Using Entry-Only Automatic Fare Collection Data to Estimate Linked Transit Trips in New York City

James J. Barry[1]

Robert Freimer[2*]

Howard Slavin[2]

[1]*MTA New York City Transit*

2 Broadway, New York, NY 10004 (Phone: 646-252-5631, Fax: 646-252-5648, e-mail: james.barry@nyct.com)

[2]*Caliper Corporation*

1172 Beacon St Ste 300, Newton, MA 02461 (Phone: 617-527-4700, Fax: 617-527-5113, e-mail: robert@caliper.com, howard@caliper.com)

[*] *Corresponding author*

Submitted on 1st August 2008

Word count: 5359 + 6 Figures = 6859

Submitted for presentation at the 88th annual meeting of the Transportation Research Board and publication in Transportation Research Record

ABSTRACT

Automatic fare collection (AFC) systems are used by many large transit systems. Most were designed solely for revenue management, but they contain a wealth of customer usage data that can be mined to create inputs to operations planning and demand forecasting models for transportation planning. More detailed information is potentially available than could ever be collected by any travel survey, assuming the transactional data can be processed to produce the desired information. Previous work in this field focused primarily on rail transit, since boardings at fixed stations are easier to locate than buses that move around. In this paper, we present a case study for MTA New York City Transit (NYCT), a transit system where a rider swipes a fare card only to enter a station or board a bus. Our work is the first to include all transit modes trips in an entry-only system, which is significantly more challenging because all the alighting locations need to be inferred and the boarding bus locations estimated. No location information (AVL or GPS) was available for their buses. We created software that processes the 7+ million daily transactions and creates a dataset of linked transit trips that can then be analyzed using GIS-based query software to create reports, maps, origin-destination matrices, load profiles and new datasets. Subway journeys are assigned using a schedule-based shortest path algorithm.

INTRODUCTION

Many large transit systems use automatic fare collection (AFC) systems, where a rider swipes a card to enter a subway station or to board a bus. Once AFC systems came into use, transportation planners realized that they contained a wealth of customer usage data that could be mined to create inputs to operations planning and demand forecasting models. Since most AFC systems were designed solely for the efficient collection of fares from a huge number of riders per day, special and often difficult processing of their transactional datasets is required to derive useful information for transportation planning (*1*).

Early work by Barry, Newhouser, Rahbee and Sayeda (*2*) showed that AFC transactions could be used to estimate subway origin-destination patterns in New York City. Rahbee then moved to Chicago, where he applied similar techniques to Chicago Transit Authority (CTA)'s system (*3*). These works focused on rail transit trips and did not include all bus trips. MIT, working with CTA and Transport for London has conducted detailed analysis of fare card data (*1, 4, 5 and 6*), typically with more and better quality information than available in New York City.

AFC transit systems can be classified by whether or not they require a swipe to leave the system. Exit swipes are required by distance based fare systems (e.g., San Francisco and London) and facilitate a more straightforward approach to recover transit trips made by riders. Other cities (e.g., New York and Chicago) use a flat rate fare system and only collect boarding information. Such entry-only transit systems require additional logic to impute the alighting locations of transit trips.

This paper presents our work for MTA New York City Transit (NYCT) to extend Barry et al.(*2*)'s methods to include all transit modes: subway, local and express buses, ferry and tramway. An early description of this effort and its relevance to demand forecasting appeared in (*7*). The goals included improving the estimation process for subway trips, generating origin-destination patterns by traffic analysis zone, creating load profiles for bus and subway routes, and extending the analysis beyond the AM peak period.

MTA's MetroCard system is an entry-only AFC system on which a rider swipes a MetroCard to enter a subway station or dips a MetroCard to board a bus, generating a unique transaction that is logged to a mainframe computer database. No transaction occurs when a rider exits a station or a bus. The system was designed in the early 1990s to collect fares efficiently from millions of riders per day with a flat rate fare structure. It was not set up to capture the details of the trips made by each rider. Saved transaction times are truncated to tenths of an hour.

The basis of our approach is that for each transaction, we attempt to identify the route and the specific boarding and alighting stops that define a trip leg. Multiple trip legs are combined into a linked trip, when it is inferred that a rider uses his/her MetroCard two or more times to complete a single journey.

The organization of the remainder of this paper is in three parts. The first part describes our methods to process the available AFC and schedule data to produce datasets of geo-referenced linked transit trips, known by NYCT as the *Citywide Transit Travel Database*. Each dataset contains all the transactions for a single day. The second part describes the user-oriented GIS-based query and analysis tools that can be used to study and analyze the trip data by creating reports, maps, origin-destination matrices, load profiles and new datasets. The final part describes how the system is being used in practice by NYCT.

CREATING TRANSIT TRIPS

This section discusses how the MetroCard transactions were processed to create geo-located, linked transit trips. Figures 1 and 2 provide a schematic flowchart of the extensive data processing required. Other inputs used include a geographic representation of transit routes, the log of actual bus trips and the subway and bus schedules.

