Proceedings of the Eastern Asia Society for Transportation Studies, Vol. 5, pp. 374 - 389, 2005

OPTIMIZING BUS TRANSIT NETWORK WITH PARALLEL ANT COLONY ALGORITHM

Bin Yu
Doctor student
Dept. of Civil Engineering
Dalian University of Technology
2 Linggong Road, Ganjingzi, Dalian
116024, China
Fax: +86-411-8470-8768
E-mail:yubin@student.dlut.edu.cn

Zhongzhen Yang
Professor
Transportation Program
Dalian Maritime University
1 Linghai Road, Ganjingzi, Dalian
116026, China
Fax: +86-411-8472-6756
E-mail: yangzhongzhen@263.net

Chuntian Cheng
Professor
Dept. of Civil Engineering
Dalian University of Technology
E-mail: ctcheng@dlut.edu.cn

Chong Liu
Graduate student
Transportation Program
Dalian University of Technology
E-mail:liuchong1980@163.com

Abstract: This study develop an optimization model for bus transit network based on road network and zonal *OD*. The model aims at achieving minimum transfers and maximum passenger flow per unit length with line length and non-linear rate as constraints. The coarse-grain parallel ant colony algorithm (CPACA) is used to solve the problem. To effectively search the global optimal solution, we use a heuristic pheromone distribution rule, by which ants' path searching activities are adjusted according to the objective value. Parallel ACA is carried out for shortening the calculation time. The model is tested with survey data of Dalian city. The results show that an optimized bus network with less transfers and travel time can be obtained, and the application of CPACA effectively increases the calculation speed and quality.

Keywords: Bus network optimization, Direct-through passenger flow density, and Coarse-grain parallel ant colony algorithm

1. INTRODUCTION

Passenger transport in large- and medium-sized cities mainly relies on transit system. The rationality of the transit network planning, therefore, directly influences the travel time and transfer rate of the passengers, and the overall running cost of the transport systems. An ideal transit network, which is featured by large service area, high direct-through trips, small non-linear rate, short travel time, and high accessibility, should be able to meet the needs of the majority. However, as urban layout and population distribution change, the service level of the transit network may be gradually reduced, which has adverse impact on the development of public transportation and the benefits of transit enterprises. To solve this problem, the existing transit network must be adjusted. One of the most adoptable techniques is to artificially change partial routes regardless of the transit network as a whole. In addition, such manual adjustment is largely dependent on the practical experience of the designer(s).

Transit network optimization has been receiving constant concern. Some representative researchers, Dubois *et al* (1979), have subdivided the transit network optimization problem into three sub-topics, 1) identifying the road sections in need of layout, 2) laying out the routes, and 3) optimizing the departure frequency of each line. Ceder and Wilson (1986) first introduced a three-stage model, i.e. trip distribution, route design and frequency setting, into transit network design. Hasselstrőm (1981) proposed a two-stage model which simultaneously optimizes the line and the frequency. Baaj and Mahmassani (1995) argued that the transit network can be generated by optimizing the line and the departure frequency at the same time, after which the network analysis could be undertaken. J. F. Guan (2004) proposed the utilization of integer programming to optimize both line layout and trip distribution in the Minimum Spanning Tree network. Sonntag (1979) approached the line plan problem in railways by a heuristic elimination method without appropriate simultaneous consideration of passenger line assignment. Simonis (1981) took a different approach from Sonntag (1979) by starting with an empty line configuration. Lines with the most direct-through travelers on their shortest paths will be successively added. Wang (2001) has proposed a model that calls for gradual lay out and optimized networking, which is similar to that of Simonis. In other words, Wang's model aims at achieving the maximum direct-through passenger flow volume in a vacant network

Basically there are two types of models for urban transit network design. One is by combining the transit route design and the departure frequency. The other is solely concentrating on the transit network optimization, based on which the departure frequency of each transit line is studied. Both network design and frequency planning are vital to the transit operating cost and passenger travel convenience. However, in comparison with the departure frequency, the route network is much more stable and less liable to external influences, and does not easily get changed once it is established. It is in this sense that the transport network design calls for the utmost circumspection. The departure frequency, on the other hand, is highly sensitive to factors such as passenger flow, weather and road conditions, and therefore needs to be adjusted in accordance with the different situations. Therefore, the quality of the network design may be adversely influenced if transit network and departure frequency are simultaneously optimized, when the network design determines the efficiency of the entire transit system. The apparent ignorance of the departure frequency here does not equal to neglecting the benefits of the transit enterprises. The utmost goal of transit network design is to facilitate passenger trips and to reduce the operating costs by some constraints; in addition, the unstableness of the departure frequency can lead to some uncertain factors in the optimization process. Therefore, transit network deign is prioritized in our research.

It is noteworthy that albeit there are already many established models on transit network design, and most of them are complicated in structure and are not resolvable without simplification. It becomes even more complicated when it comes to resolving the optimization models taking transfers into consideration. Researchers have been delicately seeking solutions to solve the complicated models. For example, Steenbrink (1974) proposed the traditional mathematical programming method; Baaj and Mahmassani (1995) proposed a hybrid arithmetic which combines the path-searching heuristic algorithm in artificial intelligence and the transit system analysis method in operational research. Dubois *et al* (1979) identified the routes layout by a heuristic arithmetic; Chakroborty *et al* (1995) solved the scheduling problem by adopting inherit genetic algorithm. All these methods have employed implicit enumeration methods, which has become the only solution to solving this kind of problems. To sum up, the solutions to the transit network optimization can be divided into five categories: analytic method, heuristic arithmetic, hybrid algorithm, experience-based arithmetic (Dashora *et al* 1998; Fernandez 1993), and simulation model (Senevirante 1990).

