Chapter 6

USING AUTOMATED FARE COLLECTION DATA, GIS, AND DYNAMIC SCHEDULE QUERIES TO IMPROVE TRANSIT DATA AND TRANSIT ASSIGNMENT MODELS

Howard Slavin, Andres Rabinowicz, Jonathan Brandon,
Giovanni Flammia, Robert Freimer
Caliper Corporation, Newton, Massachusetts

1. INTRODUCTION

This paper provides an interim report on a novel research effort aimed at developing improved data and models for demand prediction for very large transit systems. We explore the nature and use of automated fare collection (AFC) system data which has great potential for characterizing and forecasting transit use, but it requires a considerable effort to make it useful. The material discussed is motivated by work that we are performing in New York City, but the research should be transferable to many other large systems.

Although there are a variety of theories and mathematical models for transit route choice, little is known about traveler behavior in large transit systems that are characterized by many alternatives for the same trip. One reason is that there has been greater emphasis on mode choice than on route choice for transit users. Another reason is that there is little empirical information available. Onboard surveys can be difficult in large systems and often may be limited in the scope of data obtained making inferences about route choice impossible. Even when available, service changes and external events often render survey data quickly obsolete.

There are important outstanding questions about pathfinding, route choice, and assignment methods that involve both behavioral and mathematical issues. As with much transport modeling, highly simplified assumptions typify transit models. Many of these assumptions have not been investigated empirically in complex environments.

Automated fare collection (AFC) system data offers the promise of providing more information about demand and traveler behavior than has heretofore been available. AFC data are generated every day and, with some effort, can be processed for use in planning and modeling efforts.

Previous studies have used AFC data to enhance planning and modeling for rail transit (Barry et al. 2001; Rahbee, 2002; Zhao, 2004), but have not been comprehensive and did not include all bus trips. Processing of AFC data can be extremely difficult as AFC systems have usually been designed for revenue management and not for other purposes (Zhao, 2004), and the issues associated with bus data are much more complex than those for rail systems where boarding and alighting locations account a fixed location. Automated vehicle location systems simplify the problem considerably, but these are not yet widespread or ubiquitous. Consequently, developing an alternative approach for locating bus boardings was a major aspect of our research.

The advent of copious quantities of AFC data makes it important and timely to consider how to use AFC data to develop an improved understanding of route choice in large transit systems. However, it is important to understand the nature and limitations of AFC data.

AFC data typically provide data records for each boarding that is made subsequent to a swipe or dip of a farecard or other digital medium. There is a time stamp associated with each swipe which may be to the nearest second or in the case of New York, truncated to six minute intervals. On a bus, the identity of the bus will also be.recorded and for subways, there will be a record of which entrance or turnstile was used. Note, that in the latter case, the route boarded may not be uniquely determined.

To be useful AFC data must be geocoded accurately which, in the absence of automated vehicle location (AVL) technology, is by itself rather challenging. For rail facilities, the geocoding will be straightforward. For buses, this problem is vastly more difficult because it is the rare system that has widespread AVL at the present time. By linking the schedules with the geography of the bus routes, approximate geocoding can be performed based upon the scheduled location of a bus for each boarding. Interpolation can be used to handle some cases in which buses are off-schedule and this has been done.

Importantly, the location of bus to rail, rail to bus, and bus to bus transfers helps anchor the locations of boardings at other stops even when there is less than perfect schedule adherence.

AFC data often does not include data on the exit or alighting stop for each trip. When this is the case, the destination of each trip must be imputed or inferred from information on other trips made by the same farecard user. The typical imputation method is to assume that the destination of the first trip is the boarding location of the next trip. This can be refined in various special cases to provide better information, but there is evidence that even the simplest logic works fairly well in matching system aggregate boardings and alightings by location.

An important special case is that of linking trip segments that are actually part of the same trip. These will commonly be observed when there is a mode to mode or bus to bus transfer that results in two swipes within an appropriate time interval. This technique improves the imputation of trip destinations considerably.

GIS is an enabling technology for making use of AFC data and is a pivotal technology for visualizing transit pathfinding and assignment results. Accurate geographic depiction of transit routes and the location of stops is necessary for success in reverse geocoding of locations from schedules. Because there are many route patterns, use of a representative pattern, as is common in some modeling efforts, can lead to errors in trip table development and forecast route utilization. Note, too, that the specific route configuration is a determinant of the capacity provided between specific origins and destinations.

With the vastly increased computing capabilities that we now have, it is perfectly reasonable to have GIS databases containing every trip during a particular time period for a large system. For one day in New York City, there were approximately 7.5 million AFC transactions. Once geolocated, these can be tabulated and queried at will and further processed. These databases can be extended to hold multiple days of travel and to make it possible to analyze the trips made by the same individual (farecard) across several days or time periods.

Similarly, it is equally feasible to store all paths (or hyperpaths) that are generated in the course of a transit assignment and to perform similar queries. These should all meet some minimal standards for reasonableness and conformance to input assumptions such as the maximum number of transfers or the maximum length of a trip.

