# Taxi Resource Allocation in the Era of "Internet+"

Xiaodong Xu

# North China Electric Power University, Baoding 071000, China 513178577@qq.com

### Abstract

In view of the problem A, first based on a large number of analysis reports and our own understandings, we defined "supply and demand matching degree" as  $\beta$  value can be quantified and obtained. And then by looking up the related document literature, through summary can get five groups of operational indicators easy to be obtained from the data as: population density, travel frequency, the selected proportion of the taxi, starting price, average waiting time, and the different areas and different time of target data represent different space-time concept in problem A. In the five groups of indicators data of Beijing time 8:00 to be correlation analysis with  $\beta$ , further selected four groups of strong correlation major indicators as population density, travel frequency, selected taxi ratio, average waiting time; then using selected these four groups of main indicators and  $\beta$  to build multivariate linear regression model, and the regression equation between  $\beta$  and every indicator is obtained; in order to verify the rationality of the established regression equation, with Beijing time 22:00 four groups of indicators into the model, get a set of  $\beta$ predictive value, and the actual value of the  $\beta$  using Pearson correlation coefficient to analyze the correlation, found the correlation is stronger, shows that the established model with strong degree of coincidence of the actual situation; so it can use this model to analyze the influence of different space and time indictor on the matching of supply and demand. In view of the problem B, it requires to design a reasonable and effective subsidy scheme, and verify the rationality of the scheme. An effective and reasonable subsidy scheme can inevitably reduce the problem of "difficult to take a taxi", by searching data, found that the taxi daily total mileage is a certain value, so the taxi daily effective mileage, can directly affect the difficulty level of taking a taxi. So we reflect the difficulty level of taking a taxi by the daily effective mileage of the taxi, and take it as a goal, combined with Module of Preventing Increase for the establishment of the model, through choosing the middle variables such as orders, software market shares, the frequency of taking a taxi, etc., to establish the relation function between daily mileage of the taxi with subsidies amount. It can be found from the observation function images, with the increase of the subsidies amount, the daily effective mileage of a taxi shows the tendency of increase first then decrease. Through consulting data, we can regulate the effective mileage above 250 kilometers can ease the difficulty of taking a taxi at the greatest extent. Then substitute it into the known data, compared with data obtained from function, the difference of both is not large, then it can be concluded that the scheme is reasonable.

Keywords: Regression model Correlation analysis Module of Preventing Increase.

# **1. Problem restatement**

Taxi is one of the important means of vehicles for citizen's activities, so "difficult to take a taxi" is a social hot spot people concerned. With the advent of the era of "Internet +", there are many companies have established the software service platform for taking a taxi by relying on mobile Internet, to realize the communication between the passengers and the taxi drivers, at the same time introduced a variety of taxi subsidy schemes. Please search relevant data for mathematical modeling to study on the following problems:

Try to establish reasonable indicator, and analyze the "matching degree of demand and supply" of taxi resources in different space and time.

If you want to establish a new software service platform for taking a taxi, what kind of subsidy scheme should be designed and how to verify its rationality.

# 2. Model Hypothesis

Assuming all the collected data are reliable, and can reflect the actual situation;

Assuming a linear correlation between the quantity demanded of taxi and each indicator;

Assuming all taxi drivers regard making profits as object, and the maximize profit;

Assuming the factors of policy are not considered, for policy is likely to change the taxi software development;

Assuming the empty rate is main influencing factor of the problem of "difficult to take a taxi", without regard to the influence of other factors.

Symbol	Description
Q	Total population
J	Every mileage of a taxi
М	Taxi order
F	Frequency of taking a taxi
Ν	Taxi amount
С	Coverage rate of taxi software
β	Reflect matching degree of demand and supply

### **Symbol Description**

# 3. Problem Analysis

# **3.1 Problem an Analysis**

Problem A requires we establish a reasonable indicator, and analyze the matching degree of demand and supply of taxi resource of different space and time. To analyze the matching degree of supply and demand based on the established indicator, obviously, each indicator is different in different area and different space and time, therefore, we can analyze the matching degree between supply and demand amount by analyzing the relations between supply and demand and each indicator.