Despite the prolific optimization models and solutions in previous researches, one problem lies in that most of them are theory-oriented and not practically implemented. Bearing this context in mind, we focus on the integrated transit network and are based on passenger demands. Here the objective is to facilitate passengers' trips as well as to foster the transit enterprises' profits. An optimization model is developed, which aims to maximize the direct-through passenger density. A vacant network is first established, followed by adding routes to the network according to the principle of maximizing the direct-through passenger density, until all passengers are distributed to the network or some given constraints are overrun. This method differs from previous researches (Simonis, 1981), Michael *et al*. (1997), Wang (2001) in that most of the previous researches firstly identified the shortest path between the origin station and the terminal, and then sought the route bearing the largest through passenger density among these shortest routes, while this study is not limited to the shortest paths between the origin station and the terminal, but seeks the through passenger density maximization path in all possible routes. This can be explained by two reasons. First, the passenger flow is not always the largest on the shortest path (In fact, it cannot possibly be.), which means it is not reasonable to lay all routes along the shortest path. Although this can simplify the models and reduce the calculations, it unfortunately damages the quality of the design. On the other hand, when the objective is set to be maximizing through passenger flow, due to the increase of the passenger flow in accordance with the increase of the length of the transit line, a certain route may be abandoned because it is comparatively shorter, even if this short route abounds in passenger flow. This can result in a deviation of the route from the needs of the passenger flow. In addition, the overall length of the entire network increases where longer line is laid, which consequently increases the operating costs without full utilization of the network or fleet. Therefore, this paper employs direct-through passenger flow density maximization as the optimization objective, so as to enhance the network utilization rate as high as possible. The densest routes are laid first to facilitate passenger trips and to benefit the transport enterprises.

Network design is an NP-hard problem difficult to get solved via traditional approaches. In recent years, a large number of researches have showed that Heuristic Algorithm is suitable to large scale optimization problems. Based on the previous researches, this study also adopts a heuristic algorithm: Ant Colony Algorithm. ACA is a new optimization algorithm proposed by Italian researchers Dorgo M. *et al*. It is a colony based optimization approaches inspired by food-seeking actions of ant colonies. The algorithm is not tied to the mathematical descriptions of specific problems, but enjoys the overall optimization capacity and the parallel capacity in essence. In the meantime, it benefits from stronger robustness, shorter needed time, and easier computerization than earlier evolution algorithm such as genetic algorithm and simulated annealing algorithm. ACA has been successfully applied to solving some classic compounding optimization problems, e.g. TSP, QAP and Job-shop. In order to achieve the optimum solution more effectively this study adopts a new objective-function-based heuristic pheromone assignment approach, and adjusts the path-seeking activities of the ants in light of the target functions. To expedite the calculations, multi-computerized ant colony algorithm is applied under the cluster situation.

2. OPTIMIZATION MODELING

2.1 Urban Transit Network Design

Urban transit network planning and design (UTNPN) must be based on passenger *O-D*, and should aim at facilitating trips as well as fostering the transit enterprises' profits. The network design ought to meet certain criteria other than solely leans on experience. Experience can then be referred to help adjust and optimize the outcomes of the design. Urban transit network is based on the road network, each transit line is a continuous path that connects the adjacent sections; the aggregate of these transit lines make up on the transit network. To meet the passengers' needs, an effective transit network carries the following characteristics:

1) Reach ability, i.e. making most of the capacity to meet the demand of the entire network;
2) Low transfer rate, i.e. providing the passengers with as much direct service as possible;
3) Short travel time, i.e. laying out the transit lines according to distances to reduce the overall passenger travel time of the whole service area;
4) High network efficiency, i.e. prioritizing the layout of those transit lines with the densest passenger flow to utilize the network and the vehicle capacity.

2.2 Transit Network Optimization Model

Albeit there are many techniques on transit network optimization, most of them are confined to theoretical research and are practically infeasible, whereas "Designing line by line, and optimizing the lines into network" (Wang, 2001) is practical and convenient. In his previous research, however, direct-through passenger flow was mostly regarded as the optimization objective, line distance was as the constraint, or, as in some cases, the shortest distance between the origin and terminal was applied to lay out the transit line. The single objective of achieving direct-through passenger flow maximization may cause the transit lines in the network excessively long, while the insufficiency in setting the layout constraints may restrict the alternative options and affect the ultimate optimization quality. The trend of the line and the direction of the majority passenger flow are therefore inconsistent. In order to overcome these disadvantages, this paper adopts *maximum through passenger flow density* as the optimization objective, i.e. minimum transfer rate and maximum passenger flow per unit length. Moreover, a series of constraints are adopted.

The ultimate goal of the transit network is to facilitate passenger trips. Then the network design must be grounded on trip demand. To identify a transit line, i.e. to identify the stops and road sections it covers, there could be some transit lines between two adjacent zones, shown as Fig.1. Because inter-zone passenger flow is considered, the passenger flow is invariable, i.e. SP_{ij}, whichever path is selected. Therefore, ideally, the shortest path, hereby α

should be selected to shorten the passenger travel distance and to increase the efficiency of the line. In this way the inter-zone path selection problem can be simplified to an inter-zone shortest path problem, while the passenger flow remains unchanged. The inter-zone passenger flow is correlated only with the zones. As a result, once the sequence of the zones is fixed, the transit line is accordingly established, which means that the transit line optimization problem can be simplified to a zone-sequence identification problem. Unlike laying the transit line with the shortest *od* path, the partially shortest path between adjacent zones is adopted. Fig. 1 illustrates the difference between these two methods. Assume that *o,d* is the original-terminal zones of the line to be searched, and *l*. *okld* is the shortest path between *od, oijd* represents another path. It can be seen that Q_{od}^{oikjd} is obviously longer than Q_{od}^{okd}. According to the shortest path method, *okd* should be selected. However, this may result in deviation. This paper is based on the objective to achieve maximum direct-through passenger flow density, wherefore both the length of the line and the corresponding through passenger flow are taken into account. Consequently, the selected path may well be *oikjd* other than *okd*. Compared with the shortest path method, the method adopted here is more consistent with the direction of the maximum passenger flow.

