An Internet-based Prompted Recall Diary with Automated GPS Activity-trip Detection: System Design

Sean T. Doherty and Dominik Papinski
Department of Geography & Environmental Studies
Wilfrid Laurier University
Waterloo, Ontario
Canada N2L 3C5
Email: sdoherty@wlu.ca

and

Martin Lee-Gosselin
École Supérieure d'Aménagement du territoire et de Développement
Université Laval
Québec, Canada, G1K 7P4
Email: Martin.Lee-Gosselin@crad.ulaval.ca

November 2005 (Conference CD Rom Draft)

Word count: 5474 + 8 Figures = 7474 words

To be presented at the 85th Annual Meeting of the Transportation Research Board, January 22-26, 2006.
Abstract

The use of Global Positioning Systems (GPS) to detect certain aspects of human activity and travel patterns is well established. This is especially true for extracting start and end times of activities and trips that are difficult for people to recall accurately. Some attempts have been made to develop automated routines for such detection, although most involve by-hand analysis to complete the picture. Even then, not all aspects of human activity are traceable via GPS, such as activity types and involved persons, and there is growing demand for other activity attributes related to underlying decisions processes. It is inevitable that some form of interaction with subjects is needed to obtain these attributes and gain a better understanding of behaviour. Conducting a “Prompted Recall” survey is one approach to filling this gap. Three possible approaches are discussed in this paper, including sequential, temporal/tabular, and spatial. The detailed design of a tabular diary-like prompted survey is then presented. It is unique in that it is administered via the Internet, designed explicitly for person-based GPS tracking, incorporates wireless data transmission to a central server, and utilizes a fully automated activity and trip detection algorithm. It also has a compact “diary” style format and utilizes an intuitive means for correcting and updating detected attributes to arrive at a complete schedule. Example results are presented that demonstrate the potential of the system to handle simple and complex patterns, and for organizing prompts for detailed attribute acquisition. Several key avenues for future development and application are discussed.
INTRODUCTION

Over the past several decades, travel “diaries” have persisted as the main data collection tool for understanding and forecasting travel behaviour. Along the way, many enhancements, supplements and alternative delivery mechanisms have been developed to increase their accuracy, reduce respondent burden, and extend the types of information gathered. This includes the move to more activity-based and time-use diary methods. Increasing attention is also being paid to issues of data quality, standards, and non-response \([1]\). Wholly new techniques have also emerged that incorporate or extend diaries in key ways to capture people’s motives, decisions processes, and future behavioral responses. These include stated preference/adaptation surveys \([2, 3]\), gaming and simulation \([4-6]\), panel surveys \([7-9]\) cognition experiments such as “think aloud” protocols \([10]\), activity scheduling process surveys \([11-13]\), and a wide variety of qualitative approaches \([14, 15]\).

Increasingly, new technologies such as computers, the Internet, and passive tracking technologies are also being used to support these efforts. In particular, Global Positioning Systems (GPS) are receiving considerable attention and experimentation as a means to supplement and potentially replace travel diary methods. Most off-the-shelf GPS devices are capable of providing location accuracy (longitude, latitude, and altitude) within 5 to 10 meter accuracy. After a warm-up of 30-90 seconds, such devices can provide such information on a second-by-second basis if desired. When combined with a data logger such as a hand-held computer, GPS can passively provide a highly accurate trace of personal or vehicular movements over long periods of time (e.g. \([16-18]\)).

Geographic Information Systems (GIS) are an ideal tool for storing, processing, and visualizing this data. Recent research has shown that it is possible through post-processing to accurately determine the vast majority of travel routes, traffic flows, trip start/end times, and trip stops (or activity locations) from GPS data \([19-22]\). For example, Bullock et al. \([20]\) utilized a combination of automated rules and by-hand techniques to successfully detect 91% of trips. A basic rule for identifying trip ends was first used, consisting of at least 120 seconds worth of stationary GPS points. Additional rules dealt with situations such as signal loss or vehicle reversals. Doherty et al. \([23]\) utilized GIS (Geographic Information Systems) to develop algorithms that detected trip routes by matching GPS points to the nearest link using a “buffering” technique, and identified trip-ends (or “activity” nodes) based on the clustering of GPS points, leading to over 90% accuracy without further by-hand processing. When compared to self-reports, GPS has been shown to increase trip detection by 20-30%, especially for shorter more discretionary trips \([20, 24, 25]\).

