



OPEN

## Bi-objective bus scheduling optimization with passenger perception in mind

Shuai Liu<sup>1</sup>, Lin Liu<sup>1✉</sup>, Dongmei Pei<sup>1</sup> & Jue Wang<sup>2</sup>

With the development of big traffic data, bus schedules should be changed from the traditional "empirical" rough scheduling to "responsive" accurate scheduling to meet the travel needs of passengers. Based on passenger flow distribution, considering passengers' feelings of congestion and waiting time at the station, we establish a Dual-Cost Bus Scheduling Optimization Model (Dual-CBSOM) with the optimization objectives of minimizing bus operation and passenger travel costs. Improving the classical Genetic Algorithm (GA) by adaptively determining the crossover probability and mutation probability of the algorithm. We use an Adaptive Double Probability Genetic Algorithm (A\_DPGA) to solve the Dual-CBSOM. Taking Qingdao city as an example for optimization, the constructed A\_DPGA is compared with the classical GA and Adaptive Genetic Algorithm (AGA). By solving the arithmetic example, we get the optimal solution that can reduce the overall objective function value by 2.3%, improve the bus operation cost by 4.0%, and reduce the passenger travel cost by 6.3%. The conclusions show that the Dual\_CBSOM built can better meet the passenger travel demand, improve passenger travel satisfaction, and reduce the passenger travel cost and waiting for cost. It is demonstrated that the A\_DPGA built in this research has faster convergence and better optimization results.

The public transportation system is a vital social welfare undertaking in a city and a livelihood project that improves people's overall quality of life<sup>1</sup>. Public transit, as a primary form of urban transportation, significantly enhances the quality of urban life<sup>2</sup>. The bus scheduling plan is the essential criterion for bus operation. A scientific bus scheduling plan can effectively adjust social and economic benefits and the personal interests of passengers<sup>3</sup>. With the accelerated development of urbanization and the diversification of travel destinations, traditional scheduling methods based on a fixed period, a fixed departure frequency, and the experience of the workers will result in a waste of bus costs and a longer waiting time for passengers. How balance the revenue of the bus operation department and the efficiency of passenger transit has become an inevitable problem in bus schedules.

There are many studies on bus scheduling optimization at home and abroad. They have established bus scheduling optimization models based on computer vision simulation theory<sup>4</sup>, uncertainty theory<sup>5</sup>, and stochastic simulation theory<sup>6</sup>. The objective setting for the optimization model is divided into two categories: (1) the minimum cost of passengers<sup>4,7</sup>; and (2) the maximum benefit of public transport companies<sup>6,8</sup>. Passengers' travel costs are divided into passengers' waiting for time cost<sup>9</sup>, passengers' traveling for time cost<sup>10</sup>, and passengers' transfer waiting for time cost<sup>7</sup>. The operating costs of the bus company include the travel time of the buses<sup>9</sup>, the travel cost of the buses<sup>5</sup>, and the dynamic remaining capacity of the buses<sup>11</sup>. With the continuous advancement of technology, scholars in recent years have integrated the costs of both passengers and bus companies<sup>5,10</sup>. They make the goal of the model is to minimize the sum of costs for passengers and bus companies<sup>9,12,13</sup>. However, they all ignore passenger congestion when constructing their bus scheduling optimization models, which does not improve passenger satisfaction.

In terms of solving the scheduling optimization model, a variety of methods and algorithms have been proposed to solve the model by adjusting the departure interval, to allow the bus schedule to be reasonably adjusted to suit the passenger's travel purpose. Traditional methods: such as periodic activity scheduling method<sup>14</sup>, Branch & Bound (B&B) methods<sup>15,16</sup>, and standard linear programming method<sup>17</sup>. These methods are only suitable for solving the scheduling models with small data dimensions and require high accuracy of data. When the data dimension increases, a lot of calculation costs will be incurred. At present, heuristic algorithms such as genetic algorithms are mainly used in the field to solve scheduling models. Liu et al. used the classical GA and calculated the optimal departure interval of buses to solve the bus dispatching optimization model<sup>7</sup>. Zuo, Niu, and others

<sup>1</sup>College of Geodesy and Geomatics, Shandong University of Science and Technology, Qingdao 266590, China. <sup>2</sup>QingDao ZhenQing Bus Group Co., Ltd, Qingdao, China. ✉email: liulin2009@126.com

improved the classical GA and proposed an NSGA-II algorithm to solve the bus dispatching model<sup>12,13</sup>. Feng et al. proposed a GA with an "Elitist Preservation" strategy combining the economic method of "dynamic scoring" to solve the multi-objective scheduling model<sup>18</sup>. Niu and Zhang divided a day into several time blocks on average, calculated the scheduling model of each time segment separately, and designed a hybrid GA related to hierarchical crossover and mutation operations to solve the scheduling optimization model<sup>19</sup>.

In summary, bus scheduling optimization is currently based on objective objectives, such as minimizing the traveling cost for passengers or minimizing the operating costs for the bus company. They seldom take into account the soft objective, such as the passenger's riding feeling, and do not take this soft objective as the influence factor of establishing the public transport scheduling optimization model, so they cannot improve the passenger's riding experience. Secondly, when GA is used to solve the model in the existing research, the crossover and mutation probability of the algorithm are generally set as fixed values. The two probabilities cannot be automatically adjusted with the fitness of individuals and the dispersion of the population, which will affect the behavior and performance of the model and directly affect the convergence of the algorithm.

This research aims at the above two defects. First, soft objectives such as passenger perception are introduced based on considering the objective objectives in the public transport scheduling optimization model. The soft objectives are divided into two aspects: passenger perceptions of waiting at stops and passenger congestion perceptions within the bus. We constructed the Dual\_CBSOM model by taking into account the interests of both bus operators and passengers so that the optimization results would increase the revenue of the bus company and improve passenger perception. Second, we improved the defect that the crossover and mutation probability in classical GA are fixed values, and constructed the A\_DPGA algorithm. We designed a new adaptive crossover and mutation operator, and the crossover and mutation probability of the algorithm can be adjusted adaptively with the individual fitness during the optimization of the model.

## Data

**Data sources.** In the process of building and solving the model, it is necessary to know some basic parameters of the target route. These parameters can be calculated by taking the bus and stop surveys, as well as through bus GPS positioning data combined with IC card swiping data. In this paper, we mainly use calculated data supplemented by observed data as the basic data to solve the model. These data are shown in Table 1.

**Analysis of crowded feeling for passengers.** The percentage of passengers in the bus reflects the level of overcrowding inside the bus and is the basis for classifying the level of bus service. The percentage of passengers is the number of passengers in the bus divided by the authorized number of passengers in the bus, and the level of crowding in the bus is also an important factor in whether or not passengers get on the bus at a stop. According to the reference<sup>4</sup>, we classify the crowded feeling of passengers in the car according to the percentage of passengers, and the results are shown in Table 2.

Data classification	Data type	Advantages	Disadvantages
Observed data	Bus arrival and departure time Station and number of passengers getting on and off Waiting time of passengers Location and distribution of bus stations	Accurate, detailed, and realistic	High cost and heavy workload
Calculated data	Probability of passengers getting off Number of passengers in the bus Transfer station of passenger Time of passengers getting on	Low cost and wide coverage	Low accuracy and difficulty in obtaining

**Table 1.** Basic data of the model.

Grade	Percentage of passengers	Feelings of passengers
A	0–25%	Pick a seat at will
B	25–40%	Blandly pick a seat
C	40–50%	Reluctantly pick a seat
D	50–61.7%	Passengers' personal space still exists
E	61.7–75%	passengers have physical contact
F	75–100%	Passengers feel a little crowded
G	> 100%	Crowded and oppressive among passengers

**Table 2.** Classification of the crowded feeling of passengers.

## Establishment of Dual\_CBSOM

**Dual\_CBSOM assumptions.** Model assumptions are made to capture the essence of the problem, ignore secondary factors, and achieve simplification of the actual problem, to facilitate the establishment, solution, analysis, and transplation of the model<sup>20</sup>. The model assumptions not only reflect the actual situation to a certain extent but also save the computational cost of the model.

