A Study on Cybercrime Policy Analysis Based On K-Means++ Clustering and TOPSIS Modeling

Xinming Cheng [1, *, #], Shunkai Wang [2, #], Zhipeng Lv [3, #]

[1] School of Environment, Liaoning University, Shenyang, China, 110036

[2] School of Mathematics and Statistics, Liaoning University, Shenyang, China, 110036

[3] Faculty of Information, Liaoning University, Shenyang, China, 110036

* Corresponding Author Email: chengxm2025@163.com

[#]These authors are contributed equally.

Abstract. As modern technology advances, more and more of our world is interconnected through the Internet, but it has also increased our individual and collective risk of cybercrime. This paper uses modelling to find policies that are effective in curbing cybercrime. Firstly, this article utilized the K-Means++ clustering algorithm to classify countries into five categories based on the cybercrime rates, then this article uses Logistic regression to study the impact of prosecution and reporting on the cybercrime situation. Secondly, in order to explore the model that can identify the effectiveness of relevant policies, this article constructed a Combined Empowerment - TOPSIS model for assessing the effectiveness of the policies and finally get the most effective policies to solve cybercrime in different categories of countries. After that a Pearson correlation analysis of the different policies is performed to explore synergies between policies. Finally, in order to make theory complete and scientific, this article looked for four demographic features: internet coverage (%), wealth (GDP per capita), education level, and population happiness index, predicted the above four indicators as well as the cybercrime rate through the LSTM algorithm, and combined them with the historical data to perform a comprehensive Spearman correlation analysis, and then united the results of the analysis into theory.

Keywords: K-Means++ Clustering, Logistic Regression, TOPSIS, LSTM, Correlation Analysis.

1. Introduction

Since the 1990s, when the Internet entered the phase of globalization and popularization, it has brought great convenience to human society. However, the Internet has also given rise to a new mode of crime -- cybercrime. In recent years, cybercrime has become frequent, with expanding economic losses and security risks. The low-cost and hidden nature of cybercrime has brought great difficulty to the detection of cases, and its unique cross-regional nature has undoubtedly further increased the difficulty of combating it. Exploring ways to effectively defend against cybercrime is of great significance to building a global cybersecurity pattern and maintaining social and economic stability.

The current defense measures against cybercrime are mainly divided into two categories, one is to use computer science to combat cybercrime by using firewalls or antivirus software; the other is to change the social factors, such as increasing penalties and publicity, and encouraging the victims to report the cases. Karim et al. extracted the phishing-related URLs from the dataset library, and constructed a phishing detection system by using machine learning, which improves the cybercrime identification efficiency [1]; Chang, Ching-Chun et al. automated reversible steganography coding and improved message encryption through nonlinear discrete optimization [2]. They mainly use innovations in computer technology to defend against cybercrime. Others study the macro-level cybersecurity situation and propose strategies at the societal level: Carvalho et al. studied laws and regulations and their strategies to combat cybercrime at the national level [3]. Ma Guang et al. explored international regulations on transnational cybercrime from multiple perspectives and drew recommendations for governance [4]. Comparatively, the latter has more flexibility and influence, but may be more dependent on the efficiency of policies and international cooperation.

Scholars in the same field have conducted research on cyber defense policies, which can be broadly categorized into two types, one exploring the drivers of cybercrime (including economic, political, geographic, educational, etc.) and proposing efficient ways to deal with them, and the other exploring the correlation between cybercrime and national policies, such as legislation, to analyze the effectiveness of the policies. Althibyani conducting an interview survey of more than 600 students in Saudi Arabia found that IT-related education usually has a significant impact on cybercrime awareness and prevention [5]. In addition, Bruce found that cybercrime rates are disproportionately high in some countries, suggesting that cybercrime may be related to geographic factors [6]. Chen further demonstrates that different economic and technological conditions between countries may be important drivers of their spatial heterogeneity [7]. In addition, the improved TOPSIS method is innovatively applied in this paper for policy effectiveness assessment, and many studies have proved its efficiency and accuracy; Zhang et al. used the hierarchical analysis method to determine the subjective weights and then combined with the objective weighted entropy method to form the combined weighting method, and constructed the combined weighting-TOPSIS model of the water-surge risk assessment system for karst tunnels [8]. The improved integrated TOPSIS method showed better evaluation results in the application compared to the traditional evaluation methods.

