ABSTRACT

Title of Document:	EXPLORING THE LINKAGES BETWEEN TRAVEL BEHAVIOR AND HEALTH WITH PERSON-LEVEL DATA FROM SMARTPHONE APPLICATIONS.
	Sapeksha Virinchi Vemulapati, M.S., 2013
Directed By:	Associate Professor, Lei Zhang, Department of Civil and Environmental Engineering

In the past, scholars have explored different variables and linked them with the individual's travel behavior. This study explores the linkages between an individual's health and his/her everyday travel behavior. In order to capture accurate and comprehensive travel behavior information, a smartphone application is developed that can track user location for long periods without the need of user intervention. Focus is placed on designing the application to have minimum respondent burden and long-standing battery life of the smart device. Subjects are recruited through a web survey designed to collect information about the individual's healthy living habits. Data from the application is regressed against the health measure data acquired from the survey. Results show that active modes of travel are positively associated with the person's general health measures. The feasibility of this platform as a data collection

method is highlighted while explaining the limitations due to the sample distribution and size.

EXPLORING THE LINKAGES BETWEEN TRAVEL BEHAVIOR AND HEALTH
WITH PERSON-LEVEL DATA FROM SMARTPHONE APPLICATIONS.

By

Sapeksha Virinchi Vemulapati

Thesis submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Master of Science
2013

Advisory Committee:
Associate Professor, Dr. Lei Zhang, Chair
Professor, Dr. Paul Schonfeld
Associate Professor, Dr. Cinzia Cirillo

Acknowledgements

I would like to sincerely thank my advisor, Dr. Lei Zhang, for his support throughout my study. His guidance has helped me develop research oriented skills as well as self-learning capability towards the projects I took under him and otherwise. His approach to problem solving and research has motivated me to put in a lot of enthusiasm into this project. The intellectual discussions with him have opened many a doors of thought for me and this project would not have been as smooth and as interesting without him.

I would like to particularly acknowledge my co-students under Dr. Zhang, especially Cory, Mostafa, Chenfeng and Shanjiang. Our project and group meeting sessions gave me a fundamental knowledge of the subject in a short span of time. Their equal participation in the papers that I was part of could not have been substituted with anyone else.

Friends have contributed immensely through informal discussions outside the class-room and lab. They have helped me clarify concepts in my own mind. Working in the lab has been fun with colleagues at all times of the day. I would also like to thank all lab mates, Masters and Ph.D. students, who have been helpful throughout.

Lastly, I would like to thank my parents and roommates for their overall support and encouragement. This project would not have been possible without them. I thank everyone for his or her priceless contribution.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

1.1 Statement of the Problem

The expanse of travel is growing at an immense rate, making transportation multimodal and increasing the time that people spend on travel. Cities and regions around the world are taking more and more initiatives focused on curbing congestion and motivating commuters to move towards public transport. Understanding traveler behavior is essential in the process of developing and designing these initiatives. Researchers have explored a variety of variables and linked them with the individual's travel pattern. Some of these include trip purpose, urban form and land use, socio-demographic factors, destination image, environmental awareness and psychosocial variables. Studies have largely focused on automobile travel rather than active travel. There is enough evidence that suggests the positive association of healthy-living with travel behavior. However there is little, if any, empirical research that explores the nature of the relationship between general health of an individual and his/her travel behavior.

In order to examine such a relationship, especially the influence of health and travel behavior on each other, there is a need to capture accurate and complete travel pattern of commuters. Transportation planners and modelers have waited over a decade for large-scale (Global Positioning System) GPS-based person and household travel

surveys. The advantages of GPS-based surveys, where devices are used to collect objective and accurate travel details with seemingly little respondent burden, are widely known to researchers. The landscape of travel surveys has been revolutionized with the emergence of smartphone applications. In recent years, a number of studies have been done on the use of smartphone applications to collect travel behavior data and conduct travel surveys. However, no attempt has been made to standardize the structure of conducting and collecting data from these surveys.

To address the above-discussed problems, this research explores a potential relationship between health and travel behavior of an individual. In order to collect accurate and complete travel pattern of individuals, a smartphone application is developed that can track user location for long periods without user intervention. An online survey is also designed and conducted to collect general health information of individuals. The purpose of this thesis is to develop a smartphone application to collect comprehensive travel behavior data and to use this platform in trying to find a potential relationship between an individual's general health and travel behavior.

1.2 Background

1.2.1 Health and Travel Behavior

Transportation systems impact physical activity in a variety of ways. Prior to 1890, walking was the primary mode of transportation (Ford, 1987). Other options were available such as horse cars, trolleys, and even early automobiles, but walking

remained a major component of all trips. That transportation paradigm has changed with technological improvements and the transportation system of today looks and acts very differently than at any prior time period (Muller 2004).

From 1980 to 2003 the United States experienced a 40% increase in the number of "overweight" residents. The rates have increased over the years leading to 65-73 percent of the U.S. population being currently overweight (Center for Disease Control-CDC 2004). This increase in overweight and obesity has been partially attributed to a lack of physical activity (Department of Health and Human Services-DHHS 1996). Lack of physical activity has a far more pronounced effect on public health than obesity. This is because individuals who are relatively physically fit may still be technically overweight due to other reasons (i.e. chemical imbalance, medications, muscular build, etc.). Today lack of physical activity is thought to be the primary factor in more than 200,000 deaths per year in the United States (Blair & Powell, 1994). Research has shown that increasing physical fitness plays the largest role in improving health regardless of other factors (CDC 2004).

Active transportation is one easy way that physical activity can be included in people's daily lives. "Active transportation" refers to the use of any mode that requires using human physical power (Saelens, Sallis, and Frank 2003). For the purposes of this thesis, however, active transportation will refer to animate modes of walking and bicycling. According to many researchers, integrating additional walking and biking into daily routines may prove to be a better public health strategy than traditional structured and organized programs (Handy 2004, Litman 2003b, and Saelensminde 2002). The basic assumption is that changing trip-making behavior to

include more non-motorized trips can translate into favorable public health consequences (Killingsworth, De Nazelle & Bell, 2003).

Even though the relationship between an individual's health and travel behavior have been realized, few empirical studies have been done on this subject. There are many studies, however, on the influence of land use, physical activity on the metropolitan level. For example, residents of sprawling counties were likely to walk less and weigh more than residents of compact counties (Ewing et. al., 2003). Obesity and physical fitness have been widely recognized as critical components of an individual's health and there is a need to explore this at the person-level.

1.2.2 GPS Trackers, Smartphones and Apps

For more than a decade, transportation planners and modelers have waited for large-scale GPS-based person and household travel surveys in which GPS-based data logging devices are used to passively collect objective and accurate travel details with seemingly little respondent burden. Since the first test sponsored by the U.S. Federal Highway Administration in Lexington Kentucky in 1996, slow yet steady progress has been made, with efforts made around the globe, towards this goal. The study was designed to help transportation planners understand the extent of respondent error in self-reported data. The majority of GPS-based travel surveys conducted since in the United States have involved the deployment of passive GPS logging devices in tandem with travel diaries. The immediate focus of this technology application was to improve the quality of travel survey data, with a long-term goal of eventually

replacing respondent-reported data with travel details collected passively through these devices (Bricka & Bhat, 2008).

Due to technology limitations in earlier GPS studies, vehicle-based approaches were the primary method of obtaining objective, passive travel data. Given major enhancements in wearable (or battery operated) GPS data loggers in the past few years, person-based study options are now available. It is worthwhile to note that the same GPS devices can be used for either approach. The use of one approach versus the other can and should be dictated by the study objectives and approach tradeoffs. For example, in the 2007-2008 Washington DC regional travel survey, the primary focus was on vehicle miles traveled (VMT) and so an in-vehicle approach was implemented; whereas the Chicago regional household travel survey wanted to focus on transit users, so a person-based GPS approach was implemented. Vehicle-based approaches may result in better information regarding transportation supply-side information (such as route choice, roadway speeds, etc.), but do not capture other travel behavior aspects such as mode choice. Person-based approaches may show significant promise in capturing mode choice, but are subject to other limitations such as higher respondent burden (associated with the need to carry and charge the device) and quality (wearable GPS data can be scattered or 'noisy' due to incorrect usage or significant sky blockage).

The rise in the availability of location-enabled devices has greatly expanded transportation data collection options. Although GPS can record accurate time and geographic information of travel, participants must still provide detailed attributes

such as trip purpose and mode. GPS suffers from other limitations, which can be overcome by using smartphones. The advantages of using smartphones over GPS trackers will be discussed in the next chapter.

In the United States, smartphone adoption has been extremely rapid. In February 2012, the Pew Internet Project found that nearly half of American adults – 46% – own Smartphones, compared to 35% in May 2011 (now estimate to be 51%). 88% of US adults have a cell phone. GPS has become the standard approach to meeting the E911 requirements for locating people with a cellphone under emergency conditions. In addition to incorporating GPS, most Smartphones also consist of accelerometers, cameras, keyboards and other features. Smartphone sensors can be grouped into 3 categories according to their use in data collection applications:

1. Motion Sensors

    a. Accelerometer, measures the device linear acceleration

    b. Gyroscope, measures the angular rate of change (i.e. rotation velocity)

    c. Magnetometer (i.e. compass), measures magnetic field strength

2. Location Sensors

    a. GPS which is commonly used in outdoor settings

    b. Network-based location services which uses cellular network and Wi-Fi to determine the location (using triangulation)

3. Ambient Sensors

    a. Light sensor

    b. Microphone

c. Proximity sensor

Mobile Application Development is the process by which application software is developed for hand-held devices especially mobile phones and personal digital assistants. Since 2007, the market of applications for smartphones on various platforms has exploded to a great value. Today, there are over 900,000 applications on the App Store (Application marketplace for the IOS platform) and just a little bit lower on the Google Play Store (marketplace for the Android platform). Through advanced Application Programming Interfaces (APIs), context aware applications can be built in fields ranging from productivity to entertainment. The application developed as part of this research only utilizes the category of location sensors. Many Smartphone apps take advantage of positioning (GPS or Wi-Fi) information, including basic web searching such as Bing. Some notable applications are:

1. Foursquare: Check-in to locations

2. Google Maps: Get directions to your destination

3. Next Bus: Know when the next bus arrives/ departs.

4. Tile App: Track down lost items
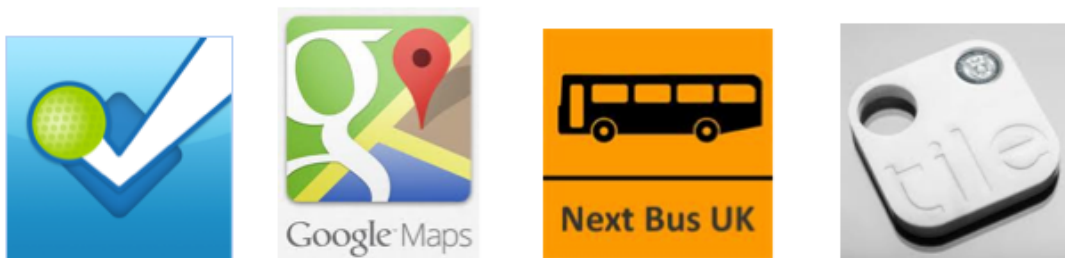
Figure 1: Existing Mobile Applications

1.3 Objective

The objective of this study is two-fold:

1. To develop a Smartphone Application that would serve as a platform to passively collect comprehensive person-level travel behavior data,

2. To design a health survey to collect information about an individual's general healthy-living habits

3. To build a model that will help understand and analyze the relationship between an individual's travel behavior and his/her general health measures.

The smartphone application will help standardize the method of collecting person-level travel behavior data by automating the process, hence not requiring any user intervention after the installation. The application's survey feature will allow researchers in customizing surveys and pushing them to respondents in real-time. The Health and Travel Survey will collect person-level general health information and will help the development of future health-related surveys. Results, though not significant, show that there is a positive association between travel behavior and healthy-living and that researchers should focus on understanding this relationship better.

1.4 Organization of the Thesis

The contents of the rest of this thesis are divided into five chapters. Chapter 2 provides the literature review that covers the evolution of travel behavior studies, emergence of GPS and Smartphone-based surveys as well as previous studies that

have linked healthy-living and travel behavior. Chapter 3 conveys the research design and methodology. The overall process of the application development is explained, from the specifications of the requirements through the system design until the deployment. The design and the implementation of the Online Health and Travel Survey are also described. Finally, chapter 5 presents results and discussions, and chapter 6 ends with conclusions and implications.

