RAPID TRANSIT NETWORK DESIGN USING SIMULATED ANNEALING

Shahab Kermanshahi, Department of Civil Engineering, Sharif University of Technology, Azadi Ave., Tehran 11365-9313, Iran. {kermanshahi@civil.sharif.ac.ir}

Yousef Shafahi, Department of Civil Engineering, Sharif University of Technology, Azadi Ave., Tehran 11365-9313, Iran. {shafahi@sharif.edu}

Mostafa Mollanejad, Department of Civil Engineering, Sharif University of Technology, Azadi Ave., Tehran 11365-9313, Iran. {m.mollanejad@mehr.sharif.ir}

Mahmood Zangui, Department of Civil Engineering, Sharif University of Technology, Azadi Ave., Tehran 11365-9313, Iran. {m.zangui@sharif.edu}

ABSTRACT

Rapid Transit Network Design (RTND) is a relatively new type of Transit Network Design (TND) that aims to find the best set of rapid transit routes. It considers special characteristics of rapid transit routes, and can thus fully exploit the potential of transportation networks for creating them. In this study, a meta-heuristic algorithm is proposed for RTND, namely a simulated annealing algorithm (SA), which has been adapted for RTND. The objective of the algorithm is to maximize rapid transit network coverage. For achieving this goal, the route generation procedure first specifies a set of candidate routes. Some practical assumptions are made about the conditions of an acceptable route in this step. Next, SA is applied to perform the search process through the candidate set. Some innovative approaches are applied in order to speed up the SA algorithm. The algorithm is implemented in C# and a benchmark problems is used to evaluate its efficiency. Comparison between this algorithm and three previous studies proves the reliability and effectiveness of the method. Moreover, the most important aspect of this method is that it depicts excellent efficiency and accuracy as a practical method for being used for large-scale networks.

Keywords: rapid transit, network design, simulated annealing.

INTRODUCTION

Having an effective public transportation system is essential in achieving a sustainable transport system, simply because satisfying the increasing demand, just by construction of new streets and improvement of the existing ones, is not a cost-effective and practical approach. As a result, expansion and improvement of the public transportation system is a common strategy for all transportation planners. There are, however, many obstacles in the path, the most important one of which is transit network design.

Network design is the first stage of transit service planning process (Ceder and Wilson, 1986). The main goal of Transit Network Design (TND) is to define a set of routes, each route being determined by a set of consecutive links and stations. The most frequent among the objectives used in this stage are minimizing passengers travel time (Lee and Vuchic, 2005), minimizing operational costs (Wan and Lo, 2003), combination of these two objectives (Borndorfer et al., 2005; Tom and Mohan, 2003), and combination of these two and other objectives like maximizing service coverage and minimizing fleet size (Ceder, 2003). Besides, some studies put the focus on the special aspects of this area. For instance, Zhao and Ubaka (2004) consider number of transfers and routes directness, and Lee and Vuchic (2005) consider variable demand. Due to the direct effect of service frequencies on passengers travel time and required fleet size, frequencies can also be set in this stage in many studies.

Because of the complexities of transit network design problems, considering all of its aspects is impossible. Therefore, different objective functions and simplifying assumptions are used in various studies and thereby, achieving the exact solution has been possible in some of them. Examples of these are Wan and Lo (2003), and Guan et al (2003). Borndorfer et al. (2005) and Barra et al. (2007) are other examples that, respectively, apply column generation and constraint programming to find the exact solution for TND. Despite the simplification in these studies, exact solution only can be found in small and medium-sized networks. This issue has led the researchers to employ approximate methods to solve the problem.

Generally, approximate methods for solving transit network design problems are classified to heuristic and meta-heuristic methods. Examples of the heuristic approaches can be seen in the work of Shih and Mahmassani (1998), and Mauttone and Urquhart (2009). Along with heuristic methods and in more studies, meta-heuristic methods are applied. Genetic Algorithm (Tom and Mohan, 2003; Nqamchai and Lovell, 2003), Tabu Search (Fan and Machemehl, 2008), and Simulated Annealing (Fan and Machemehl, 2006) are some meta-heuristic methods which are commonly used for TND problems.

Rapid Transit Network Design (RTND) is a type of transit network design that considers special characteristics of rapid transit networks to design them to be more practical and more specialized. One of the most important characteristics of rapid transit networks is having more direct and high-frequency routes. RTND includes defining a set of rapid transit routes that maximize an essential index such as network coverage.

In recent years, rapid transit network design has been debated in various studies. Laporte et al. (2005), which is one of the first of such studies, develop as an integer programming problem and solves it for small networks. The main objective of the model is to maximize the transit network coverage. Garzon et al. (2005) is another study that analyzes

RTND problem in regard to transfers. In this paper, the start and the end of all transit routes (terminals) are assumed to be predetermined. Marin (2007) extends the RTND problem and solves it for a larger network (9 nodes and 15 links). In this study, some constraints of previous studies are relaxed and problem is solved more comprehensively.

Because of the intractable computational time that was required to solve RTND problem, developing faster methods became one of the main goals of subsequent studies. For instance, Marin and Jaramilo (2009) use accelerated Benders Decomposition to reduce the solution time. Moreover, Escudero and Munoz (2009) apply a two-stage approach and use some simplifying assumptions to decrease computational time. In a later study, Kermanshahi et al. (2010) propose a specialized implicit enumeration algorithm that, in spite of considering the effect of transfers that makes the problem bigger and more complicated, have a significantly shorter solution time.

