Problem Chosen	2025 MCM/ICM Summary Sheet	Team Control Number
F		2500434

Hello Summary!

INDEX

1 Introduction

1.1 Background

In the digital age, the speed and scale of global connectivity have reached unprecedented heights. However, with technological advancements, **cybercrime** has also become increasingly complex and diverse. These crimes pose significant threats and challenges to personal privacy, corporate assets, national security, and social stability. The transnational and covert nature of cybercrime makes it difficult to address effectively. Attackers often exploit legal differences and technical vulnerabilities across countries to evade accountability. Additionally, many businesses and institutions, concerned about their reputation and commercial interests, often choose **not to publicly report** cyberattacks, opting instead to pay ransoms or handle incidents privately. This further exacerbates the hidden nature of cybercrime. Developed countries, with their highly digitized economic and social structures, are often prime targets for cybercrime, while developing countries face their own unique challenges.

To address this global challenge, countries have introduced national cybersecurity **policies** aimed at enhancing their defensive capabilities through legal, technical, and organizational cooperation. The **effectiveness** of these policies varies significantly across nations, and these differences may be closely related to factors such as policy design, enforcement, technological infrastructure, education levels, economic development, and internet penetration rates. In this context, understanding which factors make certain countries' cybersecurity policies more effective has become a critical issue. By analyzing the global distribution of cybercrime, national cybersecurity policies, and their outcomes, we can identify which policies and laws are particularly effective in preventing, prosecuting, and mitigating cybercrime. This data-driven analysis not only helps countries improve their cybersecurity policies but also provides valuable insights for global cybersecurity cooperation.

1.2 Restatement of the problem

We are required to identify patterns that can inform the data-driven development and refinement of national cybersecurity policies and laws, which focused on those that have demonstrated effectiveness. Our goal is to develop a theory explaining what constitutes a strong national cybersecurity policy and then support it with a data-driven analysis.

Several key aspects are listed below:

- Patterns in cybercrime distributed worldwide.
- Assessment to effectiveness of National Cybersecurity Policies.
- Correlation between national demographics and our distribution analysis.

1.3 Analysis of problems

We have divided each problem into several different steps:

1.3.1 Cybercrime distribution across the world

- Process the JSON files published on VCDB .

- Use GCI as the primary indicator to assess countries with disproportionately high levels of cybercrime.
- Group countries based on GCI Tier measure, and visually represent the distribution of cybercrime in each group using heatmaps.

1.3.2 Effective policy or law analytical approach

- Identify a representative country from each of the T1 to T5 national clusters and collect cybersecurity-related policies enacted by these countries.
- Plot time-series line graphs depicting the trends of cybercrime over time and analyze which policies have been effective in curbing criminal activities.
- Conduct a targeted analysis of specific policy models' effectiveness in certain countries, focusing on particular indicators.
- Integrate temporal factors and national contexts to comprehensively evaluate the impact of policies.

1.3.3 National demographics correlation

- TODO

2 Symbol and Assumptions

2.1 Symbol Description

Symbol	Description
D_i	Cybercrime distribution in each country
P_i	Population of each country
y	Transformed value using $y = \log(1 + x)$.

Abbreviation	Full Form
GCI	Global Cybersecurity Index[1]
VERIS	the Vocabulary for Event Recording and Incident Sharing[2]
VCDB	the VERIS Community Database[3]

2.2 Assumption

- Countries with a population below 5% are excluded from the consideration of cybercrime distribution because a small change of number could bring a significant difference to statistical analysis.
- In the statistical analysis of the global distribution of cybercrime, factors such as national population growth, internet access, wealth levels, and education levels are assumed to have no

significant impact on the quantitative distribution of cybercrime incidents. This study hypothesizes that the distribution of cybercrime can be more intuitively understood by focusing solely on the number of incidents, independent of these socio-economic variables. This assumption is based on all available data recorded since the inception of cybercrime statistics, aiming to isolate the distribution patterns of cybercrime from other potential influencing factors.

- The impact of newly enacted laws or policies on cybercrime is not instantaneous; there is a time lag before their effects become evident.

3 A Data-Driven Model for Global Cybercrime Hotspot Mapping

Despite the continuous evolution of national cybercrime since the inception of data collection, along with changes in policies, legal frameworks, and population demographics, we can create a global cybercrime hotspot map by leveraging crime data recorded by VERIS over the years. This not only facilitates the analysis of cybercrime volumes by country but also allows for fitting the data against policy and population variables to assess their influence on cybercrime trends.