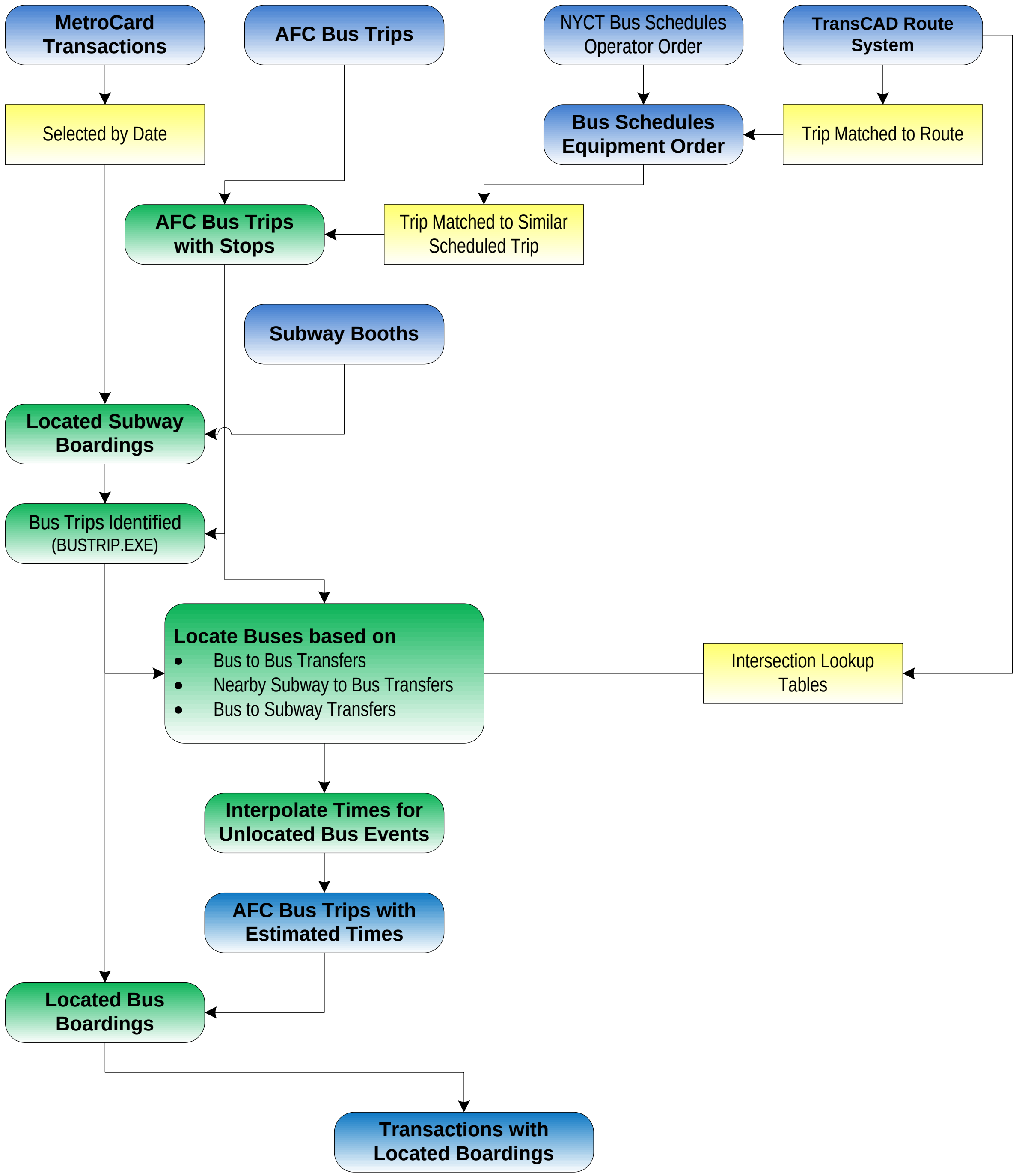


FIGURE 1 Data flow (part 1).

