WEP Warwick Evans Publishing

Transactions on Computer Science and Intelligent Systems Research
ISSN: 2960-1800, eISSN: 2960-2238 | Volume 11
TPCEE 2025

Research on the Spatial Distribution Characteristics of Global Cybercrime and Prevention Strategies Based on Geographically Weighted Regression

Xin Yu [1], Haofei Li [2,*], Jiangyu Sun [3]

[1] School of Business Administration, Henan Polytechnic University, Jiaozuo, China, 454000

[2] School of Computer Science and Technology, Henan Polytechnic University, Jiaozuo, China, 454000

[3] School of Physics and Electronic Information, Henan Polytechnic University, Jiaozuo, China, 454000

* Corresponding Author Email: 15538905773@163.com

Abstract. Against the backdrop of deep global digital integration, the transnational nature and spatial heterogeneity of cybercrime have intensified governance challenges. This study employs a Geographically Weighted Regression (GWR) model to integrate cybercrime data from 129 countries (including 12 indicators such as crime incidents, success rate, and prevention rate) from the International Telecommunication Union (ITU) 2023 Global Cybersecurity Index (GCI) and the VERIS Community Database (VCDB), constructing a three-dimensional analytical framework of "Spatial Distribution - Influencing Factors - Policy Efficacy". The results show that four countries including the United States and the United Kingdom are high-incidence target countries for cybercrime (crime incidents ≥ 138 times), while five countries such as Singapore and South Korea achieve a crime prevention rate of over 89%. Economic level (GDP), internet penetration, and cybersecurity policy efficacy (GCI index) exhibit significant spatially heterogeneous impacts on crime distribution, with well-developed and strictly enforced legal policies accounting for 80% of the crime inhibition effect. Geospatial proximity and technical investment intensity show positive inhibitory effects, and regional collaboration mechanisms can enhance prevention efficiency by over 30%. The study reveals the dual laws of cybercrime—"economic centers attracting crime" and "spatial differentiation of policy efficacy"—providing data support for optimizing global cybersecurity policies.

Keywords: Cybercrime; Geographically Weighted Regression (GWR); Spatial Heterogeneity; Transnational Prevention; Policy Efficacy.

1. Introduction

The globalization of digital technology has given rise to the industrialization of cybercrime. According to the 2023 report by the United Nations Office on Drugs and Crime (UNODC), the economic losses from cybercrime reached 6.4 trillion US dollars, with over 70% of cases involving transnational collaboration. There are significant disparities in protective capabilities among countries: Nordic countries allocate 1.2% of their GDP to cybersecurity budgets, resulting in a crime success rate of only 18%; in contrast, African countries have an average budget of 0.15%, with a success rate as high as 76%. Addressing the challenge of "inefficient governance in high-risk areas" requires in-depth analysis from a spatial perspective. These global trends and disparities underscore the urgent need for effective cybersecurity strategies, which this study aims to address using specific data from the International Telecommunication Union (ITU) and the VERIS Community Database (VCDB) to model cybercrime dynamics.

Traditional Ordinary Least Squares models often overlook spatial non-stationarity, leading to limitations in explaining complex regional crime patterns[1]. While the Geographically Weighted Regression (GWR) model has proven effective in regional crime analysis by accounting for spatial heterogeneity[2], global-scale studies frequently lack the integration of policy efficacy. This research

addresses this gap by being the first to incorporate the five dimensions of the ITU GCI, constructing a "space-policy" interaction model.

By constructing a Geographically Weighted Regression (GWR) model, this study aims to reveal the spatial distribution characteristics of global cybercrime, identify high-incidence target countries and highly effective prevention countries, analyze aspects such as crime types, occurrence time, location, contributing factors, and impacts of cybercrime, and extract key influencing factors. A key objective is also to integrate the spatial effects of policy efficacy, providing a scientific basis for formulating transnational cybersecurity policies.

2. Materials and Methods

2.1. Data Acquisition and Preprocessing

2.1.1. Data Composition:

(1) Crime Data: From the VERIS Community Database (VCDB,

https://verisframework.org/vcdb.html), containing 213,457 cybercrime records with fields including case type, country of occurrence, and success status, at the national spatial resolution.

(2) Policy Data:

The 2023 Global Cybersecurity Index (GCI) from the International Telecommunication Union (ITU, https://datahub.itu.int), including 5 first-level indicators such as Laws, Tech, and Cooperation, covering 194 countries (score range: 0–100).

