GPS Pilot Project

PHASE THREE: "MIXED MODE" DATA COLLECTION AND IDENTIFICATION OF SPECIAL POPULATION SEGMENT

Prepared for:

New York Metropolitan Transportation Council

June 2008

Contractor – STV

Principal

Dr. Catherine T. Lawson, PhD – University at Albany
Dr. Cynthia Chen, PhD – City College New York
Dr. Hongmian Gong, PhD – Hunter College
Sowmya Karthikeyan, Alain Kornhauser – ALK Technologies

Graduate Research Assistants:
Jamie Cepler, Maria Chau, University at Albany, SUNY
Haiyun Lin, CC CUNY, Salvatore Dziekan, Hunter College CUNY

Table of Contents

Project Overview: Phase Three

The **GPS Pilot Project** is designed to test the feasibility of using Global Positioning Systems (GPS) technologies for the upcoming NYMTC Household Travel Survey.

In Phase One, the research team examined previous applications of GPS technology to collect ‘passive’ data on travel behavior and integrate the data into household travel survey efforts by (1) conducting a scan and literature review on the experiences of agencies and organizations that have used GPS technology in travel surveys; (2) evaluating the most recent ‘off-the-shelf’ person-based GPS units available; and (3) researching GPS and GIS software interface and hardware data exchange.

In Phase Two, the research team conducted a series of “controlled” experiments using two “off-the-shelf” GPS units: the GlobalSat and the i-Blue. These two units, compared to other available units, appear to have sufficient capabilities to operate in the Manhattan environment, a highly dense area with an urban canyon effect.

In Phase Three, the research team evaluates the ability of GPS technology to collect data for “mixed mode” travel. In addition, an analysis using data from the 1997/98 Regional Travel Household Interview Survey attempts to identify special population segments for future GPS unit deployment. Various applications of GPS are discussed.

Picture 1.1 View from Brooklyn Bridge walkway overlooking BQE and Manhattan Bridge (Chau 2007)

Section One: GPS Performance Analysis

Introduction

In the previous report 'field test one' examined the accuracy of GPS loggers in Manhattan amidst areas of potential urban canyon effect. Since travel options in New York City journey underground, aboveground, and over water it was important for the research team to evaluate the performance of the GPS loggers en route between various modes of travel. A second field test was devised to analyze the functionality of the GPS loggers in these circumstances. In this section 'field test two' analyzes the accuracy, signal resiliency, and data capturing capabilities of the GPS logger while traveling between transit modes connecting New York City's five boroughs. Description of the field test strategy, route, equipment and details of the GPS loggers' performance are documented along with the findings from the analyses.

Field Test Two - GPS by Different Travel Modes in New York City

Description

A second field test was conducted to see how well the two types of GPS loggers work in subways, buses, railway transit, bridges, and ferries. Subway routes traveling from underground to above-ground and vice versa were also selected to measure the response time of the GPS units. The test was performed by five students using six i-Blue and six GlobalSat units during the weekend of August 18 and 19, 2007 (See Appendices 1 and 2 for *Field Test Instructions* and *Data Sheets and Reports*).

The GPS units were tested using the following routes *(Figure 1.1)*:

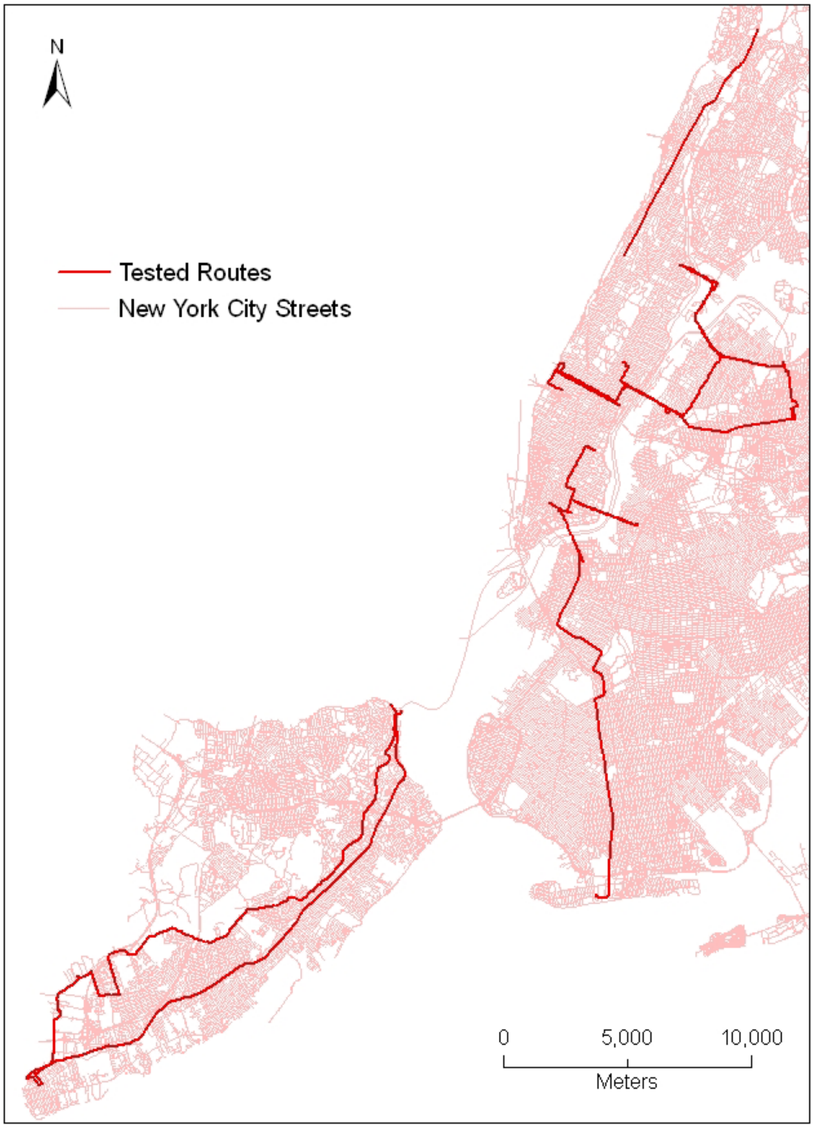
1. The '1' train from *116 Street* in Manhattan to *242 Street* in the Bronx *(Figure 1.2)*;
2. Brooklyn portion of the 'F' train *(Figure 1.3)*;
3. Ferry connecting Manhattan and Staten Island;
4. Staten Island Railway;
5. Cross town bus 'M50' in Midtown Manhattan, bus 'S74' in Staten Island;
6. Crossing the Triborough Bridge, Queensboro Bridge, and Manhattan Bridge by bus; Crossing the Williamsburg Bridge by bus and train.

Details and Analyses

Except underground in the subways, both i-Blue and GlobalSat received satellite signal on buses, elevated trains, bridges, and ferries. *Table 1.1* indicates that the accuracies of i-Blue and GlobalSat are very similar. Over half of the signal points from both i-Blue (56.54%) and GlobalSat (57.45%) fall within a 10-meter buffer zone *(Table 1.1)*. The '0 to 25' meter buffer accounts for 87.16% and 88.11% of the signal points from i-Blue and GlobalSat, respectively. In *Table 1.2*, the mean deviation of the recorded locations for i-Blue was 13.27 meters, slightly higher than the one for GlobalSat at 12.78 meters, indicating that the accuracy of GlobalSat is slightly better than i-Blue. A 'T-test' of the mean deviation for both GPS units reveals that it is significant at the 0.01 level *(Table 1.3)*. However, the maximum value was much lower for i-Blue (170.43 meters) compared to GlobalSat (2006.42 meters). The standard deviation was also lower for i-Blue.

As discussed in the report from the first field test, the software interface of the GlobalSat logger does not provide variables to exclude signal points with poor quality. All signal points from GlobalSat in the second field test were therefore included in the calculation, including a few strayed points as far as 2006 meters away from the tested route. On the other hand, poor signal points from i-Blue were excluded that were indicated as "no fix" or "estimated" for the variable VALID with less than '4' satellites used and '10 or above' for the variable HDOP. This substantially reduced the maximum value and standard deviation for i-Blue.

Figure 1.1: Routes in the 2nd Field Test



Compared with the first field test, the accuracy of i-Blue and GlobalSat improved significantly in the second field test. While only '41% to 45%' of the signal points in the first test fell within the '0 to 25' meter buffer, the results of the second test increased to '87% to 88%'. The mean deviation in the first test ranged from '40 to 52' meters; in the second test that range decreased to '12 to 13' meters. The considerable differences in accuracy between the first and second test reveal that the urban canyon effect is insignificant in areas outside of Downtown and Midtown Manhattan.

The mean deviation in the second test could have even been lower if the subway GIS file was as accurate as the street GIS file. *Figure 1.4* is a screenshot representing the southern end of the 'F' train route in Brooklyn which was used to calculate the mean deviation. Most of the stray signal points deviating over 100 meters were measured using this GPS file. This illustrates that the geographic quality of GIS files used to analyze and model GPS data is capable of significantly affecting the results.

Table 1.1 - Number of GPS Signal Points in Each Buffer

	Buffer (meter)						
	0-25	**26-50**	**51-75**	**76-100**	**> 100**	**Total**	**0-10**
i-Blue	50,087	5,683	861	351	481	57,461	32,490
i-Blue %	87.16	9.89	1.50	0.61	0.84	100.00	56.54
GlobalSat	44,572	4,761	661	250	341	50,585	29,062
GlobalSat %	88.11	9.41	1.31	0.50	0.67	100.00	57.45

Table 1.2 Deviation of the recorded locations

Distance (meter)	i-Blue 747	GlobalSat
Mean	13.27	12.78
Count	57,463	50,585
Sum	762,506.47	646,310.29
Minimum	0.00	0.00
Maximum	170.43	2006.42
Standard Deviation	17.16	18.76

Table 1.3. T test on the deviation for i-Blue and GlobalSat

	Mean	Levene's Test for Equality of Variances		t-test of Equality of Means Assuming Equal Variances		
		F	Sig.	t	df	Sig. (2-tailed)
i-Blue	13.27	40.980	0.000	4.509	108046	0.000
GlobalSat	12.78					

In terms of the response time of the GPS units, students were asked to record the time when their subway train traveled from underground to above-ground or above-ground to underground. *Table 1.4* shows the results of the difference in time between a student's recorded time and the GPS time stamp for i-Blue and GlobalSat. A positive number indicates that the student recorded their time before the GPS unit received signals. A negative number indicates that the GPS unit received signals before the student recorded the time.

In *Table 1.4,* column '1 Train' represents the travel route, from *116 Street* station in Manhattan to *242 Street* in the Bronx, *(Figure 1.2),* (see Appendix 6 for data sheet and report). The column 'F Train' represents the other travel route in Brooklyn from *York Street* station to *Coney Island* station *(Figure 1.3)*, (see Appendices 2 and 4 for data sheets and reports). Students were instructed to carry both i-Blue and GlobalSat units and travel along the same route during their return trip. On average, both i-Blue and GlobalSat received signals 39 second after the student recorded the time. The average for i-Blue is 39.05 seconds which is slight better than the GlobalSat on average at 39.13 seconds.

Figure 1.2: Locations Recorded by GlobalSat loggers on Subway '1' Train in Manhattan and the Bronx

