Proceedings of the 1992 Winter Simulation Conference
ed. J. J. Swain, D. Goldsman, R. C. Crain, and J. R. Wilson

TAXI MANAGEMENT AND ROUTE CONTROL: A SYSTEMS STUDY AND SIMULATION EXPERIMENT

William A. Bailey, Jr.

Belk College of Business
UNC at Charlotte
Charlotte, NC 28223

Thomas D. Clark, Jr.

College of Business
Florida State University
Tallahassee, FL 32306

ABSTRACT

The structure and behavior of a typical urban taxi system are addressed in this study. It is approached as a complex, first-order feedback system with network queuing structure. A computer simulation model of a characteristic municipal system is proposed and alternative taxi station and routing patterns addressed. An analysis of variance research design is employed to assess model results, and a method for stationing resources and controlling taxi movement is suggested. The simulation model is written in the Simulation Language for Alternative Modeling (SLAM).

1 PROBLEM DESCRIPTION

Historically, the taxicab industry has played an important but little recognized role in the provision of public transportation services in urban areas. In many small urban and rural areas, taxicabs have provided the only public transportation services available, and have been an essential service for many community residents without convenient access to an automobile (Kirby, 1981). The form of the taxi service is such that any member of the general public may hire a vehicle and driver to convey the person directly by road from one point to another that is specified by the passenger(s).

Taxicabs constitute a form of transportation called paratransit. Kirby et al (1974) define it as "those forms of intraurban passenger transportation which are available to the public, are distinct from conventional transit (scheduled bus and rail), and can operate over the highway and street system." It includes those types of transportation in-between the private automobile and conventional mass transit, such as car pools, rental cars, dial-a-ride, jitneys, limousines, and, of course, taxicabs.

Paratransit can be conveniently grouped into three categories: those you hire and drive, those you hail or phone, and those for which you make prior arrangements with other travelers. Hail or phone service is often called demand-responsive transportation (DRT) and includes taxicabs and dial-a-ride.

In the dial-a-ride problem, customers call a dispatcher or scheduler requesting service. Each customer specifies a distinct pickup and delivery point and a desired time for pickup and delivery. Vehicles (usually small busses or vans) are allowed to deviate from their direct route to serve other passengers. This type of system is usually called shared-ride transit (SRT) and is often classified as a vehicle routing and scheduling problem. This area has received considerable attention.

In a conventional taxicab system, once a vehicle picks up a customer, it proceeds non-stop to the customer's destination. This is often referred to as exclusive-ride transit (ERT). The operation of an exclusive-ride taxi system is not as easily classified as a scheduling and routing problem. It seems to better be represented as a queuing problem, specifically an M/G/s queuing problem. An M/G/s queuing problem is one in which calls arrive according to a Poisson distribution (M), calls are serviced in accordance with some general service distribution (G), and there are s parallel servers (taxis). Unfortunately, the complexity of the service distribution makes the problem too difficult to address using existing queuing theory unless very restrictive assumptions are made. Even if the service distribution could be derived, exact queuing formulae have not yet been developed for the M/G/s problem. Nonetheless, queuing concepts do provide a useful framework for discussing the problem.

A coherent, continuing stream of research that addresses the operation of a taxi fleet apparently does not exist. This is due primarily to the fact that the taxi industry is so fragmented and private, that it is often overlooked by government as a viable means of public transit. Thus, little funding for research has been provided, especially when compared with other transportation problems.

Only three pieces of research specifically addressing the operation of an exclusive-ride taxi fleet were identified. These are Meyer and Wolfe (1961), McLeod (1972), and Gerrard (1974). All of these attempt to confront the problem via the analytical/queuing approach by making certain restrictive assumptions, although, in two cases, simulation was used as a validation tool. McLeod's (1972) contribution is the most extensive to date.

The major operational control function of a taxicab system is dispatching, which involves the activities of demand forecasting, taxicab/labor scheduling, telephone answering, information storing, communications distribution, and vehicle operation. This group as a whole constitutes the dispatching function. A number of issues stemming from the system's inherent complexity also must be addressed and decisions relative to them made prior to a dispatching operation. For example:

- How many cabs should be in operation at any point in time (fleet size)?
- Should a cab return to "base" after serving a customer? Should it remain at a given location? Should it relocate to a position where another call may be more likely?
- What is the expected frequency of calls from various parts of the community at various times of the day and week?
- Should a customer be assigned only to idle cabs or would it be more efficient to assign the person to a cab currently in service but in a position to respond more quickly?

