



# Robust Passenger Itinerary Planning Using Transit AVL Data

Mark HICKMAN

*Abstract*—Automatic vehicle location (AVL) data can be used for a variety of operational and planning purposes at a transit agency. One possible application is for passenger itinerary planning, where the passenger seeks to go from an origin to a destination in some “optimal” way. The AVL data give actual vehicle arrival and departure times, which in turn can give more robust itineraries to passengers. A method to determine such robust itineraries is described, and a small example using the bus network in Tucson, Arizona (USA) is used to illustrate the method. Advantages of this method over traditional itinerary planning methods are also outlined.

*Index Terms*—automatic vehicle location, itinerary planning, routing, transit itinerary

## I. INTRODUCTION

Service reliability is a continual problem in the public transit industry in the United States. Services that run off schedule can cause frustration to passengers, in addition to greater waiting time, transfer time, and total origin-to-destination travel time. There are several ways of addressing service reliability, most notably in improving scheduling and operations control techniques to bring the schedules more into line with operational requirements. Nonetheless, passenger frustration about uncertain travel times, and the timing of particular origin-to-destination trips, are often a reality.

One way of helping passengers to work within the existing service environment is to provide them with information about the reliability of the service. As an example, many transit agencies in the US are exploring passenger information systems that provide current bus location and/or next bus arrival information [1]. This can have the effect of reducing passenger frustration on the current trip by keeping them informed about the service.

Another source of traveler information is itinerary planning. In the itinerary planning problem, a prospective passenger seeks to find a path from their origin to their destination, at a given time of day. The problem is to determine the most

desirable path, and then to provide the passenger with an itinerary consisting of when to meet the bus, where to get on the bus, and how long it will take to reach the destination (or an intermediate stop, if a transfer is required). Most commonly, variables that are used to determine the “most desirable” path include travel times, access and egress times, waiting times, the number of transfers, and fares.

Previous research, and most commercial software, use the service schedule to generate these itineraries. The itinerary planning problem is often modeled as either a traditional shortest path problem, or a time-dependent version of the traditional shortest path problem (for example, [2], [3], [4] and [5]). Small adaptations of these traditional methods are necessary to accommodate waiting times and transfer times in the network, since these measures are clearly time-dependent (i.e., they correspond if even in a small way with the published schedule).

Using the existing schedule is a good option for itinerary planning if all services ran according to schedule; however, in many cases, the schedule may not accurately reflect the variability in travel times on a day-to-day basis. As a result, the itinerary plans that are produced may not be accurate, and may unfortunately leave the passenger with considerable uncertainty in their travel. As an example, a passenger may simply want to know how likely it is that they may actually make a certain transfer.

However, transit vehicle location technology can provide information for the passenger about service reliability. Technologies such as the global positioning system (GPS) can be used to locate buses on a very frequent basis, in order to establish if and by how much the bus may be off schedule. Moreover, if these data are archived, they can be used to tell passengers about historical schedule adherence information. Such information can be used directly in the itinerary problem to advise passengers on the most reliable services.

This paper defines and illustrates a method to use archived automated vehicle location (AVL) data to provide more robust itinerary planning. By using historical AVL data, a picture of the system reliability can emerge that allows the passenger to make a more informed choice of an itinerary. To describe this method, the next part of this paper presents some of the dimensions of service reliability that affect itinerary planning, and their effect on existing methods for transit itinerary planning. The third section describes an algorithm to generate

This work was supported in part by the National Science Foundation, CAREER Program.

Mark Hickman is with the University of Arizona, Department of Civil Engineering and Engineering Mechanics, P.O. Box 210072, Tucson, AZ 85721-0072, USA (e-mail: [mhickman@engr.arizona.edu](mailto:mhickman@engr.arizona.edu)).

passenger itineraries using archived AVL data. In the fourth section, the methodology is illustrated on a simple example from Tucson, Arizona. Finally, the last section offers conclusions on the value of this method for robust itinerary planning.

## II. CHALLENGES TO ROBUST ITINERARY PLANNING

### A. AVL Data

Automatic vehicle location (AVL) technology typically involves some technology on board the bus that, in conjunction with either terrestrial or satellite beacons, determines where the bus is at a given time. This location information may in turn be communicated to a central control location, either on a periodic basis ("polling") or on an exception basis. These data can be used for managing operations, but they can also be stored for later off-line planning purposes. Also, the data can be pre-processed before archiving into particular summary measures; for data management purposes, this avoids storing all the raw data.

