

Ugur Cetin – Nurbanu Ince – Pelin Bayindir – Secil Savasaneril \*

## ROUTE SELECTION AND SCHEDULING FOR URBAN PUBLIC BUS TRANSPORTATION

*In Turkey, municipalities provide a major portion of the city public transportation service. Since the service has to be provided potentially to every resident, and at a low price, municipalities are often left struggling with negative balance. This implies that operating under low cost is essential for survivability. We conduct a study to analyze the public bus transportation system in Ankara which is the capital city of Turkey. The analysis reveals that there exist inefficiencies in the service due to fleet allocation, routing, dispatch frequency, and scheduling of the vehicles. Considering the daily and hourly transportation needs of the passengers, we determine routes, dispatch frequencies and schedules through a mixed integer programming model. The solution of the model is compared with the current practice under performance measures such as demand coverage, transit duration, and cost. Results show that there is potential for improvement.*

**Keywords:** Integer programming, public transportation, route selection, bus scheduling.

### 1. Introduction

This study discusses the decision support system designed by the authors for inner city public transportation for Ankara municipality (in Turkey) in 2010. In Turkey, the municipalities are responsible for certain services to the residents, one of which is the public transportation. The decisions taken by the municipalities regarding public transportation are at the strategic (intermodal transportation network design, determining the fleet size), tactical (allocation of vehicles to counties, constructing the routes) and operational levels (pick-ups and deliveries, maintenance). Public transportation service is provided in several modes, such as bus, metro, commuter rail, and private public transportation vehicles. As of 2009, 33% of all the passengers are transported by buses of the municipality. The city of Ankara is divided into five regions considering the geographical and demand characteristics (high or low private car ownership, inner-region demand, etc.) of the districts. Our focus in the study is on Region-5. In this region, transportation service is provided either through buses or commuter rail (with a single line). The scope of the study is redesigning the part of the public transportation service that is provided by buses in Region-5. The scope involves fleet allocation to departure points, designing the routes, and determining the departure times and frequency of the departures.

There is a large body of work in the literature that addresses the public transit network design and scheduling. [1] makes a comprehensive review and a classification of the studies. Accordingly, studies on public transportation system design address problems in (i) transit network design, (ii) frequency setting, and (iii) timetabling.

[2] also consider deadhead timing, and driver scheduling as components of the whole process. In the literature, the studies either address each of these problems independently, or in combinations. Ideally, a systemic view should be taken and all three problems must be addressed together, however the complexity and difficulty of the problem makes this impossible. Actually, even the individual problems are mostly NP-hard and are difficult to deal with. Furthermore, there are authors such as [3] that claim these problems must be addressed independently, since for instance, the network design is a strategic level decision whereas timetabling is an operational level decision. In our study, we jointly address the transit network design (i.e., node and route construction) and frequency setting problems. We do not address the timetabling or driver scheduling problem. In the literature the studies that address these two problems jointly, either take a mathematical approach by formulating the problem as a mixed integer problem or take heuristic approaches such as neighborhood search heuristics, simulated annealing or evolutionary algorithms. Objectives might vary from providing short travel times to passengers, to minimizing cost, or to demand fulfillment. See the references in [1]. In our study, we formulate the problem as a mixed integer program with the objective of maximizing net revenues.

In this study, firstly, the problem symptoms of the current system are analyzed and the causes of these symptoms are identified. The major problem symptom is the budget deficit, which indicates possibly low percentage of satisfied demand (which leads to low revenues) and possibly long routes (which leads to high transportation cost and low service levels). After the causes of the problems are identified, the necessary steps towards alleviating

\* Ugur Cetin, Nurbanu Ince, Pelin Bayindir, Secil Savasaneril

Department of Industrial Engineering, Middle East Technical University, Ankara, Turkey, E-mail: ssecil@metu.edu.tr

them are taken. These steps lead to redesign of the public bus transportation system in Region-5. The resulting improvements are evaluated through comparison of performance measures.

## 2. Analysis of the system

The region studied has certain characteristics. The location of the districts within the region is relatively far from the city center. Therefore the region has a “dormitory” type demand for transportation service, in that the demand generated within the region is quite small with respect to the demand generated outside the region (to the city center). Furthermore, the residents transport-out in the early morning hours and transport-in in the evening hours. Finally, due to long transportation distance to the city center the percentage of residents that prefer public transportation to private driving is high. The socio-economic profile of the residents hints that the cheapest transportation mode (which is bus transportation among the alternative modes) is highly preferred and the demand is elastic to price. This means, if the bus transportation service by municipality is timely and adequate there is a significant potential in terms of demand satisfaction.

