

## PROCEEDINGS PAPER

# How to Find a Comfortable Bus Route - Towards Personalized Information Recommendation Services

Danhui Guo<sup>1</sup>, Ziqi Zhao<sup>2</sup>, Wei Xu<sup>2</sup>, Jinsong Lan<sup>1</sup>, Tao Zhang<sup>2</sup>, Shuguang Liu<sup>3</sup>, Jianhui Li<sup>1</sup> and Yuanchun Zhou<sup>1</sup>

<sup>1</sup> Computer Network Information Center, Chinese Academy of Sciences, Beijing, 100190, China  
guodanhui@cnic.cn, lan.jinsong@qq.com, lijh@cnic.cn, zyc@cnic.cn

<sup>2</sup> School of Information, Renmin University of China, Beijing, 100872, China  
zhaoziqi@ruc.edu.cn, weixu@ruc.edu.cn, zt2011202547@ruc.edu.cn

<sup>3</sup> Chongqing Institute of Green and Intelligent Technology, Chinese Academy of Sciences, Chongqing, 400714, China  
liushuguang@cigit.ac.cn

With the generation of massive data from bus IC cards, how to effectively and efficiently recommend comfortable bus routes to bus passengers is a challenging and complex task. In this paper, waiting time, crowded time, and driving time between different bus stations on different bus routes at different times of the day are calculated from bus IC cards data history. Then, a multi-objective program with various constraints is suggested to recommend comfortable bus routes for bus passengers, and a genetic algorithm is developed to search for expected solutions. The proposed method is implemented using bus IC cards data from Chongqing, China and will be a promising tool for bus passengers when choosing comfortable bus routes.

**Keywords:** Online information services; Bus routes; Recommendation; Multi-objective programming; Genetic algorithm

## 1 Introduction

With rapid economic growth, city sizes are expanding, and city traffic flow is surging. Thus traffic congestion, traffic accidents, and serious air pollution are becoming more common. Priority has been given to the development of urban public transportation policy by using GPS, GIS, the internet, and communication technology to realize the collection, transmission, storage, and processing of the massive historical and real-time data gathered from bus IC cards, in order to build an intelligent, modern city public transport dispatching platform that will solve social problems such as urban traffic congestion, energy shortages, and air pollution. In addition, when passengers choose a reasonable travel plan, they can lessen not only their travel costs but can also increase the efficiency of public transport vehicles. Therefore, the choice of a reasonable bus route is of great importance to the daily operation and management of urban traffic.

In recent years both in China and abroad, urban public transport has been the subject of extensive research and discussion. However, there are still many complaints about crowded buses, the length of waiting time, and the unpunctuality of the buses. How to effectively and efficiently recommend comfortable bus routes to bus passengers is a challenging and complex task. To address this critical challenge, we must consider factors reflecting the passengers' demands such as waiting time, crowded time, and driving time. To recommend comfortable bus routes for bus passengers, we suggest using multi-objective programming with various constraints and have developed a genetic algorithm to search for solutions. As a result, a bus route according to the differing requirements of passengers can be recommended.

The rest of the paper is organized as follows. In Section 2, we review related work on the data processing of bus IC cards and personalized information recommendation services. In Section 3, we present a multi-objective program with various constraints to recommend comfortable bus routes for bus passengers and use a genetic algorithm to search for solutions. In Section 4, we discuss our implementation and empirical analysis of the proposed method. Finally, we conclude the paper and point out future research directions.

## 2 Literature review

In recent years, much attention has been paid to programming city bus routes and information recommendation services. Most studies focus on bus scheduling and planning (Cheng, 2010). Bo et al. (2002) presented an optimal method of peak curve according to the Fisher algorithm of serial specimen clustering and set a random optimum model by the theory of random service system. Ge et al. (2005) proposed an improved discrete method to produce a bus timetable that takes both the cost of dispatching a vehicle and the value of passengers' waiting time into consideration under the constraint of the vehicle capacity. Avishai (1984) put forward four methods of passenger flow statistics that are used to determine the departure time and intervals between buses. Haghani and Chiang (2003) took bus lines as a network and completed the route of the journey for every passenger to manage the vehicles centrally. The goal was to make the passengers' waiting time shorter and the cost of vehicles lower under the conditions of limited parameters.

