A Procedure of Solving Transit Network Design Problem

Based on a Non-Dominated Sorting Algorithm of NSGA II

Thesis

Presented to the Faculty of the Graduate School

Of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Master of Science

by

Qinyuan Xiong

Conferral Date: May 2017

Abstract

This thesis focuses on the transit network design problem, where the focus is on understanding the trade-off between user costs and system owner cost. Given the focus on understanding this trade-off, a solution procedure based on the use of the Non-Dominated Sorting Genetic Algorithm (NGSA II) is developed. An illustrative example is presented.

Biographical Sketch

Qinyuan Xiong was born and raised in Nanjing, China. She is a M.S. candidate at Cornell University in Transportation System Engineering, with a minor in Operational Research and Information Engineering. Before entering graduate school, she earned her Bachelor' Degree in Transportation Engineering at Southeast University, with an exchange experience at National Cheng Kung University. She was awarded National Scholarship for her excellent academic performance during her time as an undergraduate.

Qinyuan has a passion do improving passenger transportation. Her undergraduate study focused on Transportation Planning, with a graduate thesis focused on Signal Timing Optimization Method for Pedestrians. During her stay at Cornell, she has focused on solving the transit network design problem. She had her first internship in Parsons Brinkerhoff, which is a consulting company in Civil Engineering. She is looking forward to become a research-oriented engineer in the field of Transportation Engineering.

Acknowledgements

I would first like to thank my thesis advisor Professor Linda K. Nozick of the School of Civil and Environmental Engineering at Cornell University. The door to Prof. Nozick's office was always open whenever I had a question for my research. She is not only equipped with top scientific literacy, but also a role model in ethics. I am indebted to her and her valuable comments on this thesis. I could not imagine my graduation or completion of this paper without her generous help.

I would also like to acknowledge Professor Shane G. Henderson of the school of Operational Research and Information Engineering at Cornell University as the second reader of this thesis. With his recommendation of coursework, I have improved my understanding and practical skills in statistical analysis and optimization. I would like to thank Mr. Monty Du, the major referee for my first internship. I succeed in gaining practical skills in Civil Engineering with his help.

In addition, I would like to thank for my friends at Cornell and old friends at Southeast University. Thank you for your technical and psychological support, during the difficult days.

Finally, I would like to thank my parents, for their encouragement. Thank you for keeping me going.

Table of Contents

List of Figures and Tables

Chapter 1 Introduction

1.1 Transit Network Design Problem Introduction

A well-designed transit network is known to benefit society in many ways. For example, transit systems reduce congestion by reducing private transportation as well as fuel. [1] As an illustration, in 2011, U.S. public transit use saved 450 million gallons of fuel in 498 urban areas. [2] Also, transit networks benefit the more vulnerable members in the society, including the elderly, disabled, children and the economically disadvantaged. [3]

The transit network design problem (TNDP) has been explored for decades, including designing routes, determining route frequencies and signal settings, etc. [4] Because of the difficulty on this problem, in practice it is commonly addressed using the following five steps: 1) route design; 2) frequency setting; 3) the timetabling; 4) vehicle scheduling and 5) the crew scheduling. [5] This thesis focuses on the use of optimization to integrate the first two steps.

1.2 Outline

This thesis is organized as follows. Chapter 2 reviews the literature on the TNDP using mathematical modelling and heuristic methods for solving TNDP. In Chapter 3, a mathematical model for TNDP is described. Then, in Chapter 4, we introduce an Evolutionary Multi-Objective Algorithm (EMOA) using the Non-Dominated Sorting Genetic Algorithm II (NSGA II) and its application to the TNDP. Finally, in Chapter 5, the solution method is applied to a 20-node transit network design sample problem instance. Conclusions and opportunities for further study are described in Chapter 6.

Chapter 2 Literature Review

2.1 Literature Review on TNDP

The TNDP has been studied for decades. Guihaire and Hao presents a literature review and includes literature focused on transit network route design, transit network frequency setting and transit network timetabling as well as the solution method employed. They explicitly compare the range of objectives used as well as the variety of constraints employed. Given the focus on three inter-related sub-problems (route design, frequency setting and timetabling) they propose a terminology to describe the relationship between the sub-problems that highlights how particular papers fit in the context of global transit planning. [5]

Konstantinos Kepaptsoglou, M.ASCE; and Matthew Karlaftis, categorized the literature by the different objectives used, the underlying decisions to be made, the input parameters (e.g. network structure, demand, etc.) and the solution method employed. [4] Reza Zanjirani Farahani, Elnaz Miandoabchi, W.Y.Szeto, Hannaneh Rashidi, also discusses the TNDP in the more general context of urban transportation network design problems, which includes the TNDP with a focus on the tactical and operational problem. For the TNDP, they summarize papers by the constraints represented, the objectives pursued and the solution method developed. [6]

Solution procedures developed and applied to the TNDP can be divided into two categories: exact search methods and heuristic methods. In the domain of exact search method, Yan, Shangyao, and Hao-Lei Chen (2002) develop an intercity bus carrier time-scheduling model and solution method. The formulation is constructed based on a mixed integer multiple commodity network flow problem where the objective focuses on minimizing costs and the passenger fares are represented as negative costs. An algorithm based on Lagrangian relaxation, is developed to efficiently solve the problem. A case study focused on a major Taiwan inter-city bus operation is used to illustrate the model performance. [7] Ralf Bordorfer, Martin Grotschel, Marc E. Fetch develops a mathematical model for finding routes and corresponding frequencies for at least two objectives. The resultant multi-commodity network flow model is applied for route optimization, with the main feature that passenger paths can be freely routed and routes are dynamically generated. The model is applied using the city of Postdon, Germany. [8] For realistic scale problems, computational burden is often an issue; hence the literature describes the development and application of many heuristic methods to pursue these solutions.

Heuristic methods that have been employed include genetic algorithms, tabu search, simulated annealing, and ant colony. Our interest is in understanding the trade-space for design considering user and operator costs; hence we focus on using evolutionary algorithms which are particularly effective at estimating the trade-off frontier in the context of a single optimization procedure. Hence the remainder of the discussion of solution procedures focuses on literature that makes use of evolutionary-based methods.

Population based heuristic approaches to the solution of the TNDP are usually divided into three steps. First, candidate routes based on network characteristics and constraints are generated. Second, a collection of initial solutions based on the candidate route set are created. Third, Cross over and mutation are then used over a sequence of iterations (generations) to identify either an optimal solution or a trade-off frontier.[9]

Tom, V,M. and S. Mohan construct a bus transit network with the goal of optimizing total system costs considering a collection of operational constraints. System cost is expressed as a function of bus operating costs and the passengers' cost. The solution procedure has two steps. In the first step, a feasible candidate route set is generated. In the second step, a genetic algorithm is used to generate solutions. The major innovation of this paper lies in adding frequency to the coding scheme besides route selection. [10] More recently, Fang Zhao and Ike Ubaka focuses on the TNDP with the objective of minimizing transfers and minimizing route circuity while maximizing service coverage. The research is composed of three parts: 1) the representation of transit network solution space; 2) representation of transit route and network constraints; and 3) the solution search process. [11] Borja Beltran, Stefano Carrese Ernesto Cipriani, Marco Petrelli proposed a method to optimize a transit network of low emission green vehicles with limited route number. First, a heuristic method is used to generate a large candidate set of feasible routes. Second, a genetic algorithm is applied to find the optimal sub-set of routes with associate frequencies. [12] Ernesto, Cipriani, Stefano Gori, Marco Petrelli developed a methodology to optimize a multi-modal public transit system in an urban area. This network has many-to-many demand transit system demands. A heuristic method is used to generate a set of feasible routes. Then, a genetic algorithm is used to generate an optimal route set. After the optimization, the associated frequencies are determined. The optimization is tested on the city of Rome, with 50% reduction on the number of lines and 20% reduction on the operating cost. [13] Saeed Asadi Bagloee, Avishai proposed a study on the optimization of real world-scale system with multiclass transit vehicles. The stops in the network are selected using gravity model. A heuristic method is used to generate shortest paths among the stops. A metaheuristic method is implemented under budgetary constraints to construct the final transit network. This TNDP method is tested on the data of city of Winniperg and produced a reduction of 14% in travel time compared to the existing network. [9]

Antonio Mauttone, Maria E. Urquhart implemented a heuristic method called GRASP in transit network optimization. This is a non-dominated sorting method, balancing the two objectives of user's cost and operators' cost. The optimization result reveals that the non-dominated result of GRASP can generate more non-dominated solutions than the weighted sum method in the Mandl and antoerh real test case under same computational effort. Ties Brands, Luc. J. J Wismans, and Eric C. van Berkum apply algorithm ε-NSGA II in optimizing infrastructure planning problem, the objectives include total travel time, public transport system of operating depicts, and the climate change in the transit system. This algorithm allows different objectives to be non-dominated sorted. [14]

2.2 Thesis Contribution

This paper contributes to transit network design in two areas. First, this paper presents a multi-objective mathematical model of the TNDP under the dual objectives of minimizing user's and operator's costs. The formulation focuses on generating the optimized route pattern with associated frequency, under constraints focused on flux, network capacity, route length and frequency limits. Second, a non-dominated heuristic method of NSGA II is applied in this TNDP solving procedure. This non-dominated sorting method would enable the trade-off between the objectives and would generate a non-dominated front of solutions rather than a ‘best’ solution.

Chapter 3 Mathematical Modelling of TNDP

The transit system is often represented as an undirected graph, with nodes representing the transfer stations and links indicating the connection between the nodes. Transit systems in urban areas often have multiple routes and in large areas, many routes. Each route runs from a fixed origin to a fixed destination through a series of nodes and links with a fixed frequency based on the time of day and day of week.

From network route information and the trip demand matrix, we can identify how much of the demand can be satisfied. For each trip that can be accommodated, there is both an 'in-vehicle-s impedance' and 'out-of-vehicle impedance. In-vehicle impedance may include the link time cost of the traveler's path, the vehicle stop time at each station, etc. Out-of-vehicle impedance may include traveler wait time at the transfer station and walking time to the stops. Here, in-vehicle impedance is calculated by the summing up the link time cost on each path, while out-of-vehicle time is calculated by summing the waiting time at each transfer station. For the purposes of this research, the transfer time is assumed to be equal to half the headway of the route.

3.1 Problem Formulation

Set Explanation:

- φ- Set of origin and destination nodes
- $\omega(k)$- Set of all routes that compose path k
- $\Omega(r)$- Set of all nodes on route r
- S_{od}- Set of all possible paths for origin O and destination D

Variable explanation:

- $\hat{r}$- The optimized routes under objectives of Z_1 and Z_2
- $\hat{f}$- The optimized frequency settings for the corresponding route pattern under objective values
- r- The network solution route pattern

- $x_{ij,r}^{od,k}$- The flux volume on the link ij, on route r, in the k^{th} path from origin O to destination D,
- E_{od} - The unsatisfied demand for the origin O and destination D
- f_r- The frequency of route r
- $x_r^{od,k}$- The flux on route r, in the path k from origin O to destination D
- $t_{in-vehicle}^{od,k}$- The travel time in the path k from origin O to destination D
- $t_{wait}^{od,k}$- The wait time in the path k from origin O to destination D
- L_r- The length of route r
- T_r - The time cost of the route r

Parameter:

- C- The capacity of each vehicle
- d_{od} - The demand for origin O and destination D
- $t_{penalty}$-The time penalty for unsatisfied demand by person
- C_{km} - The parameter of the operating cost based on kilometer
- C_{hr} - The parameter of the travelling personnel's cost based on hour

Functions:

- Z_1- The function of user's cost
- Z_2- The function of operator's cost
- I_c- The function of a multi-commodity model to assign flux on the network based on link impendences

3.2 Transit Network Mathematical Modelling

The instance of the TNDP, the focus is on an optimization problem. As mentioned previously, the goals are to identify the design trade space between the minimization of the users' and the operator costs. Thus, the TNDP can be formulated as follows:

$$(\hat{r}, \hat{f}) = argmin\{Z_1, Z_2\} \tag{1}$$

Subject to:

$$\{E_{od}, x_{ij,r}^{od,k}, f_r\} = I_c(r, d_{od}, C),\ \forall(o,d) \in \varphi \tag{2}$$

$$\sum_k x_r^{od,k} + E_{od} = d_{od}, \quad \forall(o,d) \in \varphi, k \in S_{od} \tag{3}$$

$$\sum_k x_{ij,r}^{od,k} \leq C * f_r,\ \forall(o,d) \in \varphi, k \in S_{od} \tag{4}$$

$$L_{min} \leq L_i \leq L_{max}, \mathrm{i} \in r \tag{5}$$

$$f_{min} \leq f_i \leq f_{max}, \mathrm{i} \in r \tag{6}$$

The dual objectives are given in equation (1). It also indicates that the model generates routes and associated frequencies. Equation (2) indicates that a multi-commodity formulated to generate the flux across the network, the unsatisfied demand between each OD pair and the frequency for each route. Equation (3) says that the flows across the route-based paths for each OD pair and the unsatisfied demand for that OD pair must equal the demand for that OD pair. Equation (4) conserves the flux on the network to be within the capacity of the route. Equation (5) limits the length of the route to within the minimum length and maximum length constraint. Equation (6) bounds the frequency of each route. It is important to note that this formulation assumes a route set for the model to pick from.