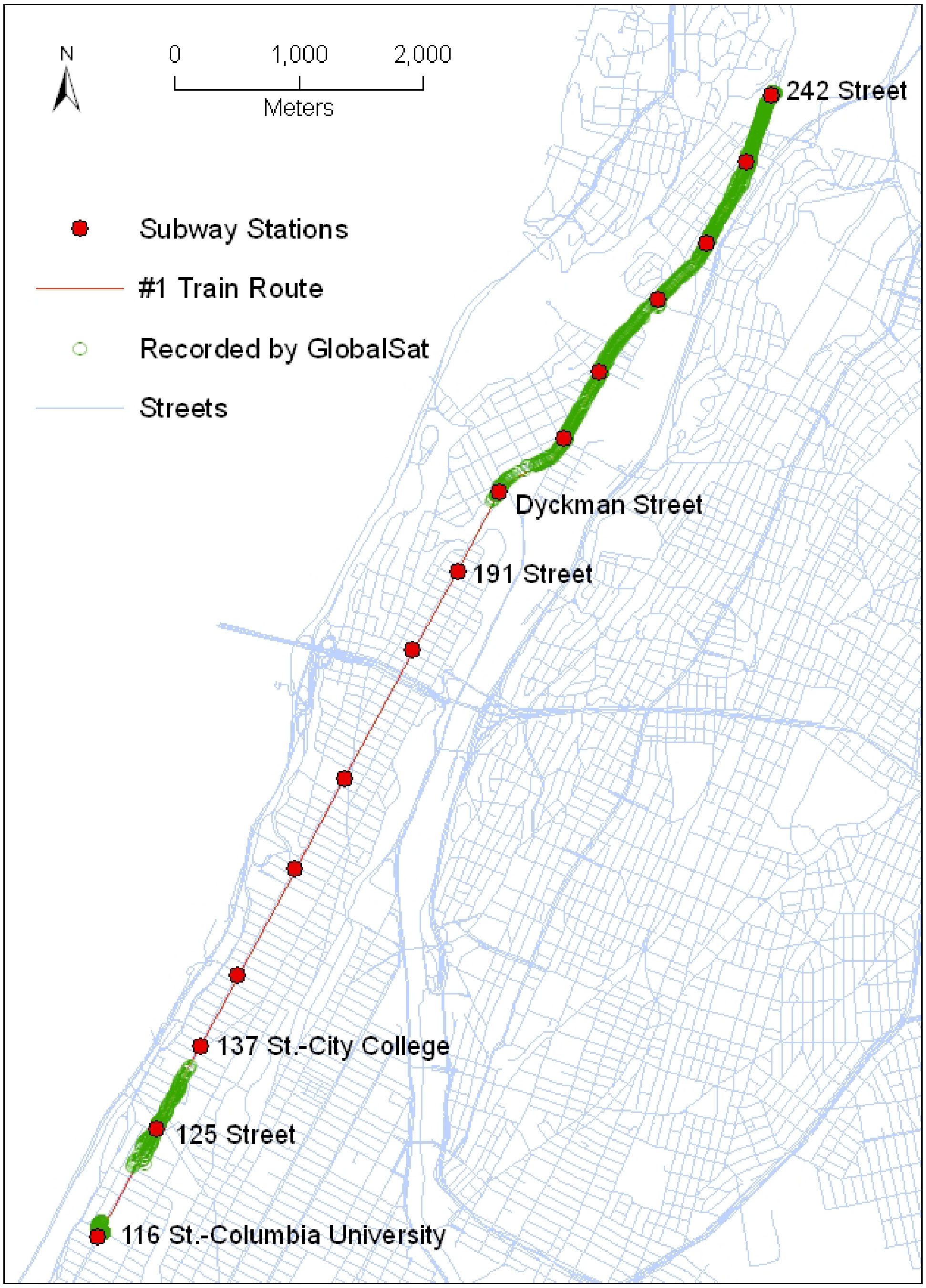


Figure 1.3: Locations Recorded by i-Blue loggers on Subway 'F' Train in Brooklyn

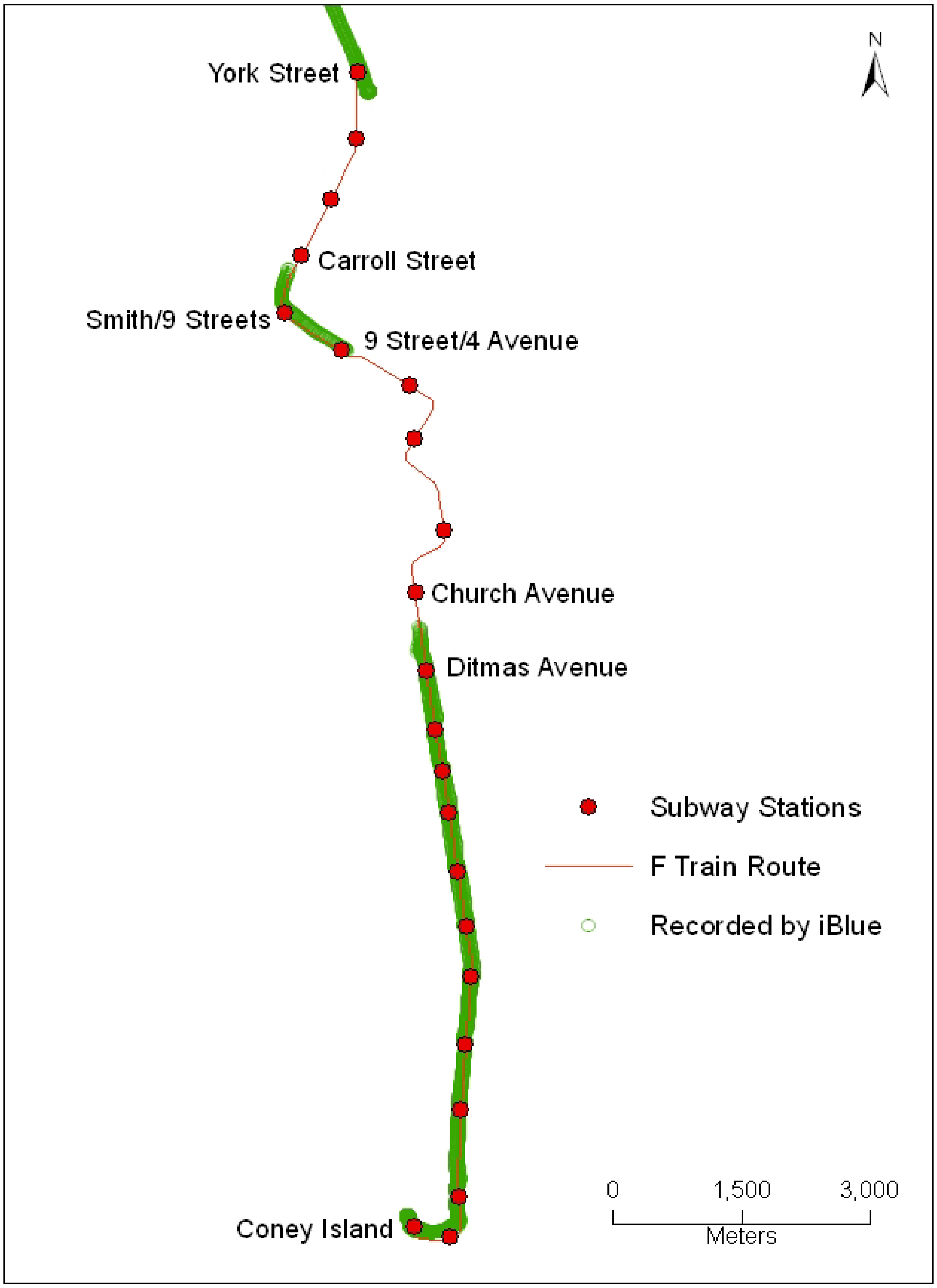
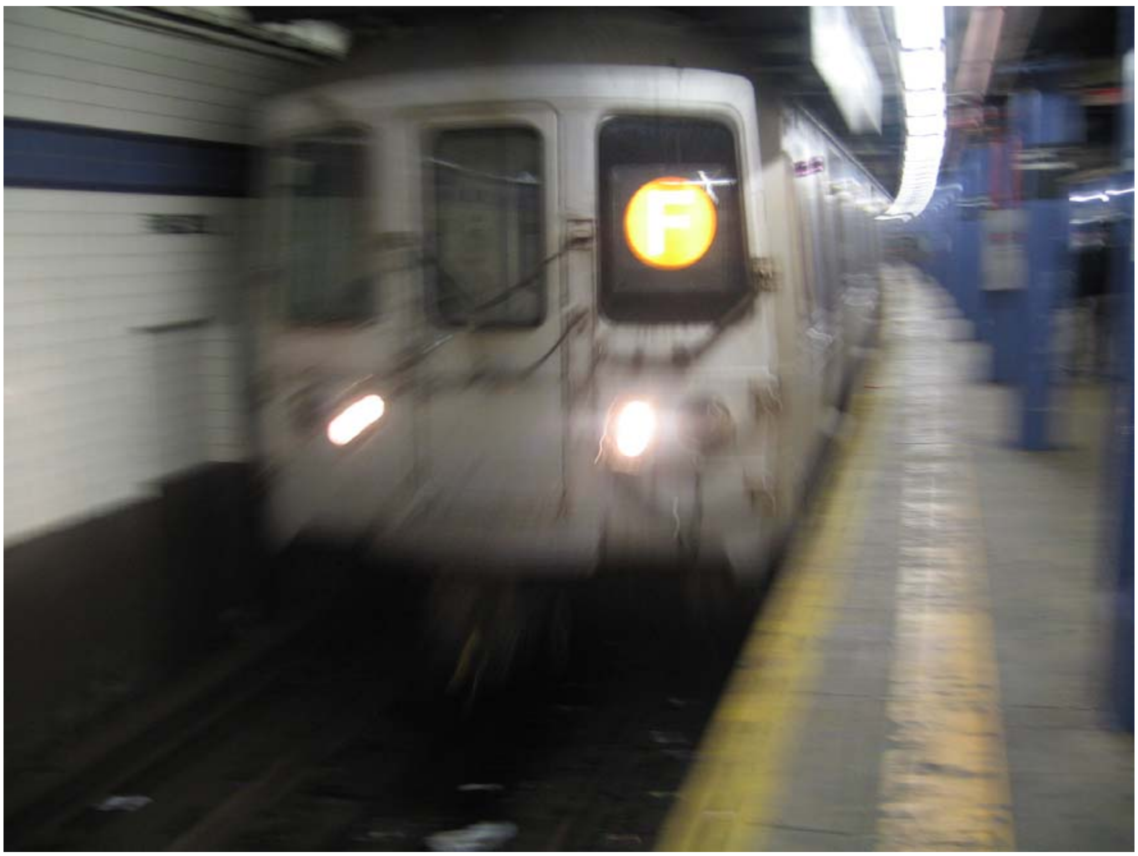


Table 1.4 Difference between student's recorded time and GPS time stamp

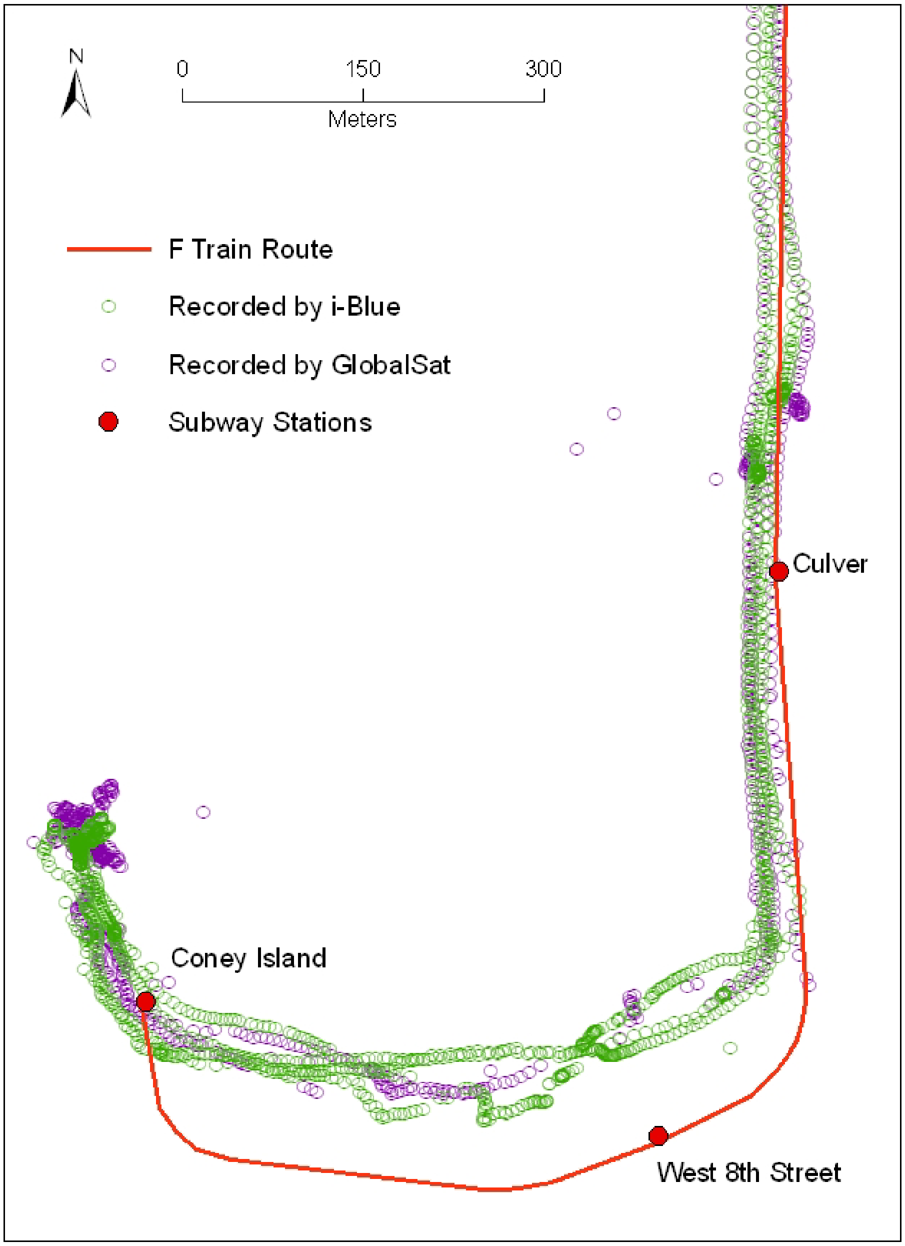
	Difference in Time (second)							
	i-Blue Unit #				GlobalSat Unit #			
	'1' Train**		'F' Train***		'1' Train**		'F' Train***	
Traveling to	4003	4007	4009	4035	1034	1035	1038	1042
above ground	0	3	37	16	0	3	88	105
underground	23	21	26	48	20	17	8	-18
above ground	19	-7	42	24	24	-1	18	38
underground	82	79	31	-8	74	78	-13	29
above ground	103	99	*	10	106	103	15	38
underground	83	82	*	46	79	80	43	5
Average difference in time	39.05				39.13			

** No data was recorded. The student participant forgot to charge the GPS unit the night before the trip and turned off the GPS toward the end of the trip.*
***Number 1 train from '116th Street' station in Manhattan to '242 Street' station in the Bronx*
***'F' subway train in Brooklyn from 'York Street' station to 'Coney Island' station*



Picture 1.2 'F' Train on East Broadway Station Platform (Chau 2007)

Figure 1.4: The Inaccurate Subway GIS File



Conclusion

In the second field test, i-Blue and GlobalSat were tested extensively using different travel modes in New York City. Both GPS loggers worked well on buses, elevated trains, bridges, and ferries. Although i-Blue and GlobalSat were not able to receive satellite signals underground in the subways, they both were able to detect satellite signals within 39 seconds on average, after emerging above ground. The accuracies of both GPS units in the second field test had improved significantly compared to the first field test, indicating that the urban canyon effect is insignificant in areas outside of Downtown and Midtown Manhattan. The accuracy from the second field test could have been better if the subway GIS file used in the calculation were as accurate as the street GIS file. Comparing i-Blue and GlobalSat, GlobalSat is slightly more accurate than i-Blue, whenever the satellite signals is good enough to be logged. However, i-Blue consistently recorded more location points than GlobalSat and is preferable especially in areas where GlobalSat is unable to log data.

Section Two: Identifying Special Population Segments

Introduction

Transit Users

A major problem of the 1997/1998 household travel survey was the considerable number of misreported transit records. During the post-processing phase of the survey, two updates were performed on the transit data records to correct the issue. The types of issues and the procedure applied to correct them were not reported in any of the known documents. Despite this, the research team was able to identify the subjects who misreported their transit records and compare them to those who did not. The research team was able to discern the socio-economic and demographic attributes associated with this population segment.

Transit users are categorized as persons who had used public transit at least once on the survey day. In order to identify the socio-demographics of those transit users who misreported their transit records, a binary logit model was estimated. Table 2.1 reports the findings of these attributes. The model uses '*transitflagpass12upd*' as the dependent variable. This variable was designated the value of '1' if the subject's transit record was updated during pass 1 or pass 2, otherwise it was designated the value of '0'.

Table 2.1. Binary logit model for transit users who tend to misreport transit trips[1]

Variables	Variable Definition	Estimate	t-ratio
Intercept		-1.31	-6.93
Hhsize	Household size	-0.08	-2.43
Worktransitmodeshare	Census tract's share of transit for work trips	2.34	7.32
Male	If male, male = 1; else, male = 0	0.29	3.52
Licensed	If has a driver license, licensed = 1; else, licensed = 0	0.51	5.36
Full-time	If has a full-time job, full-time = 1; else, full-time = 0	0.31	3.14
Finance	If works in finance, finance = 1; else, finance = 0	0.33	2.28
Inc25_35	If 25k < household income < 35k, then inc25_35 = 1; else, inc25_35 = 0	-0.45	-3.23
Black	If black and non-Hispanic, black = 1; else, black = 0	-0.29	-2.20
White	If white and non-Hispanic, white = 1; else, white = 0	0.17	1.52
Asian	If Asian/Pacific Islander, asian = 1; else, asian = 0	0.68	3.19
Manhattan	If lives in Manhattan, Manhattan = 1; else, Manhattan = 0	0.78	4.81
Bronx	If lives in Bronx, Bronx = 1; else, Bronx = 0	0.05	0.27
Kings	If lives in Kings, Kings = 1; else, Kings = 0	0.55	3.03
Queens	If lives in Queens, Queens = 1; else, Queens = 0	0.52	2.46
Regtransit	If the subject is a regular transit user (the subject answers that his/her usual mode to work is transit), Regtransit = 1; else, Regtransit = 0	0.90	7.67
Agebelow20	If the subject is younger or equal to 20, agebelow20=1; else, agebelow20=0	-0.18	-1.17
Age21to35	If the subject is between 21 and 35, then age21to35=1; else, age21to35=0	0.23	1.91
Age36to55	If the subject is between 36 and 55, then age36to55=1; else, age36to55=0	0.26	2.18
Totveh	Total number of vehicles owned by the household	0.07	1.57
	Log-likelihood of the constant only		**-2368.36**
	Log-likelihood of the above-described model		**-1933**
	Number of observations		**1912**

This model proved satisfactory. The 'log-likelihood of the above-described model' is -1933, representing a 19% improvement over the constant-only model.

[1] Only transit users are selected as part of this analysis.

Table 2.1 reveals two significant variables which characterize the transit-flagged respondents: '*worktransitmodeshare*', measures the percentage of people in the census tract who uses public transit to commute to work, and '*regtransit*', measures if the respondent's usual mode to work is public transit. Both estimates are positive, indicating that respondents who reside in a transit-oriented neighborhood[2] or use public transit regularly are more likely to misreport transit records. Both of these two variables can be viewed as 'exposure effects'. A possible cause contributing to the 'exposure factor' may be due to the fact that the call center was located physically outside of the New York area. Since the call center may have been unfamiliar with the complex transit system that exists in the city they might have been more likely to accept incorrect responses rather than probing the respondent for an accurate location. The remote survey-taker may have mis-recorded or misunderstood the local participant's response. This could have affected the number of transit-flagged records added by errors from an 'exposure factor'. One recommendation in response to this is to have the survey firm work with local call centers to reduce the extent of this problem.

Other characteristics of transit-flagged respondents in the model suggest that 'being male', 'being between 36 and 55 years old', 'having a driver's license', 'having a full-time job', 'working in the financial industry', or 'being Asian' increases the probability of misreporting transit records. While 'having a larger household size', 'having a household income between 25k and 35k', or 'being Black (non-Hispanic)' decreases the probability of misreporting transit records. There is also a '*Significant Area*' effect, even after controlling for exposure and socio-economic/demographic factors. The model reveals that subject living in Manhattan are more likely to be transit-record flagged, followed by those living in Kings (Brooklyn), and then Queens.

Young Males

Existing literature suggests that young males have a particularly low response rate[3]. Since the sampling plan for the 1997/98 household travel survey was not based on socio-demographics, it was impossible for the research team to determine whether a particular population segment was under-sampled or not. A comparison was performed on the proportion of young males (ages 16 to 24) using the 97/98 sample and the 2000 Census. The tabulation revealed that for the 97/98 sample 1,017 out of 27,369 people (~3.7%) surveyed were young males (ages 16 to 24), compared to about 6.2% from the 2000 Census. This result suggests young males were under-sampled, if the sample is compared to the 2000 census, however the binary logit regression analysis described above does not suggest that young males tend to misreport transit records.

Conclusion and Recommendation

From the findings and analysis on the 1997/98 dataset and the literature review, the following population segments are identified where the use of GPS survey might improve the response rate:

From analysis on the 1997/1998 household travel survey data:

- Male
- 36 to 55 years old
- Full-time job
- Driver license
- Working in the financial industry
- Asian

From literature review (Murakami, 2008)

- Young males

---

[2] Measured by a high share of public transit users for work trips

[3] Presentation by Elaine Murakami on January 11, 2008 in the Survey Workshop at NYMTC.

Sensitivity Analysis

Based on the outcome from the binary logit model, a sensitivity analyses was also conducted estimating the probability of misreporting transit records, given a person's socio-demographics and neighborhood attributes. The results of this analysis are shown in Table 2.2. For example, the probability of misreporting transit records is 0.99 for single Asian males between the age of 36 and 55, living in Manhattan, with driver's license, working full-time in the financial industry, and uses transit as a regular mode to work. While the probability of misreporting transit records is only 0.34 for young females, under 20 years old, living in a 3-person household in the suburbs, with driver's license, working full-time. A similar sensitivity analysis may be considered by NYMTC or the survey consultant in deciding who should be included as part of the GPS survey.

Table 2.2: Sensitivity Analysis: Probability of Misreporting Transit Records based on Socio-demographic/Neighborhood Attributes

	Probability of misreporting	**0.99**	**0.98**	**0.93**	**0.91**	**0.85**	**0.70**	**0.57**	**0.34**
Attributes of those whose probability of misreporting	**Household size**	1	1	1	1	1	1	3	3
	Share of transit for work (census tract level at home location)	0.9	0.7	0.5	0.5	0.5	0.5	0.5	0.1
	Male	1	0	0	0	0	0	0	0
	With a driver license	1	1	1	1	1	1	1	1
	With a full-time job	1	1	1	1	1	1	1	1
	Working in finance industry	1	1	0	0	0	0	0	0
	Having a household income bet. 25-35k	0	0	0	0	0	0	0	0
	Black	0	0	0	0	0	0	0	0
	White	0	0	0	0	0	0	0	0
	Asian	1	1	0	0	0	0	0	0
	Living in Manhattan	1	1	1	0	0	0	0	0
	Living in Bronx	0	0	0	0	0	0	0	0
	Living in Brooklyn	0	0	0	1	0	0	0	0
	Living in Queens	0	0	0	0	0	0	0	0
	Using transit as a usual mode to work	1	1	1	1	1	0	0	0
	Below 20	0	0	0	0	0	0	1	1
	Between 21 and 35	0	0	0	0	0	0	0	0
	Between 36 and 55	1	1	1	1	1	1	0	0
	Total number of vehicles	0	0	0	0	0	0	0	0

Section Three: Applications for GPS

Introduction

Compared with the traditional CATI surveys, GPS-equipped data surveys can provide accurate information on the location of the subject's second to second travel route, however, the trip's *purpose* or *mode of travel* isn't explicit. These two pieces of information are essential in transportation modeling. An additional modified survey is often required in accompanying the GPS-equipped survey. In this section, an initial review of existing modified surveys for GPS data collection effort are examined to assess the equipment and methodologies used.