The archived data from the transit agency in Tucson consists of the times when buses were recorded passing scheduled time points on their runs. Specifically, the data includes the scheduled time (to the nearest minute) the bus was to arrive at the timepoint, the actual time the bus arrived, and the deviation from schedule (again in minutes). It also contains various other data such as the bus identification number and the route number.

From this archival data, some simple data processing can be done to compute: (1) the probability mass function of the arrival of each bus to each time point on the route; and (2) the probability mass function of the running time from one time point to another, given the time when the bus arrived to the first time point. With these probability mass functions, various measures can be derived, such as means, variances, and other distributional properties of both bus arrival times and bus running times. An example of the probability mass function of bus arrivals to a particular time point is given in Figure 1. In this case, the bus is most commonly 2 min late (about 58% of the time), but is generally between 1 min early and 3 min late.

Figure 2 shows the probability mass function of the travel time from one time point to another, at a given time of day. What one may note in many cases is that if the bus is running ahead or behind schedule at a timepoint, the driver may then modify their driving behavior (by waiting, or slowing down, or speeding up) over the following part of the route. As a result, the distribution of running times over the route is dependent on the time when the bus arrives at the first timepoint. In this case, the bus arrival at 7:32 (two min early) results in a density function that is skewed to longer travel times, as the driver often slows down to avoid being early further down along the route.

Histogram of Arrival Times, Route 9  
7:32 am at Tanque Verde and Sabino Canyon

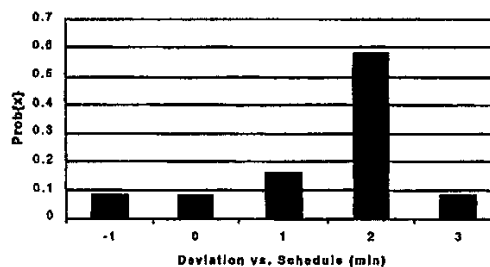


Fig. 1. Probability mass function for bus arrivals at a timepoint.

Histogram of Travel Times, Route 9  
7:32 am Tanque Verde & Sabino to Downtown, 2 min early

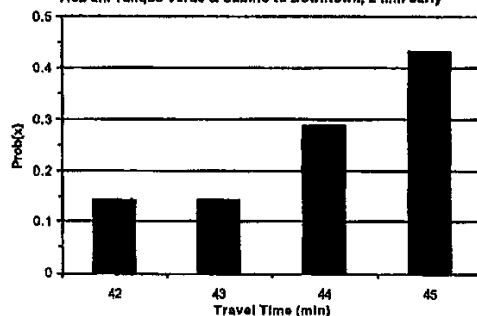


Fig. 2. Probability mass function for bus running times from one timepoint to another.

### B. Path Determination

The fact that bus arrival times and running times are both defined in terms of time-dependent probability density functions has significant implications for how an itinerary can be generated. The travel times in the network, both in terms of the passenger waiting times and the bus running times, are random and time-dependent. Finding a "shortest" or minimum travel time path in this context is considerably more difficult than the simple shortest path methods described previously.

There has been considerable research over the past several years on shortest path problems of this nature. The most notable initial effort into this problem was the work of Hall [6], who examined the problem of finding the *a priori* shortest expected travel time path from an origin to a destination, with a network of stochastic and time-dependent travel times. This research unearthed a number of very critical issues. First, this research was the first to point out that algorithms based on Bellman's principle of optimality are not guaranteed to find an optimal path in stochastic and time-dependent networks. That

is, the shortest path cannot be found simply by concatenating shortest paths within the network. Second, Hall identified a heuristic that could be used to identify a shortest path in the stochastic and time-dependent network. Essentially, the heuristic begins by identifying the shortest path from the origin to the destination using the *minimum possible* travel time on all arcs. The algorithm then continues, in a manner similar to a k-shortest path algorithm, to enumerate paths from the origin to the destination, using the *minimum possible* travel time. The algorithm terminates when the minimum possible travel time on the last path found exceeds the least *expected* travel time on the paths that have already been enumerated. The minimum expected travel time path is then selected. Finally, this algorithm provides a good starting point for this problem. At the same time, since it requires explicit path enumeration, its computational complexity can be considered to be exponential in the size of the network, in the worst case.

