HIGH-SPEED RAILWAY BASED ON GENETIC ALGORITHM FOR PREDICTION OF TRAVEL CHOICE

Long Chen-xu, Li Jing and Gao Yue
School of Economics and Management, Information Management and Information System, Beijing Jiaotong University
Yuen Estate, No. 3, 100044, Haidian District, Beijing, China

Keywords: Genetic algorithm, High-speed railway, Forecast analysis, Modal split.

Abstract: Genetic algorithm is a new optimizing searching method based on biology evolutionary theory. Just as evolution deals in populations of individuals, genetic algorithms mimic nature by evolving huge churning populations of code, all processing and mutating at once. With the frequency of passenger travel speeding up and passenger's demand to the quality of life higher and higher, passengers have higher and higher demands to the travel. Especially in environment, comfort and quick aspect, different passenger focuses on different aspects, therefore, before the research, we must classify the passenger. This paper applies British Sheffield university GA toolbox, with the application of matlab, finally makes a forecast analysis. Because the forecast analysis is based on the questionnaire during the Spring Festival, so the emphasize on the travel choice made by the passengers in this certain circumstances is necessary. In addition, the forecast analysis will be more or less different with other forecast analysis normally.

1 INTRODUCTION

Genetic algorithm (GA) is a random search method. It decreases the effect of original values greatly through crossover and mutation operations, and it can easily find out the global optimal results.

WU Qun-qi and XU Xing (2007), thought that passengers choose transportation mainly rely on time, economy and feeling these three factors. The constituent of the value of trip time of the specific travel subject has something to do with the travel interests: full economic relevance, totally not related to the economy, some economic-related. According to the value of travel time and the correlation of travel interests, it suggests the mechanism of passengers' travel choice. Simple genetic algorithm uses binary encoding, this approach is simple, easy to implement crossover and mutation operations, in line with the principle of the minimum character set encoding. As we all know, the simple genetic algorithm is poor in the local search capabilities and has a slower speed in optimizing the global, therefore, it is necessary to establish the crossover operator which can simultaneously search both in feasible and infeasible solution space, the mutation operator which has a capability of fast search in the earlier stage and the capability of maintaining the optimal solution in the later stage, and the selection operator which has the capability of maintaining the "elite". MA Yong-jie, MA Yi-de, JIANG Zhao-yuan, SUN Qi-guo (2009) propose to take use of the solutions which has been searched for to avoid the "throwback" of the offspring and the degradation is obligatory. GONG Gu, ZHAO Xiang-jun, HAO Guo-sheng, CHEN Long-gao (2009) propose to divide the search space, we can inherit the excellent allele to the next generation. While using the taboo domain and the active domain can quickly improve the algorithm for performance.

The core problem of this paper is to use the collected initial data from the questionnaire based on high-speed railway passenger travel, the influencing factors theory related to high-speed railway passengers travel choice, and the knowledge about genetic algorithm to analyse the factors influencing passengers choosing transportation and make predictions. The objective function is maximizing the passengers' travel utility. However, different types of passengers have different preferences and different emphasis on economy, time and feeling. Therefore, each type of passenger takes different aspects into consideration and makes travel choice. Obviously, it is very important to establish a

Chen-xu L., Jing L. and Yue G..
HIGH-SPEED RAILWAY BASED ON GENETIC ALGORITHM FOR PREDICTION OF TRAVEL CHOICE.
DOI: 10.5220/0003433400260031
In *Proceedings of the 13th International Conference on Enterprise Information Systems* (ICEIS-2011), pages 26-31
ISBN: 978-989-8425-54-6

simulation model.

This paper used Matlab to solve the genetic algorithm problem, with binary encoding, and made predictions for passengers' travel choice.

2 CONCEPTUAL FRAMEWORK

Passengers choose transportation mainly rely on time, economy and feeling these three factors. Dividing these three dimensions, we classify the factors influencing passengers making travel choices.

Passengers make travel choices will be influenced by some objective, potential factors. Through analysing the passengers' characteristics and different factors' effects on passengers, we got a flow chart of passenger travel.

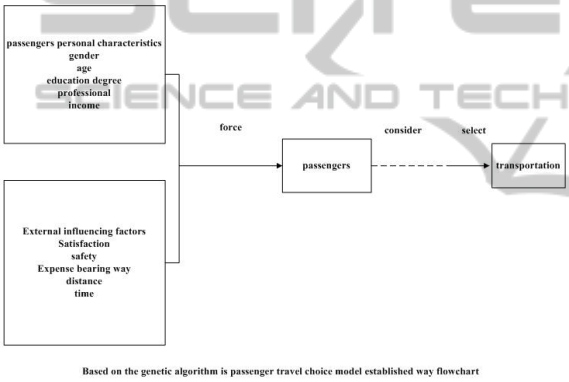

Figure 1: Flowchart passenger travel mode choice.

