**The Stochastic School Transportation Problem**

Submitted December 15, 2004

David Ripplinger
Small Urban and Rural Transit Center
Upper Great Plains Transportation Institute
North Dakota State University
430 IACC Building
P.O. Box 5074
Fargo, ND 58105
Phone (701) 231-5265
Fax (701) 231-1945
david.ripplinger@ndsu.nodak.edu

The efficient transportation of elementary and high school students within the United States has long been of interest to academics and practitioners alike. Research in the field has closely followed the development of solution techniques to general network routing problems by appropriately altering these models and the associated algorithms to address the subtleties of a particular bus routing problem. In this paper, a mathematical model that accommodates many of the sources and impacts of uncertainty is presented. Due to the impact of time windows on routing, traditional solution heuristics are unsuitable. However, a number of regular stochastic events that arise in the design and operation of pupil transportation are discussed and general solution outlines are presented.

The efficient transportation of elementary and high school students within the United States has long been of interest to academics and practitioners alike. Research in the field has closely followed the development of solution techniques to general network routing problems by appropriately altering these models and the associated algorithms to address the subtleties of a particular bus routing problem. Absent; however, is the impact of random events on system operation despite the fact that the generalized problem, referred to as the Stochastic Vehicle Routing Problem (SVRP), has received much attention in the literature. The SVRP has practical value where in many cases the application of heuristic solutions to the Vehicle Routing Problem (VRP) fare quite poorly when certain events occur. In the case of school transportation this might include fluctuations in student ridership, travel time, travel cost, and vehicle reliability.

## SOURCES OF UNCERTAINTY IN SCHOOL TRANSPORTATION

While uncertainty is often ignored in both the literature and application of vehicle routing problems, its impacts are very real and oftentimes quite significant. Research in the field saw great progress in the 1980's and 1990's primarily through theoretical papers such as those by Bertsimas (1,2) and Jaillet (3,4,5), which are summarized along with other contributions by Gendreau, Laporte, and Seguin (6). These papers focus specifically the optimal routing of vehicles when demand and customer presence is stochastic. Though research in the area continues, the theoretical rigor of the problems and computation demands of algorithms make complex problems very difficult if not impossible to solve.

This may explain why the literature on school bus routing has adapted other modern solution techniques such as simulated annealing (7), and tabu search(8), while ignoring the impact of uncertainty. It does clearly not result from the school transportation problem having

relatively less uncertainty or importance than other fields. In fact, as uncertainty often affects the level of service provided to pupils and the political ramifications that may result, the problem is certainly relevant.

As stated previously, the traditional stochastic vehicle routing problem literature focuses on the changing presence of customers and their level of demand. Parallel to this in the context of pupil transportation is realized ridership on a particular day. This could be alternatively stated as how many students, if any, are to be picked up or dropped off at a particular stop. Students may have irregular riding patterns due to reasons such as illness, weather, vacation, alternative transportation, truancy, student relocation, or behavioral problems. Many of these causes, though stochastic, may affect a large number of students similarly resulting in the individual realizations not being independent. This would be the case, for example, if a contagious disease affected attendance or if warm weather resulted in a large number of students deciding to walk to school rather than ride.

Though not part of the traditional stochastic vehicle routing problem, the cost to travel certain paths, which may be temporal or financial, may vary over time. Events such as the unexpected change in the cost of fuel or oil, the weather, or road conditions may affect the optimal solution to the problem, by changing the routes followed by a school district's vehicles. Some occurrences could be presumed to proportionally affect all routes evenly, as might be the case if fuel costs rose. On the other hand, construction on a particular street or highway might only have a local effect.

The final concern is that of vehicle reliability. Though built for hundreds of thousands of miles of service, vehicles often become inoperable and require the vehicle to be pulled from service for a considerable period of time for repair. Another possibility is that a breakdown

occurs while in a vehicle is in transit, a situation demanding that a substitute vehicle be readily available to complete the route. Situations like these are not uncommon and in areas where fleets have been aging they occur at increasing rates. In the next section, a generalized mathematical model representing the school transportation problem with uncertainty is presented.

