Simulation on Resident Individual Trip Choice Decision Making: Based on Modeling of Rough Set and Genetic Algorithms

Jing Li, Xin Zhu, Dan Chang

Simulation on Resident Individual Trip Choice Decision Making: Based on Modeling of Rough Set and Genetic Algorithms

Jing Li[1] Xin Zhu[2] Dan Chang[3]

1School of Economics and Management, Beijing Jiaotong University, Beijing, China 100038, jingli@bjtu.edu.cn
2School of Economics, Peking University, Beijing, China 100871, 86369771@163.com
3School of Economics and Management, Beijing Jiaotong University, Beijing, China 100038, dchang@bjtu.edu.cn

Abstract. Urban Residents are the evaluation subjects for the transportation policy. It is very important for the scientific development of transportation policy to research on the travel behavior from the perspective of quantitative modelling. Urban residents scientific of individual travel decisions are the prerequisites for group urban residents travel behaviors. This paper analyses the general population travel behavior based on the basis of the actual survey data and uses rough set theory to reduce the residents travel decision influencing factors. The univariate weight matrix and the corresponding weight matrix could be obtained. On this basis, the genetic algorithms objective function and fitness function could be optimized. Finally, an example simulation is given in this paper to verify the validity of the travel decision-making simulation model.

Keywords: Travel decision ·Rough Set · Genetic Algorithms · Modeling

1 Introduction

Citizens are the most important parts of urban transportation formations. The choice of citizen travel mode will be influenced by the individual competency and preference; meanwhile, it will be influenced by the transportation environment. The aim of the urban transportation policy is to allow the majority of the population could reach their destinations with a faster, safer, more comfortable and more affordable travel mode. Moreover, both the localized and the whole domain liquidity established by the citizens traveling activities should be taken into consideration. A valid transportation policy is determined by the social satisfaction, economic benefit and level of environmental protection. Scientific and rational analysis on various citizen travelling behaviors, such as choosing public transport or non-public transport will help to find the problems in the existing transport policy system and give fair recommendations.

At the present stage, research on resident individual trip mainly involves following fields: commuter and structure of urban space, character of commuter behavior basing on traveling investigation, character of trip chain and trip behavior modelling.

adfa, p. 1, 2011.

The researches on resident individual trip behavior theory are very fresh and mainly focus on several fields.

The first field is how to confirm factors that influent individual trip behavior. Iseki Hiroyuki and Smart Michael (2012) invested 2122 residents and analyzed the factors such as age, sex, income, race, religion, private car travel mode and frequency of public facilities using. They found that security is the most important factor for all the residents and individuals with different background will have different attitudes to different factors which will be changed with trip experience.

The second field is to develop the individual trip rules. Abou-Zeid Maya and Ben-Akiya Moshe (2011) proposed a framework basing on dynamic selection to catch the society influence on trip choice. They defined comparison of happiness and proved that comparison of happiness rising from comparison of the individual and others. Hence, the overall satisfaction or utility commute would be increased.

The third field is about the resident individual trip information. Son Sanghoon and Khattak Asad (2011) comparatively analyzed the university students' access and use of information about travel. The results showed that the internet and a wide range of information have great impact on the students' trip choice.

Although research on individual trip behavior has been carried out for several years, the theory supports and application explorations are still weak which could be formulated as following:

In the aspect of subject investigated: there is few study on simulation of individual trip decision. They always focus on degree of satisfaction more than personal preference.

In the aspect of research method: They are always qualitative and learned from other existing results. Most studies focus on the travel factors of specific travel groups and only extract the importance of factors even if using the quantitative methods. Moreover, the method always is a statistical analysis of the factors. The application of modern mathematics is only a minority. Meanwhile, the comparison and analysis of different methods basing on results is lack.

Therefore, this study hopes to achieve a breakthrough in the above areas. The research content focus on the simulation of travel decisions basing on personal preference. In the research methods, combined with statistical theory, the modern mathematical theory, accurately extracting of certain population preferences through rough set theory to establish individual travel decision model based on genetic algorithms and simulation analysis.