GPS Modified Surveys

Modified surveys for GPS-equipped data collection have been conducted in at least three locations: Chicago, Japan[4], and New South Wales, Australia. An examination of each survey is reviewed in the subsequent paragraphs.

GPS Pilot Survey Project, University of Illinois, Chicago

The Chicago project was a pilot survey conducted by Dr. Kouros Mohammadian from the University of Illinois, Chicago. This project is a preliminary example of how maps can be used in a household travel survey. This review is based on a conversation between Dr. Chen (CCNY) and Dr. Mohammadian (UIC) on November 15, 2007 and an assessment of the survey process through a video program provided by Dr. Mohammadian.

Respondents were required to carry a GPS unit on the survey day. At the end of day, the respondent was asked to download the GPS data to a pre-configured computer. This computer was pre-loaded with the necessary software and drivers to process the downloaded information. The downloaded data was then connected to Google Earth. A route traversed by the respondent on the survey day was displayed in Google Earth. The pre-configured program in Google Earth consolidated the points generated by the GPS unit into a set of discrete locations that the software thought the respondent visited on the survey day. For each location shown in Google Earth, the program asks whether or not it was a valid activity. The respondent had the option of selecting 'yes', 'no', or 'add to the previous activity'. If the respondent selected 'yes' the program would ask for the following information[5]:

- activity type
- the time when this activity was decided
- others who participated with the respondent
- the alternative locations available to this activity
- the reason and the time this particular location was chosen
- the time when the starting time and the duration of the activity was decided
- the level of flexibility for the starting time and the ending time of the activity

For each trip made by the respondent, the program displayed the route in Google Earth, with information on distance, time and speed. The respondent was then asked whether or not the route and the associated

---

[4] Probe Person Survey

[5] The video demo received from UIC does not describe the situation when the respondent selects "no". In addition to "yes" or "no", the respondent can also select the option of "add to the previous activity". However, no information is provided on this option.

characteristics displayed, accurately described the trip from the previous location to the current activity location. If the respondent agreed, the program then asked them to select the mode for the trip and when and why he/she chose this mode. Similarly, information on when and why a particular route was chosen was also solicited. The respondent was also asked to rank the relative importance of each of the following factors: travel time, travel distance, amount of traffic, safety/comfort, cost, and scenery/sightseeing.

Probe Person Survey

The Probe Person Survey was designed and carried out by Dr. Yasuo Asakura and Dr. Eiji Hato in Matsuyama City during 2003 - 2005, using a Probe Person (PP) survey system. The PP survey system is a hybrid travel data collection system combining a GPS-assisted mobile phone and an Internet web diary for GIS. This system is called 'MoALs'. The components of the system and the survey flow are shown in *Figures 3.1* and *3.2*.

Before the actual survey day, each cell phone is pre-loaded with a list of locations from which the respondent can select on the survey day. There are two kinds of locations registered. The first type includes popular landmarks in Matsuyama (e.g., city hall and department stores). The second type includes locations the respondent visits often, such as home, school, workplace, friends' houses etc. Registration of the second type of locations is completed through the internet.

The GPS-equipped cell phone is pre-programmed and installed with a travel survey application. On the survey day, a survey respondent carries a GPS-equipped cell phone. The respondent is required to report their activities on the cell phone on the survey day. *Figure 3.3* illustrates the display screen of the GPS-equipped cell phone. The main menu of the cell phone presents the respondent with three options to choose from: departure, transfer, and arrival. At the start of the each trip, the respondent selects "departure", which then asks the respondent to choose the departure point, destination, and mode of transportation. If the respondent changes a mode of transportation, he/she selects "transfer" first and then selects the mode he/she will use along with the location of their transfer. At the end of the trip, the respondent selects "arrival" and enters the arrival location. To operate the cell phone correctly on the survey day, the respondent must remember to execute these operations on the cell phone whenever he/she starts a trip, changes a mode of transportation, or ends a trip. The locations on the menu include home, office, and other locations registered by the respondent in advance. The modes of transportation include walking, bicycle, passenger car, and bus. At the end of the day, the positions and times collected by the GPS-equipped cell phone are automatically "stamped" with the trip characteristics, i.e. the action/mode/location inputted by the respondent.

A blog-type internet survey called "weblog" is used to further edit the collected trip information. The blog, short for weblog, is a website designed for participants to write a description of their activities in diary format on internet. In each blog, a calendar is shown on the webpage where the respondent's diary is sorted chronically. Users can read the diary by clicking on a specific date on the calendar. Each respondent is required to create a username and a password in advance. Once they log in, they will be given a webpage which contains his/her activity and travel diary. After returning home, the respondent edits his/her WEB diary. The first step is to log in to the diary site on the internet. A list of trip records with the departure/transfer/arrival information from the GPS-equipped cell phone will be displayed on the web. Clicking on a trip will generate an actual route map for that trip on the screen. The respondent can check his/her trip records and make corrections if necessary. They can also enter other trip related information, such as trip purpose and the number of fellow passengers etc[6].

[6] It is unclear from the paper how the survey will handle trip related information when a respondents visits a location or uses a mode that is not pre-loaded.

Figure 3.1. The outline of the PP system

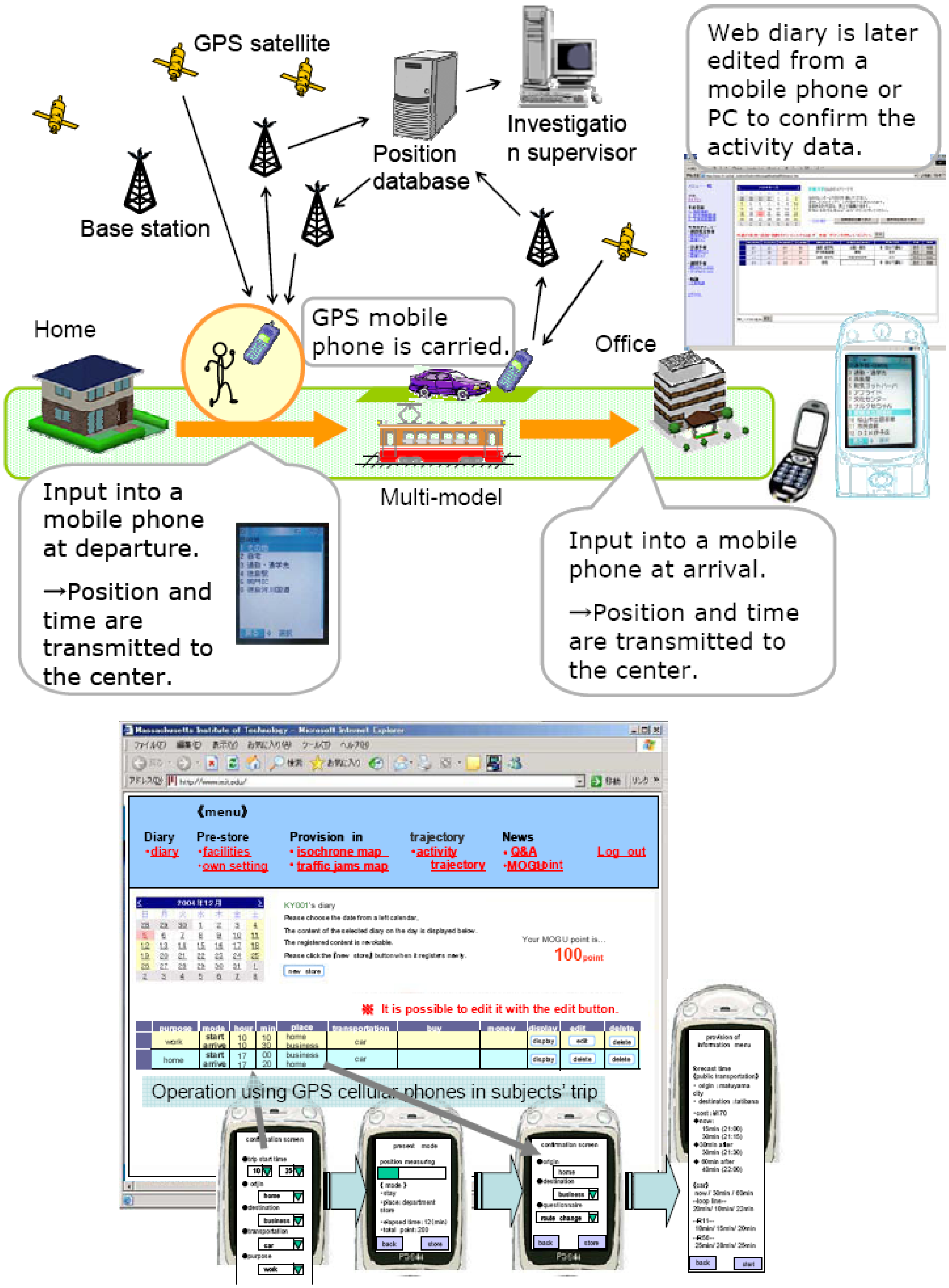


Figure 3.2: Survey flow of the PP system MoALs

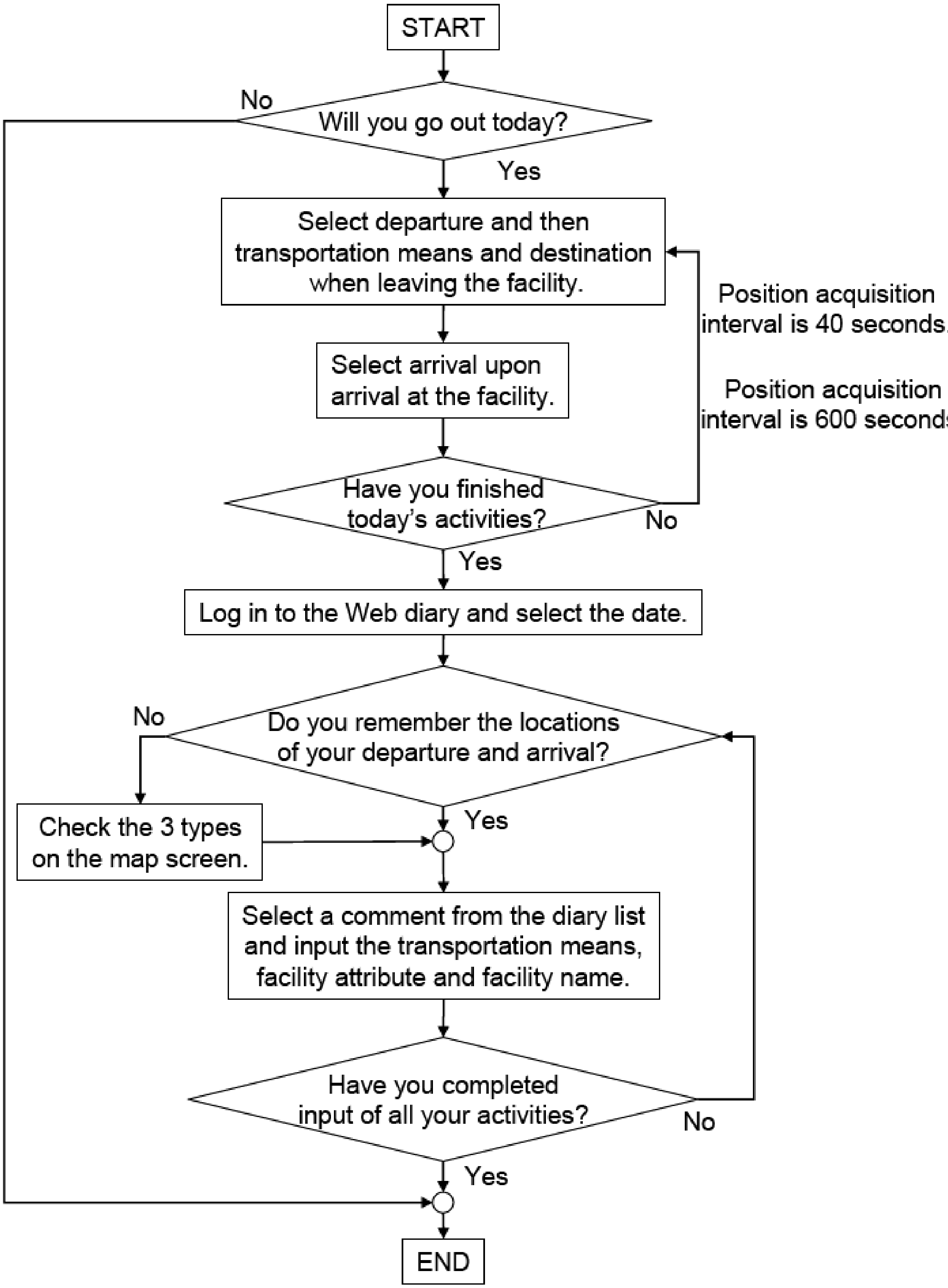
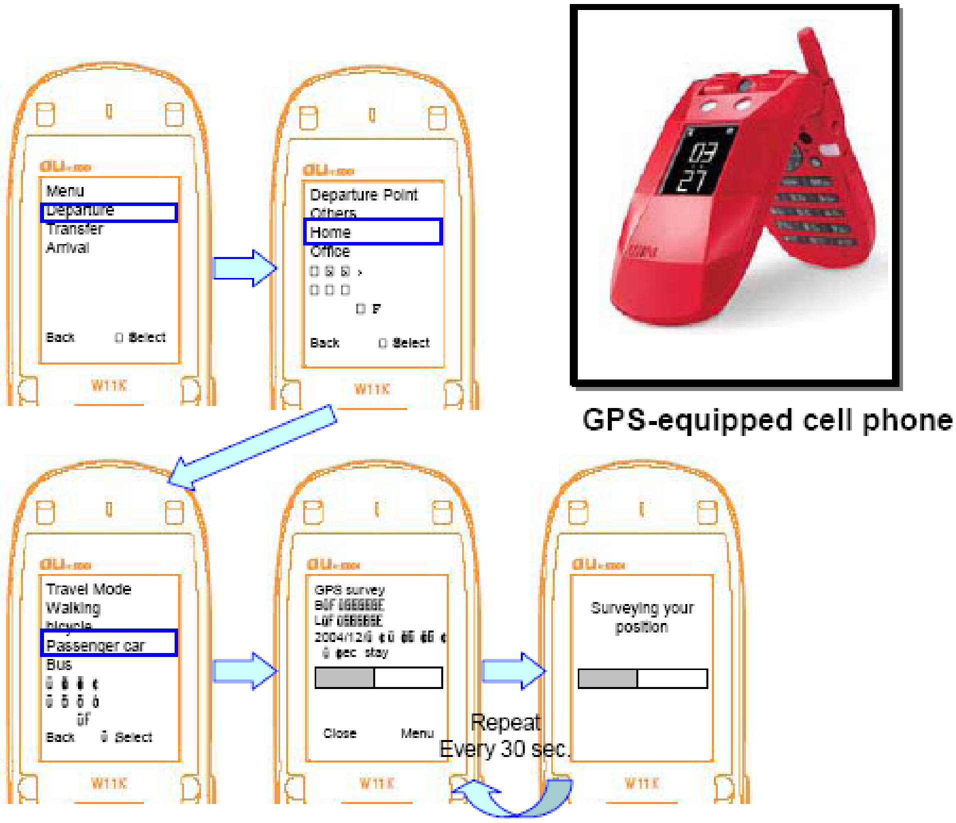


FIGURE 3.3. GPS Equipped Cell Phone Operating Screen and the Exterior of GPS-equipped Cell Phone



The time interval of collecting consecutive location information is set to be 30 s during a trip and 600 s between trips. Even if the participant forgets to operate on the phone, the location data are automatically collected at an interval of approximately 600 s. More than 20 surveys using PP survey have been conducted in Japan since 2003.

The longitude, latitude, and time of day are automatically collected through the PP survey system without the operation on the part of the respondent. Speed and distance are not directly reported from this system. It is likely that such data is calculated from the location and time of day data.

The authors also developed a procedure to identify respondents' location and route information. The procedure is described as follows. First, a Stay-or-Move Identification (SMI) is conducted, so that each location point in the data set is labeled as either a "moving point" or a "stay point". A sequence of staying dots indicates that an activity was performed at a particular location, or the respondent had stopped. Then a map matching process is used to identify the location of data points on a coded map of a transport network. The route of the trip is also identified at the same time. The authors stated that the SMI and map matching algorithms are essential data processing methods before a survey can be conducted.

The proposed PP survey system required a certain level of technological literacy on the part of respondents. This may result in an unrepresented sample (*Itsobo and Hato 2006*). An examination of the survey respondents who participated in their survey revealed that the elderly people were less likely to be involved in the PP survey, while younger people participated without any difficulties[7]. This suggests that technical/computer literacy must be considered in the GPS survey and might be a concern if a GPS/GIS/Google Earth web survey is proposed.

[7] The authors did not describe in detail how the information in the participation of the elderly and the younger participants is obtained.

New South Wales, Australia GPS Studies

Researchers at the Institute of Transport and Logistics Studies at the University of Sydney in New South Wales, Australia, have modified their GPS units to improve data quality. The device used for their studies was evaluated by a group of volunteers (Swann and Stopher 2007). Focus groups were used to learn how different types of individuals responded to the use of GPS for data collection. The results from discussions with the users were used to redesign the GPS device. For example, to respond to the descriptor "too obstructive" – the device was made slimmer. For those who were unable to determine if the device was on, a set of bright lights were added to indicate the status of the equipment. A longer life battery was installed to respond to complaints about having to recharge the unit. In addition to the changes in the physical elements of the equipment, a user friendly interface that "talks" was included. It provides information when the power is low, when the satellites are over head, etc. All of these features appear to improve the data quality by responding to the needs of the users (Swann and Stopher 2007).

To further improve the data collection process using GPS, Stopher et al. (2007) recommend a longer period of data collection using a GPS device. In their study, they found that a one-day sample was highly unstable. They suggest using a 15 to 20 day survey period.

Synthesis of and Statistics for Recent GPS-enhanced Travel Surveys

Dr. Jean Wolf from *GeoStats* shared with Dr. Chen (CCNY) the abstract written for International Conference on *Survey Methods in Transport: Harmonisation and Data Comparability*, France 2008[8]. This abstract is briefly summarized below, including a table from their document.

