



Data Analysis and Optimization Strategies for the Enrollment Structure of International Students in China Based on Computer Technology

Qiaoqiao Li*, Qianqian Yu

School of International Education, University of International Business and Economics, Beijing, China

DOI: 10.32629/jher.v5i1.2133

Abstract: With the continuous deepening of globalization and educational internationalization, the number of international students coming to China has been steadily increasing. Optimizing the enrollment structure of these students is significant for enhancing the internationalization level of China's higher education. This paper analyzes the data on the enrollment structure of international students in China using computer technology, discusses the existing issues, and proposes corresponding optimization strategies. The aim is to further promote the healthy development of international education in China.

Keywords: computer technology, international students in China, enrollment structure, data analysis

1. Introduction

In recent years, with the acceleration of China's educational internationalization process, the number of international students coming to China has been continuously increasing. Their diverse cultural backgrounds have injected new vitality into China's international education efforts. To better provide educational services to international students, it is crucial for educational institutions to deeply understand data concerning their nationalities, academic levels, and major choices[1]. Analyzing the general information of international students using computer technology can accurately grasp their needs and characteristics. Based on the results of enrollment structure data analysis, more targeted teaching and management strategies can be formulated. This can improve the education quality and life satisfaction of international students in China, support further optimized enrollment expansion, and attract more outstanding international students to study in China[2].

2. Analysis of the Current Enrollment Structure of International Students in China

2.1 Analysis of the Scale and Growth Trend of International Students

Since 2014, China has become the third largest host country for international students, following the United States and the United Kingdom, accounting for 8% of the global share of international students. Between 2016 and 2018, the number of international students in China continued to grow, covering more than 200 countries and regions worldwide. Although there have been fluctuations in recent years due to the pandemic, in the long run, the Chinese study abroad market still holds strong growth potential (see Table 1).

Table 1. Total Number of International Students in China from 2016 to 2022

Year	Total Number of International Students
2016	442,773
2017	489,200
2018	492,185
2019 (registered)	488,200
2019 (actual)	178,000
2020	392,000
2021	210,000
2022	440,000

Based on the current growth trend of international students in China, it can be predicted that during the future construction of the national education system, brand building will need to deepen openness and integration and improve education quality. It is essential to standardize the admission mechanisms to ensure the influx of high-quality students, improve the legal system

to protect the rights of international students, and optimize the socio-economic environment to provide a solid foundation for brand building. In the future, the “Study in China” brand will need to enhance the multi-subject educational capacity, offer diverse and high-quality educational resources, improve the study abroad experience, and strengthen online services.

2.2 Characteristics of the National Distribution of International Students

Analyzing the national distribution characteristics of international students in China reveals the broad influence and attractiveness of Chinese higher education on a global scale. The national distribution of international students is shown in Table 2.

Table 2. National Distribution of International Students in China

Rank	Country/Region	Number of International Students
1	South Korea	70,540
2	Thailand	28,608
3	Pakistan	23,198
4	USA	20,996
5	India	18,717
6	Russia	14,971
7	Japan	15,057
8	Indonesia	13,689
9	Vietnam	13,038
10	Kazakhstan	9,565

From the data, it is evident that Asian countries are the main sources of international students, with South Korea, Thailand, Pakistan, and India ranking at the top. This reflects the close cooperation and exchanges in education between China and its neighboring Asian countries. South Korea, as the largest source country, has more than 14% of its students in China, indicating the strong educational ties and high recognition of Chinese higher education among Korean students. Cultural and educational exchanges between China and countries like Thailand, Pakistan, and India are becoming more frequent, leading to a significant number of students from these countries choosing to study in China. Additionally, students from Western countries such as the United States and Russia also choose to study in China, demonstrating the expanding international influence of Chinese higher education. According to statistics from the Comparative and International Education Society (CIES), Asian students account for 9.84% of international students in China, Europeans for 16.11%, Africans for 13.91%, Americans for 8.60%, and Oceanians for 1.54%.

