An Automated Activity Identification Method for Passively Collected GPS Data

Vetri Venthan Elango

Randall Guensler

School of Civil and Environmental Engineering

Georgia Institute of Technology

Atlanta, GA

May 10th, 2010

The 3rd Conference on Innovations in Travel Modeling

Overview

- Introduction
- Problem statement and study scope
- Background
- Methodology
- Case study
- Conclusions and future research

Introduction

- Traditional two-day travel diaries have limitations in representativeness, completeness, and accuracy
  - Household travel variability
  - 25% of trips are missed
  - Various reporting accuracy issues
- Cross-sectional travel data have limited usefulness in disaggregate travel behavior analysis
- Longitudinal travel data are necessary for studying variations in travel behavior over time
- GPS devices that collect travel data passively can be excellent instruments for longitudinal surveys

Limitations of GPS Data

- **Limitations of passively collected GPS data include the failure to directly capture:**
  - **Trip purpose**
  - **Travel mode**
  - **Driver identification**
  - **Number of people involved in the activity**

Overview

- Introduction
- Problem statement and study scope
- Background
- Methodology
- Case study
- Conclusions and future research

Problem Statement and Scope

- **Develop a methodology to identify trip end activity with passively-collected GPS data**
- **Use only the household data that would normally be collected during participant recruitment**
- **Use a commercially-available mapping software that is location independent of the GIS data**
- **Assume that data cleaning, and processing of the passively collected GPS data has been completed and individual trips have been identified**

Overview

- Introduction
- Problem statement and study scope
- Background
- Methodology
- Case study
- Conclusions and future research

Longitudinal travel surveys

- Panels of traditional travel diary surveys
- GPS passive data collection
- Examples
  - Puget Sound Panel Study
  - Commute Atlanta Study

Travel Survey Methods using GPS

- **Handheld devices designed to replicate traditional cross-sectional travel diaries**
- **Longitudinal studies with vehicle-based GPS devices installed in participants' vehicles**
- **A hybrid of longitudinal passive GPS data collection coupled with intermittent online travel diary surveys**

Identifying Activity type from GPS data

- **Wolf, et al., 2001**
  - **Elimination of the Travel Diary: Experiment to Derive Trip Purpose from Global Positioning System Travel Data**
- **Schönfelder, et al., 2002**
  - **Exploring the Potentials of Automatically Collected GPS Data for Travel Behaviour Analysis**

Overview

- Introduction
- Problem statement and study scope
- Background
- Methodology
- Case study
- Conclusions and future research

Assumptions

- **Assumptions, based upon passively collected data from instrumented vehicles:**
  - **The radius within which people tend to park their vehicles and walk to a destination is 0.2 miles**
  - **The points of interest that are closest to the trip end are the most likely locations visited by the individual**
  - **The search radius for the home location is 500 feet from the trip end**
  - **The search radius for work and school locations is 1000 feet from the trip end**
  - **If no points of interest within 0.2 mile of the trip end, the trip purpose is classified as 'Unknown'**