3.3 Objective Function Formulation

The objectives to minimize user's and operator's costs are given in Equations (7) and (8).

$$\min Z_1 = \sum_{(o,d)\in\varphi} \sum_{k\in P_{od}} x_r^{od,k} * \left(t_{in-vehicle}^{od,k} + t_{wait}^{od,k}\right) + t_{penalty} * \sum_{(o,d)\in\varphi} E_{od} \tag{7}$$

$$\min Z_2 = C_{km} * \sum_{r\in R} L_r * f_r + C_{hr} * \sum_{r\in R} T_r * f_r \tag{8}$$

The user's cost is formulated in (7) with the unit of time [12]. Notice that the user's cost is the sum of the transit user's cost and a penalty per passenger trip not accommodated. Without considering unsatisfied demand, the transit network may include a few routes with low frequencies. [15] The transit user's cost includes two parts: the traveler's in-vehicle travel cost and the user's wait time cost. In defining the penalty of the transit user's cost, since the transit system is used to benefit the transit user and carry demand, the penalty of the transit user should be much larger than any traveler's cost. The user' cost have the unit of hour.

The operator's cost is calculated in Equation (8). [12] Notice that the route cost is a function of the length of the route and the frequency of the route as well as the duration of the route and, again the frequency of the route. The first term stem from vehicle operating costs, which depend on the total bus distance travelled. This term includes the fuel, tries, maintenance, repair cost, etc. [16] The second term stems from travelling personnel's costs, stems from the total travel time of vehicle service. This cost refer to the operating people' salary. The parameters of C_{km} and C_{hr} would guarantee that the operating cost and travelling personnel's cost take a proper percentage in the operator's cost separately. These parameters would also enable the operator's cost to be estimated in the unit of dollar, according to the transit method of bus, metro, BRT, etc.

Chapter 4 TNDP Solving Procedure based on NSGA II

The procedure developed to address this problem is given in the flow chart below. Notice that it begins with a route set generation process to provide the algorithm with qualified routes. Moreover, there is an auxiliary module for network solution evaluation, where a collection of routes produces the frequencies as well as the unsatisfied demand and finally the computation of user's and operator's costs. Once the routes are identified and the network solution can be evaluated, the NSGA II optimization would start with population initialization, after which genetic operator and non-dominated sorting used over a sequence of generations to produce the trade-off frontier.

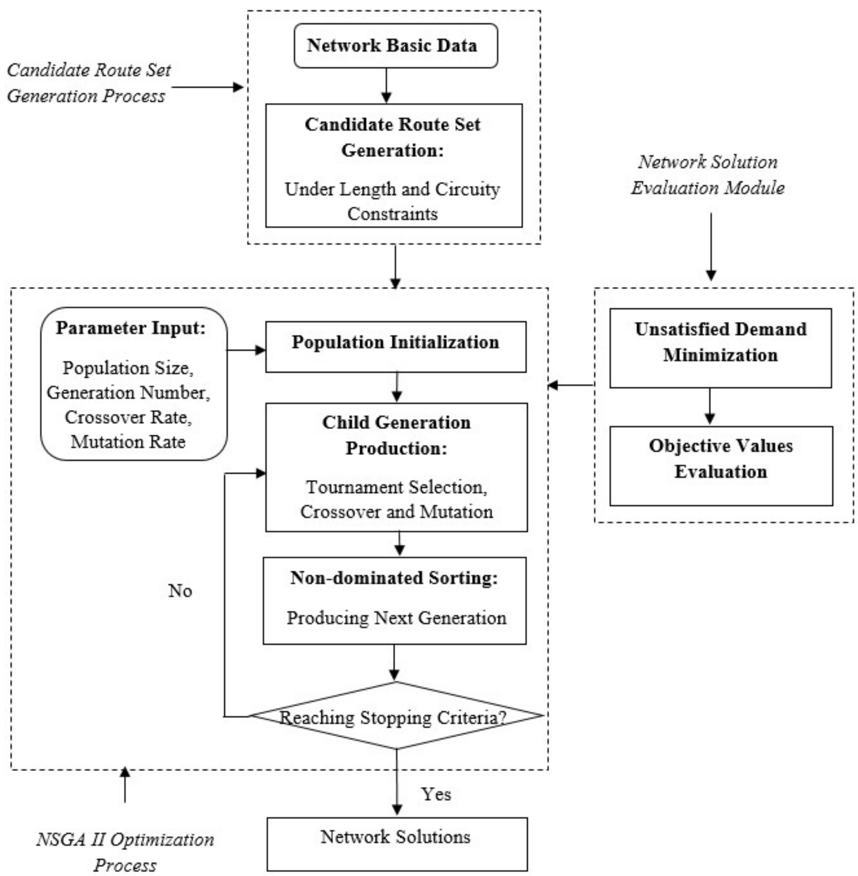


Figure 4.1: Flow Chart for NSGA II based TNDP Procedure

4.1 Candidate Route Set Generation Process

A candidate route set is required for the NSGA II optimization process. This process need the input of network basic data, route length constraints, route circuity constraint and a parameter k for controlling the candidate route number. The network input data includes demand matrix, potential transfer stations and potential link impedances. Also, route length constraints are set to avoid routes to be too long or too short. A short route is not desirable because it is likely to run through few nodes and is therefore unlikely to attract sufficient passengers. Very long routes are to be avoided since the path covers too many nodes and therefore will be desirable by too many passengers. Further, maximum circuity constraint would eliminate very circuitous routes, which would avoid corresponding operational difficulties. Finally, at most k shortest paths would be selected for each qualified OD pair, since there may be enormous shortest paths per sorted OD pair. The specific steps to generate this route set is as follows.

Step 1: Generate the shortest path and calculate the path length for each OD pair;

Step 2: Identify the OD pairs meeting two criteria: 1) shortest path length is less than the maximum length and 2) the shortest path multiplied by the maximum circuity is more than the minimum length.

Step 3: From the identified OD pairs, find all possible routes and select the k-shortest paths, where k is parameter that is user supplied. Sort out the qualified paths in k-shortest paths with maximum and minimum length constraint, as well as the circuity constraint. Put qualified routes into the candidate route set.

In the first step, the generation of the shortest path gives the shortest length between each OD pair; In the second step, the possible OD pairs with the qualified routes identified. Finally, at most k qualified paths are selected for each qualified OD pair, in order that the size of the candidate route set can be limited.

4.2 Network Solution Evaluation

Once the routes to be included in a solution have been identified, an optimization model is used to determine the route frequencies as well as the demand not satisfied. The next subsection gives that model. Once this has been done, the user's and operator's costs can be computed.

The assignment model of the transit network is commonly formulated as a multi-commodity network flow problem. [17] This paper essentially takes the same approach but both the route frequencies as well as the flows by OD pair and link are decision variables. Once the solution is obtained to this model the user's and operator's costs can be computed. This model is a possible formulation of the network flux assignment model as mentioned in equation (2).

The optimization model is formally defined as follows:

$$\min \sum_{(o,d)\in\varphi} E_{od} + \frac{1}{2} * C * \sum_r f_r \quad (9)$$

Subject to

$$x_{it,r}^{od,k} - x_{tj,r'}^{od,k} = 0 \quad r \in \omega(k), t|(i,t) \in A_r, t \in \Omega(r)\ \forall(o,d) \in \varphi, k \in S_{od} \quad (10)$$

$$\sum_k x_{ot,r}^{od,k} + E_{od} = d_{od} \quad \forall(o,d) \in \varphi, k \in S_{od} \quad (11)$$

$$\sum_k x_{ij,r}^{od,k} - C * f_r \leq 0 \quad \forall(\mathrm{o},d) \in \varphi, k \in S_{od} \quad (12)$$

$$x_{ij,r}^{od,k} \geq 0 \quad \forall(o,d) \in \varphi, k \in S_{od}, r \in \omega(k), (i,j) \in A_r \quad (13)$$

$$E_{od} \geq 0 \quad \forall(o,d) \in \varphi \quad (14)$$

$$f_r \geq f_{min} \quad r \in \omega \quad (15)$$

$$f_r \leq f_{max} \quad r \in \omega \quad (16)$$

The minimization objective formula (9) consists of two parts. The first part is the penalty of the unsatisfied demand. The second term associated a cost with the route frequencies. This second term is done to avoid frequencies that generate low vehicle occupancies.

For the constraints, (10) represents that for each path for any OD pair, the flux conserves along the path. Constraint (11) represents that for a certain OD, the sum of satisfied and unsatisfied flow should be equal to the total demand. Constraint (12) limits the link flux to be under the route capacity. Constraint (13) and (14) limits the flow and unsatisfied demand in network to be equal to or more than zero. Constraint (15) and (16) impose restrictions on the frequencies selected by route.

Based on the solution obtained above, Equations (7) and (8) are used to evaluate the two objective functions.

4.3 NSGA II Description

NSGA II (Non-dominated Sorting Genetic Algorithm II) is an Evolutionary Multi-Objective Algorithm, developed based on NSGA. [18] It is a population-based heuristic algorithm, and would perform non-dominated sorting to make the trade-off between multi objectives. These characteristics enable NSGA II to produce a set of non-dominated final solutions with wide diversity among its population, rather than a "best" solution.

The NSGA II algorithm is constructed in the following sequence. First, a parent population of solutions is created. Then each child is created selecting the two parents and performing genetic operations. Finally, non-dominated sorting is used on the combined set of a parent and a child population to produce the next parent generation. This population production process continues until it reaches the end the of iterations. [18] Figure 4.2 below illustrates the algorithm generation production procedure for NSGA II. According to the features of NSGA II, the algorithm is divided to modules of population initialization, solution fitness evaluation, tournament and genetic operations, and non-dominated sorting and elitism.

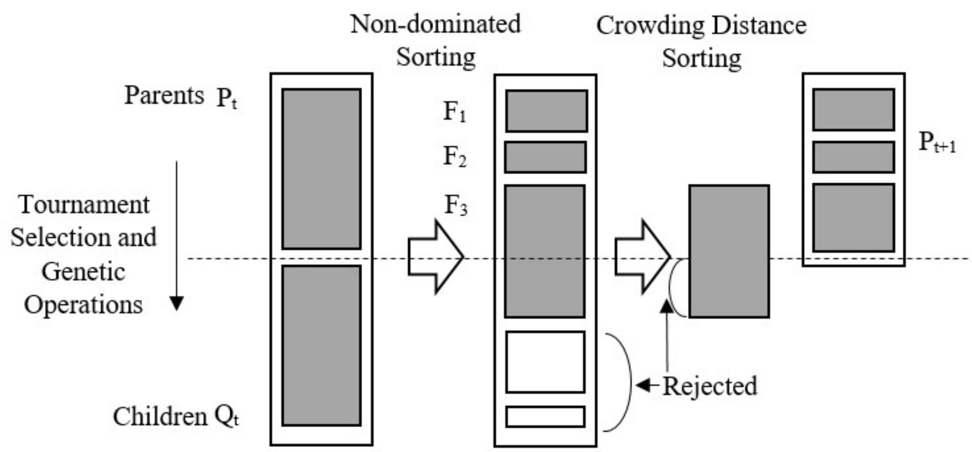


Figure 4.2: NSGA II Algorithm Generation Production Procedure [18]

4.3.1 Population Initialization

NSGA II is a form of genetic algorithm with a solution set called population. Everyone in the population contains the gene of decision variables, stored in the matrix, in the format of decimal or binary. During the population initialization process, solutions are generated randomly within the feasible range. After generating the decision variables of the solutions, the objective values are calculated and stored with each solution.

