Genetic Algorithm Approach for Transit Route Planning and Design

By Steven Chien,[1] Zhaowei Yang,[2] and Edwin Hou[3]

(Reviewed by the Urban Transportation Division)

Abstract: The problem of determining an optimal feeder bus route, feeding a major intermodal transfer station (or a central business district), in a service area is considered. Subject to geographic, capacity, and budget constraints, a total cost function, consisting of user and supplier costs, is developed for determining the optimal bus route location and its headway considering intersection delays, irregular grid street patterns, heterogeneous demand distributions, and realistically geographic variations. The criterion for the optimality is to minimize the total cost objective function. The number of feasible bus routes increases drastically with the increased number of the links (streets), and thus this problem is computationally intractable for realistic urban networks. This paper presents examples and demonstrates that the proposed genetic algorithm efficiently converges to the optimal solution, which is validated by the optimal solution obtained by applying an exhaustive search algorithm.

INTRODUCTION

One of the most important elements in planning a bus system is to determine bus route locations. Generally, bus routes are usually located on main direct-through streets. However, considering heterogeneous demand distributions in most service areas, such direct-through route locations may not be cost effective. Therefore, relocating bus routes and optimizing their headways should be considered to reduce the operating cost as well as improve the route accessibility. Transit operators and users may prefer short routes to reduce operating cost and in-vehicle time, respectively. Nevertheless, passengers also wish that a bus route traversed near their origins/destinations (e.g., a tortuous route) in order to reduce the access impedance. However, the tortuous route will increase the user in-vehicle time and bus operating cost. Transit agencies should consider this problem when either planning a new bus route or extending an existing bus route into a service region.

In the past 30 years, similar problems have been analyzed while considering many-to-one travel patterns by using analytical methods [i.e., Holroyd (1967), Byrne and Vuchic (1971), Chang and Schonfeld (1993), and Chien and Yang (2000)]. The many-to-one problem can be simply defined as how to select zones, line spacings, and headways to carry people from distributed origins to a single destination or vice versa. The analyzed bus networks in those studies were usually simplified as a set of parallel routes feeding a trunk line or a single point, such as a central business district (CBD) or a transportation terminal. They also assumed that bus routes are straight lines because of the assumed demand homogeneity and the convenience to formulate models.

An iterative algorithm was developed by Bryne and Vuchic (1971) to minimize the total cost and find the optimal bus route spacings and headways for a feeder bus system with parallel routes serving a linear CBD that was perpendicular to the bus routes. Holroyd (1967) analyzed a grid bus network with a given origin-destination demand that was uniformly distributed over an infinite plane. The optimal route spacing and headway were determined by minimizing the total cost. Chien and Schonfeld (1997) optimized a grid transit system in a heterogeneous urban area without oversimplifying the spatial and demand characteristics. The optimal route spacings, headways, and stop spacings were jointly optimized. Discrete demand distributions, which realistically represent the geographic variations in demand, were considered.

Each public transit mode has its own operating characteristics and thus provides a different level of service. Unlike paratransit systems, conventional transit operation is restricted by fixed routes and schedules to serve areas with substantial demand densities. Therefore, transit agencies face a problem of how to provide the most cost-effective service considering given geographical environments. A recent method for improving service of a fixed-route bus system is the out-of-direction travel technique developed by Welch et al. (1991), where the accessibility of a transit route was improved. Since the out-of-direction technique only increased the accessibility of certain segments, the total access cost for the whole service area was not minimized.

It is worth noting that bus route planning problems discussed in previous studies were solved analytically while simplifying route structure or demand distribution over space and time. However, the complexity of the problem increases significantly, while geographic (e.g., irregular service region and street pattern) and demographic (e.g., heterogeneous demand distribution) conditions are considered (Chien and Yang 2000). This precludes the use of analytical methods and other novel optimization techniques, such as genetic algorithms (GAs) should be considered. GAs are stochastic search algorithms that mimic the process of natural selection and genetics. It has been demonstrated to be an effective tool in solving various optimization problems (Hou et al. 1994).

Bus route selection and scheduling is a large combinatorial optimization problem. The solution space increases more than exponentially with the problem size (i.e., the density of streets and the complexity of the demand distribution). This paper presents two methods, a GA and an exhaustive search algorithm, to optimize a feeder bus route location and its operating headway while considering intersection delays and realistic street patterns. GAs are very efficient and effective in solving complex optimization problems (Gen and Cheng 1997). Recently, GAs have attracted research interest from different disciplines. The most important advantage is that GAs do not require specific structures in the objective functions and their performance for optimizing complex functions is quite prom-

[1]Asst. Prof., Dept. of Civ. and Envir. Engrg., New Jersey Inst. of Technol., Newark, NJ 07102.

[2]Transp. Engr., TransCore, 1250 Broadway, 34th Floor, New York, NY 10001.

[3]Assoc. Prof., Dept. of Electr. and Comp. Engrg., New Jersey Inst. of Technol., Newark, NJ 07102.

Note. Discussion open until November 1, 2001. To extend the closing date one month, a written request must be filed with the ASCE Manager of Journals. The manuscript for this paper was submitted for review and possible publication on December 16, 1998; revised November 7, 2000. This paper is part of the ***Journal of Transportation Engineering***, Vol. 127, No. 3, May/June, 2001. ©ASCE, ISSN 0733-947X/01/0003-0200–0207/$8.00 + $.50 per page. Paper No. 19883.

J. Transp. Eng. 2001.127:200-207.

ising. For a bus route selection problem, the measure of effectiveness (e.g., total cost) is unlikely to have a closed form. Hence, the GA approach for such a problem seems favorable.