One purpose of the research presented in this paper is to develop a method to address these questions and others with the intent of producing heuristic-based solutions to the operational problem. The research produced a simulation model of a taxi system that can be implemented under various operating constraints. A second important goal is to illustrate a viable method for approaching a problem for which management science has not produced a solution. A structural model of the system under study will be presented first, followed by a description of the problem and the experimental design used to address it. A summary of the computerized parametric model will be presented to describe how the system has been simulated, and to delineate the problems of verification and validation in this and similar systems. Finally, the statistical results of the simulation will be presented to describe model operation and support conclusions concerning improved system behavior.

2 SYSTEM STRUCTURE

In order to focus first on just the variable structure of the system, and not the complexity of the elements, interactions, and states of the taxi system itself, a black box model of the taxi system is presented in Figure 1.

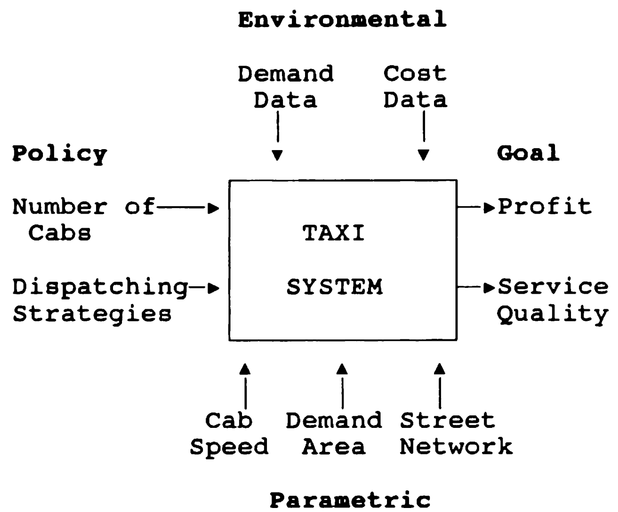

Figure 1: Variable Structure

There are four sets of variables. The goal variables describe the two objectives of the taxi system: maximize profitability and service quality (measured by customer waiting time). These two goals are partially conflicting. The policy variables specify the means available to control the operation of the taxi fleet. These variables will be manipulated in order to identify the "optimal" combination with respect to the performance measures (goal variables). The environmental variables identify the various conditions under which the policies are tested. Finally, the parameters describe the constants under which the system will operate. These values are assumed to be fixed and provide the necessary infrastructure in order to experiment with the model.

The state diagram in Figure 2 illustrates the important events that occur during system activity and clearly identifies the customer waiting time objective. From a queuing theory perspective, "waiting time" is identical to "response delay," since, technically, the cab begins "service" as soon as the response is initiated. Therefore "service time" is the sum of "response time" and "transit time." However, from the customer's (and taxi dispatcher's) perspective, service does not begin until the customer is picked up. Therefore, "waiting time" is the sum of "response delay" and "response time." this is a major reason why queuing theory cannot effectively address the taxi problem; "waiting time" has a different meaning.

```
Customer
  View ┌───waiting time────┐
        response  response  transit
         delay      time     time
TIME=▶•──────────▶•──────────▶•──────────▶•
      t0          t1         t2         t3
     cust        cab        cust       cust
     calls     responds    pickup     deliv
Queuing          └────service time────┘
 View
```

Figure 2: System State Diagram

3 EXPERIMENTAL DESIGN

The variable structure and system state diagram form the basis for parametric model development and suggest the variables of importance for a research design to control experimentation with the model. The objective of the simulation model is to provide data that will indicate which policies are best either to increase profitability or decrease customer waiting time. The research design focuses on the policy variables (factors) under management's control, stochastic or environmental variables (blocks), and goal variables (performance measures) with analysis directed toward the combination of various policies that produce the most favorable output measures. The design of an experiment that focuses on the two policy variables of fleet size (number of cabs) and dispatching strategy will be discussed.

3.1 Dispatching Strategies

The development of a dispatching strategy requires at least four decisions to be addressed. These decisions concern cab selection, cab relocation, cab eligibility, and customer selection and revolve around answers to several questions. After delivering a customer, should the cab: 1) return to a centralized "base", 2) remain at the delivery location, or 3) relocate to another "optimal" location based on some relocation algorithm? When a call arrives, should it be assigned: 1) to only idle cabs, or 2) to any cab (idle or busy)? Five combinations of these options were selected for the purpose of experimental analysis. These are summarized in Table 1.