Most previous research on Cybercrime Prevention Policies has focused on the national level, ignoring the differences between countries, and the conclusions have a small scope of application. This study synthesizes two aspects of cybercrime driver exploration and policy effectiveness assessment, analyzes the effectiveness of different response methods against cybercrime globally based on the basic situation of multiple countries, and then makes targeted recommendations for countries with different social situations, and the conclusions are of more general and realistic significance. The process framework diagram of this paper is shown in Figure 1.

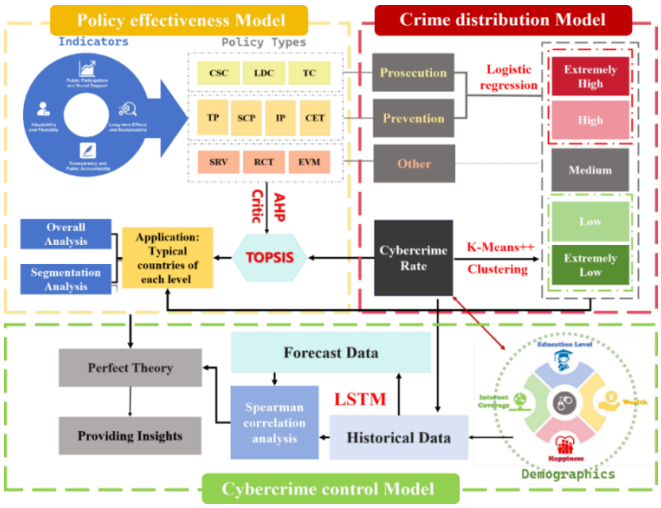

Figure 1 Framework of the article

2. Methods

2.1. K-Means++ Clustering Methods

K-Means algorithm is a kind of unsupervised learning and also a clustering algorithm based on division, generally using the Euclidean distance as a measure of similarity between data objects, similarity is inversely proportional to the distance between data objects, the greater the similarity, the smaller the distance. Since the classification results of the K-Means algorithm can vary depending on the selection of the initial points, we utilize a modification of this algorithm: K-Means++.

Step 1: A randomly selected sample from the dataset is used as the first clustering center c_i .

Step 2: Calculate the shortest distance between each sample and the current existing class clustering center (i.e., the distance to the most recent clustering center), denoted by D(x);the larger this value is,

the greater the probability of being selected as a clustering center; Next, the probability of each sample being selected as the next clustering center is calculated:

$$p = \frac{D(x)^2}{\sum_{x \in X} D(x)^2} \tag{1}$$

Finally, use the roulette wheel method to select the next clustering center c_i . Repeat this step until k clustering centers are selected.

Step 3: Assign clusters. Clusters are assigned to each point in the data set by calculating the distance of each data point from the center of mass and assigning it to the closest center of mass. In the k-means algorithm, the Euclidean distance is generally used to calculate the distance from the sample point to the center:

$$d(x, c_i) = \sqrt{\sum_{j=1}^{d} (x_j - c_{ij})^2} \tag{2}$$

Where x is a data point, c_i is the ith clustering center, d is the dimension of the data object, x_j and c_{ij} are the values on the jth dimension of x and c_i , respectively.

And the objective function of K-Means is:

$$k = \arg\min \| x_i - c_k \|^2 \tag{3}$$

Where w_{ik} is the indicator function. If the data point is in the cluster, w_{ik} =1; otherwise, w_{ik} =0.

And the objective needs to minimize J to reach the optimal value, so we need to differentiate the two variables w_{ik} and c_k separately. We easily get: if $k = \arg\min \| x_i - c_k \|^2$,then the result is 1; otherwise, the result is 0. That is, in this step, the data points x_i are assigned to the closest center of mass c_i according to the Euclidean distance.

Step 4: Update center of mass. Recalculate the center of mass for each cluster, i.e., take the average of all data points in the cluster as the new center of mass. After the data points are assigned to a particular cluster, we need to recalculate to find the best center of mass. In one step, we differentiate the c_k in the objective function from step 2. The arithmetic gives us:

$$c_k = \frac{\sum_{i}^{m} w_{ik} x_i}{\sum_{i}^{m} w_{ik}} \tag{4}$$

This formula shows that the new cluster center of mass, c_k is the weighted average of all data points xi within that cluster, with a weight of w_{ik} . in other words, it is the centroid of all the data points in cluster k. It is an averaged representation, which ensures that the center of mass moves towards the center of the data points within the cluster during the update process.

Step 5: Repeat the above steps until the clustering center no longer changes or the maximum number of iterations is reached and then stop.

