# Where were you? A time-geographic approach to activity destination re-construction

M. W. Horner[1], J. A. Downs[2]

[1]Florida State University, Tallahassee, FL
Email: mhorner@fsu.edu

[2]University of South Florida, Tampa, FL
Email: jdowns@usf.edu

## 1. Introduction

The trend towards spatially disaggregate modelling and computation has manifested itself across a host of scientific fields, particularly that of transportation (Shaw and Wang 2000). With the proliferation of individual-level datasets describing the detailed activities of households and their constituent members, it is possible to analyze human movements at a high degree of precision (Goodchild 2010). Moreover, these data are becoming better in quality as transportation surveys increasingly take advantage of digital technologies (e.g. GPS receivers) to effectively capture information on people's daily activities at spatially dispersed origin and destination locations (Stopher et al. 2007).

However, despite these advances, travel survey data are not without quality issues (Wolf et al. 2001). First, many daily activity travel surveys still use methods that require a sampled individual's reported origin or destination location to be physically geocoded. In other words, the survey data capture method (e.g. paper survey, telephone interview) does not directly collect the exact spatial coordinate of an activity origin and/or destination (e.g. 123 Pine Street, Sunnyville, FL, 31234). Potential exists for origins and destinations of reported trips not to be geo-referenced, perhaps due to mis-reported information, and/or inconsistencies in spatial address databases, which can limit the usefulness of the survey data. Given an account of trips made by an individual during a typical day, the analyst may be able to confirm that a survey respondent took a particular trip, but may be unable determine the exact origin and/or destination of the trip. From a transportation analysis standpoint, this is a serious problem because it in effect renders the record useless, especially in cases where analyzing the chain of activity locations is of interest (Hensher and Reyes 2000, Horner and O'Kelly 2007).

This paper explores a computational approach for recovering unlocateable activity locations from travel surveys. Derived from recent work in time geography, the method reconstructs the most probable location(s) of missing origins and destinations that were unable to be determined via georeferencing procedures. We adapt a recently developed probabilistic time-geographic approach (Downs 2010) that incorporates individuals' known origin and destination locations, and the time they spent at these (and the unmatched) locations.

## 2. Background and Method

Facilitated by improvements in computational power and geographic information systems (GIS) technology, there has been renewed interest and subsequently many recent developments in the field of time geography. Haagerstrand's pioneering work of the 1970's, which was re-visited and extended in the 1990's by Miller and others, set the foundation for the sustained stream of research that continues today (Haagerstrand 1970, Miller 1991, Kwan 1998, O'Sullivan et al. 2000). Time geographers concern themselves with examining and applying the classical constructs (e.g. space time prisms, cones, geo-ellipses), and research has also proceeded along several other related lines including addressing uncertainty and representations of space, including adapting metrics for use on networks (Kuijpers et al. 2010, Neutens et al. 2011).

One recent area of interest has been in the idea of developing a ‘probabilistic’ time geography (Downs 2010, Winter and Lin 2010, Winter and Yin 2010). In work by Downs (2010), the traditional geo-ellipse representation is improved to visualize likely area(s) a mobile object could have travelled given a time budget. Known as time geographic density estimation (TDGE), the method does not focus solely on the outer polygon depicting the maximum space that could be consumed (i.e., the geo-ellipse), but rather it incorporates a form of data smoothing to interpolate a surface within the polygon showing the most likely places that object could have been located. Of course, this likelihood is determined not only by the individual’s available travel budget but also by their other known points.

Per Downs (2010) the formulation for the TDGE estimator is

$$\hat{f}_t(x) = \frac{1}{(n-1)[(t_n - t_1)v\ ]^2} \sum_{i=1}^{n-1} G\left(\frac{\|x - x_i\| + \|x_j - x\|}{(t_j - t_i)v}\right), \tag{1}$$

where $\hat{f}_t(x)$ is the time geographic density estimate at any point $x$ in a map and $G$ is a distance-weighting function of the geo-ellipse. The number of control points is indicated by $n$, with each point having a time stamp $t$. Sequencing of points is governed by the ordering of subscripts $i$ to $j$. Effectively this formulation in equation 1 fits a distance-weighted geo-ellipse function to each consecutive pair of control points in a space-time path. In this paper we adapt this method to be used with travel survey data, where the interest is in identifying missing origin and destination locations, given other known spatial and temporal information about a respondent’s activity locations.

