PRELIMINARY REVIEW COPY

1. Report No. 0-5176-2 Preliminary Review Copy	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Conversion of Volunteer-Collected GPS Diary Data into Travel Time Performance Measures: Algorithm for Extracting Travel-Diary Data from GPS Streams and GPS-TDG Software Design		5. Report Date August 31, 2005
7. Author(s) Sivaramakrishnan Srinivasan, Prabuddha Ghosh, Stacey Bricka, and Chandra R. Bhat		6. Performing Organization Code
		8. Performing Organization Report No. 0-5176-2
9. Performing Organization Name and Address Center for Transportation Research The University of Texas at Austin 3208 Red River, Suite 200 Austin, TX 78705-2650		10. Work Unit No. (TRAIS)
		11. Contract or Grant No. 0-5176
12. Sponsoring Agency Name and Address Texas Department of Transportation Research and Technology Transfer Section/Construction Division P.O. Box 5080 Austin, TX 78763-5080		13. Type of Report and Period Covered Research Report (1/1/05–8/31/05)
		14. Sponsoring Agency Code
15. Supplementary Notes Project conducted in co-operation with the Federal Highway Administration.		

16. Abstract

Conventional travel-survey methodologies require the collection of detailed activity-travel information, which imposes a significant burden on respondents, thereby adversely impacting the quality and quantity of data obtained. Advances in the global positioning system (GPS) technology have provided transportation planners with an alternative and powerful tool for more accurate travel-data collection with minimal user burden. The data recorded by GPS devices, however, do not directly yield travel information; the navigational streams recorded by GPS devices have to be processed and the travel patterns derived from them. The focus of this research project is to develop software to automate the processing of raw GPS data and to generate outputs of activity-travel patterns in the conventional travel-diary format. The software will identify trips and characterize them by several attributes, including trip-end locations, trip purpose, time of day, distance, and speed. Within the overall focus of the research, this report describes the algorithm developed for the detection and characterization of trips from GPS navigational streams. The software design details are also presented.

17. Key Words Household travel surveys, global positioning system (GPS), GPS-based travel surveys, GPS data recording formats, processing GPS navigational data	18. Distribution Statement No restrictions. This document is available to the public through the National Technical Information Service, Springfield, Virginia 22161.		
19. Security Classif. (of report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of pages 38	22. Price

Form DOT F 1700.7 (8-72)   Reproduction of completed page authorized

CONVERSION OF VOLUNTEER-COLLECTED GPS DIARY DATA INTO TRAVEL TIME PERFORMANCE MEASURES:

ALGORITHM FOR EXTRACTING TRAVEL-DIARY DATA FROM GPS STREAMS AND GPS-TDG SOFTWARE DESIGN

Sivaramakrishnan Srinivasan

Prabuddha Ghosh

Stacey Bricka

Chandra R. Bhat

Research Report 5176-2

PRELIMINARY REVIEW COPY

Research Project 0-5176

Conversion of Volunteer-Collected GPS Diary Data into Travel Time Performance Measures

Conducted for the

TEXAS DEPARTMENT OF TRANSPORTATION

in cooperation with the

U.S. DEPARTMENT OF TRANSPORTATION
Federal Highway Administration

by the

CENTER FOR TRANSPORTATION RESEARCH
THE UNIVERSITY OF TEXAS AT AUSTIN

August 2005

DISCLAIMERS

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the Federal Highway Administration or the Texas Department of Transportation. This report does not constitute a standard, specification, or regulation.

There was no invention or discovery conceived or first actually reduced to practice in the course of or under this contract, including any art, method, process, machine, manufacture, design or composition of matter, or any new and useful improvement thereof, or any variety of plant, which is or may be patentable under the patent laws of the United States of America or any foreign country.

NOT INTENDED FOR CONSTRUCTION,
BIDDING, OR PERMIT PURPOSES

Chandra R. Bhat
Research Supervisor

ACKNOWLEDGMENTS

TABLE OF CONTENTS

CHAPTER 1 INTRODUCTION

For nearly fifty years, household travel surveys have been used to document the travel behavior of regional households as part of long-range transportation planning efforts. The survey data are used for general planning and policy analysis, as well as to serve as the foundation for regional travel demand models. Technology advancements have resulted in changes in household travel survey data-collection procedures, the most recent being the introduction of global positioning systems (GPSs) to record travel patterns. In this procedure, travel data are collected by equipping the respondents' automobiles with GPS receivers and recording the position and velocity of the vehicles at set intervals of 1–5 seconds. In addition to the use of in-vehicle GPS devices for travel data collection (which focus on vehicle travel specifically), newer studies have developed and deployed wearable, personal, or handheld GPS units in travel surveys to collect data on personal travel using any mode of travel. A comprehensive synthesis of the state of practice in GPS-enhanced travel surveys was presented in Research Report 1.

In the current TxDOT-funded research project, our focus is on the use of passive in-vehicle GPS devices in which the user intervention is limited to, at most, turning the device on and off. Such passive in-vehicle GPS-based travel surveys offer several advantages over conventional data-collection approaches: (1) Data collection is passive, thereby minimizing respondent burden; (2) The locations of activity and travel are automatically determined to great spatial accuracy, that is, to the level of latitude and longitude; (3) The GPS-recorded timing data are more accurate than the self-reported estimates; (4) Travel route information is automatically collected; (5) Trip speeds are recorded as actual observations rather than being calculated as part of the post-collection processing; and (6) Trip distances can be computed accurately using the detailed position data along the length of the trip.

Given the advantages of the GPS survey approach, it offers substantial promise for the development of trip data to support modeling. However, there are several issues that need to be addressed for the effective use of this new technology for household travel data collection. These include: (1) The GPS data are collected in the form of navigational streams (i.e., recordings of position and velocity at set intervals) and require substantial processing to convert these streams into the conventional travel-diary format for subsequent use for modeling purposes; (2) Equipment specifications and operational shortcomings impact the quality and quantity of the

GPS data collected; (3) The trip purpose information is unknown; this needs to be elicited from the respondent directly or imputed using the GPS data in conjunction with supplemental land-use and network data; (4) The vehicle occupancy levels are unknown, and the driver and the passengers in the vehicle cannot be identified; such information, if needed, has to be elicited from the respondents directly; and (5) If in-vehicle GPS devices are used, the derived trip diary is an accurate record of the sequence of vehicle trips and not person trips.

In summary, the GPS technology shows promise to minimize costs, while maximizing the volume and accuracy of travel data collected. On the other hand, it is also evident that the use of passive GPS devices for data collection shifts considerable burden from the respondent to the analyst. Therefore, the success of this new technology as a travel data collection instrument depends on the ability of the analyst to derive meaningful trip information from the navigational data streams of GPS devices. Research Report No. 1 documented several issues to be addressed in extracting trip-diary information from GPS navigational streams and discussed the limitations of the procedures developed by other studies.

In the current TxDOT-funded research, the objective is to develop a prototype software tool, GPS-Based Travel-Diary Generator (GPS-TDG), which automates the process of converting navigational data streams collected passively from in-vehicle GPS devices into an electronic activity-travel diary. The research is motivated by a need to design robust/efficient algorithms and software for analyzing GPS data streams from several hundreds of vehicles over multiple days.

Within the overall objective identified above, this report describes the automated procedure developed for extracting the trip-diary information from GPS navigational streams and other supplemental data. The rest of this report is organized as follows. Chapter 2 presents a conceptual overview of the software and also discusses the assumptions made in software development. Chapter 3 presents the detailed travel-diary extraction algorithm. Chapter 4 discusses the software implementation details. Chapter 5 presents a brief summary of this report.