GeoStats was involved in six large-scaled GPS enhanced travel surveys, three of them being large metropolitan household travel surveys with an in-vehicle GPS unit for trip auditing purposes. These three surveys were conducted by the Metropolitan Washington Council of Governments (MWCOG), the Baltimore Metropolitan Council, and Chicago Metropolitan Agency for Planning. Based on available information, table 3.4 below summarizes some basic parameters for each of the household travel survey efforts.

Table 3.4: Basic Parameters of the three Metropolitan Household Travel Surveys (Source: GeoStats)

Client	**MWCOG**	**BMC**	**CMAP**
Area	DC Region	Baltimore Region	Chicago Region
Length	12 months	9 months	8 months
Timeline	02/07 through 01/07	05/07 through 01/07	03/07 through 10/07
Deployment Period	2 days, with diary on day 1	2 days, with diary on day 1	7 days, with diary on day 1
Number of Vehicles	Up to 3 vehicles per household	Up to 3 vehicles per household	Up to 3 vehicles per household
GPS Sample Size	816 households	400 households (429 with MWCOG sample)	300 households
Total Sample Size	10,000 households	3800 households (4500 with MWCOG sample)	5800 1-day diary households 5800 2-day diary households
% of Total Sample	8.16 %	10.5% BMC sample (9.5 with MWCOG sample)	5.2 % of 1-day diary households (none of the 2-day diary households)

[8] Jean Wolf, Marcelo Oliveira, and Michelle Lee, *Geostats*

Section Four: Post Processing and Mode Identification Tests

Introduction

Post-processing techniques offer an opportunity to gain value-added from GPS data streams. The GPS data collected during the student experiments from the GlobalSat and i-Blue units were analyzed by the research team at ALK for usability to identify the sequence of trips made by an individual throughout the course of a day. GPS technologies have evolved such that, given an unobstructed line-of-sight view of at least four (4) GPS satellites, the GPS recording devices should be capable of providing data sufficiently accurate with respect to position, velocity, and time for documenting trip histories. This portion of the research is an assessment of the robustness to which one can identify individual trip segments, including the start and stop location and time, the mode and the purpose. Unfortunately, the use of GPS, and any other mechanism for that matter, for the determination of position-velocity-time trip histories, introduces errors in those data. The scope of the analysis also includes an assessment of the error characteristics of the GPS recordings.

Two approaches were used: human inspection of GPS trip data and algorithmic analysis. The algorithmic analysis of GPS data was constructed from consistent patterns recognized by human inspectors. Future investigations may look at the use of neural network training techniques for a more automated approach to the identification of individual trip segments.

It became obvious early in the research that the display of the GPS data superimposed on high resolution Google Earth images greatly enhanced the quality of human inspection. For example, GPS points at which an individual entered a subway station or boarded a vehicle, was walking down a street, entered a building, boarded a bus or a car, was riding in a bus or a car or an elevated train was obvious, even in areas of rather poor GPS data, were useful for analysis. Visual human inspection also clearly identified data that contained substantial error causing the data to jump erratically or drift in a biased direction.

The human inspection technique revealed that “bad” GPS data had characteristic patterns that allow it to be readily identified and separated from “good” GPS data. Specifically, the time-space pattern of “good” GPS data (the “GPS Signature”) of one mode of travel is significantly different from the GPS signature of another mode of travel. Analysis of the GPS signature allows for the mode identification and the transition point from one signature to another identifies the time and place of the transfer from one mode to another. Thus, one can identify the home location and the home departure time, the walk to a bus stop, the location of the bus stop, the duration of the wait for the bus, the time of bus boarding, the route taken by the bus, the alighting location and time from the bus, the walk to the subway station, the location and time of entrance to the subway station, the location and time of emergence of the subway above ground (the subway route is not directly identifiable), the route taken by the elevated train, the location and times of each station stop, the location and time of descent from the elevated, the walk to the office and the location and time of entry into the workplace (trip purpose is not directly observable from the data). Such sequences are also humanly discernable in the presence of substantial “urban canyon effects” mentioned previously in Section One.

Data comparison of alternative GPS recording devices

Both GlobalSat and i-Blue output data in KML format and are compatible with Google Earth software. Additionally, each unit also output the data in text formats that could be used for analysis with Microsoft Excel (Comma delimited or Tab delimited). However, neither unit provided the data in standard NMEA format, which would allow for easy interface with ALK's transportation network data and navigation software.

The data items logged and output by both units were:

- Record number (serial number for each observation starting at 1 when the unit is turned on)
- Date (format yyyy-mm-dd by GlobalSat; yyyy/mm/dd by i-Blue)
- Time (format hh:mm:ss)
- Latitude (format ±ddmm.mmmm by GlobalSat, dd.dddddd, N/S by i-Blue)
- Longitude (format ±ddmm.mmmm by GlobalSat; dd.dddddd, E/W by i-Blue)
- Speed (km/hr)
- Altitude (meter)

In addition to the above data items, i-Blue also logged and output the following:

- RCR
- Valid (indicates quality of GPS observation – No fix/ Estimated Mode/ DGPS/SPS)
- Heading
- HDOP (Horizontal Dilution of Precision)
- Number of Satellites Used & Number of Satellites in View

The i-Blue unit is preferred to the GlobalSat unit because the former provides additional data elements necessary for our analysis. The Valid tag allows us to eliminate observations with incorrect position logs (such as "No fix" when the unit does not get a signal). The number of satellites used in the calculation of position coordinates is also a key indication. Positions calculated using four or more satellites are more accurate (3D fix) compared to those calculated with three satellites (2D fix). The data points from just one or two satellites should be ignored. Another unit of measure that suggests accuracy of the data point is the Horizontal Dilution of Precision (HDOP), which indicates the geometric strength of the satellite configuration, and consequently positional accuracy. The heading information may be useful in distinguishing between train and bus modes, as trains change their direction more gradually than a bus or a car that can make sharp angular turns.

For the above reasons, the data collected using the i-Blue units during the student experiments were used in the formal analysis.

Methodology for identifying modal trip segments

It is postulated that one of the best differentiators of different modes is the speed of travel. While in extremely heavy congestion, one may be able to walk faster than a bus or a cab, in most situations along the same route, the subway or elevated is faster than a car or cab, which is faster than a bus which is faster than walking which is, of course, faster than waiting. Thus speed is a clear differentiator of mode as is its equivalent: cumulative distance traveled. Moreover, each mode tends to have an identifiable time history signature of cumulative distance even in the presence of urban canyon error. Waiting accumulates very little distance. Walking is a very slow steady rise. Buses have more frequent stops than cars and cabs. Elevated trains have fast accumulations interspersed with very discernable stops. Subways have long stretches of no data, with then a sudden large jump in cumulative distance. Furthermore, large jumps in cumulative distances over very short periods of time can be readily identifying a GPS error. Since GPS position is more precise than GPS speed, the research team chose to focus on the time history signature of cumulative distance traveled as the mode identifier in the analysis. A change in the cumulative time history signature identifies the time and location of the modal interface. Urban canyon noise may obscure the precise time and location of a modal interface; however, map-matching to the actual physical location of subway entrances and bus stops can subsequently resolve the most likely modal interface location.

The research included a series of processing steps beginning with the preparing the data for analysis and concluding with the development of an algorithm. The following procedures were followed:

1. *Prepare data for analysis*. Import comma-delimited file from one student experiment to Microsoft Excel, and arrange into individual columns. Format Latitude, Longitude, Date and Time fields for use with formula.
2. *Identify usable data points*. Filter out data points that are either “No fix”, have HDOP values greater than 5, or with one or two satellites used. The remaining data points are from 3D fix or 2D fix and can be used with reasonable confidence in positional accuracy.
3. *Compute Great Circle Distance from each good data point to the next one*. Set distances to/from bad points to zero. Compute cumulative distance traveled from start of trip.
4. *Plot cumulative distance over time*. Color code the output by data quality (green for good 3D fix data, yellow for 2D fix, gray for bad data, and gaps in white indicating no data when GPS unit was switched off).
5. *Create a new KML file with just the good data points and including the TimeStamp reference*. This new KML file can be used with the Time Scale feature in Google Earth, allowing us to view the GPS points within a timeframe that could be changed using a timescale.
6. *Review the intermediate output*. Study the graph and the Google Earth plots for preliminary inferences that can be made from human visual inspection.
7. *Correlate output with experimental narratives*. From the student report that describes the trip, look for examples for each mode of travel (walking, bus, underground subway, elevated subway). Study the Google Earth plot, the graph of cumulative distance over time and other data elements for the corresponding time intervals.
8. *Identify modes*. Look for patterns in the data that are unique to each mode (GPS mode signatures).
9. *Develop an algorithm to analyze the GPS data*. This process makes it possible to obtain trip information (start and end location and times, modes of travel, mode transfer locations and times) in an automated manner.
10. *Apply this algorithm to data collected from other students*. Document how the results from the analysis compare with those described in the student reports.

1.1 **Preliminary inferences**

Figure 4.1 is a display of the cumulative distance traveled over time for the entire multi-modal trip in the New York Metropolitan Region. It is important to discern the regions where there is good and useable GPS data (green and yellow) and no useable data (gray). The ‘no useable’ data may be caused by the individual being essentially completely obstructed (going into a building or a subway entrance or a very narrow street, such as Wall Street).

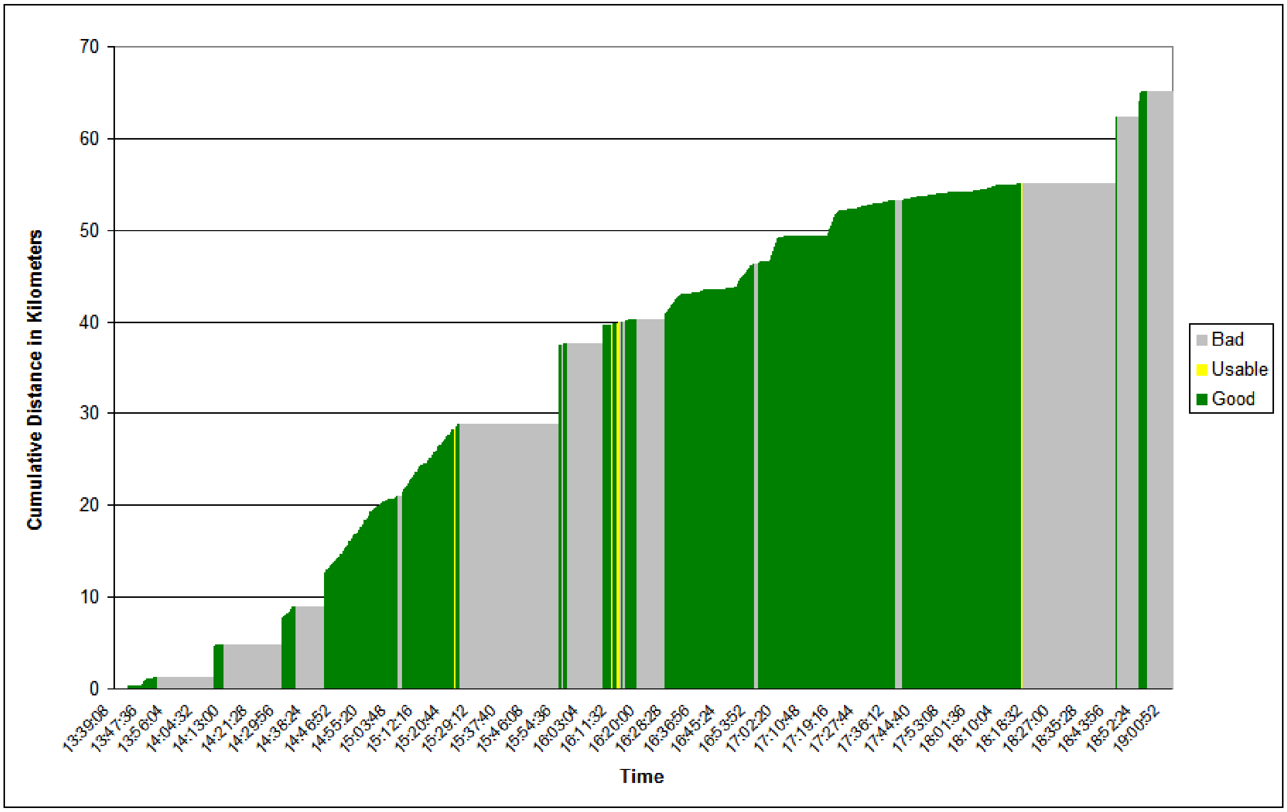


Figure 4.1 Graph of cumulative distance traveled over time

The research team made the following interpretations from studying the graph:

- The trip started around 1:45 pm and ended around 7:00 pm.
- There are no wide white gaps, which mean that the unit logged data almost continuously through the entire trip.
- The green flat areas typically correspond to person being stationary (like waiting for a bus). However, it may also correspond to traveling in a vehicle that is stuck in traffic.
- The degree of slope represents how fast the person is moving. Mild (closer to the horizontal) slopes indicate walking, while steeper slopes indicate traveling in a vehicle (bus, car, taxi, or elevated subway). Occasionally, mild slopes may also be due to slow speeds due to traffic congestion.
- Short breaks in slopes at repetitive intervals indicate stops. These can help differentiate between traveling in a bus/elevated train and in a car/taxi.
- The gray horizontal portions correspond to parts of the trip where the GPS unit could not get good signal. These would indicate person was in an underground subway or inside a building. The width of the gray column and the vertical rise at the end of the gray can help to differentiate between the two.
    i. If there is substantial vertical rise and the slopes at either end are steep (greater than walking speed), then the person is probably inside a moving vehicle going through a bad signal area (tunnel or underground)
    ii. If the width of gray is substantial, and there is a vertical rise indicating the person moved since the last good signal, and the slopes at either end are mild or flat, then it would indicate an underground subway trip, originating at the start of gray and finishing at the end of gray.
    iii. If there is no vertical rise at the end of the gray, it means the person did not travel since the last good signal; they probably entered a building and came out the same way.

1.2 **GPS mode signatures**

Based on the previous general observations, a collection of trips was analyzed. Trip segments that correspond to each travel mode were individually analyzed in order to identify their distinguishing aspects.

Each trip segment has its start and end points in time and space. The mode signatures were identified by observing the location plots on the Google Earth and the characteristics of the graph and other data elements for that time period.

The data elements that contribute to the trip segment details are:

- Time at each end
- Location (Latitude & longitude) at each end
- Cumulative distance traveled from segment start to finish
- Time interval for the segment
- Calculated average speed for the segment
- Observed speeds at the start, finish and intermediate points
- Calculated average speeds for a specified time interval (say 2 minutes) before segment start and after segment end
- Change in heading (direction) over short time intervals (say 10 seconds)

This section lists the observable (and measurable) data characteristics for each mode with illustrative examples at the end.

A. Stationary

- Cumulative distance traveled is zero
- Average speed for the segment is zero
- Segment time interval is at least 5 minutes

B. Walking

- Cumulative distance traveled is low
- Average speed, and observed speeds at any time are within 4 km/hr

C. Underground Subway

- No good signal for at least 10 minutes
- Cumulative distance traveled is more than 1 km
- Segment start or end is at a subway stop or on a train line
- Mild slopes before segment start or after segment end indicate walking to/from the subway stop.
- Steep slopes after segment start or before segment end indicate person on moving train that shifts between underground and above ground
- Calculated speed (distance covered over elapsed time) over 10 km/hr

D. Elevated train

- Reasonably good signal due to travel above ground
- Cumulative distance traveled is more than 1 km
- Segment start or end is at a subway stop or on a train line
- Mild slopes before segment start or after segment end indicate walking to/from the train stop.
- Steep slopes after segment start or before segment end indicate person on moving train that shifts between underground and above ground
- Calculated speed (distance covered over elapsed time) over 10 km/hr
- No sharp change of heading (only gradual change in direction possible in trains)
- This portion of the trip should also be checked for bus and car/taxi modes
- Discernable intermittent stops indicating elevated train station stops.

E. Bus

- Good signal due to travel above ground
- Average speed greater than walking speed (4 km/hr)
- Some observed speeds greater than walking speed (4 km/hr)
- Mild slopes before segment start and after segment end to indicate walking to/from the bus stops.
- Observed speed is zero for at least 10 seconds at segment start and end for getting on/off the bus.
- The locations at stationary segments should be at bus stops
- Frequent stationary segments at no more than 5 minute intervals are characteristic of this mode (bus stops). This would further distinguish between this mode and car/taxi
- One or more sharp change in heading would further differentiate between this mode and elevated subway (buses can make left or right turns at sharper angles within shorter intervals, while trains cannot)

F. Car/Taxi

- Good signal due to travel above ground
- Average speed greater than walking speed (4 km/hr)
- Some observed speeds greater than walking speed (4 km/hr)
- Mild slopes before segment start and after segment end
- Observed speed is zero for at least 10 seconds at segment start and end for getting on/off the vehicle.
- Segment start or end at locations other than bus stops
- Not many stationary segments that coincide with bus stop locations
- One or more sharp change in heading would differentiate between this mode and elevated subway (cars can make left or right turns at sharper angles within shorter intervals, while trains cannot)

G. Walking through a building or visiting in a building (store, bank, etc.)

- Temporary loss in good signal. Longer gaps would mean destination reached, or an underground subway trip
- Mild slopes before segment start and after segment end
- Cumulative distance traveled is low
- Average speed for the segment is low

1.3 **Examples for Walking, Subway, Elevated Train and Bus**

Walking Example: Student 2 exited the B39 bus at around 5:03 pm and got back on another B39 bus around 5:09 pm. The student stayed at the bus stop for the time in between. In the graph, which also shows speed (km/hr) in orange with the secondary scale on the right, we can observe that the student was in a vehicle coming to a stop from about 75 km/hr and got off the bus at about 5:02:40 pm and started walking for a few minutes. The student seemed to have stayed stationary from 5:05 pm, probably waiting to take the B39 bus back across Williamsburg Bridge for the remainder of the experiment.