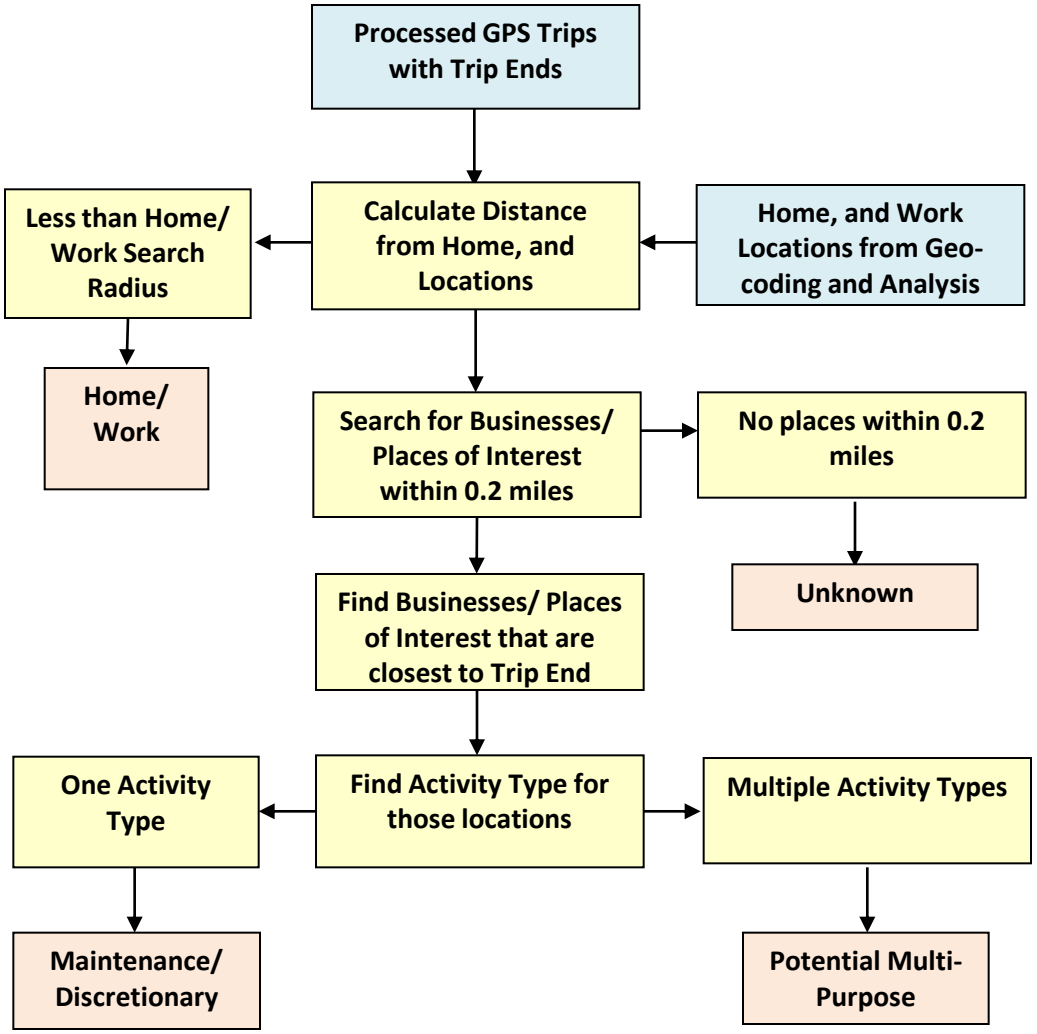
Activity Types

- Home
- Work
- Maintenance (shopping, services, school and dining)
- Discretionary (sports, recreation, social visit, landmarks etc.)
- Multi-purpose (potential)

Coded Points of Interest

- Home
- Work
- School and preschool
- Dining locations
- Shopping and service locations
- Discretionary activity locations
- Locations identified by the household during the recruitment process

Algorithm



Georgia Tech

Overview

- Introduction
- Problem statement and study scope
- Background
- Methodology
- Case study
- Conclusions and future research