CHAPTER 2 GPS-TDG: CONCEPTUAL OVERVIEW

A conceptual overview of the GPS-TDG is presented in Figure 2.1. The software is designed to operate either in a basic analysis mode or in an enhanced analysis mode. The basic mode is designed to convert the GPS data into a simple trip file that distinguishes among home-based work, home-based other, and non-home-based trips. The enhanced mode utilizes additional data to derive a more refined trip purpose classification taxonomy. As a result, the basic analysis mode is less demanding in terms of input data compared to the enhanced analysis mode.

There are four input data components that are common to both the basic and enhanced analysis modes, as indicated in Figure 2.1. In addition, the enhanced analysis mode requires three additional inputs. Section 2.1 of this report discusses all software inputs. Among these inputs, the GPS streams require pre-processing to be converted into a format that can be readily used by the software. This pre-processing step is described in Section 2.2. Section 2.3 focuses on the main processing module of the GPS-TDG software. This is the core of the software that implements the travel-diary extraction algorithm and creates the trip files. The primary output of the GPS-TDG software is the derived travel diary, which is discussed in Section 2.4. This travel-diary data can then be easily aggregated to compute several inter-zonal highway performance measures, such as average travel time, speed, and distance, in the overall, and by trip purpose and time of day. The final software product will support the computation of the above-mentioned measures. The implementation details of these computations are, however, not the focus of this report.

Inputs for both basic and advanced analysis

Additional inputs for enhanced analysis

Raw GPS navigational streams	TAZ boundaries (GIS layer)	Home and work locations of primary drivers	Algorithm parameters

- Land-use data (GIS layer)
- Individual and HH demographic characteristics of primary drivers
- Activity-type MNL model specification

Pre-Processor

Processed GPS navigational streams

GPS-TDG Process Module
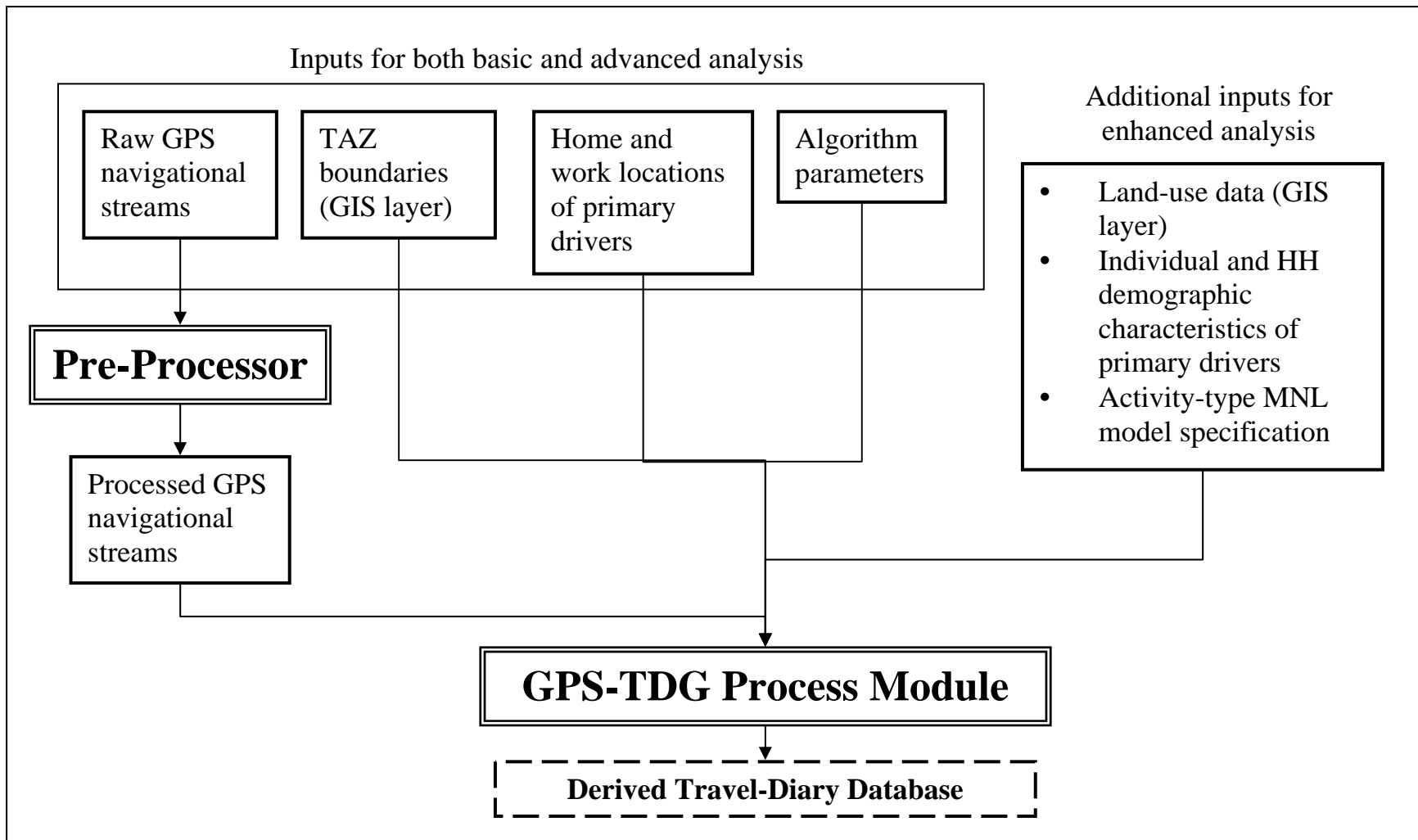
Derived Travel-Diary Database

Figure 2.1 Conceptual Overview of the GPS-TDG

2.1. Inputs

As discussed, there are four main inputs that are required for both the basic and enhanced analysis modes. These are

1. The raw GPS navigational streams in the GPRMC format. The reader is referred to Research Report 1 for a detailed description of the GPRMC format.

2. The boundaries of the traffic analysis zones (TAZs) as a geographic information system (GIS) layer.

3. The home and work (if employed) locations (as latitudes and longitudes) of the primary driver for each household vehicle. In the development of the GPS-TDG algorithms, it is assumed that each GPS-equipped vehicle is used predominantly by one person and this person is designated as the primary driver of the corresponding vehicle.

4. Parameters (such as dwell-time threshold) to control the travel-diary extraction algorithm. These parameters enable the analysts to better customize the software to different study regions and for different study objectives.

In order to run the software in the enhanced analysis mode, three additional inputs are required. These are:

1. Land-use or zoning data. These data are provided as a GIS layer in which the study area is divided into a number of land-use parcels, and each parcel is assigned a specific land-use type.

2. Individual and household demographic characteristics (such as gender, employment/student status, and number of children in the household) of the primary driver associated with each GPS equipped vehicle.

3. Specifications for the multinomial logit (MNL) model for a disaggregate activity-type classification. This includes the activity type classification, the list of explanatory variables, and the values of the model coefficients on these explanatory variables.

2.2. Pre-Processor

The GPS units record vehicle movement at 1–5 second intervals. Each of these records contains the vehicle's position in terms of latitude and longitude, as well as velocity and other

position indicators. In addition, each record also contains a flag to indicate whether sufficient detail is present to consider the record "valid" for locational identification purposes. In the pre-processing step the invalid records are removed, and the number of invalid records removed immediately prior to each valid record is recorded. The valid records are then restructured into a format that is readily usable for trip-diary extraction processing. Specifically, the latitude and longitude are converted into decimal degrees, the universal time coordinates (UTC) dates and times are converted to the local dates and times, and speeds are converted to metric units from knots.

2.3. Process Module

Once the data have been pre-processed, the next step is to create the trip file. This is accomplished through the process module, which forms the core of the GPS-TDG software and implements the travel-diary extraction algorithm. As discussed in detail in Research Report 1, operational characteristics (user-flagged versus purely passive systems), data-collection protocols (switched power versus continuous power systems), and data logging rules (such as speed-checked data recordings) define the structure of the GPS data streams recorded. This in turn impacts the processing methods required for travel-diary extraction. Our software is being developed primarily for purely passive systems (i.e., the driver does not manually "flag" the end of trips) that are switched-power (i.e., the recording of GPS data stops when the engine is powered off), and without any data logging rules (i.e., data points are recorded continuously as long as the engine is powered-on, irrespective of the travel speed or other characteristics).