This research uses bus following survey and online search statistics to obtain the following model assumptions:

- (1) The buses travel at a uniform technical speed, without regard to special circumstances such as traffic accidents;
- (2) The travel rules of 5-day bus IC card users and mobile payment users can represent the travel rules of all bus passengers;
- (3) The time-consuming technical operations such as deceleration when entering the stop, acceleration when leaving the stop, and opening and closing doors are all fixed values;
- (4) The operating cost per unit time of the bus is not affected by factors such as the number of passengers;
- (5) The cost per unit time of bus passengers is the average wage per unit time of local residents.

**Parameters solving.** *Fitting function of passenger flow probability density.* By processing the bus GPS positioning data, bus stop location data, and passenger IC card swiping data, we can calculate the number of people getting on the bus at different times at each stop<sup>21</sup>. Then combined with the bus scheduling schedule for further processing, we can get the number, time, and station of people getting on the bus in each driving direction. Combined with the proportion of the number of people who swipe the card, we can calculate the actual number of people getting on the bus at each stop in a certain driving direction. Finally, according to the calculated data, based on the period to fit the number of people getting on the bus at each stop, then get the passenger flow fitting function  $f(t)$ . Since there are several peak and low periods in a day for passenger travel, this paper uses multiple sine functions to fit the number of passengers, as illustrated in Formula (1).

$$f(t) = \sum_{i=1}^{fn} a_i \sin(b_i t + c_i) \quad (1)$$

For the selection of the order  $fn$  of the above functions, although the larger the  $fn$  is, the better the fitting effect of the scatter points are, but it will inevitably lead to the phenomenon of "over-fitting" and the increase of calculation amount. We set  $fn = 4$  and the function fitting effect is better. Through the fitting function of passenger flow probability density, we can calculate the number of passengers getting on the bus at any time at each station in each direction. If the fitting function of passenger flow probability density in the upward direction of a certain stop is  $f_e(t)$ , and the total number of people getting on the bus in the upward direction of the stop on that day is  $r$ , then the number of people getting on the bus from  $t_0$  to  $t_1$  time is  $x_e$ , it is calculated by Formula (2).

$$x_e = r \int_{t_0}^{t_1} f_e(t) dt \quad (2)$$

If  $t_0$  is the time when the last bus left the stop, and  $t_1$  is the time when the next bus enters the stop, the average waiting time  $t_{wait}$  of passengers at the stop is given by Formula (3).

$$t_{wait} = x_e \times (t_1 - t_0) / 2 \quad (3)$$

*Prediction of the probability of getting off.* The prediction and analysis of passenger flow and passenger flow direction for public transportation is the basis and premise of road traffic planning<sup>22</sup>. The parameters affecting the bus departure time are simplified to be influenced only by the time spent by passengers getting on and off the bus, i.e. the stopping time of a bus at a station is determined by the number of people getting on and off at that station. This requires the dynamic calculation of the number of people getting on and off at the station. The number of people getting on at the station can be calculated using Formula (4), and the number of people getting off at station  $k$ th is the cumulative sum of the number of people who get on the bus at previous  $k - 1$  stations and then get off at the  $k$ th station. Under the premise of knowing the number of people getting on the bus at the previous  $k - 1$  stations, the number of people getting off at the  $k$ th station can be calculated by the attraction weight of the reverse direction station and the Poisson distribution rate of passenger travel<sup>23,24</sup>. The Matrix  $W$  of the probability of getting off the bus can be defined by the attraction weight of the reverse direction station and the passenger trip Poisson distribution, as illustrated in Formula (4).

$$W_{m \times m} = \begin{bmatrix} 0 & w_{1,2} & \cdots & w_{1,m} \\ 0 & 0 & \cdots & w_{2,m} \\ \cdots & \cdots & \ddots & \cdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} \quad (4)$$

where  $m$  is the total number of stations,  $W$  is an upper triangular matrix, the  $k$ th row indicates the probability of passengers getting on the bus at the  $k$ th station and getting off at the subsequent stations, and the  $k$ th column indicates the probability of passengers getting on the bus at the previous  $k - 1$  stations and getting off at the  $k$ th station, e.g. the probability of getting on at the 2nd station and getting off at the  $k$ th station is  $W_{2,k}$ . The number

of people getting off at a station can be calculated dynamically by  $W$ . For example, if the number of passengers getting on at stations 1st ~  $k - 1$  in the upward direction is  $x = [x_1, x_2, \dots, x_{k-1}]$ , then the number of people getting off at station  $k$  can be calculated by Formula (5):

$$y^+ = [x_1 \ x_2 \ \dots \ x_{k-1}] \times [W_{1,k} \ W_{2,k} \ \dots \ W_{k-1,k}]^{-1} \tag{5}$$

*The time when the bus arrives and leaves the stop.* The time  $t_{jz}$  of a certain bus entering the stop and the time  $t_{lz}$  leaving the stop are important parameters for simulating and dynamically calculating the number of people getting on and off at the stop. From model assumption (1): we know that the time  $t_{jz_k}$  when the bus arrives at the  $k$ th stop is the time  $t_{jz_{k-1}}$  when the bus leaves the  $k-1$ th stop plus the time it takes to travel at the technical speed from the  $k - 1$  stop to the  $k$  stop. The time  $t_{lz_k}$  of the bus leaving the stop at the  $k$ th stop is the time  $t_{jz_k}$  that the bus arrives at the stop plus the time required for technical operations such as slowing down when arriving at the stop, accelerating when leaving the stop, opening and closing the doors, and the time it takes for passengers to get on and off the bus.

*Bus travel speeds.* The speed at which buses travel during operation depends largely on the congestion on the road. We classify and discuss the speed of buses according to the congestion in the study area. The results are shown in Table 3.

In Table 3, the average bus speed is 20.5 km/h in the morning peak from 6:30 to 8:30 a.m. and 20.2 km/h in the evening peak from 17:00 to 20:00. The average bus speed is 30 km/h at other times of the day.

*Description of other important parameters.* In order to express the Dual\_CBSOM more conveniently and quickly, we set  $i$  as the serial number of the bus departure and set  $k$  as the serial number of the bus stop. The settings of important parameters in the model are shown in Table 4.

**Dual\_CBSOM building.** For a certain line, the operating cost of the bus company and the travel cost of passengers are directly affected by the bus scheduling plan. During the same period, the more frequent the bus departures, the higher the operating cost of the bus company and the lower the waiting cost of passengers, and vice versa. The objective of this model is to find a scheduling schedule that balances the interests of both the bus company and bus passengers.

*Establishment of the objective function.* We established a bus scheduling optimization model with the first departure time of 6:00 and the last departure time of 22:00 within one day. The optimization objectives are the lowest cost of bus company and the lowest cost of passenger travel, respectively. There are the following three objective functions  $f_1$ ,  $f_{a2}$ , and  $f_{b2}$ :

Time	Speed (km/h)
6:00–6:30	30
6:30–8:30	20.5
8:30–17:00	30
17:00–20:00	20.2
20:00–22:00	30

**Table 3.** Classification of bus speed.

Parameters	Specific forms	Description of parameters
$T$	$[t_1, t_2, \dots]$	Departure schedule optimized by bus schedules, where $t_i$ is the departure time of the $i$ th bus
$S$	$\{T_1, T_2, \dots\}$	This is the population, each individual in the population is a departure timetable $T$
$ST$	$\{ST_1, ST_2, \dots\}$	Populations with binary encoding
$F, P$	$[f_1, f_2, \dots], [p_1, p_2, \dots]$	The fitness and selection probability of each individual in the population
$m$	28	Total number of stops on a line
$X, Y$	$[x_1, x_2, \dots], [y_1, y_2, \dots]$	The number of passengers getting on and off at the stop
$D$	$[d_1, d_2, \dots]$	The actual distance between adjacent stops of a line, the unit is km
$\alpha, \beta$	$\alpha + \beta = 1$	Weight coefficients for converting a dual objective function to a single objective function
$o$	16 s	Time-consuming for the bus to slow down when arriving at the stop, accelerate when leaving the stop, and open and close the doors
$r_k$	$[r_1, r_2, \dots]$	The total number of passengers taking the bus in a certain direction of each stop for 1 day

**Table 4.** Description of important parameters.

- (1) The operating cost of the bus company  
Knowing that a departure schedule is T and from model assumptions (3) and (4), we can calculate the operating costs of the bus company by Formula (6):

$$f_1(T) = a_1 \sum_{i=1}^{Len\_T} (t_i^{end} - t_i^1) \tag{6}$$

In this Formula,  $a_1$  is the bus operating cost per unit time,  $Len\_T$  is the total number of bus departures,  $t_i^{end}$  is the time when the  $i$ th bus arrives at the terminal, and  $t_i^1$  is the departure time of the  $i$ th bus.