(3) Economic and Technical Data:

2023 GDP per capita from the World Bank (http://data.worldbank.org) and internet penetration rates from Internet World Stats (http://www.internetworldstats.com).

2.1.2. Processing Procedures:

(1) Spatial Matching: Geographic coordinates were matched based on ISO 3166-1 country codes, and boundary errors were corrected using Google Earth Engine to ensure spatial analysis accuracy.

(2) Data Cleaning: Countries with over 30% missing crime data were excluded, retaining 129 valid samples.

(3) Standardization: The GCI index was standardized using the Z-score method:

$$Z_{GCI} = \frac{X_{GCI} - \mu}{\sigma} \tag{1}$$

where μ is the global mean and σ is the standard deviation.

2.2. Methodology Introduction

Geographically Weighted Regression (GWR) Model: GWR is a local regression model for spatial non-stationarity, allowing regression coefficients to vary with spatial location [3-4]. The formula is:

$$y_i = \beta_0(u_i, v_i) + \sum_{k=1}^{m} \beta_k(u_i, v_i) \cdot x_{ik} + \varepsilon_i \tag{2}$$

where:

y_i: Cybercrime rate in country *i* (number of cases per million population);

x_i: Independent variables (GDP per capita, Tech index, Law index);

(u_i, v_i): Geographic centroid coordinates of country *i*;

$\beta_k(u_i, v_i)$: Local regression coefficient of the *k-th* independent variable, reflecting its impact on the crime rate at location (u_i, v_i);

3. Model Construction and Solution

3.1. Model Development

The development of the Geographically Weighted Regression (GWR) model is crucial for understanding the spatial heterogeneity of cybercrime. This process primarily involves two key steps: the precise design of the spatial weight matrix and the optimal selection of bandwidth. The spatial weight matrix, as further detailed in Section 3.1.1, quantifies the influence of neighboring countries on each other's cybercrime rates based on their geographical proximity. Bandwidth optimization, discussed in relation to the Gaussian kernel function, determines the extent of this spatial influence, ensuring that the local regression models accurately capture the spatially varying relationships between cybercrime and its influencing factors. Through these rigorous spatial econometric methods, the GWR model effectively achieves a heterogeneous analysis of crime distribution, allowing for a more accurate and localized understanding of global cybercrime patterns and the effectiveness of prevention strategies.

3.1.1. Spatial Weight Function

The Gaussian kernel function is used to characterize the spatial correlation strength between countries, with the formula:

$$w_{ij} = \exp\left(-\frac{d_{ij}^2}{2h^2}\right) \tag{3}$$

where:

$d_{ij}$is the spherical distance between country *i* and *j* , reflecting the impact of geographical proximity on crime propagation;

h is the bandwidth parameter controlling the spatial scope of local regression.

3.1.2. Independent Variable Selection

Three categories of core influencing factors are included to construct the explanatory variable matrix:

(1) Economic Level: GDP per capita (log-transformed to eliminate dimensional differences);

(2) Technical Capacity: ITU GCI Technical Measures Index (Tech, reflecting the intensity of technical deployments such as firewalls and intrusion detection systems);

(3) Policy Efficacy: ITU GCI Legal Completeness Index (Law, covering the integrity of cybersecurity legislation and enforcement strength).

This study deeply discusses the five indicators of the GCI index—Average Legal Measures, Average Technical Measures, Average Cooperation Measures, Average Organizational Measures, and Average Capacity Development Measures[5-7]. Through ITU's GCI scores for each country, the weight relationships of the five indicators in national cybersecurity policies are analyzed, as shown in Fig. 1.