The present study makes use of a model similar to the one by Kermanshahi et al (2010), and proposes an approximate method to find an acceptable solution for RTND problem in a reasonable time. The main input of the model is a graph with nodes demonstrating major transit stations and arcs standing for potential links, which can be selected as a part of rapid transit routes and connect major stations together. The sequence of links form routes and the set of all routes build up a rapid transit network. Major stations, also, can be assumed as origins and destinations of all passengers and the origin-destination demand matrix is another input of the model used. As the other inputs can be mentioned construction cost of links and stations, private car travel time matrix, and rapid transit travel time on each link.

The objective of this study is to maximize the rapid transit network coverage, which was also used in Laporte et al. (2005), Marin (2007), Marin and jaramilo (2009), Escudero and Munoz (2009), and Kermanshahi et al. (2010). For an OD pair, if rapid transit has an acceptable travel time when compared to auto, network coverage is available. This acceptability is determined with the help of Congestion Factor (μ). In this study, it is assumed that for each OD pair, when transit time is not greater than predetermined multiplier (μ) of auto travel time, the transit system is acceptable and this pair is covered. Another assumption is the independence of auto travel times from transit travel times. Because of this assumption, therefore, private car travel times between OD pairs are predefined as one of the model inputs.

For achieving the above-mentioned objective, first, the feasible routes which are susceptible of presence in the optimal solution are selected and introduced as the set of candidate routes. And then, the optimal combination is selected from these feasible routes. The feasibility is defined by some constraints on the length and directness of routes. Some other constraints are imposed in selecting the optimal route combination.

The subsequent sections of this paper are organized as follows: The next section presents the mixed integer programming model for RTND problem. The objective function and related constraints are also described. Then, the simulated annealing-based approach is proposed to search for the optimal route combination. Next, the proposed solution methodology is applied to a sample networks and the numerical results are also discussed. Finally, the work is summarized and some future plans are discussed.

PROBLEM FORMULATION

In this study, a linear mixed-integer formulation is presented for the rapid transit network design problem. First, the parameters and variables are defined.
The parameters of the model are:

$\mathcal{G} = (\mathcal{N}, \mathcal{A})$ denotes a transportation network, comprised of a set of nodes, $\mathcal{N}$, and a set of undirected links, $\mathcal{A}$. ($|\mathcal{N}| = \mathrm{n}$ and $|\mathcal{A}| = \mathrm{e}$);

W = set of OD pairs, denoted by $\mathrm{w} = 1, \dots, \mathrm{n}^2$;

$\mathrm{o(w)}$ = origin of OD pair w;

$\mathrm{d(w)}$ = destination of OD pair w;

P = set of candidate transit routes, used as an input to the model, denoted by $\mathrm{p} = 1, \dots, \mathrm{k}$;

N (i) = set of neighbors of node i, ($\mathrm{N(i)} \subseteq \mathrm{N}$ and $\mathrm{k} \in \mathrm{N(i)} \Leftrightarrow \mathrm{ki} \in \mathrm{A}$);

$\mathrm{T_{ij}} = \left[\mathrm{t_{ij}}\right]_{1\times e}$ = rapid transit travel time on each network link;

$\mathrm{C} = [\mathrm{c^p}]_{1\times k}$ = construction cost of transit routes;

$\mathrm{U_{pri}} = \left[\mathrm{u^w_{pri}}\right]_{n\times n}$ = private car travel times between OD pairs;

$\Delta = \left[\delta^p_{ij}\right]_{e\times k}$ = transit routes' incidence matrix (δ^p_{ij}=1 if route p traverses edge ij; =0 otherwise);

$\mathrm{D} = [\mathrm{d^w}]_{1\times n^2}$ travel demand matrix;

L = maximum allowed number of routes;

μ = congestion factor;

TP = transfer penalty which is added to transit travel time for every transfer; and

M = a big enough number.

The variables of the model are:

$\mathrm{Y} = [\mathrm{y^p}]_{1\times k}$ = selected routes vector. $\mathrm{y^p} = 1$ if route p is selected; =0 otherwise.

$\mathrm{f^w_{ij}} = 1$ If link ij is used by OD-pair w in transit network; =0 otherwise.

$\mathrm{x^{w,p}_{ij}} = 1$ If link ij is used by OD-pair w in transit network, and is a part of route p; =0 otherwise.

$\mathrm{tr^w_i} = 1$ If OD-pair w transfer at node i; =0 otherwise.

$\mathrm{T} = [\mathrm{t_w}]_{n\times n} = 1$ If OD-pair w uses transit mode; =0 otherwise.

$\mathrm{u_{pub}} = \left[\mathrm{u^w_{pub}}\right]_{n\times n}$ = OD pairs transit travel time matrix, $\mathrm{u^w_{pub}}$ is sum of link travel time and transfer penalties (if any).

$\mathrm{z(Y)}$ = transit network coverage of vector Y.