It is also useful to note here that the objective of this software is to completely automate the travel-diary extraction procedure. However, identification of the network links actually traversed by the vehicle and hence the determination of travel route is not within the scope of this software.

2.4. Derived Travel Diary

The fundamental output of the GPS-TDG software is the derived travel diary. This output contains the trips extracted from the GPS streams for each of the equipped vehicles. Each trip is characterized by the following attributes: trip timing, duration of activity at trip destination, trip-

end locations (latitude/longitude and TAZ), trip distance, average trip speed, measure of variation in speed during the trip, activity purpose at trip-end location (aggregate classification scheme in the case of basic analysis and a disaggregate classification scheme in the case of enhanced analysis), and trip purpose. In addition accuracy measures are also generated to capture the impacts of signal loss or equipment malfunction on the identification and characterization of trips.

CHAPTER 3 TRAVEL-DIARY EXTRACTION ALGORITHM

This chapter presents a detailed description of the methodology developed to process the GPS navigational streams to derive the travel-diary data. Figure 3.1 presents a conceptual overview of the travel-diary extraction algorithm. The first step of this algorithm accepts all required inputs. These inputs include the pre-processed GPS navigational streams, household and individual demographic data, a GIS layer of the TAZ boundaries, and other optional GIS land-use data. In addition, the software allows for the analyst to specify several parameters used for trip detection and characterization. A complete description of all inputs is provided in Section 3.1.

The algorithm is designed to extract the trip-diary information for one vehicle at a time. Correspondingly, all data for a single vehicle are read in (Step 2), and Steps 3, 4, and 5 are performed on this data. Step 3, described in detail in Section 3.2, is trip detection, which involves scanning through the GPS stream for the vehicle under consideration until a potential trip end (or stop) is detected. In Step 4 the trip characteristics such as start and end times, distance, and average speed are computed for the potential trip detected in Step 3. A complete list of all trip attributes is computed, and the corresponding methods are presented in Section 3.3. In Step 5, the reasonableness of the trip attributes determined in the previous step is examined by comparing it to user-provided thresholds. A potential trip that fails these checks (described in Section 3.4) is classified as a "false" trip. On the other hand, if the potential trip end is found to be reasonable, the trip characteristics are written out to the output file. If the end of the GPS stream has not been reached, the processing proceeds by repeating Steps 3 through 5 until the entire GPS stream has been processed. This completes the trip-diary extraction for the vehicle under consideration. The data for the next vehicle in the input are then read in and the Steps 3, 4, and 5 are executed in a loop until the diary data have been generated for all vehicles.
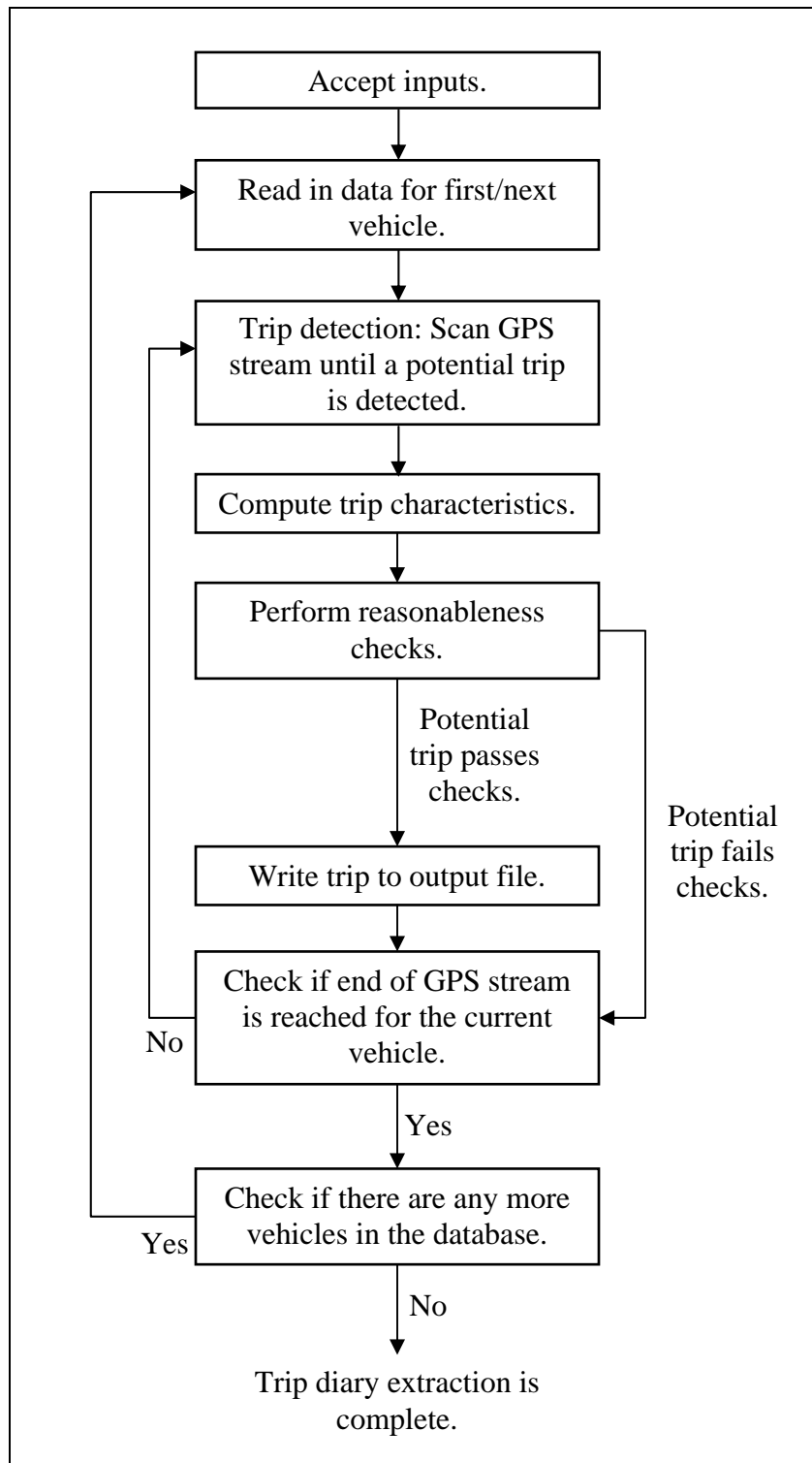
```
                    ┌─────────────────────────┐
                    │      Accept inputs.      │
                    └────────────┬────────────┘
                                 │
                                 ▼
         ┌───────────▶┌─────────────────────────┐
         │            │ Read in data for first/next │
         │            │        vehicle.          │
         │            └────────────┬────────────┘
         │                         │
         │                         ▼
         │   ┌────────▶┌─────────────────────────┐
         │   │         │ Trip detection: Scan GPS │
         │   │         │ stream until a potential trip │
         │   │         │     is detected.         │
         │   │         └────────────┬────────────┘
         │   │                      │
         │   │                      ▼
         │   │         ┌─────────────────────────┐
         │   │         │ Compute trip characteristics. │
         │   │         └────────────┬────────────┘
         │   │                      │
         │   │                      ▼
         │   │         ┌─────────────────────────┐
         │   │         │  Perform reasonableness  │──┐
         │   │         │        checks.           │  │
         │   │         └────────────┬────────────┘  │
         │   │              Potential             │  │
         │   │            trip passes             │  │ Potential
         │   │              checks.               │  │ trip fails
         │   │                      ▼              │  │ checks.
         │   │         ┌─────────────────────────┐ │
         │   │         │ Write trip to output file. │ │
         │   │         └────────────┬────────────┘ │
         │   │                      │              │
         │   │                      ▼              │
         │   └─────────┌─────────────────────────┐◀┘
         │        No   │  Check if end of GPS stream │
         │             │  is reached for the current │
         │             │        vehicle.          │
         │             └────────────┬────────────┘
         │                        Yes
         │                          ▼
         │             ┌─────────────────────────┐
         └─────────────│  Check if there are any more │
              Yes      │  vehicles in the database. │
                       └────────────┬────────────┘
                                  No
                                    ▼
                       ┌─────────────────────────┐
                       │  Trip diary extraction is │
                       │        complete.         │
                       └─────────────────────────┘
```

Figure 3.1 Conceptual Overview of the Travel-Diary Extraction Algorithm

3.1. Step 1: Accept Inputs

The inputs to the GPS-TDG software are:

1. The pre-processed GPS navigational streams. Each record (or line) of this file contains the following information: household and vehicle identifiers, local date and time, latitude, longitude, speed, heading, and the number of invalid records immediately prior to this record in the raw file.