Currently the buses operate between 6.00 – 23.00 (17 hours = 1020 minutes). There exist 5 departure points in Region-5 and departing buses arrive at 3 major destination points in the city center. The routes that span these departure and destination points cover 80% of the demand. Our study focuses on these routes. In the following, we present the problem symptoms in the current transportation system.

### 2.1 Capacity utilization

The capacity of the buses cannot be efficiently utilized due to characteristics of the demand during the day. There exist two peak-hours (in the morning and in the evening), during which the number of passengers exceeds the bus capacity. In the day time (when the demand is low), the utilization of the buses is low. Low utilization is also partly due to the requirement that certain departure frequencies must be ensured. For example, headways, i.e., the time between two consecutive trips should not exceed 30–40 minutes, or otherwise service would be regarded as very low. When capacity utilization of the buses for a certain route is examined, it is observed that the utilization rate may top 140% and might be as low as 50%. It is also seen that capacity usage highly fluctuates over time. The average number of overloaded trips in the morning peak hours (6am–9am) corresponds to 42.7% of total overloaded trips per day. Similarly, in the evening hours (4pm–7pm) 37.3% of the buses are overloaded. The correlation coefficient for “number of overloaded buses” in a given half-hour interval and “total number of passengers traveled” in that half-hour interval is calculated as 0.956. This is an indicator that the allocation of the buses to the routes over time is not responsive to fluctuations in demand.

### 2.2. Capacity allocation over the routes

Capacity usage not only fluctuates over time, but also over the routes (the lines). The analysis shows that if the number of passengers traveling on a route is high, then the number of overloaded buses on that route is also high. The densest routes are the ones that carry the passengers from the residential areas to the city center. Daily line frequencies for those routes might be well over 100. For the least dense lines, the daily frequencies are 1 or 2. The high number of overloaded buses in the densest lines is an indicator that capacity is not allocated to the routes appropriately.

### 2.3. Deviations from the ring duration

It is observed that on some of the lines, deviations (in minutes) occur from the scheduled ring durations. When these deviations are analyzed, the root causes are identified as the long route length (to the city center), excessive number of stops along the routes, and high demand the routes face. The analysis shows that all of these variables are highly correlated. When redesigning the routes, and frequency of the lines, the route length, number of stops, and the capacity of the line must be taken into consideration so as to eliminate the potential causes of the deviations.

### 2.4. Fleet utilization

We have also conducted qualitative analysis through interviews with the passengers, community, drivers and people responsible at the municipality. The general belief is that the fleet size is not sufficient to serve the area, and the main symptoms are indicated as high unsatisfied demand, and over-utilization of buses in peak-hours. However, when the number of passengers transported during the peak hours is contrasted with the total active capacity (on the average 90% of the fleet is active, while the 10% is under maintenance), it is seen that the utilization in the morning peak hours is around 85%. This indicates a need for better allocation of resources and proper bus schedules.

## 3. Solution approach

The problem symptoms indicated that there is a need for better allocation of resources. This problem involves three major decisions, constructing the routes, locating the bus stops, and determining the line frequencies, and dispatch times considering the demand and available fleet capacity. The objective is to maximize the total operating net revenue (revenue from tickets minus all operating cost). We first formulate the problem as a mathematical programming model. However, a mathematical programming model that addresses all three decisions would be large in size, and be non-linear. Thus, we decompose the problem into three (which lead to a heuristic approach): we first determined the bus stops, then generated the routes considering the demand and existing routes, and finally through a mixed integer programming model selected the least

costly routes to meet demand and determined the line frequencies.

### 3.1. Preparing the input to the model

When constructing the stops and the routes, the road network and the demand must be taken into consideration. The demand for the routes was not readily available as data. The reason is not all records on all buses and lines were kept properly. Consequently, we worked on constructing the demand nodes on the network. In the model, demand nodes correspond to potential bus stops. According to the specifications of public transportation, each bus stop must be within at most 400–500 m walking distance of the residential locations. Furthermore, the bus stops must be easily accessible by residents and located on the roads. If a certain area is highly populated, then several stops might be necessary.

