A bilevel model for optimizing station locations along a rail transit line

S. Samanta & M. K. Jha
Department of Civil Engineering, Morgan State University, Maryland, USA

Abstract

The population growth and increase in number of commuters in urban areas give rise to the need to either build or extend a rail transit system requiring: (i) optimization of the alignment of a rail transit line and (ii) optimization of station locations along rail transit lines. The positions of the stations depend on many factors, such as total cost for locating them, proximity from the residential neighborhoods, feasibility studies, and environmental, and political factors. In this paper a bilevel programming approach, which minimizes the total cost, is proposed to solve the station location problem. At the lower level, the potential ridership generated from the major cities or CBDs are estimated by dividing the study area in optimum number of zones, to maximize the usage by the potential riders. At the upper level, the number and location of intermediate stations are determined by minimizing the sum of user, operator, and location costs. The primary components of costs are the costs associated with traveling to the station and in-vehicle travel time, land-acquisition cost (also known as right-of-way cost), and cost of operating the train, construction of stations and parking facilities. The total cost is minimized using a Genetic Algorithm (GA). A Geographic Information System (GIS) is used in order to work directly with maps of the proposed rail-line, existing road networks and transit lines, and land and property boundaries. The population and passenger distribution in the study area as well as the travel times are obtained using a GIS which is integrated to the GA, to obtain the best station locations. The model is applied to decide on the optimum positions of stations along a transit corridor of an artificial case study. The results indicate that one can optimally locate station locations with improved precision if GIS data with sufficient accuracy were available.
Keywords: bilevel programming, rail transit station location, genetic algorithm.

1   Introduction

The population growth and the increase in the number of commuters in urban areas give rise to the need of extending the facilities of the existing transportation system. Transit system has been considered to be an important and efficient transportation mode for the mobility of mass population. It is gaining preference these days as it discourages the congestion and pollution, generally caused by automobiles. A rail transit system comprises of a transit line and stations along it. Once the transit line is aligned, the issue of locating the stations functionally and optimally along it arises. The positions of the stations depend on a number of factors, including the total cost for locating them, a feasibility study, and environmental and political factors. But, the most significant factor which influences the location of a station is the travel demand at the particular zone surrounding the station. In order to estimate the travel demand for a transit system potential rail transit riders have to be identified. The transit ridership or the travel demand generated at a particular station or zone is governed by the socioeconomic and demographic factors of that geographic area. The characteristics of the regional population play important roles while estimating the travel demand.

The research problem in the paper is to find the optimal number and locations of stations along a rail transit line provided the source, destination and the alignment of the transit line are given. The model is developed based on the potential ridership along the transit line. The objective function of a station location problem is developed as a cost minimization problem, which is a non-linear, and non-convex programming problem. The problem is formulated as a bilevel programming model. The concept of bilevel programming is applicable when two different stages of the problem share the common set of resources, but control different sets of decision variables to reach to the final solution. The station location problem is modeled such that the clustering of population is done homogeneously at the lower level, in order to obtain the optimal services from the stations, and at the upper level the cost is minimized, to optimize the number and location of stations. The population data can be obtained using GIS database. A Genetic Algorithm (GA) is used to solve the optimization problem at the upper level. The model is applied on a small artificial case study, where the optimal locations of stations are obtained after the total population along the transit line is clustered in two groups.