## MATHEMATICAL MODEL

There are many possible sources of uncertainty in the school transportation problem, a single flexible model that adequately accounts for each random event would likely be large and cumbersome. In most cases, many of these stochastic events are extremely rare or have such a small impact that when they do occur they may be ignored. As few systems exist in the absence of uncertainty, a deterministic model can be viewed as an extreme case where either no uncertainty exists or is ignored altogether. Before the mathematical model is presented a generalized school transportation system is described.

The system consists of a network, students, vehicles, and schools. The directed network, G, consists of a set N of n nodes which may be student homes, schools, or intersections, and a set A of m undirected arcs, which are the streets and roads traversed by vehicles transporting students to and from school. Though important, it is assumed that students have been pre-assigned to stops. While this can be solved as part of a larger problem, it is often ignored, as it is here; due to the additional complexity and computational demands that result. The system also includes a set S consisting of s students each of which are assigned a stop $\pi(s)$, and school, $\sigma(s)$, and occupy one unit of space in his or her assigned vehicle. As stated in the previous section, realized ridership is a common source of uncertainty in school transportation.

The mathematical representation of the problem is described as follows:

Minimize $z = E_o\left[\sum_{i=0}^{n}\sum_{j=0}^{n}\tilde{c}_{ij}(\omega_t)x_{ij}^{k}(\omega_t)\right] + \sum_{k=1}^{K} c_k y_k$

(1)

Subject to

$$\sum_{k=1}^{K}\sum_{i=0}^{n} x_{i,j}^{k} = 1 \qquad \text{for j=1,…..,n} \qquad (2)$$

$$\sum_{i=0}^{n} x_{i,p}^{k} - \sum_{j=0}^{n} x_{p,j}^{k} = 0 \qquad \text{for k=1,…..,K}$$

$$\text{for p=1,…..,n} \qquad (3)$$

$$\sum_{s=1}^{S} o_{ij}^{s} \le c \qquad \forall k \in K$$

$$\text{for i,j=1,…..,n} \qquad (4)$$

$$\sum_{i=0}^{n}\sum_{j=0}^{n} t_{i,j}\delta_{i,j}^{s} \le t_{\max} \qquad \forall s \in S \qquad (5)$$

Where

$c_{ij}$ =cost to travel between nodes i and j

$$x_{ij}^{k} = \begin{cases} 1 & \text{if vehicle k travels between nodes i and j} \\ 0 & \text{otherwise} \end{cases}$$

$c_k$ =fixed cost of owning a vehicle k

$$y_k = \begin{cases} 1 & \text{if vehicle k is available} \\ 0 & \text{otherwise} \end{cases}$$

$t_{i,j}$ =time needed to travel between nodes i and j

$$x_{i,j}^{\pi} = \begin{cases} 1 & \text{if route travels between nodes i and j after visiting stop } \pi \\ 0 & \text{otherwise} \end{cases}$$

$$x_{i,j}^{k} = \begin{cases} 1 & \text{if a vehicle k travels between nodes i and j} \\ 0 & \text{otherwise} \end{cases}$$

$$x_{i,p}^{k} = \begin{cases} 1 & \text{if a vehicle k travels between nodes i and p} \\ 0 & \text{otherwise} \end{cases}$$

$$x_{p,j}^{k} = \begin{cases} 1 & \text{if a vehicle k travels between nodes p and j} \\ 0 & \text{otherwise} \end{cases}$$

$c^{k}$ = capacity of vehicle k

$$o_{ij}^{k} = \begin{cases} 1 & \text{if student s occupies a seat on vehicle k between nodes i and j} \\ 0 & \text{otherwise} \end{cases}$$

$t_{\max}$ =maximum student ride time

$$\delta_{i,j}^{s} = \begin{cases} 1 & \text{if student s travels from node i to node j} \\ 0 & \text{otherwise} \end{cases}$$

Before continuing, a brief review of probability space and the timing of decisions will be conducted to better understand the model at hand. A probability space is a triple $(\Omega, F, P)$ where Ω is the set of all possible outcomes, F is the set of events, and P is the probability that an event will occur. For example, let Ω represent the all possible combinations of temperature and precipitation for tomorrow. If the event of importance is if the temperature is below freezing then one would consider all outcomes where the temperature was at or below freezing. One would ignore if it was snowing or not.