4.3.2 Solution Fitness Evaluation

In NSGA II, the fitness of everyone in the population is estimated by two attributes: non-domination rank (i_{rank}), and crowding distance ($i_{distance}$). For domination rank calculation, solution A is defined to dominate solution B if solution A has at least one objective value better than solution B with other objective values no worse than solution B. The solutions with no domination relationship are defined to be on the same front. In NSGA II, the solutions are separated by rank number (i_{rank}). We first find a non-dominated front and give it rank 0 and take all the solutions from the set. Then we find the front that are only dominated by these sets as rank 1. Then the rank of 2 and 3.

For crowding distance, its calculation uses the idea of Euclidian distance between individuals. This distance is calculated based on their m objectives in the m dimensional hyper space. The individuals in the same front (with the same rank number) are compared together to calculate the crowding distance. We would first assign the value of infinite to the points on the boundary. Then solutions in the middle are calculated for their crowding distances. The calculation procedure for the distance calculation is demonstrated as follows:

1) For j^{th} solution in the same front i, initialize the distance to be F_i (d_j)=0

2) For each objective function m=1,2,…,M, sort the solution set within the same front according to an ascending order;

3) For each objective function m=1,2,…,M, assign infinite to the boundary solutions, $I(d_1)=I(d_l)=\infty$. For all other solutions i=2 to i=l-1, assign distance based on the following formula. This distance is the sum of all the objective distance.

$$\mathrm{I}(\mathrm{d}_k) = \mathrm{I}(\mathrm{d}_k) + \frac{I_m^{k+1} - I_m^{k-1}}{f_m^{max} - f_m^{min}} \tag{17}$$

where $f_{I[i]}^{m}$ represents the fitness value of objective function m of the i^{th} solution in non-dominated set I. f_{min}^{m} and f_{max}^{m} are the minimum and maximum values of objective function m. For each front F_i, n is the number of individuals.

We define a solution to be better than another solution once it satisfied one of the following two criteria: 1) the solution is on a lower ranked front ($i_{rank\ 1}<i_{rank\ 2}$); 2) the solution is on the same front ($i_{rank\ 1}=i_{rank\ 2}$) but have a larger crowding distance. The different of rank means that the dominated solution is better than other solutions in all objective values. Also, a larger crowding distance represents a better diversity within the same front. Thus, a solution with better fitness has better objective values or better diversity.

4.3.3 Tournament Selection and Genetic Operations

Tournament selection is an optional operation performed before the genetic operations. Its existence allows parents with better fitness to be selected. First, grant all parents with an equal probability of being selected. Then, randomly pick two parents S_i and S_k and select the parent with the higher fitness to be the first parent. The process is repeated to generate the second parent.

Crossover and mutation are two genetic operations for producing child generation. The crossover can be single-digit or multi-digits. For single digit crossover, gene is switched after the selected digit; For multi-digit crossover, we switch the gene in between the selected digits. choosing the following the digits and switch the gene in between the selected digits. For mutation, we randomly choose a digit and turn this into another number. The feasibility of the changed gene should be tested. The genetic operations should be repeated until the child solution is feasible.

4.3.4 Non-dominated Sorting and Elitism

For non-dominated sorting, this mechanism would select a subset of solutions with better fitness from the original solution pool. This would enable the next generation to have better objective values and better diversity. For elitism, this action allows the best parents to be saved in the population, by letting the parent generation compete with the child generation. NSGA II includes the process of elitism by combining the parent generation and the child generation, and doing sorting to create a new population of the next generation.

4.3 Application of the TNDP Procedure

During the application of the TNDP procedure, we need to define the parameters of the three modules of candidate route set generation, network solution evaluation and NSGA II optimization algorithm. Also, the programming method for realizing this process and the coding scheme should also be determined. Finally, the feasibility of network solution is defined, and method of eliminating the infeasible solutions is introduced.

First, the parameters for generation the candidate route set need to be determined. Users need to define the maximum length, the minimum length, the maximum circuity and the parameter k of shortest paths for each feasible OD pair. These parameters should be decided according to the length of network links and the expected size of candidate route set. After the decision of the parameters, the candidate route set can be generated to prepare the optimization process with choices of routes.

Next, parameter for the TNDP mathematical model need to be determined for the objectives and constraints. For user's cost, the parameter t_{penaty} of the time penalty for each unsatisfied user should be set to be much larger than the sum of travel time and wait time on any feasible path. For operator's cost, the parameter of system operating cost of C_{km} and the parameter of travelling personnel's cost of C_{hr} need to be defined in order that the operating cost and travelling personnel's cost cover a proper percentage within the operator's cost. These two parameters should also transfer the operator' cost into the unit of dollar according to the transportation method of bus, rail, metro, etc. When it comes to the constraints, the capacity C of each vehicle, the maximum and minimum limits of frequency needs to be defined to constrain the route capacity. From the mathematical model of with parameters, the flux on each path and frequency for each route can be generated under route pattern and transit demand. Objective values of user' cost and operator's cost can then be calculated. As for the genetic operations, single digit crossover and mutation is used.

Then, it comes to the parameter determination of the NSGA II. The parameter of population size, generation number, crossover rate and mutation rate needs to be defined for the parameters. The population number should be large enough to enable diversity of combination of different routes. The generation number should be large enough to perform enough iterations in order that a steady pattern is generated in the final generation for the objectives. For crossover and mutation rate, a sum of 1 should be given in order that no child is a direct duplication from parent generation. This result from the non-dominated sorting, which asks the parent and child generation to be sorted together the next generation.

After defining the parameter, the programming language and the coding scheme are should be determined. For programming, there are some commonly choices of programming language of Matlab, Python and Visual C++ for this TNDP procedure. Also, the programming files is usually split into the modules of candidate route generation, network solution evaluation and NSGA II algorithm, with a post-optimization module doing the final result analysis and graph drawing. For the coding scheme, the solutions are composed of route numbers, with these numbers coded in the binary array. The route number should be a decimal number within the range between 1 and the number of candidate routes. Then, it shall be translated into a binary string, whose length l is the binary array length of the largest number in the candidate route set. We use the candidate route set matrix to translate the solution gene into the network routes. Moreover, there may be varied route numbers in population, with the same gene length. Thus, a continuous digit of zeros would reveal an empty route number. For any gene array, the non-empty digits representing route number would always be on the front, with empty digits representing no route number always recorded after those digits.

Here, a feasible solution is defined by two criteria: 1) no route number appear twice in the same solution; 2) no route number is larger than the number of candidate route set. These criteria guarantee that the route number is the same as the gene appears, with all route numbers can be translated into network routes. Since, the genetic operations of crossover and mutation would change the gene, in feasible solutions may be produced. If a solution with a variable of route more than the candidate route set number is generated, the crossover or mutation point is chosen again until a feasible solution is produced. If a child solution with overlapped route number existed in the gene is generated, this solution would be completed deleted, with genetic operator starting from the decision of crossover and mutation, choosing new parents and new crossover or mutation digit.

With the input parameter, well-programmed files, the TDNP solving procedure can realized. A set of population would be produced and iterations would be performed until reaching final generation number. Only the solutions on the non-dominated front would be chosen to be considered for the network route choice.

Chapter 5 Transit Network Numerical Example Results

5.1 Network Basic Information Description

The illustrative transit network is assumed to be a bus-based transit system, which consists of 20 nodes, 43 links. The nodes represent potential transit transfer stations, while the links represent the potential connection lines between transfer stations. The vehicle speed is assumed to be 15km/hr.

This network structure is illustrated below with nodes and link costs. The transfer nodes are divided into two groups a central group and non-central group. The central group consists of node G, J, K, L and O. The non-central group contains all other nodes. All trips either begin at a non-central node to a central node or between a pair of central nodes.

