Cybersecurity Assessment Based on Entropy Weight Method and Polynomial Regression Modeling

Yibo Zhang *

School of Mechanical Engineering, Northwestern Polytechnical University, Xi'an, China

* Corresponding Author Email: 916881864@qq.com

Abstract. The aim of this paper is to assess cyber security stability and identify effective strategies. The study collects multi-dimensional data such as crime rate and crime success rate and uses entropy weight method to construct a stability assessment model. On the basis of the traditional entropy weighting method, the weights are set manually to combine domain knowledge and subjective judgment, the normalization anomalies are avoided by checking and fine-tuning the data, random noise is introduced to ensure the stability of the entropy weighting method, and the multi-dimensional comprehensive assessment is carried out and the explanatory nature of the weighted scores is enhanced, which improves the shortcomings of the traditional method. In terms of policy assessment, data related to cybercrime policies are collected and quantified, the crime reduction rate is calculated by differencing, the entropy-weighting method is applied to analyze the effects of prevention, prosecution, and mitigation intensities on the crime reduction rate, and a polynomial regression model is developed to study the effects of policy implementation time, while random fluctuations are added to simulate the rebound for more realistic prediction results. The model was evaluated for goodness-of-fit and cross-validated for stability and good predictive ability. This study provides an effective method for cybersecurity stability assessment and policy formulation.

Keywords: Cybersecurity; entropy weight method; polynomial regression modeling.

1. Introduction

This paper is based on entropy weight method [1] and polynomial regression model [2] to accurately assess cyber security stability and formulate effective strategies. On the one hand, the entropy power method is used to construct a cybersecurity stability assessment model, standardize the collected multidimensional data such as crime rate and crime success rate, and calculate the index weights and entropy values, to derive the cybersecurity coefficients of different countries [3]. On the other hand, in the policy assessment, the crime rate reduction rate is calculated by difference, the entropy weight method is again applied to analyze the relationship between the intensity of policy measures and the crime reduction rate, and the effect of the time of policy implementation on the crime rate is also investigated with the help of polynomial regression model [4]. Through the combined use of this series of methods, the assessment of network security stability and the identification of effective policies are studied [5]. This study improves the limitations of the traditional entropy weight method, with strong model stability and good predictive ability, which provides strong support and reference for cyber security research and practice [6].

2. Network Security Stability Assessment

2.1. Stability Assessment Modeling

To construct the target model, the data were first standardized to include collected crime rates, crime success rates, crime failure rates, crime prosecution rates, and crime reporting rates.

$$x'_{ij} = \frac{x_{ij} - \min(x_j)}{\max(x_j) - \min(x_j)} \tag{1}$$

Divided into five indicators, the weights of each sample value in each indicator are calculated:

$$p_{ij} = \frac{x'_{ij}}{\sum_{i=1}^{n} x'_{ij}} \tag{2}$$

Noting that the five national security indicators have different impacts, the study derives weights for crime rate and crime success rate, which are then back standardized.

Next, the entropy of each indicator is calculated by applying the following formula:

$$e_j = -k \sum_{i=1}^{n} p_{ij} \ln p_{ij}, k = \frac{1}{\ln n} \tag{3}$$

After obtaining the entropy value, the following formula is applied, and then the entropy weight of each indicator is calculated:

$$w_j = \frac{1-e_j}{\sum_{j=1}^{m} (1-e_j)} \tag{4}$$

2.2. Safety Stability Analysis Solving

Based on ordinary EWM, the study was adjusted according to the actual situation.

2.2.1. Flexibility in Setting Weights

Traditional entropy weighting methods rely solely on the degree of discretization (information entropy) of the data itself to calculate the weights and cannot incorporate domain knowledge and subjective judgment. In this study, the weights were set manually so that the weights could be adjusted. For example, the weights for crime rate and crime success rate were set high, while the weights for crime reporting rate, prosecution rate, and failure rate were set low.

2.2.2. Exception Handling in Data Normalization

During data normalization, if the minimum and maximum values of a column are the same, this may result in a denominator of zero, which can lead to calculation errors. The study added checking whether the minimum and maximum values are the same and fine-tuned the minimum and maximum values of these columns (by adding or subtracting a very small value), thus avoiding anomalies in the normalization process.

(1) Introduction to Random Noise

The presence of identical values in the data may cause the information entropy calculated by the entropy weighting method to be zero, thus preventing the correct assignment of weights. The study introduces a small random noise in the data to avoid this situation and ensure the stability of the entropy weighting method.

(2) Comprehensive Multi-Dimensional Assessment

The traditional entropy weighting method usually focuses on the assessment of a single dimension and cannot cope with the complexity of multiple dimensions at the same time. This paper combines multiple indicators (crime rate, crime success rate, crime reporting rate, prosecution rate and failure rate) for comprehensive assessment to make the results more comprehensive and reasonable.