2   Literature review

The station location problem along the transit line has been approached by various researchers in different times. The problem of locating stations along transit rail lines in urban areas seems to be interesting and challenging to the researchers from the earlier days. Vuchic and Newell [1] studied the station location problem considering the linear transit line and uniform passenger distribution along it. They solved a set of partial differential equations to minimize the total cost. Wirasinghe [2] considered rail-station location as one of

the parameters to design nearly optimal parameters for a rail/feeder-bus system on a rectangular grid. The approach to solve the problem was to develop an analytical model to minimize the travel time and the operating costs by considering a rectangular grid, which is not applicable in practical life. Wirasinghe and Seneviratne [3] tried to optimize rail line length in an urban transportation corridor by minimizing total transportation costs. Formulation of the various cost components, e.g., user and operator costs were described vividly in their study. Spasovic and Schonfeld [4] developed an optimization model for transit service coverage using a penalty method. The stop spacing was obtained by considering a differentiable cost function. Kikuchi and Vuchic [5] developed a process which provides a theoretical basis for the selection for transit stopping policy and the number of stops. This study does not specify the spacings between them or the locations along a transit line. Horner and Grubesic [6] proposed a planning approach for locating rail terminals based on GIS data. It is mainly based on the demand parameter rather than the cost consideration. In another study by Lutin and Markowicz [7] the potential ridership was estimated considering the corridors based on the residential landuses. The demand was considered to be uniform. The cluster was developed based on the density and the proximity of the transit station in that study.

Bilevel Programming concept, which is fairly new, is applied on various transportation problems. Le et al. [8] formulated a bilevel model for the terminal location for a logistics problem. They designed the lower level model as per the behavior of the individual vehicle using the terminal minimizing its own cost and the upper level as the behavior of the planner for minimizing the total cost of the logistics system and the optimal size and location of the terminals. A GA is used to solve the problem optimally. Huang and Liu [9] proposed a bilevel approach for a logistics distribution network with balancing requirements. It promises to give a flexible network. The upper level is to minimize the total cost and the lower level is designed to obtain the balanced workload for the network. A GA is used to solve the problem. A bilevel program model is designed for toll optimization on a multicommodity transportation network [10]. At the upper level, the objective is set to maximize the revenue and the lower level is to minimize the total travel cost, i.e., to obtain the shortest path in the network. The model is applied on two examples. Murray-Tuite and Mahmassani [11] developed a bilevel formulation to determine the vulnerable links in a transportation network. At the lower level, the traffic assignment is done optimally and the vehicles are assigned accordingly. At the upper level, the disruption is maximized in the network based on the vulnerability index. The classical Traveling Salesman Problem (TSP) is also solved by bilevel programming [12]. The toll optimization problem is reduced to a TSP problem and is solved by relaxing the bounds. Hejazi et al. [13] and Yin [14] proposed a GA based solution tool for bilevel programming models. Mathur and Puri [15] studied the bilevel programming model for a bottleneck of non-convex problem. They studied the nature of the problem and proposed a formulation and solution methodology. By reviewing the literature the motivation to apply the bilevel

programming on the station location problem is drawn in order to incorporate the potential ridership in the model.

3   Model formulation

The bilevel programming model is designed as a two-stage optimization problem, where the optimal values of one set of variables obtained at one level are used to determine another set of variables at another level. The representation of a typical bilevel programming can be found in the literature [9]. The station location problem is formulated as a bilevel programming model as given below.

    The formulation of the objective function and constraints are shown in two different levels, such as lower and upper levels. The description of the formulation of the model is represented by fig. 1. The structure of the bilevel programming model is shown in fig. 2.
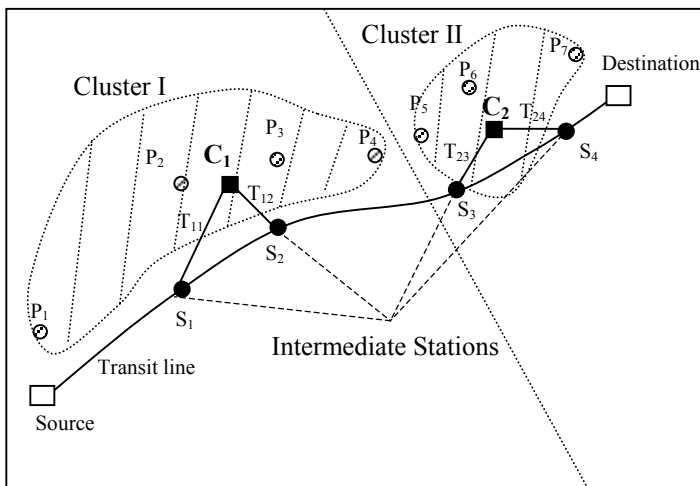

Figure 1:   An example showing the locations of intermediate stations, clustering of population and the centroids.