Moreover, a group of researchers focused on how to plan bus routes from both the passengers' and bus companies' perspectives. Chen et al. (2004) analyzed urban bus dispatching in terms of service, quality, and efficiency and built a bus service frequency optimal model with a multi-target optimization model to fit passengers' satisfaction as well as the benefit of bus corporations. Li et al. (2009) set up a bus scheduling optimization model with bus revenue as the object function and chose passengers' satisfaction with waiting time and congestion using a simulated annealing evolutionary algorithm based on a self-adaptive operator.

In terms of data collection and data analysis, both domestic and overseas scholars have made great progress. Dai et al. (2006) have done research on an analysis of passengers' travel information and traffic statistics through the bus IC card records of passengers. Bagchi (2004) examined the impact of smart-card data in relation to two case studies: one was concerned with the impact on the data collection process and the other explored the impact on travel behavior.

Meanwhile, several works exist that discuss personalized information recommendation services, which have been applied into various contexts such as job search (Al-Otaibi, 2012; Malinowski, 2006; Paparrizos, 2011), expert finding (Cao, 2005; Fang, 2007; Macdonald, 2008), travel planning (Park, 2008; Staab, 2002; Yu, 2009), and restaurant recommendation (Lee, 2006; Liang, 2008; Mui, 2001). For example, Al-Otaibi and Ykhelf (2012) built recommender systems to provide job information for job hunters. Balog et al. (2009) proposed language models to find appropriate experts to help users solve tricky problems. In tourism, a personalized route recommendation service was designed for theme parks, and RFID information and tourist behaviors were explored to support decisions (Park, 2008). In recent years, online information service has become more and more personalized. As a representative of applications in the field of online maps, a real-time bus route can be easily and quickly discovered on Google maps. In individual cities such as Beijing and Xiamen, citizens can get information about the location and arrival time for buses through a mobile phone app.

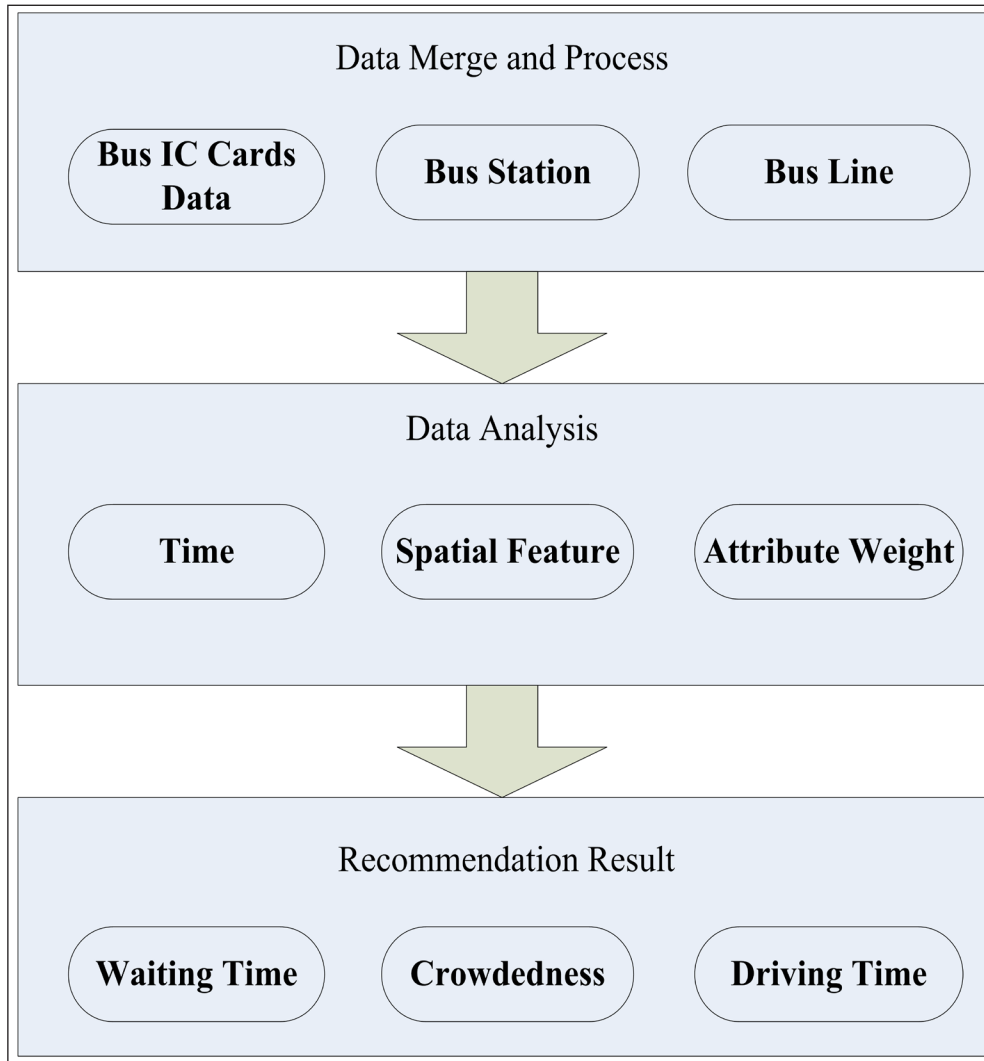
However, most of this research and these applications focus on convenience when choosing a bus route while few are concerned with the information recommendation services. To cover the research gap discussed above, a multi-objective program with various constraints is suggested to recommend comfortable bus routes for bus passengers, and a genetic algorithm is developed to search for the expected solutions.