2.3 Preferences in Academic Disciplines and Majors of International Students

When discussing the major choices of international students in Chinese universities, it is evident that their interests span a wide range of fields, from science and engineering to humanities and social sciences, as well as medicine, agriculture, law, and the arts. Many students choose to study Chinese language and culture due to their love for China, aiming to gain a deeper understanding of this ancient civilization. There is a significant interest in traditional Chinese culture majors, such as Chinese classical literature, calligraphy, and painting. Additionally, some international students focus on China’s economy, management, and law within the social sciences to gain insights into China’s social and economic environment. Science and engineering majors like computer science, mechanical engineering, and electronic engineering are also popular, providing valuable learning opportunities for those intending to work or start businesses in China[3].

2.4 Distribution of Educational Levels among International Students

Although China ranks third in the world for the total number of international students, the proportion of non-degree students is relatively high. In terms of educational levels, China ranks fifth globally in the number of undergraduate international students. However, there is a significant gap compared to the United States and the United Kingdom, which hold the first and second positions respectively. For postgraduate students (masters and doctoral), China ranks only sixth, behind the United States, the United Kingdom, Germany, Australia, and Russia. Specifically, the proportion of postgraduate students in China is only 17.3%, which is considerably lower compared to 53.7% in Germany and 48.6% in Australia in 2018.

Overall, the combined proportion of undergraduate, master’s, and doctoral students among international students in

China is only 49.9%, similar to Japan but far below the United States' over 70% and more than 90% in Russia, Australia, and the United Kingdom. This indicates that there is substantial room for improvement in the structure of international students in China, particularly in increasing the number and proportion of high-level talents to further optimize the international student composition.

3. Issues in the Enrollment Structure of International Students in China

3.1 Homogeneous National Distribution, Lack of Diversity

Based on the statistical analysis mentioned above, the national distribution of international students in China is primarily concentrated in Asian regions, especially neighboring countries. This structure results in a homogeneous cultural background among international students, which is not conducive to cultivating talents with a global perspective and intercultural communication skills. In other words, the overly concentrated structure limits diversity among international students and affects the depth and breadth of the internationalization of Chinese higher education. Additionally, a single-source national structure limits interaction between international and Chinese students, impacting international students' understanding and appreciation of Chinese culture.

3.2 Over-Concentration in Academic Disciplines and Majors, Lack of Balance

In terms of major choices, international students in China tend to favor humanities and social sciences, with relatively fewer students in science, engineering, and medical fields. This imbalance is not conducive to cultivating well-rounded and innovative talents and affects the overall development level of Chinese higher education. If international students do not adjust their choices of majors, it will inevitably lead to intense competition in popular majors and difficulties in enrolling students in less popular ones, wasting educational resources and lowering the quality of teaching. Moreover, an over-concentration in certain majors affects the employment prospects of international students, exposing them to greater competition in the job market.

3.3 Low Educational Level, Insufficient Proportion of High-Level Students

The educational level of international students in China is generally low, with an insufficient proportion of high-level students, which is a prominent issue. Compared to developed countries, the proportion of high-level students, such as master's and doctoral students, is still low. This limits the academic development of international students, particularly in cultivating innovation and research skills. Furthermore, the low educational level somewhat restricts the internationalization level of Chinese higher education and affects its attractiveness to outstanding international students[4].

4. Data Analysis Methods Based on Computer Technology

4.1 Data Collection and Preprocessing Techniques

Data collection is the foundation of data analysis. For analyzing the enrollment structure of international students in China, data needs to be gathered from sources such as the Ministry of Education, universities, and study abroad agencies. This data should include information on students' nationalities, major choices, educational levels, scholarship status, and post-graduation destinations. Utilizing computer technology enables efficient collection, integration, and cleaning of large datasets, ensuring data quality and accuracy. The subsequent data preprocessing steps involve cleaning the data to remove duplicates, errors, or invalid entries, converting the data into a format suitable for analysis, and integrating data from different sources into a unified dataset. During the preprocessing stage, feature values are scaled to have a mean of 0 and a standard deviation of 1. The formula for standardization is:

$$z = \frac{x - \mu}{\sigma} \quad (1)$$

Here, z is the standardized data, x is the original data, μ is the mean, and σ is the standard deviation.