Chapter 2: Literature Review

2.1 Active Travel Behavior

Past research has proven that a variety of personal factors make one individual behave differently than another (Golledge and Stimson 1997). These different factors also allow individuals to make personal decisions when it comes to their travel behaviors. Travel behavior can generally be referred to as the study of what people do over space and how people use transportation (Hayes 1993). Goulias (2000) gave a more comprehensive definition stating that travel behavior is "the modeling and analysis of travel demand on the basis of theories and analytical methods from a variety of scientific fields. These include but are note limited to, the use of time and its allocation to travel and activities, the use of time in a variety of time contexts and stages in the life of people, and the organization and use of space at any level of social organization, such as the individual, the household, the community, and other formal or informal groups". As Handy (2005) well stated, however, the majority of travel behavior research to date has focused on automobile travel rather than animate or active travel. It is especially important to consider the following concepts with regard to active travel.

2.1.1 Time Allocation

The initial concept of spatial and temporal capacities and constraints on individual behavior were proposed in the 1960's. Hagerstrand (1970) originally emphasized the importance of time in human activity. He noted that "time has a critical importance when it comes to fitting people and things together for functioning in socio-economic systems." So even if a given location is near an individual, they may not be able to allocate enough time to travel to it. Spatial proximity alone does not inherently make a difference.

Temporal constraints also play a large role in active mode choice. Transportation systems reduce the amount of time required for movement across space. A person must trade time for space through movement or communication to participate in activities (Golledge and Stimson 1997). Greater separations inherently imply a lower level of accessibility. This especially holds true with regards to active mode choice. When destinations are located further apart the time required to reach those destinations increases. Choosing an active mode may not allow travel as quickly as other available modes resulting in a large capability constraint.

Travel behavior research also asserts that all individuals have a limited time-budget to allocate among flexible activities, such as shopping or other errands, which may not have stringent time restrictions and predetermined starting or ending times. The time-budget and the ability to trade time for space using transportation technologies or modes determine an individual's accessibility to opportunities that exist in relatively few places for limited durations. Sallis, Frank, Saelens, and Kraft (2004) argue that

there is a threshold of time at which time spent traveling is perceived as no longer reasonable. This is especially relevant considering active trips generally require more time for travel, which may exceed any existing threshold.

Until recently, traditional modeling of travel behavior and time allocation regarded trips as the primary focus of analysis. Travel diaries would inherently leave out "trips" which began and ended in the same location with no stopping points in the middle (i.e. recreation trips such as a walk around the neighborhood), and trips, which were used as feeders to other modes (i.e. bicycling to the bus stop). Using this type of trip-based analysis inevitably left out a large number of active trips leading to drastic underreporting, since many participants in research did not consider recreation outings as "trips", as outlined by the traditional research definition. Many researchers analyzing time allocation were disappointed by the drawbacks of trip-based analysis and turned instead to a different type of measurement framework; activity-based. According to Ettema and Timmermans (1997) "activity-based approaches typically describe which activities people pursue, at what locations, at what times, and how these activities are scheduled given the location and attributes of destinations, the state of the transportation network, aspects of the institutional context, and their personal and household characteristics." Activity-based approaches help researchers identify activity patterns that more accurately reflect how people plan and organize their days.

2.1.2 Demographic Characteristics

Young people (under age 18) and older individuals (age 65+) are the groups most likely to utilize active modes of transportation (Burbidge, Goulias and Kim 2006, Ewing et al 2003, and Pucher and Renne 2003). It should be noted that one likely reason for this is that both the young and elderly are often captive to specific modes of transportation. For example, prior to age 16 individuals in the United States cannot legally obtain a driver's license. Older individuals may lose the ability to operate an automobile as they age due to vision loss, decreased response reflex, or other degenerative conditions. This makes both groups reliant on other drivers, active modes, or transit for transportation.

Socio-economic status has been found in various studies to affect active mode choice as well. Giles-Corti and Donovan used a cross-sectional survey to study 1803 adults near Perth, Australia, and found that survey respondents in low socio-economic areas had superior spatial access to many recreational facilities but were less likely to use them when compared with those living in high socio-economic areas (2002a). After adjustment, respondents living in low socio- economic areas (not explicitly defined in the research) were 36% less likely to undertake vigorous activity. Research has also shown that lower income individuals utilize active modes of transportation less than those with higher income, even when both groups live within the same neighborhoods with similar infrastructure available (Brownson et al 2001, and Pas and Koppelman 1986). One would generally assume that lower income individuals traveling shorter distances would utilize active modes more often, but this does not seem to be the case, even when considering low income individuals who are captive and may not

have access to an automobile or do not have the ability to drive a car.

Gender has also proven to influence travel behavior. Brownson et al's (2000) cross-sectional survey of 1269 adults in rural Missouri showed that women are significantly more likely to participate in physical activity and utilize neighborhood trails than men. When individuals do travel actively, women are more likely to walk for transportation, but men are more likely to utilize a bicycle for active travel (Pucher and Renne 2003). Also notable for its social affect on travel behavior is education level. Burbidge, Goulias, and Kim (2006), and Coogan (2003), showed that individuals with higher levels of education walk significantly more than those with lower levels of education.

2.2 Smartphone-Based Surveys

GPS devices can record accurate time and location information of travel. However participants must still provide details such as the trip mode and purpose. To collect information that cannot be derived from GPS data alone, prompted recall methods may be used, including paper-based (Doherty et. al., 1999), mobile-based (Chen & Fan, 2012) and web-based (Auld & Mohammadian, 2009).

While largely successful when used as a supplement to household surveys, GPS trackers have some disadvantages. Researchers or agencies conducting these surveys have the burden of purchasing and distributing these devices to survey recruits. Potential for loss or damage to these devices is also considerate. While the costs of these devices have dropped significantly since first introduced, they still represent a

considerable investment. This limitation is overcome in the case of smartphones, as they belong to the survey subject. In addition, the devices are an additional burden to carry around and most of the times are forgotten at the household.

The landscape of travel surveys has been revolutionized with the emergence of smartphones. Smartphones are ubiquitous and versatile loggers. Besides GPS, each smart device includes a variety of sensing technologies such as accelerometer, Wi-Fi and Bluetooth. University of Minnesota's UbiActive application uses data captured from the accelerometer, GPS and 3-dimensional magnetic sensors to determine location, speed and orientation in order to calculate physical activity duration and intensity (Chen, Bierlaire & Flotterod, 2011). Being personal devices, they are rarely forgotten and the respondent doesn't feel an additional burden in carrying it. Costs involved in conducting travel surveys through Smartphones are only towards developing and deploying software applications through online app stores, which is significantly lesser than GPS-based surveys. Deployment of modifications in the software (updates), reporting of erroneous data and feedback to respondents can be provided in real-time because the devices are always connected to the internet. Overall, developments in this field suggest that location-enabled technologies shall reduce the occurrence of erroneous data, improve accuracy of reported trips, locations and reduce respondent burden.

In recent years, many studies have been done on the use of smartphone applications to collect travel behavior data and conduct travel surveys. However, during this time

no attempt has been made to standardize the structure of conducting and collecting data from these surveys. Some of these applications have been discussed below:

2.2.1 Quantified Traveller

UC Berkeley developed the "Quantifiable Traveler" application. This application was recently piloted with about 80 people tracking their travel for a survey period of more than two weeks. The Smartphone application acts as a passive data collector, relying on WIFI (80%) and GPS (20%) for imputing a travel route. Through a combination of both using predominantly WIFI instead of GPS and minimizing the amount of data collected, the battery life was extended.

One trade-off that is often discussed in relation to passive location tracking is how much detail about travel speed and route is needed to impute route and mode. In this app, there was an average of 30 location points per hour. Using prompted recall, respondents were asked to review their travel and to make modifications as needed to the imputed route and travel mode. For vehicle drivers, 94% of trip routes were verified as correct, but respondents corrected over 50 percent of routes in BART trips. About 14 percent of imputed modes were changed to another travel mode during the respondent verification stage. This app is designed to provide feedback to respondents about calories burned, greenhouse gas emissions, travel time and travel cost.
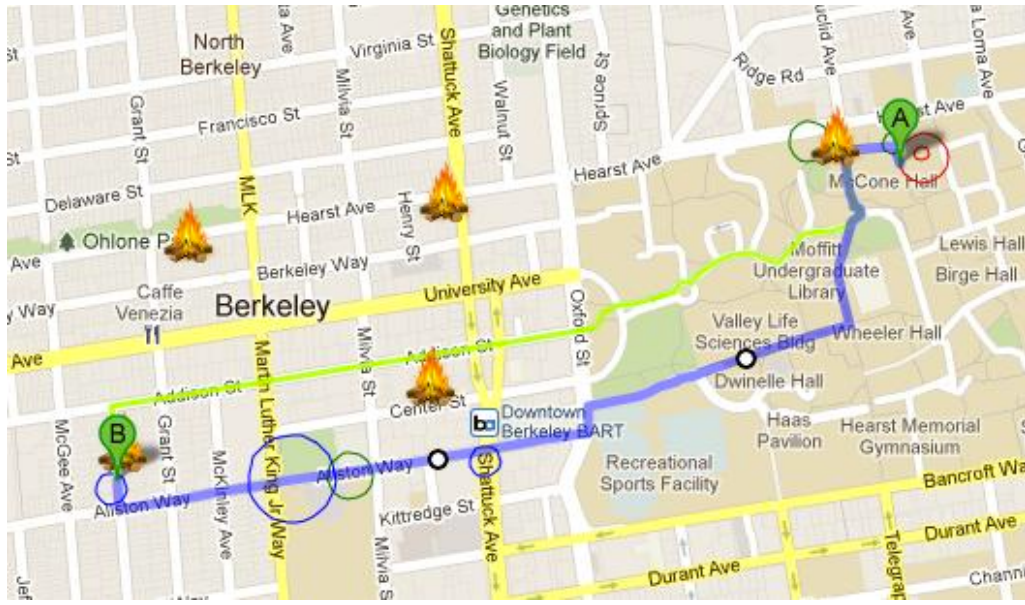
Figure 2: Imputing route from limited GPS and Wi-Fi points
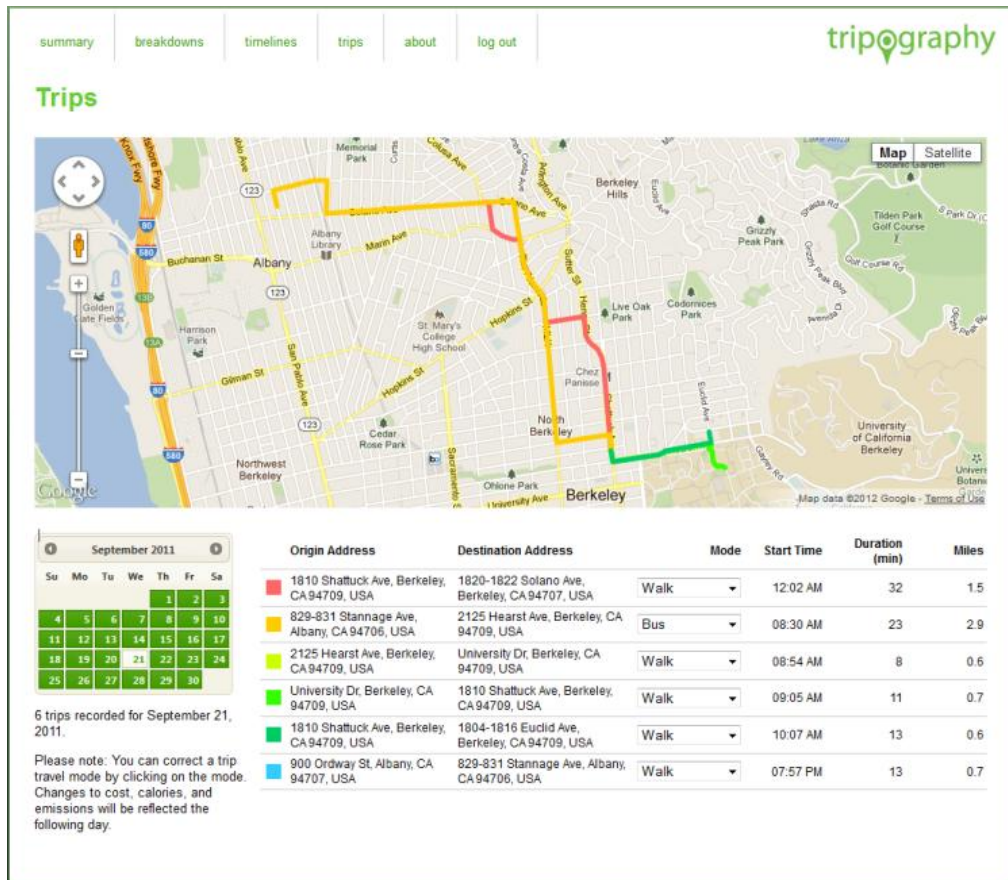
Figure 3: Summary of imputed travel modes and locations
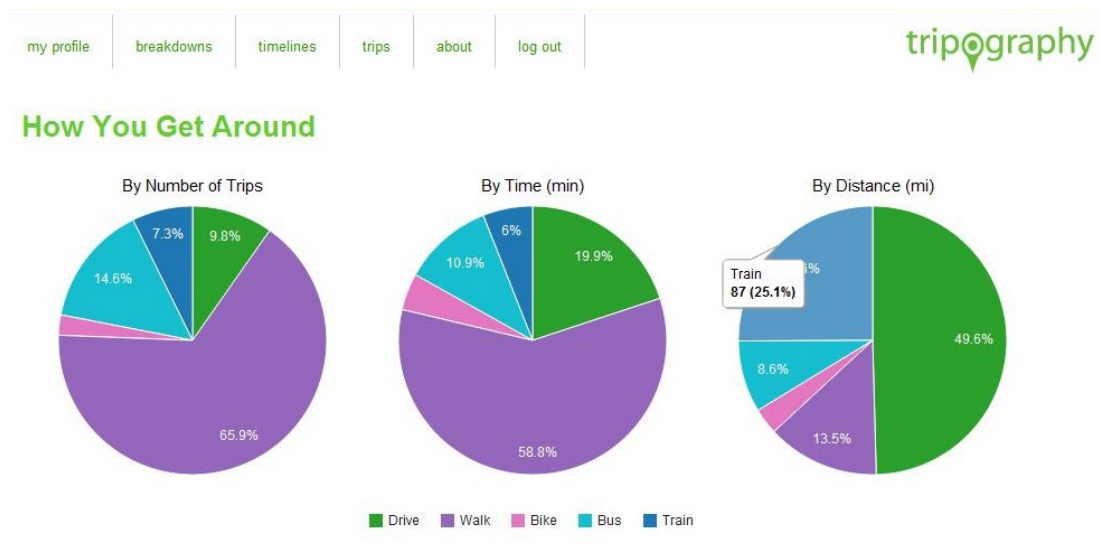
Figure 4: Feedback screen provided to participants

