Application of Bee Colony Optimization Algorithm in Warehouse Facility Location of Rail Transit Network

Kun Zhang[1], Hui Liu[1], and Kunlei Lian[2]

[1]China Railway Siyuan Survey and Design Group Co., Ltd. Wuhan, China. 430063

[2]State Key Laboratory of Digital Manufacturing Equipment & Technology, Huazhong University of Science and Technology, Wuhan, China 430074

Abstract

This paper proposes a mathematical model for the warehouse facility location problem in a rail transit network. A bee colony optimization algorithm is developed to identify the optimal number of warehouses as well as their locations. Problem-specific encoding scheme and crossover operator are designed to facilitate adaption of the algorithm. In addition, local search algorithm is developed to improve the algorithm's capability in searching for the global optimal solutions.

Key words: Rail transit, warehouse facility location, bee colony optimization

1 Introduction

In rail transit network systems, warehouse facilities play a key role in delivering the required equipment and materials in a accurate and timely manner, based on the needs of each lane. Warehouse facility locations therefore greatly affect whether a rail transit network system can operate efficiently at minimal costs. Moreover, the decision has a long-lasting effect on operation costs in future years as these facilities are hard to change once they are determined. An optimal warehouse facility location solution is desired in order to satisfy the requirements of each lane and minimise the total operation costs. It is common in current practices to set up a dedicated warehouse for each rail lane, which not only consumes more lands and investment, but also leads to duplicated and wasted stock of equipment and materials. With the advance of rail transit systems in China, it is anticipated that they will be built in more cities, with major ones witnessing as many as seven or eight rail lanes. It remains an urgent task to identify the best warehouse facility locations for these systems.

2 Warehouse facility location model

There exist many constraints in warehouse facility location problems, which is NP-hard to solve due to the fact that its complexity increases exponentially with increases in scale. Numerous approaches have been proposed to address this problem, including genetic algorithm, simulated annealing, fuzzy comprehensive evaluation, shortest path algorithm. Bee colony optimization (BCO) is a new metaheuristic algorithm proposed in recent years that has received more and more attentions from the research community and has been applied to solve various industrial optimization problems. This paper builds the formal mathematical model for the warehouse facility location problem, followed by detailed descriptions of adapting BCO to solve this problem. Then, performance of BCO is validated using the real problem encountered in the rail transit facility location scenario in Wuhan, China.

It is known that, in a rail transit system, car depots consume around 70% of equipment and materials, while the rest of them is used to meet demands from maintenance sites that are usually built close to car depots. Therefore, we assume that all the potential warehouse facility locations are known in advance for all the eight rail lanes in Wuhan, and the problem is to decide the optimal number of warehouse facilities and their corresponding locations subject to the constraint of satisfying demands from all the rail transit network. The objective is to minimize total warehouse facility costs, which consists of construction cost, maintenance cost and delivery cost. Note that the construction cost is a fixed cost for each warehouse facility, and the rest two costs are called operational cost, which accumulates gradually. It is possible that one warehouse facility can supply multiple rail lanes and the cost increase coefficient is determined based on established rail transit systems.

The three cost components are explained below:

- Construction cost. This cost consists of the basic construction cost of a warehouse facility and the extra cost incurred by servicing multiple rail lanes. A cost coefficient is associated with each of the four subcategories, namely, house-building, equipment, track-laying and transportation. The cost can be described mathematically as

$$C_1 = C_{10} + \sum_{i=0}^{8} C_{11}(\sum_{j=0}^{8} x_{ij}) \tag{1}$$

and

$$C_{11}(\sum_{j=0}^{8} x_{ij}) = y_i \times C_{build} \times (1 + (\sum_{j=0}^{8} x_{ij} - 1) \times C_{Mbuild}) \tag{2}$$

$$= C_{build} \times (C_{Mbuild} \times \sum_{j=0}^{8} x_{ij} + (1 - C_{Mbuild}) \times y_i) \tag{3}$$

where x_{ij} is a binary variable equal to 1 if warehouse facility i covers the demands of rail lane j, and 0 otherwise; y_i is a binary variable equal to 1 if the rail lane i is selected to build a warehouse facility; $\sum x_{ij}$ is the number of rail lanes that warehouse i serves. C_{build} is the construction cost of a single warehouse facility; C_{Mbuild} is the extra cost associated with serving multiple rail lanes; C_{10} is the total construction cost of a warehouse facility; $C_{11}(x)$ is the function to calculate extra costs, which is estimated using a linear function based on experience.