(3) Interpretability of Weighted Scores

The traditional entropy weighting method only calculates weights and scores and lacks a clear explanation of the meaning of the scores. In this study, the code was rewritten to directly reflect the security of each country through the weighted scores, an improvement that makes the results more interpretable in practical applications.

After cracking the code, the study obtained the national cybersecurity coefficients of different countries and concluded that Japan is the most secure country among them.

From the above analysis, it can be understood that regions with high crime rates also have high control rates over cybercrime. The following main points can be summarized in this section:

First, the high incidence of cybercrime enables local governments to recognize the serious harm of cybercrime, thus focusing on the formulation of relevant policies and the implementation of actions; second, the relevant institutions in high-crime regions can better accumulate practical experience so as to formulate appropriate and effective strategies and carry out the upgrading of relevant

technologies; due to the trans-regional nature of cybercrime, high-crime regions will actively carry out international cooperation, and with other countries Because of the trans-regional nature of cybercrime, crime-prone regions will actively carry out international cooperation, share intelligence with other countries, and take joint measures to combat cybercrime more effectively; finally, strengthening public awareness and social supervision in crime-prone regions also provides a solid foundation for building a harmonious cyberenvironment. Because of the cross-regional nature of cybercrime, crime-prone regions will actively engage in international cooperation, share intelligence with other countries and work together to combat cybercrime, thus making measures to curb cybercrime more effective; finally, strengthening public awareness and social supervision in cybercrime-prone regions will also provide a solid foundation for building a harmonious cyberenvironment.

3. Effective Policy Identification Analysis

3.1. Data Collection and Quantification of Policy Evaluation Indicators

Using data collected from various websites related to cybercrime policies, this paper assesses and quantifies the presence of measures related to prevention, prosecution and mitigation in the policies and, if corresponding measures exist, translates their corresponding implementation intensity into objectively quantifiable metrics. Where 0 indicates that the measure is missing from the policy and 1 indicates that the measure is present. The intensity of the measures was evaluated using an ascending scale of 1 to 5. The processed data are shown in Fig. 1.

	prevention	prosecution	mitigation	strength		
USA	1	1	1	4	4	3
China	1	1	1	1	5	4
UK	1	1	1	3	4	3
Canada	1	1	0	4	4	0
Australia	1	1	1	4	4	4
German	1	1	1	5	5	4
French	1	1	1	4	4	3
India	1	0	1	3	0	4
Japan	1	1	1	4	3	4
Brazil	1	0	1	2	0	3
South Africa	1	1	0	3	3	0

Fig 1. Quantification of Policy Measures and Advantages.

By collecting data on the policy from various websites, this paper obtained crime rates before and after the enactment of the policy. After analyzing it, it can be concluded that the crime rate has decreased in all the countries after the enactment of the policy.

3.2. Modeling and Solving for Policy Evaluation

At first, this paper differenced the crime rates collected before and after the policy was implemented to obtain the rate of decrease in crime rates. Then, the relationship between the prevention, prosecution, and mitigation intensities of each policy and the crime reduction rate is plotted separately, as shown in Fig. 2.

Fig 2. Relationship between prevention, prosecution and mitigation intensities and crime reduction rates.

Preliminary results suggest that these three intensities do not have the same, or even very different, effects on crime reduction rates. Prevention intensity had almost no effect on crime reduction rates, while prosecution and mitigation intensities had a slightly greater effect on crime reduction rates. The three intensity values were then summed and compared to the crime reduction rate, and it was found that some policies that summed to very high intensities did not reduce crime very well, while some policies that summed to lower intensities did reduce crime very well. The study then proceeds to identify the proportion of each of the three intensities that will lead to the best reduction in crime rates.

Still applying the entropy power method, this section calculates the weights of each sample value of each of the three indicators, then calculates the entropy value of each indicator, and then calculates the entropy power of each indicator to obtain the results of the solution.

Utilizing the entropy power method, the result obtained is that increasing policy efforts will not achieve good results, and the related preventive measures are essential although they have little impact. Prosecution efforts should also include international cybersecurity cooperation, which can effectively strengthen the fight against transnational crime. Such cooperation is often effective in reducing cybercrime, especially in cases involving transnational cybercrime. Effective policies should also have feedback.

Thus, preliminary analysis suggests that a good policy should include both prevention, prosecution, and mitigation, with slightly lower weights for prevention and correspondingly higher weights for prosecution and mitigation.

In analyzing the data, it was again found that the timing of policy implementation and the impact of socioeconomic factors also had a significant impact on crime reduction rates. To investigate the effect of the time of policy implementation on the rate of crime reduction, the study developed a polynomial regression model to predict the change in the crime rate five years after the enactment of the policy as shown in Fig. 3, and to analyze the trend of the long-term effect of the policy after implementation.

