

Assisting a Public Transit Agency with Planning of Bus Schedules Using Computer Simulation: A Collaborative Project between Cornell University and Tompkins Consolidated Area Transit (TCAT) in Ithaca, NY

Nirav Shah

Operations Research and Industrial Engineering
Cornell University
204 Rhodes Hall
Ithaca, NY 14853

Jonathan Helm

Operations Research and Industrial Engineering
Cornell University
204 Rhodes Hall
Ithaca, NY 14853

Abha Dubey

Operations Research and Industrial Engineering
Cornell University
204 Rhodes Hall
Ithaca, NY 14853

Wenshan Xu

Operations Research and Industrial Engineering
Cornell University
204 Rhodes Hall
Ithaca, NY 14853

Francis Vanek

Civil & Environmental Engineering
Cornell University
220 Hollister Hall
Ithaca, NY 14853-3501
Corresponding Author (Fmv3@cornell.edu)

Rod Ghearing

Tompkins Consolidated Area Transit (TCAT)
737 Willow Ave.
Ithaca, NY, 14850
Currently with Grand Valley Transit, Grand Junction, CO

Dwight Mengel

Tompkins Consolidated Area Transit (TCAT)
737 Willow Ave
Ithaca, NY, 14850
Currently with City of Ithaca Department of Social Services, Ithaca, NY

Abstract -- In this paper we present results from a service learning project carried out by four Master of Engineering students in the School of Operations Research and Industrial Engineering at Cornell University during the 2003-2004 academic year, on behalf of the local transit agency, TCAT. The project participants developed a simulation model to evaluate schedule changes and applied it to a proposed shortening of the time between bus arrivals, or “headway”, on TCAT’s Route 81 circulator bus that serves the Cornell campus in Ithaca, NY. The model was developed by adapting a commercial simulation software package called ProModel that is usually used to simulate the layout of manufacturing facilities. Use of the simulation model helped TCAT planners better understand the implications of the proposed schedule change to segment-by-segment passenger counts at stops and on-board vehicles. Based on our experience, we discuss a number of issues both with the development of simulation models for transit operations and the practice of collaboration between university students and faculty and public transit agencies.

Index terms – computer simulation, geocoded data analysis, transit system performance, university and public agency collaboration

INTRODUCTION

Planning of both new and existing transit routes is by its nature a complex undertaking. The transit agency must consider the level of passenger demand that varies by time of day, day of week, season, weather conditions, etc., as well as the relative cost of adding new service versus the increased revenue obtained from attracting new riders, among other things. Use of computer software, from spreadsheets to optimization or simulation packages, to support this planning process, is therefore an attractive proposition.

In this paper we consider the application of one approach in particular, simulation, using a case study of developing and applying a simulation model to a proposed service change to a specific transit route in the network of Tompkins Consolidated Area Transit (TCAT) in Ithaca, NY. TCAT was formed in 1998 from three transit agencies that served the region, including Cornell University Transit, which specifically served the Cornell campus and immediate surroundings. Thus there was already a close connection between the University and the transit agency, which facilitated the creation of a service learning project. The work in this paper builds on projects carried during the 2001-2002 and 2002-2003 academic years, also involving teams of Master of Engineering (M.Eng) students working with TCAT to use Information Technology (IT) to help improve the level of service.

The parts of the paper include 1) background on simulation and a review of other applications in the transit industry, 2) the development of the simulation program for TCAT, 3) presentation of results from the simulation model as applied to TCAT, and 4) a discussion of observations about applying simulation to transit.

BACKGROUND

Although the term “simulation” can be applied in a number of different contexts, in this paper it refers to a computer model that reproduces the movement of buses along a route, incorporating scheduled movements of buses and the boarding and alighting of passengers. Some aspects of the simulation are random or “stochastic”, such as the arrival of passengers at a stop in the model; these activities are generated using a statistical distribution for the mean and variance of the number of activities in a given time period. The simulation can be carried out without any use of graphics, i.e. input is given to piece of computer code, which then runs and outputs certain measure of performance that are of interest. However, the use of a graphical “animation” of the simulation, as demonstrated in this project, can facilitate both the design of the simulation model and the interpretation of its function and output.

Two recent advances in technology and computing have the potential to facilitate the use of simulation models of transit operations. The first is the Global Positioning System, or GPS, device. This device makes it possible to give the location of a bus as it travels along its route to within a few meters. On-board GPS devices can be added to individual vehicles in a bus fleet for a modest cost compared to the capital cost of the vehicle itself. With the GPS device in place, it is possible to record the location of each transaction in the fare collection device, allowing the operator to accumulate data about how many passengers are boarding at each stop along a route by time of day, along with the type of fare payment and other attributes. Thus a detailed travel

behavior data set becomes available to the operator for a modest cost. This use of “archival” geographic data is distinct from the use of GPS devices to give the location of buses in real time; the latter use of GPS is outside the scope of this paper.

