Frontiers in Business, Economics and Management
ISSN: 2766-824X | Vol. 19, No. 3, 2025

Global Trends of Cybercrime and National Cybersecurity Strategy Research

Zekai Deng [#], Yujing Xu [#], Wenjie Hu [#]

School of Geophysics and Information Technology, China University of Geosciences (Beijing), Beijing, 100083, China
#These authors contributed equally.

Abstract: With the increasing global internet connectivity, cybercrime has increasingly become a severe challenge faced by the international community. This paper aims to deeply analyze the impact of cybercrime and evaluate the effectiveness of national cybersecurity policies. By collecting and analyzing data on cybercrime incidents from the VERIS Community Database (VCDB), this paper reveals the geographical distribution, type characteristics, and development trends of cybercrime. The study finds that developed countries such as the United States, the United Kingdom, and Australia are the main targets of cybercrime, and there is a significant correlation between the frequency of cybercrime and the demographic data of countries. Based on these findings, this paper puts forward targeted policy recommendations to assist governments of various countries in formulating more effective cybersecurity strategies. This research provides important theoretical and practical bases for understanding the complexity of cybercrime and formulating countermeasures.

Keywords: Cybercrime; Global Trends; Cybersecurity Strategy; Demographic Data.

1. Introduction

Cybercrime encompasses a wide range of illegal activities where computers or networks are utilized as instruments, targets, or platforms for unlawful actions. This includes activities like electronic hacking and denial-of-service attacks [1]. As global internet usage has surged, so has the incidence of cybercrime (C. Lu and Chang 2007). These crimes have led to substantial financial damage for individuals, organizations, and nations. The FBI reported that cybercrime resulted in global losses of 4.2 billion in 2020, with projections reaching 4.2 billion in 2020, with projections reaching 6.9 billion by 2021 (Cldy 2023). These figures underscore the urgent need for further research on cybercrime [2]. Our goal is to offer specific recommendations for policies aimed at combating cybercrime.

Cybersecurity involves the coordination of resources, processes, and frameworks aimed at safeguarding cyberspace and its systems from events that disrupt the alignment between legal and actual property rights [3]. A report by Steve Morgan in Cybercrime Magazine, titled “Cyberwarfare in the C-Suite,” projected that cybercrime would inflict $6 trillion in damages globally in 2021. This amount would rank cybercrime as the third-largest global economy and one of the most pressing challenges facing humanity. To counter this, various strategies have been implemented, including political, legal, social, economic, and technological measures [4]. Despite substantial government investments over the past decade to enhance cybersecurity for both public and private systems, there is a lack of comprehensive understanding regarding the policies adopted worldwide and their efficacy in reducing cybercrime risks. Tsakalidis et al. [5] advocate for policy monitoring methods, similar to those used in public health and education, to better understand and control cybercrime. Georgiadou et al. [6] have developed a framework to assess an organization's cybersecurity readiness, emphasizing the human element. They argue that information security culture must adapt to new realities by offering relevant policies and solutions. Sarker et al. [7-9] explore how these adaptations can support data-driven decision-making to protect systems from cyber threats.

This study employed multiple linear and ARIMA models to predict cybercrime rates and created an evaluation model for cybersecurity policies. The Difference-in-Differences (DID) algorithm was used to measure the effectiveness of national cybersecurity policies. Additionally, Pearson correlation analysis was applied to examine the influence of factors like internet penetration, GDP, and average education levels on cybercrime, with results quantified through P value and correlation coefficient matrices.

2. Methodology

2.1. Dataset Description

In order to obtain basic data on cybercrime cases, this paper used the publicly available VCDB data source from https://verisframework.org/index.html and https://verisframework.org/vcdb.htm, and obtained 9,898 cybersecurity-related cases worldwide from 1971 to 2023. These data can be used for fitting and prediction.

Futhermore, this paper collected population data of the United States, the United Kingdom, Canada, and Australia from the United Nations Population Department, obtained the GDP data of these four countries for each year through the World Bank, acquired the Internet penetration rate of these four countries from the International Telecommunication Union, and also obtained the high school graduation rate of people aged 25 and above in these four countries from UNESCO. These data can be used for correlation analysis.