2. The demographics data file identifying the individual and household characteristics of the primary drivers of the GPS-equipped vehicles. If only the basic analysis is to be performed, the data required are only the home and work locations. If enhanced analysis is to be performed, all demographic data required to predict activity types using the activity type classification model are included. Such additional demographic data required in the context of the Laredo application are employment and student status of the individual, gender, and the number of children in the household.

3. A GIS layer identifying the TAZ boundaries.

4. A GIS land-use layer (required only for enhanced analysis). Locally available parcel-level data with a rich land-use classification scheme are suitable. Alternately, the land-use information provided in the Census TIGER files can also be used. The Laredo application of the software uses parcel-level zoning data.

5. Specification of the MNL model for activity type classification (required only for enhanced analysis). This includes the activity type classification scheme employed, the list of explanatory variables to predict the disaggregate activity purpose for each trip, and the vector of model parameters. See Appendix A for the MNL model specification for the Laredo application.

6. Algorithm parameters. The trip detection and characterization algorithm implemented in the GPS-TDG software are controlled by several user-specified parameters. The values of these parameters can be varied as required by the analyst to suit the requirements of different study areas and for various analysis objectives. (The final software product will also include built-in default values for all these parameters.) A description of all the user-defined parameters is presented in Table 3.1 below. The use of these thresholds in the trip detection and characterization algorithm is discussed in the next section.

Table 3.1 List of User-Specified Algorithm Parameters

Parameter	Description
Enhanced Analysis Flag	Determines whether the enhanced analysis components of the algorithm, such as disaggregate activity purpose determination are implemented or not
Engine Power-off Dwell-Time Threshold	Minimum dwell-time gap between successive valid GPS records for signaling an engine power-off stop
Non-Engine Power-off Dwell-Time Threshold	Minimum duration for which the speed should be less than the speed threshold for signaling a non-engine power-off stop
Speed Threshold	Value of instantaneous speed below which the vehicle is assumed to be at rest
Home Location Distance Threshold	Maximum distance between a trip-end location and home for classifying the trip purpose as "Home"
Work Location Distance Threshold	Maximum distance between a trip-end location and work for classifying the trip purpose as "Work"
Work Duration Threshold	Minimum activity duration at a trip end for classifying a trip end as "Work"
Average Update Rate	Average duration between successive navigational points recorded by the GPS instrument used in the survey
Trip Distance Computation Time Threshold	Minimum time between pairs of GPS points for which the distances are computed and summed to determine the trip length
Minimum Trip Duration Threshold	Minimum trip duration for a potential trip to be classified as a real trip
Minimum Trip Speed Threshold	Minimum trip average speed for a potential trip to be classified as a real trip

3.2. Step 2: Detect Potential Trip Ends

The detection of a potential trip involves identifying a pair of GPS records within the overall navigational stream, with one record corresponding to the end of a trip and the other

corresponding to the start of the next trip. The first valid GPS record in the stream corresponds to the start of the first trip. Based on prior GPS trip research, there are two types of trips to be detected: those involving an engine power-off (parking at a location for work, shopping, at home, etc.) and those in which the engine is not turned off (dropping someone off at a location, going through a drive-through bank or fast food location, etc.). The detection procedure for each of these cases is discussed below. All user-specified algorithm parameters referenced henceforth are underlined.

3.2.1. The engine power-off case

In the engine power-off case, a pair of successive GPS records, one corresponding to the end of a trip and the next corresponding to the start of a subsequent trip are detected by examining the dwell times between successive valid records in the GPS stream. The dwell time is computed by subtracting the time due to signal loss from the total time between successive valid GPS records. Thus, the procedure guards against classifying time gaps due to extended periods of signal loss as trip ends. The procedure is as follows:

1. Compute the total time difference (*TotTimeDiff*) between successive GPS records as the difference in the time stamps of the two records.

2. Compute the signal loss time (*SigLossTime*) between the successive GPS records as the product of the number of invalid records removed (during pre-processing) immediately prior to the second GPS record and the Average Update Rate. The reader is referred to Section 3.1 for the discussion on preprocessing and the removal of invalid records.

3. Compute the Dwell Time (*DwellTime*) as the difference between *TotTimeDiff* and *SigLossTime*.

4. If the *DwellTime* between a pair of successive GPS records exceeds the Engine Power-off Dwell-Time Threshold, a potential trip end is flagged. The first of these two successive GPS records corresponds to the end of a trip and the second corresponds to the start of the next trip.

In addition, a potential trip end is also flagged at the end of the GPS stream for the vehicle, and the last GPS record represents the trip end of this last trip.

3.2.2. The non-engine power-off case

Stops without engine power-off cannot be detected using the dwell time concept, because the GPS points are continually recorded, even during the period when the vehicle is stopped. Hence, we examine the speed data to identify non-movement and a non-engine power-off trip end. Specifically, if the instantaneous speed recorded in the GPS stream is less than the Speed Threshold continuously for a period greater than the Non-Engine Power-off Dwell-Time Threshold, a potential trip end is detected. The GPS record for which the speed continually remains below the threshold value is taken to represent the end of a trip. The first subsequent GPS record with a speed above the threshold value represents the start of the next trip.

3.3. Step 3: Compute Trip Attributes

Once a potential trip end is detected using the procedures described in the previous section, several trip attributes are computed. These attributes are (1) trip start and end times, (2) trip-end locations, (3) trip duration and length, (4) trip speed, (5) activity duration at trip destination, (6) activity type at trip destination, and (7) accuracy flags. The methods for computing each of these attributes are discussed below.

3.3.1. Trip start and end times

The start time of a trip is computed as the time stamp on the first GPS navigational stream record corresponding to a trip and the end time of a trip is computed as the time stamp on the last GPS navigational stream record corresponding to the same trip.

3.3.2. Trip-end locations

The position information (i.e., latitude and longitude) on the first and last GPS records of a trip determines the most detailed trip-end locations. In addition, the trip-end locations are also specified in terms of TAZ. Specifically, the latitude and longitude of the trip destination-end is overlaid on the TAZ boundaries GIS layer to determine the TAZ in which the trip destination lies, and a spatial "join" procedure is invoked. The TAZ layer of the trip origin is simply

determined as the TAZ of the destination of the previous trip. The only exception to this is in the case of the first trip, where the trip origin TAZ is determined using the overlay procedure.

