Global Spatial Analysis and Policy Impact Evaluation of Cybercrime Based on Multimodal Modeling

Yutong Jiang[1,a,*], Xi Wang[2,b]

[1]School of Accounting, Southwestern University of Finance and Economics, Chengdu, China

[2]School of Finance and China Institute of Financial Studies, Southwestern University of Finance and Economics, Chengdu, China

[a]1724045399@qq.com, [b]1071640634@qq.com

*Corresponding author

Keywords: Cybercrime, Spatial Analysis, DID Model, Pearson Correlation, LSTM, Cybersecurity Policy Evaluation

Abstract: In the context of increasing global interconnectivity, cybercrime has emerged as a pervasive and borderless threat, necessitating effective international cybersecurity policies. This study investigates the spatial patterns and policy impacts of cybercrime using data from the VERIS Community Database (VCDB), encompassing over 10,000 cybersecurity incidents from 2006 to 2024.First, we employ descriptive statistics and heatmap analysis to examine the global distribution of cyber attacks, their success and failure rates, and reporting frequency across regions, highlighting disparities in defense capabilities and judicial systems. Second, to assess the effectiveness of national cybersecurity legislation, we construct a Difference-in-Differences (DID) model comparing treatment and control groups across 14 countries. This includes a baseline model, policy type and intensity effect models, and a dynamic time effect model, offering comprehensive insights into policy performance. Third, we integrate demographic indicators and use Pearson correlation analysis alongside an LSTM predictive model to explore and forecast the influence of population-level variables on cybercrime trends. Our findings provide valuable guidance for tailoring future cybersecurity strategies and contribute to the policy discourse at global forums such as the ITU Cybersecurity Summit.

1. Introduction

The accelerating digitalization of the global economy has brought about tremendous benefits in terms of efficiency, connectivity, and innovation. However, this rapid expansion of cyberspace has also given rise to escalating cybersecurity threats, with cybercrime becoming a pervasive and highly damaging global issue [1]. Its transnational nature, coupled with low detection and reporting rates, poses significant challenges to law enforcement, particularly in data-sensitive sectors such as finance, healthcare, and government services [2]. These issues threaten not only individual and corporate interests but also national security and global economic stability.

Despite the proliferation of cybersecurity frameworks and regulations implemented by various countries, there remains a lack of comprehensive, data-driven evaluations of their actual effectiveness [3]. Furthermore, global cybercrime patterns are unevenly distributed, and regional capabilities in cyber defense and judicial response vary significantly [4]. Understanding these spatial dynamics and policy outcomes is critical for designing effective and scalable intervention strategies [2][5].

This study leverages the VERIS Community Database (VCDB), which contains over 10,000 documented cyber incidents from 2006 to 2024, offering a unique opportunity to explore the geographic and temporal dimensions of cybercrime. We examine the global distribution of cyber incidents, analyze the effectiveness of national cybersecurity laws using a Difference-in-Differences (DID) approach, and explore how demographic and socioeconomic characteristics influence cybercrime trends. Finally, we integrate machine learning, specifically an LSTM model, to predict future cybercrime patterns based on these factors. By combining spatial analysis, econometric

  DOI: 10.25236/iiicec.2025.008

modeling, and predictive analytics, this paper aims to provide actionable insights for policymakers, reinforce global cyber governance, and contribute to ongoing efforts toward building a more resilient and secure digital ecosystem.

This paper makes the following key contributions:

(1) Global Spatial Profiling of Cybercrime: We provide a detailed spatial analysis of cybercrime incidents using descriptive statistics and heatmaps based on the VCDB dataset, uncovering geographic disparities in attack types, success rates, and reporting frequencies.

(2) Quantitative Evaluation of Cybersecurity Policies Using DID: By applying a Difference-in-Differences (DID) model to compare countries with and without cybersecurity legislation, we assess the real-world effectiveness of policy interventions. The model includes analyses of policy types, intensity, and temporal effects.