## 3 The proposed approach

Based on the existing research results in Section 2, this section will present the proposed approach. It is composed of three steps, which are shown in **Figure 1**. First, data from IC cards are merged and processed. The information about bus stations and bus lines is integrated with the IC card data. Second, the average driving time between different bus stations by different bus lines is computed. The number of passengers on a bus can also be estimated although only the number of passengers who get on the bus is available. Finally, to make a personalized recommendation, the preference of the individual passenger must be taken into account by allowing passengers to give different weights to different factors. Through the model proposed, bus lines have integrated waiting time, crowdedness, and driving time according to passengers' needs. The method provides a tool for bus passengers when choosing comfortable bus routes.

### 3.1 Data merge and process

Because the data of bus IC cards is massive, some indicators need to be further processed. The following are the calculation methods.



**Figure 1:** Procedure of the proposed approach.

### 3.1.1 Calculating at which bus station to get off

Bus passengers don't swipe cards when getting off the bus in several cities. Therefore, we cannot directly find out the number of passengers getting off a bus at any bus station. The original method used to discover this number can be described like this. If a passenger swiped an IC card at bus station A and then swiped an IC card at bus station B and bus stations A and B belong to the same bus route, where the passenger got off can be inferred. However, the above method may not work when the same card is reused. Then the attraction rate of the bus station can solve the problem. Attraction rate reflects the attractive intensity of a bus station to passengers. Considering the symmetry of passenger flow from up-link and down-link, the rule can be deduced.

### 3.1.2 Calculate the times when a bus is crowded

Mathematically, a bus's crowded time is formulated as

$$c_{ij} = \begin{cases} 0 & (0 < h_i \leq s_1) \\ d_{ij} & (s_1 < h_i \leq s_2) \\ d_{ij} * (h_i / s_2) & (s_2 < h_i) \end{cases} \quad (1)$$

where  $d_{ij}$  is the driving time from bus station  $i$  to bus station  $j$ ;  $c_{ij}$  is the crowded time from bus station  $i$  to bus station  $j$ ;  $h_i$  is the number of passengers from the bus station  $i$ ;  $s_1$  is the first threshold value of number of crowded passengers on the bus.  $s_2$  is the second threshold value of the number of passengers on the

crowded bus. From Eq. (1), we can deduce that the more passengers on the bus, the more complaints will be made by the passengers.

### 3.2 Building the model

First, the decision variables and parameters to be used in the model are defined as follows:  $x_{ij}^{pq}$  is the decision variable if the bus route  $p$  is recommended from bus station  $i$  to bus station  $j$  at time period  $q$  then  $x_{ij}^{pq} = 1$ ; otherwise  $x_{ij}^{pq} = 0$ ;  $S$  is the total number of bus stations,  $Q$  is the total number of time periods and  $B$  is the total number of bus routes;  $w_{ij}^{pq}$ ,  $d_{ij}^{pq}$ ,  $c_{ij}^{pq}$  stand for the waiting time, driving time, and crowded time from bus station  $i$  to bus station  $j$  at time period  $q$ ;  $w_{\max}$ ,  $d_{\max}$ ,  $c_{\max}$  are the maximum values of  $w_{ij}^{pq}$ ,  $d_{ij}^{pq}$ ,  $c_{ij}^{pq}$ ;  $T_1$ ,  $T_2$ ,  $T_3$  are three objectives standing for waiting time, driving time, and crowded time.

