

UC Santa Barbara

Specialist Research Meetings—Papers and Reports

Title

GPS Tracking and Time Geography: Applications for Activity Modeling and Microsimulation, Position Papers

Permalink

<https://escholarship.org/uc/item/5z40v7j3>

Authors

Goulias, Kostas
Janelle, Donald G.

Publication Date

2005-10-01

GPS Tracking and Time-Geography Applications for Activity Modeling and Microsimulation

Position Papers

FHWA-sponsored Peer Exchange and CSISS Specialist Meeting

10-11 October 2005
The Upham Hotel
Santa Barbara, California

This document contains the position papers prepared by participants in relationship to the goals for the peer exchange and specialist meeting.

The specific goals of this Peer Exchange were to:

- a. Discuss methods and techniques for using the GPS data to be applied to activity modeling and microsimulation.
- b. Increase practitioner interest in the potential of GPS data for activity models.
- c. Encourage academics and their graduate students to pursue research activities with these datasets.
- d. Develop priorities for research that could be conducted using the Commute Atlanta dataset after it has been anonymized for public release.
- e. Develop priorities for research that could be conducting using other GPS data collected from other household travel/activity surveys.

The position papers from participants were posted on the CSISS web site in advance of the meeting. The Peer Exchange took place over one and one-half days (starting on the morning of 10 October and ending with lunch on 11 October). Presenters and participants from state DOTs, MPOs, consultants, and academics gave short presentations and discussed the variety of issues that emerged directly for the goals above. The presentation outlines and related graphic displays are available in pdf format at <http://www.csiss.org/events/meetings/time-geography/participants.html>.

For access to the **final report** on the meeting and its outcomes, see:
<http://escholarship.org/uc/item/1x41q271>

GPS and Travel Behavior: A Few Research Questions

Terry C. Bills

Environmental Systems Research Institute (ESRI)

It has been a number of years since the possibility of utilizing GPS enabled data collection techniques for travel behavior has been recognized. Initially conceived of as a way to enrich the collection of traditional travel diaries required by travel demand models, a much larger agenda has emerged for the role of GPS data collection techniques in helping to analyze travel behaviors. There are however, a number of conceptual and methodological questions which remain before this promise can be realized. The following attempts to lay out a few of these issues.

An often noted problem with traditional diary-based travel surveys is the number of trips which go unreported, and the considerable error often found through post analysis. Passive GPS collection techniques could significantly reduce the number of unreported trips, and help unambiguously determine travel paths undertaken. Together with detailed GIS land use data (including residential and business location data), it should be possible to passively collect most information required for more detailed tour based models currently under development.

In most metropolitan areas, geo-coded business establishment data (commercial, retail, etc) exists which allows various trip ends to be imputed with a high degree of accuracy. The same is true for residential and other land uses. GIS land use data together with GPS collected individual travel path tracings should be able to provide most of the information required for quite detailed travel diaries (including non-motorized trips, provided the transmission device is on the individual, and not the vehicle). For the purposes here, I take this as a given.

In addition to the ability to enrich the collection of travel diaries, a number of analysts have pointed to GPS technology as a way of collecting more accurate speed and network data required by travel demand models. It is now common practice to use GPS equipped probe vehicles for speed data collection, as well as to capture more detailed roadway characteristic data. Such data can serve to better calibrate and validate existing urban models, which have often had to rely on quite sketchy speed data in their model parameters.

Less well known is the use of GPS collected speed (and other) data for performance assessment, and the establishment of highway based performance measures. Most major metropolitan areas include Highway (segment based) congestion improvement among their regional performance measures. GPS collected speed data provides a much more accurate measure of such travel improvement, over imputed speeds derived from traditional travel models. A number of MPOs around the country have begun to move in this direction for the collection of data associated with highway performance improvement. Again such use of GPS falls within well accepted current practice.

For GPS collected travel data to move beyond these current initiatives however, a number of conceptual and methodological issues have to be addressed. First and foremost concerns the large amount of data generated by GPS based tracing data. For anyone who has viewed airport based flight path tracings, the question of the appropriate unit of analysis (and for what purpose) is not trivial. The same is true with individual travel behavior generated from GPS based collection. Clearly some level of aggregation is appropriate from the nearly limitless number of individual collection points of data.

How, and into what units are the many points of data combined to become meaningful for analysis? While this answer will vary by research purpose, for the purposes here, let me suggest trip segments as the smallest basic unit of analysis. Collection points should be detailed enough to determine the exact route taken through the transportation network, and these points should then be aggregated into specific trip segments. Trip purpose and other associated information (including time stamping) can then be associated with that trip segment, and can become a single record. That record of course can become one segment in a larger trip chain, but does allow for analysis of individual trip segments and paths. Statistical algorithms should be able to determine trip segment ends based upon measures of non-movement, although such algorithms should be empirically validated.

Depending upon the research purpose, such travel segments could be aggregated into an individual's single day travel activity, into a household travel activity, or in some cases, the aggregate travel behavior of much larger "communities". Holding privacy issues in abeyance, it may well be possible to collect extremely large samples (from cell phone sources) for urban areas from passive GPS collection techniques. This raises questions concerning the appropriate sample size if the goal is a full activity based or microsimulation model.

It also raises questions about the appropriate level of analysis. Is it individual behavior we seek to understand and model? Household level? Or it is really the aggregated behaviors of individuals as they dynamically traverse congested transportation networks? Obviously the data can support all levels of analysis, but with the "density" of the data now theoretically possible, we can much more closely replicate actual collective travel behaviors. This may also require new data mining techniques to discover why particular paths were chosen, and how to explain collective outcomes. We have only begun to think about such data mining techniques, and the appropriate questions to ask. It may well be that what seemed so "far-reaching" about urban microsimulation models a number of years ago, may not have been far reaching enough.

But even in the shorter term, passively collected GPS travel behavior data holds promise to significantly advance our understanding of travel behavior, and to serve as empirical validation of current microsimulation models. While full microsimulation models have mostly been applied for traffic operations and corridor level analysis, there have been few ways to accurately calibrate and validate such models. Clearly rich GPS collection holds significant promise in this regard. Even beyond microsimulation models, dense GPS data may allow us to empirically test hypotheses advanced concerning land use and transportation interactions, an area woefully short of empirical testing.

Mr. Bills is the Transportation Industry Manager for ESRI. He has over twenty years experience in Transportation and GIS, and previously worked for a large Metropolitan Planning Organization, and served as President of GIS/Trans, a private company engaged in transportation related GIS. He holds two Masters Degrees from UCLA, and was also a Doctoral Candidate at UCLA.

GPS-Assisted Data Collection to Support Transportation Planning

Larry Blain, Puget Sound Regional Council

At one time, all a transportation planner needed to know was the number of vehicles which could be expected to use a given facility. With the availability of computers and the design needs of the Interstate Highway System, travel demand models were developed in the 1960s to generate, distribute, and assign auto trips based on the projected locations of population and employment. The primary data sources upon which these models were built were home travel surveys, roadside interview surveys, and traffic counts. Supply was the road system, and demand was the choices of destinations and routes.

“During the late 1960s and the 1970s there occurred in transportation systems analysis a shift from unimodal analysis of large capital decisions toward multimodal systems and considerations of pricing, operational policies, and the construction of new facilities.”
(Discrete Choice Analysis by Ben-Akiva and Lerman: MIT Press, 1985. Page 1)

To get observations of people’s travel behavior, the Puget Sound Regional Council (PSRC) has conducted several rounds of travel surveys, including the ten waves of its panel survey from 1989 to 2002. In 1999 the PSRC conducted a household activity survey which included a two-day travel diary for each trip maker. These diaries were used to record locations and activities throughout the day, as well as mode of travel, arrival and departure times, and family members traveling together. Departure times, routes, and stops were recorded by the survey participants.

People’s travel decisions are influenced by their knowledge of and perception of the transportation “environment” – the travel times, reliability, costs, characteristics, comfort, and safety of the various modes. The more we know about the environment in which people make their travel decisions, the better we can understand and model their decisions. In 2000 the PSRC used GPS equipped cars to travel the region’s roads and record data points from which travel speeds could be calculated. Interns in the cars also noted posted speed limits, numbers of lanes, etc.

At the same time that technology can collect and process greater quantities of data, people have become more reluctant to share personal data and/or be inconvenienced. A GPS device can be used to record passively the location by time of a person or a vehicle. The data can identify trips which were omitted from a diary, and the recorded trip ends may enable a researcher to infer the purpose of a trip. Some day the units also may be used to implement a road pricing scheme in which rates vary by facility.

With funding from the Federal Highway Administration, the PSRC currently is conducting a “traffic choices” demonstration project to gather data about people’s responses to paying for access to faster roads. For eleven months 260 households will have 420 GPS units permanently installed in their vehicles. During the first two months and the last six weeks, the units will simply record the movements of the vehicles, to

establish the baselines. During the middle 7 ½ months the drivers will “pay” between \$0.10 and \$0.50 per mile to use freeways (depending on the time of day), and half that amount to use major arterials. The project is using real money. Each household has an endowment account which will be spent down according to their choices. At the end of the project, each household gets to keep the money remaining in the account.