Hall's work was also revisited by Wellman, Ford and Larson [7]. In developing an algorithm to solve for the *a priori* shortest path with stochastic and time-dependent travel times, they used the concept of stochastic dominance to identify paths that may dominate other paths in the network. The general idea of stochastic dominance can be stated as follows: if the cumulative probability distribution of the cost of one path A is less than or equal to that of a path B, then path A is stochastically dominated. Graphically, this can be shown in Figure 3. Assuming one is seeking the minimum value of a particular variable  $x$  (e.g., travel time), then path B stochastically dominates path A.

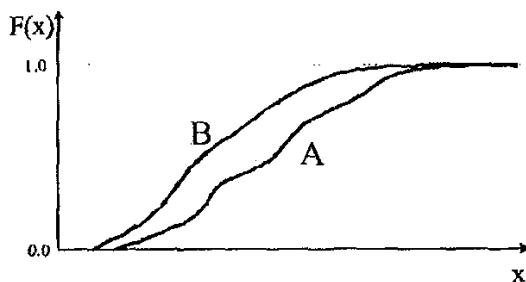


Fig. 3. Pairwise stochastic dominance between paths.

The solution algorithm presented in [7] enumerates paths from the origin to the destination, but "prunes" paths that are stochastically dominated. More formally, path B is considered to be stochastically dominated by path A if and only if, for a given stochastic cost of path A,  $c_A$ , and a cost of path B,  $c_B$ , the following holds [8]:

$$\Pr\{c_A \leq z\} \geq \Pr\{c_B \leq z\} \quad \forall z, \text{ and} \\ \exists \text{ some } z \text{ such that } \Pr\{c_A \leq z\} > \Pr\{c_B \leq z\}$$

By pruning stochastically dominated paths, the algorithm

intends to decrease the number of paths that must be enumerated in the solution procedure.

Finally, Miller-Hooks and colleagues have completed perhaps the most extensive study of the shortest path problem in stochastic and time-dependent networks. Of most relevance to this work on the *a priori* shortest path is the original work of Miller-Hooks [9] and related papers [10], [11], and [12]. In considering the enumeration of paths, the results in [10] and [12] suggest that different types of dominance can lead to efficient pruning techniques during algorithm execution. In particular, for *first-order stochastic dominance*, path pruning techniques can significantly reduce the number of paths enumerated in the solution. Pairwise comparisons are made between the current "best" path and any candidate best path from a given node to the destination. All non-dominated paths are stored in the solution. Finally, despite the clearly poor theoretical properties of the *a priori* shortest path problem, the computational results on randomly generated networks in [10], [11] and [12] suggest computational time is empirically linear in the size of the network.

It is important to note that the results from Hall [6] and Wellman, Ford and Larson [7] relate to both continuous and discrete representations of time. On the other hand, the stochastic and time-dependent problems addressed by Miller-Hooks [9] [10] [11] [12] are strictly problems in discrete time, for which more efficient algorithms can be developed. This fact is exploited in this work as well. Since the SunTran AVL data are given in terms of minutes, a discrete time interval of  $t=1$  min is used for the modeling below.

### III. AN ALGORITHM

The proposed algorithm for this problem uses the basic ideas developed in [6], [7], [10] and [12]. At a more general level, one would like to generate a set of paths that are not stochastically dominated. This set can be shown to correspond to the set of paths that are not dominated under any monotonic passenger utility function (see [9] and [13]). For this specific problem, we assume that the passenger has a utility function, either known or unknown, that is a function of the total travel time. One may also assume that this function is monotonic decreasing with the travel time.

Conceptually, the algorithm proceeds as follows, using the notion of stochastic dominance:

1. Generate an initial path from the origin to the destination that is likely to be non-dominated (i.e., in the final set of non-dominated paths). Calculate the cumulative distribution function for travel time on this path.
2. Proceed by generating paths in the network, pruning out paths based on the stochastic dominance criterion. At any step, compare any newly-generated path with the "best" existing path found to date, and prune any paths that are stochastically dominated.