3.1 Lower level

At this level the aim is to cluster the population of the study area into a number of homogeneous subgroups using the concept of classification [16]. The centroids are obtained by clustering. The equivalent clustered population of each subgroup is supposed to act from the respective centroid. The travel time of the clustered population is calculated from the centroid of each subgroup to the stations which remain within the range of the particular cluster. The study area and the length of the rail transit line are divided into the equal sub areas or segments, respectively as that of the number of subgroups. The mapping is done to the centroid of a subgroup and one or more stations, which remain within the

respective segment of the transit line, respectively and the travel time is calculated from the centroid to the respective stations. The total population is partitioned in *K* exclusive, and collectively exhaustive subgroups. The homogeneity of the subgroups will depend on the variance of their population values. After creating the subgroups the equivalent population of the subgroup and the positions of the centroids of the subgroups are calculated by the centre-of-gravity method.
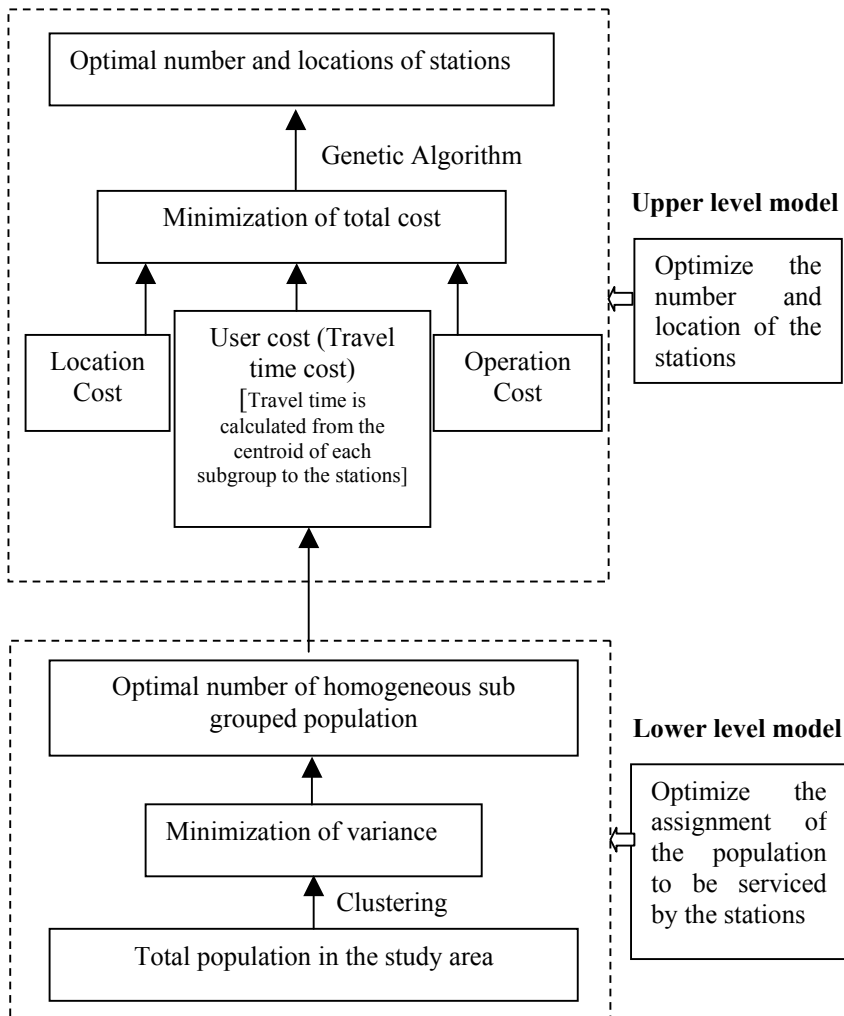

Figure 2:   Structure of the bilevel programming model.