After deciding on where the nodes (stops) should be, the demand is generated and assigned to the nodes. For this the following procedure is adopted. The current population of the districts within the region is available. Also how the demand is distributed within the region is available (to the detail level that shows the number of floors in the apartments and the locations). Considering the homogeneity or heterogeneity of the population within a district, the population is fractioned and assigned to the nodes that span the district. However, the population value is not used in the raw form as the demand for transportation service. There is a proportion of the population that makes trips in a day, and a certain portion of these uses public transportation. This provides us with a “trip generation coefficient”. These coefficients are available for origin-destination pairs within Region-5. However, the latest coefficient figures (which are obtained through household surveys) are due to 1992. Those figures are used by extrapolating to year 2010. The demand figures obtained are used for modeling week-day demand. The demand is disaggregated into hours using the sample values for the districts. Figure 1 shows how demand is disaggregated for one of the routes. Each value shows the proportion of demand that occurs in one-hour intervals. Note these proportion values add up to 1. Since demand in this region is “dormitory type”,

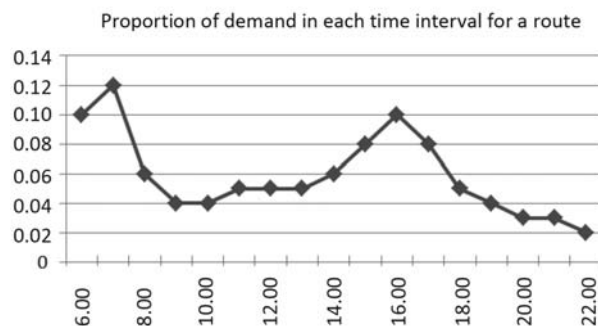


Fig. 1 The demand is disaggregated into one-hour time intervals. The hourly demand for the node-destination pairs (and vice versa) is inferred from the total demand for each of the routes

it is assumed that all demand is to the 3 destination points (in the city center), and demand to other destinations en-route is negligible. Furthermore, since the route is a round-trip and symmetric, and all passengers get off at the destination in the morning hours and get on at the destination in the evening hours, the two-way demand is aggregated to reflect the demand to the route. This is an approximation to simplify the mathematical model.

For constructing the routes, the departure points of the current transportation system were used as they are. The destination points are already few and well established. Therefore the routes are constructed considering the existing departure points and destinations. All routes are in the form of round-trips. Furthermore, the concerns of the municipality were taken into consideration. For instance, a route should be as straight and short as possible, in that the distance of a route is at most 120% of the distance that can be traveled over in same time period. A round trip of a route should be completed in at most two hours. In Fig. 2 we show part of a route that covers the major portion of the residential areas. The remaining part (which is not shown) follows a highway directly to the destinations.



Fig. 2 The route formed by the nodes 40, 42, 29, 30, 33, 36 and 60 (map taken from Google Earth and Basarsoft). In the map, the nodes 16–20–24–25–60 is on the major highway that leads to the destination points

### 3.2. The mathematical model

We formulate a mathematical model specifically designed for this problem. The mathematical model takes a set of candidate routes, the stops on these routes, the demand generated at these stops and the destination points as input. It gives the best routes, the corresponding line frequencies and dispatch times as outputs. We first introduce the notation for the mathematical model in Table 1.

$$\begin{aligned} \text{Max } (ef) \sum_{i,n,a} X_{ina} - (cng \sum_{i,t} 2c_i Y_{i1t} + cg \sum_{i,t,b \neq 1} c_i Y_{ibt}) - \\ - ((dr)s + cm) \sum_{o,b} F_{ob} \end{aligned} \quad (1)$$