(3) Integration of Demographic Variables and Predictive Modeling: We explore the relationship between cybercrime and demographic factors using Pearson correlation coefficients. Furthermore, we develop an LSTM-based predictive model to forecast the potential impact of demographic changes on cybercrime trends.

(4) Policy Guidance for Global Governance: We present data-driven recommendations and a policy memorandum for international cybersecurity forums, such as the ITU Cybersecurity Summit, supporting global collaborative efforts in combating cyber threats.

2. Methodology

2.1. Cyber Attacks

Cyber attacks demonstrate complexity and global characteristics. Developed countries like the United States have become major victims due to technological, economic, and geopolitical factors. Additionally, countries such as Russia are more frequently identified as sources of attacks. However, a large number of attack sources and victim situations remain difficult to attribute specifically, indicating that the anonymity and technical complexity of cyber attacks continue to be significant global challenges [6].

Furthermore, regional analysis reflects the general pattern of cyber attacks worldwide. Through the analysis and visualization of victim country statistics, we can clearly demonstrate the geographical distribution characteristics of global cyber attacks. The regional distribution of cybersecurity incidents is displayed in Figure 1.

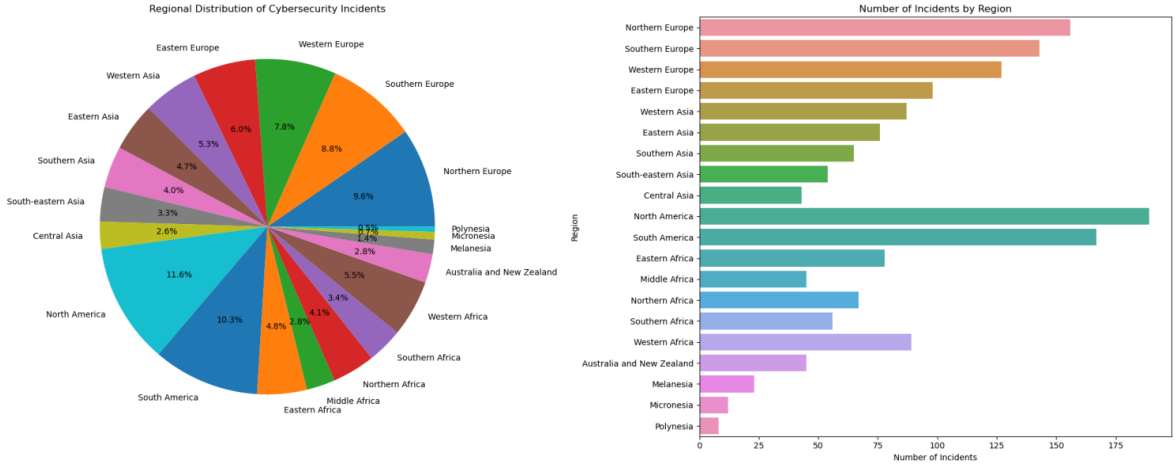

Figure 1 Regional Distribution of Cybersecurity Incidents.

The results clearly demonstrate the significant geographical disparity in cyberattacks across the globe. North America and Europe emerge as the primary target regions for cyber attacks worldwide. In contrast, African regions report relatively fewer cybercrime incidents. This pattern reveals the uneven geographical distribution of cybercrime activities.

In the global cybersecurity landscape, developed and developing countries demonstrate distinctly different roles. Developed countries have extremely high internet penetration rates, which has led to

deep digitalization across all sectors of society, thus creating vast attack surfaces. Meanwhile, the changes in developing countries are equally significant, as they substantially influence the evolution of global cybersecurity. Therefore, investigating the situation in developing countries holds considerable significance, as shown in Figure 2:

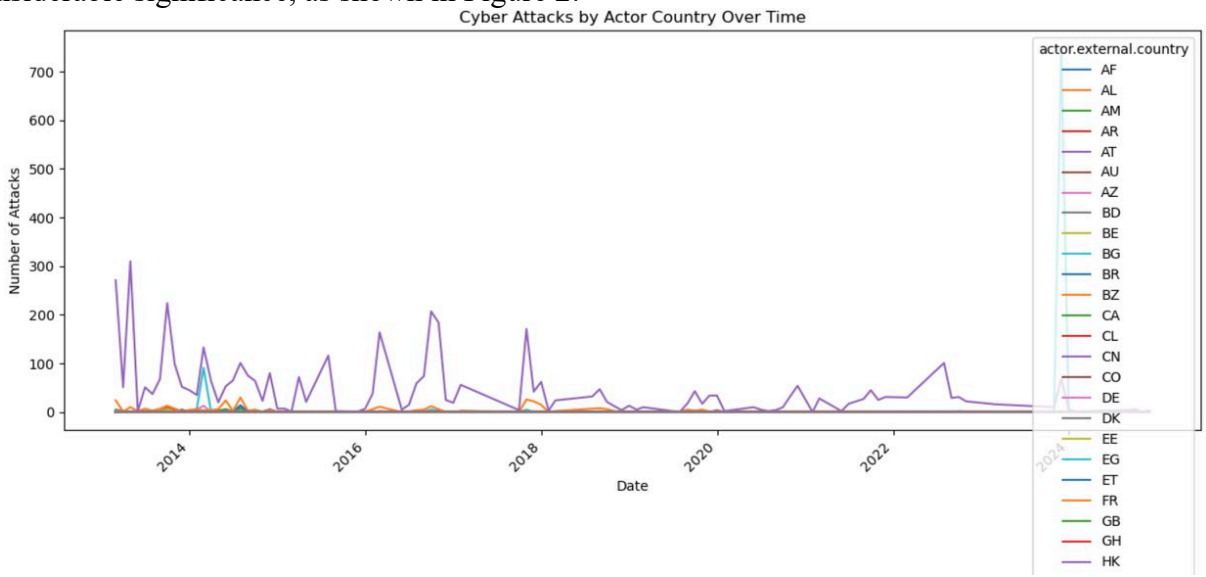

Figure 2 Cyber Attacks by Actor Country Over Time (2014-2024).

According to Figure 2, the origins of global cyberattacks show a trend of concentration in developing countries. From 2014 until before 2024, attackers primarily originated from developed countries. However, in 2024, there was a sudden surge in the number of attackers from developing countries. This shift is closely related to several factors: These developing countries have produced many individuals with computer expertise during their rapid digital transformation. Yet, due to their relatively weak economic foundation and inadequate cybersecurity regulations, some of these skilled individuals have turned to cybercriminal activities.

2.2. The Effectiveness of Cybersecurity Policies on Cybercrime Prevention and Control

Policy Implementation: The research sample consists of 13 countries divided into two groups: the treatment group comprises 8 countries that have implemented cybersecurity policies, while the control group includes 5 countries that have not yet implemented such policies [7]. The basic time effect model is used to capture the overall changes in effectiveness before and after policy implementation. The model expression is as follows.

$$Y_{it} = \beta_0 + \beta_1(Treatment_i \times Post_t) + \beta_2 Treatment_i + \beta_3 Post_t + \gamma X_{it} + \alpha_i + \lambda_t + \varepsilon_{it} \quad (1)$$

The binary variables Treatment_i and Post_t indicate whether country i has implemented cybersecurity policies and whether it is in the post-implementation period (1=yes, 0=no). β_1 reflects the net effect of cybersecurity policy implementation on cybercrime levels, while β_2 and β_3 control for the inherent differences between treatment and control groups and the impact of common time trends, respectively. Here, Y_{it} represents the cybercrime level in country i during period t. The country fixed effects α_i control for time-invariant country characteristics, and the time fixed effects λ_t capture common time trends affecting all sample countries. The control variables X_{it} include key factors that may influence cybercrime levels.

