Optimization in Dial-a-Ride System Analysis:

A comparison of recent modelling and an expected value model.

August 1, 1978
Technical Report Documentation Page

1. Report No. UM-HSRI-78-32
2. Government Accession No.
3. Recipient's Catalog No.
4. Title and Subtitle Optimization in Dial-a-Ride Systems Analysis: a comparison of recent modelling and an expected value model
5. Report Date August 1, 1978
6. Performing Organization Code
7. Author(s) Nancy E. Wallace
9. Performing Organization Name and Address Michigan Transportation Research Program (MTRP) Highway Safety Research Institute (HSRI) 2901 Baxter Road The University of Michigan Ann Arbor, Michigan 48109
10. Work Unit No. (TRAIS)
11. Contract or Grant No.
12. Sponsoring Agency Name and Address Michigan Department State Highways & Transportation (MDSH&T) State Highway Building, P.O. Box 30050 Lansing, Michigan 48909
13. Type of Report and Period Covered Final Task Report June 1 - August 1, 1978
15. Supplementary Notes

16. Abstract

A literature review of optimization models for small bus system designs is presented. Maximization algorithms are briefly discussed for project CARS at Massachusetts Institute of Technology, project BUSTOP at Northwestern, a model developed by Tapan Datta at Wayne State University, and a mean value model developed at Ford Motor Company and programmed at The University of Michigan. The design features of the models are presented and suggestions are offered as to the applicability of the models to the transit authority planning environment.

17. Key Words
Dial-a-Ride Design Models
18. Distribution Statement

The opinions, findings, and conclusions expressed in this publication are those of the authors and not necessarily those of the Michigan State Highway Commission.
1. Preface
2. Introduction ............................................. 1
3. The Nature of the Dial-a-Ride Problems Literature Review .... 1
5. Planning in the Transit Authority: Need of a Middle Ground ... 9
   5.1 Mean Value Model ................................ 10
   5.2 Model Flow and Data Inputs ......................... 12
6. Conclusions .............................................. 14
   Bibliography ............................................ 15
1. PREFACE

This report stems from a request by Mr. Charles Uray, Jr., Chief Deputy Director, Michigan Department of State Highways & Transportation, to the Michigan Transportation Research Program for identification and description of models which could be used by the department staff in estimating small bus fleet requirements for future dial-a-ride system operations. Upon identifying these models, it became obvious that there were significant differences in their structure, input variables, and general approaches. It also became obvious that they were designed for different users (e.g. academic, transportation authority planner, etc., etc.).

Rough gross projections of fleet requirements can be made manually by using such factors as population density and area size to be served. However, if more precision in the estimates is desired, then operational characteristics and levels-of-service, among other things, must be taken into account. Therefore, systems must be "designed" at least on a preliminary basis. For this purpose, the models reviewed here are dial-a-ride system design models.

This paper also addresses the question of what would be the most practical model government planners could use for the development of system designs and reasonably accurate fleet size projections.

1. Letter to Dr. Charles G. Overberger, Vice President for Research, The University of Michigan, Ann Arbor, Michigan, from Mr. Charles Uray, Jr., Chief Deputy Director, Michigan Department of State Highways & Transportation, Lansing Michigan, December 14, 1977.
2. INTRODUCTION

Despite the growing interest in demand-responsive (dial-a-ride) bus systems, the debate over optimal design structures for such systems remains largely unsettled. Generally, the existing models rely on a network problem or allocation problem formulation of the system dynamic. However, there is considerable variation between the models due to differing emphasis on the demand, supply, or cost modelling frameworks. There are also enormous differences in the problem formulation as defined by the academic or market development researchers and the transportation system planners within the local transit authorities. These differences, though partly due to variable computer capabilities, are also a reflection of different spatial and scale orientations. The state-of-the-art researcher tends to concentrate on models driven by stochastic estimations of demands, whereas the planner in the local authority tends to concentrate on rule-of-thumb relationships between sector divisions and demand rates as a function of known residential density.

This analysis is an attempt to speak to the disparity between the modelling methodologies in academia and those methods most commonly used by bus system planners. The initial question asked in the paper is what kind of problem is a dial-a-ride system analysis? The literature review attempts a partial response to that question as well as a consideration of the various linear optimization techniques used to model dial-a-ride design. The literature critique then considers a simpler and more usable model, again using linear techniques, which provides a more manageable middle ground between the complicated optimization models that currently exist in academia and the general lack of similar techniques in the actual planning environment.