Passenger travel is based on different purposes, such as going home, for business, for tourism, so passengers based on different trip purposes may make different choices in travel mode. In addition, the economic capacity of travelers and who will bear the cost of travel will have a great impact on travel mode choice.

2.1 The Division of Passenger Types

2.1.1 According to the Travel Purpose

According to passengers' respective trip purpose, all passenger traffic in the channel can be divided into:

①business; ②tourism; ③work; ④school; ⑤home; ⑥interchange; ⑦others.

Among them, passengers for business always focus on convenience and comfort, but have low sensitivity of the cost; passengers for tourism often focus on comfort and fare levels, and have high sensitivity of the cost; work, school and other commuter passenger traffic often takes fare for the primary consideration, and has certain requirements on punctuality; those for home always have low requirements for comfort and higher sensitivity of the costs, and take fare levels for main consideration.

2.1.2 According to the Cost Mode

According to who will bear the cost, passenger traffic can be divide into:

①travel at public expense ②travel at their own expense.

Travel at public expense means that passenger does not need to pay for the travel because of the social production and the need to work, and travel costs are included in the cost of social production. Travel at their own expense is to meet passengers' own needs and travel costs are included in private consumption. Because of the existence of these two different ways of bearing the cost of travel, there will be some differences when the two parts of passenger traffic choose the travel mode.

Because the travel time is included in the cost of production, passenger traffic at public expense will pay more attention to time and look for convenient, fast and punctuality while they are choosing the travel mode. So they have higher selection bias of civil aviation and high-speed railway. Moreover, they have higher requirements about the frequency of the mode of transportation, departure time and arrival time. And they are less sensitive to the travel cost. However, passenger traffic at their own expense will have lower requirements for the quality of transport and higher sensitivity for the travel costs because they must pay for themselves.

2.1.3 According to Income

According to monthly income, the overall passenger traffic is divided into:

①lower-income travelers: 1000 yuan; ②low-income travelers: 1000 yuan to 3,000 yuan; ③middle-income travelers: 3000 yuan to 5000 Element; ④high-income travelers: 5000 yuan to 10,000 yuan; ⑤higher-income travelers: more than 10,000 yuan.

Due to a difference in income level, different types of passengers have different capacity of bearing the travel costs and have different sensitivity of costs. The previous data showed that middle-income and less income stream of passengers have a preference for the traditional existing rail or road.

And for the travel costs are the main considerations, they have a relatively low requirement for the transport quality such as comfort, convenience and punctuality. In addition, they have a relatively high degree of sensitivity for the fare level, so the fluctuations in fares will cause great changes in the passenger traffic distribution.

Passengers who have high and higher income will take comfort, convenience and punctuality into consideration because they have high abilities to pay, and they always select high quality transportation services of transportation, such as high-speed railway, civil aviation. What's more, these passengers are less sensitive to the cost, so a certain range of fluctuations in travel mode will have a little influence on them.

3 RESEARCH MODEL

The problem was described as: There were m travel modes and n batches of passengers(category) waiting to be distributed.

Before the target allocation, the key considerations of each batch of the target and each travel mode's weight on each target has been evaluated and sorted. J-approved visitors' "travel value" is w_j, i-approved travel mode's weight on j-approved target is p_{ij}, each travel mode's "trial" benefit value on each target is $u_{ij} = w_j * p_{ij}$. Among them, u_{ij} stands for each batch of the passenger's size of the degree of the effectiveness of the "trial". The purpose is to meet the basic principles of the target allocation and pursuit of the overall effectiveness of the best, that is seeking $\max(\sum_{j=1}^{n} u_{ij})$.

4 RESEARCH METHODS

This paper used binary encoding and the number of individuals was 40. In addition, the max number of generations was 50 and the generation gap was 0.9.

This paper used PN instead of passenger numbers, and TV instead of travel value.

Based on the numerical analysis of questionnaires, the standard value of the price dimension is 3.16, the standard value of the time dimension is 2.69 and the standard value of the environmental dimension is 4.47.

Choices for passengers to choose for travel:

①EMUs ②Direct train ③Coach ④Aircraft ⑤MICE.

