Problem Chosen	2021 ShuWei Cup Summary Sheet	Team Control Number
C		202111046885

Research on Tourism Development Planning

Summary

It is necessary to study the future development planning of COVID-19's tourism industry.

For question one, firstly, collect the 204 5A scenic spots distributed in various provinces across the country, Secondly, we use **text mining** to crawl the tourist comment data, visualize the data, and get the visual processing map through word frequency analysis, Finally, through data classification and preprocessing, the distribution characteristics of scenic spots in China are obtained intuitively. The density of 5A scenic spots **in East China is the highest**; The density of 5A scenic spots **in North China is also relatively high**, The density of 5A scenic spots **in South China and central China is also relatively concentrated**; However, there are **relatively few 5A scenic spots in the northwest, northeast and southwest.**

For question two, Firstly, using the reception capacity of 5A scenic spots in various provinces in China collected in question 1, the daily carrying capacity and the current evaluation level of scenic spots are standardized, and a two-dimensional score data set is obtained by transforming the two groups into an interval value of 0-10, Then combined with **k-means** algorithm to divide the universe, quantitatively analyze the two groups of data characteristics, classify these scenic spots, get the results of **K-means cluster**, and divide the scenic spots into three grades: high, medium and low.

For question three, taking the Beijing Forbidden City Museum as an example for quantitative analysis, **NSGA-II multi-objective optimization** is selected as the algorithm to solve the problem. in which the epidemic level and scenic spot personnel flow are the two constraints of the multi-objective optimization problem. The two main optimization objective functions, tourist safety factor and scenic spot income, are determined. It is concluded that when the epidemic level is **level I**, the flow limit of the scenic spot is **60%**; When the epidemic level is **level II**, the flow limit of the scenic spot is **50%**; When the epidemic level is **level III**, the flow limit of the scenic spot is **30%**.

For problem four, by constructing **the model of cellular automata**, it can better simulate the flow restriction of scenic spots during business under the epidemic situation. That is, the flow limiting degree of the known termination conditions reaches the maximum value, and the mathematical model considering not only the safety factor of tourists, but also the passenger flow and the income of the scenic spot is calculated through the influence of the monitoring number, distance and environmental factors.

For problem five, flow restriction is the basic premise for the opening of scenic spots under the normalized epidemic prevention and control, and it is also the basis of appointment management. On the premise of ensuring that the scenic spots do a good job in the epidemic prevention and control of scenic spots, adhere to the principle of zoning and grading, and put forward differentiated management schemes for different scenic spots for the government management in view of the existing problems at this stage, hoping to have reference significance for the economic recovery of local scenic spots.

Therefore, in the context of the epidemic situation, it is of great significance to make reasonable planning and correct prediction of tourism and how to weigh the advantages and disadvantages of all parties to promote the development of tourism and cultural exchange.

Keywords: Text mining; Data visualization processing; k-means clustering; NSGA-II multi-objective optimization algorithm; Cellular automata;

Contents

1. Introduce

1.1 Background

With the development of society, tourism has become the largest economic industry in the world, and China is one of the countries with the fastest development of tourism in the world, and has rich tourism resources. The novel coronavirus pneumonia is the most serious public health emergency in China since its establishment. Tourism was strongly impacted by the epidemic.

Looking forward to the future, we should give consideration to the long-term. Taking the opportunity of dealing with the epidemic, we should further strengthen, optimize and refine relevant supporting policies, strengthen the guidance of the post epidemic tourism market, promote the stable and standardized development of the tourism employment market, attach great importance to the domestic tourism market, guide the return of national outbound tourism consumption, and make a systematic plan for revitalizing the inbound tourism market, And promote a new round of reform, opening up and high-quality development of tourism through a series of measures. The comparison of the number of scenic spots before and after COVID-19 is shown in figures 1-1-1 and 1-1-2.

Therefore, the rational planning and correct prediction of tourism demand and how to weigh the advantages and disadvantages of all parties under the background of the epidemic are of great significance to promote the development of tourism and cultural exchange.

Figure 1-1-2 before the epidemic

Figure 1-1-2 post epidemic situation

1.2 Work

① Collect and analyze the basic data of 5A scenic spots in various provinces in China.

② Collect and count the current evaluation level and daily reception capacity of each 5A scenic spot.

③ Assess the reception capacity of each scenic spot and classify it.

④ Provide a reasonable flow restriction model for scenic spots with different epidemic levels.

⑤ A quantitative model considering epidemic prevention, tourism income and tourism experience is proposed and simulated.

⑥ Provide differentiated management schemes for different scenic spots during the epidemic.

2. Problem analysis

2.1 Data analysis

According to the data analysis, the statistical map of the number distribution of 5A tourist attractions in China is shown in Figure 2-1:

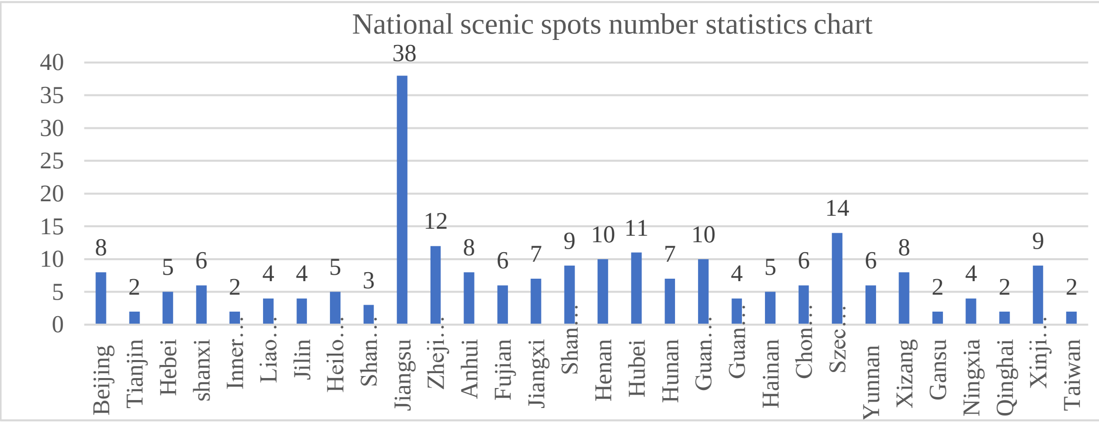

Figure 2-1 distribution statistics of 5A tourist attractions in China

See Appendix 1 for the basic data of scenic spots in all provinces, the current evaluation level and daily maximum reception capacity of each scenic spot.2.2.

2.2 Analysis of problem one

For question one, first, collect 5A scenic spots distributed in various provinces across the country,there are differences in the spatial distribution of 5A scenic spots in various provinces across the country, including 7 subordinate regions of Northeast China, North China, Northwest China, East China, central China, South China and southwest China.

Secondly, text mining is used to visualize the data, eliminate meaningless words, retain words that can intuitively reflect the subjective evaluation of tourists on different scenic spots, and then analyze the word frequency of the reserved words, finally get the visual processing map and get the current evaluation level of the scenic spot.

Finally, through data classification and preprocessing, the distribution characteristics of scenic spots in China are obtained intuitively.

2.3 Analysis of question two

For question two, the daily carrying capacity and the current evaluation level of each 5A scenic spot collected in problem 1 are standardized. By transforming the two groups of processing into an interval value of 0-10, a two-dimensional score data set is obtained, and then the unequal division of the domain is carried out in combination with the k-means algorithm to quantitatively analyze the characteristics of the two groups of data, and classify these scenic spots.

2.4 Analysis of question three

For problem 3, the NSGA - Ⅱ multi-objective optimization algorithm is used to analyze and summarize two main objective functions, different epidemic levels and two constraints of scenic spot personnel flow, so as to realize the mathematical modeling of multi-objective optimization problem. Then take the Beijing Forbidden City Museum as an example for quantitative analysis, and provide a reasonable flow restriction model for the scenic spot according to different epidemic levels. Under the background of epidemic situation, the scenic spot should maximize the income on the premise of fully considering the safety factor of tourists.

2.5 Analysis of question four

For problem 4, the possible and ongoing phenomena are simulated and analyzed by building a model considering epidemic prevention, tourism revenue and tourism experience. Among them, the model based on cellular automata can better simulate the current limit of scenic spots during business under the epidemic situation. This paper mainly studies the relationship between flow restriction and tourism income and tourism experience by using cellular automata to establish the corresponding model under the condition of a certain epidemic level.