3.3.3. Trip duration and length

The trip duration is computed as the difference between the start and end times of the trip. The trip length (or distance) is determined using the point-to-point sum of distances approach (the PP approach; see Research Report 1 for a detailed discussion on alternate methods). Broadly, this method involves computing the distance between successive pairs of recorded locations, the two points in each pair being spaced apart by at least the <u>Trip Distance Computation Time Threshold.</u> The reader will note that, by computing the distances between GPS points which are 5 or 10 seconds (the <u>Trip Distance Computation Time Threshold</u>) apart rather than 1 second apart, it is possible to minimize the over-estimation of the trip distance due to the positional errors associated with the GPS devices (see detailed discussion in Research Report 1). These distances are then summed over the entire trip to obtain the trip length. Further, the summing of the distances is not performed over segments of the trip where the speed is less than the <u>Speed Threshold</u> (i.e., short stretches when the vehicle is not moving).

Further, two different methods are used to compute the distance between pairs of locations in the PP approach. In the first method the latitude and longitude information for the two points are used to determine the distance using the following formula:

$$D = \frac{1}{360 \cdot 2\pi \cdot R} \arccos[K]$$
$$K = \cos(a_1)\cos(b_1)\cos(a_2)\cos(b_2) + \cos(a_1)\sin(b_1)\cos(a_2)\sin(b_2) + \sin(a_1)\sin(a_2)$$

where

a_1 and a_2 represent, respectively, the latitude and longitude of the first point (in decimal degrees),

b_1 and b_2 represent, respectively, the latitude and longitude of the second point (in decimal degrees), R is the radius of the earth (3959 miles), and D is the distance between the two points (in miles).

In the second method the distance is computed as the product of the recorded instantaneous speed and the time between the pair of GPS points. Thus, the software provides two estimates of the trip distance, one based on position information and the other based on

speed information collected by the GPS devices. If actual trip distance measurements were also available (say from odometer readings) along with GPS streams, these two methods can be validated to determine the better method.

3.3.4. Trip speed

The instantaneous speeds are averaged over all the GPS records corresponding to a trip to compute the average trip speed. The standard deviation of the instantaneous speed measurements is also computed to provide a measure of variation in speed along the trip length.

3.3.5. Activity duration

The duration of activity at a trip end is computed as the difference between the end time of the current trip and the start time of the next trip. The reader will note that the activity duration at the origin of the first trip and at the destination of the last trip cannot be determined.

3.3.6. Activity type at trip destination and trip purpose

The activity type undertaken by the driver of the vehicle at the trip-end location is determined next. This varies according to whether or not the user is in the basic or enhanced mode of the software.

In the basic analysis mode, the trip-end activities are classified into one of the following three aggregate types: home, work, and other. As already discussed in Section 3.1, the locations of home and work in terms of latitude and longitude are provided as inputs to the algorithm. A trip-end activity is classified as "home" if the distance of the trip-end location from home is less than the Home Location Distance Threshold (This is necessary to account for the difference between where the vehicle is parked and where the home is located.) A trip-end activity is classified as "work" if (1) the distance of the trip-end location from work is less than the Work Location Distance Threshold, and (2) the activity duration at the trip end is greater than the Work Duration Threshold. (Again, there is a need to consider vehicles being parked in garages and lots off-site from the main work location.) Trip ends which are not classified as home or work are classified as "other" by default.

The activity types at the origin and the destination of the trip are then used to determine the trip purpose using the following rules:

1. If the activity type at one of the trip ends (i.e., origin or destination) is "home" and the activity type at the other trip end is "work," then the trip purpose is "home-based work" (HBW).

2. If the activity type at one of the trip ends (i.e., origin or destination) is "home" and the activity type at the other trip end is not "work," then the trip purpose is "home-based non-work" (HBNW).

3. If the activity type at both trip ends is not "home," then the trip purpose is "non-home based" (NHB).

In the enhanced analysis mode, the trip-end activities classified as "other" are further classified into disaggregate types if land-use data and the explanatory variables and parameters of the multinomial logit model for the activity type classification are also provided as inputs. For ease in presentation, the probabilistic procedure for disaggregate activity type determination is described here in the context of three activity types: shopping, leisure, and serve-passenger. This procedure can be extended in a straightforward manner to any number of activity types.

1. Let X_S, X_L, and X_{SP} be the values of the explanatory variables corresponding to shopping, leisure, and serve-passenger respectively.

2. Let β be the vector of coefficients (i.e., model parameters) estimated for the above explanatory variables.

3. Compute the probability that the current trip-end activity type is shopping as

$$\Pr(Shop) = \frac{\exp(\beta X_S)}{\exp(\beta X_S) + \exp(\beta X_L) + \exp(\beta X_{SP})}.$$

4. Similarly, compute the probability that the current trip-end activity type is leisure and serve-passenger, respectively, as

$$\Pr(Leisure) = \frac{\exp(\beta X_L)}{\exp(\beta X_S) + \exp(\beta X_L) + \exp(\beta X_{SP})}$$

$$\Pr(ServePax) = \frac{\exp(\beta X_{SP})}{\exp(\beta X_S) + \exp(\beta X_L) + \exp(\beta X_{SP})}$$

5. Draw a random number, U, from the uniform [0,1] distribution.

6. The assignment of the trip-end activity type is accomplished as follows:

   a. If $0 \leq U \leq \Pr(Shop)$, assign shopping as the activity type.

   b. If $\Pr(Shop) \leq U \leq \big(\Pr(Shop) + \Pr(Leisure)\big)$, assign leisure as the activity type.

   c. If $\big(\Pr(Shop) + \Pr(Leisure)\big) \leq U \leq 1$, assign serve-passenger as the activity type.

In the context of the Laredo GPS survey, we have classified the "other" activities into nine different purposes. This activity-type classification and the corresponding MNL model estimation results are presented in Appendix A.

The reader will note that with the computation of disaggregate activity purposes at trip-ends, it is also possible to further disaggregate home-based non-work and non-home-based trips into finer categories. This is not performed within the scope of this software. However, the output file containing the derived trip diaries can be easily manipulated within any spreadsheet software to compute the disaggregate trip purposes of interest.

3.3.7. Accuracy flags

Two measures of accuracy are computed for each trip detected. The first measure, *NRecRatio,* is a measure of the extent of missing/invalid GPS data for a trip. This is computed as follows:

$$N\operatorname{Re}cRatio = \frac{N_{Valid}}{N_{Valid} + N_{Invalid}}$$

where N_{valid} is the number of valid GPS points for the trip and $N_{Invalid}$ is the number of invalid GPS points for the trip. As discussed in Section 2.2, the preprocessing removes the invalid records from the raw GPS stream and records the number of such invalid records removed immediately prior to each valid GPS record. Smaller values of this measure indicate that a significant fraction of the complete GPS records corresponding to this trip were invalid and hence the trip attributes computed are less accurate than those records with higher values.

The second measure of accuracy, the *MaxSuccInv*, is the maximum number of successive invalid records in the GPS stream corresponding to the trip. A large value of this measure possibly indicates a long period of signal loss or equipment malfunctions, thereby making the trip attribute computations suspect.

3.4.   Step 4: Reasonableness Checks

One could introduce reasonableness checks on each attribute and also could examine combinations of attributes (say for example, trip timing and purpose) to ensure that the trip characteristics predicted are intuitive. Currently, based on preliminary examinations of the trips detected and characterized using the procedures described in Sections 3.2 and 3.3, we have identified two checks. The first check ensures that the trip duration is of at least a certain minimum value by comparing the computed duration against the <u>Minimum Trip Duration Threshold</u>. The second check ensures that the trip average speed is of at least a certain minimum value by comparing the computed average speed against the <u>Minimum Trip Speed Threshold.</u> Potential trips identified using the procedures described in Section 3.2, which have trip durations lower than the <u>Minimum Trip Duration Threshold</u> or trip speeds lower than the <u>Minimum Trip Speed Threshold,</u> are classified as false trips.