4.1 According to the Purpose

Passengers for business always focus on convenience and comfort, but have low sensitivity of the cost, so the weight for the price dimension is 0,for the time dimension is 0.6 and for the environment is 0.4, that is, for the purpose of business, "travel value " is 0.6 * 2.69 +0.4 * 4.47 = 3.402. Passengers for tourism often focus on comfort and fare levels, and have high sensitivity of the cost, so the weight for the price dimension is 0.3 and for the environment is 0.7, that is for the purpose of tourism, "travel value " is 0.7 * 4.47 +0.3 * 3.16 = 4.077. Work, school and other commuter passenger traffic often takes fare for the primary consideration, and has certain requirements on punctuality, so the "travel value" for the passengers whose purpose is going to work is 0.6 * 2.69 +0.4 * 3.16 = 2.878, passengers to school's "travel value " is 0.2 * 2.69 +0.8 * 3.16 = 3.066. Passengers for home have low requirements for the comfort, often focus on the price level and have high sensitivity for the costs, so the "travel value" is 1 * 3.16 = 3.16. Passengers for transfer have high requirements for time, so the "travel value" is 1 * 2.69 = 2.69. Other passengers' weight are similar, that is, its "travel value" is 0.3 * 2.69 +0.3 * 4.47 +0.4 * 3.16 = 3.412.

Table 1: Travel Value - According to the Travel Purpose.

PN	1	2	3	4	5	6	7
TV（w_j）	3.402	4.077	2.878	3.066	3.16	2.69	3.412

4.2 According to the Cost Mode

Because the travel time is included in the cost of production, passenger traffic at public expense will pay more attention to time and look for convenient, fast and punctuality while they are choosing the travel mode. So they have higher selection bias of civil aviation and high-speed railway. Moreover, they have higher requirements about the frequency of the mode of transportation, departure time and arrival time. And they are less sensitive to the travel cost. So the weight for the time dimension is 0.42 and for the environment dimension is 0.58, that is, the "travel value" is 2.69 +4.47 * 0.42 * 0.58 = 3.7224. However, passenger traffic at their own expense will have lower requirements for the quality of transport and higher sensitivity for the travel costs

Table 2: Weights - According to the Travel Purpose.

i-j(p_{ij})	1	2	3	4	5
1	0.44	0.01	0.01	0.53	0.01
2	0.19	0.23	0.31	0.22	0.05
3	0.23	0.36	0.18	0.11	0.12
4	0.15	0.42	0.26	0.12	0.05
5	0.15	0.52	0.19	0.13	0.01
6	0.31	0.41	0.26	0.01	0.01
7	0.25	0.25	0.25	0.25	0.25

Table 3: Travel Value - According to the Cost Mode.

PN	1	2
TV（w_j）	3.7224	3.422

Table 4: Weights - According to the Cost Mode.

i-j(p_{ij})	1	2	3	4	5
1	0.28	0.39	0.05	0.27	0.01
2	0.12	0.43	0.11	0.17	0.17

because they must pay for themselves. So the "travel value" is 0.8 +0.2 * 3.16 * 4.47 = 3.422.

4.3 According to Income

The previous data showed that middle-income and less income stream of passengers have a preference for the traditional existing rail or road. And for the travel costs are the main considerations, they have a relatively low requirement for the transport quality such as comfort, convenience and punctuality. In addition, they have a relatively high degree of sensitivity for the fare level, so the fluctuations in fares will cause great changes in the passenger traffic distribution.So the "travel value" for the lower-income is 0.9 * 3.16 +0.09 2.69 +0.01 * 4.47 = 3.1308. And the "travel value" for the low-income passenger is 0.9 * 3.16 +0.082 * 2.69 +0.02 * 4.47 = 3.1486.

The "travel value" for the middle-income passenger is 0.5 * 3.16 +0.25 * 2.69 +0.25 * 4.47 = 3.37.

Passengers who have high and higher income will take comfort, convenience and punctuality into consideration because they have high abilities to pay, and they always select high quality transportation services of transportation, such as high-speed railway, civil aviation. What's more, these passengers are less sensitive to the cost, so a certain range of fluctuations in travel mode will have a little influence on them. So the "travel value" for those high-income is 0.21 * 3.16 +0.39 * 2.69 +0.4 * 4.47 = 3.5007, and the "travel value" for the higher-income is 0.15 * 3.16 +0.36 * 2.69 +0.49 * 4.47 = 3.6327.

Table 5: Travel Value - According to Income.