2.6 Analysis of question five

For question five, flow restriction is not only the basic premise for the opening of scenic spots under the normalized epidemic prevention and control, but also the basis of appointment management. The long-term depressed travel demand of tourists and the urgent recovery of the tourism industry have exacerbated the difficulties of epidemic prevention and control. Under the premise of ensuring the stable control of the epidemic situation, the scenic spots adhere to the principle of zoning and grading on the premise of doing a good job in the epidemic prevention and control of scenic spots, and put forward differentiated management plans for different scenic spots for the government management in view of the existing problems at this stage, hoping to have reference significance for the economic recovery of local scenic spots.

3. Symbols and Assumptions

3.1 Symbol Description

Serial number	Symbol	significance
1	n	Safety factor

2	m	Scenic spot income
3	a	Daily carrying capacity
4	$a_{\max}$	Maximum daily carrying capacity
5	c	Per capita consumption
6	k	Per capita service cost of scenic spot
7	a_x	Bearing capacity after current limiting

* Other unspecified symbols are described in the text

3.2 Fundamental assumptions

1. Assuming that the economic and political environment inside and outside China is relatively stable.

2. It is assumed that China's policy will not change in the short term.

3. Assuming that the re demand development of China's tourism industry is gradual, that is, the epidemic situation is controllable and the development is stable.

4. It is assumed that China's tourism resources will not change greatly in recent years.

5. It is assumed that the influence of various factors follows its development law.

4. Model

4.1 Model I

4.1.1 Modeling process

(1) Analysis on spatial pattern of regional 5A scenic spots

According to the terrain, climate, humanities, economy and politics, China is divided into seven geographical regions: Northeast, North China, East China, central China, South China and southwest. The spatial distribution characteristics of scenic spots are mainly to analyze the location, commonness and individual differences of each geographical region of scenic spots in different regions.

(2) Three detailed rules for the current evaluation level of 5A scenic spot

Rule 1: scoring rules for service quality and environmental quality

Rule 2: scoring rules for landscape quality

Rule 3: scoring rules for tourists' opinions

(3) Maximum daily reception capacity of 5A scenic spot

The maximum reception capacity of the scenic spot refers to the maximum number of tourists that the scenic spot can accommodate under certain time conditions and on the premise of ensuring the personal safety of tourists and the safety of tourism resources and environment in each scenic spot. Different types of scenic spots have different estimation methods or basic spatial bearing standards of the maximum bearing capacity, and the numerical magnitude of the corresponding bearing capacity is also different.

4.1.2 Establishment and solution of model

Firstly, text mining is used to collect the current evaluation level of the scenic spot. The flow chart is shown in 4-1-1:

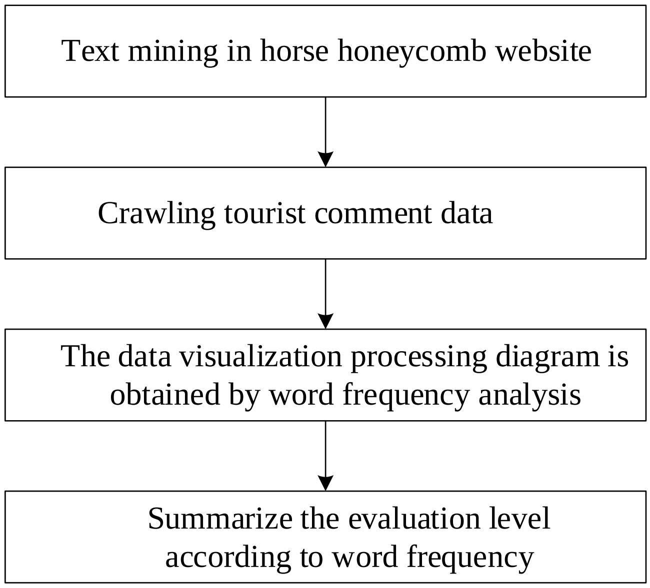

Figure 4-1-1 flow chart of text mining

Step 1 Get text: crawl and record the data of tourists' comments in different scenic spots on the horse honeycomb website. To get the text in the network to the text database (dataset). Write a crawler program to grab information from the network

Step 2 Visualize the data.

Step 3 Quantify the data without outline and eliminate words with no practical significance. After the above steps, you will get relatively clean materials. Some words play a key role in the text, and even the main words can determine the text orientation.

Step 4 Retain the valuable words after analyzing the original data, that is, the words that can intuitively reflect the subjective evaluation of tourists on different scenic spots, then analyze the word frequency of the retained words, and finally obtain the visual processing diagram, as shown in figure 4-1-2:

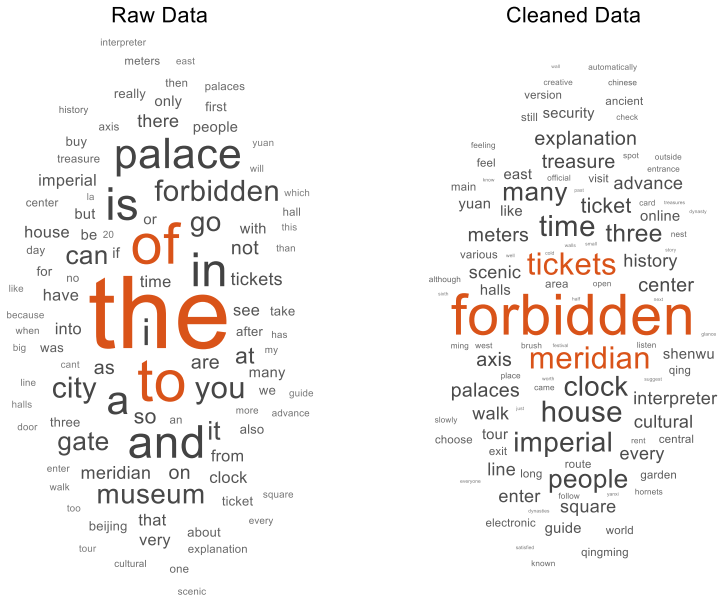

Figure 4-1-2 data visualization processing

By the end of 2020, 204 5A level scenic spots in all provinces of China were located in 34 provinces in Chinese mainland (including 5 autonomous regions, 4 municipalities and 2 special administrative regions). The location distribution of 5A scenic spots across the country in the whole country is shown in figure 4-1-3.

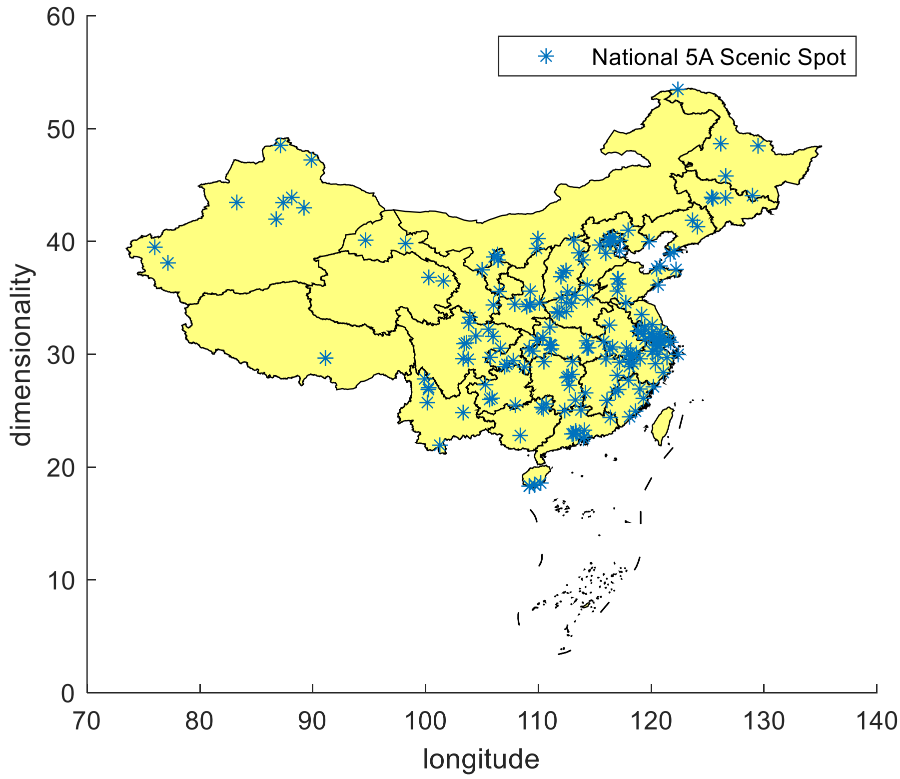

Figure 4-1-3 distribution characteristics of 5A scenic spots in China

First of all, the density of 5A scenic spots in East China (Jiangsu and Zhejiang) is the highest. The reason is obvious. Jiangsu in East China has been a fertile land and a land of fish and rice since ancient times. Its economic strength ranks among the top five in China, and forms the Yangtze River Delta urban agglomeration together with Zhejiang and Shanghai. Secondly, the density of 5A scenic spots in Beijing Tianjin Hebei region with Beijing as the center in North China is also relatively high, especially in the Yangtze River Delta; South China and central China, such as 5A scenic spots around Guangdong in South China and Southeast China centered on Jiangxi in Central China, are also relatively concentrated; In addition, relatively speaking, there are relatively few 5A scenic spots in the northwest, northeast and southwest.