In most of these applications, GPS has been used largely to detect observed attributes of vehicular trips, including start, end times, routes and trip ends. In context of emerging travel behaviour and activity-based approaches, these represent only a subset of the types of information commonly sought in diary-based surveys, albeit the most difficult for people to recall in most cases. Person-based GPS tracking holds the potential for extending the extent of detection to include trips by other modes (walk, bike, bus, train, etc.), and more complex activity patterns. Overall, GPS appears to offer unprecedented potential to improve the data quality and extent of travel surveys, and thus inherently reduces respondent burden.
As an alternative to GPS, Asakura and Hato [26] showed how cellular phones can be used to track people at 2-minute intervals to within 20–150 meter accuracy. Instead of satellites, location is estimated using the antennas of service providers located about every 100 meters in Osaka, Japan. In a test, Asakura and Hato use a basic “move or stay” rule to post-process such data to detect trips by start and end time, and demonstrate with examples how these compare to a trip diary for the same person. They conclude that such a system holds potential for monitoring travel behaviour in a real world environments, but that much further research is needed. In particular, they suggest much further development of trip, mode and route detection algorithms based on statistical analysis of the observed location data rather than ad-hoc rules. They also suggest that future tracking surveys incorporate a cell-phone based questionnaire to capture undetectable trip attributes such as trip purpose.

Despite all these key advances, it is largely recognized that GPS or any other tracking technologies will never provide for the ability to completely replicate a persons’ activity-travel patterns, nor capture all the typical elements included in activity-trip diaries. Well known GPS signal outages, positional inaccuracies, cold-start issues, and other technical problems prevent 100% GPS tracing. These problems are even more acute for person-based GPS tracking, as people enter buildings/tunnels frequently resulting in GPS signal loss or just get their body in the way of the GPS antennae. Even if these problems are overcome through improved technologies such as Assisted-GPS [27], it is likely that no algorithm will ever be able to 100% detect all the complexities of the activities and trips made by people – multi-stop activities, short drop-off activities, and other random-like patterns are likely to remain difficult to detect automatically.

Perhaps more importantly, there are activity-trip diary elements that are difficult if not impossible to trace with GPS, including involved persons (unless perhaps, everyone was GPS traced), activity types (or trip purposes), and the underlying motives and decision processes that underlying observed patterns captured by diaries. The demand for these latter elements has been ever increasing in our quest to better understand and forecast travel behaviour [28-31]. Even then, some researchers are going beyond the detection of observable patterns based on GPS, and have demonstrated potential for automated detection of attributes such as trip purpose as inferred from underlying land-use data [32] and even underlying activity scheduling decisions such as activity-travel modifications and impulsive decisions [33].

In light of this, a hybrid class of GPS supported diary surveys is emerging that combine passive GPS tracing for the detection of a portion of peoples’ activity-trip pattern, followed by an active and explicit attempt to prompt users to recall undetected elements and supplemental information. This hybrid design can be termed a “Prompted Recall” (PR) survey [34]. Such surveys recognize that, rather than relying on time-consuming manual inspection and interpretation of GPS-traced data to determine a persons activity-travel pattern (including interpreting gaps and detecting activities and trips), the real “expert” on the patterns is approached to do so – the persons themselves.

OBJECTIVES

The objectives of this paper are to explore the different types of Prompted Recall methods that are possible (including reference to those few published applications that exist), identify future challenges and opportunities, and present the design and initial testing of an automated on-line
prompted recall system that builds upon these concepts. This experience is also used as a basis to speculate about future design enhancements.

PROMPTED RECALL METHODS

There are several possible methods/interfaces for conducting prompted recall surveys, including sequentially, spatial, temporal/tabular or some combination of these. Sequential methods would involve some sort of systematic querying of subjects for missing or supplemental activity-trip pattern attributes, such as a series of questions over the phone, or regular prompting for attributes simultaneous to location tracking. For example, Murakami and Wagner’s [16] used GPS along with a hand-held computers to prompt for trip purpose and passenger information immediately prior to each auto trip. If location-enabled cell phones are used, a similar approach could be adopted.

Temporal/tabular prompted recall methods would involve some form of time-ordered display of GPS-detected activities and trips, including start and end times (activities and trips), travel times/modes, and activity locations. Subjects would be asked to correct, confirm, and/or add supplemental information to such a display. The most logical display would be a table or calendar format, using paper-and-pencil or an interactive computer display (see also the “illustrative examples” section below). Such a system was envisioned by Doherty et al. [19], but otherwise the authors are aware of no published applications.