So for 11 months there will be 420 vehicles with GPS units traveling the streets of the region, recording their locations every 250 meters - sufficient accuracy to determine which facility they are on, and when. More than 300,000 vehicle trips will be made under conditions simulating tolling, and another 150,000 in “normal” conditions. Since many trips during those months will be repeat trips, the resulting database will be able to be analyzed to show variations in departure time and route choice relative to changes in congestion and – the primary subject of this project) – the imposition of pricing. In this project no effort will be made to identify the purpose of each trip or the number of people in the car.

The PSRC has made extensive use of both the panel survey and the 1999 household activity survey in recent updates to its travel demand model. The next major modeling initiative is anticipated to be the development of an activity-based model to simulate the day-to-day decisions of the households synthesized and located within the UrbanSim land-use model which is currently being implemented. A new large household survey being planned for some time in 2006 is an ideal opportunity for PSRC to acquire data using a combination of traditional and GPS recording methods.

POSITION PAPER

Michael F. Goodchild

University of California, Santa Barbara

Potential of dense-tracking data

There seems to be no doubt that densely sampled tracks have many powerful applications in transportation and social science. We know already that even with standard GPS (no differential correction) the average of a small number of tracks provides the cheapest and fastest method of maintaining currency in street centerline databases (compare the *Los Angeles Times* story on problems of map currency, 21 August 2005), and this potential is already being exploited by some vendors. We know that averaging tracks is capable of providing positional accuracies at the meter and even sub-meter levels, which would be sufficient to support lane-level modeling and guidance. We know that dense tracking provides the most rapid way of detecting and responding to accidents and other obstructions. We also know that tracks can be parsed to determine activities (particularly mode) with some accuracy, and that comparison of tracks can provide useful information on interactions. So what impediments make this potential problematic?

The IRB problem

The potential for tracking to compromise privacy is clearly a major issue, and institutional review boards are likely to make it very difficult to conduct research in this area without stringent safeguards. A recent article (VanWey *et al.*, 2005) argues that all current methods for protecting confidentiality in microdata, such as aggregation, positional distortion, and firewalls are flawed, and that major funding is needed to support the exploration and examination of possible alternatives. At the same time, there is a conspicuous difference between the stringency likely to be exhibited by IRBs, and the environment in which the private sector and many government agencies operate.

Analytic techniques

Dense tracks represent a novel type of data that is not compatible with any of the current analytic environments—GIS, statistical packages, OLAP, etc. Simple null hypotheses, such as CSR for cross-sectional point data, have no equivalent for tracks. We need to build a toolkit of simple analysis techniques, to bring the analysis of tracking data to somewhere near the level of support available for cross-sectional data. We need a library of metrics of similarity between tracks, to support a basic equivalent of cluster analysis, and to identify anomalies and outliers—what do we mean by saying that two tracks are “similar”? We also need better methods of visualization, as anyone who has tried to make sense of the Hagerstrand plot of large numbers of tracks will attest.

Traditional analytic techniques have tended to focus on the “double negative” approach of null hypothesis rejection. Thus although it is normally uninteresting, for example, to establish that a point pattern is distinctly different from the random pattern of CSR, we tend to be satisfied with analyses that result in a rejection of the null hypothesis, and to move only rarely to a more positive acceptance of a well-defined alternative. In part this

is the result of lacking tests with adequate power, since there will always be an infinity of acceptable alternatives, but only one rejected null.

With today's GIS and massive computing power there is the potential to adopt a more positive approach, as is being demonstrated by various types of dynamic models. But we need a toolkit of techniques appropriate for tracking data—stochastic models of tracks, for example.

Although the idea of an average track is compelling, in reality there is no obvious and widely accepted method for averaging tracks, just as there is no widely accepted method for averaging lines in two dimensions (Goodchild, Cova, and Ehlschlaeger, 1995). We need a standard model of uncertainty in tracks, comparable to the Gaussian distribution for scalar measurements. Other simple geometric tools are also needed, such as the polyline-to-arcs tool developed by Noronha and Church for preparing GIS data to support Paramics simulation, and should become a standard part of a computing environment for tracks.

In short, the widespread availability of dense-tracking data has exposed the lack of suitable tools, models, and theories in the current scientific apparatus, and has created an urgent need for fundamental research.

References

- Goodchild, M.F., T.J. Cova, and C.R. Ehlschlaeger, 1995. Mean objects: extending the concept of central tendency to complex spatial objects in GIS. *Proceedings GIS/LIS '95, Nashville, TN*, pp. 354–364.
- VanWey, L.K., R.R. Rindfuss, M.P. Gutmann, B. Entwisle, and D. Balk, 2005. Confidentiality and spatially explicit data: concerns and challenges.

Trace Data for Activity-based Travel Demand Forecasting Models

Kostas Goulias

Department of Geography

University of California, Santa Barbara

A position statement for the FHWA-sponsored Peer Exchange and CSISS Specialist Meeting

GPS and Time-Geography Applications for Activity Modeling and Microsimulation

Santa Barbara, CA

October 10-11, 2005

In travel demand analysis and forecasting we finally see progress in advancing a variety of ideas that were originally proposed in the 1960s and 1970s. One field of research and practice that dominates proposals for new techniques is the activity-based approach with three basic ingredients:

- Explicit treatment of every day life as a time allocation exercise. One expression of this is to treat travel as derived demand, i.e., participation in activities such as work, shop, and leisure motivate travel but travel could also be an activity as well (e.g., taking a drive). Time allocation is viewed as management of episodes (events defined by their starting time, duration, and ending time) and they are arranged in a sequence forming a pattern of behavior that can be distinguished from other patterns (a sequence of activities arranged in a chain/tour of episodes). These episodes and their chains are not independent and their interdependency is accounted for in the theoretical framework(s). For example, one can identify episodes that belong to a “project” such as prepare an evening meal for friends. This project includes a trip to a wholesale store and stay at the store, a trip to a gourmet grocery store and stay at the store, a variety of meal preparation tasks at home, arrival of the friends at home and so forth. More complex projects may focus on children education, career advancement, caring for older parents, and so forth.
- The household is considered to be the fundamental social unit and treated as a decision making unit. Interactions among household members are explicitly modeled to capture task allocation and roles within the household. The relationships and change in these relationships as households move along their life cycles are also modeled. In this way, individual’s motivations, commitments, constraints, and distributed intelligence are explicitly modeled and they are simulated. This allows to model the behavioral process underlying scheduling of activities to identify potential for change under varying policy scenarios;
- Explicit consideration of constraints by the spatial, temporal, and social dimensions of the environment is also given a primary focus. Consistent with and inspired by time-geography these constraints can be explicit models of time-space prisms or reflections of these constraints in the form of model parameters and/or rules in a production system format.

Operational activity-based model systems follow four theoretical and computational traditions that are combinations of:

- Microeconomic models based on Becker's time allocation and McFadden's conditional Logit formulations that develop maximum utility models to simulate policies using appropriate variables within each model;
- Production system/computational process models following Newel's and Simon's rule-based approaches simulating behavior using statements of the type - if X= A then Y=B type of rules;
- Cellular automata models following von Neumann's ideas that also contain rule-based models; and
- Statistical pattern recognition and transition probability approaches to create "data-driven" models.

The input to these models are data of social, economic, and demographic information of potential travelers and land use information to create schedules followed by people in their everyday life. The output are detailed lists of activities pursued, times spent in each activity, and travel information from activity to activity (including departure time, travel time, mode used, and so forth). This output is very much like a "day-timer" for each person in a given region and contains as a subset output usually produced by four-step applications. Unlike the four-step model that produces summaries of coarse behavioral description at the level of small geographic regions (called traffic analysis zones), activity-based approaches produce simulated activities for individuals called *synthetic schedules*. Moreover, given the statistical/econometric maturity of the field we also begin to see outputs that in addition to the averages produced are also accompanied by measures of variability.

To produce these outputs all activity-based approaches use some kind of computational evolutionary engine that is called *microsimulation*. Each model system operates in a somewhat different way but all aim at recreating a virtual microcosm in which individuals and their households are the focus. Their environment is represented by parcels of land where they live and they visit and stay. Their movement is represented by vehicles on highways. Time progression depends on the application but one aim is to develop second by second microsimulations to match other model systems developed for traffic operations.

Many recent model reviews find that applications use hybrid paradigms that are combinations of statistical/econometric models and computational process models to represent behavior. Other use statistical models embedded into microsimulation frameworks to evolve either individuals and/or households over time. Modeling and (micro)simulation appear to be concentrating at two poles. They are either designed for long horizons such as 25 years in the future with yearly cycles aiming at long range forecasting or for shorter periods aiming at reproducing within a day activity and travel patterns. Attempts to reconcile and coordinate different time scales are also starting to appear in the literature and emerging theoretical frameworks.

In this context, traces (i.e., time sequenced paths in space) of individuals and/or vehicles using geospatial technologies can serve a variety of purposes. These traces can be produced by any type of technology and do not necessarily need to be produced by GPS devices alone. The ultimate objective of activity-based approaches is second by second and parcel by parcel simulation of individuals and their social units (households, firms, professional groups, and so forth) on existing highways and pathways. Assuming provision of data will be at that level of spatial and temporal resolutions, there are at least six beneficial roles for GPS data in this context:

1. Use as one source of data to build household time-space prisms.

Travel demand forecasting models use the Hagerstrand-type time-space prism idea to model constraints for individuals. Ultimately we would like to model interactions with other persons via coupling constraints but also the dynamics of everyday communication and task allocation within social groups. These scheduling dynamics, however, may change dramatically shaping each individual's time space rapidly and making the prism itself a much more elastic entity than originally envisioned by Hagerstrand (and all subsequent formulations - see Janelle's position paper) much less restrictive in its influence on travel behavior. Richer longitudinal (many subsequent days) GPS data of all the persons in a household and activity diaries matched to the data may provide the first source of information about household patterns and their change from one day to the next. This information can be used to identify household-based time-space prisms and the rapidly changing action spaces implied by these prisms.

