9-1997

Issues in the Design of a Stop-Level Transit Patronage Model

James G. Strathman
Portland State University

Kenneth Dueker
Portland State University

Zhongren Peng
Portland State University

Issues in the Design of a Stop-Level
Transit Patronage Model

James G. Strathman
Kenneth J. Dueker
Zhongren Peng

Interim Report

Introduction

Tri-Met, the transit agency serving the Portland, Oregon metropolitan area, is implementing a new bus dispatching system (BDS) that includes automated vehicle location (AVL) technology. An important objective in the agency's decision to invest in a new dispatching system was to maintain or improve service reliability in an urban environment characterized by rapid growth and worsening traffic congestion.

With the new system, dispatchers and field supervisors will be able to employ operations control measures in response to real time information on the actual location of buses in relation to their scheduled location. If these measures are indeed productive, we can expect to observe improvements in headway maintenance and running time variation in the short run, and patronage gains in the longer term. In the present phase of the research project, the authors are conducting a pre-BDS baseline analysis of service reliability on selected routes, as well as a survey to document transit riders' perceptions of service quality. Following implementation, operations control measures will be assessed in terms of their effects on service reliability and on riders' perception of service quality.

With the adoption of the AVL component of the new system, Tri-Met will be able to locationally reference buses along routes and compare actual and scheduled times at time points. While the use of real time operations control measures is expected to result in better quality service, it could also result in benefits for planning and scheduling activity. The agency's new locational referencing capability will provide the means for validating boarding and alighting data collected by Automatic Passenger Counters (APCs) at specific stops, whereas before AVL, passenger data could not be consistently resolved to the bus stop level. This gain has important planning implications because service changes frequently are focused at the route segment or stop level, where route level totals are less useful.

Previous work by the authors has lead to the development of a route-segment level patronage model for Tri-Met's bus system, based on passenger census data the agency

periodically collects coupled with level of service information and route corridor data generated by a Geographic Information System (GIS). In this report we discuss issues in designing a transit patronage model using stop level data.

Unit of Analysis: Route, Route-Segment or Stop Level Models

Transit ridership varies at the route, route segment and bus stop level, while transit level of service varies at the route or route-segment level, at least from the viewpoint of service planning. Models that estimate transit demand and service supply thus can be developed at the route, route-segment and stop level.

Previous work has used the route or route segment as a basic unit of analysis in developing transit demand models. These studies represent an improvement over system-wide models by accounting for variation in ridership and service supply among routes and route segments. The underlying assumption of these route or route-segment level models is that the social-demographic characteristics along the route or route segment are homogeneous, but this is invalid in most cases. Transit ridership fluctuates among different route segments and stops because of the different land uses or trip generators surrounding transit stops. It is difficult for route and route-segment level models to take into account important stop-level characteristics, such as the pedestrian environment, residential or employment locations, or park and ride lots. Furthermore, adjacent routes compete with each other for passengers. At the route or route-segment level, these competing effects are controlled by the percentage of overlapping service area over the entire route or route segment. This is only a surrogate because transit services compete with riders at the stop level.

Bus stops seem to be the ideal unit of analysis for a transit demand model. With stop-level models, the stop-specific variables can be incorporated. These variables include

amenities such as shelters, lighting and patrols, and the pedestrian environment (e.g., sidewalks and slopes).

The lack of stop-level data, particularly ridership, has hindered the development of transit ridership models, but technological developments offer prospects for improvement. Geographic Information Systems (GIS) provide a means to allocate and integrate social-demographic data and transit service data at the stop level. AVL technology provides a means of locational referencing of passenger activity and other operational data recovered by APCs.

On-time performance, a key service quality indicator, cannot be accurately assessed at each bus stop because there is no official scheduled arrival time for every stop. Therefore, on-time performance at non-time point stops may need to be inferred from on-time performance at a previous time point. Alternatively, a model can be developed at the time point level rather than at the stop level. But time points usually represent major intersections and land use attractions. Ridership at those stops may not represent those in other non-time point stops. Thus modeling transit demand at the time-point may introduce bias. One way to reduce this bias is to aggregate ridership of stops that surround a time point to that time point. This will downgrade the stop-level model to the route-segment level, with the number of segments equaling the number of time points.