Using the above definition, the model formulation is as follows:

(P1)

$$Maximize\ z(\mathrm{Y}) = \sum_{w \in W} d_w \times t_w \qquad (1)$$

$$\sum_{p} y^p \leq L \qquad (2)$$

$$\sum_{p} y^p \times c^p \leq B \qquad (3)$$

$$\sum_{i \in N(k)} f_{ik}^w - \sum_{j \in N(k)} f_{kj}^w = \begin{cases} -1 & k = o(w) \\ 1 & k = d(w) \\ 0 & \text{otherwise} \end{cases} \qquad \forall k \in N, \forall w \in W \qquad (4)$$

$$x_{ij}^{w,p} \leq \delta_{\mathrm{ij}}^{\mathrm{p}} \times \mathrm{y}^{\mathrm{p}} \quad \forall (i,j) \in \mathcal{A}, \forall w \in W, \mathrm{p} \in P \qquad (5)$$

$$f_{ij}^w \leq \sum_{\mathrm{p}} x_{ij}^{w,\mathrm{p}} \quad \forall (i,j) \in \mathcal{A}, \forall w \in W \qquad (6)$$

$$\textstyle\sum_{k \in N(i)} x_{ki}^{w,p} + tr_i^w \geq \sum_{k \in N(i)} x_{ik}^{w,p} \quad \forall i \in N, \forall w \in W, \mathrm{p} \in P \qquad (7)$$

$$u_{pub}^w = \sum_{ij \in E} \left(t_{ij} \sum_{p} x_{ij}^{w,p} \right) + \mathrm{TP} \sum_{\mathrm{i} \in \mathrm{N}} tr_i^w \quad \forall w \in W \qquad (8)$$

$$u_{pub}^w - \mu u_{pri}^w - M(1 - t_w) \leq 0 \quad \forall w \in W \qquad (9)$$

The main deciding variables of the model are y^{p}, which determine whether each route is present in the optimal route combination or not. As mentioned before, the objective of this model is to maximize the rapid transit network coverage. In the objective function, t^{w} are binary variables that take value 1 if and only if for passengers of OD pair w desirability of rapid transit is higher than a private car. This desirability is defined by constraints (9). Besides, in the objective function, d_w refers to the travel demand.

Constraints (2) force the number of routes in the optimal route combination to be equal to or less than a predefined value. Constraint (3) is the budget constraint. In this constraint, c^p is route construction cost which is determined with regard to the costs of links and stations present in a route and given to the model as a parameter. Constraints (4) are conservation of flow constraints for a unit flow per each OD pair. Indeed, these constraints demonstrate continuity of routes between each OD pair. Constraints (5) show that a link can be accessed by passengers as a part of route p only if route p is in the optimal combination and contains this link. Constraints (6) state that a link has been used by an OD pair only if at least one of the routes which are used by this OD pair across the link. With the assistance of constraints (7), the nodes in which OD pair w transfer are identified and the corresponding variables tr take value 1. Furthermore, constraints (8) calculate the travel times in the transit network and store them in variables $\mathrm{u_{pub}}$.

If the values of vector Y are determined, constraints (2) and (3) can be left out. Consequently, because objective function and remaining constraints are decomposable based on each OD pair, model P1 can be decomposed into n^2 model like P2. In each of these models, the vector Y is predetermined, and index w has a constant value.

(P2)

$$Maximize \quad d_w.t_w \qquad (10)$$

$$\sum_{i\in N(k)} f_{ik}^w - \sum_{j\in N(k)} f_{kj}^w = \begin{cases} -1 & k = o(w) \\ 1 & k = d(w) \\ 0 & \text{otherwise} \end{cases} \qquad \forall k \in N \qquad (11)$$

$$x_{ij}^{w,p} \leq \delta_{ij}^{p} \ . \ y^{p} \quad \forall (i,j) \in \mathcal{A}, p \in P \qquad (12)$$

$$f_{ij}^w \leq \sum_{p} x_{ij}^{w,p} \quad \forall (i,j) \in \mathcal{A} \qquad (13)$$

$$\sum_{k\in N(i)} x_{ki}^{w,p} + tr_i^w \geq \sum_{k\in N(i)} x_{ik}^{w,p} \quad \forall i \in N \ , p \in P \qquad (14)$$

$$u_{pub}^w = \sum_{ij\in E} \left(t_{ij} \sum_{p} x_{ij}^{w,p} \right) + \text{TP} \sum_{i\in N} tr_i^w \qquad (15)$$

$$u_{pub}^w - \mu u_{pri}^w - M(1 - t_w) \leq 0 \qquad (16)$$

For more simplification, Constraint (16) is represented as below:

$$Mt_w \leq -u_{pub}^w + \mu u_{pri}^w + M \qquad (16')$$

Since the objective function is multiple of t_w, this variable is only present in constraint (16), and in this constraint, upper bound of t_w defined by $-u_{pub}^w$, instead of maximizing $d_w.t_w$ in objective function, u_{pub}^w can be minimized. Thus, Model P2 can be rewritten as:

(P3)

$$Minimize \quad \sum_{ij\in E} \left(t_{ij} \sum_{p} x_{ij}^{w,p} \right) + \text{TP} \sum_{i\in N} tr_i^w \qquad (17)$$

(11), (12), (13), and (14)

As it can be seen, if vector Y is determined, the RTND problem decomposes into n^2 problem in form of P2, each of which is a shortest path problem. Considering that shortest path problems, owing to their specific solution methods, are solved in a significantly lower time, the decomposition can be used for solving rapid transit network design problem in a lot less time. In this approach, feasible values of Y are searched in order to find the optimal solution. In each step of the searching process, after choosing the route combination, the objective function is calculated by solving the P3 problems and using constraints (16).