2 Methodology

The travelers are considered absolutely rational in the traditional transportation trip model. However, more and more research and practical experience shows that the assumption of absolutely rational is incorrect. Hence, the conception of bounded rationality in Behavioral Economics provides new ideas for research on individual trip behaviors and decisions. Bounded rationality means that the travelers will not chase the best travel utility those are always affected by the factors such as trip experience

and travel habits. They want to guarantee the difference between actual travel utility and best travel utility to be in an acceptable range. The conception of bounded rationality is of course accepted in this paper. However, this paper is still taking conception of absolutely rational as foundation in order to carry out the research procession. The assumptions are as following,

1. The decisions are at least 2 when a resident will start travel. That is $n_d \geq 2$;
2. The residents confirm all the decisions before one travel, ignoring all the second decision on the road;
3. The travel utility U_d is different for individual preference. Each resident is always chasing their best travel utility. The best travel utility is marked as U_{opt}.
4. The resident decision set D is the independent variable. Travel utility U is the dependent variable. The merits of the decision-making travel decide the effectiveness of travel utility.
5. In this section, the resident decisions set are divided into two categories: good set and bad set. At the same time, more vivid examples could be used to illustrate the relationship between residents travel decisions and travel utility. To make:

$$\begin{cases} D_B = 1 \\ D_G = 2 \end{cases}, \qquad \begin{cases} U_d = 1 \\ U_{opt} = 2 \end{cases}$$

However, it is necessary to cost for resident making travel decision, for example time costs and experience costs. Therefore, d is the explanatory variables of C. Good decision-making cost should be higher than the cost of inferior decisions because the merits of the decision-making will lead to changes of the decision-making cost. Therefore, to make:

$$\begin{cases} C_{D_B} = 1/2 * d \\ C_{D_G} = d \end{cases}$$

Residents are always wandering between the worst decision and the optimal decision. They do not want to make the worst decision; Meanwhile, they are also likely unable to make optimal decisions. The reason lays in the uncertainty of complex travel environment. Therefore, residents who have bounded rationality are always making decision between optimal and worst. It is an acceptable decision set for residents. However, this paper only focuses on the personal preference because of the travel environment's uncertainty and complexity. The assumption in this paper is the resident is absolutely rational and chasing best travel utility.

Therefore, this paper will start with the resident individual travel behavior. Firstly, rough set theory will be applied to research on travel decision affecting factors and the contribution of each factor weight or the contribution weight of any two factors could be confirmed. Secondly, simple genetic algorithms will be used to simulate individual residents travel decisions and each genetic operator specific parameters and corresponding fitness function will be confirmed. Finally, modified single urban resi-

dent travel decision model basing on rough set and simple genetic algorithms will be established, compared and explained.

Hutchison individual residents y, individual resident's travel decisions affecting factors x supposing there are n factors affecting the individual a trip. The complete factor set which are affecting individual resident trip decision could be defined as $X = \{x_1, x_2, x_3, \cdots, x_n\}$. The travel utility of individual y is U_d. Individual y is always chasing the best travel utility U_{opt}. Therefore, y may choose a group of feasible decisions D. He could achieve the best travel utility through the comparative analysis. At the same time, the assumption above for the residents travel decisions here is still applicable.

2.1 Pre-processing the model based on rough set theory

(1) Select important factors.

There are some other factors may influent the individual decision expect for some common factors such as safe, comfort, cost and time which are always influencing the trip decision because of the differences in individual preferences, the inequality of individual travel decision quantity and the difference in purposes. Those factors could be defined as $\tilde{\lambda}$

Define the decision $D = f(X) = f_1(Xc) + f_2(\tilde{\lambda})$. X_c is common factor. Meanwhile, X_c and $\tilde{\lambda}$ establish the complete factor set. That is $X_c \cup \tilde{\lambda} = X, X_c \cap \tilde{\lambda} = \Phi$.