- (2) The waiting cost of passengers  
In the operational planning process of public transport, the time a passenger spends on waiting is a very critical element for judging passenger service<sup>25</sup>. Reference<sup>26</sup> pointed out that the psychological feelings of passengers will change with the increase in waiting time. From 0 to 6 min, the passenger's state of mind is stable; from 6 to 15 min, the passenger begins to be anxious; from 15 to 35 min, the passenger's anxiety level increases, and the passenger considers giving up waiting; after 35 min, the passenger gives up waiting and changes the travel mode. To ensure that no passengers change their travel mode, this paper sets the bus departure interval to be less than 20 min, and the passenger's waiting cost per unit time is  $a$ . The waiting cost for different waiting times is shown in Table 5.

The calculation method of passenger waiting cost is shown in Formula (7):

$$f_{a2}(T) = a_2 \sum_{i=1}^{Len\_T} \sum_{k=1}^m (t_{-jz_i^k} - t_{-jz_{i-1}^k})/2 \times r_k \times \int_{t_{-lz_{i-1}^k}}^{t_{-lz_i^k}} f_k(t)dt \tag{7}$$

In the Formula (7),  $a_2$  is the unit time cost of passengers at different waiting times, and  $f_k(t)$  is the probability density fitting function of the passenger flow at stop  $k$ .

- (3) The crowding cost of passengers  
The mood of passengers on the bus will change according to the degree of crowding in the bus, thus changing the cost of passengers per unit time to take the bus. As the level of crowding in the bus increases, passengers feel irritable and perceive time to be slower. According to reference<sup>27</sup> we set the travel cost per unit time for passengers according to the percentage of passengers on the bus. The results are shown in Table 6.

The calculation method of the additional crowded cost for passengers is shown in Formula (8):

$$f_{b2}(T) = (a_3 - a) \sum_{i=1}^{Len\_T} \sum_{k=1}^m (t_{-lz_i^k} - t_{-lz_i^{k-1}}) \times (r_k \times \int_{t_{-lz_i^{k-1}}}^{t_{-lz_i^k}} f_k(t)dt - y) \tag{8}$$

In this Formula,  $a_3$  is the crowded cost per unit time of passengers with different percentages for passengers on the bus, and  $y$  is the number of passengers getting off the bus,  $y$  is calculated according to Formula (5).

The dynamic calculation method of the time when the bus arrives at and leaves the stop is shown in Formula (9):

Waiting time (min)	Waiting cost per unit time (yuan)
0-6	a
6-15	1.5*a
15-35	1.9*a
> 35	2.4*a

**Table 5.** Passenger waiting costs.

Percentage of passengers	Unit time cost (yuan)	Crowding cost ratio (%)
0-40%	a	0
40-50%	1.08* a	7.41
50-61.7%	1.19* a	15.97
61.7-75%	1.28* a	21.88
75-100%	1.45* a	31.03
> 100%	1.57* a	36.31

**Table 6.** Passenger crowded costs.

$$\begin{cases} t_{lz_i^k} = t_{jz_i^k} + \max(F[x_i^k], F[y_i^k]) + o/2 \\ t_{jz_i^k} = t_{lz_i^{k-1}} + d_{k-1}/30 \\ t_{jz_i^1} = t_i^1 \\ x_i^k = r_k \times \int_{t_{lz_i^{k-1}}}^{t_{lz_i^k}} f_e(t) dt \\ y_i^k = \sum_{j=1}^{k-1} x_i^j \times W_{j,k} \\ F[c] = 1.2 \times \text{round}(c + 1) \end{cases} \tag{9}$$

In Formula (9), *round* indicates rounding to the nearest integer, *F* is a defined operator, and the operator indicates the waiting time of the bus at the stop. According to the survey, the average time for each passenger to get on the bus is 1.2 s, and the bus driver needs to wait for the last passenger to get on the bus and sit or support the bus before starting the bus, that is, the stop time of the bus at the stop is the actual number of passengers to get on the bus plus one person's time. For example, *F*[3] means that this bus needs to stop and wait for  $1.2 \times (3 + 1) = 4.8$ s at this stop before leaving for the next station.

*Constraint conditions of the model.*

(1) Minimum stock constraints for buses at stops

In the actual operation of the buses, the buses will wait at the origin stop and the last stop for departure. Assuming that the origin stop is  $A^+$ , there are  $p_1$  buses, the last stop is  $A^-$ , there are  $p_2$  buses, and the driving distance between the two stops is  $dist = \sum dj, j = 1, 2, \dots, m - 1$ . To ensure that at least one bus can be dispatched freely at both the origin stop and last stop, the departure schedule needs to satisfy the following constraint:

$$\sum_{i=2}^{p_1+p_2} (t_i^1 - t_{i-1}^1) \geq 2 \times \frac{dist}{v} \tag{10}$$

(2) The constraint of adjacent departure interval

To fully ensure the utilization efficiency of public resources and protect of the interests of bus passengers, it is mandatory to stipulate the maximum and minimum departure interval of two adjacent buses as illustrated in Formula (11):

$$t_{min} < t_i^1 - t_{i-1}^1 < t_{max} \tag{11}$$

In summary, the bus scheduling optimization model is shown in Formula (12):

$$\begin{aligned} \min & \begin{cases} f_1(T) = a_1 \sum_{i=1}^{Len\_T} (t_i^{end} - t_i) \\ f_2(T) = f_{a2}(T) + f_{b2}(T) \end{cases} \\ s.t & \begin{cases} \sum_{i=2}^{p_1+p_2} (t_i^1 - t_{i-1}^1) \geq 2 \times \frac{dist}{v} \\ t_{min} < t_i^1 - t_{i-1}^1 < t_{max} \end{cases} \end{aligned} \tag{12}$$

**Use A\_DPGA to solve the Dual\_CBSOM**

In the above optimization model, the objective functions are the operating cost  $f_1(T)$  of the bus company and the total cost  $f_2(T)$  of passenger travel, respectively. By introducing the weight coefficients  $\alpha$  and  $\beta$ , we can obtain the overall objective function  $f(T) = \alpha f_1(T) + \beta f_2(T)$  and transform the multi-objective optimization function into a single-objective optimization function. The variable to be optimized in the model is the departure timetable  $T$ , which is composed of discrete vectors to form the solution space, so the optimization problem is a combinatorial optimization problem. Theoretically, the combinatorial optimization problem can be solved by the B&B method developed based on the enumeration method. But in fact, this method cannot be applied to the model solution with a large data scale, and the time it takes to solve the bus scheduling problem using this method cannot be estimated. The GA is widely used in optimization problems with complex data and a large amount of computation. Using GA to solve combinatorial optimization problems can quickly search for acceptable optimal solutions and finally generate global optimal solutions.

Although the classical GA is suitable for solving complex optimization problems, it has slow convergence speed and poor local search ability, which cannot guarantee that the contemporary optimal value will be inherited by the next generation. Classical GA contains parameters such as crossover and mutation probabilities that are determined empirically and cannot be changed during the iterative process of the algorithm. AGA can improve the defects of classical GA, but AGA tends to produce a local optimal solution, and the population tends to fall

into "premature maturity" in the process of solving the model and cannot find the global optimal solution. In this research, the AGA is improved and an A\_DPGA is proposed to solve the Dual\_CBSOM model.