4.2 Model II

4.2.1 Modeling process

Step 1 Standardize the daily carrying capacity (x) and the score (y) of the current evaluation level of the scenic spot.

$$x_t = 10 \cdot \frac{x_{\min}}{x_{\max} - x_{\min}} \tag{1}$$

$$y_t = 10 \cdot \frac{y_{\min}}{y_{\max} - y_{\min}} \tag{2}$$

Step 2 The two groups of processed data are transformed into interval values of 0-10 to obtain a two-dimensional score data set.
Step 3 K-means clustering is performed, and the flow chart is shown in 4-2-1:

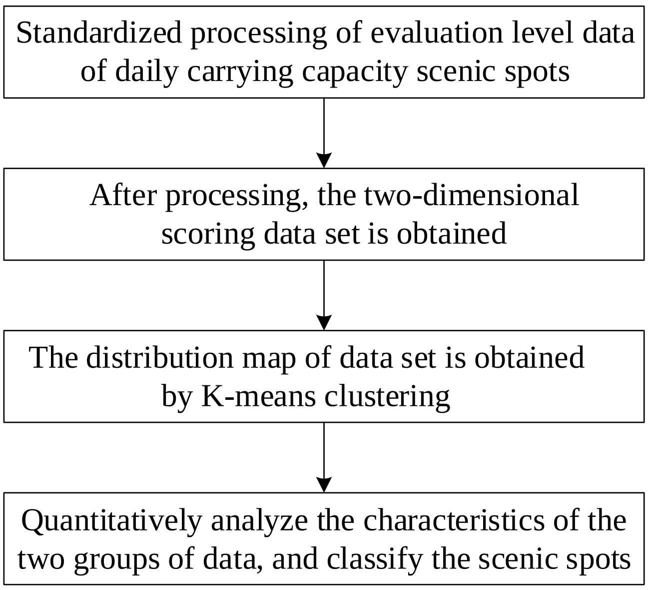

Figure 4-2-1 flow chart of model II

4.2.2 Establishment of model

Firstly, the data is preliminarily processed according to an automatic clustering algorithm, and then the non equal division of the universe is carried out combined with the k-means algorithm.

Among them, the basic steps of K-average clustering method are:

(1) Suppose K classes are clustered to determine K class centers $Z_1(1)$, $Z_2(1)$, …, $Z_k(1)$.

(2) In iteration K, the sample set $\{Z\}$ is classified as follows:

For all $i=1, 2, \ldots, K$, $i \neq j$

If $\|Z - Z_j(k)\| < \|Z - Z_i(k)\|$, therefore $Z \in S_j(k)$.

(3) Let the new class center of $S_j(k)$ obtained from (2) be $Z_j(k+1)$

Make $J_j = \sum_{Z \in S_j(k)} \|Z - Z_j(k+1)\|^2$ minimum, $j=1, 2, \ldots, K$

therefore $Z_j(k+1) = \frac{1}{N_j} \sum_{Z \in S_j(k)} Z$, number of samples in $N_j : S_j(k)$.

(4) For all $j=1, 2, \ldots, K$, if $Z_j(k+1) = Z_j(k)$, then terminate.

Pay attention to the selection of K at the beginning, which will affect the iteration results; The clustering method needs many iterations and revisions;

The schematic diagram of cluster analysis is shown in figure 4-2-2:

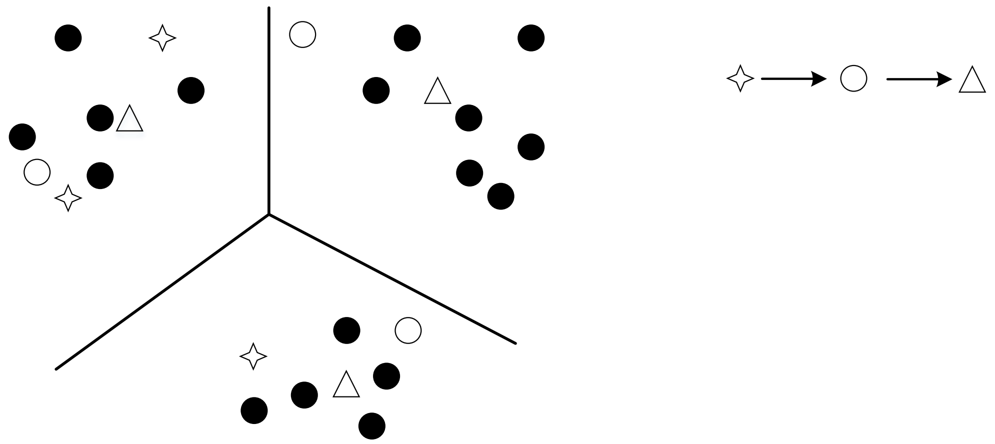

Figure 4-2-2 schematic diagram of cluster analysis

4.2.3 Solution of model

By comparing the results of the original dataset in figure 4-2-3 with the K-means cluster in figure 4-2-4, it can be seen that the score of the blue area is the highest, the score of the green area is high, and the score of the red area is average.

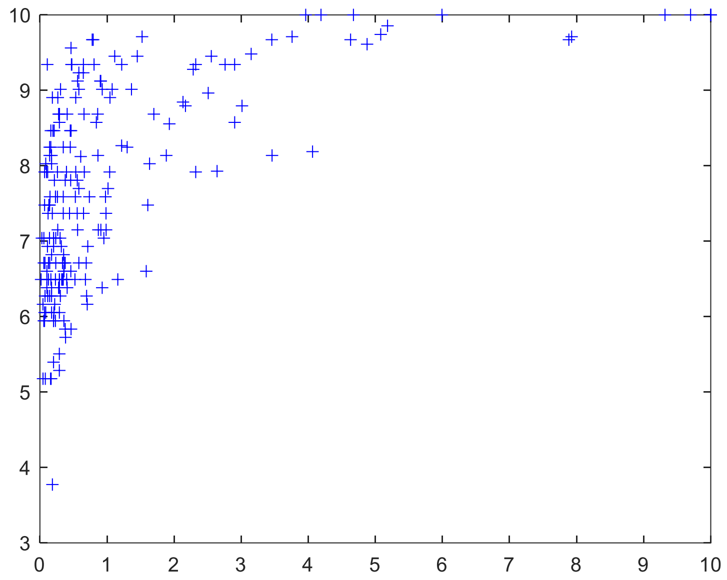

Figure 4-2-3 original data set

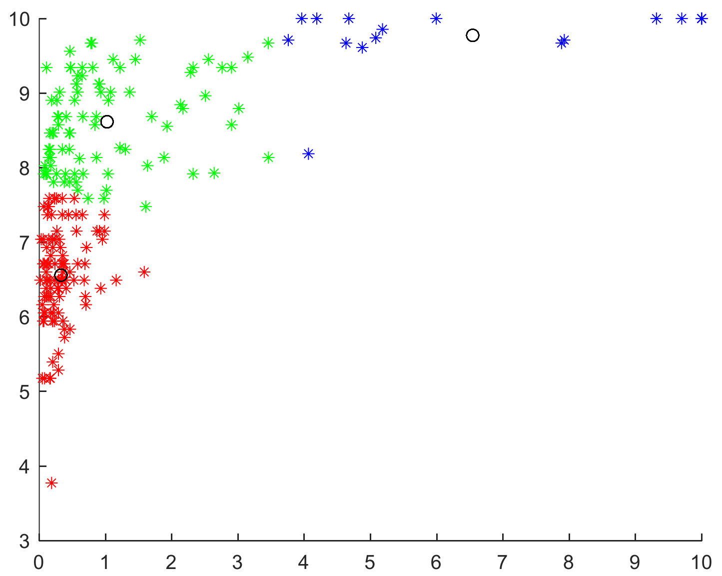

Figure 4-2-4 results of K-means cluster

4.3 Model III

4.3.1 Modeling process

NSGA- II is selected as the algorithm basis for solving the multi-objective optimization problem. Two main objective functions, different epidemic levels and two constraints of scenic spot personnel flow are analyzed and summarized, and the mathematical modeling of the multi-objective optimization problem is realized.

NSGA - II algorithm flow is shown in figure 4-3-1:

First, the first generation parent population P1 is generated, and then the decoding process is performed.

Then the individual fitness is determined by non dominated sorting and crowding degree calculation. According to the results of fitness calculation, the first generation population is selected, crossed and mutated.