2.2. Logistic Regression Model

Logistic Regression is a widely used statistical learning method for classification problems, which is actually an algorithm for binary or multiple classification problems. Logistic regression predicts

the probability of an event occurring by mapping the output of a linear regression to between 0 and 1 using a logistic function (also known as a Sigmoid function).

Step 1: Build a regression model. The goal of logistic regression is to predict a binary outcome y ∈, which is modeled by the following equation:

$$p(y=1|X)=\sigma(w^{T}X+b) \tag{5}$$

Where X is the input feature (which can be a vector of multiple features); w is the weight vector; and b is the bias term; $\sigma(z)=\frac{1}{1+e^{-z}}$ is Sigmoid function.

Step 2: Build the loss function. The loss function for logistic regression is the log loss function, which takes the following form:

$$J(w,b)=-\frac{1}{m}\sum_{i=1}^{m}\left[y^{(i)}\log(h_{\theta}(x^{(i)}))+(1-y^{(i)})\log(1-h_{\theta}(x^{(i)}))\right] \tag{6}$$

Where m is the number of training samples; $h_{\theta}(x)=\sigma(w^{T}X+b)$ is predictive probability of logistic regression.

Step 3: Solve it using gradient descent method. Like linear regression, logistic regression typically uses gradient descent to optimize the loss function and solve for the parameters w and b. The gradient update rule for logistic regression is as follows:

To the gradient of w:

$$\frac{\partial J(w,b)}{\partial w}=\frac{1}{m}\sum_{i=1}^{m}(h_{\theta}(x^{(i)})-y^{(i)})x^{(i)} \tag{7}$$

To the gradient of b:

$$\frac{\partial J(w,b)}{\partial b}=\frac{1}{m}\sum_{i=1}^{m}(h_{\theta}(x^{(i)})-y^{(i)}) \tag{8}$$

Keep updating w and b until the loss function converges.

2.3. TOPSIS Model

Assume that there are n evaluation objects and m indicators.

Step 1: The values of each attribute for multiple scenarios are formed into a decision matrix, where the rows represent each scenario and the columns represent each attribute.

Step 2: The decision matrix was normalized by forward normalization, scaling all attribute values into the [0,1] interval.

For positive indicator:

$$z_{ij}=\frac{x_{\max}-x_{ij}}{x_{\max}-x_{\min}} \tag{9}$$

For negative indicators:

$$z_{ij}=\frac{x_{\max}-x_{ij}}{x_{\max}-x_{\min}} \tag{10}$$

Step 3: Determine the weight vector: CRITIC (70%) and AHP (30%). The CRITIC weighting method is an objective assignment method for determining the weights of multiple evaluation indicators. This method takes into account not only the volatility of the data, but also the correlation between indicators. Its core idea is to comprehensively measure the objective weights of indicators by evaluating the comparative strength and conflict of the indicators.

Step 4: Calculate the optimal and worst solutions:

optimal solution : $\left(Z_1^+, Z_2^+, \cdots, Z_m^+\right)$

worst solution : $\left(Z_1^-, Z_2^-, \cdots, Z_m^-\right)$

Step 5: Calculate the gap between each evaluation metric and the optimal and worst vectors:

$$D_i^+ = \sqrt{\sum_{i=1}^{m} W_j (Z_j^+ - z_{ij})^2}, D_i^- = \sqrt{\sum_{i=1}^{m} W_j (Z_j^- - z_{ij})^2} \tag{11}$$

Where W_j is the weight of each indicator.

Step 6: Measure the proximity of the evaluation object to the optimal solution:

$$C_i = \frac{D_i^-}{D_i^+ + D_i^-} \tag{12}$$

The larger the value of C_i, the better the evaluator is.

Since the ratio of the two weighting methods in the combination assignment is artificially determined, the impact of this on the results needs to be considered, by varying this ratio we find that the final weighted results fluctuate within 5 per cent from the results at 7:3, as shown in Figure 2, which does not significantly affect the results. Therefore, the ratio of CRITIC (70%) and AHP (30%) is appropriate.