**Proposal of A\_DPGA.** AGA. In the classical GA, the crossover probability and mutation probability of genes are set according to people's experience and these two probabilities remain unchanged during the optimization process. In this paper, the AGA is used to solve the model, and the adaptive adjustment function is introduced, so that the crossover probability  $p_c$  and mutation probability  $p_m$  are automatically adjusted with the individual fitness and the dispersion degree of the population. When the population tends to fall into the local optimal solution,  $p_c$  and  $p_m$  are increased accordingly, and when the population diverges in the solution space,  $p_c$  and  $p_m$  are correspondingly decreased. In AGA, the solution process of  $p_c$  and  $p_m$  is shown in Formulas (13) and (14).

$$p_c = \begin{cases} k_1 \frac{f_{max} - f_1}{f_{max} - f_{avg}}, & f_1 \geq f_{avg} \\ k_2, & f_1 \leq f_{avg} \end{cases} \tag{13}$$

$$p_m = \begin{cases} k_3 \frac{f_{max} - f_2}{f_{max} - f_{avg}}, & f_2 \geq f_{avg} \\ k_4, & f_2 \leq f_{avg} \end{cases} \tag{14}$$

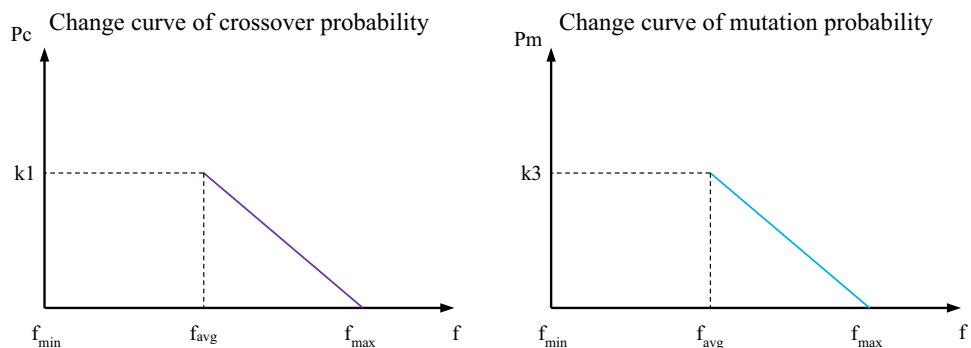
In the Formulas,  $f_{max}$  is the largest fitness value in the group,  $f_{avg}$  is the average fitness value in the group,  $f_1$  is the larger fitness value of the two individuals to be crossed, and  $f_2$  is the fitness value of the individual to be mutated.  $k_1$  and  $k_2$  are crossover probability constants, and  $k_3$  and  $k_4$  are mutation probability constants. The adaptive changes in crossover probability and mutation probability are shown in the following Fig. 1.

The two charts in Fig. 1 show the process of adaptive change of probability with the fitness value of individuals in the population in the maximum optimization problem. The left chart shows the adaptive change curve of crossover probability, and the right chart shows the adaptive change curve of mutation probability.

**A\_DPGA.** Due to the shortcomings of AGA, it is easy to fall into premature convergence and destroy optimal individuals, which reduces AGA's effectiveness. In this paper, we propose an A\_DPGA. The two probabilities should vary gradually at the average fitness  $f_{avg}$  of the population, that is, the two probabilities decrease slowly with the individual's fitness changes. This increases the probability of crossover and variation of individuals at the mean fitness and reduces the likelihood of the population falling into a local optimum solution. Around the maximum fitness  $f_{max}$  of the population, the two adjustment curves should be as smooth as they can be so that the two probabilities gradually approach zero as the individual fitness increases. This enables the better fitness individuals to be retained as much as possible in order for the results to converge more quickly to the global optimum solution. The two probabilities of IDPGA are calculated as shown in Formula (15) and (16).

$$p_c = \begin{cases} \frac{p_{c\ max} - p_{c\ min}}{1 + \exp(\frac{2(f' - f_{avg})}{f_{max} - f_{avg}} - 1)} + p_{c\ min}, & f' \geq f_{avg} \\ p_{c\ max}, & f' < f_{avg} \end{cases} \tag{15}$$

$$p_m = \begin{cases} \frac{p_{m\ max} - p_{m\ min}}{1 + \exp(\frac{2(f - f_{avg})}{f_{max} - f_{avg}} - 1)} + p_{m\ min}, & f \geq f_{avg} \\ p_{m\ max}, & f < f_{avg} \end{cases} \tag{16}$$



**Figure 1.** Adaptive change of crossover probability and mutation probability at AGA.



Formulas (15) and (16) are the adaptive change process of crossover probability and mutation probability in the IDPGA. The population's maximum and minimum crossover probabilities are known as  $p_{cmax}$  and  $p_{cmin}$ , respectively. The population's maximum and minimum mutation probabilities are known as  $p_{mmax}$  and  $p_{mmin}$ , respectively. Figure 2 displays the changes in the two probabilities.

The left and right charts in Fig. 2 are the adaptive change curves of crossover and mutation probability in the IDPGA, respectively. At the average fitness  $f_{avg}$  of the population, the two probabilities change slowly. At the fitness close to the maximum fitness  $f_{max}$  of the population, the two adjustment curves become smooth.

**Solve the Dual\_CBSOM.** A\_DPGA is an iterative calculation process. Its basic process includes initial population generation, encoding–decoding, fitness calculation, individual selection probability allocation, selection, crossover, mutation, termination condition judgment, and other processes.

Assuming that the total number of individuals in the initial population is  $sn$ , the initial population can be obtained by generating  $sn$  individuals. Define the departure interval in minutes. For the generation of each timetable  $T$ ,  $t_1 = 6 : 00$ ,  $t_{end} = 22 : 00$ , and constraints (1) and (2) must be satisfied.

**Encoding–decoding.** We sequentially code the Individuals for the entire population. The coding principles are as follows: first, generate 960 elements coded as 0. For each  $t_i$  in  $T$ , if the time  $t_i$  is  $x$  minutes away from 6:00, code the  $x$ th 0 elements as 1.

The decoding principle is as follows: in the 960 binary codes, find the position of code 1 in turn, then the position of code 1 found this time can be converted into the departure time of this time. For example, if the position of code 1 is found to be 212 bits from the starting point, and it is 212 min away from 6:00, it can be decoded to obtain  $t = 9 : 32$ .

**Fitness calculation.** In the GA, we can evaluate each individual in the population by the fitness function and select the individuals entering the next generation in the population by the fitness value of the individual, so the selection of the fitness function is very important. According to the DCBSOM, the optimization objective is to minimize  $f(T) = \alpha f_1(T) + \beta f_2(T)$ . Then define the fitness function  $Fit(f(T))$ :

$$Fit(f(T)) = 1/f(T) \quad (17)$$

The fitness value  $Fit_j$  of each individual  $j$  can be obtained by calculation.

**Allocation and selection.** The selection operation refers to the selection of individuals from the population according to a certain selection probability, and the selected "excellent" individuals with high fitness are used as parents to reproduce the next generation of individuals. The commonly used selection methods include roulette selection, random traversal sampling, local selection, truncated selection, competition selection, etc. In this paper, the roulette selection method combined with the optimal preservation strategy is used for the selection operation, and the probability of each individual being selected is as shown in Formula (18):

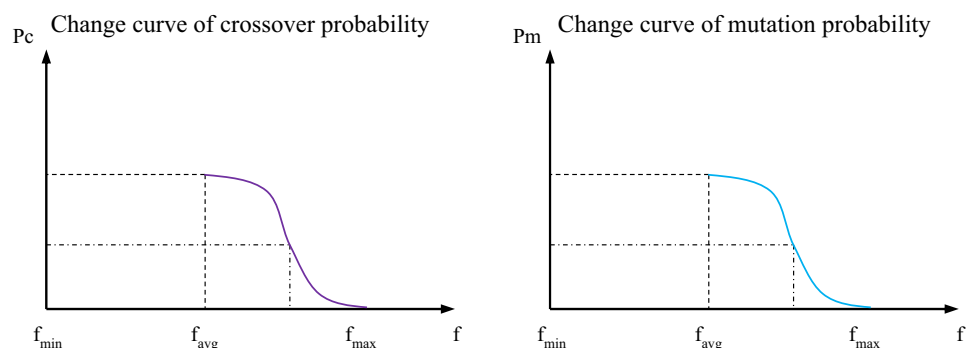
$$p_j = \frac{Fit_j}{\sum_{j=1}^{sn} Fit_j} \quad (18)$$

In the formula,  $p_j$  represents the probability that the  $j$ -th individual is selected;  $Fit_j$  represents the function fitness value of the  $j$ th individual;  $sn$  is the population size, that is, the number of individuals in the population.