Gen represents the number of iterations, Gen = 1, forming the sub population qgen of Gen generation. The sub population is further decoded. Then, according to the requirements of NSGA-II algorithm, the sub population and parent population are combined to form a large population P=pgen ∪ qgen. According to the calculation results, individuals with high fitness are selected to form a new generation of parent population pgen+1. Gen=Gen+1, and then, based on the new generation parent population pgen, carry out genetic operations such as selection, crossover and variation to form a new generation child population qgen. Cycle until the algorithm termination condition is reached.

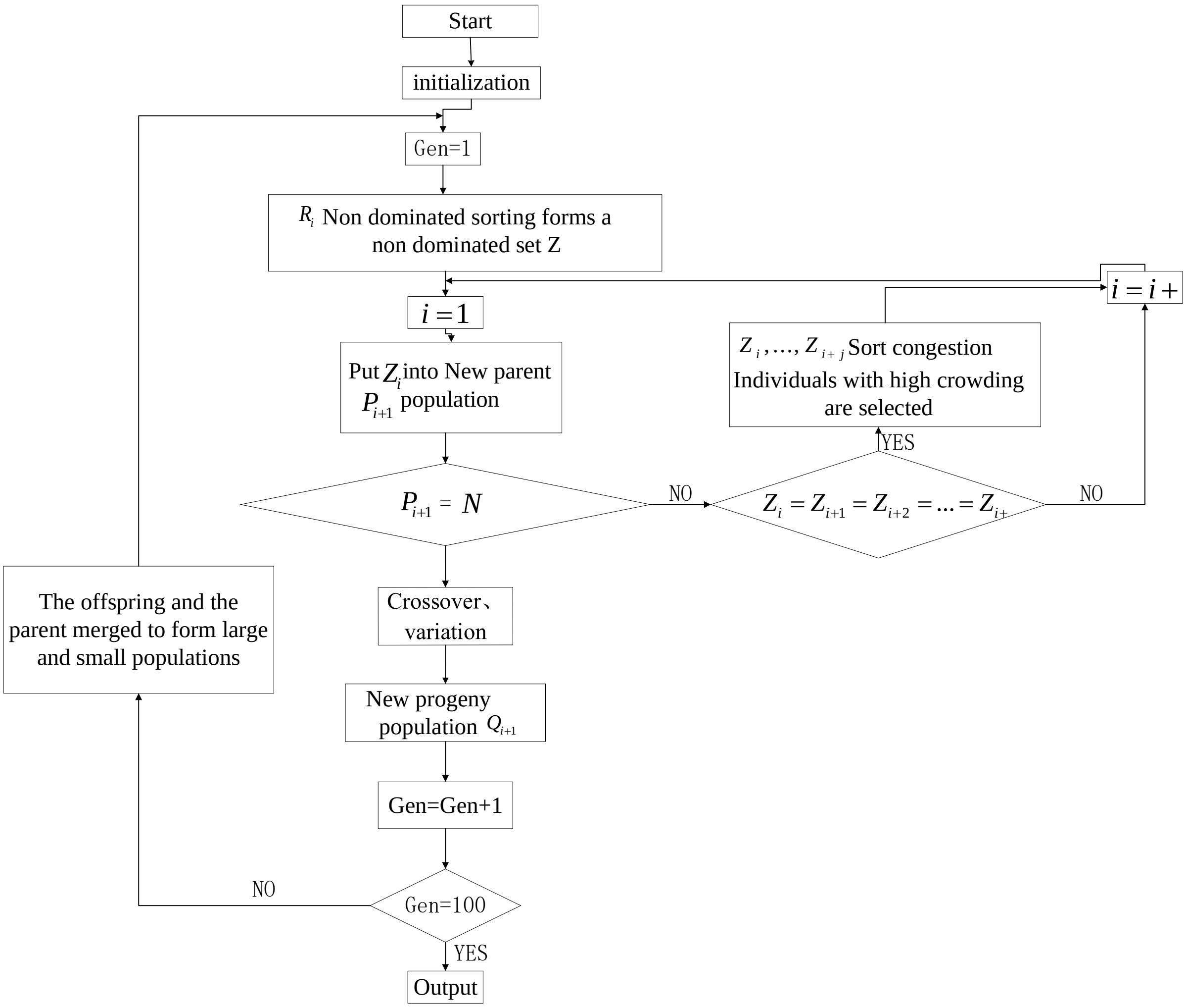

Figure 4-3-1 algorithm flow of model III

4.3.2 Establishment of model

NSGA - II multi-objective optimization algorithm uses agglomeration distance to calculate congestion. The aggregation distance of an individual is the sum of the distance difference between the abscissa and ordinate of the sum of two adjacent individuals.

p_4 gathering distance: $p_4 - cd = \left|\frac{f_1(3) - f_1(5)}{f_1^{\max} - f_1^{\min}}\right| + \left|\frac{f_2(3) - f_2(5)}{f_2^{\max} - f_2^{\min}}\right|$ 。

p_7 gathering distance: $p_7 - cd = \left|\frac{f_1(6) - f_1(8)}{f_1^{\max} - f_1^{\min}}\right| + \left|\frac{f_2(6) - f_2(8)}{f_2^{\max} - f_2^{\min}}\right|$ 。

The larger the agglomeration distance, the smaller the congestion. $CD_7 > CD_4$ .That is, the agglomeration distance of p_7 is larger than that of p_4, but the crowding degree of p_7 is smaller than that of p_4, which means higher fitness. Because p_1 and p_9 are at the edge, their crowding degree is infinite.

4.3.3 Solution of model

Taking the Beijing Palace Museum as an example for quantitative analysis, NSGA - II is selected as the algorithm to solve the multi-objective optimization problem. Under the background of epidemic situation, the scenic spot should maximize the income on the premise of fully considering the safety factor of tourists.

The epidemic level is set as the parameter α , and the two main optimization objective functions, tourist safety factor and scenic spot income, are determined as follows:

$$n = \frac{a}{a_{\max}} \tag{1}$$

$$m = cx - kx\,, \quad k = 1,2,3...... \tag{2}$$

The epidemic level and scenic spot personnel flow are the two constraints of the multi-objective optimization problem, and the constraint equation is as follows:

$$a_x = ba_{\max} \tag{3}$$

$$k = p\alpha \tag{4}$$

It is concluded that when the epidemic level is level I, the flow limit of the scenic spot is 60%; When the epidemic level is level II, the flow limit of the scenic spot is 50%; When the epidemic level is level III, the flow limit of the scenic spot is 30%.

4.4 Model IV

4.4.1 Modeling process

The flow chart of comprehensive thinking of question 4 is shown in figure 4-4-1:

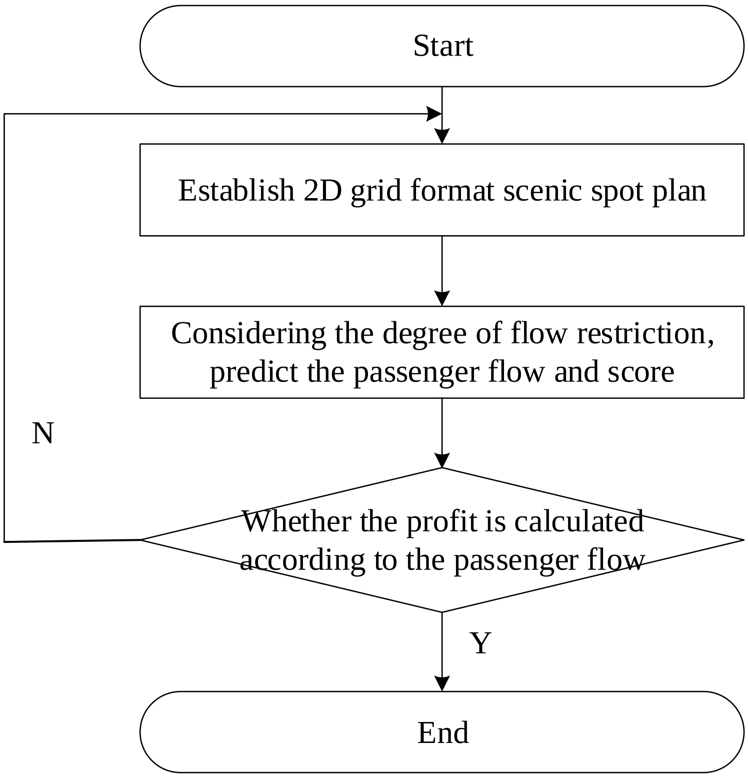

Figure 4-4-1 flow chart of comprehensive thinking of question IV

4.4.2 Establishment of model

① Cell side length

Cell is the basic unit of cellular automata. Each cell has the function of memory and storage, and all cells in the cellular automata system are constantly updated according to the dynamic rules. Generally speaking, in the simulation process, the smaller the cell size, the more details of diffusion can be displayed. However, the cell size is too small, which will increase the calculation time and the amount of data. The purpose of this model is to quickly and dynamically simulate the business situation of scenic spots under epidemic situation. Therefore, the appropriate cell size is selected according to the actual situation.