- Maintenance cost. Similar to the aforementioned construction cost, the cost consists of both the basic maintenance cost of a single warehouse facility and the extra cost incurred by serving multiple rail lanes. Maintenance costs come from various general expenses including personnel salaries, utilities and service costs. The extra cost of serving multiple lanes is also estimated using a linear function, but the coefficients are different. The cost can be described mathematically as

$$C_2 = C_{20} + \sum_{i=0}^{8} C_{21}(\sum_{j=0}^{8} x_{ij}) \tag{4}$$

and

$$C_{21}(\sum_{j=0}^{8} x_{ij}) = y_i \times C_{operate} \times (1 + (\sum_{j=0}^{8} x_{ij} - 1) \times C_{Moperate}) \tag{5}$$

$$= C_{operate} \times (C_{Moperate} \times \sum_{j=0}^{8} x_{ij} + (1 - C_{Moperate}) \times y_i) \tag{6}$$

where C_{20} is the basic warehouse facility maintenance cost; $C_{21}(x)$ is the function to calculate the extra cost; $C_{operate}$ is the maintenance cost coefficient of one single warehouse facility; $C_{Moperate}$ is the cost coefficient of one warehouse servicing multiple rail lanes.

- Delivery cost. This cost is decided by unit delivery cost, delivery distance and whether a rail lane is covered by a particular warehouse facility:

$$C_3 = T_{times} \times \sum_{i=0}^{8} \sum_{j=0}^{8} Ct_{ij} \times D_{ij} \times x_{ij} \tag{7}$$

where delivery distance and unit cost are both random variables following certain distributions. T_{times} is the number of delivery times within one year. Ct_{ij} is the delivery cost between two locations and $T_{times} = 365/P_{period}$ where P_{period} is the delivery cycle in days.

The model has the following constraints:

1. There can be one and only one warehouse facility servicing a given rail lane:

$$\sum_{i=0}^{8} x_{ij} = 1, \forall j = 1, 2, \cdots, 8 \tag{8}$$

2. The total operational cost should decrease after optimization:

$$C_2 + C_3 < C_{2before} \tag{9}$$

   where $C_{2before}$ is the total operational cost before optimization.

3. The total construction cost should be smaller after optimization:

$$C_1 < C_{1before} \tag{10}$$

   where $C_{1before}$ is the construction cost before optimization.

4. Warehouse facility size limit which limits the maximal number rail lanes a warehouse facility can serve at the same time:

$$\sum_{j=0}^{8} x_{ij} \leq scale, \ \forall j = 1, 2, \cdots, 8 \tag{11}$$

5. The number of warehouse facilities that can be built:

$$\sum_{i=1}^{8} y_i = W_{num}, \ \forall i = 1, 2, \cdots, 8 \tag{12}$$

6. Preset warehouse facilities must be built:

$$x_{kk} = 1 \tag{13}$$

7. Decision variable constraints:

$$x_{ij} \in \{0,1\}, y_i \in \{0,1\} \tag{14}$$

8. Relations between x_{ij} and y_i:

$$x_{ij} \leq y_i,\ \forall i, j = 1, 2, \cdots, 8 \tag{15}$$

3 Application in equipment selection

3.1 Algorithm workflow

The warehouse facility model is subject to various constraints defined in the previous section, and exact methods like branch-and-bound and mathematical programming take very long time to reach optimal solutions. On the other hand, metaheuristic algorithms have been successfully applied in many engineering optimization problems. Bee colony optimization (BCO) is a metaheuristic algorithm that imitates the mating process of bee colonies. A bee colony consists of four types of bees, namely, queen bee, drone, worker bee and bee larva. There usually exist only one queen bee within a colony and it is often the one who lives the longest. The queen bee aims to mate with drones to produce larvae. A drone is a haploid that mates with the queen bee and dies after the mating process. Genes of the drone enter the queen bee and passe along to larvae. Worker bee exists to produce honey to provide food for larvae, which grow up to be either queen bee or drones. Abbass proposed a bee colony optimization algorithm based on the observation of a bee colony's mating and reproduce process, and successfully applied it to various engineering optimization problems.

In BCO, bees, including the queen bee, drones and larvae, represent individual solutions to the problem at hand and all the constraints are embedded in the solution representation. In other words, a bee represents a feasible solution to the warehouse facility location problem depicted by the mathematical model described in the previous section. In BCO, the number of times that the queen bee mates is determined by the number of drones, the results of the mating process decide the size of larvae. The size of the queen bee's ovary indicates the storage size of drones' genes. The number of iterations indicates the number of mating process. The workflow of BCO can be summarised as follows:

- Initialize parameters of the BCO.
- Create the initial population and calculate their fitness values. Selected the bee with the best fitness value and use it as the queen bee, the rest bees are used as drones.