Notation for the mixed integer programming model

Table 1

<b>Indices and sets</b>	
i:	candidate routes, $i \in I = \{1, 2, \dots, 78\}$
n:	Nodes (stops), $n \in N = \{1, 2, \dots, 67\}$
o:	origin (departure) nodes, $o \in O = \{1, 2, 3, 4, 5\}$
d:	destination nodes, $d \in D = \{1, 2, 3\}$
N(i):	set of nodes on route $i$
I(o):	set of routes originated from $o$ ,
I(d):	set of routes destined to $d$ ,
I(n):	set of routes that go through node $n$
a:	hourly time intervals, $a \in A = \{1, \dots, 17\}$ . $a = 1$ denotes the interval of 6am – 6:59am.
b:	types of buses, $b \in B = \{1, \dots, 9\}$ .
t:	discrete time points for a work day (in minutes), $t \in T = \{1, \dots, 1020\}$ .
T(a):	set of time points in time interval $a$
<b>Scalars</b>	
cng :	cost of gas per km
cg :	cost of gasoline per km
disc:	% of passengers using discounted fare
df:	fare for discounted pass
rf:	fare for regular pass
ef:	expected fare per demand: $(df)disc + rf(1 - disc)$
s:	daily salary
dr:	average number of drivers per bus
cm:	cost of maintenance per bus\
<b>Parameters</b>	
p <sub>da</sub> :	fraction of total daily demand for destination $d$ realized in time interval $a$
D <sub>dn</sub> :	total daily demand for destination $d$ at node $n$
D <sub>dna</sub> :	$(p_{da})D_{dn}$ , demand generated at node $n$ in time interval $a$ for destination $d$
K <sub>b</sub> :	capacity of type $b$ bus
N <sub>b</sub> :	the number of type $b$ buses in the fleet
N:	the number of all buses
c <sub>i</sub> :	length of route $i$ (one-way)
t <sub>i</sub> :	completion time of route $i$
<b>Decision variables</b>	
X <sub>ina</sub> :	passenger demand of node $n$ to be transported on route $i \in I(n)$ , in time interval $a$ , $X_{ina} \geq 0$ .
Y <sub>ibt</sub> :	1 if there is a departure of type $b$ bus at time $t$ for route $i$ , $Y_{ibt} \in \{0, 1\}$
A <sub>obt</sub> :	number of type $b$ buses available at origin $o$ , at the beginning of time $t$ , $A_{obt} \in \mathbb{Z}^+$
F <sub>ob</sub> :	the number of type $b$ buses assigned to origin $o$ , $F_{ob} \in \mathbb{Z}^+$

subject to

$$\sum_{i \in I(d) \cap I(n)} X_{ina} \leq D_{dna} \quad \forall n, d, a \quad (2)$$

$$\sum_{n \in N(i)} X_{ina} \leq \sum_b (K_b \sum_{t \in T(a)} Y_{ibt}) \quad \forall i, a \quad (3)$$

$$A_{ob(t+1)} = A_{obt} - \sum_{i \in I(o)} Y_{ibt} + \sum_{i \in I(o)} Y_{ib(t-comp(i))} \quad \forall o, b, t \quad (4)$$

$$A_{ob1} = F_{ob} \quad \forall o, b \quad (5)$$

$$\sum_{b \in B} Y_{ibt} \leq 1 \quad \forall i, t \quad (6)$$

$$\sum_{i \in I(o)} Y_{ibt} \leq A_{obt} \quad \forall o, b, t \quad (7)$$

$$\sum_o F_{ob} \leq N_b \quad \forall b \quad (8)$$

$$X_{ina} \geq 0, Y_{ibt} \in \{0, 1\}, A_{obt}, F_{ob} \in \mathbb{Z}^+ \quad (9)$$

In the model, the objective function in equation (1) is the sum of energy (gas or gasoline) cost, driver salaries, and the bus maintenance cost subtracted from the total revenue generated. The constraint in equation (2) ensures that demand satisfied by the selected routes and dispatch frequencies does not exceed the demand generated by each node for a certain destination at any given time interval, (3) ensures the satisfied demand does not exceed the capacity of a route in a time interval, (4) ensures the flow balance of the buses throughout the time intervals. Constraint (5) ensures available buses depart from the origin nodes at the beginning of the day. At any point a single bus departs from an origin node is ensured by (6). And remaining constraints ensure that available buses are allocated to the routes at any given point and total number of buses allocated does not exceed the fleet size. The objective function denotes the net profit obtained from revenues from the ticket and the cost of operating the fleet. The cost includes driver cost, maintenance cost, gasoline and natural gas consumption cost incurred in proportion to the length of the routes.

There are a total of 88,842 non-zero, 716,040 binary, and 45,945 integer variables. The number of constraints is 222,012. We used CPLEX 11.2.1 as the solver and used GAMS 23.0.2 as the modelling software. The optimal integer solution is found in 1190 secs.

#### 4. Results

The model gives outputs such as fleet size at each departure point, trip schedules, which type of buses to serve which routes,

total distance traveled, energy consumption figures, amount of unsatisfied demand, coverage ratio, and utilization rates. Comparison of those performance measures with the measures of the current system (in Table 2) indicates that there exist potential for improvement.