2.2.2 NuStats PaceLogger

As part of NuStats role as prime contractor conducting the Oregon Travel and Activity Study (OTAS), NuStats worked in close collaboration with Portland METRO staff to develop a Smartphone application for deployment alongside the Portland, OR portion of the statewide household travel study. The PaceLogger application was developed as a demonstration as a specific proof-of-concept—notably, that 1) recruited household travel study participants would be willing and able to collect GPS data via Smartphone during their travel date and 2) this GPS data could be successfully linked back to travel diaries to provide additional analysis and modeling value.

The ultimate product (PaceLogger) successfully captured detailed GPS trip data for over 1,000 study participants, demonstrating that Smartphone applications could serve as an inexpensive, easily deployable means of gathering accurate supplemental

trip level GPS data for adding value to traditional household travel study trip data.

From a development standpoint, NuStats chose to modify an existing open-source licensed smartphone application originally designed to capture cyclist route choices. This application, CycleTracks (see summary below or visit the CycleTracks webpage) was developed by the San Francisco County Transportation Authority and is currently being used by several agencies across the country. CycleTracks application provided the ability to leverage Smartphone technology to collect accurate GPS traces on a trip-level basis and upload these GPS records to a NuStats database. NuStats adapted the program by building an interface for capturing traveler information so that the record of GPS traces could be easily tied to participant recruitment and retrieval data. In addition to this respondent data (basic demographics, personalized PIN number), NuStats also added instructional text, FAQ's, and other features to facilitate program use by participants.
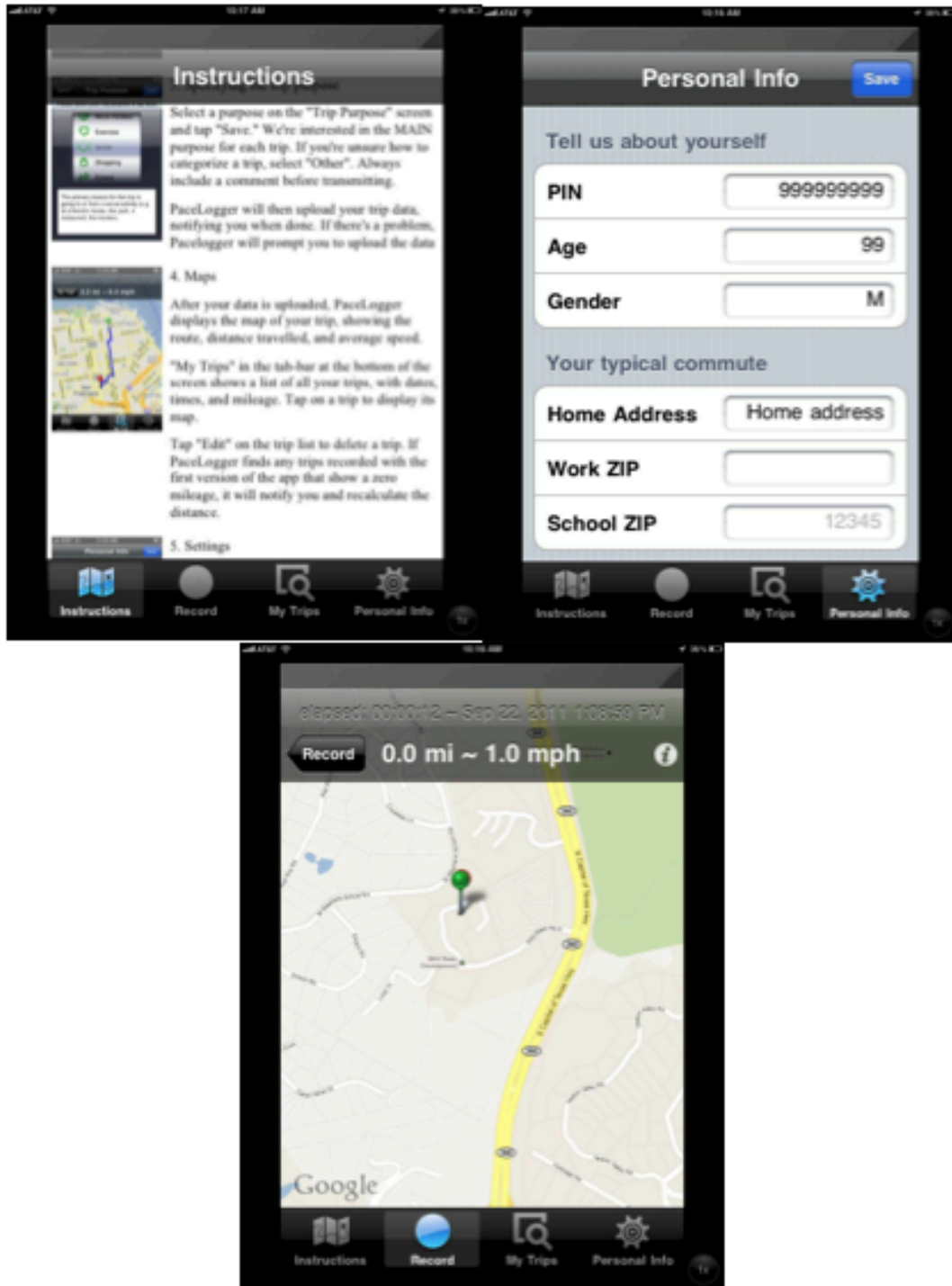
Figure 5: CycleTracks User Interface

Compared to the current practice of providing GPS data loggers to respondents by

mail, the use of respondent-owned Smartphones reduces the hardware costs of conducting a household travel survey and there is a significant improvement in the timeliness of travel survey data retrieval, because the data is sent via cellular phone networks, rather than relying on respondents to return equipment.

In this study, the Smartphone survey was limited to only one day, to coincide with the traditional paper diary. The Smartphone app did not request travel mode from the respondent.

PaceLogger proved highly functional and widely used by study participants. No issues were reported with regards to battery drain or other technical problems. The quality of data in general was excellent, with less than 10% of submitted person-level records failing QC checks and removed from delivery dataset—a percentage that is bound to improve with additional experience with the program. While NuStats did provide a $10 per household member incentive for using the PaceLogger application, recruitment staff felt that this incentive was not the primary driver of respondent participation, but rather that curiosity, novelty, and respect for technological efficiency underscored the effort. In fact, the most frequently received feedback on PaceLogger was a wish for the software to replace the traditional diary and retrieval stages of the household travel study. Currently, NuStats is working to improve the functionality of PaceLogger and examining means of expanding the application's use.

Other studies have developed applications that conduct online surveys through the smartphone as well as use this information in diabetes and human exposure research.

However, most of these applications have still not overcome two major drawbacks of this system: Battery Life and User burden. It is necessary to build a system that can collect travel behavior information about the individual all through the day for months together without any user intervention, once setup during installation.

2.3 Health and Travel Behavior

A number of studies have investigated the linkages among health and leisure travel. Their reports have mainly focused on specific aspects of air travel (Katz et. al., 2002; Low & Chan, 2002) and the contribution of the leisure experience to one's stress level or psychological aspects (Caldwell & Smith, 1988; Chalip, Thomas & Voyle, 1992). Other studies have looked at the time spent on physical activities and healthy modes of transportation such as walking/biking having a considerate impact on health. Studies provide evidence of a dose-response relationship between physical activity and healthy living, i.e., as a person engages in larger amounts of physical activity, the resultant health benefits are larger. The built environment has come under scrutiny as a potential contributor to the involvement in active travel (Sandvik et. al., 1993). Walking for transportation can be an important source of physical activity for otherwise non-active individuals (Handy, Suzan & Clifton, 2007). Results from the SMARTRAQ program have shown that residents in high dense neighborhoods spend more time in walking and physical activities (Sallis et. al., 2004). However, there is a lack of research that analyzes person-level health data to evaluate potential transportation policy and planning applications. Scholars should

evaluate individual travel behavior information against person-level health data to understand the overall impact on a community or a region.

In 2007, the US population took an estimated 10.3 billion public transportation trips, an increase of 32% from 1995; a behavioral change if sustained could impact health favorably (Chapman & Frank, 2004). Usage of public transportation can potentially generate health benefits from the persistent aerobic activity that results from walking and climbing stairs when one is riding buses and trains, and from moving to, from and within stations. The benefits achieved from switching to public transport outweigh the risks associated with it (Morabia et. al., 2009). The primary concerns that transportation researchers have are the traffic-related pedestrian injuries and various health-related effects of automobile-related air pollution (Morabia et. al., 2010). However, physical unhealthiness is a far more serious problem and the linkage with transportation means that health needs to become a central concern for the transportation field. "There is a clear need for transportation and public health professionals to collaborate closely on research, policy, and practice that will lead to joint efforts to meet societal needs" (Handy, Suzan & Clifton, 2007).