Then, a model is built as follows:

$$\min T_1 = \sum_{p=1}^B w_{ij}^{pq} x_{ij}^{pq} \quad (2)$$

$$\min T_2 = \sum_{p=1}^B d_{ij}^{pq} x_{ij}^{pq} \quad (3)$$

$$\min T_3 = \sum_{p=1}^B c_{ij}^{pq} x_{ij}^{pq} \quad (4)$$

subject to

$$0 < w_{ij}^{pq} < w_{\max} \quad (5)$$

$$0 < d_{ij}^{pq} < d_{\max} \quad (6)$$

$$0 < c_{ij}^{pq} < c_{\max} \quad (7)$$

$$\sum_{p=1}^B x_{ij}^{pq} = 1 \text{ for } i, j = 1, 2, \dots, S \quad (8)$$

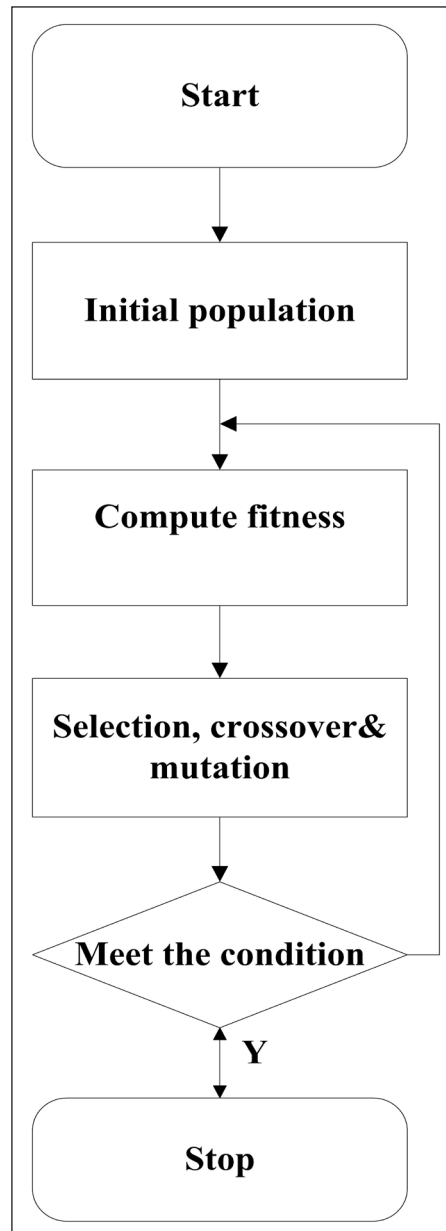
$$x_{ij}^{pq} \in \{0, 1\} \text{ for } q = 1, 2, \dots, Q \text{ and } p = 1, 2, \dots, B \quad (9)$$

In the above model, there are three objectives. Objective function (2) minimizes the waiting time, objective function (3) minimizes the driving time, and objective function (4) minimizes the crowded time. Aside from the objectives, constraint functions (5), (6), (7) guarantee that waiting time, driving time, and crowded time are within certain limits. Constraint function (8) guarantees that a comfortable bus route between any two bus stations can be recommended.

In order to solve the model more conveniently, we next transform a multi-objective programming problem to a single objective programming problem as follows:

$$\min \text{score} = \sum_{s=1}^3 \omega_s T_s \quad (10)$$

The score is the general performance of the bus route, taking  $T_1$ ,  $T_2$ ,  $T_3$  into account.  $\omega_s$  is the weight of the object.  $\omega_s$  is assigned between 0 and 1. The sum of  $\omega_s$  is 1. In a real world application, the weights totally depend on decisions made by passengers. They can choose the significance among different factors and offer different weights for these factors. As a result, the recommended bus lines may change as the weights of different factors change.



**Figure 2:** The main procedure of GA.

#### 4 The proposed algorithm for solving the model

There are many different methods used to solve problems with multiple objectives (Steuer, 1992; Vassilew, 1993). The most common one is the linear weighting method (Hwang, 1981; Yu, 1985). In this, multiple objectives are transferred into an ordinary linear program by adding them with different weights. Although it is easy to understand, it is still a little bit subjective because of its man-made weight (Deb, 2001). Because more and more researchers are focusing on multi-objective optimization, there are many effective methods to use, in particular the multi-objective genetic algorithm (GA). Multi-objective GA has almost a similar main procedure as GA. The GA solution process to solve the model includes computing fitness, selection, crossover, mutation, and stopping criteria.

The main GA procedure is shown in **Figure 2**.

##### 4.1 Initial population

In this part, before the initialization, we do encoding for the chromosome that is a specific solution vector, usually recognized as  $x$  in the solution set  $X$ . Genes, which always represent certain features, are distributed in a chromosome. Encoding is designed for matching the solution set and the chromosomes (Konak, 2006).