CHAPTER 4 SOFTWARE IMPLEMENTATION DETAILS

This chapter discusses the structure of the GPS-TDG software that implements the travel-diary extraction algorithm described in Chapter 3. The software is being developed using the Object-Oriented Programming Paradigm in the Java programming language. ArcGIS 9.0 is used as the underlying GIS platform. Correspondingly, all GIS inputs to the software are in the ESRI shape file format. All non-GIS data inputs are provided as flat files (i.e., simple comma delimited ASCII files). The derived trip diary is also output as a flat file. The structures of all the input data and the output file are first discussed in Section 4.1. Section 4.2 describes the object-oriented design of the software by identifying all the classes and the relationships among the classes. Finally, Section 4.3 provides an overview of the process flow.

4.1.   I/O Formats

As already discussed, all GIS inputs to the software (i.e., the TAZ boundaries layer and the land-use data) conform to the ESRI shape file formats. Further, the software also assumes that all geographic data have been converted to ArcView shape files and are in the Texas co-ordinate system (NAD 1983 Texas Statewide Mapping System). There are three main non-GIS data items input to the software. These are (1) the pre-processed GPS streams, (2) the home and work locations and demographic information file, and (3) the link file. Each of these is discussed in Sections 4.1.1 through 4.1.3 below. In addition to providing data inputs, the analyst also provides the parameters to be used in the algorithm and the specification of the MNL model for activity type determination (if the software is run in the enhanced analysis mode) through a graphical interface. The travel diaries extracted from the GPS streams are stored in an output file. Section 4.1.4 presents the structure of this output file.

4.1.1.   Pre-processed GPS streams

The pre-processed navigational streams for each vehicle reside in a separate file. Each record (or line) of these files has the following format:

"GREC," GPSID, HHID, VehID, Time, Latitude, Longitude, Speed, Heading, NumInvRecs,

where

1. "GREC" is a record-type identifier indicating that this is a record in a GPS stream

2. GPSID is the identifier for the GPS device used.

3. HHID is the identifier for the household.

4. VehID is the identifier for the GPS-equipped vehicles in a household.

5. Time is the instantaneous date and time of data recording stored as time in milliseconds from 1970 (the internal Java representation).

6. Latitude is the instantaneous latitude position recording in decimal degrees.

7. Longitude is the instantaneous longitude position recording in decimal degrees.

8. Speed is the instantaneous speed recoding in meters per second.

9. Heading is the direction of movement in degrees.

10. NumInvRecs is the number of invalid records immediately prior to this GPS record that were removed during pre-processing.

4.1.2. The locations and demographics file

The location and demographics file contains information on the household locations of all the GPS-equipped vehicles and the work locations of all primary drivers who are employed. In addition any other demographic data to be provided to the software for enhanced analysis are also included here. This comma-delimited file has two types of records, the household record and the person record. The household record has the following format:

"DREC," RecType, HHID, HomeLocLat, HomeLocLong, OtherHHData1, OtherHHData2,

where

1. "DREC" is a record-type identifier indicating that this is a record in a demographics file.

2. RecType is the identifier used to distinguish between household and person records in the location and demographics file. Household records have an entry 1 and person records have an entry 2.

3. HHID is the identifier for the household.

4. HomeLocLat is the latitude of the home location in decimal degrees.

5. HomeLocLong is the longitude of the home location in decimal degrees.

6. OtherHHData1 and OtherHHData2 are any other household characteristics data that may be required for the enhanced analysis. For example, the number of children in the household is one such data item required for the Laredo implementation of the enhanced analysis.

The person record has the following format:

"DREC," RecType, HHID, PersID, EmpFlag, WorkLocLat, WorkLocLong, OtherPersData1, OtherPersData2,

where

1. "DREC" is a record-type identifier indicating that this is a record in the locations and demographics file.

2. RecType is the identifier used to distinguish between household and person records in the location and demographics file. Household records have an entry 1 and person records have an entry 2.

3. HHID is the identifier for the household.

4. PersID is the identifier for a person (primary driver) in the household.

5. EmpFlag is a binary flag used to indicate the employment status of the individual (1 if employed and 0 if not).

6. WorkLocLat is the latitude of the work location in decimal degrees.

7. WorkLocLong is the longitude of the work location in decimal degrees.

8. OtherPersData1 and OtherPersData2 are any other person-level characteristics data that may be required for the enhanced analysis. For example, the gender of the person and his/her student status are two such data item required for the Laredo implementation of the enhanced analysis.

4.1.3. The link file

The link file serves as a means to associate each vehicle with its primary driver. Hence, this file provides a link between the files containing the GPS streams and the person-level information contained in the locations and demographics file. Each record of this file has the following format:

"LREC," GPSFilename, HHID, VehID, PersID,

where

1. "LREC" is a record-type identifier indicating that this is a record in the link file.

2. GPSFilename is the name of file containing the pre-processed GPS stream corresponding to the vehicle VehID.

3. HHID is the identifier for the household.

4. VehID the identifier for the GPS-equipped vehicles in a household.

5. PersID is the identifier for the primary driver in the household associated with the vehicle VehID.

4.1.4. The output file

The primary output from the GPS-TDG software is a comma-delimited file containing the trips derived from the GPS streams of all the vehicles in the survey. The first record in this file is a header record characterized by the entry "HREC" in its first field. This header record sequentially includes the values of all the threshold parameters used in the travel-diary extraction algorithm. A list of these parameters has been provided in Table 3.1. The very last record in the output file is a trailer record with the following format:

TREC, TripCount, VehCount,

where

1. "TREC" is a record-type identifier indicating that this is a trailer record in the output file.

2. TripCount is the total number of trips derived from GPS streams contained in the output file.

3. VehCount is the total number of vehicles in the output file.

The rest of the output file is comprised of three types of records: the vehicle header records, the vehicle trailer records, and the trip records. The vehicle header and vehicle trailer records, respectively, are included immediately before and after the trip records corresponding to a particular vehicle and thus help to distinguish the trips of one vehicle from another. The vehicle header record has the following format:

24

"VH," HHID, VehID, PersID,

where

1. "VH" is a record type identifier indicating that this is a vehicle header record.

2. HHID is the identifier for the household.

3. VehID the identifier for the GPS-equipped vehicles in a household.

4. PersID is the identifier for the primary driver in the household associated with the vehicle VehID.

The trailer record has the following format:

"VT," HHID, VehID, NumTrips,

where

1. "VT" is a record type identifier indicating that this is a vehicle trailer record.

2. HHID is the identifier for the household.

3. VehID the identifier for the GPS-equipped vehicles in a household.

4. NumTrips is the total number of trips detected for this vehicle. This is equal to the number of trip records included between the header and the trailer records of a particular vehicle.

Each trip record in this file contains the characteristics of a single trip in the following format:

TR<Trip #>, StartTAZ, EndTAZ, StartLat, StartLong, EndLat, EndLong, StartDateTime, EndDateTime, StartActType, EndActType, TripPurp, EndActDur, TripLength1, TripLength2, AvSpeed, VarSpeed, NRecRatio, MaxSuccInv,

where

1. TR<Trip #> is the trip number identifier. The trips are sequentially numbered starting from 1.

2. StartTAZ is the TAZ identifier for the trip-start location.

3. EndTAZ is the TAZ identifier for the trip-end location.

4. StartLat is the latitude of the trip-start location.

5. StartLong is the longitude of the trip-start location.

6. EndLat is the latitude of the trip-end location.

7. EndLong is the longitude of the trip-end location.

8. StartDateTime is the date and time of the start of the trip.

9. EndDateTime is the date and time of the end of the trip.

10. StartActType is the activity type pursued by the driver of the vehicle at the origin of the trip (an aggregate classification scheme is used in the basic analysis and a disaggregate analysis scheme is used in the enhanced analysis).

11. EndActType is the activity type pursued by the driver of the vehicle at the destination of the trip (the classification scheme is same as that for StartActType).

