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.