Policy Types: To thoroughly investigate the impact of cybersecurity policies on cybercrime, this study develops models based on two dimensions: policy types and policy intensity. First, from the perspective of policy types, cybersecurity policies are classified into three main categories: preventive measures, punitive measures, and technical requirements. The policy type effect model is established as follows:

$$
\begin{aligned}
\mathrm{Y}_{\mathrm{it}} = \ \beta_0 \ + \ \beta_{11}(prevent_{edu_i} \times \ Post_t) + \\
\beta_{12}(prevent_{riski} \times \ Post_t) \ + \beta_{13}(prevent_{monitori} \times Post_t) + \\
\beta_{21}(punish_{admini} \times Post_t) \ + \beta_{22}(punish_{criminal\ i} \times Post_t) + \\
\beta_{31}(tech_{standardi} \times Post_t) + \beta_{32}(tech_{securityi} \times Post_t) \\
+ \ \beta_{33}(tech_{auditi} \times Post_t) + \gamma \mathrm{X}_{\mathrm{it}} + \alpha_{\mathrm{i}} + \lambda_{\mathrm{t}} + \varepsilon_{\mathrm{it}}
\end{aligned} \tag{2}
$$

$\beta_{1m}, \beta_{2n}, \beta_{3m}$ respectively measure the effectiveness of preventive, punitive, and technical policies after their implementation (where m=1,2,3 and n=1,2).

Policy Intensity: To further evaluate the completeness and strength of policy implementation, we construct policy intensity indicators. We assign weights to specific measures within each policy category and sum them up to obtain the corresponding policy intensity index.

$$
\begin{cases}
Intensity_{prev_i} = w_{11} \times prevent_{edu_i} + w_{12} \times prevent_{risk_i} + w_{13} \times prevent_{monitor_i} \\
Intensity_{puni_i} = w_{11} \times punish_{admin_i} + w_{12} \times punish_{criminal\ i} \\
Intensity_{tech_i} = w_{11} \times tech_{standard_i} + w_{12} \times tech_{security_i} + w_{13} \times tech\ _{audit_i}
\end{cases} \tag{3}
$$

Therefore, the policy intensity effect model is expressed as follows.

$$
\begin{aligned}
\mathrm{Y}_{\mathrm{it}} = \ \beta_0 \ + \ \beta_1(\mathrm{Prevention}_i \times Intensity_{previ} \times \ Post_t) \\
+ \ \beta_2(\mathrm{Prosecution}_i \times Intensity_{punii} \times Post_t) + \\
\beta_3(\mathrm{Technical}_i \times \ Intensity_{techi} \times \ Post_t) \\
+ \gamma \mathrm{X}_{\mathrm{it}} + \alpha_{\mathrm{i}} + \lambda_{\mathrm{t}} + \varepsilon_{\mathrm{it}}
\end{aligned} \tag{4}
$$

$\mathrm{Prevention}_i$, $\mathrm{Prosecution}_i$ and $\mathrm{Technical}_i$ where the variables respectively represent the overall implementation status of the three types of policies.

2.3. LSTM Model

The model uses 64 LSTM units to process and analyze complex nonlinear relationships between input features. Each LSTM unit contains control mechanisms, including input gate, forget gate, and output gate components. To simplify the subsequent data processing, we set the return_sequences parameter to False. This means the LSTM layer only outputs the hidden state of the last time step instead of the complete sequence output. After the LSTM layer, we implemented a regularization layer using dropout mechanism to control overfitting. For the fully connected section, the model uses a dense layer with 32 neurons and applies ReLU activation function for nonlinear transformation. The output layer has 2 neurons, which corresponds to the binary classification problem.

The model training uses the Adam optimizer to minimize the loss function for given parameters. For each parameter, Adam calculates first-moment estimate (mean) m_t and second-moment estimate (variance) v_t of the gradients. This helps adjust the learning rate for each parameter to better suit their scale:

$$
m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t \tag{5}
$$

$$
v_t = \beta_2 v_{t-1} + (1 - \beta_1) {g_t}^2 \tag{6}
$$

Here, g_t represents the gradient at time step t, while $\beta_1$andβ_2 are decay rates that are typically close to 1.