12. TripPrup is the aggregate trip purpose classification for the trip (HBW, HBNW, or NHB).

13. EndActDur is the duration of activity pursued by the driver of the vehicle at the trip destination.

14. TripLength1 is the trip length computed using the instantaneous position information.

15. TripLength2 is the trip length computed using the instantaneous speed information.

16. AvSpeed is the average speed of the trip computed using instantaneous speed data.

17. VarSpeed is the variance is the speed over the length of the trip.

18. NRecRatio is an accuracy measure computed as the ratio of the number of valid GPS records corresponding to this trip to the total number of records for this trip.

19. MaxSuccInv is a second accuracy measure computed as the maximum number of successive invalid records in the raw (before pre-processing) GPS stream corresponding to this trip.

4.2. Classes and Relationships

The object-oriented design of the software system resulted in the identification of the following major classes: (1) Household, (2) Person, (3) Vehicle, (4) LinkRec, (5) Trip, (6) Location, and (7) GPSRec. The attributes and methods encapsulated within each of these classes and the relationships among these classes are briefly discussed next.

The household class contains the location and demographic information of the household whose GPS records are currently being processed by the software. Similarly, the person class contains the work location and other demographic information of the primary driver of the vehicle whose GPS records are currently being processed.

26

The vehicle class contains a pointer to the pre-processed GPS file corresponding to the vehicle currently being processed. The vehicle class also encapsulates the method for trip detection. Specifically, this method implements the algorithm discussed in Section 3.2 to detect potential trips from a GPS navigational stream. Finally, the vehicle class also contains the method to write out the set of trips detected to the output file.

The data in the LinkRec class are associated with those in the Link File. As a first step to the processing of data from any vehicle, a LinkRec object is created, which is then used to retrieve appropriate data to the create household, person, and vehicle objects.

The trip class contains all data required to characterize a single trip. An instance of a trip class is created each time a potential trip is detected by the method included in the vehicle class. The methods for computing the various attributes of a trip (as discussed in Section 3.3) are encapsulated within this class. Thus, when provided with a stream of GPS records corresponding to a single trip, the trip class can completely characterize this trip.

The trip class also contains two instances of the location class, one corresponding to the start location of the trip and the other corresponding to the end location of the trip. The advantage of constructing a separate class for the locations is in the encapsulation of the GIS overlay procedures within this class. Specifically, the location class contains the procedure to determine the attributes of the zone within which the current location lies when provided with the underlying GIS layer.

Finally, in addition to all the classes discussed above, the software structure also includes a GPSRec class. An instance of this class is created to store each record of the GPS stream corresponding to the trip currently being extracted.

4.3. Process Flow

The main steps in the processing logic implemented within the GPS-TDG software to extract the travel-diary data from the GPS navigational streams and other data are as follows:

1. Set the algorithm parameters based on user-specified inputs.

2. Load the link file and the locations and demographics file into main memory as hashtables for easy access.

3. Open the output file and write the header record.

4. Perform the following steps looping over the elements in the hashtable containing the link file data.

   a. Read in the first/next record from the link file hashtable and create an instance of the LinkRec class.

   b. If end of file is reached, proceed to Step 5.

   c. Create a vehicle object for the vehicle contained in the LinkRec and assign the appropriate GPS file containing the navigational streams for this vehicle.

   d. Using the person-to-vehicle association provided in the LinkRec object, retrieve the data for the appropriate person (primary driver) and household from the locations and demographics hashtable and create instances of person and household classes.

   e. Write the vehicle header to the output trip file.

   f. Perform the following steps looping over the GPS navigational stream records corresponding to the current vehicle.

      i. Read the first/next record from the GPS navigational stream, create an instance of the GPSRec class, and add this GPSRec object to a buffer in the vehicle class.

      ii. If the end of the GPS navigational stream has been reached, proceed to Step ix.

      iii. Invoke the trip detection procedure embedded within the vehicle class.

      iv. If a potential trip is not detected, revert to Step i.

      v. If a potential trip is detected, create a trip object and pass the buffer of GPS records corresponding to this trip to this object.

      vi. The trip object computes all the characteristics of the current trip.

      vii. Add the newly created trip to the buffer of trips contained in the vehicle class.

      viii. Flush the buffer of GPS records contained in the vehicle class and revert back to Step i.

ix. Create a trip object and pass the buffer of GPS records corresponding to the last trip.

x. The trip object computes all the characteristics of the last trip.

xi. Add the newly created trip to the buffer of trips contained in the vehicle class.

g. Perform reasonableness checks on each trip contained in the vehicle object and write out only the reasonable trips to the output file.

h. Write out the vehicle trailer record to the output file and revert to Step a.

5. Extraction of the trips from all GPS-equipped vehicles is completed. Write the trailer record, close the output trip file and quit.

CHAPTER 5 SUMMARY

Household travel survey data constitute a fundamental input to travel-demand model development for use in transportation planning and policy analysis. Although the design of surveys has been enhanced over the years to facilitate a more complete recollection of travel, and the advancements in the field of computers and telecommunications are being exploited to make data recording easier and more accurate, there are still several concerns about the data quality obtained from conventional household travel surveys. These concerns are primarily associated with the challenge of requiring respondents to record details about all trips and activities for the designated travel survey period.

Advances in GPS technology have provided transportation planners with a very powerful tool for accurate data collection with minimal user burden and thereby address the concerns of the conventional household travel surveys. The data recorded by GPS devices, however, do not directly yield the travel information. Rather, the data are recorded as navigational streams that have to be processed to derive the travel information.

The objective of this research project is to develop software, called GPS-TDG, to automate the processing of the raw GPS streams and to generate outputs of activity-travel patterns in the conventional travel-diary format. Within the overall focus of this research, this report presented the data processing algorithms and the software design details.

Chapter 2 presented a conceptual overview of the GPS-TDG system, identifying the various inputs, the pre-processing of the GPS streams, the assumptions in the development of the process module, and the various attributes of the trips computed by the software. This chapter also highlighted the two modes of application of the software, that is, the basic mode and the enhanced mode.

Chapter 3 presented the detailed algorithm for extracting the travel-diary information from the GPS streams and other supplemental data. The overall procedure broadly comprises a trip-detection component, a trip characterization component, and a reasonableness checks component. The trip detection component includes methods for the detection of both engine power-off and non-engine power-off trips. In the enhanced analysis mode, trip characterization also includes a probabilistic procedure for disaggregate activity type classification.

Finally, Chapter 4 focused on the software design details. The software is being developed using the Object-Oriented Programming Paradigm in the Java language. ArcGIS 9.0 is used as the underlying GIS platform.

Appendix

An MNL Model for Activity-Type Classification

This appendix presents the MNL model estimated for activity-type classification using travel-survey and land-use data from the Laredo region. This appendix comprises two sections. Section 1 presents the data sources, sample formation procedure, and sample characteristics. Section 2 discusses the empirical model results.

1. Data Sources, Sample Formation, and Sample Characteristics

Data for the model estimation were obtained from two major sources. The first source is the Laredo household travel survey data. This survey was administered by NuStats on behalf of the Texas Department of Transportation and collected one-day travel data for about 2000 households in Webb County. The second source of data is the parcel-level land-use/zoning data (as a GIS layer in the ESRI shape file format) obtained from the City of Laredo. In this GIS database, the Laredo region was divided into over 68,000 land-use parcels, and each parcel was assigned one of eleven land-use types. These are (1) Agricultural, (2) HDR, (3) Heavy industrial, (4) Institutional, (5) LDR, (6) MDR, (7), Park/Rec. Open Space, (8) Retail/Office, (9) Warehouse, (10) Water, and (11) "blank" or no land-use assignment.

The raw travel survey data were processed in several steps to generate the final estimation data sample. First, the trips for non-home and non-work purposes were extracted from the complete raw trip file. Second, the trips made using transit, non-motorized modes, or as a passenger in a car were removed (i.e., only those trips that were undertaken by driving an automobile were retained). Third, trip records that did not have a geo-coded destination within Webb County were removed. After the above three data-cleaning steps, 3,549 trip records are available for model estimation. Fourth, the trip-end locations were overlaid (using ArcGIS 9.0) on the GIS land-use layer, and a spatial join was used to determine the land-use classification of the parcels in which the trip ends lie. Fifth, individual and household demographic characteristics were appended to the trip file appropriately. Finally, the trip-end activity types were aggregated into nine categories. Table 1 identifies these disaggregate activity types and presents the shares of the trips in the sample corresponding to each of these types.