Clearly then, one can compare measured trips with modeled trips as a validation measure and importantly as a source of diagnostic information that can lead to improved models. Measured trips with attributes derived from accurate GIS-based routes and schedules can be used to generate estimation datasets for route choice models that should have more realistic and empirically-based coefficient estimates.

The trips derived from the AFC data provide a rich source on information on temporal as well as spatial patterns of use, permit characterization of boardings at stops by time interval and can be used to ascertain peak hour travel characteristics. When combined with a dynamic and/or schedule based assignment or simulation, peak load points and other measures that are useful for operations planning can be derived.

The GIS-based software tools are designed to automate the processing of the AFC data so that transit agencies can process these data on their own. After the data have been processed, user-oriented GIS-based query and analysis tools can be used to study and analyze the trip data. While the implementation has been specific to New York, the software tools developed can be customized and applied for a wide range of transit systems with either more or less data.

The organization of the remainder of this paper is in two parts. First, we describe the AFC data and the processing necessary to make it useful for planning and modeling.

Second, we describe the use of this data in analysis of transit assignment models.

2. AFC DATA PROCESSING

Extensive data processing was performed of AFC and other transit data including routes and schedules. A schematic of what is involved is shown in Figure 1. Many different data sets are utilized to accomplish the task of generating geographically referenced trip origins and destinations.

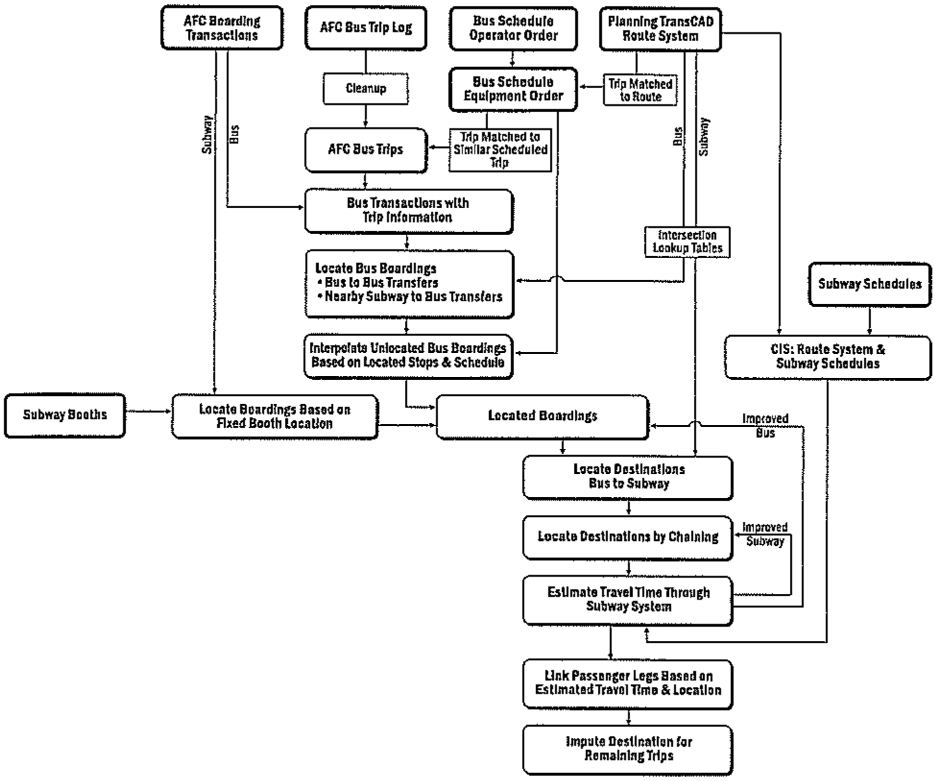

AFC Boarding Transactions
AFC Bus Trip Log
Bus Schedule Operator Order
Planning TransCAD Route System
Cleanup
Bus Schedule Equipment Order
Trip Matched to Route
Subway
Bus
AFC Bus Trips
Trip Matched to Similar Scheduled Trip
Bus
Subway
Bus Transactions with Trip Information
Intersection Lookup Tables
Locate Bus Boardings
• Bus to Bus Transfers
• Nearby Subway to Bus Transfers
Subway Schedules
Interpolate Unlocated Bus Boardings Based on Located Stops & Schedule
GIS: Route System & Subway Schedules
Subway Booths
Locate Boardings Based on Fixed Booth Location
Located Boardings
Improved Bus
Locate Destinations Bus to Subway
Locate Destinations by Chaining
Improved Subway
Estimate Travel Time Through Subway System
Link Passenger Legs Based on Estimated Travel Time & Location
Impute Destination for Remaining Trips

Figure 1. Schematic representation of AFC data processing

There are several species of farecards. They can be purchased for different time periods and dollar amounts.

The computers that record farecard swipes keep information on the farecard used and the boarding date, time and mode among other data.