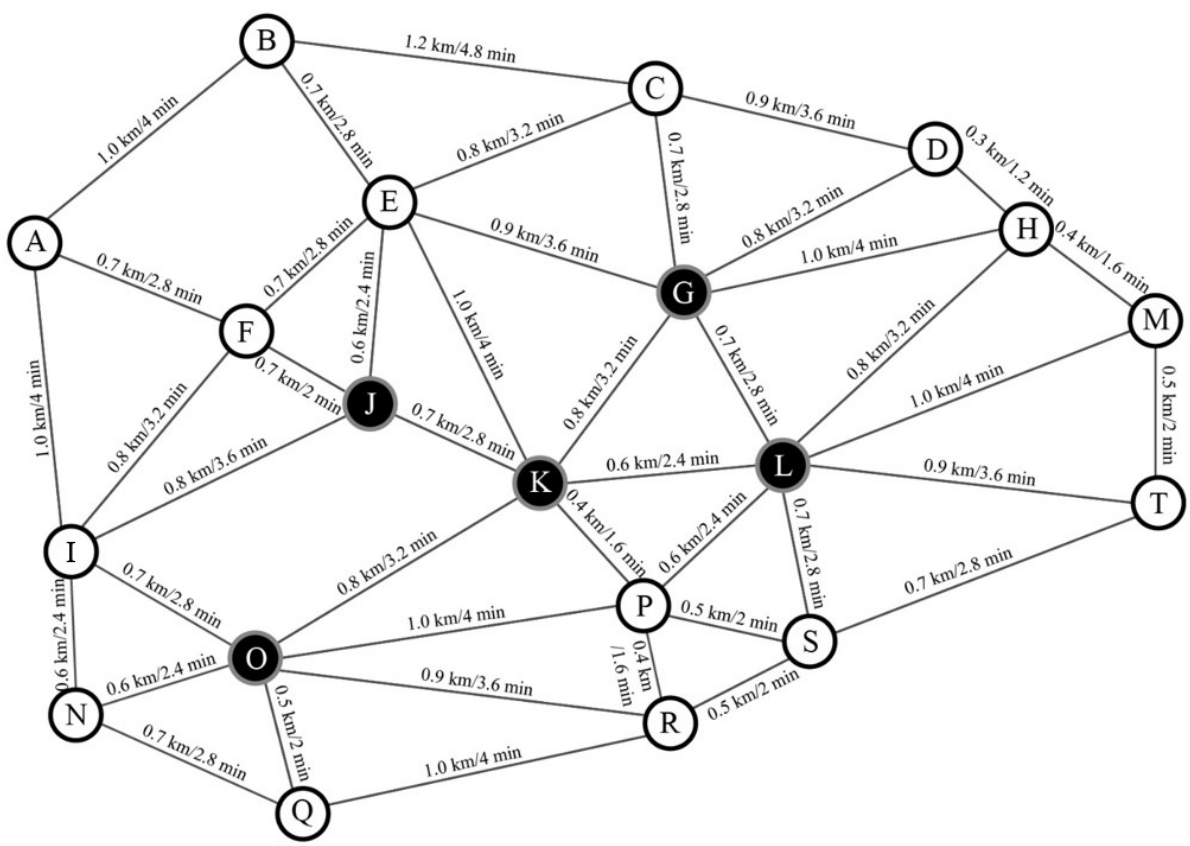


Figure 5.1: The Road Topology and Link Impedances of the Experimental Network

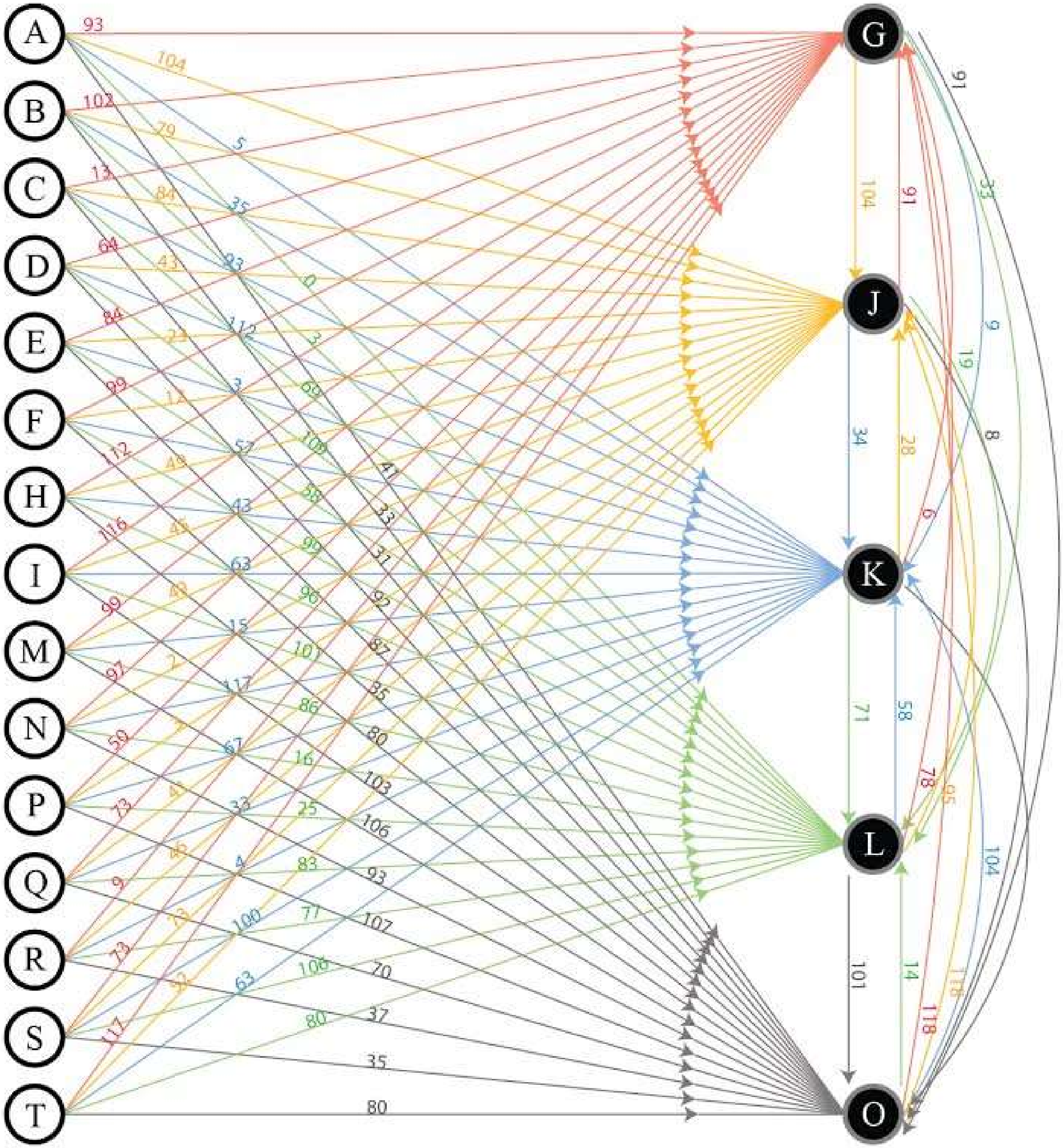


Figure 5.2 Demand for the Transit Network per Hour

5.2 TNDP Solving Procedure Application

The implementation of the TNDP solution procedure on this specific transit network, requires the selection of the parameters, the definition of the coding scheme and the evaluation of solution feasibility, which are given below.

The parameters for the TNDP procedure are defined. First, the parameters for candidate route set generation process are defined, according to the illustrative network information. We define that candidate routes must be between 3 km and 6 km, with a circuity less than 1.5. Further, the ten shortest paths are used for any feasible OD pair. This leads to a candidate route set of 700. Second, the parameters for network solution evaluation are determined according to transit type and network size. For the constraints of mathematical model, we assume vehicle capacity to be 50, with the route frequency between 1 and 12 per hour. For the user's cost evaluation model, the penalty $t_{penalty}$ for unsatisfied demand for the user's cost is set as 600 minutes, which is much larger than the sum of travel time and wait time cost on any feasible path. For operator, we assume an hourly personnel cost of around 100 dollars for each bus. This assumption is established based on the network size and the transportation method of bus. Also, travelling personnel cost occupies 70% in the operating cost. [19] Thus, we assume the parameter C_{km} of operating costs to be 3 dollars per kilometer, with the parameter C_{hr} of travelling personnel's cost to be 100 dollars per hour. Third, for the optimization algorithm of NSGA II, the population number is set as 250 with 100 generations. The crossover rate is set as 0.95 and the mutation rate is set as 0.05. The network solutions are initialized to be between 4 and 8 routes.

The TNDP procedure is coded in Matlab 2014a. With the candidate route set number of 700, each route is translated into a binary string with the length of 10. Since there are solutions, whose route number is less than the maximum route number, a continuous digit of 10 zeros would reveal an empty route number. The binary code of route number is shown in Table 5.1. In this application, systems with between 4 to 8 routes are considered. This reveals a result of 80-digit binary code for each gene. The non-empty digits representing route number would always be on the front, with empty digits representing no route number always recorded after those digits. According to the coding scheme, the crossover point must be chosen in non-empty digits for both parents and the mutation point must be chosen in the non-empty digits for the parent during the genetic operations.

Table 5.1: Example of Candidate Route Number in Binary Code

Candidate Route Number	Binary Code for the Route	Nodes Visited on Route
1	0000000001	AFEGC
2	0000000010	ABEGC
3	0000000011	AFEBC
4	0000000100	AIFEC
5	0000000101	AIJEC
null	0000000000	null

In this transit network example, a feasible network solution is defined to contain no route number larger than the candidate route number of 700, as well as no overlapped route number. After the genetic operation, infeasible solutions are dealt with in order that only feasible children are recorded in child population. If a solution has a route number larger than 700, the crossover or mutation point would be chosen again. If a solution with an overlapped route numbers appears, the decision of crossover or mutation would be made again and parents would be re-chosen.

5. 3 Non-dominated Front Solutions Analysis

In this example, solutions on the non-dominated front of 10 runs with population size of 250 and 100 generations is used. Indexes of user's cost, operator's cost, satisfied demand rate are calculated. The relationships for these parameters are analyzed below. The non-dominated front contains 755 network solutions and is sorted in an ascending order by satisfied demand rate.

Figure 5.3 illustrates the relationship between user's cost and operator's cost. Each point on the frontier is labeled with the number of routes in the solution, with the dot size linearly connected with the satisfied demand rate of the network. The solutions illustrate the negative relationship that exists between the user's cost and the operator's costs, while solutions with a larger number of routes tend to have a higher operator's cost and a lower user's cost. Although solutions with between 4 and 8 routes are considered, solution with route number 6 does not appear on the non-dominated front. This reveals that network with 6 routes has less competence in user's cost and operator's cost than networks solutions with other route numbers. Networks with route number of 7 or 8 would have less user's cost than networks with route number 6, while networks with route number of 4 or 5 would have less operator's cost than networks with route number 6. Moreover, networks with route number of 4 covers the majority of the network solutions with the percentage of 54.71%, while network with route number of 8 covers a larger percentage of 37.88%. This occurs since network of 4 routes would generates a low operator's cost and networks of 8 route would generates a low user's cost, which would grant advantages to these network solutions. Futher, the network solutions with larger route number tend to have higher satisfied demand rate, since the expansion of network would increase the capacity of network.

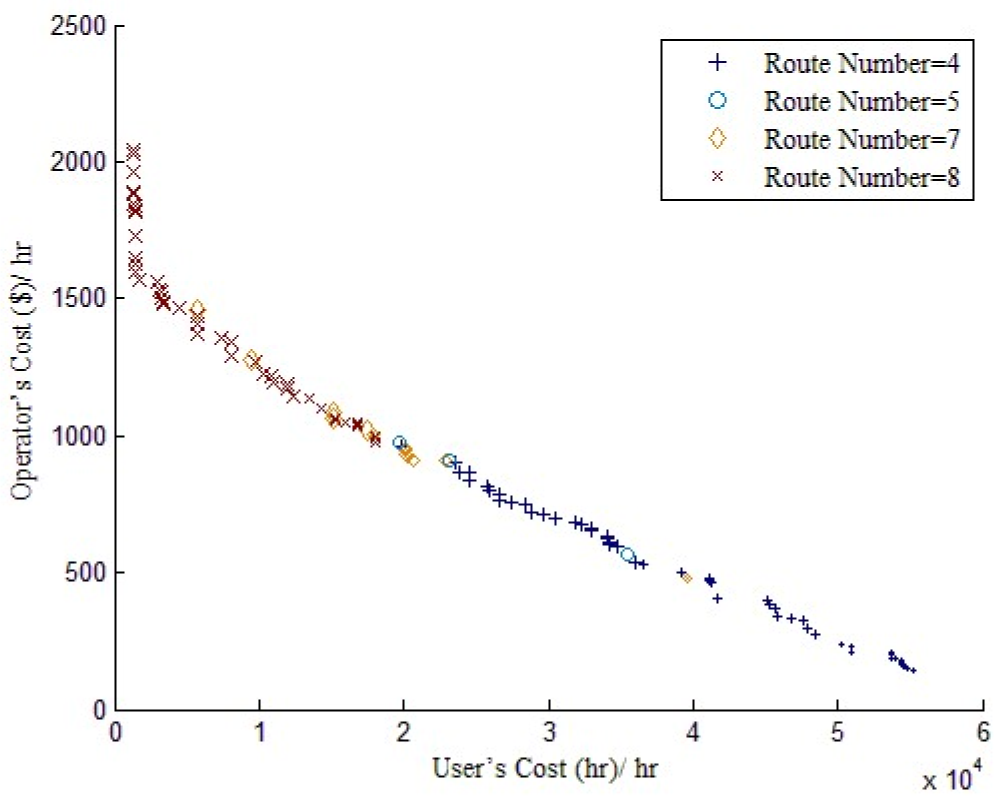


Figure 5.3: User's Cost and Operator's Cost of Network Solutions on the Pareto Front

Figure 5.4 illustrates the amount of demand satisfied in the solutions on the front, ordered in an ascending sequence. It is useful to observe that amount of satisfied demand ranges from around 7.56% to

100% on the non-dominated front. This difference between the lowest rate and highest rate is 92.44%, indicating that the ability of the solutions to satisfy the demands varies a lot. Generally, solutions that fail to accommodate a substantial number of trips have fewer routes and lower frequencies. This leads to low operator costs and high users' costs due to the penalties associated with unsatisfied demands.

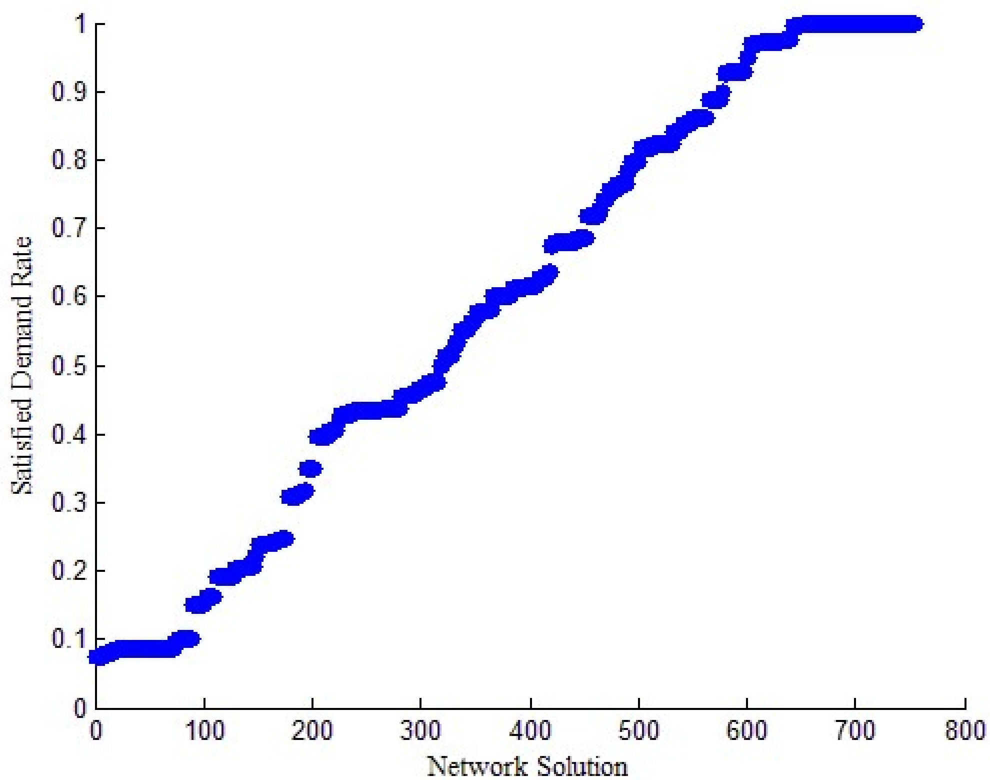


Figure 5.4: Satisfied Demand of Network Solutions

The Figures of 5.5 (a), 5.5 (b) and 5.5(c) implement the independent variable of the route number for the network solution, with the variables of user's cost, operator's cost and satisfied demand rate. From this graph, it can be observed that the operator's cost and satisfied demand rate rise with the route number increase, while the user's cost decrease with the rise of route number. This relationship can be well-explained, since increase of route number would give an expansion to the transit network, which would give traveler a better chance of travelling in the transportation network. Also, with the expansion of network, the system tends to have longer routes with higher frequencies, which would increase the operator's cost. Thus, there is lower unsatisfied demand penalty, lower user's cost, more satisfied demand rate and more operator's cost with larger numbers of routes.