PN	1	2	3	4	5
TV（w_j）	3.1308	3.1486	3.37	3.5007	3.6327

Table 6: Weights - According to Income.

i-j(p_{ij})	1	2	3	4	5
1	0.01	0.37	0.11	0.01	0.01
2	0.01	0.57	0.31	0.02	0.01
3	0.01	0.03	0.41	0.05	0.01
4	0.68	0.02	0.07	0.42	0.43
5	0.28	0.01	0.09	0.50	0.54

5 FORECAST ANALYSIS

Firstly,because the background is under the spring festival, some predictions may not match exactly with the normal. During the spring festival, different passenger may have different considerations with the usual, such as the price, the time and the environmental dimension. Many people may take "as long as arrive the destination" into main consideration in the pessimistic circumstance, but not for the other factors.

5.1 According to the Purpose

According to Table 7, business travelers will choose plane for the next trip,passengers to tourism and work will choose EMUs for the next trip,passengers to school will choose long-distance bus and passengers to home and transfer will choose direct trains for the next trips. While the others will choose EMUs.

Because the passengers for tourism have higher requirements for the time and environment, moreover, most of them are free trips, their sensitivity to the price level is relatively lower. The predicted result showed they would choose plane for their next trips, this is more objective. Passengers for home and transfer have higher requirements for

Table 7: Predictions - According to the Travel Purpose.

PN	1	2	3	4	5	6	7
Predictions	4	1	1	3	2	2	1

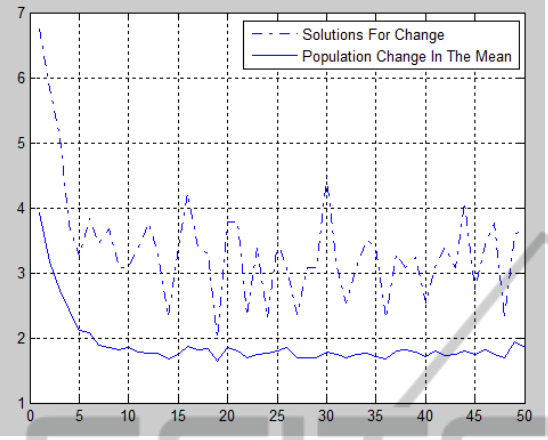

Figure 2: Predictions - According to the Travel Purpose.

time, so the predicted results are more objective.

Figure 2 is the application made by matlab. It is a change tracking map of the total effective value and the mean of the population.

5.2 According to the Cost Mode

According to table 8, passengers who travel at public expense will choose plane next trips and the others who travel at their own expense will take direct trains next trips.

Table 8: Predictions - According to the Cost Mode.

PN	1	2
Predictions	4	2

Passengers who travel at their own expense have low sensitivity for the price, because they do not pay the fees, they will take time or comfort into consideration. So the predictions are more objective. However, the train is less expensive compared with other mode of travel, the passengers who travel at their own expense choosing the direct train is more realistic.

Figure 3 is the application made by matlab. It is a change tracking map of the total effective value and the mean of the population.

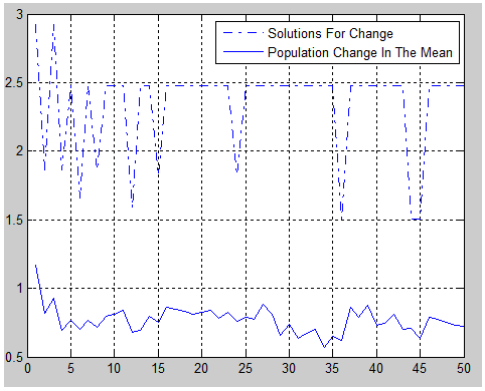

Figure 3: Predictions - According to the Cost Mode.

5.3 According to Income

According to Table 9, low and lower income passengers will take direct trains as their next trips, middle income passengers will choose long-distance bus, high-income passengers will choose EMU and higher-income passengers will choose planes.

Table 9: Predictions - According to Income.

PN	1	2	3	4	5
Predictions	2	2	3	1	4

In real life, most low-income passengers will choose direct trains as their travel mode for the price is lower, and this is very beneficial to those low-income passengers. However, middle-income passengers will be on a more balanced consideration of all aspects, so the predictions are more realistic. High and higher income passengers are less sensitivitive to the price, what they care about are time, comfort, so the predictions is also close to reality.

Figure 4 is the application made by matlab. It is a change tracking map of the total effective value and the mean of the population.

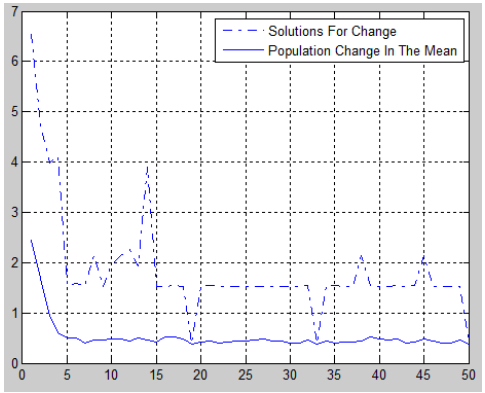

Figure 4: Predictions - According to Income.