2.2. Data Preprocessing

Before 2000, there was a large amount of invalid data mixed in our database. Therefore, this project conducted manual screening of the data in the database through data retrieval and chose the 21st century as the key period for analysis. Moreover, due to the over - concentration of cybersecurity cases in space, there is a problem that although the number of cases varies greatly around the world, the

differentiation is not obvious in the sections with a small number of cases. In order to better reflect the relative attributes of data among countries, this paper took the logarithm of the number of cyber -crime cases in each country. These efforts will enhance the data visualization process. During the process of logarithmic processing, our projcet conducted data cleaning again, removing the data in the database that would result in a logarithm of 0 or a negative number, and ensuring that the number of cybersecurity cases in each country is limited to between 5 and 30 for easier comparative observation:

$$bubbleSizes = 5 + 25 \times \frac{\log_{10}(cases)}{\log_{10}(\max(cases))} \quad (1)$$

In Equation 1, "bubbleSizes" represents the "size of bubbles" in the heat map Figure 1. "cases" denotes the absolute number of cyber - attack cases in each country, and "max(cases)" represents the number of cases in the country with the largest number of cyber - attack cases.

The above work is the preliminary step of data processing. During the feature engineering and subsequent modeling stages, our project conducted targeted investigations and research based on the cleaned database.