The first and second moment estimates undergo bias correction to ensure greater accuracy during early training. This is necessary because m_t and v_t may be underestimated initially. The bias-corrected expressions are as follow.

$$
\widehat{m}_t = \frac{m_t}{1 - \beta_1^t} \tag{7}
$$

$$\hat{v}_t = \frac{v_t}{1-\beta_2^t} \tag{8}$$

The parameters are updated using these corrected first and second moment estimates.

The initial learning rate is set to η=0.01, and the regularization term is determined through cross-validation. The dropout rate is set to $p_{dropout} = 0.5$. In the model training, the L2 regularization term and the loss function with regularization can be expressed as follow.

$$\Omega_{(\theta)} = \frac{\lambda}{2}\sum_i {\theta_i}^2 \tag{9}$$

$$L = \frac{1}{N_{batch}}\sum_{i=1}^{N_{batch}}(y_i - \hat{y}_i)^2 + \Omega_{(\theta)} \tag{10}$$

In these equations, y_i represents the true value of the ith sample, $\hat{y}_i$ is the model's predicted value, and N_{batch} is the batch size.

The model is trained for 200 epochs with a batch size of 32, using Mean Squared Error (MSE) as the evaluation metric.

$$MSE = \frac{1}{N}(y_i - \hat{y}_i)^2 \tag{11}$$

Where N is the total number of samples in the test set, y_i is the true value of the i-th sample, and $\hat{y}_i$ is the corresponding model prediction. The training effect and iteration process are shown in Figure 3.

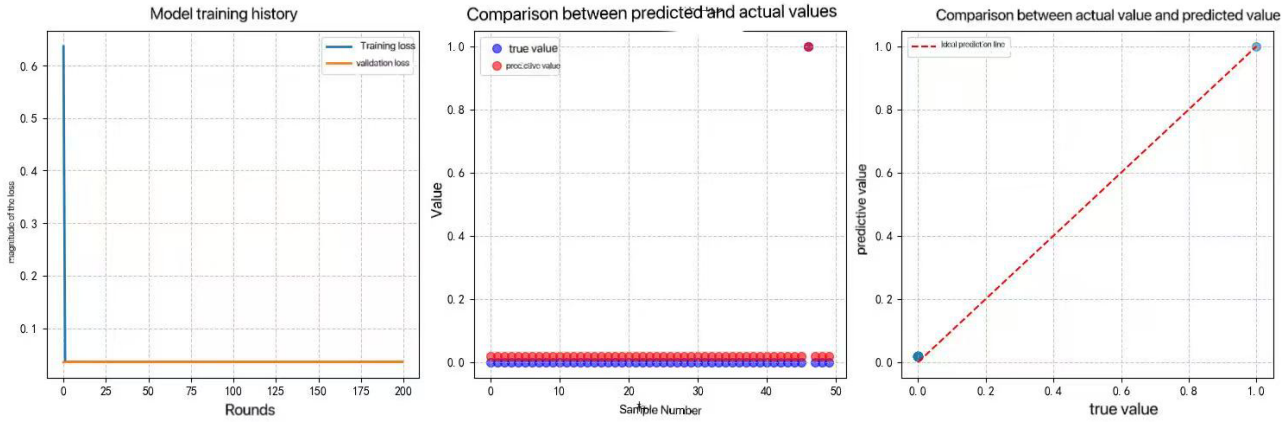

Figure 3 Correlation Heatmap of Socioeconomic Indicators and Cyberattacks in Major Counties.

3. Results

3.1. Incident Type Distribution

As shown in the Figure 4, the heatmap analysis reveals significant cyber infiltration activity and high attack success rates in Russia, likely due to lenient law enforcement and geopolitical factors. Similarly, "Unknown" regions show high levels of data exfiltration and infiltration. In contrast, China, Colombia, and the UK exhibit minimal attack incidents, attributed to their strong cybersecurity defenses and robust regulatory frameworks that effectively mitigate cybercrime.