MODEL FORMULATION

The primary purpose of this research is to optimize the location of a feeder bus route serving an irregular shaped suburban area with a generally irregular grid street pattern, as shown in Fig. 1, where the dotted and solid lines represent the zone boundaries and actual streets, respectively. The objective total cost, consisting of user and supplier costs, is dependent on the bus route location and its operating headway. Subject to geographic, capacity, and budget constraints, the objective of this study is to determine the best bus route and headway that minimize the total cost. The basic assumptions made for developing the cost model are as follows:

- The irregularly shaped service area (Fig. 1) can be divided into many rectangular zones according to the street spacing and demand distribution.
- A feeder bus route is proposed to provide service for connecting the area and a CBD or a major transfer terminal. Thus, the travel demand pattern of the area is either many-to-one or one-to-many in the morning or afternoon peak periods, respectively.
- A line-haul distance J, connecting a CBD (or a major transfer station) and the service area at an entry point, is assumed to be constant.
- The demand is inelastic with respect to service quality or price and is uniformly distributed within each zone, but differs among zones within the analyzed time period.
- Buses can stop anywhere along the route (e.g., demand stopping) whenever a boarding or an alighting is requested by a passenger. Thus, the location of bus stops can be neglected.
- The value of time is assumed to be additive. This assumption can be relaxed as long as the function of time value can be developed.
- The intersection (or node) delay, incurred by buses transversing intersections, is constant regardless the sizes of the buses, but may vary at different intersections.

In Fig. 1, zones differ in sizes due to the nature of the street pattern in the service area. To formulate a cost model for this network, the zones are divided into small ones that have identical width and length. Thus, the street network can be simplified, as shown in Fig. 1, where light dash lines represent dummy links that buses cannot traverse.

The street network is represented by an m-row by n-column matrix with $m \times n$ intersections or nodes. A link, connecting two neighboring nodes, is either a vertical or a horizontal way. Fig. 1 also shows the relationship between zero (i, j) and four

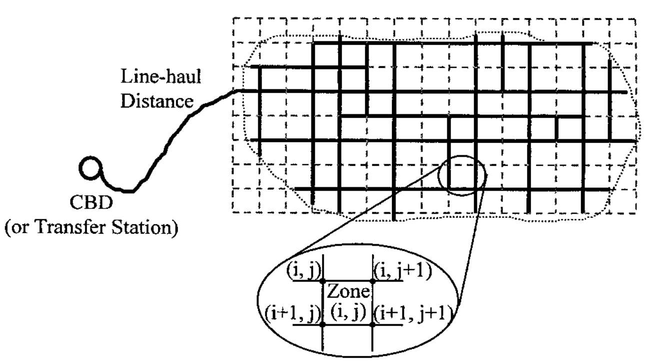

FIG. 1. Configuration of Service Area and Street Network

adjacent nodes (i, j), $(i + 1, j)$, $(i, j + 1)$, and $(i + 1, j + 1)$. Note that the coordinates of the upper-left node are used to identify the zone.

The buses are traveling from left to right (or vice versa) in the horizontal direction and up or down in the vertical direction. Any routes with branches and reversal are infeasible in this study. A feasible bus route is constituted by a sequence of links and nodes. To identify the links and notes of a bus route, the link and node incident matrices (e.g., $\mathbf{A}_{ij}^{Y}$, $\mathbf{A}_{ij}^{X}$, and $\mathbf{B}_{ij}$) are introduced and defined as follows:

$$\mathbf{A}_{ij}^{Y} = \begin{cases} 1, & \text{vertical link connecting nodes } (i, j) \\ & \quad \text{and } (i+1, j) \text{ is of the bus route} \\ 0, & \text{otherwise} \end{cases}$$

$$\mathbf{A}_{ij}^{X} = \begin{cases} 1, & \text{horizontal link connecting nodes } (i, j) \\ & \quad \text{and } (i, j+1) \text{ is of the bus route} \\ 0, & \text{otherwise} \end{cases}$$

$$\mathbf{B}_{ij} = \begin{cases} 1, & \text{node } (i, j) \text{ is of the bus route} \\ 0, & \text{otherwise} \end{cases}$$

The total cost objective function consists of supplier and user costs. The supplier cost C_S is a function of average vehicle round trip travel time and headway H_B. The bus travel time is the summation of round trip link travel times and node delays, if the layover time is neglected. Thus, the supplier cost C_S can be formulated as

$$C_S = \frac{2u_B}{H_B}\left(\sum_{i=1}^{m-1}\sum_{j=1}^{n}\frac{\mathbf{A}_{ij}^{Y}W}{V_B} + \sum_{i=1}^{m}\sum_{j=1}^{n-1}\frac{\mathbf{A}_{ij}^{X}W}{V_B} + \frac{L_J}{V_J} + \sum_{i=1}^{m}\sum_{j=1}^{n}B_{ij}T_{ij}\right) \tag{1}$$

where u_B, W, L_J, V_B, V_J, and T_{ij} represent the bus operating cost, the zone width and length, the line-haul distance, the average bus operating speeds in the service area and in the line-haul segment, and the average intersection delay incurred by buses passing through node (i, j), respectively.

The user cost C_U, constituted by the user access cost C_A, the user wait cost C_W, and the user in-vehicle cost C_V, is formulated in (2)

$$C_U = C_A + C_W + C_V \tag{2}$$

Although the derivations of cost functions C_A, C_W, and C_V were discussed by Chien and Yang (2000), a brief description regarding these cost functions are stated below. The access cost C_A, incurred by passengers walking to the bus route, is

$$C_A = \sum_{i=1}^{m-1}\sum_{j=1}^{n-1} a_{ij}u_A \tag{3}$$

where a_{ij} = total access time for passengers originating from zone (i, j) to the bus route; and u_A = value of user access time. The access time is determined based on the locations of zones to the bus route. Since the optimal bus route location is mainly determined by street pattern and demand distribution of the service area that have been discussed by Chien and Yang (2000), a_{ij} is difficult to formulate in one direction. Fourteen cases are categorized and shown in Fig. 2. The boundary lines for identifying the fractions of passengers accessing the bus route are required to be determined in Cases 3, 4, 5, 7, 8, and 9, while considering that buses are heading to the CBD. The total access times for all the cases are formulated and summarized in Table 1.