In 1976, there were more than sixty demand-responsive transportation systems in the United States and Canada. These systems are designed to offer the desirable characteristics of automobile or taxi travel at a cost that is only slightly higher than conventional bus systems. The systems offer door-to-door service with a goal of guaranteed wait and ride times for the consumer. Dial-a-ride system designs are usually categorized into four service types: route deviation, many-to-one, many-to-few, and many-to-many.

---

The route deviation service is similar to the more conventional large bus system, however, vehicles may operate either on fixed schedules or on telephone dispatch. The many-to-few system differs from the route deviation in that it provides service between many points (neighborhoods) and a single activity sector (a shopping area). Many-to-few service structure is more complex, in that it provides transportation to several activity areas and many sectors, all on a single vehicle tour. This service type requires that the system designer consider more complicated tour designs that can remain optimal for various vehicle dispatch policies. The least structured system is the many-to-many. Here activity areas become the same as other origins and destinations. In the many-to-many design the best level of service is sought such that the wait time of new demand is minimized without excess inconvenience to the passengers already on the vehicles.

Various linear optimization techniques apply to the analysis of dial-a-ride system structure. The consideration of the dial-a-ride system design first poses the question of minimal trip time flows between two specified nodes, or linked pairs of nodes, in the network. This would imply that a multi-terminal flow problem in which several nodes are designated as (source-terminus) pairs, could be used to minimize time flows through the network. Intersector travel time estimates and vehicle tour planning could then be determined by coding demand sectors by arcs and nodes and by assigning average speed estimates to each arc. Shortest path algorithms could be used to sequence vehicle stops subject to assignment constraints such as consumer wait time, consumer ride time, and the nature of dispatch triggers as measured in time. The traveling salesmen problem (Elmaghraby argues that the problem can also be viewed as a longest-path problem) also pertains once a travel time matrix has been defined, $D_{ij}$, between every ordered pair (i, j), so that in a sequence of nodes $(i_1, i_2, \ldots, i_n)$:

1. a maximum number of nodes up to an assigned constraint appears in the sequence
2. the total length of the sequence is minimal.

---

4. Ibid. p. 13
A literature search of current dial-a-ride analysis models indicates a fairly consistent usage of the network flow analysis and shortest path, or traveling salesman problem formulation. There are however, interesting differences among the models.

In April of 1967 transportation researchers at Northwestern University developed a generalized dial-a-ride system analysis model (BUSTOP) for service to diffused origins and destinations (many-to-many). Network flow analysis was not used in this problem formulation due to the assumption that the service area was a rectangular grid network with constant travel times on each link. Maximum travel time in the model was a linear function;

$$\text{Max Travel Time} = \begin{cases} kn & \text{for } n \leq n' \\ T + en & \text{for } n > n' \end{cases}$$

where

- $n$ = the number of links between origins and destinations
- $n'$ = a control parameter constant set equal to 10 links
- $k$ = a control parameter constant that equals 1 minute per link
- $T$ = a control parameter constant that equals 5 minutes
- $e$ = a control parameter constant that equals one-half minute per link

Minimum pick-up time was a constant equal to one minute. The maximum pick-up time guaranteed the passenger a pick-up before six minutes had elapsed.

Vehicle assignment in the model was achieved using the "nearest neighbor" principle that is, the vehicle which is nearest (by distance) to the origin of the call was the first one considered for service. Passengers were assigned to the closest vehicle that could service them without violating the time constraints of any other committed passengers. A new vehicle was generated only if there was not another vehicle in the system that could service a call. When a vehicle had delivered its last assigned passenger, it would return to the terminal and then become the first to be regenerated.

The BUSTOP demand assignment was not optimal. In fact, no attempt was made to assign a passenger to a bus so that the efficiency of the system was maintained at an optimum. Once a passenger was assigned to a bus no attempt was made to reassign, regardless of the actions of the other buses in the system.