Developed regions like North America, the EU, and East Asia report higher cybercrime cases due to better reporting systems, specialized law enforcement, and greater cybersecurity awareness among citizens, who are more likely to report incidents. In contrast, other regions face challenges such as weak legal frameworks, poor reporting mechanisms, and limited law enforcement, leading to significant underreporting and a "dark figure" in cybercrime statistics.

Effective cybercrime prosecution faces key challenges, requiring three main elements: (1) comprehensive cybercrime laws, specialized judicial procedures, and enforcement agencies; (2) professional technical teams and advanced equipment for handling electronic evidence; and (3) strong international law enforcement cooperation to address jurisdictional conflicts arising from the transnational nature of cybercrime. Regions with robust prosecution capabilities, such as the United

States, China, the EU, Russia, and Singapore, meet these criteria due to their advanced legal, technical, and cooperative frameworks.

Based on the above information, we can identify several patterns: different countries face different types of cybercrime, and due to the transnational nature of cybercrime, the reporting location and prosecution jurisdiction are frequently inconsistent.

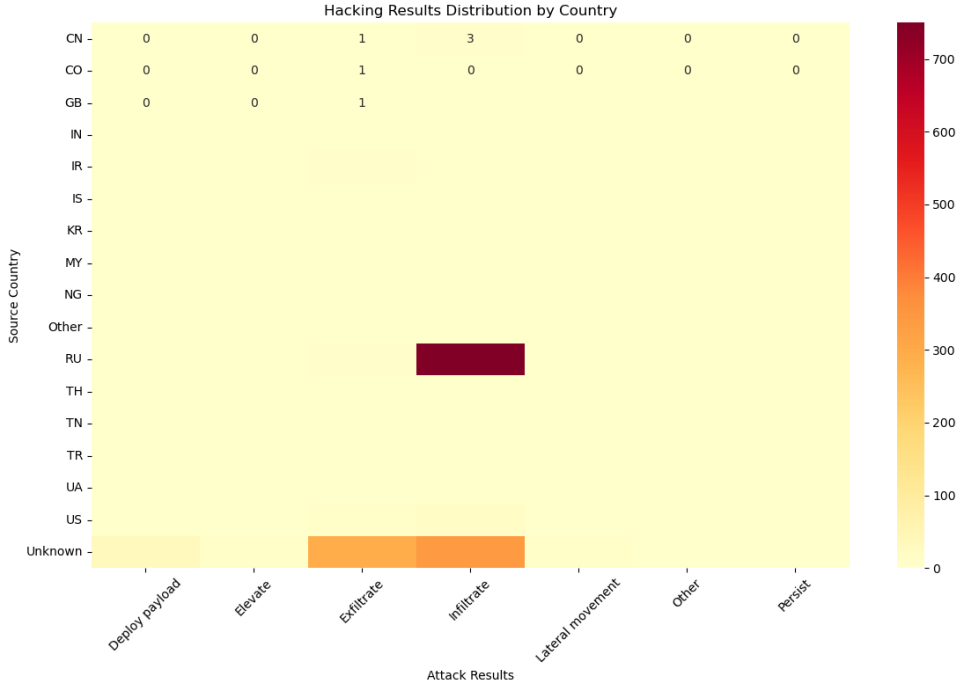

Figure 4 Heatmap of Hacker Results Distribution by Country.

3.2. Regression Output Analysis

Policy Implementation Impact: The interaction term coefficient β_1= -16.2178 demonstrates that policy intervention has a significant inhibitory effect on cybersecurity incidents. After controlling for other factors, compared to the control group, the treatment group implementing cybersecurity policies experienced an average reduction of approximately 16.32 cyber attacks. Moreover, with $p<0.001$, β_1 is highly significant at the 1% level, indicating that the observed reduction is extremely unlikely to be caused by random factors. Additionally, we selected a 95% confidence interval to verify the reliability of the policy effect. This interval lies entirely in the negative range (-24.789, -7.897) and does not cross zero, further confirming the reliability of the policy intervention effect.