FEASIBLE ROUTES GENERATION PROCESS

Considering the characteristics of the SA method, which is going to be elaborated further, it is necessary to first identify the feasible routes. In the problem we are dealing with, a route is considered feasible if it meets the two conditions of having an acceptable length and not entailing a tour. The permissible range for the length of each route is defined by the two parameters of minimum and maximum acceptable lengths, which are determined by the planner with due attention to the network characteristics. The necessity for not having tours is also considered because of the special rapid transit network features.

SIMULATED ANNEALING

Simulated Annealing (SA) is a meta-heuristic algorithm based on the local search that has the capacity for avoiding local optimums. Furthermore, its ease of use and convergence nature has made it into one of the most popular techniques in the field. In general, simulated annealing is utilized for discrete problems and less commonly, for continuous optimization problems.

In each iteration of the mentioned algorithm, the objective function is compared on two solutions (namely, the current and candidate solutions). All of the improving solutions and only a part of the non-improving solutions are accepted (with the aim of avoiding local optimums and reaching a global optimum). The probability of non-improving solutions being accepted depends on the temperature parameter, which usually decreases with the proceeding of the algorithm.

In this study, the SA algorithm is employed as a sequence of homogeneous multi-chain Markov, in a manner that the temperature, T, is kept constant for a certain number of iterations (internal loops) and then reduced on a pre-determined schedule. The number of these reductions (external loops) is pre-determined as well. The various steps of the SA algorithm are displayed in Figure 1.

Neighborhood structure and initial solution

In the process employed in this study, only one type of move is defined and that is the exit of a route from a route combination and the entry of a candidate route into the combination. In this manner, the mere difference of each two neighbor solutions would be only one route. Considering the fact that in this method, the set of feasible routes is introduced before the commencement of the SA algorithm, the definition of the neighborhood structure on the basis of routes replacement is reasonable, and despite the inherent advantages of small neighborhoods in the SA performance, defining such a big neighborhood in the problem is inevitable. It must be mentioned, however, that this definition has become an advantage of the SA method in the candidate solution section due to its proposal of a suitable strategy. This neighborhood structure is, furthermore, influential on the method for choosing the initial solution.