The cumulative probability for each individual is then calculated as shown in Formula (19):

$$Q_j = \sum_{i=1}^j p_i \quad (19)$$

In the formula,  $p_j$  represents the probability that the  $j$ th individual is selected, and set  $p_0 = 0$ ;  $Q_j$  represents the cumulative probability of the  $j$ th individual.



**Figure 2.** Adaptive change of crossover probability and mutation probability at A\_DPGA.

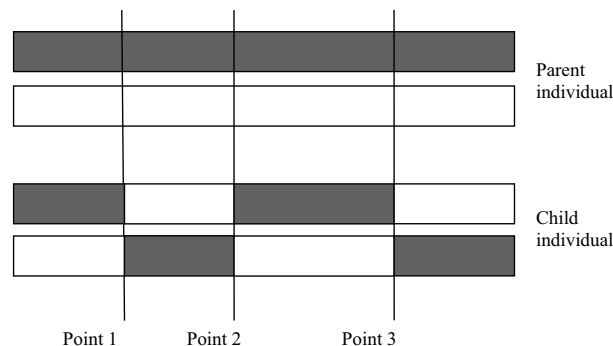


A random number  $z$  uniformly distributed between 0 and 1 is generated one at a time, and when  $Q_{j-1} < z < Q_j$ , the  $j$ th individual is selected as the parent. Repeat the above process, select  $sn-1$  individuals, and then select the best individual in the population as the last individual in the parent.

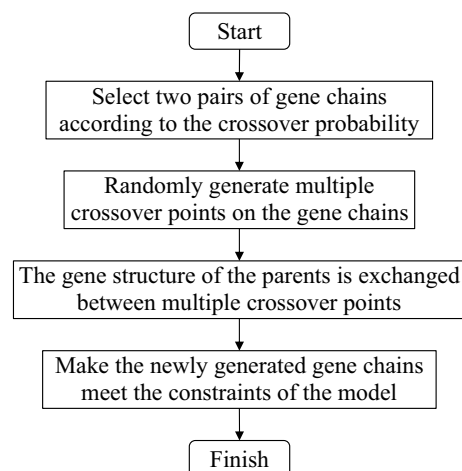
**Genetic crossover.** Gene crossover is the process of crossing parts of the structures in two individuals to produce two new sub-individuals. The selection of crossover probability of the GA will affect the performance of the algorithm and directly affect the convergence of the algorithm. The higher the crossover probability, the more likely it is that the pattern of old individuals will be destroyed, and the faster new individuals will be created. However, too high a crossover probability may destroy excellent individual patterns, while too low a crossover probability may delay the generation of new individuals and lead to premature maturation of the algorithm. The common crossover methods for binary coding are single-point crossover, multi-point crossover and uniform crossover, etc. This paper uses the multi-point crossover method to solve the model. In this paper, the design of the multi-point crossover method schematic diagram is shown in Fig. 3, and the design of the crossover technique route is shown in Fig. 4.

As shown in Fig. 3, three crossover points of the genetic crossover were generated in the parent individual. The structure before point 1 and between point 2 and point 3 remains unchanged without crossover transformation. The interchange between point 1 and point 2 and between point 3 and the final structure to generate new individuals. Compared with the single-point crossover method, the multi-point crossover method has greater damage to the structure of the parent individual, allowing the algorithm to have more solutions during the genetic crossover. It can promote the search for solution space and avoid premature convergence of the algorithm.

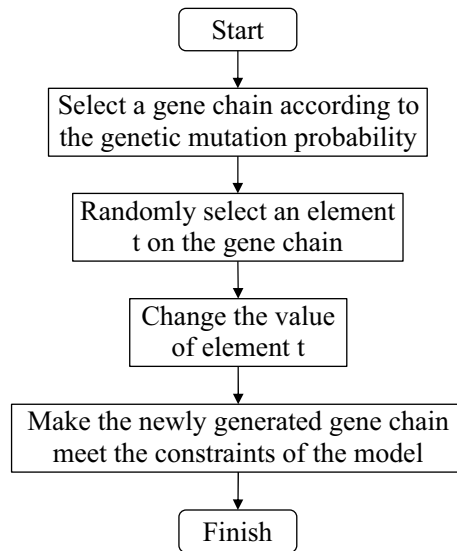
**Genetic mutation.** Gene mutation is the change of the code for some individuals in the population after genetic crossover. This operation is a local search, which makes the population after genetic crossover more diverse, and can avoid the algorithm from falling into the local optimal solution. For the probability of genetic mutation, the value should not be too large, otherwise it will degrade the A\_DPGA into a random search algorithm. The technical route of genetic mutation designed is shown in Fig. 5.



**Figure 3.** Schematic diagram of multi-point crossover.



**Figure 4.** Technical route of genetic crossover.



**Figure 5.** Technical route of genetic mutation.

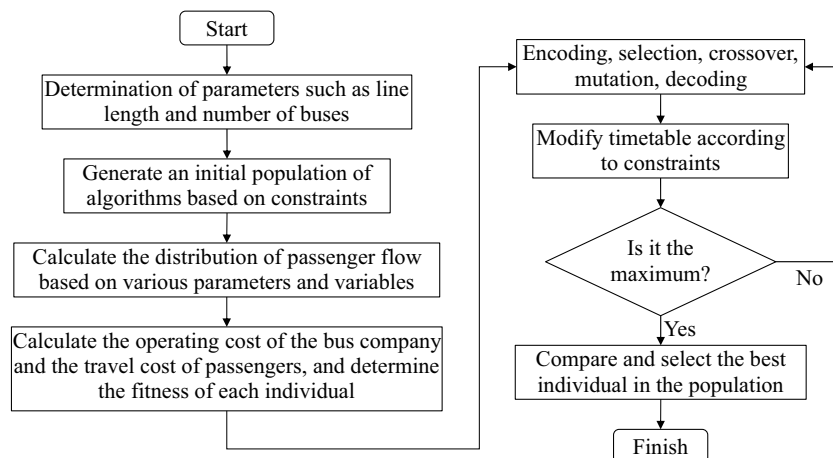
*Termination condition of the algorithm.* In general, the termination condition of a GA can be satisfied by one of the following three conditions: (1) the entire GA solution iterates a specified number of times; (2) in an iteration process, the population fitness is less than the minimum threshold we set; (3) the minimum fitness and the average fitness of the population changes for several consecutive times is less than a certain threshold. The A\_DPGA used in the article sets the termination condition according to condition (1). The technical route of the A\_DPGA in this paper is shown in Fig. 6.

