A Data Collection Framework for Exploring the Dynamic Adaptation of Activity-Travel Decisions

Jessica Y. Guo  
Department of Civil and Environmental Engineering  
University of Wisconsin – Madison  
jyguo@wisc.edu

Sunil Nandam  
Department of Civil and Environmental Engineering  
University of Wisconsin – Madison  
snandam@wisc.edu

Teresa Adams  
Department of Civil and Environmental Engineering  
University of Wisconsin – Madison  
adams@engr.wisc.edu
1 Introduction

The activity-based approach to travel demand forecasting entails modeling individuals’ travel decisions within the context of their activity episodes, with explicit consideration of decision makers’ socio-demographic characteristics, personal tastes, and constraints. Most of the presently operational activity-based travel demand models (ABTDM) are based on the assumption that individuals are rational decision-makers under a utility maximization principle. These modeling systems typically consist of a suite of econometric models representing the outcome of individuals’ decision-making of which activities to pursue, where, at what time, for how long, with whom, in what sequence, while traveling by what mode and what route. The econometric approach has been criticized for not explicitly modeling the decision process underlying activity engagement and travel, limiting the behavioral richness of the model. In particular, it has been argued that the way people adapt their activity-travel decisions in response to policy variables is rooted in the process of rescheduling their lives across time and space, and in relation to other household members and social contacts (Doherty and Miller, 2000). While some activities are planned far in advance and form the initial schedule structure, other scheduling decisions are made closer to execution of the activity or even impulsively. Rescheduling may be triggered by factors such as time pressure, schedule conflicts, congestion charging, real-time travel information, weather, and event cancellation. Most utility-based models are inadequate in explicitly representing the behavior dynamics in response to these factors, rendering such models better suited for long-range planning applications but not for short-range planning.

The other class of ABTDM can be characterized as computational process models (CPM), which focus on the process of decision-making. The CPM approach offers more flexibility than econometric models in representing the complexity of travel decision making. In particular, the CPM approach has been used to explicitly model within-day rescheduling behavior in an event-based simulation framework. Most of the CPM assume that activity agenda of an individual is defined on a daily basis and consists of the planned activities and durations of these activities. Given a deterministic activity agenda, the sequence, locations and start times of activities and transport modes used for traveling are determined in a subsequent scheduling phase. Subsistence activities such as work and school are assumed fixed; maintenance, social, and leisure activities are considered as more flexible. Individuals determine on a daily basis whether or not to participate in these flexible activities and, if they do, these flexible activities are scheduled around the fixed activities. On the day when the planned/scheduled activities are to be executed, details of the activities determined during the scheduling phase may be altered in response to personal and external factors. The adjustment and adaption of activity attributes are said to occur during the rescheduling phase.

This planning-scheduling-rescheduling framework that provides the foundation for most CPM of 1-day activity-travel pattern is limited in at least the following three aspects:

- Dichotomous classification of activities: Activities are classified based on the dichotomy between fixed and flexible activities. This strict dichotomy, although useful for algorithmic development, is artificial (Kuijpers et al., 2010). In reality, activities may have varying degrees of fixity versus flexibility. Recent research showed that the degree of fixity depends
on activity purpose, time-of-day, persons involved, location, use of internet, and gender (Schwanen et al., 2008).

- Dichotomous classification of decision-making process: The decision-making process is artificially divided into the pre-day planning and scheduling stage versus the on-the-day rescheduling stage. In reality, the planning-scheduling-rescheduling of activities occurs on a time continuum and may overlap (that is, a pre-planned activity may be rescheduled while a new activity is being planned).

- Assumption of plan completion and certainty: Attributes of activities are assumed to be determined with certainty during the planning and scheduling phases. In reality, activities are often planned only partially, leaving certain attributes vague or undetermined until much closer to activity execution.

In order to further advance CPM development in the abovementioned three areas, the present study sets out to design and pilot test a methodology for collection planning-scheduling-rescheduling behavioral data. The methodology is designed with the goal of reducing the resources typically needed to collect such type of behavioral data, thereby allowing the systematic collection of data from a sample size large enough to support statistical modeling. This goal is accomplished through the combined use of person-based passive location tracking and an internet-based survey instrument that collect both revealed and stated adaptation data.

The survey instrument has been implemented and the methodology is currently being pilot-tested to collect data from a sample of 50 individuals. This paper describes the details of the data collection methodology. Lessons learnt from the data collection process and findings from the empirical analysis of the data will be added to the presentation at the ITM2012 Conference if this paper is selected.