Spatial prompted recall methods would involve use of a GIS to generate a map showing a person’s trip routes, activity stops/location, and an array of text boxes or other map attributes depicting such items as mode, speed, location name, start/end times, trip and activity sequence (in the day). All this overlaid on the road and land-use network for context. Again, the task of the user would be to correct, confirm, and/or add supplemental information either directly on the map or via some alternative input means. This could be done using a paper-and-pencil map, or interactively on a computer screen. An early unpublished example is by Marca [35] who used a simple map generator coupled with an HTML form for input of attributes such as destination name and involved persons for each trip segment displayed in the map. A more recent and thoroughly tested application is reported by Stopher and Collins [34] who used both maps and a tabular (temporal) presentation of vehicle-based GPS detected trips in order prompt for passengers, trip purposes, and travel costs. Partially automated routines were initially used to detect trips, followed by manual preparation of tables and maps mailed to subjects some days after the GPS data are collected. Tabular displays listed trips by end time, travel time, distance, location of stop, and a reference to a map depicting the trip. The GIS-produced maps displayed color-coded trips by route, direction of travel, and stop location, overlaid on the transport and land-use network. An eight-page questionnaire was then used to systematically prompt for stop purposes and travel modes for each trip detected, as well as queries for undetected stops and/or incorrectly identified stops. As an improvement over the paper-and-pencil version, an innovative Internet-based survey was also pilot-tested. This provided a map-based animation along each GPS trip detected, stopping at the end to prompt for supplemental information via text boxes. Users could also stop the animation to add undetected stops. Whilst this approach avoided the need for detailed paper-and-pencil instructions, it was pilot tested with a very small sample. Further data collection is currently underway led by Stopher, including person-based tracking projects.
Prompted Recall Challenges and Opportunities

At this very early stage, there are many challenges and opportunities for developing prompted recall surveys, and it is clear that much testing is needed before they can be applied on a wide scale as a replacement to traditional diary surveys. In particular, the choice of sequential, temporal and/or spatial interface for prompting requires more testing. Presentation of paper-and-pencil maps, and a questionnaire for prompting has the advantage of familiarity, but requires a substantive delay in acquiring responses (which may affect recall), and a fairly lengthy questionnaire requiring a tedium of instructions. Animated maps offer the advantage of seamless organization of prompts (to the user), more control over how much information is displayed at one time, and the potential for quicker response times. Both however, assume the user is capable of interpreting maps. It is also unclear how well users will be able to interact with maps to make corrections and/or fill in gaps that will inevitably occur. Close attention to user-friendliness will be needed.

By contrast, temporal displays are likely more familiar to people, resembling calendars that many people use on a daily basis to organize and schedule their day. Maps on the other hand are occasionally used for navigation purposes, and rarely for day-to-day behaviour of people in familiar environments. Temporal displays are also much more compact in presentation when presented sequentially, whereas a map may cover large areas, and traces frequently double-back on each other making for some confusion. Spatial data can either be listed in the table as another column (by address or intersection), or a link provided to a map for reference. A means to provide temporal information in a map is less obvious, and certainly less familiar. Tabular displays also present a logical way to display missing stops/activities through blank table entries, whereas gaps in spatial data are much less obvious when viewed on a map, especially if there are many, and continuity between trips and stops is lost.

The means by which activities and trip patterns are derived from raw GPS traces is also an ongoing development task. Currently, mostly manual methods are used wherein a human operator must reduce a large set of GPS points to a sequential set of trips and stops (activities), then proceed to create temporal or spatial displays. If an algorithm is used, it is designed to provide only an initial set of detected attributes that are then further processed manually.

Person-based tracking offers the potential to track more than just auto trips, and provides for a much more thorough accounting of behaviour. It also introduces the potential for many more instances of signal outage and complex patterns, than do vehicular traces. Whereas vehicles have a consistent high power source, fixed antennae with clear view of the sky, and tend to stick open areas and known routes (streets), people will invariable carry a GPS device on a variety of location on their body, travel along unusual paths, enter buildings where the signal is lost, and have to re-charge batteries if the device is to remain small. The automatic detection of person-based patterns, and how these are handled in a PR survey, are relatively untested areas of development.
AN AUTOMATED ON-LINE PROMPTED RECALL SYSTEM

Given what has been learned so far, and the unique challenges and opportunities identified above, a GPS-based prompted recall system was developed that expands on past techniques with the following unique features:

- Completely Internet-based and computerized
- Designed explicitly for person-based GPS tracking
- Incorporates live, wireless transmission of the GPS data to a central server
- Utilizes a fully automated activity and trip detection algorithm (activity/trip start times, geocoded locations, location type derived from land-use, travel modes) run on a daily basis (ready for prompting the next day)
- Has a compact temporal interface presented in a “diary” style format, but includes interactive map for viewing detected locations (or for specifying new ones)
- Utilizes an intuitive means for confirming and correcting detected attributes, specifying supplement attributes, and filling gaps, in order to arrive at a complete schedule.
- Is easily extendable in terms of supplemental attributes that can be prompted for (pilot test supplemental prompts included activity types, involved persons by name)