2.1 Generating an Accurate Geographic Route System

A key aspect of our strategy for processing the AFC data was to georeference boardings based upon a good representation of the transit system. In TransCAD, stops can be located at their true geographic locations and need not be at intersections of streets.

The fullest use of AFC data requires an accurate route system that captures all of the scheduled variations in transit service. This is not commonly available for travel demand forecasting but can be developed in a GIS using suitable input data.

A previous route system had been developed and has been in use for developing transit networks for forecasting. This route system needed to be further refined to handle route variations by time of day and to represent the patterns identified in the schedule. This is an ongoing process which is enhanced by detailed attention to schedules and stop locations. Altought perfect fidelity is not required, it is the ultimate goal.

A matching procedure was developed by comparing sequences of stops to assign a TransCAD route ID to each scheduled bus and subway trip. All missing subway patterns were added to the route system using an automated procedure. The AM route system has 61 subway routes and 767 bus routes (Figure 2).

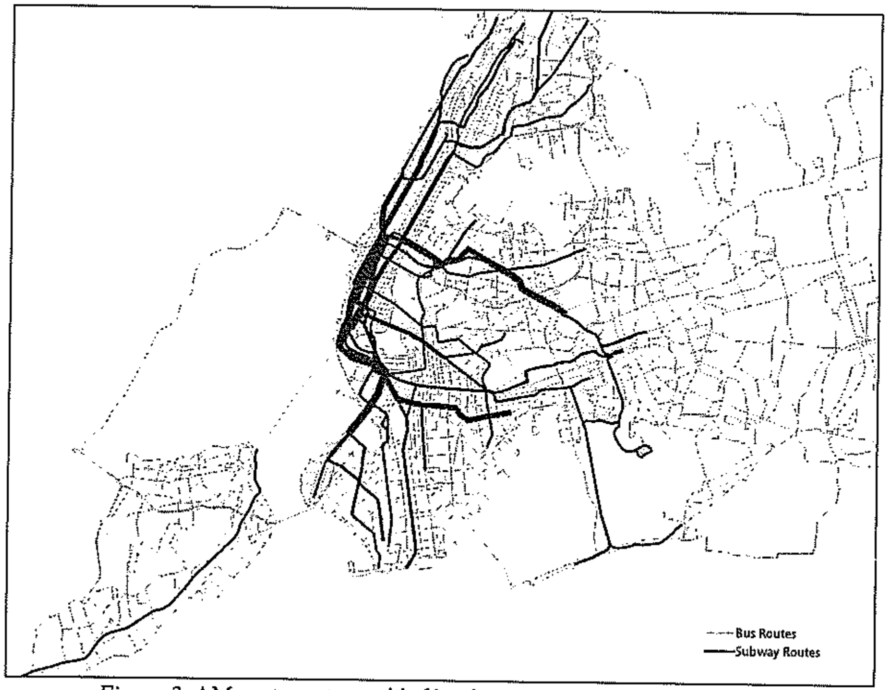


Figure 2. AM route system with 61 subway routes and 767 bus routes

GIS queries make it possible to identify feasible transfer points. We use the route system to pre-compute intersection lookup tables between pairs of bus routes. The table identifies the nearest stops on the intersecting routes. A similar table from subway stations to bus routes is computed. Entries are provided between a station and a bus route when all the subway routes from that station intersect the bus route at a single location. Both types of lookup tables are used in determining bus boarding locations.

2.2 AFC Boarding Transactions

The MetroCard transaction file contains one record for each entry swipe or dip. Entry transactions are recorded by subway turnstiles and bus fare boxes. These data are of high quality, with little cleanup required. Unfortunately, the transaction times are truncated at each farebox to six minute accuracy to save storage space. Fourteen days of data (3am–3am) comprise almost 95,000,000 records, which is about 8GB.

A utility was created to convert the files from the mainframe COBOL format to TransCAD's binary format, discarding irrelevant data fields. The utility also changes any unreasonably large fare deductions to missing values and sorts the transactions by serial number, transaction date and time. Transactions outside the two week study period were discarded. Records corresponding to multiple passengers using a value-based MetroCard were consolidated into a single record with a corresponding entry for the number of riders.

2.3 AFC Bus Trip Log

The AFC bus trip table logs information from bus trips and one or more records should be present for each physical trip. Certain trips may be broken into multiple records due to various events, which include sign changes made by the driver, fixed times during the day (midnight, 6am, 9am, 4pm and 7pm), crew changes, etc. Sometimes, a record may represent more than one trip if a driver failed to sign on or update the overhead sign code.

The AFC bus trip tables were received in a text format, one file for each bus operator. The file format is fixed-format ASCII once the report heading at the beginning is removed. We applied a dictionary so that the files could be used with TransCAD. This was then saved as a TransCAD binary file, keeping only the immediately useful fields. The starting and ending dates and times were each converted into a single numeric field to facilitate matching with the AFC transactions later in this project. Records for events falling outside the study period were discarded. The resulting file has 1,200,000 records.