With respect to timing, at time 0, which in our case could be considered the time when initial routes and other long term decisions are made, a decision maker knows the probability of a given event occurring. They will not know the actual outcome until the future at which time they may make changes to their plan, a process referred to in the literature as recourse. This could be the morning following a surprise snow storm which might affect travel times, and routes to be serviced.

Returning now to the model, the objective of the model is to minimize the expected cost of traversing the bus routes as well as the cost of having vehicles available as described in equation (1). It is assumed that both costs and routes traveled can be altered in the future so that their values are not known with certainty at the present. In the notation used, both the cost and routes are a function of the outcome, $\omega_t$, in $\Omega$, where the subscript t denotes the time. The first term calculates the expected cost of fleet operations by summing across the expected cost of a vehicle traveling between nodes i and j across all arcs, and vehicles. The second term in (1) calculates the availability cost of operating by summing across all vehicles, where it is assumed that these decisions are made and the associated costs incurred at time 0. Constraints (2) and (3) ensure that each stop is visited once only once and that a vehicle that arrives at a given stop also leaves from it. Stops have the flexibility of being added or removed in the future, so that unpopulated ones are avoided if they add to the cost of operation. There are also capacity and ride time constraints (4,5).

Availability of vehicles, accounted for in the second part of equation (1), is of importance as certain irregular events such as vehicles reaching capacity, running behind schedule, or breaking down may require replacements to satisfy the demands of the system. These replacements may be in the form of additional vehicles that are not regularly used or alternatively the cost of a contract to have replacements available at short notice. It may be the case that no event occurs such that replacement vehicles must be utilized. This does not; however, eliminate the sunk cost ofhaving them available.

It should be noted that the optimal value of the objective function is not guaranteed to be equal to the realized value. In robust models that account for a great number of sources of

uncertainty or that provide precise financial calculations, equivalent values would likely be as much due to coincidence as to any objective factor.

*Time Windows*

The impact of time windows on the school transportation problem makes it much different than other stochastic vehicle routing problems. Following the finalization of bus routes, pupils are notified of approximate pickup times. Currently, a large number of school districts send precise pickup times generated by commercial routing software to parents.

As with public transit, a primary measure of level of service in pupil transportation is time spent waiting at a stop. While buses will dwell to avoid passing a stop early, delays are regular occurrences. Prolonged delays may put the students at increased danger due to continued exposure to the elements or corrupt individuals. When delays are foreseen it may be possible to send out additional vehicles to maintain the integrity of the route or to contact parents to inform them of possible delays. The latter; however, may require a great deal of time, especially in larger districts. While such events are likely to occur at some point in time, they should be avoided due to the large political costs they demand.

Outside the realm of school transportation, rerouting vehicles daily or in real-time might be feasible. However, in this context dynamic programming such as that presented by Powell (7) would likely fail to meet the time window constraints. It would also impose tremendous costs in most cases. The most pragmatic solution of addressing the importance of providing service within the window is to focus primarily on ensuring the availability of an adequate number of vehicles and drivers to service the routes.

*System Breakdown*

Certain extreme circumstances may result in a breakdown of the system. In the context of the stochastic school transportation problem, the system is assumed to breakdown when any student is not picked up or dropped off within the time window. These events should be rare when there is an adequate solution to the transportation problem. Steps should be taken; however, to establish policies that help mitigate such situations before they occur, based on prior experiences. These steps should be part of the decision making process that occurs at time 0. For example, certain events that present a high likelihood of breakdown and that can be foreseen may merit cancellation of transportation service.

## PROBLEMS AND SOLUTION TECHNIQUES

When the impacts of time windows are taken into account, heuristic methods used to solve traditional stochastic vehicle routing problems are of no practical value. However, at least three options exist to aid managers in their pre-school year routing and vehicle procurement process. These include routing under capacity vehicles, partitioning the school year, and partitioning routes. Each technique has unique merits and alleviate the problem differently.

