Bus Dispatching Optimization Based on Genetic Algorithm

JiaMei Wang
College of Transportation Engineering
Tongji University
Shanghai, China
2012wjm@tongji.edu.cn

DongXiu Ou , DeCun Dong , Lun Zhang
College of Transportation Engineering
Tongji University
Shanghai, China
ou.dongxiu@tongji.edu.cn

***Abstract*— In this paper, genetic algorithm is used to optimize the bus dispatching problem, which coordinates with the arrival of the passengers and improves service level by reducing the average passenger waiting time. The arrival distribution of the transfer passenger associates with the former transport modes in hub, and the discrete stochastic arrival distribution can be depicted by simulation. Firstly, the initial scheme should be chosen considering the bus operational schedule and search speed to seek an optimal solution. Then reasonable fitness function is build to select the fitter solution and ameliorated genetic operators--- crossover and mutation are used to generate a second generation population of solution from those selected. Finally, the optimized schedule can be generated by these procedures and an example will be used to analyze the effectiveness of GAs.**

Keywords- GAs; bus dispatching; coordination; optimization

I. INTRODUCTION

Bus fleet dispatching and routing are essential to enhance the level of service, operating efficiency and cost. Vehicle dispatching is a process to find the result for the function of aim under some restriction condition and to make sure the best sequence of behave based on finites resource. So an optimal schedule is characterized by minimal fleet size and waiting time for achieving maximal service. But how to optimize the bus dispatching is more difficult than we thought, because it should be considered with the people's arrival which is a stochastic and dynamic process that the evaluation item (people waiting time) will change with the variety of departure time. Approaches proposed to deal with this complex problem ranges from simple to sophisticated mathematical theories and innovative heuristics.

A modeling frame together with a heuristic based solution procedure has been proposed which focuses on optimizing the level of service for a given number of buses [1]. A procedure for solving real-world large-scale multiple depot vehicle scheduling problems considering the route time constraints has also been proposed [2]. Stochastic disturbances of daily passenger demands that occur in actual operations have also been incorporated within a model to establish a stochastic-demand scheduling model solved using two heuristic algorithms [3]. In consideration of the complexity of vehicle scheduling problem with route and fueling time constraints, a multiple ant colony algorithm (ACA) inspired by the foraging behavior of real colonies of ants has been proposed and used to solve a traveling salesman problem [4]. Applying a multiple response optimization approach based on the response surface methodology within an integrated simulation model built upon two Markov models is proposed for optimizing the dispatch rules of public bus service [5].

A genetic algorithm is a global search heuristic used in computing to find exact or approximate solution to optimization. Developing with the computer simulations of evolution which started as early as in 1954 with the work of Nils Aall Barricelli, the evolution algorithm became more common. Genetic algorithms (GAs) became popular through the work of John Holland in the early 1970s, and particularly his book Adaptation in Natural and Artificial Systems [6].

Genetic algorithm is effective for solving the problem of public traffic vehicle's schedule [7]. In this paper, an ameliorated genetic algorithm is proposed for operational dispatching optimization which considers the passenger waiting time and the coordination with the other transport modes. The framework of implementation procedure is represented in Chapter 2.The content of genetic algorithm consists of mainly 4 parts (details in Chapter 3) for optimizing the schedule. The result of an example is analyzed in Chapter 4 and the conclusion and outlook are expressed in Chapter 5.

II. IMPLEMENTATION FRAMEWORK

A. Genetic algorithm flowchart

The flowchart of GAs describes the process how to search an optimal schedule in Figure 1. Every iteration optimizes the schedule.

The research is supported by the Natural Science Foundation of Shanghai, China grant 10ZR1432200