A great number of encoding schemes have been used due to the features and identities of problems (Fan, 2011). Binary encoding and real encoding are the most familiar techniques. Encoding takes a vital role in GA programming as it determines the mode and efficiency of the algorithm operation. Deciding the population size and initial population are the next step. Aside from special demands, the population is generally randomly initialized.

#### 4.2 Computing fitness

The fitness function is calculated according to its objective function and its constraint condition, in order to choose the fitter chromosomes to be parents in the next step, selection. The smaller the value of the fitness function and the higher the rank of the chromosome, the better the resulting vector is.

#### 4.3 Selection

As mentioned above, chromosomes are ranked with a fitness value. There are several methods with which to select fitter ones to be the parents, such as roulette wheel selection, stochastic universal sampling, local selection, and tournament selection. For example, the roulette wheel is based on a randomly generated percentage. Each chromosome has a probability based on its fitness value (Fan, 2011).

#### 4.4 Crossover & mutation

In the crossover step, parents swap some of their genes with each other and generate new offspring (Mirazavi, 2001). The resulting mutation increases the diversity of the genes and chromosomes by changing the genetic value with a given probability (Fan, 2011). There are several crossover and mutation methods. The easiest one is to employ the existing function in a GA toolbox, such as the *gatbx*, which is made by Sheffield University in the UK.

In our model using the multi-objective GA, we choose the parallelism selection genetic algorithm to solve the model. The parallelism selection genetic algorithm is a typical measure used in multi-objective problems, which need to divide the population into several groups to solve the respective objectives. The number of groups is determined by the quantity of the objective. Each group has its own fitness function and selection results. We merge all groups into one and implement the crossover and mutation step afterwards.

Considering that the three objectives in the model have different scales and dimensions, we normalize them using objective attainment (Chen, 2008). The definitions of the three are as follows:

$$u(T_1) = 1 - \frac{T_1^{\max} - T_1}{T_1^{\max} - T_1^{\min}} \quad (11)$$

$$u(T_2) = 1 - \frac{T_2^{\max} - T_2}{T_2^{\max} - T_2^{\min}} \quad (12)$$

$$u(T_3) = 1 - \frac{T_3^{\max} - T_3}{T_3^{\max} - T_3^{\min}} \quad (13)$$

where the values of  $u(T_1)$ ,  $u(T_2)$  and  $u(T_3)$  are between 0 and 1.

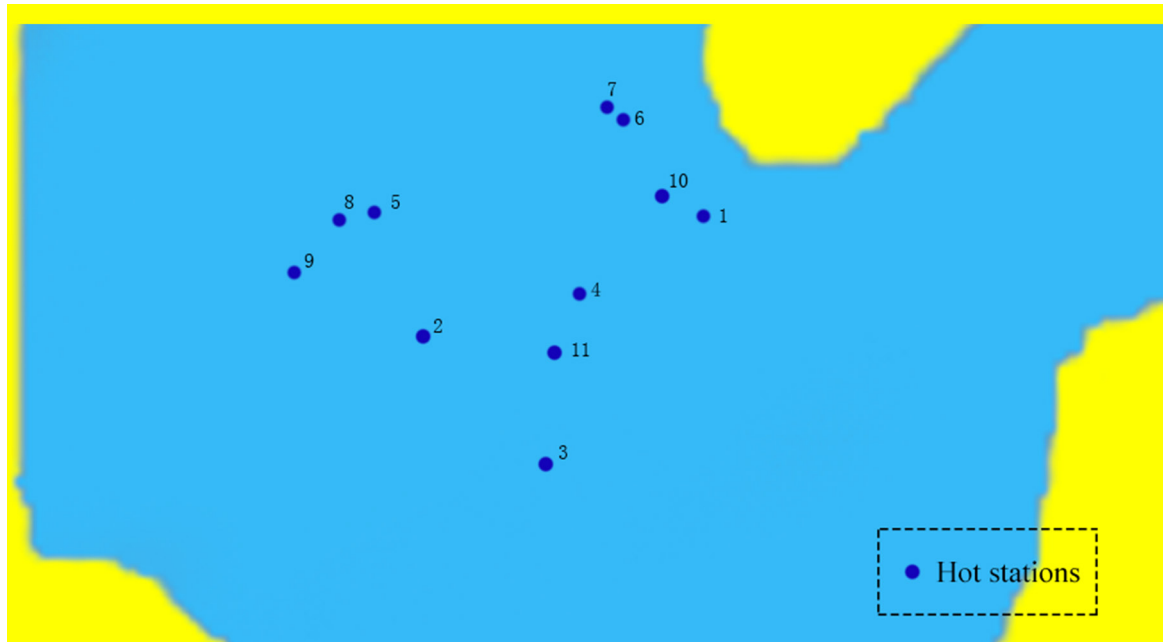
We encode the chromosome just like the encoding methods in the 0–1 integer programming problem. Each gene represents a decision variable that either chooses a bus route or not. The binary code is necessary and suitable here. The notation for chromosome representation is the following:

$$x_1 x_2 x_3 \dots x_{m-1} x_m$$

where  $m$  is the number of decision variables (alternative routes). One chromosome indicates a one solution scheme.