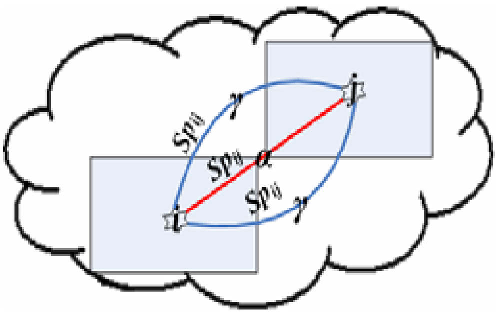

Fig.1 Path selection between two adjacent zones

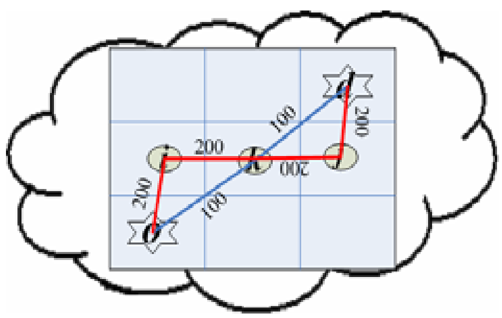

Fig.2 Path selection between O&D zones

According to the network design principles in the previous section and in light of the model for transit network design by Wang (2001), we establish a transit network optimization model.

$$\max \quad D_{od} = \frac{\sum_{i\in N}\sum_{j\in N} SP_{ij}x_{ij}}{\sum_{i\in N}\sum_{j\in N}\Delta_{ij}l_{ij}x_{ij}} \qquad s.t. \begin{cases} L_{min} \le L_{od} \le L_{max} \\ Q_{od}^{sum} > Q_{min} \\ q_{od}^{x} \le q_{max}^{x} \\ b_{od}^{n} \le b_{max}^{n} \\ Q_{od}^{kl} < Q_{max}^{kl} \\ \forall l_{ij} > 0.5km \\ \forall m \ne \forall n \qquad m,n \in S_{od} \\ NTR > 50\% \\ o,d \in N \end{cases} \tag{1}$$

where D_{od} = direct-through passenger flow density;

o = the origin zone; d = the destination zone;

SP_{ij} = the direct-through passenger flow between zone i to zone j within the network;

$L_{od} = \sum_{i\in N}\sum_{j\in N}\Delta_{ij}l_{ij}x_{ij}$ = the length of the transit line;

L_{min} / L_{max} = the minimum/maximum length of the transit line;

l_{ij} = the length of the road between i and j;

$$\Delta_{ij} = \begin{cases} 1 & \text{when } i, j \in S_{od} \text{ and } i, j \text{ are adjacent dots on } S_{od} \\ 0 & \text{other} \end{cases}$$

q^{x}_{od} / q^{x}_{max} = the non-linear rate / the maximum non-linear rate;

Q^{kl}_{od} = the passenger flow of section k, l;

Q^{kl}_{max} = the maximum allowable passenger flow of section k, l;

Q^{sum}_{od} = the total passenger flow of the line;

Q_{min} = the threshold passenger flow of the transit line;

b^{n}_{od} / b^{n}_{max} = section non-equilibrium factor of passenger flow / the maximum non-equilibrium factor;

NTR = non-transfer ratio the average transfer times of the entire network.

This is a non-linear integer programming problem, with constraints as follows:
Length constraint: appropriate length of the line is 30-40*min* per single-trip, with a minimum of 20*min* and a maximum of 45*min* (for small and medium cities) or 60*min* (for large cities). If the average trip speed is 15*km/h*, the minimum line length (L_{min}) is 5*km*, and the maximum line length (L_{max}) is 11.25*km* (for small and medium cities) or 15*km* (for large cities). Generally, a transit line cannot be approved if its length does not fulfill the constraint of the minimum and maximum lengths, i.e. $5km \le L_{od} \le 15km$.

Minimum line-setting passenger flow constraint: in order to ensure higher efficiency and higher economic benefits to the transit enterprises, a transit line should not be opened between zones of very low passenger flow.

Non-linear constraint: the smaller the non-linear rate, the better the line. The appropriate rate of a generic city is 1.15~1.20, and should be less than 1.5 for a single transit line.

Section non-equilibrium factor of passenger flow constraint: the section non-equilibrium factor of passenger flow is the proportion of the maximum section passenger flow to that of the average section passenger flow. Generally, the factor should not exceed 1.5.

Section passenger flow constraint: the section passenger flow of a line must be less than the section capacity. Otherwise only part of the passenger flow can be serviced, while the over ceiling passenger flow remains surplus.

Station spacing constraint: according to the suggested value in the *Code for Transport*

Planning on Urban Road, the average station spacing shoube be l_{ij}=0.5~0.6*km*.

Trend constraint: **a** transit line should not contain a loop; namely, the same line should not pass the same station twice or more than twice.

Network passenger non-transfer ratio constraint: once the transit network is formed, actual non-transfer ratio can be identified. The equation is as:

$$\text{Non-transfer ratio} = \frac{\text{Total through passenger flow}}{\text{Total passenger O-D of the planned area}} \tag{2}$$

If there is a comparatively big discrepancy between the calculated non-transfer ratio and the adopted non-transfer ratio, the network should be re-designed by using the calculated ratio, until the adopted non-transfer ratio approximates to the calculated ratio.