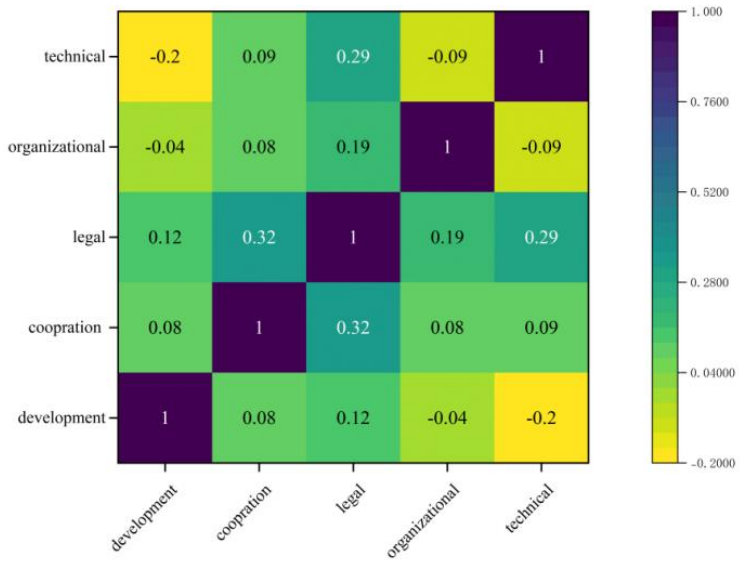

Figure 1. Weight relationships of the five GCI indicators in national cybersecurity policies

This figure illustrates the correlation coefficients between the five GCI indicators, providing insights into their interdependencies within national cybersecurity policies.

Based on this, we statistically analyzed the number of countries where each indicator served as the most influential and least influential factor in cybersecurity policies, with results as shown in Fig. 2.

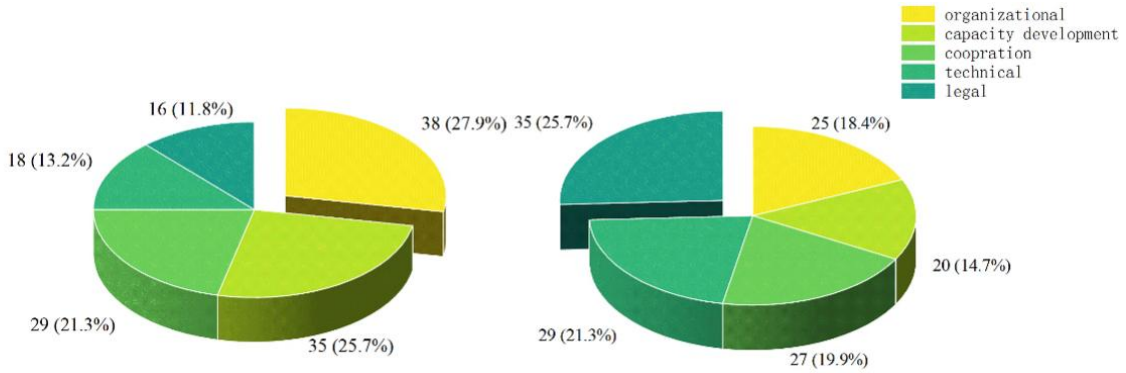

Figure 2. Statistical count of countries where the five GCI indicators serve as the most and least influential factors

This figure visually represents the frequency with which each GCI indicator is considered the most or least influential factor in countries' cybersecurity policies, highlighting the varying emphasis placed on different policy dimensions globally.

Inferences from the statistical results:

(1) Average Organizational Measures more frequently serve as the most influential factor, indicating that organizational measures play a prominent role in the cybersecurity affairs of most countries and have a relatively more significant impact on cybersecurity outcomes.

(2) Average Capacity Development Measures act as secondary influential factors, suggesting that in some countries, capacity development measures have a notable effect.

(3) The largest number of countries consider Average Legal Measures as the least influential factor, indicating that legal measures relatively play a minor role among these influencing factors in many countries.

This provides an important reference for countries to formulate and improve cybersecurity policies, prompting them to attach importance to organizational measures and capacity development measures while optimizing legal measures to enhance overall cybersecurity levels.

3.2. Variable Definition and Matrix Construction

Table 1 provides a clear definition of the x_{ik} variables used in the GWR formula, ensuring clarity and reproducibility of the model.

Table 1. Definition of x_{ik} variables in the GWR formula

x_{ik} Value	Definition
x_{i1}	GDP of the *i-th* country
x_{i2}	Internet penetration rate of the *i-th* country
x_{i3}	ITU's GCI score for the *i-th* country

Weight Matrix Construction:

$$W = \begin{bmatrix} w_{11} & \cdots & w_{1n} \\ \vdots & \ddots & \vdots \\ w_{n1} & \cdots & w_{nn} \end{bmatrix} \tag{4}$$

And

$$w_{ij} = \exp\left(-\frac{d_{ij}^2}{2\times 8000^2}\right) \tag{5}$$

where n=129, and the diagonal elements of the matrix satisfy $w_{ii} = 1$.