- Check whether stopping criteria is met, stop if so, go to the next step otherwise.
- Initialize the ovary, energy and speed of the queen bee. Repeat the following steps if the energy is bigger than the threshold and ovary is not full: 1) choose a drone and let it mate with the queen been if mating criteria is met, put its genes into the ovary; 2) decrease the energy and speed of the queen bee.
- The queen bee selects sequentially the genes in its ovary to produce larvae, which is then raised by different worker bees. If the resulting bee has a better fitness value than the queen bee, it becomes the new queen bee. Similarly it becomes a drone if its fitness value is better than any of the drones.
- Output the queen bee.

3.2 Encoding scheme

This paper uses a integer-based encoding scheme which consists of two parts. Take the rail transit system of Wuhan as an example, part two of a solution uses a binary value to indicate whether a warehouse facility is chosen at the corresponding rail lane or not, part one of the solution decides the warehouse facility servicing the corresponding rail lane. Figure ? shows that rail lanes 1, 3, 4, 6, 7, 8 are selected to build warehouse facilities and the eight lanes will be serviced by the warehouse facilities built on lanes 1, 1, 3, 4, 7, 6, 7, 8, respectively.

3.3 Fitness function

The fitness value of an individual in the bee colony is indicated by the total cost computed from the underlying solution it represents. As discussed in earlier section, the cost consists of three components, namely, construction cost, maintenance cost and operational cost:

$$F_{fitness} = M - (C_1 + C_2 + C_3) \tag{16}$$

where $F_{fitness}$ is the fitness value of a bee in the colony, and it is computed using the three cost components. Note that M is a big number and the more cost a solution takes, the worse its fitness value is.

3.4 Crossover operator

This paper uses the single-point crossover operator depicted in figure ?. In figure ?, two parent bees $P1$ and $P2$ creates a new offspring solution $O1$ through crossover. The crossover operator involves two steps, focusing on two parts of a solution respectively. The crossover operation on part two employs a two-point crossover operator.

3.5 Local search algorithm

The worker bees in the BCO algorithm refer to various local search algorithms that are used to improve the new offspring solutions generated by the crossover operator. These larvae grow up to be either queen bee or drones and the same larva may grow up into either role if raised by different worker bees. On the other hand, different larvae may grow up into either role even raised by the same work bees. This paper uses three different local search heuristics, namely, point insertion, random position swap and adjacent position swap.

4 Experiments

4.1 Problem description

In order to validate the performance of the proposed algorithm, the warehouse facility location problem in Wuhan rail transit system is used as the benchmark instance in this paper.

There are eight rail transit lanes and each lane's material demands are known in advance. The problem is to decide the best number of warehouse facility locations as well as the rail lanes they service correspondingly. Table ? contains the relative positions of the eight possible car depot locations that can be used to build warehouse facilities. Eight warehouses need to be built before optimization, meaning each warehouse services one and only one rail lane. For example, the position of 1(181, 231) indicates the first warehouse facility located at x-axis of 181 and y-axis of 231.

Parameters of the proposed algorithm are set as follows: number of drones 100, number of larvae 200, size of queen bee ovary 100, number of iterations 200. The warehouse facility location model described in this paper is implemented in C++ and 20 independent replications are run to obtain the optimal warehouse locations. The computational results show that 4 warehouse facilities are enough to cover the demands from all rail lanes. The warehouse coverage is illustrated in table ?. Figure ? gives the graphical demonstration of the computational results.

4.2 Computational result comparison

The average cost in 20 runs is 294,100,000, compared to original cost of 341,600,000. A total of 47,500,000 is saved from the optimization exercise. Figure ? shows the cost comparison before and after optimization.

4.3 Algorithm stability

Figure ? shows the algorithm stability for the 20 runs. The x-axis represents the number of each individual run, and the y-axis indicates objective function value related to total cost. The star in the graph indicates the best solution in the 20 runs, and the line represents the average value throughout the 20 runs. It can be seen from the graph that the algorithm demonstrates great level of stability.

5 Conclusions

This paper proposes a warehouse facility location model in a rail transit system and a bee colony optimization algorithm is applied to identify both the best number of warehouses and their corresponding locations. Problem-specific encoding scheme, crossover operator and various local search algorithms are designed to improve the searching efficiency of the algorithm. In future research, it is worthwhile to consider construction cycle and urban planing in the analysis and conduct multiple rounds of the optimization exercise in order to obtain global optimal solutions.

References