The user wait cost C_W is equal to average wait time multiplied by the total boarding demand and the value of user wait time u_W. Note that at small headways passengers do not follow schedules and their average wait time is half of the headway. Thus

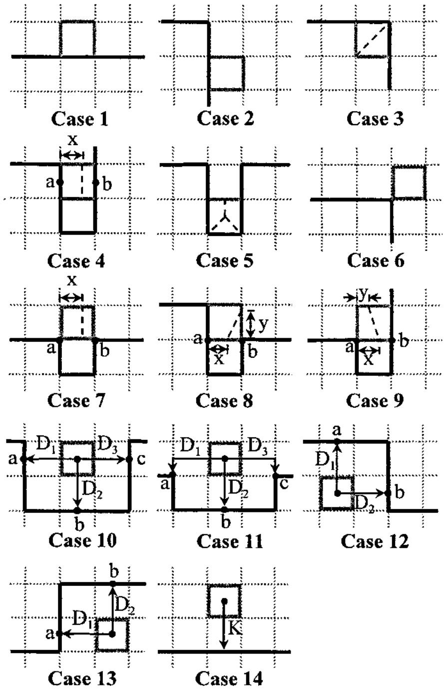

FIG. 2. Configurations of Zones to Bus Route

TABLE 1. Total Access Times for All Cases in Fig. 2

Case number	Total access time
1, 2	$\dfrac{q_{ij}W^3}{2g}$
3	$\dfrac{q_{ij}W^3}{3g}$
4	$\dfrac{q_{ij}W(W^2 - 2Wx + 2x^2)}{2g}$
5	$\dfrac{5q_{ij}W^3}{24g}$
6	$\dfrac{q_{ij}W^3}{g}$
7	$\dfrac{q_{ij}W(W^2 - Wx + x^2)}{g}$
8	$\dfrac{q_{ij}(3W^3 - W^2y + Wy^2 - Wxy - xy^2)}{6g}$
9	$\dfrac{q_{ij}(3W^2 + x^2 - 2Wx - Wy + xy + y^2)}{3g}$
10, 11	$q_{ij}W^2\left[\min\left(\dfrac{D_1}{g}, \dfrac{D_2}{g} + \dfrac{Z_1}{V_B}, \dfrac{D_3}{g} + \dfrac{Z_1 + Z_2}{V_B}\right)\right]$
12, 13	$q_{ij}W^2\left[\min\left(\dfrac{D_1}{g}, \dfrac{D_2}{g} + \dfrac{Z_1}{V_B}\right)\right]$
14	$\dfrac{Kq_{ij}W^2}{g}$

$$C_W = \frac{1}{2} H_B \sum_{i=1}^{m-1} \sum_{j=1}^{n-1} q_{ij}W^2u_W \quad (4)$$

The user in-vehicle cost is the product of the total in-vehicle time and the value of user in-vehicle time. The total in-vehicle time is the sum of the total link travel time, the total intersection delay, and the total line haul travel time. Thus, the user in-vehicle cost can be formulated as

$$C_V = u_I \left(\sum_{i=1}^{m-1} \sum_{j=1}^{n} \frac{Wf_{ij}^Y}{V_B} + \sum_{i=1}^{m} \sum_{j=1}^{n-1} \frac{Wf_{ij}^X}{V_B} + \sum_{i=1}^{m} \sum_{j=1}^{n} T_{ij} f_{ij}^B + \sum_{i=1}^{m-1} \sum_{j=1}^{n-1} \frac{L_J q_{ij} W^2}{V_J} \right) \quad (5)$$

where u_I = value of user in-vehicle time; and f_{ij}^X, f_{ij}^Y, and f_{ij}^B represent the passenger flows traversing links $[(i, j), (i, j + 1)]$ and $[(i, j), (i + 1, j)]$ and node (i, j), respectively. The passenger flows f_{ij}^X and f_{ij}^Y are equal to the average flow passing through two adjacent nodes $(f_{ij}^B + f_{i,j+1}^B)/2$ and $(f_{ij}^B + f_{i+1,j}^B)/2$, where f_{IJ}^B for any node (I, J) at different bus route configurations are shown in Fig. 3. Finally, the total user cost can be obtained by substituting (3)–(5) into (2).

As mentioned before, the total cost objective function C_T is the summation of the supplier cost C_S and the user cost C_U. According to the supplier and user costs derived from (1) and (5), the optimizable variables of the total cost function include $\mathbf{A}_{ij}^Y$, $\mathbf{A}_{ij}^X$, $\mathbf{B}_{ij}$, and H_B. Thus, the objective function can be formulated as

$$\min \quad C_T(\mathbf{A}_{ij}^Y, \mathbf{A}_{ij}^X, \mathbf{B}_{ij}, H_B) = C_S + C_U \quad (\mathbf{A}_{ij}^Y, \mathbf{A}_{ij}^X, \mathbf{B}_{ij}) \in \{0, 1\} \quad (6)$$

Although the total cost function can provide valuable insights into the relationship among the decision variables and exogenous parameters, capacity, and budget constraints [(7)] should be considered to make the model more realistic

$$H_m \le H_B \le H_M \quad (7)$$

where H_B is bounded by the minimum headway H_m to ensure that the supplier cost will not exceed the budget and by the maximum headway H_M to ensure that the service capacity satisfies the demand. Thus, H_M and H_m can be derived as

$$H_M = \frac{\rho}{\sum_{i=1}^{m-1} \sum_{j=1}^{n-1} q_{ij}W^2} \quad (8)$$

$$H_m = \frac{2u_B}{b} \left(\sum_{i=1}^{m-1} \sum_{j=1}^{n} \frac{\mathbf{A}_{ij}^Y W}{V_B} + \sum_{i=1}^{m} \sum_{j=1}^{n-1} \frac{\mathbf{A}_{ij}^X W}{V_B} + \frac{L_J}{V_J} + \sum_{i=1}^{m} \sum_{j=1}^{n} \mathbf{B}_{ij} T_{ij} \right) \quad (9)$$

where ρ represents the vehicle capacity; and b represents the budget to operate the transit route.