We joined the trip table to our bus depot database to add the depot and carrier to each bus trip record. Similarly, the sign text and route pattern were added by joining with the sign lookup table. Significant effort was expended to cleanup the bus trip log to attempt to have one record per physical bus trip.

2.4 Bus Schedules

For buses, there are up to three separate schedule files, corresponding to Weekdays, Saturday and Sunday. Each file details the location of stops, the list of trips and the sequence of stops which comprise a trip along the route.

We developed a TransCAD GISDK script to convert the schedule files into a TransCAD binary table. The procedure opened each schedule file in turn and expanded the data into a single table with a total of 5,270,891 records, each corresponding to a schedule event. These schedule events occur at 63,049 unique route-stops, corresponding to fewer physical locations.

We reordered the schedule records to reflect the equipment order instead of the operator order that is used in the schedule file. The equipment order schedule was used to assign a reasonable scheduled trip to each AFC bus trip based on pattern code, sign code and time.

2.5 Subway Schedules

The subway schedules are similar to the bus schedules. For each line, there are up to three separate files, corresponding to Weekdays, Saturday and Sunday. Each file details the stations, the list of trips and the sequence of stations which comprise a trip along the route.

We developed a GISDK script to convert the subway schedules into a TransCAD binary table. The procedure opened each schedule file in turn and expanded the data into a single table with a total of 595,876 records, each corresponding to a schedule event. These schedule events occur at 2,549 unique route-stops, corresponding to fewer physical locations.

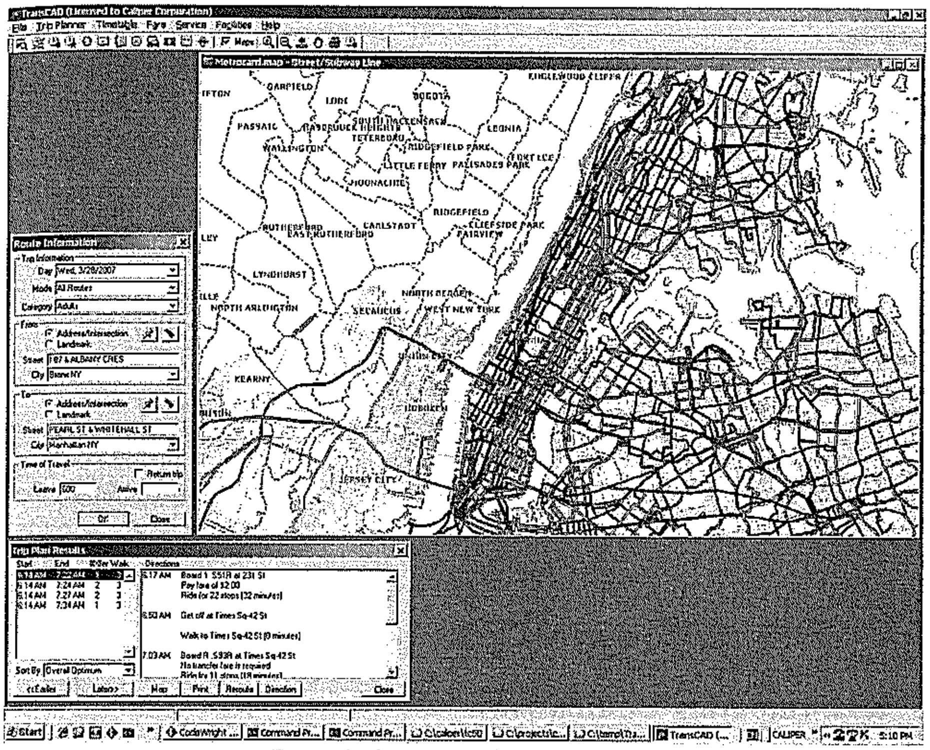
2.6 Subway Booths

We created a master list of subway station booths with coordinates for use as a reference layer. We combined the station stop information from NYCT's route system with their list of fare control locations. This created a list of subway booths, with information on the station, the serving subway lines and coordinates that can be joined to the boarding transactions using the booth code.

2.7 Customer Information System (CIS)

We extracted the subway, tram and ferry portion of the route system and loaded the corresponding schedules into a TransCAD-based CIS module that generates schedule-based shortest path queries.

These were used to estimate subway destinations and travel times and also as a passenger trip visualization tool.



Route Information

Trip Information
Day: Wed, 3/28/2007
Mode: All Routes
Category: Adult

From
● Address/Intersection
○ Landmark
Street: E 87 & ALBANY CRES
City: Bronx NY

To
● Address/Intersection
○ Landmark
Street: PEARL ST & WHITEHALL ST
City: Manhattan NY

Time of Travel
□ Return trip
Leave: 600 Arrive:

OK Close

Trip Plan Results

```
Start     End       #Xfer Walk    Directions
6:14 AM   7:22 AM   3             6:17 AM   Board 1 SS1R at 231 St
6:14 AM   7:24 AM   2     3                 Pay fare of $2.00
6:14 AM   7:27 AM   2     3                 Ride for 22 stops (32 minutes)
6:14 AM   7:34 AM   1     3
                                  6:50 AM   Get off at Times Sq-42 St

                                            Walk to Times Sq-42 St (0 minutes)

                                  7:03 AM   Board R SS3R at Times Sq-42 St
                                            No transfer fare is required
                                            Ride for 11 stops (18 minutes)
Sort By: Overall Optimum
```

<<Earlier Later>> Map Print Results Direction Close

Figure 3: Customer Information System

2.8 Locating Subway Boardings

We located subway boardings by using the subway booth layer. This is more precise than the physical location of the subway station. However, it does not always uniquely identify the line that was boarded or the direction of travel.