Chapter 3: Methodology

3.1 Application Development

3.1.1 Specification of Requirements

A well-functioning application would have to collect travel or location information about the individual in a timely manner and make this information accessible to the researcher when needed. In line with the objective of the research, the requirements of the application were specified in three simple tasks. It shall be able to:

1. Track the travel information about the individual, all-day, for long periods without any user intervention.

2. Store this information locally on the device, without losing it.

3. Send the stored information to the researcher at regular intervals

3.1.2 System Design

The first requirement was achieved through Background Location Updates, a feature available with most smartphone platforms (implemented on IOS and Android). The sensing technologies are kept running on the smartphone even when the user is performing other tasks on it or turns the screen off. This meant that the application would need to be optimized to consume as low power as possible. Smartphones have three sensing technologies that are sources of location in applications: Cell Signal, Wi-Fi and GPS, in the order of lowest to highest accuracy. High accuracy consumes

high power and hence, the three sensing technologies also fall in the same order from least to most power consumption.
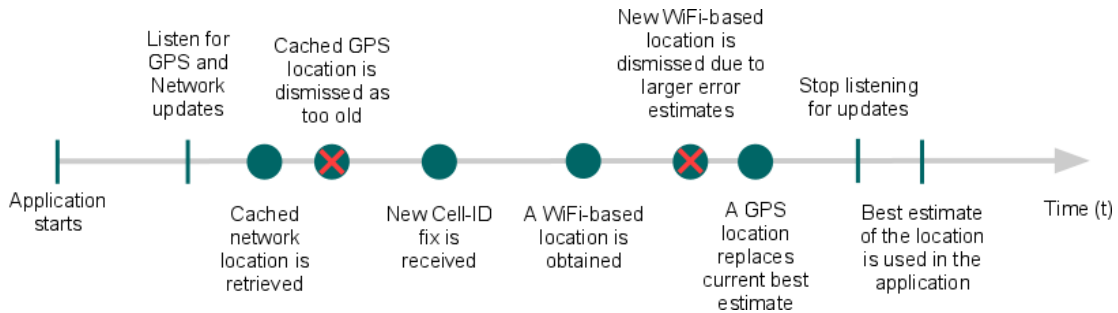
Figure 6: Location Request Timeline

Using GPS for a long time would drain the battery of the device quickly. Therefore, the GPS chip would remain switched off, until the movement of the user was detected through the other location sources. Imaginary circular regions around the user's current location would serve as triggers for the switching the GPS chip on and off. The application would request location information from these sources based on two parameters, *d* and *t*, which stand for Minimum Distance and minimum time. Minimum Distance refers to the distance in meters the user has to travel before another request is sent. Similarly, Minimum Time refers to the time in milliseconds that have to pass, for the application to make another location request. The values of *d* and *t* can be changed to suit specific needs.
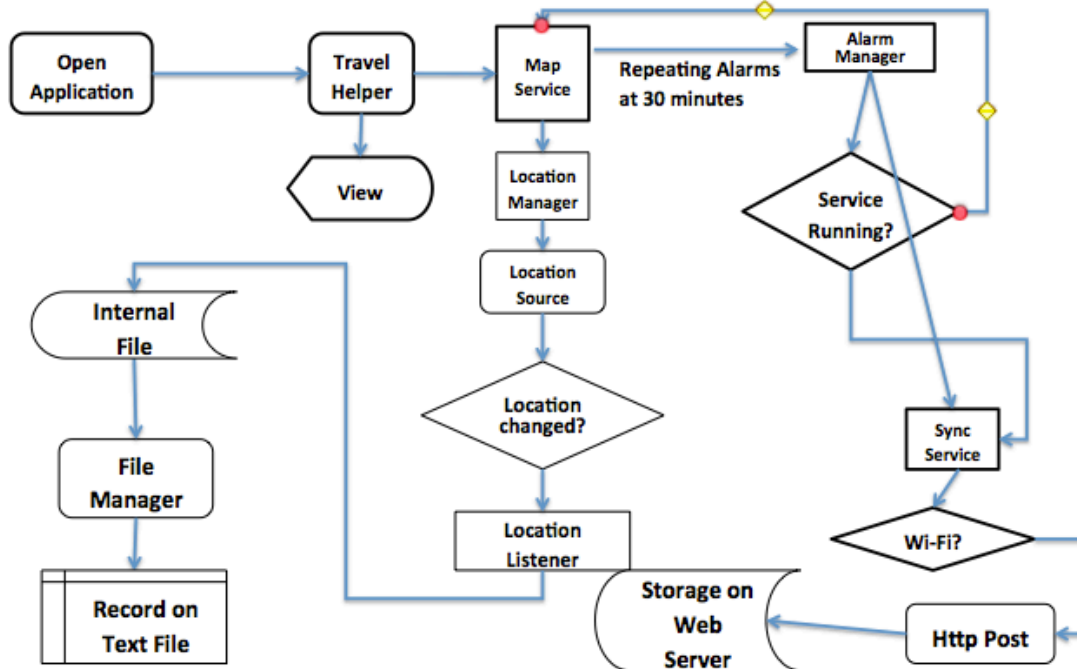
Figure 7: Travel Helper Application Design

Received location information is stored on a single comma separated text file locally on the device. Each location point is recorded on a new line for easy data processing at a later stage. The information available from each location point is described in the Data Product section. The text file is then transmitted to a pre-setup server at regular intervals. Currently, this interval is set to 24 hours. Once the transmission receives a successful response, the text file is cleared, to save space locally on the disk.

3.1.3 Implementation Issues

Even though there are many advantages in having a mobile phone (or smartphone) application that collect location data, this system is a long way from being adopted in

a large scale. It is worth discussing some of the major problems in smartphone-based surveys through the issues that this research has experienced.

Battery Life: Battery life continues to drive much of the design of portable computing devices combined with GPS. These devices are used on a daily basis for various kinds of activities such as calling, entertainment, gaming etc. A viable travel survey app should work on the public's Smartphones and be capable of gathering required data without disabling a device's calling, texting or browsing functions or rapidly depleting its battery. Hence the power consumption of Travel Helper is kept as low as possible while not hindering the accuracy of the information collected.

Varying Smartphone Platforms: In the US, over 50 percent of smartphones are running on the Android operating system, with smaller shares of IOS, RIM and several other platforms. Though most smartphones carry the same features, the interaction between the system and sensing technologies such as GPS varies across different platforms. Travel Helper on Android works completely different when compared to that on the IOS. With the increase in adoption of smart devices in the transportation industry, the functioning of these sensing technologies will be standardized.

Data Storage and Transmission: Another important concern is the storage of data on the mobile phone. Interactive tracking apps that store trip data on the device before uploading to a server may exhaust the device's available memory. The impact of this

issue is particularly relevant to older smartphones and those that offer limited database support. Hence, Travel Helper regularly sends data to the cloud and clears the data on the device after successful transmission. This method then needs to avoid using limited yet costly Internet data of the phone.

3.1.4 Deployment and Maintenance

The Travel Helper application has been developed for the IOS and Android operating systems. The Android version can be downloaded from the Google Play Store. The IOS version, however, is still in the testing phase and only chosen individuals have access to the application. Applications on both platforms were designed to give the best user experience while not sacrificing on the objectives of the research. Snapshots of the application are shown below:

Figure 8: Travel Helper Application Snapshots

3.2 Health and Travel Survey

3.2.1 Survey Design

In order to recruit subjects for installing the Travel Helper application, a Health and Travel Survey questionnaire was prepared. It was designed primarily based on the annual health survey Behavioral Risk Factor Surveillance System (BRFSS) and other health surveys by World Health Organization (WHO) and Healthy People. A brief description of these three surveys is provided below:

The Behavioral Risk Factor Surveillance System is a collaborative project of the Centers of Disease Control and Prevention (CDC) and U.S states and territories. The BRFSS is an ongoing data collection program to measure behavioral risk factor for the adult population (18 years or older) living in households. The BRFSS was initiated in 1984 collecting data on risk behaviors through monthly telephone

interviews. The BRFSS Questionnaire has three parts, however, the Health and Travel Survey is limited to the first part, also known as the Core Component. Questions about the involvement in regular exercise, physical activity, alcohol consumption and smoking are included in the Core Component of the survey.

3.2.2 Recruitment of Subjects

Considering the above factors, a Health and Travel survey was designed and hosted online at the Travel Survey website (http://www.travel-survey.org/survey-participation-11). Subjects were recruited by announcing a reward of $100 for 10 random participants of the survey. Users had to answer the questionnaire online. They would receive instructions to install the application on their phones. They had to keep the application running until the application was successful in transmitting at least 4 weeks (28 days) of data.

Chapter 4: Model Specification and Analysis

4.1 Location Data Processing

4.1.1 Data Product Features

A pre-pilot study was done to test the reliability and accuracy of the information collected by the application. The study involved a sample of University of Maryland students downloading and using the application. Data was successfully received from 16 subjects (out of 23 installations). The study also helped detect primary errors in the application and in improving the user experience. The observed application usage and drop out helped re-design the application to have minimum burden on the user.

Only 46 subjects installed the application and provided data for the expected period of a minimum of 4 weeks. All subjects are current or recent undergraduate and graduate students of the University of Maryland. Descriptive statistics of the data collected from these subjects is shown below. All data was collected in the months of May – July 2013. The individuals have been ordered based on the number of days of data received from them, highest being first. The order has been kept the same across all graphs.
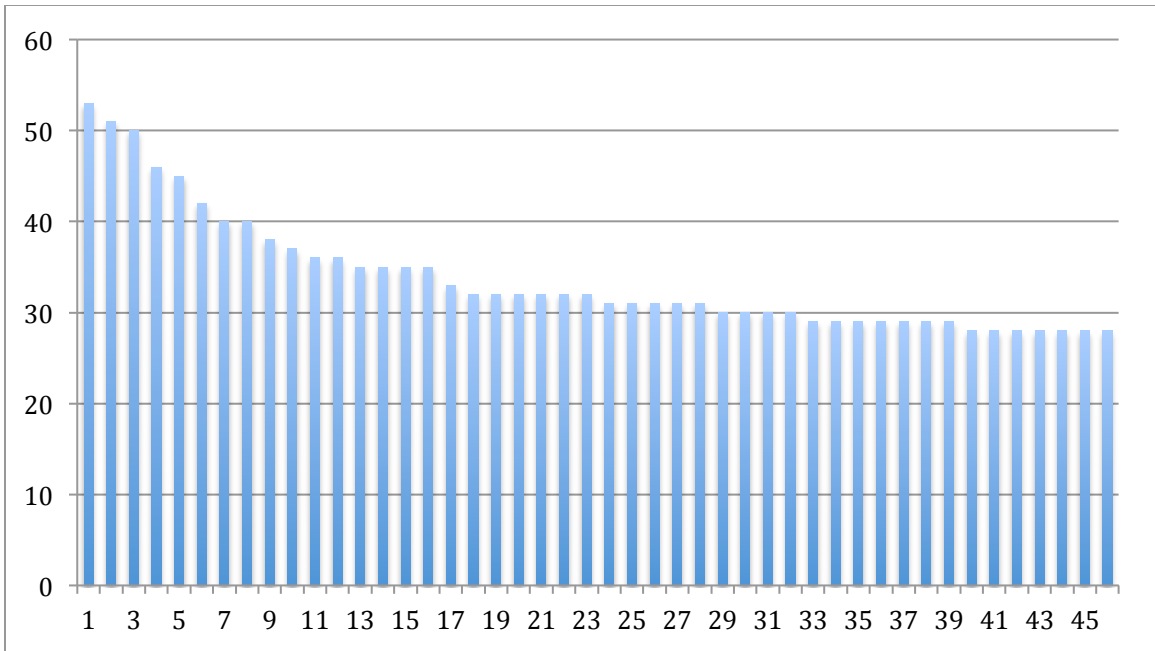
Figure 10: Number of Days of Data Collected
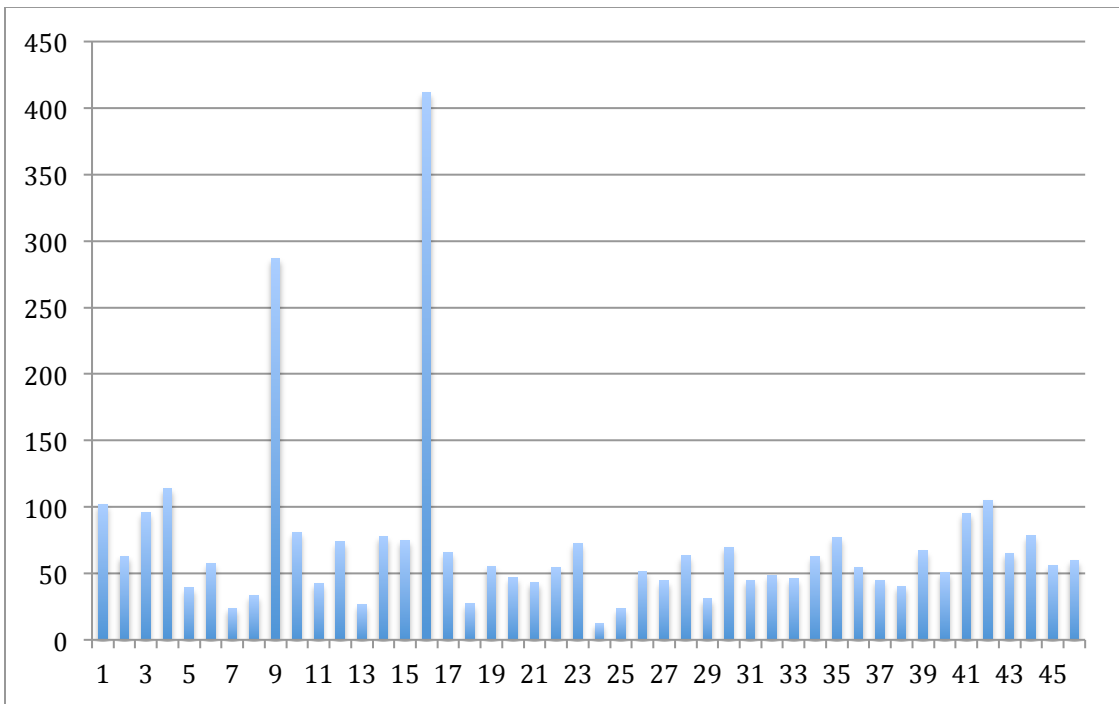
Figure 11: Number of Data Points Recorded Per Day