Each group is represented by:

$$X_k : (G_{C_k}, D_{C_k}); k = 1,2,...K$$

where, G_{C_k} = *equivalent population of kth subgroup*

D_{C_k} = *position of the centroid of kth subgroup*

The objective function is expressed as

$$\text{Minimize } V_K \tag{1}$$

Subject to

$$K \le N \text{ ; if } D_i - D_{i-1} \ge \frac{L}{N} \tag{2}$$

$$K \le \frac{L}{N - J} \text{ ; otherwise,} \tag{3}$$

where, J = *number of times when* $D_i - D_{i-1} \le \dfrac{L}{N}$

where, V_K is the variance of the population values of the K subgroups given by,

$V_K = Var(\overline{G}_C)$

where, $\overline{G}_C : \{G_{C_k}\}; k = 1,2,....K$

3.2 Upper level

At upper level, the total cost is minimized based on the division of the study area in optimal number of subgroups and the travel time cost of the population of the study area. The objective function can be written as:

$$\text{Minimize } TC = UC + OC + LC \tag{4}$$

Subject to

$$\Delta S_i \ge \Delta S_{min} \tag{5}$$

where, UC = User cost = unit travel cost (UTC) $\times$ total travel time

OC = operator cost = unit operator cost (UOC) $\times$ vehicle travel time (t_v)

LC = Location cost (property value)

ΔS_i = Spacing between i^{th} and $(i+1)^{th}$ stations

ΔS_{min} = Minimum spacing

The above costs have to be formulated as a function of the decision variables which are the coordinates of the station points. Mathematically, the cost can be formulated as,

$$TC(D_1, D_2,.....,D_N) = \sum_{k=1}^{N} TC(D_k) = \sum_{k=1}^{N} \left[\sum_{i=1}^{K} UC(D_k, P_i) + OC(D_k) + LC(D_k) \right] \tag{6}$$

where, $D_k = k^{th}$ station

P_i = centroid of ith subgroup

Now the passenger access time (t_a) is calculated as it plays the most important role in the total cost calculation. For i^{th} subgroups and N^{th} station, the total access time of the passengers can be written as,

$$t_{a_i} = \sum_{j=1}^{M} p_i \times t_{ij} ; i = 1,2,3,....K \tag{7}$$

where, p_i = population of ith cluster

t_{ij} = travel time from the centroid P_i to j^{th} station

M = the number of stations which exist within the range of the ith cluster

The operator cost is a function of the distance traveled by the vehicle, and the location cost is composed of the cost of right-of-way acquisition.

4   Solution methodology

The optimization problem is solved using a GA based algorithm, which is modified from the previous work by Jha and Oluokun [17].

4.1 Steps of the genetic algorithm

Step I.   Generate   a   random   number   between 1 and N: $r_d(1,N) = k$; k is a positive integer. This is the number of intermediate stations between the source and destination.

Step II. Generate k random numbers $(r_{d1}, r_{d2},....,r_{dk}) = [\lambda_1, \lambda_2,.....,\lambda_k]$ in the interval $(\Delta S_{min}, L - \Delta S_{min})$, which represent the interstation spacings. The initial population consists of $\lambda_1, \lambda_2,....\lambda_n$ which represent the distances of the possible intermediate stations from the starting station.

Step   III. Calculate fitness (total cost function) of the population members.

Step IV. Apply mutation and crossover operators. Mutation operator is developed by randomly selecting a gene and replacing it by a randomly selected real number within the limiting values. The mutated chromosome becomes $[\lambda_1, \lambda_2,...\lambda_k^{'},...\lambda_n]$, where $\lambda_k^{'}$ is the replaced value (changed distance from the starting station).