2.9 Adding Trip information to Bus Transactions

We matched the bus boarding transactions with the bus trip file using the bus number, location code and transaction time, taking the time truncation into account. In the case of transitions between bus trips, we choose the later trip for the boarding. We are able to identify a bus trip 99% of the boarding transaction.

2.10 Locating Bus Boardings

The six minute time truncation degrades the accuracy of our bus boarding locations in many cases, since the bus can service many stops during six minutes.

To minimize this inaccuracy, we first locate those transactions which are transfers from some other bus route or subway. We then interpolate times for the non-transfer stops. To apply this strategy, we process all the boarding transactions for a particular bus trip together.

For MetroCards, with two consecutive bus transactions within thirty minutes (sixty minutes for an express bus), we locate the transfer onto the second bus using a pre-computed bus route intersection table based upon the route system. A similar intersection table is used for subway to bus transfers, when the subway station is within five miles and only has routes which intersect the subsequent bus route at a single location.

The stops located are used to assign actual transaction times to some of the stops along the trip being processed. These times along with the starting and ending times for the trip, which are not truncated, allow us to update the schedule times to estimate when the bus was at each stop. These estimate times are truncated to six minute accuracy. Each bus boarding that has yet to be located is randomly assigned to one of the stops with an estimated time equal to the transaction time based upon a uniform distribution.

2.11 Imputing Destinations

Imputation of trip destinations has been performed using the simple method previously applied by Barry et al. (2001) with enhancements. We apply a simple rule that the destination of a leg is the origin of the next leg, unless there is no feasible destination for that bus route or subway system nearby. When multiple passenger trips occur on the same day, then the first origin is used as the final destination.

The linking of segments into trips is a major enhancement that gives a more accurate O-D table. Some of the linking is done prior to destination imputation and then additional linking is performed later. One-way trips are presumed to have the same destinations as other trips from that boarding location. For bus trips made by persons whose home location is known, the first trip origin is taken to be the nearest bus stop unless it is not nearby or it is the only trip made that day. Other, more complex imputation is envisioned for specific activity patterns and durations and may exploit the multi-day nature of the data.

2.12 Estimating Travel Times and Linking Trips that Involve Multiple Segments with Dynamic Schedule Queries

A fair number of transit trips involve multiple segments connected by transfers including some which involve walking from one platform or stop to another. When consecutive boardings within a specified time period, such as an hour, are present in the AFC data, the segments are candidates to be tested as linked trips. In the case of New York City, these will typically involve bus to bus transfers or bus-subway combinations in either order. Rail to rail transfers will not typically require a second swipe or dip of the farecard so these will not be explicit in the transaction data. Some trips may involve more than two segments, and these can be investigated using the same methods.

Travel times are used for linking or not linking trips. For bus trips, we estimate the travel time by using the difference in estimated times for traveling between the bus stops.

For subway trips, we estimate the travel time by using a query to the schedule-based shortest path procedure (CIS), starting at the entrance station and ending at a station near the next transaction. This also provides a reasonable path within the subway system and destination station, which may not be the closest to the next transaction. This is illustrated in Figure 4 which shows two trips–one that can be linked and one that cannot.