Person-based Passive GPS Tracking Equipment

The equipment package for this study consisted of four main components assembled in a small shoulder pack worn by subjects, including:

1. **Cellular phone** (Bluetooth-wireless enabled, Symbian OS, 208 x 320 display, 48 MB storage capacity, sized 115 x 57 x 24 mm, weighing 150 grams)
2. **Lightweight GPS receiver** (Bluetooth-wireless enabled, 12 channel receiver, built-in antennae, sized 77.9 x 56.9 x 22.8 mm, weighing 70 grams, accuracy within 2.2m WAAS enabled territories)
3. **External booster antenna** (water resistant, sized 30 x 25 x 8 mm, cable 196cm)
4. **Supplemental Battery** (custom made, 2 USB power outlets, rechargeable 6 cell battery sized 38 x 70 x 56 mm, battery sled sized 20 x 87 x 44 mm, total weight 329 grams)

The shoulder pack was assembled such that the small booster antenna is located on the shoulder strap so as to have a clearer view of the sky when worn. The battery and GPS receiver are placed within the pack, whereas the cell phone is placed in its own pocket on the shoulder pack sleeve for ease of use. For this study, users did not interact with the phone, although it is possible to complete the prompted recall diary using an Internet browser on the phone.

The GPS receiver was setup to wirelessly broadcast various GPS sentences in the NMEA (National Marine Electronics Association) 0183 standard, at one second intervals. Among other fields, this sentence includes the date, time, latitude, longitude, speed, and satellite strength. A custom-made Java program was designed for the cell phone to receive, compress, and then send this data wirelessly to a central server for storage and further processing. Data transmission from the cell phone was via GPRS (General Packet Radio Service). A full day of one-second interval compressed GPS data ranged in size from 1.2 to 1.7 MB. Unlimited Internet service for each phone was obtained at a cost of $65 Canadian dollars per month.
Prompted Recall Diary Design

The first step in the prompted recall diary is to process the GPS data (see Figure 1 or Figure 2 for example GPS data) with an automated server-based algorithm designed to detect the following activity-travel pattern attributes:

- Event type (activity, trip, unknown gap)
- Event start/end times
- Trip modes (auto, walk, bike, bus)
- Activity location geo-code (longitude and latitude)
- Land-use classifications in vicinity of activity location, if available (Residential, Commercial, etc.)

The algorithm is based on a systematic set of rules that operates with no human intervention. At first, the algorithm took an average of 214 seconds to process a single day, but a fundamental shift in the way spatial data was processed resulted in a reduction to 30 seconds. Details on the algorithm are currently the topic of a separate publication; the objective of this paper is the design of prompted recall diary.

The next step in the survey process is to display the automatically detected attributes in a tabular diary-style display, as shown in Figure 3a or Figure 4a. Example data is displayed as it would appear prior to any user inputs (see next section for discussion on the example data). The date and ability to scroll to other days is provided at the top of the diary. Events are listed horizontally in sequence by time, event type, location, and “other” attributes (e.g. involved persons). Color is used to provide a natural means to convey which event attributes require attention, wherein red indicates wholly unknown attributes, yellow indicates attributes that require confirmation or updating with more detail, and green indicates attributes that are confirmed and completed. Additionally, white gaps indicate periods that were undetermined. Users are instructed as follows:

1. Click on any yellow attributes to modify if incorrect, or update with more detail (in the case of event types and locations)
2. Specify any missing red items (such as involved persons)
3. Fill in gaps with one or more new events
4. Delete wholly incorrect events, but only after trying to modify first.

They are instructed to do above until everything has turned green and there are no gaps in their day. Example completed diaries are shown on the right-hand side of Figure 3b and Figure 4b.

Specifying event attributes (time, event type, location, involved persons) involves clicking the attribute on the main diary screen, bringing the user to a new page for each attribute, as described in Figure 5 (time, event type, involved persons) and Figure 6 (location). When finished specifying the attribute, users click “Save” and are returned to the main screen. This approach was essential to providing enough space to specify the attribute as many attributes involve pull-down lists or radio buttons to speed entry, options to specify “new” items, and even an
interactive map in the case of locations. All this cannot be accommodated on the main screen, nor is this desirable from an aesthetic perspective. The design is such that only an amended description of key attributes are displayed on the main screen sufficient for the event to be easily distinguished by the user – namely the start/end time, event type and location name. This retains the compactness of the display. Involved persons are not even shown, and are instead represented by a single icon in the “other” column (the happy face) that serves as a reminder to click on and specify them. Any additional attributes, such as travel costs or when the event was planned, would be similarly displayed with an icon in the “other” column on the main screen.