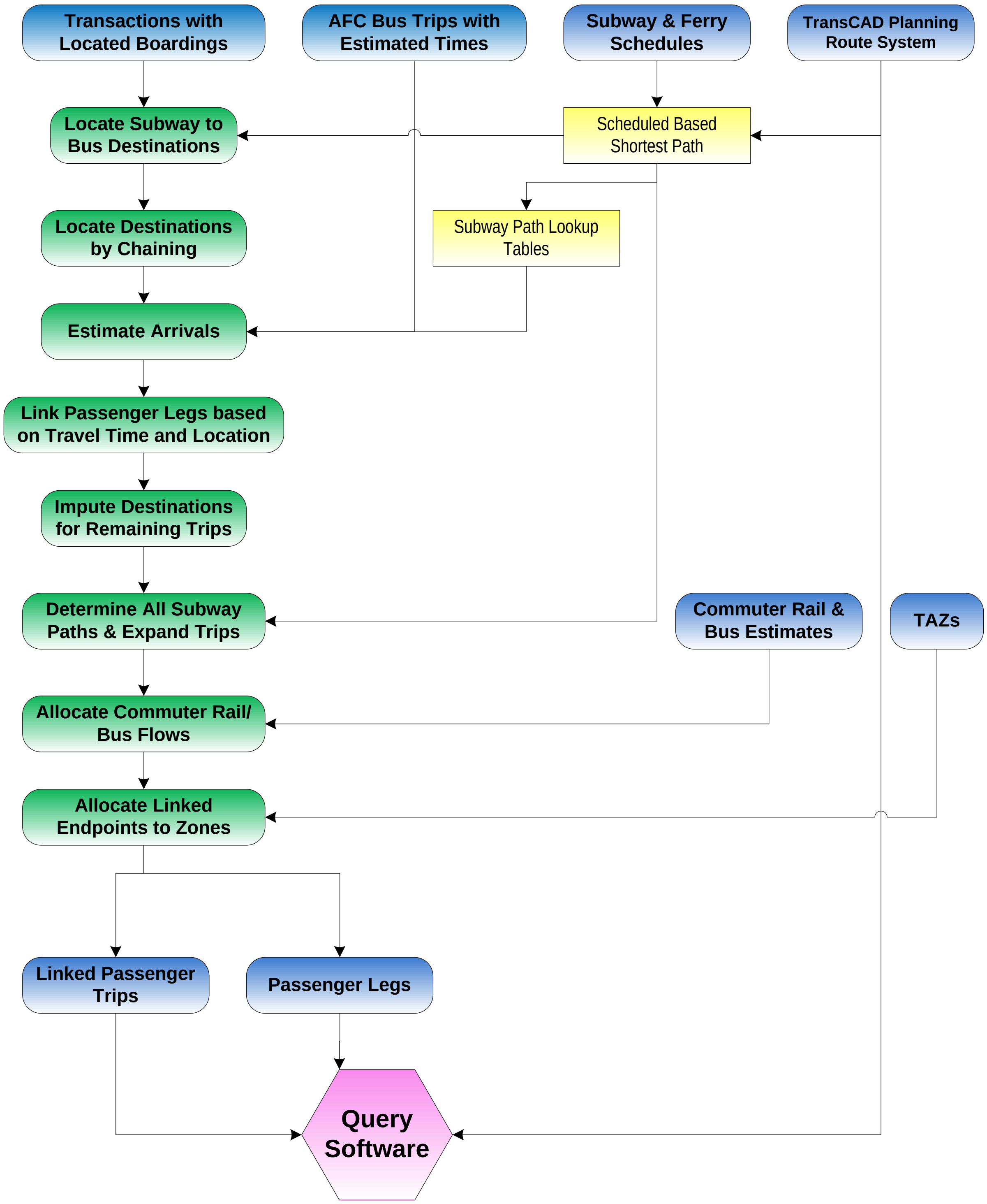


FIGURE 2 Data flow (part 2).

The procedure generates a table of linked passenger trips for a given day, each trip with an origin and destination that is either a subway station or bus stop. A second table details the component legs for each linked trip. These two tables are the principal inputs used by the query software.

AFC Transactions

The development project was based on a sample dataset of MetroCard transactions provided for a two week study period in late April 2004. Almost 95 million records were provided in the raw mainframe format and converted into a fixed-format binary table on a PC.

Unlike many cities, NYCT never shuts down completely and runs 24 hours a day, so 3am was chosen as the start and end of a day to minimize the overlapping trips. Our procedure processes a day's worth of transactions at a time. Weekdays include over 7 million transactions, corresponding to over 9 million transit legs, since many riders transfer within the subway system.

MetroCard transactions are generated for a high percentage of boardings: 97% for the subway and 86% for buses. The remaining trips are single-use MetroCards or cash fares that are not captured.

Accurate Geographic Route System

Many aspects of our processing rely on an accurate representation of the transit system, with a route for each pattern and stations/bus stops being correctly located. These include the schedule-based shortest path (SSP) algorithm used to estimate the paths through the subway system, the location of transfers between routes and estimating arrival times so that trips can be linked.

Prior to this project, NYCT had been using a route system with only the major patterns for the AM peak period. In conjunction with this project, NYCT migrated over to a new route system based on a more accurate base-map and extended to include all the subway patterns, bus patterns for other time periods and routes for other bus systems that also use the MetroCard system. Geographically accurate bus stops were added with a standardized identifier so that schedules could be matched to routes. Creating such a route system required a lot of time and effort.

AFC Bus Trip Log

The MetroCard system records a log of events for each bus with event times. These events include destination sign changes than can usually be used to determine the route pattern followed by a bus. It does not record any location information for a bus or a trip identifier that can be synchronized with the schedule. Significant effort was required to cleanup this dataset because frequently multiple records need to be combined into a single trip and some records are missing or incorrect since they are manually generated by the driver changing the sign, which does not always happen as planned.

Subway Schedules

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

These were converted into a fixed-format binary table with 596,876 records, each corresponding to a schedule event. These events occur at 2,549 unique route-stops, corresponding to fewer physical locations.

The schedules were used to update the route system to contain all the subway patterns, including short turns. They are also used as inputs to the SSP algorithm, used to determine paths through the subway system.

The MetroCard system is also used by the Staten Island Railway (SIR), the Roosevelt Island Tramway (RIT), parts of the PATH system under the Hudson and the JFK AirTrain. All of these function similar to the subway in that they have fixed stations. Routes were added for all these operators. Schedules for the SIR and the free Staten Island Ferry were added so that the SSP algorithm could be used to assign trips between Staten Island and Manhattan.

Bus Schedules

The NYCT bus schedules are similar to the subway schedules. For each bus route, there are up to four separate schedule files, corresponding to Weekdays (school open), Weekdays (school closed), Saturday and Sunday. These were converted into a fixed-format binary file with other 5 million schedule events, occurring at 63,049 unique route-stops. The schedules had to be converted from service schedule order into equipment order, so that they could be synchronized with the actual trips from the AFC bus trip log.

MetroCard transactions were also provided for the other bus operators within New York City, including Long Island Bus, NYCDOT franchised routes, Atlantic Express and Metro-North Hudson Rail Link. Each of these provided their schedules in a different format that had to be converted and integrated separately. For a variety of reasons the schedule information for the DOT routes was sometimes incomplete or inaccurate and during this project these routes were being taken over by the MTA and grouped into the MTA Bus Company.

Boarding Locations

Subway and tramway boardings are located using the turnstile fare collection information. This provides an immediate identification of the station, but not the route(s) boarded when multiple lines serve the same station complex or the rider switches subway trains. Even the direction traveled is usually unavailable, since both platforms are often accessible from the same entrance.

Bus boardings are located by estimating the location of the bus at the trip boarding transaction time. This is rather challenging in New York City, since their bus fleet did not employ automated vehicle location (AVL) technology and the transactional times are truncated to six minute intervals. Instead, the AFC Bus Trip log is combined with the bus schedules and positional information derived from certain MetroCard transactions to obtain approximate bus locations for most trips, interpolating using the distance between stops.

The most difficult portion of this project was trying to figure out a method to clean up the bus trip tables and match them against the schedule, so that they could be used to locate the bus boardings. The principal difficulty was the inconsistent number of records corresponding to an actual trip. In order to create a single trip record, there is a frequent need to combine multiple trip records. In other cases, records need to be split to correct for a driver’s failure to change the sign. We had several false starts before we settled on our final strategy.

Our original strategy to locate bus transactions involved cleaning up the AFC bus trip log to contain one record per actual trip. The scheduled trip, along with the TransCAD route identified using the schedule files, would control the cleanup process. The scheduled stop times would be adjusted to reflect the actual times observed and passenger dips would be located based upon interpolation along the trip route.