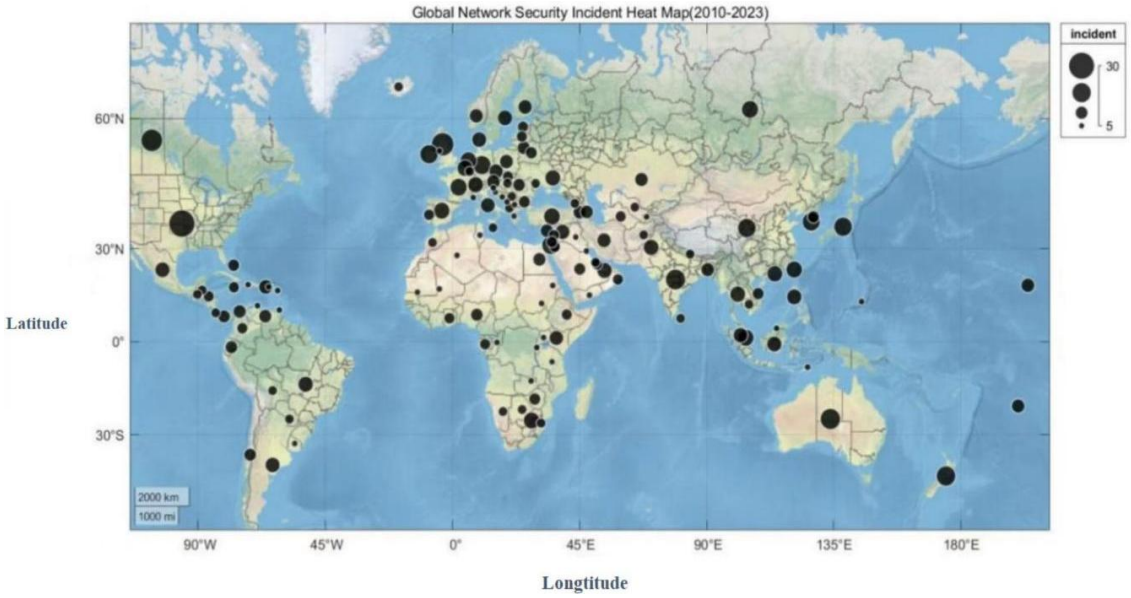

Figure 1. Global distribution heatmap of cybercrime

2.3. Analysis Methods

The Difference-in-Differences (DID) method is a widely used approach in the existing literature for assessing the impact of a certain policy. It typically involves dividing the samples into a treatment group (those affected by the policy) and a control group (those not affected by the policy). By comparing the differences in outcomes between the treatment and control groups both before and after the policy intervention, the DID method aims to isolate the causal effect of the policy. This approach combines insights from both before-after studies and treatment-control comparisons, thereby controlling for time-invariant unobservable characteristics and other confounding factors that may influence the outcomes [10]. This paper studied the publicly released national security policies, the number of cyber crime cases, and the Global Cybersecurity Index (GCI) data of four countries: the United States, the United Kingdom, Canada, and Australia. Based on this, this project established an effective evaluation model for cybersecurity policies, and used the Difference - in - Differences (DID) algorithm in it to evaluate the effectiveness of national cybersecurity policies.

PCC has been extensively used for variable selection, due to its simplicity and as it assists in recognizing the degree of correlation between input and output variables [11]. To explore the impact of specific national conditions on cybercrime, this project first quantified demographic data: internet access rate, national GDP, and education level. Subsequently, this project used the Pearson correlation analysis model to evaluate the degree of influence of each factor, and quantified the correlations using the p value matrix and the correlation coefficient matrix.

2.4. Policy Evaluation Framework

This paper analyzed the national security policies issued by different countries and compared them with the distribution of cybercrime. It was found that some policies and certain laws were particularly effective in prevention, prosecution, or other mitigation measures. The International Telecommunication Union's Global Cybersecurity Index (GCI) is a relatively good existing cybersecurity scoring standard. This paper obtained the case rate data (the percentage of the number of cases in a country in a given year to the total number of cases in that country within six years) of the United States, the United Kingdom, Canada, and Australia from 2018 to 2023.

Regarding the above visualization results, this paper suspect that there may be a certain correlation between the case rate and two factors: the year and the GCI score.

For the multiple linear regression model, this paper selected 70% of the data for model fitting and the other 30% of the data to test the fitted curve. It was found that the curve error was within an acceptable range. Therefore, this project retained this model and predicted the number of cases in the four countries for five years. The results are shown in Figure 2:

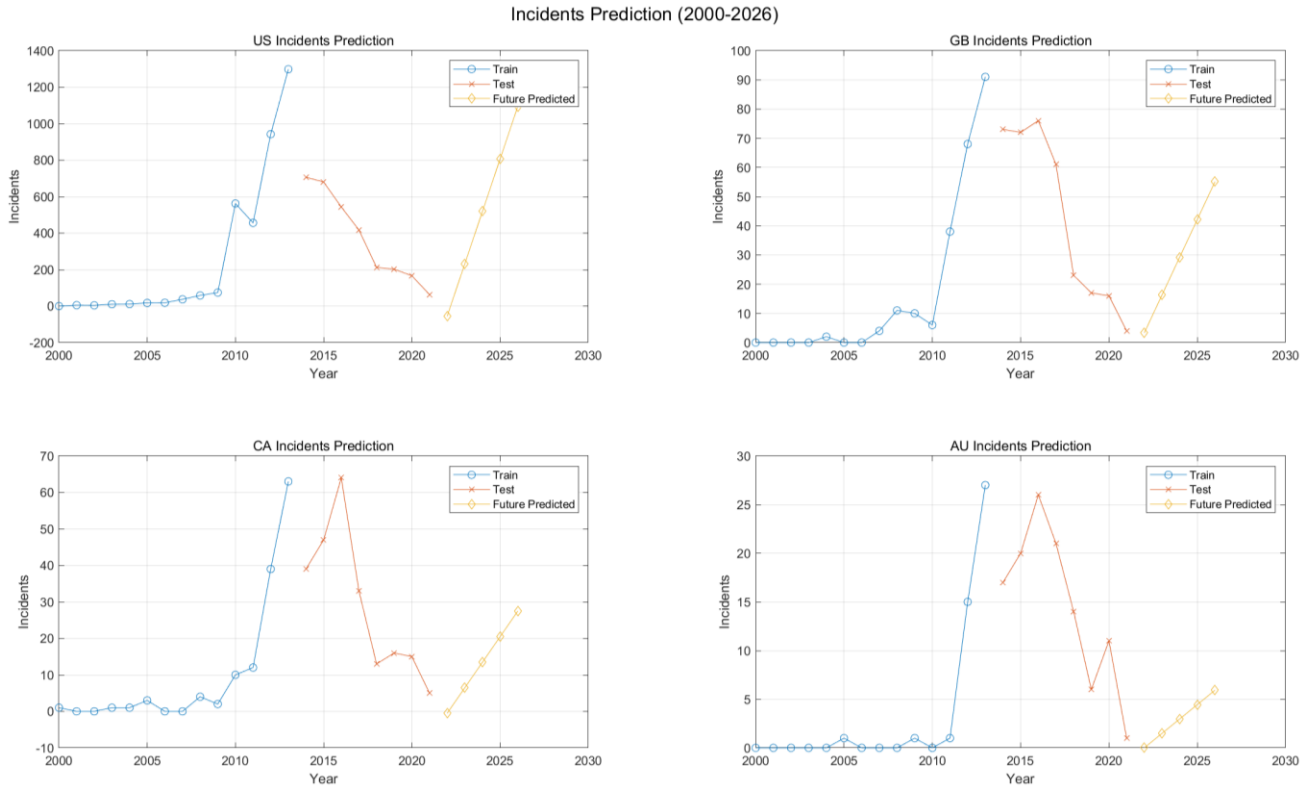

Figure 2. Multiple linear regression prediction

To further explore the impact of policies on cybersecurity attacks, when establishing the ARIMA model, this paper marked the promulgation time of relevant policies of each country in Figure 3. The four countries, namely the United States, the United Kingdom, Canada, and Australia, promulgated new policies in 2009, 2015, 2016, and 2019 respectively. Similar to the multiple regression model, this paper also used 70% of the data for fitting and 30% of the data for testing.

It is worth noting that, since the absolute number of cases in the United States is much larger than that in the other three countries, in order to appropriately display the correct patterns in the graph, this projct focused the research factors on the number of cases per capita each year. This paper found that, as the ARIMA model takes more into account the influence of time, its fitting effect is much worse than that of the multiple regression model.

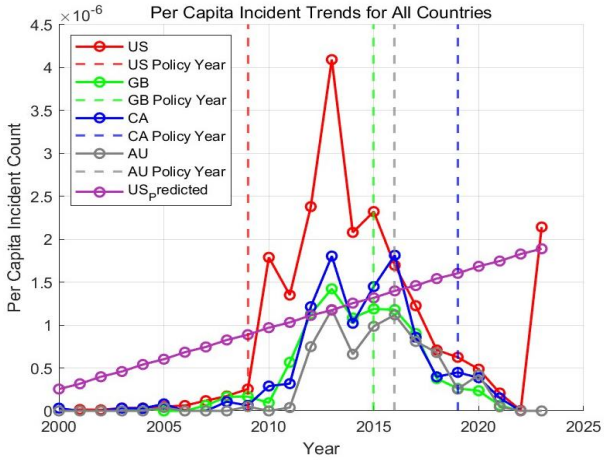

Figure 3. ARIMA model

Before the policy evaluation, this paper used polynomials of orders 2 to 6 to fit the number of cases per capita, and employed the Akaike Information Criterion (AIC) to select the optimal fitting order (the smaller the AIC value, the higher the fitting accuracy):

Table 1. AIC

order	2	3	4	5	6
AIC	35.8203	18.1583	19.6442	14.5764	51.2615

As can be seen from the table, the fifth - order polynomial is the optimal fitting degree. Based on the fifth - order polynomial fitting, our project used the DID (Difference - in - Differences) algorithm to evaluate the effectiveness of the policy. The treatment group is the number of cases per capita in the four countries. To eliminate the interference of the policy on the control group, this paper used the fifth - order polynomial prediction curve of the total number of cases before the policy was promulgated (from 2000 to 2008) as the control group.