Overall, the PR system is designed to automatically detect seven attributes (activity events, trip events, event start, event end, trip mode, activity location geo-codes, and location land-use classification) and prompt for four additional attributes not currently detected by the algorithm (activity specific type, activity-involved persons, trip passengers, location-unique labels).

The approach of using a tabular-diary combined with specific attribute pages, allows for a highly compact display in “calendar/schedule” style that is likely familiar to many people. The computer seamlessly manages the flow of queries without need for complex redirects commonly needed in paper-and-pencil surveys. Note also that Help pages are provided throughout, including on the main screen and for each attribute screen. This allows instructions to be provided specifically where and when needed, rather than having to provide a comprehensive manual that can be cumbersome if not overwhelming for users in paper-and-pencil surveys. An Options menu is also available where users can edit their lists of locations, people and events. This is more applicable to longer-term studies, when such lists may become long and need occasional paring.

Although the main interface is tabular, users do have the option to view (but not interact with) their daily diary in spatial form, as shown in Figure 7 and Figure 8. Activities are depicted as red circles, and trips as straight red lines with chevrons indicating direction of travel. The straight-line approach to trips was adopted not only to simplify the map, but to avoid frequent cases where trip routes disappear due signal loss which can be handled by the algorithm logic, but could make map interpretation difficult for the user. Figure 7 depicts the option to display labels at activity locations indicating their start/end time and the activity type. Displaying labels in Figure 8 resulted in a cluttered display with labels that overlapped and obscured other information – one indicator of the difficulty in designing a spatial PR survey.

**Illustrative Examples**

Two example sets of data are shown here to illustrate the functioning and design of the PR system. The first is a simple home-work-home pattern of a full-time employed person, whereas the second is a more complex multi-tour pattern of a full-time student. These subjects were instructed to carry the GPS tracking pack everywhere they went, but otherwise did not interact with it.

---

1 A subsequent phase of research will report on a larger pilot test of the system currently underway in Waterloo, Ontario at the time of writing.
The raw person-based GPS traces are provided in Figure 1 and Figure 2. In the figures, note the very obvious travel routes, as well as clusters of points were activities are conducted. Note also that the signal is occasionally lost along route (far left of Figure 1 when user entered a tunnel) and at destinations (although this is difficult to see at the resolution in the figures). Note also that, by nature, the travel routes do not always stick to roadways, as when the subject in Figure 1 walked through a park.

The results of the GPS algorithm prior to any user inputs is shown in tabular form in Figure 3a and Figure 4a. In the simple home-work-home example, there algorithm had trouble dealing with the multi-mode trip to work in the morning (walk-bus-walk). The algorithm correctly detected a walk trip, but interpreted the wait time at the bus stop as a “Residential” activity, followed by another walk trip. In addition, likely due to signal loss, the algorithm did not completely detect all the hours spent in the home at the beginning and of the day, resulting in gaps. As a result, of the 41 attributes of the 3 activities and 4 trip segments, 26 were correctly detected automatically. Users had to make 5 changes to detected attributes, and 10 updates of new attributes, to arrive at their final diary. If one ignores the attributes that the algorithm does not currently even attempt to detect (i.e. activity specific type, activity involve persons, trip passengers, location unique labels), then 26 of 31 attributes were correctly detected.

The more complex multi-tour example included 8 activities and 7 auto trips, in two tours. All activities and trips were correctly identified, but the mode was incorrect in two cases, and the start time of the first event was incorrect (and thus a gap was evident before it). There were no incorrectly identified events that the user did not do. Four of the land-use classifications of locations also did not match the eventual imputed location/event type, although users do not update this information (it is used only as an initial indicator of the location – users supply a unique label of their own to replace it). Overall, of the 95 attributes, 61 were correctly detected automatically. Users had to make 7 changes to detected attributes, and 23 updates of new attributes, to arrive at their final diary. Again, ignoring the attributes that the algorithm does not currently even attempt to detect, then 61 of 68 attributes were correctly detected.

Taken together, the person-based GPS tracking and automated algorithm appear to have correctly detected about two-thirds of the attributes that would be included in a detailed activity diary. Of the attributes that it was capable of detecting, it got almost 90% correct. Granted, this is based on only two examples, but it does provide considerable promise. A more in-depth analysis of the algorithm and its performance based on a larger pilot study is expected in a subsequent publication.