Table 1: Dispatching Policies

Policy	Cab Location	Cab Selection
1	Return to Base	Idle Cabs Only
2	Remain	Idle Cabs Only
3	Relocate	Idle Cabs Only
4	Remain	All Cabs
5	Relocate	All Cabs

3.2 Fleet Size

The number of cabs creates another dimension and can be any integer value, but to keep the combinations to a reasonable number, the fleet sizes used in the study were 25, 30, 35, and 40. The range of numbers was chosen based on data from a 1985 International Taxicab Association survey of 141 taxicab companies (1986). The survey showed that 71% of the cab companies operate with an average fleet size of 18 to 30 cabs. The figures used in the study are consistent with the reported data and represent, therefore, a range of reasonable alternatives. The actual numbers for a given company would be determined by the interaction between a desired level of customer service and the associated costs.

3.3 ANOVA Design

The four fleet sizes combined with the five dispatching strategies yields and design matrix that contains 20 cells (4x5). A two way analysis of variance (ANOVA) with five treatments (dispatching strategy) across four levels (fleet size) is used to test this design. Each of the twenty factor combinations were samples ten times giving a total of 200 independent replications of the simulation. The parametric statistical model, therefore, is:

$$Y_{ijr} = \mu + C_i + N_j + CN_{ij} + \epsilon_{ijr}$$

where
$i = 1,2,3,4,5$ and
$j = 1,2,3,4$ and
$r = 1,2,3,4,5,6,7,8,9,10$

Y_{ijr} is the observation of the response variable Y for replication r of level i of the first factor (policy) and level j of the second factor (fleet size).

A completely randomized design resulted as all replications both within a treatment and over all treatments are made statistically independent by use of correlated sampling through control of the random number streams. The null hypothesis is that there is no difference in output from the various taxi policies (while fixing the effect of the number of cabs) versus the alternative hypothesis that at least one policy is different. If the null hypothesis is false (which is almost surely has to be), then Tukey's multiple comparisons test may be used to detect which policy is significantly different from the others.

Correlated sampling for each factor combination creates the same interarrival times, customer origins and destinations. All experiments, therefore, will yield the same average trip distances and times. It seems clear that, in general, the policy which results in the quickest response time also will minimize the distance traveled without a customer (when no revenue is being earned). Since revenue is held constant across all factors, profit is maximized when costs are minimized. Costs are minimized (given a fixed number of cabs) when the distance traveled in response to a call is minimized since no revenue is generated until the customer is picked up.

One problem persists, however, as it is conceivable that the number of taxis can be reduced with little effect on response time (the time between when the cab begins to respond to a call and the time it picks the customer up) but with a large effect on customer waiting time because of the increase in response delay (the time between when a call arrives and the cab can begin responding). See Figure 2 again. In other words, an increase in the fleet size results in system congestion, but since response time is measured from the time a taxi begins responding to a call and not when the call arrives, it may be affected only slightly by the congestion (in fact it will probably go down), yet a considerable amount of time may go by before a cab begins responding, especially if all cabs are busy. Although this may be of no concern to the taxi manager in the short run, damage to the reputation and profitability of the company could occur in the long run.

The simulation model was operated under two different incoming call distribution. Under one scenario, the incoming calls were exponentially distributed with a constant mean of thirty minutes between calls in each of the 25 operating demand zones included in the model. This results in 50 calls per hour in the system or about .83 per minute. This demand pattern is maintained over the entire length of the simulation. Under the second scenario, the incoming calls were exponentially distributed with a nonstationary mean that is a random number between 0 and 60 minutes in each zone and is altered every three hours. This results in a different demand rate in each of the 25 zones. These two conditions represent extremes of the demand distribution spectrum. If the same policy is best under both conditions, one may assume that it is robust under a variety of distributional demand conditions.

4 PARAMETRIC SIMULATION MODEL

The FORTRAN-based, discrete event features of the SLAMII simulation language were used for the mathematical model of the system. The model is composed of ten subroutines that generate and manage system flow. (A detailed description of the model and a program listing are available upon request.) The model has a discrete event structure with an event calendar controlling its behavior. The three major events are a request for service, allocation of a cab to a customer, and completion of service.

