with Baidu Index suddenly rising to 5072. On August 19, 2016, Yuyu Xu, a college student, was defrauded and ended up dead. Her death caused social shock and triggered public condemnation of telecom fraud. The fourth was from September 26 to October 2, 2016, with Baidu Index, 3499; The fifth was from November 27 to December 3, 2017, with Baidu Index rising slowly to 2,178 and fluctuating slightly since then.

There had not been many changes in the public concern about "online fraud", with only one prominent peak from March 23-29, 2015. The Index rose to 2006, as shown in Fig. 1.

Netizens' concern in 2019 weren't the same as that in 2016 with surges. However, the 44th Statistical Report on China's Internet Development released by CNNIC showed that "online fraud" accounted for 28.1 percent in the "Security Threats to Internet Users". This type of crime remained a concern of the general public.

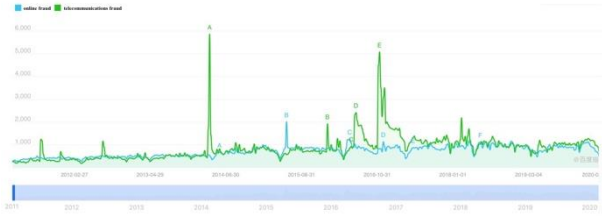

Figure 1. The extent of concern of netizens

2.2. The extent of Media Concern

The study focused on the number of news relating to the keywords "telecom fraud" and "Internet fraud", reported by major Internet media and included in Baidu News Channel. The data was from January 1, 2011 to March 16, 2020. The Media Index for "online fraud" was 15 and the Index for "telecom fraud", 55. Therefore, the media was also very concerned about this issue.

Media concern had also gone through several peaks. The peaks for "telecom fraud" were: Baidu index showed 365 from May 21-27, 2012; 284 from January 13-19, 2014; 411 from November 9-15, 2015; 732 from April 11-17, 2016, and 952 from August 29 to September 4, 2016. It could be seen, especially through several peak pointss, media concern and netizens concern were roughly the same. For example, media concern and netizens concern from January 13-19, 2014 and from August 29 to September 4, 2016 were similar, as shown in the same highest values. This indicated that the concern of the media guided that of the netizens, while the latter, at the same time, retroacted to the former. The two echoed each other, and the overall trends were convergent, as shown in Fig. 2.

For example, the aforementioned Yuyu Xu Case was first published by Yimeng Evening News and then followed up by mainstream media like The Paper (Pengpai News) and Southern Metropolis Daily. The follow-up reports presented to netizens the illegal fraud chain behind the telecom fraud and sales of personal information. This heated topic continued to produce effects of depth and width. Netizens also reposted and commented a lot on it. Under the agenda setting of the media, the public did not focus solely on Yuyu Xu Case. Not only were they dissatisfied with the simple arrests of fraudsters by public security departments, they also appealed to relevant departments to thoroughly investigate the sources of the information disclosure and the overall illegal profit chain. [4]

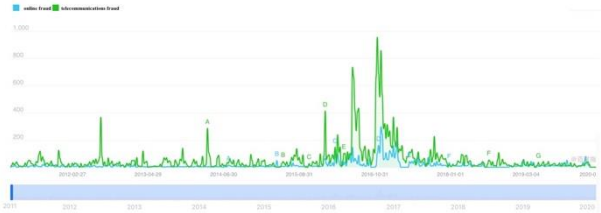

Figure 2. Media Index

2.3. Age Attribute

Researchers collected data from February 1-29, 2020 and the analysis, as shown in Figure 3, indicated most of the netizens concerned about "online fraud" and "telecom fraud" were younger than 40 years old, of whom almost half were 20-29 years old. The elderly were less concerned. This was of course related to the age structure of netizens. According to the 44th Statistical Report on China's Internet Development Status released by CNNIC, netizens over the age of 50 accounted for 13.6% in China, and those between the ages of 20 and 40 accounted for 48.3%. The elderly, lack of knowledge about financial investment, health care, and network technology, could easily be defrauded by criminals and suffer heavy losses. This, to some extent, showed that great concern about telecom fraud could help reduce the risks of being defrauded.