The basic regression form adopted by the DID model is:

$$Y_{it} = \gamma_t POST_t + \gamma_i TREAT_i + \delta POST_t \times TREAT_i + \varepsilon_{it} \quad (2)$$

Through the fitting of the regression equation, our group obtained the coefficient δ of the DID interaction term as - 0.8. This result indicates that the policy has a certain impact on the overall regression model. Next, this paper analyzed the significance of the result. Our group will calculate the p - value of the interaction term coefficient for judgment, and its calculation formula is:

$$p = 2 \times \min(\phi(-|t|), 1 - \phi(|t|)) \quad (3)$$

Among them, Φ is the cumulative distribution function of the standard normal distribution, and t is t the statistic, with the calculation formula:

$$t = \frac{\hat{\beta}_3 - \beta_3}{SE(\hat{\beta}_3)} \tag{4}$$

Where SE represents the standard error.

This paper generally set the significance level at 0.05. A p value less than 0.05 indicates that the correlation is statistically significant. This paper finally obtained a comprehensive p value of 0.0456, which is less than 0.05, indicating that the impact of our policy factor on the model is significant.

In conclusion, our project's analysis shows that the policies implemented by the Australian government in 2016, such as the "Australian Cyber Security Strategy" which strengthened the cultivation of cybersecurity professionals through cybersecurity education and training, the "National Cyber Security Action Plan 2019 – 2024" of Canada in 2019, and the "UK Cyber Security Strategy (2015 - 2018)" of the UK government in 2015, which enhanced cybersecurity education and the cultivation of cybersecurity professionals, have proven to be effective.

2.5. National Condition Analysis Framework

Analyzing the relationship between national characteristics and cybersecurity is a multi - dimensional and complex issue. It requires us to establish a comprehensive analytical framework to explore the impact mechanisms of various national characteristics on cybersecurity. This paper has introduced multiple characteristic dimensions such as internet penetration rate, economic development level, and education level, and studied their interactions with cybercrime.

This project employed the Pearson correlation analysis model. This model is utilized to measure the linear correlation between two continuous variables. It requires that the data follow a normal distribution, and to some extent, a linear relationship should exist between the variables. The model determines the strength of the correlation by calculating the correlation coefficient and assesses whether the correlation is significant by computing the p value.

The calculation formula for the correlation coefficient r is:

$$\mathrm{r} = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum (X_i - \bar{X})^2 \sum (Y_i - \bar{Y})^2}} \tag{5}$$

The p value is calculated using the t distribution, and the formula is:

$$t = r\sqrt{\frac{n-2}{1-r^2}} \tag{6}$$

Where n is the sample size.

Our group created a correlation coefficient matrix and a p value matrix based on the Pearson correlation analysis model, as shown in Figure 4 and Figure 5.