② Cell space

Cell space is divided into one-dimensional, two-dimensional and three-dimensional cell space, which is the spatial grid of cell distribution. The simulated business situation of the scenic spot under the epidemic situation corresponds to a two-dimensional cellular space. Cell and cell space are used to represent the static components of the system, so only these two static components can not support the normal operation of cellular automata. Therefore, it is necessary to introduce the evolution rules of cellular automata, the dynamic component of the scenic spot current limiting system.

③ Neighborhood

The evolution rule of cellular automata is defined in the local range. The state of a cell at（t+1）time depends on the state of the cell itself at time t and the state of its neighborhood cell.

Generally, two-dimensional cellular automata contains three neighborhoods. The two-dimensional grid diagrams of three neighborhoods are shown in Figure 4-4-2 to 4-2-4 below. The first is the von Neumann neighborhood, which is composed of a central cell and four neighbors in the southeast, northwest and northwest. The second is Moore neighborhood, which also includes four directions: Northeast, northwest, Southeast and southwest. The third is to extend Moore neighborhood.

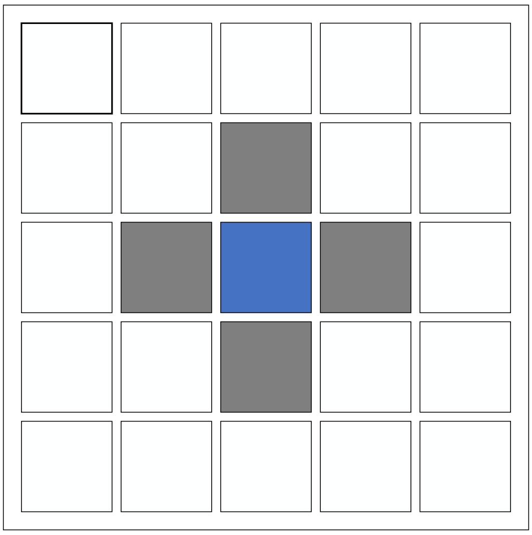

Figure 4-4-2 von Neumann neighborhood

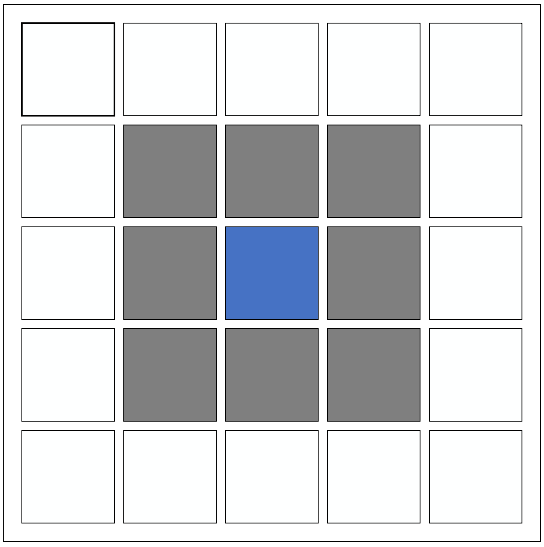

Figure 4-4-3 Moore neighborhood

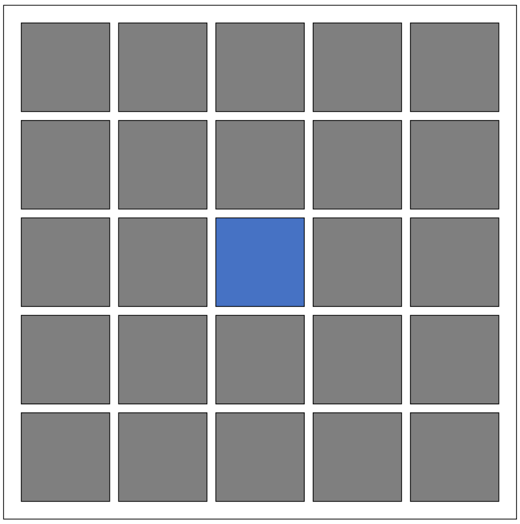

Figure 4-4-4 extended Moore neighborhood

In order to quickly simulate the flow restriction of the scenic spot during operation under the epidemic situation, the first von Neumann neighborhood is selected for simulation.

④ Cell morphology

In the business flow restriction model of scenic spots under simulated epidemic based on cellular automata, the cellular state s defines three values: 0, 1 and 2. When there is no epidemic situation and flow restriction is not required in the scenic area, the cell is defined as 2; When an epidemic situation occurs and a reasonable flow restriction plan is implemented, the cell is defined as 1; At the end of the epidemic, the business situation of the scenic spot gradually recovered, the flow restriction degree gradually decreased, and the cell is defined as 0.

4.4.3 Solution of model

Based on the above analysis, the solving steps of the problem are described as follows:

Step 1 Initializing the grid matrix to store the cell population as a cell space;

Step 2 Initialize the cell population occupying a certain proportion of the total number of grids, which is used to express the state when the epidemic situation with state 1 occurs and a reasonable current limiting plan is implemented;

Step3 Draw the cell group on the graphical interface and mark the passenger flow of the scenic spot under normal operation with red status 2; Green mark the passenger flow of reasonable flow restriction plan in case of epidemic situation with status 1; Black indicates the passenger flow of the scenic spot whose business condition gradually recovers at the end of the epidemic with the status of 0, and the flow restriction degree gradually decreases.

Step 4 obtains the next time state of the cell according to the rules;

Step 5 Calculate the total number of 5A scenic spots in all provinces in China. It is assumed that the passenger flow of a reasonable flow restriction plan in the event of an epidemic with state 1 is implemented in advance, and on this premise, the cell state of a scenic spot is calculated and obtained;

Step 6 Redraw the cell population to the screen.

Step 7 Each process is defined as a function, which is executed circularly for 4-6, and can sleep for 1.5s in the middle, so as to observe the change process of the results.

Cellular automata carries out the interaction between local cells according to the rules, so as to cause the global change of the system. The evolution rules dominate the whole dynamic behavior of cellular automata. Therefore, the next cell state is determined according to the state of the target cell determined by the model and its critical domain state. The simulation of the model is realized by MATLAB programming of cellular automata. Combined with the above, it can be understood that cellular automata simulation needs to understand three points. One is cell. In MATLAB, it can be understood as a square block composed of one or more points in the matrix, and a point in the matrix represents a cell. The second is the change rule, which determines the state of the cell at the next moment. The third is the cell state. The cell state is user-defined and usually opposite. In this model, the cell state is the current limit and normal business conditions of the scenic spot. The following dynamic search process diagrams are obtained, as shown in figures 4-4-6 to 4-4-9:

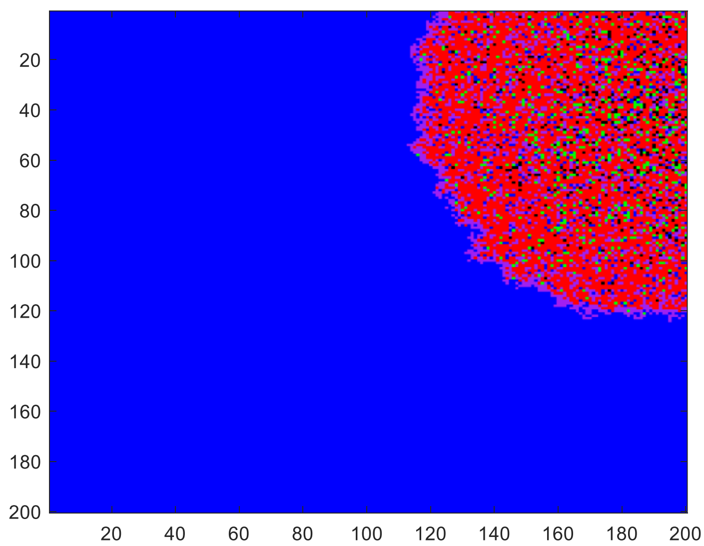

Figure 4-4-6 cellular automata simulation I

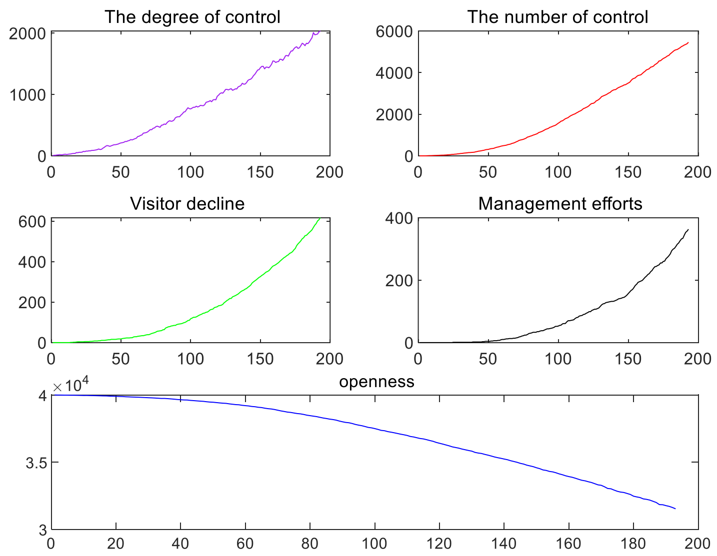

Figure 4-4-7 change trend of various factors I

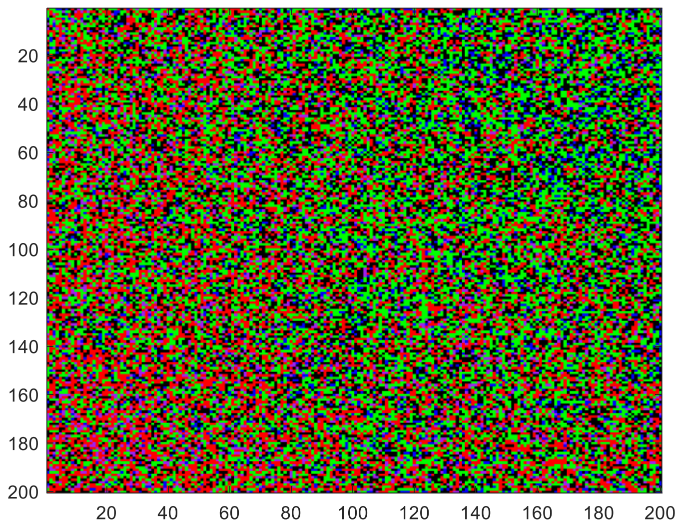

Figure 4-4-8 cellular automata simulation II

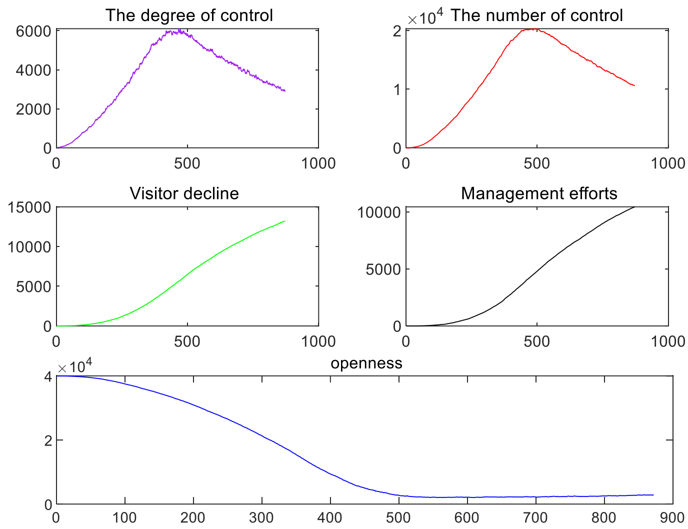

Figure 4-4-9 change trend of various factors I

It is concluded that the model is a process of searching the root cause according to the conditions in the termination state. That is, the flow limiting degree of the known termination conditions reaches the maximum value, and the mathematical model considering not only the safety factor of tourists, but also the passenger flow and the income of the scenic spot is calculated through the influence of the monitoring number, distance and environmental factors. The starting conditions are determined by the above conditions.

In the quantitative scenic spot management model, strictly control the tourist flow, scientifically and reasonably set the carrying capacity of the scenic spot, prevent personnel from gathering, and fully consider the notification of epidemic prevention requirements to ensure that the scenic spot is profitable. To give tourists a better experience, the three restrictions provided in this model are epidemic prevention standard, tourism income and tourism experience. The epidemic level, maximum carrying capacity of tourists and current evaluation level of different tourist attractions are different. Therefore, if the cellular automata model is used to simulate and analyze the data, an evolution rule needs to be added, This rule needs to realize the dynamic change of two-dimensional cell space. Therefore, we can choose to add constraints to further simulate the cellular automata model.

5. Differentiated management scheme

Recently, a large number of tourists gathered and crowded in some scenic spots during the gradual resumption of opening up, increasing the risk of epidemic spread. In order to ensure the orderly operation of scenic spots, effectively prevent, timely control and properly dispose of the new coronal pneumonia virus, strictly standardize the management of scenic spots after the lifting of the ban, and ensure the safe and orderly opening of scenic spots, the following differentiated management schemes for different scenic spots are formulated:

1 Adhere to prevention and control first and implement limited opening-up.

On the premise of doing a good job in the epidemic prevention and control of scenic spots, the scenic spots adhere to the principle of zoning and grading, strictly implement the requirements of the guide for epidemic prevention and control measures for the restoration and opening of scenic spots, achieve limited and orderly opening, and strictly prevent disorderly opening. During the epidemic prevention and control period, indoor scenic spots should be closed, and the number of outdoor scenic spots should not exceed 30% of the approved maximum carrying capacity. Before implementing the temporary preferential policies, the charging scenic spots shall carefully evaluate to prevent the passenger flow from exceeding the limit.

2 Strengthen flow management and strictly prevent personnel from gathering.

The free outdoor scenic spots shall establish and improve the reservation system, and promote the tour reservation by time through various channels such as instant messaging tools, mobile client, scenic spot official website and telephone reservation, so as to guide tourists to enter the park at intervals and travel off peak. Strictly limit the number of on-site tourists.

The outdoor area of the charge should do well in the registration of tourist information. Visitors who have made an appointment or on-site ticket and tickets should visit the tourist area to provide the ID number and contact information. It is necessary to do a good job in the registration of tourist information. Tourists who have made an appointment or on-site tickets and tickets should visit the tourist information and provide necessary information such as body ID number, contact information and so on. The identity information should be used according to the law. Avoid leakage. Travel agencies and tourism passenger transport operators shall strictly implement the requirements of relevant prevention and control guidelines and properly control the passenger capacity of tourism vehicles. In the conditional area, we should give full play to the role of the local Internet plus tourism service platform, and adopt various new technologies, such as big data analysis, to promote intelligent tourism, scientific diversion, and guide tourists, so that we can make the traffic management of tourist attractions front and strictly control the flow of passengers.

Scenic spots should fully understand the positive significance of flow restriction and avoid being influenced by economic benefits. It is important to restart the economy, but we still need to put tourist safety first. The flow restriction measures in the scenic spot effectively protect all kinds of tourism resources, equipment and facilities in the scenic spot, which is conducive to the rational operation and management of the scenic spot. At the same time, the dense population is more likely to cause the rapid spread of the epidemic.

Taking the Potala Palace scenic spot in Tibet as an example, online booking is limited to the afternoon, which fully reflects the advantage of time flow. When tourists buy tickets, the computer will automatically allocate a visit period. Tourists can pass through the gate with their ID card without changing tickets during this period. If they miss this period, the tickets will be automatically invalidated. Share the passenger flow per unit time, avoid the concentration of passenger flow and reduce the damage to the scenic spot.

Whether it is a restricted or unrestricted scenic spot, there may be bottleneck scenic spots in the scenic spot – bottleneck scenic spots need to adopt relevant facilities and management personnel to ensure that the total number of tourists does not exceed the instantaneous tourist reception of the scenic spot. Measures that can be considered include one-way traffic, nine curved railings, etc. the number of flow restrictions should not be considered according to the maximum number of flow restrictions in the scenic spot, but should be calculated according to the fact that the entrance is not too crowded. Taking Huangshan as an example, 20000 people are scattered in the whole scenic spot, but it is a great potential safety hazard to appear at the same time.

3 Refine management measures and standardize tour order.

Scenic spots should be equipped with necessary personnel and equipment, strengthen cleaning and disinfection, strictly implement prevention and control measures such as body temperature screening, and use “Health Code” verification and other means in combination with reality. If suspicious persons are found, they shall be dissuaded from entering, temporarily isolated, and immediately notified to the local health department for timely disposal. We should optimize the setting of tourist routes to prevent tourists from getting together and crowding due to unreasonable route planning. We should strengthen patrol inspection, guide tourists to do a good job in safety protection, and maintain the personnel spacing in places such as ticket purchase, sightseeing, rest and catering. We should strengthen the investigation and treatment of potential safety hazards of various tourism equipment and fire-fighting equipment, and strengthen the control of field fire sources. Special personnel shall be arranged at the entrances and exits of scenic spots, important visiting points and other areas where people are easy to gather to strengthen dredging, avoid congestion and ensure that prevention and control are in place.