A second development is the emergence of new simulation software packages that facilitate the creation, validation, and application of simulation modeling by using graphical user interfaces, pre-packaged statistical distributions, and flexible import/export functions to streamline the simulation modeling process. We use one such manufacturing software package, ProModel, for the model used by TCAT. The simulation software has applications in a number of industries. One such industry is manufacturing, where industrial engineers can use simulation packages specifically designed for their field to lay out floor plans for manufacturing plants. To simplify the development of a simulation model, the software packages have built-in objects that represent models of manufacturing equipment, which can be pasted into a graphical interface. For our work, we have adapted the standard factory layout map in ProModel so that it instead presents a map of the portion of the City of Ithaca where the bus routes of interest are located.

Literature Review

A review of the literature and the internet revealed a limited number of applications of simulation modeling to transit planning, with none replicating exactly the work that was carried out in this project. For example, Wong et al (2004) carried out a simulation that focused on the effect of bus reliability and bus failures during service on overall system function¹. Ding et al (2000) carried out a simulation test of a transit route in the state of New Jersey to verify that the simulation model could accurately reproduce real-life bus operations². Morgan (2002) developed a microscopic simulation model of transit operations for use in evaluating of Advanced Public Transportation Systems (APTS)³. Also, a model was developed in order to assist the city of Rome, Italy, with the development of a nighttime public transit service to alleviate traffic congestion in the city center, including convenient frequency, number of vehicles, number of customers transported, and expected revenues⁴.

Other studies in the literature are closer to our work in that they focus specifically on the interaction between quantity of passenger, frequency of service, and overall service quality. In the first example, a simulation model was developed for Middle Tennessee State University in the USA in order to determine the optimal number of buses for a new route that was to be added to an existing system with one route⁵. A second example is a model of the movement of pilgrims by bus between two locations along the traditional route of the Hajj pilgrimage in Saudi Arabia, which like our model uses ProModel as a basis, and evaluates total time for moving all passengers, average length of trip, and congestion level on buses⁶. Lastly, Zargari and Khan simulate a bus “Transitway” (i.e., corridor available only to buses for more rapid service) in order to predict the level of performance, with a specific interest in transit time on the Transitway⁷.

We conclude from the literature review that while there have been some papers on applying simulation to transit systems, they are much fewer than the number of papers on simulation applied to traffic flow generally, and that our work makes a unique contribution in a number of ways. First, we use geocoded boarding data as an input to creating statistical distributions to represent passenger boarding patterns; this approach has the potential to more accurately represent where and when passengers board, and we did not encounter it in the literature review. Secondly, we encountered only one other example of the use of ProModel to simulate a transit

system, so we believe that more experience can be gained by continuing to work with this package and explore its potential. The use of this software or other non-transportation simulation packages, may provide a lower cost solution than some of the commercial transportation simulation packages, which are quite expensive and therefore out of reach of many transit agencies. Lastly, our paper follows the complete life cycle of the simulation project, from model development, to recommendations, to implementation of the recommendations, to ex post facto evaluation of whether the recommendations were correct. We did not encounter another such paper related to transit system simulation.

Reflecting on the general scarcity of published work in this area, it appears that transit agencies are not using simulation as a tool to the extent that other entities that manage complex systems (highway departments, manufacturers, freight logistics providers, etc.). A possible explanation for this situation is that transit agencies often have limited budgets for adopting information technology, and many have not been in the position to make the investments in hardware or software beyond the need to pay for buses, fuel, wages, and other essential costs. Yet, focused application of engineering skills combined with these new technologies can, over time, pay for itself in terms of reduced cost and increased revenues. A goal of this paper is therefore to help build up the body of literature on simulation of transit that will lead to a more modern transit industry.

Cornell-TCAT Collaboration as a Service Learning Opportunity

In conclusion to this section, it is notable that the collaboration between TCAT and the Cornell School of OR & IE provided an excellent opportunity for service learning. In our view, the transit agency delivers a valuable social service, in terms of reducing traffic on the roads, saving energy, and especially giving mobility to the disabled, elderly, young, and those of low income who cannot afford a car. Thus the students worked through the transit agency in this project to help protect the environment and deliver better service conditions to the disadvantaged in the community by helping to improve bus system performance (= "service"), and at the same time developed new engineering knowledge (= "learning").

The management of TCAT was challenged by the need to understand the implications of operating decisions on transit service quality, and was aware of the potential value of the geocoded boarding data and the potential for using computer-based analysis to model the system in order to better understand it. At the same time, they did not have the financial means to pay professional consultants to carry out a simulation analysis, and they also did not have the skills in-house or the available person-hours to carry out such a project. The Cornell students who joined the project (Nirav, Jonathan, Abha, and Wenshan) not only had the skills and the motivation to apply their learning to a real-world problem, but they and their peers are themselves users of the bus system and were able to experience first-hand its strengths and weaknesses. During the course of the project, the students took advantage of opportunities to improve their skills in refining recommendations with the partner organization, and presenting their findings in both a public final presentation and a full technical report. In discussions with the project advisor (Francis) at the end of the project, the students expressed their satisfaction with being able to apply their operations research and computer programming skills to a real-world problem and to have real impact, but also to have the chance to gain experience interacting with a partner outside of Cornell, as represented by the TCAT General Manager and Service Development Manager (Rod and Dwight).