**Example verification and analysis**

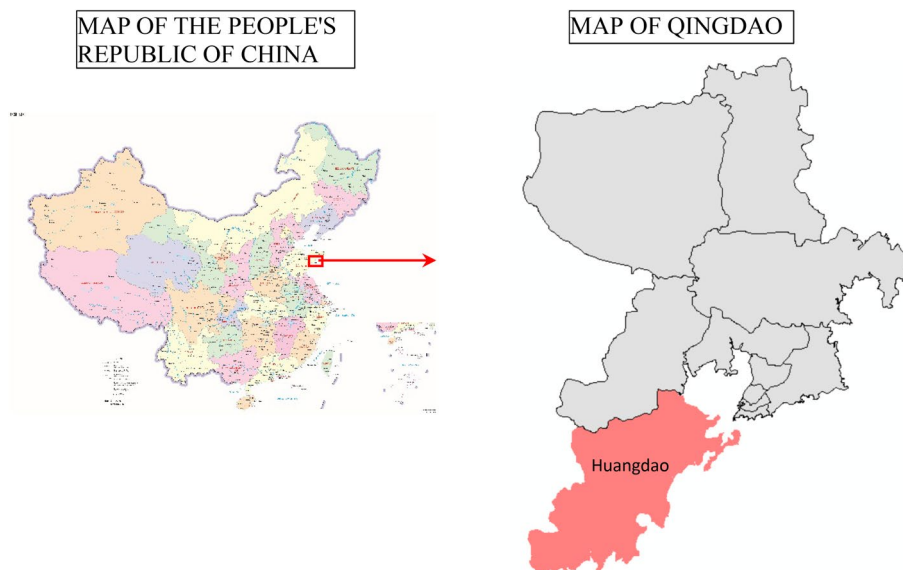
To verify the performance of the Dual\_CBSOM and A\_DPGA, taking the actual route in the Huangdao District of Qingdao City as an example, and studying the passenger flow data of bus schedule in a certain direction on a certain day. The map of the study area is shown in Fig. 7.

Qingdao is part of Shandong Province and is located in the southern part of the province. The red area on the left shows the location of the city for Qingdao in China, which is part of the coastal cities of eastern China. Huangdao District, also known as Qingdao West Coast New District, is located in the southwest of Qingdao and is the municipal district of Qingdao City, Shandong Province. The roads in Huangdao are generally in good condition, with high speed limits, wide lanes, and mostly tarmac roads. However, in recent years, due to the increase in the number of private cars, there are often traffic jams.

The bus line number is D9; there are 18 buses at the origin stop and the last stop, and the buses depart from these two stations. There are 28 stations in the line, and the stations passing in the upward direction are Longhu Yuanshan, Zhiyuan Middle School, ..., West Coast East Bus Station, and Tangdaowan Community. The distances between stations are: [366, 704, 280, 630, 386, 530, 635, 644, 300, 778, 612, 486, 230, 479, 709, 921, 420, 718, 400,



**Figure 6.** Technical route of the A\_DPGA.



**Figure 7.** Map of the study area. Sources: the image on the left is a map downloaded from the Standard Map Service website of the National Bureau of Surveying, Mapping and Geographic Information, and the URL is <http://bzdt.ch.mnr.gov.cn/index.html>. The image on the right is a vectorised map generated by the first author [liu] using ArcGIS 10.7. The left map is based on the standard map with the review number of GS(2022)4316 downloaded from the standard map service website of the National Bureau of Surveying, Mapping and Geographic Information, with no modifications to the base map.

714, 892, 320, 620, 691, 677, 355, 374], unit: M. The number of IC card swiping passengers at each station in the upward direction of the line on August 27, 2018 are [11, 529, 611, 223, 218, 210, 156, 126, 191, 174, 54, 33, 251, 158, 111, 31, 67, 27, 20, 27, 44, 13, 33, 62, 18, 7, 2, 0], unit: person-times. This data is calculated by combining the IC card swiping data and bus GPS data for 5 days from August 27, 2018, to August 31, 2018.

After investigation and actual debugging, we preset the parameters of the Dual\_CBSOM as shown in Table 7:

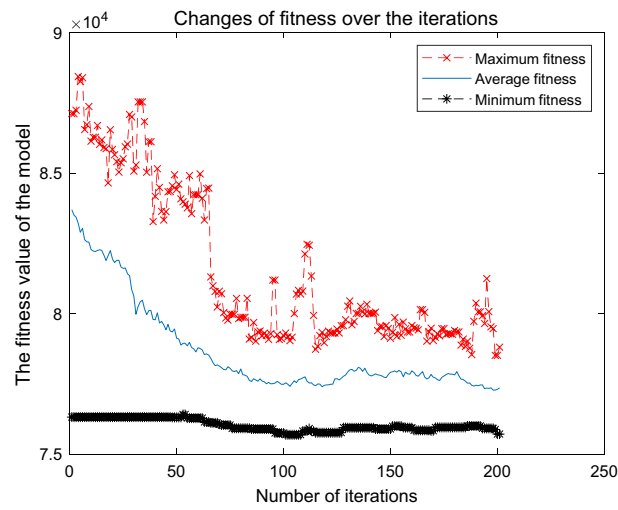
To solve the Dual\_CBSOM, we each write the GA, AGA, and A\_DPGA in MATLAB language. During the iteration process of the algorithm, the maximum fitness, average fitness, minimum fitness, and their corresponding individual solutions of each population are recorded. Figures 8, 9, and 10 depict the model-solving procedures for GA, AGA, and A\_DPGA, respectively. Figure 11 shows the three algorithms' best solution procedures.

Figure 8 shows the change process of three fitness values with iteration times when the GA is used to solve the Dual\_CBSOM. It is impossible to guarantee that the population's best individuals will be passed down to the following generation because the crossover probability and mutation probability of the population is fixed in the GA. When iterating 53 and 109 times, the population's ideal solution is destroyed, which results in a worse outcome for the current generation than for the one before it. The algorithm iterates 101 times and gets the optimal solution. The fitness of the original departure schedule is 77,085. The optimal solution of the population obtained by the GA is 75,686, which reduces 1.81%.

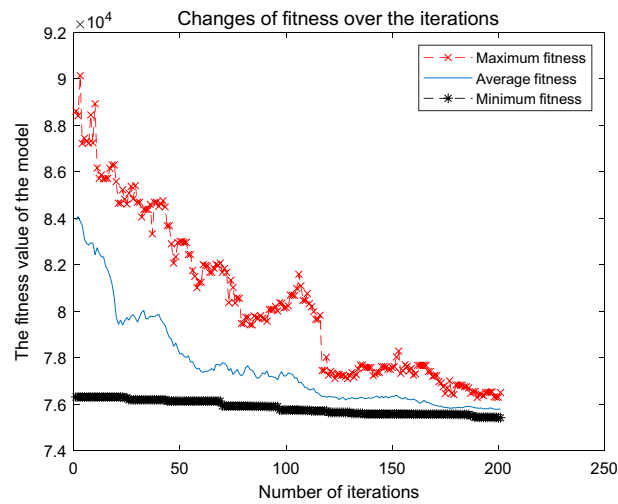
Figure 9 shows the changing process of fitness when solving the Dual\_CBSOM with AGA. The minimum fitness of the population leveled off after 100 iterations, and the average fitness of the population oscillated and stabilized after 120 iterations. The optimal solution is 75,422 after 200 iterations, which is 2.16% less than the original departure schedule and 0.35% less than the optimal solution of GA. This shows that the AGA can better retain the excellent individuals in the population, and the fitness of the population becomes smaller and smaller with the increase of iteration times. However, the findings of this model may eventually converge to a locally

Model parameter	Value	Model parameter	Value
Unit time operating cost of the bus	540 yuan/h	Crossover probability constant $k_1$	0.5
Unit time cost of bus passengers	22 yuan/h	Crossover probability constant $k_2$	0.9
Weight coefficient $\alpha$	0.75	Mutation probability constant $k_3$	0.01
Weight coefficient $\beta$	0.25	Mutation probability constant $k_4$	0.05
Population size $S_n$	100	Maximum departure interval $t_{\max}$	20 min
Maximum number of iterations $G_{\max}$	200	Minimum departure interval $t_{\min}$	5 min
The ratio of IC card-swiping passengers to total bus passengers	1:2		

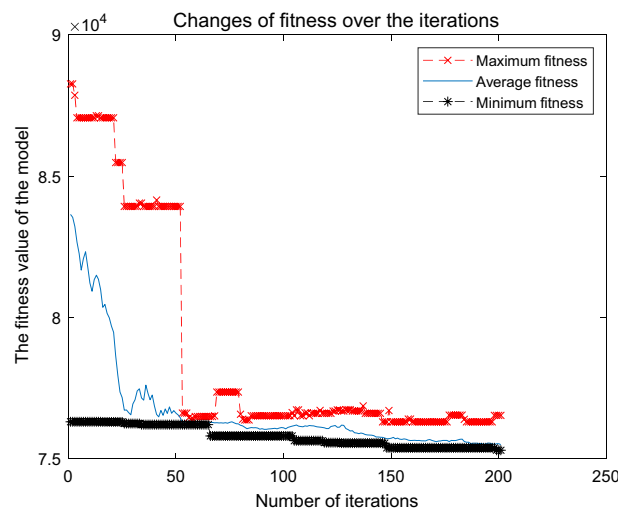
**Table 7.** Parameter settings of the Dual\_CBSOM.