3.1 Cybercrime distribution across the globe

We made use of a world-wide map to represent all cybercrime occurred around the world. In the map, the color filled in each country represents the total number of cybercrime incidents recorded since the beginning of the statistics. The color gradient, ranging from dark blue to dark red, corresponds to eight severity levels (1 to 8). Countries marked in blue indicate a low frequency of cybercrime incidents, while those marked in red represent a high density of such incidents. For instance, the United States, where the VERIS concept was first proposed, has the highest number of recorded incidents (7,236), whereas many other countries have only 1 or 2 recorded incidents. To address this significant disparity in data distribution, we applied a logarithmic transformation to the data using the formula

$$y = \log(1 + x)$$

where x here represents D_i. This transformation was implemented using the function

$$np.log1p()$$

in Python to ensure computational precision and stability, particularly for small values. The final results are visualized in Figure 1.

3.2 High-prevalence regions

We obtained population data P_i for various countries over recent decades from the World Bank Group's website[4]. Simultaneously, we processed data from the VCDB to tabulate the annual number of cybercrime incidents D_i for each country from 2000 to 2025. However, due to discrepancies in the specific countries reported by the World Bank Group and those listed in the VCDB, we had to exclude certain countries to ensure that only those appearing in both datasets were retained. Ultimately, 109

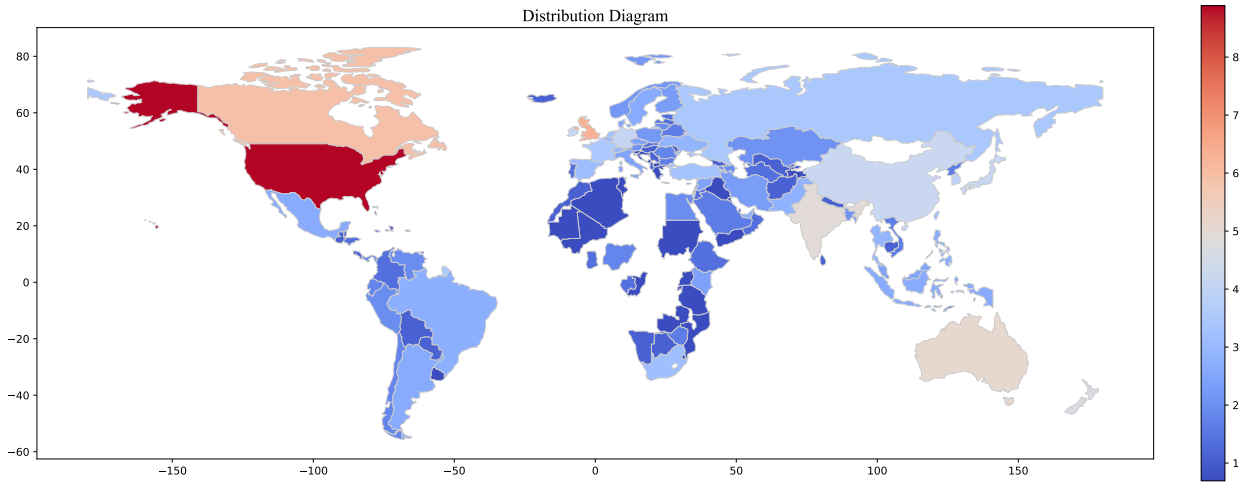


Figure 1: Crime distribution

countries were included in the model. To represent the average number of cybercrime incidents per capita, we calculated the ratio D_i/P_i for each year from 2000 to 2025. Since the resulting values were too small for practical analysis, we scaled them by a factor of 10^8 to express the data as the number of cybercrime incidents per 100 million people, denoted as hmD/P_i:

$$hmD/P_i = \frac{D_i}{P_i} \times 10^8$$

According to the GCI (Global Cybersecurity Index) standards, countries are classified into five tiers, denoted as T1 to T5. We used this classification as the basis for K-means clustering analysis, dividing the 109 countries into five groups based on the percentiles published on the GCI website: the top 10%, the next 20%, the following 25%, the subsequent 25%, and the bottom 20%. For each group, the annual average of hmD/P_i (the number of cybercrime incidents per 100 million people) was calculated. To visualize the results, we constructed a 3D clustering heatmap of cybercrime trends, where the x-axis represents the five tiers (T1 to T5), the y-axis represents the time span from 2000 to 2025, and the z-axis represents the average hmD/P_i values. This visualization is presented in Figure2.