TCAT CASE STUDY

TCAT is a small- to medium-sized operator, serving approximately 2.5 million passenger boardings per year using 60 vehicles (a mixture of full-sized and minibuses) on 41 routes. The development of a simulation model came out of TCAT's interest in testing the potential effect of reducing headways (i.e., time between consecutive arrivals of a given bus route at a stop) on bus crowding on a particular route, namely the Route 81 serving the Cornell University campus. TCAT also had a longer-term interest in making greater use of geo-coded passenger data and in developing a capacity to use simulation modeling. Route 81 travels between two satellite parking lots on the edge of the campus. It runs along main streets through the campus center, thus serving both university faculty and staff traveling from parking lots and students who live in peripheral dormitories. In a second phase of the model, Route 81 was modeled jointly with Route 30, which connects the Cornell campus with the Ithaca city center in one direction and a retail area in the other, so the latter is considered during the model development as well.

Development of Boarding Data Distributions for Simulation Model from GPS Data

In order to run the simulation model used in this project, the modeler requires randomly generated passengers for each stop and for each run of the bus. This in turn requires the development of a database, from the boarding of the passenger to the archiving of raw GPS data to the creation of a data set that is readable by the simulation software. The following steps are required:

1. Passenger boards a bus
2. Fare collection system receives GPS signal and records geocoded boarding data
3. Cumulative geocoded data is downloaded from bus to IT system at end of day
4. Archived geocoded data is processed into boarding data by date, time, fare type, route, and location
5. For each data point, location is converted from longitude and latitude to association with specific stop
6. Statistical distributions generated from boarding data by stop for uploading into simulation model

For step 5, contiguous boxes were drawn around the bus stops on the route denoting the region associated with that stop. All latitude and longitude coordinates falling within a given stop's boundaries were assigned as passengers having boarded the bus at that stop. Variation in boarding location may be due to the fact that the GPS data is only precise to within a few meters. In addition, the bus driver may only record an arrival after the bus has started moving although the actual arrival occurred while the bus was at the stop. The map of the route with stop boundaries is shown in Fig.1. Each zone is associated with a single stop in each direction on the route. The passenger boardings geocoded within the boundaries of the zone are considered to have occurred at that stop. The black dots on the map represent major landmarks on the Cornell campus.