2  Past Data Collection Efforts

One of the earliest studies on how people make activity-travel decisions is done by Hayes-Roth and Hayes Roth (1979) where subjects were asked to “think-aloud” how they go about planning their day given a set of activities to pursue in an imaginary city. Ettema et al. (1993) developed MAGIC, an interactive computer experiment to identify the types of steps people used to construct a one-day schedule. As observed by Clark and Doherty (2010), the existing methodologies used to observe the rescheduling decision process typically focus on either revealed adaptations or stated preference.

Revealed adaptation surveys involve observation of actual decisions made in context of actual daily life. For example, the computerized household activity survey elicitor (CHASE) records household activity scheduling decisions over a period of 7 days. It is designed to track the sequence of steps taken by individuals to form weekly activity schedules using a calendar format. Other similar survey efforts include OPFAST, EXACT, and REACT. All these surveys include observation of actual rescheduling decisions that occurred in reaction to everyday scenarios of change. They typically involve a preliminary survey of preplanned activities/trips and their attributes (start time, end time, duration, location, activity type, and involved persons) for an upcoming period. On subsequent days, subjects are asked to update their schedule by adding new activities and modifying/deleting others. Subjects are often prompted for information on why and when rescheduling decisions occurred. In CHASE, this was done interactively.
during the week via computer generated prompts, whereas in OPFAST, an in-depth follow-up interview was used to probe for further details. The advantage of the revealed adaption surveys is that the observed rescheduling behavior are inherently realistic; the disadvantage is that the rescheduling events may be limited in number and scope since they arise naturally throughout the day.

An example of the stated preference approach was Nijland et al (2009), who developed a web-based questionnaire to assess how subjects respond to a reduction in available time for a planned activity in a number of hypothetical situations. A major shortcoming of this approach is that the subjects are unfamiliar with the hypothetical setting and may wrongly estimate the time required for the travel/activity. Roorda and Andre (2008) attempted to alleviate this concern by citing a reported activity episode and asking the subjects what they would do in the hypothetical scenario of an unexpected 1-h delay at the activity.

Availability of GPS devices led to the development of surveys that collect the spatial-temporal information of executed activities, thereby reducing the burden on the subjects. EXACT (Rindsfüser et al., 2003) uses a hand-held PDA that is used for both collecting the spatial data and reporting the scheduling decisions. Most recently, Clark and Doherty (2010) developed a “passive” data extraction system using a GPS receiver linked wirelessly to a BlackBerry which sends spatial-temporal data to a remote server for storage and processing. This scheduling data collection framework has six phases: (1) preplan interview; (2) coding of the preplan schedule; (3) passive GPS tracking; (4) internet-based prompted recall diary; (5) comparison of planned versus executed schedules; (6) In-depth rescheduling interview. The survey method was reported to result in a substantial decrease in respondent burden and an increase in rescheduling events captured. However, this approach is resource intensive and therefore not feasible for collecting a moderately large sample size.

3 Proposed Data Collection Methodology

Drawing from lessons learnt from past data collection efforts, the present study developed a multi-media data collection methodology, featuring the use of person-based passive location tracking and an internet-based survey instrument that collect both revealed and stated adaptation data. The data collection process consists of four stages: (1) recruitment, (2) activity-travel plan retrieval, (3) GPS tracking and internet-based prompted diary recall, and (4) rescheduling questionnaire and stated adaption experiment. As illustrated in Figure 1 and described in more details below, the internet-based instrument is designed to allow the survey participants to self-administer the multi-phase survey. The survey website is implemented using PHP and JavaScript.
Figure 1. The main page of the internet-based survey instrument allows study participants to self-administer the data collection process.
3.1 Recruitment and Registration

The present pilot study focuses on individuals 18 years or older. Participants are recruited through campus-wide emails to the faculty and staff at the University of Wisconsin – Madison. This is supplemented with putting fliers up around campus and advertising through word-of-mouth. The recruitment print materials and the consent forms indicate that GPS wear, computer/internet access, and up to 1 hour of their time will be required for participation. In return, each participant receives $15 cash upon when they return the GPS device after the successful completion of the survey. Interested individuals are asked to contact the research team and the first 50 qualified individuals are selected to participate. Upon returning the signed consent form for the study, the participants are visited by a member of the research team and given a 5 minute introduction to the GPS data logger and the online survey website. The participants are then asked to set up a user account on the survey website. As part of this registration process, they are asked to provide socio- demographic information such as household structure, employment status, age, and income. The user is then assigned a “survey day” for which their planning-scheduling-rescheduling data are collected.