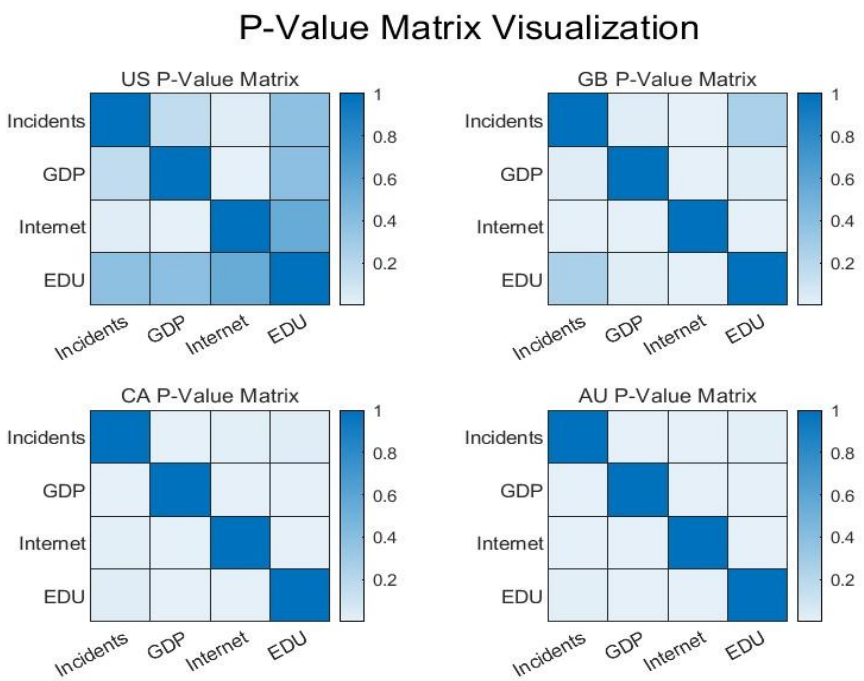

Figure 4. P-Value Matrix Visualization

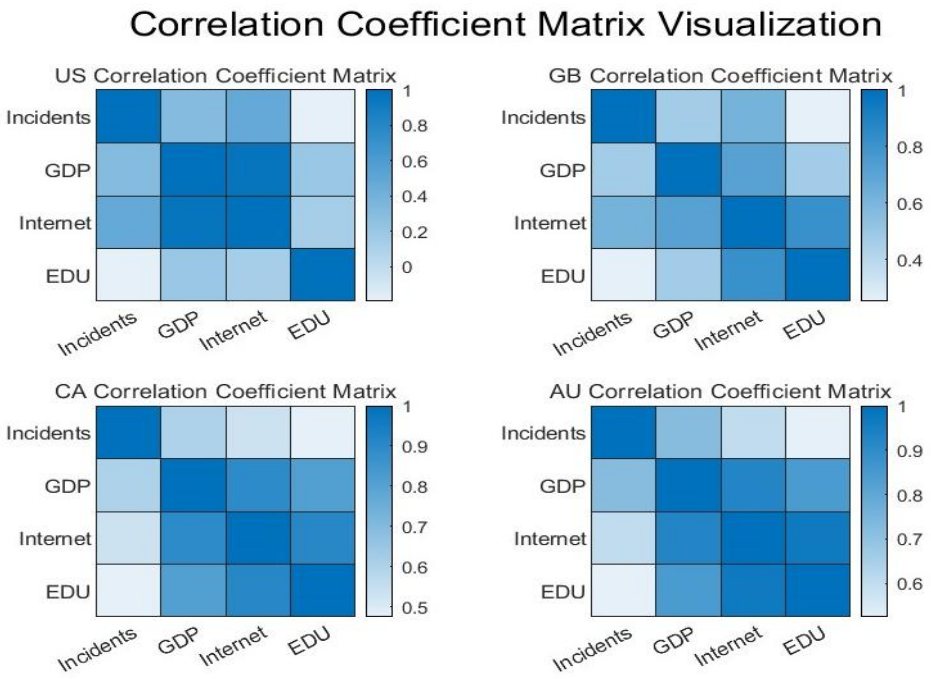

Figure 5. Correlation Coefficient Matrix Visualization

P-Value less than 0.05 indicates that the correlation is statistically significant. The larger the absolute value of the correlation coefficient, the stronger the linear correlation. A positive value indicates a positive correlation, while a negative value indicates a negative correlation. Both the p value and the correlation coefficient range from [0, 0.1].

Combining the relationship between the correlation coefficient and the p value of variables, this paper can draw the following conclusions: The number of cyber cases in these countries is positively correlated with the countries' internet access rate and GDP level. Except that in the United States, the education penetration rate is negatively correlated with the number of cybercrime, in the other three countries, there is a positive correlation. We speculate that the reason for this situation is that the United States achieved a relatively high education penetration rate many years ago, and fluctuations in the number of cybercrime have a greater impact on the final result.

3. Results and Discussion

3.1. Analysis of Cybercrime Trends

In this study, it was found that developed countries such as the United States, the United Kingdom, Canada, and Australia are high - risk countries for cybercrime. Developed countries have advanced information and communication infrastructure that is widely covered, a more complex network environment, and their citizens have more opportunities to access the Internet. Moreover, cross - border exchanges in developed countries are more frequent than those in developing countries. All these are important reasons why developed countries become the main targets of cybercrime.