3. Terminate when no additional paths can be found that are non-dominated.

The concept for this algorithm is based on the previous methods outlined in [6], [7], [10], [12]. However, it is superior to the methods in [6] and [7] in the sense that it provides a way of generating an initial path that is likely to be non-dominated, from which efficient pruning can occur. Also, in comparison with [10] and [12], the problem is defined with a single origin-destination pair, rather than all nodes to a single destination. As a result, this method is computationally more efficient, as it is relatively easy in this case to generate a single path that is likely to be non-dominated. The algorithm then proceeds as in [7], [10] and [12] to compare additional generated paths to the original non-dominated path, as a means of pruning the solution space. Again, this step considerably reduces the computation time, particularly in large networks and/or when many time intervals are considered.

#### A. Initial path generation

In the first procedure, an initial path  $p^*$  from the origin to the destination is generated that is likely to be non-dominated. This path is identified as the *least possible* time-dependent path from the origin to destination for each time interval  $t$ . This is consistent with [6], which began with the least possible time path in the network.

For each arc  $(i,j)$  in the network, the minimum *possible* time-dependent travel time  $m_{ij,t}$  is identified. This minimum possible travel time has an associated positive probability of occurrence, denoted as  $\rho_{ij,t}(m_{ij,t})$ . Then, the minimum *possible* travel time on each arc is used to find the time-dependent shortest path from the origin  $s$  to the destination  $q$ , for each time interval  $t$ . It turns out that this time-dependent shortest path can be generated using existing label-correcting algorithms in polynomial time; in our case, we use the decreasing order of time (DOT) algorithm described in [14]. As a note, the DOT algorithm is applicable when time is represented in discrete intervals  $t$ , but may not be applicable for continuous time representations.

In generating the shortest time-dependent path, we also maintain probability labels associated with each path. The probability label  $\Pi$  associated with each path  $p$  is equal to the product of the probabilities of the minimum possible travel time on each arc on the path. Mathematically,

$$\Pi_p = \prod_{(i,j,t) \in p} \tilde{\rho}_{i,j,t}(m_{i,j,t})$$

This assumes independence of travel times across links in the network. Correlation and dependence result in more complicated calculations, but do not invalidate the results presented here.

In case there are multiple paths with the same minimum possible travel time, the path with the higher probability  $\Pi_p$  of being the minimum is chosen. If both the minimum possible

travel times and the probabilities  $\Pi$  are equal for two or more paths, a single path is chosen arbitrarily. Curiously, if the path with the minimum possible travel time is unique, it will not be stochastically dominated. Even if the path with the minimum possible travel time is not unique, if a path can be shown to have the uniquely highest probability of the path being minimum, the path will not be stochastically dominated. Hence, at the end of the first step, there is an initial path  $p^*$  from the origin to the destination for each time interval  $t$  that is not likely to be stochastically dominated. The full cumulative distribution function (cdf) for the travel time on  $p^*$  is also derived.

#### B. Generation of other non-dominated paths

In the second procedure, additional paths in the network are enumerated, and dominated paths are pruned from the solution space. Using the probability mass functions for travel times in the network, the calculation of the cumulative density function for a path involves simple convolutions of these probability density functions. In the pruning step, a pairwise comparison of paths is performed. If the cumulative density function (cdf) of the travel time on a given path is dominated by another (second) path, then that first path is pruned. This dominance is determined by comparing the cdf of the new path with the cdf of the travel time of an existing non-dominated path – i.e., the minimum possible time path determined previously.

The algorithm terminates when there are no more paths to be evaluated, either because the paths have been added to the non-dominated set or because the paths have been pruned away. Again, it is easily shown that the set of non-dominated paths remaining after this pruning technique includes a path that is optimal under any monotonic passenger utility function [9][13].

## IV. EXAMPLE

As an illustration of both the concept and the methodology, consider the following example taken from the bus agency, SunTran, in Tucson, Arizona. A traveler is interested in going from their home near the intersection of Fort Lowell Road and Alvernon Way to downtown. The trip, a work trip, is to be taken at 6:45 am. A simple schematic of the possible routes and connections to make this trip is shown in Figure 4. In this example, there are four feasible paths that can be taken. All four paths involve a transfer.

If one were strictly to use the printed schedule in determining a path, the path taking Route 11 (Alvernon) to Route 8 (Broadway) is the shortest travel time path, with the bus leaving at 6:49 am and the second bus arriving downtown at 7:20 am. This trip takes 31 minutes, according to the schedule. Additional schedule information on all paths is reported in Table 1.