978-1-4244-5046-6/10/$26.00 ©2010 IEEE

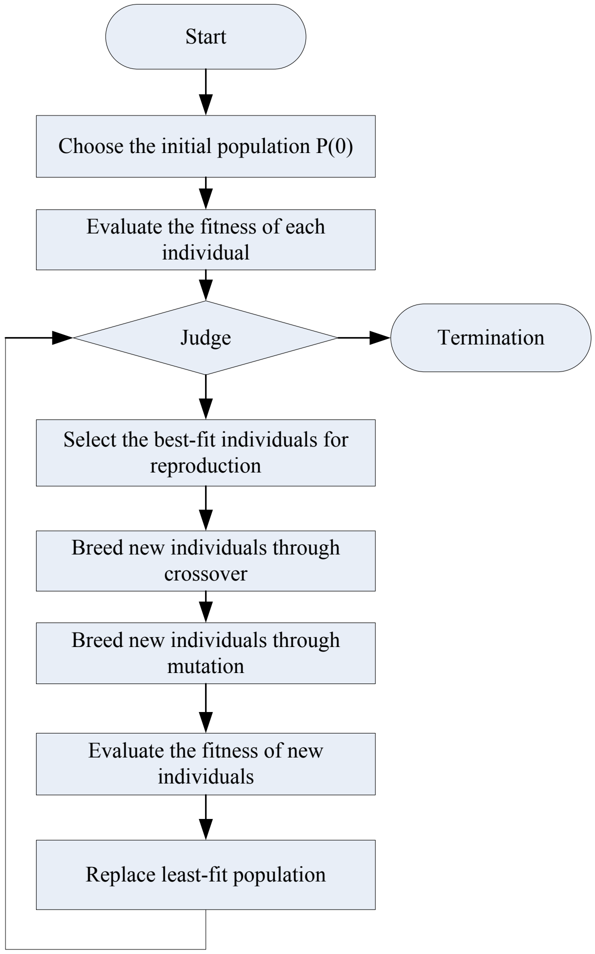

Figure 1. The genetic algorithm flowchart

B. Genetic algorithm assumption

In order to use genetic algorithms to optimize the bus schedule, the assumptions and constraint conditions of the problem must be set for allowing a more true reflection of reality.

Firstly, the interval of the departure time should be constrained in a range considering the limits of the available vehicles number in the station and the requirement of the passengers. Secondly, the numbers of the passengers will not change even if the schedule is changed and the passenger who cannot take the recent bus because that a full load maybe take the next bus. Finally, only the waiting time in the terminus is considered, in this paper, without regard to the passenger flow in the intermediate site in order to simplify the algorithm.

III. THE DESIGN AND REALIZATION OF THE ALGORITHM

A. The arrival distribution of passenger

Bus passengers can be divided into two categories in Hub, some come from the other modes of transport such as rail transport, railway, bus, and others from surrounding areas. Therefore, the distribution of passengers in these bus stations not only contains the randomness but also has some certainty. In other words, it is composed of a random Poisson distribution of arrival passenger flow and a pulse transfer passenger flow from large-capacity transport. In order to describe the distribution of passenger from different directions more accurately, this paper adopts the distribution of passenger through analyzing the result obtained by a simulation system PedSim which simulates the behavior of passengers in Shanghai South Railway Station.

The diagram describes the arrival distribution of passengers from 7:14am to 9:30am on NO.973 bus stops in Figure 2, and the vertical axis is the number of passenger. It shows that the passenger flow approximates to a random Poisson arrival from 7 am to 8 am, and the two peaks appear at around 8:24 am and 9:18 am due to a large number of passenger coming down from the large-capacity transport to transfer to the bus. Analyzing the pedestrian behavior, the arrival time has a great relationship with the transfer time.

Therefore, bus dispatching in a large hub station should take the distribution characteristics of passenger flow into account to determine the start time and start interval to meet the needs of passengers, which can not only shorten the average waiting time of passengers, but also improve the efficiency of bus operations. It is guessed that the bus departure time should be close to the later period of passengers' arrival peak, while the intervals should also fit time that peak sustains.

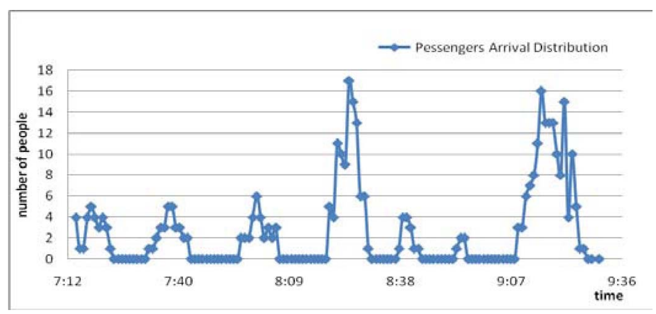

Figure 2. The arrival distribution of passengers taking NO.973 bus

B. Building fitness function

Bus dispatching optimization is a dynamic process, so it is difficult to use static, linear model to describe. Actually, it is an NP-hard problem. Any change of start time of a vehicle will affect the total passenger waiting time as well as the change of time interval. To solve such problem, we should enumerate infinite variety of possible options, and then select one of the best, but it will expend huge time and so inefficient. So it is necessary to find a more efficient way to get the exact or approximate optimization. In this paper, the genetic algorithm, which is a random search algorithm, always seeks the optimal solution, and owns the ability to jump out of local optimal solution, can complete the dispatching optimization with a faster search speed.