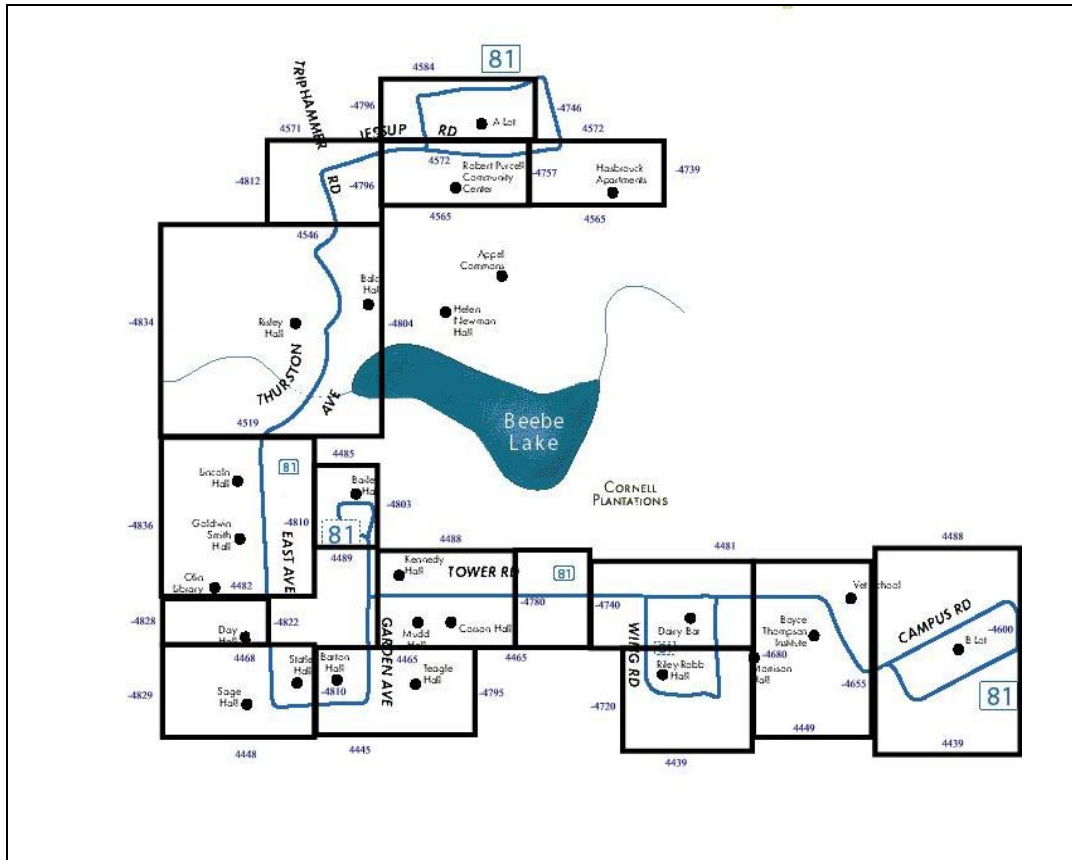


FIGURE 1

DIVISION OF ROUTE INTO ZONES FOR ASSOCIATING LONGITUDE/LATITUDE DATA WITH STOPS

Three months of data were used to give sufficient information regarding the number of passengers boarding the bus at a given time, day of the week, and bus stop, to be able to estimate statistical distributions for boardings using a spreadsheet program. A table of statistics was then exported into the simulation. For this model, a curve was first fit to the distribution of passengers over the course of the day, and then the total number of passengers for the stop was generated from a Poisson distribution, such that the mean and variability of passenger boardings matched the observed behavior from the GPS data. The data requirements of the simulation software package favored this approach, although in another environment it would be possible to estimate the distribution for each stop and time of day in isolation. As an example, Fig.2 gives the average Monday boardings per hour at a peripheral stop on the route (Parking lot “B”); the stop serves many inbound staff members on their trip to work, so demand is highest in the early hours.

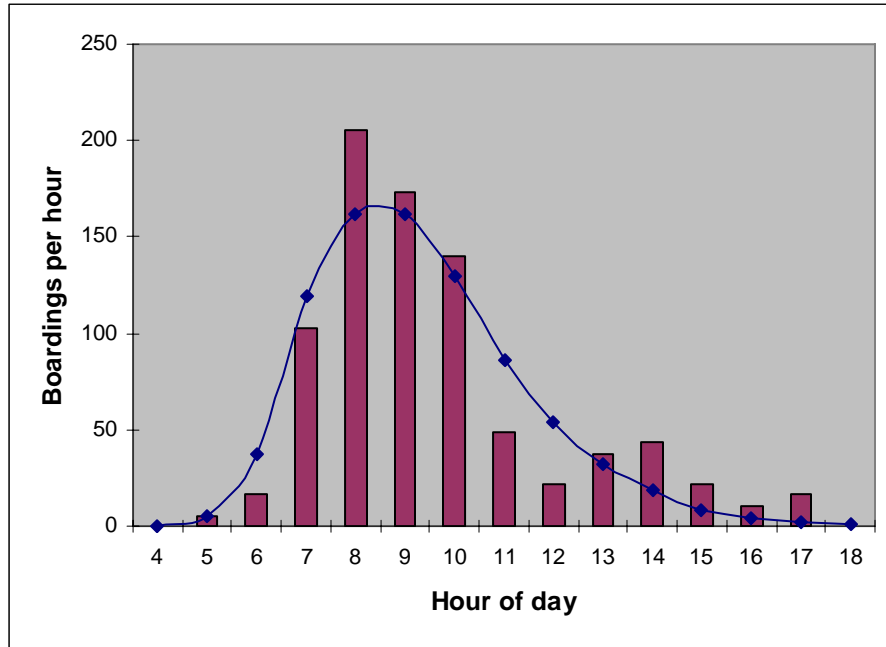


FIGURE 2
PASSENGER ARRIVAL DISTRIBUTION FOR B LOT STOP ON MONDAYS

Manual Gathering of Alighting Data

Because electronic alighting data were not available and TCAT keeps no record of passenger alightings for this route, we collected a small quantity of alighting data manually. Alighting data were gathered during 4 time slots:

- 7:30 am - 10.15 am (Morning)
- 10.15 am - 2.30 pm (Mid-day)
- 2.30 pm - 4.15 pm (Afternoon)
- 4.15 pm - 6.30 pm (Evening)

The “morning”, “afternoon”, and “evening” slots are considered to be “peak” periods, meaning that the number of passengers riding is at or near the maximum amount; other periods are called “off-peak.” For each time slot, at least 5 runs were carried out on the entire length of Route 81 ensuring that for every time slot, there would be at least 5 different sets of data for each bus stop. For Route 30, data samples were collected only for stops that the two routes had in common. Route 81 actually starts operation at 4:30 A.M, however, no alighting data was gathered during this period so a pattern similar to other off-peak periods was assumed.

Alighting patterns are measured in the model in terms of percentages of riders on board, not absolute values. The patterns were assumed to be the same for any bus traveling within a given time slot, regardless of the exact time at which the bus was running. Also, no distinction was made in the alighting pattern regarding day of the week. In other words, the number of boardings and hence average number of passengers on board might rise or fall depending on the day of the week or the time of the run within a given time slot, but the fraction alighting at each

stop would remain the same. This assumption was necessary since otherwise the alighting data gathering requirement would have greatly increased. However, it makes the simulation less realistic, since in fact the possibility exists that certain destinations, such as the gym or library, have a stronger attraction at certain exact times or on certain days of the week, and therefore change the alighting pattern.

The volume of passengers is also usually larger in the middle of the week, so the alighting data is used in the model

Other Modeling Considerations

Bus movements are deterministic in the model, as traffic congestion and delays are in general not a significant factor for the Route 81. (Delays do occur around the times that students change class on campus, however this consideration was deferred for future research.) Dwell time due to passenger boarding and alighting is also fixed. If the model were applied to a different route in the future, either of these parameters could be randomized.

Passengers are modeled as a total number of persons either waiting at a stop or on board a bus, rather than being modeled discretely. There is currently no origin-destination data available for specific passenger trips, such as someone who boarded at stop A at time x and alighted at stop B at time y , which would allow simulation and tracking of unique passengers.