The first step is to determine the appropriate and persuasive indicators, we here by looking up relevant document literature, and make classifies and summaries, can get a set of operational, easy data acquisition and strong correlation indicators.

The second step is to collect the relevant data, for "the matching degree of supply and demand" under the different time and space, in the aspect of region, dividing according to different administrative regions in different cities, in the aspect of time, dividing according to day section and night section in a day, carry out data collection of demand and supply amount of taxi on smart travel platform of DiDi. While the corresponding indicator at different space and time of different areas is collected through the Internet and literature.

The third step is to establish a linear regression model, using the collected data to obtain the regression equation of the demand and the indicators, and do regression test, after model complete, and test with the data, it is found that the results obtained

from model are close to the actual value, thus the rationality of the model is verified.

The fourth step, after the corresponding taxi demand amount obtained through each indicator in random time and region, then the matching degree of demand and supply of taxi is defined, and analyze the matching degree of the fitting demand amount and supply amount, and then obtain the final conclusion.

#### **3.2 Problem B Analysis**

Problem B asked us to design effective and reasonable subsidy scheme for a taxi software service platform, and verify the rationality of the scheme. An effective reasonable subsidy scheme should be inevitably relieve the difficulty of taking a taxi, even the taxi effective mileage X increases, here establishing a model for this purpose. And to prove the rationality of the subsidy scheme, substitute the collected data into the established subsidy model, if it can effectively increase the effective mileage X, then it can prove the model is reasonable.

First is the establishment of subsidy scheme model, through a series of economic relation derivation, by MATLAB data fitting analysis first, it considers that the coverage rate of taxi software and subsidies is conformed to Module of Preventing Increase; then set up Module of Preventing Increase, here only to establish the amount of subsidies as the independent variables, and the relation of effective mileage X with the subsidy amount a is obtained, and then the corresponding subsides scheme at the maximum mileage is the best solution. Next is the validation of rationality of subsidy scheme model, substitute the subsidies of different stage into the formula for analysis, compared the obtained value with the actual value, the small difference degree means the established subsidy scheme model with higher rationality.

# 4. Model Establishment and Solution

#### 4.1 Model Establishment and Solution of Problem A

The study of the matching degree of demand and supply of taxi resources, and taxi supply and demand balance has very important significance and strong social significance for optimal allocation of resources in the industry, the coordination and sustainable development of the whole urban traffic system, etc. When the demand is less than the supply, there will be driver refuse to take passengers, service quality declining, can't meet the requirements of urban residents travel; when supply exceeds demand, there will be the relative surplus situation of taxi supply, not only the taxi industry cannot get the maximum benefit, but also bring the urban traffic congestion, etc.

In order to make a comprehensive analysis of the matching of demand and supply of taxi resources in different time and space, you need to establish a reasonable indicator, systematically and comprehensively consider about various factors. Here, we establish related indicators by searching data all round, then quantize the matching degree of supply and demand, next use the regression analysis model for the solution of the weight of each index on the match degree of supply and demand, and finally with the established model for the validation of the collected data.

### 4.1.1 Select Reasonable Indicators with Principal Component Analysis

The concept of taxi supply and demand is a pair of both associated and distinguished. The taxi supply is set by the government for taxi development

planning and operation management measures. And demand is the payment ability provided in social and economic life in terms of people and cargo space displacement, which are associated with a variety of factors.

There are many factors influencing the taxi passenger transport demand, here, by looking for related information<sup>[1]</sup>, and making analysis and summary, we selected several common indicators that influencing the demand of taxi, including external factors and internal factors, while population density selected as external factor, and the internal factors including the frequency of travel, proportion of travel, starting price and passenger's average waiting time, a total of five indicators.

The indicators affected taxi demand are shown in Fig. 4.1.1



Because the influence of selected indicators in this problem on the matching degree of supply and demand cannot be ensured, therefore, it makes further screening for these five indictors with the principal component analysis model.