Step   V. Develop the crossover operator by selecting two solutions from the population randomly. A part containing one or more than one station is identified from each of these two solutions and they are exchanged maintaining the feasibility conditions to produce two new solutions.

Step VI. Develop a selection-replacement scheme to ensure efficient convergence towards the global optimum.

Step   VII. Iterate for the specified number of generations from steps I-VI.

Step   VIII. Stop when searching through specified number of generations is complete or improvement in the objective function value is negligible (within 1%). Obtain the optimal station sequence and associated optimal cost.

Table 1:   Population of cities.

City	Population
P_1	10,000
P_2	5,000
P_3	25,000
P_4	10,000
P_5	18,000
P_6	16,000
P_7	13,000

Table 2:   Location cost for the stations along the transit line.

Distance from the source (ft)	Cost ($)
0	10,000
2	10,000
4	25,000
6	17,200
8	13,000
10	15,600
12	12,000
14	30,000
16	22,000
18	15,300
20	14,500
22	16,900
24	15,000

5   Numerical example

The model is applied in a small artificial case study. It is assumed that seven cities e.g., $P_1, P_2,..., P_7$ are present in the study area (Figure 3). For simplicity, the transit line is assumed to be linear and the coordinates of the source and destination are given. Based on the population (Table 1) and the location of the cities we cluster the total population in two mutually exhaustive subgroups. The clusters are obtained by minimizing the variance of the total population of the two subgroups at the lower level. We find out the centroids of these two subgroups, which are represented by C_1 consisting of the clusters of P_1, P_2, P_3 and P_4 and C_2 consisting of the clusters of P_5, P_6 and P_7. The populations of the two clusters are 50,000 and 47,000 respectively. The coordinates of C_1 and C_2 are given by (10.4, 18.5) and (18.9, 21) respectively. The total population of each cluster is assumed to act from the centroid and the positions of the centroids are shown in Figure 3. Thus, we obtain two homogeneous subgroups of population from which the travel time cost is calculated to the stations along the transit line. The stations are allocated to the subgroups for the service based on their locations and the range of the cluster under which they fall. The unit travel time cost and unit operator cost considered for the example are 1000 dollars/ft

and 2000 dollars/ft respectively. The location costs are given in Table 2. We calculate the travel time to the stations from the centroid of the cluster within each subgroup. We add the operation cost and location cost along with the calculated travel time cost to constitute the total cost. At the upper level we minimize the total cost using the Genetic Algorithm. We obtain the location of four stations S_1, S_2, S_3 and S_4 (Figure 3) which gives the minimum cost of $\$57.603 \times 10^5$.
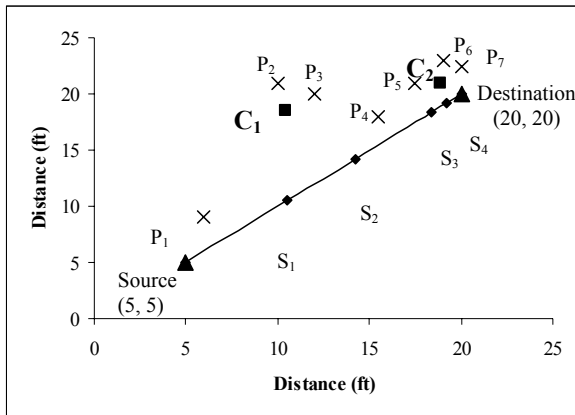

Figure 3:   Optimal locations of stations.