Bus capacity, i.e., the number at which the bus is considered to be full and the model no longer allows additional passengers on, is a random variable in the model, rather than using the rated capacity of the bus. Even though the rated capacity maybe be X , a bus driver cannot track the exact number of passengers on the bus relative to the rated capacity, and therefore must use discretion to determine when a bus is full. A bus may also be declared full by the driver even though not every seat or aisle space is occupied.

RESULTS FROM MODELING THE ROUTE 81

Validation of Route 81 Using Existing Schedule

The model was first validated by simulating the Route 81 with the existing schedule and confirming that the simulated route replicated the characteristics of the actual route. The results showed that the simulation was accurate, i.e. an average of many runs of the simulation reported approximately the same number of boardings as was observed by TCAT in their actual boarding data. Also, from observation of the animation, the behavior of the simulation was predictable. For example, during the morning rush, passenger counts on board the Route 81 were highest at the outer ends of the run for buses inbound to central campus, and dropped once they reached central campus and began to unload staff and students on their way to jobs and classes. A screen shot of the simulation in action is shown in Figure 3.



FIGURE 3
 SCREEN SHOT OF SIMULATION MODEL DURING RUN

Results from Comparison of Simulated Bus Operation Before and After Schedule Change

Per the goals of the project, the simulation model was used to test the effect of changing from a variable headway schedule during the day (10 minute peak headways and 15 minute headways off-peak) to a fixed 10 minute headway. The motivation for the schedule change was to alleviate overcrowding at certain times, and also to make the service more convenient, since it would both be more frequent and easier to remember the schedule. Demand is held fixed for both schedules, so demand is generated in both runs of the simulation from the same statistical distributions. The simulation then tracks total passenger count on the bus, number of passengers turned away due to full buses, and cost performance measures (e.g. cumulative net revenue during the run of the simulation). Later runs of the simulation test the effect of raising passenger demand in response to shorter headways, and the effect of modeling routes 81 and 30 simultaneously. The transit operator was aware that if the demand truly did remain fixed, comfort on the buses would improve, but net revenues would decrease (more bus runs with the same number of fares collected). However, they were financially prepared for this possibility; also, they anticipated that there was a good chance that the ridership would increase after the schedule change, a prediction that was borne out by events, as discussed below.

Taking average passengers on board buses leaving a stop in the central campus (“Statler Hall”) shows that starting in the middle of the day, there are fewer passengers on board with the

modified schedule. Figure 4 shows the average loading across 100 runs of the simulation; inspection of the individual simulation runs revealed that there were days with up to three full buses during the peak period (2 to 5 PM). With the modified schedule, buses would be less likely to run at or near the full capacity, leading to increased comfort on the buses and ease of entry and exit, which TCAT customers were thought to appreciate.

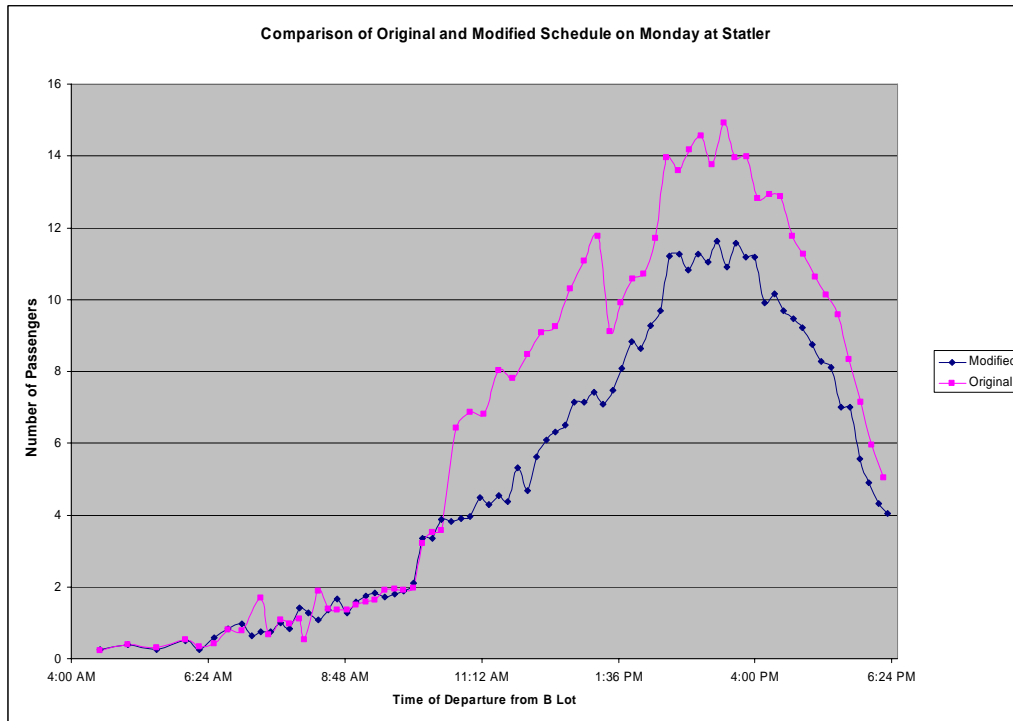


FIGURE 4

COMPARISON OF AVERAGE RIDERSHIP AT STATLER HALL STOP ON MONDAYS FOR ORIGINAL AND MODIFIED SCHEDULE

The assumption of demand remaining constant is conservative, since there is a realistic chance that the presence of additional buses may lead to an increase in passengers. To test this scenario the simulation was rerun with demand increases of up to 20% for the modified schedule to compare before and after performance of the route. It was found that the modified schedule would be able to handle 15% more passengers with a relatively similar level of customer service (i.e. average occupancy, level of crowding on buses, number of full buses) to that of the existing schedule and level of ridership.