2. Use as a cognitive verification machine about perception of time and distance

Earlier contributions to travel behavior research in 1970s demonstrate that traveler behavior is a function of knowledge about spatial opportunities and time. The striking majority of models use space and time as measured and observed by analysts instead of the persons traveling. GPS and related data offer a unique opportunity to compare these two sources of information in terms of distances (measured and perceived), time durations (measured and perceived), and location sets (observed and considered by the traveler).

3. Serve as external validation/verification devices for synthetic schedule generators

The synoptic measures described by Janele are strategically important devices for synthetic schedule validation when we simulate a geographic area. They provide a richer database than traffic counts because they contain a longitudinal record of movement for a substantial number of persons. One can envision a variety of validation exercises using different aggregation schemes to perform validation/verification of models ranging from individuals to entire cities.

4. Serve as route choice base estimation data

Very few activity-based models address the "assignment" of people and vehicles on the transportation system. As we learn more about these new approaches, route choice emerges as the next important step in modeling and simulation. One way to build route choice models is to use GPS traces that are complemented by other variables as the fundamental model estimation database.

5. Main source of data for the study of habit persistence and anchor points

Limited past research shows people repeat a few basic patterns of behavior over time. It also shows there are anchor points (locations) around which activities and interactions with other persons take place (e.g., home, work, school). Availability of traces offers a unique opportunity to study cycles in behavior and the circumstances under which cycles and habit are broken and replaced with new patterns.

6. Complement travel activity data to produce more accurate data and statistics

A more mundane task in activity data collection is imputation/augmentation of missing data to identify short forgotten trips and missing from activity/travel diaries and hidden behavioral aspects the diaries fail to capture (e.g., walk from parking lot to the office).

Although there are other data collection and data capture needs for activity-based approaches, GPS traces and related data should become a standard tool to complement activity/travel diary surveys. Vehicle-mounted instruments are the first step in procuring the data needed to inform modeling and simulation but we need to move toward wearable devices that can collect spatio-temporal traces covering entire households and other social groups for all their movements in a day.

References

Pribyl O. and K.G. Goulias (2005) Simulation of daily activity patterns. Chapter 3 in *Progress in Activity-Based Analysis* (Ed. Harry Timmermans), Elsevier, pp. 43-65.

Goulias K.G. (2004) Transportation planning and modeling, Chapter 29. In *Handbook of Transportation Engineering* (Editor M. Kutz), McGraw-Hill, New York, NY. Pp. 29.3-29-25

Activity modeling with GPS tracking data: new assumptions for the age of ICT

Helen Couclelis
Department of Geography
University of California
Santa Barbara
cook@geog.ucsb.edu

Introduction

One highly promising use of dense-tracking data captured by GPS is in the area of activity-based modeling of travel behavior. A few pilot studies (e.g., Bradley et al., ND; Wolf et al. 2005) have already demonstrated the validity of the approach in helping reconstruct daily activity schedules, or in complementing travel diary data. Much of the interest of the transportation studies community in the activity-based approach is based on a small number of basic assumptions about the relationships holding in industrialized urban societies among space, time, and individual or household activity. Time geography and its signature model, the space-time prism, have long provided a fruitful conceptual framework around which much of that research has developed. While these assumptions have proved very robust in the past, the increasing penetration of ICT in virtually all aspects of everyday life invites us to rethink some basic premises of the activity-based modeling approach. Three of these assumptions, and the challenges these pose in the information age, are outlined below.

Space and Activity in the Information Age

A basic aspect of time geography that needs to be reexamined in the age of ICT is the underlying assumption of a strong correspondence between a person's movement across space and time, and that person's activity schedule over the same period. It used to be that the location of a person in daily urban space was a reliable proxy for the activity that person was engaged in at the time: if at home then personal and family activities; if at the workplace then work; if at the mall then shopping; if in a car then moving between activities; and so on. Tell me where you are, and I'll tell you what you are doing (Wolf et al., 2005). Conversely, knowledge of a person's or group's activity schedule, along with knowledge of the spatial structure of the corresponding facilities and land uses, used to provide strong clues as to where that person or group might be found at different times of the day. For decades that premise was reliable enough to drive urban planning and land use modeling. For the vast majority of people and activities this is still by and large the case. However, ICT-induced anomalies are proliferating so rapidly that they can no longer be ignored. There is a growing disconnect between activities and adapted spaces as people can shop from their home or workplace, carry on business transactions from their car, socialize while walking alone down the street, or work for a living while sipping cappuccinos at the corner café.

Time and Activity in the Information Age

As long as it is possible to assume a reliable correspondence between activities and adapted spaces, it naturally follows that switching from one activity to the next usually involves some form of physical travel, so that peoples' daily space-time trajectories as

represented in the space-time prism correspond fairly closely to the flows from activity to activity in their daily time budgets. Moreover, to the extent that travel is a cost, switching between activities must be kept at a minimum. This makes for lengthy continuous blocks of time when people stay at the same place doing the same thing. By driving the cost of switching between certain kinds of activities to near zero, ICT remove one major reason (there are many others, of course) for the eight-to-five workday, or the Saturday morning shopping expedition. This does not mean that we have already reached the age of ‘anything, anywhere, any time’: there are practical constraints of place, and social requirements of presence, that limit a person’s freedom to shuffle around his or her activities at will. No one doubts though that the flexibility to do so is increasing, and that people are taking advantage of the newly available opportunities in creative and often unexpected ways.

Households or individuals?

Another assumption that may need to be partially reconsidered is that the household, rather than the individual, is the fundamental decision-making unit. The assumption is largely based on the availability of cars and drivers within the household, and the need for the activity schedules of other household members to adapt to those of the car drivers. However, these days dad can do the shopping on-line from home while mom is driving the car to work, and little Susie can check the references she needs for her homework without having to wait for either parent to drive her to the library. While there are many physical, practical and social reasons why household members will continue to coordinate their daily schedules, this is another area where certain taken-for-granted constraints are likely to become less and less relevant as we move deeper in the information age.

The fragmentation of activity

In principle, then, as ICT-supported activities spread in a society, we can expect to see three things. First is a weakening association between activity and place, meaning that the same activity may be performed at many different places, and the same place may temporarily harbor many different activities. Second, we expect to see a weakening association between activity and time, as in the absence of a need to travel to another place, the cost of switching between activities is drastically reduced. Third, we may see the individual, rather than the household, emerging as the basic decision making unit in activity-based travel modeling. With the increasingly commonly available option of travel-less access to information, goods, and services, people are free to schedule certain activities – and especially the discrete tasks and subtasks that make up the main daily activities – according to any number of other considerations. An activity thus becomes more likely to be performed individually, in discontinuous chunks at arbitrary times. For example, the activity of shopping consists of the tasks of gathering information about a product, comparing prices, trying things out, making a purchase, transferring the item home, exchanging or returning the item, and so on. Each of these tasks may be carried out relatively independently of the others either by physically visiting stores or remotely, with the help of ICT, and with or without the cooperation of other household members. When, how and where these different tasks will be carried out can obviously be very different for a person with access to ICT than for a person without such access. These

observations form the basis for the ‘fragmentation of activity’ hypothesis proposed by Couclelis (2000, 2003). Instead of occupying compact chunks inside the daily prism, ICT-aided activities tend to disintegrate into sets of discrete tasks that get spread out across places and over time.

Some research challenges for GPS-based tracking in the age of ICT

To the extent that these speculations have an empirical basis, they suggest that GPS tracking would have been a perfect data-capture device for travel behavior research some thirty years ago, when ICT was not yet a major player in people’s daily lives. What about today? I would be interested in discussing questions along the following lines:

- How could we test the modified assumptions proposed in this abstract?
- If there is some basis to them, what are the implications for transportation research in general, and for GPS tracking in particular?
- Is the space-time prism still an appropriate model for conceptualizing the relationships between people’s spatio-temporal trajectories and their activities?
- What can GPS tracking data tell us about choices people make between ICT and travel, for different purposes?

References

Bradley, M., Wolf, J., and Bricka, S. (ND) ‘Using GPS Data to Investigate and Adjust for Household Diary Data Non-Response’

Couclelis, H. 2000. From sustainable transportation to sustainable accessibility: can we avoid a new *tragedy of the commons*? In D. G. Janelle and D. C. Hodge, editors, *Information, Place, and Cyberspace: Issues in Accessibility*. Berlin, Germany: Springer, 342-56.

Couclelis, H. 2003. Housing and the new geography of accessibility in the information age. *Open House International*, Vol. 28, no. 4: 7-13.