3. PARALLEL ANT COLONY ALGORITHM

The model is a large-scale combination optimization problem, and then this paper adopts ACA to solve it. The ACA is a heuristic search algorithm applied to combinatorial optimization. Dorigo *et al* (1996) applied the ACA to solve the TSP, and extended the technique to solving non-equilibrium TSP, QAP, and Job-shop. To overcome the possible stagnation in Ant-Q, Stützle *et al* (1999)proposed the MX-MIN AS, i.e. MMAS, which improved the fundamental ant algorithm from three aspects: 1) The initial pheromone value is set to its maximum value, $\tau_{\max}$, to foster more adequate optimization search; 2) Only those ants who modify their shortest path after a cycle can alter or add the pheromone; 3) to avoid prematurely converging the overall optimal solution, the pheromone density of each path is constrained within $[\tau_{min},\tau_{max}]$. Gambardella *et al* **(**2000**)** put forward a hybrid ant system (HAS), in which the ants establish their own solutions in every cycle and use these solutions as the start points to search local optimal solutions by some local search algorithm as the corresponding ants' solutions. This can enhance the quality of the solutions in a short time. Botee *et al* **(**1999**)** further studied the selection of parameters m, α, β, and ρ by genetic algorithms.

3.1 Principles of the ant colony algorithm

ACA is essentially a system inspired by studies of the behavior of real ant colonies. The principles of the algorithm can be illustrated by examining the food searching process of an ant colony. Along their way from the food source to the nest, ants communicate with one another with pheromone (a chemical substance). The pheromone gradually evaporates over time. As the ants move, a certain amount of pheromone is deposited on the ground along the path they follow, marking the path with a trail of substance. The ants, then, determine their movements by judging the density of the chemical substance on a path. This process can be described as a loop of positive information feedback, in which the more ants follow a given trail the more pheromone is left on that trail, and the larger probability that this trail will be followed by other ants. This selection process is the ant's self-catalyzing activities, by means of which at last the ants find the optimal path to the destinations.

3.2 Ant colony algorithm analysis

By simulating the ants' real behavior, artificial ants are made by supposing 1) an artificial ant has a certain degree of memory capability enabling it to memorize the paths it passes; 2) a artificial ant is abandoned once it finds a path leading its nest to the food source; 3) an artificial ant is not blindfold in searching for the next path but is doing so consciously with a specific purpose; 4) the environment in which the artificial ant stays with is discrete. ACA usually consists of four elements: initialization, transition rule, update rule, and terminate condition. The ants herein below, unless indicated in particular, refer to artificial ants.

(1) Transition rule

An ant's movement is not blind but is in accordance with certain transition rule. When the ants move, there are some nodes which do not satisfy the constraints and will not be visited. For instance, in the UTNPN problem, the nodes which have already been visited are the unfeasible nodes. The transition rule refers to the probability for an ant choosing a feasible node. For the k_{th} ant at the i_{th} node, the probability to choose the next node, j, is as follows:

$$p_{ij}(k) = \begin{cases} \dfrac{\tau_{ij}^{\alpha} \times \eta_{ij}^{\beta}}{\sum_{h \notin tabu_k} \tau_{ih}^{\alpha} \times \eta_{ih}^{\beta}} & j \notin tabu_k \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

where τ_{ij} = the pheromone density of edge (i, j);

η_{ij} = the visibility of edge (i, j);

α, β = the information heuristic factor, the visibility heuristic factor;

$tabu_k$ = the aggregate of the unfeasible nodes for the k_{th} ant.

The pheromone provides an indirect means of communication among the ants. In other words, the ants can communicate with one another by sensing the pheromone density. The pheromone density on edge *(i, j)* reflects the previous experience of the ant colony on the edge. It reflects the overall information accumulated in the ants' moving process, i.e. the residual information $\eta_{ij.}$ On the other hand, the value in terms of visibility is derived from a greedy method which is relevant to the original problem. This method only takes the local information on edge *(i, j)* into account, and therefore reflects the heuristic information (visibility η_{ij}) in ants' movements, e.g. the length. These two aspects are interdependent and closely correlated regarding their influences upon the performance of the ant colony algorithm. In sum, the task of the transition rule is to find a balance between the random and the certainty (Botee *et al*, 1999).

(2) Update rule

The ants travel along different nodes in accordance with the transition rule, until a solution to the original problem is established. For example, in the UTNPN problem, one solution for the ants is to search a path leading the start point to the end point. A cycle of the ACA is defined as when all ants have established their own solutions. The pheromone density on each trail will be updated after each cycle is completed.

$$\begin{aligned} &\tau_{ij}(t+1) = \rho \times \tau_{ij}(t) + \Delta\tau_{ij} \qquad \rho \in (0,1) \\ &\Delta\tau_{ij} = \sum_{k=1}^{p} \Delta\tau_{ij}^{k} \end{aligned} \tag{4}$$

where $\Delta\tau_{ij}$ = the amount of the added information in the cycle;

ρ = the residual pheromone coefficient;

$\Delta\tau_{ij}^{k}$ = the amount of added pheromone of the k_{th} ant on edge *(i, j)* in the current cycle.

In a real ant colony system, the pheromone density tends to be higher on a shorter path. Similarly, in the ACA, the path that most approximates the optimal scenario carries more pheromone, which makes it more attractive in the next cycle.