Addition of Route 30 to Simulation Model

As an extension on the model, TCAT's Route 30 was added to the model in order to determine how this overlapping bus route would affect the current, isolated model of Route 81, since Route 30 is the other high-volume bus route serving the campus. The main purpose was to determine the effect, if any, on the results obtained from the initial simulation of the isolated Route 81, in terms of changes in average passengers on board, bus overcrowding, or financial performance.

In order to combine the two routes in a single model, the model needs to determine for each passenger boarding whether they are traveling to a stop served by both routes. The alighting percentages from above were used to calculate the fraction of passengers that would alight at a stop served by both routes. In the simulation model, these passengers could take either a 30 or 81, whichever arrived first. Other passengers traveling to destinations served only by Route 81 would wait for the first 81 bus. (Passengers traveling on the 30 to points outside of the Route 81 corridor are included in the passenger count on the 30 within the geographic scope of the model, but their onward trips to point of alighting are not considered in the model.)

The results of the combined route model confirmed that the interplay between the two routes would not greatly alter the results from the simulation of the Route 81 in isolation. The average occupancy on board the Route 81 buses at different points along its route was similar to the model of the 81 in isolation, with slightly lower average values across all days and time periods.

Exporting Data to Spreadsheets and Use of Conditional Formatting to Find Crowding

The model includes a macro to automate export of ridership data during the simulation of individual days for a more detailed view of ridership patterns by the hour and day in a spreadsheet. These data sets recorded how many riders were on the bus at each bus stop at each pick up time during the day for a given simulation run. The simulation was run 100 times to find the average values, so each dataset was averaged to create a single worksheet of results.

From this worksheet, a decision support tool was developed to show points in the route where there was a potential for overcrowding. Passenger counts were divided into different brackets (0-5 passengers, 5-10, 10-20, 20+), and then conditional formatting was used in the spreadsheet, i.e. green = low, yellow = somewhat crowded, red =crowded.

Impact on TCAT's Scheduling Decision

Shortly after the results of the simulation model were made available in Spring 2004, TCAT management took the decision to change the Route 81 schedule to the shorter headway. The comment was made at the time that the projection of reduced passenger levels on the buses in return for a slight decline in net revenue (assuming constant demand) was thought to be a desirable change. Also, the ability to project stop-by-stop numbers of passengers waiting and on-board, and also see the buses operating in the simulation model, gave management greater confidence in the decision. Discussion with TCAT after the change confirmed that, anecdotally, both crowding and instances of full buses decreased. The 10 minute headway has been maintained since 2004, and in the intervening time, annual boardings on the route have grown from 477,000 in 2003 to 636,000 in 2005, an increase of 25%. The anticipated connection between better service and more riders was therefore proven correct. This expansion occurred after the route grew only 0.8% from 2002 to 2003. While these statistics do not prove a causal link between increasing frequency in 2004 and growth in boardings, they do suggest that improving the convenience of using the route may have contributed to the growth.

Possible Extensions to the Model

The work with simulation for TCAT's route 81, as well as other routes in the network, can be extended in a number of ways. Since passenger demand during the middle of the day

(approximately 10 AM to 2 PM) is driven in large part by students traveling to and from class, and since peak student class enrollment has different peaks on different days of the week, it is of interest to test a modified bus schedule that matches class start times in the simulation model. Adverse weather conditions, especially from November to March, are known to increase ridership, so a modified winter schedule could be tested in the model as well. In addition, there are a number of less frequent bus services that serve part of the route 81 corridor, so their addition to the model would make the results more accurate. The effect of deploying articulated buses with a higher maximum capacity on the route can be tested as well.

DISCUSSION

The principal motivation for developing simulation applications for transit routes is to be able to better study in detail the impact of changing service parameters on bus occupancy, given demand that is random but whose average and standard deviation is known. There are two elements that support the improved level of detail, namely 1) the use of geo-coded data, which permits the analysis of passenger counts at all stops along the route, and 2) the simulation model itself, which allows the practitioner to both observe the number of passengers on the bus as moves in the simulation environment, and analyze passenger counts by location in detail.

One benefit of this added information is the improved understanding of the relationship between peak demand, bus capacity, and overcrowding. In conventional analysis of bus ridership, the total number of riders and number of route miles operated is known. From this information, it is possible to infer the peak number of riders on the busiest section of the route, and establish that the passenger capacity of the bus is adequate. However, use of average ridership numbers based on aggregate passenger data in this way overlooks the impact of variability. For example, even if average occupancy is acceptable compared to seats provided on the bus service with a given headway, if there is wide variation in demand, the number of instances where in actual practice the demand exceeds capacity and passengers are forced to wait for the next bus may be unacceptable. Use of simulation takes this variability into account.