The common factors and other factors influence the individual y's decision. However, they could not decide the decision combination of y. The final decision of y could determine the individual travel utility U_d in the end.

Define $U_d = g(D)$, in which the best travel utility $U_{opt} \geq U_d$.

Obviously, there are n factors in the complete factors set X. There are k factors in common factors set X_c. Of course, $k \leq n$. The function of rough set is to reduce the original complete factors set to get an important factors set. Hence, the important factors set is basing on common factors set. The important factors set could be marked as X_I, $X_I = \{x_j \mid x_j \in X_c\}$. That means there are P important factors selected.

(2) Expression of knowledge System.

Generally speaking, a knowledge system could be expressed as an ordered quad $S = \{U, R, V, f\}$ among which, $U = \{x_1, x_2, x_3, \cdots, x_n\}$ is on the field. It is the set of whole samples. $R = C \cup D$ is the set of attributes, among which sub set C is condition attribute set reflecting the characteristics of an object and D is decision attribute set reflecting the categories of objects. $V = \cup V_r$ is the set of attributes values. V_r represents the range of attributes. $f : U \times R \to V$ is an information function to confirm the attribute of every object X in set U. That is, if $x_i \in U$, $r \in R$, then $f(x_i, r) = V_r$. Hence, R is a rough set. $\alpha_R(X)$ could be considered as the accuracy which is approximating the set X under the equivalence relation R.

The original data could be filtered and the invalid data and the inapplicable data will be deleted. Furthermore, a two-dimensional decision information table will be established to describe urban travel choice factors knowledge expression system.

(3) Calculate the univariate weighting matrix.

The univariate weighting matrix could be obtained by extending the reduction important factors set. $I = \{w_1, w_2, \cdots, w_p\}$.

(4) Calculate the correlation factor weighting matrix.

The correlation factor weighting matrix could be obtained by the univariate weighting matrix,

$$R = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1p} \\ w_{21} & w_{22} & \cdots & w_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ w_{p1} & w_{p2} & \cdots & w_{pp} \end{bmatrix}$$

Meanwhile, R is a $p \times p$ real symmetric matrix. The weight value on the main diagonal equal the corresponding weight values in the univariate factor matrix, which is $R.w_{11} = I.w_1, R.w_{22} = I.w_2, \cdots, R.w_{pp} = I.w_p$.

The univariate factor matrix and correlation factor matrix will be explained and analyzed. The Genetic Algorithms will be introduced after the confirmation of univariate factor matrix and correlation factor matrix.

2.2 Genetic Algorithms modeling

The decision is not always single when the individual resident y is travelling. Therefore, the set of decision variables is $D = \{d_1, d_2, d_3, \cdots, d_m\}$.

The m is the number of decisions and $d_i = \begin{cases} 0 \\ 1 \end{cases}$

Define the trip objective function of individual resident y is $U(D) = RX_I$. It represents the best travel utility that y is chasing. The fitness function could be marked as $fitness$ and $U_{opt} = \max U(D)$.

The fitness evaluation criteria are:

1. The value of fitness is the larger the better. The larger means the individual travel utility is better and the adaptability is better.
2. The individual travel optimal effectiveness could be selected by selecting the superior individual fitness.
3. The best travel utility $U_{opt} = \max U(D)$ is the maximum value of the objective function and it is positive number.

3 Data modelling, calculation and analysis

3.1 Research data sources

Data in this paper is obtained by questionnaires. Questionnaires were issued in September 2010 to about 1000 and recovered 486 valid questionnaires. The investigation time covered three daily time paragraphs: 7:00~11:00, 11:00~13:00 and 15:00~19:00. The survey subjects covered all sectors of society in Beijing as entirely as possible including different ages and different occupation. The survey locations covered the major urban areas in Beijing including bus stations, railway stations, shopping street, business office, park attractions, transit hub, as well as subway cars and other places. The questionnaire questions reflected the urban travel choice influence factors. These factors constituted the index system of urban travel choices, as shown in Table 1.The five factors were travel time, travel prices, travel distance, travel safety, traveling comfort.