DISCUSSION AND FUTURE DIRECTIONS

The use of GPS to detect certain aspects of human activity and travel patterns is well established. This is especially so for start and end times of activities and trips, trip routes, and activity locations, which are difficult for people to recall accurately in traditional diary surveys. When GPS data is logged regularly (every $x$ seconds or minutes) without too much signal outage, it can be displayed on a map where the bulk of a person’s movements (trips) and stops (activities) become readily apparent to the naked eye, especially if road network and land-use information is included on the map. Some attempts have been made to develop automated routines for such detection, although most involve by-hand analysis to complete the picture. Even then, not all
aspects of human activity are traceable via GPS, such as activity types and involved persons, and there is growing demand for other activity attributes related to underlying decisions processes. It is inevitable that some form of interaction with subjects is needed to obtain the complete picture, and gain a better understanding of behaviour.

Conducting a “Prompted Recall” (PR) survey is one approach to filling this gap. Three possible approaches were discussed in this paper, including sequential, temporal/tabular, and spatial, each with their own merits. The detailed design of a computerized temporal, diary-like PR survey, is presented in this paper. It is unique in that it is administered via the Internet, designed explicitly for person-based GPS tracking, incorporates wireless transmission of the GPS data to a central server, utilizes a fully automated activity and trip detection algorithm, has a compact “diary” style format (but includes interactive map for viewing detected locations), and utilizes an intuitive means for correcting and updating detected attributes to arrive at a complete schedule. In all, it is designed to automatically detect seven attributes (activity events, trip events, event start time, event end time, trip mode, activity location geo-codes, and location land-use classification) and prompt for four additional attributes not currently detected by the algorithm (activity specific type, activity-involved persons, trip passengers, location-unique labels). The computerized design was essential to maintaining a compact display, managing the capture of a high level of detail on event attributes, and in providing specific help files where needed.

Example results under real-world conditions are presented that demonstrate the potential of this system for collecting highly detailed and accurate information on human activity and travel. In the examples (one simple home-work-home, one complex multi-tour day), about two-thirds of those attributes normally included in a detailed activity diary were correctly detected automatically by the system. Excluding those attributes that cannot currently be automatically detected (activity type labels, location labels, involved persons), the system was capable of detecting close to 90% of attributes correctly without user intervention. Granted, this is based on only two examples, but it does prove considerably promising with respect to greatly improving the quality of the diary data, and substantially reducing respondent burden.

The temporal interface design proved particularly easy to amend in order to handle simple and complex human activity-travel patterns, and to organize prompts for detailed attribute acquisition, including activity types at two levels (group and specific type), locations via an interactive map, etc. The same level of success in a spatial PR survey would take considerable development. Organizing the same information on a map can become difficult and confusing very quickly, especially if re-trace of routes is evident or if multi-stop tours are evident in a compact area. Additionally, it is important to display at least some event attributes on screen such as start/end times or event types (for instance, see the labels in Figure 7) so that users can easily distinguish one event from another. However, as was experienced with Figure 8, this can become very difficult, as labels overlap for locations close to each other, or for locations visited more than once in a day (like home, or children’s schools). A tabular display has the capacity for displaying complex multi-stop tours, since events can scroll horizontally indefinitely, and a natural sorting order (time) is available to keep things in perspective. Maps do not have an equivalent obvious means to sort out what is displayed on screen and aid in interpretation, although the animated map approach by Stopher and Collins [34] appears promising. But even then, it is difficult to assume that people will have the requisite map interpretation skills to
conducted such a task, whereas most people are quite familiar with the more calendar-like display of the current design.

There are several key avenues for future development and application at this early stage in development of GPS prompted recall surveys. Improvements in the frequency and accuracy of attribute detection would have obvious design benefits, and should be an on-going task. For instance, over a multi-day period, the algorithm could incorporate “learning” rules. In the event that entire households are GPS tracked, perhaps even involved persons could be detected. Conducting a larger pilot study involving a questionnaire on users’ reactions and experience with the diary would also be very valuable with respect to design improvements. A larger sample will also allow more definitive statements on accuracy gains and respondent burden savings. The growing availability of GPS-enabled cellular phones, including those with Internet browser capabilities, could substantially reduce costs and make for wider application potential.

There is also much potential to expand the system. Simple expansion could involve addition of “other” attributes to the tabular design, such as when the event was planned, whether it was modified before execution, who they communicated with regarding the decisions, or how flexible it is in time and space. A more complex expansion could involve deployment of the on-line diary in advance to capture the pre-planned event attributes of subjects. These could then be compared to GPS-detected event attributes in order to explore underlying decision processes in more depth.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the generous financial support received for this project from GEOID (Geomatics for Informed Decisions) Network of Centres of Excellence Program of the Canadian federal research councils, and the Social Sciences and Humanities Research Council of Canada MCRI program. Thanks go to Luke Cwik, Mikhail Markin and Nick Gommans for their creative programming work on this project. Thanks also go to the diligent field work of Erika Nemeth, and all those who generously supplied their time in completing the survey.