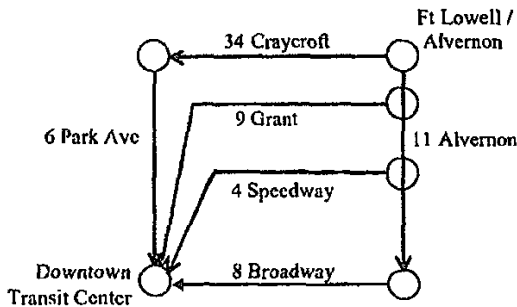


Fig. 4. Schematic of example trip options from Fort Lowell and Alvernon to downtown.

TABLE 1: SCHEDULED SERVICE ON EXAMPLE PATHS

Itinerary	Scheduled Departure	Scheduled Arrival Time	Scheduled Travel Time
34 to 6	6:52 am	7:30 am	38 min
11 to 9	6:49 am	7:45 am	56 min
11 to 4	6:49 am	7:35 am	46 min
11 to 8	6:49 am	7:20 am	31 min

A passenger receiving this information may have several questions, including:

- How likely am I to arrive at the destination on schedule?
- What is the latest I might arrive at the destination?
- What happens if I miss the transfer?
- How likely is it that I miss the transfer?

With the AVL data, these questions can be answered, at least from the basis of historical system performance. The transit agency can use the actual service data to identify reasonable (i.e., non-dominated) paths in the network, and provide the passenger with information about the reliability of service on these paths. Some paths that may be very reliable, but with higher expected travel time, may also be comparable in a passenger's utility to paths with shorter expected travel time but higher variability.

For this example, one month of service data from Tucson's transit agency was processed to derive the bus arrival time and running time distributions described previously. In turn, these distributions were used to generate passenger waiting time and total travel time distributions. These could then be compared directly to determine the non-dominated set of paths. In this example, this is straightforward. However, with larger networks, the enumeration of paths may be more computationally burdensome.

In following the steps of the algorithm, the minimum possible travel time path is also the Route 11 to the Route 8, with a minimum possible travel time of 30 minutes. Using this path as a non-dominated path, in this case the other three paths are all dominated. As a result, this path is optimal for any utility function that is monotonic (decreasing) in travel time.

The cumulative distribution functions of the arrival time downtown are shown in Figure 5. In this case, it is clear that Route 11 to Route 8 is a dominant path. Route 34 to Route 6, however, dominates the other two options (Route 11 to Routes 9 or 4). There is no stochastic dominance between the paths that cover Routes 11 to Route 9 or to Route 4.

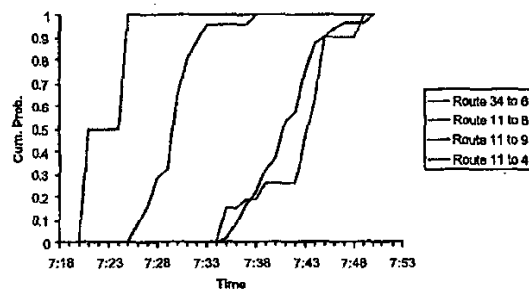


Fig. 5. Cumulative distribution functions of arrival time to downtown.

Once the non-dominated paths are determined, the utility of each of these paths can be determined. If the passenger's utility function is known, the utilities can be calculated directly. Alternately, selected performance measures for each path could be presented to the passenger, allowing the passenger to determine their best path based on their individual preferences.

As examples, the passenger may want to know the earliest departure time from the origin (to avoid missing the bus!), the probability of making a transfer, or the actual expected arrival time, or the earliest or latest possible arrival time. Selected performance measures for the four paths in this example are shown in Table 2. To begin, note that the itinerary using Routes 34 and 6 may run considerably early: its earliest recorded departure is 6:43 am, while its scheduled departure time is 6:52 am. The Route 11 bus seems to depart closer to its intended time, with an earliest departure time of 6:47 am. If a traveler wanted to use the Route 34 to Route 6 itinerary, this information would advise them to arrive very early to the stop. Note also in Table 2 that all the paths have a probability of 1.0 of making the transfer. This may be comforting to the passenger, knowing that they are likely to make each possible transfer.