Unfortunately, a variety of data imperfections and matching issues combined to foil our cleanup and matching strategy. After several false starts and some valuable suggestions made by Larry Hirsch, we developed a successful approach based on the following observations:

1. The most important goal of matching the bus trip to the schedule is to determine the route pattern, and the location of and relative times between stops. Thus, it is sufficient to find a similar trip made by a bus, as long as it uses the same pattern. It is no longer problematic if two trips match the identical scheduled trip, since only the pattern of stops and their relative times are used by the location procedure.

2. Bus positions can be localized using transfers observed in the data. For example, if a passenger takes two bus legs within a short period of time and the routes intersect then the second bus must be near the intersection point at the transaction time. Similarly, some subway to bus locations can be used as well.

A two pass method is used to locate bus transactions. During the first pass the primary goal is to pinpoint the location of the bus at several stops during each trip by using bus to bus and subway to bus transfer points. Interpolated times are then assigned to intermediate stops and these are used to assign locations to the remaining bus dips during the second phase. Adjustments to the boarding stop used are made during the trip linking procedure.

The boarding locations developed are geographically less accurate than for subways, due to the six minute truncation of recorded boarding times, the estimation process, and the cleanup required to use the AFC Bus Trip log.

Alighting Locations

Two assumptions are made that allow us to determine alighting locations:

1. Most riders start their next trip at or near the destination of their previous trip.
2. Most riders end their last trip of the day at or near the start of their first trip of the day.

These were shown by NYCT (*2*) to be reasonable for subway riders by using a travel diary survey to validate the destinations. The methodology resulted in 90% valid destinations.

An additional assumption is made that the pattern of single-fare card users is similar to that of multiple-fare card users at a given boarding location.

A chaining procedure is applied to determine the likely alighting locations for riders with two or more MetroCard transactions on a particular day. Each transaction defines a *leg* (unlinked trip) in the passenger's journey. The location of the next transaction by a MetroCard is used to find a nearby bus stop or subway station for the alighting location of the current leg, assuming a consistent one exists. This assumes that many riders start their next movement near the conclusion of their prior movement. After the final leg, the procedure loops back to the first transaction, unless the passenger only makes a single linked trip during the day. This logic is justified by the fact that many riders return to their origin at the end of the day. Impossible destinations, those that are unreachable by the subway system or bus in question, are discarded. Destinations for single trips and other trips with no chained destination are assigned using random sampling of distributions derived from other riders who have the same trip origins.

The arrival time for each imputed alighting location is estimated using by running a SSP query for each subway journey or using the estimated stop times computed for the bus trip. Legs are combined into a single linked trip if the expected arrival time of the first leg is within eighteen minutes of the start of the next leg. The location consistency is ensured by the fact that alighting locations have not yet been assigned for inconsistent legs.

Expansion

Not all transactions will have their alighting locations determined by the chaining procedure, since it fails to assign an alighting location when either the rider made a single trip (which could possibly be a multi-modal trip) or no alighting stop is consistent with the next boarding location. Two expansion procedures that use sampling to assign alighting locations are applied. For subway transactions, an alighting stop is assigned by uniformly sampling based on the observed distribution from riders boarding at the same station with assigned alighting stations. For bus passengers, a similar approach is used based on a distribution of alighting stops for all passengers boarding at the same stop during the day for that route pattern.

Linkage

Two or more movements for a rider are linked together into a single trip, when they occur within a short period of time. The alighting times for subway trips are determined using the SSP algorithm, which utilizes the complete subway schedule and geographic representations of all route patterns, to predict the route traveled through the subway system and the time of arrival. This method also handles the use of the Staten Island Ferry, which is a free connector between the NYCT subway system and the Staten Island Railway. Bus alighting times are determined using the estimated arrival time for the bus at the alighting stop.

Zone Allocation

The procedure generates a table of linked passenger trips for a given day, each trip with an origin and destination that is either a subway station or bus stop. Nearby origin and destination zones (2000 Census Block Groups) were assigned to each trip, using a logit allocation procedure that distributes the trips to nearby zones based on a weighting of walking distance, and population and/or employment depending on the time of day. Special handling was added to assign the trips starting and ending on commuter rail or bus.

Validation

A variety of methods were used to validate the location and linkage procedure while it was being developed. In particular, we implemented some automated procedures to tabulate the located transactions so that the results can be compared against other available information for validation purposes.

For subway transactions, the entrance and exit counts were tabulated by subway station complex for the full day and by four hour periods, matching the periods used at each station to match the polling of the registers; these periods vary in their starting hour. These results can then be used for:

- Comparisons with subway register counts. Exit counts should be treated as a lower bound, since not all exiting passengers are counted.
- Checking the balance between the entrance and exit total for a station, since usually they should be close.

For bus transactions, ride check data is an alternate source of information. Counts of boardings, alightings and overall load are provided for each stop along each bus trip. Only a couple of routes were checked during our two week study period. The available ride check data was imported and successfully matched it to the scheduled trips. Then the bus transactions were tabulated by trip stop, so that the ride check data could be compared against our located transactions.

In order to compare the results from the automated location procedure with some known trips, some MetroCards were purchased and ten predetermined tours were taken, logging the boarding/alighting times and locations. The derived results were compared to the actual trips. By in large, the results were quite promising given the quality of some of the datasets. A couple of subway journeys used different pairs of trains to get between the same pair of stations, which happens when actual choices do not match the SSP query expansion.

QUERY SOFTWARE

Powerful query software was created for this project that allows almost any conceivable query to be answered; the software works in two steps: trip/leg selection and output creation. Queries can be made on either the linked trips table or the unlinked legs table.

The tool is set up to work with multiple days of data, processed from the two week study period: April 19 to May 2, 2004 that was selected as a representative period for data collection and was free from exceptional events, such as snow storms, disasters and school vacations.

The software is a custom application of TransCAD that is accessed through a new custom toolbox. The software was designed to be easy for anyone to use to get the information they desire. No special knowledge or training is required to use the system. Users can interactively select the geographic location and time period of interest.