REFERENCES


List of Figures

Figure 1  Person-based Raw GPS traces: Simple home-work-home example ......................... 16
Figure 2  Person-based raw GPS traces: Complex multi-tour example ............................... 17
Figure 3  On-line prompted recall diary main interface: Simple home-work-home example ...... 18
Figure 4  On-line prompted recall diary main interface: Complex multi-tour example ............. 19
Figure 5  Pages for specifying event attributes in the diary – time, event type, and involved persons ......................................................................................................................... 20
Figure 6  Specifying locations in the diary ............................................................................. 21
Figure 7  Spatial summary of GPS detected activities available to user: Simple home-work-home example ................................................................................................................................. 22
Figure 8  Spatial summary of GPS detected activities available to user: Complex multi-tour example ................................................................................................................................. 23
Figure 1  Person-based Raw GPS traces: Simple home-work-home example

Note: some points have been edited to protect subject identity.
Figure 2  Person-based raw GPS traces: Complex multi-tour example

Note: some points have been edited to protect subject identity.
Figure 3  On-line prompted recall diary main interface: Simple home-work-home example

a) Initial results of event detection algorithm, showing color-coded event attributes (see text for explanation)

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
<th>Location</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>7:00 AM</td>
<td>Activity</td>
<td>Residential</td>
<td>✗</td>
</tr>
<tr>
<td>7:07 AM</td>
<td>Trip: Walk</td>
<td>N/A</td>
<td>✗</td>
</tr>
<tr>
<td>8:00 AM</td>
<td>Activity</td>
<td>Residential</td>
<td>✗</td>
</tr>
<tr>
<td>8:05 AM</td>
<td>Trip: Walk</td>
<td>N/A</td>
<td>✗</td>
</tr>
<tr>
<td>8:10 AM</td>
<td>Activity</td>
<td>UNIVERSITY</td>
<td>✗</td>
</tr>
</tbody>
</table>

b) Completed diary after filling gaps, confirming all attributes, updating location/activity labels, and adding involved persons/passengers.

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
<th>Location</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>7:00 AM</td>
<td>Activity: AT HOME</td>
<td>Home</td>
<td>✗</td>
</tr>
<tr>
<td>7:07 AM</td>
<td>Trip: Walk</td>
<td>N/A</td>
<td>✗</td>
</tr>
<tr>
<td>8:00 AM</td>
<td>Trip: Bus</td>
<td>N/A</td>
<td>✗</td>
</tr>
<tr>
<td>8:03 AM</td>
<td>Trip: Walk</td>
<td>N/A</td>
<td>✗</td>
</tr>
<tr>
<td>8:05 AM</td>
<td>Trip: Walk</td>
<td>N/A</td>
<td>✗</td>
</tr>
<tr>
<td>8:10 AM</td>
<td>Activity: Work/School: Work</td>
<td>The University</td>
<td>✗</td>
</tr>
<tr>
<td>8:15 AM</td>
<td>Trip: Walk</td>
<td>N/A</td>
<td>✗</td>
</tr>
<tr>
<td>9:00 AM</td>
<td>Activity: AT HOME</td>
<td>Home</td>
<td>✗</td>
</tr>
</tbody>
</table>
Figure 4  On-line prompted recall diary main interface: Complex multi-tour example

**(a)** Initial results of event detection algorithm, showing color-coded event attributes (see text for explanation)

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
<th>Location</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>8:10 AM</td>
<td>Activity</td>
<td>Residential</td>
<td>X</td>
</tr>
<tr>
<td>8:20 AM</td>
<td>Trip: Bus</td>
<td>N/A</td>
<td>X</td>
</tr>
<tr>
<td>8:44 AM</td>
<td>Activity</td>
<td>UNIVERSITY</td>
<td>X</td>
</tr>
<tr>
<td>9:05 AM</td>
<td>Trip: Auto</td>
<td>N/A</td>
<td>X</td>
</tr>
<tr>
<td>9:15 AM</td>
<td>Activity</td>
<td>Commercial</td>
<td>X</td>
</tr>
<tr>
<td>9:33 AM</td>
<td>Trip: Auto</td>
<td>N/A</td>
<td>X</td>
</tr>
<tr>
<td>9:55 AM</td>
<td>Activity</td>
<td>Resource and Industrial</td>
<td>X</td>
</tr>
<tr>
<td>10:14 AM</td>
<td>Trip: Bike</td>
<td>N/A</td>
<td>X</td>
</tr>
<tr>
<td>11:00 AM</td>
<td>Activity</td>
<td>Residential</td>
<td>X</td>
</tr>
<tr>
<td>5:40 PM</td>
<td>Trip: Auto</td>
<td>N/A</td>
<td>X</td>
</tr>
<tr>
<td>6:00 PM</td>
<td>Activity</td>
<td>Resource and Industrial</td>
<td>X</td>
</tr>
<tr>
<td>6:14 PM</td>
<td>Trip: Auto</td>
<td>N/A</td>
<td>X</td>
</tr>
<tr>
<td>6:27 PM</td>
<td>Activity</td>
<td>OTHER</td>
<td>X</td>
</tr>
<tr>
<td>6:57 PM</td>
<td>Trip: Auto</td>
<td>N/A</td>
<td>X</td>
</tr>
<tr>
<td>7:13 PM</td>
<td>Activity</td>
<td>Residential</td>
<td>X</td>
</tr>
<tr>
<td>11:59 PM</td>
<td>Activity</td>
<td>Residential</td>
<td>X</td>
</tr>
</tbody>
</table>