3.2 Activity-Travel Plan Retrieval

On the day before the designated survey day, participants are reminded via email to log onto the survey website to provide information regarding any plans they may have for the next day. Differing from most past efforts of collecting planning and rescheduling data, our instrument recognizes that individuals do not necessarily plan all of the attributes of preplanned activities, nor do they plan the activities with the same degree of certainty. Respondents are allowed to state uncertainty by indicating, for example, activity start time as “5pm-7pm”, “in the late afternoon”, or “after dinner”. Planned location(s) for activity pursuit can be stated as a set of specific sites or a neighborhood such as “Downtown”.

3.3 Actual Activity-Travel Information Retrieval

On the day before the survey day, reminder emails are sent to the participants regarding the need to wear their GPS device during the course of the next day. The GlobalSat DG-100 GPS Data Logger is used in this study. It measures 3.3” x 0.5” x 1.7”, weighs 60g, and stores up to 22 hours of tracking data.

At the end of the survey day, the participant is asked to connect the GPS device to a computer and upload the track data to the study’s web server, where the GPS data is processed in real-time and probable stop locations where the participant might have undertaken activities are identified. These stop locations, together with the GPS track, are then shown to the participant in an interactive environment using Google Maps API V3 technology. The participant is promoted to go through each stop marked on the map, verify its location accuracy, and provide information about the activities undertaken at each location. The participant is also asked whether each of the executed activities corresponds to which of the pre-planned activities. Furthermore, the participant can add stop locations that might have been missed by the computational algorithm. Stops that might have been erroneously flagged can also be deleted. This map-based approach is believed to improve activity recall while reduce data-entry burden.
3.4 Rescheduling Questions

Once the executed activity-travel data is retrieved, an application on the server compares the reported activities against the activity plan to identify inconsistencies that represent possible rescheduling events. The application then presents these inconsistencies back to the participant as a form of memory jogger. The application also formulates questions – using artificial intelligence techniques – regarding the rescheduling events and the underlying decision-making process. Participants are asked of questions such as “What made you decide to change the location for meeting up with friends?” and “When did you become certain of the exact time to meet up with friends?” If an activity is identified as cancelled, the participant is asked of the open-ended question of “What was the most important factor that led to the cancellation of this activity?”.

Since the travel data is collected for only a day, it is likely that the revealed adaption survey questions would only pick up a limited number and scope of real-world rescheduling events. In order to improve our understanding of rescheduling behavior and to support the future development of empirical models, our survey instrument also includes a stated adaptation experiment component to increase the amount and the range of data collected. The experiment entails developing hypothetical scenarios based on the executed activity episodes as reported by the study participants. This approach builds on the work by Roorda and Andre (2008). However, our experiment goes beyond just considering a hypothetical 1 hour delay in getting to a reported activity. Rather, our server-end intelligent algorithm formulates, in real-time, “what-if” questions around “alternative realities” of the reported activities. Each participant is asked of a set number of randomly selected “what-if” questions. The following is a sample stated adaptation survey question:

You worked out at the university gym at 5:30pm. What would you have done if you found out that the university gym was closed after having arrived at the location?

a) Go to a different location for the same activity.

b) Move on the next activity that has been planned for the day.

c) Do something else that has not been planned.

d) Other – please explain:__________________________________________.

4 Research Contributions and Conclusions

The ability of existing CPM to properly represent travelers’ rescheduling dynamics has, in part, been limited by the lack of empirical data to provide sufficient insights into the planning-scheduling-rescheduling process. In the absence of empirical data and knowledge about the true decision making behavior, the activities are often ‘rescheduled’ to mechanistically resolve any temporal conflicts and/or inconsistencies introduced by the modeling system’s internal scheduling rules. Activity rescheduling becomes modelers’ way of artificially fitting the simulated patterns to the assumed constraints, as opposed to truly representing the underlying behavior.

This study develops a methodology to collect activity-travel planning-scheduling-rescheduling data in more depth and breadth as compared to past data collection efforts. In particular, the survey instrument explicitly probes the participants about the flexibility and uncertainty of
various aspects of their activity-travel plans. To the best of the authors’ knowledge, this is also the first instrument that is capable of generating stated adaption questions on-the-fly based on reported activities. The data collected is envisaged to better inform modelers of the dynamic nature of the decision mechanism in the generation of household members' activity agendas and in the within-day activity adjustment and rescheduling. Furthermore, the internet-based survey management system reduces the cost of data collection and, more importantly, offers the possibility of collection data from a much larger sample of individuals.

ACKNOWLEDGEMENT

This work is partially funded by the National Center for Freight and Infrastructure Research and Education (CFIRE) based at the University of Wisconsin. CFIRE is a national university transportation center supported by the US Department of Transportation.

REFERENCES