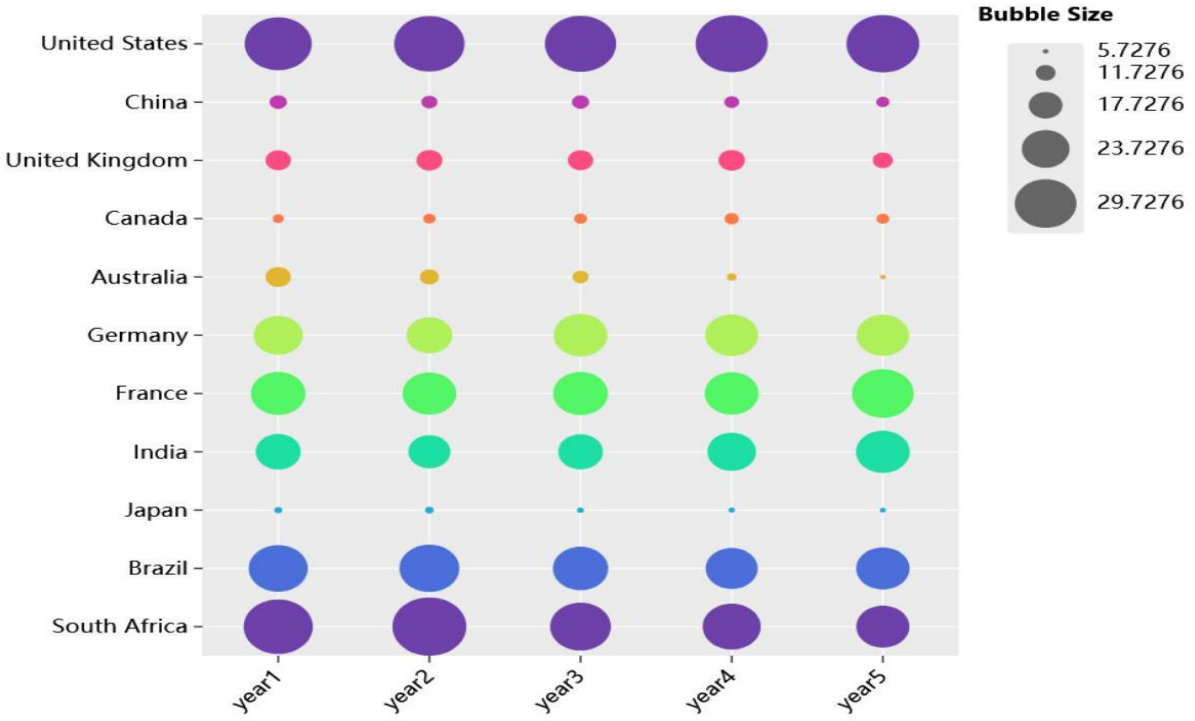

Fig 3. Changes in crime rates in the next five years.

Considering the initial period after the implementation of the new policy, the effect may not be obvious due to the adaptation period of the public and law enforcement agencies. Therefore, we added random fluctuations, i.e., simulated rebound, to the polynomial regression model to better obtain more realistic prediction results.

Modeling sub-polynomial regression:

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \cdots + \beta_n x^n + \epsilon \tag{5}$$

y Is the observed value of the dependent variable; $\beta_0, \beta_1 \ldots \beta_n$ is the regression coefficient to be estimated; ϵ is the error term, representing random fluctuations.

Next a simulated bounce is introduced, assuming a trend reversal at $x = a$, an adjustment term can be introduced to simulate this reversal:

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \cdots + \beta_n x^n + \alpha \cdot H(x - a) \cdot f(x) \tag{6}$$

$H(x - a)$ Is a Heaviside step function, that is?

$$H(x - a) = \begin{cases} 0, & x \le a \\ 1, & x > a \end{cases} \tag{7}$$

a Is the adjustment factor, which is used to control the strength of the inversion?

$f(x)$ Is a function on to characterize the change in trend of x after a reversal.

Decoding yields relevant results: among them, China and Australia show a downward trend from year to year, while other countries are fluctuating up and down and rebounding. Comparison with the previous section can further validate the conclusion that a good policy should include all three intensities of prevention, prosecution and mitigation at the same time, with a slightly lower weighting for prevention and a correspondingly higher weighting for prosecution and mitigation.

3.3. Model Testing

This section continues with the testing of the model: performing a goodness-of-fit assessment and a simple cross-validation to test the stability and predictive power of the model.