The optimal bus headway can be determined if the elements in matrices $\mathbf{A}_{ij}^Y$, $\mathbf{A}_{ij}^X$, and $\mathbf{B}_{ij}$ are treated as exogenous variables. The analytical solution for the optimal bus headway can be obtained by setting the first derivative of the total cost function with respect to the decision variable H_B equal to zero. Thus, the optimal bus headway H_B^* can thus be derived as

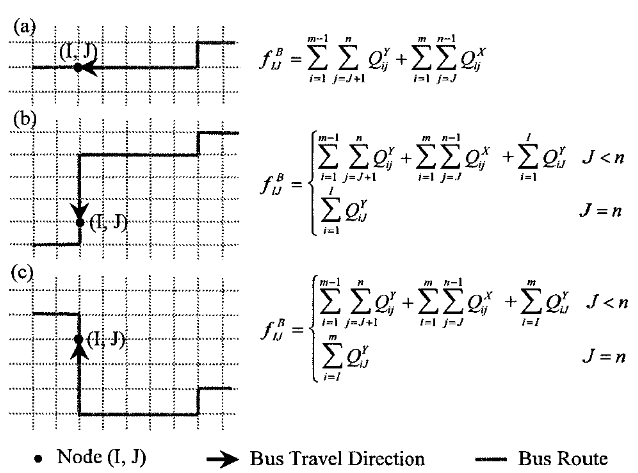

FIG. 3. Passenger Flows at Node (I, J) for Various Configurations

J. Transp. Eng. 2001.127:200-207.

$$H_B^* = \sqrt{\frac{4u_B\left(\sum_{i=1}^{m-1}\sum_{j=1}^{n}\frac{\mathbf{A}_{ij}^{Y}W}{V_B} + \sum_{i=1}^{m}\sum_{j=1}^{n-1}\frac{\mathbf{A}_{ij}^{X}W}{V_B} + \frac{L_J}{V_J} + \sum_{i=1}^{m}\sum_{j=1}^{n}\mathbf{B}_{ij}T_{ij}\right)}{\sum_{i=1}^{m-1}\sum_{j=1}^{n-1} q_{ij}W^2u_W}} \tag{10}$$

In (10), all the variables are nonnegative, and the second derivative of the total cost function with respect to H_B is positive. Thus, the objective function $C_T(\mathbf{A}_{ij}^Y, \mathbf{A}_{ij}^X, \mathbf{B}_{ij}, H_B)$ is convex, which implies that there is a unique optimal headway for any given matrices $\mathbf{A}_{ij}^Y$, $\mathbf{A}_{ij}^X$, $\mathbf{B}_{ij}$. Therefore, the minimum total cost for the bus route can be obtained by substituting the optimal headway, obtained from (10), into (6). If the optimal headway obtained from (10) violates either the capacity or budget constraint, the optimal headway can be directly obtained from (8) or (9), respectively.

EXHAUSTIVE SEARCH (ES) ALGORITHM

While the street network of the service area is represented by an $m \times n$ matrix, it consists of dummy and actual streets. Only the feasible routes composed by actual streets will be evaluated in the proposed ES algorithm for increasing computation efficiency. Thus, all infeasible routes will be screened out before applying ES to search the optimal solution.

The number of total cost functions (e.g., the number of feasible routes) is dependent on the configuration of the geographic (street pattern) condition and demographic (demand) distribution of the service area. The global minimum total cost can be found by comparing the minimum total costs for operating all feasible routes, and then the optimal bus route location and headway can be identified. A five-step ES procedure is designed for finding the optimal solution and stated as follows:

1. Initialization—Establish the link and node incident matrices $\mathbf{A}_{ij}^Y$, $\mathbf{A}_{ij}^X$, and $\mathbf{B}_{ij}$, respectively, based on the street network. All the dummy links cannot be traveled by buses and are given as the value 0 in the link incident matrices $\mathbf{A}_{ij}^Y$ and $\mathbf{A}_{ij}^X$.
2. Identification of a bus route—Select a feasible bus route that has not yet been evaluated, based on the incident matrices established in Step 1. The routes containing any dummy links will be automatically screened out here.
3. Solution of the optimal headway—Determine the optimal bus headway H_B^* of the identified bus route by using (10). If it violates the capacity or budget constraint, the optimal headway is obtained from (8) or (9).
4. Determination of the total cost—Substitute the optimal bus headway obtained from Step 3 and the associated link incident matrices into (6). The route location can be determined based on the node incident matrix $\mathbf{B}_{ij}$. If all feasible routes have been evaluated, go to Step 5; otherwise, go to Step 2.
5. Search the optimal solution—Search the minimum total cost C_T^* by comparing the solutions generated in Step 4, while the optimal route location can be also identified.

The developed ES procedure can be successfully applied to find the optimal feeder bus route on a given network because the optimal solution is identified after all feasible solutions have been found. However, if the given network is complicated in terms of the greater number of the streets, the number of feasible routes becomes huge. Therefore, the ES is computationally expensive to solve the route-planning problem.

GA

The proposed solution algorithm is characterized as a GA. GAs search for the optimal solutions by manipulating a set of feasible solutions, called a population of strings, using operations based on the natural evolution process. In each iteration, a new population is selected based on the current population using genetic operations that mimic the natural evolution process. The strings undergoing genetic operations are selected according to their performance. Better performing solutions have a higher probability of being selected, while poorly performing solutions have a smaller chance.

The transit route-planning problem discussed in this study is well suited to be solved by GAs because of the following reasons:

- The problem is similar to the functional optimization problem. The link and node occupancy matrices $\mathbf{A}_{ij}^Y$, $\mathbf{A}_{ij}^X$, and $\mathbf{B}_{ij}$, which represent bus route locations, are very similar to the binary strings that are the basic element of GAs.
- The route location is affected by the demand distribution as well as street pattern of the service area. By combining the segments of two different routes, the newly generated route may be more cost effective than its previous generation. According to this characteristic, an efficient crossover operator can be easily devised for the GA.