The *Query Builder* step defines a query, combining one or more selection primitives that conceptually select a set of trips/legs (see Figure 3). The query can require that either any or all of the primitives be matched. For linked trips, the primitives include selecting by the mode, route, or pattern of a particular leg in the trip sequence; by the origin and/or destination of the trip by specifying stops, zones, Census Tracts or Boroughs; by the inclusion of a particular type of transfer between modes within the trip; or by a general SQL query. For unlinked legs, the primitives include selecting by the mode, route or pattern of the leg; by the origin and/or destination of the leg by specifying stops, zones, Census Tracts or Boroughs; by specifying the mode of either the proceeding or following leg used as part of a linked trip; by a general SQL query; or by the position of the leg within its linked trip. Queries can be saved for later reuse.

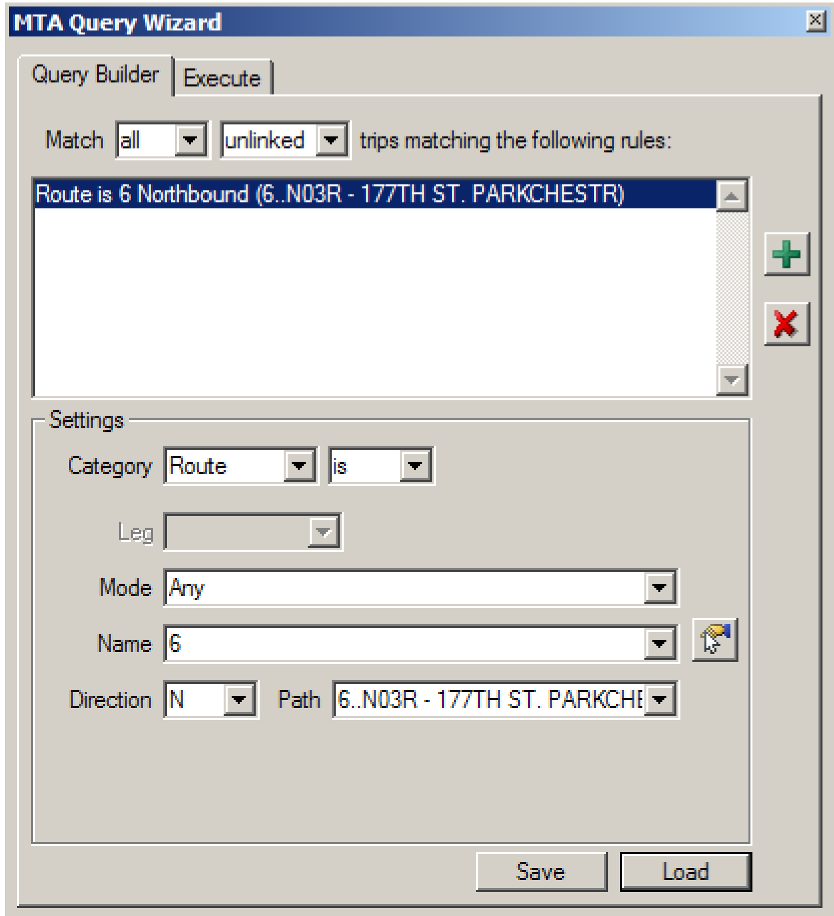
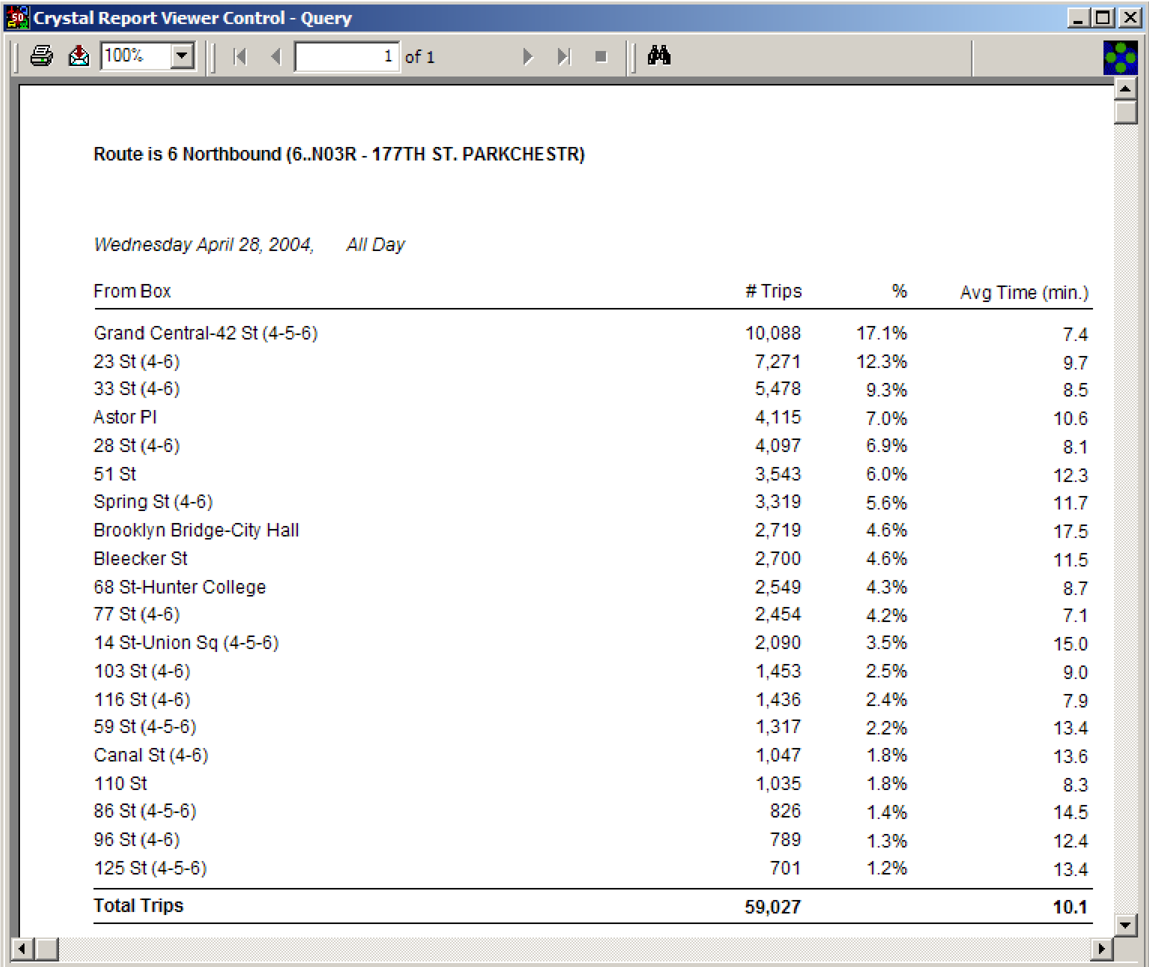