6   Conclusion and future scope

The bilevel programming model is applied on a small artificial case to determine the optimal number and locations of stations along a rail transit line. The clustering of the population is done at the lower level. The efficiency of the model depends on the accuracy of the clustering process. The intent of clustering is to distribute the population homogeneously to the stations for obtaining the optimal service from the stations. For the example studied, we obtain two clusters of population at the lower level and four stations along the transit line at the upper level as optimal solutions. It has been observed that as the clustering varies, it gives different solutions. As we increase the number of clusters, the variance decreases but the total cost increases. Additional research has to be carried out to explore other variables which can make the clustering process perform better and more efficiently, other than the variance. The parametric study has not been considered in the scope of the present study. The sensitivity analysis of the parameters has to be performed to make the model robust. The model has to be validated on a real life case as in future works.

References

[1]   Vuchic, V.R. & Newell, G.F., Rapid Transit Interstation Spacing for Minimum Travel Time. *Transportation Science*, **2(4)**, pp. 359-374, 1968.

[2]   Wirasinghe, S.C., Nearly Optimal Parameters for a Rail/Feeder-Bus System on a Rectangular Grid. *Transportation Research-A*, **14**, pp. 33-40, 1980.

[3]   Wirasinghe, S.C. & Seneviratne, P., Rail Line Length in an Urban Transportation Corridor. *Transportation Science*, **20(4)**, pp. 237-245, 1986.

[4]   Spasovic, L. & Schonfeld, P., Method for Optimizing Transit Service Coverage. *Transportation Research Record,* 1402, pp. 28-39, 1997.

[5]   Kikuchi, S. & Vuchic, V.R., Transit Vehicle Stopping Regimes and Spacings. *Transportation Science*, **16(3)**, pp. 311-331, 1982.

[6]   Horner, M. W. & Grubesic T.W., A GIS-based planning approach to locating urban rail terminals. *Transportation*, **28**, pp. 55-77, 2001.

[7]   Lutin, J. M., & Markowicz, B. P., Estimating the effects of residential joint-development policies on rail transit ridership. *Transportation Research Record*, 908, pp. 7-12, 1983.

[8]   Le, Y., L. Miao, Wang H. & Wang C., A Bilevel Programming model and a solution method for public logistics terminal planning. Transportation Research Board, Washington DC, 2006.

[9]   Huang B. & Liu N., Bilevel Programming Approach to Optimizing a Logistic Distribution Network with Balancing Requirements. *Transportation Research Record,* 1894, pp. 188-197, 2004.

[10]   Brtcorne L., Labbe M., Marcotte P. & Savard G., A Bilevel Model for Toll Optimization on a Multicommodity Transportation Network. *Transportation Science*, **35(4)**, pp. 345-358, 2001.

[11]   Murray-Tuite P.M. & Mahmassani H.S., Methodology for Determining Vulnerable Links in a Transportation Network. *Transportation Research Record,* 1882, pp. 88-96, 2004.

[12]   Marcotte P., Savard G. & Semet F., A Bilevel Programming Approach to the Travelling Salesman Problem. *Operation Research Letters,* **32**, pp. 240-248, 2004.

[13]   Hejazi S. R., Memariani A., Jahanshahloo G. & Sepehri M. M., Linear Bilevel Programming Solution by Genetic Algorithm, *Computers and Operation Research,* **29***,* pp. 1913-1925, 2002.

[14]   Yin Y., Genetic-Algorithms-Based approach for Bilevel Programming Models. *Journal of Transportation Engineering*, pp. 115-120, 2000.

[15]   Mathur K. & Puri M.C., Theory and methodology: A Bilevel Bottleneck Programming Problem. *European Journal of Operation Research*, **85**, pp. 337-344, 2005.

[16]   Ben-Akiva & M., Lerman S. R., *Discrete Choice Analysis*, The MIT Press, Cambridge, Massachusetts, London, England, 1985.

[17]   Jha, M.K. & Oluokun C., Optimizing Station Locations along Transit Rail Lines with Geographic Information Systems and Artificial Intelligence, in *Computers in Railways IX (COMPRAIL 2004),* J. Allan, C.A. Brebbia, R.J. Hill, G. Sciutto, S. Sone (eds.), WIT Press, Southampton, U.K., 2004.