---

The objective of the BUSTOP model was to determine the supply and demand operating characteristics of a dial-a-ride system. Cost considerations were not explicitly included in the model, although unit cost data could have been calculated based on the system results. The model designers further suggested that a new algorithm be included to aid in keeping track of routed buses by their assignment via shortest paths. Such a model development would entail the inclusion of a "next-node matrix" that would be generated through the tree generations. Entries in the matrix would contain the identification number of the next node through which the vehicle passed on its shortest path to the next node on the bus list.

Another important dial-a-ride system design model is project CARS at MIT. From 1967 to 1971, the potential for the use of computers in the control of demand-responsive transportation systems was explored by researchers at MIT. The early work on project CARS was based on a rectilinear grid system. The more recent work, however, has greatly simplified this structure and now uses only airline distances between points. The justification for this change in the new generation of models is that a specific street network is not relevant to the sensitivity of the fundamental algorithm concepts.

The original CARS algorithm was designed to minimize total service time (for actual and future passengers) subject to fixed constraints on:

- a) wait time (constant)
- b) ride time (linear function of trip length)
- c) total service time constraint (linear function of trip length)

The feasible assignments were evaluated using a linear objective function measured in time. The objective function that was used described "consumer disutility" as well as incremental increases in vehicle tour length. The feasible assignment that also showed the smallest value for the objective function was always selected as "best." If no feasible assignment existed, infeasible assignments were evaluated by using the objective function, and the best assignment was chosen - that is, the assignment that minimized the objective function. Any assignment that did not violate any constraint was preferable to an assignment involving a violation, independent of the value of the solution when used in the objective function.

The modeling policy was intended to reduce the number of passengers experiencing unreasonably long service times, and therefore it implied some increase in the mean service time. Such an increase in the mean service time necessarily reduced the over-all level of service (a measure of the ratio of dial-a-ride transit to auto travel time) for the system. Thus it appeared in this formulation that assignments that were superior from the viewpoint of the objective function were rejected if a constraint was violated. (The hard constraints thus appeared to introduce perturbations in the service performance and led to decisions that tended to underutilize system resources.)

The problem of resource misallocation arose from conflicts in the dual goal of optimizing assignment allocation while creating tours. As Bellmore and Nemhauser note in their survey of analytical work using the traveling salesman problem:

"Let t be any time in which each node is visited exactly once and $x_{ij}=0$, if $(i,j)$ is not in the tour and $x_{ij}=1$ if $(i,j)$ is in the tour. The $x_{ij}$ is a feasible solution to the assignment problem

$$z(t) = \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} x_{ij} = w.$$ 

Unfortunately the converse may not be true; feasible solutions to the assignment problem may not be optimal." 8

A further difficulty in finding optimal solutions in the problem design stemmed from the lack of an algorithm to restrict vehicles from serving the whole system. There was no mechanism in the model that limited vehicles to a finite set of origins and destination pairs.

In view of the problems with the first-generation CARS algorithms, the researchers at MIT finally concluded the "classical optimization solutions to the customer-to-vehicle assignment problem are not feasible." 9 The central problem then identified was that the older algorithms were designed to work in extremely complex and difficult environments that were stochastic in several respects. The demands, which were designed to appear randomly in time, were designated to origins distributed probabilistically in geographic space. Theoretically, although the models could be formulated and solved for small problems, there was no realistic size.

There was thus a re-assessment of the project and a new generation of algorithms were developed using a heuristic rule to assign the demand to any given vehicle and then to appropriately insert the new origins and destinations in the vehicle's route.

The present project CARS algorithms consider the cost to the user of the system and the service time (waiting time plus travel time) by the user. The objective of the second generation CARS algorithm is to minimize both conflicting variables (conflicting because lower service time necessitates more vehicles in service which implies a cost increase). The new algorithm thus "optimizes" in the two dimensional space of cost and time.