FIGURE 3 Query Builder example for the 6 train.

The *Execute* step specifies the day to use for the query, selected from the list of available days. The trips/legs can be further restricted based on a time period and/or an existing selection set of linked trips. The choices for the output include reports (see Figure 4), maps (see Figure 5), O-D matrices or a TransCAD selection set that can be used to create external tables/spreadsheets. The reports can be summarized by arrival time, departure time, mode or route of a particular leg, origin, destination, and/or O-D pair. A ridership report by either route pattern or route segment can also be produced. Maps can depict the origins and/or destinations of the trips/legs selected. Maps can also include a scaled theme depicting the ridership by street/track segment. The O-D matrices summarize the trips/legs by stop, zone, Census Tract or Borough. Besides creating export tables, the selection set can be used for general analysis in TransCAD or for examining individual linked trips in a customized trip browser that depicts each leg of the linked trip on the map (see Figure 6).



Route is 6 Northbound (6..N03R - 177TH ST. PARKCHESTR)

Wednesday April 28, 2004, All Day

From Box	# Trips	%	Avg Time (min.)
Grand Central-42 St (4-5-6)	10,088	17.1%	7.4
23 St (4-6)	7,271	12.3%	9.7
33 St (4-6)	5,478	9.3%	8.5
Astor Pl	4,115	7.0%	10.6
28 St (4-6)	4,097	6.9%	8.1
51 St	3,543	6.0%	12.3
Spring St (4-6)	3,319	5.6%	11.7
Brooklyn Bridge-City Hall	2,719	4.6%	17.5
Bleecker St	2,700	4.6%	11.5
68 St-Hunter College	2,549	4.3%	8.7
77 St (4-6)	2,454	4.2%	7.1
14 St-Union Sq (4-5-6)	2,090	3.5%	15.0
103 St (4-6)	1,453	2.5%	9.0
116 St (4-6)	1,436	2.4%	7.9
59 St (4-5-6)	1,317	2.2%	13.4
Canal St (4-6)	1,047	1.8%	13.6
110 St	1,035	1.8%	8.3
86 St (4-5-6)	826	1.4%	14.5
96 St (4-6)	789	1.3%	12.4
125 St (4-5-6)	701	1.2%	13.4
Total Trips	**59,027**		**10.1**

FIGURE 4 Report generated for the 6 train example.