Comparison of the performance measures under the current and the proposed system

Table 2

Number of Routes	26	69
Number of Dispatches	1607	999
Number of Buses	398	164
Number of Passengers Transported	45123	97987
Average Length of Routes (km)	65.23	29.32
Number of Drivers	795	378
Revenue (Turkish Lira)	153104	159530
Fuel Costs (Turkish Lira)	73988	58913
Driver and Maintenance Costs (TL)	148362	26843
Profits without fixed costs (TL)	-69246	73473

We observe that the solution proposes an increase in the number of routes, whereas a decrease in the number of dispatches, compared to the current practice. Currently the routes are excessively long, and this results in poor performance, especially at the peak hours. By decreasing the length of the routes we aim to achieve an increased customer satisfaction (although we cannot measure it directly). Since the route lengths are lower, to provide adequate coverage the number of routes increases. The decrease in the number of dispatches is due to effectively allocating the dispatches to time intervals, and the routes to the nodes. This results in an increase in satisfied demand, and thus in the revenue generated, while a decrease in the fuel costs.

When we analyze the routes that are selected, we observe that 69 out of 78 potential routes are picked by the model. The eliminated routes are the longest ones. Actually, the longest routes could be the ones that potentially generate the highest revenues if there is sufficient demand. So if a high capacity bus can be assigned to those routes, the model does so. However, since the fleet size is limited, some of the routes have to be eliminated, and those are the ones that generate low demand per unit length. In the solution, mostly each area is covered by several lines (routes) where the lines differ only by their destination. Since the demand is of dormitory type, even if the destinations differ, the majority of the routes are the same in terms of the nodes covered. In several of the eliminated routes, an area is covered at least by one line. There is only one area which is not covered at all. In the model, placing a lower bound on demand satisfied for each node-destination pair might lead to a fairer service provision. Furthermore, the proposed solu-

tion gives a positive value as the objective function value. If the optimal solution yielded a negative value, the model would suggest that the transportation service shouldn't be provided. This is not a viable option for the municipality, and placing a lower bound on satisfied demand would avoid such complications as well.

## 5. Conclusions

In this study, we addressed the public bus transportation service problem faced by the Municipality of Ankara in Turkey. This project is conducted as a part of the senior level Systems Design course in Middle East Technical University, Department of Industrial Engineering in the 2009–2010 academic year. The problem symptoms indicated that there are inefficiencies in the allocation of the resources. To alleviate the inefficiencies, redesign of the public bus transportation service were undertaken. The data requirements of the project were overwhelming, while the availability of data was very low. Due to limitations on data, the scope of the problem were narrowed, some simplifying assumptions and approximations were made. The problem is stated as the selection of routes and deciding on the dispatch frequencies at each line as well as the dispatch times, with the objective of maximizing net revenues. A mixed integer programming model is formulated to address the problem. The results of the model are contrasted with the current practice based on some key performance measures.

The results show that taking a systemic approach through operations research has potential to lead to a better system. However, there is still much work to do to improve the whole transportation system in Ankara. First of all, there should be a systematic approach to keep track of demand information for each origin, node and destination. Currently, it is possible to only record the time to get on the bus, and for the route that the bus serves. Obviously, there is a need to that the transportation system provides much richer information. The passengers can be informed dynamically through smart bus stops, and/or smart phones on the whereabouts of the vehicles. Through vehicle tracking system the time spent on each stop and on the route could be measured and even the traffic signals can be designed accordingly. Those steps are slowly taken by the municipality nowadays.

Secondly, the capacity of the buses must be more flexible. Increasing the number of buses is not a good alternative due to the rigid and large capacity of the buses. Thirdly, the proportion of public transportation is around 60%, and 33% of that is due to bus transportation. This ratio is low for a large metropolis like Ankara. Low public transportation service utilization results in increased traffic problems, which in turn affects the quality of public transportation. To improve the overall transportation system in the city, the attraction points must be planned carefully considering the potential demand points and the accessibility. The underground metro system must be constructed to span a larger portion of the city, and the current underground usage ratio of 5% must be increased.



## References

- [1] GUIHAIRE, V., HAO, J. K.: Transit Network Design and Scheduling: A Global Review, *Transportation Research: Part A*, 42, 2008, pp. 1251–1273
- [2] CEDER, A., WILSON, N. H. M.: Bus Network Design, *Transportation Research: Part B*, 20, pp. 331–344, 1986.
- [3] YU, B., YANG, Z., CHENG, C., LIU, C.: *Optimizing Bus Transit Network with Parallel Ant Colony Algorithm*, Proc. of the Eastern Asia Society for Transportation Studies, vol. 5, pp. 374–389.