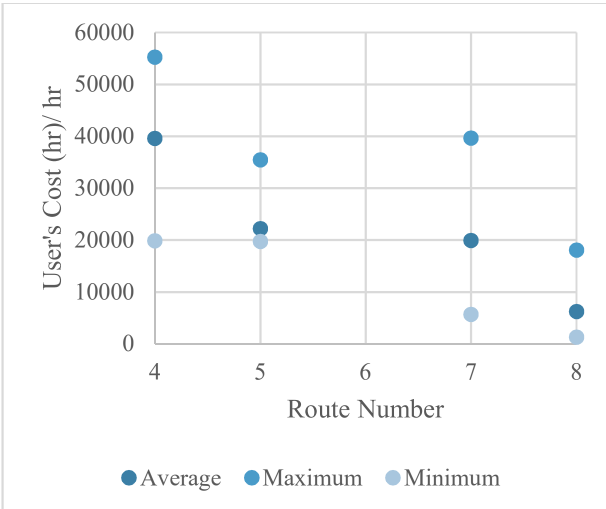


Figure 5.5 (a): The User's Cost as a Function of the Number of Routes

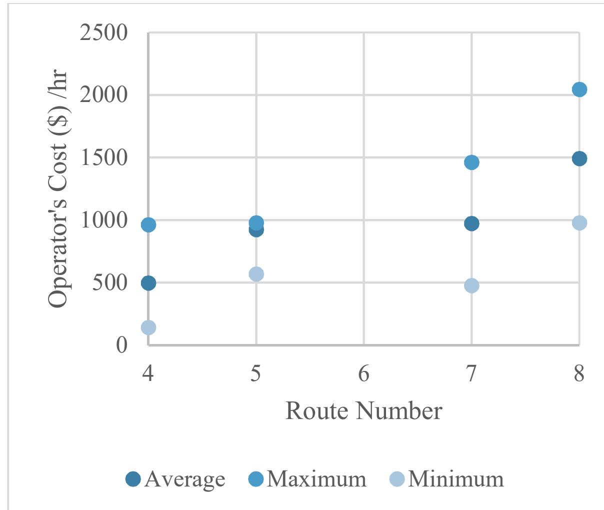


Figure 5.5 (b): The Operator's Cost as a Function of the Number of Routes

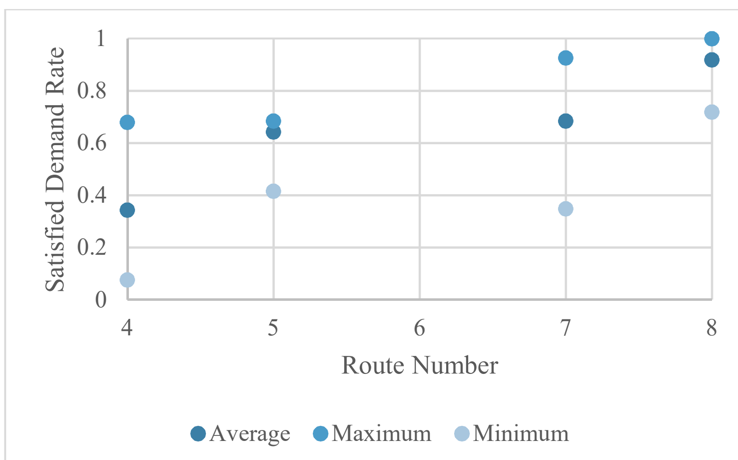


Figure 5.5(c): The Satisfied Demand Rate with Network Solution Route Number

Figures 5.6 (a) and Figure 5.6 (b) illustrate the relationship between the satisfied demand rate versus the two costs. From the first graph, we would observe that the satisfied demand rate decreases approximately linearly with the user's costs. It is very close to linear because the penalty per trip not accommodated is so large giving the relationship a strongly linear character. From the second graph, it can be observed that the satisfied demand rate rises with the operator's cost. The rise of the operator's cost occurs because the service network is expanding to provide service to more trips. These are some solutions reaches the 100% satisfied demand rate with different operator' costs, where solutions with more operator' cost tend to have less user's cost.

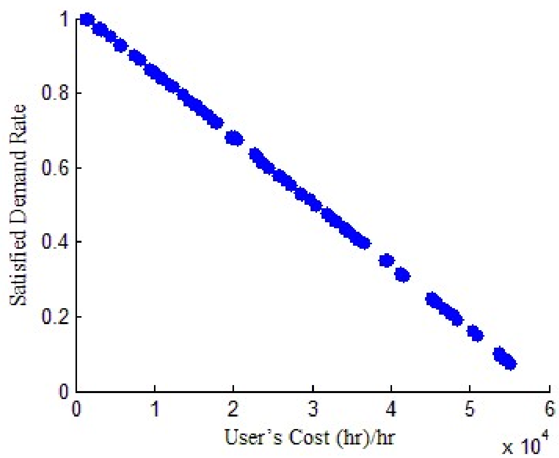


Figure 5.6 (a): Relationship between Satisfied Demand Rate and User's Cost

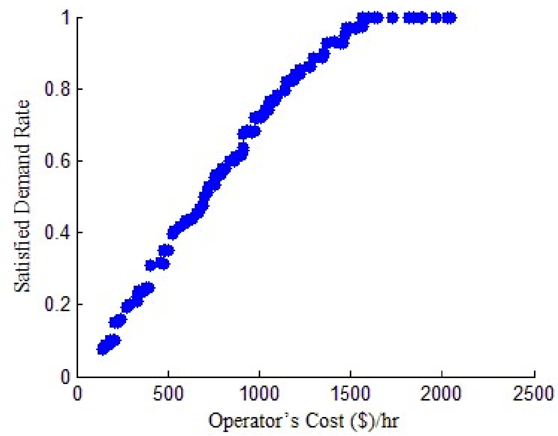


Figure 5.6 (b): Relationship between Satisfied Demand Rate and Operator's Cost

5. 4 Satisfied User's Time and Monetary Costs Analysis

Once the transit system is established, the accommodated users are the major participants of the transit system. Thus, the total network time impedance (without penalty for unsatisfied demand) and the operator's monetary cost are averaged over satisfied user number for the non-dominated solutions, which would provide a clear pattern of satisfied traveler's average time and monetary cost. Relationship for the two average indexes and satisfied demand rate are studied.

Figure 5.7 below gives the relationship between the user's average time cost (for those trips accommodated) and the satisfied demand rate. It is useful to notice that, in these two dimensions, as the satisfied demand rate increases the average trip time per passenger served decrease. This occurs because as more and more demand is accommodated, routes and higher frequencies are needed. Thus, with the expansion of the network, a better route choice may be provided and less wait time can be realized, which would cause the reduction of average trip time.

It is important to realize that the points that are dominated in these two dimensions have lower operator costs. For example, the points of A, B and C are highlighted, and their indexes are listed in Table 5.2. Point A dominates point B with the same satisfied demand rate and lower travel time for accommodated users. However, point B has an advantage over A because the operator's costs are lower. Point C is dominated by the other two solutions with lower travel time for satisfied users and a higher satisfied demand rate.

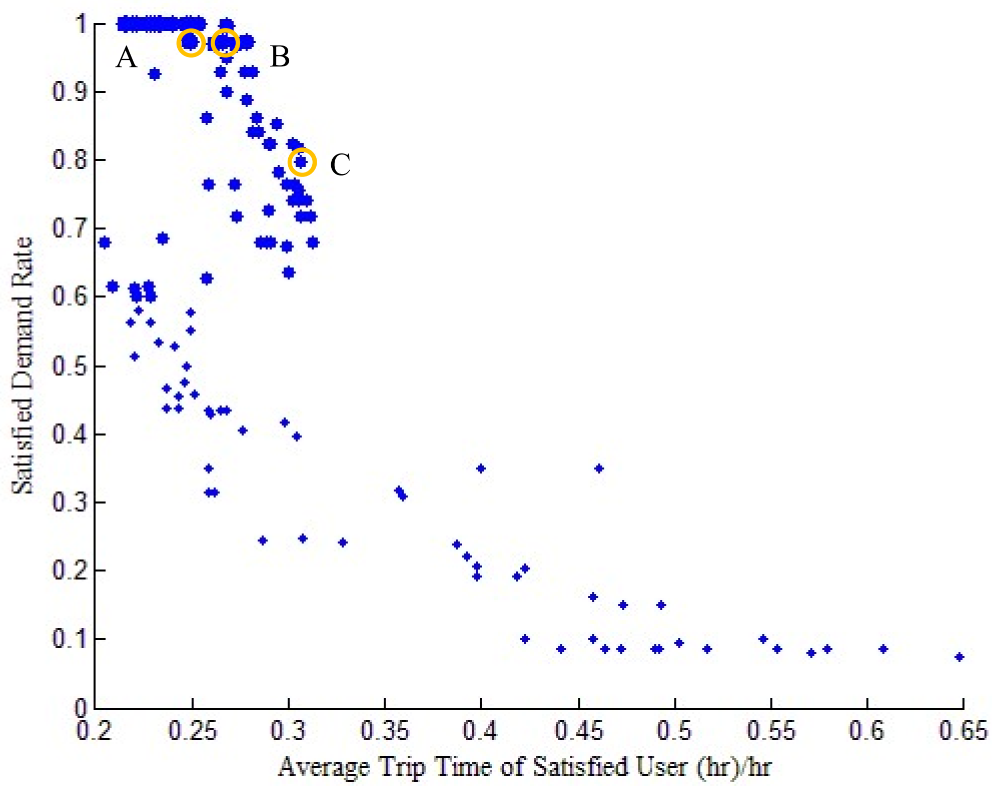


Figure 5.7: Satisfied User's Trip Time Cost with Satisfied Demand Rate

Table 5.2: Indexes for Point A, Point B and Point C

	User's Cost (hr)/ hr	Operator's Cost ($)/hr	Satisfied Demand Rate	Average8 Trip Time of Satisfied User (hr)/ hr	Domination Relationship
Point A	3,167.71	1532.205	97.09%	0.2493	Dominates B and C
Point B	3,333.68	1476.62	97.09%	0.2780	Dominated by A, dominates C
Point C	13,530,84	1140.25	79.67%	0.3065	Dominated by B and C

Figure 5.8 illustrates how the cost per tip incurred by the operator decrease as the number of trips accommodated increases with the range of 122.96 to 0.25, which reveal a huge difference.

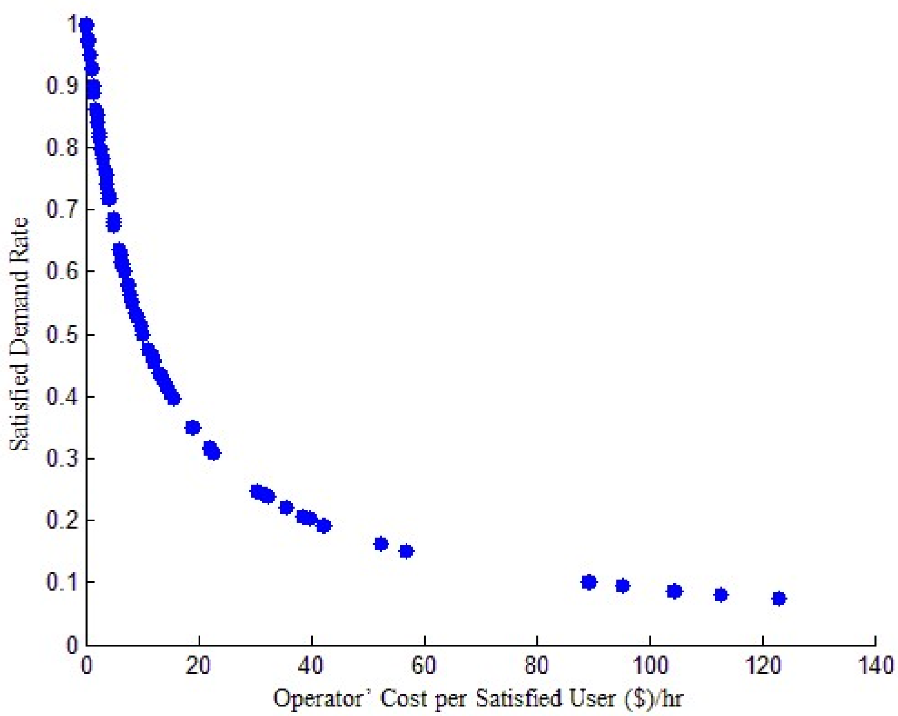


Figure 5.8: Operator's Cost per Satisfied User and Satisfied Demand Rate

5.5 Network Solution Examples Analysis

To study the characteristics of network solutions, three solution networks are selected according to their satisfied demand rate, as listed in Table 5.3. This table describes three network solutions including their routes and associated frequencies, users' costs, operator's costs, satisfied demand rate and average trip time cost per traveler. Network solution 200 has 4 routes with an averaged frequency of 4.21. Network solution 450 has 5 routes with an averaged frequency of 6.69. Network 700 has 8 routes with an averaged frequency of 7.59. The rise in the number of routes and the increased frequency indicates the expansion of transit network, which results in more satisfied demand rate, lower user's cost and higher operator's costs.