Policy Type Impact: After establishing the reliability of policy intervention, the influence of each policy type is as shown in Table 1:

Table 1 Differential Analysis Table of Cybersecurity Policy Implementation Effects.

Policy Effect	$Policy_{Type}$	Avg_{Effect}	$Significant_{Positive}$	$Significant_{Negative}$	$Total_{Countries}$
0	$prevent_{edu_{effect}}$	-4.180091	0	4	9
1	$prevent_{risk_{effect}}$	-4.252361	0	4	9
2	$prevent_{monitor_{effect}}$	-3.732896	0	3	9
3	$punish_{admin_{effect}}$	-4.252361	0	4	9
4	$punish_{criminal_{effect}}$	-0.604115	0	2	9
5	$tech_{standard_{effect}}$	-4.252361	0	4	9
6	$tech_{security_{effect}}$	-3.687672	0	2	9
7	$tech_{audit_{effect}}$	0.000000	0	0	9

Impact of policy intensity: In addition to the type of policy affects the frequency of cyber-attacks, so does the intensity of the policy. After summarizing and counting the PDFs, the intensity metrics can be obtained, and some of the metrics are first shown in Table 2 below:

Table 2 Differential Analysis Table of Cybersecurity Policy Strength Effects.

Policy Effect Summary	prevent_monitor _amount	punish_admin _amount	punish_criminal _amount	tech_standard _amount	tech_security _amount
US	5	3	0	1	65
GB	0	57	23	124	0
CA	0	14	4	12	0
AU	1	1	0	18	2

The conclusion can be drawn after solving and visualization using python:Except for the two metrics “tech_security_amount_effect” and “tech_audit_amount_effect”, which have no significant contribution to the number of attacks, the remaining six metrics have a significant reduction effect on the number of cyber attacks. The metric “prevent_edu” is selected for visualization and can be obtained in Figure 5 below:

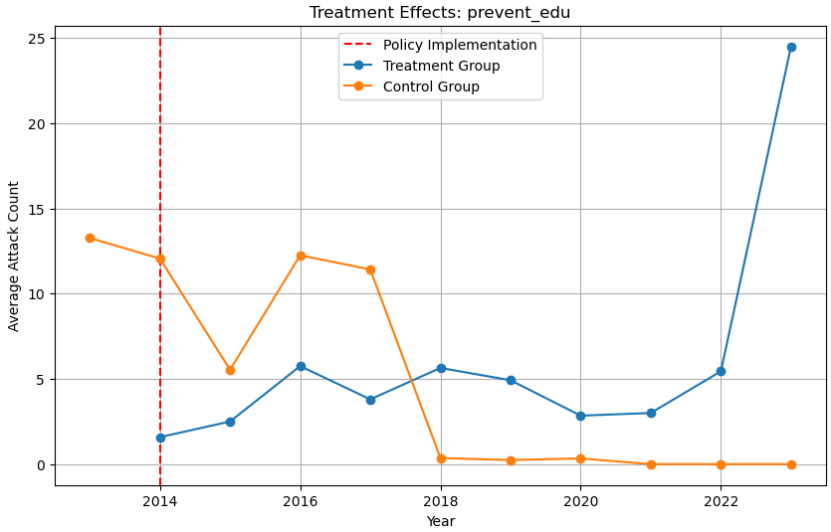

Figure 5 Effects of Cybersecurity Prevention Education Policies (2014-2023).

Control group (blue line): the average number of assaults in the treatment group gradually increases after a brief decline following policy implementation (2014), and shows a significant upward trend especially after 2022.

Treatment group (orange line): the average number of assaults in the control group decreases significantly after the policy is implemented and falls to almost zero around 2018, remaining low in the following years.