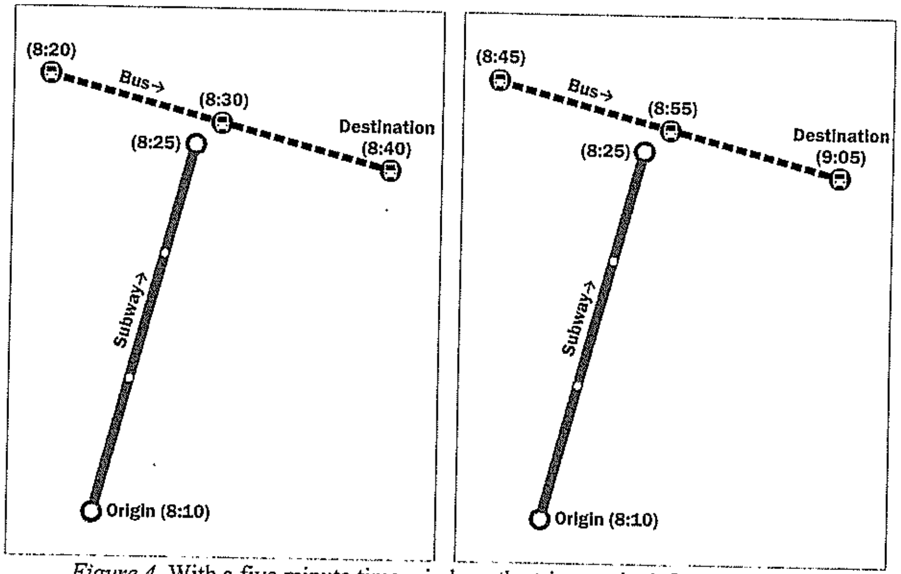


Figure 4. With a five minute time window, the trips on the left can be linked

We use the estimated travel times to decide whether a sequence of two or more passenger legs, delineated by AFC boarding transactions, should be linked together into a single passenger trip. The transaction location and time of day affects the amount of time that is allowed for walking and waiting for the next bus.

2.13 Assigning Destinations for Remaining Trips

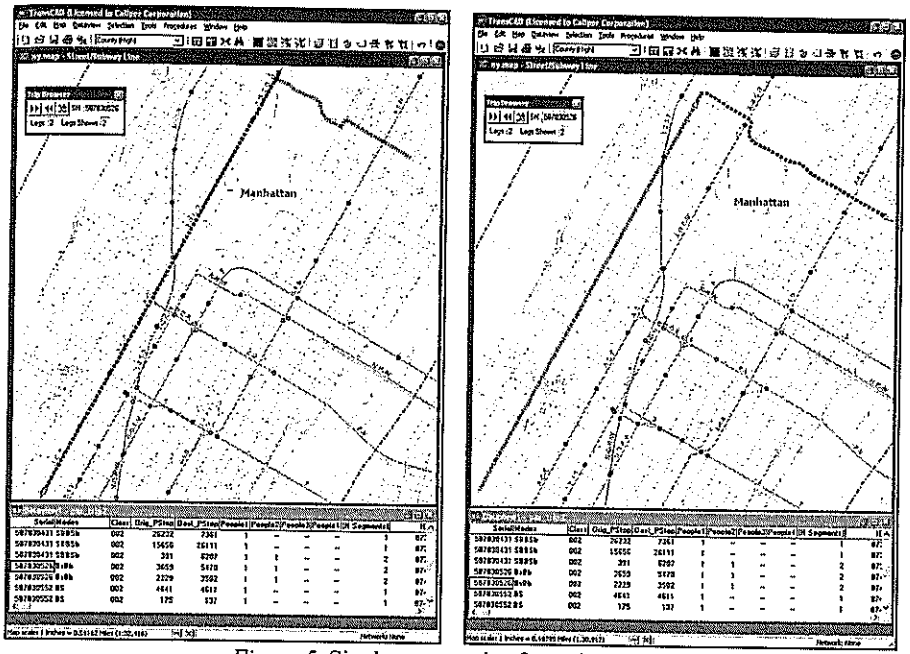
For passenger trips which do not yet have a destination assigned, we uniformly sample from all the trips starting at the same origin with destinations to impute a destination.

At the conclusion of this process, we have assigned origin and destination stops for most of the AFC transactions. A further post processing allocates each trip to an origin and destination transportation analysis zone (TAZ).

2.14 Trip Queries, Summaries, and Analysis

Once geocoded, linked, and destinations imputed, the file of trips can be queried in any fashion to develop O-D tables by time period, and arrival and departure profiles by location. The arrival time can be imputed from the destination and the schedule or use a network based measure for the elapsed time between the boarding and the destination stop.

Trip patterns can be followed for a single person for a whole day. For example, Figure 5 shows the two trips made by one person in a day.


TransCAD (Licensed to Caliper Corporation)
ny.map - Street/Subway Line
Trip Browser
Legs: 2 Legs Shown: 2
Manhattan
Map scale: 1 Inches = 0.51162 Miles (1:32,416)
Network: None

Serial/Modes	Class	Orig_PStop	Dest_PStop	People1	People2	People3	People4	N Segments	I[
587030431 5885b	002	26222	7361	1	–	–	–	1	07:
587030431 5885b	002	15656	26111	1	–	–	–	1	07:
587030431 5885b	002	391	6207	1	1	–	–	2	07:
587030526 0i0b	002	3659	5470	1	1	–	–	2	07:
587030526 0i0b	002	2229	3502	1	1	–	–	2	07:
587030552 BS	002	4641	4612	1	–	–	–	1	07:
587030552 BS	002	175	537	1	–	–	–	1	07:

Serial/Modes	Class	Orig_PStop	Dest_PStop	People1	People2	People3	People4	N Segments	I[
587030431 5885b	002	26222	7361	1	–	–	–	1	07:
587030431 5885b	002	15656	26111	1	–	–	–	1	07:
587030431 5885b	002	391	6207	1	1	–	–	2	07:
587030526 0i0b	002	3659	5470	1	1	–	–	2	07:
587030526 0i0b	002	2229	3502	1	1	–	–	2	07:
587030552 BS	002	4641	4615	1	–	–	–	1	07:
587030552 BS	002	175	137	1	–	–	–	1	07:

Figure 5. Single person trips for a single day

AFC data give a time-dependent stop boarding profile. This can be used to define the peak hour more precisely and to support dynamic models of route choice. Figure 6 illustrates PM subway boardings by subway station.