3.3 Other Cybercrime Incidents

Using additional data obtained from the VCDB, we constructed heatmaps on a global scale based on the number of successful cybercrimes, thwarted cybercrimes, and reported cybercrimes, respectively. Due to the disproportionately high volume of data from the United States, we applied the same logarithmic transformation ($y = \log(1 + x)$) as in Figure 1 for consistency, where x represents successful attacks, thwarted attacks, and reported attacks, resulting in the three sub-figures presented in Figure3.

In sub-figure (a), the number of successful attacks closely aligns with the total number of attacks in most countries. For instance, the United States recorded 7,189 successful attacks out of 7,236 total attacks, yielding a success rate of $\frac{7189}{7236} \approx 99.35\%$. Similarly, the United Kingdom reported 569 successful attacks out of 574 total attacks, with a success rate of $\frac{569}{574} \approx 99.13\%$.

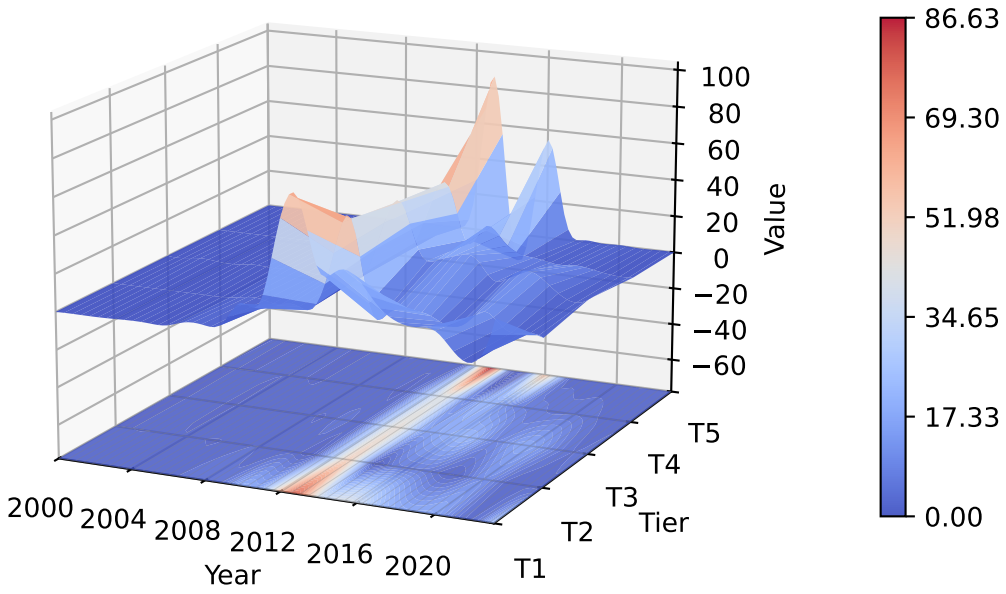


Figure 2: 3D with Spaced Projection

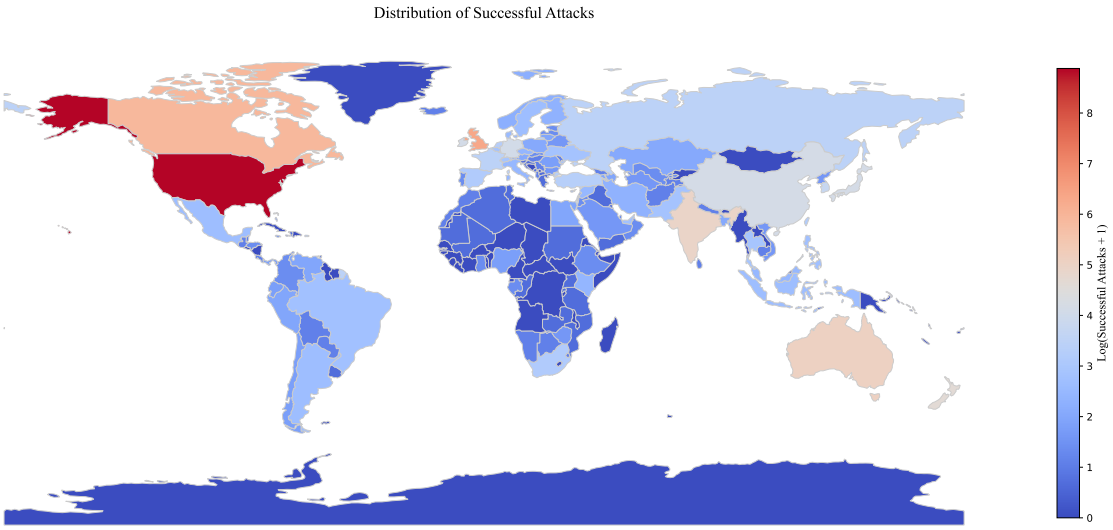
In contrast, countries with lower attack volumes did not show significant differences between the total number of attacks and the number of successful attacks, indicating that almost every attempted attack was successful.