Table 1. index system of urban travel choices

Aim	Factors	Index
Analysis on urban travel options factors	Characteristics	Age
		Sex
		Occupation
		Travel options
		The average monthly disposable income
		The average monthly transport expenditure
		Car or not
	Travel time accuracy	The accuracy of the destination arrival time
		Peak or off peak
		Waiting time
	Travel costs	Ticket price
		Parking price
		Fuel costs
		tolls
	Travel distance	Total distance
		Transfer distance
		Number of transfers
	Travel safety	Driving safety
		Property security
	Traveling comfort	Seat or not
		Crowded or not
		Facilities convenience
		Bulky items or not

3.2 The knowledge expression of urban travel factors and weight calculation

As seen from Table 1, urban travel choice factors are not only including time, piece and other factors known to public, but also including comfort, safety and other factors.

The resident survey questionnaire orders could be used as research object collection according to the rough set theory. The invalid data and inapplicable data are deleted. The remaining data set is $U=\{x_1, X_2, x_3, x_4, x_5...\}$, a, b, c, d, e represents the importance of the accuracy of the time, the importance of the travel expense , the importance of the total travel distance, the importance of travel safety, and the importance of travelling comfort. It is the public factors set $X_c=\{a,b,c,d,e\}$, which establishing the condition attribute set of urban travel choice influence factors $c=\{a,b,c,d,e\}$The decision attribute collection set D is established by the chosen travel mode. Hence, the corresponding figures are used to represent the items in condition attribute as following,

{1=very unimportant；2=relatively unimportant；3=general；4=more important ；5=very important}

Similarly, in the decision attribute, the corresponding set is

{0=bicycle；1=bus；2=underground；3=taxi；4=car}

Therefore, dimensional decision information table could be established to describe the urban travel choice influencing factors knowledge expression system. As shown in Table2 (Experts):

Table 2. urban travel choice influencing factors knowledge expression system

No.	Condition Attributes					Decision
	Time accuracy	Cost	Distance	Safety	Comfort	Travel mode
x_1	1	4	4	5	5	2
x_2	1	5	4	5	3	1
x_3	2	3	2	3	3	2
x_4	2	4	3	3	3	1
x_5	3	1	3	5	4	1
x_6	3	2	3	5	4	1
x_7	3	2	5	5	5	0
x_8	3	3	2	5	5	3
x_9	3	3	2	4	2	4
…	…	…	…	…	…	…

The univariate weight matrix could be calculated result as $I=\{4/27,8/27,7/27,2/27,6/27\}$ according to the rough set theory. And the corresponding weight matrix could be calculated result as

$$R=\begin{bmatrix} 4/27 & 15/126 & 17/126 & 4/126 & 14/126 \\ 15/126 & 8/27 & 18/126 & 10/126 & 15/126 \\ 17/126 & 18/126 & 7/27 & 10/126 & 14/126 \\ 4/126 & 10/126 & 10/126 & 2/27 & 9/126 \\ 14/126 & 15/126 & 14/126 & 9/126 & 6/27 \end{bmatrix}$$

On this questionnaire concerned, the crowd comprehensive collective preferences are $\text{cost} \succ \text{distance} \succ \text{comfort} \succ \text{time accurancy} \succ \text{safety}$ (symbol $\succ$ represents perference and means better than)

3.3 Rough- Genetic Algorithms Modelling

The cost (C), distances (L), comfort (C, F), time accuracy (T) as well as cost and time accuracy combination(C, T), time accuracy and distant (T, L) are selected as parameters according to the calculation result basing on rough set theory.

$$U(D) = WX = \sum \begin{bmatrix} W_C H(C) + W_L K(L) + W_{C,F} Q(C,F) + W_T G(T) + \\ W_{C,L} P(C,T) + W_{T,L} V(T,L) + \varphi \end{bmatrix}$$

In the formula, φ represents the additional satisfaction. It is the utility value of other factors λ. These factors are generally classified as one group due to the unobserved character.