As previously mentioned, the area serviced by the taxi system is divided into a grid-like structure of 25 operating zones. Each zone represents an identifiable market where calls are generated using an appropriate random variate generator.

Each customer is assigned a destination and only one cab is selected to provide service. The closest cab when the call is received is always selected. If only idle cabs are eligible and there are not idle cabs available when a call arrives, then the first idle cab is selected.

For each experiment, the model is operated to simulate one week of service for the 25 zones that form the grid of the referent city. The call arrival distribution discussed previously are employed. When cabs travel from zone to zone, the distance is based on the rectilinear or "Manhattan metric" distance between the zones (as opposed to the linear distance). The point of origination and destination can be anywhere within each of the four square mile (2x2) zones. The actual travel distance is computed using a random variate with a triangular distribution with parameters dependent on the nearest and farthest points between the two zones. It can be shown that this distribution provides a reasonable approximation of cab travel distances. In order to simplify time computations, the average rate of speed between all zones is uniform random variable distributed between 20 and 30 miles per hour. Finally, all cabs are assumed to directly transport the customer without any intermediate delay.

5 EXPERIMENTAL RESULTS

Due to space restrictions the tabulated results are not presented, but the basic conclusions can be summarized.

First, as expected, the treatment effect of dispatching strategy and fleet size were highly significant under all operating scenarios. In each case, there was also a significant interaction effect among the strategies and fleet size, making interpretation of the results more difficult.

To focus more clearly on the difference between policies, therefore, a series of one-way ANOVA tests were performed. These tests supported the conclusion that there was a significant difference between the dispatching strategies but that differences were less discernible as fleet size increased.

To determine exactly which strategies were different from others, Tukey's multiple comparisons test was employed at the 1% significance level. Strategy 1 (cab returns to base; idle cabs only) was as bad or worse than all other strategies at all times. This was expected. Strategy 4 (remain at drop-off location; all cabs eligible) was as good as or better than any other strategy at all times. Strategy 5 (relocate to high demand areas) was not as effective as anticipated and, in fact, it appeared to rank worse as the fleet size increased. A significant improvement (>50%) occurred when the number of cabs increased from 25 to 30; however, the improvement after that was slight. If a critical minimum number of cabs for adequate service is not available, it appears that the cab allocation policy (dispatching strategy) chosen makes little difference.

Having identified strategy 4 as superior, a second stage of the analysis was performed to identify the profit maximizing fleet size for this strategy. Profit is computed by subtracting costs from revenue. Revenue can be approximated by total service miles multiplied by some average price per mile. Costs can be represented by total miles multiplied by some average expense per mile (gas, maintenance, and so forth) plus a fixed capital cost per taxi. The mileage totals for each of the ten simulation runs for strategy 4 were averaged for each of the four different fleet sizes.

When only 25 cabs were employed, all calls could not be serviced, so the system was unstable (i.e. utilization was greater than 1). Stability was achieved when 30 or more cabs were utilized. As fleet size increased, the average distance required to respond to any call clearly is reduces due the greater dispersion of cabs throughout the region, thereby increasing the likelihood of a cab being nearer to a call. This reduces overall variable costs. The capital costs associated with the purchase (or rental) of each additional cab and the labor costs incurred as more cabs are employed somewhat offset these savings. The cost minimization problem, therefore, involves a tradeoff between operating, capital, and labor costs.

The average total mileage values are displayed in Table 2 and were used to estimate the total mileage function using a simple linear regression. Since there is clearly an inverse relationship between total mileage (y) and the fleet size (x), a reasonable regression model would be:

$$y = f(x) = \alpha_0 + \alpha_1/x^n + \epsilon$$

The value for n must be determined by experimentation. Using least-squares estimation and the data in Table 2, the following equation yielded the best results:

$$y = f(x) = 85813.22 + 3.96023\text{x}10^{12}/x^6 \quad (R^2=99.7\%)$$

The data associated with 25 cabs was included in the regression since total mileage increases rapidly as fleet size in reduces from 30 to 25. Leaving it out would have resulted in an equation which would have underestimated mileage at fleet sizes below 30.