There are three pheromone updating models: a) Ant-density; (b) Ant-quantity; (c) Ant-cycle. Based on them, we adopt a new update models: Ant-Weight. This update rule ensures that the pheromone distribution of the searched paths is direct proportion to the solution optimization. The more favorable the path, the more pheromone allocated to it, and more accurate directive information is provided for later search. But because each link obviously contributes a different proportion to the objective function, more pheromone increment should be allocated to favorable links, while less pheromone increment should be allocated to the less favorable. In this way, the valid information obtained from the previous search can be retained for further and more careful search in a more favorable area, which helps speed up the convergence of the algorithm. At the same time, the effectiveness of a large-scale search can be ensured to facilitate the algorithm to find the overall optimal path.

$$\Delta\tau_{ij}^{k} = \begin{cases} \dfrac{Q}{f^{k}} \times \dfrac{f_{ij}^{k}}{\sum\limits_{i,j\in S_{od}} f_{ij}^{k}} & \text{when the kth ant visits edge}(i, j) \text{ in the tth cycle} \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

where Q remains a constant.

f_{ij}^{k} = the target function value of the k_{th} ant on edge *(i, j)*.

f^{k}= the target function of the k_{th} ant in the entire path.

In order to prevent the non-optimal path in the network from being visited by mass ants within a short time, aggregating pheromone to excess, and taking a dominant position in the network, and to prevent the optimal path from being unvisited and its probability to be chosen being lowered as the pheromone on it evaporates, upper and lower limits [τ_{min},τ_{max}] are introduced to the pheromone on each edge (Stützle *et al*, 1999) to avoid local optimization and to enlarge the probability of gaining a higher-quality solution. These mechanisms provide the ACA with a strong capability in terms of finding better solutions. In this paper we adopt the heuristic pheromone update rule for the objective function value to improve the ants' searching quality. In addition, a parallel ACA is adopted to decrease the search time.

3.3 Coarse-grain parallel ant colony algorithm

A huge amount of calculation is generated from implementing the ACA in practice, thus parallel implementation of ACA exerts an important role. There are three kinds of parallel ACA: independent parallel ant colony algorithm, Master-Slave parallel ant colony algorithm, and coarse-grain parallel ant colony algorithm. The CPACA is quite similar to the impendent ant colony parallel algorithm mentioned above. The difference is that it can bring information exchange among the sub-colonies once it has come through a fixed evolution process. The initialized ant colonies are divided into several sub-colonies according to the number of the processors. The sub-colonies are then distributed to corresponding processors to evolve independently. Later, as the rule works, the sub- colonies will transfer the "outstanding ants" to other sub-colonies. The transition can employ multi-topological structures. Here we employ the ring topology. By exchanging the "outstanding ants" between sub-colonies to introduce their best pheromone information, the path selection of the ants is diversified to effectively prevent the premature convergence.

In the implementation of parallel ACA, if an algorithm involves substantive communication between the processors, the parallel effect will be greatly sacrificed. All of the algorithms discussed above take place in a distributed environment other than share the same memory system. This is because the distributed structure is more widely seen and used. By comparing the communication cost and solution precision of each model, we can see that the coarse-grain parallel algorithm is comparatively applicable in that it requires little communication and can speed up the convergence while ensure the quality of the solution. Therefore, it is more suitable to adopt in the distributed environment.

4 THE ALGORITHM PROPOSED IN THIS PAPER

The essence of the transit network design is the optimal setting between the origin station and the terminal. Different settings can form different transit lines with different direct-through passenger flow densities (D_{od}). This is very similar to ACA described as above. If we take the origin station as the *nest* and the terminal station as the *food*, the UTNPN problem can be simplified as a process by which the ant colony searches for *food* starting from the *nest* by means of the pheromone deposited, namely, searching for an optimal transit line from the origin station to the terminal judging by the passenger flow density. Because the Enumeration Method is the only effective solution to this problem and it involves a huge amount of calculation when it comes to large-scale practical problems, we are adopting the parallel ACA in this paper. In addition, as the algorithm will be exerted under the cluster environment, the CPACA is adopted to complete the communication between the sub-ant colonies by MPI. The algorithm first generates m initial sub-ant colonies, each consisting of p ants. All of the sub-ant colonies maintain their own pheromone matrix independently, and independently search for the optimal transit line between the origin station and the terminal according to the "pheromone", i.e. the passenger flow density, along the given network. When all of the p ants in each sub-colony have completed the search task, the line with the largest passenger density among all alternative viable lines is chosen as the line to be laid in this cycle. The search process repeats until all lines are completely laid out. The specific steps of the algorithm are as follows:

Step1 Initialization

The ant colony should be initialized after completing the transit network initialization by generating m sub-ant colonies and identifying their topological structures, transition intervals, and transition scales, followed by initializing the communication environment between the sub-ant colonies.

First an appropriate initial weight needs to be allocated to all edges (Stützle *et al*, 1999). Since the passenger density is regarded as the "pheromone" deposited by the ants on the path, we use the average passenger flow density to initialize the pheromone matrix, specifically written as:

$$\bar{\tau} = \sum SP_{ij} / \sum l_{ij} \tag{6}$$

where $\bar{\tau}$ refers to the average passenger flow density of the entire transit network. At last, the forager ants can be distributed to the graph. Because each path carries the same amount of pheromone, the ants can be randomly distributed to the nodes nearby the nest.

Step2 Transition rule

The transition rule refers to the probability of each path to be chosen by the ants during their path-searching process. It is determined by the pheromone density τ_{ij} and the visibility value η_{ij} of the corresponding edges. In the algorithm in this paper, τ_{ij} is derived by updating the pheromone on each edge upon completion of a cycle, based on the transition rule discussed below; while η_{ij} is derived by a greedy method which encourages the ants to visit the locally optimal path. As the target function value is derived by maximizing the passenger flow density of each line, let $\eta_{ij} = d_{ij}$, where d_{ij} is the passenger flow density of link ij: $d_{ij} = SP_{ij} / l_{ij}$. We define the probability for an ant to transfer from zone i to zone j as formulation 3.

