Using GPS Technology to Measure On-Time Running of Scheduled Bus Services

Philip Bullock, Qingjian Jiang, and Peter R. Stopher
The Institute of Transport Studies, The University of Sydney

Abstract

Assessing bus service running times has been a difficult and expensive task for many urban bus operators. This has restricted the ability of operators to collect adequate data to identify problems and improve service levels. Passive Global Positioning System (GPS) devices offer a low-cost means of collecting large amounts of highly accurate data, to be used in an ongoing performance assessment program. Some programming skills are required to break continuous GPS data into information that is meaningful to a scheduler. This article provides an overview of a software application developed to process and analyze GPS datasets collected by a bus operator in Sydney, Australia, in 2002-2003. The data collection procedure and processing algorithms are described, and examples are presented of output produced by the software. The algorithm developed to process the GPS data worked well. We conclude that passive GPS is a cost-effective method of collecting data on performance. For operators running buses on five or more routes, system development costs could be recovered within two to three years.

Introduction

This article provides an overview of a pilot Global Positioning System (GPS) research project undertaken in late 2002/early 2003 by the Institute of Transport Studies at the University of Sydney and a bus operator in Sydney, Australia. The aim of the project was to develop a cost-effective Geographic Information System (GIS) based program to process and analyze GPS data collected on buses operating on a specific route.

The article presents an overview of the steps taken to collect the input data used in the project, and details the trip-processing and timetable query program developed for processing and analyzing the GPS data. Some examples are presented of output produced by the main trip-processing and timetable query program, as well as some of the ways it can be used by schedulers. It is concluded that for operators of most sizes, passive GPS is an attractive method of collecting data on performance.

Background: Difficulties of Measuring On-Time Running

Assessing running times of bus services has traditionally been a difficult and expensive task for the majority of bus operators in Australia and in other parts of the world (Kharola, Gopalkrishna, and Prakash 2003). Until recently, travel times have generally been collected manually by timekeepers positioned at key points along a given route or service corridor. The time-consuming nature of this process restricts the ability of operators to collect large and meaningful samples of data, which could be used to improve timetables and levels of service. It is also difficult, if not impossible, to identify congestion points from such data, and to evaluate the impact that they might have on overall service levels.

Automatic vehicle location (AVL) technology offers a means of collecting large samples of travel time data, which can be used as part of an ongoing performance assessment program. The rapid pace of change in AVL systems, however, can make investment decisions difficult for many bus operators. In the past few years, a number of sophisticated on-line systems have been developed for providing information to customers about bus arrival times, allocating priority at traffic lights, and enabling bus operators to respond to traffic problems in real time (GPS Online 2000; Morehead 2001; Infodev 2003; NextBus 2003). Such applications are not cheap to develop, and may cost in the vicinity of hundreds of thousands, or even millions of dollars. In Auckland, New Zealand, a large-scale real-time pas-

senger information (RTPI)/bus priority system is being developed that will involve fitting more than 700 buses with GPS equipment, providing on-street variable passenger information displays, and modifying traffic lights. The estimated cost of this project is NZD $7 million (Auckland City Council 2003). The system developed for the London bus network is probably one of the largest AVL systems set up to date. There are currently more than 8,000 buses using the system, and passenger information screens will be fitted to 4,000 stops by the time the rollout is complete. To date, almost GBP $50 million has been spent developing the system (GIS Development 2004). AVL systems are also being integrated with automatic passenger count (APC) systems to provide information on boardings and alightings and passenger kilometres of travel (Rossetti and Turitto 2000).

Although relatively little work has been done to evaluate the benefits of these systems, there is evidence to suggest they can have a positive impact on operational efficiency. Strathman et al. (2000) examined a computer aided dispatching and AVL system developed in Portland, Oregon, and found that the system improved on-time performance and reduced total running times. While real-time systems indubitably have a range of benefits, much less is known about the effects they ultimately have on patronage (which is why they are developed in the first place).

The high costs of integrated AVL systems require them to be largely funded by transport authorities, as opposed to individual operators. This is especially the case when systems involve modifications to state-owned assets such as bus stops and roads. Passive or off-line GPS technology, operating independently of other systems, represents a practical, low-cost method for collecting travel time data. Over the past few years, GPS technology has improved markedly and accurate GPS data loggers have become very affordable, and can be purchased for as little as USD $200 to $300. The appeal of this technology lies in its simplicity and affordability. In many situations, operators may only require information to help determine whether their buses are running on time, and where problems might be occurring on the network. Such information does not need to be available in real time to be useful.

One of the key advantages in using data loggers is that they are portable, and can be moved easily between buses operating on different routes and in different regions. Other than a major study undertaken by Kharola et al. (2003) in Bangalore, India, it appears that little work has been done to date using off-line systems to collect GPS data on buses.

Despite the advantages of passive GPS, there are some practical difficulties that need to be overcome when using portable data loggers. These difficulties stem from the fact that low-cost data loggers collect GPS data independently of other systems within the bus, such as on-board ticketing systems. Output files from passive data loggers provide continuous streams of spatial and temporal data (i.e., geographic coordinates, time and date), but no other meaningful reference information (e.g., the route the bus was operating on, trip start and end times, shift changes).