The objective function is transformed into the fitness function.

$$f(D_j) = \sum_{d_1}^{d_n} [w_t g(t_{d_j}) + w_c h(c_{d_j}) + w_l k(l_{d_j}) + w_{c,f} q(c_{d_j}, f_{d_j}) + w_{c,t} p(c_{d_j}, t_{d_j}) + w_{t,l} v(t_{d_j}, l_{d_j}) + \varphi_{d_j}]$$

A specific fitness function could be given according the univariate weight matrix and the corresponding weight matrix which are confirmed basing on rough set theory.

$$f(d_1, d_2) = 8/27 * d_1^2 + 7/27 * d_2^2 + 18/126 * d_1^3 + 17/126 * d_2^3 - 15/126 * d_1 * d_2 - 4/27 * d_1 - 6/27 * d_2$$

4 Examples of simulation, results and analysis

The model in this paper is for individual. The data is randomly selected from actual questionnaire in order to fit the simulation results more realistic. This questionnaire is answered by a lady living in Beijing. The average monthly disposable income is 2000 to 3000 RMB. The average monthly transportation expense is from 200 to 300RMB. And she did not have a car.

The simulation follows the travel situation: (1) the travel aim is tourism; (2) no bulky items; (3) no travel this line before. A calculation shows that the cost of travel traffic scope of its disposable income is [6.7%， 15%]. The proportion of expenditure is relatively high.

The trip will be converted to a travel decision model based on genetic algorithms. The decisions that should be confirmed including (1) Travel satisfaction should achieve highest; (2) Line i select sub way, bus or taxi? (3) Line ii select subway, bus or taxi? Conventions decision variable D is taken from [1, 4]. If the value of D is [1, 2） the selection is bus. If the value of D is [2, 3） the selection is underground; if the value of D is [3, 4] the selection is taxi. $d_1, d_2 \in [1,4]$, which are two decision variable. Matlab tool box is used to do program. The following figures are the results of the operation.

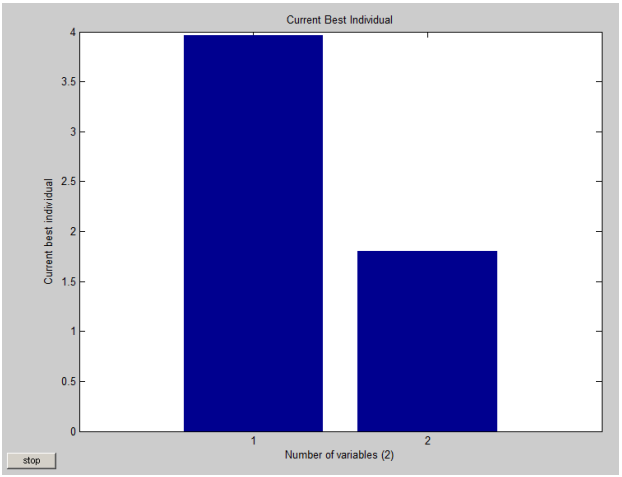

Fig. 1. rough- genetic algorithms simulation results

Fig.1 shows the individuals which has the best fitness value of each generation. According to fig. 1, the final selections of the individual are: 3.6474 and 3.14209. The corresponding phenotypes are both taxi.