In sub-figure (b), only the United States and Canada reported thwarted attack cases, with 6 and 2 instances, respectively.

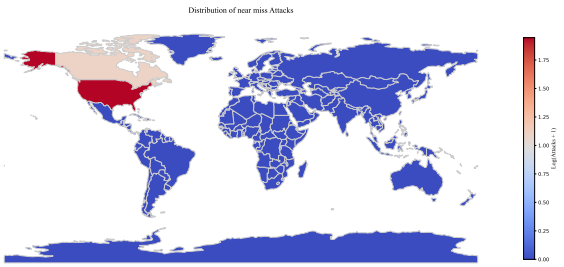
In sub-figure (c), the number of successfully reported attacks and the number of countries involved were significantly higher than in sub-figure (b). This suggests that while many attacks were successful, a portion of them were detected and reported.

3.4 Pattern Discovery

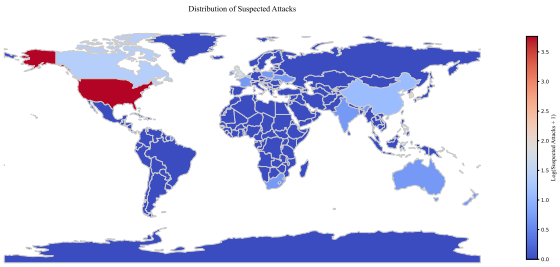
As illustrated in Figure 3, the majority of countries worldwide lack adequate defensive capabilities against cyberattacks. Among the disclosed incidents, only the United States and Canada have recorded successful defense cases, with the United States reporting 6 instances and Canada reporting 2. This disparity can be attributed to the early development and technological maturity of the United States' cybersecurity infrastructure. The advanced technologies employed by both the U.S. government and private enterprises have enabled a certain level of resilience against some cyberattacks. Additionally, the disproportionately high volume of cybercrime in the United States plays a significant role. The extensive record of cyber incidents increases the likelihood of successful defenses, as it provides more opportunities for defensive mechanisms to be tested and refined. In contrast, other countries lack both the advanced defensive technologies and the high volume of cybercrime records that the United States possesses. Consequently, no successful defense cases have been reported for these countries in the



(a) Successful Cybercrime Incidents



(b) Mitigated Cybercrime Attempts



(c) Reported Cybercrime Incidents

Figure 3: Other Cybercrime Incidents

VCDB .

Cybercrime exhibits distinct patterns both across countries categorized by the GCI (Global Cybersecurity Index) and over time. To analyze these trends, we calculated the proportion of cybercrime incidents and visualized the results in Figure 2. This figure illustrates the annual average hmD/P_i for each GCI tier (T1 to T5) from 2000 to 2025, providing a comprehensive overview of the spatial and temporal distribution of cybercrime.

- **2012 to 2016** Around 2012, both T5 and T1 countries exhibited the highest number of cybercrime incidents. Specifically, the hmD/P_i index for T5 countries reached 86.63, marking the highest average number of cybercrime incidents among all tiers over the 25-year period. In contrast, T2, T3, and T4 countries reported significantly fewer incidents compared to T1 and T5 during this timeframe. The high cybercrime rates in T5 countries can be attributed to their poor performance across all five GCI assessment metrics: Legal, Technical, Organizational, Capacity Development, and Cooperation. These deficiencies likely contributed to their low GCI rankings and inadequate cybersecurity infrastructure, which in turn made them more vulnerable to cybercrime. On the other hand, T1 countries, despite their high GCI rankings, experienced a surge in cybercrime incidents around 2012. This suggests that even these nations were not fully prepared for the rapid evolution of cyber threats during this period. It is important to note that the fifth edition of the GCI assessment data was only released in 2024, meaning that the high rankings of T1 countries in 2024 do not necessarily reflect their cybersecurity capabilities in 2012.

