Oper Res Int J (2015) 15:25–49
DOI 10.1007/s12351-015-0168-y

ORIGINAL PAPER

# An optimization method for planning the lines and the operational strategies of waterbuses: the case of Zhoushan city

**Bin Yu · Zixuan Peng · Keming Wang · Lu Kong · Yao Cui · Baozhen Yao**

Received: 4 March 2014 / Revised: 12 January 2015 / Accepted: 19 January 2015 / Published online: 7 February 2015


**Abstract** Marine transportation can provide more comfortable transit service compared to road-based public transportation. This paper proposes an optimization method for planning the lines and the operational strategies of waterbuses. In this method, the candidate hub ports are selected first, and then a two-stage optimization model is constructed. The model comprehensively considers the interests of both the passengers and the operators by optimizing the lines and the operational strategies of the waterbuses. To solve the model, a shuffled genetic algorithm is proposed. Furthermore, Zhoushan city in China has been chosen as the case study to test the proposed method. The results show that the optimized waterbus lines and operational strategies are better than those that are currently used by the water transportation system in Zhoushan city.

**Keywords** Waterbus · Line planning · Operational strategy · Shuffled genetic algorithm

## 1 Introduction

In China, many people have relocated to the urban areas as a result of increasing urbanization. Because of this urban population growth, the city traffic demands have increased considerably. More motor vehicles on the road make congestion a frequent occurrence in many Chinese cities. The typical method of solving traffic

B. Yu · Z. Peng · K. Wang · L. Kong · Y. Cui
Transportation Management College, Dalian Maritime University, Dalian 116026, People's Republic of China

B. Yao (✉)
School of Automotive Engineering, Dalian University of Technology, Dalian 116024, People's Republic of China
e-mail: yaobaozhen@hotmail.com

congestion is to construct a large-scale transportation infrastructure, such as a road network. However, the development practices of many cities, both here and abroad, have shown that increasing the supply of transportation alone is not sufficient to meet a growth in transportation demand. In fact, developing a large-capacity transportation mode, such as public transportation, is an effective way to provide large-scale transportation for an urban area. Most public transportation systems are road-based, and therefore, the development of public transportation cannot adequately increase the traffic supply and decrease road congestion.

For coastal and riverside cities, developing waterborne public transportation is a suitable technique for solving the urban traffic congestion problem, particularly for Zhoushan city in China, which is located on an archipelago. In Zhoushan city, there is both road-based and marine transportation. A few main islands are connected by bridges, and they make it possible for motor vehicles to travel. However, there is also congestion because of the limited capacity of the bridge at the entrance of the city. Previous studies (Barth and Boriboonsomsin 2008) have shown that the emissions of motor vehicles are highest in cases of stop-and-go traffic and in high-speed situations. From the perspective of environmental protection and emissions reduction, emissions will be a serious problem as the vehicle ownership of Zhoushan city continues to increase.

The most commonly used transportation mode for residents in Zhoushan city is point-to-point inter-island waterborne public transportation. To reduce the operating costs associated with waterborne public transportation in Zhoushan, the departure frequency is low for inter-island routes that have relatively few passengers. This low departure frequency also makes the waterborne public transportation system inconvenient for daily travelling.

Waterbuses are introduced in this paper to solve and optimize the public transportation problems associated with an archipelago area and to improve the public transportation services for coastal and riverside cities. It is hoped that more convenient and comprehensive services can be provided to the passengers by optimizing the lines and operational strategies of the waterbuses. The introduction of waterbuses can attract more passengers to this public transportation method, which may indirectly alleviate traffic congestion at the cross-sea bridge entrance; it may also further reduce traffic emissions. Therefore, the optimization of the waterbuses' lines and operational strategies, which are proposed in this paper, are of significance for archipelagos and for coastal and riverside cities.

## 2 Literature review and contributions

Due to the narrow range of waterbus applications (limited to coastal and riverside cities), there are few studies of waterbuses. Keiji and Keiki (2007) used a Pittsburgh-style learning classifier system to study the robustness of the waterbus system. Waterbuses should continue to serve passengers even when catastrophic weather impacts a city. Keiki et al. (2007) studied an urban traffic network composed of light rail trains, subways, and waterbuses using a quantitative analysis method. The results showed that having diverse modes of transportation in an urban network was conducive to the development of new services for a transportation sector. Ye et al. (2007) chose Shanghai and Huai'an as examples to analyze the attractiveness of waterbuses to passengers using a disaggregate model. Overall, most research of waterbuses has been fixed on using a qualitative analysis of the operating strategy. Waterbuses have been assumed to be a supplementary part of the cities' transportation systems, and only the route selection has been analyzed. There has been a lack of complex waterbus network building and no concrete operating strategy.

There are many similarities between waterbuses and bus/train operations. The operation of waterbuses also requires consideration of site spacing and frequency. Therefore, this paper refers to Dubois et al. (1979), Ceder and Wilson (1986), Agrawal and Mathew (2004), Park (2005), and Zhou and Zhong (2005) for the study of waterbus line or frequency. Pattnaik et al. (1998), Van Nes et al. (1998), Fu et al. (2009), Chien et al. (2001), Yu et al. (2012a, b) who optimized a bus transit route network subject to such issues as geography, capacity, direct demands and transfers. However, there are differences in the characteristics of waterbuses and road-based buses. In actual operations, the number of waterbus passengers must be smaller than the rated capacity, whereas the capacity restrictions of road-based buses are not as strict. Additional factors should be included in the design of waterbus lines, such as the sailing distance, sea conditions and port-anchored order. In terms of scheduling and departure frequency, scholars such as Shih and Mahmassani (1995) and Chakroborty et al. (2001) have developed optimized models from the perspective of vehicle and fleet size. Scholars such as Yu et al. (2010, 2011), and Taketoshi et al. (2012) have directly optimized the timetables and frequency of departures. To minimize the total travel time of the passengers or the total cost of system, Tom and Mohan (2003), Schöbel and Scholl (2006), Yan and Tang (2008), Yu et al. (2012a, b) and Yao et al. (2014a) all studied the bus frequency and the bus scheduling. These research studies about bus/train scheduling and departure frequency can be applied to waterbuses.

Additionally, many studies have jointly considered the bus/train networks and scheduling. Jeremy and Tom (2011) searched for a set of bus routes and schedules to solve the transit route network design problem using a wide variety of heuristic and meta-heuristic approaches. Baaj and Mahmassani (1995) designed a route-generation algorithm (RCA) to optimize bus lines and their frequency. Cordeau et al. (1998) classified recent optimization models for optimized train routing and scheduling by focusing on a model's structure and algorithmic aspects. Similar research problems have been studied by Chakroborty (2003), Bell and McMullen (2004), Zhao (2004) and Guan et al. (2006). The models described above optimized the network and frequency simultaneously, and they considered the interests of both the passengers and operators. Due to the complex nature of these problems, a hybrid algorithm was used to solve the problems. Joint optimization of bus/train networks and scheduling is relatively mature; in contrast, waterbuses lack quantitative research, and they have constraints that are unique to this transportation method. Therefore, the joint optimization of a waterbus network and its operational strategy is a challenging task.

In a waterbus system, there are many factors that should be considered when the operator operates the network. The total cost for the operator is one of the key points of interest. At the same time, for passengers who receive the service, the costs, including the travel time, the times of day of the transfers, the fees and so on, must be considered when they choose a waterbus. There is a conflict of interest between the operators and passengers. To meet the needs of the passengers in terms of public transportation convenience, the service frequency should be as high as possible. However, a high service frequency will result in some wasting of resources and rising operating costs, and it does not meet the desire of the operator for economic efficiency. Therefore, a balance of interests between the operators and passengers is critical to building a stable waterbus system. This paper optimizes both the lines and operational strategy (such as the ship number arrangements and departure frequency) of the waterbuses.

There are two main contributions in this paper. First, a two-stage model is proposed to optimize the waterbus system for archipelago areas and some coastal and riverside cities. The solution, which includes a tradeoff between the operators and passengers, is used to ensure the usability of the optimization. Then, a shuffled genetic algorithm (SGA) algorithm is proposed to solve the model. The newly proposed algorithm provides better results than the solutions that can be found using a traditional GA and SA. Second, Zhoushan city in China is used to test the optimization model based on real data. It provides a reference for the future development of waterbuses in Zhoushan city, and it has practical significance.