## 3. Research Structure

We provide a detailed review of time geography, including related developments with respect to uncertainty and probabilistic issues. We also discuss disaggregate travel methods in regards to transportation surveys and activity analysis. From there, we modify the TGDE approach to work with empirically observed travel survey data from a smaller Midwestern U.S. city. Several adaptations to TDGE are suggested, including incorporating a traditional transportation network structure into its estimation (Neutens et al. 2008). We also compare various approaches for re-creating missing survey points by experimenting with alternative weighting functions as well as exploring the whether using more than two known points does a better job of predicting intermediate unknown locations. To get at the idea of ‘better’ we can simulate missing data simply by dropping out a known intermediate point between two other known activity points, use the TDGE technique to ascertain how well we predict the known point. This will act as a calibration procedure to indicate the preferred combination of parameters for applying TDGE to recover missing destination locations.

## 4. References

Downs J, 2010, Time-geographic density estimation for moving point objects. *Lecture Notes in Computer Science* 6292:16-26.

Goodchild M, 2010, Twenty Years of Progress: GIScience in 2010. *Journal of Spatial Information Science* (1):3-20.

Haagerstrand T, 1970, What about people in Regional Science. *Papers of the Regional Science Association* 24:7-21.

Hensher D, and A Reyes, 2000, Trip chaining as a barrier to the propensity to use public transport. *Transportation* 27 (4):341.

Horner M, and M O'Kelly, 2007, Is Non-work Travel Excessive? *Journal of Transport Geography* 15 (6):411-416.

Kuijpers B, H Miller, T Neutens, and W Othman, 2010, Anchor uncertainty and space-time prisms on road networks. *International Journal of Geographical Information Science* 24 (8):1223.

Kwan M, 1998, Space-time and integral measures of individual accessibility: A comparative analysis using a point-based framework. *Geographical Analysis* 30 (3):191-216.

Miller H, 1991, Modelling accessibility using space-time prism concepts within geographical information systems. *International Journal of Geographical Information Science* 5 (3):287-301.

Neutens T, T Schwanen, and F Witlox, 2011, The Prism of Everyday Life: Towards a New Research Agenda for Time Geography. *Transport Reviews: A Transnational Transdisciplinary Journal* 31 (1):25 - 47.

Neutens T, N Van de Weghe, F Witlox, and P De Maeyer, 2008, A three-dimensional network-based space–time prism. *Journal of Geographical Systems* 10 (1):89.

O'Sullivan D, A Morrison, and J Shearer, 2000, Using desktop GIS for the investigation of accessibility by public transport: an isochrone approach. *International Journal of Geographical Information Science* 14 (1):85-104.

Shaw S, and D Wang, 2000, Handling Disaggregate Spatiotemporal Travel Data in GIS. *GeoInformatica* 4 (2):161.

Stopher P, C FitzGerald, and M Xu, 2007, Assessing the accuracy of the Sydney Household Travel Survey with GPS. *Transportation* 34 (6):723.

Winter S, and Z Yin, 2010, The elements of probabilistic time geography. *Forthcoming in GeoInformatica*, DOI 10.1007/s10707-010-0108-1.

Winter S, and Z Yin, 2010, Directed movements in probabilistic time geography. *International Journal of Geographical Information Science* 24 (9):1349 - 1365.

Wolf J, R Guensler, and W Bachman, 2001, Elimination of the Travel Diary: Experiment to Derive Trip Purpose from Global Positioning System Travel Data. *Transportation Research Record* (1768):125-134.