Wolf, J., Guensler, R., and Bachman, W. (2005). ‘Elimination of the Travel Diary: Experiment to Derive Trip Purpose from Global Positioning System Travel Data’

Dr. Helen Couclelis is Professor of Geography at the University of California, Santa Barbara. She holds a Doctorate from the University of Cambridge, England, a Diploma in Urban and Regional Planning from the Technical University of Munich, Germany, and an MA. equivalent in Architecture from the Technical University of Athens, Greece. Prior to joining UCSB in 1982, she spent several years as a professional planner and policy advisor in Greece. She also represented Greece on several European Union and UN meetings and task forces on urban issues. Her research interests are in the areas of urban and planning theory and modeling, in behavioral geography and spatial cognition, in geographic information science, and increasingly, in the geography of the information society. She was Associate Director of the National Center for Geographic Information and Analysis (NCGIA) from 1993 to 1996 and currently serves on the executive committee of the Center for Spatially Integrated Social Science (CSISS, see www.csiss.org).

Synoptic Analysis of Space-time Activity Patterns

Donald G. Janelle

Center for Spatially Integrated Social Science

University of California, Santa Barbara

A position statement for the FHWA-sponsored Peer Exchange and CSISS Specialist Meeting

GPS and Time-Geography Applications for Activity Modeling and Microsimulation

Santa Barbara, CA

October 10-11, 2005

Introduction

Global positioning systems (GPS) and related spatial technologies will soon be embedded as ubiquitous features of the nation's automobile pool and mobile phone networks, creating a dense systems of mobile sensors with the potential to document in real-time the continually changing patterns of vehicle occupancy on roads and human occupancy of space within urban regions. These developments provide a foundation for the dynamic synoptic mapping of urban activity systems and new ways to explore issues and solutions in transportation planning.

Synoptic analysis has been an important research area in climatology, meteorology, and oceanography for decades, providing valuable insight about processes of both long and short duration over geographical space. The key to success in these disciplines is associated with large sets of distributed information sensors that provide continuous coverage over space and time, widely accepted procedures for standardized measurements, and visualization tools for rendering analytical and descriptive results in a short time frame. Although analogy with human activity is not straightforward, researchers over the past half-century have invoked synoptic concepts to better understand changes in urban and regional activity systems. Examples include Chapin and Stewart's 1953 experiments in mapping population densities in selected U.S. cities for different times of the day; Stewart and Warntz's 1958 conceptualization of population potential fields—continuous non-uniform surfaces that change over time as a consequence of population movements and spatial variations in growth rates; Angel and Hyman's 1976 modeling of urban velocity fields as a basis for estimating travel times; Taylor and Parkes' 1975 analysis of plausible but fictitious census-like variables for eight different times of the day for a stylized city; and Janelle, Klinkenberg, and Goodchild's 1998 use of Halifax space-time diary data to link conceptually and empirically individual space-time paths with diurnal changes in ecological patterns of human activity.

The more explicit focus on individual space-time paths and space-time prisms (pioneered by Hägerstrand 1970 and demonstrated in the work of Lenntorp 1978, Burns 1979, Kwan and Lee 2004, and Miller and Wu 2000) are best seen in the context of the opportunities and constraints that exist in the unfolding time ecologies of urban spaces and behavioral settings. The ability to link tracking data with activity diaries and social surveys would provide a more complete contextual base for interpreting individual activity behavior and for defining the time ecology of urban regions. This would add important new understanding of human activity patterns across diurnal and weekly cycles and permit more encompassing measures of accessibility based on Hägerstrand's (1970) time geography model of society.

The Vision

GPS or cell-phone based dense tracking data provide the foundation for the dynamic modeling of behavior in contemporary cities and for applications that draw on the rich theoretical insight of time-geography. Access to dense tracking data permits dramatically new ways of measuring and representing human activity patterns. Researchers will be able to zoom in and out spatially and at any level of temporal aggregation to produce measures and maps of dynamic shifts in population densities. Other measures potentially include traffic densities, ratios of cars to trucks, proportions of within-region to pass-through traffic, surface representations of average travel speeds, congestion indexes, and weekend to weekday

ratios, along with other indicators for exploratory research and diagnostic use. In combination with social and diary surveys, these dense tracking data can yield temporal variations in social group integration, dynamic measures of spatial concentration, at-risk populations to threatening events, and a wide range of other indicators of human occupancy in space and time. Such data could be used to display a dynamic representation of population potential surfaces (Warntz 1966) and velocity fields (Angel and Hyman (1976), illustrating their utility for determining minimum paths and for allocating public services that account for diurnal variations in the serviced populations. They might also be used to assess the significance of such concepts as accessibility and opportunity for answering questions about health risk, social exclusion and equity, and optimality in facility or service allocation.

What is Needed?

To bring about applications of responsive synoptic analysis of space-time activity patterns, there is a need to set in place methods for collecting, analyzing, and modeling changing patterns of behavior over diurnal weekly, and other periodicities. In contrast to decennial census data and travel survey methodologies, the synoptic approach relies on data that are continuous over time and space, and it requires procedures for the automated real-time visualization of traffic and human movement behavior in space. Thus, through use of global positioning systems (GPS) or cellular technologies for dense tracking of subjects, and through data visualization tools, a dynamic conversion of these data into mapped patterns and index measures is possible. Although this approach opens up a wide variety of new opportunities for plan evaluation in urban and transportation planning, it also poses significant challenges about how to manage large complex datasets that simultaneously reference space and time, how to use effectively such data for inference assessment, and how to protect the location privacy rights of individuals while minimizing the resultant loss of scientific and social value to the information. In this position statement, I am suggesting the need to develop proof-of-concept research based on the use of existing data sets on vehicle movements to:

- (1) demonstrate a synoptic approach for analyzing movement behavior in selected American cities;
- (2) document the value of this approach for modeling space-time patterns of activity behavior;
- (3) address issues of how to manage extensive space-time data sets;
- (4) define and test procedures for protecting individual location privacy rights;
- (5) assess the feasibility of linking GPS tracking data with respondent activity diaries in urban regions; and
- (6) establish strategies to implement dense tracking for continuous synoptic mapping for urban regions. A plan for the deployment of a large-scale, real-time implementation of these methods in a major American city is needed.

References

- Angel, S., and G.M. Hyman 1976. *Urban Fields: A Geometry of Movement for Regional Science*. London: Pion.
- Burns, L. 1979. *Transportation, Temporal and Spatial Components of Accessibility*. Lexington, MA: Lexington Books.
- Chapin, Jr. F.S., and P.H. Stewart 1953. Population densities around the clock. *The American City* (October). Reprinted in H. M. Mayer and C.F. Kohn (eds.). *Readings in Urban Geography*. Chicago: University of Chicago Press, 180-182.
- Hägerstrand, T. 1970. What about people in regional science? *Papers of the Regional Science Association* 24:1-12.
- Janelle, D.G., B. Klinkenberg, and M.F. Goodchild, 1998. The temporal ordering of urban space and daily activity patterns for population role groups. *Geographical Systems* 5:117-137.
- Kwan, M-P., and J. Lee. 2004. Geovisualization of human activity patterns using 3D GIS: A time-geographic approach. In M. Goodchild and D. Janelle (eds.) *Spatially Integrated Social Science*. New York: Oxford University Press, 48-66.

- Lenntorp, B. 1978. A time geographic simulation model of individual activity programmes. In T. Carlstein, D.N. Parkes, and N.J. Thrift (eds.). *Human Activity and Time Geography*. London: Edward Arnold, 162-180.
- Miller, H.J., and Y. Wu 2000. GIS software for measuring space-time accessibility in transportation planning and analysis. *GeoInformatica*, 4, 141-159.
- Stewart, J.Q., and W. Warntz 1958. Physics of population distribution. *Journal of Regional Science* 1:99-113.
- Taylor, P.J., and D.N. Parkes 1975. A Kantian view of the city: A factor ecological experiment in space and time. *Environment and Planning A*. 7:671-688.
- Warntz, W. 1966. The topology of socio-economic terrain and spatial flows. *Papers, Regional Science Association* 17:47-61.

Time -Geographic Methods For Analyzing GPS Data

Mei-Po Kwan

Department of Geography, Ohio State University

kwan.8@osu.edu

For transportation researchers and geographers interested in understanding human behavior in space-time and its complex relationship with the urban environment, the possibility of collecting and using GPS data offer new opportunities and pose many challenges at the same time. Recent studies that use GPS data of individual trips indicate that GPS may allow the accurate reconstruction of individuals' activity-travel patterns with considerable space-time details (Murakami and Wagner, 1999). The reconstructed patterns (as revealed preference of travelers) can be used for various purposes, including the study of route planning and route choice behavior, activity scheduling involving complex space-time trade-offs, the daily and weekly temporal rhythms of activities or trips (e.g. cycles and repetitions), and responses to real-time information as in the context of Intelligent Transportation Systems (Kwan, 2000; Zhou and Golledge, 2000). Further, as the precise space-time trajectories of walking trips can be recorded, transport problems of the mobility-impaired population subgroups - such as the vision-impaired or the elderly people - can be analyzed in great space-time details.

GPS data may also enable the development of a whole range of new analytical and modeling methods useful to transportation researchers and geographers. Given the recent development and application of time-geographic GIS-based methods in many areas of human spatial behavior, it is now possible to analyze and model individual travel behavior in relation to a realistically represented urban environment - in the form of a comprehensive geographic database of the study area that includes information about all land parcels and all street segments (including segment-specific travel speed and turn restrictions). GIS-based geocomputation and 3D geovisualization of activity-travel diary data are examples of the methods useful for the analysis of GPS data (Kwan and Lee, 2004).