The CARS algorithm attempts to insert a new demand origin and destination in all possible positions in all possible vehicle tours. With each attempt, a function is used to measure the disruption caused by the insertion of the new origin and destination within a route. The objective function thus selects an assignment by considering the incremental degradation of the service for all users as a function of the service parameters applicable to any given user class. The objective function in this model is:

\[
\text{Minimize } Z = \sum_{i} \sum_{j=i}^{M_i} (U_{ij} - U'_{ij}) + U_{\text{new}} + E \Delta TL
\]

Where:

- \(N\) = number of classes of passengers
- \(M_i\) = number of passengers is class \(i\)
- \(U_{ij}\) = disutility before the trial assignment of passengers already assigned to a tour
- \(U'_{ij}\) = disutility after the trial assignment of passengers already assigned to a tour
- \(U_{\text{new}}\) = disutility of the new passenger
- \(WT_{ij}\) = wait time for passenger \(j\) in class \(i\)
- \(RT_{ij}\) = ride time for passenger \(j\) in class \(i\)
- \(PTD_{ij}\) = pick-up time deviation for passenger \(j\) in class \(i\)
- \(DTD_{ij}\) = delivery time for passenger \(j\) in class \(i\)
- \(A_i, B_i, C_i, D_i\) = class specific weighting
- \(E\) = weighting factor for tour length increase
- \(\Delta TL\) = tour length increase caused by trial assignment.

---

The function thus sums the marginal disutility due to each trial insertion. The sum has three further components: the marginal disutility to a passenger already on a tour, the disutility of the passenger currently being assigned, and the disutility attributable to the added load on the system resources. The $A_i$ and $B_i$ reflect the trade-off between wait time and travel time and service reliability (ability to meet guaranteed pick up and drop-off times) is controlled through $C_i$ and $D_i$.

The intent of the algorithm is to determine the balance that must be struck between passenger utilities (good service to the customers on the system) and system resources. The model can be used to evaluate the degree to which service can be controlled through varying the parameters:

1) number of vehicles in operation
2) area size (in square miles)
3) demand rate (in passenger miles per hour)
4) the objective function parameters.

The model can also be expanded to consider the many-to-many dial-a-ride dispatch service.

Still another attempt to model the dial-a-ride system structure has been carried out by Dr. Tappan Datta at Wayne State University. The Datta effort focused on the cost-effectiveness of various system configurations. The research hypothesis guiding this work was "that a generalized study approach demonstrating a viable system selection strategy of various discrete system alternatives for a wide range of demand was needed for decision-makers to optimize system selection." The model developed by Datta defined system attribute functions for each system in terms of ridership levels, roadway network configuration, pooling policy, and service area. The work was initially designed to consider the many-to-one service type. However, there were adaptation options so that the many-to-many service could also be considered. The basic objective of the research was to evaluate the trade-offs between attribute functions and thus to determine bounds of demand and user characteristics. It was hypothesized that if system attributes under various demand conditions were reduced to cost functions, the suitability of various system alternatives could be evaluated.

11. IBID, Wilson, p. 48
The model is initiated by generating strings of collections and distributions for each specific area of dial-a-ride operations, with no limit set on the size of the area. A point-to-point travel distance matrix is generated from input sector sizes and demand rates. The distance matrix is converted to a travel time matrix for a given speed. The model examines the demand list of each sector at a specified interval of time and tests the maximum headway constraint and dispatch logic. If the dispatch triggers are not satisfied the model moves to the next increment of time and tests again. This operation is done concurrently in all sectors. When a dispatch criterion is satisfied, a bus is dispatched. The distribution sub-routine produces the "best tour en route to minimize passenger travel time of the string of executable demands." The statistics produced include wait time, ride time, number of vehicles required, and vehicle productivity (number of passengers carried per vehicle). The simplifying assumption is then presented that if demand for service and operating costs are all inelastic, the model output could be considered as a production function. Such a production function when used as the constraint on a simple linear social welfare function of the form \(U_j = f\) (system performance, planning policies)\(^{14}\) could hypothetically be used to maximize welfare.

Other models that are beginning to appear in the dial-a-ride literature tend to devote less time to the actual system interaction and more time estimating disaggregated consumer preferences for use in the objective functions cited in the preceding models. Such models seek to identify the socio economic, demographic, and attitudinal parameters that affect the usage level of dial-a-ride systems.\(^{15}\)

13. Ibid p.62


-8-
4. CONCLUSIONS: CONCERNING STATE-OF-THE-ART

It would appear that the following conclusions could easily be drawn from the literature just cited. The limited sensitivity analysis attempted in current dial-a-ride models indicates a sensitivity of system performance to waiting time, trip time, and link travel time. It appears that stochastic modeling of dial-a-ride systems leads to suboptimal use of vehicle resources due to the trade-offs between optimal assignment and optimal tour formation. Finally there is some indication that a heuristic traveling salesman algorithm will work for routing in most of the models.