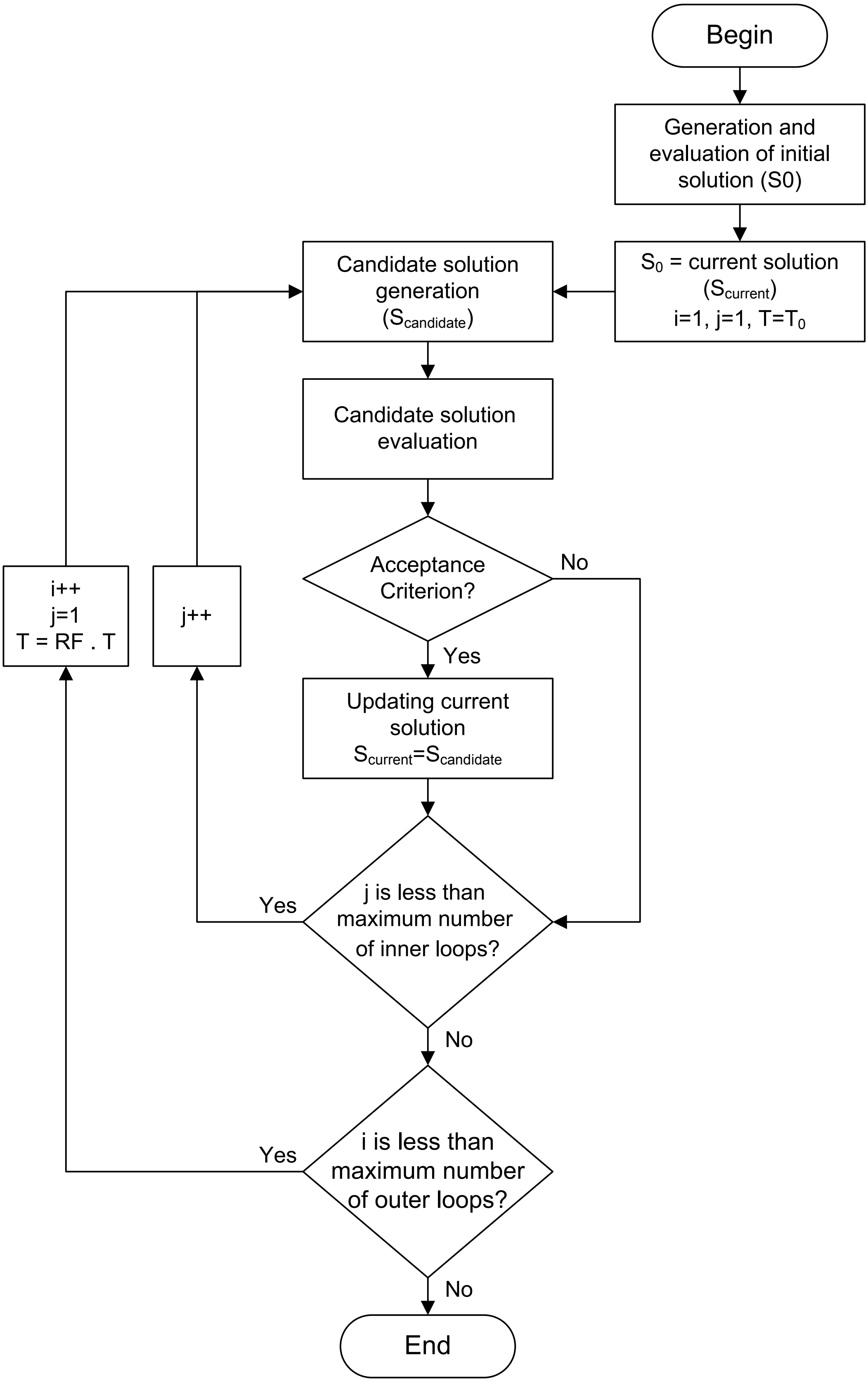

Figure 1 – SA algorithm steps

In the present SA algorithm, the initial solution is chosen in a completely random manner due to the rapid access to all the points of the feasible region. Faster access to different points of the feasible region lowers the influence of smart choice of the initial solution and in the solving method proposed here, the neighborhood structure is in a manner that in a problem with 10-route combinations, access to each point of the feasible region would be possible at most after 10 iterations. This complicates the cost-effectiveness justification of choosing a suitable initial solution.

Candidate solution generation

In order to choose a neighborhood as a candidate solution, a route from inside the current route combination and one from outside must be chosen, so that with replacing the old route with the new one, a new solution is achieved in the previous solution neighborhood. In this method, the route to be replaced is chosen randomly while the new route to replace would be selected in a smart approach.

Due to the high number of routes outside the solution, there exist a lot of choices for the entering route. Randomly choosing the entering route would allow the nomination of only one route in each iteration and thus each route must pass through numerous iterations. This would finally lead to a reduction in the convergence speed toward good solutions, since suitable routes must wait for being randomly chosen just like the others. In this study, the entering route selection process attempts to create the influence of the features of each route in the selection process and thus resolve the above mentioned shortcoming. This process and the other elements of the candidate solution generation process are displayed in Figure 2.

In this figure, it can be seen that the entering route may be chosen using criterion 1, 2 or just randomly. The probability of each of these options is determined by the planner and kept constant throughout the SA algorithm. If one of the two criteria is employed for the choice, the entering route would be proposed through assessment of a significant number of routes (In this study, 5 percent of feasible routes is assessed) and choosing the best. It is worth mentioning that criterion 1 considers network surface expansion and criterion 2 attempts to establish better coverage in the more important parts of the network from demand viewport.

In criterion 1, the new route must be capable of connecting new stations to the network which were not connected to it before. In order to assess this capability, the desirability index D1 is calculated for a large number of routes and the route with the highest desirability is chosen.

$$D1_k = \left[\sum_{i \in N_k} \sigma_i . \left(\sum_j d_{ij} + \sum_j d_{ji} \right) \right] . R_k$$

in which
N_k = the set of stations in the network crossed by route k;
σ_i = 1 if station i is connected to the network; = 0 otherwise; and
d_{ij} = demand from node i to node j.

$$R_k = \frac{\text{number of OD pairs on route k not directly connected before}}{\text{total number of OD pairs on route k}}$$

R_k is a parameter to show the extent of novelty of a proposed route to the network. For instance, if the new route is a subset of one of the routes present in the network, R_k would equal zero. This parameter is multiplied into the two criteria employed and includes the impact of the new route's novelty in them.