Several difficulties still remain for future studies that attempt to use GPS data. First, discontinuities or gaps in GPS data cannot be entirely eliminated due to the problem of "loss of fix" that can happen under certain circumstances (e.g. loss of signal inside certain areas of a building). As past studies indicate, dealing with this data discontinuity problem may require considerable time and effort during the data preparation phase of a study. Second, although comprehensive data can be collected if all of the activities and trips an individual undertakes throughout the survey period are recorded, users of GPS may not keep their devices on throughout the day for various reasons. This will introduce gaps in the data and make it difficult to achieve completeness in GPS data. Third, the activity a user is doing at a particular time cannot be easily recorded with GPS devices. Although linked devices or integrated devices (e.g. PDA) had been used to record and link activity data with the location data generated by GPS devices, collecting complete and reliable data

is still very difficult. Fourth, the ability of GPS devices to reveal and record people's geographical position raises serious concern about privacy when using the technology.

References:

Kwan, Mei-Po. 2000. "Evaluating gender differences in individual accessibility: A study using trip data collected by the Global Positioning System." A Report to the Federal Highway Administration (FHWA), U.S. Department of Transportation, 400 Seventh Streets, S.W., Washington, D.C. 20590.

Kwan, Mei-Po and Jiyeong Lee. 2004. "Geovisualization of human activity patterns using 3D GIS: A time-geographic approach. In *Spatially Integrated Social Science*, 48-66, Michael F. Goodchild and Donald G. Janelle (eds). New York: Oxford University Press.

Murakami, Elaine and David Wagner. 1999. Can using global positioning system (GPS) improve trip reporting? *Transportation Research C*, 7:149-165.

Zhou, Jack and Reginald Golledge. 2000. "An analysis of household travel behavior based on GPS." Paper presented at the Ninth International Conference on Travel Behaviour Research, the International Association of Travel Behavior Research (IATBR), Gold Coast, Australia, July 2000.

Mark Bradley

Bio

Mark is an independent consultant based in Santa Barbara, California. He has a B.S. in Operations Research from Cornell University, an M.S. in Systems Simulation Modeling from Dartmouth College, and an M.S. in Mythological Studies and Depth Psychology from Pacifica Graduate Institute. For more than twenty years, he has carried out consulting projects to apply state-of-the-art travel demand modeling methods. He spent ten years working in Europe with Oxford University and Hague Consulting Group before returning to the United States in 1995. Since then, he has helped to create activity-based travel demand model systems for use in Portland, San Francisco, Sacramento, Columbus, and Atlanta.

Position Statement

The position statements by Goulias and Pendyala contain an overview of various issues and topics in GPS for activity-based travel demand modeling. Since my consulting focus is on designing and creating activity-based models for immediate implementation by regional MPOs, I will limit my focus to specific issues related to improving activity-based models in the near term.

Data Collection

The most common use of GPS data collection in travel surveys to date has been to validate and adjust diary-based surveys. This is NOT a very good use of GPS data, and the analyses I have done suggest to me that it would be much better to move to GPS as the *primary* data collection method, using diary-based methods only to provide additional details. This approach is being adopted for the upcoming Oregon Statewide Travel Survey. The key features planned for that survey include:

- A statewide sample of about 3000 households per year, with about 1000 in a longitudinal panel and the rest interviewed just once (a repeated cross-section).
- Seven day vehicle-based GPS data collection for every vehicle.
- Additional diary-based data collection for one or two days, in order to get additional trip details, especially for walk and transit trips.
- Use of the retrieval telephone call to validate and correct GPS trace data in real time (identify missing stops, extra stops, etc.)
- Later use of person-based wearable GPS devices when the technology improves sufficiently. (The pilot study indicated too much missing data with the current technology.)

Data Analysis

Activity-based models systems attempt to simultaneously predict several dimensions of travel behavior, including:

- Activity purpose and frequency
- Trip chaining
- Activity duration
- Departure time choice
- Destination choice
- Mode choice
- Route choice

Clearly, these choices are all interrelated, but it is not feasible to treat a day's activity pattern as one complex simultaneous choice. Thus, the main practical challenge in activity-based modeling is to derive a statistically tractable yet behaviorally realistic way of defining a hierarchy across these choices. GPS data can be very valuable in this research. Some of the key questions to be answered are:

Which choice dimensions are most stable and "fixed" over time? Which activities affect the characteristics of other activities carried out by the same person?: One might assume that activities that are always done at the same place and at the same time of day have higher priority in activity decisions than those activities that are more flexible. Conversely, if we find that regular activities are done at different times and/or locations when another specific activity is present, then we can assume that the second activity has priority. For example, someone may always arrive and leave work at the same time, except for Thursdays when they pick their child up from school. The activity of picking up the child would have priority in a scheduling sense. Similarly, someone might always do grocery shopping at a certain store, except when combined with other personal business and entertainment activities, in which case they choose a supermarket towards the city center. In that case, shopping would have a lower priority in terms of location choice. A long-term GPS data set such as that collected in Atlanta can be very useful in looking at distinctions of this type.

What is the extent of variability observed over time? With a longer term GPS data set, we can measure variability in both the "exogenous" variables in our model systems (e.g. what is the variability in travel time between home and work at a certain time of day?). We can also measure the variability in the endogenous behavioral variables. For example, on days when a person performs the same sequence of activity types, how much variation is there in the times and places that those activities are performed. Also, how much variety is there in the routes that a person drives between the same set of locations?

So, we need to look at both regularity and variability in the activity patterns carried out by the same individuals over time. These are two sides of the same coin and can be analyzed at the same time, but will each provide unique information to inform future model development.

Michael McNally
Institute of Transportation Studies
University of California, Irvine

The Merging of Travel Forecasting and Traffic Management Data and Models

A FHWA-sponsored Peer Exchange and CSISS Specialist Meeting
GPS and Time-Geography Applications for Activity Modeling and Microsimulation
Santa Barbara, CA October 10-11, 2005

Position Statement

My experience with travel forecasting in general and activity-based models in particular has led me to the conclusion that we are facing some fundamental data problems. It is not simply an issue of obtaining travel behavior data that is both richer and more accurate; rather, it is the belief that data and models that can address the forecasting of travel behavior tomorrow are the same constructs that are needed in addressing traffic management today.

The fields of travel forecasting and traffic operations have grown closer over the last decade, in large measure due to the increased availability of microsimulation software and the promise of real-time travel data. Several recent projects in each of the two traditional fields have supported this contention. My work with Ming Lee developing the REACT! web-based activity survey and in developing the TRACER GPS-based in-vehicle data logger (with James Marca and Craig Rindt) was conducted in parallel with the development of CARTESIUS, a multi-agent traffic management decision support system. The cross-fertilization was quite significant and led to two other projects, AUTONET and PTC.

AUTONET is an architecture for an interoperable information technology infrastructure that features a mobile, ad-hoc, dynamic, peer-to-peer network that integrates vehicles, information, and communication systems with a fixed transportation infrastructure to attain a comprehensive distributed transportation management system. The primary result thus far may be the realization that the technology of AUTONET is achievable but the application itself seems to neither satisfy the needs of forecasting nor management. This led to the development of the PTC project that addresses the problem of collecting, storing, and utilizing AUTONET-type data from each vehicle in a traffic network, where the data can be trusted by transportation system operators while simultaneously ensuring traveler privacy. To achieve this goal, each vehicle stores its own travel history, under the consent of the driver, by accepting authenticated information from roadside controllers (persistent traffic cookies, PTC) using short-range wireless communication. The authenticated data stored in each vehicle forms a distributed database of historical travel patterns. The central hypothesis of the project is that these historical travel patterns can be used to predict the movement of vehicles currently in the system, which can, in turn, be used for traffic management applications.

The gap between operations and planning data is shrinking and we should be planning for a future when spatial data is almost universally available from mobile service providers in ways analogous to how web click streams are converted into marketing data for businesses today.

High-resolution measurement of time geographic entities

Harvey J. Miller
Department of Geography
University of Utah
260 S. Central Campus Dr. Room 270
Salt Lake City, UT 84112-9155 USA
harvey.miller@geog.utah.edu

Time geography is a powerful perspective for understanding human spatial behavior, in particular, constraints and tradeoffs in the allocation of limited time among activities in space. The development of *location-aware technologies* (LAT) raises the potential value of time geography in social research and in applications such as transportation engineering and planning. LATs, such as the global positioning system (GPS) and radiolocation methods can allow measurement of basic time geographic entities and relations at high spatio-temporal resolutions, as well as massive data volumes. These data have potentially enormous scientific value to time geographers, urban researchers, transportation analysts and social scientists.

A problem is that classical time geography is not ready for the measurement tasks demanded by LATs. Classical time geography is conceptual: analytical statements of its basic entities and relations do not exist beyond the informal descriptions of Burns (1979) and Lenntorp (1976). These are not effective for inferring time geographic entities and relationships from high-resolution measurement of mobile objects in space and time. They are not sufficient for analyzing the propagation of uncertainty when measuring these objects imperfectly. The looseness and incompleteness of these descriptions means that comparisons of detailed data and time geographic analyses across different studies may be difficult since the entities and relations are not strictly comparable.