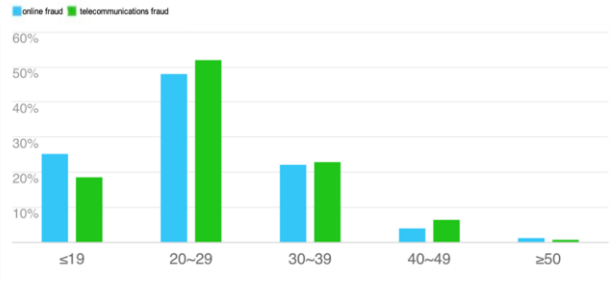

Figure 3. Age distribution.

2.4. Gender Attribute

An equal percentage of men and women were concerned about "online fraud", and the percentage of men concerned about "telecom fraud" was higher than that of women, as shown in Fig. 4. There were slightly more men than women in general. In terms of victim gender, men and women accounted for 63% and 37% respectively. Among the top ten types of frauds and scams, except for rebate scam, part-time scam and free-gift scam, men outnumbered women in the rest of pornography scam, dating scam, financial and credit fraud, transaction fraud, account hacking scam, low-price lure scam, and counterfeit fraud. Ninety-eight percent of the victims of pornography scam were male. [5]

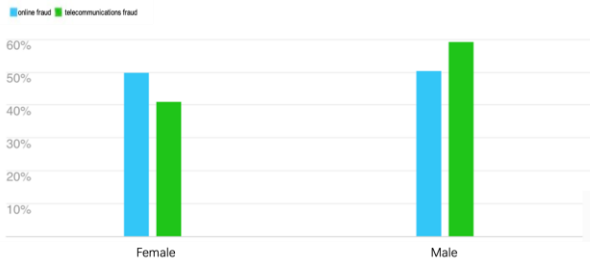

Figure 4. Gender distribution.

2.5. Issues Concerned

Netizens paid attention to the buzz word "telecom fraud" from March 9-15, 2010. It could be seen that netizens were increasingly concerned about online fraud, campus loans, anti-telecom fraud manual, financial fraud and credit card fraud. They tended to care less about Wei Tang Telecom Fraud Case and the success rate of fraudulent transfer recovery. This could be interpreted by the actual occurrence of telecom fraud recently, as shown in Fig. 5. The top three types of frauds and scams in the first half of 2019 were transaction fraud, part-time job scam and dating scam. [5]

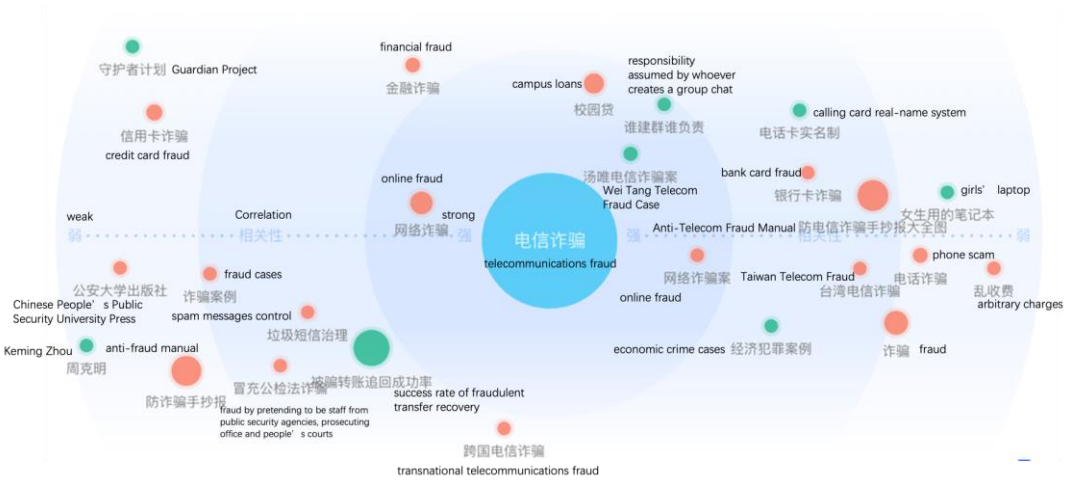

Figure 5. Keyword distribution.