First to standardize data, calculate the correlation coefficient matrix, then calculate the eigenvalue and eigenvector, and choose the main component, the main composition analysis of five indicators is shown in table 4.1.1:

	population density	travel frequency	travel proportion	taxi starting price	average waiting time
correlation index	-0.886	-0.959	-0.989	0.803	-0.934
significance	0.045	0.010	0.001	0.101	0.020

Table 4.1.1 Main Composition Analysis of Five Indicators

It is known from the table, the significant of taxi starting price is 0.101 > 0.05, and a significant level of 0.05 is not relevant, so this indicator can be removed. Thus four indicators after the principal component analysis: population density, travel frequency, travel proportion and the average waiting time.

4.1.2 The Definition of the Matching Degree of Supply and Demand: We collect the taxi quantity and taxi demand in different regions and different time on our Kuaidadi smart travel platform<sup>[2]</sup>, and the corresponding empty-loading ratio of taxi of space and time. Empty quantity of different time in different regions can be classified as:

### Empty quantity = taxi quantity $\times$ empty rate

In order to analyze the matching degree of taxi demand and supply, here, we set a quantitative value  $\beta$  can be obtained, used to measure the matching degree of taxi supply and demand and make definition and classification as follows:

$$\beta = \frac{\text{Empty quantity}}{1}$$

 $p = \frac{p}{\text{the demand of taxi}}$ Refer to relevant information<sup>[3]</sup>, we know that the taxi empty rate is an important factor to measure the relationship between demand and supply of the taxi. According to the investigation and analysis of our country city taxi association on domestic and foreign typical city, when the empty rate of the taxi at about 30%, city taxi passenger supply and demand achieve a balance basically; when taxi empty rate lower than 25%, there is obviously taxi passenger transport demand exceeds supply,

passengers' waiting time getting longer, taxi service quality reduced; When the taxi empty rate is higher than 40%, and the city taxi passenger for too need phenomenon, it is obviously taxi passenger transport supply exceeds demand, and the passengers' waiting time shortened, taxi service quality improved, but the taxi industry profits down, and bring a lot of pressure to urban traffic.

Based on the above information it can make theory criterion of classification of  $\beta$  value, the classification criteria of  $\beta$  on matching degree of supply and demand are shown in table 5.1.2, and the greater of  $\beta$ , the degree of the matching of supply and demand is higher.

B value	matching degree of supply and demand		
$\beta \ge 1.3$	better		
$1.1 \le \beta \le 1.3$	general		
$\beta \le 1.1$	worse		

 Table 5.1.2 Matching Degree Grade of Supply and Demand Definition

Here take relevant data of taxi in different time of six different administrative districts of Beijing and Shanghai as an example, the collected data are shown in table 5.1.3:

District		taxi demand	taxi quantity	taxi unloaded ratio	empty quantity	β
	Chaoyang District	2054	9073		1996	0.971762
Beijing	Xicheng District	698	2859	22	629	0.901146
	Haidian District	1502	6858		1509	1.00466
	Minhang District	4930	17402		5220	1.058824
Shanghai	Huangpu District	651	6211	30	1863	2.861751
	Pudong new area	3617	12605		3781	1.045341

### 4.1.3 The Establishment of the Multiple Regression Model: (I) Model establishment.

After the matching degree of taxi supply and demand are defined and the establishment of a reasonable indicator, we use the collected data, establish regression analysis model to analyze the relationships between reflecting the matching degree of taxi supply and demand of  $\beta$  and the various indicators. The following is the example of the relevant data of taxi at different time and six different administrative regions in Beijing and Shanghai, using multiple linear regression models for fitting, and regression equation between  $\beta$  and various indicators are obtained.

