Creating Bus Timetables Under Stochastic Demand

Afshin Shariat Mohaymany* & S.M. Mahdi Amiripour

Afshin Shariat Mohaymany, is with the School of Civil Engineering, Iran University of Science and Technology. S. M. Mahdi Amiripour, is a PhD student at the same Department.

KEYWORDS
Timetable, Urban bus network, Stochastic demand, Variable demand, Simulation

ABSTRACT
Local bus network is the most popular transit mode and the only available transit mode in the majority of cities of the world. Increasing the utility of this mode which increases its share from urban trips is an important goal for city planners. Timetable setting as the second component of bus network design problem (network route design; timetable setting; vehicle assignment; crew assignment) have a great impact on total travel time of transit passengers. The total travel time would effect on transit utility and transit share of urban trips. One of the most important issues in timetable setting is the temporal coverage of service during the day. The coverage of demand is an objective for setting timetables which has not been well studied in the literature. In this paper a model is developed in order to maximize the temporal coverage of bus network. The model considers demand variation during the day as well as the stochastic nature of demand. A distribution function is used instead of a deterministic value for demand. The model is then implemented to an imaginary case.

1. Introduction
One of the most important components of urban bus network design is setting bus timetables or timetable setting. This part has not been well studied in the literature. The main objective of setting timetables for bus network is to have a punctual service. This punctuality would mainly benefit the passengers of the service which were not so important for transit agencies till now. The main goal for transit agencies was always been decreasing the costs of the system so they made a lot of effort on vehicle and crew assignment. But today, as the demand increases the utility of transit system became so important. A well timetable results lower travel time in network and better temporal coverage. Subsequently the utility of bus system and its share from urban trips increases. As a matter of fact timetable setting can play an important role in increasing transit system utilization. Besides, a good timetable can increase the reliability of urban transit system as well.

The history of setting bus timetables returns to 1967 when Lampkin and Saalmans [8] formulate bus network design and setting service frequencies for routes using mathematical programming techniques. After that some researches have been done on this subject. Ceder (1984) [1] developed four methods for setting frequency on routes using passenger count data. He used maximum load station passengers and route load profile. These frequency determination methods became the basis for a timetable setting method, later developed by Ceder (1986) [2]. In this method the timetable was developed in three steps, making alternative timetable options; comparing timetable according to utilization measures; and smoothing procedures for bus departure times for the case of evenly spaced headways. Voss (1992) [10] formulated the problem of network design in schedule synchronization, minimizing the waiting time of passengers at the transfer nodes. His study refers to the cases where each bus route is jointed by a set of possible departure times. The problem was modified to a case where different routes partly use the same tracks, implying that security instances must be observed. Desilet and Rousseau (1992) [5] described a model, which selects a starting time for each route from a set of possible starting times. Their objective was minimizing the total penalty associated with transfers from one line to another for all the lines. The penalty function, which could be calculated in various...
ways, takes into account the random nature of traveling times. De Palma and Lindsey (2001) [6] analyzed the optimal timetable for a given number of public transport vehicles on a single transit line. The riders of these lines differ with respect to the times at which they prefer to travel and the schedule delay costs they incur from traveling earlier or later than desired.

As clearly can be seen, the waiting time have a great impact on transit utility so most of the researchers used waiting time as their objective for setting a timetable. These researchers aim at minimizing waiting time at transfer stations. Ceder et al. (2001) [3] based their timetable setting model on maximizing simultaneous arrivals of two buses to a transfer node at the same time in a network in order to minimize the waiting time and consequently the total travel time. They increase synchronization of bus timetables. They formulated the problem as a mixed integer linear programming problem and solved it by a heuristic algorithm. Eranki (2004) [7] proposed a model succeeding the Ceder’s model, which considers a time interval for deciding if two runs are at the same time or not. The idea proposed by Ceder used in some other researches like Broomely and Currie (2005) [4] and Quak (2003) [9] as well. These models consider waiting time just at transfer stations.

These researches considered the demand as a constant deterministic parameter. They did not take into account the stochastic nature of demand and variability of it during a day. Yan et al. [11] (2006) studied the problem of routing and timetable setting of inter-city buses under stochastic demand. Although this study aims at inter-city bus lines but it can be referred as an early study in the field for considering the stochastic nature of travel demand.