3. ANALYSIS OF SPATIAL CHARACTERISTICS ON THE EXTENT OF USER CONCERN [6]

The study then analyzed the extent of user concern about "telecom fraud" in different areas based on IP addresses that Internet users were assigned.

3.1. Regional Differences in the Extent of User Concern

Researchers collected data from July 1, 2013 to March 16, 2020 and it showed that netizens concerned about "online fraud" and "telecom fraud" were mainly located in East China, followed by South China, Central China, and relatively few in Northeast and Northwest China. This was highly consistent with the public reporting frequency. Netizens in East China, Central China, and South China topped the list in terms of reporting frequency. Guangdong ranked first which indicated that the public in this area were usually concerned about such issues, so they were likely to have certain anti-fraud awareness and experience.

3.2. Inter-provincial Differences in the Extent of User Concern

Users in different provinces showed quite different extent of concern and the top three provinces were Guangdong, Jiangsu, and Zhejiang, as shown in Fig. 6.

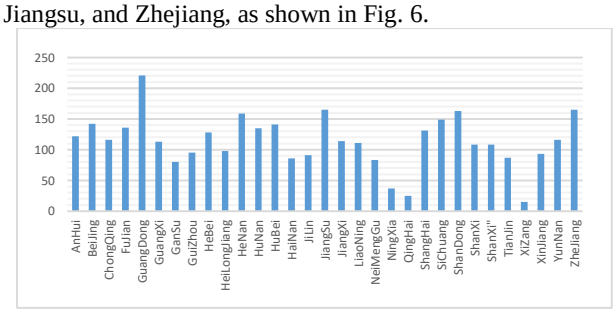

Figure 6. Distribution of telecom fraud concern by Provinces in 2019

Provinces could be categorized into four types according to the extent of concern: deep concern, growing concern,

common concern and less concern about telecom fraud, as shown in Table 1.

Table 1. Distribution of Provinces Concerned about Telecom Fraud in 2019

Types	Provinces
deep concern about telecom fraud (extent of concern≥150)	Guangdong, Henan, Jiangsu, Shandong, Zhejiang, Sichuan
growing concern about telecom fraud (150＞extent of concern＞100)	Anhui, Beijing, Chongqing, Fujian, Guangxi, Hebei, Hunan, Hubei, Jiangxi, Liaoning, Shanghai, Shaanxi, Shanxi, Yunnan
common concern about telecom fraud (100＞extent of concern＞50)	Gansu, Guizhou, Heilongjiang, Hainan, Jilin, Inner Mongolia, Tianjin, Xinjiang
less concern about telecom fraud (extent of concern＜100)	Ningxia, Qinghai, Tibet

4. MODEL CONSTRUCTION OF THE EXTENT OF USER CONCERN

The study had so far discussed the characteristics of the population concerned about telecom fraud in different regions. Then it needed to explore the factors that affected user concern. Researchers found through survey and literature review that the factors affecting user concern included the amount of call interception, time for detection, illiteracy rate, GDP per capita, number of netizens, and phone number attribution. Multiple linear regression was employed to analyze user concern factors, and variables with no significant impact were eliminated.

4.1. Construction of Regression Model [7]

There was a lot of public concern over "telecom fraud" in 2016 from Figs. 1 and 2, with high fluctuations. Therefore, data in 2016 were chosen as a sample for analysis. Researchers selected interception, time for detection, illiteracy rate, GDP per capita, number of netizens, and phone number attribution as explanatory variables, and user concern as explained variable. Based on the sample data of mobile security software intercepting fraudulent calls, the call interception variable was the number of fraudulent calls intercepted by 360 Mobile Security in each province, which reflected the attack area of telecom fraud. Time for detection variable was the ability to identify fraudulent calls, namely, the average time it took for all users in a region to receive an unidentified fraudulent call (abbreviated as average fraud detection time) before hanging up the call and marking it as a fraudulent call. The shorter average fraud detection time, the higher ability users in the region had to identify fraudulent calls, and on the contrary, the longer time, the lower ability. The illiteracy rate variable referred to the proportion of the illiterate population over the age of 15, standing for the cultural and educational development of the region. The per capita GDP variable represents the economic development of a region. The variable of the number of netizens referred to the number of people surfing the Internet in a region, showing the network development in that area. The variable of phone number attribution reflected to some extent the preference of criminals for hiding themselves through the choice of phone numbers for telecom fraud, which had a certain significance for the revelation of criminals' actual hiding places.