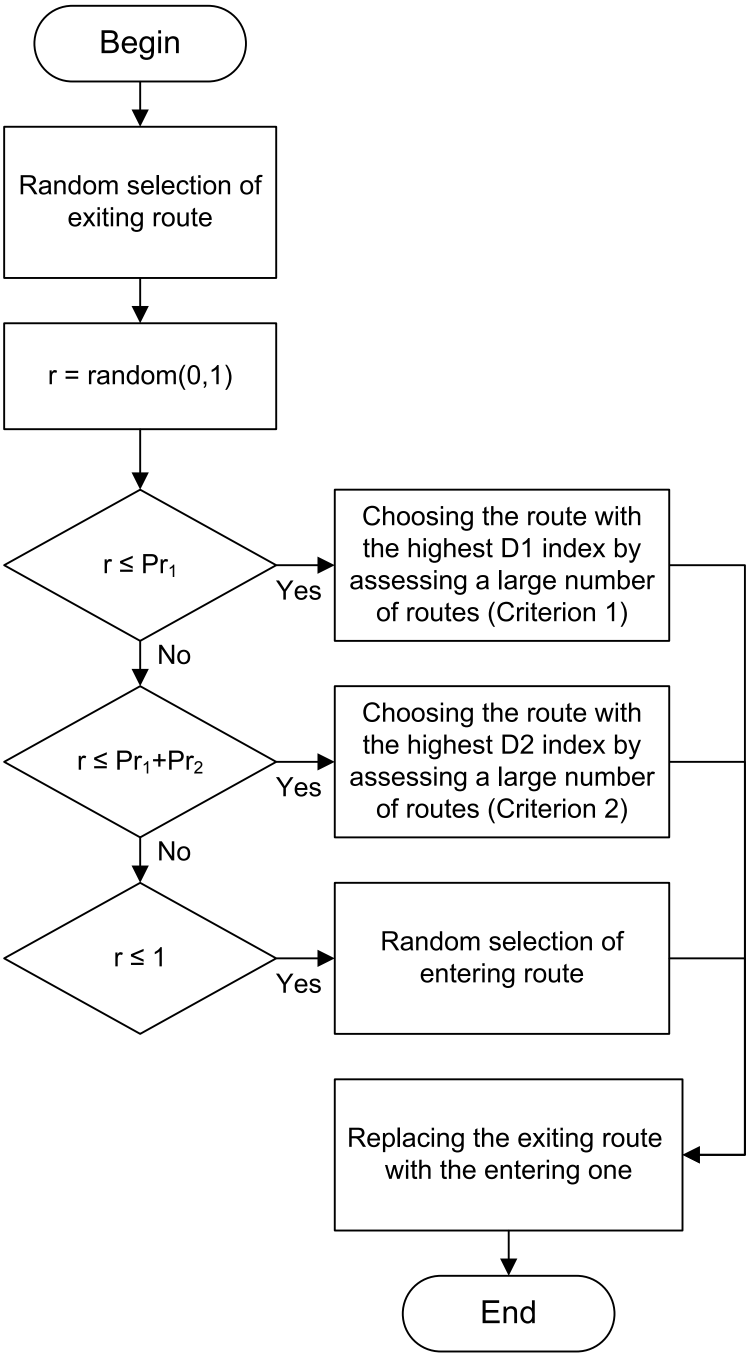

Figure 2 – Elements of candidate solution generation process

Criterion 2 identifies the routes that cover the more-demanded points. The desirability index D2 is defined as follows for this purpose:

$$D2_k = \left[\sum_{i \in N_k} \left(\sum_j d_{ij} + \sum_j d_{ji} \right) \right] . R_k$$

Index D2 is directly related to the number of more important stations it connects which were not connected before. Like criterion 1, if this criterion is employed in choosing the entering route, the route with the highest D2 is chosen and nominated from among a large number of routes.

Solution evaluation process

As mentioned before, if the routes present in a solution (Y) are given, the problem can be decomposed for any OD pair. Thus, the objective function can be calculated through solving the problems resulting from the decomposition, which are the shortest path problems for the OD pairs of ensuring network. In this study, the Floyd-Warshall algorithm is employed to solve the shortest path problems. In order to consider the transfers, a network representation as shown in figure 3 is used. In this representation, each node is replicated for L times and each of the routes in the route combination can only pass through its related node. The construction cost of virtual links between the virtual and base nodes is zero and the travel time from the base node to these nodes equals the transfer penalty. As a result, for each transfer, a time equal to the transfer penalty would be added to the passengers' travel time. Existence or absence of coverage for each of the OD pairs would also be determined after the calculation of the shortest path in the transit network and its comparison to the travel time using private cars (constraints 16). In each iteration of the SA algorithm, the routes present in the combination are given and the amount of the objective function can be calculated by the above method and employed in the acceptance criterion.

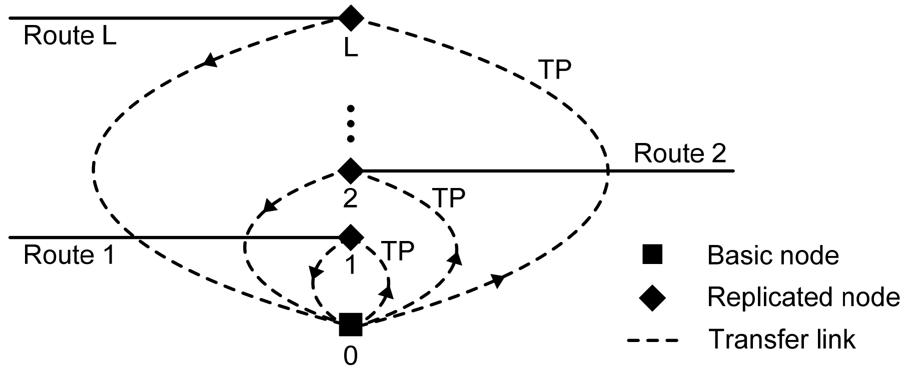

Figure 3 – Representation of nodes to consider transfer penalties

Acceptance criterion and cooling schedule

In the SA algorithm, the acceptance criterion for the candidate solutions is based on Metropolis Criterion. If f is the objective function for an optimization problem, after the generation of a candidate solution in each iteration, the extent the objective function of the candidate solution increases in comparison to the present solution, Δf, is calculated. The probability of acceptance for a candidate solution equals $\exp(\Delta f\ /\mathrm{T})$, if the objective function decreases, and equals one otherwise. With the continuation of this process, the system would progress towards equilibrium and after a large number of iterations, the probability distribution of the solutions would converge to the Boltzmann Distribution.

The cooling schedule in the simulated annealing includes initial temperature, a plan for lowering it and the termination criterion. The development of an effective cooling schedule plays a critical role in minimizing the time needed by the algorithm to reach an acceptable solution. In the proposed SA method, the initial temperature is determined in a manner that in the beginning, around 60 percent of the inferior solutions are accepted too. As depicted in figure 2, the cooling schedule includes multiplying the reduction factor (RF) to the temperature at the end of each outer loop. This factor may have a value between 0.8 and 0.99, which would remain constant throughout the algorithm. The termination criterion is also reaching the determined number of outer loops.