Table 5.3: Three Network Solutions

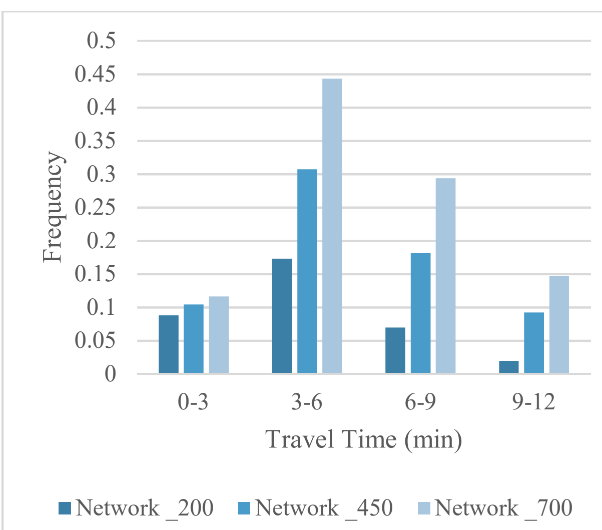
Network Number	Routes, Length and Frequencies	User's Cost (hr) /hr	Operator's Cost ($)/hr	Satisfied Demand Rate	Average Trip Time (hr)/ hr
200	[123, BEKLH, 3.1km, Freq: 4.38/ hr] [664, NOKPLS, 3.1 km, Freq: 3.77/ hr] [669, NORPLS, 3.2 km, Freq: 1.62/ hr] [581, IOKLT, 3 km, Freq: 7.05/ hr]	39167	498.8	34.98%	0.2585
450	[64, AFJKLP, 3.1 km, Freq: 8.23/ hr] [407, AFEJKL, 3.2 km, Freq: 3.42/ hr] [622, LSPKON, 3 km, Freq: 5.05/ hr] [510, HDGKOI, 3.4 km, Freq: 12/ hr] [689, ORPKLT, 3.2 km, Freq: 2.71/ hr]	19713	977.1	68.42%	0.2343
700	[32, AFJKLH, 3.3 km, Freq: 7.54/ hr] [142, BEGKL, 3 km, Freq: 9.64/ hr] [284, CEKPRS, 3.1 km, Freq: 8.24/ hr] [389, EJKPON, 3.3 km, Freq: 6.50/ hr] [419, FJKLMH, 3.2 km, Freq: 5.12/ hr] [510, HDGKOI, 3.4 km, Freq:10.44/hr] [557, IOKPSL, 3.1 km, Freq: 5.36/hr] [695, QOKPLT, 3.2 km, Freq: 7.88/ hr]	1356	1880.2	100%	0.2283

Figure 5.9 shows histograms for travel time, wait time, circuity and occupancy and the corresponding frequencies. For travel time and wait time and circuity histogram, the corresponding frequency is defined as the number of passengers within the variable range divided by the total demand. This would reveal a sum of frequency as satisfied demand rate for each network solution. Frequency of occupancy is stated afterwards, since it has more to do with the definition of occupancy.

For Figure 5.9 (a) and Figure 5.9 (b), travel time wait time and their frequencies are shown. Travel time is defined as the sum of link time cost for the traveler, while wait time is calculated by the sum of wait time at the starting or transfer station. Within the satisfied demand, most travel time lie within 9 minutes for the transit network, namely 94.42%, 86.56% and 85.31% for Network 200, Network 450 and Network 700. Moreover, wait time for most trips are under 15 minutes. Within the satisfied demand, 81.95%, 89.56% and 97.43 can be satisfied with no more than 10 minutes of wait time in solutions 200, 450 and 700, respectively.

Circuity is defined as the path travel time divided by the shortest path travel time (supposing all 43 links are included). Figure 5.9(c) shows the circuity with associated frequency of the three networks. The average circuity for the networks are 1.05, 1.13 and 1.08. Within the satisfied demand, 96.45%, 92.49%, and 94.68% are below 1.5 circuities in solutions 200, 450 and 700, respectively.

Given a directed path, occupancy is calculated by link flow volume divided by link capacity for each route. The frequency is calculated by the link capacity for a specified route divided by the network capacity. Thus, the sum of frequency would be added to one for each network solution. The average occupancy is 41.24% for Network 200, 35.41% for Network 450 and 31.61% for Network 700. Figure 5.9 (d) shows the occupancy for link with associated frequency of the three networks. From the graph, we would observe for around one third of the occupancy would be 0, where the operator would pay for operation on these links when no passengers has need to travel on the links. Also, suppose the occupancy of 0-0.75 to be a comfortable occupancy for passenger, there are 40.68%, 39.92% and 48.21% occupancies are in this range for Network 200,450, and 700 separately.



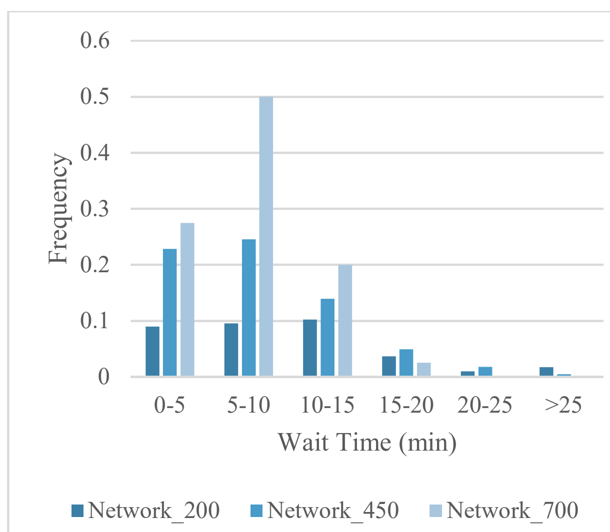


Figure 5.9 (a): Travel Time for Three Network

Solutions

Figure 5.9 (b): Wait Time for Three Network

Solutions

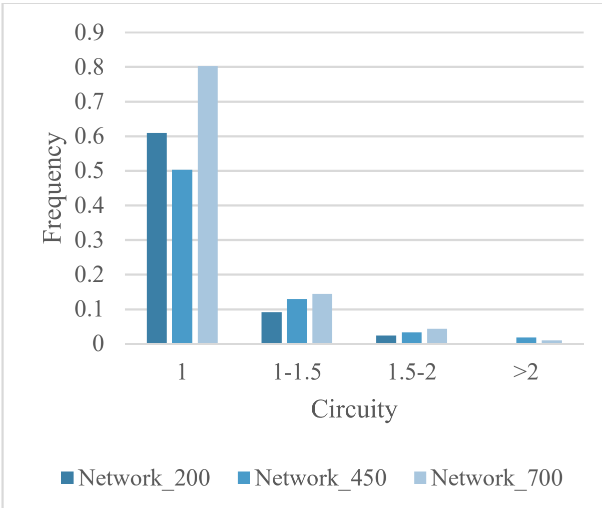


Figure 5.9 (c): Circuity for Three Network

Solutions

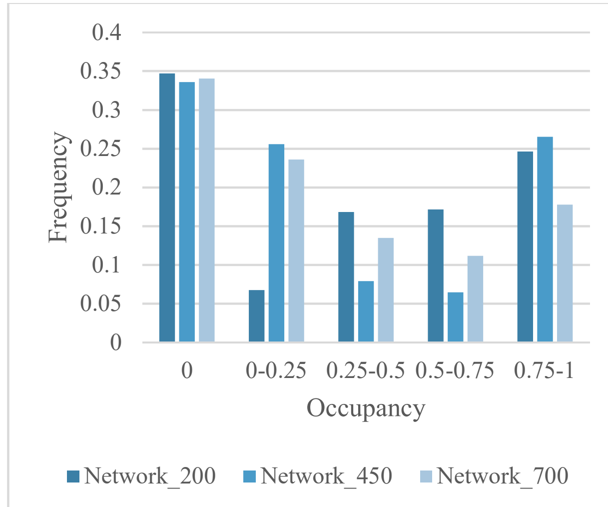


Figure 5.9 (d): Occupancy for Three Network

Solutions

Chapter 6   Conclusion and Future Extension

In this thesis, a mathematical model for generating the optimized route pattern with associated frequencies is formulated, with two objective function formulas of user's and operator's cost. Then, a TNDP solving procedure based on NSGA II is developed, which generate a trade-off frontier between the two objectives. Further, this procedure is applied to an illustrative example which contains 20 nodes and 43 links. The non-dominated solutions from 10 runs of 250 population 100 generations are sorted out. The indexes of the network solutions are studied for their values and their relationships.

There are at least a couple of areas for future research. First, there is substantial opportunity to make use of parallel computing. For example, different parameters for crossover and mutation can be carried out simultaneously allowing the solution procedure to dynamically pursue those parameters more heavily that lead to better child solutions. Second, research focused on the application of this TNDP procedure to an actual-size city network is valuable. Comparison can be made between the city's current transit network situation and the result of the transit network design application. Third, additional efforts are also warranted in expanding the objectives considered. Illustrative additional objectives include environmental impacts and congestion reduction impacts. Finally, the fare of bus and the actual operator's cost can be compared, and the time for reclaiming the construction fee of bus system can be estimated.

Reference

[1] Pattnaik, S. B., S. Mohan, and V. M. Tom. "Urban bus transit route network design using genetic algorithm." *Journal of transportation engineering* 124.4 (1998): 368-375.

[2] American Public Transportation Association, 2014. *2014 Public Transportation Fact Book*.

[3] American Public Transportation Association, 2015. Public Transportation Benefits. http://www.apta.com/mediacenter/ptbenefits/Pages/default.aspx (accessed July 2015)

[4] Kepaptsoglou, K., and M. Karlaftis. "Transit route network design problem: review." *Journal of transportation engineering* 135.8 (2009): 491-505.

[5] Guihaire, Valérie, and Jin-Kao Hao. "Transit network design and scheduling: A global review." *Transportation Research Part A: Policy and Practice* 42.10 (2008): 1251-1273.

[6] Farahani, Reza Zanjirani, et al. "A review of urban transportation network design problems." *European Journal of Operational Research* 229.2 (2013): 281-302.

[7] Yan, Shangyao, and Hao-Lei Chen. "A scheduling model and a solution algorithm for inter-city bus carriers." *Transportation Research Part A: Policy and Practice* 36.9 (2002): 805-825.

[8] Borndörfer, Ralf, Martin Grötschel, and Marc E. Pfetsch. *A path-based model for line planning in public transport*. Konrad-Zuse-Zentrum für Informationstechnik Berlin, 2005.

[9] Bagloee, Saeed Asadi, and Avishai Avi Ceder. "Transit-network design methodology for actual-size road networks." *Transportation Research Part B: Methodological* 45.10 (2011): 1787-1804.

[10] Tom, V. M., and S. Mohan. "Transit route network design using frequency coded genetic algorithm." *Journal of Transportation Engineering* 129.2 (2003): 186-195.

[11] Zhao, Fang, and Ike Ubaka. "Transit network optimization-minimizing transfers and optimizing route directness." *Journal of Public Transportation*7.1 (2004): 4.

[12] Beltran, Borja, et al. "Transit network design with allocation of green vehicles: A genetic algorithm approach." *Transportation Research Part C: Emerging Technologies* 17.5 (2009): 475-483.

[13] Cipriani, Ernesto, Stefano Gori, and Marco Petrelli. "Transit network design: A procedure and an application to a large urban area." *Transportation Research Part C: Emerging Technologies* 20.1 (2012): 3-14.

[14] Brands, Ties, Luc JJ Wismans, and Eric C. Van Berkum. "Multi-objective transportation network design: Accelerating search by applying ε-NSGAII." *2014 IEEE Congress on Evolutionary Computation (CEC)*. IEEE, 2014.

[15] Fan, Wei, and Randy B. Machemehl. "Using a simulated annealing algorithm to solve the transit route network design problem." *Journal of transportation engineering* 132.2 (2006): 122-132.

[16] http://bca.transportationeconomics.org/benefits/vehicle-operating-cost

[17] Shevchik, V. I. "Multicommodity flow problem." *Journal of Mathematical Sciences* 65.1 (1993): 1462-1464.