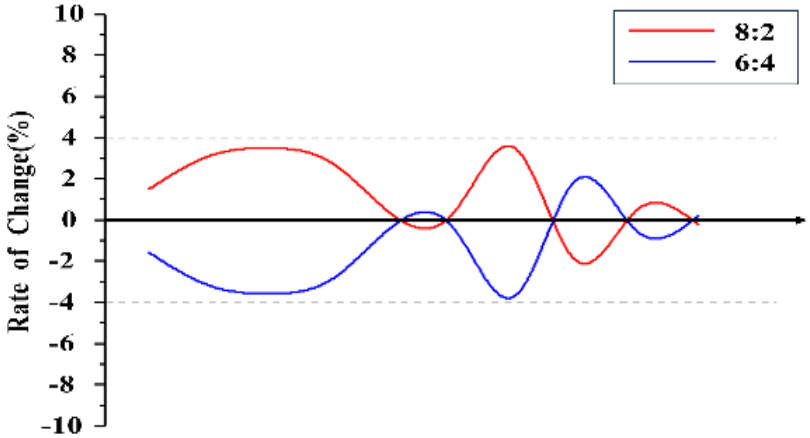

Figure 2 The change rate of policies

2.4. LSTM prediction algorithm

LSTM (Long Short-Term Memory Network) is an improved recurrent neural network (RNN), which solves the long-term dependency problem of RNN by introducing a gating mechanism, was proposed by Hochreiter and Schmidhuber in 1997, and is widely used in the fields of natural language processing, image recognition and so on. Its core realizes selective retention and updating of information through forgetting gates, input gates, cell states and output gates.

Forget Gate: Decides whether to discard or retain historical information, outputting a probability value between 0 and 1 (1 means completely retained, 0 means completely forgotten). the forgetting gate is computed as follows:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{13}$$

Where W_f is the weight matrix of the forget gate, b_f is the bias term, x_t is the input of the current time step, and h_{t-1} is the hidden state of the previous time step. σ is the sigmoid function that maps the input values to probability values between 0 and 1.

Input Gate: Determines the level of retention of new input information, outputting a value between 0 and 1 (1 being fully retained, 0 being completely ignored), the input gate is computed as follows:

$$i_t = \sigma(W_i[x_t, h_{t-1}] + b_i) \tag{14}$$

Where W_i is the weight matrix of the input gate, b_i is the bias term, x_t is the input of the current time step, and h_{t-1} is the hidden state of the previous time step.

The formula for the candidate cell state is given below:

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \tag{15}$$

Where W_C is the weight matrix of the candidate cell states, b_C is the bias term, x_t the input of the current time step, and h_{t-1} is the hidden state of the previous time step.

Cell State: The cell state can be viewed as the core of the entire LSTM network, which stores and transmits information, as well as controls the flow and update of information. The cell state of the LSTM is updated and passed to the next time step. At each time step t , the cell state is updated with the following equation:

$$c_t = f_t \cdot c_{t-1} + i_t \cdot \tilde{c}_t \tag{16}$$

Where f_t indicates the weight of forgetting the cell state; i_t indicates the weight of updating the cell state; and $\tilde{c}_t$ indicates how much the new input at the current time step can affect the cell state.

Output Gate: The output gate determines which information needs to be output via a sigmoid function. At each time step t, the output gate is calculated as follows:

$$o_t = \sigma(W_o[x_t, h_{t-1}] + b_o) \tag{17}$$

Where: W_o is the weight matrix of the output gate, b_o is the bias term, x_t is the input of the current time step, and h_{t-1} is the hidden state of the previous time step. σ is the sigmoid function that maps the input values to probability values between 0 and 1.

Next, the LSTM processes the cell state c_t through a tanh function to obtain the hidden state h_t for the current time step:

$$h_t = o_t \cdot \tanh(c_t) \tag{18}$$

Where: tanh is a hyperbolic tangent function that maps input values to values between -1 and 1.

Through output gates, LSTM is able to adaptively control the output of information based on the cellular and hidden states at the current time step. In this way, important information can be automatically filtered out and passed to the next layer or output layer in the LSTM network. The role of the output gate is to control which information should be output in the hidden state h_t of the current time step, thus improving the accuracy and effectiveness of the LSTM network.

3. Results and discussion

3.1. 3.1 Analysis of the global distribution of cybercrime and clustering results

For the distribution of cybercrime in the world, we collected information on 100 countries with valid cybercrime data and plotted the results on a world map (data from https://www.itu.int/). For the

distribution of cybercrime in the world, we collected information on 100 countries with valid cybercrime data and plotted the results on a world map, as shown in Figure 3. It is worth noting that the results of our study perhaps differ from the results of some previous studies on the distribution of cybercrime. This may be due to the fact that the metric they use is the number of cases of cybercrime in each country[9].

However, in fact this approach does not give a good indication of the cybersecurity situation in a country because the population bases of different countries are supposed to be different. This paper innovatively uses the cybercrime rate as the basis for measuring the cybercrime situation, i.e., the ratio of the number of cybercrime cases to the total number of all crimes in the country, which eliminates the influencing factor of the difference in population bases and makes the results more scientific and reliable.