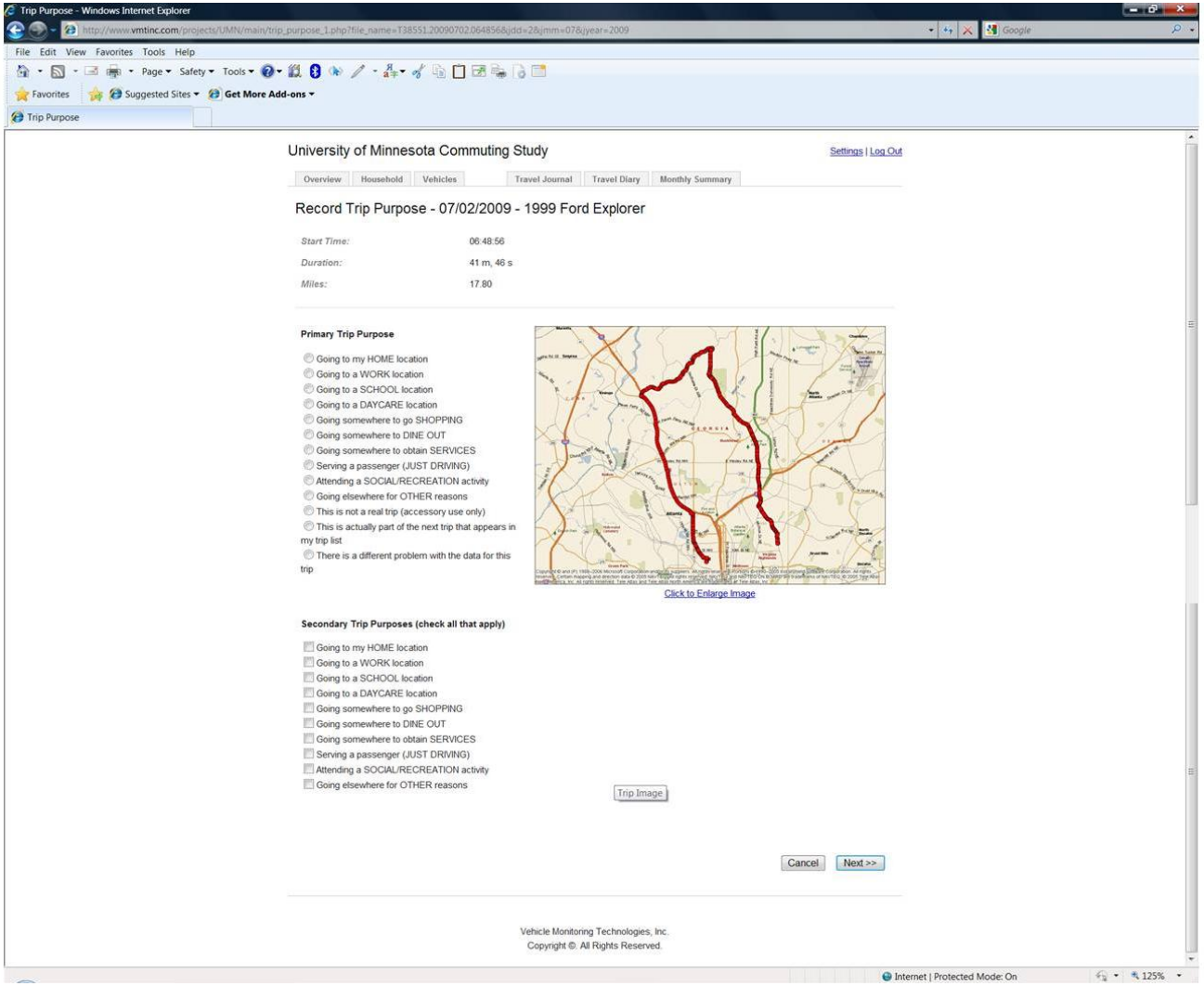
Data

- The University of Minnesota conducted a travel behavior study on the use of the I-35W Bridge
- 46 participants I-35 Bridge commuters
- Participant's commute vehicle instrumented with GPS device September - December 2008
- GPS device transmitted second-by-second data to the server using GPRS/GSM
- Entire household demographics not collected, only participant's information

Data Collection Method

- Hybrid travel survey with passive GPS data and interim online travel diaries
- Each participant was asked to complete 6 to 14 days of travel diary during the study
- The participants completed 94% of received travel diary requests
- Participants voluntarily completed extra travel diary days that were not requested
  - Participants reported trip purpose details for twice as many trips as they were asked to provide data
  - 200% reporting rate