Step3 Path search

Before the path search, the viable origin o and terminal d should be chosen. For instance, an OD pair is regarded as non-viable if it lies on a line that has already been laid out or does not satisfy the constraints. After choosing the origin and destination terminals, each sub-ant colony starts searching paths between the OD pair independently. To carry out the path search, the zones whose distances to the current zone k are less than the pre-scheduled station spacing (0.5-0.8) are first chosen to form an aggregation of alternative zones, whereafter the next zone l is chosen according to the ant transition rule. Each line is valuated in turn after the calculation.

Step4 Pheromone update rule

The network needs to be updated after a line is identified. Namely, the pheromone in the network needs to be updated.

1)Pheromone distribution rule

The network optimization is completely dependent on the direction of pheromone τ_{ij} and visibility η_{ij}. The visibility η_{ij} is comparatively stable, so the pheromone τ_{ij} becomes particularly important in terms of searching for new lines and better solutions. When updating the pheromone, a basic ACA distributes the incremented pheromone with equilibrium for all paths the ants have visited. Due to the increment of the pheromone, some unfavorable paths are subject to large probabilities to be chosen in the next search cycle, while if the optimal path is not visited, the pheromone on it will evaporate gradually, which leads to a lowered probability for it to be chosen in the next search. This may result in false directive information and a large amount of invalid searches. Therefore, we have adopted a pheromone distribution rule based on the target value: Ant-Weight. Specifically, the pheromone update rule is as formulation 4 and formulation 5.

2) Migration strategy

In order to diversify the selection scope of paths, new paths are continuously exploited to get better solutions. In addition, an “ant transition” is introduced, in which some outstanding ants are transferred to other sub-ant colonies after some time of search, stimulating the ants to find more favorable paths.

It is very important to identify the transition epoch and the transition scale (n_m). If the either epoch is too long or the scale is too small, the transition will not be reacting on the sub-ant colonies to exploit new paths, and the immigrant ants’ survival tends to be few. On the other hand, if either the epoch is too short or the scale is too large, the diversity of the sub-ant colonies will be ruined, which causes a premature convergence. Hence, the transition epoch and scale are determined according to specific situations. By simulation, a transition rule with an epoch of 5 and a scale of 1 is found to be able to ensure the solution quality as well as the fastest convergence speed. In addition, the topology of connection between the sub-ant colonies needs to be identified before the transition is implemented. Matsumura (1998) has found that unidirectional ring topology not only facilitates the excellent genes to diffuse in the colony, but separates the sub-colonies effectively so that the diversity of the colonies is protected. This paper also adopts the unidirectional ring topology to connect the sub-ant colonies. Thus, the search task of an individual ant is completed. *Step 4* is retrieved to continue the search cycle, until all ants of the sub-colonies have completed the search tasks.

Step5 Laying lines
After all ants of the sub-ant colonies have completed a search cycle, the line with the largest passenger flow density between OD is chosen to be an alternative line. The stops it covers are sequentially added into aggregate S_{OD}. Thus, the search task for lines between OD is completed, and we return to Step 3 for the next OD path search. This process repeats till all valid lines between OD are searched out. In this wise, no more than one alternative line is generated between each valid OD. The collection of these alternative lines is written as S. The collection of S is derived after the search tasks for all valid lines between OD are completed, whereafter the lines are laid. The key to the transit line layout lies in identifying the direct-through passenger flow density (D_{OD}) after matching the origin and destination terminals. D_{OD} is related to the stops the line covers and the line length. The line with the largest passenger flow density in the entire S is then chosen to be added into the network, after which relevant data are updated.

Step6 Revising the network
After line layout, the passenger flow carried by a line needs to be subtracted from the original passenger flow matrix. The passenger flow matrix revision is carried out according to the method as follows. Firstly, the total passenger flow of each line section (including the existent passenger flow V^{kl}) and the carrying capacity are calculated. If the carrying capacity of every section of a line outweighs its passenger flow, all passengers on the line would be carried by that same line, and the inter-station passenger flows should be subtracted from the passenger flow matrix. If the section passenger flow is less than the total section passenger flow, the line can only carry part of the passengers, and the residual passenger capacity needs to be calculated followed by subtracting the carried passenger flow from the original passenger flow matrix. Lastly, the OD pairs of the lines laid are deleted from the terminal matrix. So far,

the line with the largest passenger flow density in this cycle is laid to the network. The cycle completes after revising the network. The next step is to determine whether or not the cycle terminates. The outcomes are exported if it is to be terminated; otherwise, *Step 2* is retrieved to continue the layout. The search process is not terminated until there are no lines complying with the set constraints in the network, or the cycled times have reached the pre-scheduled level.

5. NUMERICAL TEST

To examine the model and the efficiency of the algorithm, Data in Dalian city is used for numerical test. The population in Dalian is 2 million, and the build-up area is 180 km^2. Dalian's road network consists of 3200 links and 2300 nodes. At present, there are 89 bus lines and 1500 bus stops. Fig.3 illustrates the situation of the bus network. We partition Dalian city into 800 zones and get bus OD matrix through on-board survey of all of the 89 lines. Then by taking the zonal centers as bus stops, we design the bus network with the developed model and algorithm. The parameters in the ACA are estimated through simulation (Table 1). There are 8 ant colonies, within each colony there are 30 ants. Then we carry out our model and algorithm with Microsoft Visual C++.Net 2003 on the cluster environment formed by 8 computers. Fig.4 illustrates the results of the CPACA. Total 61 bus lines are designed, which extend 692km. The designed bus network can basically satisfy the trip demand. By comparing the optimized with current bus networks, we get following results.