Normalization is a process that scales data to a range between 0 and 1. The formula for normalization is:

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (2)$$

Here, x_{min} is the minimum value of the data, and x_{max} is the maximum value of the data.

For skewed distributed data, logarithmic transformation can be used to make it closer to a normal distribution. The formula is:

$$y = \log(x) \quad (3)$$

The Box-Cox transformation is a power transformation method used to stabilize variance and make data closer to a normal distribution. The transformation formula is:

$$y = \begin{cases} \frac{x^\tau - 1}{\tau}, & \text{if } \tau \neq 0 \\ \ln(x), & \text{if } \tau = 0 \end{cases} \quad (4)$$

In the formula, τ is the parameter of the Box-Cox transformation.

4.2 Statistical Analysis and Visualization

During statistical analysis, computer technology allows for the rapid completion of descriptive statistics. By synthesizing frequency distributions, means, medians, etc., and utilizing correlation analysis to explore relationships between variables, differences between different groups can be compared through variance analysis. The distribution of international students can be displayed using bar charts, pie charts, line graphs, etc., while heat maps, dendrograms, etc., can be used to reveal relationships between variables.

4.3 Data Mining and Association Analysis Techniques

For the enrollment structure data of international students in China, classification techniques can be used to categorize student groups, clustering techniques to discover similarities or differences among students, and anomaly detection techniques to identify abnormal data. Association analysis is a key technique used to discover relationships between variables, such as analyzing major choices and educational levels to reveal the attractiveness of specific majors to high-level students or the major preferences of students from specific countries. Taking statistical data from a university as an example, the association data between the nationality of international students, major preferences, and scholarship rates are shown in Table 3:

Table 3. Association Data between the Nationality of International Students, Major Preferences, and Scholarship Rates

Nationality	Engineering Majors	Arts Majors	Scholarship Rate
USA	43	27	62%
South Korea	31	41	54%
India	57	13	48%
Russia	29	35	59%

4.4 Prediction Models and Trend Analysis

Combining historical data to establish time series analysis models, regression analysis models, etc., to predict the enrollment structure of international students in China, helps relevant departments and schools formulate more reasonable enrollment plans and education policies. Time series analysis models are typically used to analyze data sequences that change over time. For predicting the number of international students, commonly used time series models include Autoregressive Moving Average (ARMA) models, Autoregressive Integrated Moving Average (ARIMA) models, etc. The general form of an ARMA model is denoted as (p, d, q), where p is the order of the autoregressive terms, d is the differencing order, and q is the order of the moving average terms. The mathematical expression of the ARIMA model is more complex, but the basic idea is to make a non-stationary series stationary through differencing, and then fit an ARMA model:

$$(1 - \phi_1 B - \dots - \phi_p B^p)(1 - B)^d X_t = (1 + \theta_1 B - \dots + \theta_q B^q) Z_t \quad (5)$$

Here, B is the lag operator, ϕ_i and θ_i are the parameters of the model, Z_t is the white noise series.

$$Y = \alpha + \beta_1 X_1 + \beta_2 \ln(X_2) + \beta_3 \ln(X_3) + \dots + \beta_n \ln(X_n) + \delta \quad (6)$$

In formula 6, α is the constant term, β is the regression coefficient, \ln is the natural logarithm, $\hat{\epsilon}$ is the random error term, and X_n represents n influencing factors.

In trend analysis, besides using prediction models, data mining techniques can also be employed to identify key factors influencing the changes in the number of international students (such as policy changes, international situations, etc.), and further analyze how these factors affect the future trends in the number of international students. To validate the application effectiveness of algorithms based on computer technology in prediction, and to elucidate future development trends, historical total enrollment data of international students from a certain university is collected and organized, as shown in Table 4:

Table 4. Historical Total Enrollment Data of International Students

Year	Total Enrollment
2017	387
2018	412
2019	439
2020	466
2021	494
2022	523

It is necessary to check whether the residuals of the model are random, ensuring that there are no obvious patterns or trends. Additionally, the model should pass the Box-Ljung test. If actual data is available, the model's accuracy should be evaluated by comparing the predicted values with the actual values and calculating prediction errors, such as Mean Squared Error (MSE) or Mean Absolute Error (MAE). The predicted total enrollments for the years 2023 to 2026 are 550, 578, 607, and 637 respectively. The prediction results are influenced to some extent by various external factors such as policies, international situations, and educational quality. Therefore, caution should be exercised when interpreting the prediction results.