A second, and related, advantage of the use of simulation is the ability to study levels of occupancy on the bus at all points on the route. Since the use of GPS data and building of a simulation model automates the calculation of average, peak and minimum occupancy on the bus in operation, all stages of the route can be conveniently viewed and analyzed.

Along with reproducing the existing schedule in the simulation model, it is also possible to test alternative headways in the model; an example of this analysis was shown in the TCAT case study above. For example, it may be of interest to test the benefit of shortening headway on a crowded route where passengers are routinely left at stops due to buses being full to capacity. By running the simulation model with the new schedule, it is possible to ascertain the benefit of adding capacity. Alternatively, an operator may consider increasing headways in order to shift resources to other routes in their network. The simulation model can be used to verify that reduction in bus service will not lead to overcrowding.

Alternative levels of simulation modeling

Given the time involved in building a simulation model, especially for multiple routes, it is desirable to be able to model transit operations at a level of detail appropriate to the results desired. Accordingly, we delineate three levels of transit simulation:

1. *Single-route model*: a single transit route in isolation.
2. *Corridor model*: multiple routes operating in a single transit corridor, possibly on the same street or parallel streets in close proximity. Where certain routes operate in only part of corridor, it is possible to include only the part of the route in the corridor, as was done with the Route 30 in the TCAT study.
3. *Network-wide model*: the complete network in a single model, explicitly modeling passengers transferring between routes.

For many situations, the corridor model may provide the appropriate level for modeling transit operations in order to gain useful insights for manageable efforts. A single-route model may fail to capture the tendency of transit riders to adapt to changes in the schedule of one route by switching to another. A network-wide model may require a great deal of additional effort, and in the case of testing changes on a single route, the effect of routes operating in other parts of the network may not add much insight to what can be learned in the corridor model. It should be noted that in the case study above, we have modeled a corridor in a small transit system where such a corridor can be isolated from the rest of the system with relative ease. In the case of a large urban transit system operating a large number of routes in a dense urban grid, it remains an open question as to whether such corridors can be isolated, or whether a “sub-network” model (for example, a network of routes radiating outward from the Central Business District) may be more appropriate.

Data considerations

In order to accurately simulate bus transit routes, the operator or analyst must be able to gather detailed boarding and alighting data for the route or routes under study. The boarding data are the easiest to gather, as they are tied to the electronic fare collection system that are in widespread use. With such a system on the bus, it is a relatively small additional cost to add GPS in order to geocode the data, and then store them in an archive for later analysis. Our experience indicates that challenges still remain with the complexity and tedium around the multi-step process of converting raw geocoded data into ridership statistics. If interest in simulation of transit systems were to gain momentum, however, we do not foresee any inherent difficulty in the creation of a streamlined process for making this conversion.

Alighting data are more difficult to gather because there is less incentive to gather these data electronically. Devices for recording the alighting of each passenger as they leave the vehicle are available, but not in widespread use, as transit management often cannot justify the additional cost. Thus in the near term, the alighting data available for simulation models will likely remain the small, hand-gathered data sets of the type used in the TCAT case study. (It should be noted that hand-gathering of alighting data need not be seen as wasted effort, as it serves the dual purpose of gaining hands-on experience with the practicalities of how a given transit route operates.) Here again, growing interest in simulation may help to justify investment in electronic gathering of alighting data in the future.

In addition to boarding and alighting data, information about transit riders’ origin to destination travel patterns, and especially door-to-door travel (e.g. from place of residence to workplace) would allow further evolution of simulation modeling of transit. This data would enable the discrete modeling of individual traveler’s trips through the system, as opposed to the aggregating of passengers at stops and on buses seen in the case study above. It would also

make possible not only the simulation of frequency changes on geographically fixed routes, but simulation of the effect of rerouting. Clearly such data, while desirable, would also be difficult to obtain. If gathering data from the direct surveying of transit passengers proves prohibitive, it may be possible to adapt origin-destination data from the four-step models used by transportation planners.

Alternative Platforms for Simulation

In our work we have used an existing manufacturing design software package as the platform for development of our simulation model, since it provided a convenient way to model the operation of a transit route by taking advantage of commonalities between a transit vehicle moving through a community and a guided vehicle moving through an assembly line. Alternative platforms are also possible. One option is the use of a general simulation programming language such as SLX; another is to use a high level programming language such as C++ to code the simulation model from the ground up. The latter option gives the greatest flexibility in model design, and also has the lowest software cost, but also requires the greatest effort to launch a model. Pursuit of simulation modeling of transit operations using any of these three types of platforms might lead to the evolution of dedicated transit simulation platforms that would be both powerful and easy to use.

Discussion of Service Learning Experience

The experience of working with a transit agency on improving bus service proved valuable to the students, to the agency, and to the broader community served by the bus system. Some of the ridership served by transit buses includes those not able to drive, such as the elderly, the disabled, or those with insufficient economic means to be able to afford car ownership. Transit buses also improve energy efficiency and reduce the number of private cars circulating in the urban area, helping the city to become more sustainable. Improvements to the performance of the transit system allow the operator to carry more passengers at lower cost, furthering these social and environmental goals.