The aim of this paper is to take into account the stochastic nature of travel demand as well as its variability during a day. This paper develops a model to decrease the waiting time with respect to demand variation during day, considering waiting time at all stations (not just transfer stations). This consideration can lead us to have a well temporal coverage network. Besides, the stochastic nature of transit demand would guarantee an amount of passengers during day. Also would cause to avoid overutilization or underutilization in transit system. The model proposed in this paper is also considers the fleet size constraint that would makes it more applicable to urban transit systems.

The paper is organized as follows:

On part 2 after this introduction the model is defined and the formulation of objective is presented. Part 3 introduces the solution algorithm framework which would be expanded in parts 4 through 7. Part 4 defines the first step in solution algorithm which is creating timetable alternatives. Part 5 which is the second step of solution algorithm assigns passengers to bus routes by using a shortest path method. After creating alternative timetables and assignment, in part 6 the simulation process which calculates utilization measures is presented. Part 7 presents a decision making technique for choosing the best timetable from feasible solutions with respect to system constraints. On part 8 the model is implemented to a sample network and on part 9 conclusions and propositions for future works are presented.

2. Formulating Model

As previously mentioned, the main objective in this model is reducing waiting time, which originates from demand variation and dispatch times of bus runs in the network. Consider G(A, v), where A is the network of arcs (arterials) in the network and v is the number of nodes (bus stations). On this network:

M: Total number of bus routes in the network
T: Total Interval of the Design
Fk: Number of dispatches in interval T in route k
Nmaxk: Number of stops in route k
Wnk: Existing passengers on stop n on route k in time t

So the objective of the model can be formulated as Eq. 1:

$$\text{Min} \sum_{k=1}^{F} \sum_{r=1}^{T} \sum_{n=1}^{N_{\text{max}}} W_{nk}$$

Also we have a fleet size constraint which reflects the budget limitations. This constraint would be discussed in part four.

For simplicity of developing the model these assumptions are considered in the paper:

1- Travel time on network arcs are constant and are a multiple of simulation step.
2- Passengers choose the shortest route as their choice, considering penalty for transfer. Assignment method is all-or-nothing method.
3- Demand has a normal distribution function.
4- Origin-destination matrices would be normal in intervals and would have uniform distribution within intervals.
5- The capacity of all buses is the same.
6- Passengers won’t change their mode of transport as waiting time increases.
7- Demand would not change by changing the timetable.

This formulation and assumptions are only reflecting the big picture of the model. Other assumptions and constraints that have been used in the simulation process are obvious and not mentioned here (e.g. buses cannot board more than their capacity).

3. Solution Algorithm

A four step process (Fig. 1) is defined as the solution algorithm for the model. At the first step after collecting required input data, according to the system policies some alternative timetables would be created. These alternatives would be used as the basis of
simulation. The second step is transit assignment and can be run parallel to the first step. Transit assignment has two separate levels. The first level is finding the shortest path between every two stations in the network, and second level is assigning passengers to the routes (which would be done in step 3, simulation). At step 3 the process of bus runs through network in design interval T and the process of boarding and alighting passengers are simulated in order to calculate the utilization measures for every alternative timetable. Step 4 defines a methodology for choosing the best alternative timetable according to system policies and constraints.

Fig. 1. The solution algorithm for the model

Parts 4 to 7 define these steps completely.

4. Timetable Alternatives

The model proposed in this paper is a simulation based model. This model tests alternatives by simulating the processes of dispatching buses, boarding and alighting passengers and finally calculating utilization measures. In such model the first step is to create some alternatives. All these alternatives would be a potential timetable for the network. The feasible solution set in timetable setting problem is so big and finding the best timetable among all feasible solutions is practically impossible. Consequently the feasible solutions should be delimited and some of nonoptimum ones should be wiped out. Three constraints have been defined in order to shorten the feasible solutions:

1-Splitting the design interval T into three types and considering different minimum and maximum headways for each type of time interval.

2-Ranking existing routes into three types according to their importance in the network and considering different minimum and maximum headways for each rank.

3-Considering different steps in alternative making process for each type of time interval and route rank.