The data were taken from 2016 Analytical Report on Telecommunications Fraud in China, China Statistical Yearbook 2017 and Baidu Index.

By using telecom Fraud Baidu Index－extent of user concern as the dependent variable, and per capita GDP, call interception, detection time, illiteracy rate, phone number attribution and the number of netizens as explanatory variables, the study constructed in the following a multiple regression model:

$$\mathrm{BI} = \beta_0 + \beta_1\mathrm{GDPPC} + \beta_2\mathrm{NP} + \beta_3\mathrm{ILP} + \beta_4\mathrm{INQ} + \beta_5\mathrm{TIME} + \beta_6\mathrm{IU} + \varepsilon$$

Among them, β_0 was the intercept, and β_1, β_2, β_3, β_4, β_5 and β_6 represented the regression coefficients of per capita GDP, number attribution, illiteracy rate, call interception, detection time, and the number of netizens respectively.

4.2. Data Analysis

The study applied SPSS software to perform multiple linear regression on the data, making analyses of model goodness of fit, residual analysis, significance analysis of regression equation, missing value analysis, and significance analysis of regression coefficients of retained variables and excluded variables. The regression equation fitting process in Table 2 and Table 3 showed that the stepwise screening strategy was used to eliminate the two variables of the number of netizens and the phone number attribution. Finally, the regression equation was established after fitting, and the DW test value of the equation, 2.262, was obtained. The residual had no autocorrelation, indicating that the established equation could fully explain the characteristics and varying pattern of the explained variable, and the variables in the equation were more comprehensive and reasonable.

After screening the variables, the regression model in Table 2 was the final equation. When the significance level was 0.05, the P-value of the partial F test of the regression equation was 0, less than the significance level α.

Therefore, the linear relationship between the explained variable Y and the explanatory variables was significant, and it was appropriate to establish a linear model. In addition, the F statistic ratio was 58.735, and the significance test result was more significant, indicating that the regression had a significant effect on the interpretation of variation, which also reflected the significant effect of the regression equation.

Table 2. Model Fit Analysis

Model	R	R-squared	Adjusted R-squared	Standard Error of the Estimate	DW
1	.949 [a]	.900	.885	20.121	2.262

Table 3. Significance Analysis of Regression Equation

Model		Sum of Squares	Degree of Freedom	Mean Square	F	Significance
1	Regression	95112.233	4	23778.058	58.735	.000 [b]
	Residual	10525.767	26	404.837		
	Total	105638.000	30			

[a.] Dependent variable: Telecom Fraud.
[b.] Predictor variables: (constants), illiteracy rate, call interception, detection time, GDP per capita

Through model testing, it could be known that the analyzed samples and the established models met the needs of research. On this basis, the regression coefficients of the sample data were analyzed, and the regression coefficient significance T test was used to obtain the non-standardized regression coefficients and standardized regression coefficients of the retained variables of the model, as shown in Table 4. The regression contribution of each variable and its significance: sig<0.05, significant. VIF was used for collinearity diagnosis (when 0<VIF<10, there was no multicollinearity), there was no collinearity in the independent variables. Finally, the equation for user concern of telecom fraud was obtained:

$$\mathrm{BI} = 230.287 + 0.001\mathrm{GDPPC} - 1.527\mathrm{ILP} + 5.872\mathrm{INQ} - 4.311\mathrm{TIME}$$

By establishing the regression equation, the study obtained the quantitative equation of the concern of telecom fraud, which reflected more clearly the degree of influence of each influencing factor on the concern of telecom fraud. By adjusting the four independent variables, it can predict the degree of telecom fraud. The accuracy rate of fraud attention is 90%.