Evolutionary Process of Proposed GA

The proposed GA, developed in this section for searching the optimal bus route, includes two major operations: route generator and genetic operators. The GA starts with an initial population in which routes are randomly generated by the route generator based on the prespecified population size and the street pattern of the service area. Based on the objective value (total cost) of each route [obtained from (6)], all of the routes in the population are evaluated by the reproduction operator and the routes with lower costs are selected to form a new population. Thus, the quality of the new generation, in terms of cost, is always superior (cheaper) to the previous one. After reproduction, new routes generated by the crossover and mutation operators in the new population replace some old routes. This process will repeat until one of the following conditions are met: (1) the maximum number of iterations has been reached; (2) the current ''best'' solution is an acceptable solution; and (3) the current ''best'' solution cannot be improved over a large number of iterations.

Route Generator

As discussed in the ES, only feasible routes will be considered while searching for the optimal solution. An $m \times n$ street direction matrix (SDM) is established to eliminate the possibility of generating bus routes with dummy links. The SDM encodes the possible traveling directions (a combination of right, up, and down) from any node (i, j). For example, if the (right, up, down) travel direction at node (i, j) is (1, 1, 0) where the code [0, 1] represents that the node [has not, has] such a traveling direction, SDM (i, j) is equal to 6 ($1 \times 2^2 + 1 \times 2^1 + 0 \times 2^0 = 6$). Fig. 4 illustrates a 5-row $\times$ 7-column street network and the corresponding SDM.

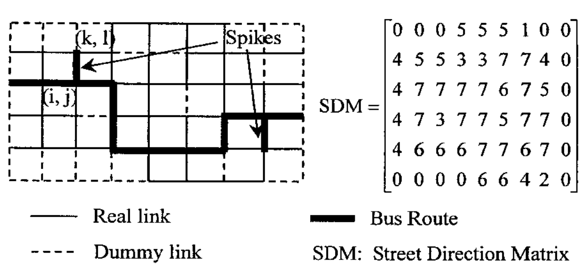

FIG. 4. ''Spikes'' on Generated Route and SDM Example

J. Transp. Eng. 2001.127:200-207.

Based on the SDM, a random feasible bus route can be generated. The steps are described below:

1. Selection of the starting node—Randomly select a valid starting node (i, j), where $j = 1$, while the traveling direction of the node is set to ''right.''
2. Determination of the next node—Determine the preceding node (k, l) according to the traveling direction from node (i, j). For example, if the traveling direction from node (i, j) is ''right,'' (k, l) can be $(i, j + 1)$.
3. Determination of the traveling direction—Determine the traveling direction from node (k, l) by randomly selecting a direction according to its corresponding SDM value. For example, if the corresponding SDM value of node (k, l) is 6, the traveling direction can be either ''up'' or ''right.'' Let $(i, j) = (k, l)$, and repeat Steps 2 and 3 until $j = n$.

A bus route generated from the above procedure guarantees that it will not have any dummy links. However, it is possible that ''spikes'' as shown in Fig. 4 may occur. These spikes can be easily identified by examining the nodes traversed along the route. A spike exists if the route has a sequence of nodes of the following form: . . . $(i, j) - (k, l) - (i, j)$. . . . This can be eliminated by simply removing nodes (k, l) and (i, j) from the sequence of nodes that represents the bus route. Based on the prespecified population size, a number of bus routes (initial population) are randomly generated by implementing the previous procedure.

Genetic Operators

The proposed GA consists of three genetic operators (reproduction, crossover, and mutation), which incorporate the ideas of survival of the fittest and genetic selection. These three operators form the search engine of the GA. First, the reproduction operator is used to select a new population (a small subset of feasible routes) based on the current population. The selection criteria are dependent on the notion of survival of the fittest, where the route with a lower cost will be selected into the new population. The procedure for implementing the reproduction operator is summarized below:

1. Identification of the route with the minimum cost—Find the route $R_{\min}$ by minimizing cost C_T in the current population. The costs of routes can be obtained by substituting their corresponding incident matrices and optimal headways, obtained from either (8), (9), or (10) (like Step 3 in ES), into (6). If all iterations are completed and the solution converges after this step, the bus route with the minimum cost, identified from the current population is the optimal solution.''
2. Calculation of cost differences—For each route in the current population, except for the one with the minimum cost found in Step 1, calculate the cost difference c_i between the pair of route R_i and $R_{\min}$ by using (11)

$$c_i = C_T(R_i) - C_T(R_{\min}) \quad \text{for } 1 \le i \le P \qquad (11)$$

where P represents the population size. Thus, the average cost difference c_a can be obtained from (12)

$$c_a = \frac{1}{P} \sum_{i=1}^{P} c_i \qquad (12)$$

3. Selection of appropriate routes to the new population—Any route R_i with $c_i \le c_a$ is selected into the new population. Certainly, the route with the minimum cost obtained from Step 1 is a member of the new population.
4. Selection of additional routes—If the number of routes selected by Step 3 is N, additional $(P\text{-}N)$ routes are randomly selected from the current population into the new population (routes in the current population may be reselected into the new population). Thus, the new population with size P is generated.

It is worth noting that the population size has to be chosen prior to operating the proposed reproduction procedure. Clearly, a larger population size P would lead to a larger number of routes being evaluated and thus increases the likelihood of finding the ''optimal'' solution. However, the computation (CPU) time needed for a single iteration is proposed to the P (P multiples average time required to process one route). Therefore, a balance between the P and computational burden must be found.

After the new population has been generated, the crossover operator can proceed. The function of the crossover operator is to generate new routes based on the routes in the existing population. The crossover operator follows the classical crossover operator used for binary strings (Ansari and Hou 1997), where the binary strings are divided into two parts based on an arbitrary selected crossover site and recombined to form two new strings. To perform the crossover operation on a pair of routes, the crossover site must be an identical node on both routes. Based on the sequence of nodes that form each of the routes, a crossover site can be identified, and new routes can be generated by swapping the remaining part of the routes, as shown in Fig. 5.