*Routing Under Capacity Vehicles*

A common method used to deal with volatility in ridership is to adjust the number of students assigned to stops or routes and thus vehicles. While this may result in vehicles operating under capacity and increasing costs, when factoring the political toll of system breakdown it may be the best management alternative. Returning to probability theory, this can be achieved by deciding

upon a tolerance, α, for the percentage of trips during which buses reaching capacity is acceptable, as expressed by equation (6). Since exceeding capacity is not acceptable, additional vehicles will be required to complete the route.

$$P(s_\beta > Cap_\beta) \leq \alpha \tag{6}$$

Though this technique has not been addressed in the literature, any attempts to properly solve the problem must maintain integer values of students. For example, if the probability of any student be present at a particular stop is .95 one cannot have 4.5 out of five possible students present. As a result of this requirement, the problem may best be addressed with Monte Carlo simulation. Also of note is the standard size of buses in the United States. Though they may be configured in a number of different ways, transportation managers may not be able to find buses with the precise number of seats that they desire. This is even more obvious in areas with shallow vehicle markets or for those districts facing financial constraints.

*Partitioning the School Year*

Another method to address uncertainty is to partition the school year into disjoint periods. This may be a practical alternative for school districts located in areas that experience inclement weather during discernable periods. Each period can then be solved using traditional heuristic methods. In parts of the United States that experience significant snowfall, buses could be routed differently for the early fall and late spring when the possibility of severe events is much less than during winter months. One real world challenge to this approach might arise in attempting to procure additional vehicles for high demand periods or dispensing of surplus ones

during periods of low demand. Consideration would also need to be given to the costs of determining the new routes and of notifying parents.

*Partitioning Routes*

When a bus reaches capacity while in route, it may be possible to send another to complete its tour within a time window. The practical challenge to such a technique is that a transportation manager is often required to 'balance' the lengths of rides. In the case of a dramatic increase in ridership on a given day, each route may have more students than seats on the vehicle assigned to it. This would result in the need for vehicles and drivers equal in number to those already employed, a clearly illogical possibility.

For the partitioning of routes to be an effective management tool, the initial routes must be designed with this feature in mind. For school districts with single schools, this might be accomplished quite easily by quickly transporting students to the school upon reaching capacity and then picking up the other students located close to the school, a condition that can be stipulated when developing the routes. For larger school districts, the staggering of school times may allow similar options.

*Registration and Ridership Fees*

A great deal of volatility in ridership could be managed by requiring registration of all riders before the school year begins when buses are routed and other one transportation management decisions are made. Alternatively, students could be charged a bus fee which may have the effect of increasing interest in school transportation management, accountability for riders, while raising additional funds for the school district. This fee would guarantee each student a seat on

the bus, perhaps only during certain parts of the school year. The added income could be used to defray the cost of often empty seats or for vehicles that would only utilized during certain times of the year.

## SUMMARY

Though uncertainty plays a strong role in the school transportation problem, the impacts of time constraints prevents the use or modification of existing methods to solve the problem. Though techniques using methods from operations research are described here, the cost of developing and implementing them may be cost prohibitive. For many school districts, the continuation of the status quo of either addressing the challenge of transporting students by routing students either by hand or by computer and then altering the deterministic results using subjective value judgments as to the actual operation of a system may be best.

**REFERENCES**

1. Bertsimas, D.J. "Probabilistic Combinatorial Optimization Problems." Operations Research Center Report. No. 193, Massachusetts Institute of Technology, 1988.

2. Bertsimas, D.J. "A Vehicle Routing Problem with Stochastic Demand." *Operations Research* 40 (1992): 574-585.

3. Jaillet, P. "Probabilistic Traveling Salesman Problem." Operations Research Center Report No. 185, Massachusetts Institute of Technology, 1985.

4. Jaillet, P. "Stochastic Routing Problems." *Stochastics in Combinatorial Optimization*. G. Andreatta, F. Mason, and P. Serafini, eds., pp. 197-213. Newark NJ: World Scientific, 1987.

5. Jaillet, P. "A priori Solution of a Traveling Salesman Problem in which a Random Subset of the Customers are Visited." *Operations Research* 36 (1988): 929-936.

6. Gendreau, M., Laporte, G., and R. Seguin, "Stochastic Vehicle Routing." *European Journal of Operations Research* 88 (1996): 3-12.

7. Powell, W.B. "Transportation and Logistics." *Handbook of Applied Optimization*, Panos Pardalos and Mauricio Resende, ed., Oxford, GB:: Oxford University Press, :679-688, 2002.