3.2. Correlation Analysis of Demographic Data and Cybercrime Data

The study also found that the cybercrime situations in the countries under main study are greatly influenced by Internet usage. Besides the United States, the cybercrime situations in the other three countries are significantly affected by wealth. However, only in Canada and Australia are cybercrime situations strongly influenced by the education level.

Based on the correlation analysis of Internet usage, wealth, education level and the distribution of cyber crime, it can be seen that in high - risk countries for cyber security with a high GCI (Global Competitiveness Index, assume it refers to this if not specified otherwise) score, policies for popularizing the safe use of the Internet among the public should be implemented. Additionally, some economic - friendly policies can be carried out to increase the per capita GDP, so as to reduce cyber crime.

3.3. Evaluation of National Cybersecurity Policies

Based on the research on the differences in national security policies, it has been found that implementing policies to strengthen the cultivation of cybersecurity talents in high - risk cybersecurity countries with a high GCI score is highly conducive to reducing cybercrime. These countries can establish cybersecurity related majors at the university level and provide numerous opportunities for further study to students majoring in these fields, so as to enhance the overall level of national talents in the field of cybersecurity.

4. Conclusions

This paper reveals the geographical distribution of cybercrime and its correlation with demographic data by analyzing the impact of cybercrime and national cybersecurity policies. The study finds that developed countries are the primary targets of cybercrime, and the frequency of cybercrime is significantly correlated with demographic data such as population size and education level. Based on these findings, this paper puts forward policy recommendations such as strengthening cybersecurity education, enhancing technological capabilities, and intensifying international cooperation to assist governments in formulating effective cybersecurity policies and reducing the occurrence of cybercrime. However, due to limitations in data sources and the complexity of cybercrime, future research could expand data sources, delve deeper into the mechanisms and evolution laws of cybercrime, and evaluate

the effectiveness of policy interventions.

References

[1] Das S, Nayak T. Impact of cybercrime: Issues and challenges [J]. International journal of engineering sciences & Emerging technologies, 2013, 6(2): 142-153.

[2] Erdoğan M, Akmeşe Ö F. Bibliometric Analysis of Studies on Cyber Crimes Between 2000-2023 [J]. ADBA Computer Science, 2025, 2(1): 19-29.

[3] Craigen D, Diakun-Thibault N, Purse R. Defining cybersecurity [J]. Technology innovation management review, 2014, 4(10).

[4] Chinedu P U, Nwankwo W, Masajuwa F U, et al. Cybercrime Detection and Prevention Efforts in the Last Decade: An Overview of the Possibilities of Machine Learning Models [J]. Review of International Geographical Education Online, 2021, 11(7).

[5] Tsakalidis G, Vergidis K, Petridou S, et al. A cybercrime incident architecture with adaptive response policy [J]. Computers & Security, 2019, 83: 22-37.

[6] Georgiadou A, Mouzakitis S, Askounis D. Working from home during COVID-19 crisis: a cyber security culture assessment survey [J]. Security Journal, 2022, 35(2): 486-505.

[7] Sarker I H, Kayes A S M, Badsha S, et al. Cybersecurity data science: an overview from machine learning perspective [J]. Journal of Big data, 2020, 7: 1-29.

[8] Bharadiya J. Machine learning in cybersecurity: Techniques and challenges [J]. European Journal of Technology, 2023, 7(2): 1-14.

[9] Sarker I H. Machine learning for intelligent data analysis and automation in cybersecurity: current and future prospects [J]. Annals of Data Science, 2023, 10(6): 1473-1498.

[10] Yang Z, Fan M, Shao S, et al. Does carbon intensity constraint policy improve industrial green production performance in China? A quasi-DID analysis [J]. Energy Economics, 2017, 68: 271-282.

[11] Jayaweera C D, Aziz N. Reliability of principal component analysis and Pearson correlation coefficient, for application in artificial neural network model development, for water treatment plants [C]//IOP Conference Series: Materials Science and Engineering. IOP Publishing, 2018, 458(1): 012076.