The first constraint wipes out those timetable alternatives which have a low frequency for crowded time intervals or a high frequency for uncrowded time intervals. The second one wipes out those alternatives with high frequency for unimportant routes and vice versa. As a matter of fact these two constraints wipe out alternatives which we are sure about their optimality and we are sure there are alternatives much more appropriate than them in feasible solution set.

The third constraint limits alternatives for choosing a number as a time interval headway for a specified route, e.g. if the minimum headway for a time interval and a specified route be 5 and the maximum headway be 15, then for a step equals to 1, headway for this interval can be chose from set {5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15} and for a step equals to 2 headway should be chosen from set {5, 7, 9, 11, 13, 15}. This constraint may wipe out some optimal alternatives but those wiped out alternatives were not much more different than the ones stayed in feasible solution set.

These three constraints help to reduce the size of feasible solution set and subsequently increase the applicability of the model to a large scale network.

5. Assignment

Passenger assignment to routes is based on origin-destination matrix of transit. This process has two different levels. At the first level, the shortest paths between every two stations are found, using Floyd-Warshall algorithm in the network. The shortest path method takes into account the penalty for transferring from one route to another. The first level would be done before the simulation process. At the second level, the passengers are assigned to the routes during the simulation process by an all or nothing method to the shortest path between stations. The input data for assignment would be the origin-destination matrix between every two stations during the day. The output of assignment would be the shortest path between every two stations in the network which would be used in simulation process.

6. Simulation Process

For calculating utilization measure for each timetable alternative the process of dispatching buses, boarding and alighting passengers has been simulated. A C++ program is developed to implement the process. Two important issues during simulation are discussed:

6.1. Stochastic Demand

During the simulation, at each station passengers are arriving according to the time interval. The amount of passengers in a station has a distribution function instead of a deterministic value. The distribution function reflects the stochastic nature of demand at the stations, e.g. a normal distribution, showed by \(N(\mu, \sigma)\) in this paper. Considering a distribution function instead of a deterministic value for demand may cause some difficulties in adding and reducing demand. For example at station 1 there are \(N(5, 1)\) passengers and in station 2, \(N(6, 1.5)\). If a bus boards all the passengers at these two stations then how many passengers are exist in the bus. For calculating this amount, a Monte Carlo simulation is applied. To do so we take some random variables from the first distribution function and add it.
up with random variables from the other distribution function, if this process continues we would have a distribution function for the amount of passengers exist on the bus. Getting random numbers from a distribution function is a major part of Monte Carlo simulation. There are some simple ways to get random numbers from a normal distribution function but they would not applicable to other distribution functions. This means that the procedure for getting random numbers from distribution functions depends on the types of these functions. For normal distribution functions the procedure would be an easy and fast one. But for another functions such as lognormal that is a possible function for the demand the procedure would be time consuming and cumbersome. Using Monte Carlo simulation technique is so time consuming especially when the number of distribution functions increases, and the problem defined in this paper is such. So according to the fact that two distributions would added up so many times during the simulation process a precomputation technique is used to reduce the time of simulation. In this technique all the required additions of distribution functions are first calculated and stored in a file and every time we need that adding instead of computing, the precomputed adding result would be used.

![Flow chart for simulation process](image)

**6.2. Simulation framework**

All elements of the bus system have a predefined activity which should be done at the mean time. The basics for the simulation process are as follows:

**Simulation step:** On simulation process, at the beginning of the planning interval $T$, ($t=t_0$) all stations are empty and all the buses are at their depots. Activities for all elements are defined. When $t=t_1$, passengers come to stations and buses start their trip from the first station. This process continues until $t=T$ which is the end of the simulation. On all simulations we call $t_n$, simulation step, which shows how much and how the activities of elements should go through the time. A low simulation step would cause a higher precision for the simulation process but it increases the simulation time so much. The simulation step can be
found out during a trade off between precision and simulation time.

**Passengers, Stations and buses structure:** The origin-destination matrix with the precision required in this model can not be collected in small intervals (e.g. simulation step) but buses can run in this step. Thus, in order to simulate this process we should break the origin-destination matrix into smaller intervals. Every interval from the origin destination matrix is break into intervals equal to simulation step and we call the passengers in those smaller intervals a “bunch”. Arrivals of passengers to the stations always would be in bunches, so the stations are filled with different bunches of passengers who came at different times and have different destinations. Buses also have the same structure as stations. They board and alight bunches at stations, in other words they pick or put bunches at stations.