The objective of bus coordination dispatching is to reduce the time for passenger to transfer bus. If a bus fleet is in accordance with the optimized schedule, it will achieve higher operation efficiency and service levels. In order to facilitate comparison of experiment, we select the NO.973 bus as an example in Shanghai South Railway Station. NO.973 bus bears the transfer passengers and passengers coming from surrounding, whose terminus is located at the Shanghai South Railway Station transfer Square.

Decision variable (individual) is expressed as $X=[x_1,x_2,\cdots,x_m]$, indicating the period between departure time of the n^{th} bus and the first bus. According to the actual condition, the departure time of the first bus and the last bus can be set, and the total number of runs is m, the total operating time is T_l, the constraint of the optimization problems is as follows:

$$\begin{cases} x_i \in Z \quad and \quad x_i \geq 0, \quad i=1,2,\cdots,m \\ x_1=0, x_m=T_l \\ x_1<x_2<\cdots<x_n<x_{n+1}<\cdots<x_m \end{cases} \tag{1}$$

In this paper, the objective function is the minimum average waiting time of passengers, assuming the arrival distribution of passenger as $f(t)$, in this bus station all passengers waiting time as follows:

$$\begin{aligned} T_w &= \int_0^{x_1} f(t)\cdot(x_1-t)dt+\int_{x_1}^{x_2} f(t)\cdot(x_2-t)dt \\ &+\cdots+\int_{x_n}^{x_{n+1}} f(t)\cdot(x_{n+1}-t)dt+\cdots+\int_{x_{m-1}}^{x_m} f(t)\cdot(x_m-t)dt \end{aligned} \tag{2}$$

The average waiting time is $T_{aw}=\dfrac{T_w}{\int_0^{x_m} f(t)dt}$. Smaller the objective function value is, shorter the passengers waiting time is and the higher service level is. But it is difficult to use the successive distribution to describe the passengers' arrival. So we use the data according to the result of the simulation. In this discrete arrival process, the number of passengers who arrive at t is marked as $N(t)$. The equation (2) changes to (3):

$$\begin{aligned} T_w &= \sum_0^{x_1} N(t)\cdot(x_1-t)+\sum_{x_1}^{x_2} N(t)\cdot(x_2-t) \\ &+\cdots+\sum_{x_n}^{x_{n+1}} N(t)\cdot(x_{n+1}-t)+\cdots+\sum_{x_{m-1}}^{x_m} N(t)\cdot(x_m-t) \\ s.t. \quad & \sum_{x_i}^{x_{i+1}} N(t)\leq C_{load} \qquad i=1,2,\cdots,m-1 \end{aligned} \tag{3}$$

C_{load} is the capacity of the bus. The formula means that the passenger number arriving in every interval should be less than the number the bus can load. Otherwise, the residual waiting time T_r should be added assuming that the residual passengers wait for the next bus when the number is beyond the load in (x_i,x_{i+1}).

$$T_r=\sum_{x_i}^{x_{i+1}} (N(t)-C_{load})\cdot(x_{i+2}-x_{i+1}) \tag{4}$$

The number of passengers does not change when selecting the optimal dispatching scheme, so the total passenger waiting time can be regarded as the objective function.

The fitness is used to measure the each individual of population which is close to the optimal solution in genetic algorithm. If an individual has a high fitness, it will have a higher probability to copy to the next generation. A fitness function is a particular type of objective function that quantifies the optimality of a solution, which is marked as $F(X)$. An ideal fitness function correlates closely with the algorithm's goal, and yet may be computed quickly. To afford facilities for the select operator, fitness is generally non-negative. So formula (5) as follow is regarded as fitness function to seek the optimal solution in this genetic algorithm.

$$F(X)=C_{\max}-T_w \tag{5}$$

$C_{\max}$ is the maximum of the objective function value in the same generation. Thus the waiting time is smaller, the individual fitness is greater.

C. *The initialization of the algorithm*

The genetic algorithm is designed using the "true value encoding", that is, the genes of the chromosome are equal to true values of the decision variable. If the bus departure time is defined as a gene, the chromosome X corresponding with the population $X=[x_1,x_2,\cdots,x_m]$ can be encoded as $x_1,x_2,\cdots,x_m$. The character of "true value encoding" is suitable to optimize the bus dispatching, because good genes can completely be copied to a new generation because the genes of the optimal individual are comparatively good.