The rest of this paper is organized as follows. In Sect. 3, we introduce the two-stage model of optimizing the lines and the operational strategies of the waterbuses; In Sect. 4, the SGA algorithm is described; In Sect. 5, case study of Zhoushan city, the convergence of the algorithm and the results analysis are presented; Finally, some conclusions and direction for future research are provided in Sect. 6.

## 3 Optimization model for waterbus lines

### 3.1 Problem description

Waterbuses are a mode of public transportation that is similar in form to road-based public transportation. In the current system, the most commonly used travel mode in Zhoushan city is a point-to-point inter-island waterborne public transportation system that has a small number of trips trip. In this paper, a hub-and-spoke system is introduced to optimize the planning of the waterbus lines by integrating the short lines. In a hub-and-spoke system, each feeder port is connected directly to a hub port, and connections between the feeder ports are not allowed. Then, a network of routes is formed among the hub ports. The optimized network integrates the original scattered lines. The passenger trip flows converge on the hub ports. The number of operational lines decreases, but the passenger load factor is larger than before. These two aspects make it possible for the network to achieve economies of scale and to avoid excess capacity. Therefore, the introduction of a hub-and-spoke system can reduce the number of ships that are added to the waterbus lines, and it can also reduce the addition of new fixed costs.

To solve the waterbus line optimization problem, the hub ports should be selected, and then, the order of the waterbus berthing, the travelling lines, the number of ships and the departure frequencies can be optimized. After the optimization, the waterbuses can be operated using a hub-and-spoke approach. The hub-and-spoke mode of operation is shown in Fig. 1.

## 3.2 Assumptions

The following assumptions are introduced to simplify the construction and the calculation of the waterbus optimization model:

- All ships used in the waterbus system are of the same type.
- The traveling speed of the ship remains unchanged; it is not subject to external factors;
- The period of waterbus operations is from 7:00 to 20:00;
- The number of passengers arriving at each site follows a uniform distribution;
- The shipbuilding price, speed and fuel prices are based on the average levels for 2012.

## 3.3 Waterbus line optimization model

A summary of the model construction is as follows. First, the hub ports are selected in case the waterbus system is too large to optimize. The hub ports are the stops that are directly served by the waterbus lines, and the remaining ports are feeder ports, which are connected to the hub ports. Then, the best connections between the remaining ports and the hub ports are determined according to passenger demand

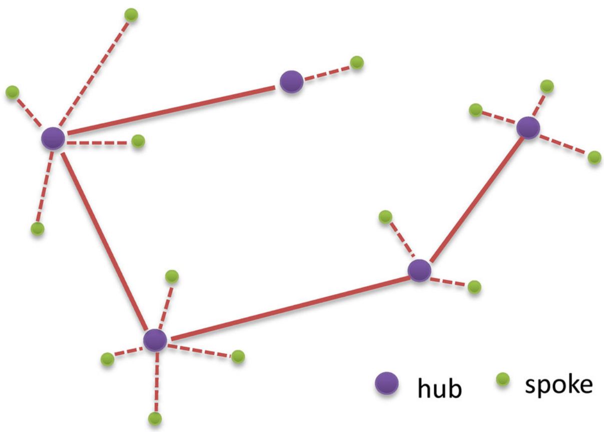


**Fig. 1** Hub-and-spoke route schematic diagram

and sailing time. Finally, the lines and operational strategies of the waterbuses for the hub ports are optimized by the waterbus line optimization model.

### *3.3.1 Determination of the hub ports*

To shorten the travelling time of the passengers, which will make waterbus travel more attractive, there should not be an excessive number of berthing ports (the hub ports mentioned above) because the waterbus berthing process takes a long time in an actual operating environment. Therefore, the hub ports should be selected first from all candidate ports. The remaining ports are feeder ports, and they will be connected to the hub ports to form a waterbus network with an aggregation-based clustering method. In this paper, the hub ports and feeder ports are determined before any model calculations. Then, they are input into the two-stage model as known quantities.

The hub port selection process is based on the rule of importance. A port with a greater number of passengers and lines is more important, and this port will likely be chosen as a hub port. The hub port is chosen according to the highest values of $y_t$:

$$y_t = \frac{x_t w_t}{\sum_{i=1}^{n} x_i} \tag{1}$$

where $y_t$ is the evaluation result of the $t$-th evaluation object; $x_i$ is the evaluation index value of the $i$-th evaluation object, and $w_t$ is the evaluation weight of the $t$-th evaluation object.

$$w_t = a_t / \sum_{i=1}^{n} a_i \tag{2}$$

where $a_t^2 = \sum_{k=1}^{r} (x_t - \overline{x_i})^2 / (k-1)$ is the variance of the evaluation index $x_t$ and $k$ is number of the calculating base period.

After choosing the hub ports, the aggregation-based clustering method is used to determine the connections between the hub and feeder ports:

$$P_{a,b} = \frac{\theta(s_{a,b}, q_{a,b})}{\sum_i \theta(s_{i,b}, q_{i,b})}, \quad a \in i, b \in j \tag{3}$$

where $p_{a,b}$ is the probability of feeder port $b$ being connected to hub port $a$. $s_{a,b}$, $q_{a,b}$ are the distance and passenger flow values between feeder port $b$ and hub port $a$, respectively, and $i, j$ is the set of hub ports and feeder ports.

$$\theta(s_{a,b}, q_{a,b}) = \frac{1}{s_{a,b}} \times q_{a,b} \tag{4}$$

where $\frac{1}{s_{a,b}}$ represents the distribution situation of the feeder port and $\theta(s_{a,b}, q_{a,b})$ considers both the distance and passenger flows in determining the connection between the hub port and feeder port.

### *3.3.2 Upper-stage model*

The upper-stage model uses the minimum of the total travel time of waterbus passengers as its objective function to optimize the waterbus lines. The total travel time consists of the waterbus travelling time among the hub ports, the passengers' wait time and the ferry travel time between the hub ports and feeder ports. The specific formula is shown as follows:

$$[\mathrm{UP}]: \mathrm{Min} \sum_{\substack{i,j \in N \\ k \in K}} \frac{S_{ij}^k}{v} \times q_{ij}^k + \sum_{\substack{i \in N \\ k \in K}} \frac{U_i^k}{2 \times f_k} + \sum_{a,b \in A} \varepsilon \times t_{ab} \times Q_{ab} \times \delta_{ab} \tag{5}$$

where $N$ is the set of hub ports; $K$ is the set of waterbus lines; $A$ is the set of all ports, including the hub and feeder ports; $v$ is the average running speed of the waterbus, which is an empirical value; $S_{ij}^k$ is the distance traveled from port $i$ to the next port $j$ through path $k$; $U_i^k$ is the number of passengers boarding at port $i$ through path $k$; $\varepsilon$ is the conversion coefficient for the transfer time; $t_{ab}$ is the time from port $a$ to the next port $b$; and $Q_{ab}$ is the number of passengers from port $a$ to the next port $b$.

$$\delta_{ab} = \begin{cases} 1 & \text{one of port a and port b is hub port, the other is feeder port} \\ 0 & \text{otherwise} \end{cases}$$

$$f_k = \frac{v \times n^k}{\sum_{i,j \in N} S_{ij}^k \times 2} \tag{6}$$

where $f^k$ is the departure frequency of path k and $n^k$ is the number of vessels on path $k$.

$$q_{ij}^k = \sum_{m=1}^{i} U_m^k - \sum_{n=2}^{i} D_n^k \tag{7}$$

where $q_{ij}^k$ is the number of passengers from port $i$ to the next port $j$ through path k and $\mathrm{D}_i^k$ is the number of passengers disembarking at port $i$ through path $k$.

$$S_{ij}^k = l_{ij} \times x_{ij}^k \times \Delta_{ij}^k \tag{8}$$

where $l_{ij}$ is the distance to sail from port $i$ to port $j$.

$$x_{ij}^k = \begin{cases} 1 & \text{port i and port j are on path k} \\ 0 & \text{otherwise} \end{cases}$$

$$\Delta_{ij}^k = \begin{cases} 1 & \text{port i and port j and path k are adjacent} \\ 0 & \text{otherwise} \end{cases}$$

### 3.3.3 Lower-stage model

The objective function of the lower-stage model is to find the lowest total operational cost. This paper considers the ship acquisition costs and fuel costs, which equal the largest proportion of the total cost. Penalty costs are also considered. The specific formula is as follows:

$$[\mathrm{LP}] : \mathrm{Min} \sum_{k \in K} (p_{vessel} \times n^k + n^k \cdot B^k \times T) + C_{punish} \tag{9}$$

where $p_{vessel}$ is the vessel price; $B^k$ is the cost of fuel consumption per day on path k; and $T$ represents the base period of the calculation. In this paper, the life cycle of the ship is the base period, and $C_{punish}$ is the penalty cost.