The crossover probability (the frequency of applying the crossover operations) controls the search for the optimal solution. Since new routes are generated using the crossover operator, a higher crossover probability would lead to more routes being examined and thus increases the chances of finding the optimal solution, but the CPU time also increases. The value of the crossover probability is highly problem dependent and can be determined through simulations. After conducting simulation analysis, it was found that the appropriate crossover probability in this study is between 0.8 and 0.9.

The mutation operator is used to create a random change in a route. If all the routes in the population are identical, new routes cannot be created using the crossover operator. The mutation operator allows the GA to escape from such a situation. Clearly, this can disrupt the routes generated by the crossover operator and should be applied infrequently. Like crossover probability, the mutation probability is problem dependent as well and can be determined with simulation analysis. It was found that the mutation probability that handing the frequency of mutation operation should be in the range of 0.05–0.1. Ways for determining parameters (population size, crossover probability, and mutation probability) has been discussed by

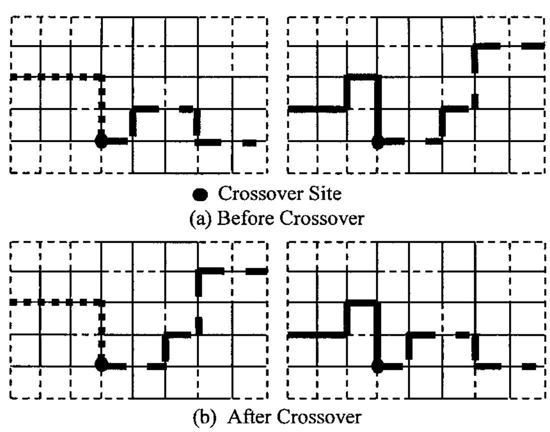

FIG. 5. Before and After Crossover Operation

Goldberg (1989), Gen and Cheng (1997), and Ansari and Hou (1997).

NUMERICAL ANALYSIS AND RESULTS

A service area, 4-km long and 10-km wide, is divided into 160 zones (based on given street pattern and demand distribution) whose sizes are identical (0.5 × 0.5 km). The line-haul distance, connecting the service area with a CBD, is 10-km long. The street pattern within the area is an irregular grid, as shown in Fig. 6. The average passenger access speed and bus operating speed in the service area and in the line-haul segment are 3, 20, and 80 km/h, respectively, while the average delays at intersections are identical and equal to 1 min. The values of user access, wait, and in-vehicle times are 10, 10, and 5 dollars/passenger-hour, respectively. The average bus operating cost is 50 dollars/vehicle-hour, the bus capacity is 50 passengers/vehicle, and the budget for operating the transit system is 900 dollars/h. The demand distribution over the area is summarized in Table 2.

The optimal bus route will be identified based on the minimum total cost. After searching all 489,888 feasible routes for the area by the ES, the optimal bus route location is found and shown as Route 1 in Fig. 6. On the other hand, after executing 40 GA iterations with a population size of 50, a crossover probability of 0.9, and a mutation probability of 0.1, the same optimal solution is achieved. Fig. 7 shows the optimal bus route location was found at the 12th iteration, while the total cost function was minimized when performing the GA. Since the optimal solutions found by the ES and the GA are identical, the optimal solution obtained by the GA is validated.

The optimal bus headway is 0.13 h, achieving the minimum total cost of 3,211.61 dollars/h. The resulting supplier cost is 27.1% of the total cost, and the user cost is 72.9% of the total cost including 7.7% for waiting cost, 28.5% for access cost, and 36.7% for in-vehicle cost. It is noted that the user in-vehicle and access costs make up more than 65% of the total cost, which dominate the optimal route location. Because the demand distribution over the whole service area is heterogeneous, the optimal bus route tends to balance the user access cost and the sum of the supplier and in-vehicle costs. Thus, it traverses high demand zones so as to increase the route accessibility.

To determine the effects of capacity and budget constraints to the optimal route location and minimum total cost, sensitivity analysis, considering various vehicle capacities from 38 to 115 passengers/vehicle, was conducted. No solution exists as the vehicle capacity is below 48 passengers/vehicle. The feasible region of optimal headway is bounded by budget constraint and capacity constraint as shown in Fig. 8. Route 1 achieves the minimum total cost if vehicle capacity is less than 71 passengers/vehicle; otherwise, Route 2 (Fig. 6) is the optimal one. In addition, as vehicle capacity exceeds 71 passengers/vehicle, the headway bounded by the budget constraint jumps from 0.125 to 0.134 h due to the increased route length by shifting service from Route 1 to Route 2. Generally, it was

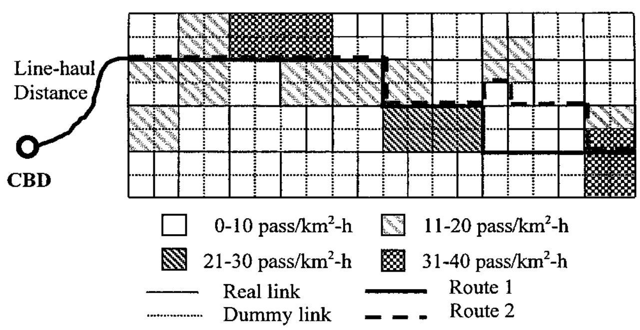

FIG. 6. Configuration of Service Area

TABLE 2. Boarding Demand Density q_{ij} (passengers/km^2-h) of Zone (i, j)

	i							
j	1	2	3	4	5	6	7	8
1	0	0	3	1	3	3	1	1
2	0	0	3	1	3	3	1	1
3	5	5	3	3	1	1	0	0
4	5	5	3	3	1	1	0	0
5	9	9	2	2	2	1	0	0
6	9	9	2	2	2	1	0	0
7	9	9	3	3	1	1	1	1
8	9	9	3	3	1	1	1	1
9	2	2	4	4	1	1	1	1
10	2	2	4	4	1	1	1	1
11	1	1	4	4	7	7	0	0
12	1	1	4	4	7	7	0	0
13	0	0	1	1	7	7	1	1
14	0	0	1	1	7	7	1	1
15	1	3	3	1	1	1	0	0
16	1	3	3	1	1	1	0	0
17	2	2	0	0	2	1	1	1
18	2	2	0	0	2	1	2	2
19	0	0	1	1	4	8	8	8
20	0	0	1	1	4	8	8	8