**Utilization Measures:** During the simulation process utilization measures which would be used as the criteria for comparing alternatives are calculated. So many measures can be calculated during simulation such as waiting time (which is the model objective), passenger-kilometer of travel, passengers who have not been transported, load factor of buses, waiting time for transfer and serviced passengers. At the first step \((t=t_0)\) waiting time equals to the number of existing passengers on all stations, is saved. This would be zero at \(t=t_0\). At the second step passenger bunches are added to the stations and on step three buses are dispatched from their start interval \((T)\). At the end of simulation, utilization measures are calculated from the data collected during simulation.

**7. Choosing Appropriate Timetable**

After simulation process and calculating utilization measures it is time to choose the best timetable alternative according to the criteria and system constraints.

As mentioned earlier, there are two criteria for choosing appropriate timetable from feasible solutions which are the total waiting time that has been calculated from the simulation process and fleet size that can be easily excavated from the timetable itself.

The total waiting time has distribution function according to the stochastic nature of demand which was considered in the model and it has been shown by \(N(\mu, \sigma)\) (if demand has a normal distribution). Consequently there are three criteria for decision making:

1- Total waiting time mean value
2- Total waiting time standard deviation value
3- Fleet size

The best timetable is chosen from a three level decision making process. The first level considers system policies on the maximum waiting time and fleet size allowed. As shown in Fig. 3-a by adding these two constraints the feasible solutions are getting smaller and so many alternatives are wiped out because of the system policies.

At the second level some of the alternatives are chosen according to the benefit over cost ratio. In this level an interval for this ratio is considered and each alternative has a B/C ratio in this interval is chosen and named as set C. (Fig. 3-b shows the fact)

![Fig. 3. a) Level one: applying system constraints. b) Level two: using benefit over cost ratio](image)

The third level which chooses the final alternative deals with the standard deviation of total travel time which reflects the stochastic nature of demand and certainty about demand. In this level the alternative with the minimum standard deviation is chosen as the best alternative.

**8. Sample Network**

In order to prove the feasibility of proposed model, it has been applied to a sample network. Fig. 4 shows the sample network. This network has 80 arcs which
are city arterials and 49 nodes which are bus stops. Assumptions for this problem are as follows:
1. Simulation step is 1 minute.
2. Travel time on network arcs are constant and are a multiple of 1.
3. Planning interval T is 1000 minutes.
4. Passengers choose their shortest route as their choice considering a 5 minutes penalty for transfer.
5. Origin-destination matrices are in 10 minutes intervals which have uniform distribution within the 10 minutes interval.
6. Demand has a normal distribution function.
7. Passengers won’t change their mode of transport as waiting time increases.

Eleven bus routes were designed for this sample network. The network, routes and demand are three main inputs for the model. Beside that, some system policies are needed especially for creating alternative timetables. The policies which have been considered in this problem are:
1. Splitting the planning horizon into three different levels.
2. Ranking bus routes into three different ranks.
3. Considering different steps in alternative making process for each type of time interval and route rank.

The minimum and maximum headway for each level of time interval (1 to 3) and route rank (A to C) are as follows:

\[
\begin{bmatrix}
1 & 5 & 8 & 10 \\
2 & 8 & 10 & 12 \\
3 & 10 & 12 & 15
\end{bmatrix}
\]

\[
\begin{bmatrix}
1 & 10 & 15 & 20 \\
2 & 15 & 20 & 25 \\
3 & 20 & 25 & 30
\end{bmatrix}
\]

Also steps in creating timetable alternative is as follows:

\[
\begin{bmatrix}
A & B & C \\
1 & 3 & 4 & 6 \\
2 & 4 & 6 & 7 \\
3 & 6 & 7 & 7
\end{bmatrix}
\]

In the first step of running the model 768 timetable alternatives have been created according to the policies mentioned before and they have been stored in a text file for being used in simulation process. The required fleet size for every timetable is also calculated in this step. These values will be used in step 4 (choosing appropriate timetable). For creating alternative timetables a MATLAB program has been developed which creates alternatives in a reasonable time.