A higher number of data points per day indicate that the subject travels a lot. This happens because the application is optimized to record location information only when the user is travelling. It can be observed that for most individuals for whom daily commute may be the only trip for the day, the number of points recorded is between 50 to 100 points per day.
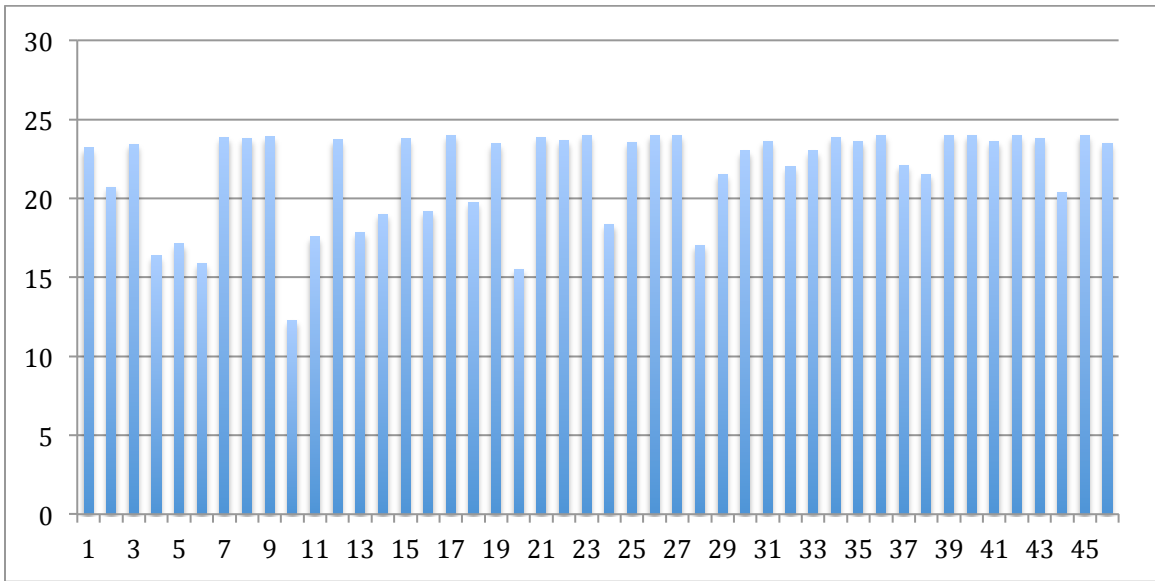
Figure 12: Number of Hours Per Day Captured

For most of the individuals, on average the application is able to capture travel throughout the day. This shows that the data collected from the application is good enough to use for travel behavior analysis.

The data collection and transmission methodology in the current version of the application has been designed while keeping in mind the battery consumption of the application as well as the required location information to be able to realize the travel pattern of the user.

The application requests location information from the GPS chip in the device at regular time intervals. The frequency of these requests is dependent on whether the user is understood to be moving or stationary. The chip then returns location data to the application. The initiation of this request is made when a pre-defined distance is traversed. As shown in the figure, when a location outside of the circle (with pre-defined distance radius) is crossed, data is recorded onto a local file system on the device. Currently, the time and distance intervals are defined as per the user's state:

1. Moving: Time interval = 2 minutes, Distance interval = 75 meters

2. Stationary: Time interval = 5 minutes Distance interval = 250 meters

The application is capable of collecting location data continuously all through the day and for long periods without any user intervention. The full-scale study hence aims at conducting the survey for periods of 2 – 3 months. All recorded data is stored on a single file on the device. Once a day, this file is sent to a server, consequently clearing the content for saving space on the device. Data on the server is then downloaded and analyzed for processing.

Each location point recorded has information on all of the following variables. Information on the altitude, as an exception, may not be available for all data points:

1. Location: Latitude/Longitude of the point

2. Instantaneous Speed: Speed of the device when recorded

3. Date: Date of the recorded point

4. Time: Time of the recorded point accurate to millisecond

5. Accuracy: Accuracy of location information recorded

A sample of the location dataset received from the application is shown below:

Table 1: Sample Travel Helper Location Dataset

Latitude	Longitude	Speed	Accuracy	Time	Date
38.984957	-76.949055	1.041166	5	19:05:42	05/19/2013
38.98487	-76.949133	0.807995	5	19:05:46	05/19/2013
38.984813	-76.949239	0.995833	5	19:06:11	05/19/2013
38.98478	-76.949354	1.03863	5	19:06:19	05/19/2013
38.984733	-76.949472	1.055842	5	19:06:25	05/19/2013
38.984767	-76.949596	0.80183	5	19:06:36	05/19/2013
38.984779	-76.949721	0.975941	5	19:06:43	05/19/2013
38.984762	-76.949842	1.089352	5	19:06:50	05/19/2013
38.984747	-76.949961	0.904442	5	19:07:02	05/19/2013

4.1.2 Travel Mode Detection

The travel mode is identified through a hierarchical process. This process is based on the hierarchical process discussed by Stopher (2007). The hierarchy begins with walk trips, which are identified based on average speed and maximum speed. Following this, off-network public transport trips are identified, using the appropriate network information. The next mode to be identified is bus, which is identified through average speed, beginning and ending of the trip on a bus route. The remaining trips

will be car and bicycle. Bicycle trips have proved to be the most difficult to pin down, origin of the trip being at home or at a location previously identified as the destination of a bicycle trip and average speed. The remaining trips are assumed to be by car. If the destination of the previous trip was not served by car, then it is assumed that the trip is by car passenger. Otherwise, all trips are assumed at this point to be car driver. In the future, further tests can be introduced to determine if there are other household members with an identical trip by time, route, and origin and destination location, which would allow identification of additional car passenger trips.

4.1.3 Analysis Methodology

The data collected from the mobile application is then analyzed to identify travel behavior attributes that are critical for individual behavior analysis. Closely spaced latitude longitude points hugely represent locations where the user has spent a lot of time or has visited frequently. Hence clustering algorithms are used to greatly improve the time required to process the dataset. Locations where the user has spent more than 6 hours, or in other words been stationary or hasn't moved significantly, are considered as important places of interest. Duplicate and closely spaced points are clustered around these important places of interest. A step-by-step illustration of the methodology used for analysis is shown below:
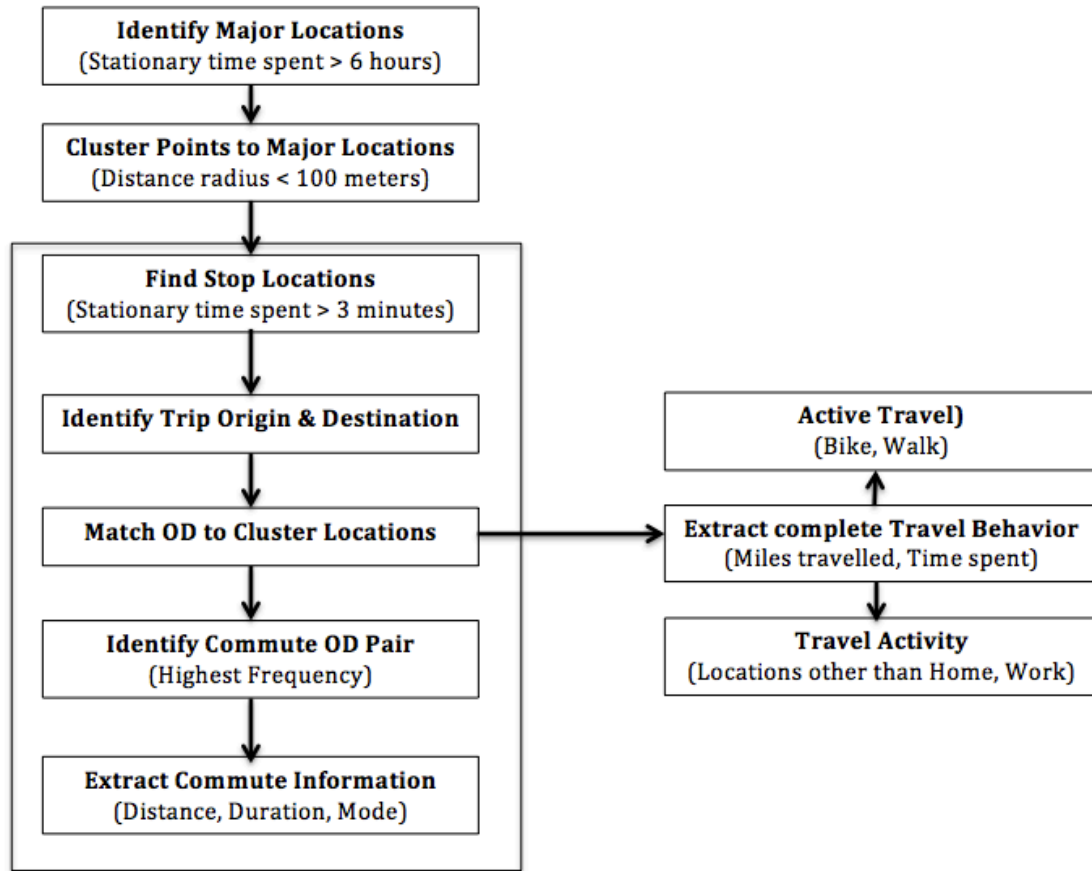
Figure 13: Location Data Analysis Methodology

4.2 Modeling Travel-Behavior Health Linkages

4.2.1 Standard Health Measures

The World Health Organization implemented World Health Survey in 2002-2004 in partnership with 70 countries to generate information on health of adult populations and health systems. Healthy People provides science-based, 10-year national objectives for improving the health of all Americans. According to Healthy People, there are four foundation health measures that are used to monitor promoting health, preventing disease and disability, eliminating disparities and improving quality of

life: General Health Status, Health-Related Quality of Life and Well-Being, Determinants of Health and Disparities.

The following measures of the General Health Status provide information on the health of a population:

1. Physically Unhealthy Days

2. Self-assessed health level

3. Chronic Disease Prevalence

4. Limitation of Activity

4.2.2 Model Specification

The model is specified as a linear regression model. It is designed to help explore the dependency of the four general health measures (listed above) on travel behavior, demographics, and diet and exercise habits. An additional variable Body Mass Index will be measured to observe the "Obesity level" in the subjects. Body Mass Index is calculated using weight (in kilograms) and height (in meters) using the formula:

$$BMI = weight / (height)^2$$

The entire list of dependent and independent variables is shown below. Information on General Health Status variables, Individual Healthy Behavior variables and Demographics are collected through the survey. Travel Behavior variables are imputed from data collected by the application.

Dependent Variables	Independent Variables
General Health Status	*Individual Healthy Behavior*
Physically Unhealthy Days	Diet
General Health Level	Physical Activity
Chronic Disease Prevalence	Regular Exercise
Limitation of Activity	Alcohol Use
Obesity Level (BMI)	Tobacco Use
	Demographics
	Age
	Gender
	Race
	Income
	Education Level
	Travel Behavior
	Commute Duration
	Commute Distance
	Commute Mode
	Time Spent in Travel
	Trips made per Day

Physically Unhealthy Days refers to the number of physically unhealthy days that the subject has experienced in the last one month. General Health Level is a self-assessed health status on a scale of 1 to 10 as reported as part of the online survey. Chronic Disease Prevalence and Limitation of Activity are binary variables that describe the existence of disability or a chronic disease in the subject. Obesity Level is measured as the body mass index of the individual.

Diet habits are categorized into how frequently the subject follows a diet. He/she might rarely follow a diet or follow a diet every day. Hence, Diet is specified as multiple dummy variables in the model. Race, Income and Education Level are also specified as multiple dummy variables. Physical Activity refers to the number of

hours spent on physical activities and sports in a week, while Regular Exercise is a binary variable describing whether the subject spends consistent time on exercise.