#### 4.5 The proposed GA algorithm

On the basis of the model, the fitness function considers the constraints mentioned above. The constraint is that there is only one alternative route that can be the final recommendation result. The pseudo-code of our fitness function is as follows:



**Figure 3:** The geographical distribution of hot stations in Chongqing, China.

```

while each row
  if(constraint)
    result=inner product of decision variables & coefficient
    (refer to waiting time, driving time...)
  else
    result=max value
  end if
end

```

As for the selection, crossover, and mutation methods, we adopt the existing function in gaptx with a population size of 100, generation gap of 0.9, maximum number of generations of 100, and a crossover rate of 0.7.

## 5 Empirical analysis

### 5.1 Data description and evaluation criteria

The data resources are the bus IC card records for three days in June 2012 in Chongqing, China. In order to achieve the visualization experimental results, the experiment used records in the period from 8:00–9:00, which is in the morning rush hour. According to the bus IC card records, bus stations with high passenger flow volume are considered as hot stations. The geographical distribution of hot stations in Chongqing, China is shown in **Figure 3**.

According to the distribution of hot stations, 20 hot lines between two hot stations are chosen randomly for the experiment.

### 5.2 Experimental results

The three objects of waiting time, driving time, and crowded time are given equal weight. Using the data provided and the model above, we finally arrive at the empirical results in **Table 1**. In the meanwhile, computing time for each recommended result between the original bus stop and the destination is less than one second.

To do the comparison, we search on Google (<https://www.google.com/maps/>) to compare the recommendation results. Because the data for the experiments were produced in 2012 and the search results were done in 2014, bus line changes might have occurred, leading to inconformity (**Table 2**).

To evaluate the degree of satisfaction for the recommendation results, the questionnaire presented in the appendix was designed for passengers to complete (**Table 3**). In the questionnaire, the recommendation results from the proposed method and Google Map are randomly shown for a fair comparison. The final

Origin	Destination	The alternative bus routes	The most comfortable route	Score of the best route
1	2	128,118,815	128	1.758934
3	2	341,823,148,263	823	1.108361
1	3	416	416	2.875463
4	3	403,416,463,473,471	403	1.047812
5	2	454,231,148	148	1.705109
3	5	454,148	148	1.151822
6	2	128,118,107,815	815	1.198029
4	6	818,128,118,815	815	1.366495
5	6	113,181,863	181	1.967453
7	1	132,118,605,181,114,872	114	1.171361
2	7	139,118,106,815,148	815	1.479173
5	7	113,821,270,181,202,863,201,148,801	863	0.924748
8	6	113	113	2.280954
7	8	113,270,202,201	201	1.684108
9	2	467,464,209,472	464	1.289752
5	9	133,270,166,236,807	236	1.078292
10	4	138,818,128,416,118	128	0.937499
11	1	818,416,411	818	1.194276
2	11	839,873	839	1.568126
10	11	818,416	416	1.578496

**Table 1:** The recommended comfortable bus route results.

Origin	Destination	The alternative bus route	The fastest route
1	2	815,118,421,138	815
3	2	823,469,148,454	823
1	3	411,412,416	411
4	3	403,476,471	403
5	2	231,148,875	231
3	5	148,454,504	148
6	2	128,118	128
4	6	818,815	818
5	6	863,201,270,202	863
7	1	118,132,112	118
2	7	202,815	202
5	7	201,863,202,270	201
8	6	863	863
7	8	201,113,163,270	201
9	2	467,464,819	467
5	9	807,166,464	807
10	4	416,412,411	416
11	1	411,416,606,873	411
2	11	231,839,429,836	231
10	11	606,416,412	606

**Table 2:** Google Map search results.



results show that the performance of our proposed method outperforms the one using Google Map in terms of waiting time, driving time, crowdedness, and other factors that influence bus route selection.