Table 4. Analysis of Regression Coefficients

Model		Unstandardized Coefficient		t	Significance	Collinearity	
		B	Standard Error			Tolerance	VIF
1	(Constants)	230.287	34.888	6.601	.000		
	Call interception	5.872	1.674	3.508	.002	.347	2.881
	Detection time	-4.311	1.236	-3.488	.002	.553	1.809
	GDP per capita	.001	.000	3.082	.005	.329	3.035
1	Illiteracy rate	-1.527	.582	-2.625	.014	.810	1.235

5. CONCLUSION

First of all, there was a positive correlation between the amount of call interception and the extent of concern over telecommunications fraud. The more fraudulent calls received by users in one province, the more serious telecommunications fraud attacks on the region and hence, the higher public concern about telecom fraud. Users in Guangdong for example, intercepted the most number of fraudulent calls, accounting for 20.0% of the total number of fraudulent calls intercepted in China. It also had the highest concern among netizens on telecom fraud, which was far ahead of other provinces.

Secondly, there was a negative correlation between detection time and telecom fraud concern. The longer users in various provinces detected fraudulent calls, that is, the lower their detection ability, the less concern over telecommunication fraud. Tibet, Ningxia, Qinghai had the lowest level of detection of telecom fraud, so there were also less concerns about telecom fraud.

Thirdly, there was a negative correlation between the illiteracy rate and telecommunication fraud concern. The higher the illiteracy rate, the less concern people in the area had over telecommunication fraud. The existing low level of education and lack of concern about telecom fraud could easily lead people to cognitive blind spots and fall into the trap of telecom fraud.

Finally, there was a positive correlation between per capita GDP and telecom fraud concern. The economic development would also affect the extent of public concern about telecom fraud, but overall the effect was not much.

The degree of telecom fraud attacks, public detection ability, education level, and local economic development

could all have an impact on the extent of concern netizens had over telecom fraud. The higher the public concern about telecom fraud, the more likely it is to reduce the occurrence of the fraud. Therefore, the publicity of anti-telecom fraud should be tailored to local people and conditions to attain desirable results.

ACKNOWLEDGMENT

This study was supported by Public Security Theory and Soft Science Research Key Project of the Ministry of Public Security of China (project number 2020LLYJJSST004), Key Construction Funding Project of the "Thirteenth Five-Year Plan" of Jiangsu Province in China, Jiangsu University Superior Subject Construction Funding Project (issued by Office of the People's Government of Jiangsu Province [2011] No.137), Jiangsu University Brand Major Construction Funding Project, Jiangsu Province "333" Project Funding.

REFERENCES

[1]2020 Report analysis of research on search engines. Retrieved from https://www.tang-seo.com/5113.html.(In Chinese)

[2]N. Feng, and D.K. Wang, "Analysis of the spatio-temporal characteristics of social concern about the Belt and Road Initiative—based on the survey of Baidu Index big data," Southeast Communication, 2018. (In Chinese)

[3]W.B. Xiao, "Who is concerned about China's government information publicity: from the perspective of Baidu Index," Intelligence Magazine, 2013. (In Chinese)

[4]The way the spread of social hidden diseases ignites public opinion: from Yuyu Xu's telecom fraud case and other incidents. Retrieved from https://www.yangfenzi.com/shehui/64979.html. (In Chinese)

[5]Tencent: research report on telecom network fraud governance in the first half of 2019. Retrieved from http://www.199it.com/archives/918236.html. (In Chinese)

[6]X. Li, "Temporal and spatial characteristics and influencing factors of online concern about cruise tour—a research based on Baidu Index," Statistics and Information Forum, 2016. (In Chinese)

[7]S.S. Chai, Y.M. Cao, and C.F. Long, "On the evaluation of influencing factors of tourists risk perception based on multiple regression model analysis," Journal of Ocean University of China (Social Science Edition), 2011. (In Chinese)