COMPUTATIONAL RESULTS

The SA algorithm has been coded and compiled in C# and then used for the RTND problem in an example network in order to be assessed in terms of efficiency. This example is a medium-sized network with 9 nodes and 15 links that was first proposed by Marin (2007) and whose optimal routes were identified through different combination of parameters. After Marin, Escudero & Munoz (2009) and Kermanshahi et al (2010) also solved RTND problem on this network. The calculations for this example were carried out on a system using a 3-GHz Pentium 4 processor with a 1-gigabyte internal memory (in order to compare with the previous studies). Figure 4 shows the transportation network of this example. The numbers written on the links are their construction cost and length respectively, while the same cost for nodes are written beside them. Moreover, demand and travel time for private cars, separately for each OD pair, have been determined by D and U^{priv} matrices respectively.

$$D=\begin{bmatrix} 0 & 9 & 26 & 19 & 13 & 12 & 13 & 8 & 11 \\ 11 & 0 & 14 & 26 & 7 & 18 & 3 & 6 & 12 \\ 30 & 19 & 0 & 30 & 24 & 8 & 15 & 12 & 5 \\ 21 & 9 & 11 & 0 & 22 & 16 & 25 & 21 & 23 \\ 14 & 14 & 8 & 9 & 0 & 20 & 16 & 14 & 12 \\ 26 & 1 & 22 & 24 & 13 & 0 & 16 & 14 & 12 \\ 8 & 6 & 9 & 23 & 6 & 13 & 0 & 11 & 11 \\ 9 & 2 & 14 & 20 & 18 & 16 & 11 & 0 & 4 \\ 8 & 7 & 11 & 22 & 27 & 17 & 8 & 12 & 0 \end{bmatrix}, U^{priv}=\begin{bmatrix} - & 1.6 & 0.8 & 2 & 2.6 & 2.5 & 3 & 2.5 & 0.8 \\ 2 & - & 0.9 & 1.2 & 1.5 & 2.5 & 2.7 & 2.4 & 1.8 \\ 1.5 & 1.4 & - & 1.3 & 0.9 & 2 & 1.6 & 2.3 & 0.9 \\ 1.9 & 2 & 1.9 & - & 1.8 & 2 & 1.9 & 1.2 & 2 \\ 3 & 1.5 & 2 & 2 & - & 1.5 & 1.1 & 1.8 & 1.7 \\ 2.1 & 2.7 & 2.2 & 1 & 1.5 & - & 0.9 & 0.9 & 2.9 \\ 2.8 & 2.3 & 1.5 & 1.8 & 0.9 & 0.8 & - & 1.3 & 2.1 \\ 2.8 & 2.2 & 2 & 1.1 & 1.5 & 0.8 & 1.9 & - & 0.3 \\ 1 & 1.5 & 1.1 & 2.7 & 1.9 & 1.8 & 2.4 & 3 & - \end{bmatrix}$$

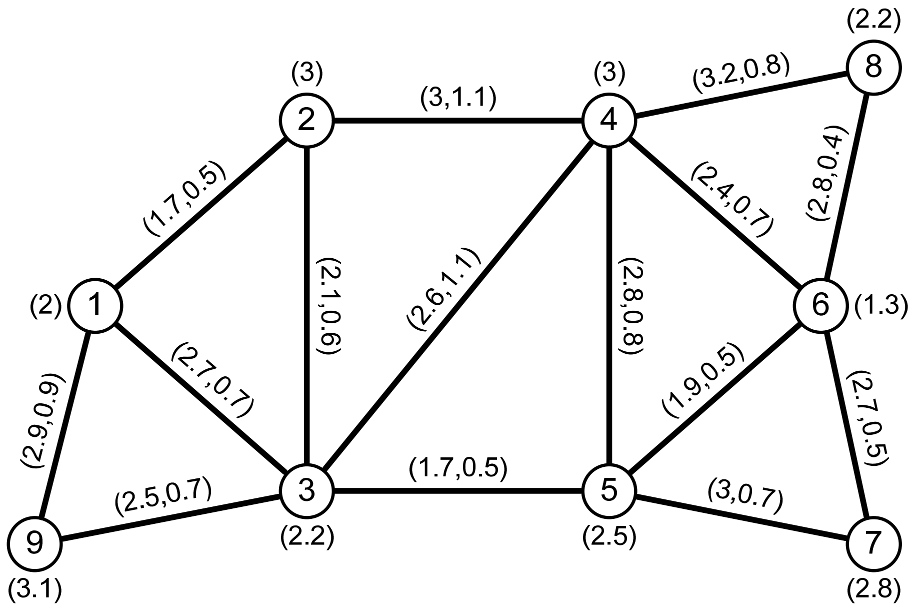

Figure 4 – Transportation network in example 1

Table 1 compares the results of this study with those of Marin (2007), Escudero & Munoz (2009) and Kermanshahi et al (2010). The employed SA algorithm has 30 inner and 30 outer loops. The initial temperature and the reduction factor were chosen at 200 and 0.9 respectively, while Pr_1 and Pr_2 were 0.4 and 0.5 respectively. The SA results presented in table 1 are average of 20 sample runs.