## 6 Conclusion

This paper proposes an approach for the problem of how to effectively and efficiently recommend comfortable bus routes to bus passengers. In this approach, multi-objective programming with various constraints and a genetic algorithm are utilized to solve the problems. Using the method proposed, waiting time, crowded time, and driving time between different bus stations in different bus routes at different times are calculated using data from bus IC cards. The proposed method was implemented using data from bus IC cards in Chongqing, China. In order to make the result of experiment visual, 20 hot lines between two hot stations were chosen randomly to verify the results. An empirical analysis was carried out. The most comfortable bus route between the starting point and destination was recommended, which integrated the factors of waiting time, crowdedness, and driving time that passengers were most concerned about. In summary, the proposed approach allows bus passengers to effectively and efficiently choose a comfortable bus route according to their own needs.

Some research issues remain for further studies. First, the massive amount of bus IC cards data needs to be processed in a more detailed fashion. More relationships and knowledge may be found by further data analysis. In addition, presenting the results in the form of a mobile app would be more effective. Finally, the method proposed would be improved by reading evaluations done by bus passengers.

## 7 Acknowledgments

This research work was partly supported by the National Natural Science Foundation of China (Grant No. 71001103, 91224008, 91324015, 91324019), the Humanities and Social Sciences Foundation of the Ministry of Education (No. 14YJA630075), the Beijing Natural Science Foundation (No. 9122013), the Beijing Social Science Fund (No. 13JGB035), the Beijing Nova Program (No. Z131101000413058), and the Program for Excellent Talents in Beijing.

## 8 APPENDIX

Number of the Origin	Number of the Destination	The alternative bus routes	Which one do you feel is the fastest?	Which one do you feel is the most comfortable?	Please put the alternatives in order according to the degree of your satisfaction.
1	2	128,118,815,421,138			
3	2	341,823,263,469,148,454			
1	3	411,412,416			
4	3	403,416,463,473, 476,471			
5	2	454,231,148,875			
3	5	454,148 ,504			
6	2	128,118,107,815			
4	6	818,128,118,815			
5	6	113,181,863,201,270,202			
7	1	132,118,605,181,114,872,112			
2	7	139,118,106,815, 202,148			
5	7	113,821,270,181,202,863,201,148,801			
8	6	113 ,863			
7	8	113,270,202,201,163			
9	2	467,464,209,472,819			
5	9	133,270,166,236,807,464			
10	4	138,818,128,416,118,412,411			
11	1	818,416,411,606,873			
2	11	873,231,839,429,836			
10	11	818,416,606, 412			