TABLE 2. PERFORMANCE MEASURES FOR EXAMPLE PATHS

Itinerary	Earliest Departure	Earliest Arrival	Latest Arrival	Prob(Making Connection)	Actual E[Arr Time]
34 to 6	6:43 am	7:26 am	7:38 am	100%	7:30 am
11 to 9	6:47 am	7:35 am	7:50 am	100%	7:46 am
11 to 4	6:47 am	7:35 am	7:49 am	100%	7:43 am
11 to 8	6:47 am	7:21 am	7:25 am	100%	7:23 am

Also, the preferred path, Route 11 to Route 8, has an expected arrival time downtown of 7:23, or 3 min after the scheduled time of 7:20 am. The earliest and latest observed arrival times on this path are 7:21 and 7:25 am, respectively, showing considerable reliability. For the other paths, however, one notes a much larger range in the distribution of arrival times, covering 12 minutes (7:26 to 7:38 am) for Route 34 to Route 6, and 14 to 15 minutes for the other paths using Route 11. This may significantly affect whether a traveler might choose one of these paths.

#### V. CONCLUSIONS

Vehicle location data can be used as a resource for a number of passenger information purposes. Particularly in cases where service reliability is a concern, providing passengers with credible information on the timeliness of service may improve the perceived quality of that service.

The methodology described in this paper provides a general framework for using AVL data to improve passenger information. More extensive experimental results are planned with the Tucson transit agency, particularly for demonstrating the computational aspects of the proposed algorithm.

#### REFERENCES

- [1] Robert Casey et al., *Advanced Public Transportation Systems: The State of the Art - Update 2000*, Federal Transit Administration, Report FTA-MA-26-7007-00-1, December 2000.
- [2] Dial, Robert B., "Transit pathfinder algorithm," *Highway Research Record* # 205, pp. 67-85, 1967.
- [3] Tong, C. O., and A. J. Richardson, "A computer model for finding the time-dependent minimum path in a transit system with fixed schedules," *Journal of Advanced Transportation*, Vol. 18, No. 2, pp. 145-161, 1984.
- [4] Bander, J., and C.C. White, "A new route optimization algorithm for rapid decision support," *Proceedings of the Vehicle Navigation and Information Systems Conference*, Part 2, P-253, pp. 709-728, 1991.
- [5] Koncz, N., J. Greenfold, and K. Mouskos, "A strategy for solving static multiple-optimal-path transit network problems," *ASCE Journal of Transportation Engineering*, Vol. 122, No. 3 (May/June), pp. 218-225, 1996.
- [6] Hall, Randolph W., "The fastest path through a network with random time-dependent travel times," *Transportation Science*, Vol. 20, No. 3, August 1986, pp. 182-188.
- [7] Wellman, Michael P., Matthew Ford, and Kenneth Larson, "Path planning under time-dependent uncertainty," *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence*, Montreal, Quebec, August 18-20, 1995, pp. 532-539.
- [8] Fishburn, Peter C., and Raymond G. Vickson, "Theoretical foundations of stochastic dominance," pp. 39-113 in Whitmore, G.A., and M.C. Findlay, *Stochastic Dominance*, Lexington (MA): Lexington Books, 1978.
- [9] Miller-Hooks, Elise, *Optimal routing in time-varying, stochastic networks: Algorithms and implementations*, PhD Dissertation, University of Texas at Austin, Department of Civil Engineering, 1997.
- [10] Miller-Hooks, Elise, and Hani Mahmassani, "On the generation of nondominated paths in stochastic, time-varying networks," *Proceedings of TRISTAN III*, Puerto Rico, June 1998.
- [11] \_\_\_\_\_, "Least expected time paths in stochastic, time-varying transportation networks," *Transportation Science*, Vol. 34, No. 2, May 2000, pp. 198-215.
- [12] \_\_\_\_\_, "Path comparisons for a priori and time-adaptive decisions in stochastic, time-varying networks," *European Journal of Operational Research* (in publication), 2002.
- [13] Hadar, Josef, and William Russell, "Rules for Ordering Uncertain Prospects," *American Economic Review*, Vol. 59, No. 1, 1969, pp. 25-34.
- [14] Chabini, Ismail, "Discrete dynamic shortest path problems in transportation applications: Complexity and algorithms with optimal run time," *Transportation Research Record* 1645, 1998, pp. 170-175.