Chapter 5: Results and Discussion

5.1 Profile of Survey Respondents

All subjects of the pre-pilot study as well as the ongoing pilot are undergraduate and graduate students of the University of Maryland. A total of 149 students signed up through the online survey through the entire period of the pre-pilot as well the current ongoing pilot. However, only 46 students were able to collect the minimum required data for the project. The percentage of males was 65% (females were 35%). The median age of the population was 22 years. The following tables contain income and race distribution statistics of all entries received through the web-survey.
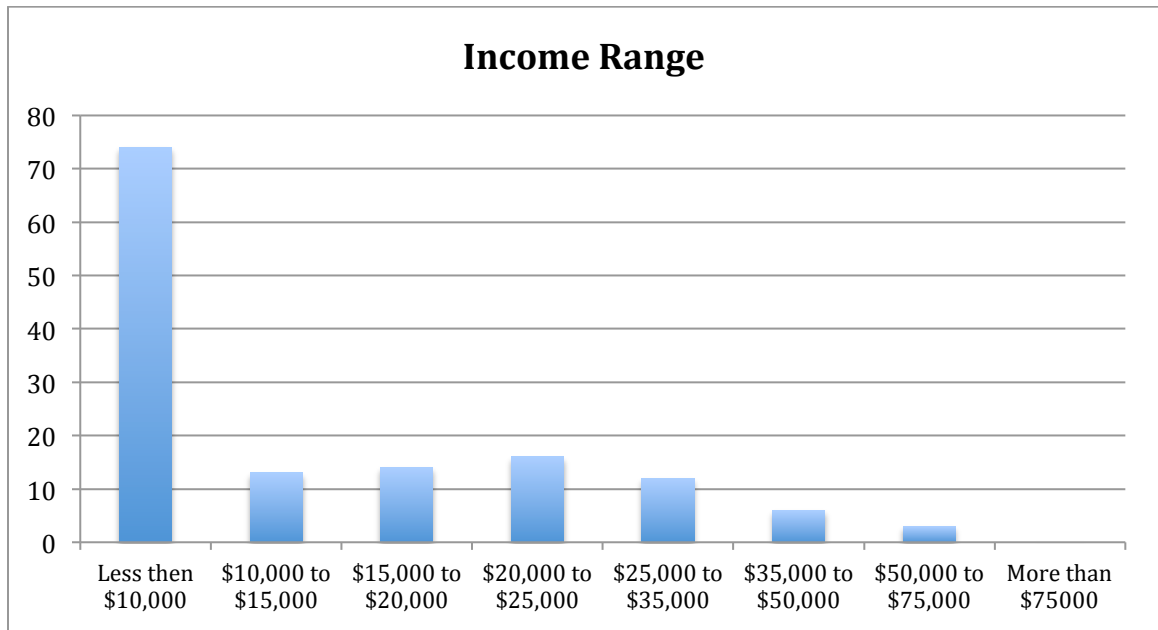
Figure 14: Income Distribution of Survey Respondents

Race

Frequency	Non-Hispanic White	Hispanic White	Black or African American	Asian	Native Hawaiian or Other Pacific Islander	American Indian or Alaska Native	Other
	49	9	6	76	0	1	8

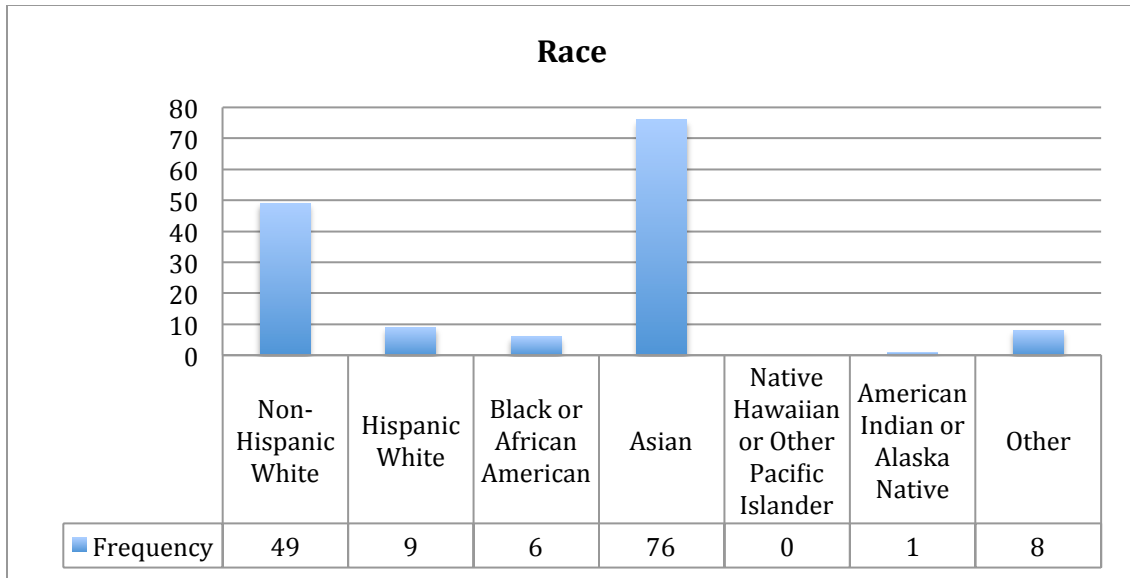**Figure 15: Race Distribution of Survey Respondents**
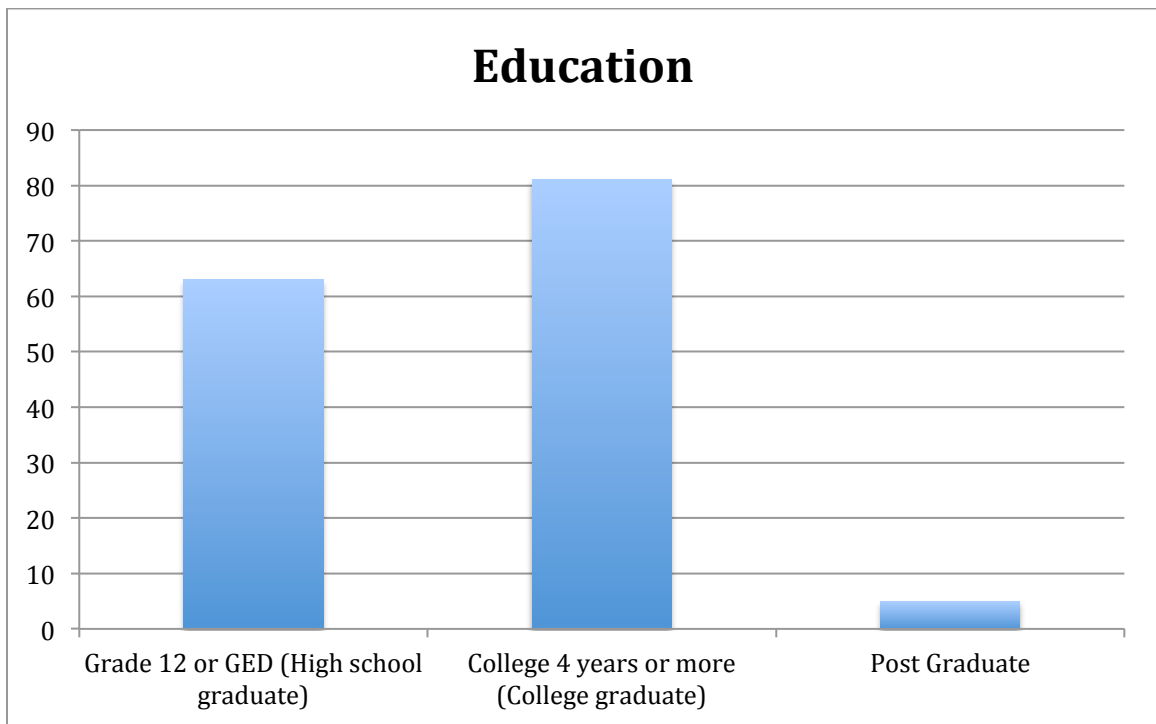
Education

Figure 16: Education Level Distribution of Survey Respondents

5.2 Multiple Regression Analysis

Presented below are the results of the multiple regression model specified in the Health and Travel Survey section with travel behavior data from the pilot study of University of Maryland student subjects. All of them being students, the income levels, education and age are pretty much similar across the whole sample. As a consequence of the sample's prevalent homogeneity and small size, most of the coefficients are not significant. The coefficients only for some variables are presented in the table below. Variables not presented, were either insignificant or were irrelevant with regards to the particular sample of the study.

Data on Chronic Disease Prevalence and Limitation of Activity (Disability) variables was collected through the Online Health and Travel Survey. However, the sample did not consist of subjects with any chronic disease and/or disability. Therefore, these variables were omitted from the model used here. Commute Distance, on the other hand was found to have a high correlation with Commute Duration and hence, Commute Distance was removed from the model.

Table 3: Regression Results (Unhealthy Days)

	Coefficients	Standard Error	LCL	UCL	t Stat	p-level
Intercept	-1.88	5.39	-15.15	11.38	-0.35	0.73
Bike	-0.42	1.44	-3.98	3.13	-0.29	0.77
Bus/Rail	0.08	1.34	-3.21	3.37	0.06	0.95
Auto	0.50	1.72	-3.73	4.72	0.29	0.77
Commute Time	**0.09**	**0.04**	**-0.02**	**0.19**	**2.05**	**0.05**
Trips / Day	**0.87**	**0.34**	**0.04**	**1.71**	**2.57**	**0.02**
Non-Idle Hours / Day	**-1.50**	**0.86**	**-3.62**	**0.62**	**-1.74**	**0.09**
Alcohol Use	**0.33**	**0.16**	**-0.07**	**0.72**	**2.03**	**0.05**
Physical Activity Hours / Week	**-0.07**	**0.05**	**-0.19**	**0.04**	**-1.56**	**0.13**
Rare Diet	0.59	1.03	-1.93	3.12	0.58	0.57

	Coefficients	Standard Error	LCL	UCL	t Stat	p-level
1 Day Diet	-0.04	1.10	-2.76	2.67	-0.04	0.97
3 Day Diet	-0.39	1.16	-3.24	2.47	-0.33	0.74
Everyday Diet	1.58	1.37	-1.80	4.95	1.15	0.26
Age	0.00	0.25	-0.63	0.62	-0.01	0.99
Gender	0.52	0.97	-1.88	2.92	0.53	0.60
Graduate	0.58	1.96	-4.25	5.40	0.29	0.77
Post Graduate	-0.25	2.49	-6.39	5.88	-0.10	0.92

Table 4: Regression Results (General Health)

	Coefficients	Standard Error	LCL	UCL	t Stat	p-level
Intercept	5.91	3.03	-1.54	13.36	1.95	0.06
Bike	0.30	0.81	-1.70	2.30	0.37	0.71
Bus/Rail	-0.09	0.75	-1.94	1.76	-0.12	0.91
Auto	-0.77	0.96	-3.14	1.61	-0.79	0.43
Commute Time	**-0.05**	**0.02**	**-0.11**	**0.01**	**-2.06**	**0.05**
Trips / Day	-0.19	0.19	-0.66	0.28	-0.99	0.33
Non-Idle Hours / Day	0.76	0.48	-0.43	1.94	1.57	0.13
Alcohol Use	-0.01	0.09	-0.23	0.21	-0.14	0.89
Physical Activity Hours / Week	**0.04**	**0.03**	**-0.03**	**0.10**	**1.45**	**0.16**
Rare Diet	-0.30	0.58	-1.72	1.12	-0.52	0.61
1 Day Diet	0.37	0.62	-1.16	1.89	0.59	0.56
3 Day Diet	0.89	0.65	-0.72	2.49	1.36	0.18
Everyday Diet	0.69	0.77	-1.20	2.59	0.90	0.37
Age	0.10	0.14	-0.25	0.45	0.67	0.51
Gender	0.36	0.55	-0.98	1.71	0.66	0.51
Graduate	-0.22	1.10	-2.93	2.49	-0.20	0.84
Post Graduate	-0.16	1.40	-3.61	3.29	-0.11	0.91

Table 5: Regression Results (Body Mass Index)