Figure 42 Plot on Google Earth 5:03 pm - 5:09 pm (waiting at the bus stop)

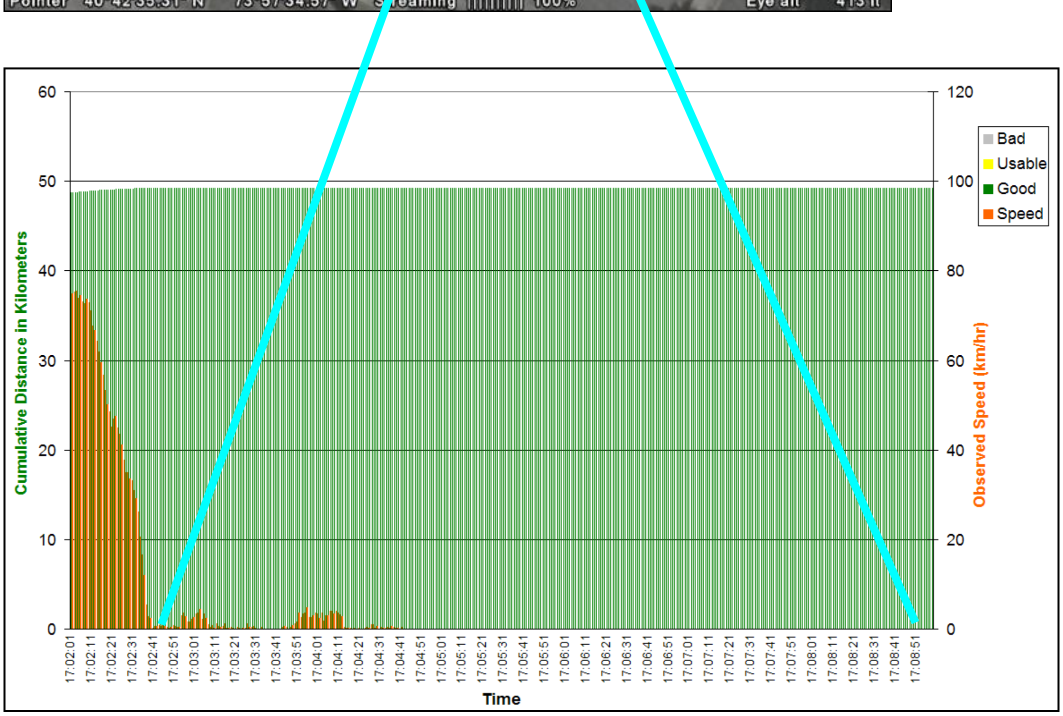
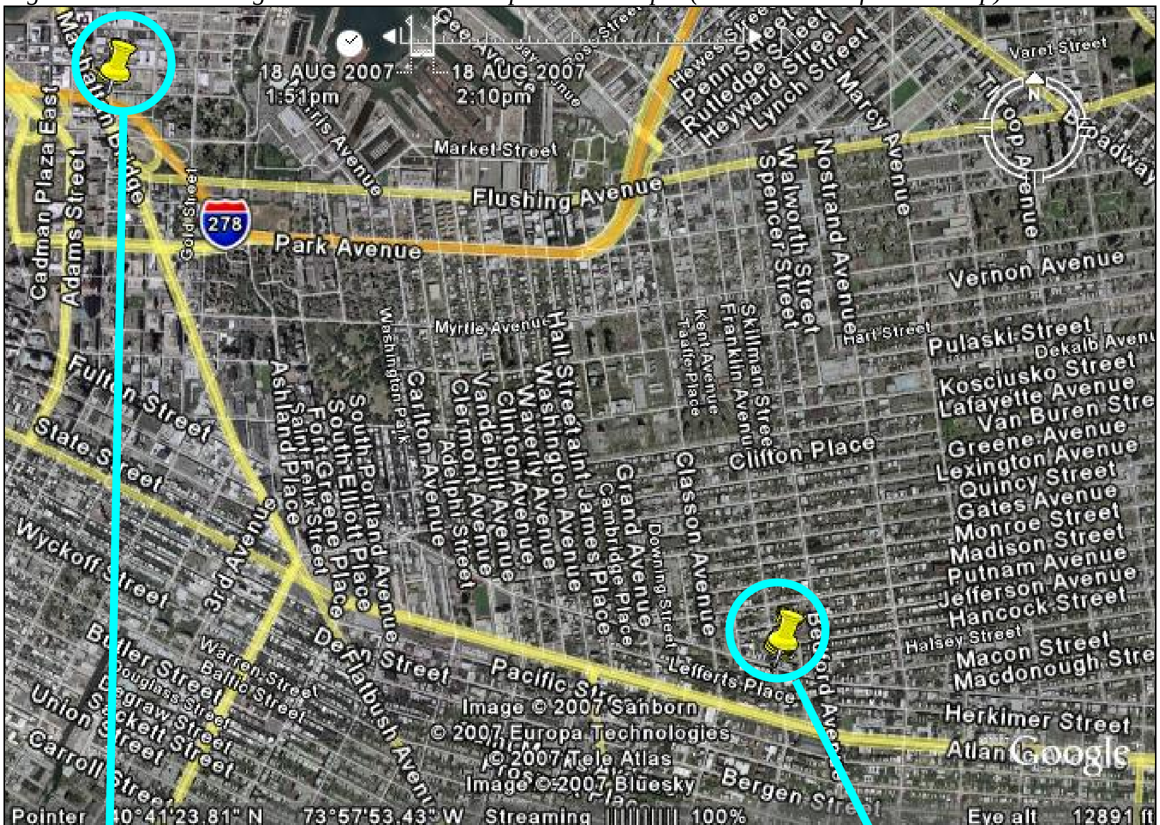


Figure 4.3 Graph of cumulative distance and observed speed over time (5:02 pm – 5:09 pm

Underground Subway Example: Student 2 took F train from Franklin Ave at 1:52 pm and reached York Street at 2:09 pm.

Figure 4.4 Plot on Google Earth between 1:51 pm and 2:10 pm (start and end of F train trip)



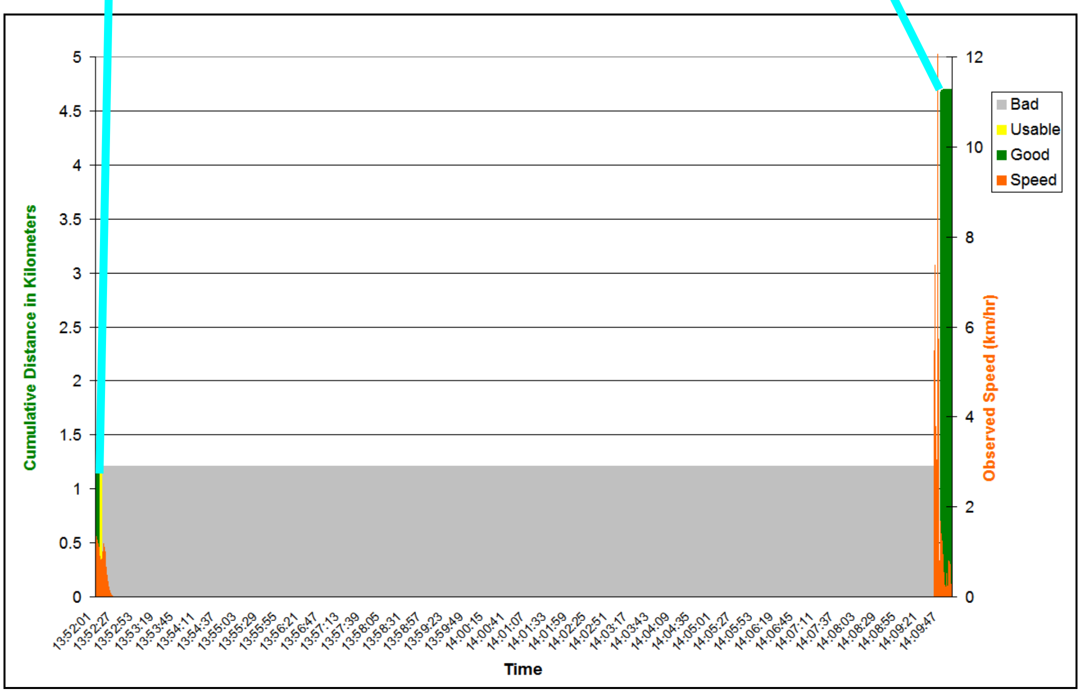
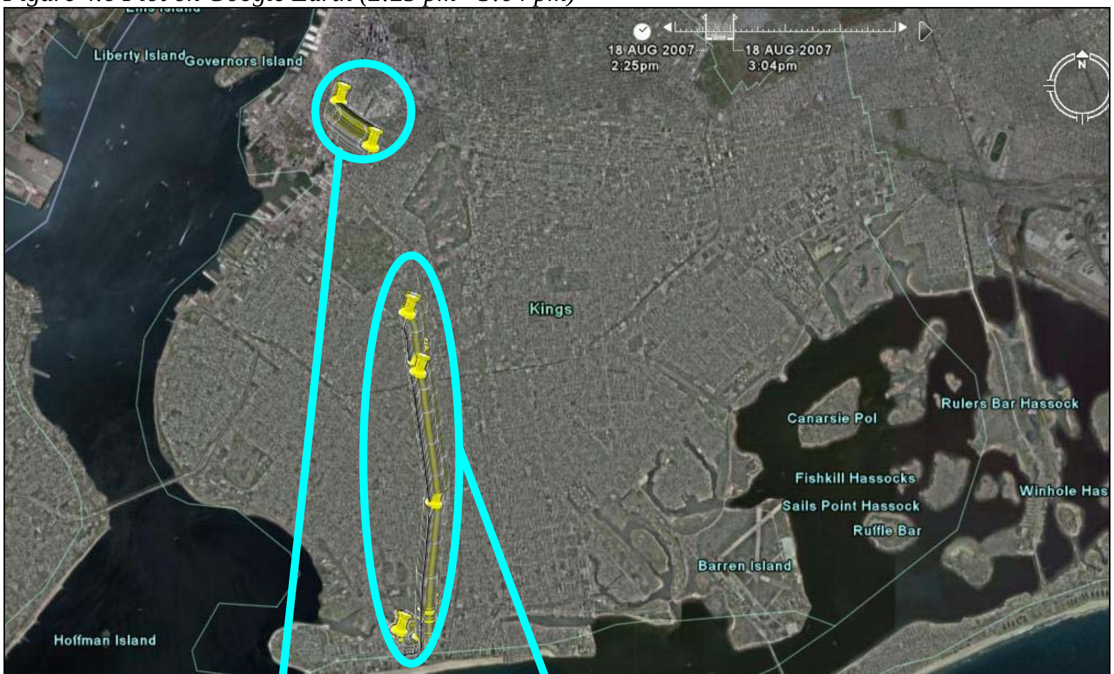


Figure 4.5 Graph of cumulative distance and observed speed over time (1:52 pm – 2:10 pm)

Elevated Train Example: Student 2 was on the G train that started underground at Hoyt/ Schemerhorn 2:25 pm, emerged above ground at 2:30 pm, went back underground at 2:34 pm and again emerged above ground for the second time at 2:43 pm. The student stayed on the train till 3:03 pm to exit at Coney Island.

Figure 4.6 Plot on Google Earth (2:25 pm - 3:04 pm)



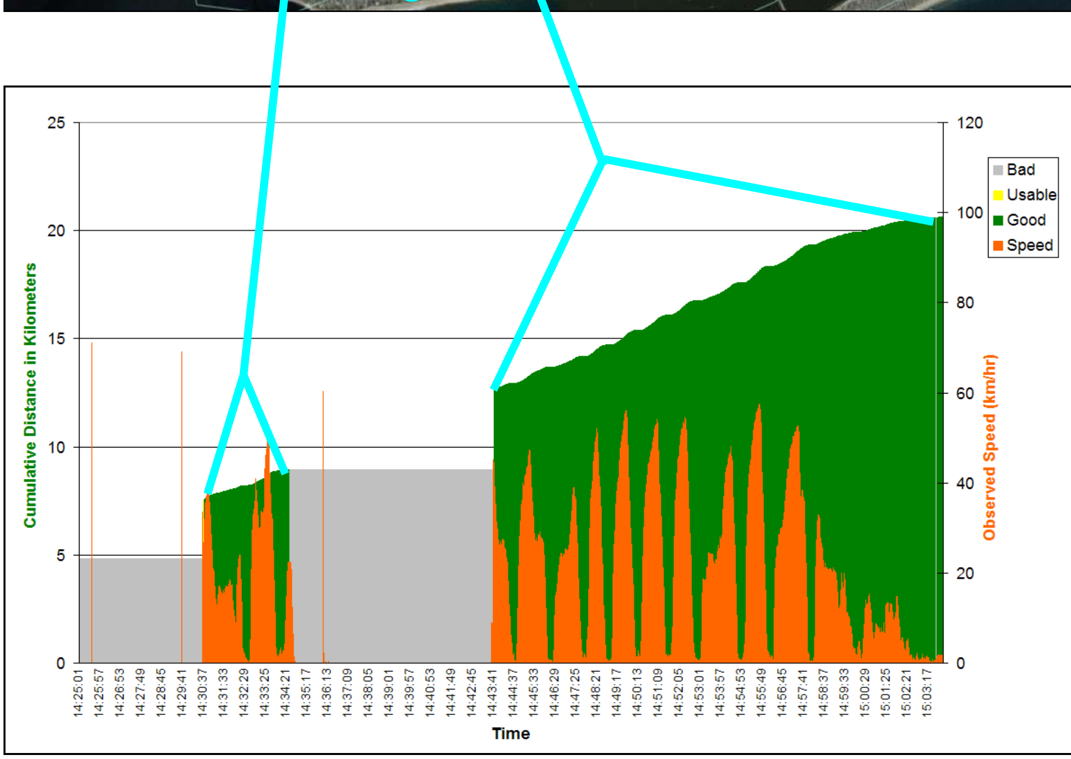
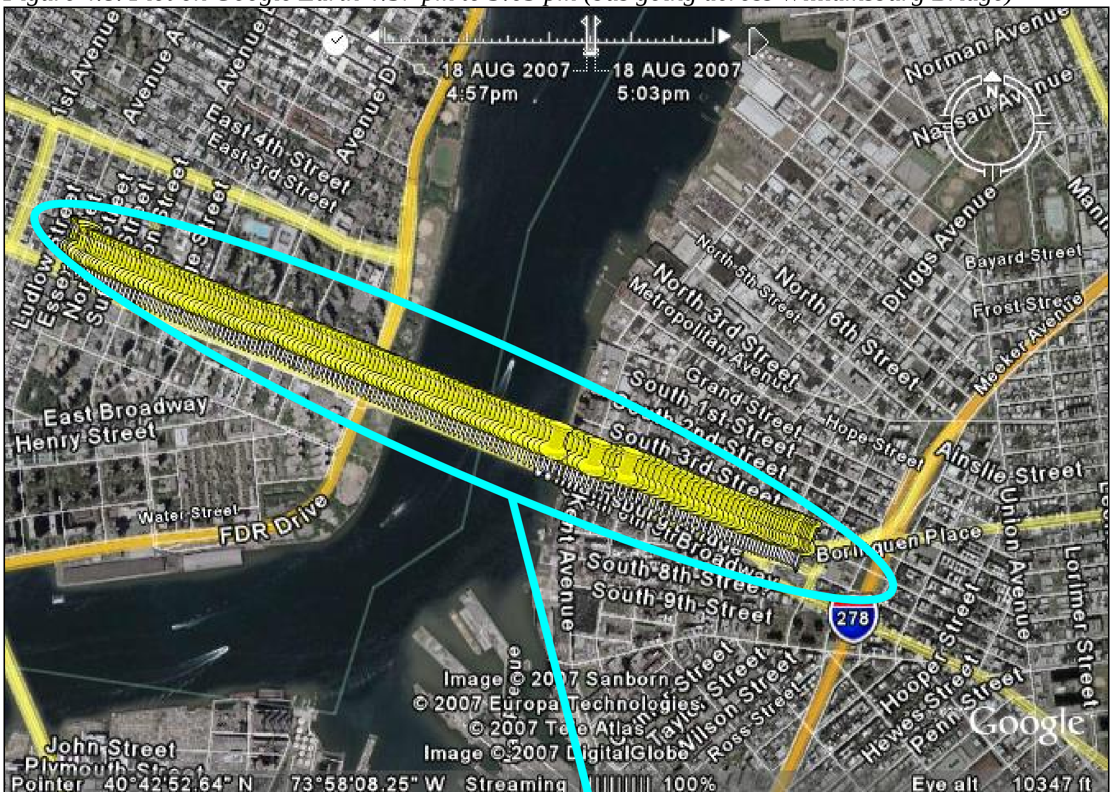


Figure 4. 7. Graph of cumulative distance and observed speed over time (2:25 pm – 3:04 pm)

Bus Example: Student 2 took the B39 bus across the Williamsburg Bridge. The student got on the bus at 4:59 pm, and exited at 5:03 pm after crossing the bridge.