3.3. Model Results and Spatial Feature Analysis

3.3.1. Identification of High-Incidence Target Areas

Through GWR model fitting, the global distribution of cybercrime cases exhibits a "core-periphery" spatial differentiation feature, as depicted in Fig. 3.

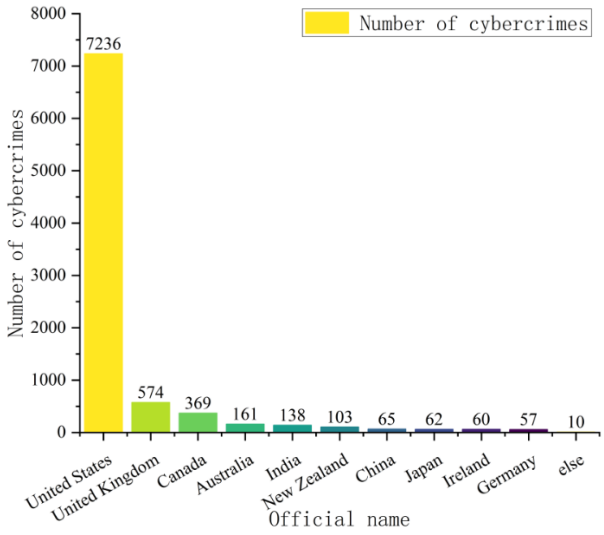

Figure 3. Histogram of global cybercrime numbers

This histogram vividly depicts the disproportionate distribution of global cybercrime incidents, with a significant concentration in a few countries, underscoring the "core-periphery" spatial differentiation feature identified by the GWR model.

High-Incidence Core Areas: North America (the United States, Canada), Western Europe (the United Kingdom, Germany), and South Asia (India) show dark colors (crime rate > 100 cases per million population), concentrating 58% of global cybercrimes, reflecting the "crime gravitational effect of digital economic centers".

Table 2 presents a comparison of key indicators for high-incidence countries in 2023, specifically highlighting the United States and India.

Table 2. Comparison of key indicators for high-incidence countries (2023)

[1] Country	[2] Crime Rate	[3] Success Rate	[4] Law Score	[5] Tech Investment
[6] USA	[7] 837	[8] 78	[9] 72/100	[10] 0.8
[11] India	[12] 138	[13] 55	[14] 48/100	[15] 0.3

Prevention Periphery Areas: East Asia (Singapore, South Korea) and Northern Europe (Sweden, Finland) (crime rate < 20 cases per million population) build prevention barriers through the dual dimensions of "technical defense + legal deterrence"[8-9], verifying the spatial inhibitory effect of policy efficacy.The strategic effects of prevention demonstration areas are further elaborated in Table 3, showcasing the success of countries like Singapore and Germany in their prevention efforts.

Table 3. Strategy effects of prevention demonstration areas

[16] Country	[17] Prevention Rate	[18] Tech Score	[19] Proportion of International Cooperation Cases
[20] Singapore	[21] 92%	[22] 91/100	[23] 45%
[24] Germany	[25] 87%	[26] 88/100	[27] 58%

3.3.2. Spatial Heterogeneity of Influencing Factors

(1) Economic Level: Per capita GDP in North America is positively correlated with the crime rate (β=0.12, $p<0.01$), while in East Asia, the high technical investment leads to a negative correlation (β=-0.08, $p<0.05$)

(2) Legal Measures: In European countries, every 10-point increase in the Law score reduces the crime rate by 18% (β=-0.18); in Africa, the effect is insignificant due to legal fragmentation (β=-0.03).

(3) International Cooperation: In ASEAN countries, each annual increase of 4 joint law enforcement operations reduces the success rate by 4% (β=-0.04, Fig. 4),as illustrated in Fig. 4.