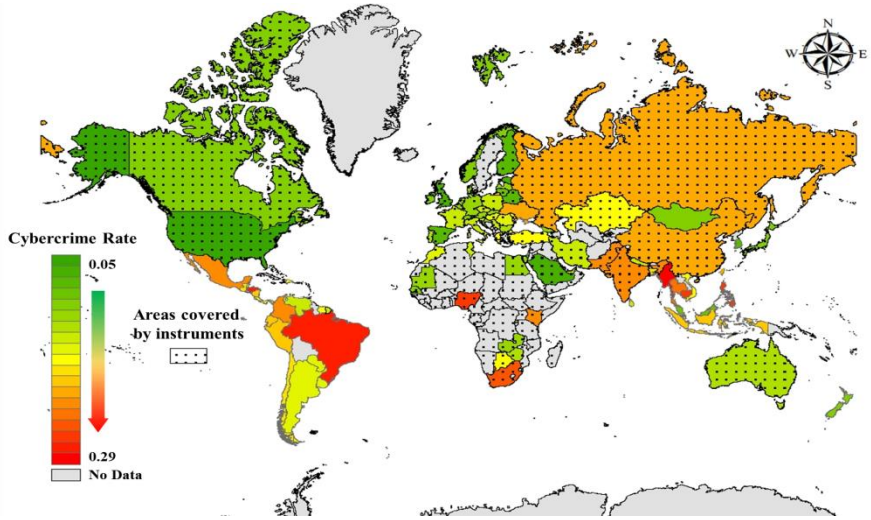

Figure 3 Geographical distribution of cybercrime rates

Through the software we clustered these 100 countries based on cybercrime rates and the results are shown in Figure 4:

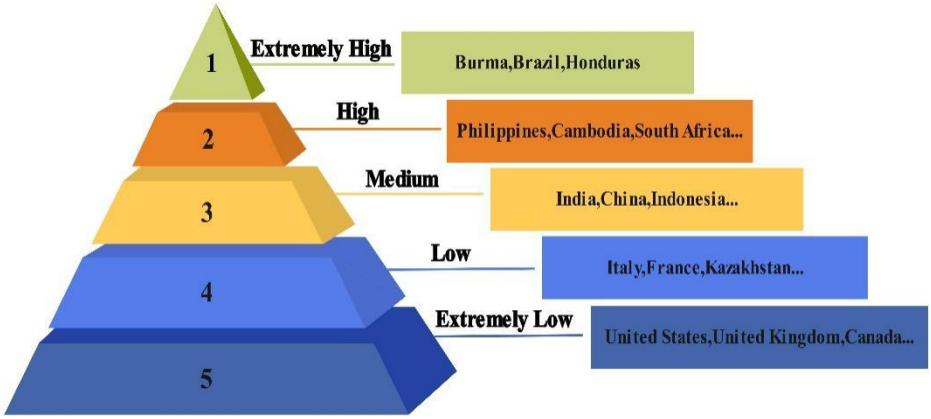

Figure 4 Results of cluster analysis

It can be seen that the countries with high rates of cybercrime are mainly concentrated in some countries in South-East Asia and Africa and South America, which is consistent with the information of GCI in ITU. This could be because the relatively low level of technology in those countries, coupled with the lack of a sound legal system, has led to a situation in which cybercrime cannot be effectively curbed.

3.2. Results of logistic regression analysis

We select four countries in level 1,2 and consider that cybercrime is not prevented in these countries, similarly we select four countries in level 4,5 and consider that cybercrime is prevented in

these countries. For countries selected from level 3, we believe that cybercrime may be successful or prevented. The results of the logistic regression with whether cybercrime is deterred or not as the dependent variable and the rate of prosecution and reporting of cybercrime cases in each country as the independent variable are shown in Figure 5:

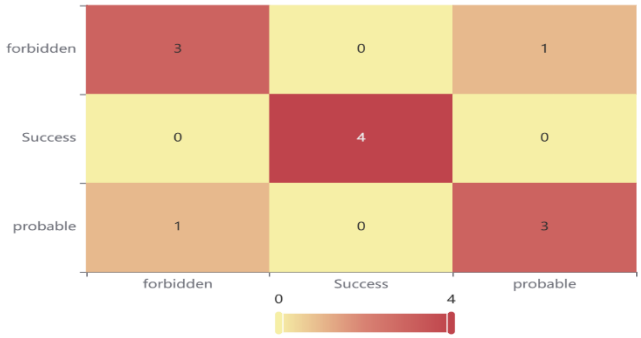

Figure 5 Heat map of confusion matrix

The evaluation indexes extracted from the confusion matrix show that the overall performance of the classification model is excellent: Accuracy reaches 83.3%, indicating that a high proportion of the samples predicted to be positive are actually positive; the AUC value is as high as 98.6%, reflecting that the model has a strong capability of distinguishing between positive and negative samples; and the recall rate is 83.3%, reflecting that a high proportion of the actual positive samples have been correctly identified.

