A Binary Logistic Regression Model to Unprosecuted Cases of Assisting a Cybercrime

Wang Yadi¹, Bei Yuxin¹
¹People’s Public Security University of China, Beijing, China
162241185@qq.com

Abstract. Nowadays, there are more and more cases of assisting a cybercrime, and the cost of investigation is relatively high. After the suspects were brought back to the case, a large number of such cases were deemed non-prosecution for various reasons, which caused a certain waste of judicial resources. Through the empirical analysis of 100 non-prosecutions and 100 judgments of cases of assisting a cybercrime, this article explores the influencing factors of the review and prosecution results of such cases, and builds a binary logistic regression model to predict them, with a view to provide reference for the public security organs to better handle such cases.

1. Introduction
In recent years, our country's cybercrime has continued to be high and frequent. Cybercrime has a fine division of labor, a high degree of professionalism, and an increasingly large scale of crime, which breeds a black-gray industrial chain associated with it. "Criminal Law Amendment (9)" adds Article 287 bis of the "Criminal Law" establishes the offence of assisting a cybercrime. The perpetrator, knowing that others use the information network to commit crimes, who provides technical support such as Internet access, server hosting, network storage, communication transmission, etc., or provides assistance in advertising promotion, payment and settlement, etc. will constitute this crime when the circumstances are serious, which realizes the independent conviction of the help behavior to the cybercrime benefit chain.

In practice, a large number of cases of assisting a cybercrime were not prosecuted during the review and prosecution stage of the procuratorate. The public security organs have invested a lot of manpower, material resources, and financial resources, but the actual effect of fighting crime still needs to be improved. This article adopts regression analysis method to study the factors that the cases of assisting a cybercrime not to be prosecuted, and builds the corresponding regression model, in order to provide scientific guidance for the investigation department to transfer related cases in practice and save investigation costs.

2. Literature review
From a research perspective, the domestic research on cybercrime assistance behavior mainly has a legal perspective and a criminological perspective. There are many research results from the perspective of law, mainly researching its legal nature, imputation model [1], judicial application and other issues, focusing on normative research, and also part of empirical research [2][3]. Criminology research mainly focuses on the related cybercrime black industry chain [4], exposing the patterns of its crime, and discussing how to fight.
From the perspective of the research content, it mainly focuses on two aspects. On the one hand, the nature of the criminalization of this type of behavior is explored. There is a fierce debate between the two views of accomplices becoming principal offenders and principal offenders becoming accomplices [5][6], producing a large number of research results; on the other hand, it studies the judicial application of the offence of assisting a cybercrime, and strives to solve the difficult problems in the conviction and sentencing of the help behavior to the cybercrime [7][8]with fruitful results.

At the same time, relevant research results abroad mainly involve: investigation and evidence collection of cybercrime [9] [10], identification of virtual identity [11], legislative comparison and application of law for new crimes [12], etc.

To sum up, the domestic research of legal aspects of the help behavior to the cybercrime are relatively in-depth, but in the existing empirical research, the research ideas, research content, and research methods are still relatively monotonous, and the guiding significance in judicial practice is relatively limited. Foreign research results are practical and closely related to crime. In this research field, this paper innovates the method, uses the binary logistic regression analysis to assisting cybercrime cases not to prosecute the influencing factors, and establishes the regression model.

3. Analysis on the influencing factors and correlation of unprosecuted cases of assisting a cybercrime

3.1. Data sources
In the Chinese judicial documents website (wenshu.court.gov.cn, time node: 11:30, June 15, 2021), select advanced search, set the case type as "criminal case", the document type as "judgment", and the cause of action as "assisting a cybercrime". According to the reverse order of time, nearly 100 judgments and 100 non prosecution decisions in cases of assisting cybercrime are selected as the original samples.