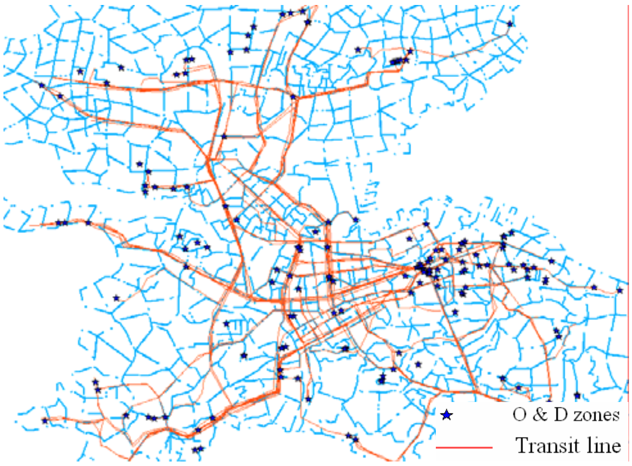

Fig.3 Current bus network

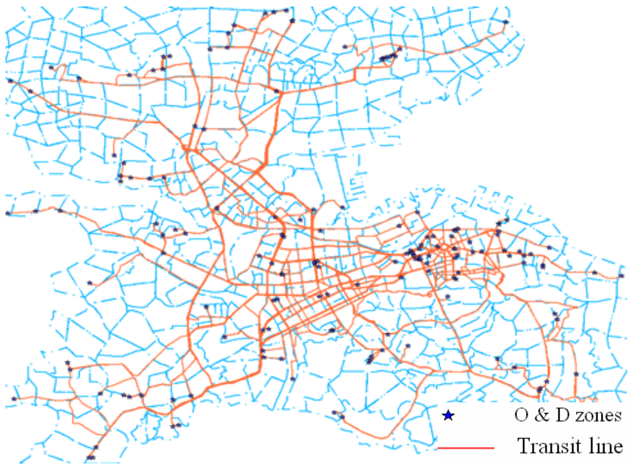

Fig.4 The optimized bus network

Table 1 Parameters in CPACA

m	p	α	β	Q	*epoch*	n_m
6	30	2	1	1000	10	1

1) Fig.5 illustrates the direct-through passenger densities of optimized and current bus network. We can see that direct-through passenger density of current bus network is about 28.8, much lower than the optimized 42.7. This is mainly because that current bus lines overlap each seriously to disperse the trip flow, thus lowers the efficiency of the network.

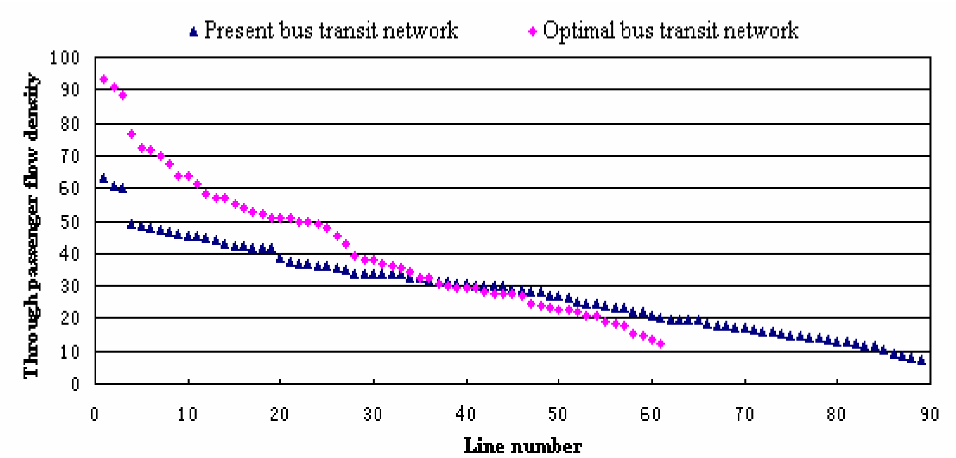

Fig.5 The direct-through passenger densities of optimized and current bus network

2) In optimized case, direct-through passengers shares 51%, while in current network it is about 41%.
3) As Fig.6 shows all of the optimized bus lines are shorter than current ones.
4) As Fig. 6 shows no-linear rates of many current lines are high, some of them are even over 3.4, the averaged non-linear rate of current bus lines is about 1.75, while the averaged one of the optimized lines is about 1.52.

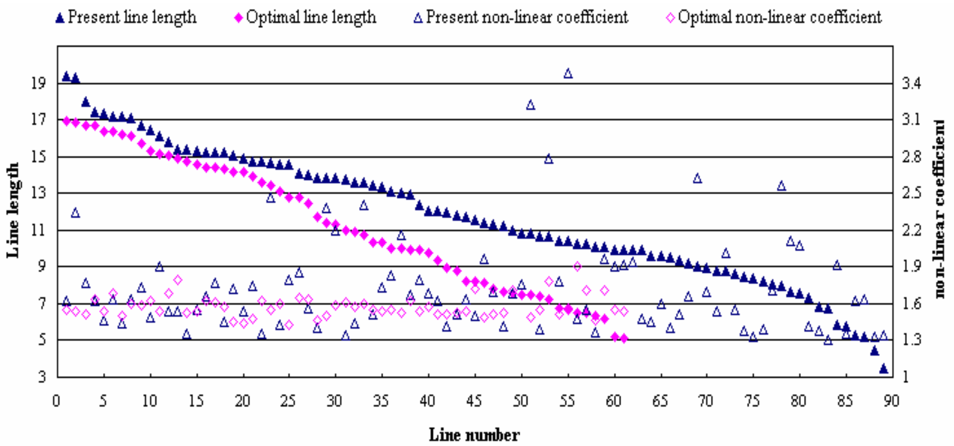

Fig.6 Line length and non-linear coefficient of optimized and current bus network

At last, efficiencies of CPACA and ACA are compared. We test 10 times for the two algorithms with the same road network. From Fig.7, we found that the optimized result of CPACA is better than that of ACA.