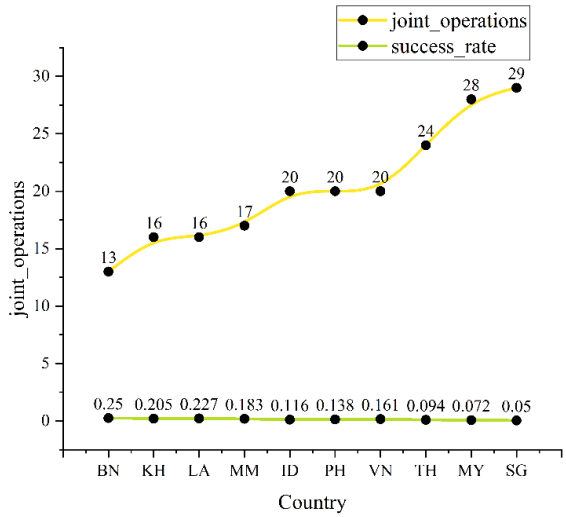

Figure 4. Correlation between joint law enforcement and crime success rate in ASEAN countries

4. Discussion

Experiences from highly effective prevention countries indicate that the synergistic effect of technological innovation (such as AI monitoring) and legal improvement (such as the EU GDPR) can significantly enhance prevention efficacy. While regional collaboration mechanisms can enhance prevention efficiency by over 30%, the combined effect of strong technical investment and well-developed legal policies, as observed in prevention leaders, contributes to an even greater crime inhibition effect. Specifically, well-developed and strictly enforced legal policies alone account for 80% of the crime inhibition effect , demonstrating the critical role of multi-faceted prevention strategies[10].

5. Conclusions

This study reveals the spatial heterogeneity patterns of global cybercrime through the GWR model, confirming the significant spatial effects of economic level and policy efficacy. Regional proximity collaboration can enhance prevention efficiency by over 30% , and a 1% increase in technical investment relative to GDP can reduce the crime rate by 12%. The research proposes that global governance focus on "legal coordination in hotspots, technical empowerment in weak areas, and intelligence sharing across all regions," providing spatially precise support for cybersecurity policies. Future studies could integrate the temporal dimension, such as a Geographically and Temporally Weighted Regression (GTWR) model, to analyze policy lag effects and dynamic coordination mechanisms.

References

[1] SARKAR G, SHUKLA S K. Behavioral analysis of cybercrime: Paving the way for effective policing strategies[J]. Journal of Economic Criminology, 2023, 2: 100034.

[2] PHILLIPS K, DAVIDSON J C, FARR R R, et al. Conceptualizing Cybercrime: Definitions, Typologies and Taxonomies[J]. Forensic Sci, 2022, 2: 379-398.

[3] JIANG F, LI C Y. Study on Geographically Attribute Neural Network Weighted Regression[C]//2022 3rd International Conference on Geology, Mapping and Remote Sensing (ICGMRS). Zhoushan, China, 2022: 213-218.

[4] MURSITO T P A P, BERLIANTO M C, AN'AMTAADININDRA Y, et al. Modeling the Amount of Poverty in Central Java using Geographically Weighted Regression[C]//2022 International Conference on Science and Technology (ICOSTECH). BatamCity, Indonesia, 2022: 1-6.

[5] AMARULLAH A H, RUNTURAMBI A J S, WIDIAWAN B. Analyzing Cyber Crimes during COVID-19 Time in Indonesia[C]//2021 3rd International Conference on Computer Communication and the Internet (ICCCI). Nagoya, Japan, 2021: 78-83.

[6] WOODS D W, WALTER L. Reviewing Estimates of Cybercrime Victimisation and Cyber Risk Likelihood[C]//2022 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW). Genoa, Italy, 2022: 150-162.

[7] HAMISU M, IDRIS A M, MANSOUR A, et al. Analysis of Cybercrime in Nigeria[C]//2020 IEEE 2nd International Conference on Cyberspac (CYBER NIGERIA). Abuja, Nigeria, 2021: 73-79.

[8] SAINI I S, KAUR N. The Power of Predictive Analytics: Forecasting Crime Trends in High-Risk Areas for Crime Prevention using Machine Learning[C]//2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT). Delhi, India, 2023: 1-10.

[9] GAJJAR V R, TAHERDOOST H. Cybercrime on a Global Scale: Trends, Policies, and Cybersecurity Strategies[C]//2024 5th International Conference on Mobile Computing and Sustainable Informatics (ICMCSI). Lalitpur, Nepal, 2024: 668-676.

[10] KUNAL M R, SHARMA D, ANURAG. Understanding Cyber-Attacks and their Impact on Global Financial Landscape[C]//2023 International Conference on Circuit Power and Computing Technologies (ICCPCT). Kollam, India, 2023: 1452-1456.