[18] Deb, Kalyanmoy, et al. "A fast and elitist multi-objective genetic algorithm: NSGA-II." *IEEE transactions onevolutionary computation* 6.2 (2002): 182-197.

[19] http://publictransport.about.com/od/Transit_Vehicles/a/How-Much-Does-A-Bus-Cost-To-Purchase-And-Operate.htm

Appendix

Appendix A: A Transit Network Analysis Example

This transit network analysis example is aimed at presenting the network solution evaluation process under routes information and traveler's demand. It is constructed in the sequence of network base information statement, unsatisfied demand minimization model result presentation and objectives calculation process illustration. First, the network information and traveler's demand is demonstrated in the Figure A.1 and Table A.1. Next, the maximum multi-commodity flow model in formula (9) – (16) is implemented to generate the traveler's flux on each path, unsatisfied demand for each OD pair and the frequency of each route. Results are shown in Table A.2 and A.3 for the details of routes and passenger's flux information. Finally, user's cost and operator's cost calculation process is demonstrated in Table A.4 and Table A.5.

Figure A.1 illustrated the example network, with nodes, link impendence and routes. This network contains 5 nodes, namely 'A', 'B', 'C', 'D', 'E'. Three routes are included in the network, illustrated with the color of white, grey and black in the figure. Vehicles are assumed to travel on the network with the speed of 15 km/h, with the capacity of each vehicle of 50.

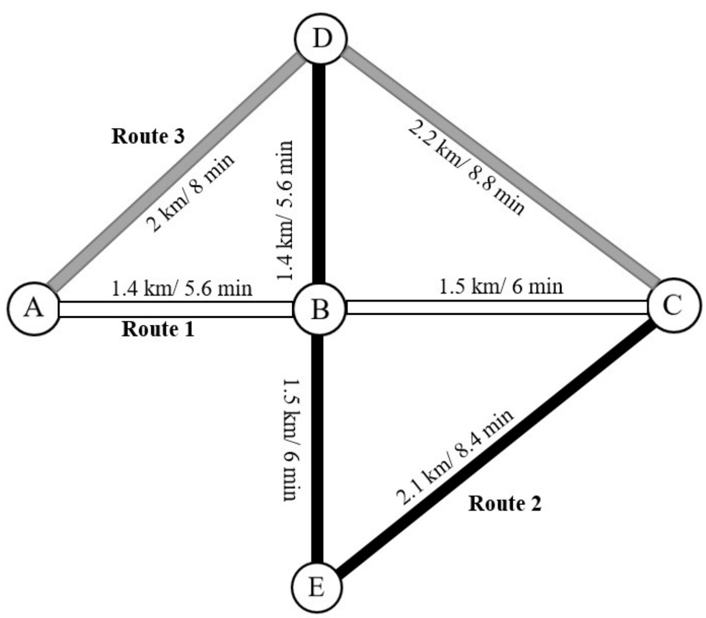


Figure A.1: An Example Transit Network

For the passenger's travel demand, Table A.1 illustrates the averaged travel demand from origins to destinations in an hourly rate.

Table A.1: Demand Matrix of the Example Network per hour

OD	A	B	C	D	E
A		170	150	180	200
B	100		110	50	160
C	50	150		30	70
D	180	50	150		80
E	30	80	120	100	

With the input of the network base information and the traveler's demand, the optimized frequency for each route can be calculated by formula (9) - (16). The minimum frequency is set as 1, while maximum frequency is set as 12 during this calculation. Details of the routes as listed below in Table A.2.

Table A.2: Details of Routes in the Example Network

Name	Nodes	Frequency/ hr	Capacity of Vehicle
R1	A - B - C	6.29	50
R2	D - B - E - C	6.89	50
R3	A - D - C	5.60	50

The shortest path for each OD pair, the travel time cost, wait time cost and the flux along each path are also generated from the maximum multi-commodity model with formula (9) –(16). Wait time is counted as half the headway here. The results are listed in Table A.3 below.

Table A.3: Traveler's information between Node Pairs of the Example Network

Origin Node	Destination Node	The Shortest Path	Route Number	Path Travel Time (min)	Path Wait Time (min)	Flux on each path/ hr
A	B	A - B	R1	5.6	4.77	70
A	C	A - B - C	R1	11.6	4.77	50
A	D	A - D	R3	8	5.36	80
A	E	A- B - E	R1+ (b) + R2	11.6	4.77	60
B	A	B - A	R1	5.6	4.77	100
B	C	B - C	R1	6	4.77	40
B	D	B - D	R2	5.6	4.36	50
B	E	B - E	R2	6	4.36	160

C	A	C - B - A	R1	11.6	4.77	50
C	B	C - B	R1	6	4.77	40
C	D	C - D	R3	8.8	5.36	230
C	E	C - E	R2	8.4	4.36	70
D	A	D - A	R3	8	5.36	280
D	B	D - B	R2	5.6	4.36	50
D	C	D - C	R3	8.8	5.36	150
D	E	D - B - E	R2	11.6	4.36	80
E	A	E - B - A	R2 + (b) + R1	11.6	4.36	164.42
E	B	E - B	R2	6	4.36	80
E	C	E - C	R2	8.4	4.36	120
E	D	E - B - D	R2	11.6	4.36	100

* " + (node) +" between route numbers indicates that passengers need to transfer from one route to another at this node.

In the network solution evaluation model, the objective of user's cost and operator's cost are calculated by formula (7) and (8). In user's cost evaluation, the time penalty for the unsatisfied demand is defined 10 hour for a traveler. In operator's cost evaluation, the parameter for maintenance cost C_{km} is defined as 3, while the parameter for the operating cost C_{hr} is defined as 100. The user's cost is presented in Table A.4, while the operator's cost is illustrated in Table A.5.

Table A.4: User's Cost Calculation

OD Pair	Path Flow /hr	Path Cost (min)	Total Path Time for each OD pair (hr)/ hr
A-B	70	10.37	12.10
A-C	50	16.37	13.64
A-D	80	13.36	17.81
A-E	60	16.37	16.37
B-A	100	10.37	17.28
B-C	40	10.77	7.18
B-D	50	9.96	8.3
B-E	160	10.36	27.63
C-A	50	16.37	13.64
C-B	40	10.77	7.18
C-D	230	14.16	54.28
C-E	70	12.76	14.89
D-A	280	13.36	62.35
D-B	50	9.96	8.3
D-C	150	14.16	35.4
D-E	80	15.96	21.28
E-A	164.42	15.96	43.74
E-B	80	10.36	13.81
E-C	120	12.76	25.52
E-D	100	15.96	26.6
Satisfied Demand:	2024.42	Total Travel Time (hr)/ hr	447.30
Unsatisfied Demand:	65.58	Total Penalty for Unsatisfied Demand (hr)/ hr:	655.8
Total Demand:	2090	User's Cost:	1103.1

* Path cost is defined as the sum of path travel time and path wait time given in Table A.3.

Table A.5: Operator's Cost Calculation

Route Number	Route Distance (km)	Route Time (min)	Frequency /hr	System Operating Cost ($)/ hr	Travelling Personnel's Cost ($)/ hr
R1	2.9	11.6	6.29	54.71	121.57
R2	5	20	6.89	103.33	229.61
R3	4.2	16.8	5.6	70.56	156.8
			Sum Cost ($)/hr:	228.59	507.99
			Operator' Cost ($)/ hr:	736.58	

* System Operating Cost is defined as the sum of the product of the route distance and frequency. System Operating Cost is defined as the sum of the product of the route time cost and frequency.

Appendix B: Accepted Potential Routes Set of the Experimental Transit Network

Table B.1: Candidate Route Set Detail

Route Number	Nodes	Route Number	Nodes	Route Number	Nodes	Route Number	Nodes
1	AFEGC	36	ABEGH	71	AIONQ	106	AFEGLT
2	ABEGC	37	ABECDH	72	AFINOQ	107	AFEKLT
3	AFEBC	38	AFJEGH	73	AFJKOQ	108	AFJKLST
4	AIFEC	39	ABEGDH	74	AFJIOQ	109	AIOPST
5	AIJEC	40	AFJKGH	75	AFJINQ	110	AFEKPST
6	AFJEGC	41	AFEGL	76	AFIONQ	111	BECGD
7	AFJKGC	42	AFEKL	77	AIORQ	112	BCGHD
8	AFJEBC	43	AIOKL	78	AFJKPRQ	113	BEGCD
9	AFJKEC	44	AIJKL	79	AFEKOQ	114	BEKGD
10	AIFJEC	45	ABEGL	80	AFJKPOQ	115	BEGLHD
11	AFEGD	46	ABEKL	81	AFIOR	116	BEKLHD
12	ABCD	47	AFEJKL	82	AINOR	117	BECGHD
13	AFECD	48	AIOPL	83	AIOPR	118	BEJKGD
14	ABECD	49	AFJEGL	84	AFEKPR	119	BCGLHD
15	ABEGD	50	AFEKPL	85	AIOQR	120	BEJKLHD
16	AFJECD	51	AFJKLM	86	AFJKPSR	121	BCGDH
17	AFJEGD	52	AFEGHM	87	AINQR	122	BEGLH
18	AFJKGD	53	AFJKLHM	88	AIOKPR	123	BEKLH
19	AFEGHD	54	AFEGDHM	89	AIJKPR	124	BECGH
20	AFJKLHD	55	ABCDHM	90	ABEKPR	125	BECGDH
21	ABECG	56	AFECDHM	91	AIORS	126	BCGLH
22	AFEKG	57	AFJKLTM	92	AFJKPRS	127	BEJKLH
23	AFJKLG	58	AFJKPLM	93	AFJKLS	128	BEGCDH
24	AFJECG	59	ABEGHM	94	AIOPS	129	BEKGH
25	AIOKG	60	AFEGLM	95	AFEKPS	130	BEKPLH
26	AIJKG	61	AIORP	96	AIOKPS	131	BEFAI
27	AIFEG	62	ABEKP	97	AIJKPS	132	BAFJI
28	AIJEG	63	AFEJKP	98	AIORPS	133	BEKOI
29	ABEKG	64	AFJKLP	99	AFJKPLS	134	BEKJI
30	AFEJKG	65	AFJEKP	100	ABEKPS	135	BCEFI
31	AFEGH	66	AFIOP	101	AFJKLT	136	BCEJI
32	AFJKLH	67	AINOP	102	AFJKPST	137	BEJFAI
33	AFEGDH	68	ABEJKP	103	AFJKPLT	138	BEJKOI
34	ABCDH	69	AFIOKP	104	AIORST	139	BEKJFI
35	AFECDH	70	AIFJKP	105	AFJKPRST	140	BEKONI

Route Number	Nodes	Route Number	Nodes	Route Number	Nodes	Route Number	Nodes
141	BEJKPL	176	BEFJKO	211	BEGLS	246	CGKOQN
142	BEGKL	177	BEJINO	212	BEKLS	247	CGKPON
143	BEFJKL	178	BEJKPO	213	BEKPRS	248	CEFION
144	BCDHL	179	BEKPRO	214	BCGLS	249	CEJION
145	BEKGL	180	BEFINO	215	BEJKLS	250	CGKOIN
146	BEKPSL	181	BCGKP	216	BEJKPRS	251	CEFIO
147	BCGKL	182	BCGLP	217	BEGKPS	252	CEJIO
148	BEGHL	183	BEJKLP	218	BEGLPS	253	CGLPO
149	BEGKPL	184	BEGLKP	219	BEKLPS	254	CGKPRO
150	BECDHL	185	BAFJKP	220	BEKPLS	255	CEKPO
151	BEGHM	186	BCEKP	221	BEGLT	256	CGLPKO
152	BECDHM	187	BECGKP	222	BEKLT	257	CDGKO
153	BEGDHM	188	BEGLSP	223	BCDHMT	258	CEGKO
154	BCGHM	189	BEKLSP	224	BEKPST	259	CGLPRO
155	BEGLM	190	BEKOP	225	BCGLT	260	CGLKPO
156	BEKLM	191	BEKOQ	226	BEJKLT	261	CEKOQ
157	BCGDHM	192	BAIOQ	227	BEGHMT	262	CGKPRQ
158	BEGLHM	193	BAINQ	228	BECDHMT	263	CGLKOQ
159	BEKLHM	194	BEJKOQ	229	BEJKPST	264	CEJKOQ
160	BCGLM	195	BEJIOQ	230	BEGDHMT	265	CGKPOQ
161	BAFIN	196	BEFIOQ	231	CGKOI	266	CGLPRQ
162	BEKON	197	BEFINQ	232	CGEFI	267	CEFIOQ
163	BEJFIN	198	BEJINQ	233	CGEJI	268	CEJIOQ
164	BAION	199	BEKPRQ	234	CGKJI	269	CGLPOQ
165	BEFJIN	200	BEKPOQ	235	CBAI	270	CEFINQ
166	BEJKON	201	BEKPSR	236	CEFAI	271	CDHLPR
167	BEJION	202	BEGKPR	237	CEKOI	272	CGLPSR
168	BEFION	203	BEGLPR	238	CBEFI	273	CGKLPR
169	BEFAIN	204	BEKLPR	239	CBEJI	274	CGKOR
170	BAFJIN	205	BEFJKPR	240	CEKJI	275	CDHLSR
171	BEKPO	206	BEJKPSR	241	CEKON	276	CEKPSR
172	BAFIO	207	BEKOR	242	CEJFIN	277	CDGKPR
173	BAINO	208	BCGKPR	243	CGLKON	278	CDHMTSR
174	BEGKO	209	BEGLSR	244	CEJKON	279	CEGKPR
175	BEJFIO	210	BEKLSR	245	CEFJIN	280	CGKLSR