It can be said that since too many zones exist and the constrains are very complicated, thus although we set upper and lower limits for the phenomena, searching work still tends to be trapped in local optimal solution. However, introduction of Migration operation in CPACA diversified the ant colony and widened the searching space, problem of being trapped in local optimal solution

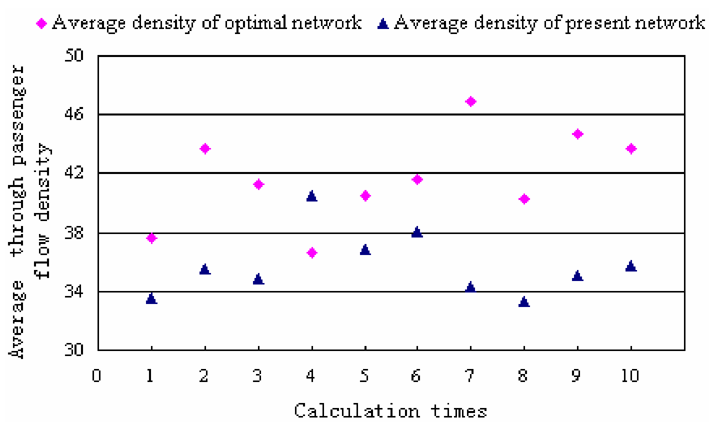

Fig.7 The comparison of CPACA with ACA

is improved somewhat. Moreover, since searching time is much more than commuting time between two ant colonies, we think CPACA can improve the optimization quality with same numbers of the ant colonies.

6 CONCLUSIONS

Model for optimizing urban transit network has been developed, which takes maximum direct-through passenger density as goal and considers benefits of both passengers and transit companies. It means that minimizing the average trip time, namely to make as many as passengers to travel between original and destination without transfer. On other hand, the model maximizes the profit of bus companies, namely to raise operation efficiency and shorten the total bus lines. Based on ACA we developed an algorithm for the model. In order to improve the optimization and quick the calculation, CPACA is adopted. We found that coarse-grain model with less communication is more suitable in cluster environment. Finally, with data of Dalian city we test our model and algorithm and compare some indices between the optimized and current bus networks. And merits of CPACA are also illustrated.

Improvement of the searching efficiency is our further study. Since searching the route is time-consuming, then the algorithm runs a long time. We can make the improvement from two aspects: 1) simplifying the road network, 2) improving efficiency of ACA. Moreover, generation of only one line in each an iteration humped the line design afterwards, and then the final result may not be the global optimized one, some areas with less passenger density may become un-served area of transit network.

REFERENCE

Baaj, M. H. & Mahmassani, H.S. (1995), Hybrid route generation heuristic algorithm for the design of transit networks. Transp. Res., Part C (Emerg. Technol.). 3(1),31-50.

Botee HM, Bonabeau E. (1998) Evolving ant colony optimization. Complex Systems, 1(2):149~159.

Ceder, A. & Wilson, N. H. M. (1986), Bus network design, *Transportation Research—B*, 20B(4), 331–44.

Chakroborty, P., Deb, K. & Subrahmanyam, P. S. (1995), Optimal scheduling of urban transit system using genetic algorithms, *ASCE Journal of Transportation Engineering*, 121(6), 544–53.

Dashora, M., and Dhingra, S.L. (1998). Expert system for bus routing and scheduling. Computers in urban planning and urban management, Vol. 2 P.K. Sikdar et al., eds., Narosa Publishing, New Delhi, India, 841-853.

Dorigo,M. (1992) Optimization, learning and natural algorithms, Ph.D. Thesis, Politecnico di Milano, Italy.

Dorigo, M., Maniezzo, V., Colorni, A., (1996). The ant system: Optimization by a colony of cooperating agents. IEEE Transactions on Systems, Man and Cybernetics 1 (26), 29– 41.

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.

Fernandez, R. (1993). An expert system for the preliminary design and location of high capacity bus stops. Traffic Eng. Control, 34,533-539.

Gambardella LM, Dorigo M. (2000) An ant colony system hybridized with a new local search for the sequential ordering problem. INFORMS Journal on Computing, 12(3):237~255.

Hasselstrőm, D. (1981). Public Transportation Planning - A Mathematical Programming Approach. Doctoral Dissertation, University of Gőteborg, Sweden.

J. F. Guan (2004) presents an early stage model of simultaneous optimization of transit line configuration and passenger line assignment in a minimum spanning tree network.

MatsumuraT, NakamuraM, OkechJ, etal. (1998) A parallel and distributed genetic algorithm on loosely-coupled multiprocessor system [J]. IEICE Trans. Fundam. Electron. Commun. Comput. Sci., E81A(4),540-546.

Michael R. Bussieck, Peter Kreuzer, Uwe T. Zimmermann, (1997). Optimal lines for railway systems. European Journal of Operation Research, 96(1) 54-63.

Senevirante, P.N. (1990) Analysis of on time performance of bus services using simulation. J. Transp. Eng., 116(4),517-531.

Simonis, C., (1981). Optimierung von omnibuslinien. Berichte stadt – region – land, Institut für Stadtbauwesen, RWTH Aachen.

Sonntag, H., (1979). Ein heuristisches verfahren zum entwurf nachfrageorientierter inienführung imöffentlichen personennahverkehr. Z. Oper. Res. Ser. A-B, 23: B15–B31.

Steenbrink.P.A (1974). Optimization of transport networks, Wiley,London.

Stützle,T. H. Hoos, (1999) *MAX-MIN* Ant System and local search for combinatorial optimization problems, in: S. Voß, S. Martello, I.H. Osman, C. Roucairol (Eds.), Meta-heuristics: Advances and Trends in Local Search Paradigms for Optimization, Kluwer Academic, Boston, MA, pp. 313–329.

Wang Wei, Yang Xinmiao(2001). Planning and Management Technology for Urban Transit System, Science Press, Beijing.