In summary, the model has good validity and accuracy. In other words, whether or not cybercrime is reported and prosecuted has a prominent impact on the cybercrime situation in that country.

By analyzing the data, it is easy to see that countries where cybercrime is frequently reported and prosecuted generally have lower cybercrime rates than countries where cybercrime is not affected by these factors. This is consistent with the fact that citizens in countries where cybercrime is frequently reported and prosecuted generally have higher levels of cybersecurity awareness and better judicial systems, allowing cybercrime to be effectively curtailed[10].

3.3. Ranking of policy effectiveness and identification of synergistic relationships

Using the bases of policy adaptability and flexibility, long-term impact and sustainability, transparency and public accountability, public participation and societal support, we found 10 specific policies affecting cybercrime to be: cybersecurity education and training, specialized cybercrime courts, transnational judicial cooperation, and others. They are shown in Figure 6.

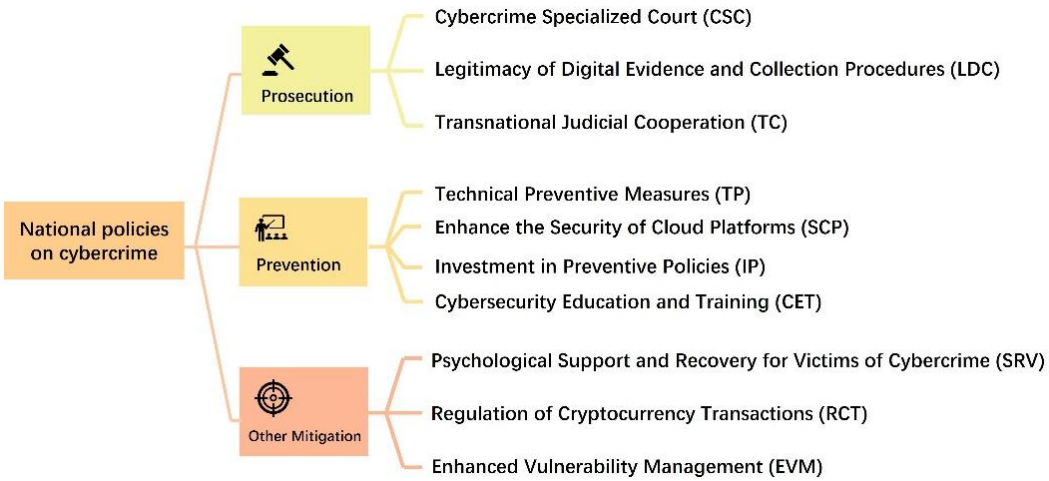

Figure 6 National policies on cybercrime

After that we choose a representative country in each category to analyze (The results of the subsequent analysis of these countries can be generalized to such countries). They are Brazil, the Philippines, Russia, Italy and the UK. For the UK. We find that the policy "Enhance Vulnerability Management" has a composite score of 0.9197, which is much larger than the other policies, while the policy "Establish Specialized Cybercrime Courts" has a composite score of 0.2668, which is much lower than the other policies. This may be due to the fact that the UK has a high level of economic development, a relatively well-developed cybercrime system and a low cybercrime rate, making the establishment of a specialized cybercrime court redundant and not worth the cost[11]. For the UK, the most important thing now is to improve the loopholes of cybercrime.

For the other four countries, we present the results through the Figure 7:

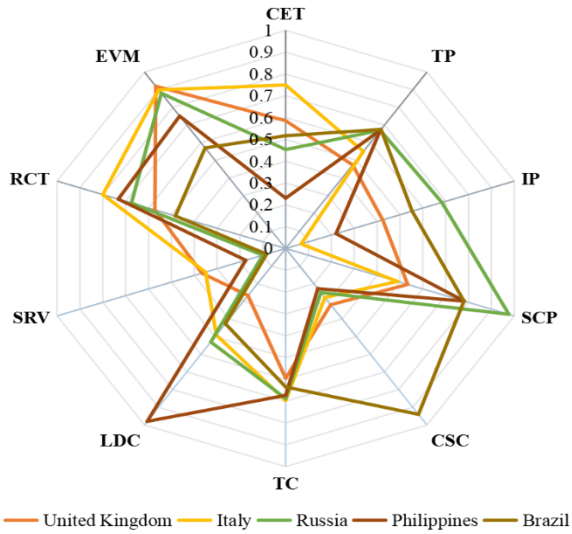

Figure 7 Radar charts of policy effectiveness

From Figure 7 we can generalize the following conclusions: for countries with low rates of cybercrime, the most effective policy for managing cybercrime is to strengthen loophole management and improve existing policies; for countries with medium rates of cybercrime, the security of online platforms should be improved on top of the above; for countries with high rates of cybercrime, the most effective policies are the establishment of specialized cybercrime courts and the creation of bills on the reasonableness of digital evidence, as well as other rule of law and prosecution measures on cybercrime.

The following is still using the UK as an example to analyze the relevance of each policy, and using Pearson's correlation analysis the following heat map can be obtained:

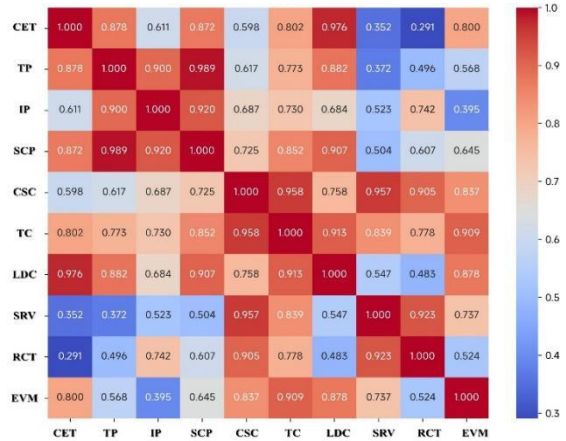

Figure 8 Heat map of correlation analysis

The Figure 8 shows a strong correlation between cybersecurity education and training technical preventive measures, investment in preventive policies, and enhancing the security of cloud

computing platforms, probably because these policies are mainly used for the prevention of cybercrime, which exhibits a policy synergy, and therefore their adoption times are correlated[12]. Similarly, the strong correlation between the policies on cybercrime-specific courts, transnational judicial cooperation, and digital evidence legitimization and collection procedures is mainly due to the fact that these policies mainly serve cybercrime-related judicial procedures, and thus their adoption times are correlated. The remaining policies do not show a clear correlation with other policies, as they are mainly complementary to existing policies.

3.4. LSTM prediction and correlation analysis results

We looked for four demographic characteristics: Internet coverage (%), wealth (GDP per capita), education level, happiness index of the population. In order to make the results of the correlation analysis more reliable and generalized, and to make some corrections to the historical data, we used the LSTM prediction model to predict the above four demographic characteristics as well as the data on cybercrime rates. For the sake of generalization of the results, we remain with the five countries selected in 3.3, and due to the similarity of the process, only Italy is used as an example for the detailed analysis below. Then we combine the historical data with the predicted data for Spearman correlation analysis. The following bubble matrix plot is obtained:

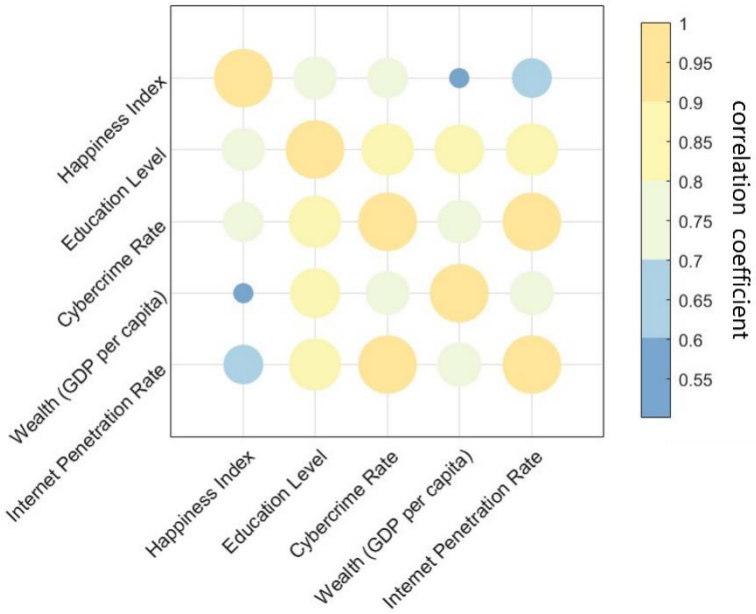

Figure 9 Bubble matrix plot of Spearman correlation analysis

From the Figure 9, we can see that the correlation coefficients of Internet penetration, education level and cybercrime are 0.997,0.885 respectively, which are all greater than 0.8, indicating that these two demographic characteristics are strongly correlated with the rate of cybercrime, i.e., they will have a greater impact on cybercrime. The other four countries have the same results as Italy, which reinforces our results. This is consistent with our previous conclusions, which are mainly reflected in the fact that higher Internet penetration will lead to more and more cybercrime information being reported, thus affecting the occurrence of cybercrime; the higher the level of education, the greater the awareness of the rule of law, which will make those victims of cybercrime defend their legitimate rights and interests by means of lawsuits and other means, thus curbing the occurrence of cybercrime.

4. Conclusions

The current cybersecurity situation is very serious, and cybercrime has become increasingly frequent and diversified, threatening every corner of the Internet era, both in developing and developed countries. Therefore, extracting the key drivers of cybercrime and constructing a policy

effectiveness evaluation model can propose efficient policies to curb cybercrime in a short period of time in line with national conditions, which can not only solve the pain point problems quickly and precisely, but also guide the deployment of resources to maximize the benefits. By analyzing the global incidence of cybercrime and its geographic distribution, we find that cybercrime can be effectively curbed when countries increase the publicity of cybercrime cases and call on victims of cybercrime to actively defend their rights through legal means.

The results of the run of the policy effectiveness evaluation model we developed indicate that for countries with low cybercrime rates, the most effective policy is to improve the loopholes in the existing policy system; for countries with medium cybercrime rates, the security of online platforms should be strengthened based on the improvement of the loopholes; and for countries with high cybercrime rates, the current task is to establish a sound judicial system to combat cybercrime, and, if necessary establish specialized cybercrime courts. We also find that further expanding Internet penetration and improving the level of education in countries can also help to curb the occurrence of cybercrime. In addition, synergies between different policies should be considered when implementing policies in order to maximize their effects. We did not consider the special national conditions of some countries in the course of our study, such as religious customs, social nature, war turmoil and other special factors that may create resistance to the implementation of some policies. However, this does not affect the fact that the research results are accurate and reliable in most cases and can effectively help countries with different status quo to formulate efficient policies.

References

[1] Karim A, Shahroz M, Mustofa K, et al. Phishing Detection System Through Hybrid Machine Learning Based on URL [J]. IEEE Access, 2023, 11: 36805-36822.

[2] Chang C C. Automation of reversible steganographic coding with nonlinear discrete optimisation [J]. Connection Science, 2022, 34(1): 1719-1735.

[3] Carvalho J V, Carvalho S, Rocha A. European strategy and legislation for cybersecurity: implications for Portugal [J]. Cluster Computing-the Journal of Networks Software Tools and Applications, 2020, 23(3): 1845-1854.

[4] Ma Guang Z S. Study on the international law regulation of transnational cybercrime -- from the perspective of international conventions, jurisdiction, and soft law norms [J]. Korea Law Review, 2020, 98: 163-198.

[5] Althibyani H A, Al-Zahrani A M. Investigating the Effect of Students' Knowledge, Beliefs, and Digital Citizenship Skills on the Prevention of Cybercrime [J]. Sustainability, 2023, 15(15).

[6] Bruce M, Lusthaus J, Kashyap R, et al. Mapping the global geography of cybercrime with the World Cybercrime Index [J]. Plos One, 2024, 19(4).

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

[8] Zhang X, Wang S Q, Chen H, et al. Risk Assessment of Karst Tunnel Water Inrush Based on Combined Weighting Method [J]. Tehnicki Vjesnik-Technical Gazette, 2025, 32(1): 157-164.

[9] Khan S, Saleh T, Dorasamy M, et al. A systematic literature review on cybercrime legislation [J]. F1000Research, 2022, 11: 971.

[10] Popham J, Mccluskey M, Ouellet M, et al. Exploring police-reported cybercrime in Canada variation and correlates [J]. Policing-an International Journal of Police Strategies & Management, 2020, 43(1): 35-48.

[11] Joo M, Hun-Yeong K, In L J. Cyber Security Governance Analysis in Major Countries and Policy Implications [J]. Journal of The Korea Institute of Information Security and Cryptology, 2018, 28(5): 1259-1277.

[12] Hong Y, Neilson W. Cybercrime and Punishment [J]. Journal of Legal Studies, 2020, 49(2): 431-466.