In the second step the shortest path between every two station is found, considering a 5 minutes penalty for transfer. To do so, a C++ program has been developed which generates a text file containing the shortest path data, readable by the next level programs.

According to part 5, Floyd-Warshall algorithm has been used for finding the shortest path between every pair of station.

At the third level the outputs of first two levels, network matrix, route data and demand matrices are become the inputs of the simulation process. On the simulation process the dispatching, boarding and alighting of passenger bunches are done as described in Fig. 2 and utilization measures are calculated. This process which was implemented in C++, calculates the utilization measures in a reasonable time in a 1000 minutes planning interval.
In addition to total waiting time, the simulation process has the ability to calculate so many measures which are useful for planning the system. Traveled passenger-kilometer, serviced trips, not serviced trips and buses load factors are the most popular ones which can be published as the output of the simulation process. The program developed in this research can handle all of them. Table 2 shows the results of the simulation process for best seven and the worst four timetable alternatives in waiting time.

As can be seen in table 2 the difference between the alternative with the minimum waiting time and the one with maximum waiting time is 100%, in case the fleet size difference of them is only 30% that shows the influence of temporal coverage in timetable setting problem. This fact is also clear when alternative #55 has waiting time equals to 6882 minutes with fleet size equals to 125 while alternative #2 has a waiting time 6975 minutes with fleet size equals to 131.

As can be seen in table 2 the difference between the alternative with the minimum waiting time and the one with maximum waiting time is 100%, in case the fleet size difference of them is only 30% that shows the influence of temporal coverage in timetable setting problem. This fact is also clear when alternative #55 has waiting time equals to 6882 minutes with fleet size equals to 125 while alternative #2 has a waiting time 6975 minutes with fleet size equals to 131.

After calculating utilization measures in the simulation process it’s time to choose the best alternative for this sample problem. As mentioned in part 7 a three level decision making process would choose the best alternative. At first level a fleet size constraint equals to 110 and a waiting time constraint equals to 10000 minutes wipes out 683 alternatives and leave 83 alternatives in feasible solutions. The second level of decision making process, considers an operation cost equals to 700000 Rials for each bus in a day and a value of 20000 Rials for one hour of waiting time and wipes out 63 other alternatives and leaves only 20 alternatives in feasible solutions. Finally by the criterion of standard deviation the best alternative is chosen which is Timetable #276. The headways of every cycle of this timetable are shown in table3.

Fig. 6 shows the place of chosen alternative among other ones. serviced trips, not serviced trips and buses load factors are the most popular ones which can be published as the output of the simulation process. The program developed in this research can handle all of them. Table 2 shows the results of the simulation process for best seven and the worst four timetable alternatives in waiting time.

As can be seen in table 2 the difference between the alternative with the minimum waiting time and the one with maximum waiting time is 100%, in case the fleet size difference of them is only 30% that shows the influence of temporal coverage in timetable setting problem. This fact is also clear when alternative #55 has waiting time equals to 6882 minutes with fleet size equals to 125 while alternative #2 has a waiting time 6975 minutes with fleet size equals to 131.
9. Conclusions and Extensions

The hypothesis of this paper was the fact that we can reach to better temporal coverage of demand and a more certain amount of passengers, using the stochastic variable demand as the objective of timetable setting model. The simulation based model developed in this paper can set all types of timetables (headway-based and schedule-based) and does that in a reasonable time. According to the results of implementation, it can be said that considering demand variation and using maximum coverage -reflected by the waiting time measure- in the model lead us to better timetable in the mean of temporal coverage.

Although considering the stochastic nature of transit demand is a realistic assumption and makes the model more real. But the results shows the fact that according to the complexity that this consideration has brought to the model and solution of the model, it has not caused so much difference in standard deviation of utilization measures. However this conclusion should be studied more because the demand which was used in this paper was an imaginary demand with normal distribution and it is so possible that if a real data be used for the model the stochastic nature of demand have more influence on timetable setting.

This study presented the first timetable setting model under stochastic demand for urban bus networks and so many assumptions were considered to make the problem solvable. Any assumption other than what has been done in this paper can be a topic for a future work. Besides, we propose to replace the alternative making step of the model with a genetic algorithm based program and use the simulation process as a fitness function for genetic algorithm which would result a better timetable.

References