While it would seem practical to use a GPS device that would allow drivers to enter additional reference information, this would probably just make the system unreliable. Bus drivers work in a relatively stressful environment and it is likely that they would often forget to indicate when they started and finished routes or arrived and departed from the depot. Likewise, fully automatic or integrated systems may not be an option because of expenses involved in modifying or upgrading ticketing systems.

Some programming skills are required to convert continuous points into records that are more useful to an operator. Several important tasks need to be undertaken before analysis can take place. First, periods of in-service or out-of-service running need to be defined, and routes need to be identified. This can be a complicated task because operators often design shifts so that buses may switch between different areas and routes, from trip to trip, to maximize vehicle utilization. Once routes are identified, individual trips must then be matched with a timetable to compare scheduled and actual running times.

Input Data

Three main sources of data were required to develop the programs: bus stop coordinates, timetable information, and in-vehicle GPS data collected from the study route. The following sections describe the methods used to collect and edit the input data. All GPS data used in this project were collected using GeoLogger® passive nondifferential GPS data loggers, produced by GeoStats. The Geologgers were fitted with Garmin GPS receivers which have an accuracy rating of ±15 meters, although the experience of the Institute of Transport Studies is that on average it is closer to ±5 meters. All GIS programs were developed using the GISDK™ programming language in Caliper Corporation's TransCAD® package.

Bus Stop Data

Although a database of bus stop locations is held by the NSW State Government, these data were not considered to be sufficiently accurate for this project. Bus stop coordinates were, therefore, collected by the bus operator in late November 2002, using a data logger and a company vehicle. Arrival times at major timing points along the route were recorded, and the GPS data were downloaded and put into separate layers for inbound and outbound stops.

Timetable Data

Timetable data were generated from the scheduling software used by the bus operator and saved in Excel spreadsheets. Minor modifications were required to convert the data into a format that could be recognized by the GIS program. Numerical values stored as times were converted to integers, and columns and rows were transposed, so that each row of the table represented a trip, with columns representing the scheduled arrival times at timing points along the route.

In-Vehicle Bus Data

Data were collected from four buses, starting in late December 2002 and finishing in mid-March 2003. Four buses operating principally on the study route were fitted with data loggers. A formal sampling plan was not considered necessary because the project was mainly focused on development of methodology, and because only one route was considered.

Because the devices were designed to be plugged into the cigarette lighter outlet of an ordinary motor vehicle, some modifications were needed so that the power cords could be plugged into the AV accessory outlets of the buses. Other than this, the devices were relatively easy to install. GPS antennas were easily attached to the roofs of the buses because of their magnetic bases. It was not known what polling rate would be most suitable, so two of the devices were set to record data on one-second intervals, while the other data loggers were set on five seconds. Data were collected 24 hours a day during the study period because the accessory outlets in the buses were constantly powered. As a result, the data loggers needed to be downloaded and cleared every few days.

Trip-Processing Algorithm and Timetable Query

The trip-processing algorithm and timetable query was the core program designed to generate travel time output from the GPS data files. There are essentially three main tasks performed by the algorithm within the program. First, continuous GPS

records are broken into separate blocks of records, or basic trips (trip definition). Next, the program examines these basic trips, determines the type of trip made, and analyzes running times and travel times between timing points. In the third part of the program, GIS maps, layers, and selection sets are created so that pro-cessed data can be viewed and analyzed by the user.

Trip Definition

Three criteria were used to break continuous data into basic trips. In deciding where to insert a break point or trip end, the program examines:

1. Whether records appear in one of three areas: the depot (Depot), and the two end points of the route (stop 1 and stop 17)
2. The number of bus stops traveled through on the study route
3. Any reversal in the direction of travel

In the first step taken in the trip definition process, coordinates of the depot and bus stops are loaded into a temporary array. The location of each GPS record is examined and an additional array is created identifying GPS records that are located within 50 meters of a bus stop, and 120 meters of the depot. When more than one point is located within the radius, the identification of the closest point to the center is recorded in the array. GPS records are then sorted into separate groups within the GIS layer (selection set) for each day.

For each day's worth of records, the program searches for points that start or end at the depot, stop 1 (ST1), or stop 17 (ST17). That is, if the first record of the day is found within the depot, the program then looks for the next location that subsequent records appear in. There are three possible locations considered—the depot, ST1, or ST17 (if more routes were defined within the program structure, more end points would be searched). If a bus drove from the depot in the morn-ing to one end of the route, stopped briefly to pick up passengers and then made a scheduled trip along the study route, the depot would be the first location marked, ST1 the second, and ST17 the third.

Within each combination of the three locations (depot–ST1, depot–ST17, ST1–ST17, etc.), a separate series of subcommands examines the number of stops passed to determine the likely trip ends. Figure 1 provides an illustration of how this process works. If a bus traveled from ST1 to ST10 and passed through 15 to 17 stops, this would mean that the bus traveled along the study route without deviation, and the trip end would therefore be defined as ST17. If the number of stops was less than 15, this would mean the bus traveled only part of the route

and so the trip would be coded as a "special trip" (e.g., if a bus ran back to the depot via some alternative route to save time). If the number of stops is greater than 17, then it was likely that the bus has made more than one trip, and the trip end would be defined as the point that the bus changed its direction of travel (a change from inbound to outbound). A similar sequence of commands is used to examine records between the three main locations.