Figure 4.8. Plot on Google Earth 4:57 pm to 5:03 pm (bus going across Williamsburg Bridge)



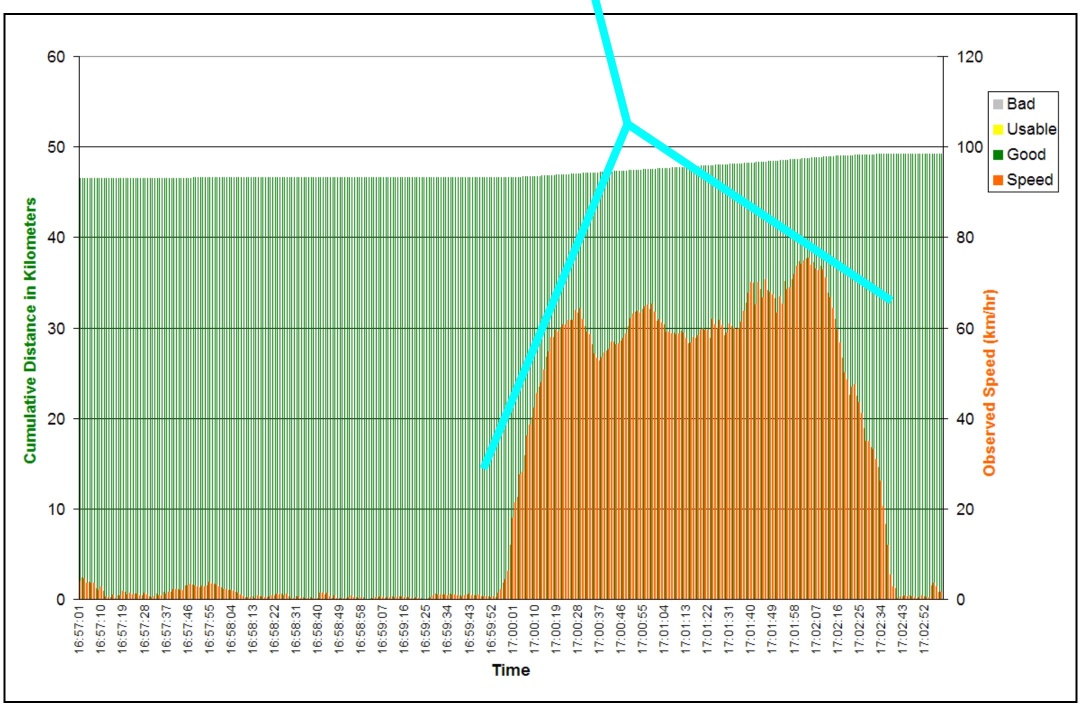


Figure 4.9 Graph of cumulative distance and observed speed over time (4:57 pm – 5:03 pm)

1.4 **Mode Identification Algorithm**

A simple algorithm was developed to determine the travel modes (bus, train, walking) of a trip in an automated manner. Visual Basic version of the algorithm was applied to the GPS data from student experiments that were imported to MS Excel.

The main steps of the algorithm are listed below:

1. Loop through the data points sequentially. Start at the first good data point. Note time of start point. Set default mode "walking" for start of trip.
    a. Look for the segment end, ignoring breaks in signals of 20 seconds or less:
        i. If current mode is "walking": Find the data point where slope (computed as avg. speed over 2 minutes from a given data point) changes to above 4 km/hr, or when the signal disappears.
        ii. If current mode is "train" and current data point is bad, find the next good data point.
        iii. If current mode is "train" and current data point is not bad (not underground), or if the current mode is "vehicle" (bus or train not yet determined), find the data point where slope reduces to 2 km/hr, and then further determine the exact data point where speed is zero for at least 20 seconds, or when the signal disappears.
    b. Compute time difference, cumulative distance traveled and average speed for segment.
    c. If current mode is "train", and speed at segment end is not zero, then go to step a. to look for a new segment, end. The train has gone from underground to above ground, but the mode is still "train".
    d. If current mode is "vehicle", then track change in heading over 5 second intervals for the entire segment. If the heading changes sharply (less than 50 degrees (or greater than 310 degrees) over 5 seconds, then the vehicle is not a train. Set mode to "bus".
2. Record current segment mode, start and end times. Set current segment end as next segment start and proceed with the loop.
3. Once the entire data set has been looped through, check the output trip summary for completeness and accuracy.
4. If multiple sequential segments are of the same mode and have less than 1 minute intervals between them, then combine them into one mode.
5. Add additional data elements to trip summary (such as actual time of day (instead of time in seconds), start and end locations for each segment. The start and end locations that are in Latitude-Longitude in the GPS data can be converted to more understandable format (such as a subway stop or street address) using a GIS module.

For future analysis, this algorithm could be modified to include "car or taxi" and "ferry" modes. Also, it can be fine tuned and calibrated to improve accuracy of mode identification. Pattern recognition modules may be incorporated into the algorithm to better recognize changes in speed and slopes that are characteristic to bus or train stops, instead of relying on average speed calculations.

This algorithm could also be modified to incorporate GIS elements to further enhance accuracy of mode identification. By creating an intermodal transportation network model for the region that includes separate sub networks for individual bus and train networks, it is possible to use the GPS data to provide additional trip details such as bus routes. Also, by comparing the sequential set of links where the GPS points fall, it may be possible to distinguish the mode between a bus from "car/taxi" or even an elevated "train", which may otherwise be difficult or impossible to determine with data analysis without the GIS component.

1.5 **Results of mode identification algorithm**

The algorithm was applied to the data from student experiments. The results from the program were compared with the trip description from the student reports. Following are the comparisons:

Student 2:

From Student Report					From Algorithm		
Mode	**Start Location**	**End Location**	**Start Time**	**End Time**	**Mode**	**Start Time**	**End Time**
Shuttle	Park Place	Franklin Ave	1:45	-	Vehicle	1:47	1:49
Train	Franklin Ave	York	1:52	2:09	Train	1:52	2:09
Train	York	Coney Island	2:12	3:03	Train	2:12	3:03
Train	Coney Island	York	3:04	3:56	Train	3:04	3:55
Train	York	Delancey/Essex	3:56	4:07	Train	3:56	4:07
Train	Essex	Marcy	4:18	4:33	Vehicle	4:28	4:32
Train	Marcy	Essex	4:49	4:56	Vehicle	4:49	4:54
Bus	Across Williamsburg Bridge		4:59	5:02	Vehicle	4:59	5:02
Bus	Back across Williamsburg Bridge		5:09	5:23	Vehicle	5:16	5:23

The algorithm is able to decipher the mode transfer times within 1-2 minutes of accuracy. It is able to identify train modes pretty consistently, especially if the entire or a part of the train trip was underground. However, the algorithm has not been able to clearly distinguish between above ground modes from only the speed and heading patterns. However, this can be easily overcome by including a GIS module to the algorithm.

Following are the results from implementing the algorithm on other GPS data collected from student experiments. It may be noted that the algorithm was developed from just one student's (Student 2) experiment. This algorithm was then applied to the remaining GPS data streams. For future analysis, it may be better to calibrate and fine-tune the algorithm using a broader and larger set of experimental data before implementation.

Student 3:

From Student Report					From Algorithm		
Mode	**Start Location**	**End Location**	**Start Time**	**End Time**	**Mode**	**Start Time**	**End Time**
Bus	83rd St. 34th Ave	Astoria Blvd & 83rd St	2:42	2:49	Vehicle	2:42	2:53
Bus	Ditmars Blvd. & 81stSt.	2nd Ave. &125th St.	2:58	3:16	Vehicle	2:58	3:17
Train	Lexington & 125th	Bowling Green	3:36	4:08	Train	3:36	4:08
Ferry	Battery Ferry Terminal	Staten Island Ferry Terminal	4:30	4:57	Vehicle	4:34	4:53
Train	Ferry terminal	Tottenville	5:02	5:40	Vehicle	5:03	5:40
Bus	Craig & Main	St. George Ferry Terminal	6:55	8:12	Vehicle	6:57	8:11
Ferry	Ferry Terminal	Ferry Terminal	9:08	9:36	Vehicle	9:02	9:33
Train	Whitehall	Roosevelt Ave	9:40	10:31	Train	9:41	10:34
Bus	Roosevelt Ave.	83rd St. & 35th Ave	10:45	10:53	Vehicle	10:41	10:49
Walk	83rd & 35th	Home (3 blocks away)	10:53	10:58	Walk	10:49	10:53

Student 3 (Day 2):

From Student Report					From Algorithm		
Mode	**Start Location**	**End Location**	**Start Time**	**End Time**	**Mode**	**Start Time**	**End Time**
Bus	82nd	Roosevelt Ave	2:31	2:37	Vehicle	2:33	2:35
Bus	Roosevelt Ave.	2nd & 3rd Ave 60th St	2:42	3:15	Vehicle	2:43	3:16
Bus	61th and 3rd	68th and 3rd	3:19	3:22	Vehicle	3:20	3:25
Walk	68th and 3rd	69th and Lexington Ave	3:22	3:27	Walk	3:25	3:30

Student 4:

		From Student Report			From Algorithm		
Mode	**Start Location**	**End Location**	**Start Time**	**End Time**	**Mode**	**Start Time**	**End Time**
Walk	Wyckoff St	Pacific St.	10:15	-	Vehicle	10:15	10:21
Train	Pacific St	23rd st	-	-	Train	10:22	10:54
Walk	-	-	-	-	Vehicle	12:26	12:32
Bus	Allen & Delancy	Across Williamsburg B.	1:15	1:18	Vehicle	1:16	-
Bus	Back across Williamsburg Bridge		1:18	1:22	-	-	1:24
Train	Essex St.	Marcey Ave	1:36	1:43	Vehicle	1:38	1:43
Train	Marcey Ave.	Essex St.	1:46	1:55	Vehicle	1:47	1:52
Walk	Essex St.	York St.	1:55	2:48	Vehicle	2:01	2:40
Train	York St.	Coney Island	2:48	3:34	-	3:05	3:33
-	-	-	-	-	Vehicle	3:42	4:00
Train	Coney Island	Hoyt/ Schermerhorn	-	4:17	Vehicle	4:09	4:12
Walk	Hoyt/Sch.	Wycckoff St	4:17	4:30	Vehicle	4:24	4:27

Student 5:

		From Student Report			From Algorithm		
Mode	**Start Location**	**End Location**	**Start Time**	**End Time**	**Mode**	**Start Time**	**End Time**
Walk	187th & Ft. Washington	Broadway & 168th	10:53	-	Walk	10:53	11:14
Train	-	42nd	-	-	Train	11:14	12:11
Bus	42nd St. & 12th Ave	50th St. and 2nd Ave.	12:28	12:46	Bus	12:28	12:45
Walk	50th St	58th St	12:46	1:00	Walk	12:45	1:09
Bus	58th, across Qnsboro Br.	25th & Queens Plaza	1:00	1:15	Bus	1:09	1:15
Walk	-	-	-	-	Walk	1:15	1:22
Train	-	Astoria	-	-	Bus	1:22	1:28
Bus	31st & Hoyt	125th & 2nd	1:48	1:55	Bus	1:48	1:55
Bus	125th & 2nd	32nd & Astoria	2:04	2:12	Vehicle	2:06	2:11
Train	Astoria	Queens Plaza		-	Bus	2:25	2:32
Bus	Queens Plaza	2nd & 60th	2:35	2:39	Bus	2:35	2:39
Walk	2nd & 60th	49th & 1st	2:39	3:12	Walk	2:39	3:12
Bus	49th & 1st	49th & 12th	3:12	3:35	Bus	3:12	3:36
Walk	49th & 12th	50th & 11th		-	Bus	3:40	3:46
Bus	50th & 11th	8th Ave	4:24	4:30	Bus	4:25	4:28
Train	8th Ave	168th	-	-	Bus	5:06	5:10
Walk	168th	Home	-	5:25	Walk	-	5:25

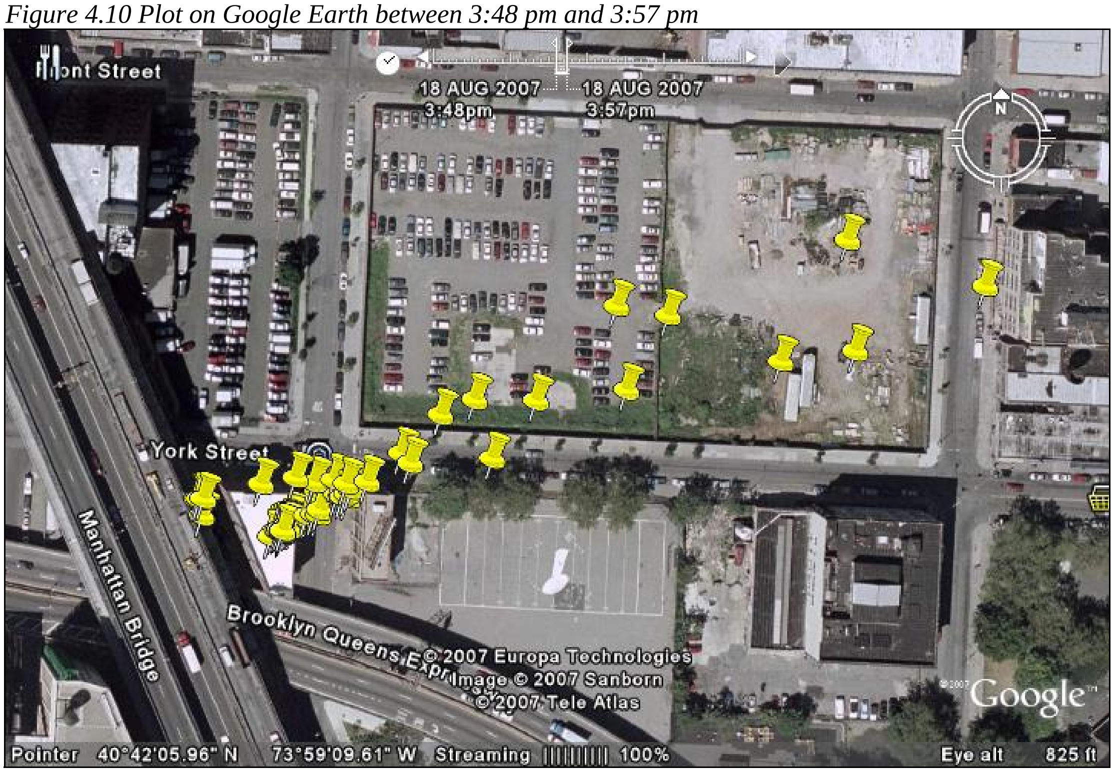
Student 6:

		From Student Report			From Algorithm		
Mode	**Start Location**	**End Location**	**Start Time**	**End Time**	**Mode**	**Start Time**	**End Time**
Walk	-	116th St	7:45	7:50	Walk	7:45	7:57
-	-	-	-	-	Bus	7:57	7:59
Train	116th St	242nd St	7:56	8:20	Train	7:59	8:20
-	-	-	-	-	Bus	8:21	8:30
Train	242nd St	125th	8:21	8:45	Train	8:30	8:42

Issues with GPS data

There were a few problems with using the GPS data for mode identification purposes. For example, in several cases there was incorrect location information: Several GPS data points that were thought to be good (3 or more satellites in view, HDOP less than 5) actually turned out to have improbable location information. For example, the following is a Google Earth plot of some good GPS points during the student's walk around the block near York Street station. Figure 4.10 illustrates this problem. Observe that some of these points are inside a parking lot and blocked off areas. While this example does demonstrate the GPS position measurements do drift from their actual location, the magnitude of the drift may be acceptable for many purposes.

Figure 4.10 Plot on Google Earth between 3:48 pm and 3:57 pm



Another issue arose due to inaccurate heading information. Heading and change-in-heading were a parameter used to help in differentiating between bus and train modes. Since trains cannot make sharp changes in direction over short time intervals, any above ground trip segment that had a significant quick change in direction is deemed to be a bus or an automobile, unless it is a walk segment (very low speeds). However, there were some segments that were actually train segments (as noted in student reports) that the program categorized as bus segments because they had one or more such quick changes in direction. Thus, there needs to be a refinement in the change-in-heading parameter. Given that much of the New York street system is a grid, one needs to look for changes in heading that correspond to turns to perpendicular streets or lines. Train tracks do not permit such turns to be made as sharply as they can be made on the street system. Thus an intervening sharp turn preceded and followed by a reasonably long straight segment is an essentially certain eliminator of elevated trains as a possible mode. Both small and large errors in heading tend to be characterized by random oscillating fluctuations.

The research technique currently lacks the ability to distinguish between a slow moving vehicle in traffic congestion and walking. This leads to inaccurate results for some bus-to-walk or walk-to-bus mode transfers. The main ambiguity ends up being the determination of the actual start or stop location of a particular intermodal walk-vehicle sequence. This is particularly worrisome for short trips that could be taken as walk trips. For longer trips, it seems only to affect the precise start and/or end of the vehicle segment.

In cases where a taxi follows a bus route, the research team may not be able to distinguish between the two modes. However, these could be overcome by incorporating a GIS component to the model, and analyzing the mode options at exact segment start and end locations, and along the routes. Taxi start and stop points tend to be at locations other than bus stops and taxis tend to not follow bus routes. No attempt was made to distinguish between car and taxi modes; however, since most trips are made in personal cars (average car occupancy is very close to one, few trips are "chauffeured"), a subsequent person trip must originate at the destination location of the current car trip. It also may be assumed that a car trip incurs a longer walk segment to the ultimate destination, than does a taxi trip.

To date, the research team has not assessed the capability to infer trip purpose information from the GPS data points. It is believed that the arrival and departure time from a particular location tend to correlate well with the activity taking place at that location. Work, shopping and recreation tend to take place at unique locations. Length of stay tends to differentiate the workers from the patrons at a restaurant. Sequences of stops tend to differentiate delivery personnel from linked personal trips.

In cases where individuals are carpooling, if each person carries a GPS unit, then it should be possible to determine if they are traveling together as a carpool for a portion of their trip. In such situations, as observed when a student carried two of the same i-Blue units during the entire trip, all of the data will not be identical. In such instances, the heading information tends to have large variations; however, the position information (distance between each receiver) has a very small variance most of the time. It is very unlikely that they could be so close together so often without traveling together.

Summary

With the development of post-processing techniques, person-based GPS units can provide a very reasonable and effective way of capturing information that precisely defines multi-modal and linked trips even in the dense environment of the New York Metropolitan Region. Wearing a GPS unit while riding in elevated trains, buses, cars and taxis and walking and waiting in the Manhattan street environment, does not substantially deteriorate the quality of GPS data. With some rather simple data analysis, the GPS information provides the start time and start location, end time and end location of individual modal segments. Given the complete loss of data during subway trip segments and the unique locations of subway entrances, the start and stop of individual subway trips are also identifiable. When the GPS data is processed and superimposed on high resolution imagery available in media such as Google Earth, one can readily identify trip origins, destinations, modes and one can readily infer many trip purposes.