Table 1 – Results comparison for example 1

b	μ	z^*				CPU time			
		Ma	Sc	KE	SA	Ma	Sc	KE	SA
28	0.75	361	361	361	361	8 hour	21 s	0.3 s	0.3 s
28	1	466	466	466	464	8 hour	111 s	0.3 s	0.3 s
28	1.5	522	522	522	520	8 hour	290 s	0.3 s	0.3 s
48	0.75	672	672	672	663	233 min	3.6 s	1.3 s	0.6 s
48	1	912	912	912	904	898 min	30.3 s	1.3 s	0.7 s
48	1.5	1035	1035	1035	1026	165 min	3.5 s	1.3 s	0.6 s

As shown in Table 1, the SA algorithm provided acceptable results compared to the previous studies. This method, despite reasonable accuracy, takes less time. The reduction of time needed in this method was not significant in comparison to the study by Kermanshahi et al (2010), but, since the computational time increases exponentially with the network size, using other methods would not be possible for larger networks. This is while SA gives satisfactory results with any time budget.

CONCLUSION

In this study, a meta-heuristic method based on simulated annealing was used for the rapid transit network design, and a benchmark problem was also employed to compare this method with the methods already exist in the literature. The proposed SA algorithm is able to find a solution with satisfactory accuracy, in a shorter time compared to the previous methods. This is possible, in the first place, due to the suitability of the problem characteristics for the utilization of local search, and in the second, to the smart method used for the selection of the routes entering the route combination in the algorithm.

The computational study shows that the SA algorithm gives extremely accurate (error of less than 1 percent) approximate solution more than 250 times faster than Escudero & Munoz (2009), and 1.5 times faster than kermanshahi et al. (2010). As a result, the proposed algorithm can tackle the bigger problems more easily.

Future research can involve investigations of possible improvements of the SA algorithm in terms of generation of the candidate solutions, optimal combination of parameters, as well as cooling schedule. Comparing proposed SA algorithm to the other meta-heuristics can be considered too. Moreover, hybrid meta-heuristics can be suggested in future studies.

REFERENCES

Barra, A., L. Carvalho, N. Teypaz, V. D. Cung and R. Balassiano (2007). Solving the transit network design problem with constraint programming. In: Proceedings of the 11th World Conference in Transport Research, University of California, Berkeley, USA, June 24–28.

Borndorfer, R., M. Grotschel and M. E. Pfetsch (2007). A Column-Generation Approach to Line Planning in Public Transport. Transportation Science, 41, 123-132.

Ceder, A. (2003). Public transport timetabling and vehicle scheduling. Chapter 2 in Advanced Modeling for Transit Operations and Service Planning, Edited by W. Lam and M. Bell. Pergamon Imprint, Elsevier Science Ltd. Pub., 31-57.

Ceder, A., and N. H. M. Wilson (1986). Bus network design. Transportation Research Part B, 20, 331-344.

Escudero L. F. and S. Muñoz (2009). An approach for solving a modification of the extended rapid transit network design problem. TOP, 17, 320-334.

Fan, W. and R. B. Machemehl (2006). Using a simulated annealing algorithm to solve the transit route network design problem. Journal of Transportation Engineering, 132, 122-132.

Fan, W. and R. B. Machemehl (2008). A Tabu Search Based Heuristic Method for the Transit Route Network Design Problem. Computer Aided Scheduling of Public Transport, Book by Springer Series, Lecture Notes in Economics and Mathematical Systems.

Garzón-Astolfi, A., J. A. Mesa and F. A. Ortega (2007). A rapid transit network design model in regard to transfers. In: Advanced OR and AI Methods in Transportation, Publishing House of Poznan University of Technology, Poznan, 212-217.

Guan, J. F., H. Yang and S. C. Wirasinghe (2003). Simultaneous optimization of transit line configuration and passenger line assignment. Transportation Research Part B, 40, 885–902.

Kermanshahi, S., Y. Shafahi, M. Zangui and M. Mollanejad (2010). Rapid transit network design: a new model and solution algorithm. In: 89th Transportation Research Board Annual Meeting, Washington D.C., USA.

Laporte, G., J. A. Mesa and F. A. Ortega (2005). Maximizing trip coverage in the location of a single rapid transit alignment. Annals of Operations Research, 136, 49-63.

Lee, Y. J. and V. R. Vuchic (2005). Transit network design with variable demand. Journal of Transportation Engineering, 131 (1), 1-10.

Marín, Á. (2007). An extension to rapid transit network design problem. Top, 15, 231-241.

Marín, Á. and P. Jaramillo (2009). Urban rapid transit network design: accelerated benders decomposition. Annals of Operations Research, 169, 35-53.

Mauttone, A. and M. Urquhart (2009). A route set construction algorithm for the transit network design problem. Computers & Operations Research, 36, 2440 - 2449

Ngamchai, S. and D. Lovell (2003). Optimal time transfer in bus transit route network design using a genetic algorithm. Journal of Transportation Engineering, 129 (5), 510–521.

Shih, M. , H. S. Mahmassani and M. Baaj (1998). A planning and design model for transit route networks with coordinated operations. In Transportation Research Record: Journal of the Transportation Research Board, 1623, 16-23.

Tom, V. M. and S. Mohan (2003). Transit route network design using frequency coded genetic algorithm. Journal of Transportation Engineering, 129 (2), 186-195.

Wan, Q.K. and H. K. Lo (2003). A mixed integer formulation for multiple-route transit network design. Journal of Mathematical Modelling and Algorithms, 2 (4), 299-308.

Zhao, F. and I. Ubaka (2004). Transit network optimization – minimizing transfers and optimizing route directness. Journal of Public Transportation, 7 (1), 67-82.