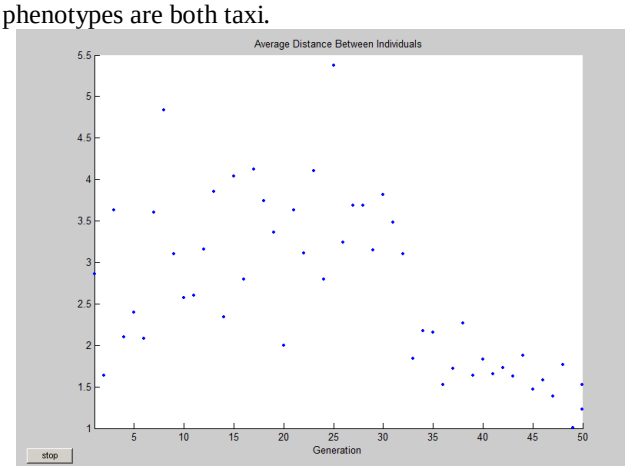

Fig. 2. the average distance among individuals in rough-genetic algorithms model

Fig.2 describes the average distance between individual in each generation. According to Fig.2, the average distance among individuals from far to near and from big to small are eventually nearly coincides through the rough –genetic algorithms optimization. That represents the individual selected eventually by the rough-genetic algorithms are high quality.

Therefore, the lady is advised to choose taxi both on line i and on line ii according to the results given by the rough-genetic algorithms. The final prediction is that the lady should choose a taxi directly on this trip and she will achieve the optimal effectiveness.

5 Conclusions

In this paper, the single factor weight matrix and various factors related to weight matrix are brought out by analyzing the factors of individual residents travel behavior and decisions which are restricting their travel. Furthermore, this paper utilizes the advantage of Genetic algorithm to simulate the individual residents travel decision behavior. An example simulation is given in this paper to verify the validity of the travel decision-making simulation model.

It is contributing for improve the accuracy of the study by combining the rough set and genetic algorithm together in this paper. However, most of the residents could not make the best travel decisions during the travel time. They would choose a group of satisfaction with the decision-making under common situation. That is to allow the difference of travel utility and optimal utility in the acceptable range. Therefore, in process of the genetic algorithm modeling, the objective function and fitness function should be changed. This is the further research contents.

Acknowledgments: This paper is supported by the National Natural Science Foundation (No. 71103014) and Beijing Philosophy and Social Science Fund (No.14JGC095).

Reference

1. Iseki Hiroyuki, Smart Michael, 2012. How Do People Perceive Service Attributes at Transit Facilities? Examination of Perceptions of Transit Service by Transit User Demographics and Trip Characteristics. Transportation research record, pp. 164-174.
2. Abou-Zeid Maya, Ben-Akiva Moshe, 2011. The effect of social comparisons on commute well-being. Transportation research part A-Policy and Practice, pp. 345-361.
3. Son Sanghoon, Khattak Asad, Chen Ju-Yin, 2011. Comparative analysis of university students'acquisition and use of travel information.Transportation research record, pp. 46-54.
4. Karl Kim, Pradip Pant, Eric Y. Yamashita, 2008. Hit-and-Run Crashes: Use of Rough Set Analysis with Logistic Regression to Capture Critical Attributes and Determinants. Transportation Research Record, pp. 114-121.
5. Stephen Clark, 2009. Characterising and predicting car ownership using rough sets. Transportation Research Part C: Emerging Technologies, pp. 381-393.
6. Min Fan, Ping Zou, Dong Zhu Xing, 2010. Transportation choice model based on rough set theory. Proceeding of 2010 Asia-Pacific youth conference on communication, pp. 982-984.
7. Ceylan Huseyin, Ceylan Halim, 2012. A Hybrid Harmony Search and Transyt hill climbing algorithms for signalized stochastic equilibrium transportation networks. Transportation research part C-Emerging technologies, pp. 152-167.
8. Roca-Riu Mireia, Estrada Miquel, 2012. Trapote Cesar. The design of interurban bus networks in city centers. Transportation research part A-Policy and Practice, pp. 1153-1165.
9. Hong Wei-Chiang, Dong Yucheng, Zheng Feifeng, Wei Shih Yuhg, 2011. Hybrid evolutionary algorithms in a SVR traffic flow forecasting model. Applied mathematics and computation, pp. 6733-6747.
10. Dubois D, 1990. Prade H. Rough Fuzzy Sets and Fuzzy Rough Sets. International Journal of General System, pp. 191-209.