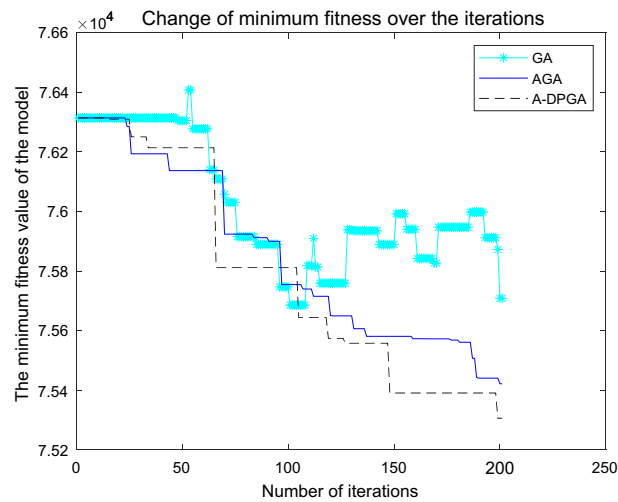
**Figure 8.** Model solving process of the GA.



**Figure 9.** Model solving process of the AGA.



**Figure 10.** Model solving process of the A\_DPGA.



**Figure 11.** Optimal value solving process of three algorithms.

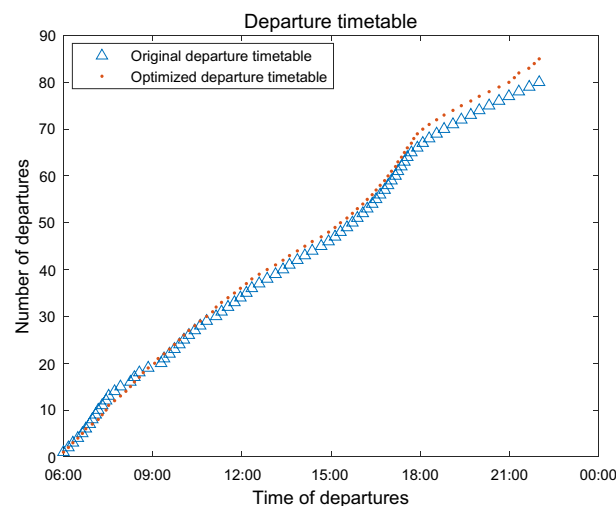
optimal solution, because the method optimizes too slowly, requiring dozens of iterations to optimize a relatively minor degree of fitness value.

Figure 10 shows the solution process of the A\_DPGA to the Dual\_CBSOM. This figure shows that the average fitness of the population oscillates and stabilizes after 60 iterations, and the minimum fitness of the population obtains its optimal value after 199 iterations, with an optimal value of 75,306. Compared with the solution result of the original departure plan, GA and AGA, the optimal value is reduced by 2.3%, 0.5%, and 0.15%, respectively.

The three curves in Fig. 11 show the change process of the minimum fitness when the GA, AGA, and A\_DPGA solve the Dual\_CBSOM, respectively. The GA can not guarantee that the current optimal value can be successfully inherited by the next generation population, so the optimal value will suddenly increase in some iterations. Both the AGA and the A\_DPGA can make the optimal value in the population to be inherited by the next generation population by adaptively adjusting the parameters, but the solution speed of the AGA is too slow and the time cost can be high. The A\_DPGA can find the optimal value rapidly and make sure that it is more appropriate for the actual situation by smoothing the probability around the average fitness and the maximum fitness.

Since the optimal value obtained in this study by solving the model with the A\_DPGA is better, we choose the departure timetable corresponding to this optimal value for comparison with the original departure timetable, as shown in Fig. 12.

From Fig. 12, we can see that the difference between the optimized departure timetable and the original departure timetable is mainly concentrating on three time periods: 6:30–9:00, 11:30–15:00, and 17:30–22:00. For example, the optimized departure plan increases the number of bus departures during the morning rush hour from 6:30 to 8:30 a.m. At the same time, there are other times when the number of bus departures is reduced



**Figure 12.** Comparison of departure timetables.

	Original departure schedule	GA	AGA	A_DPGA
Total fitness value	77,085	75,686	75,422	75,306
Bus company cost (yuan)	39,164	40,849	40,100	40,798
Passenger waiting cost (yuan)	96,637	86,191	87,036	84,552
Passenger crowding cost (yuan)	94,212	94,005	94,350	94,280
Total cost of passenger travel (yuan)	190,849	180,196	181,386	178,832
Number of iterations		101	199	198
Number of departures	80	85	83	85

**Table 8.** Summary of results.

compared with the original departure timetable. The optimized departure timetable more closely follows the passenger flow distribution law.

Table 8 shows a summary of the original departure schedule and the results solved by the three algorithms. The overall objective function value of the acceptable solution obtained by solving the model with the A\_DPGA is 75,306, while the overall objective function value of the original departure timetable is 77,085, which is reduced by 2.3%. The bus operation cost  $f_1$  increased from 39,164 yuan to 40,798 yuan, which is a 4.0% increase in costs due to 5 more bus dispatches. Passenger cost  $f_2$  dropped from 190,849 yuan to 178,832 yuan, a decrease of 6.3%. The total waiting cost of passengers decreased from 96,637 yuan to 84,552 yuan, which declined by 12.51%. The waiting time per capita of passengers decreased by 1.07 min. This demonstrates that the results obtained by solving the model with the A\_DPGA improve passenger perceptions, reduce passenger waiting times, reduce the occurrence of passenger crowding, and provide decision support for bus optimization.

## Conclusion

To address the shortcomings or knowledge gaps identified in the above literature, this paper addresses bus scheduling in two main parts: First, use the original bus data to analyze and calculate the basic information of passengers, and establish a Dual\_CBSOM according to the travel rules of passengers; Second, improve the classical AGA and design a new adaptive crossover operator and mutation operator, so that the crossover probability and mutation probability can automatically adjust their values according to the fitness of individuals and the degree of dispersion of the population. We use the A\_DPGA to solve the Dual\_CBSOM, and take the actual line data as an example to verify the feasibility of the Dual\_CBSOM and A\_DPGA.

The example's results show that the above method is used to optimize the bus dispatch timetable in Huangdao District, Qingdao. Compared with the original scheduling schedule, the solved scheduling schedule has higher passenger satisfaction, less waiting time, and lower extra congestion. After optimizing the bus departure schedule, passengers will be more inclined to choose the bus as their travel mode and the number of passengers will increase, which will increase both the interests of passengers and the bus company. This shows that the use of Dual\_CBSOM and A\_DPGA is effective for bus optimization in Qingdao and can provide a basis for bus scheduling departure plans. This can make bus schedules more intelligent, promote the construction of smart transportation and advance the development of smart cities. Therefore, it can improve the overall operational efficiency of the urban transportation system and play a very important role in the wave of smart city construction. According to the travel patterns of passengers in different regions, the models and algorithms can also be applied to the research of public transport schedules in other cities, for example, they can provide a reference for metro scheduling plans. However, the method and results cannot be applied to the scheduling of long-distance transport such as trains and airplanes.

This paper does not consider the impact of unforeseen circumstances such as road traffic accidents on bus operations. Subsequent research can consider more realistic situations, and we can incorporate the influence of traffic lights and road accidents on the time of vehicles entering and leaving the station into the scheduling model. It is also possible to quantify the unit time cost of passengers according to the different occupations of passengers, therefore the scheduling model is more suitable for the actual situation.

## Data availability

The data that support the findings of this study are available from the first author, [Liu], upon reasonable request.