Figure 1. Process Used to Define Basic Trips
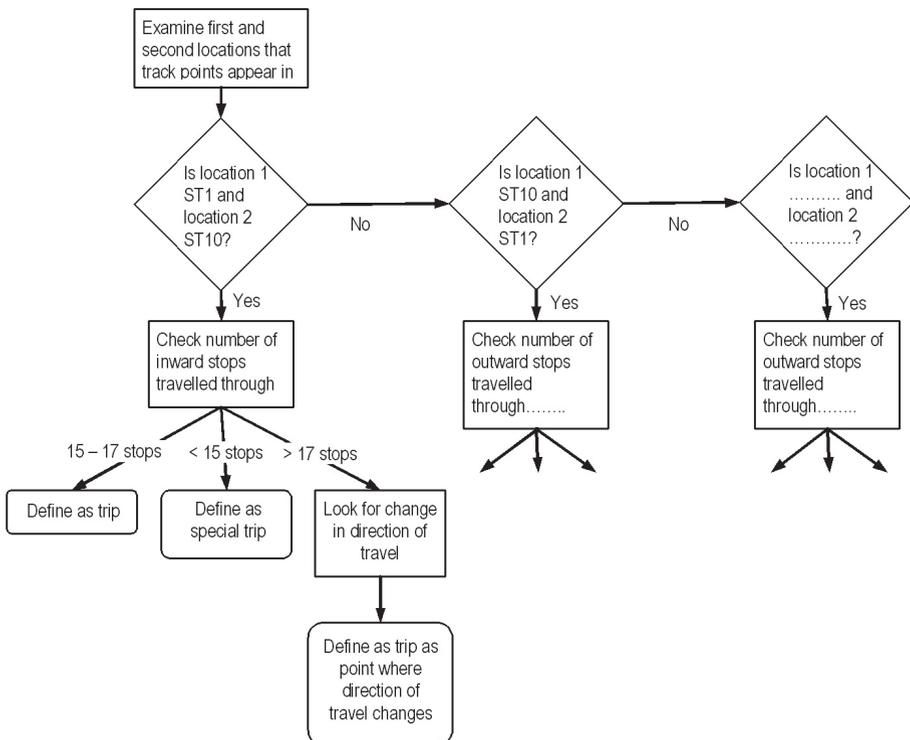
Trip Type Definition and Timetable Query

Once the program has flagged the likely start and end points of trips, the algorithm then defines the type of trip made. Beginning with the first trip of the first day, the program examines each set of records and classifies them into one of the following categories: Route A (main study route), route B, route C, trips out from the depot (O_Depot), trips into the depot (I_Depot), trips made out from the

depot and straight back to the depot without stopping (D_2_D) and unknown trips (UNKNOWN). The direction (inbound or outbound) is also determined for each route.

Whenever the program detects a trip made along the study route, a subroutine assesses on-time running and measures the time taken to travel between timing points. On-time running is measured by comparing the time the bus arrived at a timing point (the time recorded by the data logger) with the time that the bus was scheduled to arrive (the time shown on the timetable). This requires each GPS trip made along the study route to be correctly matched with trips shown on the timetable. From the data observed as part of the validation procedure, it was noted that most buses tend to start within just a few minutes of their scheduled start time; thus, in most cases, it appeared quite easy to determine which GPS trip belonged to which timetable trip.

Once the GPS and timetable start times have been matched, the program then examines the time the bus arrived at each timing point, and calculates the difference between the GPS arrival time, and the scheduled arrival time. Travel times are also calculated between each set of timing points.

Creation of Maps and Output Files

The program opens a base map stored in the specified directory and imports the GPS data in the form of a single GIS point layer. Within this layer, each trip is marked within a selection set. A number of different output files are produced, including a trip summary file and timing check output files for both inward and outward directions. Table 1 shows a selection of data contained in the trip summary output file. Start and end times are shown for each trip as well as the time that the bus was stationary between trips (lay-up time). Scheduled travel times are shown for trips that were made along the study route (Route A).

Table 2 shows a sample of output generated from the timetable query. The columns with single timing point names (ST1, ST2, etc.) show the difference between the scheduled arrival time, and the actual GPS arrival time for each of the timing points along the study route. Columns with multiple timing points (ST1_ST2, ST2_ST3 etc.) show travel times recorded by the GPS between timing points.

GIS is a very powerful tool for visualizing spatial data; however, the data query features in most standard GIS packages are relatively simple and do not allow users to specify multiple attributes or conditions within a single query. A data selection set toolbox was designed as a visualization tool to allow people not overly familiar