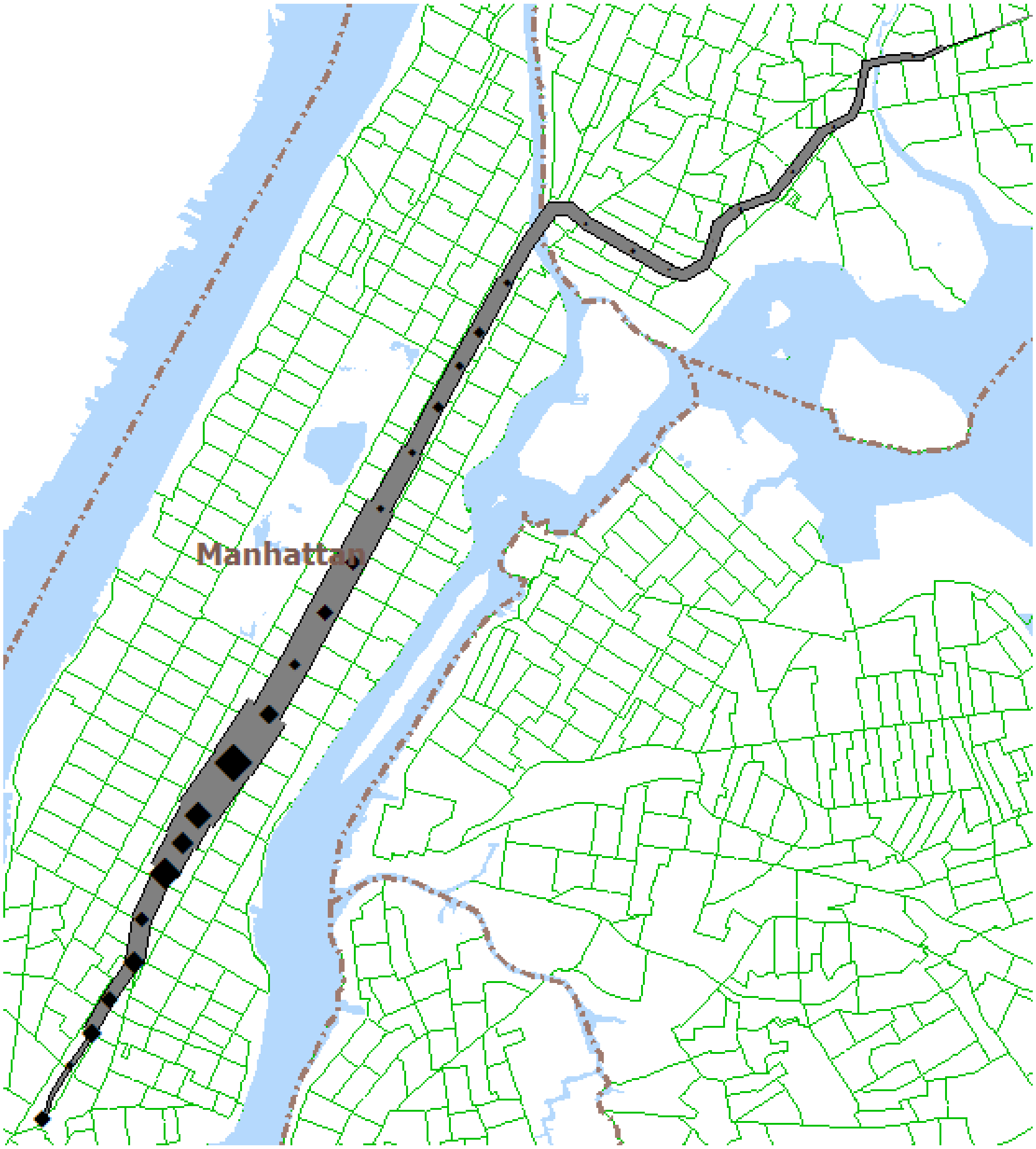


FIGURE 5 Ridership map generated for the 6 train example.

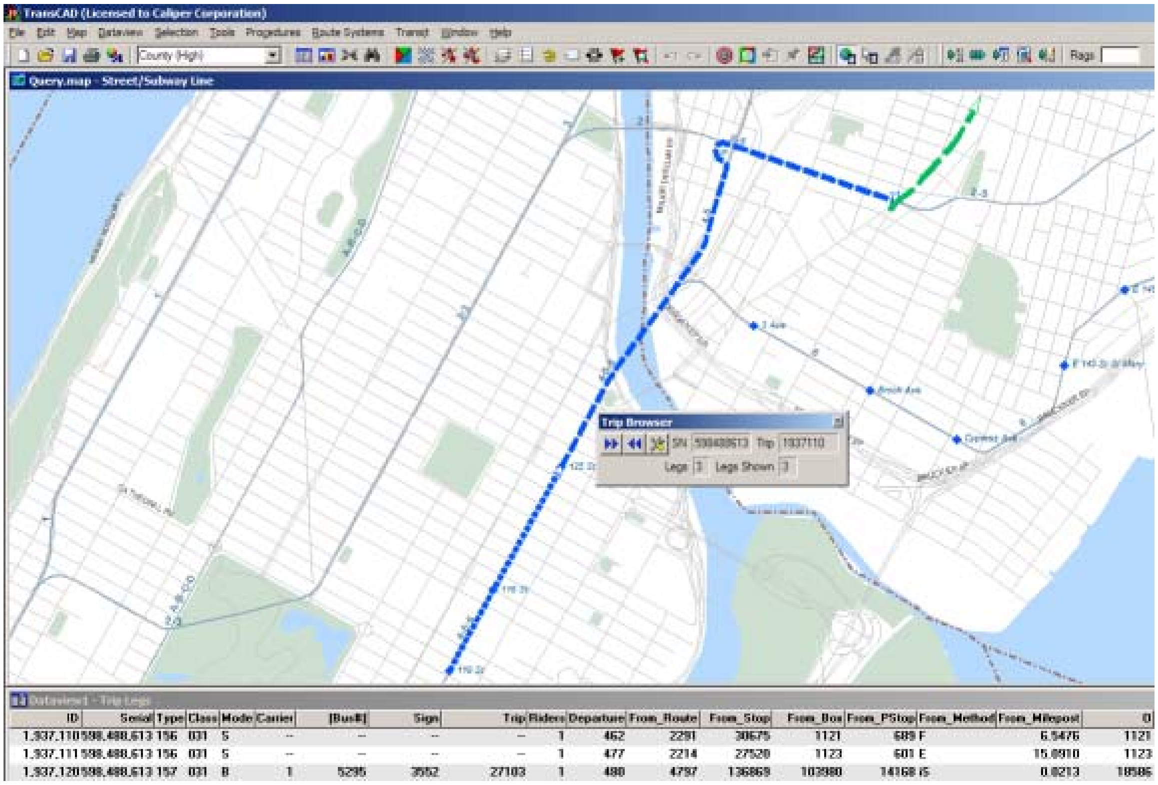


FIGURE 6 A 3-legged subway-bus trip being browsed.

CURRENT USE OF THE CITYWIDE TRANSIT TRAVEL DATABASE

The Citywide Transit Travel Database with its query tool in TransCAD provides a unique resource for short and long term transportation planning in New York City as well as for the planning and scheduling of subway and bus services by NYCT. The millions of transaction records produced daily by the MetroCard fare collection system provide unprecedented details on the spatial and temporal aspects of transit usage. In addition to New York City Transit, MetroCards are accepted on Long Island Bus, MTA Bus, Atlantic Express, Metro-North Hudson Rail Link, the Bee Line bus system in Westchester County, Staten Island Railway, Roosevelt Island Tramway, PATH and JFK AirTrain.

Obtaining basic origin and destination information by route for transit users in New York City is a daunting task. On an average weekday there are 5 million weekday subway entries and 2.5 million bus boardings. This database provides origin and destination information by route including the key distinction between linked and unlinked trips that is essential for understanding existing travel patterns and developing forecasts. This trip information at the subway station and bus stop level was also converted into traditional zone (block group) to zone trip tables for use in various travel demand models. NYCT maintains a detailed transit network model for analyzing the effects of subway and bus service alternatives on route choice, travel time savings and convenience. For regional transit analysis focused on commuter rail and subway usage, the MTA maintains a model for mode choice, trip assignment and benefit analysis (RTFM). The MPO for the New York portion of the region (NYMTC) maintains a full fledged travel demand forecasting model that incorporates features from both the MTA and NYCT models. Application examples are as follows.

Long and Short Term Service Planning

1. The zone-to-zone trip tables are essential for the demand forecasting tasks required for major additions to the subway system such as the Second Avenue Subway. Zone level trip tables are needed for projecting future travel as population, employment and land use changes occur.