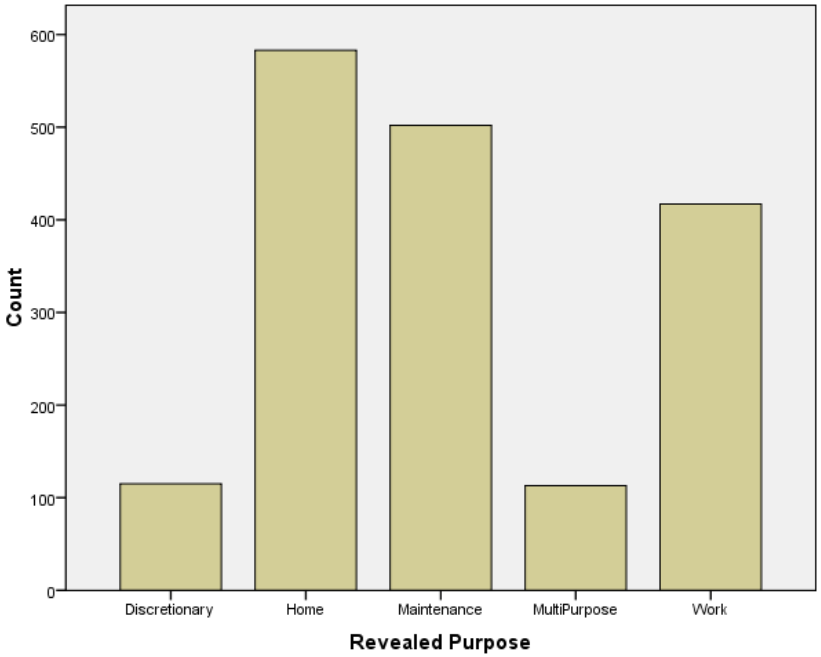
Online Survey

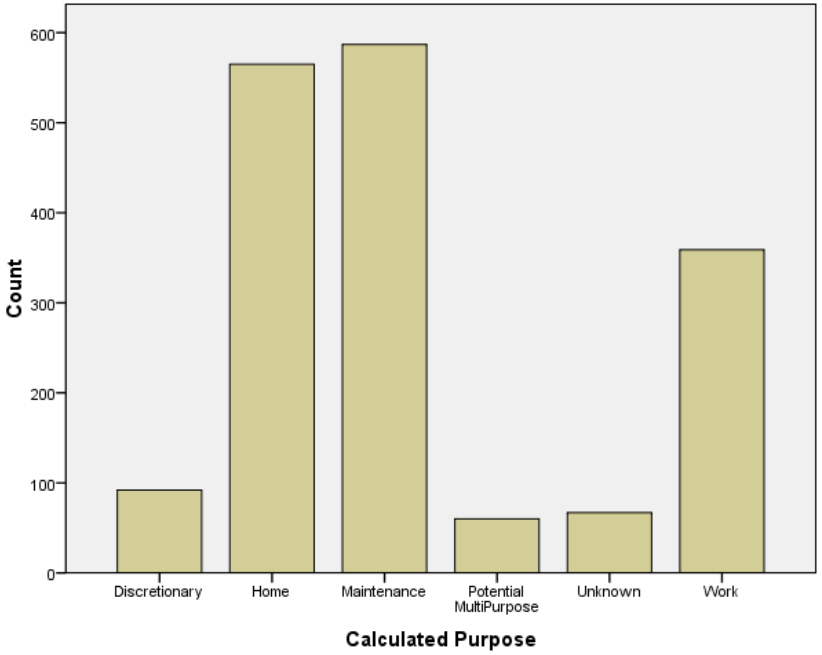


Data QA/QC

- Participants completed diaries for 4300 trips
- The initial data set included the 2185 trips for which travel diary were requested
- Some data were eliminated:
  - 150 trips with GPS data issues
  - 250 trips with trip purpose coded as "Other"
  - One household that reported using their vehicle for commercial purposes
  - One household that clearly reported frivolous data
- Retained 1730 trips in the analytical data set