Generally, the initial of population $P(0)$ should be random. However, a fixed schedule is be used in regular operation, and its interval is constant. The distribution of the passenger has a large effect to the optimal process of dispatching. When the large capacity transport arrives, the number of passengers reaching the station will increase, and the start time should be appropriately adjusted to meet need of passengers in this peak time. Firstly, an individual of the initial population is encoded according to the actual schedule, and then other individuals should be generated by the adjustment of the encoding according to the transfer passengers in hub. The arrival time of the large capacity transport is set as x_h, then the peak time of the transfer passenger arriving the bus station is x_h+T_c, and T_c is the time passenger changes to the bus. If setting it as a departure time and the interval does not change. The adjustment of the encoding is as follow:

$$x_i{}' = x_i+(x_h+T_c-x_c) \tag{6}$$

x_i' is the value of x_i after the adjustment. x_c is the departure time which is most close to the peak time. The formula is represented that the total departure time has been moved to a direction to fit one peak. Certainly, the other individual can be generated by this means, only selecting the different value of $x_h + T_c$. The initialization discussed ahead can enlarge the probability that the individual contains good genes, while do not contrary to the principle of the dispatching, and play a certain role to improve the search performance.

D. The design of genetic operation

Genetic operator is an operation of chromosome, concluding the selection, crossover, mutation which will produce a "child" solution, typically sharing many of the characteristics of its "parents". New parents are selected for each new child, and the process continues until a new population of solutions of appropriate size is generated.

1) Select operator

Proportional model is used to describe that the probability the individual is selected is in direct proportion to its fitness.

$$P(X_j) = \frac{F(X_j)}{\sum_{j}^{N} F(X_j)} \tag{7}$$

$P(X_j)$ means the probability to select the chromosome X_j, N is the number of chromosome in this generation. If the fitness is larger, the chromosome will be reproduced and a high quality gene will come to next generation.

2) Crossover operator

The measure of crossover is preferential crossover, that is, gives the priority to high quality genes. If there are N copied chromosomes X'_j, each chromosome has m genes x_i, the optimal fitness of the chromosome depends on the character of genes. The idea is that the good genes in the chromosome combine together to optimize the chromosome. From objective function (3), the formula $\sum_{x_{i-1}}^{x_i} N(t)\cdot(x_i - t) / \sum_{x_{i-1}}^{x_i} N(t)$ can be used as standard to scale the quality of gene x_i, and the small value means the less average passenger waiting time. Considering the value of former formula is related to gene x_{i-1}, genes segment can be used to do the crossover operation.

3) Mutation operator

The measure of mutation is uniform mutation. In this paper, the range of any one gene is (x_{i-1}, x_{i+1}), so the new gene in mutation chromosome is represented as:

$$x_i' = x_{i-1} + r(x_{i+1} - x_{i-1}) \tag{8}$$

r is a random number with the uniform probability distribution in(0,-1).