Table 1. Activity Type Classification and Sample Shares

Activity Type	Freq.	%
Work-related	248	6.99
School	198	5.58
Grocery Shopping	397	11.19
Other Shopping	203	5.72
Personal services	476	13.41
Social, recreational, civic	347	9.78
Eat out	312	8.79
Serve-passenger at school	980	27.61
Serve-passenger at other location	388	10.93
Total	3549	100

2. Empirical Model Results

The empirical model results are presented in Tables 2(a) and 2(b). Table 2(a) presents the parameters for "work-related," "school," "grocery shopping," "other shopping," and "personal services" activity types. Table 2(b) presents the parameters for "social," "eat out," "serve-passenger at school," and "serve-passenger at other location" activity types. The explanatory variables are broadly classified into land use, activity-travel characteristics, and demographic characteristics. Each of these categories of variables is discussed in subsequent sections below. In the overall, the reader will note that the model results are fairly intuitive.

2.1 Land Use

The land use at the trip-end location is found to have a very strong impact on the activity type at the destination. Trips ending at parcels with an industrial or warehouse land use are most likely to be for work-related purposes. Trips ending in institutional land-use zones are most likely to be for the purposes of school or serve-passenger at school. If the land use at the trip end is retail/office space, then the trip purpose is most likely to be shopping. In this case, the trip purpose can also be for eat-out or personal services. Trips ending in parcels with a residential land use (LDR, or MDR/HDR) are most likely for social purposes or for pick-up or drop-off of

children at school. Finally, trip-ends that end in parcels without a land-use assignment are found to be more likely for grocery shopping or for pick-up or drop-off of children at school.

2.2 Activity-Travel Characteristics

The next set of variables captures the impacts of the characteristics of the trip and the activity duration on the type of activity at the trip destination. The following empirical results were obtained:

- Trips of shorter duration are less likely to be for eat-out purposes (see the negative coefficient on the "trip duration" variable for the "eat-out" purpose).

- The coefficients on the "activity duration" variable are negative for all the activity types except for the "work-related" purpose, for which it is fixed to zero. This indicates that, in general, longer the activity duration, greater is the likelihood of the activity type being "work-related" compared to other purposes. The reader will also note that the coefficient on the activity duration variable is most negative for the two serve-passenger purposes. This implies that activities of shorter durations are more likely to be for pick-up/drop-off purposes.

- The next set of variables captures the impacts of activity start time (or, equivalently, trip-end time) on the activity type. As would be expected, trips ending between 7:00 and 9:00 A.M. are more likely for "work-related," "school," or "serve-passenger at school" purposes. Trips undertaken during the mid-day period, that is, 9:00 A.M. to 2:00 A.M., are likely for eat-out or for personal services. Trips ending between 2:00 and 6:00 P.M. are most likely to be pick-up of children from school. Grocery shopping trips are also found to be likely to be undertaken during this period. Finally, activities starting after 6:00 P.M. are most likely to be for social or eat-out purposes.

- If the origin of the trip is home, then the destination is more likely to be school or serve-passenger at school compared to other activity types. This result is intuitive, as these two activity types can be expected to be undertaken as the first stop in a trip chain in lieu of the significant space/time constraints associated with these purposes. Trips originating at work are likely to be for eat-out purposes. Possibly these are lunch trips undertaken during the mid-day from work.

2.3 Demographic Characteristics

The individual and household demographic characteristics that significantly impact the activity type at the trip destinations are employment status, student status, gender, and number of children in the household. Employed persons are more likely to make work-related trips, whereas students are more likely to make school trips. Men are estimated to be more likely to undertake eat-out trips and less likely to chauffeur children to/from school. Finally, the greater the number of children in a household, the greater is the likelihood that an individual from this household will undertake serve-passenger trips.

Table 2(a) MNL Model for Activity Type Classification

	Work-related		School		Grocery shopping		Other shopping		Personal services	
	Param.	t stat	Param.	t stat	Param.	t stat	Param.	t stat	Param.	t stat
Land use										
Industrial or warehouse	1.5104	3.071	-	-	-	-	-	-	-	-
Institutional	-	-	0.7022	2.291	-	-	-	-	-	-
Retail / office space	-	-	-	-	1.9651	7.518	2.0343	11.562	0.3002	2.192
LDR	-	-	-	-	-	-	-	-	-	-
MDR or HDR	-	-	-	-	-	-	-	-	-	-
No land use	-	-	-	-	1.2656	5.142	-	-	-	-
Activity-travel characteristics										
Trip duration	-	-	-	-	-	-	-	-	-	-
Activity duration	-	-	-0.0051	-6.635	-0.0165	-13.878	-0.0115	-10.843	-0.0123	-14.165
Activity start time is …	-	-	-	-	-	-	-	-	-	-
7 to 9 AM	0.4378	1.673	1.1321	4.821	-	-	-	-	-	-
9 AM - 2 PM	-	-	-	-	0.501	3.478	-	-	0.7788	5.458
2 PM - 6 PM	-	-	-	-	-	-	-	-	0.4191	2.875
after 6 PM	-	-	-	-	0.6922	4.504	-	-	-	-
Trip origin is home	-	-	0.977	4.498	-	-	-	-	-	-
Trip origin is work	-	-	-	-	-	-	-	-	-	-
Demographic characteristics										
Employed	3.0294	7.748	-	-	-	-	-	-	-	-
Student	-	-	3.265	15.611	-	-	-	-	-	-
Male	-	-	-	-	-	-	-	-	-	-
Number of children in HH	-	-	-	-	-	-	-	-	-	-
Constant	0	fixed	1.4004	2.849	3.9153	7.961	3.7517	8.416	5.0575	11.482

Table 2(b) MNL Model for Activity Type Classification (continued)

	Social		Eat out		Serve-passenger at school		Serve-passenger at other location	
	Param.	t stat	Param.	t stat	Param.	t stat	Param.	t stat
Land use								
Industrial or warehouse	-	-	-	-	-	-	-	-
Institutional	-	-	-	-	1.8989	8.432	-	-
Retail / office space	-	-	0.8477	5.629	-	-	-	-
LDR	1.4129	8.495	-	-	1.7472	7.607	0.9835	5.369
MDR or HDR	1.1106	2.955	-	-	1.9238	5.246	-	-
No land use	-	-	-	-	0.9117	5.348	-	-
Activity-travel characteristics								
Trip duration	-	-	-0.0157	-2.585	-	-	-	-
Activity duration	-0.0083	-11.347	-0.0129	-11.948	-0.0294	-17.344	-0.0384	-14.574
Activity start time is …	-	-	-	-	-	-	-	-
7 to 9 AM	-	-	-	-	2.6581	17.313	-	-
9 AM - 2 PM	-	-	1.2642	7.699	-	-	-	-
2 PM - 6 PM	-	-	-	-	1.6574	11.445	-	-
after 6 PM	0.9557	6.623	1.1763	6.325	-	-	-	-
Trip origin is home	-	-	-	-	0.8614	7.949	-	-
Trip origin is work	-	-	1.2323	7.958	-	-	-	-
Demographic characteristics								
Employed	-	-	-	-	-	-	-	-
Student	-	-	-	-	-	-	-	-
Male	-	-	0.2752	2.16	-0.3425	-3.029	-	-
Number of children in HH	-	-	-	-	0.4816	10.014	0.1509	2.8
Constant	4.4005	10.146	4.0095	8.764	3.3162	6.947	5.8298	13.401