Route Number	Nodes	Route Number	Nodes	Route Number	Nodes	Route Number	Nodes
281	CGLTS	316	DGEJI	351	DHLKOQ	386	EJFION
282	CEKLS	317	DHLKOI	352	DHLPRQ	387	EKJIN
283	CDHLPS	318	DHLKJI	353	DHLPOQ	388	EJKOQN
284	CEKPRS	319	DHLPOI	354	DHLSRQ	389	EJKPON
285	CDGLS	320	DGKJFI	355	DGKPRQ	390	EKPRON
286	CEGLS	321	DHMLKJ	356	DHMTSRQ	391	EGKOQ
287	CGHLS	322	DCGEJ	357	DGLKOQ	392	EJKPRQ
288	CGKLPS	323	DCGKJ	358	DHGKOQ	393	EKONQ
289	CGKPLS	324	DGKEJ	359	DHLPKOQ	394	EJFIOQ
290	CDHMLS	325	DGLPKJ	360	DGKPOQ	395	EJKPOQ
291	CGKLT	326	DHGLKJ	361	DGKPSR	396	EKPROQ
292	CGKPST	327	DHLGKJ	362	DHGLPR	397	EFINOQ
293	CGLHMT	328	DHLGEJ	363	DHMLKPR	398	EFJKOQ
294	CGLPST	329	DHLKEJ	364	DGLPSR	399	EJFINQ
295	CEKLT	330	DHLSPKJ	365	DGLSPR	400	EJINOQ
296	CDGLT	331	DGKON	366	DHLKPSR	401	EGLST
297	CEGLT	332	DHLKON	367	DHMTLPR	402	EKLST
298	CGHLT	333	DHLPON	368	DCGKPR	403	EKPRST
299	CGKPLT	334	DGLKON	369	DGKLPR	404	ECGLT
300	CDHLST	335	DHGKON	370	DHGLSR	405	EGLMT
301	DGCEF	336	DHLPKON	371	EGLTM	406	EKLMT
302	DCGEF	337	DGKPON	372	EKLTM	407	EJKPLT
303	DGKEF	338	DHLPRON	373	EKPLM	408	EGKLT
304	DGLKJF	339	DGKOQN	374	EKPSTM	409	EGKPST
305	DHGKJF	340	DHMLKON	375	EJKLHM	410	EJKLST
306	DHLPKJF	341	DGKPO	376	EGCDHM	411	FJEGH
307	DHGEJF	342	DHLPRO	377	EKGHM	412	FJKGH
308	DGCEJF	343	DHMLKO	378	EKPLHM	413	FJKPLH
309	DHLGEF	344	DGLPO	379	ECGLM	414	FEGLH
310	DHLKEF	345	DHLKPO	380	EJKLTM	415	FEKLH
311	DGKOI	346	DHLSRO	381	EFAIN	416	FJEGDH
312	DGKJI	347	DCGKO	382	EKOQN	417	FJKGDH
313	DCEFI	348	DGKPRO	383	EKPON	418	FJECDH
314	DCEJI	349	DGLPKO	384	EGKON	419	FJKLMH
315	DGEFI	350	DHLSPO	385	EKOIN	420	FECGH

Route Number	Nodes	Route Number	Nodes	Route Number	Nodes	Route Number	Nodes
421	FEGHM	456	FIOKPS	491	GKPROQ	526	HLKPON
422	FECDHM	457	FEJKLS	492	GLPKOQ	527	HMLKON
423	FEGDHM	458	FEJKPRS	493	GLKPRQ	528	HLKOQN
424	FJKLTM	459	FIORPS	494	GLPROQ	529	HLPRQN
425	FJKPLM	460	FJKLPRS	495	GLKPOQ	530	HLSRON
426	FEGLM	461	FJKPLT	496	GEKOQ	531	HLSPO
427	FEKLM	462	FEGLT	497	GKPSRQ	532	HMLPO
428	FJKPSTM	463	FEKLT	498	GLPSRQ	533	HMTSRO
429	FJEGHM	464	FJKLST	499	GLSPRQ	534	HGLKO
430	FJKGHM	465	FJKPRST	500	GLSROQ	535	HLGKO
431	FAINQ	466	FEKPST	501	HLPOI	536	HLKPRO
432	FEKOQ	467	FJKLMT	502	HGKOI	537	HMTSPO
433	FJKPRQ	468	FEJKLT	503	HLPKOI	538	HDGLKO
434	FJINOQ	469	FJKLHMT	504	HGEFI	539	HGKPO
435	FJKPOQ	470	FEGHMT	505	HLKJFI	540	HLSPKO
436	FEJKOQ	471	GCEFI	506	HLPROI	541	HLSRQ
437	FJKONQ	472	GCEJI	507	HGEJI	542	HLPKOQ
438	FEJIOQ	473	GEFJI	508	HGKJI	543	HGKOQ
439	FJIONQ	474	GLPOI	509	HLPKJI	544	HMTSRQ
440	FAINOQ	475	GKPROI	510	HDGKOI	545	HLKPRQ
441	FIOQR	476	GLPKOI	511	HGLKJ	546	HLPROQ
442	FJKLSR	477	GLPKJI	512	HLGEJ	547	HDGKOQ
443	FINQR	478	GEFAI	513	HLGKJ	548	HLKPOQ
444	FIOKPR	479	GKEFI	514	HLKEJ	549	HMLKOQ
445	FEKPSR	480	GKEJI	515	HDGLKJ	550	HLPSRQ
446	FEGKPR	481	GEJIN	516	HGEFJ	551	IFJKPL
447	FIJKPR	482	GKJIN	517	HLSPKJ	552	INOKPL
448	FAIOR	483	GEFIN	518	HMLPKJ	553	IOKGL
449	FEGLPR	484	GKPRON	519	HMTLKJ	554	IORPKL
450	FEKLPR	485	GLPKON	520	HGCEJ	555	IFEGL
451	FEGLS	486	GLPRON	521	HLPON	556	IJEGL
452	FEKLS	487	GKPRQN	522	HLPKON	557	IOKPSL
453	FEKPRS	488	GLKOQN	523	HGKON	558	IFEKL
454	FIOPS	489	GLKPON	524	HLPRON	559	IJEKL
455	FJEKPS	490	GEKON	525	HDGKON	560	INORPL

Route Number	Nodes	Route Number	Nodes	Route Number	Nodes	Route Number	Nodes
561	HLPON	596	IJEGL	631	MLKON	666	NOKLPS
562	HLPKON	597	IOKPSL	632	MLPON	667	NQOPRS
563	HGKON	598	IFEKL	633	MTSRON	668	NIOPRS
564	HLPRON	599	IJEKL	634	MHLKON	669	NORPLS
565	HDGKON	600	INORPL	635	MTSPON	670	NIOQRS
566	HLKPON	601	JFINOQ	636	MLPKON	671	NOKPST
567	HMLKON	602	JKPSRQ	637	MTLKON	672	NORPST
568	HLKOQN	603	JEFIOQ	638	MHLPON	673	NOPLT
569	HLPRQN	604	JFIONQ	639	MTSRQN	674	NOPRST
570	HLSRON	605	JKLPRQ	640	MLPRON	675	NOKPLT
571	HLSPO	606	JEFINQ	641	MHGKO	676	NOQRST
572	HMLPO	607	JEKPRQ	642	MHLPKO	677	NQORST
573	HMTSRO	608	JKLPOQ	643	MLKPO	678	NQRPST
574	HGLKO	609	JKORQ	644	MTLPO	679	NIORST
575	HLGKO	610	JFAIOQ	645	MTSPRO	680	NOKPRST
576	HLKPRO	611	JKLPST	646	MHDGKO	681	HLPON
577	HMTSPO	612	JEGLT	647	MHLPRO	682	HLPKON
578	HDGLKO	613	JEKLT	648	MLSRO	683	HGKON
579	HGKPO	614	JKGLT	649	MTSRPO	684	HLPRON
580	HLSPKO	615	JKPLST	650	MHLKPO	685	HDGKON
581	HLSRQ	616	JKPSLT	651	MLPRQ	686	HLKPON
582	HLPKOQ	617	JEKPST	652	MLPOQ	687	HMLKON
583	HGKOQ	618	JKPLMT	653	MTSPRQ	688	HLKOQN
584	HMTSRQ	619	JEGHMT	654	MTSROQ	689	HLPRQN
585	HLKPRQ	620	JKGHMT	655	MHLKOQ	690	HLSRON
586	HLPROQ	621	LPKOQN	656	MHLPRQ	691	HLSPO
587	HDGKOQ	622	LSPKON	657	MLSRQ	692	HMLPO
588	HLKPOQ	623	LKPRQN	658	MTSPOQ	693	HMTSRO
589	HMLKOQ	624	LPROQN	659	MLPKOQ	694	HGLKO
590	HLPSRQ	625	LPKOIN	660	MTLKOQ	695	HLGKO
591	IFJKPL	626	LPRQON	661	NOQRPS	696	HLKPRO
592	INOKPL	627	LPSRON	662	NQORPS	697	HMTSPO
593	IOKGL	628	LSPRON	663	NIORPS	698	HDGLKO
594	IORPKL	629	LKJFIN	664	NOKPLS	699	HGKPO
595	IFEGL	630	LPROIN	665	NIJKPS	700	HLSPKO

Appendix C: Transit Demand for the Example Network

Table B.1: Transit Demand for the Example Network (per hour)

OD	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
A	—	0	0	0	0	0	93	0	0	104	5	0	0	0	41	0	0	0	0	0
B	0	—	0	0	0	0	102	0	0	79	35	3	0	0	33	0	0	0	0	0
C	0	0	—	0	0	0	13	0	0	84	93	69	0	0	31	0	0	0	0	0
D	0	0	0	—	0	0	64	0	0	43	112	109	0	0	92	0	0	0	0	0
E	0	0	0	0	—	0	84	0	0	23	3	58	0	0	87	0	0	0	0	0
F	0	0	0	0	0	—	99	0	0	12	57	99	0	0	35	0	0	0	0	0
G	0	0	0	0	0	0	—	0	0	104	9	33	0	0	91	0	0	0	0	0
H	0	0	0	0	0	0	112	—	0	49	43	96	0	0	80	0	0	0	0	0
I	0	0	0	0	0	0	116	0	—	45	63	101	0	0	103	0	0	0	0	0
J	0	0	0	0	0	0	91	0	0	—	34	19	0	0	8	0	0	0	0	0
K	0	0	0	0	0	0	6	0	0	28	—	71	0	0	18	0	0	0	0	0
L	0	0	0	0	0	0	78	0	0	95	58	—	0	0	101	0	0	0	0	0
M	0	0	0	0	0	0	99	0	0	49	15	86	—	0	106	0	0	0	0	0
N	0	0	0	0	0	0	97	0	0	2	117	16	0	—	93	0	0	0	0	0
O	0	0	0	0	0	0	118	0	0	118	104	14	0	0	—	0	0	0	0	0
P	0	0	0	0	0	0	50	0	0	3	67	25	0	0	107	—	0	0	0	0
Q	0	0	0	0	0	0	73	0	0	41	33	83	0	0	70	0	—	0	0	0
R	0	0	0	0	0	0	9	0	0	46	4	77	0	0	37	0	0	—	0	0
S	0	0	0	0	0	0	60	0	0	73	100	106	0	0	35	0	0	0	—	0
T	0	0	0	0	0	0	117	0	0	54	63	80	0	0	80	0	0	0	0	—