Subject to:

$$B^k = \left( 24 \times 10^{-6} \times P_{fuel} \times g \times \frac{\xi^{\frac{2}{3}}}{c} \right) \times v^3 \tag{10}$$

$$C_{punish} = \theta \times \left( \sum_{\substack{i,j \in N \\ k \in K}} \frac{S_{ij}^k}{v} + \sum_{k \in K} f_k \right) \tag{11}$$

where $P_{fuel}$ is the fuel price; $g$ is the fuel consumption rate of the main engine; $\xi$ is the vessel's dimension; $c$ is the admiralty coefficient; and $\theta$ is the conversion coefficient of the time cost. The travelling time and transfer time are converted into a penalty cost, which is a part of the objective function. It can constrain the line lengths and transfer times.

$$\sum_{d \in F} o_{di}^k + \mu_i^k = U_i^k \tag{12}$$

$$q_{ij}^k \le E \tag{13}$$

$$O_{budge} \le \bar{O}_{budge} \tag{14}$$

where $F$ is a set of feeder ports; $o_{di}^k$ is the number of passengers transferring from feeder port $d$ to hub port $i$ by path $k$; $\mu_i^k$ is the number of passengers travelling from the hub port to the other ports directly, excluding the passengers from the feeder ports; $E$ is the capacity limit of each waterbus; $O_{budge}$ denotes the budge of operator, and $\bar{O}_{budge}$ denotes the budge of the current system.

In the upper-stage model, the decision variables are $x_{ij}^k$ and $\Delta_{ij}^k$, and they are based on the passenger perspective. Several waterbus lines are chosen through a combination of decision variables. Then, both $x_{ij}^k$ and $\Delta_{ij}^k$ become input parameters for the lower-stage model. From the operator perspective, the lower-stage model is the point at which $n^k$ can be achieved. $n^k$ is also regarded as the input parameter for the upper-stage model. The upper and lower models iterate until they reach the termination condition.

From a passenger perspective, a higher departure frequency of the waterbuses results in shorter waiting times and thus more convenient travel. However, from the operator perspective, a higher frequency with the determined routes involves equipping more ships and consuming more fuel, which does not reflect the economics needs of a waterbus operator. Therefore, both the economic efficiency requirements of the operators and the convenience of passengers must be considered simultaneously to balance the interests of both sides and thus achieve optimization.

## 4 Solution algorithm

Many literatures suggest that heuristic algorithm is often the first choice to solve this kind of complicated problems (Yao et al. 2010, 2011, 2013; Yao et al. 2014b, c). Genetic algorithm (GA) is one of the heuristic methods which are widely used in many literatures (Yu et al. 2010, 2011). To optimize the lines and the operational strategies of the waterbuses, SGA is introduced to solve the upper-stage model and the lower-stage model is solved by lingo (Haase and Kolisch 1997).

GA is a randomized search method that evolved from learning the laws of evolution (survival of the fittest) of the biosphere. It was first proposed by Holland (1975). Its main characteristic is that it operates directly on the structure object without the limit of derivation and function continuity. With random searches, GA can access and guide a model to help it to optimize the search space automatically, and it can adaptively adjust the searching direction without pre-determined rules. GA has been successfully applied in many fields.

Wren and Wern (1995) solved the public transport driver scheduling problem with GA. Both Pattnaik et al. (1998) and Bielli et al. (2002) used GA algorithm to solve the bus transit network optimization model. Tom and Mohan (2003) proposed a model to optimize bus routes and frequencies based on GA. Park (2005) presented a model to optimize both headways and slack time with a simulation-based GA. Chakroborty et al. (1995, 2001) and Chakroborty (2003) also solved bus scheduling problem with GA. Tongchim and Chongstitvatana (2002) even improved the GA with parameter adaptation and information processing letters. Bin et al. (2013) settled the two-phase model with GA.

### 4.1 Upper-stage model algorithm

The basic components of the SGA algorithm are explained below. The basic idea of a shuffle is a combination of deterministically based complex search technology and biological competition, which is present in nature. The $n$ samples constitute a generation, and each generation is divided into many sub-populations. Each sub-population can independently evolve in different directions to find the solution space. After reaching a certain generation, each sub-population will converge to find the best solution. Then, the sub-populations will generate new offspring, which will be divided into new sub-populations according to the rule described above. The speed of global optimization and convergence can be improved by sharing information about the search space with all sub-populations. A shuffle system was first proposed by Duan et al. 1994, and it has been successfully used in the area of hydrology to solve rainfall-runoff models. Afterwards, other researchers have also used this approach to present near-optimal programs at the network level (Fwa et al. 1994; Ferreira et al. 2000).

In a SGA, the sample size (which is also the number of chromosomes) is $S_{SGA} = g \times m$, where $g$ is the number of sub-populations and $m$ is the number of sample points in each sub-population. Each sample point corresponds to a single chromosome of an offspring in a traditional GA. It is represented by $x_1, \ldots, x_s$, and the fitness function value for each sample point is $F_{(x_i)} = F_i, i = 1, \ldots, s$. After the fitness function of each sample is calculated, we create an array of all fitness function values in descending order, which we denote as $D = \{(x_i, F_i), i = 1, \ldots, s\}$, where $F_1 \le F_2 \le \cdots \le F_s$. we partition $D$ into $g$ sub-populations $A^1, A^2, \cdots A^g$, where $A^k = \left\{ \left( x_j^k, F_j^k \right) | x_j^k = x_{k+m(k-1)}, F_j^k = F_{j+m(k-1)}, j = 1, \ldots, m \right\}$, $k = 1, 2, \ldots, g$. Each sub-population iterates. After a certain number of iterations, sub-populations $A^1, A^2, \ldots, A^g$ will be re-combined into the new D. We continue the partition, evolution and shuffle procedure until the convergence criteria is satisfied.

### 4.1.1 Coding

In this paper, the waterbus routes must be generated before building a waterbus network. The waterbus route programs are determined through a combination of all hub ports. For example, $n$ hub ports are chosen from all candidate ports, and the final waterbus network will be built based on these $n$ hub ports. In this case, there should be $H^n = \sum_{i=2}^{n} C_n^i$ waterbus route programs. Furthermore, the length of the coding is in accordance with the number of chosen hub ports. If there are $n$ ports, the coding length is $H^n$; that is, each code represents one possible waterbus route program, which is connected by several ports. For example, there are four hub ports. Using $H^n = \sum_{i=2}^{n} C_n^i$, eleven waterbus route programs can be determined as shown in Table 1. Each code on the chromosome indicates whether the corresponding

**Table 1** Waterbus route programs

| Combinatorial number | Serial number | Route program |
|---|---|---|
| $C_4^2$ | 1 | 1, 2 |
| | 2 | 1, 3 |
| | 3 | 1, 4 |
| | 4 | 2, 3 |
| | 5 | 2, 4 |
| | 6 | 3, 4 |
| $C_4^3$ | 1 | 1, 2, 3 |
| | 2 | 1, 2, 4 |
| | 3 | 1, 3, 4 |
| | 4 | 2, 3, 4 |
| $C_4^4$ | 1 | 1, 2, 3, 4 |

route program has been selected; one means that it belongs to the waterbus network, and zero means that it does not belong. In Fig. 2, the codes, which correspond to $C_4^2(1)$, $C_4^2(2)$ and $C_4^2(5)$, are equal to one. This means that there are three waterbus lines in the network, lines 1–2, 1–3 and 2–4, respectively.