Table 1. GPS Data Summary File

Date	Route	Dir.	Start Time	End Time	Layup Time	Travel Time	Scheduled Travel Time
161202	O_Depot	Out	6:39:58	6:54:23		00:14:25	
161202	Route C	In	6:54:32	8:16:23	0:00:09	01:21:51	
161202	Route C	Out	8:40:13	10:05:36	00:23:50	01:25:23	
161202	I_Depot	In	10:06:11	10:15:16	0:00:35	0:09:05	
161202	O_Depot	In	16:26:29	17:18:11	06:11:13	00:51:42	
161202	UNKNOWN	Out	17:19:06	18:33:47		01:14:41	
161202	I_Depot	In	18:34:28	18:43:08	0:00:41	0:08:40	
171202	O_Depot	Out	5:12:08	5:22:08	10:29:00	00:10:00	
171202	Route A	In	5:22:13	6:33:19	0:00:05	01:11:06	01:18:00
171202	Route A	Out	6:57:29	8:30:35	00:24:10	01:33:06	01:43:00
171202	Route A	In	8:59:56	10:50:58	00:29:21	01:51:02	01:38:00
171202	Route A	Out	11:00:23	12:45:10	0:09:25	01:44:47	01:43:00
171202	Route A	In	13:02:11	14:45:53		01:43:42	01:38:00
171202	Route A	Out	14:54:38	16:38:55	0:08:45	01:44:17	01:43:00
171202	Route A	In	17:02:55	18:43:56	00:24:00	01:41:01	01:38:00
171202	Route A	Out	18:57:21	20:30:57	00:13:25	01:33:36	01:33:00
171202	I_Depot	In	20:35:12	20:53:37	0:04:15	00:18:25	

Table 2. GPS Travel Time Output File

TRP_ ID_S	ROUTE_ NO	DATE	WEEK DAY	TT_ DAY	S_TIME	S_TIME _S	ST1	ST1 ST2	ST2	ST2_ ST3
5	Route A	281202	Saturday	2	6:33:46	6:30:00	3:46	7:24	1:10	3:06
9	Route A	281202	Saturday	2	12:51:28	12:50:00	1:28	10:09	:37	3:46
11	Route A	281202	Saturday	2	16:51:58	16:50:00	1:58	14:30	5:28	5:00
13	Route A	281202	Saturday	2	20:52:22	20:47:00	5:22	6:15	3:37	2:05
20	Route A	291202	Sunday	3	9:51:12	9:50:00	1:11	8:42	-1:07	5:00
23	Route A	291202	Sunday	3	16:02:30	15:50:00	12:30	14:25	15:55	4:21

with GIS to run advanced queries on a large dataset. Figure 2 shows the selection set toolbox in TransCAD®. The toolbox allows GPS records to be filtered using any combination of the following six criteria: speed, day of week, time of day, route, timing points, and direction of travel. Once selected, the user can apply color themes on average speeds to highlight points of congestion along the route.

Figure 2. GIS Query Tool
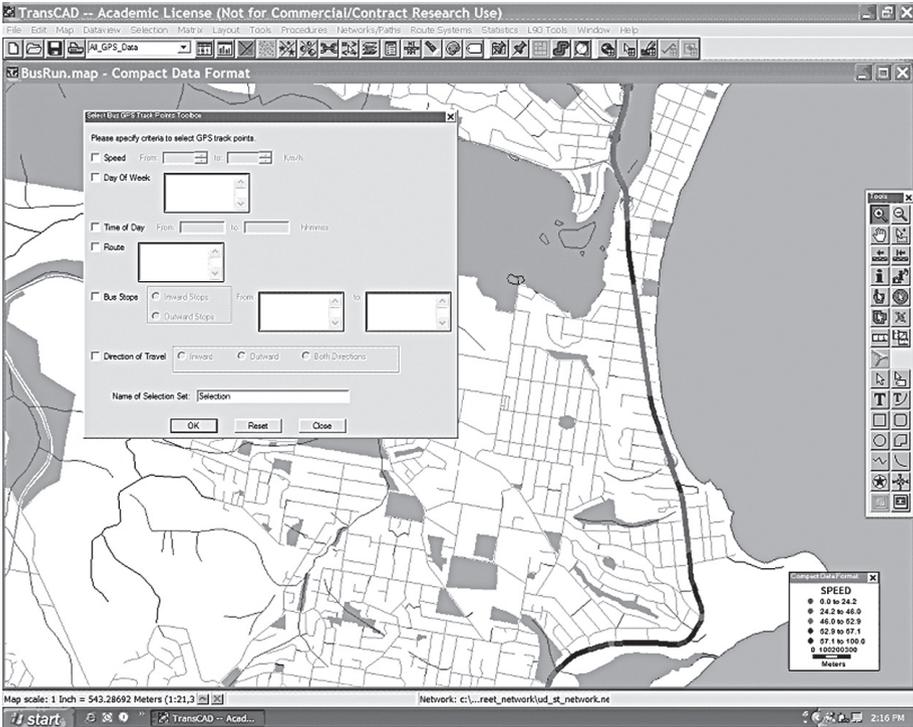
Validation

The programs were validated using eight GPS data files collected from December 2002 to March 2003. Two files were selected from each of the four buses that collected data, to ensure an even spread of dates and a balance between the various polling rates. GPS summary files and timing check files were compared with fare collection data reports provided by the bus operator. These reports were generated from data downloaded from driver smart cards, and represent a record of actual schedules (as opposed to planned schedules developed in the scheduling software). Although there are a number of limitations of using fare collection data (shift times are shown but not bus operation times, in-depot and out-depot movements are not specifically identified), they provide a reasonable record with which to compare the GPS data. Two main tasks were performed as part of the validation process. First, summary files were checked to ensure that trip types (i.e.,

the route) were correctly defined. Second, timing check files were examined to make sure that GPS trips were correctly matched with the timetable.