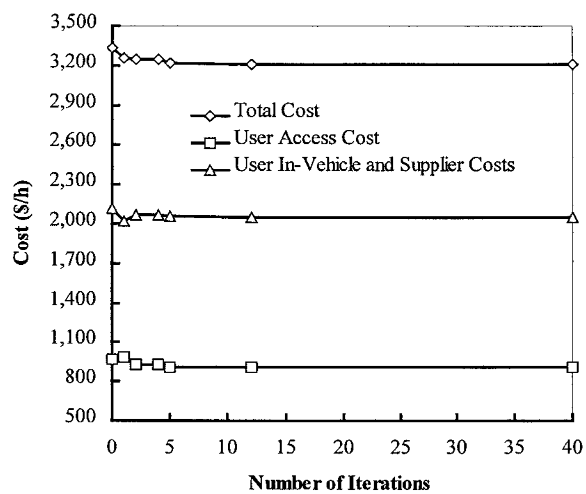

FIG. 7. Cost versus Number of Iterations in GA

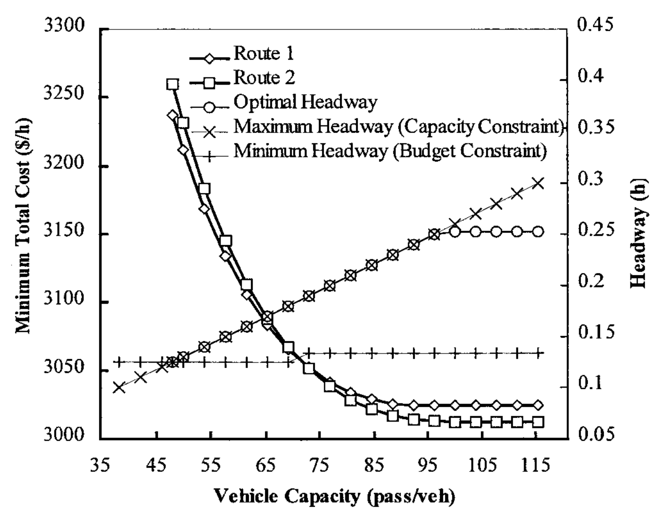

FIG. 8. Minimum Total Cost and Optimal Headway versus Vehicle Capacity Subject to Capacity and Budget Constraints

J. Transp. Eng. 2001.127:200-207.

TABLE 3. Experimental Networks and CPU Times (Network Size 9×21)

Network number	Total number of feasible routes	Percent dummy/total links	CPU time (min)	
			ES	GA
1	26,244	73.6	25.5	4.4
2	59,049	72.4	53.1	4.2
3	98,415	71.8	102.5	5.5
4	174,960	69.8	162.1	5.7
5	419,904	68.9	389.5	5.6
6	489,888	68.3	451.0	4.8

found that the minimum total cost decreases and the headway increases significantly, while vehicle capacity increases up to 98 passengers/vehicle. In the meantime, the optimal headway is dominated by the capacity constraint. While vehicle capacity exceeds 98 passengers/vehicle, the optimal headway is around 0.25 h.

It is worth noting that the values of the GA parameters, including the population size, the number of iterations, the crossover probability, and the mutation probability, have an enormous effect on the efficiency of finding a ''good'' solution and the quality of the solution. It was found that the optimal solution can be achieved when the parameters are properly determined. For example, if the selected population size is 40, only a near-optimal solution can be found.

In the previous analysis, it was also found that the difference between the CPU times consumed by the GA (excluding the simulation time for finding GA parameters) and the ES is huge (e.g., 4.8 min and 7.5 h, respectively). Thus, the benefit of using the GA in this application emerges. To validate the accuracy of results and explore the performance of the GA, additional experiments were conducted and discussed below.

Six networks, as shown in Table 3, with different street patterns were designed for the purposes of model validation and comparative analysis between the ES and the GA, while the demand distribution and the baseline values were kept the same as those in the previous numerical example. All GA parameters were also kept the same except that the population size in experiments 3–5 were set as 60, in order to find the optimal solution.

To fairly compare the computational efficiency between the ES and the GA, both algorithms are programmed in C computer language and performed on a Pentium 200-MHz computer, whose CPU times required for all experiments are summarized in Table 3. The CPU time for the ES increases proportionally as the number of feasible routes increases. However, the CPU time for the GA is independent of the number of feasible routes, but highly related to the selected population size. Obviously, the CPU time for the GA is significantly less than that for the ES, especially as the number of feasible routes is large.

CONCLUSIONS

Subject to geographic, capacity, and budget constraints, the objective total cost function including supplier and user costs has been minimized. The most cost-effective bus route and the corresponding headway, validated by comparing the results obtained from ES, are determined by using the GA, considering heterogeneous demand distribution in a given time period and realistic geographic environments (e.g., irregular grid street patterns). Since the service area can be divided into arbitrarily small zones, the proposed model can practically deal with demand variability. Developing a method to optimize transit route location considering time-varying demand, named temporally integrated optimization (Chang and Schonfeld 1993), would be a significant extension. With this extension, one may either schedule the timing for operating different optimal routes or optimize a fixed transit route while considering the aggregate effects of demand distributing over multiple time periods in a day.

The bus route selection problem in this study is a combinatorial type optimization problem. The ES is very expensive to solve the problem in terms of the CPU time. However, the GA offers a much more efficient way to solve this problem, especially for large and complicated networks, as shown in Table 3. It is worth noting that the proposed GA can be applied to optimize bus route locations for any conventional bus system (not limited to a feeder service) if the access and through-flow demands are formulated based on non-CBD-based travel patterns. A good geographic information system also could be valuable in supporting such analysis and optimization.