$H^n$ will be larger as the number of hub ports increases. This means that the length of each code will be specified, and the computing time will increase accordingly. To reduce the computing time, some clearly unreasonable waterbus route programs should be removed first, and the scale of $H^n$ will decrease accordingly. In this paper, long and short waterbus routes were removed from the program collection because of our desire to minimize the total operating costs and travel times of the passengers. If a waterbus route is overly long, the route should be equipped with more waterbuses to ensure a reasonable trip frequency. Similarly, if a waterbus route is not sufficiently long, which occurs when the route only connects two closed stops that have a small trip flow, it will also be uneconomical. The unreasonable waterbus route programs can be summarized as following.

Assuming that there are 7 hub ports to optimize for a waterbus system and that the port locations are shown in Fig. 3, there are three possible situations (a–c) according to the combinations of all hub ports. As shown in Fig. 3a, all hub ports are connected (this is only one of the possible connections; the best program will be determined by the shortest path method, which is commonly used to solve an Assignment Problem). If a waterbus route is overly long, as shown in Fig. 3a, the demands of the operators and passengers cannot be satisfied at the same time. For the operators, the departure frequency cannot be achieved using an economical amount of vessel equipment, and to ensure a convenient departure frequency for the passengers, there should be more vessels, which may lead to higher operating costs. Another extreme situation is shown in Fig. 3b; in this case, the waterbus route is not sufficiently long. When the route length is short and the trip flow is small, it will be uneconomical to add this route into the final waterbus network, and therefore, it should be removed from the programs before the GA is encoded. As shown in Fig. 3c, the route moves directly from stop 1 to stop 3, and it ignores stop 2. There may be other routes that connect stop 2 in the final waterbus network. However, although a network, which includes the route shown in Fig. 3c, can ensure that all

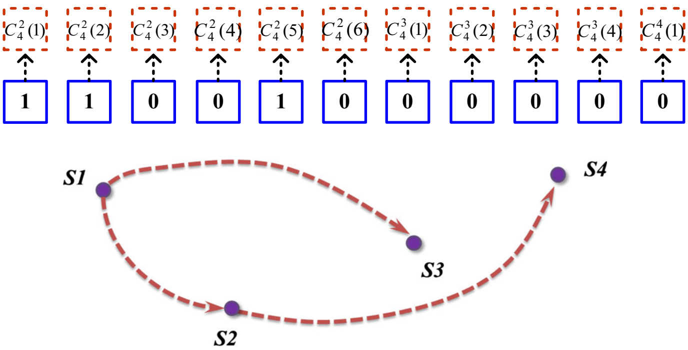


**Fig. 2** Coding of a waterbus line

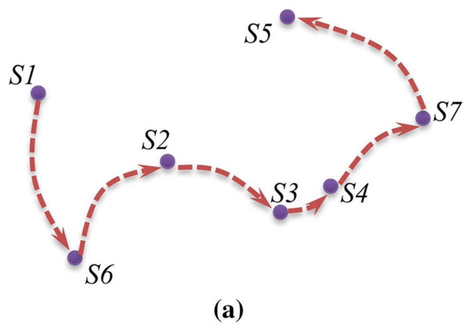


**Fig. 3** Examples of unreasonable waterbus route programs, **a** A waterbus route program that is overly long, **b** A waterbus route program that is not sufficiently long, **c** A waterbus route program that is unreasonable

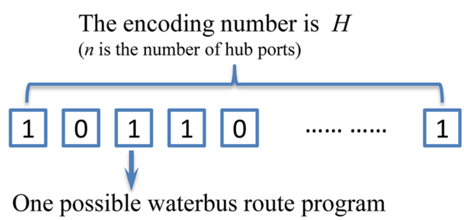


**Fig. 4** Example of GA coding

stops are connected, if the trip flows between stops 1 and 2 and between stops 2 and 3 are overly large, then this route is also unreasonable. Therefore, these three situations should be removed from the waterbus route programs before encoding the SGA. The coding length will be $H$, which is determined by filtrating $H^n$. The meaning of code in this paper is shown in Fig. 4. The code shown in Fig. 4 means that waterbus route programs 1, 3, 4 and $H$ have been chosen to build the final waterbus network.

Therefore, some new constraints should be added to the upper-stage model to simplify the code and to reduce the computing time.

$$r_{Hi} \leq R \tag{15}$$

$$O_{ikj} \times S_{ikj} \geq W \tag{16}$$

Constraint (15) can be used to remove the waterbus route programs that are overly long; these routes increase the operating costs of the entire waterbus network; $r_{Hi}$ represents the length of the $i$-th waterbus route program; $R$ is the acceptable route length, and it can be set at a default value according to practical application. Waterbus route programs that are too short and that have small traffic flows will be removed by constraint (16); $O_{ikj}$ and $S_{ikj}$ are the trip flows and lengths of the route program ($i$, $k$, $j$), respectively; $k$ is a point set, which represents the stops between start stop $i$ and end stop $j$. $W$ is also set at a default value according to practical application.

### 4.1.2 Fitness function

The objective function of an upper-stage model is the lowest total time. Therefore, a constant $G$ is introduced to make a conversion to facilitate the calculation of the fitness function. To ensure the connectivity of the entire network (a counter example case is shown in Fig. 3), a penalty function is quoted. If there are two ports that cannot be connected, a sufficiently large number will be chosen as the penalty cost; otherwise, the penalty cost is zero. In this case, the fitness function can be calculated by the following equations:

$$F = \frac{G}{\sum\limits_{\substack{i,j \in N \\ k \in K}} \frac{S_{ij}^k}{v} \times q_{ij}^k + \sum\limits_{\substack{i \in N \\ k \in K}} \frac{U_i^k}{2 \times f_k} + \sum\limits_{a,b \in A} \varepsilon \times t_{ab} \times Q_{ab} \times \delta_{ab} + \Phi(t)} \tag{17}$$

$$\Phi(t) = \begin{cases} M & \text{if there is no link between i and j} \\ 0 & \text{otherwise} \end{cases} \tag{18}$$

where $F$ represents the fitness function; $G$ is a constant; $\Phi(t)$ indicates the penalty function of the connectivity; and $M$ is chosen to be sufficiently large.

### 4.1.3 Genetic operation

*4.1.3.1 Selection operation* Selection is the operation that is used to choose superior individuals and to phase out inferior individuals from a population. The purpose of selection is to bring optimum individuals directly to the next generation or to produce a new individual by a paired cross, which will be inherited to the next generation. The roulette selection method is used in this paper. If the group size is $n$, the fitness of an individual $i$ is $F_i$, and the selection probability of $i$ is $P_i = F_i / \sum_{j=1}^{n} F_j$. For a crossover operation that occurs afterward, individuals can randomly compose mating pairs after being selected.

*4.1.3.2 Crossover operation* Crossover refers to the operation that causes part of the structure of the two parent individuals to replace and restructure, and it also generates new individuals. The crossover operator in this paper is single-point crossover. An example of single-point crossover is given below (Fig. 5).

*4.1.3.3 Mutation operation* The mutation operator changes the gene locus value of some individual strings in the groups. In general, a decision of whether to mutate all the individuals in a population is first needed, and then, we randomly select the gene locus to conduct a mutation of the selected individuals. Since it is encoded in binary, the mutation operation is as follows (Fig. 6).

## 4.2 Lower-stage model algorithm

The lower-stage model is a simple integer programming model. It can be solved by general software. In this paper, we use the lingo program to solve it (Fig. 7).

## 4.3 The process of algorithm

The two-stage optimization model is solved by the following process.