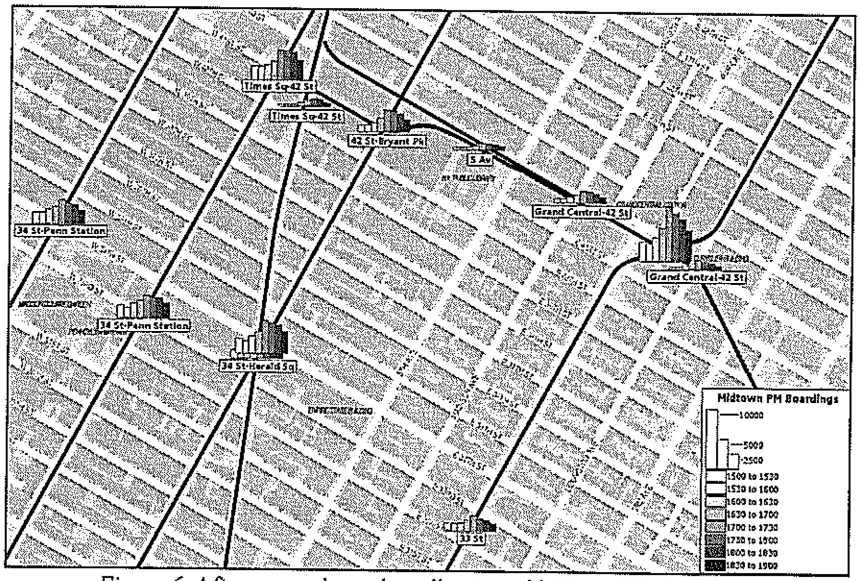


Figure 6. Afternoon subway boardings at midtown subway stations

Day to day AFC data is available for the same travelers permitting an understanding of the variability patterns of transit use and route choice. In the future it may also be possible to observe dynamic day-to-day adjustments in behavior.

3. USING AFC DATA TO IMPROVE TRANSIT ASSIGNMENTS

An important thrust of this research has been to explore the use of AFC data in producing more realistic transit assignment models. We have only just begun this effort, so we will outline some of our thinking about the most useful approaches.

3.1 Background on Transit Assignment for New York City Transit Systems

New York city transport has been using a stochastic user equilibrium assignment in TransCAD since 1989 and reported good success (Slavin et al. 1991). The model for the 5 boroughs of the city performs sub-mode split by allocating trips to a combined bus and subway network. The core methodology was developed by Caliper and followed two smaller applications of SUE for commuter rail forecasting in the New York Region.

Transit service in New York City is characterized by a multiplicity of modes and services, most of which are overlapping and relatively high frequencies. SUE was the preferred option because the transit system is heavily utilized and capacity limitations are a significant determinant of route and system utilization. SUE also produces a reasonable split of trips between local and express subway routes which is also important for demand forecasting and operations planning. In this regard, it should be noted that many prevalent transit assignment methods are heavily dependent on service frequencies rather than overall path utilities. Local service is more frequent but slower and therefore may not be modeled effectively with more traditional methods.

An AM peak period assignment is used for operations planning forecasting. It is a generalized cost assignment that includes the fare weighted by the value of time and numerous travel time and waiting time components. The weights were derived from professional judgment and trial and error.

3.2 Using AFC Data and GIS to Improve Transit Assignments

A major thrust of our effort to improve the SUE transit assignment model is to investigate its performance more fully with the advantage of using the AFC data in different ways. This is just at a preliminary stage but we can outline our thinking about possible improvements. Broadly, we envision better data inputs, improved calibration and validation, and sufficient empirical support for the implementation of dynamic models of various types.

3.3 Improved Input Trip Tables

Using an AFC trip table or averages of AFC trip tables should be an improvement because accurate trip tables have not been previously available. This is especially true with respect to estimates of bus trips and combined bus and subway trips which have not previously been available.

To the extent that AFC data need to be scaled up to account for other trips or modified to correct destination imputations, the AFC trip tables can be used as a seed for O-D matrix estimation based on additional counts.

3.4 Improved Measures of Effectiveness of the Assignment Process

The AFC data provide the possibility of assessing the accuracy of the assignment model in predicting the shares of subway and bus trips and the share of trips that use both modes in combination. These measures can be computed for various geographies and time periods.

When there is a significant service change, AFC data should make it possible to detect impacts and compare them to those predicted by a forecasting model. Longitudinal analysis of assignment models would be a welcome step forward from purely static, cross-sectional analysis.

3.5 Improved Understanding of Traveler Behavior and Choice

AFC data enable analysis of the variability of transit tripmaking by the same travelers when the same farecard is utilized over a time frame longer than one day. One example of this is shown in Table1, which lists the departure time differences for travelers who boarded the 7th avenue IND subway station in Brooklyn between 7 and 8 am on two consecutive days.