3.2. Variable description
In this paper, the explanatory variables are selected as the indicators to measure the results of case processing in the stage of review and prosecution. Through the analysis of 100 decisions of no prosecution and 100 judgments of China procuratorial network, the explanatory variables that may have an impact on the results of review and prosecution are summarized, including: criminal record, first offense, casual offense, crime form, illegal income, return the illegal income, compensate victims and obtain understanding of the victims, confession attitude, repentance performance, suspect's identity, time to participate in crime, clearly know, surrender, meritorious service, confess, type of accomplice, plead guilty, lack of evidence, contradictory evidence and type of criminal acts, totalling 20 indicators. The definition and basic statistics of each variable are shown in Table 1.

| Variable | Variable definitions | value | Mean | S.D. |
|----------|---------------------|-------|------|------|
| prosecution | Whether to prosecute or not | 0= no prosecution, 1= prosecution | 0.50 | 0.501 |
| criminal record | have criminal record or not | 0= no, 1= yes | 0.48 | 0.501 |
| first offence | whether it is first offence or not | 0= no, 1= yes | 0.50 | 0.501 |
| casual offence | whether it is casual offence or not | 0= no, 1= yes | 0.14 | 0.343 |
| crime form | wether the crime is accomplished or not | 0= other, 1= accomplished | 0.99 | 0.100 |
| illegal income | amount of illegal income | 0= other, 1= small | 0.42 | 0.495 |
| return illegal income | whether to return illegal income | 0= no, 1= yes | 0.27 | 0.445 |
| understanding of victims | compensate the victims and obtain understanding of victims | 0= no, 1= yes | 0.05 | 0.218 |
| confession attitude | confession attitude | 0= other, 1= good | 0.21 | 0.408 |
| repentance performance | whether have repentance performance | 0= other, 1= yes | 0.14 | 0.343 |
3.3. Pearson correlation analysis of dependent variables and independent variables

The Pearson correlation analysis was conducted between dependent variables and independent variables. The conclusion was as follows: 1) criminal record $x_1$, first offence $x_2$, illegal income $x_3$, return the illegal income $x_4$, repentance performance $x_5$, clearly know $x_6$, confess $x_8$, type of accomplice $x_9$, plead guilty $x_{10}$, lack of evidence $x_{11}$ and types of criminal acts $x_{12}$ is highly correlated with whether the case is finally not prosecuted, and Pearson correlation coefficient is 0.01. 2) surrender $x_7$ is related to the prosecution in the value of 0.05.

| Index | y   | x1  | x2  | x3  | x4  | x5  | x6  | x7  | x8  | x9  | x10 | x11 | x12 |
|-------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| y     | 1   |     |     |     |     |     |     |     |     |     |     |     |     |
| x1    |     | -551**| 1   |     |     |     |     |     |     |     |     |     |     |
| x2    |     | 580” | 911”| 1   |     |     |     |     |     |     |     |     |     |
| x3    |     | 344” | -262”| 284”| 1   |     |     |     |     |     |     |     |     |
| x4    |     | 225”| -0.127| 0.113| 0.281”| 1   |     |     |     |     |     |     |     |
| x5    |     | 190”| -259”| 278”| 0.138| 0.155”| 1   |     |     |     |     |     |     |
| x6    |     | -513”| 452”| -426”| -171”| -165”| -171”| 1   |     |     |     |     |     |
| x7    |     | 145”| -167”| 169”| 0.111| 0.058| 0.002| -0.114| 1   |     |     |     |     |
| x8    |     | 750”| -503”| 524”| 278”| 227”| -536”| 0.123| 1   |     |     |     |     |
| x9    |     | 277”| -203”| 215”| 0.127| 0.121| 0.101| 0.022| 0.085| 0.182”| 1   |     |     |
| x10   |     | 608”| -617”| 586”| 335”| -294”| -240”| -0.573”| 241”| 0.676”| 0.122| 1   |     |
| x11   |     | -482”| 507”| -482”| -324”| -261”| -190”| 547”| -256”| -609”| -0.055| -0.792”| 1   |
| x12   |     | -399”| -328”| 341”| -300”| 234”| -152”| -264”| 187”| 384”| 0.368”| 0.404”| -265”| 1   |

Note: * and ** indicate significant correlation at P < 0.05 level and extremely significant correlation at P < 0.01 level respectively
4. Regression model construction and empirical analysis

4.1. model construction

Logistic regression model is an important classification model, which is widely used because of its simple algorithm, high efficiency and easy promotion [13]. Whether not to prosecute cases of assisting a cybercrime belongs to the problem of two categories. The binary logistic regression method of logistic regression is applied to predict whether the case of assisting a cybercrime is not prosecuted.

Firstly, regression analysis was conducted for all variables. The regression results are shown in Table 3. Among the factors with sig value less than 0.05 (significance), the most influential factors to predict whether the case will be prosecuted are: the amount of illegal income, attitude to plead guilty, performance of repentance and whether to confess.