1. The concrete steps of the SGA algorithm in the upper-stage model are as follows:

Step 1 Randomly generate initial population to represent the waterbus route programs. Each program is generated by the shortest path method based on the waterbus hub ports. The number of chromosome is the sample size of shuffled procedure;

Step 2 Calculate the fitness function value which is objective function of the upper model and arrange them in descending order;

Step 3 Partition the initial population into $g$ sub-populations according to descending order of fitness function values and each sub-population has $m$ sample points. $S_{SGA} = g \times m$ is the chromosome number of each generation;

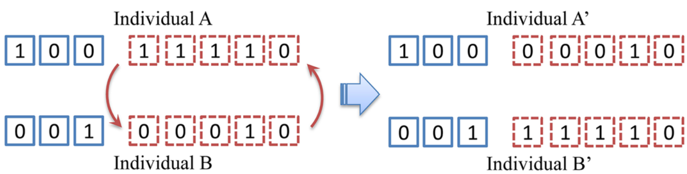


**Fig. 5** Example of a single-point crossover


**Fig. 6** Example of a mutation operation

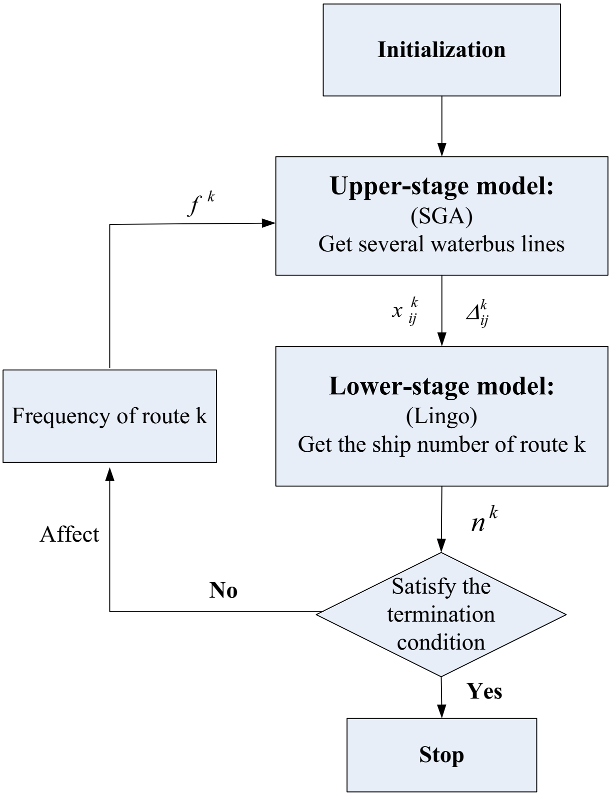


**Fig. 7** Process of algorithm, Notes: Data source: ‘2012 Zhoushan Comprehensive Transportation Planning Report’

Step 4 Perform selection, crossover and mutation operations of SGA to create new populations in each sub-population;

Step 5 Shuffle the g sub-populations and check the fitness function value of each chromosome;

Step 6 Continue to repeat the partition, calculation and shuffle procedure until the algorithm meets the termination condition that the fitness function value no longer decreases or the algorithm reaches the maximum number of iteration. If meet, output the optimal individual and jump out of the algorithm; otherwise perform Step 2.

2. Put the waterbus network that is gotten from the upper-stage model into the lower-stage model as the input variables. The lower-stage model is solved by lingo.
3. Judge whether the algorithm meets the termination condition that the travel time of passengers and the operator cost do not change or the results switch between the two schemes (jumping back and forth). If meet, output the optimal individual and jump out of the algorithm; otherwise put the solution getting from the lower-stage model into the upper-stage model as the input variables. Perform Step (1).

## 5 Case study

There are a total of 1,390 islands and reefs in Zhoushan city; therefore, Zhoushan city is known as "the city of a thousand islands". Currently, Zhoushan city has a land shortage problem; it is a densely populated area. Therefore, whether due to natural conditions or practical needs, it is necessary to open a waterbus system in Zhoushan city. Twenty-two already existing ferry terminals in Zhoushan city have been selected as candidate ports for the waterbus line optimization. The specific locations are shown in Fig. 8.

### 5.1 Hub port selection and feeder port aggregation

The hub port selection results are determined using Eqs. (1) and (2) for the 22 candidate ports. They are the result of the aggregation-based clustering method. A result with more promising evaluation results will be more likely to be chosen as a hub port. In Fig. 9, the evaluation results for S1, S2, S3, S4, S10, S13 and S18 are significantly better than the other ports. Therefore, these seven ports have been selected as the hub ports for the waterbus line optimization model. The remaining 15 ports connect to the hub ports as feeder ports. The calculation results that define the connection between hub ports and feeder ports are determined by Eqs. (3) and (4). Each feeder port is assumed to connect with only one hub port; the specific connection is shown in Fig. 10.

### 5.2 Results of the waterbus line optimization

Before conducting the optimization modeling effort, a considerable amount of data must be obtained. To facilitate the model calculation, the vessel type and sailing speed are set in accordance with Zhoushan's existing vessels: the rated number of passengers is 200, and the speed of the ships is 8.1 *kn*. The existing operational ships are used as waterbuses to utilize existing resources, and only the depreciation costs of these vessels are considered. The optimization model is calculated considering the actual situation with the existing vessels, and therefore, the base period is extended to the life cycle of the vessel: $T = 10$ years. Considering the changeable sea weather, it is assumed that the vessels operate for 300 days per year. The fuel