- **2016 to 2020** During the period from 2016 to 2020, the number of cybercrime incidents declined across all tiers, from T1 to T5. Notably, T1 countries experienced a sharp decrease in cybercrime rates. This trend can be attributed to several factors. First, T1 countries, with their advanced technological infrastructure and robust cybersecurity policies, were better equipped to adapt to emerging threats. During this period, many T1 nations implemented comprehensive cybersecurity strategies, including stricter regulations, enhanced public-private partnerships, and increased investment in cybersecurity education and training. Additionally, international cooperation played a significant role in mitigating cross-border cyber threats. For example, initiatives such as information sharing through organizations like VERIS contributed to a more coordinated global response to cybercrime. However, despite the overall decline, T5 countries remained the most affected by cybercrime, likely due to their low performance in the GCI assessment metrics. These deficiencies continued to hinder their ability to build effective cybersecurity defenses, leaving them more vulnerable to cyber threats.

- **2020 till Present** Around 2020, a remarkable shift has occurred in the distribution of cybercrime incidents across the GCI tiers. Notably, the number of reported cybercrime incidents in T5 countries has dropped to almost negligible levels, a phenomenon that warrants further investigation. Similarly, T4 countries have maintained cybercrime rates comparable to those of T5. In contrast, T1, T2, and T3 countries continue to report a portion of cybercrime incidents, although these numbers have been steadily decreasing year by year. This trend may reflect the cumulative impact of sustained investments in cybersecurity infrastructure, policy improvements, and international collaboration among higher-tier countries. However, the sharp decline in T5 countries' cybercrime records remains puzzling and could be attributed to underreporting. Further research is needed to explore these potential explanations.

4 Policy Identification and Analysis

To identify the effectiveness of national cybersecurity policies, it is essential to analyze the correlation between the implementation of these policies and the subsequent trends in cybercrime. By examining the distribution of cybercrimes and comparing it with the timing and content of various national policies, we can discern patterns that highlight which measures are particularly effective or ineffective. This analysis will focus on key metrics such as the reduction in cybercrime incidents, the success rate of prosecutions, and the overall resilience of national cybersecurity infrastructures. Through this data-driven approach, we aim to provide actionable insights for the development and refinement of cybersecurity policies.

4.1 Selection of Representative Centroid Countries

Having constructed a clustering model to categorize countries into five clusters (T1 to T5) based on GCI and other relevant metrics, we now proceed to analyze the effectiveness of cybersecurity policies within each cluster. To ensure a representative and data-driven analysis, we will select one central country from each cluster that meets the following criteria:

- **Representativeness:** The selected country should typify the overall characteristics of its cluster, reflecting the general trends and patterns observed within that group.
- **Data Availability:** The country must have sufficient historical data on cybersecurity policies and legislation enacted over the past two decades, allowing for a comprehensive analysis of policy impacts.

By focusing on these representative countries, we aim to draw meaningful insights into the effectiveness of various cybersecurity policies and laws, which can then be generalized to other countries within the same cluster.

Implementation Steps

To identify the representative country for each cluster, we first calculate the average GCI for each cluster. The average GCI, denoted as $\overline{GCI}$, is computed as follows:

$$\overline{GCI} = \frac{\sum_{i=1}^{n} GCI_i}{n}$$

where n is the number of countries in the cluster. Next, we compute the absolute deviation of each country's GCI from the cluster average:

$$\{a | a = |GCI_i - \overline{GCI}|\}$$

where a is the approach to the average $\overline{GCI}$. The country with the smallest deviation is considered the most representative of its cluster. After this initial selection, we further filter out countries with insufficient or incomplete legal and policy documentation.