It should be noted that for a given route, variations of service do not occur at the stop level. Transit service usually varies at the route or, at most, the route segment level. Therefore, the service variable may not have much variance, and may thus turn out to be insignificant in the model output. Even the effective head-way at stops may not have enough variation for the variable to be statistically significant. On-time performance is mostly a measure of route performance rather than the service performance at the stop level. Furthermore, on-time performance is often negatively associated with ridership or, more specifically, ridership variation. Therefore, there are confounding effects of on-time performance and ridership. On the one hand, better on-time performance may increase

ridership while, on the other hand, a ridership increase may negatively effect on-time performance. So the coefficient of on-time performance in the model may be positive, negative or insignificant.

The stop-level model is more useful to relate transit demand with demographic and land use characteristics, and it is of more interest with regard to operations control. Better understanding of transit demand and land use and demographic characteristics can help transit service planning to better plan services at the route or route-segment level. A well-designed of stop-level model is essential to serve both service planning and operations control needs. But, the above discussion implies different levels of observation may be needed for operations (time point) than for service planning (bus stop). this issue is discussed further below.

Models for Service Planning and Operations Control

Service planning is most concerned with relating transit ridership and level of service (LOS). The level of service is designed to vary across routes and time in responding to passenger demand variations. The level of service is usually planned at the route or route-segment level, but demand is realized at the stop level. Operations control, on the other hand, is concerned with maintaining reliable service. One goal of operations control is to reduce the variation of effective headway between buses. For this goal the appropriate unit of analysis is the time point.

Different models are needed to represent the behavior of transit service planners and operation controllers. For service planning, the model is best estimated at the route or route segment level. It is specified in the following general form:

Ridership = F(LOS, riders' social-demographic data, land use data)
LOS = F(ridership, land use data)

For operations control, the model is best estimated at the time point level, or at the stop level if interpolated stop level arrival times can be used, and is of the following general form:

Ridership = F(effective headway, on-time performance, riders' socio-demographic data, land use data)

Effective headway = F(load factor, ridership, on-time performance of the previous bus)

A composite model may be constructed to consider the needs of both service planning and operation control as follows. Notice this is a recursive model between Ridership and LOS. The model has to be estimated on at least two stages. That is, the first stage is to estimate the simultaneous model between ridership and effective headway at the time point or stop level. The second stage is to aggregate the estimated ridership to the route level, and to estimate the LOS model at the route level.

LOS = F(ΣRidership, Land use data) (-- at the route level)

ΣRidership = F(Effective headway, on-time performance, riders' social-demographic data, land use data) (-- at the time point or stop level)

Effective headway = F(Load factor, ridership, on-time performance of the previous bus) (-- at the time point or stop level)

Stop level models will be explored first. If model development does not prove viable at this level, we will fall back to the time point level.

Model Structure

With abundant ridership data at the stop level, models can be developed for individual routes. But the question is whether we need to develop models for individual

routes. For the purpose of service planning, models for each route typology, such as cross-town route, feeder routes, radial routes and express routes may be sufficient.

For each route typology, there need to be models for different time periods and directions (inbound and outbound) to capture the temporal variations of ridership. These variations include time of day (peak, midday, evening and night) and day of week (weekday, Saturday and Sunday). Furthermore, separate models may be needed to estimate boardings and alightings.

Ridership Data Sampling

The model development by route typology and time of day involves spatial and temporal data sampling.

Spatial Sampling

A decision has to be made as to what routes are to be selected as representatives of a route typology. Once the representative routes are chosen, stops need to be selected. One option is to use all stops for every sampled route in the model estimation, but the disadvantage is that the service area of adjacent stops overlap and more allocations would be needed for socio-demographic variables. Another option is to sample stops. The advantage of sampling at the stop level is the reduction of allocation of socio-demographic variables. But the disadvantage is that there may be some sampling errors. A carefully designed sampling technique such as stratified sampling needs to be developed to make sure that the sample stops represent the stop population.

Temporal Sampling

To capture the temporal variations (hourly, daily and seasonal variations) of ridership, data need to be selected from different time periods. The hourly variation of the same time period such as midday, morning peak and afternoon peak can be captured using

the ridership distribution (e.g., mean and standard deviation) of that time period. The daily variations of the same time period can be captured using the ridership distribution of days of the week. And the seasonal variations of ridership are captured using the ridership distribution of the months in seasons.