Overall, the trip detection algorithm worked well. Table 3 shows a breakdown of trips detected by the trip-processing algorithm for the eight data files. Of the 251 trips detected, 96 were Route A trips. Route B and Route C trips comprised 11 percent of the trips detected by the program, while around 100 trips were made to and from the depot. Three trips were made where the bus left the depot, drove two or three blocks, and then drove straight back to the depot.

Table 3. Trips Detected by Trip-Processing Algorithm

Trip Destinations	No. of Trips	Trips as % of Total	No. of Undetected/ Misclassified Trips	Undetected/ Misclassified Trips as % within Group
Route A	96	38.2%	0	0.0%
Route B	2	0.8%	0	0.0%
Route C	25	10.0%	0	0.0%
In Depot	50	19.9%	7	12.0%
Out Depot	51	20.3%	5	9.8%
Depot - Depot	3	1.2%	0	0.0%
Unknown	24	9.6%	24	100.0%
Total Trips	251	100.0%	35	13.9%

No errors or inconsistencies were found in any of the 123 trips coded as Route A, B, or C, which suggests the program interpreted the data very well. Table 3 also shows the number of trips that went undetected or were misclassified by the program. A total of 12 trips were misclassified as either in-depot or out-depot. Of the 7 trips within the in-depot group, 4 were actually Route A trips which appeared to end prematurely. The remaining 3 trips, misclassified as in-depot, incorporated travel made on routes not defined within the program structure and could not be correctly interpreted. Five out-depot trips also incorporated travel on a number of routes which were not defined within the program structure.

A total of 24 trips went undetected by the program and were coded as unknown. Table 4 provides an explanation of what actually took place in the case of each of

these trips. Eight of the trips were Route A trips which were missed for a variety of reasons. Seven of these 8 trips were missed because of signal loss, principally around the final stop that the bus traveled to in the CBD (ST17). Interestingly, 5 of these events (and one third of the total unknown trips) occurred in one data file, which suggests there may have been some power problems with the memory storage unit. (When the battery of a memory storage unit becomes low, data often become patchy.) Four Route C trips were also missed because of signal problems. A total of 10 trips were coded as unknown because the routes were not defined within the program structure. Two other trips were missed because of a data logger malfunction (duplicate time values), the cause of which was probably low power or a bad signal.

Table 4. Explanation of Unknown Trips

Trip Description	No. of Occurrences
Route A, with loss of signal	8
Route B and C, with loss of signal	4
Undetectable routes	10
Other Data logger fault	2
Total	24

Urban canyon effects degraded the quality of CBD-based travel time output and, unfortunately, these problems could not be fully resolved. Travel times between CBD stops were often coded as missing in output files because no points would be recorded within the buffer areas, despite the fact the bus would have passed the stops. Because the study routes ended just outside the city, urban canyon problems generally caused no problems in the trip definition component of the program. If the route finished in some other part of the city, urban canyon effects would have caused significant problems because in many cases, track points may not have appeared in the first or last stops. This would have resulted in a lot more trips being coded as unknown.

The algorithm developed to compare GPS times with the timetable also worked well. Start times appeared to be correctly matched against all 96 trips made along

the study route. In almost all cases, it was fairly obvious that the correct time was selected by the program because the GPS start time was no more than a few minutes before or after the scheduled start time (headway was 30 minutes for most times of the day).

No evidence was found to suggest that a one-second polling rate was superior to a five-second rate. There was no observable difference in travel time output produced from one- and five-second files and a one-second polling rate appears to offer no benefits to offset its greater memory storage requirements. If the data loggers used for this project were set to five seconds and only recorded while the bus was in motion, the memory storage units could probably have been left in the buses for around two or three weeks before they needed to be changed.

Assessing On-Time Running

A specialized Excel spreadsheet was developed to allow users to manipulate output files produced by the programs developed in TransCAD®, and to generate statistics on travel times and differences between scheduled times and actual running times. According to Strathman et al. (2000), these are probably the most widely recognized indicators of service reliability. The spreadsheet was designed to allow GPS data to be filtered according to date, day of week, time of day, route, bus number, and travel times.

Table 5 shows a summary of the output data generated by the timetable query. Differences are shown between GPS travel times and scheduled travel times for all inbound trips made along the study route in the validation files. These statistics could also be generated for specific time periods such as peak/off peak and weekday/weekend; however, the focus here is to provide an overview of what the output looks like and how it might be used by the operator. Positive numbers represent late running, while negative figures indicate that the bus arrived early. For Route A trips observed in the validation files, buses arrived an average of 3 minutes and 59 seconds late to the final stop (ST17). As always, care needs to be taken interpreting output because the numbers may be influenced by one of two outliers. In this case, it can be seen that the maximum value column shows at least one Route A trip was more than 45 minutes late to ST17. Almost all of the maximum values were attributable to this one Route A trip made on a Sunday, which started 7 minutes late and became increasingly late as the trip went on. Before any

meaningful analysis can be done with the spreadsheet, it is obviously necessary to search for outliers like these and flag them, or exclude them from the dataset.