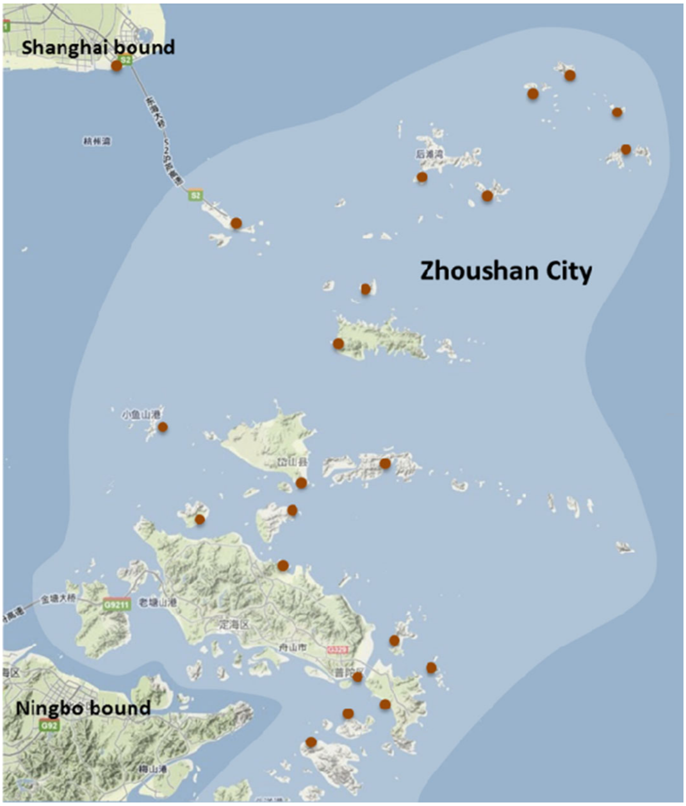


**Fig. 8** Current transportation situation in Zhoushan city, and the specific location of 22 candidate ports

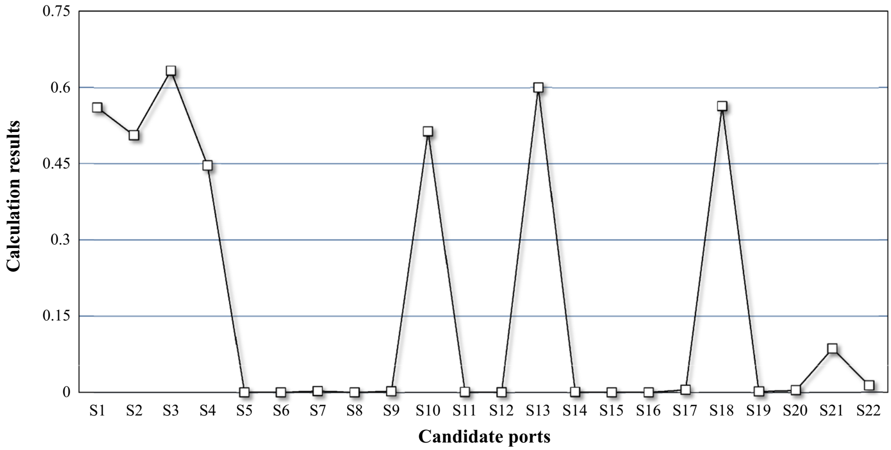


**Fig. 9** Selection results for the hub ports

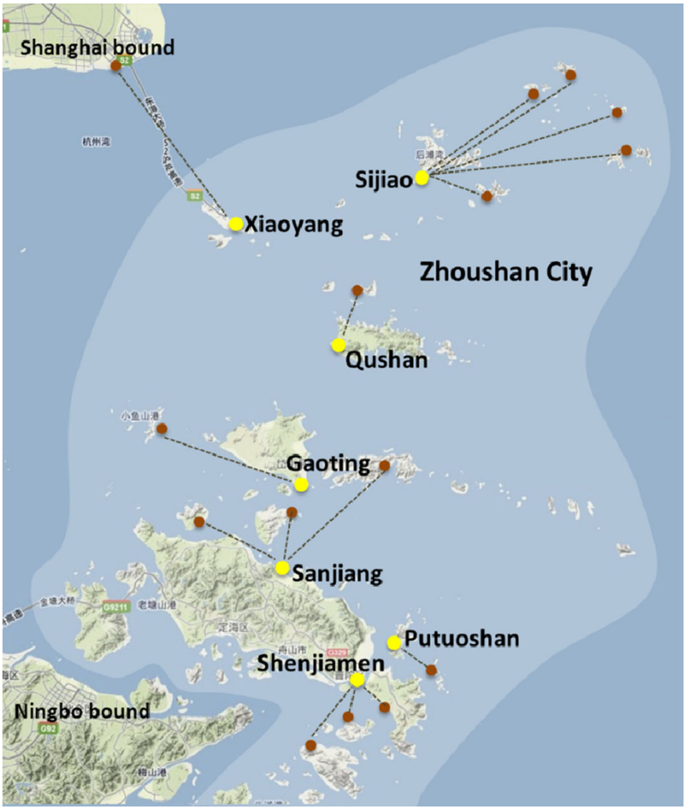


**Fig. 10** Selection results for the hub port and feeder port connections

cost is 7.5 RMB/liter, or \$7,500/ton. In addition, the crossover rate of the SGA algorithm is set as 0.7; the mutation rate is 0.1; and the maximum iteration number is 1,000 generations. M is $10^9$. Each generation has 30 chromosomes, and they will be separated into five sub-populations, each of which has six chromosomes.

### *5.2.1 Calculation results*

The results of Zhoushan's waterbus line optimization are shown in Fig. 11. The seven hub ports are connected by four lines: S1–S18–S10, S18–S3–S2–S13–S4, S1–S2–S4, and S18–S3–S4.

The result of the waterbus operational strategy is $n^k = \{21, 26, 24, 32\}$. There should be a total of 103 ships for four optimized waterbus lines. The ship number and departure frequency of each line are shown in Table 2. The data shown in parentheses in Table 2 have been optimized by the model, and the data that are out of the brackets have been adjusted for convenience during the actual operation. To account for uncertainty in the waterbus's berthing time and to use a realistic

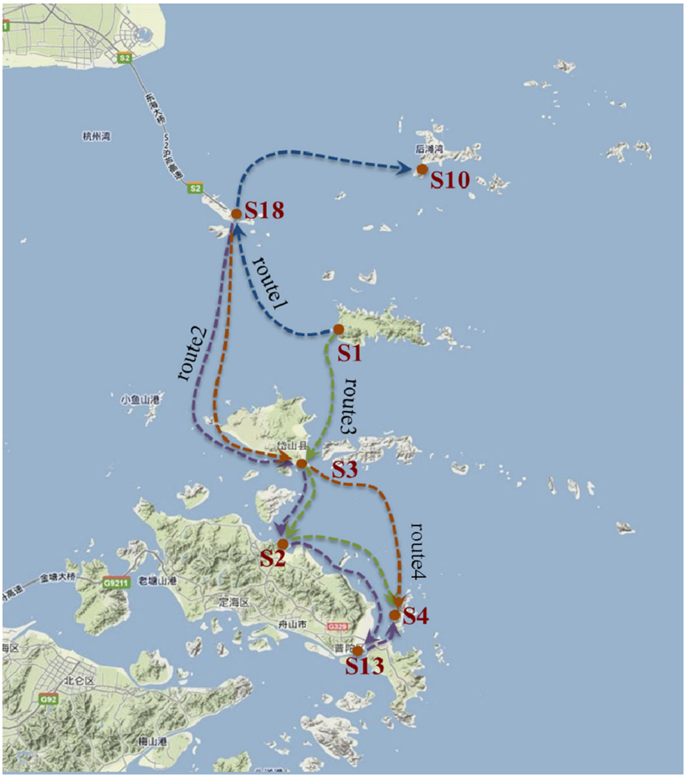


**Fig. 11** Final results for the upper model of Zhoushan's waterbus optimization

operational environment (such as sea waves), the departure frequency has been appropriately lowered to ensure the punctuality rate of the waterbuses.

### *5.2.2 Convergence of the algorithm*

The program has been run ten times to test the convergence of the SGA, as shown in Fig. 12. The results of the fitness function are in a state of sharp decline. During the first 150 generations, there are large fluctuations. After another 150 generations, the

**Table 2** Ship number arrangement and departure frequency

| Waterbus lines | Ship number | Departure frequency (time/hours) |
|---|---|---|
| R1. S1-S18-S10 | 21 | 2 (2.25) |
| R2. S18-S3-S2-S13-S4 | 26 | 3 (3.47) |
| R3. S1-S3-S2-S4 | 24 | 2 (2.66) |
| R4. S18-S3-S4 | 32 | 5 (5.37) |

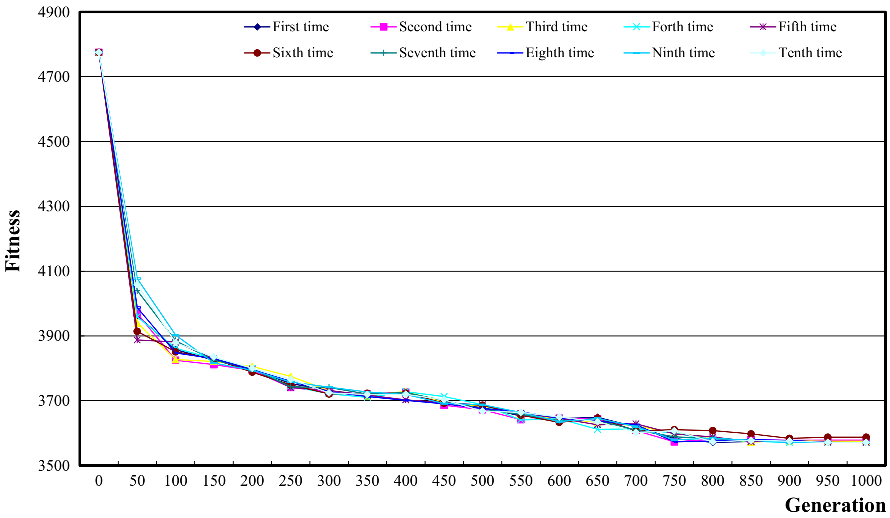


**Fig. 12** Convergence of the calculation

changes begin to flatten. The optimal solution appears in approximately 850 generations. The differences among the results of the ten calculations are small. This shows that the convergence of the algorithm is good, and the optimal solution can be found after approximately 850 iterations.

Finally, the SGA, the traditional GA algorithm (which does not include any shuffling) and the SA are compared to test the performance of the newly proposed algorithm. In the SA, the probability of accepting a worse solution is $\exp\left(\frac{-\Delta}{T}\right)$, and the initial temperature $T_0$ can be determined using the initial status, which can be accepted in the beginning of the annealing process with a probability $P_0$. Pilot runs are performed, and the mean cost of increasing $\bar{\Delta}$ is then computed. In the calculation, $T_0$ is calculated as follows: $T_0 = \frac{\bar{\Delta}}{\ln(P_0^{-1})}$. Each algorithm is tested ten times with the same data. As shown in Fig. 13, the running time of the SGA is shorter than the running time using a traditional GA and SA due to the introduction of a partition, which diversifies the population. Additionally, the SGA provides a better solution than the traditional GA and SA as shown in Fig. 13. The running time of the SA is shorter than the running time of the traditional GA; however, both solutions are close. Therefore, the proposed SGA can improve the optimization quality, and it can also have a good performance.