Given the objective nature of the GPS data and the ability to readily obtain precise trip measures such as travel time, distance and actual space-time path traversed GPS provides a uniquely effective way of capturing such data. It has been demonstrated that a rather simple analysis focused on the computed cumulative distance traveled as a function of time, modified by intermittent information provided by speed and heading yields sufficient information to identify the start and end time and location of modal segments even in the presences of the errors caused by the urban canyon environment. The simple algorithm described in Section 3 provided an automated means of reducing the voluminous GPS data to fundamental modal segments: 'mode', 'start time', 'start location', 'end time' and 'end location'.

Based on visual observations of the GPS data using Google Earth, it is believed that given a detailed digital map database of the New York City subway system, bus routes, streets and highways, a map-matching algorithm could readily assign the captured GPS data to specific bus routes, bus and transit stops and stations and subway entrances.

Correlation of trip end with available point-of-interest information, time-of-day information and trip-end duration can be expected to readily identify trip purpose. The success of this procedure suggests a strong potential for a broader, more encompassing study of GPS as a source of travel survey data be undertaken.

To advance the mode signature technologies, it will be necessary to involve a large number of individuals involved in normal daily activities from all groups and included a particular focus on normally unrepresented groups. The purposes of the current study focused on a targeted sequence of trips so as to obtain as much information as possible from each mode, in particularly difficult "urban canyon" situations. Future studies should involve many more individuals involved in normal daily behavior rather that scripted behavior.

Data streams from a comprehensive GPS deployment would make it possible to assess the correlations of location, time, duration and purpose. The cumulative distance analysis should be refined using the broader data analysis. The analysis should incorporate a detailed digital map database of both the transit system and the street and highway system. It should document the quality of map-matching techniques to properly identify transit on-off locations and specific bus and transit routes. It is expected that advanced uses of GPS-based travel survey data would improve metropolitan travel models, including issues of specific path choice and detailed time-of-day issues.

Section Five: Discussion

From the findings in this third report on collecting data for “mixed mode” travel and identifying special population segments for future GPS survey deployment, the research team is able to build on the responses to questions presented by NYMTC. The research team also introduces additional questions addressing concerns on GPS surveys. These questions are indicated in this section by an italicized “*R:*” to differentiate between question that NYMTC poses and those raised by the research team.

Issues Regarding the Use of GPS

Should a person-based-GPS survey be a part of the upcoming regional household travel survey?

Technically **–** the use of the existing GPS products available on the market is feasible in Manhattan and elsewhere in the region. In terms of time and location information, GPS generated data is superior to any self-reported data gathered using a traditional survey methodology because it has no rounding errors, no location confusion, and reduces the burden[9] on respondents. The first experiment demonstrated the accuracy and reliability of GPS generated data in a dense urban area and the second experiment demonstrated additional benefits with respect to distinguishing various modes of travel.

Practically **–** the GPS survey should be a part of the upcoming regional household travel survey only if it improves the response rate for certain population segments, the data quality, or both. The improvement in response rate can only be achieved if the GPS unit or survey entices those non-responding to participate. As for the question on whether the GPS data will improve the data quality, the answer is positive. As mentioned before, information on time and location will be exact. The trip purpose information will be obtained from a modified survey that goes with the GPS survey and its quality is likely no worse than that of traditional surveys. Mode information can either be obtained from the modified survey or from the GPS data directly. The results of our second experiment indicated that in many cases, mode information can be correctly generated. If given sufficient amount of time, the research team is confident that it is possible to develop an algorithm within the GIS environment to deduce the mode interchange information. Chung and Shalaby (2005) reported that their GIS analysis was able to correctly detect 91.7% of all trip modes in downtown Toronto.

R: If the GPS component is included in the travel survey, who should be give GPS units?

From the binary logit model analysis on the transit-flagged records in ‘Section Two’ and given that the use of GPS will improve data quality, it makes sense that GPS units should be given to people who tend to have problematic transit records. Over sampling the transit population should also be considered in addressing these transit issues.

[9] Even when the respondent must fill out a diary accompanying the GPS data collection, it is expected that the respondent burden is less than that of a traditional survey.

What would be the minimum recommended sample size of the person-based GPS component of the regional household travel survey?

The appropriate sample size is dependent on the future use of the data. Stopher et al. (2008) provide some guidance on sample size for GPS surveys – since the sample size depends on the intended use of the data. They strongly recommend the collection of multi-day data as it allows for the analysis of day-to-day travel behaviors, while reducing sample sizes and survey costs. In addition, the lack of trip purpose in the current version of the GPS technology limits the possibility of selecting a sample of GPS-users only. The rest of this section will offer guidance on how to determine minimum sample size for these population segments considering that GPS units are only provided to certain population segments.

> Given the variable of interest, sample size calculation depends on two important elements: the variability, over the population, in the parameter to be measured and the degree of precision required for the parameter estimation. The size of population does not play an important role in general, unless the population size is very small[10].
>
> Suppose we are interested in estimating two key variables for young males: the daily trip rate and percentage of transit user[11]. From the 1997/98 household travel survey data, the mean daily trip rate for young males is 4 with a standard deviation of 2.28. The mean proportion of transit users for the young male population is 0.25 with a standard deviation of 0.43. For trip rate, suppose we have 95% confidence that the sampling error is no more than 5% of the sample means. Thus, the margin of error equals 0.2. The required sample size is: $N = (\frac{z_{\alpha/2} \times S_x}{E})^2 = (\frac{1.96 \times 2.28}{0.2})^2 = 100$. In other words, a sample size of 100 is required to allow us have a trip rate estimate such that there is a 95% probability that the sampling error will be no more than 5% of the sample mean. The calculation for the share of transit users in the young male population is similar, except the estimate now is a proportion variable. Assume that we desire to have the sampling error no greater than 5 percentage points. The required sample size is: $N = \hat{p}(1-\hat{p})(\frac{z_{\alpha/2}}{E})^2 = 0.25 \times (1-0.75) \times (\frac{1.96}{0.05})^2 = 288$. Therefore, a sample of 288 young males is required to estimate the proportion of transit users with sampling error less than 5%.
>
> The sample size calculation based on different variables of interest will result in different numbers. An easy solution is of course to select the maximum required sample size. However, in reality, this is often infeasible due to cost considerations. In actual surveys, compromises are often made. One can also conduct multiple stage sample size calculations during the survey process, such that the precision level can be controlled in real time. This strategy can pose some challenges in the real world, as it requires one to collect and analyze data simultaneously.
>
> Another issue to be aware of is that the above calculations are based on the simple random sampling strategy. If stratified sampling is used, as is often the case, the required sample size will be smaller than the one derived from the simple random sampling, unless the variation in the population is the same across strata.

[10] This contradicts to the intuitive feeling that sample size is always expressed in percentage form.

[11] Here, we use the same definition for transit user as before. A transit user is someone who uses any of the available public transit modes on the given survey day.

How could the person-based GPS survey be used to improve and / or enhance current travel survey processes that have been used in the NYMTC Region?

The travel survey process includes a number of tasks, including determining the sampling frame, recruitment, data collection, and post processing of the data after it is collected. Using GPS can be seen as a trade-off between using the traditional surveying processes, which has the risk of lower quality data due to misreporting, and the risk of trying a new technology that promises to provide better data, but only if properly used and accepted for use by potential survey participants. There could be an additional benefit of improved survey participation if some groups currently not participating, would find the use of the new technology intriguing enough to now consider responding to a recruitment offer.

The research to date suggests the benefit of being able to reduce the risk of low quality data from transit riders (as occurred in the 1997 effort), thus it is very important that GPS be considered as part of the data collection effort. Since the evidence to date indicates GPS can be used as a complement, but not a substitute, for a traditional surveying process, there would not be any gains in the initial steps of conducting a travel survey. The random sampling frame would have to be assembled and initial contacts made. At this point, if people are offered an opportunity to use GPS to reduce the burden of filling out the complete diary and having to be involved in a tedious telephone experience to retell all their activities in great detail, there might be substantial improvement in participation rates. GPS has become a very popular item and people could be intrigued (and perhaps honored) to have been chosen to participate. This might apply to very busy people who can't be bothered with participating in a traditional travel surveying effort.

An additional cost saving associated with accurate data is the reduction in staff time to try and "fix" poorly collected and/or recorded data. Refinements in the data processing stage may be needed to confirm stopping behaviors. For example, there could be issues associated with the distinction between a purposeful stop and the appearance of a stop.

Using GPS for the detailed data elements will require high quality data handling techniques. The units will need to be tagged with unique identifiers, the data will need to be downloaded and properly named and stored for post-processing techniques. These skills are necessary for traditional survey work, but will need to be electronically enhanced if GPS data is part of the data system used in the survey.

How could the person-based GPS survey help in addressing non-reported trips in the travel diaries?

The first and second experiments clearly proved the value of GPS to capture all trips better than or equal to traditional surveying methods. What is not known yet is whether the operational requirements of using GPS (i.e., plugging in the unit overnight, turning it on properly, etc.) can be successfully accomplished. The first and second experiment reported issues that could result in missed trips if the equipment is not functioning correctly.

How could the person-based GPS survey help in addressing the rounding of travel times, imprecise departure and arrival times reported in the travel diaries, and bad recollection of O-D locations by respondents for geocoding?

Properly functioning GPS will ALWAYS provide better data than traditional surveys with respect to exact travel times, exact departure and arrival times, and exact locations of activities. Thus, every additional GPS unit will improve the quality of the data and avoid the loss of samples due to bad data reporting. The key will be in the training element and the ability of people to self-train.

Would the person-based GPS survey help to improve the response rate from low response groups, such as young males?

The third experiment will provide an opportunity to determine whether low response groups, such as young males, might be good candidates for using GPS as there is evidence in the market place that GPS attracts attention among young males. The issue of their willingness and their ability to use the equipment would be part of the third experiment that would provide guidance on improving response rates for targeted groups.

How could the person-based GPS help in improving the BPM modeling process?

Our discussions with the BPM modeling staff indicated that they would be able to use the higher quality origin and destination data at this time. They would also be able to use the better quality time data for estimation purposes. The actual routing data may be useful for auditing and validating model assignment in the future.

How would it be possible to track complex bus/subway transit rider paths in Manhattan?

The research completed by ALK indicates the ability to "see" individuals moving from walking to bus, to subway, etc. Their work does not incorporate GIS, however. The spatial data can be used in GIS with high accuracy of mode paths.

How to address "acquisition time" for GPS devices to start registering lat/long coordinates upon being turned on, as well as related "no signal" and "inconsistent signal" issues related to being near tall buildings, being within structures, and cloudy days?

With regards to acquisition time, it takes a few seconds to a few minutes for a GPS logger to start registering lat/long coordinates upon being turned on. It is very important to instruct the survey participants to stand still during the acquisition time and wait for signals from the GPS logger (green light blinking, for example) that it has started registering lat/long. For reacquisition time, our second test shows that it takes about 39 seconds on average for a GPS logger to register lat/long after emerging from underground subways.

GPS loggers receive signals, sometimes inconsistent signals, near tall buildings. Inconsistent signals could be removed during data processing by using parameters such as HDOP and number of satellites used. Within structures such as buildings and bridges, GPS loggers receive signals on the street level or above where the sky can be seen (with windows, for example). GPS loggers cannot receive signals in underground subways, tunnels, or on the underground levels of structures. These "no signal" issues could be addressed with the help of GIS information (such as subway lines and location of tunnels). GPS loggers receive signals even during cloudy days. The accuracy of the location points, however, decreases substantially in cloudy days. It is strongly recommended that the travel survey with a GPS component be deployed in days with good weather.

References

Asakura, Y. Hato, E. (2004). Tracking Survey for Individual Travel Behaviour using Mobile Communication Instruments. *Transportation Research C* 12 (3/4), 273–291.

Asakura, Y., Hato, E. and Sugino, K. (2005). Simulating Travel Behaviour using Location Positioning Data Collected with Mobile Phone System. In: *Simulation Approach in Transportation Analysis* (Edited by Kitamura and Kuwahara), 183–204, Springer, NY.

Asakura, Y., Hato, E., (2006). Tracking Individual Travel Behaviour using Mobile Phones: Recent Technological Development. Conference paper in11th International Conference on Travel Behaviour Research, Kyoto, Japan.

Hato, E. (2006). Evaluation of trip-activity pattern variability using probe person data. Transportation Research Board 85th Annual Meeting.

Hato, E. (2006). Development of BCALS (Behavioural Context Addressable Loggers in the Shell) for Activity Data. Transportation Research Board 85th Annual Meeting.

Hato, E. and Asakura, Y. (2001). New Approaches for Collecting Time-Space Activity Data. Transportation Research Board 80th Annual Meeting.

Itsobo, S., Hato, E. (2006). A study of the effectiveness of a household travel survey using GPS-equipped cell phones and a WEB diary through a comparative study with a paper based travel survey. Transportation Research Board 85th Annual Meeting.

Chung, E., A. Shalaby. (2005). A trip reconstruction tool for GPS-based personal travel surveys. *Transportation Planning and Technology*. Vol. 28 (5), pp. 381-401.

Stopher, P., Kockelman, K., Greaves S., and E. Clifford. (2008). Reducing Burden and Sample Sizes in Multi-day Household Travel Surveys. Presented at the 87th Annual Meeting of the Transportation Research Board in Washington, D. C., on January 13 – 17, 2008.

Swann, N. and P. Stopher. (2008). Evaluation of a GPS Survey by Means of Focus Groups. Presented at the 87th Annual Meeting of the Transportation Research Board in Washington, D. C., on January 13 – 17, 2008.

Stopher, P., Clifford, E., and M. Montes. Variability of Travel over Multiple Days: Analysis of Three Panel Waves. A paper presented at the Annual Transportation Research Board Meetings on January 13 – 17, 2008, in Washington, D. C.

Appendices: Phase Three

Appendix I: Instructions to Students for Field Test Two
Appendix II: Data Sheet and Report from Student 1 for Field Test Two

Appendix I: Instructions to Students for Field Test Two

Instructions

1. The night before the field test, make sure both the GPSs are off and recharge them overnight.

2. On the day of the field test, after you walk outside of your residence and can see the sky, follow the instructions below to turn on the two GPSs:

For the smaller GPS logger, switch the Mode Switch on the side from "OFF" to "LOG".

For the bigger GlobalSat Data Logger, press and hold the silver button on the front until a red power light and a green GPS status light are steady.

Stand still for a few minutes until the orange light on the GPS logger and the green light on the GlobalSat logger become blinking, indicating that they are receiving signals from satellites.

RECORD THE STARTING TIME OF YOUR TRIP ON THE DATA SHEET

3. During your trip, RECORD THE FOLLOWING ACTIVITIES:

Train/bus/ferry starting time and ending time
Train emerges above ground or goes underground
Train/Bus crossing a bridge -- beginning time and ending time
A 4-digit number outside of a subway train

4. When you are done with all the walking and arrive home, turn off the GPS logger by switching from "LOG" to "OFF" on the side and the GlobalSat logger by pressing the silver button until all three lights are off.

RECORD THE ENDING TIME OF YOUR TRIP

5. We prefer that you finish the field test in one day. If you cannot, make sure that you turn off both GPSs and recharge them when you get home.

6. If you have any questions during the test, you could call me at my cell phone at NNN-NNN-NNNN.

7. Write a brief report about the field test, comparing the two GPSs and detailing the time and length of any activities outside the testing routes.

DATA SHEET

1. Starting time of your trip:

Date (eg. 08/11/07):
Time (eg. 9:30 AM):

2. Mode from your house to the beginning of the testing route (eg. Walking, or taking Subway #7 and connect at Time Square to Bus # 42)

3. On the test routes, record the time for starting, emerging above ground, going underground, starting to across a bridge and ending the crossing, etc.

4. Mode from the end of the testing route to your house (eg. Walking, or taking Subway #6)

5. Ending time of your trip:

Date (eg. 08/11/07):
Time (eg. 9:30 AM):

Appendix II: Data Sheet and Report from Students for Field Test Two

DATA SHEET - STUDENT #1

DAY ONE:

1. Starting time of your trip:
Date: 8.18.07
Time: 1:39 pm

2. Mode from your house to the beginning of the testing route:
Shuttle from Park Place to Franklin Ave- 1:45 pm
Transfer to Manhattan Bound F train running on the C line at Franklin Ave - 1:52 pm
Exit F at York - 2:09 pm

3. On the test routes, record the time for starting, emerging above ground, going underground, starting to cross a bridge and ending the crossing, etc.
Testing the F line in Brooklyn
Sat by window:
Entered the F at York - 2:12 pm — Car # 6250
Switch to G running on the F line at Hoyt/Schemerhorn- 2:25 pm — Car #5689
Emerge aboveground - 2:30 pm
Go belowground - 2:34 pm
Emerge aboveground - 2:43 pm
Exit at Coney Island: 3:03

Sat away from window:
Entered the G train running on the F line at Coney Island-3:04 pm — Car # 6146
Go belowground - 3:24 pm
Emerge aboveground - 3:34 pm
Go below ground - 3:38 pm
Switch to F at Hoyt/Schemerhorn - 3:47 — Car #6024
Exit F at York - 3:53

*Took F to Manhattan for next test
Enter F at York - 3:56 — Car #5949
Exit F at Delancey/Essex - 4:07

*Walked one block to McDonalds then reentered station

Testing the J line across the Williamsburg Bridge
Sat by window:
Entered the J at Essex - 4:18 pm — Car # 4858
Emerge aboveground - 4:26
Cross Bridge - 4:29 pm
Off Bridge - 4:30 pm
Exit at Marcy - 4:33 pm

Sat away from window:
Entered the J at Marcy - 4:49 pm — Car # 4761
Cross Bridge - 4:50 pm
Off Bridge - 4:51 pm
Go belowground - 4:54 pm
Exit at Essex - 4:56 pm

*Walked south across Delancey to bus stop

Testing the B39 across the Williamsburg Bridge
Sat by window:
Entered the B39 - 4:59 pm
Cross Bridge - 5:00 pm
Off Bridge - 5:01 pm
Exit the B39 - 5:02 pm

Sat away from window:
Entered the B39 - 5:09 pm
Cross Bridge- 5:18 pm
Off Bridge - 5:19 pm
Exit the B39 - 5:23 pm

*Exited at Allen and Delancey, Walked south on Allen, East on Canal, Backtracked West on Canal, Walked south on Bowery to B51 stop….bus not running

4. Mode from the end of the testing route to your house:
Walked north on Bowery, West on Canal, Took Brooklyn bound Q at Canal St. to Prospect Park, Transferred to Shuttle to Park Place, Walked one block to apartment

5. Ending time of your trip:
Date: 8.18.07
Time: 6:55 pm

DAY TWO:

1. Starting time of your trip:
Date: 8.20.07
Time: 2:45 pm

2. On the test routes, record the time for starting, starting to cross a bridge and ending the crossing, etc.

Sat by window:
Entered the B51- 2:52 pm
Cross Bridge- 2:54 pm
Off Bridge- 2:54 pm
Exit the B51- 2:55 pm

Sat away from window:
Entered the B51- 3:03 pm
Cross Bridge- 3:03 pm
Off Bridge- 3:04 pm
Exit the B51- 3:09 pm

3. Ending time of your trip:
Date: 8.20.07
Time: 3:09 pm

REPORT- STUDENT #1

DAY ONE:

The G train was running on the F line between Hoyt/Schemerhorn and Coney Island. This did not effect the above ground route but I had to switch underground- this made the trip longer than it would be normally.