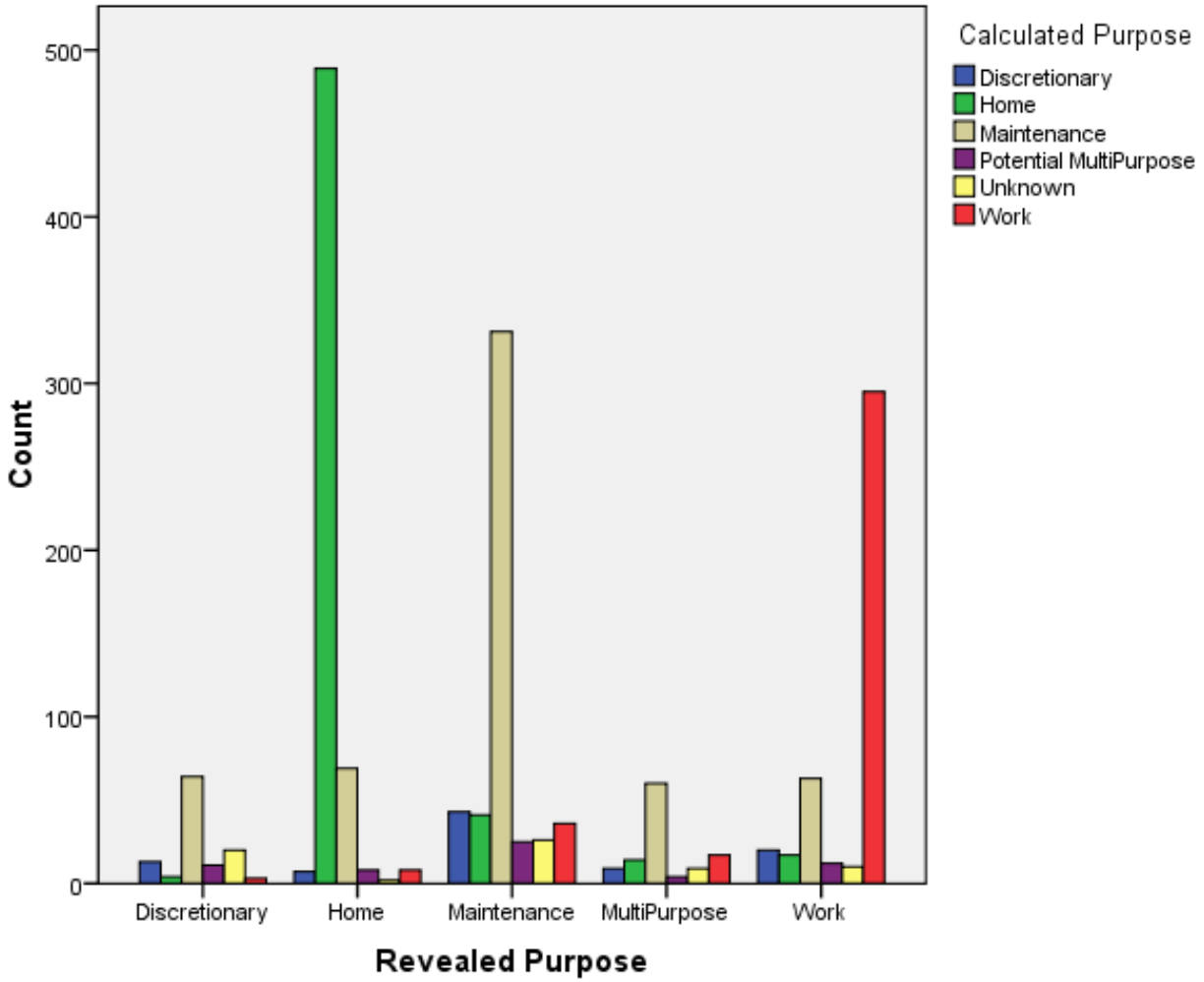
Results





Georgia Tech

Results



Georgia Tech

Results

- 67 trips (4%) fall under the 'Unknown' category
- 65.4% of the trips were identified correctly
- Home activity predicted with 84% accuracy
- Work activity identified with 71% accuracy
- Maintenance activity identified with 66% accuracy
- Discretionary and potential multi-purpose activities are predicted poorly