First, the mean square error of cross-validation is required:

$$MSE = \frac{1}{m}\sum_{i=1}^{m}\left(y_{\text{test } i} - \hat{y}_{\text{test } i}\right)^2 \tag{8}$$

Considering the complexity of the model, the coefficient of determination was adjusted by:

$$\bar{R}^2 = 1 - \frac{(n-1)}{(n-p-1)} \cdot \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{9}$$

The mean square error and the coefficient of determination are obtained for each country. The specific results are shown in Fig. 4.

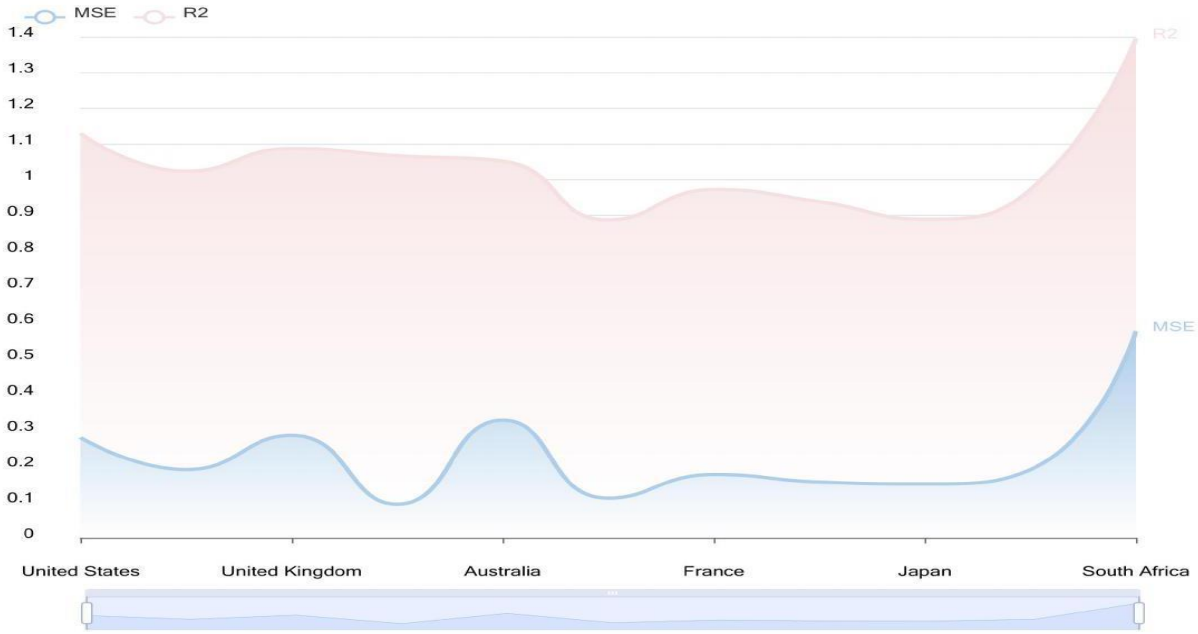

Fig 4. Test results.

From the obtained data, the mean square error for each country is very small and the coefficient of determination is close to 1, which can verify that the proposed model is more stable and has a strong predictive ability.

4. Conclusion

This paper centers on cybersecurity stability assessment and effective strategy identification and adopts methods such as entropy weight method and polynomial regression model. In stability assessment, the entropy power method is used to standardize the data of multiple indicators such as crime rate and crime success rate, calculate weights and entropy values, and conclude that Japan is the safest country, and it is also found that regions with high crime rates also have high rates of cybercrime control. In the policy assessment, the entropy weight method was applied to analyze the effects of policy prevention, prosecution and mitigation intensity on crime reduction rate, and the polynomial regression model was combined to study the role of policy implementation time, and the results showed that a good policy should contain prevention, prosecution and mitigation at the same time, and the prosecution and mitigation weights are high. The model was evaluated for goodness-of-fit and cross-validated for stability and good predictive ability. This study provides an effective method for quantitative assessment and policy formulation in the field of cybersecurity, which has important practical application value for improving cybersecurity and formulating reasonable policies.

References

[1] Chen Lu. Application of entropy weight method in information security risk assessment [J]. Information System Engineering, 2021, (09):62-64.

[2] Ren Xiaolei. Network security detection mechanism based on polynomial forward neural network [J]. Shanxi Electronic Technology, 2023, (06):77-79.

[3] Wang Xin. Simulation of effective prediction of network security defense posture [J]. Computer Simulation, 2017, 34(08):319-322.

[4] Ding Xiaoming. Policy Measures for Maintaining Information Security in Some Countries [J]. Contemporary World, 2002, (09):36-38.

[5] Wu Minfeng, Yang Yun, Sun Jianhu, et al. Network security threat identification and defense strategy based on big data analysis[J]. China Broadband, 2023, 19(10):4-6.

[6] Feng Dengguo. Research on the development trend of cyberspace security technology [J]. Information Security Research, 2025, 11(01):2-4.