There were 870 travelers on both days and more than 500 boarded at roughly the same time. Less than 10 percent had more than a one hour deviation in boarding times. Further analysis of this type may be useful in understanding the amount of variation in departure times and its consequent effect on the timing and duration of the peak loads on the transit system.

Table 1. MetroCards with first boarding at 7th Ave IND Station in Brooklyn

	7am – 7:54am on 4/27	
Δ=0	250	28.7%
Δ=6	252	29.0%
Δ=12	105	12.1%
Δ=18	51	5.9%
Δ=30	52	6.0%
Δ=60	79	9.1%
Big Δ	81	9.3%
Both Days	870	100.0%

As Zhao (2004) has shown AFC data can be used to model route choice and to develop estimates of importance weights for level of service variables. While there are difficulties in doing so, it may be possible to understand market segments better using AFC data.

Perhaps the most important use of AFC data will be in the calibration of transit assignment models. AFC data should make it possible to adjust the relative weights to generate better assignments. This might be done directly using the assignment model because, in the presence of capacity limitations, route choices are constrained. Therefore static revealed preference models may give misleading results. Alternatively, crowding may be incorporated in future discrete route choice models.

3.6 Comparison of AFC Paths and Paths Generated in the Assignment – An Illustrative Example

AFC data provide an important and possibly unique source of data for testing transit assignments and comparing alternative methods. One particularly interesting and detailed comparison is to examine the paths generated from AFC data and those generated in a transit assignment model.

To do this, we added the capability to save all of the paths generated in a transit assignment. We then compared the SUE result with the AFC data.

. We picked a popular O-D pair with an origin in Queens and a destination in Manhattan. In the AFC data we found 270 trips of which 267 took a W or N train, 2 took the W train and the M42 bus and 1 took the W train and a different bus.

For the transit assignment, the O-D pair has only 115 trips. Of these, 98 used the W or N train for some portion of the trip. However, there were many more variations generated some of involving transfers between the W and the N train which would probably never happen. While it might be possible to detect unrealistic paths without the AFC data, it is much more straightforward to assess correct behavior of the model with it.

Generally, we think that we will probably find fewer paths being utilized than those that are generated by assignment models, but this is purely speculation at this point. Further investigation of differences will be conducted in conjunction with using the AFC data to calibrate the SUE model more closely. Also, using the same (AFC) trip table for the comparisons should be helpful in understanding the results.

3.7 Dynamic Assignment Models

The logical extension of SUE is to a dynamic formulation. The motivation for dynamic traffic models is straightforward for roads (Slavin, 1996) and for congested transit systems (Nuzzolo, 2003). The case for stochastic models may be somewhat more controversial, but we will leave that discussion for another time.

AFC data facilitate implementation of dynamic models because departure time information on actual trips can be used in model development. Fabrication of this information or dynamic O-D estimation for transit may introduce errors that counteract the benefits of dynamic models.

AFC data should make it possible to test and compare alternative formulations of dynamic stochastic models and to assess the merits of schedule-based, hyperpath-based, or stochastic, path-based approaches. AFC can potentially illuminate choice sets for schedule-based approaches and can be used to adjust schedules when there are schedule adherence problems that might render schedule-based models problematic. As discussed previously, validation and comparisons of assignment methods at the path level and by time and stop locations should make AFC data a most useful discriminator of modeling methodologies.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the invaluable contributions of the NYCT project manager, Larry Hirsh, whose insights were instrumental in developing a solution to many of the difficult problems encountered in this work.

REFERENCES

1. Barry J.J., Newhouser R., Rahbee A., and Sayeda S. (2001) *Origin and Destination Estimation in New York City Using Automated Fare System Data*, Proceedings of the 2001 TRB Planning Applications Conference, Corpus Christi, Texas.
2. Nuzzolo A. (2003) Schedule-Based Transit Assignment Models in W.H.K. Lam and M.G.H. Bell eds. *Advanced Modeling for Transit Operations and Service Planning*, Chapter 5, Elsevier Science, Oxford, UK.
3. Rahbee A. and Czerwinski D. (2002) *Using Entry-Only Automatic Fare Collection Data to Estimate Rail Transit Passenger Flows at CTA*, Proceedings of the 2002 Transport Chicago Conference.
4. Slavin H., Liss M., and Ziering E. (1991) *Integrated Transportation GIS and Demand Forecasting System*, Report prepared by Caliper Corporation for the New York City Transit Authority and the Metropolitan Transportation Authority, April 1991.
5. Slavin H. (1996) *An integrated, dynamic approach to travel demand forecasting*, Transportation 23, pp. 13-350.
6. Zhao J. (2004) *Rail Transit OD Planning and Analysis Implications of Automated Data Collection Systems: Rail OD Matrix Inference and Path Choice Modeling Examples*, Masters thesis, MIT Departments of Urban Studies and Planning and Department of Civil and Environmental Engineering.