Table 3 Binary logistic regression analysis

| B                           | S.E.  | Wals  | df | Sig.  | Exp(B)  |
|-----------------------------|-------|-------|----|-------|----------|
| criminal record (1)         | -0.255| 1.419 | 0.032 | 1   | 0.858    | 0.775    |
| first offence (1)           | -1.270| 1.417 | 0.804 | 1   | 0.370    | 0.281    |
| casual offence (1)          | 0.233 | 1.232 | 0.036 | 1   | 0.850    | 1.262    |
| crime form (1)              | -23.272| 21956.000 | 0.000 | 1   | 0.999    | 0.000    |
| illegal income (1)          | -2.150| 1.054 | 4.164 | 1   | 0.041    | 0.116    |
| return illegal income (1)   | -0.312| 0.861 | 0.131 | 1   | 0.717    | 0.732    |
| understanding of victims (1)| -4.647| 4.323 | 1.155 | 1   | 0.282    | 0.010    |
| confession attitude (1)     | 4.636 | 1.252 | 13.723| 1   | 0.000    | 103.162  |
| repentance performance (1)  | -2.750| 1.335 | 4.241 | 1   | 0.039    | 0.064    |
| suspect identity (1)        | 5.119 | 3.270 | 2.451 | 1   | 0.117    | 167.225  |
| time to participate in crime (1)| 17.209| 18610.521 | 0.000 | 1   | 0.999    | 29769867.435 |
| clearly know (1)             | 0.000 | 0.000 | 2   | 1.000 | 56.214   |
| clearly know (2)             | 4.029 | 18389.147 | 0.000 | 1   | 1.000    | 9225306351.876 |
| clearly know (3)             | 25.248| 3870.710 | 0.000 | 1   | 0.995    | 17127859496.472 |
| suspect identity (1)        | 0.000 | 0.000 | 3   | 1.000 | 0.000    |
| confess (1)                 | -0.177| 0.928 | 0.036 | 1   | 0.849    | 0.838    |
| confess (2)                 | 1.881 | 1.694 | 1.234 | 1   | 0.267    | 6.562    |
| perform meritorious service (1)| 1.881| 1.694 | 1.234 | 1   | 0.267    | 6.562    |
| type of accomplice (1)       | -5.358| 1.263 | 17.984| 1   | 0.000    | 0.005    |
| plead guilty (1)             | -24.301| 7309.765 | 0.000 | 1   | 0.997    | 0.000    |
| lack of evidence             | -6.862| 9098.765 | 0.000 | 1   | 0.999    | 0.001    |
| lack of evidence (2)         | -5.251| 15553.935 | 0.000 | 1   | 1.000    | 0.594    |
| lack of evidence (3)         | 23.564| 40740.792 | 0.000 | 1   | 1.000    | 17127859496.472 |
| contradictory evidence       | -0.837| 0.719 | 1.355 | 1   | 0.244    | 0.433    |
| types of criminal acts       | -32.044| 21074.188 | 0.000 | 1   | 0.999    | 0.000    |

4.2. analysis of model results

The four variables with influence ($P < 0.05$) screened out from the above results were subjected to binary logistic regression again, and the results are shown in Table 4. Table 3 and table 4 show that confession attitude and confession ($Sig = 0.000$) have the greatest impact on whether to prosecute. Through the above analysis, we can get the equation logistic ($P = 2.14-1.27 * amount of illegal income + 3.033 * confession attitude-2.592 * penitence performance-5.284 * whether to confess$). When $P > 0.5$, assisting cybercrime case is more likely to be identified as non prosecution, on the contrary, it is more likely to be prosecuted. The accuracy of the regression model after re fitting is 91%, as shown in Table 5.