4 Do a good job in publicity and guidance and advocate civilized tourism.

The scenic spot shall release the management measures for the restoration and opening of the scenic spot, epidemic prevention and control guidelines, forest fire prevention knowledge and disaster weather early warning information through the official website, third-party platform, prompt board, radio, electronic display screen, etc., guide tourists to abide by the safety warning regulations in Tourism activities, and help tourists enhance their protection awareness and master protection knowledge, Guide tourists to consciously wear masks, abide by public order, actively cooperate with prevention and control, avoid congestion and ensure that prevention and control are in place. Do a good job in publicity and guidance and advocate civilized tourism.5、 Strengthen organizational leadership and implement the division of responsibilities.

Scenic spots should take the main responsibility for the opening of scenic spots, establish a supervision mechanism, send a supervision team to key scenic spots when necessary, and earnestly strengthen the opening of scenic spots and tourism safety inspection. We should guide scenic spots to improve the emergency mechanism, classify and improve the emergency plan, clarify the emergency measures and disposal process for epidemic prevention and control, implement the detailed prevention and control responsibilities, and strictly prevent all kinds of safety accidents caused by the resumption of operation. The scenic spot shall improve the Department linkage mechanism, improve the emergency response capacity, and deal with emergencies timely and properly to ensure that all measures are implemented in place.

6. Strengths and Weakness

6.1 Strengths of the model

- Quantitative analysis of many factors makes the results more convincing.
- The principle of K-means clustering is relatively simple, easy to implement and fast convergence; When the result clusters are dense and the difference between clusters is obvious, its effect is better.
- NSGA - Ⅱ algorithm is one of the most popular multi-objective genetic algorithms. It reduces the complexity of non inferior sorting genetic algorithm. It has the advantages of fast running speed and good convergence of solution set. It has become the benchmark of the performance of other multi-objective optimization algorithms.
- NSGA - Ⅱ algorithm uses the grade selection method to highlight good individuals, and uses congestion calculation instead of formulating shared parameters to estimate the dispersion.
- All cells in the cellular automata system are constantly updated according to the dynamic rules, which can quickly and dynamically simulate the target situation; The cell size can be customized according to the actual situation, which can display the details well.

6.2 Weakness of the model

- When collecting data, there will be slight errors, and the of data collection may be lost.
- When the sample size of cluster analysis is large, there are errors in the clustering conclusion; If the classification number is unreasonable, it needs to be recalculated. When using NSGA - Ⅱ algorithm, the units of targets with different properties are inconsistent, which is not easy to compare; The distribution of each objective weighting value is subjective; The optimization objective is only the weighted sum of each objective, and the optimization progress of each objective in the optimization process is not operable.

- In the process of building cellular automata model, many factors are simplified and ignored, which will be different from the actual situation.

6.3 Generalization of model

Cellular automata uses discrete spatial layout and discrete time interval to divide the cell into finite states. The evolution of the individual state of the cell is only related to its current state and the state of a local neighborhood. Cellular automata studies the macro behavior and law of complex system composed of a large number of parallel units similar to biological cells by means of computer modeling and simulation. L-system, finite condensation diffusion, lattice gas automata, lattice Boltzmann method and traffic flow model are the concretization of cellular automata, which have important theoretical significance and practical application value.

7. Conclusion

The spatial distribution of the maximum carrying capacity of 5A scenic spots in China is uneven, and the spatial difference characteristics are consistent with the law of population distribution. The daily maximum carrying capacity of 5A scenic spot shows a trend of "higher in the East - moderate in the middle - lower in the west", reflecting the differences in resource allocation, population, economic level and tourism supply capacity in different regions. Specifically, 5A scenic spots in the eastern region account for the highest proportion of the maximum carrying capacity, and have strong tourism supply capacity and great tourism attraction; The maximum carrying capacity of 5A scenic spots in the central and western regions is relatively low, and there is great development potential. The reasons for the inter provincial and type spatial differences in the maximum carrying capacity of 5A scenic spots are diverse and complex. Tourism resource endowment, economic development level, regional population and tourism development policies can not be ignored.

Tourism carrying capacity indirectly reflects "tourism supply capacity" and "tourism security level". Tourism supply mainly comes from tourism resources and regional climate. The world's famous tourist attractions are all suitable for climate or strange landscape. At present, China is in the stage of popular tourism. Domestic tourism demand is awesome, and the supply capability evaluation of provinces and cities or various tourism resources (scenic spots) is lagging behind. In the case of tourism supply capacity (climate suitability and landscape attraction), tourism carrying capacity can not be regarded as a reference index.

References

[1] Zhenpeng Zhang. Study on spatial distribution characteristics of Chinese scenic spots based on GIS [D]. Kunming University of Technology, 2019.
[2] Xinchun Zhu. Research on domestic tourism demand forecasting based on Fuzzy

Theory [D]. Hunan University of Technology, 2015.
[3] Yaolong Liu, Jing Liu, Jin Duan, Huaming Zhang. Spatial differentiation of maximum carrying capacity of 5A scenic spots in China [J]. Journal of College of disaster prevention and technology, 2017, 19(03): 74-80.
[4] Jian Zhao. Analysis on the spatial pattern and influencing factors of the coordination between China's tourist traffic and network attention [D]. Zhejiang University, 2020.
[5] Xiaoqing Zhao, Yufeng Jia, Pengwei Zhang, Juan Ding. Data analysis based on K-means clustering algorithm [J]. Scientific and technological innovation,2021(23): 85-86.
[6] Xiaoqing Chen, Zhongxi Hou, Liangmin Guo, Wencai Luo. Improved multi-objective genetic algorithm based on NSGA-II [J]. Computer application, 2006(10): 2453-2456.
[7] Lijuan Zhang, Fuchang Wang, Lizhen Gang. Virus propagation and system simulation model based on Cellular Automata [J]. Journal of system simulation, 2021, 33(10): 2449-2459.
[8] Yongsheng Mu, HuaMing Xiang, Qingqing Xiong. Analysis and suggestions on the current situation of flow restriction in domestic scenic spots during the epidemic -- Taking Huangshan scenic spot as an example [J]. Modern trade industry, 2020, 41(28): 37-38.
[9] Jigang Tang. Necessity and Implementation Suggestions of normalization of reserved flow restriction in scenic spots [N]. China Tourism News，2020-06-10(003).

Appendix

Appendix list:

Appendix 1	Text mining code
Appendix 2	Distribution map code
Appendix 3	K-means clustering code
Appendix 4	Cellular simulation code

Appendix 1：Text mining code

```
%Text mining word clouds
clc;
clear;
report=fileread('gugong.txt');
% Fill in filename to import text. Pay attention to modification when using.
report=regexprep(report,'[.,?()!’”“:]','');
report = erasePunctuation(report);
% Remove the punctuation
report=lower(report);
% Convert to lowercase
```

```
documents = tokenizedDocument([report]);
% Creates an array of tagged documents for text analysis
newDocuments = removeStopWords(documents)
% Remove a list of stop words
newDocuments=removeLongWords(newDocuments,15)
% Remove words with 15 or more characters
newDocuments=removeShortWords(newDocuments,3)
% Remove words with 3 or fewer characters
mystopwords =
["palace","gate","take","door","five","beijing","really","museum","city","first","second","hall","la
ke",""];
%Custom No words
newDocuments = removeWords(newDocuments,mystopwords);
cleanedBag=bagOfWords(newDocuments)
%Create a bag-of-words model.
cleanedBag = removeInfrequentWords(cleanedBag,2)
% Remove words that do not appear more than two times in the bag-of-words model.
rawDocuments = tokenizedDocument(documents);
% Create the original tokenized document without word filtering
rawBag = bagOfWords(rawDocuments)
% Create a bag-of-words model.
numWordsCleaned = cleanedBag.NumWords;
numWordsRaw = rawBag.NumWords;
reduction = 1 - numWordsCleaned/numWordsRaw
%Diagram the cloud
figure
subplot(1,2,1)
wordcloud(rawBag);
title("Raw Data")
subplot(1,2,2)
wordcloud(cleanedBag);
title("Cleaned Data")
% Compare the raw data and the cleaned data by visualizing the two bag-of-words models using
word clouds.
%customStopWords = [stopWords "thy" "thee" "thou" "dost" "doth"];
```