	Coefficients	Standard Error	LCL	UCL	t Stat	p-level
Intercept	17.65	7.90	-1.80	37.09	2.23	0.03
Bike	2.63	2.12	-2.58	7.84	1.24	0.22
Bus/Rail	2.39	1.96	-2.44	7.22	1.22	0.23
Auto	3.29	2.52	-2.90	9.48	1.31	0.20
Commute Duration	**0.09**	**0.06**	**-0.06**	**0.25**	**1.49**	**0.15**
Trips / Day	-0.24	0.50	-1.46	0.99	-0.47	0.64
Travel Hours / Day	-0.88	1.26	-3.99	2.22	-0.70	0.49
Alcohol Use	0.12	0.24	-0.46	0.70	0.51	0.61
Activity Frequency	-0.06	0.07	-0.23	0.10	-0.92	0.37
Rare Diet	-0.36	1.50	-4.07	3.34	-0.24	0.81
1 Day Diet	-0.12	1.61	-4.10	3.85	-0.08	0.94
3 Day Diet	1.07	1.70	-3.12	5.25	0.63	0.54
Everyday Diet	-1.25	2.01	-6.19	3.70	-0.62	0.54
Age	0.26	0.37	-0.65	1.18	0.71	0.48

Gender	0.75	1.43	-2.77	4.26	0.52	0.61
Graduate	-1.97	2.87	-9.04	5.11	-0.68	0.50
Post Graduate	-3.66	3.65	-12.66	5.33	-1.00	0.32

Table 6: Regression Results (Combined)

	Coefficients	Standard Error	
Intercept (Unhealthy Days)	-1.88	5.39	*R = 0.62*
Bike	-0.42	1.44	*R2 = 0.39*
Bus/Rail	0.08	1.34	
Auto	0.50	1.72	
Commute Duration	**0.09**	**0.04**	
Trips / Day	**0.87**	**0.34**	
Non-Idle Hours / Day	**-1.50**	**0.86**	
Alcohol Use	**0.33**	**0.16**	
Physical Activity Hours / Week	**-0.07**	**0.05**	
Intercept (General Health)	5.91	3.03	*R = 0.60*
Bike	0.30	0.81	*R2 = 0.36*
Bus/Rail	-0.09	0.75	
Auto	-0.77	0.96	
Commute Duration	**-0.05**	**0.02**	
Trips / Day	-0.19	0.19	
Non-Idle Hours / Day	0.76	0.48	
Alcohol Use	-0.01	0.09	
Physical Activity Hours / Week	**0.04**	**0.03**	
Intercept (BMI)	17.65	7.90	*R = 0.50*
Bike	2.63	2.12	*R2 = 0.24*
Bus/Rail	2.39	1.96	
Auto	3.29	2.52	
Commute Duration	**0.09**	**0.06**	
Trips / Day	-0.24	0.50	
Non-Idle Hours / Day	-0.88	1.26	
Alcohol Use	0.12	0.24	
Physical Activity Hours / Week	-0.06	0.07	

In the above table, coefficients that are significant are in bold. Two coefficients that have a p-value close to 0.1 have also been made bold. All the significant variables have coefficient values, which are synonymous with what one would expect theoretically.

The number of physically unhealthy days one would have if his commute duration were longer would be higher. Similarly, subjects who consume alcohol are expected to feel physically sick more number of times than who do not and the results agree with it. Number of trips taken per day, and hours spent in physical activities during the week also has supporting coefficient values. The coefficient of Time Spent in Travel, however, suggests that as the subject spends more time in travelling, the number of physically unhealthy days shall decrease. This is true when looked at smaller values of overall time spent in travel. For example, in the base case scenario of the model, the subject doesn't travel at all. The negative value of the coefficient indicates that, compared to a subject who stays at home/work and doesn't move about socializing and participating in physical activities, the individual who travels will have better physical health.

Significant variables for General Health Level also support the theoretical understanding of the relationship between the two. As the commute duration increases, the subject would feel more stressed mentally and physically, thereby bringing down his self-assessed health level. Once again, time spent in travel presents a strange case, which is true when compared with people who don't travel at all. There were no significant variables, however, for Obesity Level.

Even though the coefficients of the Commute Mode are insignificant, it is hard not to notice their values for all three dependent variables. All values are as one would expect; more number of physically unhealthy days for people who commute via

automobiles along with the least general health level and considerate increase in the body mass index. Other variables shown in the table, such as Alcohol Use, Physical Activity and number of trips taken per day also have values that support theory.

5.3 Discussion

It is important to discuss the following aspects, which further limit this exploratory research and the significance of the results. Future researchers should consider implementing models which accommodate the below thoughts, or at least, address their importance in trying to achieve realistic travel behavior models.

5.3.1 Gradient for Walking and Cycling

Cycling is more economical than walking on level ground. This ranking however does not hold when moving on a gradient. Ardigo, Sabiene and Minetti (2003) have investigated this ranking to find that cycling was the mode of choice only below 10-15% gradient, while above it walking was the least expensive locomotion type. Mode choice of travelers is hence dependent on the gradient of the route taken. This study, however, doesn't take into account this effect of gradient on the choice of mode in travelling to work or school everyday. It is important to highlight that the choice of walking and biking modes indicates that the subject may be willing to expend considerable physical effort, especially when the route includes positive gradient. Particularly in the case of University of Maryland, where this study has been conducted, many routes towards places of work and school have significant gradient

attributes. This may have influenced the choice of the subjects in choosing other comfortable modes.

5.3.2 Causality

The results indicate that subjects who commute using active modes have a greater overall health level compared to those who do not. Subjects who are observed to have alcohol and tobacco use are also observed to have lower health levels. It is important to understand the causality of these variables on the dependent health variable. Past studies have tried to understand the direction of this causal relationship. People who spend a lot of time using active travel modes may be choosing these modes because of their good health and good physical fitness. On the other hand, use of active travel modes may help increase the general health levels of individuals. This relationship needs to be understood at a deeper level before making conclusive decisions.

5.3.3 Avoiding Sample Bias

In principle, a survey could cover all individuals in the interested section of population. If all subjects being interviewed could provide perfect answers, we could measure all indicators with perfect accuracy. However, due to a limited amount of time and money, only a sample of the population is interviewed. To avoid any sample bias, a large enough sample needs to be considered for estimates to be precise. The results of this study are limited due to the sample homogeneity and size. The health survey received a total of 149 respondents from a huge student base in the University of Maryland. This low response indicates a selection bias, which may have resulted

from the type of questions asked in the survey. Health-related questions, especially those inquiring the obesity level and the use of alcohol or tobacco often result in potential subject opting out of surveys and hence creating a selection bias.

Only 46 subjects were able to provide the required amount of data of 4 weeks from the smartphone application. This low ratio of successful data responses shows that, once again, a bias may have been created among the overall survey respondents. Subjects who did not want to be tracked would have selected themselves out of the study. Monetary incentives though provided may however not be enough to attract subjects to participate in the survey.

Below are some practices to be followed to avoid any potential sample bias:
1. Do not choose subjects exclusively from particular groups
2. Do not restrict the sample to households or individuals from a certain residential location
3. Have a large enough sample to reduce the potential to have a biased sample
4. Accommodate for sampling errors that may have resulted in case of observed sample bias

5.3.4 Dependent Health Variables

The model developed in this study uses a combination of three health variables, General Health Level, Number of Physically Unhealthy Days and Obesity Level. The relative importance of these variables is however not explained. Hence, observations

cannot be made about one subject having a better overall health level than another.

Few studies, if any, have outlined standard health measures for individuals. National

surveys such as Healthy People have listed a combination of health measures which

have to be reported together to make any conclusions regarding the health status of

individuals:

1. General Health Level

2. Number of Physically Unhealthy Days

3. Chronic Disease Prevalence

4. Limitation of Activity or Disability

5. Obesity Level

Attributes 3 and 4 can be measured less frequently than the other, however, their

values have higher relative importance. Future models have to be designed keeping

this in mind so as to develop reliable health and travel behavior policies.

Chapter 6: Conclusions

The purpose of this study was to explore the potential relationship between the individual's travel behavioral patterns and his/her healthy living habits through empirical results. There is a lack of research, which links personal level health data to the individual's travel behavior, and the objective of this study to address this gap. For this purpose, a smartphone application was developed that is capable of capturing rich travel behavior data for long periods without the need of user intervention. A survey was also designed that can collect information about the individual's health behavior. Finally, a model was specified that helps explore the linkage between general health measures and everyday health habits as well as the travel patterns of users.

However, the limitations of this research are well understood. A low value for the R-squared measure shows that the model is weak. Hence, future research should experiment with models such as logarithmic and exponential to achieve a better dependence of the "y" variable on the "x" variable. The sample of the pilot study was homogenous and small. This resulted in most variables showing insignificance in contributing the health measure of the individual. This is part of an ongoing research and future studies shall overcome this aspect. A full-scale study shall be designed and conducted where participants shall be tracked for long periods of up to 2 years. Data

from this study shall be analyzed against their health information to analyze long-term negative and positive impact on health.

The application itself, serves as a great platform to generate data products as we move towards increased collaboration between the software industry and transportation industry. Improvements will be made to Travel Helper to facilitate real-time and continuous location transmission to server databases. This will enable applications in the field of security and emergency services.

Appendix A: Health and Travel Survey

Demographics (Socio-Economic Variables)

1. Age
   a. _ _ _ _ yrs

2. Gender
   a. Male
   b. Female

3. Which one of the groups would you say best represents your race?
   a. Non-Hispanic White
   b. Hispanic White
   c. Black or African American
   d. Asian
   e. Native Hawaiian or Other Pacific Islander
   f. American Indian or Alaska Native
   g. Other _____

4. What is your marital status?
   a. Married
   b. Divorced
   c. Widowed
   d. Separated
   e. Never Married
   f. A member of an unmarried couple

5. What is the highest grade or year of school you completed?
   a. Never attended school
   b. Grades 1 through 8 (Elementary)
   c. Grades 9 through 11 (Some high school)
   d. Grade 12 or GED (High school graduate)
   e. College 1 year to 3 years (Some college or technical school)
   f. College 4 years or more (College graduate)

6. What is your current employment status?
   a. Yes, I am currently employed
   b. No, I am currently unemployed

7. Which category does your income from all income sources fall in?
   a. Less than 10,000
   b. 10,000 to 15,000
   c. 15,000 to 20,000

     d. 20,000 to 25,000
     e. 25,000 to 35,000
     f. 35,000 to 50,000
     g. 50,000 to 75,000
     h. More than 75,000

8. About how much do you weigh without shoes in pounds?
     a. _ _ _ _ pounds

9. About how tall are you without shoes?
     a. _ _ feet _ _ inches

General Health (Self-assessed)

10. What would you say your general health is on a scale of 1-10, 1 being poor and 10 being excellent?
     a. 10
     b. 9
     c. 8
     d. 7
     e. 6
     f. 5
     g. 4
     h. 3
     i. 2
     j. 1

11. For how many days during the past 30 days was your physical health not good?
     a. __ Number of days
     b. None

12. Based on your physical fitness, how often do you feel tired?
     a. Twice a week or more
     b. About once a week
     c. About once a month
     d. Once in a few months
     e. Rarely
     f. Never

13. Based on your most recent doctor's visit, how often do you feel weak in your upper respiratory system like having a running nose?
     a. Twice a week or more
     b. About once a week
     c. About once a month

d. Once in a few months
e. Rarely
f. Never

14. How often do you feel stressed or have difficulty falling asleep?
    a. Twice a week or more
    b. About once a week
    c. About once a month
    d. Once in a few months
    e. Rarely
    f. Never

15. Has a doctor, nurse or other health professional ever told you that you had any of the following?
    a. Heart Attack also called myocardial infarction
    b. Angina or Coronary heart disease
    c. Stroke
    d. Asthma
    e. Skin Cancer
    f. Other types of Cancer
    g. Arthritis
    h. (COPD) Chronic obstructive pulmonary disease
    i. Diabetes

Exercise (Physical Activity)

16. During the past month, other than your regular job, did you participate in any physical activities or exercises such as running, calisthenics, golf, gardening, or walking for exercise?
    a. Yes
    b. No

17. What type of physical activity or exercise did you spend the most time during the past month?
    a. [Describe the activity here]

18. On average, how much time (in hours per week) do you spend on this activity?
    a. _ _ hours

19. What other type of physical activity gave you the next most exercise during the past month?
    a. [Describe the activity here]
    b. No other activity

Disability

20. Do you have any health problem that requires you to use special traveling equipment, such as a cane, a wheelchair, a special bed, a special telephone, or a special car?
   a. Yes
   b. No

Food and Dietary Habits

21. How often do you try to stick to a regular diet or food habit?
   a. I follow a strict diet everyday
   b. I follow a diet for more than 3 days a week
   c. I follow a diet on one day a week
   d. I rarely follow any food habit or diet
   e. I do not follow any diet

22. Would you call your diet a healthy diet?
   a. Yes
   b. No

Alcohol Use / Drinking

23. During the past 30 days, how many days per week or month did you have at least one drink of any alcoholic beverage such as beer, wine, a malt beverage or liquor?
   a. __ Days per week
   b. __ Days in past 30 days
   c. No drinks in past 30 days