Table 5. Differences Between Scheduled Arrival Times and Actual Times (Inbound Trips)

Timing Point	Count	Average	Median	Standard Deviation	Minimum	Maximum
ST1	49	0:02:31	0:01:31	0:02:56	-0:00:07	0:14:25
ST2	50	0:01:24	0:00:33	0:03:28	-0:02:23	0:15:55
ST3	50	0:00:47	0:00:02	0:03:26	-0:02:31	0:15:16
ST4	50	-0:00:26	-0:01:16	0:03:36	-0:04:34	0:14:16
ST5	50	0:01:17	0:00:43	0:03:46	-0:04:10	0:16:55
ST6	50	-0:00:15	-0:00:53	0:03:44	-0:05:47	0:13:46
ST7	50	0:00:41	-0:00:05	0:03:26	-0:06:04	0:12:55
ST8	50	0:00:28	0:00:40	0:03:23	-0:06:22	0:12:10
ST9	50	0:00:24	0:00:25	0:03:16	-0:05:19	0:11:35
ST10	50	0:01:37	0:01:37	0:03:33	-0:05:19	0:14:35
ST11	50	0:03:45	0:03:10	0:03:51	-0:02:49	0:16:26
ST12	50	0:03:38	0:03:00	0:04:00	-0:03:17	0:16:46
ST13	50	0:04:17	0:03:25	0:07:30	-0:05:58	0:46:47
ST14	50	0:05:15	0:04:46	0:07:47	-0:06:07	0:48:22
ST15	36	0:02:41	0:03:12	0:05:18	-0:09:13	0:11:58
ST16	42	0:02:12	0:02:07	0:05:18	-0:10:07	0:12:37
ST17	50	0:03:59	0:03:05	0:09:10	-0:11:33	0:45:13
TOTAL TRAVEL TIME	50	1:37:10	1:38:46	0:11:49	1:10:30	2:10:43

The counts shown for each timing point in Table 5 vary because, for some trips, there were no records located within a 50-meter radius of the stop, so it was not possible to perform a timing check. (This means that minimum and maximum values may not always correspond to the same trip, and may differ considerably.) This occurrence was most pronounced in the CBD because of urban canyon effects.

Figure 3 shows the median time differences between scheduled arrival times and actual arrival times. From the limited data observed in the validation process, it can be seen that Route A inbound services experienced their greatest general delays from ST11 to the end of the route. This information could be used by an operator in a number of ways. The first course of action would be to determine the cause of the discrepancies between the scheduled times and the actual running times. Early running is likely to occur when drivers do not stop at holding points when they arrive early, while late running can result from buses starting late, or from traffic congestion along routes. Early running, particularly where headways are half an hour or more, may be more detrimental to service quality than slight delays, because it may result in passengers having to wait for subsequent buses. When bus services are frequent, reliability may be better reflected in the ability to maintain headways, rather adhering to schedules (Strathman et al. 2000). Unless successive buses are fitted with data loggers, it would not be possible to calculate headway ratios from output files generated by this application. This suggests that the system may be most useful for routes with headways of at least 15 minutes or more. If discrepancies between scheduled and actual times are considered large enough, schedulers could adjust the timetable to more accurately reflect actual travel times. In the case of delays, travel time data could be used by bus operators to argue for improvements in traffic management (e.g., bus lanes).

Figure 3. Differences Between Scheduled Arrival Times and Actual Times (Inbound Trips)
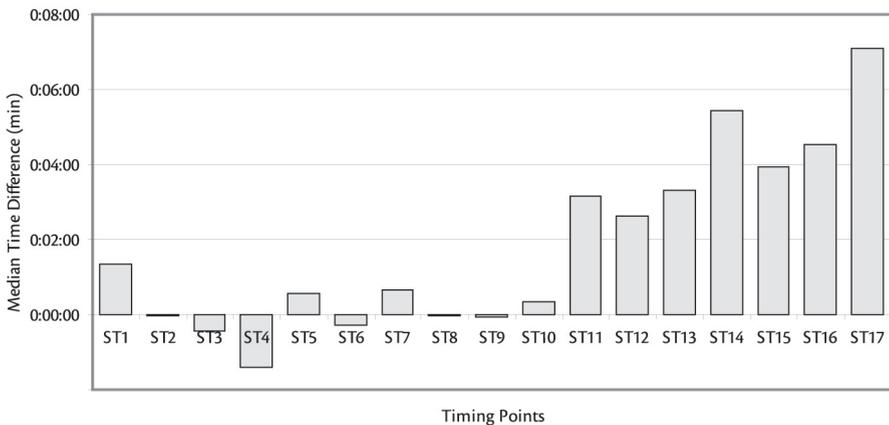
Table 6 provides descriptive statistics for travel times between timing points. A larger set of descriptive statistics is generated for general travel times because these are more commonly used by schedulers than the time differences shown in Table 5. Because good output on travel times has been very difficult to obtain up until now, it is largely unknown which statistic is the best to use for planning schedules. In travel time research literature, median times tend to be favored over averages because they are less sensitive to outliers (Quiroga 1997). For the operator of a transport service, however, a statistic such as the 85th percentile might be more appropriate. Median travel times and 85th percentile times are displayed in the graphical output generated by the spreadsheet, as shown in Figure 4. In general, the 85th percentile times are one to two minutes higher than the median times. Standard deviation is another potentially useful statistic for operators. One of the key advantages of using GPS to collect a large sample of travel times is that it provides information on the variation of travel times across different times and days of the week.