2. Detailed trip information helps address short term service planning issues such as capacity and crowding constraints, fleet additions and allocation of operating and capital resources.

Reconstruction Service Planning

3. Portions of subway lines sometimes need to be taken out of service for extended periods of time for reconstruction. Information on how passengers use the line is used to help plan for the service interruption and the network model is used to estimate how passengers will adjust their route usage and if supplementary service is needed.

4. Travel pattern information is also used for determining the passenger impacts of routine late night and weekend maintenance and replacement work that requires track outages.

Bus Route Analysis

5. Used origin and destination information to evaluated proposed splitting of long bus routes in Manhattan and Brooklyn into two shorter and more reliable routes while minimizing the need to transfer.

6. Analyzed the origin and destination locations of express routes in Staten Island.

7. For Bus Rapid Transit (BRT) corridor planning, used origin and destination information to help determine existing bus trip lengths and subway transfers in the corridor.

Route Usage

8. Provided station-to-station subway route usage information during off-peak and weekends for planning diversions during reconstruction work.

9. Extracted total number of riders using each subway route by time-of-day including transfers to improve information provided to government agencies and the press.

Model Enhancements and Updates

10. The linked trip product provides for the first time a transit trip table that includes all of the bus and subway links that form a single trip. This is important for understanding and modeling the thousands of ways passengers use the extensive bus and subway network in NYC.

11. The modeling results are used to calibrate the NYCT demand model and for validation purposes when comparing against other sources of transit usage information.

12. The zone-to-zone trips tables are incorporated into the regional trip tables used in the regional models.

Household Travel Surveying

13. As part of model updating, the MTA has been conducting a household travel survey within NYC and the project team has been using this MetroCard based information to validate the origin and destination trip tables being produced by the survey.

14. Due to logistical and financial constraints the household survey is limited to collecting detailed information from a sample of adult residents of NYC and is subject to response bias. The MetroCard based information on the other hand is based on millions of daily transactions and includes all paying passengers regardless of age and residence. These two sources of transit demand information complement each other.

CONCLUSION

This project involved creating custom software that processes MetroCard data and creates geo-located linked passenger trips. These trips can then be queried using user-friendly software to produce reports, maps, extracts and matrices as needed to support various planning and operational needs. Numerous unexpected hurdles were overcome to create a functioning system that we believe is the first to actually handle buses and mixed-mode trips for a transit system with entry-only AFC data.

The processing procedure developed to generate the O-D trip database is highly complex, involves numerous intermediate steps, and many assumptions and some sampling approximations. Undoubtedly, there are errors in the trip tables produced especially for bus trips. Nevertheless, the portrait of NYCT system utilization is probably the most accurate that has ever been available and it has many potential applications.

There are many ways the data development approach could be improved. The following improvements would improve the accuracy of the results that could be obtained from MetroCard transactions. Some of them would also allow simpler algorithms to be employed for the processing.

1. Improve the accuracy of recorded MetroCard transactions to a minute, a second or better. A bus can travel a significant distance in six minutes.

2. Preserve the orderings of the MetroCard bus boardings.

3. Improve the route system to the point where it matches the schedules exactly and also has accurate geographic locations for all the stops.

4. Improve the bus trip logging system to allow an easier recovery of trip records. This could include having the drivers enter sign codes more consistently. Even better would be to have a GPS-based system to provide bus locations.

The data provide the opportunity to perform much more sophisticated types of forecasting utilization of new services by making it possible to use dynamic (i.e., time dependent) methods of transit assignment. Further calibration of the transit assignment procedure could work hand-in-hand in making fuller use of the data generated.

The wealth of data now available also makes it possible to study non-traditional commuting patterns, which previously would have been ignored by onboard surveys. For example, commuting between or within the four boroughs outside of Manhattan instead of the more common commute into Manhattan.

ACKNOWLEDGEMENTS

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

References

(1) Zhao, J. Rail Transit OD Planning and Analysis Implications of Automated Data Collection Systems: Rail OD Matrix Inference and Path Choice Modeling Examples. Master's thesis, MIT Department of Urban Studies and Planning and Department of Civil and Environmental Engineering. Cambridge, MA, 2004.

(2) Barry, J. J., R. Newhouser, A. Rahbee and S. Sayeda. Origin and Destination Estimation in New York City with Automated Fare System Data. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1817,* Transportation Research Board of the National Academies, Washington, DC, 2002, pp. 183--187.

(3) Rahbee, A. and D. Czerwinski. Using Entry-Only Automatic Fare Collection Data to Estimate Rail Transit Passenger Flows at CTA. In *Proceedings of the 2002 Transport Chicago Conference*, Chicago, IL, 2002.

(4) Wilson, N. H. M. The Potential Impact of Automated Data Collection Systems on Urban Public Transport Planning. In *Proceedings of the Second Workshop on the Schedule-Based Approach in Dynamic Transit Modelling Theory and Applications*, Ischia Porto, Italy, 2005. Forthcoming.

(5) Utsunomiya, M., J. Attanucci and N. H. M. Wilson. Potential Uses of Transit Smart Card Registration and Transaction Data to Improve Transit Planning. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1971,* Transportation Research Board of the National Academies, Washington, DC, 2006, pp. 119—126.

(6) Zhao, J., A. Rahbee and N. H. M. Wilson. Estimating a Rail Passenger Trip Origin-Destination Matrix Using Automatic Data Collection Systems. In *Computer-Aided Civil and Infrastructure Engineering, Vol. 22, No. 5*, 2007, pp. 376—387.

(7) Slavin, H., A. Rabinowicz, H. Brandon, G. Flammia and R. Freimer. Using Automatic Fare Collection Data, GIS, and Dynamic Schedule Queries to Improve Transit Data and Transit Assignment Models. In *Proceedings of the Second Workshop on the Schedule-Based Approach in Dynamic Transit Modelling Theory and Applications*, Ischia Porto, Italy, 2005. Forthcoming.