One benefit of working with a transit agency as a partner for a service learning project is that, as a public agency receiving government funding for the purchase and operation of buses, transit agencies such as TCAT are required to keep all financial and operating data open to the public. We therefore did not encounter problems of not being able to obtain data for reasons of confidentiality, and moreover, were not prevented in any way from publicizing our findings. This transparency made it easier to carry out the project.

One challenge faced during the project was the need to convey to the partner, TCAT, the capabilities of operations research as applied to transit bus operations. TCAT management are well versed in standard transit operations techniques, including the use of computers and IT to manage and plan for bus operations. However, because they did not have background in computer simulation techniques, the TCAT management gave the Cornell team a broadly defined project target, leaving the team to decide how best to solve the problem as posed. At the end of the project, the students were challenged to not only explain their results but also give the partner a basic understanding of the underlying techniques, so that TCAT could be more confident in their work. Therefore, explaining the work of engineers in laypersons' terms became an integral part of the project. Flexibility in the scope of work also allowed the students

to use their creativity in creating the computer animation to represent the buses, as well as the spreadsheet with conditional formatting.

The Cornell team discovered that, as is often the case with a groundbreaking project of this nature, much of the person-hours of effort was expended not on the engineering analysis itself but in the necessary preparations of the data and hardware to be able to gather the data and build the simulation model. For students accustomed to working in a classroom environment where homework and exam problems arrive well pre-processed, this was an eye-opening experience. It was therefore all the more satisfying at the end of the project to finish with a simulation model that functioned correctly and reproduced the actual bus movements reasonably well.

As a final comment, we feel that, based on the experience of working with TCAT, transit agencies are a rich area for service learning projects. The focus of such projects need not be simulation modeling, since many engineering topics are of interest to transit, ranging from the mechanical function of the buses to IT to systems analysis of the potential market for new ridership. Also, many university campuses use bus systems to alleviate car congestion and transport members of the academic community between campus, dormitories, and peripheral parking facilities. Both students and transit agencies can take advantage of this synergy to provide service learning projects that give useful results to the bus operators and at the same time give the students valuable hands-on experience. We encourage engineering students and faculty at other universities to consider starting such projects.

CONCLUSION

In this paper we have presented results from our effort to develop a transit simulation model, from gathering of geocoded boarding data, through extraction of statistical distributions for ridership patterns, to the operationalizing of a simulation model and application to a schedule planning decision. Given the absence of many papers on this subject in the literature, we have also discussed general insights about transit simulation gained from the project. Our experience shows that with the advances in GPS technology and fare collection systems in recent years, a wealth of data about transit riding habits can be made available for a low cost, if an effort is made to develop analytical tools to understand and use the data. This paper represents one step in the process of developing those tools. The authors are interested that the use of simulation and the creation of other service-learning projects be created elsewhere. Please contact the corresponding author if you would like to learn more.

ACKNOWLEDGMENTS

The authors wish to thank Randy Chan, Vincent Chan, and Alienor Salin, students in the Master of Engineering Program in the School of Operations Research and Industrial Engineering at Cornell University, and Prof. Mark Eisner, director of the program, for their contribution. We also express thanks to an anonymous reviewer and to the editor, whose feedback helped to substantially improve the paper. The opinions expressed are solely those of the authors and do not represent official opinions of either Cornell University or TCAT. Also, responsibility for any and all errors rests with the authors.

REFERENCES

- ¹ Wong, W., F. Hou, Z. Shen, and K. Ikeuchi. (2004) System reliability and maintainability assessment through simulation: A case study of bus dynamic operation in transit networks. Proceedings of the Eighth International Conference on Applications of Advanced Technologies in Transportation Engineering, 2004, p 421-425.
- ² Ding, Y., S. Chien, and N. Zayas. (2000) Simulating bus operations with enhanced corridor simulator: Case study of New Jersey transit bus route 39. Transportation Research Record, n 1731, 2000, p 104-111.
- ³ Morgan, D.J. (2002) A microscopic simulation laboratory for Advanced Public Transportation System evaluation. Master's Thesis, Massachusetts Institute of Technology. Web resource, available on line at http://web.mit.edu/its/papers/Dan_Masters.pdf. Accessed August 31, 2007.
- ⁴ Schulz, T, Lemessi, M, and Filippi, F. (2001) Simulation of a night taxi-bus service for the historical center of Rome. *Proceedings of the 2001 Winter Simulation Conference*. Peters, B, Editor. 1072-1078.
- ⁵ Thweatt, M (2004). Simulation of bus route at Middle Tennessee State University. Website <http://www.mtsu.edu/~csmack/indexA.html>. Accessed July 13, 2004.
- ⁶ Andijani, A (2001): A Simulation model for the shuttle bus traffic during Nafrah period. Masters Thesis, King Fahd University of Petroleum and Minerals. Website <http://www.kfupm.edu.sa/library/theses/>. Accessed July 13, 2004.
- ⁷ Zargari, S. A., A. M. Khan. (1998) Simulation model of the bus Transitway (TRNSIM). Journal of Advanced Transportation, v 32, n 2, p. 152-174.