Both GPS loggers appeared to have a signal no matter where I sat in a subway or bus. However, about a half of the time when I would emerge aboveground both loggers would take about 2 minutes to relocate a signal. It did not seem that one logger took more or less time to re-signal than the other.

DAY TWO:

Because the B51 is a weekday only bus I had to test this route on Monday. I only turned on the GPS when I got to the bus station…I did not go to the bus station directly from my house. However, on Saturday I walked from the B39 stop to the B51 stop (before realizing the bus was not running) then proceeded from there to the Canal St. Q stop to return home.

Both GPS loggers appeared to receive a signal at all times while inside the bus.

DATA SHEET AND REPORT - STUDENT #2

Second Report: GPS Project (8.18-8.19.07)

.Key points:

1. GeoLogger again proved the better unit in most respects. Reacquisition time in emerging from subways seems to be the only place it lags. In this situation it can take a minute or more to regain signal.

.

2. Both units generally worked well on all forms of transit tested where they had direct exposure through roofs and sides, including different types of buses, the ferry, and the Staten Island RR. Exceptions were in the deep and shielded interior of the ferry, and in some cases where GeoStat had to deal with a second intervening layer, such as an overhead elevated line or corrugated roofing.

I am laying these out as a straight time line, because I think it may be easier to correlate with the details of the GPS readout. All times pm EDT.

Date: 8/18/07

2.28 Left home, walking north on 85th St. to 34th Ave, then west to bus stop at NE corner, 83rd St. and 34th Ave. Both loggers on. GeoLogger acquires signal after one minute at 2.32 at first open intersection. GeoStat does not acquire until I reach bus stop (2.36) and needs one minute plus in a stationary position to do so (2.38).

2.42 Board Q33 bus (Type RTS, #9741) going northbound to Astoria Blvd. Sitting on left side behind driver. Carrying GPS units in the pockets of my light weight jacket. Reception appears good while seated and standing in center aisle.

2.49 Off bus at Astoria Blvd. and 83rd St., walking one block west and two north across the Grand Central Parkway overpass to the M60 bus stop at Ditmars Blvd. and 81St St. Arrive here 2.57.

2.58 Board M60 bus (new hybrid model, #6374) which goes along GCP access road to Triborough Bridge. Initially standing in center aisle behind driver, moving further back to mid-bus as bus fills. At 3.07 take a seat near the rear on the aisle. Also at 3.07, bus moves on to Triborough long span. At 3.09 off long span and on to viaduct over Riker's Island, where bus slowed by traffic congestion. At 3.13 bus moves off viaduct onto approach to bridge over Harlem River. Bridge crossed at 3.14. Off bus at 2nd Ave. and 125th St. at 3.16. Both units appeared to work well at all points on the M60 transit.

3.16 Walking west on 125th St. towards Lexington Ave. Made stop in building en route on south side 125th St. between 2nd and 3rd Aves. from 3.18-3.31. Both units lose signal in building, but reacquire immediately on reemerging. Continue walking to subway station at Lexington and 125th, arriving at 3.36. Signal lost on both units immediately after leaving stairs to first level down in subway station. Take #4 train to Bowling Green station near ferry terminal.

4.08 Exist from kiosk at south end Bowling Green station. GeoStat reacquired signal immediately, GeoLogger required 30 seconds. Walked down center-west path through Battery Park (to immediate east of WW II Merchant Marine monuments) as more direct route blocked by construction. Entered Battery Ferry Terminal at 4.20, with both units losing signal within thirty feet of the door.

4.30 Boarded ferry "Spirit of America." Both units reacquired signal as soon as I was on the gangway leading to the ferry. On this voyage, and on the return trip later in the evening on the same boat, signals were received at all points on the trip with the following exceptions:

1. In the interior men's room on the second deck in the middle of the ship, neither unit.
2. In the passenger space on the main deck, the GeoStat lost signal when I was up against the steel interior wall abutting the automobile chutes (which I could not get access to.) This was about 15 feet from the outside of the ship. The GeoLogger worked here.
3. The GeoLogger momentarily lost signal as I passed along the second deck under the ship's bridge area while debarking at Staten Island.

Otherwise, signal appeared to be fine on all three decks, including the interior snack bar on the second deck.

4.57 Left ferry. Signal maintained on both units in the debarkation hall, but lost on entering the terminal proper and the passages to the busses and the Staten Island RR. Note that I had a different experience on returning through the terminal later that night.

4.59 Boarded SIRR (car 413, window seat on west side.) Train left at 5.02; no signal in either unit until train left station and train shed, then both units reacquired immediately. En route I sat in different locations in the car, including standing in the middle (from 5.22-5.30) and there appeared to be no problem with signal reception in any situation.

5.40 Train arrived at southern terminal, Tottenville. I left the station by the south entrance to Bentley St., then walked east on Bentley one block to Arthur Kill Rd., turned north two blocks to Main St., and then east again, one block to Craig Ave., the southern terminal of the S74 bus. Having just missed a bus, I stopped to eat at a restaurant on Main St. about forty feet west of Craig, intending to catch the 6.30 bus. The units appeared to maintain signal the entire time I was in the restaurant, at distances ranging from 10 to 15 feet from the large plate glass windows. The 6.30 bus was delayed and did not leave until 6.55, so I spent the entire period from 5.50 (arriving at the corner) to 6.55 within 50 feet of the intersection.

6.53 Boarded S74 bus (type uncertain, #6275) at Craig and Main. Departure 6.55, arrival at St. George Ferry Terminal at 8.12. Both units seemed to work well throughout the extended trip.

8.13 Entered Ferry Terminal. Next ferry at 9.00, so I went to the waiting room. Both units had lost signal at the bottom of the ramp leading down from the bus platforms into the building, but sometime between 8.20 and 8.50, while I was stationary, the GeoLogger reacquired signal in the waiting room. I then tested it by walking around and found it could hold the signal back into the main terminal area, near the newsstand. This is within the building structure itself, more than 100 feet from exterior of the building in any direction. GeoStat did not function at all in any of this area. I suspect that the relatively light structure of the waiting room roof played a role here, allowing signal acquisition if exposure is prolonged enough.

9.08 Boarded ferry. Both units again functioning. See notes above for details. 9.36 left ferry, walked one block north to Whitehall St. R/W station. Both units functioned as I walked quickly through the ferry terminal, and lost signal as I entered the subway at 9.40. Took R train to Roosevelt Ave. station in Queens, where I transferred to a Q33 bus. Both units reacquired as I left the subway station at 10.31 for the more open space of the bus shed, but the GeoStat did so immediately while the GeoLogger took nearly a minute.

10.39 Boarded Q33 (hybrid, #3648), which departed at 10.45 turning west on Broadway to Roosevelt Ave, then east on Roosevelt to 83rd St., where it turned north. Left bus at 83rd St. and 35th Ave. at 10.53, then walked two blocks east and on-half block north to home, arriving 10.58. Equipment shut down at 11.02. The units worked on the hybrid design, with one exception. These busses have fiberglass and aluminum lateral dividers between some seats, and when I sat against one of them, with the GeoStat in my left jacket pocket, it lost signal in the confined corner. GeoLogger had no problem here.

Date 8/19/07

This run was specifically aimed at testing the units, especially the GeoStat, in conditions where intervening structures (bridges, elevateds) might affect reception.

Both units turned on at 82nd St. and 37th Ave. in Jackson Heights. Both units acquired signal at 2.29. At 2.31 boarded Q32 bus (type uncertain, but with plush individual seating similar to long distance busses, rather than regular mass transit seats, #163) southbound on 82nd St. I sat behind the driver. One block later, at 82nd and Roosevelt Ave., now having turned west under the elevated at 2.37, the bus broke down and at 2.42 we transferred to another Q32 (RTS type, #5199.) Here I sat in the first seat on the right. We preceded west under the elevated #7 line to 50th St. and Queens Blvd, where we emerged at 2.58 onto a street close by but parallel to the railway. At 3.07, at Van Dam St., the roadway again passed under the elevated, with the bus emerging into the open space of Queens Plaza at 3.10. The GeoLogger seemed to be working well, but the GeoStat unit held signal only intermittently while under the elevated. I tried two configurations- keeping the GeoStat in my bag from 2.49-2.53 and in my hand or jacket pocket the rest of the time. While there was an intervening structure, it seemed to work only when in my hand. At 3.13 the bus entered the lower deck of the Queensborough Bridge. In the first twenty seconds on the bridge we passed through a draped construction zone that might have interfered with reception. After this, GeoLogger worked properly but GeoStat intermittently to the end of the bridge. The bus left the bridge at 3.15.

I left the Q32 between 2nd and 3rd Aves. on 60th St. and walked west to 3rd Ave., then one block north to the M101 stop at 61st and 3rd. At 3.19 I boarded an M101 Limited (articulated, #5540) which I left at 68th and 3rd at 3.22, then walked one block west to 68th and Lexington Ave., turning one block north on Lexington to 69th St., and finally, half a block west on 69th to the entrance to Hunter, where I shut the units down at 3.27. Both units were detecting throughout this segment, including the section on Lexington, which was under a standard construction scaffold. However, the last 150 feet on 69th St. was under a scaffold covered with corrugated steel, and here only the GeoLogger worked.

DATA SHEET - STUDENT #3

1. Starting time of trip:
Date: August 19, 2007
Time: 10:15 AM

2. Mode from house to the beginning of the testing route:

- Walked from 275A Wyckoff Street to the Pacific Street Subway Station. There I took the N subway uptown to 23rd Street. (NOTE: The N traveled over the bridge, went above ground at 10:33 AM and went below ground at 10:37AM)
- From 23rd Street, walked east to 1st Avenue. Had lunch there until 12:30PM. From there, walked downtown on 3rd Avenue to reach B39 on the corner of Allen and Delancy Streets.

3. Hopped on the B39 on the Manhattan side at 1:15PM and the bus entered the Williamsburg Bridge at 1:16PM. The bus got off the bridge at 1:18 PM. I then immediately caught the returning bus at 1:18 PM. It entered the Brooklyn side of the bridge at 1:20PM. The returning bus got off the bridge on the Manhattan side at 1:22 PM.

4. From the B39, I went into Essex Street Subway station (at 1:25PM) to catch the J subway*. I got into the train at 1:36PM, and it was above ground at 1:38PM. I was sitting next to the window on the trip from Essex Street to Marcey Avenue, and arrived at 1:43PM.

5. I got on the returning J subway (Train #: 4884) at Marcey at 1:47PM, and stood in the middle of the train on the return ride. The returning subway went below ground at 1:52PM, and I got out of the Essex Street station at 1:55PM.

6. From the Essex Street station, I walked towards Bowery off Canal Street for the B51 bus. I waited for a while, and realized this bus does not run on the weekends. I walked over the Manhattan Bridge to the York Street Station for the F subway.

7. At 2:48 PM I entered the York Street Subway station and immediately got on the F subway (Train #6248) heading towards Coney Island. The F did not run regularly, and I had to switch to the G subway (Train #5915) at the Hoyt-Schermerhorn station that was running on the F line. I sat next to a window on ride to Coney Island. This train went above ground at 3:05PM, then below ground at 3:08PM, and again above ground at 3:17PM. I arrived at Coney Island at 3:34 PM.

8. I took a G subway (Train #: 5914) running on the F line towards Hoyt-Schermerhorn (the last stop of this train). I sat as far from the window on the return ride as possible. The train went below ground at 4:01PM, then above ground at 4:09PM, again below ground at 4:12PM. I arrived at Hoyt-Schermerhorn at 4:17PM. From the station, I walked towards 275A Wyckoff Street.

9. Ending Time of Trip
Date: August 19, 2007
Time: 4:30PM

---

* I had forgotten to look for the number on the side of the train car, but it is similar to the subway car on the returning J train (Train #: 4884).

REPORT - STUDENT #3

The GPS logger started to track much quicker than GlobalSat. It took the latter a few more minutes to start logging. While on the subway, I also noticed that GPS logger was more sensitive and picking up more signals along the way.

Since I was off route for two hours earlier in the day, GPS logger started to have a low battery life during part of the trip on the G subway to Coney Island. To save battery, I turned GPS logger off while it was underground (Turned off at 3:15 PM, turned back on shortly after; turned off again at 3:37PM and on again at 3:42PM). Despite low battery, GPS logger held on until the end of the trip. It continued to pick up signals and track locations of the trip. GlobalSat also picked up signals along the trip, even when I sat away from the windows, but GlobalSat's lag time coming from underground to above ground took much longer than GPS logger that usually began tracking right when the train went above ground.

I had forgotten to charge either GPS overnight, and was relieved GPS logger with its warning signal flashing for the last hour made it through the entire trip. Also, I was not able to complete the B51 trip because that bus does not run on weekends.

DATA SHEET - STUDENT #4

Starting time of your trip:
Date (eg. 08/11/07): 8/19/2007
Time (eg. 9:30 AM): 10:53AM

Ending time of your trip:
Date (eg. 08/11/07): 8/19/2007
Time (eg. 9:30 AM): 5:25PM

REPORT STUDENT #4

There were a couple of problems with the smaller GPS unit. When I plugged it in to charge the indicator light didn't come on. I turned it on and then off and then I was able to charge it. When I first tried to turn it on I couldn't get a signal for nearly 10 minutes. I turned it off and then on again and several minutes later I did get a signal.

At 10:53AM I started out at Fort Washington Ave and 187th, and walked down Fort Washington Ave to 168th and then over to Broadway. I took the A to 42nd St. I walked over to 12th Ave. I got on the M50 at 12:28PM (sitting by the window) and got off 50th St and 2nd Ave at 12:46PM. I then walked up 2nd Ave to 59th St (there was a sidewalk detour between 52nd and 53rd that was enclosed) and then back to 58th St. I got on the Q101 at 1:00PM (sitting one seat away from the window), crossed the Queensboro Bridge on the upper level, and got off at 25th St and Queens Plaza at 1:15PM. I walked 2 blocks along Queens Plaza and took the N to Astoria Blvd. I got on the M60 at 31st and Hoyt at 1:48 (sitting by the window) and got off at 125th and 2nd Ave at 1:55PM.

I then crossed to the other side of 125th , got back on the M60 at 2:04 (standing in the middle of the bus), and got off at 2:12 at 32nd and Astoria Blvd. I took the N train to Queensboro Plaza and got on the Q60 at 2:35PM going over the upper level of the Queensboro Bridge. I got off at 2nd Ave and 60th St. at 2:39PM. I walked down to 49th St and over to 1st Ave. I got on the M50 at 3:122 at 3:12 (sitting away from the window) and got off at 49th and 12th at 3:35. I walked up 12th Ave to 54th, over to 11th, and then back to 50th St. I got on the M50 at 4:24PM (by the window) and got off at 8th Ave at 4:30. I then took the A to 168th. I returned home at 5:25PM.

DATA SHEET - STUDENT #5

1. Starting time of your trip:
Date: 8/25/2007
Time: 7:45 pm

2. For uptown trip:
time of the train emerging above ground: 7:57pm
time of the train going underground: 7:59pm
time of the train emerging above ground again: 8:10pm
time of the train stopped:8:20pm

3. For the downtown trip:
time of the train start: 8:21pm
time of the train going underground: 8:29 pm
time of the train emerging above ground: 8:39 pm
time of the train going underground again: 8:41pm

4. Ending time of your trip:
Date: 8/25/2007
Time: 8:45 pm

REPORT STUDENT #5

7:45 pm, I arrived the beginning of the testing route, the uptown 1 train station at 116th Street, and turned on all the GPS units. 7:50pm, all the GPS units got signal, and then I walked into the station and lost signal. Around 7:56pm, I got on the 1 train; the serial number for that cart is 2219. I sat near the door of the train.

At 7:57pm, the train went above ground when approaching 125th Street station and went underground when leaving it. At 8:10pm, the train went above ground again when approaching the Duckman Street, and kept above ground until it arrive the terminal at 242 Street at 8:20pm.

Then I walked cross the platform and got into another cart with serial number of 2207. I stood in the middle of the cart. The train started and went downtown 1 minute later at 8:21pm. At 8:29pm, the train passed Duckman Street and went underground. Nearly 8:39pm, train reached 125th Street and thus emerged above ground; it went ground again at 8:41pm.

At 8:45pm, I walked outside the train station, waited until the GPS got signal again and than turned them all off. One thing needed to be noted is that I was wearing a quartz watch, thus there might be some seconds difference between my watch and the standard time.