In multiple regression analysis, set the reflection of a certain moment area matching degree of taxi supply and demand  $\beta$  as y, the influencing matching degree indicators as:regional population density  $x_1$ , travel frequency  $x_2$ , travel proportion  $x_3$ , average waiting time  $x_{4\circ}$ 

Let regression equation as:

 $y = b_0 + b_1 x_1 + \dots + b_m x_m + \epsilon, \ \epsilon \sim N(0, \sigma^2)$ 

Where,  $b_0, b_1, ..., b_m$  are parameters to be solved,  $\varepsilon$  is normal random variable with compliance mean value of 0, variance of  $\sigma^2$ 

Here, we use SPSS to carry out multivariate regression analysis on each indicator of Beijing time 8:00 at different district and  $\beta$  value, and the final regression equation is obtained as:

 $y = 1.399 - 0.05 \times x_1 - 0.145 \times x_2 - 0.291 \times x_3 - 0.012 \times x_4$ 

District		Population density	Travel frequency	selected taxi proportion	average waiting time	β
	Chaoyang District	7530	2.8	23.08	35	0.971762
Beijing	Xicheng District	24517	2.91	23.24	38	0.901146
	Haidian District	7702	2.74	22.14	31	1.00466
	Minhang District	6653	2.48	8.95	30	1.058824
Shanghai	Huangpu District	51667	2.61	2.21	12	2.861751
	Pudong new area	4168	2.54	8.63	27	1.045341

# Table 5.1.4 Summary Sheet of each Indicator of Beijing Time 8:00 at Different District and $\beta$ Value

(II) Model test

Following we use each indictor at Beijing time 22:00 at different district and matching degree for hypothesis testing, the data summary are shown in table 5.1.5.

District		Population density	Travel frequency	selected taxi proportion	average waiting time	β
	Chaoyang District	7530	2.13	24.63	28	1.245614
Beijing	Xicheng District	24517	2.2	25.22	31	0.995249
	Haidian District	7702	1.96	23.27	21	1.067818
	Minhang District	6653	1.98	9.23	20	1.375897
Shanghai	Huangpu District	51667	2.07	2.87	13	2.826087
	Pudong new area	6653	2.24	9.17	21	1.26548

Table 5.1.5. Summary Table of Beijing Time 22:00 at Different District and  $\boldsymbol{\beta}$ 

Here, each indicator of Beijing time 22:00 at different district is substituted into fitting and regression model is obtained, and predicted value of  $\beta$ , the specific data is shown in table 5.1.6, where, the relevant data fitting degree of Huangpu district is worse, it belongs to abnormal condition should be removed.

# Table 5.1.6 The Predicted Value and Actual Value of $\beta$ at Beijing Time 22:00 at Different District

District		Predicted value	Actual value	
	Chaoyang District	1.0344	1.245614	
Beijing	Xicheng District	0.9010	0.995249	
	Haidian District	1.0121	1.067818	
Shanghai	Minhang District	1.1104	1.375897	
	Huangpu District	1.1067	1.26548	

In order to analyze the rationality of regression model from fitting, here we make correlation analysis on predicted value and actual value, the higher of the correlation, indicates the predicted value from the model is more close to the actual value. For correlation analysis, first it is assumed that there was no significant difference between the predicted value and actual value:

# $H_0: \mu_1 - \mu_2 = 0$

Here we use SPSS software to calculate the Pearson correlation coefficient of two lists of predicted value and actual value, to analyze its correlation, and the results are shown in table 5.1.7:

No.	Pearson correlation coefficient	significance (bilateral)	Ν
β	0.910	0.032	5

# Table 5.1.7 Correlation Analysis Result of Predicted Value and Actual Value of $\beta$

In correlation analysis, the level of significance is considered as  $\alpha = 0.05$ , it can be seen from the table, significance is 0.032 < 0.05, indicates the reject test assumed H<sub>0</sub> at significance level of  $\alpha = 0.05$ , and then it can estimate the significance relation existing in two variables. While the absolute value of Pearson correlation coefficient is 0.910 which is very close to 1, indicates the higher correlation degree of predicted value with actual value, so the rationality of the established regression model is stronger.