Some recent formalisms by Miller (1991) and Kwan and Hong (1998) make progress towards the goal of an analytical time geography, but these are limited to the special case of networks. Hornsby and Egenhofer (2002) develop a framework for multi-granularity representations of space-time paths, prisms and composite paths/prisms to support space-time queries (although they unfortunately invent a non-standard terminology). However, their framework uses simultaneous inequalities to describe these entities. These are cumbersome for analytical statements about high-resolution time geographic measurement and uncertainty propagation since they describe the entities implicitly.

Representation and analysis of dynamic entities measured at high spatio-temporal resolutions also arise in geographic information science. The development of LATs and *location-based services* (LBS) is inspiring a growing literature on database design for storing information on moving objects. Since digital technologies can only sample an object's location at discrete moments in time, a key problem is interpolating an object's location at any arbitrary moment based on the sampled locations (e.g., Pfoser and Jensen 1999).

This presentation will address the lack of analytical rigor at the foundation of time geography by presenting a high-resolution measurement theory for its basic entities and

relationships (Miller 2005a). Drawing from the literature on moving objects database design, this paper shows that the location or spatial extent of time geographic entities at any moment in time can be solved as compact spatial sets, or the intersection of compact spatial sets, derived from the sampled locations and auxiliary information such as a maximum travel velocity. These sets are simple geometric objects that have algebraic solutions in one or two spatial dimensions. These simple objects support evaluation of time geographic relationships such as bundling and intersections. Numeric solutions are required for some entities and relations in higher dimensions, but these are relatively tractable since they involve simple surfaces. Inferring relationships such as path bundling or path-prism and prism-prism intersections is also tractable since this involves simple geometry. An ancillary but useful benefit of the measurement theory is a generalization of time geography. The measurement theory allows analytical statement of time geographic entities in n -dimensional space rather than the strict two-dimensional space of classical time geography. The cases of direct interest in time geography are $n = 1$ (networks), 2 (planar space) or 3 (natural space), and the measurement theory can support consistent measurements and analysis of time geographic entities across these cases (e.g., a vehicle moving within and outside a transportation network). Although not addressed in this presentation, it is also easy to extend the measurement theory to include virtual interaction using information and communication technologies such as email, the world wide web and cell phones (see Miller 2005b)

Literature cited

- Burns, L. D. (1979) *Transportation, Temporal and Spatial Components of Accessibility*, Lexington, MA: Lexington Books.
- Hornsby, K. and Egenhofer, M. J. (2002) "Modeling moving objects over multiple granularities," *Annals of Mathematics and Artificial Intelligence*, 36, 177-194.
- Kwan, M.-P. and Hong, X.-D. (1998) "Network-based constraints-oriented choice set formation using GIS," *Geographical Systems*, 5, 139-162.
- Lenntorp, B. (1976) Paths in space-time environments: A time geographic study of movement possibilities of individuals, *Lund Studies in Geography Number 44*, Royal University of Lund, Sweden.
- Miller, H. J. (1991) Modeling accessibility using space-time prism concepts within geographical information systems, *International Journal of Geographical Information Systems*, 5, 287-301
- Miller, H. J. (2005) "A measurement theory for time geography," *Geographical Analysis*, 37, 17-45.
- Miller, H. J. (2005b) "Necessary space-time conditions for human interaction," *Environment and Planning B: Planning and Design*, 32, 381-401.
- Pfoser, D. and Jensen, C. S. (1999) "Capturing the uncertainty of moving-object representations," in R. H. Güting, D. Papadias, and F. Lochovsky (eds.) *Advances in Spatial Databases*, Berlin, Springer: Lecture Notes in Computer Science 1651, 111-131.

Elaine Murakami, Federal Highway Administration

Bio

Elaine Murakami is a Community Planner with the Federal Highway Administration's Office of Planning in Washington, D.C. She lives in Seattle, WA. Previous to her job at FHWA, she worked at the Puget Sound Region Council, the MPO for the Seattle area. She was instrumental in the 1996 Lexington, KY GPS pilot study that used Sony MagicLink handheld touchscreen computers along with magnetic mount GPS units for a 6-day vehicle survey (<http://www.fhwa.dot.gov/ohim/lextrav.pdf>). She has contributed toward other innovations in travel and activity surveys including the integration of web-based GIS (<http://www.fhwa.dot.gov/ohim/trb/sbir/sbir.htm>), and implementing a longitudinal travel survey (the Puget Sound Transportation Panel).

Position Statement

TITLE: Bringing Geographers and Travel Demand/Activity Modelers together to Benefit from new GPS travel data resources.

Self-reported travel and activity surveys have been the standard approach used by transportation survey researchers over the last fifty years to capture daily travel behavior used in travel demand forecasting. Data collection methods have shifted from in-home interviews to telephone interviews and mail-back survey forms. The goals for these surveys have likewise shifted, with early surveys in the 1950's and 1960's with samples of 5 percent of households in a region, to provide origin/destination matrices. More recent surveys have been in the range of 2,000 to 15,000 households for a region for a 1-day travel diary, with stratification of households by number of vehicles and number of persons in the household or number of workers in the household, with the objective of capturing trip rates by purpose, trip length distributions, and vehicle miles of travel. While sample stratification may have had a geographic component in recent surveys, e.g. quotas for specific political boundaries, or quotas for areas near transit service, the geographic distribution of trips became of less importance in these regionally conducted surveys over the past 30 years.

Using GPS may allow for further reductions in sample sizes for travel behavior surveys, but with longer data collection periods. A GPS-based survey reintroduces the concept of spatial analysis to household travel/activity surveys. The benefits of having longer data collection periods may outweigh the losses in sample size. In particular, there are at least two FHWA value pricing projects where GPS is being used for analyzing private vehicle travel behavior for "before" and "after" conditions. The Commute Atlanta data includes 365 days of "before" data, and the Puget Sound project includes 60 days of "before" data.

Some of the benefits of these long data collection periods using GPS include:

1. Ability to capture a "truer" picture of the spatial range of a person's travel

2. Ability to examine variability/stability of travel behavior over a time period such as a week or a month.
3. Ability to examine relatively infrequently made trips, e.g. those over 50 miles.

The majority of daily trips in the United States occur in private vehicles (87 percent),¹ therefore using GPS to capture the majority of travel can result in more complete (fewer errors of omission), and more accurate (time, distance) information, and provide details heretofore unavailable using self-reported methods (speed, route choice).

Research in Japan has used GPS-enabled cell phones for tracking personal travel, not limited to vehicles. Recent work in the United States to develop GPS-enabled personal tracking devices has been hindered by battery weight that increase the burden of completing the survey, and lags in 3G telecommunication services for mobile phones.

GPS surveys may be a better source than traditional travel behavior surveys for microsimulation of travel behavior because they can be directly linked to network characteristics. For example, because the route choice can be determined, use of different types of roadways (local vs. arterials vs. highways), use or avoidance of left turning lanes, use or avoidance of signalized intersections, can be incorporated into the simulation algorithms. Similarly, if detailed parcel-based land use information is available, it can also be incorporated into microsimulation algorithms.

However, there is a need for spatial analysis tools and trained analysts who can access, process, and analyze the wealth of GPS data either now available, or will be available very shortly. The transportation community can benefit from existing work in geography, but geographers must understand what applications are desired by transportation modelers. GIS statistical software is becoming more robust. This meeting will help bridge the two communities, and can establish some shared interests and efforts in the near future.

¹ “Highlights of the 2001 National Household Travel Survey” BTS-03-05, Bureau of Transportation Statistics, U.S. DOT, Washington, D.C. 2003

GPS and Travel Monitoring

Val Noronha (noronha@dgrc.ca)
University of California, Santa Barbara
September 2005

Several areas of DTD¹ in transportation are of interest to me, from the challenges of instrumentation to real-time and post-analysis of tracks, data uncertainty, map interoperability and automated update.

In the area of personal travel monitoring, five problems stand out.

Exploring microscopic driving behaviors in space-time

DTD opens up possibilities for monitoring not just origins, destinations and routes, but also microscopic behavior en route. Where do drivers slow down and accelerate? What does that reveal about congestion, road conditions and visibility, permanent or temporary traffic restrictions, the individual's familiarity with the area, or land use? There are obvious and compelling uses of such data in incident management and infrastructure analysis—for example, in preliminary tests we were able to determine whether intersections had traffic lights or stop signs, or were unregulated. One area of potential application in travel studies is to automate the detection of the least intelligible trip ends: passenger drop-offs and pick-ups. Another is perhaps to capture driver attitude as an explanatory variable in route choice.

Microsimulation

Real data on route choice, and variability in route choice, enrich traffic assignment and microsimulation algorithms. There are potential applications in other areas—reflexive response to horizontal curves and vertical gradients, given visibility and other conditions—but GPS resolution is currently insufficient and driving simulators provide superior data. In the future, when GPS is available for larger samples, and data are delivered simultaneously from other vehicle based sensors (e.g. headway distance and video), we can anticipate additional inputs to microsimulation.