## Code availability

Code name: Data and Code 1.rar. Developer: Shuai Liu. Contact address: Shandong University of Science and Technology, 579 Qianwangang Road, Huangdao District, Qingdao City, Shandong Province. Telephone number: 15275291288. E-mail: 1135774826@qq.com. Program language: MATLAB R2018b. Source code link: [https://figshare.com/articles/dataset/Data\\_and\\_Code1/22046726](https://figshare.com/articles/dataset/Data_and_Code1/22046726). <https://doi.org/10.6084/m9.figshare.22046726>.

Received: 5 January 2023; Accepted: 5 April 2023

Published online: 13 April 2023

## References

- Jiang, Y., Wang, L. & Yin, J. M. Priority development strategies of public transport in medium sized cities—A case study of Taizhou city. *J. Transp. Eng. Inf.* **17**(01), 117–121+146. <https://doi.org/10.3969/j.issn.1672-4747.2019.01.018> (2019).
- Houman, S. M., Mohamad, R., Mohamad, S. & Ezatollah, M. Optimization of main public transport paths based on accessibility—Case study: Mashhad, Iran. *J. Public Transp.* <https://doi.org/10.5038/2375-0901.19.1.8> (2016).
- Zhang, J., Li, W. Q. & Ran, B. Research on intelligent vehicles scheduling and optimizing methods for conventional public transit. 16–27 (Nanjing: Southeast University Press, 2014). ISBN: 978-7-5641-4995-6.
- Chen, L., Feng, L. & Li, Q. R. Bus departure interval model considering on-bus perceived value. *J. Harbin Inst. Technol.* **50**(03), 150–155. <https://doi.org/10.11918/j.issn.0367-6234.201704096> (2018).
- Xue, Y. Q., Guo, J., Zhong, M. & An, J. Optimization of regular bus scheduling based on uncertainty theory. *J. Transp. Syst. Eng. Inf. Technol.* **21**(06), 115–122+130. <https://doi.org/10.16097/j.cnki.1009-6744.2021.06.013> (2021).
- Li, Y. H., Xu, W. T. & He, S. W. Expected value model for optimizing the multiple bus headways. *J. Appl. Math. Comput.* <https://doi.org/10.1016/j.amc.2012.11.098> (2013).
- Liu, Q., Sun, Y. E., Huang, H. & Du, Y. Bus departure time optimization with rail transportation transfer demand prediction. *J. Chin. Mini-Micro Comput. Syst.* **43**(02), 430–437. <https://kns.cnki.net/kcms/detail/21.1106.TP20210330.1415.012.html> (2022).
- Gong, S. M., Feng, S. B. & Liu, F. M. An optimization model of bus departure interval based on genetic algorithm. *J. East Liaoning Univ. (Nat. Sci. Ed.)* **27**(04), 284–289. <https://doi.org/10.14168/j.issn.1673-4939.2020.04.12> (2020).
- Leiva, C., Muñoz, J. C., Giesen, R. & Larrain, H. Design of limited-stop services for an urban bus corridor with capacity constraints. *J. Transport Res. B-Methods* **44**(10), 1186–1201. <https://doi.org/10.1016/j.trb.2010.01.003> (2010).
- Zhang, H., Zhao, S. Z., Liu, H. S. & Liang, S. D. Design of integrated limited-stop and short-turn services for a bus route. *J. Math. Probl. Eng.* <https://doi.org/10.1155/2016/7901634> (2016).
- Torabi, M. & Salari, M. Limited-stop bus service: A strategy to reduce the unused capacity of a transit network. *J. Swarm Evol. Comput.* **44**, 972–986. <https://doi.org/10.1016/j.swevo.2018.10.009> (2018).
- Zuo, Z. Y., Niu, S., Yin, W. & Zhang, Y. Q. Optimization study of elastic launch interval based on distribution characteristics of public traffic flow. *J. Dalian Jiaotong Univ.* **40**(06), 9–14. <https://doi.org/10.13291/j.cnki.djdxac.2019.06.002> (2019).
- Niu, S. Research on optimization of elastic launch interval based on distribution characteristics of public traffic flow. *Dalian Jiaotong Univ.* <https://kns.cnki.net/kcms/detail/detail.aspx?FileName=1019039037.nh&DbName=CMFD2019> (2018).
- Serafini, P. & Ukovich, W. A mathematical model for periodic scheduling problems. *SIAM J. Discrete Math* **2**(4), 550–581. <https://doi.org/10.1137/0402049> (2006).
- Domschke, W. Schedule synchronization for public transit networks. *J. OR Spektrum* **11**(1), 17–24. <https://doi.org/10.1007/BF01721163> (1989).
- Nachtigall, K. Periodic network optimization with different arc frequencies. *J. Discrete Appl. Math.* **69**(1), 1–17. [https://doi.org/10.1016/0166-218X\(95\)00073-Z](https://doi.org/10.1016/0166-218X(95)00073-Z) (1996).
- Vansteenwegen, P. & Oudheusden, D. V. Developing railway timetables which guarantee a better service. *Eur. J. Oper. Res.* **173**(1), 337–350. <https://doi.org/10.1016/j.ejor.2004.12.013> (2004).
- Feng, J. X. *et al.* Optimising departure intervals for multiple bus lines with a multi-objective model. *J. IET Intell. Transp. Syst.* <https://doi.org/10.1049/iet-its.2017.0049> (2018).
- Niu, H. M. & Zhang, M. H. An optimization to schedule train operations with phase-regular framework for intercity rail lines. *J. Discrete Dyn. Nat. Soc.* <https://doi.org/10.1155/2012/549374> (2012).
- Jiang, Q. Y., Xie, J. X. & Ye, J. *Mathematical Modeling*. 13–15. ISBN: 978-7-04-031150-1 (Higher Education Press, 2011).
- Chen, X. W., Dai, X. & Chen, Q. Approach on the information collection, analysis and application of bus intelligent card. *J. Chin. Civ. Eng. J.* **37**(02), 105–110. <https://doi.org/10.15951/j.tmgcx.2004.02.020> (2004).
- Chen, S. Highway transportation optimization control system based on OD forecast information. *Int. J. Syst. Assur. Eng. Manag.* **12**, 748–756. <https://doi.org/10.1007/s13198-021-01071-5> (2021).
- Hu, J. H., Deng, J. & Huang, Z. Trip-chain based probability model for identifying alighting stations of smart card passengers. *J. Transp. Syst. Eng. Inf. Technol.* **14**(02), 62–67+86. <https://doi.org/10.16097/j.cnki.1009-6744.2014.02.018> (2014).
- Li, J. Y., Zhang, J., Zhang, J. W., Wang, Q. & Peng, D. P. An algorithm to identify passengers' alighting stations and the effectiveness evaluation. *J. Geoma Inf. Sci. Wuhan Univ.* **43**(08), 1172–1177. <https://doi.org/10.13203/j.whugis20160235> (2018).
- Shang, H. Y., Liu, Y. P., Huang, H. J. & Guo, R. Y. Vehicle scheduling optimization considering the passenger waiting cost. *J. Adv. Transp.* <https://doi.org/10.1155/2019/4212631> (2019).
- lv, S., Tao, L. Y. & Mo, Y. K. Level of service classification and quantification for bus waiting time on commuting trip. *J. Transp. Syst. Eng. Inf. Technol.* **15**(03), 190–195+221. <https://doi.org/10.16097/j.cnki.1009-6744.2015.03.030> (2015).
- Cao, Y. W., Ju, Y. F. & Chen, F. Optimal model of urban bus frequency. *J. Anhui Univ. (Nat. Sci.)* **31**(02), 29–32 (2007).

## Acknowledgements

The authors are thankful to the editors and the anonymous reviewers for their help, comments and suggestions. Thanks for the Qingdao ZhenQing Bus Group.

## Author contributions

S.L.: Conceptualization, Model development, designing algorithm, producing results, original draft preparation, review & editing, supervision. L.L.: Conceptualization, methodology, formal analysis, Writing—review & editing. D.M.P.: Conceptualization, formal analysis, review & editing, supervision. J.W.: Conceptualization, provide metadata.

## Funding

This work was supported by the Natural Science Foundation of Shandong Province (No. ZR 2019MD034).

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to L.L.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.





**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023