5. PLANNING IN THE TRANSIT AUTHORITY: NEED OF A MIDDLE GROUND

The modeling effort in academia is often in stark contrast to the dial-a-ride system design actually used by the transit authority. Tours are usually developed from the experience gained from the operation of the large line haul bus systems. In the real world of the transit authority it appears that decisions concerning sector size and network design are made using estimates of demand based on known housing density. One might hypothesize for instance that with the typical residential densities of 1,500 households per square mile, the demand density would probably be on the order of 20-60 demands per hour, which would imply .05-.10 round trips per household per eight-hour day. A further rule of thumb frequently used is that sector tour time generally equals the time necessary to complete the perimeter of a given zone. Other system design guidelines apart from the previously reviewed literature include such notions as,

"Zones may be small or large,....but should include enough potential riders to keep at least one vehicle busy. A typical zone will be serviced by 1 - 5 vehicles, and range in size from one square mile for a high-demand feeder service to tens of square miles in rural areas. .... Zones should lie on one side or the other of main line routes rather than straddle them...Transfer points should be at the corners of zones nearest the major destination served by the main line route."

---

Optimization in this local planning environment comes slowly on a learning curve of route adjustments to perceived demand needs. The initial planning in a large scale system is considered to be only tentative and is subject to adjustment and re-adjustment as the real nature of demand patterns reveal themselves through ridership counts.

It would thus appear that there is a need for planning tools which use a heuristic form of the optimization algorithms but are more generally usable. There is a need for a methodology that is easier to use, requires less data and less computer time, and yet provides gross estimates to central questions in the dial-a-ride design process. A model is needed that could determine the level of service that might be provided to various designated service areas at different demand densities. Such a model would provide necessary planning information in response to questions such as: if a dial-a-ride system is to be implemented to serve some given number of passengers per hour, with policy determined demand queues, sector sizes, and travel speeds, then how many vehicles will be required and what will be the average passenger travel times? If such information were forthcoming, cost information could be calculated from the estimated service characteristics. Such a model would be linear in form, as are most of the "state-of-the-art" models, and it would include an optimum heuristic that would seek to minimize overall mean passenger time by a proper allocation of the available fleet among the designated zones.

5.1 Mean Value Model

A proposed middle ground dial-a-ride system design model is the mean value model recently adapted for interactive use on the Michigan Terminal System. The model was initially developed in the early 1970's by Mr. J. R. Mumford and Mr. F. J. Mason in the Transportation Analysis Department, Ford Motor Company. The mean value model is a second-generation response to a previous generation of stochastic models that used network flow algorithms and the traveling salesman algorithm to determine optimal service characteristics of dial-a-ride system design. As Mumford and Mason suggest, the "fundamental concept of this model was the elimination of all randomness from the problem.... all random variables are replaced by their expected values." 17

The demand rates for the system design are assumed to be constant and the vehicles are assumed to be dispatched by a constant queue trigger (set by the analyst), on constant headways (time intervals between vehicles). Tour times are also constant, with

values dependent on the mean number of distributions and collection passengers carried and on the geometry of the sector sizes.

The central concept of the mean value model is the vehicle tour. The tour is a series of vehicle travel paths that must begin and end in an activity center (AC), though not necessarily the same AC. The tour design is also a limiting feature of the model, for although up to fifteen sectors can be analysed sequentially, the sectors must all be linked to the same activity center (AC). This structural limitation may prevent the analysis of complex systems with multiple activity center dispatches; however, the design is not very different from the actual dispatch of dial-a-ride service.

Vehicle tours are compositions of sets of subtours. There are three possible types of subtours in the model:

1) Tours serving an activity center. This tour type consists of travel from the previous subtour or tour to an AC where there is an off-loading and on-loading of passengers.

2) Tours serving sector $S_i$ and maintaining tour direction. This second subtour type defines a travel pattern from the previous subtour to sectors $S_i$ followed by the concurrent drop-off and collection of riders within $S_i$. The assumed position of the next subtour of this tour leads the vehicle to leave $S_i$ from the point opposite from its entry into $S_i$. The path of the vehicle within $S_i$ keeps the initial direction of the vehicle tour.