6 RESTRICTIVE

This paper relies on the background of spring, so the passengers' choices would be some different with the usual. For the Spring Festival, the biggest feature is not very easy to purchase, sometimes one even cannot buy a well-content ticket though he/she has lined up row for a few days. Even worse, he/she may buy a standing ticket. Under such circumstances, the passenger will automatically consider less about the price, time and environment, at least he/she can reach his/her destination. So, one could reduce the sensitivity to price even he/she cares a lot about the price at ordinary times, one could reduce the requirements on time even he/she concerns a lot about the time in peacetime and one could ignore the comfort of the tool even he/she demands high at ordinary times. The data obtained from the questionnaire is limited, this will lead to different degrees of restriction of the final prediction results, that is, the universality of the prediction is restrictive.

7 CONCLUSIONS

The truth "survival of the fittest" has always been throughout the genetic algorithm, it is applied to solve the optimization problem,especially for fuzzy problems and has very good robustness. However, factors influence high-speed railway passengers choosing travel mode are complex and diverse, the most important and elusive is the passenger's personal preferences. Personal preferences of passengers are difficult to be precise analysis in any case because it involves personal changes in mental activity. When predicting the travel mode, we should first analyse which factors influence travelers to choose the travel mode, such price, time and environment are factors from the external, and passengers' personal characteristics are partly on behalf of their personal preference. The basic idea of the genetic algorithm to predict the travel selection methods is that the choice of traveling factors will act on the passengers, whether subjective or subconscious, when passengers choose their travel modes, they will maximize their benefits. Thus, making the fitness function value for the effectiveness of passenger travel, that is, the bigger the value of the fitness function, the greater the benefit of the passenger, and the more satisfied the passengers. So, under the guidance of this thought, we made predictions of various different types of passengers choosing travel modes.

Passengers choice of travel mode can show whether of such services industry is doing well, and the predictions at least can provide some good advice for those services sectors who have fewer passengers, such as what to improve and how to improve from price, time and environmental comfort, etc.

REFERENCES

Zhang Qiang, Wang Bin, Zhang Rui, Xu Chun-xia, 2008. Genetic Glgorithm-Based design for DNA sequences sets. *Chinese Journal of Computers,* 2008, 31 (12) :2193-2199.

He Wei-hui, Wang Jia-Lin, Hu Long-sheng, 2009. The improvement and application of real-coded multiple-population genetic algorithm. *Chinese Journal of Geophysics*, 2009,52 (10): 2644 ~ 2651.

Ma Yong-jie, Ma Yi-de, Jiang Zhao-yuan, Sun Qi-guo, 2009. Fast genetic algorithm and its convergence. *Systems Engineering and Electronics*. 2009,31 (3) :714-718.

Ding Jian-li, Chen Zeng-qiang, Yuan Zhu-Zhi, 2003. On the Combination of Genetic Algorithm and Ant Algorithm. *Journal of Computer Research and Development*. 2003,40 (9) :1351-1356.

Xu Yonghua, Huang Lijun, 2008. Improved Genetic Algorithm in milk transport vehicles Road King optimization. *Northeast Agricultural University* .2008,39 (11) :111-115.

Gong Gu, Zhao Xiang-jun, HAO Guo-sheng, Chen Long-gao, 2009. Study and implementation of Genetic Algorithm Based on improvment of search space partition. *Journal of Henan University Narural Science.* 2009,39 (6) :631-636.

Jiang Yun-cai, Wang Qing-bo, 2004. Application of MATLAB Genetic Algorithm Toolbox. *Journal of Jiangxi Electric Voactional and Technical College.* 2004,17 (3): 42-44.

Qiang Li-xia, Yan Ying, 2006. Analysis on the Difference of demand for passenger transport of high-speed railways between home and abroad. *Railway Transport and Economy*. 2006,28(9):18-21.

Wu Qun -qi, Xu Xing, 2007. Mechanism research on travelling choices of passengers. *Journal of Chang'an University(Social Science Edition)*. 2007,9(2) :13-16. The Dvelopment of the world's high-speed railway,2006. *Railway Survey and Design.* 2006,1: 54-56.

Chen Zhang-ming, JI Xiao-feng, 2008. Study on features of railway passengers' travel activities. *Railway Transport and Economy.* 2008,30(11): 23-25.