Table 4 Binary logistic regression model results

| B                           | S.E.  | Wals  | df | Sig.  | Exp(B)  |
|-----------------------------|-------|-------|----|-------|----------|
| illegal income (1)          | -1.270| 0.519 | 5.984| 1   | 0.014    | 0.281    |
| confession attitude (1)     | 3.033 | 0.746 | 16.531| 1   | 0.000    | 20.752   |
| repentance performance (1)  | -2.592| 0.909 | 8.140| 1   | 0.004    | 0.075    |
| confess (1)                 | -5.284| 0.787 | 45.019| 1   | 0.000    | 0.005    |
| Cons                        | 2.140 | 0.684 | 9.795| 1   | 0.002    | 8.497    |
Table 5 The accuracy of the regression model

| observed | predicted | Percentage correction |
|----------|-----------|-----------------------|
|          | Whether to prosecute or not |          |
|          | no        | yes                  |
| Whether to prosecute or not | 86 | 14 | 86.0 |
| yes | 4 | 96 | 96.0 |
| Overall percentage | | | 91.0 |

5. Conclusion

Through the binary logistic regression analysis method, this paper selects 200 legal documents from China judicial document network and China procuratorial network, studies the influencing factors of whether to prosecute the cases of assisting a cybercrime, and establishes a regression model. From the conclusion of the model, the amount of suspect's illegal gains, the attitude of confession, the performance of penance and whether it is confessed is more effective in explaining whether the case is prosecuted. Combined with the practical experience of the public security organs in cracking down on the behavior of assisting a cybercrime in practice, on the basis of manual review of cases, adding model algorithm can provide scientific guidance and basis for the investigation departments and procuratorial organs to handle and review cases.

Meanwhile, this paper also has some limitations, the original sample size of this paper is too small, which may lead to some variables that will have an impact on whether the case is prosecuted or not, and the significance between whether to prosecute or not is not obvious. In the future, we need to supplement the sample size to further test and improve the model results. In the future, combined with the analysis method of this paper, we can use machine learning and other methods to evaluate the harm results of assisting cybercrime cases, and realize the classification processing and result prediction of assisting cybercrime cases.

References

[1] L Tao, “The mode of illegal liability in criminal law of the network help behavior -- from the perspective of functionalism,” Politics and law, 2020 (03), pp. 113-124.
[2] Z Haimei, L Zihao, F Jinyuan, “The judicial expansion and the limiting path of offence of assisting a cybercrime: An Empirical Study of 1737 criminal judgment documents,” Journal of Xi'an University of Electronic Technology (SOCIAL SCIENCE EDITION), 2021,31 (01), pp. 66-71.
[3] Z Ming, “‘hot’ and ‘cold’: the judicial application picture of offence of assisting a cybercrime -- Based on the empirical analysis of 72 criminal judgment documents,” Application of law, 2019 (15), pp. 23-32.
[4] Y Haisong, “The pattern and regulation of the black and gray industrial chain of cyber crime,” Journal of the National Institute of prosecutors, 2021,29 (01), pp. 41-54.
[5] L Yanhong, “Criticism on the criminization of the perpetrator of help behavior to the cybercrime becoming a principal offender,” Legal and commercial research, 2016,33 (03), pp. 18-22.
[6] Y Zhigang, “Legislative exploration and theoretical analysis of the criminalization of accomplices becoming principal offenders -- an analysis from the perspective of the legislative orientation of " offence of assisting a cybercrime,” Legal Science (Journal of Northwest University of politics and law), 2017,35 (03),pp. 83-92.
[7] L Hong, “On the nature and application of " offence of assisting a cybercrime,” Application of law, 2017 (21), pp. 33-39.
[8] H Yueliang, “Application of "knowing" in offence of assisting a cybercrime,” Research on prevention of juvenile delinquency, 2016 (02), pp. 27-35.
[9] G Natasha, “THE USE OF CRIMINAL PROFILING IN CYBERCRIME INVESTIGATIONS,” ProQuest Dissertations Publishing, 2018.

[10] E Charalambou, R Bratskas, “Email forensic tools: A roadmap to email header analysis through a cybercrime use case,” Journal of Polish Safety and Reliability Association, 2016, pp.21-28.

[11] M. C. Michel, M. Carvalho, H. Crawford and A. C. Esterline, “Cyber Identity: Salient Trait Ontology and Computational Framework to Aid in Solving Cybercrime,” 2018 17th IEEE International Conference On Trust, Security And Privacy In Computing And Communications, 2018, pp.1242-1249.

[12] SW Brenner. “Cybercrime Investigation and Prosecution: the Role of Penal and Procedural Law”. Murdoch University Electronic Journal of Law, 2001.

[13] G Wenfeng, F Chao, G Xindong, “MOOC retreat forecast based on binary logic regression model,” Computer age, 2017, (12).