Shopping choices: DTD fills in the blanks

DTD provides valuable insights on shopping choices. Due to historic structures of census data, retail location models commonly assume store choice to depend on distance to one's home. We know from introspection that the thought process is more complex. Up to now, financial firms have had some ability to track the time and location of transactions, creating a coarse breadcrumb of an individual's activities and spending habits. DTD fills in the gaps, potentially shedding light on the decision process, with plentiful data for model specification and calibration.

Improving the user interface for electronic diaries

Travel survey by trip diaries is expensive and error-prone, because of the substantial human input component. Entirely passive data gathering is technically much easier and relatively free of error, but the data do not reveal trip purpose, reasons for detours, the number of passengers traveling together, etc—these require input from the subject. To a large degree, intelligent processing can infer trip ends and purposes. The task of the user interface can therefore be reduced to prompting the user for input at points of uncertainty, and confirming inferences. Voice synthesis and recording

¹ We use the term "Dense Tracking Data" (DTD) to refer to space-time couplings (x,y,z,t) at resolutions of roughly 1~10 m in space and 0.1~10 seconds in time.

(both of which are easily implemented on common current hardware) can further simplify the user interface, decreasing the burden for the subject.

Protecting, spoofing and uncovering the truth

DTD are inherently information-rich, so the more ambitious the plans for DTD mining, the deeper the conflict with personal privacy. There are legal and institutional workarounds and avenues towards compromise, such as total informed consent. What about at the technical level? One technical compromise is to coarsen the spatial and/or temporal resolution of the data. This raises research questions: What is an appropriate resolution in space and time at which data are sufficiently anonymous, yet sufficiently useful? Because of the rich content of DTD, the block group is too small as a basic data unit, to protect the identity of individuals when the entire population is tracked; to protect privacy is smaller samples, aggregation units to be larger. What are the types of packages in which useful data can be presented to researchers so that individuals are no longer identifiable? What are the methods by which data can be spoofed, for example by swapping trajectories of residents of the same spatial aggregation unit? Can a determined sleuth reconstitute the tracks? What types of data (e.g. the GSV satellite configuration sentence in NMEA) must be suppressed? At what cost—what is the residual fitness for use of the corrupted, aggregated data?

The consequences of information in the wrong hands must be balanced against the value of information in the right hands. Perhaps as often as not, it is in an individual's interest to be tracked.

Dr. Val Noronha is Director of the Vehicle Intelligence and Transportation Analysis Laboratory at the University of California, Santa Barbara; Project Director of the National Consortium on Remote Sensing in Transportation–Infrastructure; and President of Digital Geographic Research Corporation. He has worked in GIS for more than 20 years, in Australia, Canada and the U.S., in academia and as a consultant to governments and the private sector. In transportation he has worked on location expression and interoperability standards testing, centerline mapping error, remote sensing applications, GPS, data modeling, transit and accessibility.

COLLECTION AND ANALYSIS OF GPS-BASED TRAVEL DATA FOR UNDERSTANDING AND MODELING ACTIVITY-TRAVEL PATTERNS IN TIME AND SPACE

Ram M. Pendyala, Department of Civil and Environmental Engineering, University of South Florida, 4202 East Fowler Avenue, ENB118, Tampa, FL 33620. Ph: (813) 974-1084; Email: pendyala@eng.usf.edu

Introduction

Activity-based microsimulation models aim to provide a rigorous analytical and behavioral framework for forecasting travel demand. Over the past decade, rapid advances have been made in the specification, formulation, and estimation of activity-based model systems. Simpler forms of the activity-based modeling paradigm, commonly referred to as tour-based model systems, have been developed and applied in practice with a view to enhancing the state-of-the-practice in travel demand modeling and transportation planning. Examples of such applications include Portland, San Francisco, New York, and Columbus. Several other agencies, including those of Denver, Atlanta, Dallas-Fort Worth, Sacramento, and the Puget Sound to name a few, are considering or have initiated the development of tour-based model systems for their areas.

Although tour-based model systems constitute a significant enhancement over traditional four-step trip-based models because they account for inter-dependency among trips, trip chaining, and trip scheduling, they are not full-fledged activity based model systems. This is because they do not fully reflect spatial and temporal constraints and opportunities, model interactions among agents, capture time use and allocation behavior, and consider activity participation along the continuous time dimension. In addition, both tour-based and activity-based models developed thus far have done little to account for route choice behavior and the effects of that behavior on activity participation, activity-travel durations, and scheduling patterns.

It is clear that the development of full-fledged activity-based models that fully account for and reflect the aspects noted in the previous paragraph calls for the collection, synthesis, organization, manipulation, integration, visualization, analysis, and modeling of detailed data about people's activity-travel patterns by time of day, mode and route choices, activity-sequencing and trip chaining, spatio-temporal constraints and opportunities, and interactions in time and space. Until recently, collecting such information through traditional travel surveys was virtually impossible. Therefore, activity and tour-based models developed thus far have had to make compromises and simplifying assumptions in model specification, behavioral basis, and structure. However, more recently, GPS-based data collection components/experiments, when combined with activity-travel survey efforts, are offering rich and detailed data about aspects of behavior that hitherto remained a mystery. The time is ripe for geographers and travel demand modeling researchers to work together and explore ways of fully exploiting the GPS data in combination with traditional activity-travel behavior data for developing new generations of activity-based microsimulation models.

Survey Methods: Collection of GPS-Based Activity-Travel Data

In the recent past, considerable experience has been gained in the collection of GPS data in the context of activity-travel surveys. However, there are several aspects of GPS data collection that merit further exploration so that GPS-based data collection experiments can be set up to maximize information acquisition per unit cost. These areas for exploration include, but are not necessarily limited to:

1. *Comparisons of GPS and Traditional Travel Survey Data:* While there have been a few attempts to compare travel patterns from GPS and traditional travel survey samples to understand the extent to which traditional travel survey data sets miss or under-report trips, this appears to be an area that merits further exploration. These comparisons could help in developing correction factors for traditional

travel survey data sets. More importantly, these comparisons will help determine the extent to which short and infrequent trips and trip chaining are under-reported.

2. *Passive vs. Active Data Collection:* GPS instruments are able to collect location and timing information passively on a continuous basis. In other words, the respondent or traveler does not have to enter this information. However, the traveler has to enter or record information such as trip purpose and vehicle occupants. On the other hand, it might be possible to make inferences based on GPS data. The location coordinates of the trip end, when combined with a parcel-level land use map, might provide sufficient information to figure out the most likely trip purpose pursued by the individual. The duration for which the GPS records no locational movement could be used to determine the end point of a trip without having to burden the respondent to manually indicate the end point of a trip. What duration threshold should be used to signify that a trip has been completed? In summary, it merits exploration to perform some comparisons and tests to figure out what data we can reliably gather through passive GPS instruments and what data we need to actively gather from the respondents/travelers.
3. *Resolution of Data Collection:* GPS instruments are able to provide location and timing information on a second-by-second basis. This leads to huge amounts of data/records even for a single trip. Do we really need second-by-second data? Is it sufficient to record data at 30 second intervals, 10 second intervals, 1 minute intervals, or....? Perhaps, it would be possible to construct travel itineraries using records based on different time intervals and compare the itineraries so constructed. These comparisons could shed light on the ideal temporal resolution for GPS data collection. From a spatial perspective, the GPS instruments provide coordinate data (latitude and longitude). This is exactly what we need; the only question is whether we need the latitude and longitude information for every second.
4. *Understanding GPS Data Limitations:* While there is increasing recognition of the benefits of GPS data collection, perhaps it is necessary to make sure that we also understand the limitations of GPS data. In addition to cost considerations, what are potential limitations of GPS data collection efforts in terms of technology reliability, battery life (although this is getting better by quantum leaps), missing data (say, due to travel through urban canyons), and so on? Understanding these limitations will help identify mechanisms and tools for better exploiting and using GPS data in a travel analysis and modeling context.

Analysis and Use of GPS-Based Activity-Travel Data

The application of GPS-based activity-travel data for the development of activity-based microsimulation models of travel demand requires the analysis and use of the data from a variety of new perspectives that have not been traditionally addressed by travel behavior researchers due to data and technology limitations. However, the availability of GPS data together with rapid advances in GIS software technology has made it possible to exploit and apply GPS data in ways not possible in the past. Several aspects merit consideration in the context of analyzing and using GPS data in support of activity-based travel demand modeling. These include, but are not necessarily limited to:

1. *GPS Data Reduction and Synthesis:* GPS data sets are often large and include large amounts of temporal and spatial data on travel patterns. We must develop, and have widely available, tools that can reduce and synthesize the information/records into meaningful data sets and visual displays. Tools are needed to correct for errors, fill missing gaps, decipher beginning and ending points of trips, link trip ends with secondary data such as land use and network data, and so on.