**4.1.4 Analysis of Matching Degree:** It can calculate the matching degree  $\beta$  of taxi demand and supply from the established model, here we analyze the matching degree with the grading standard defined at previous section, and the obtained matching results of Beijing time 22:00 at different district are shown in table 5.1.8.

Table 5.1.8 The Level of Matching Degree of Beijing time 22:00 at DifferentDistrict

district		β	Matching degree
	Chaoyang District	1.0344	worse
Beijing	Xicheng District	0.9010	worse
	Haidian District	1.0121	worse
Classical as	Minhang District	1.1104	General
Snangnai	Huangpu District	1.1067	general

# 4.2 Model Establishment and Solution of Problem B

**4.2.1 Establishment of Subsidy Scheme Model:** 1. Coverage rate of taxi software and subsidies relation - An effective and reasonable subsides scheme can release the problem of difficulty of taking a taxi inevitably, that is the implementation of subsidies scheme can increase the effective mileage of taxi effectively. Thus it adds effective mileage of taxi X as the objective to establish model here.

Compared the same 6649 taxies data before the promotion of taxi software (October, 2013) with the mature taxi software (October, 2014) data <sup>[6]</sup>, we found that the daily the taxi average daily operating mileage has not been affected, always stay at about 293 km/car. Accordingly, before and after the promotion of taxi software, total mileage L of taxi is basically unchanged.

Here we assume the average single mileage as fixed value J, if you want the effective mileage X increases, then the added order M should be:

$$X = J \times M$$

Passengers are divided into two groups of using taxi software and no using taxi software,  $M_1$  is the order number from group using taxi software,  $M_2$  is the order number from group without using taxi software.  $M = M_1 + M_2$ 

and order is:

$$M = N \times F$$

Define N as the number of taking a taxi, F is the frequency, it can be known that:  $M_1 = N_1 \times F_1$   $M_2 = N_2 \times F_2$   $F_1$  can be obtained from data, while  $F_2 = D$  (travel frequency) × K (selected taxi travel proportion).

The number of taking a taxi  $N_1$  and  $N_2$  is changed with coverage rate of taxi software, set total population is Q, then the selected the population of taking a taxi is W, then we have:

$$W = Q \times K$$
  

$$N_1 = W \times C$$
  

$$N_2 = W \times (1 - C)$$

While the coverage rate C will be changed with the increase of subsidies

Our collected the accumulative account number of the national taxi software and subsidies, under the constant user base in the country, the accumulative account of national taxi software is a strict proportional relation to the coverage rate C, so the relationship between the total account scale and the subsidy amount can be considered as the relationship between coverage and subsidies. The relational fitting data are shown in table 5.3.1:

 
 Table 5.3.1 The Accumulative Account Scale of Taxi Software and Subsidies in our Country

	2013	January 2014	February, 2014	March, 2014
Accumulative account scale (million)	2160	3200	5682	9828
Subsidies (Yuan)	0	10	12	16



Using MATLAB for fitting analysis on both groups of data,:

Therefore it can make rational assumption: the relation of both is consistent to Module of Preventing Increase.

2. Principle of Module of Preventing Increase

It assumed that the growth rate of taxi software coverage rate r (C) is the coverage rate C(a) function at subsidies of a Yuan, According to the practical consideration, r (C) should be the decreasing function of C.

Assumed the growth rate function r(x) is a linear function of x, and  $r(x) = r - s \times C$  (r >0, s > 0), in this equation, r is the fixed growth rate, represents the growth rate at small subsidies, in order to determine coefficient s, quote the

maximum coverage rage  $C_m$  in society, and named as coverage capacity, but coverage rate is not increase at C= $C_m$ , the growth rate is 0, so r(C) = r ×  $(1 - \frac{c}{C_m})$ .