24. One drink is equivalent to a 12-ounce beer, a 5-ounce glass of wine, or a drink with one shot of liquor. During the past 30 days, on the days when you drank, about how many drinks did you drink on the average? NOTE: A 40-ounce beer would count as 3 drinks, or a cocktail drink with 2 shots would count as 2 drinks.
   a. __ Number of drinks

25. Considering all types of alcoholic beverages, how many times during the past 30 days did you have X (X = 5 for men, X = 4 for women) or more drinks on an occasion?
   a. __ Number of times
   b. None

26. During the past 30 days, what is the largest number of drinks you had on any occasion?
   a. __ Number of drinks

Tobacco Use / Smoking

27. Have you smoked at least 100 cigarettes in your entire life?
    a. Yes
    b. No

28. Do you now smoke cigarettes every day, some days, or not at all?
    a. Everyday
    b. Some days
    c. Not at all

29. How long has it been since you last smoked a cigarette, even one or two puffs?
    a. Within the past month
    b. Within the past 3 months
    c. Within the past 6 months
    d. Within the past year
    e. Within the past 5 years
    f. Within the past 10 years
    g. 10 years or more

Appendix B: Sample Application Code

```java
/**
 * This Receiver class is used to listen for Broadcast Intents that announce
 * that a location change has occurred. This is used instead of a
LocationListener
 * within an Activity is our only action is to start a service.
 */
public class LocationChangedReceiver extends BroadcastReceiver {

    protected static String TAG = "LocationChangedReceiver";

    /**
     * When a new location is received, extract it from the Intent and
use
     * it to start the Service used to update the list of nearby
places.
     *
     * This is the Active receiver, used to receive Location updates
when
     * the Activity is visible.
     */
    @Override
    public void onReceive(Context context, Intent intent) {
      String locationKey = LocationManager.KEY_LOCATION_CHANGED;
      String providerEnabledKey = LocationManager.KEY_PROVIDER_ENABLED;
      if (intent.hasExtra(providerEnabledKey)) {
        if (!intent.getBooleanExtra(providerEnabledKey, true)) {
          Intent providerDisabledIntent = new
Intent(PlacesConstants.ACTIVE_LOCATION_UPDATE_PROVIDER_DISABLED);
          context.sendBroadcast(providerDisabledIntent);
        }
      }
      if (intent.hasExtra(locationKey)) {
        Location location =
(Location)intent.getExtras().get(locationKey);
        Log.d(TAG, "Actively Updating place list");
        Intent updateServiceIntent = new Intent(context,
PlacesConstants.SUPPORTS_ECLAIR? EclairPlacesUpdateService.class :
PlacesUpdateService.class);

updateServiceIntent.putExtra(PlacesConstants.EXTRA_KEY_LOCATION, location);
        updateServiceIntent.putExtra(PlacesConstants.EXTRA_KEY_RADIUS,
PlacesConstants.DEFAULT_RADIUS);

updateServiceIntent.putExtra(PlacesConstants.EXTRA_KEY_FORCEREFRESH, true);
        context.startService(updateServiceIntent);
      }
```

References

2008 Public Transportation Fact Book. Washington, DC: American Public Transportation Association; 2008. Available at: http://www.apta.com/resources/statistics/ Documents/FactBook/APTA_2008_Fact_Book.pdf. Accessed September 2, 2010.

Angevaren M, Aufdemkampe G, Verhaar HJ, Aleman A, Vanhees L. Physical activity and enhanced fitness to improve cognitive function in older people without known cognitive impairment. Cochrane Database Syst Rev. 2008;16(3):CD005381.

Auld, J. A. and A. Mohammadian (2009). Framework for the development of the Agent-based Dynamic Activity Planning and Travel Scheduling (ADAPTS) model. Transportation Letters: The International Journal of Transportation Research. 1(3), 243-253.

Bar-Gera, H. (2007). Evaluation of a cellular phone-based system for measurements of traffic speeds and travel times: A case study from Israel. Transportation Research Part C, 15(6):380-391.

Battelle, Transportation Division (1997) "Lexington Area Travel Data Collection Test: Final Report", Office of Highway Information Management and Office Technology Applications, Federal Highway Administration, Washington, D.C.

Besser LM, Dannenberg AL. Walking to public transit: steps to help meet physical activity recommendations. Am J Prev Med. 2005;29(4):273–280.

Bohte, W. and K. Maat (2009). Deriving and validating trip purposes and travel modes for multi-day GPS-based travel surveys: A large-scale application in the Netherlands. Transportation Research Part C: Emerging Technologies, 17(3), pp. 285 – 297.

Burbidge, S. K. & Goulias, K. G. (2008). Active Travel Behavior. Transportation Letters 2008.

Bricka, S. and C.R. Bhat (2006), "A Comparative Analysis of GPS-Based and Travel Survey-based Data," Transportation Research Record, Vol. 2, pp. 9-20.

Bricka, S. and C.R. Bhat (2008), "Using Global Positioning System Data to Inform Travel Survey Methods", Transportation Research Board, Vol. 2, Issue 42, pp 89-93.

Caldwell, L. L. & Smith, E. A, (1988). Leisure: An Overlooked Component of Health Promotion. Canadian Journal of Public Health, 79 (April/May), 44-48.

Chalip, L., Thomas, D. R., and Voyle, J. (1992). Sport, Recreation and Well Being. In D. R.

Chapman J, Frank L. SMARTRAQ: Integrating Travel Behavior and Urban Form Data to Address Transportation Problems in Atlanta. Atlanta, GA: Georgia Tech Research Institute, Georgia Department of Transportation, US Department of Transportation; 2004, Available at http://www.act trans.ubc.ca/smartraq/files/GDOT_final_report.pdf, accessed 14 August 2008.

Charlton, B., Sall, E., Schwartz, M. & Hood, J. (2011). Bicycle Route Choice Data Collection using GPS-Enabled Smartphones, Presented at the Transportation Research Board Annual Meeting 2011.

Chen, J., Bierlaire, M., and Flötteröd, G. (2011). Probabilistic multi-modal map matching with rich smartphone data. Proceedings of the Swiss Transportation Research Conference (STRC) May 11-13, 2011 , 2011.

Li Chen, Cynthia Chen, Raghavan Srinivasan, Claire E. McKnight, Reid Ewing, and Matthew Roe. Evaluating the Safety Effects of Bicycle Lanes in New York City. American Journal of Public Health: June 2012, Vol. 102, No. 6, pp. 1120-1127.

Center for Disease Control. (1996). "Vital and Health Statistics: Leading causes of death by age, sex, race, and Hispanic origin." Report No. 29, Atlanta, USA.

Chen, Q. and Y. Fan (2012). Smartphone-Based Travel Experience Sampling and Behavior Intervention. TRB Annual Meeting, January 2012.

Coleman, D. J., and Iso-Ahola, S. E. (1993). Leisure and Health: The Role of Social Support and Self Determination. Journal of Leisure Research. 25(2), 111-128.

Cottrill, C.D. et. al. (2013). The Future Mobility Survey: Experiences in developing a smartphone-based travel survey in Singapore. Transportation Research Board, January 2013

Doherty, S.T., N. Noël, M. Lee-Gosselin, C. Sirois, M. Ueno and F. Theberge (1999). Moving Beyond Observed Outcomes: Integrating Global Positioning Systems and Interactive Computer-based Travel Behaviour Surveys. Proceedings of the Transportation Research Board Conference on Personal Travel: The Long and Short of It, Washington, D.C., July 1999.

Ewing, R. et. al. (2003). Relationship between urban sprawl and physical activity, obesity and morbidity. American Journal of Health Promotion. Sep-Oct; 18(1):47-57.

Gonzalez et. al. (2009). Automating mode detection for travel behavior analysis by using global positioning systems-enabled mobile phones and neural networks. IET Intelligent transportation systems.

Handy, Suzan L., and Kelly J. Clifton. 2007. Planning and the built environment: Implications for obesity prevention. In Handbook of obesity prevention. A resource for health professionals, ed. S. Kumanyika and Ross Brownson, 167-88. New York, NY: Springer.

Hendriksen I, Zuiderveld B, Kemper H, Bezemer P. Effect of commuter cycling on physical performance of male and female employees. Med Sci Sports Exerc. 2000; 32(2):504--510.

Janelle, D.J. (2004). "Impact of Information Technologies." In S. Hanson, & G. Giuliano (Ed.), The Geography of Urban Transportation (3 ed.), Guilford Press, New York, USA, 86-112.

Jariyasunant, J., Carell, A., et al. (2011). The Quantified Traveler: Using Personal Travel Data to Promote Sustainable Transport Behavior. Earlier Faculty Research, University of California Transportation Center, UC Berkeley, 2011.

Katz, G., Knobler, H. Y., Laibel, Z., Strauss, Z., Durst, R. (2002). Time zone change and major psychiatric morbidity: The results of a 6-year study in Jerusalem. Comprehensive Psychiatry. 43, 37-40.

Leon, A. S., J. Connett, D. R. Jacobs, and R. Rauramaa. 1987. Leisure-time physical activity levels and risk of coronary heart disease and death: The Multiple Risk Factor Intervention Trial. *Journal of the American Medical Association* 258 (17): 2388-95.

Litman, T., 2002. If health matters: Integrating public health objectives in transportation planning. Victoria, B.C., Canada. www.vtpi.org/health.pdf (accessed August 2003).

Littman AJ, Kristal AR, White E. Effects of physical activity intensity, frequency, and activity type on 10-y weight change in middle-aged men and women. International Journal on Obesity (London). 2005;29(5):524--533.

Low, J. A., & Chan, D. K. (2002). Air Travel in older people. Age and Ageing, 31, 17-22.

McIntosh, I. B., Swanson, V., Power, K. G., Raeside, F., & Dempster, C. (1998). Anxiety and health problems related to air travel. *Journal of Travel Medicine*, **5**, 198-204.

Morabia A, Amstislavski PN, Mirer FE, et al. Air pollution and activity during transportation by car, subway, and walking. Am J Prev Med. 2009;37(1):72–77.

Morabia, A., Mirer, F. E., Markowitz, S. B. et. al. (2010). Potential Health Impact of Switching From Car to Public Transportation When Commuting to Work. *American Journal Of Public Health*, *100*(12), 2388-2391.

Morris, J. N., D. G. Clayton, M. G. Everitt, A. M. Semmence, and E. H. Burgess. 1990. Exercise in leisure time: Coronary attack and death rates. *British Heart Journal* 63:325-34.

Ogilvie D, Egan M, Hamilton V, Petticrew M. Promoting walking and cycling as an alternative to using cars: systematic review. BMJ. 2004;329(7469):763.

Online Travel Survey Manual: www.travelsurveymanual.org

Pas, E.I., and F.S. Koppelman. (1986). "An Examination of the Determinants of Day to Day Variability in Individuals' Urban Travel Behavior." *Transportation*. 13(2), 183-200.

Saelens, B.E., Sallis, J.F., and Frank, L.D. (2003). "Environmental correlates of walking and cycling: findings from the transportation, urban design, and planning literatures." *Annals of Behavioral Medicine*. 25, 80-91.

Saelensminde K. Cost-benefit analyses of walking and cycling track networks taking into account insecurity, health effects and external costs of motorized traffic. Transporation Research Part A: Policy & Practice. 2004;38(8):593—606.

Sallis JF, Frank LD, Saelens BE, Kraft MK. Active transportation and physical activity: opportunities for collaboration on transportation and public health research. Transport Res A: Pol Pract. 2004;38(4):249–68.

Sallis JF, Kraft K, Linton LS. How the environment shapes physical activity: a transdisciplinary research agenda. Am J Prev Med. 2002;22(3):208.

Sandvik, L., J. Erikssen, E. Thaulow, G. Erikssen, R. Mundal, and K. Rodahl. 1993. Physical fitness as a predictor of mortality among healthy, middle-aged Norwegian men. *New England Journal of Medicine* 328:533-37.

Stopher, P., C. FitzGerald and M. Xu (2007). Assessing the accuracy of the Sydney Household Travel Survey with GPS. Transportation. Vol. 34, Iss. 6. P. 723-741.

Thomas & A. Veno (Eds.), Psychology and Social Change (132-156). Palmertson North, New Zealand: Dunmore Press.

Transportation Research Board, and Institute of Medicine of the National Academies. 2005. Does the built environment influence physical activity? Examining the evidence. Washington, DC: Transportation Research Board.