Table 6. Travel Times Between Timing Points (Inbound Trips)

Timing Points	Count	Average	Median	Std. Dev.	Min.	Max.	Cumulative Median Travel Time	Segment Time as % of Total Travel Time	85th Percentile
ST1 – ST2	49	0:09:37	0:09:25	0:01:53	0:06:15	0:14:30	0:09:25	10.01%	0:11:11
ST2 - ST3	50	0:04:01	0:03:58	0:01:03	0:02:05	0:06:25	0:13:23	4.22%	0:05:01
ST3 – ST4	50	0:06:34	0:06:30	0:00:57	0:04:40	0:09:53	0:19:53	6.92%	0:07:23
ST4 – ST5	50	0:04:18	0:04:13	0:01:03	0:02:25	0:07:10	0:24:07	4.48%	0:05:26
ST5 – ST6	50	0:05:29	0:05:33	0:00:52	0:03:31	0:07:34	0:29:39	5.90%	0:06:23
ST6 – ST7	50	0:02:53	0:02:42	0:01:38	0:00:26	0:06:35	0:32:21	2.87%	0:04:38
ST7 – ST8	50	0:09:20	0:09:07	0:01:41	0:06:10	0:12:48	0:41:28	9.69%	0:11:16
ST8 – ST9	50	0:03:56	0:03:48	0:00:55	0:02:39	0:07:05	0:45:17	4.04%	0:04:45
ST9 – ST10	50	0:06:02	0:05:58	0:01:05	0:04:15	0:09:33	0:51:15	6.35%	0:06:55
ST10 – ST11	50	0:05:58	0:05:50	0:01:04	0:04:13	0:08:44	0:57:05	6.20%	0:07:03
ST11 – ST12	50	0:00:53	0:00:31	0:00:36	0:00:21	0:03:15	0:57:36	0.55%	0:01:25
ST12 – ST13	50	0:14:39	0:13:36	0:04:44	0:09:39	0:44:01	1:11:11	14.44%	0:17:03
ST13 – ST14	50	0:04:58	0:05:02	0:01:07	0:02:42	0:09:25	1:16:14	5.36%	0:05:46
ST14 – ST15	36	0:07:46	0:07:44	0:00:59	0:05:49	0:09:39	1:23:57	8.21%	0:08:48
ST15 – ST16	34	0:02:55	0:02:47	0:00:59	0:01:36	0:05:45	1:26:45	2.96%	0:03:37
ST16 – ST17	42	0:07:24	0:07:22	0:01:34	0:04:30	0:10:50	1:34:06	7.83%	0:08:56
TOTAL TRAVEL TIME	50	1:37:10	1:38:46	0:11:49	1:10:30	2:10:43			1:46:27

Figure 4. Travel Times Between Timing Points (Inbound Trips)
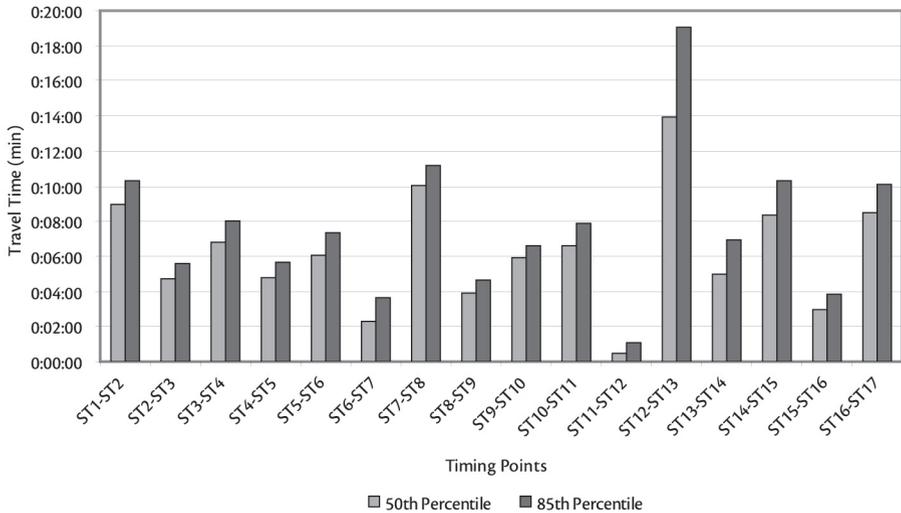
Cost Effectiveness

From the results of the pilot study, it is estimated that the cost of developing a similar set of tools for 5 to 10 routes operating from a common depot would be in the vicinity of AUD $25,000. Assuming no more than 10 routes are served from a single depot, it would probably only be necessary to invest in two or three data loggers which would be purchased for no more than a total of AUD $4,500 each. Provided units can be easily transferred between buses, the data costs themselves are negligible.

The cost of collecting two hours' worth of running times along a single route using three time keepers would probably be in the vicinity of AUD $180. Assuming that four hours' worth of observations are collected for five routes every two months, the annual costs would total about AUD $10,000 excluding data entry costs. This means that the cost of the software could probably be recovered in two to three years, conservatively. These calculations do not take into account the improved quality of the data collected by GPS, and the fact that many more observations can be collected than manually collected data.

For small operators with only a few short routes, the costs of the system may not outweigh the benefits, particularly if they are operating short feeder services

in areas where there is generally very little congestion. The system would probably be most valuable to companies operating buses on long routes (45 minutes upwards), or in areas where traffic delays are encountered.