Obtained according to the assumption:

$$\frac{dC}{da} = r(C)C$$
$$C(0) = C_0$$

Dissolved as:

$$C(a) = \frac{C_m}{1 + \left(\frac{C_m}{C_0} - 1\right)e^{-ra}}$$

Substituted into the above equation, we can obtain the relation of effective mileage and subsidies, and that is the relation of subsidies and difficulty level of taking a taxi:

$$\mathbf{X} = \mathbf{J} \times (\mathbf{Q} \times \mathbf{K} \times \frac{c_m}{1 + \left(\frac{c_m}{c_0} - 1\right)e^{-ra}} \times F_1 + \mathbf{Q} \times \mathbf{K} \times (1 - \frac{c_m}{1 + \left(\frac{c_m}{c_0} - 1\right)e^{-ra}}) \times \mathbf{D} \times \mathbf{K})$$
(1)

We only solve the maximum effective mileage as the optimal scheme.

Here, take the data of Beijing in 2014 as example, substituted the collected data into equation (1), and obtained the following relation equation:

$$X = 8 \times \left( 21148000 \times 6.5\% \times \frac{68\%}{1 + \left(\frac{68\%}{23\%} - 1\right)e^{-0.26a}} \times 0.43 + 21148000 \times 6.5\% \right) \times \left( 1 - \frac{68\%}{1 + \left(\frac{68\%}{23\%} - 1\right)e^{-0.26a}} \right) \times 2.8 \times \frac{68\%}{1 + \left(\frac{68\%}{23\%} - 1\right)e^{-0.26a}} \right)$$

After simplify and obtain  $X = 8 \times \left(\frac{401938.89}{1+1.956 \times e^{-0.26a}} + 3848936 \times \left(1 - \frac{0.68}{1+1.956 \times e^{-0.26a}}\right) \times \frac{0.68}{1+1.956 \times e^{-0.26a}}\right)$ (2)

Equation (2) is the relation model of effective mileage X and subsidies a. curve relation schema as shown in Fig. 5.3.2:



It is known from Fig. 5.3.2, the corresponding subsidies at the cure with higher effective mileage is 7 to 12 Yuan, that is to say, the monthly subsidies of Beijing in 2014 is 7 to 12 Yuan, can release the difficulty of taking a taxi at the maximum extent, is the optimal subsidy scheme.

# 4.2.2 Rationality Verification of the Subsidies Model

To demonstrate the rationality of subsidy scheme model, substitute the different numerical independent variables into the relation function, compared the obtained dependent variable with actual value, to analyze its error to judge the rationality of the subsidy scheme model.

First to verify with data at a competitive stage, the taxi software released 5 Yuan of subsidy in May, 2014, substituted a=5 into equation (2), and obtained the effective mileage X=229.0107km; while the data from " our taxi industry status analysis in 2014"<sup>[7]</sup>, the effective mileage of taxi in Beijing is 205.1 km, both data are consistent.

At the subsidy decrease stage, the Kuaidadi released 3 Yuan of subsidy in January, 2015, and not canceled in March, and substituted a=3 into equation, and obtained the effective mileage X= 199.9707 km, for the average daily total mileage per taxi is not change, is 293km, so the effective utilization of the taxi is 68.25%; while the data from "opinions on strengthen taxi management and improve operational service level" issued in 2015 of Beijing, the taxi utilization of Beijing in March, 2015 is 68%. Both data are consistent, and the error is small.

It can be known from the above verification results, the obtained value from model relation function has small difference with actual value, indicates the subsidy scheme model is much rational.

# 5. Sensitivity Analysis

# 5.1 Sensitivity Analysis of Problem A

In problem A, we selected four indicators of the population density, travel frequency, selected taxi proportion and average waiting time to measure  $\beta$ value, and the population density in these factors, we regarded as only changes with time and no changes over space, but in actual situation, the population density varies with time, we set variation as  $\Delta x_1$ , that is the population density  $x_1 = x_1 + \Delta x_1$ , and substituted into the regression equation:

 $y = 1.399 - 0.05 \times x_1 - 0.145 \times x_2 - 0.291 \times x_3 - 0.012 \times x_4$ 

Then we obtain a new regression equation:

 $y = 1.399 - 0.05 \times x_1 - 0.05 \times \Delta x_1 - 0.145 \times x_2 - 0.291 \times x_3 - 0.012 \times x_4$ 

Due the changes of population density is not too large,  $x_1 \gg \Delta x_1$ , while their weight coefficient only as 0.05, we select the population density of Beijing time 22:00 at Haidian district, assumed  $\lambda x_1 = 700$ , after normalization  $\Delta x_1 = 0.03$ , substituted into regression equation, and found y deviates about 0.00015 finally, and can be almost ignored, so the sensitivity of population density is not high.