### *5.2.3 Applicability of the results*

Currently, Zhoushan has opened five bridges to connect parts of the islands, but it still cannot satisfy the transportation demands. Thus, a waterborne public transportation system is needed to shoulder this portion of the demand. The existing waterborne public transportation system of Zhoushan consists of short-haul ferries that exist between two islands, as shown in Fig. 14. The departure frequency

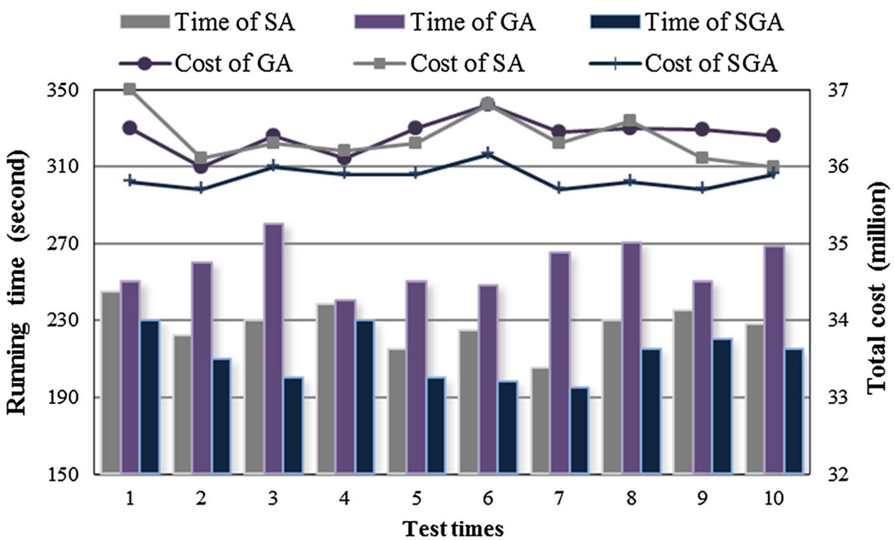


**Fig. 13** Comparison of the SGA, GA and SA

of the ferry is low, which makes passengers' daily travel method inconvenient, as shown in Table 3. Therefore, optimization of the waterbus lines and the corresponding operational strategies will be extremely useful for Zhoushan city.

The existing waterborne public transportation system of Zhoushan calls for more ships to ensure the normal operation of the lines. For the operators, additional ships mean higher fixed costs. In this case, which is not reasonably optimized, an increase in the number of ships will inevitably lead to a lack of supply for some lines, whereas some other lines may have excess capacity. Waterbus line optimization in this paper means that we have optimized the line based on the passenger trip flow, and the occurrence of the problems described above can be avoided.

For the passengers, the opening of a new waterbus has an advantage in that it increases the departure frequency compared to the situation when only the existing ferry is operating. Second, a waterbus can play a role in passenger sharing, and it can provide passengers with a wide selection of transportation alternatives that can be used when the cross-sea bridge is overcrowded in the peak periods. Furthermore, more convenient public transportation will also increase the amount of travel.

The four optimized waterbus lines connect the six islands so that the maximum passenger trip flow and departure frequency are sufficiently large to meet the demands of the passengers. The integrated routes and operational strategy can avoid excess capacity. The optimized waterbus lines can directly replace the existing ferry transportation system that exists between the six islands. Other ferry routes can be opened as feeding routes for the four waterbus lines, and eventually, a hub-and-spoke waterborne public transportation system will be formed. There are currently 137 ships serving the six islands with an annual operating cost of approximately RMB 36.26 million. The passenger travel cost is RMB 24.8 million per year. If a waterbus system is used, the operating costs will be RMB 35.7 million per year with

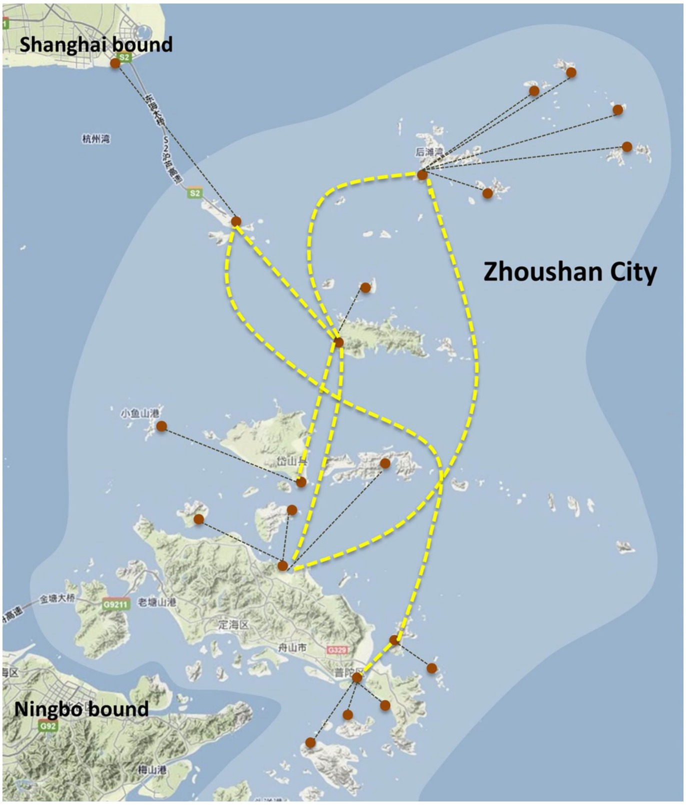


**Fig. 14** Existing waterborne public transportation system in Zhoushan

103 ships. The passenger travel cost will decrease to RMB 20.7 million per year. The optimized cost advantage is not obvious. However, due to its large departure frequency, a waterbus system can attract more passengers, and therefore, it will play an important role in marine public transportation development. After optimization, the capacity of the system will be effectively utilized, and the number of service-ships will be significantly reduced. This will reduce $CO_2$ emissions and contribute to environmental protection. This is the advantage of waterbuses compared with a road-based bus system.

In the Zhoushan case study, S18 is closer to Shanghai and has served as an external transportation hub for Zhoushan. Therefore, additional routes are expected to go through this port. S4 is an important Buddhist cultural tourist attraction; it attracts a large number of domestic and foreign tourists, and it has a large trip flow. Although it is geographically close to S13, S4 was also chosen as a waterbus hub port. S2 is the only port in the north of the island, and it connects the main island with the northern islands. It is necessary to open more routes for this port to facilitate passenger travel. The waterbus line optimization that is presented in this

**Table 3** Differences between the current solution and optimized solution

| Ferry lines | Frequency (time/days) |
|---|---|
| *Current solution* | |
| S3-S1 | 6 |
| S3-S2 | 8 |
| S3-S18 | 1 |
| S3-S10 | 3 |
| S4-S10 | 1 |
| S4-S13 | 6 |
| S2-S10 | 3 |
| Waterbus lines | Frequency (time/days) |
| *Optimized solution* | |
| S1–S18–S10 | 39 |
| S18–S3–S2–S13–S4 | 26 |
| S1–S3–S2–S4 | 65 |
| S18–S3–S4 | 36 |

paper is extremely close to the actual demand. It can solve the waterbus line planning problem, and it has a role in guiding reality.

## 6 Conclusions

In this paper, the interests of both the passengers and operators are considered. After seeking a balance between the costs of the operations and the total passenger travel times, waterbus lines and the corresponding operational strategies can be achieved. And the SGA is designed to solve the optimization model and the newly proposed algorithm is proved to have a good performance. The results of case study show that the method can provide reasonable waterbus lines and the number of vessel for costal or riverside cities. And it can take rational advantage of resources and avoid excess capacity.