Through this process, we identify the following representative countries for each cluster with their references:

- **T1: United States** [5, 6, 7, 8, 9, 10, 11, 12, 13, 14]
- **T2: Japan** [15, 16, 17, 18, 19, 20, 21, 22, 23]
- **T3: China** [24, 25, 26, 27, 28]
- **T4: Costa Rica** [29, 30]
- **T5: Namibia** [31, 32, 33]

Visualizing Policy Impact Over Time

With the selected representative countries, we proceed to visualize the impact of cybersecurity policies on cybercrime trends. For each country, we plot a line graph where the x-axis represents time (from 2000 to 2023) and the y-axis represents the annual number of cybercrime incidents. To highlight the influence of policy implementations, we mark the data points corresponding to years in which cybersecurity policies or laws were enacted with an orange color. This visualization is presented in Figure4.

This allows us to preliminarily assess the effectiveness of the policies. Specifically:

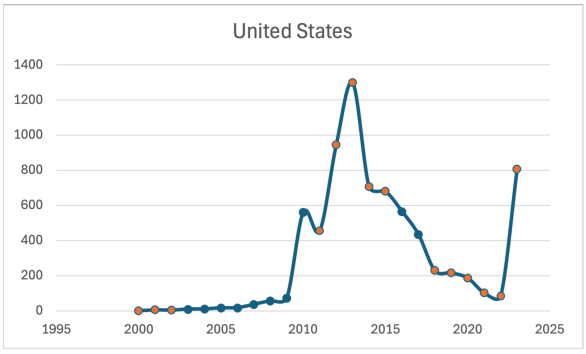
- A downward trend in the line graph following the implementation of a policy (marked in orange) suggests that the policy may have been effective in reducing cybercrime.
- An upward or unchanged trend, on the other hand, may indicate that the policy was ineffective or had unintended consequences.

This initial analysis provides a broad overview of the impact of various policies and helps identify patterns that warrant further investigation. It also serves as a foundation for more detailed analysis of specific policies, guiding future research directions.

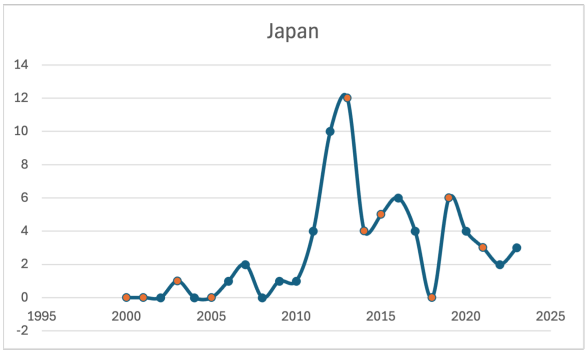
Categorizing Policies Based on Effectiveness

From the line graphs, we categorize the enacted policies into three sets based on the average cybercrime metrics in the years following their implementation compared to the year of enactment. Drawing inspiration from fuzzy set theory, we assign a value of 1 to policies that are effective, -1 to those with the opposite effect, and values between -1 and 1 to policies with varying degrees of impact. Specifically:

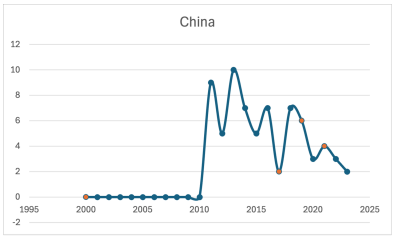
- **Effective Policies** ($Value \rightarrow 1$)**:** These are policies where the average cybercrime metrics in the years following enactment are lower than those in the year of enactment. Examples include:
  - National Institute of Standards and Technology (NIST) Cybersecurity Framework (2014)
  - National Cyber Security Centre (NCSC) Establishment (2016)
  - Investigatory Powers Act (2016)
  - Cybersecurity Strategy (2013)
  - Telecommunications Business Act Amendments (2019)
  - Cryptography Law of the People's Republic of China (2019)



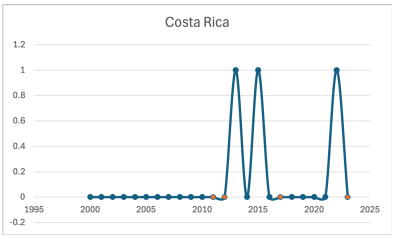
(a) US (T1)



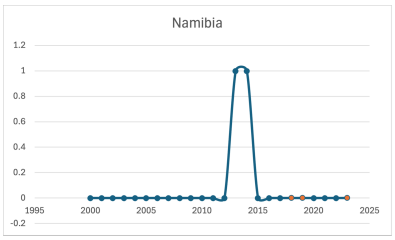
(b) Japan (T2)



(c) China (T3)