Conclusions

This pilot project has shown that it is feasible to collect accurate travel time data using simple, passive GPS devices operating independently of bus drivers and existing on-board computer systems. The approach taken for this project represents a viable, low-cost method for collecting accurate travel time data which can be used to measure on-time running and provide useful data for schedulers. One of the main shortcomings of the GPS devices and GIS processing program described here is that they cannot be easily integrated with other bus systems such as APCs. It is worth noting, however, that it would be possible to link GPS data from the data loggers with ticket sales data from on-board ticketing machines by matching times recorded in both files in a post-processing procedure.

With system development costs aside, the data collection costs associated with the approach taken in this project were very low. Hundreds of hours' worth of data were collected on the study route for little more than the cost of coordinating the movement of data loggers between the depot and head office. The challenge in using GPS to collect travel time data is no longer how accurate data can be collected, but how data can be collected and managed for buses operating in a number of different areas. If anything, GPS can collect too much information, which can make data management and interpretation difficult. Using the portable devices discussed in this article, the operator can control how much is collected.

Overall, portable data loggers appear well suited to measuring travel times and on-time running. It is not necessary to have an entire fleet of buses equipped with GPS to provide information useful to schedulers. With a small investment in just two or three data loggers, it would be possible to implement a continuous survey of many different routes. Data loggers could be rotated through different depots every few weeks and a large travel time database could be built and expanded over time.

References

Auckland City Council. 2003. Real time passenger information [Online], Accessed: *http://www.aucklandcity.govt.nz/auckland/transport/busesfirst/realtime*.asp, [2003 July 01].

GIS Development. 2004. GPS technology to find place in London buses [Online], Accessed: *http://www.gisdevelopment.net/news/viewn.asp?id=GIS:N_vsmkrtbz* [2004 August 9].

GPS Online. 2000. GPS online tracks Sydney buses: 1 September 2000 [Online], Accessed: *http://www.gpsonline.com.au/news&events/news_index/news2000.html* [2003 July 01].

Infodev. 2003. Automatic Vehicle Localization System (AVLS-GPS) [Online] Accessed: *http://www.infodev.ca/english/pagebus/avls.shtml* [2003 July 01].

Kharola, P. S., B. Gopalkrishna, and D. C. Prakash. 2003. Fleet management using GPS and GIS [Online] Accessed: *http://www.gisdevelopment.net/application/Utility/transport/utilitytr0010.htm* [2003, July 01].

Morehead. 2001. GPS keeps track of buses. Civic.Com, June 26, 2001 [Online], Accessed: *http://www.civic.com/civic/articles/2001/0625/web-gps-06-26-01.asp* [2003 July 01].

NextBus. 2003. How NextBus works [Online], Accessed: *http://www.nextbus.com/corporate/works/index.htm* [2003, July 01].

Quiroga, Cesar A. 1997. An integrated GPS-GIS methodology for performing travel time studies. Ph.D. Dissertation, Louisiana State University, Baton Rouge, LA.

Rossetti, M. D., and T. Turitto. 2000. Design of an integrated transit monitoring system based on RF/ID. International Journal of Technology Management 1 (2/3).

Strathman, J. G., K. J. Dueker, T. Kimpel, R. L. Gerhart, K. Turner, P. Taylor, S. Callas, and D. Griffin. 2000. Service reliability impacts of computer-aided dispatching and automatic vehicle location technology: a Tri-Met case study. Transportation Quarterly 54 (3).

About the Authors

PHILIP BULLOCK (*philbullock@gmx.net*) is a research analyst at the Institute of Transport Studies (ITS), University of Sydney. After completing an honors degree in human geography, Mr. Bullock managed transport and distribution for a large Sydney-based importer of footwear and other fashion goods. He began working at ITS in mid-2000 and has been involved in a number of GPS-related projects, including a GPS household travel survey and a project developing software to process and analyze GPS data collected on buses. Mr. Bullock has also been involved in work to simulate household travel survey data. In 2003, he completed received a master's degree in logistics management from ITS. Mr. Bullock is currently taking extended leave from ITS to travel overseas.

QINGJIAN JIANG (*qingjianj@its.usyd.edu.au*) has been a research analyst at ITS since December 2001. He is working with Professor Stopher on a number of projects including using passive GPS devices in household travel surveys, bus service studies, Warringah travel time studies, and travel behavior modification evaluation projects. He is responsible for developing GISDK™ programs for automation of GPS data processing and visualization of travel survey data. Mr. Jiang also participated in developing Internet surveys and other software applications for discrete choice experiments and transport research. Mr. Jiang is a Ph.D. candidate whose topic is route choice modeling using GPS HTS data.

PETER R. STOPHER (*peters@its.usyd.edu.au*) has more than 33 years of professional experience as a university teacher and professional in transport planning. A specialist in travel-demand forecasting, travel behavior research, transport survey methods, and transport and environment issues, Dr. Stopher has published more than 130 papers in leading international journals and has also published a number of books on transport-related topics. He has advised agencies around the world on various aspects of transport planning and data collection.

Dr. Stopher has made major contributions to the profession as a founding member of the TRB Committee on Traveler Behavior and Values, of which he has just been awarded Emeritus Membership, and also founded the TRB Committee on Survey Methods.