**Table 3:** The questionnaire to evaluate the degree of satisfaction with recommendation results.

## 9 References


- Al-Otaibi, S. T. & Ykhlef, M. (2012) Job Recommendation Systems for Enhancing Recruitment Process. *The 2012 World Congress in Computer Science, Computer Engineering and Applied Computing*, Las Vegas, Nevada, USA.
- Avishai, C. (1984) Bus frequency determination using passenger count data. *Transportation Research Part A: General* 18(5–6) pp 439–453.
- Bagchi, M. & White, P. (2004) What role for smart-card data from bus systems? *Municipal Engineer* 157(1), pp 39–46. ISSN 0965–0903.
- Balog, K., Azzopardi, L., & de Rijke, M. (2009) A language modeling framework for expert finding. *Information Processing & Management* 45(1), pp 1–19.
- Bo, L. J., Yao, W. P., & Wang, Y. H. (2002) The Optimum Mathematical Model on the Bus Dispatch. *Chinese Journal of Engineering Mathematics* 51, pp 67–74.
- Cao, Y., Liu, J., Bao, S., & Li, H. (2005) Research on Expert Search at Enterprise Track of Trec 2005. *Proc. of TREC*.
- Chen, Q., Niu, X. Q., Chen, X. W., & Wang, W. W. (2004) Bus Service Frequency Optimal Model. *Journal of Highway and Transportation Research and Development* 2, pp 103–105, 108.
- Chen, Y. W., Wang, C. H., & Lin, S. J. (2008) A multi-objective geographic information system for route selection of nuclear waste transport. *Omega* 36, pp 363–372.
- Cheng, G. (2010) *The Bus Scheduling Model Research based on Bus Arrival Time Prediction*. South China University of Technology.
- Dai, X., Chen, X. W., & Li, W. Y. (2006) The application of data mining technique to bus intelligent card data processing. *Computer and Communications* 1, pp 40–42.
- Deb, K. (2001) Multi-objective optimization using evolutionary algorithms. Chichester: John Wiley & Sons.
- Fan, Z. P., Chen, Y., Ma, J., et al. (2011) A hybrid genetic algorithmic approach to the maximally diverse grouping problem. *Journal of the Operational Research Society* 62(1), pp 92–99.
- Fang, H. & Zhai, C. (2007) Probabilistic Models for Expert Finding, *Advances in Information Retrieval*, Springer, pp 418–430.
- Ge, C. & Zhou, J. (2005) An improved discrete method of designing bus timetable. *Modern Transportation Technology* 6, pp 63–66.
- Haghani, A., Banihashemi, M., & Chiang, K. (2003) A comparative analysis of bus transit vehicle scheduling models. *Transportation Research Part B, Methodological* 37(4), pp 301–302.
- Hwang, C. L. & Yoon, K. (1981) *Lecture notes in economics and mathematical systems: Multiple attribute decision making, methods and application, a state of art survey*, New York: Springer.
- Konak, A., Coit, D. W., & Smith, A. E. (2006) Multi-objective optimization using genetic algorithms: A tutorial. *Reliability Engineering & System Safety* 91(9), pp 992–1007.
- Lee, B. H., Kim, H. N., Jung, J. G., & Jo, G. S. (2006) Location-based service with context data for a restaurant recommendation. *Database and Expert Systems Applications*, Berlin Heidelberg: Springer, pp 430–438.
- Li, S. B. & Cao, G. N. (2009) Bus scheduling model based on simulated annealing evolutionary algorithm. *Journal of Shandong Institute of Light Industry (Natural Science Edition)* 2, pp 83–85.
- Liang, T.-P., Yang, Y.-F., Chen, D.-N., & Ku, Y.-C. (2008) A semantic-expansion approach to personalized knowledge recommendation *Decision Support Systems* 45(3), pp 401–412.
- Macdonald, C. & Ounis, I. (2008) Voting Techniques for Expert Search. *Knowledge and Information Systems* 16(3), pp 259–280.
- Malinowski, J., Keim, T., Wendt, O., & Weitzel, T. (2006) Matching people and jobs: A bilateral recommendation approach. *Proceedings of the 39th Annual Hawaii International Conference on System Sciences*, KAUAI, Hawaii, p 137c.
- Mirrazavi, S. K., Jones D. F., & Tamiz, M. (2001) A comparison of genetic and conventional methods for the solution of integer goal programmes. *European Journal of Operational Research* 132(3), pp 594–602.
- Mui, L., Szolovits, P., & Ang, C. (2001) Collaborative sanctioning: Applications in restaurant recommendations based on reputation. *Proceedings of the Fifth International Conference on Autonomous Agents*, Montreal, Canada, pp 118–119.
- Paparrizos, I., Cambazoglu, B. B., & Gionis, A. (2011) Machine learned job recommendation. *Proceedings of the Fifth ACM Conference on Recommender Systems*, Chicago, USA, pp 325–328.
- Park, M. H., Park, H. S., & Cho, S. B. (2008) Restaurant recommendation for group of people in mobile environments using probabilistic multi-criteria decision making. *Computer-Human Interaction*, pp 114–122.

- Staab, S. et al. (2002) Intelligent systems for tourism. *IEEE Intelligent Systems* 17(6), pp 53–64.
- Steuer, R. E. (1992) Manual for the ADBASE multiple objective linear programming package. Athens: University of Georgia.
- Vassilew, V. & Narula, S. C. (1993) A reference direction algorithm for solving multiple objective integer linear programming problems. *Journal of Operational Research Society* 44(12), pp 1201–1209.
- Yu, C. C. & Chang, H. P. (2009) *Personalized location-based recommendation services for tour planning in mobile tourism applications, E-Commerce and Web Technologies*, Berlin Heidelberg: Springer, pp 38–49.
- Yu, P. L. (1985) *Multiple Criteria Decision Making: Concepts, Techniques, and Extensions*. New York: Plenum.

**How to cite this article:** Guo, D, Zhao, Z, Xu, W, Lan, J, Zhang, T, Liu, S, Li, J and Zhou, Y 2015 How to Find a Comfortable Bus Route - Towards Personalized Information Recommendation Services. *Data Science Journal*, 14: 14, pp. 1-11, DOI: <http://dx.doi.org/10.5334/dsj-2015-014>

**Published:** 22 May 2015

**Copyright:** © 2015 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 Unported License (CC-BY 3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/3.0/>.

 *Data Science Journal* is a peer-reviewed open access journal published by Ubiquity Press.

**OPEN ACCESS** 