(d) Costa Rica (T4)



(e) Namibia (T5)

Figure 4: Representative countries

- **Counterproductive Policies** ($Value \rightarrow 1$)**:** These are policies where the average cybercrime metrics in the years following enactment are higher than those in the year of enactment. Examples include:
  - Cyber Incident Reporting for Critical Infrastructure Act (CIRCIA) (2022)
  - Quantum Computing Cybersecurity Preparedness Act (2022)
  - Strengthening American Cybersecurity Act (2022)
  - CHIPS and Science Act (2022)
  - Cybersecurity Strategy (2018)
  - Cybersecurity Law of the People's Republic of China (2017)
- **Neutral or Mixed-Impact Policies (Values around 0):** These are policies where the average cybercrime metrics show no significant change, or the impact is ambiguous. Representative examples include:
  - Policy A
  - Policy B
  - Policy C
  - (Additional policies to be listed)

This categorization provides a structured framework for analyzing the effectiveness of cybersecurity policies and serves as a basis for further investigation into the factors that contribute to their success or failure.

5 The establishment and solution of problem 3 model

5.1

6 Future expected data

7 Advantages & Disadvantages

References

[1] Global cybersecurity index 2024. https://www.itu.int/epublications/publication/global-cybersecurity-index-2024, 2024.

[2] Veris - the vocabulary for event recording and incident sharing. https://verisframework.org/index.html.

[3] The veris community database (vcdb). https://verisframework.org/vcdb.html.

[4] World population. https://data.worldbank.org/indicator/SP.POP.TOTL?most_recent_year_desc=true.

[5] https://www.congress.gov.

[6] https://www.nist.gov.

[7] https://www.dhs.gov.

[8] https://www.sec.gov.

[9] https://www.whitehouse.gov.

[10] https://en.wikipedia.org/wiki/Investigatory_Powers_Act_2016.

[11] https://en.wikipedia.org/wiki/National_Cyber_Security_Centre_%28United_Kingdom%29.

[12] https://en.wikipedia.org/wiki/Telecommunications_%28Security%29_Act_2021.

[13] https://privacymatters.dlapiper.com/2024/05/uk-new-cyber-security-requirements-for-consumer-products/.

[14] https://www.skadden.com/insights/publications/2024/10/timeline-set-for-uk-cybersecurity.

[15] https://www.kantei.go.jp/jp/it/it_basic_law/index.html.

[16] https://www.ppc.go.jp/en/legal/.

[17] https://www.nisc.go.jp/eng/.

[18] https://www.mofa.go.jp/policy/page18e_000015.html.

[19] https://www.japaneselawtranslation.go.jp/en/laws/view/3677/en.

[20] https://www.nisc.go.jp/eng/pdf/cs-senryaku2015-en.pdf.

[21] https://www.nisc.go.jp/eng/pdf/cs-senryaku2018-en.pdf.

[22] https://www.soumu.go.jp/main_sosiki/joho_tsusin/eng/Telecommunications_Business_Act.html.

[23] https://www.nisc.go.jp/eng/pdf/cs-senryaku2021-en.pdf.

[24] https://en.wikipedia.org/wiki/International_cybercrime.

[25] https://en.wikipedia.org/wiki/Cybersecurity_Law_of_the_People%27s_Republic_of_China.

[26] https://en.wikipedia.org/wiki/Internet_censorship_in_China.

[27] https://iapp.org/news/a/china-issues-the-regulations-on-network-data-security-management-what-s-important-to-know.

[28] https://en.wikipedia.org/wiki/Cryptography_law.

[29] https://www.micitt.go.cr/sites/default/files/2023-06/englishEstrategia-Nacional-de-Ciberseguridad-Costa-Rica.pdf.

[30] https://www.coe.int/en/web/octopus/country-wiki-ap/-/asset_publisher/CmDb7M4RGb4Z/content/costa-rica/pop_up.

[31] https://www.coe.int/en/web/octopus/country-wiki-ap/-/asset_publisher/CmDb7M4RGb4Z/content/namibia/p

[32] https://carnegieendowment.org/research/2024/04/a-digital-odyssey-the-convergence-of-rapid-digitization-population-dynamics-and-financial-risk-in-namibia.

[33] https://www.nmt.africa/uploads/614346b1d2ebb/NMTsubmision-Reviewofnationalcybersecuritystrat

Appendix