Therefore, there needs to be weekly data from at least one month in every season, such as the second week of the month of February, May, August, and November. The weekly variations and seasonal variations (or seasonality) can be indexed using moving averages. The purpose of indexing weekly and seasonal variations is to add to variations in ridership forecasting.

For the purpose of data collection at Tri-Met, the full APC data set of APC-equipped buses on specified routes for a four-week time period is recommended. This means scheduling of APC-equipped buses to the specified routes for specified weeks during each sign-up.

Spatial and Temporal Autocorrelation

Because the data set to be used in the model is a pool of spatial and temporal data, estimation will likely need to correct for both spatial and temporal autocorrelation. Ridership at bus stops along a given route and ridership for the same stop at different time periods are probably autocorrelated. There is, however, no known method to deal simultaneously with autocorrelation over space and time. Temporal autocorrelation can be eliminated if the ridership at a stop to be used is averaged across days and months. Spatial autocorrelation cannot be eliminated unless the model is estimated at the route or zone level.

Spatial Data Allocation

Like the route and route-segment level model, the stop-level model requires that socio-demographic and land use data be allocated to individual bus stops. A fifteen-minute isochron (or a quarter mile walking distance) is defined as the service area using the allocation function in Arc/Info. Parcel or block level data (or grid data) from RLIS (Regional Land Use Information System) related to the selected street segments are also needed. Car ownership and income data will also need to be allocated from census block groups to bus stop service areas.

Special caution needs to be paid to spatial allocation of data on overlapping service areas. For service areas on adjacent stops along a route, the overlapped area needs to be equally divided among stops. However, for overlapped service areas on adjacent stops or the same stop at different route, the allocation method depends on the relationship of the routes (see below).

Competing And Complementary Effects Among Routes

It is a common phenomenon for one bus stop to serve multiple routes. Passengers have a choice among routes if these routes serve the same destination. In other words, these bus routes compete with each other for passengers. The same bus stop may also serve as a transfer point from one route to another, in which case the routes complement each other. It is difficult to determine whether two routes are complementary, competing, or independent without knowing passengers' origins and destinations.

When there is no information on passengers' origins and destinations, ad hoc criteria must be used to identify the complementary and competing status of routes. One criterion may be the sharing of the origin and destination locations. Routes that have the same destination, such as ending at a downtown transit center may be considered as

competing routes; while intersecting/overlapping routes that have different origins and destinations may be considered complements. Another criterion is the amount of overlapped service areas. If the overlapped service area of two routes (not stops) extends over 50 percent of that of the entire route, these two routes could be considered as competing routes. Service planners in Tri-Met need also to be consulted since they know the route system well.

Bus stops serving competing routes have two effects on ridership. Competition may reduce ridership on individual buses at a particular time, but it may increase the ridership on the route because the greater service frequency tends to attract more transit users (synergistic effect). So in the model, two variables can be used: the number of competing routes a stop serves and the number of complementary routes a stop serves. The sign of the coefficient of the first variable may be positive or negative depending on the relative strength of synergistic and competing effects. A bus stop that serves complementary routes should have positive effect on ridership. The use of these two variables reduces the need to proportionately allocate the served population and employment to each route that goes through the same stop.

Other Issues

Lagged Effects of On-Time Performance

If a bus is earlier than the scheduled arrival time it may miss some passengers, who then have to take the next bus. This effect must be considered in estimating ridership on a bus route at a particular time. But it can be ignored if the average ridership is estimated for a particular route at a time period, especially if a random snapshot of one bus sample is drawn from a time period. Furthermore, this effect can be ignored if route level data are used. Long term effects of on-time performance on ridership cannot be estimated without a time series model.

Load Factor

A full bus will deter passengers from boarding. A load factor, the ratio of passengers to seats, can be calculated if we have detailed data for boardings and alightings. A side benefit is that we can gain knowledge about the passengers' trip length by making some assumptions. The benefit of being able to calculate trip length is to calculate service output in terms of passenger miles. But it will have little impact on stop-level transit demand modeling.

Accessibility to and Amenities of Transit Stops

Accessibility to transit stops is represented by three variables: walking distance, slopes, and the availability and connectivity of sidewalks. Walking distance and slopes can be calculated on the street patterns using Arc/Info. Amenities at transit stops include the lighting, shelters and security patrols.