2. *Day-to-Day Variability and Weekly Rhythms in Activity-Travel Participation*: One of the great advantages of using GPS-based methods is that one can potentially collect travel data over several days (and perhaps weeks) because the burden on the respondent is lower. A lot of the spatial and temporal information is recorded automatically by the GPS device. Then, GPS-based data should be used to analyze day-to-day variability and weekly rhythms in activity-travel patterns. The Mobidrive data collected in Germany over a six week period illustrated the potential value in having data over a longer period. We can learn about inter-day interaction, inter-person interaction, task allocation, intra-person variation, and habitual and occasional activity participation. In addition, having data over multiple days and weeks can help inform the time duration that should be covered by our activity-based microsimulation models. While it may be reasonable to have daily activity patterns as an output of these models, the models themselves may need to incorporate inter-day and inter-week interactions (history dependency, anticipatory dependency, etc.) to accurately model a single day's activity-travel patterns.
3. *Space-Time Paths and Prisms*: The biggest benefit that GPS data might be able to provide is in the measurement, representation, description, analysis, and explicit incorporation of space-time prisms and paths into our microsimulation models. What makes a model an activity-based model is the explicit incorporation and recognition of space-time constraints, opportunities, action space, and paths into the model system. While some attempts are being made to do that with currently available travel survey data, these attempts can be advanced by quantum leaps through the use of GPS data that provide *detailed spatio-temporal data* at a very fine resolution in an integrated manner. Essentially, the construction of true 3-D time-space paths and prisms is possible only with GPS data. As activity-based approaches attempt to model activity-travel patterns in *time and space*, there is no doubt that GPS data have much to offer in this regard. The challenge is to develop tools to analyze and model space-time paths and develop activity-based modeling paradigms/structures/frameworks that explicitly utilize these paths and prisms to simulate activity-travel patterns of individuals.
4. *Interactions Among Agents*: GPS data should be used to advance our understanding of interactions among agents. What is the interaction of individuals with the natural and built environment over space and time? By linking GPS data with land use (parcel level) data in a GIS framework, spatial correlations and relationships can be developed. These relationships can, in turn, be incorporated into activity-based microsimulation models in the context of constructing choice sets and modeling destination choice. Similarly, inter-dependencies among activities and trips and between individuals can also be explored in greater detail with GPS data.
5. *Route Choice*: It remains a pity that the route choice modeling community has not yet embraced the use of GPS data for modeling and understanding route choice behavior. This is the best source of route choice data and can provide valuable information on inter-relationships between trip chaining, mode choice, destination choice, and route choice. In addition, one can potentially construct behavioral paradigms on route choice by examining differences between paths chosen and paths that are shorter, faster, safer, and more scenic. Perhaps it is time to bring in the route choice modeling community in the analysis of GPS-based travel data.
6. *Supply (Network) Data*: One area in which GPS-based travel data can be very useful is in developing, enhancing, updating, and correcting network (supply) data. We are all familiar with the rather poor quality network supply data that is currently available and used in our travel demand models. These variables play critical roles in trip distribution, mode choice, and traffic assignment (i.e., at least three of the four steps). However, travel times and speeds are often suspect and inaccurate. GPS-based

travel data provides accurate spatial and temporal information at the finest resolution. In other words, we can get accurate travel times by time of day, potentially build speed-flow relationships using the data, and assess speed and travel time variability by link. This is another way in which the network modeling community can benefit from the use of GPS-based travel data.

Conclusion

The above discussion illustrates the potential use, value, and applicability of GPS data to the development of activity-based microsimulation models of travel demand and understanding relationships that govern activity-travel patterns in time and space. The travel demand modeling community should move forward with an aggressive research agenda that demonstrates the usefulness and value of the data and the ease with which the data can be reduced, synthesized, analyzed, visualized, and utilized in real transportation planning applications. The Commute Atlanta GPS data set and a few others appear to provide the sample size and richness needed to help accomplish this. The research agenda should focus on developing tools, analyzing relationships, studying behavioral patterns in time and space, and informing activity-based model specification.

Time Geography for Activity Modeling with GPS Tracking Data

Shih-Lung Shaw
Department of Geography,
University of Tennessee, Knoxville, TN
Email: sshaw@utk.edu

A position statement for the FHWA-sponsored Peer Exchange and CSISS Specialist Meeting
GPS Tracking and Time Geography Applications for Activity Modeling and Microsimulation
Santa Barbara, California
October 10-11, 2005

Introduction

Global positioning systems (GPS) and related technologies have made it feasible to collect dense tracking data in an efficient way. Several studies in the U.S. and Europe have collected GPS tracking data for transportation related studies (Wolf, Guensler and Bachman 2001, Wolf, et al. 2004, Schönfelder and Axhausen 2004). Although dense GPS tracking data, with explicit spatial and temporal components, can add significant values to activity-based transportation studies, use of GPS tracking data also bring up new research questions and challenges that need to be addressed.

Activity-based Approach and GPS Tracking Data

Unlike the conventional trip-based models, the activity-based approach considers the travel/activity pattern as the basic unit of analysis to help us gain a better understanding of travel behavior (McNally 2000). GPS tracking data provide useful information on *where* and *when* people or vehicles are located. GPS data per se do not tell *what* and *why* people and vehicles are at particular locations during particular time periods. In other words, GPS can help us collect data of movements rather than activities. To analyze activity/travel patterns under the activity-based approach, it is important to also collect or derive additional information such as trip purpose, individual/household characteristics, and constraints/interdependencies of scheduling and performing activities. Consequently, it is important to ask questions such as “what GPS data can and cannot tell us?” and “what kinds of information can we derive from the raw GPS tracking data using other technologies such as geographic information systems (GIS)?” In addition, recent developments of information and communications technologies (ICT), such as the Internet and cell phones, have provided an environment that permits people to expand their activities from physical space (e.g., shopping trips) to virtual space (e.g., e-shopping). The research community has speculated that ICT could lead to important changes in human activity/travel patterns. It therefore is appropriate to ask how GPS tracking data can assist in studies of changing activity/travel patterns due to ICT.

Time Geography, GPS Tracking Data, and GIS

Hägerstrand's time geography suggests an elegant and simple conceptual framework for studying human activities in a space-time context under different types of constraints (Hägerstrand 1970). Time geography is suggested as one of the origins of activity-based approaches (McNally 2000, Jones 2003). With the dense GPS tracking data, time-geographic

concepts such as space-time path and space-time prism can be examined at finer spatial and temporal resolution levels. In addition, dense GPS tracking data offer potentials to analyze the space-time dynamics of activity/travel patterns. Although time geography offers an elegant conceptual framework, performing measurements and analyses under the time-geographic framework can be challenging (Macnab 2000, Shaw and Yu 2004). For example, how do we analyze interactions among numerous space-time paths derived from GPS tracking data? How can we measure and identify spatiotemporal clusters among the space-time paths? How can we analyze relationships between observed behavior (i.e., space-time paths derived from GPS tracking data) and potential behavior (i.e., available activity opportunities located within space-time prisms)? Geographic information systems have been used to represent the concepts embedded in time geography (e.g., Miller 1991, Kwan 2000). In addition to visualization and query of dense GPS tracking data in a GIS environment, we need a framework to help guide the development of GIS analysis functions for activity modeling (Yu and Shaw 2005). For example, how can we better integrate spatial and temporal components in a GIS environment such that it can facilitate analysis of GPS tracking data to study space-time dynamics? What are the analysis functions needed to examine spatio-temporal interactions between activities of the same individuals as well as the interactions of activities between different individuals?

GPS tracking data no doubt offer good potentials to help us gain better understanding of human activity/travel patterns. Nevertheless, this novel data type requires us think carefully what it can and cannot offer and how we can use the data to carry out useful measurements and analyses for activity modeling (on top of other critical issues such as the protection of privacy).

References

- Hägerstrand, T., 1970. What about people in regional science? *Papers of the Regional Science Association*, 24, pp. 1-12.
- Jones, P., 2003. The formation of activity-based approach to travel choice. Presentation at the 30th Anniversary Festschrift seminar, Transport Studies Unit, University of Oxford, October 15, 2003, available at <http://www.tsu.ox.ac.uk/seminars/festschrift/festschrift03.ppt>.
- Kwan, M-P. 2000. Interactive geovisualization of activity-travel patterns using three-dimensional geographical information systems: A methodological exploration with a large data set. *Transportation Research C*, 8,185-203.
- McNally, M.G., 2000. The activity-based approach, Institute of Transportation Studies, University of California, Irvine (UCI-ITS-AS-WP-00-4).
- Miller, H. 1991. Modeling accessibility using space-time prism concepts within geographical information systems. *International Journal of Geographical Information Systems*, 5, 287-301.
- Schönfelder, S. and KW Axhausen, 2004. Structure and innovation of human activity spaces, manuscript for ISTTT 16.
- Shaw, S-L. and H. Yu (2004). An extended time-geographic framework for physical and virtual activities: GIS design and implementation, presentation at the Third STELLA Focus Group 2 Workshop, Budapest, Hungary, April 22-23, 2004, available at

<http://www.stellaproject.org/FocusGroup2/Budapest2004/Presentations/STELLA-FG2-2004%20Shaw>.

Wolf, J., R. Guensler, and W. Bachman, 2001. Elimination of the travel diary: Experiment to derive trip purpose from GPS travel data. *Transportation Research Record* 1768, pp.125-134.

Wolf, J., S. Schönfelder, U. Samaga and K.W. Axhausen, 2004. Eighty weeks of GPS-traces: Approaches to enriching trip information. Presentation at the 83rd Annual Meeting of the Transportation Research Board, Washington, D.C.

Yu, H. and S-L Shaw, 2005. Revisiting Hägerstrand's time-geographic framework for individual activities in the age of instant access, paper submitted to Research Symposium on Societies and Cities in the Age of Instant Access, University of Utah, Salt lake City, November 10-12, 2005.