4. Conclusions and Future Work

This study presents a comprehensive analysis of global cybercrime patterns and the effectiveness of cybersecurity policies, using empirical data and advanced modeling techniques. By leveraging the extensive VERIS Community Database (VCDB), we identified key spatial trends in cybercrime distribution, revealing significant regional disparities in attack types, success rates, and reporting behavior [8]. Through descriptive statistics and visual analysis, we highlighted the uneven capacities of national cybersecurity infrastructures and the differentiated impacts of judicial systems.

Our application of the Difference-in-Differences (DID) model across treatment and control groups of countries provided robust evidence on the impact of cybersecurity legislation [9]. The analysis demonstrated that national-level policy interventions generally reduce cybercrime occurrences, but the degree of effectiveness varies depending on the policy type, implementation intensity, and timing. These findings contribute to a deeper understanding of which legislative measures yield the most tangible improvements in cyber resilience [10].

Additionally, by integrating demographic and socioeconomic indicators, we investigated their correlation with cybercrime occurrence using Pearson correlation coefficients. The results show meaningful associations between factors such as internet accessibility and economic development

levels with cybercrime frequency. To extend this analysis, we employed an LSTM-based forecasting model to predict future cybercrime trends, offering a forward-looking tool for anticipating policy needs and resource allocation.

In future work, we aim to refine our models by incorporating more granular policy metadata, such as specific regulatory clauses, enforcement levels, and cross-sectoral coordination mechanisms. We also plan to explore causal inference techniques beyond DID, including synthetic control methods and instrumental variable approaches, to strengthen policy impact evaluation. Finally, integrating real-time cybersecurity threat intelligence and regional institutional metrics could further enhance the responsiveness and accuracy of cybercrime prediction and policy design.

References

[1] Chen S, Hao M, Ding F, et al. Exploring the global geography of cybercrime and its driving forces[J]. Humanities and Social Sciences Communications, 2023, 10(1): 1-10.

[2] Tok Y C, Chattopadhyay S. Identifying threats, cybercrime and digital forensic opportunities in Smart City Infrastructure via threat modeling[J]. Forensic Science International: Digital Investigation,2023, 45: 301540.

[3] Wright D C S. Geographical Aspects of Cybercrime: A Literature Review[J]. Available at SSRN 4521486, 2023.

[4] Бліхар В, Цимбалюк М, Грещук Г, et al. Current state and development trends of international law in the context of economic and legal analysis of financial measures to combat cybercrime in the global environment[J]. Financial and credit activity problems of theory and practice, 2022, 6(47): 378-387.

[5] Collier B, Thomas D R, Clayton R, et al. Influence, infrastructure, and recentering cybercrime policing: Evaluating emerging approaches to online law enforcement through a market for cybercrime services[J]. Policing and Society, 2022, 32(1): 103-124.

[6] Levi M. Assessing the trends, scale and nature of economic cybercrimes: overview and Issues: In Cybercrimes, cybercriminals and their policing, in crime, law and social change[J]. Crime, law and social change, 2017, 67: 3-20.

[7] Akinbowale O E, Klingelhöfer H E, Zerihun M F. Analysis of cyber-crime effects on the banking sector using the balanced score card: a survey of literature[J]. Journal of Financial Crime, 2020, 27(3): 945-958.

[8] Saini H, Rao Y S, Panda T C. Cyber-crimes and their impacts: A review[J]. International Journal of Engineering Research and Applications, 2012, 2(2): 202-209.

[9] Roshankar R, Keyvanpour M R. GeoCrime Analytic Framework (GCAF): A Comprehensive Framework for Dynamic Spatial Temporal Crime Analysis[J]. Applied Spatial Analysis and Policy, 2025, 18(1): 1-42.

[10] Peters A, Jordan A. Countering the cyber enforcement gap: Strengthening global capacity on cybercrime[J]. J. Nat'l Sec. L. & Pol'y, 2019, 10: 487.