5. Optimization Strategies for the Enrollment Structure of International Students in China

5.1 Adjusting Enrollment Strategies to Expand the Diversity of Source Countries

Universities should intensify global recruitment efforts, enhancing the international reputation of Chinese higher education through international education fairs, international recruitment agent networks, social media, etc. Tailored recruitment policies should be developed for different countries and regions, emphasizing teaching quality and research strength to attract students from developed countries, and offering preferential policies such as scholarships and employment opportunities to attract students from developing countries. Additionally, strengthening cooperation with renowned international intermediary organizations can improve enrollment efficiency and success rates.

5.2 Optimizing Discipline and Major Offerings to Meet Diverse Learning Needs

Based on international market demand and student interests, universities should actively offer majors and directions with international characteristics, such as strengthening the integration of traditional Chinese culture, international trade and economics, and technological innovation. Simultaneously, enhancing teaching quality and research levels, providing high-quality educational resources and learning environments for international students, introducing outstanding international teachers, and strengthening international cooperation are essential. Attention should be paid to cultivating international students' cross-cultural communication abilities through relevant courses and cultural activities to help them better adapt to the Chinese environment and enhance their international perspectives and competitiveness.

5.3 Enhancing Educational Levels to Attract More High-Level International Students

To attract more high-level international students, universities should strengthen enrollment efforts at higher levels of education such as master's and doctoral programs, and provide preferential policies such as scholarships. Emphasis should be placed on cultivating international students' innovation capabilities and research literacy, conducting joint training programs with internationally renowned universities, providing research opportunities and resources, and encouraging international students to participate in research projects and academic activities.

5.4 Strengthening International Cooperation and Exchange to Improve the Quality of International Student Education

Universities should collaborate with renowned international universities to advance teaching, research, and talent development. International academic exchanges should be promoted through methods such as teacher visits and student exchanges. Actively participating in international education organizations and projects can enhance the international influence of Chinese higher education. Additionally, strengthening cooperation with international companies to provide practical and employment opportunities for international students, promoting deep integration of industry, academia, and research, and driving the innovative development and internationalization of Chinese higher education.

6. Conclusion

In conclusion, with the continuous increase in the number of international students coming to China, optimizing their source structure is crucial for enhancing the internationalization of Chinese higher education. This paper has analyzed the current status of the enrollment structure and identified issues such as a single distribution of source countries, concentration in major selection, and lower educational levels among international students in China. To address these challenges, universities need to adjust their enrollment strategies, optimize discipline settings, improve educational levels, and strengthen international cooperation and exchange. This will attract more outstanding international students, enhance the quality of education and satisfaction of international students, and further promote the internationalization of Chinese higher education, providing a more diverse range of learning opportunities for global students.

Acknowledgments

This paper is supported by “the Fundamental Research Funds for the Central Universities” in UIBE (project number: 22YB07).

References

- [1] Li, Z., Yang, H., Yang, J., et al. (2022). Practice and Research on the Construction of English-Taught Courses for International Students in China — Taking “Application of Computer Software in Engineering Technology” as an Example. *Advertising View*, (33), 109-111.
- [2] Mi, S., Yu, W., Liu, W., et al. (2022). Practice and Exploration of the Teaching of “Advanced Food Chemistry” Course for International Students in China. *Agricultural Products Processing*, (20), 114-117.
- [3] Tao, J., & Qi, H. (2022). Policy Evolution, Structural Characteristics, and Optimization Strategies of Education for International Students in China. *Heilongjiang Higher Education Research*, 40(10), 7.
- [4] Li, Z., & Li, S. (2022). How to Achieve Connotative Development of Education for International Students in China - Based on the Background of the “Belt and Road” Initiative. *Journal of Ningbo University: Education Science Edition*, 44(1), 9.