Appendix 2：Distribution map code

```
clear all;close all;
clc
geoshow('China.shp');
hold on;
x1=xlsread('Attachment.xlsx','Sheet1','D2:D203')
y1=xlsread('Attachment.xlsx','Sheet1','E2:E203')
```

```
scatter(x1, y1, '*');
xlabel('longitude');ylabel('dimensionality')
legend('National 5A Scenic Spot')
```

Appendix 3: K-means clustering code

```
clear all;close all;
clc
%import data
Dailycapacity = xlsread('Attachment.xlsx','Sheet1','F2:F205');
assessment = xlsread('Attachment.xlsx','Sheet1','H2:H205');
for i=1:204;
Dailycapacity(i,1) = (Dailycapacity(i,1)-min(Dailycapacity))/(max(Dailycapacity)-
min(Dailycapacity))
end
for i=1:204;
assessment(i,1) = (assessment(i,1)-min(assessment))/(max(assessment)-min(assessment))
end
bearingcapacity = 10.*Dailycapacity
bearingcapacity2 = 10.*assessment
dataset = [bearingcapacity,bearingcapacity2]
%Display raw data
plot(dataset(:,1),dataset(:,2),'b+');
%grid on;
data=dataset;
%Set the number of clusters
N=3;
[m,n]=size(data);
temdata=zeros(m,n+1);
typecenter=zeros(N,n);%Initialize the cluster center
temdata(:,1:n)=data(:,:);
%Generate clustering centers
for x=1:N
    typecenter(x,:)=data( randi(90,1),:);
end
while 1
distence=zeros(1,N);
num=zeros(1,N);
new_typecenter=zeros(N,n);

for x=1:m
    for y=1:N
    distence(y)=norm(data(x,:)-typecenter(y,:));%Calculation of distance
```

```
        end
        %The minimum distance
        [~, temp]=min(distence);
        temdata(x,n+1)=temp;
end
k=0;
for y=1:N
        for x=1:m
                if temdata(x,n+1)==y
                      new_typecenter(y,:)=new_typecenter(y,:)+temdata(x,1:n);
                      num(y)=num(y)+1;
                end
        end
        new_typecenter(y,:)=new_typecenter(y,:)/num(y);
        if norm(new_typecenter(y,:)-typecenter(y,:))<0.1
                k=k+1;
        end
end
if k==N
          break;
else
          typecenter=new_typecenter;
end
end
[m, n]=size(temdata);

figure;
hold on;

%Draw the graph after the final clustering
for i=1:m
        if temdata(i,n)==1
                  plot(temdata(i,1),temdata(i,2),'r*');
                  plot(typecenter(1,1),typecenter(1,2),'ko');
        elseif temdata(i,n)==2
                  plot(temdata(i,1),temdata(i,2),'g*');
                  plot(typecenter(2,1),typecenter(2,2),'ko');
        elseif temdata(i,n)==3
                  plot(temdata(i,1),temdata(i,2),'b*');
                  plot(typecenter(3,1),typecenter(3,2),'ko');
        elseif temdata(i,n)==4
                  plot(temdata(i,1),temdata(i,2),'y*');
                  plot(typecenter(4,1),typecenter(4,2),'ko');
        else
```

```
        plot(temdata(i,1),temdata(i,2),'m*');
        plot(typecenter(4,1),typecenter(4,2),'ko');
    end
end
%grid on;
```

Appendix 4 :Cellular simulation code

```
close;clear;clc
%Parameter initialization
n = 250;
inlatent = 1;
%The neighbor mode is 8 neighbor
Pro01 = 0.1;Pro02 = 0.2;Pro10 = 0.005;
Pro24 = 0.001; Pro20 = 0.04;Pro_10 = 0.06;
k = 1; %Degree of coefficient of variation
rand('state',sum(clock));
percentage = 5*ones(n+2);
percentage(2:n+1,2:n+1) = 0;
while inlatent ~= sum(sum(percentage(2:n+1,2:n+1)))
    percentage(2+floor(n*rand),2+floor(n*rand)) = 1;
end
%The time series
T = zeros(n+2);

for i = 1:10000
    U   = 5*ones(n+2);
    nqf = zeros(n+2);

    for r = 2:n+1
        for l = 2:n+1
 %change
 %diversification
            if percentage(r,l) == 0
                U(r,l) = 0;
            elseif percentage(r,l) == 1
                temp = exp(-k*(12 - T(r,l))/T(r,l));
                if temp > rand
                    U(r,l) = 2;
                    T(r,l) = 0;
                elseif rand < Pro10
                    U(r,l) = 3;
                    T(r,l) = 0;
                elseif rand <   Pro_10
```

```
            U(r,l) = 0;
            T(r,l) = 0;
        else
            U(r,l) = 1;
        end
%Neighbor influence
        if percentage(r-1,l-1) == 0
            if rand < Pro01
                nqf(r-1,l-1) = 1;
            end
        end
        if percentage(r-1,l) == 0
            if rand < Pro01
                nqf(r-1,l) = 1;
            end
        end
        if percentage(r-1,l+1) == 0
            if rand < Pro01
                ro_qianfuzhe(r-1,l+1) = 1;
            end
        end
        if percentage(r,l+1) == 0
            if rand < Pro01
                nqf(r,l+1) = 1;
            end
        end
        if percentage(r+1,l+1) == 0
            if rand < Pro01
                nqf(r+1,l+1) = 1;
            end
        end
        if percentage(r+1,l) == 0
            if rand < Pro01
                nqf(r+1,l) = 1;
            end
        end
        if percentage(r+1,l-1) == 0
            if rand < Pro01
                nqf(r+1,l-1) = 1;
            end
        end
        if percentage(r,l-1) == 0
            if rand < Pro01
                nqf(r,l-1) = 1;
```

```
                end
            end

        elseif percentage(r,l) == 2
            if Pro24 > rand
                U(r,l) = 4;
            elseif Pro20 > rand
                U(r,l) = 0;
            else
                U(r,l) = 2;
            end
            if percentage(r-1,l-1) == 0
                if rand < Pro02
                    nqf(r-1,l-1) = 1;
                end
            end
            if percentage(r-1,l) == 0
                if rand < Pro02
                    nqf(r-1,l) = 1;
                end
            end
            if percentage(r-1,l+1) == 0
                if rand < Pro02
                    ro_qianfuzhe(r-1,l+1) = 1;
                end
            end
            if percentage(r,l+1) == 0
                if rand < Pro02
                    nqf(r,l+1) = 1;
                end
            end
            if percentage(r+1,l+1) == 0
                if rand < Pro02
                    nqf(r+1,l+1) = 1;
                end
            end
            if percentage(r+1,l) == 0
                if rand < Pro02
                    nqf(r+1,l) = 1;
                end
            end
            if percentage(r+1,l-1) == 0
                if rand < Pro02
                    nqf(r+1,l-1) = 1;
```

```
                    end
                end
                if percentage(r,l-1) == 0
                    if rand < Pro02
                        nqf(r,l-1) = 1;
                    end
                end
            elseif percentage(r,l) == 3
                U(r,l) = 3;
            elseif percentage(r,l) == 4
                U(r,l) = 4;
            end
        end
    end
%matrix
    U(percentage == 0 & nqf) =    1;
%Change related data
    T(find(T ~= 0)) = T(find(T ~= 0)) + 1;
%Not last time. The timing of this change
    T(percentage == 0 & nqf) = 1;
%Update matrix
    percentage = U;

%Output of results
    R = zeros(n);
    P = zeros(n);
    G = zeros(n);
    yganz(i) = sum(sum(U == 0));
    qfuz(i) = sum(sum(U == 1));
    huanb(i) = sum(sum(U == 2));
    miy(i) = sum(sum(U == 3));
    nsixwan(i) = sum(sum(U    == 4));
    for r = 1:n
        for l = 1:n
            if U(r+1,l+1) == 0
                G(r,l) = 1;
            elseif U(r+1,l+1) == 1
                R(r,l) = 160/255;
                P(r,l) = 32/255;
                G(r,l) = 240/255;
            elseif U(r+1,l+1) == 2
                R(r,l) = 1;
            elseif U(r+1,l+1) == 3
                P(r,l) = 1;
```

```
            end
        end
      end
   figure(1);
   image(cat(3,R,P,G));
   drawnow
   figure(2);
   subplot(321),temp = plot(qfuz,'-');set(temp,'color',[160/255 32/255 240/255]),
   title(['Time: ',num2str(i),' The degree of control']);
   subplot(322),plot(huanb,'r-'),title(['Time: ',num2str(i),'The number of control']);
   subplot(323),plot(miy,'g-'),title(['Time: ',num2str(i),'Visitor decline ']);
   subplot(324),plot(nsixwan,'k-'),title(['Time: ',num2str(i),' Management efforts']);
   subplot(313),plot(yganz,'b-'),title(['Time: ',num2str(i),' openness']);
   pause(0.001);
end
```