Discussion

- It is assumed that stated trip purpose is correct
  - This is not always true
  - One household reported 3 trips starting at 16:04, 16:18 and 16:50 as trips to home
- MapPoint 2006 was used in this analysis and it may not have latest information on points of interest
- Complete household information, such as other work locations, schools, friend's homes, etc., will further improve prediction accuracy

Overview

- Introduction
- Problem statement and study scope
- Background
- Methodology
- Case study
- Conclusions and future research

Conclusions

- The proposed methodology does not require human interaction to identify the activity type
- With commercial mapping software, the methodology is applicable anywhere in the US
- Eliminates variability in land use data formats used by different organizations
- Assumptions should be re-evaluated with hand-held-GPS data
- The case study found the methodology to be 65% accurate
- Researchers noted that the stated purposes are not always 'ground truth'

Next Steps

- Incorporate activity duration, time-of-day, and day-of-week into the algorithms being used to identify trip activity
  - e.g., commercial activity occurs during business hours
- Incorporate learning algorithms that will use two-day travel diary data to automatically predict the activities that occur on other days

Cross-tabulation
Revealed Activity vs. Calculated Activity

Revealed Purpose	Calculated Purpose						
	Discretionary	Home	Maintenance	Potential Multi-Purpose	Unknown	Work	Total
Discretionary	13	4	64	11	20	3	115
Home	7	489	69	8	2	8	583
Maintenance	43	41	331	25	26	36	502
Multi-Purpose	9	14	60	4	9	17	113
Work	20	17	63	12	10	295	417
Total	92	565	587	60	67	359	1730

References

- Elango, V. V., R. Guensler and J. H. Ogle Day-to-Day Travel Variability in the Commute Atlanta, Georgia, Study. Transportation Research Record, 2007.
- Wolf, J, Oliveira M, and Thompson M, Impact of Underreporting on Mileage and Travel Time estimates. Transportation Research Record, 2003
- Yanzhi Xu and Randall L. Guensler, Advantages of Long-Term Continuous GPS-Based Survey Data For Activity-Based Travel Demand Modeling, Manuscript submitted to the 89th Transportation Research Board Annual Meeting, 2010
- Pendyala, R. M. and E. I. PAS Multi-Day and Multi-Period Data for Travel Demand Analysis and Modeling. Transportation Research Board, 2000.
- Thomas F. Golob, R. K., Lyn Long. Chapter 6 - Puget Sound Transportation Panel. In Panels for Transportation Planning, Kluwer Academic Publishers, 1997.
- Ogle, J., R. Guensler and V. Elango Georgia's Commute Atlanta Value Pricing Program: Recruitment Methods and Travel Diary Response Rates. Transportation Research Board, 2005.

References

- Xu, Y., L. I. Zuyeva, D. Kall, V. V. Elango and R. Guensler Mileage-Based Value Pricing: Phase II Case Study Implications of Commute Atlanta Project. Transportation Research Board, 2009.
- Schönfelder Stefan, A., Kay W. Antille Nicolas, Bierlaire Michel. Exploring the Potentials of Automatically Collected Gps Data for Travel Behaviour Analysis. ETH, E. T. H. Z., Institut für Verkehrsplanung, Transporttechnik, Strassen- und Eisenbahnbau IVT, 2002.
- Doherty ST, Noel N, Gosselin M-L, Sirois C, Ueno M, Moving Beyond Observed Outcomes, Integrating GPS and Interactive Comuter Based Travey Behavior Surveys, Personal Travel: The Long and Short of It, Transportation Research Board, 2001.
- Wolf, J., R. Guensler and W. Bachman Elimination of the Travel Diary: Experiment to Derive Trip Purpose from Global Positioning System Travel Data. Transportation Research Board, 2001.