# 5.2 Sensitivity Analysis of Problem B

In model of problem B, we select two variables of fixed growth rate r and the travel proportion K for sensitivity analysis, we let r, K changes small, and variable quantity is  $\Delta r$ ,  $\Delta K$ , respectively, the changed r=r+ $\Delta r$ , K=K+ $\Delta K$ , we take  $\Delta r$ =0.04,  $\Delta K$ =0.5%, so the new data is r=0.3, K=7%, and substituted into equation for calculation respectively, reflected by matlab image:



The influence of the changed functional value of K is larger than that of changed r value, thus we consider the sensitivity of K value is higher than r value.

# 6. Assessment and Promotion of the Model

### 6.1 Model Advantages

1. On the analysis of the fitting degree of regression model, using Pearson correlation coefficient to analyze the correlation of predicted value and actual value which were analyzed, and more accurately measure its relevance, thus provide convincing basis to assess the fitting degree of the model;

2. Based on the analysis of the company whether a taxi subsidy scheme can alleviate difficulty of taking a taxi, the established model fully considers the real social development and change of the market, and conducts respective modeling and analysis from three stages;

3. In the establishment of subsidy scheme model, the added correct factors to the right side of Module of Preventing Increase, vividly shows the model contained self-inhibition condition, can accurately express the relationship between the coverage rate and subsidies.

# 6.2 Model Disadvantages

The collected data in the question are all taxi data from Kuaidadi Internet platform, and a few taxies without join in Internet platform, there may be some errors.

2. There are a lot of indicators influencing the demand of taxi, only considering the operational indicators with strong correlation and easy access from data in modeling, so there is a certain limitation in application.

### **6.3 Model Improvement**

In problem A, in addition to consider the operational indicators with strong correlation and easy access from data when modeling, and other indicators of taxi demand should also be considered, such as passengers' satisfaction, etc..

In problem B, we only analyze a taxi subsidy scheme with the given subsidies, we can also establish subsidy scheme from different aspects, such as more specific subsidy scheme for accurate evaluation on different groups, different waiting time and so on.

# 6.4. Model Promotion

Subsidy scheme model in problem 2 is the determined best solution at a specific time in a particular area, to analyze the relations between the effective mileage and subsidies, while the subsidy scheme model can by changing the parameters in the model, analyze the relations of effective mileage and residents' consumption level, travel habits, the consumers, etc.

# References

- [1] Wenquan Zhang, analysis of factors influencing urban taxi demand and supply relation, Journal of Hebei Traffic Vocational and Technical College, (**2011**) 8 (1) :51-53.
- [2] Kuaidadi smart travel platform, http://v.kuaidadi.com/, September 11, (2015).
- [3] Xiaomei Feng, Taxi development scale study under supply and demand balance condition, Southwest Jiaotong University, (**2010**)
- [4] YiGuan analysis: The fourth quarter of 2014 Chinese taxi APP market pattern is set focus to the car market: http://www.enfodesk.com/SMinisite/newinfo/articledetail-id-418731. html, September 11, (2015)
- [5] Chinese Internet data information centre, Beijing traffic development research center, a taxi software makes the taxi empty rate fell by 1%:http://www. 199it.com /archives/ 322771.html, September 12, (2015).
- [6] Taxi software market analysis report of 2013-2014 of Sutu institute.http://www. sootoo. com/content/480044.shtml, September 12, (2015).
- [7] Taxi industry status analysis in 2014 of our country: http://www. chinabgao.com/k/ chuzuche/ 2067. html, September 12, (**2015**).

# Authors

Xiaodong XU. The student of North China Electric Power University