Table 2: Mileage Results for Strategy 4

Fleet Size	Avg. Total Service Mi.	Avg. Total Resp. Mi.	Avg. Total Mileage
25	52480.5	49671.5	102152.0
30	58181.2	32561.3	90742.5
35	58193.7	30154.3	88348.0
40	58197.8	28587.3	86785.1

Thus the cost minimization problem can be formulated as:

$$\begin{aligned} \text{Min } k(x) &= K_V f(x) + (k_C+k_L)x \\ &= k_V(\alpha_0+\alpha_1/x^6) + (k_C+k_L)x \\ &= k_V\alpha_0 + k_V\alpha_1/x^6 + (k_C+k_L)x \\ &= 85813.22k_V+3.90623\text{x}10^{12}k_V/x^6+(k_C+k_L)x \end{aligned}$$

where

$k(x)$ = total cost (\$/week)
k_V = variable cost (\$/mile)
k_C = fixed capital cost (\$/cab/week)
k_L = labor cost (\$/cab/week)

This formulation is similar in structure to the classical economic order quantity (EOQ) problem, and can be solved by setting the first derivative of k(x) to zero and solving for x as follows:

$$k'(x) = -6k_V\alpha_1/x^7 + (k_C+k_L) = 0$$

$$x^* = \sqrt[7]{\frac{6k_V\alpha_1}{k_C+k_L}}$$

Note: Since fleet size must have an integer value, x^* should be adjusted to the closest integer yielding the smallest cost.

A simple spreadsheet table could be created to identify the optimal fleet size (x^*) for various combinations of k_V, k_L, and k_C. The values for k_C could be based on a rental rate or a capitalized cost for owned vehicles. This methodology allows the identification of the best dispatching strategy and fleet size for a given cost structure

6 SUMMARY AND CONCLUSIONS

An approach to analysis of the dispatch process in a taxi cab company has been presented. The structure of a typical system was illustrated and a parametric model of it explained. Such a model appears to be a useful tool in analyzing the complex structures evident in taxi systems and in defining dominant policies for system operation. The methodology was illustrated using one particular theoretical system.

As illustrated, computer simulation is an appropriate and useful tool that can aid in improving efficiency in such systems. Since the cab dispatch and structure problem has received little attention in the management science literature, and an analytical solution to even some simple but similar problems has proven difficult, the approach presented in this paper takes a positive step towards providing a useful methodology. The key approach is development of a statistical research design to control policy experimentation.

The system structure is adequately captured by the model described in the paper. Some of the assumptions used were restrictive enough to limit general application in its current form, but the results indicate that the essence of actual systems is captured. For extension to other systems, several steps are necessary. First, a need exists for accurate taxi system data that can be used to derive the interarrival time distributions and to provide a statistical basis for empirical validation. This data collection effort must be extended over a long enough period to adequately represent the demand from season to season. Second, a procedure is needed for determining the optimal size and number of zones used to represent the service territory. Zone structure is a function of the of the transportation infrastructure in give cities. A methodology to classify the types of infrastructure would be valuable.

It seems that a reasonable goal for this line of research would be the development of a decision support system that could aid a dispatcher in assigning cabs to calls and allocating cabs to zones in an efficient and profitable fashion. The modeling and experimental approach presented in the paper is offered as a start in this direction.

REFERENCES

Gerrard, M. J. 1974. Comparison of Taxi and Dial-A-Bus Services. *Transportation Science* 8:2.

Kirby, R. F., K. V. Bhatt, M. A. Kemp, R. G. McGilliway, and M. Wohl. 1974. *Paratransit: Neglected Options for Urban Mobility* Washington, D. C.: Urban Institute.

McLeod, M. G. 1972. *The Operation and Performance of a Taxi Fleet* Master's Thesis. Department of Civil Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts.

Meyer, R. F. and H. B. Wolfe. 1961. The Organization and Operation of a Taxi Fleet. *Naval Research Logistics Quarterly* 8:2.

AUTHOR BIOGRAPHIES

WILLIAM A. BAILEY, JR. is an Assistant Professor in the Department of Management Information Systems and Operations Management at the University of North Carolina at Charlotte. He received a B.S. degree with majors in computer science and statistics from Florida State University in 1981, and he received a Ph.D. degree in decision sciences from Florida State University in 1989. His research interests are focussed in the area of computer simulation and decision support systems.

THOMAS D. CLARK, JR. is a Full Professor and Head of the Department of Information and Management Science at Florida State University. He has a business degree from Emory University and a DBA from Florida State University.