**(b)** Completed diary after filling gaps, confirming all attributes, updating location/activity labels, and adding involved persons/passengers.

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
<th>Location</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>8:10 AM</td>
<td>Activity: AT HOME</td>
<td>My House</td>
<td>X</td>
</tr>
<tr>
<td>8:20 AM</td>
<td>Trip: Auto</td>
<td>N/A</td>
<td>X</td>
</tr>
<tr>
<td>8:44 AM</td>
<td>Activity: Work/School: Work</td>
<td>U. of Waterloo</td>
<td>X</td>
</tr>
<tr>
<td>9:15 AM</td>
<td>Trip: Auto</td>
<td>N/A</td>
<td>X</td>
</tr>
<tr>
<td>9:33 AM</td>
<td>Activity: Shopping: Household</td>
<td>Conestoga Mall</td>
<td>X</td>
</tr>
<tr>
<td>9:55 AM</td>
<td>Trip: Auto</td>
<td>N/A</td>
<td>X</td>
</tr>
<tr>
<td>10:14 AM</td>
<td>Activity: Services: Post office</td>
<td>Post Office</td>
<td>X</td>
</tr>
<tr>
<td>11:00 AM</td>
<td>Trip: Auto</td>
<td>N/A</td>
<td>X</td>
</tr>
<tr>
<td>11:57 PM</td>
<td>Activity: Work/School: Work</td>
<td>My House</td>
<td>X</td>
</tr>
<tr>
<td>5:48 PM</td>
<td>Trip: Auto</td>
<td>N/A</td>
<td>X</td>
</tr>
<tr>
<td>6:00 PM</td>
<td>Activity: Cat sitting</td>
<td>Cat House</td>
<td>X</td>
</tr>
<tr>
<td>6:24 PM</td>
<td>Trip: Auto</td>
<td>N/A</td>
<td>X</td>
</tr>
<tr>
<td>6:27 PM</td>
<td>Activity: Shopping: Other</td>
<td>Best Buy</td>
<td>X</td>
</tr>
<tr>
<td>6:57 PM</td>
<td>Trip: Auto</td>
<td>N/A</td>
<td>X</td>
</tr>
<tr>
<td>7:13 PM</td>
<td>Activity: AT HOME</td>
<td>My House</td>
<td>X</td>
</tr>
</tbody>
</table>
Figure 5  Pages for specifying event attributes in the diary – time, event type, and involved persons

a) Start and end time

Diary: Edit Event Time  [ Help ]

Start

4 PM

End

5:45 PM

Save  Cancel

b) Event type – First level is either “Activity” or “Trip”, and subsequent levels obtain more detail.

Diary: Choose Specific Type of Event  [ Help ]

General

Activity

Group

Shopping

Specific [ New ]

Groceries

Convenience

Household

Personal

Other

Save  Cancel

c) Involved Persons/passengers

Diary: Involved Persons or Passengers  [ Help ]

Involved People [ New ]

Mary

John C.

Dad

Save
Figure 6 Specifying locations in the diary

a) User first given 3 key options – Choose a known location from the pull-down, [Edit] the automatically detected location (bringing user to b), or specify a [New] location from scratch (bringing user to b).

b) Interactive Map for location selection. Map zooms to detected geo-coordinates and user provides a unique name. If specifying from scratch, user navigates to the location (× marks the spot).
Figure 7 Spatial summary of GPS detected activities available to user: Simple home-work-home example

NOTES:
Locations shown with graduated symbols representing relative amount of time spent at the location.
Figure 8 Spatial summary of GPS detected activities available to user: Complex multi-tour example

Diary Analysis: Maps

NOTES:
Locations shown with graduated symbols representing relative amount of time spent at the location. Labels note shown.