---

**Figure 1**

1. Serving AC1
2. Serving $S$ (type 2)
3. Serving AC2
4. Serving $S$ (type 2)
5. Serving AC1

**Figure 2**

1. Serving AC1
2. Serving $S_1$ (type 2)
3. Serving $S_2$ (type 2)
4. Serving AC2
5. Serving $S_2$ (type 2)
6. Serving $S_1$ (type 2)
7. Serving AC1
3) Tours serving sector $S_i$ and reversing tour direction. The third tour type begins with travel from a previous subtour to sector $S_i$. The vehicle enters $S_i$ and then proceeds to drop off all passengers destined for $S_i$ and then picks-up all the passengers originating within $S_i$. The vehicle leaves $S_i$ at or near the original entry point.

![Diagram](image)

**FIGURE 3**

- a. serving AC1
- b. serving S1 (type 3)
- c. serving AC1

The above-defined subtours may be combined as needed by the analyst. The restriction is that all vehicle tours must begin and end with a type 1 subtour—that is, in an AC. A mean headway $H_i$ (time between tour stops) is also assigned to each tour $i$, thereby determining the level of service for that tour.

5.2 Model Flow and Data Inputs

Four basic data categories are prompted by the mean value model program: service sector data, activity center data, vehicle tour and subtour descriptions, and demand queues. The flow of the model is shown in Figure 4, which indicates the various data input and calculation components. The chart indicates the four major data inputs that are read and then checked for balanced headways. If the headway test shows that the headways in and out of AC's are out of balance, the analyst can reenter "new tour headways." The service characteristics are then computed again with optional parameter changes if the design result does not appear desirable. The possible redefinition of service inputs allows the analyst to carry out sensitivity analyses on various differences within the same system design. Cost information may be calculated from the model results. However, those calculations are not a part of the model design.

The mean value model as designed by Mason and Mumford describes the three
FIGURE 4

INPUT

SECTOR AND SERVICE DATA

VEHICLE TOUR DESCRIPTIONS

DEMAND QUEUE DATA

DWELL FUNCTION DATA

NEW TOUR HEADWAYS

NEW SECTOR SIZES

NEW LINE Haul TIMES

NEW HOURLY DEMAND RATES

START

READ INPUT DATA

CHECK HEADWAY ENOUGH

COMPUTE HOURLY DEMAND RATES BY SUBTOUR

COMPUTE PICKUPS AND DROPOFFS BY SUBTOUR

COMPUTE EVENT TIMES PER TOUR BY TOUR TYPE

COMPUTE VEHICLE TOUR TIMES

SYSTEM ADJUSTMENT

COMPUTE MEAN WAIT AND RIDE TIME

COMPUTE VEHICLE FACTORS

END

-13-
subtour types as a series of events. The occurrence of each event is calculated to be relative to the preceding event and is expressed as a function of the supply and demand variables and the sector parameters. The key calculation in the model is that of the event times for each subtour. The form of the model seeks to minimize vehicle time, which is defined as a function of on-and-off boarding time (dwell times), an acceleration constant, and an hypothesized sector division as calibrated with the previously determined travelling salesman formulation in the first-generation models.

6. CONCLUSIONS

Although the mean value model provides only expected values of dial-a-ride service, sensitivity analyses of the model parameters indicate that the model gives quite accurate indications about policy trade-offs among dial-a-ride service characteristics. The model provides good guideline information about the effects of policy-determined service sector characteristics. From the service sector characteristics, productivity and level of service measures can be derived, so that the system design could be translated into cost and supply terms.

The mean value model thus offers a viable alternative to the costly stochastic models. In its current form the model costs approximately $.30 per run and requires simply formatted input data. The model appears particularly well suited to use by transit planning authorities where limited computer facilities and lack of highly specialized personnel make a complex modelling procedure impractical.


Addendum - A

Model Sources
Model Sources

1. CARS
   Migel Wilson
   Massachusetts Institute of Technology
   Cambridge, Massachusetts 02139

2. Tappan Datta
   Department of Civil Engineering
   667 Merrick
   Wayne State University
   Detroit, Michigan 48202

3. Mean Value Model
   Michael Dewey
   SEMTA
   211 Fort Street W
   P.O. Box 333
   Detroit, Michigan 48231