Occasionally, buses cannot always operate on their routes due to incident events such as accidents, road maintenance, or special events. Considering those service disruptions, original bus routes should be rerouted. A quick and adequate solution to such a dynamic bus route planning program can reduce transit operating cost, improve the transit level of service, and attract people using public transportation systems. A series of experiments designed in this paper have demonstrated the efficiency of the proposed GA for routing a bus system.

The total cost function developed in this paper may be iteratively used with a demand reestimation model to ensure that the feeder transit system is optimized for equilibrium flows. A model that analytically integrates the supply system optimization here with the demand approach [e.g., as in Kocur and Hendrickson (1982) or Chang and Schonfeld (1993)] would be a desirable extension.

ACKNOWLEDGMENTS

This work is supported by the National Center for Transportation and Industrial Productivity and New Jersey Institute of Technology, Newark, N.J.

REFERENCES

Ansari, N., and Hou, E. (1997). *Computational intelligence for optimization*, Kluwer Academic, Norwell, Mass.

Byrne, B., and Vuchic, V. (1971). ''Public transportation line positions and headways for minimum cost.'' *Traffic flow and transportation*, Elsevier, New York, 347–360.

Chang, S., and Schonfeld, P. (1993). ''Welfare maximization with financial constraints for bus transit systems.'' *Transp. Res. Rec. 1395*, Transportation Research Board, Washington, D.C., 48–57.

Chien, S., and Schonfeld, P. (1997). ''Optimization of grid transit system in heterogeneous urban environment.'' *J. Transp. Engrg.*, ASCE, 123(1), 28–35.

Chien, S., and Yang, Z. (2000). ''Optimal feeder bus routes on irregular street networks.'' *J. Advanced Transp.*, 34(2), 213–248.

Gen, M., and Cheng, R. (1997). *Genetic algorithms and engineering design*, Wiley, New York.

Goldberg, D. (1989). *Genetic algorithms in search, optimization and machine learning*, Addison-Wesley, Reading, Mass.

Holroyd, E. (1967). ''The optimal bus service: A theoretical model for a large uniform urban area.'' *Vehicular traffic science*, Elsevier, New York, 309–328.

Hou, E., Ansari, N., and Ren, H. (1994). ''A genetic algorithm for multiprocessor scheduling.'' *IEEE Trans. on Parallel and Distributed Sys.*, 5(2), 113–120.

Kocur, G., and Hendrickson, C. (1982). ''Design of local bus service with demand equilibrium.'' *Transp. Sci.*, 16, 149–170.

Welch, W., Chisholm, R., Schumacher, D., and Mundle, S. (1991). ''Methodology for evaluation out-of-direction bus route segments.'' *Transp. Res. Rec. 1308*, Transportation Research Board, Washington, D.C., 43–50.

NOTATION

The following symbols are used in this paper:

$\mathbf{A}_{ij}^{x}$ = horizontal link incident matrix, where $i = 1, \ldots, m$, $j = 1, \ldots, n - 1$;

J. Transp. Eng. 2001.127:200-207.

$\mathbf{A}_{ij}^{Y}$ = vertical link incident matrix, where $i = 1, \ldots, m - 1$, $j = 1, \ldots, n$;
a_{ij} = total access distance for zone (i, j), where $i = 1, \ldots, m - 1$, $j = 1, \ldots, n - 1$ (km);
$\mathbf{B}_{ij}$ = node incident matrix, where $i = 1, \ldots, m$, $j = 1, \ldots, n$;
b = budget for operating transit system (dollars/h);
C_A = total user access cost (dollars/h);
C_S = total supplier cost (dollars/h);
C_T = total cost (dollars/h);
C_U = total user cost (dollars/h);
C_V = total user in-vehicle cost (dollars/h);
C_W = total user wait cost (dollars/h);
c_a = average cost difference (dollars/h);
c_i = cost difference between pair of routes R_i and $R_{\min}$ (dollars/h);
D_1 = average passenger access distance from gravity point of zone to a (km);
D_2 = average passenger access distance from gravity point of zone to b (km);
D_3 = average passenger access distance from gravity point of zone to c (km);
f_{ij}^{B} = passenger flow passing through node (i, j) (passengers/h);
f_{ij}^{X} = passenger flow passing through horizontal link connecting nodes (i, j) and $(i, j + 1)$ (passengers/h);
f_{ij}^{Y} = passenger flow passing through vertical link connecting nodes (i, j) and $(i + 1, j)$ (passengers/h);
g = average passenger access speed (km/h);
H_B = bus headway (h);
H_M = maximum headway;
H_m = minimum headway;
K = average passenger access distance from gravity point of zone to bus route (km);
L_J = line-haul distance (km);
m = number of nodes per column;
N = number of routes in which costs are less than average cost in population;
n = number of nodes per row;
P = population size;
Q_{ij}^{X} = inflow demand on horizontal link connecting nodes (i, j) and $(i, j + 1)$ (passengers/h);
Q_{ij}^{Y} = inflow demand on vertical link connecting nodes (i, j) and $(i + 1, j)$ (passengers/h);
q_{ij} = boarding demand density in zone (i, j) (passengers/km^2-h);
$R_{\min}$ = feasible route with minimum cost in current population;
T_{ij} = bus delay at node (i, j), where $i = 1, \ldots, m$, $j = 1, \ldots, n$ (h);
u_A = value of passenger access time ($/passengers-h);
u_B = average bus operating cost ($/passengers-h);
u_I = value of passenger in-vehicle time ($/passengers-h);
u_W = value of passenger wait time ($/passengers-h);
V_B = average bus operating speed in service area (km/h);
V_J = average bus operating speed in line-haul segment (km/h);
W = width or length of each zone (km);
x = distance between boundary line and bus route (km);
y = distance between boundary line and bus route (km);
Z_1 = bus traveling distance from b to a (km);
Z_2 = bus traveling distance from c to b (km); and
ρ = vehicle capacity (passengers/vehicle).

J. Transp. Eng. 2001.127:200-207.