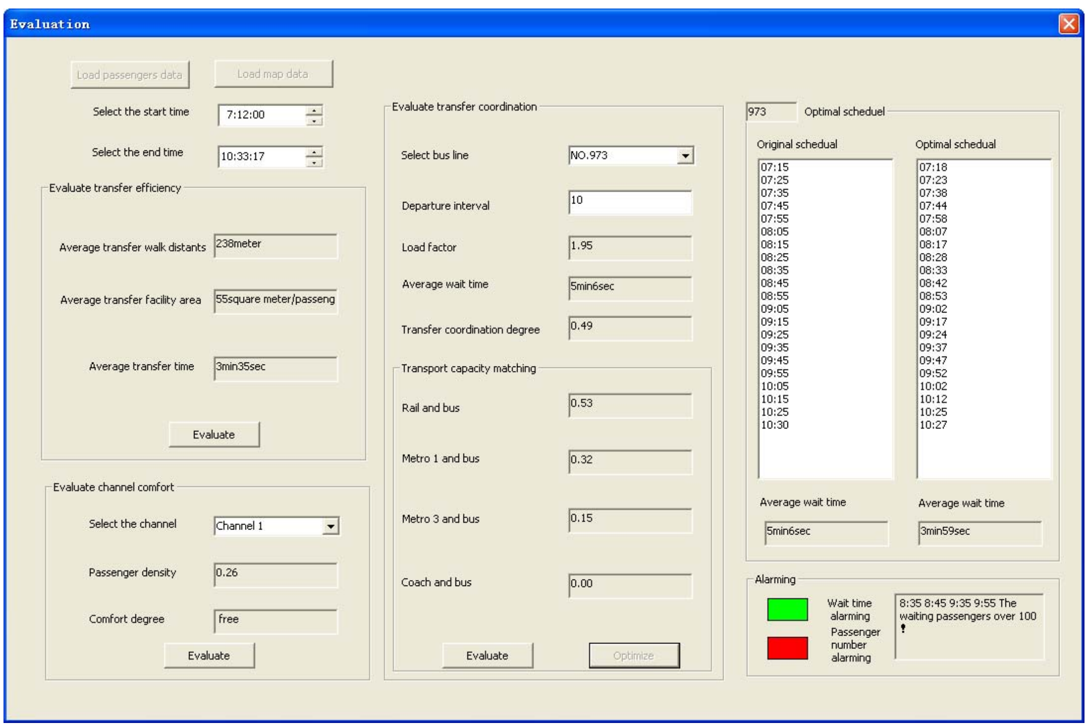

Figure 3. The evaluation interface about the service level and the optimal schedule

IV. THE RESULT OF THE OPTIMIZED SCHEME

The program based on the design of genetic algorithm and the data of transfer passengers in Shanghai South Railway Station are used to verify the algorithm. Figure 3 shows the result of the optimal schedule compared with the original schedule in right part. The average passengers waiting time in optimal schedule is effectively reduced from the original 5 minutes to less than 4 minutes. Analyzing the content of optimal scheme, the above conjecture about the departure time has been confirmed that the interval time is no longer fixed, and the optimal schedule was conducted by adjusting the density and interval, not increasing the operating frequency of the bus, because the fitting transfer passenger traffic can improve the efficiency of the bus operation.

Using the genetic algorithm, the speed of the optimal solution is faster than before. If there are n frequencies need to be adjusted and the adjustment range of every frequency is (a,b) , the complete number of operation is $(b-a)^n$. It is a large number with the exponential growth of n , but the search number of the genetic algorithm is much less than the complete search.

V. CONCLUSION AND OUTLOOK

In summary, it has a good effect to apply genetic algorithm to the bus dispatching optimization, because it can find a reliable approximate optimal solution in the huge search space to solve the dispatching problem. However there are some drawbacks in this paper to make up. The number of departure is fixed, that is, we only change the intervals. The algorithm does not take into account the additional or reduced bus number which is restricted by circumstance, and we only suppose that enough buses in station can be dispatched. In practice, in addition to the minimum average waiting time as the objective function, the other factors (the load factor and cost) also should be considered. The requirement of the passengers in the intermediate site should be considered and the load factor will affect the behavior of the passenger whether they change the vehicle or the route.

In this paper, using genetic algorithm, the bus dispatching optimize intelligently to meet the needs of passenger and to improve the efficiency of bus operation.

REFERENCES

[1] M. Spada, M. Bierlaire and T. Liebling, "Decision-aiding methodology for the school bus routing and scheduling problem", Transportation Science, **39** (4) (2005), pp. 477–490.

[2] M. Banihashemi and A. Haghani, "Optimization model for large-scale bus transit scheduling problems", Transportation Research Record, **1733** (2000), pp. 23–30.

[3] S.Y. Yan, C.J. Chi and C.H. Tang, "Inter-city bus routing and timetable setting under stochastic demands", Transportation Research Part A – Policy and Practice, **40** (7) (2006), pp. 572–586.

[4] H. Wang and J. Shen, "Heuristic approaches for solving transit vehicle scheduling problem with route and fueling time constraints", Applied Mathematics and Computation **190** (2) (2007), pp. 1236–1249.

[5] S.W. Lam, L.C. Tang, T.N. Goh and T. Halim, "Multiresponse optimization of dispatch rules for public bus services", Computers & industrial Engineering, 56(1)(2009), pp.77-86.

[6] H.John, Adaptation in Natural and Artificial Systems, MIT Press, Cambridge, MA,1992

[7] Y.P.Lee,T.An, J.M.Hung, and Y.Z.Fan, "Research on intelligent schedule of public traffic vehicles based on genetic algorithm", Journal of Transportation Systems Engineering and Information Technology , **3**(1)(2003),pp.41-50.