Considering travel time and passenger trip flow can effectively optimize waterbus line and corresponding operation strategy. It can meet the different needs of passenger (higher departure frequency) and operator (lower operating cost), and it has role in guiding reality for coastal and riverside cities. The proposed algorithm can determine the final waterbus line connection and the convergence of the algorithm is good.

For further studies, the changeable weather at sea, which may cause a suspension of the waterbus operations, should be considered, and therefore, the efficiency of waterbuses can be improved. The road-based bus system should also be considered to develop a balanced operation for the entire public transportation system of coastal and riverside cities.

**Acknowledgments** This work was supported in National Natural Science Foundation of China 51108053 and 51208079, the Trans-Century Training Program Foundation for Talents from the Ministry of Education of China NCET-12-0752, Liaoning Excellent Talents in University LJQ2012045 and the Fundamental Research Funds for the Central Universities 3013-852019.

## References

Agrawal J, Mathew TV (2004) Transit route network design using parallel genetic algorithm. Journal of Computing in Civil Engineering 18(3):248–256

Baaj MH, Mahmassani HS (1995) Hybrid route generation heuristic algorithm for the design of transit networks. Transp Res Part C 3(1):31–50

Barth M, Boriboonsomsin K (2008) Real-World CO2 Impacts of Traffic Congestion. Transp Res Rec 2008(2058):163–171

Bell JE, McMullen PR (2004) Ant colony optimization techniques for the vehicle routing problem. Adv Eng Inform 18(1):41–48

Bielli M, Caramia M, Carotenuto P (2002) Genetic algorithms in bus network optimization. Transportation Research Part C: Emerging Technologies 10(1):19–34

Bin YU, Hanbing ZHU, Wanjun CAI, Ning MA, Baozhen YAO (2013) Two-phase Optimization Approach to Transit Hub Location–the Case of Dalian. J Transp Geogr 33:62–71

Ceder A, Wilson NHM (1986) Bus network design. Transp Res Part B 20(4):331–344

Chakroborty P (2003) Genetic algorithms for optimal urban transit network design. Computer-Aided Civil and Infrastructure Engineering 18(3):184–200

Chakroborty P, Deb K, Sharma RK (2001) Optimal fleet size distribution and scheduling of urban transit systems using genetic algorithms. Transp. Plan. Technol. 24(3):209–225

Chien S, Yang Z, Hou E (2001) Genetic algorithm approach for transit route planning and design. Journal of Transportation Engineering 127(3):200–207

Cordeau JF, Toth P, Vigo D (1998) A survey of optimization models for train routing and scheduling[J]. Transportation science 32(4):380–404

Duan QY, Sorooshian S, Gupta VK (1994) Optimal use of the SCE-UA global optimization method for calibrating watershed models. J Hydraul Eng 158(1):265–284

Dubois D, Bel G, Libre M (1979) A set of methods in transportation network synthesis and analysis. Journal of Operational Research Society 30(9):797–808

Ferreira LA, Antunes A, Picado-Santos L (2000) Using genetic algorithms on a PMS segment optimization mode. Proc. 1st European Pavement Management Systems, Conf

Fu Y, Gao Z, Li K (2009) Optimization method of energy saving train operation for railway network. J Transp Syst Eng Inf Technol, 94

Fwa TF, Chan WT, Tan CY (1994) Optimal programming by genetic algorithms for pavement management. Transportation Research Record 1455, Transportation Research Board, Washington, 31–41

Guan JF, Yang H, Wirasinghe SC (2006) Simultaneous optimization of transit line configuration and passenger line assignment. Transp Res B 40(10):885–902

Haase K, Kolisch R LINGO. Operations-Research-Spektrum 19.1 (1997):1–4

Holland JH (1975) Adaptation in natural and artificial systems. University of Michigan Press, Ann Arbor, Mich

Jeremy JB, Tom VM (2011) Intelligent Agent Optimization of Urban Bus Transit System Design. Journal of Computing in Civil Engineering 25(5):357–369

Keiji S, Keiki T (2007) Waterbus Route Optimization by Pittsburgh-style Learning Classifier System. PROCEEDINGS OF SICE ANNUAL CONFERENCE 1–8:1150–1154

Keiki T, Takahiro M, Daisuke W (2007) Exploring Quantitative Evaluation Criteria for Service and Potentials of New Service in Transportation: Analyzing Transport Networks of Railway, Subway and Waterbus. Lect Notes Comput Sci 4881:1122–1130

Kunimatsu T, Hirai C, Tomii N (2012) Train timetable evaluation from the viewpoint of passengers by microsimulation of train operation and passenger flow. Elect Eng Jpn 1814

Park SJ (2005) Bus network scheduling with genetic algorithms and simulation. MS thesis, University. of Maryland

Pattnaik SB, Mohan S, Tom VM (1998) Urban bus transit route network using genetic algorithm. Journal of Transportation Engineering 124(4):368–375
Schöbel A, Scholl S (2006) Line planning with minimal traveling time. In: 5th Workshop on Algorithmic Methods and Models for Optimization of Railways
Shih MC, Mahmassani HS (1995) Vehicle sizing model for bus transit networks. Transp Res Rec 1452:35–41
Tom VM, Mohan S (2003) Transit route network design using frequency coded genetic algorithm. J. Transp. Eng. 129(2):186–195
Tongchim S, Chongstitvatana P (2002) Parallel genetic algorithm with parameter adaptation. Information Processing Letters. 82:47–54
Van Nes R, Hamerslag R, Immerse BH (1998) Design of public transportation networks. Transp Res Rec 1202:74–82
Wren A, Wern DO (1995) A genetic algorithm for public driver scheduling. Comput Op Res 22(1):101–110
Yan S, Tang C (2008) An integrated framework for intercity bus scheduling under stochastic bus travel times. Transp. Sci. 42(3):318–335
Yao BZ, Yang CY, Yao JB (2010) Tunnel Surrounding Rock Displacement Prediction Using Support Vector Machine. International Journal of Computational Intelligence Systems 3(6):843–852
Yao BZ, Yang CY, Yao JB, Hu JJ, Sun J (2011) An improved ant colony optimization for flexible job shop scheduling problems. Adv Sci Lett 4(6-7):2127–2131(5)
Yao BZ, Hu P, Zhang MH, Wang S (2013) Artificial Bee Colony Algorithm with Scanning Strategy for Periodic Vehicle Routing Problem. SIMULATION: transactions of The Society for Modeling and Simulation. International. 89(6):762–770
Yao BZ, Hu P, Lu XH, Gao JJ, Zhang MH (2014a) Transit network design based on travel time reliability. Transp Res Part C 43:233–248
Yao BZ, Hu P, Zhang MH, Jin MQ (2014b) A Support Vector Machine with the Tabu Search Algorithm For Freeway Incident Detection. Int J Appl Math Comput Sci 24(2):397–404
Yao BZ, Hu P, Zhang MH, Tian XM (2014c) Improved Ant Colony Optimization for Seafood Product Delivery Routing Problem. Promet-Traffic & Transportation 26(1):1–10
Ye L, Yang DY, He N (2007) Application of stated preference survey method in waterbus system design. Int Conf Transp Eng, Part B 40:885–902
Yu B, Yang ZZ, Yao JB (2010) Genetic Algorithm For Bus Frequency Optimization. Journal of Transportation Engineering 136(6):576–583
Yu B, Yang ZZ, Sun XS, Yao BZ, Zeng QC, Jeppesen E (2011) Parallel Genetic Algorithm in Bus Route Headway Optimization. Appl Soft Comput 11(8):5081–5091
Yu B, Yang ZZ, Jin PH, Wu SH, Yao BZ (2012a) Transit route network design-maximizing direct and transfer demand density. Transp Res Part C 22:58–75
Yu B, Yang ZZ, Shan LI (2012b) Real-Time Partway Deadheading Strategy Based on Transit Service Reliability Assessment. Transp Res Part A 46(8):1265–1279
Zhao F (2004) Transit network optimization—minimizing transfers and optimizing route directness. Journal of Public Transportation 7(1):63–82
Zhou X, Zhong M (2005) Bicriteria train scheduling for high-speed passenger railroad planning applications [J]. Eur J Oper Res 167(3):752–771