# DECIPHERING ACTIVITY PATTERNS USING THE TIME-GEOGRAPHY FRAMEWORK: A CASE STUDY OF OKLAHOMA STATE UNIVERSITY, STILLWATER CAMPUS

By

LEONARD SITJI BOMBOM

Bachelor of Science in Geography and Planning
University of Jos
Jos, Nigeria
1994

Master of Arts in Geography
University of Northern Iowa
Cedar Falls, Iowa
2005

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
December, 2013

DECIPHERING ACTIVITY PATTERNS USING THE TIME-GEOGRAPHY FRAMEWORK: A CASE STUDY OF OKLAHOMA STATE UNIVERSITY, STILLWATER CAMPUS

Dissertation Approved:

Hongbo Yu (PhD)

Dissertation Adviser

Jon Comer (PhD)

Allen Finchum (PhD)

Tanya Finchum (PhD)

ACKNOWLEDGEMENTS

Now unto the King Eternal, Immortal, Invisible, The Only Wise God, be Honor and Glory forever and ever (1 Timothy 1:17)

I would like to acknowledge my indebtedness to the Department of Geography, Oklahoma State University and its Chair, Dr. Dale Lightfoot, its faculty and staff for their support throughout the years. I am very grateful for the assistantship, the camaraderie, and the knowledge I have gained in both academics and life lessons. Thank you.

I would like to express my special thanks to my advisor, Dr. Hongbo Yu for his support, patience, encouragement, guidance and mentorship in the face of many difficulties. I wished there was a word or phrase more expressive than a mere “Thank You.”

I also wish to thank the members of my committee for their diligence, guidance, patience and support through this arduous journey. I am very grateful.

I would also like to thank my friends for their support and understanding. I want to express my thanks especially to Dr. Elisha Jasper Dung, who initiated the journey; Dr. Sunday D. Goshit and his wife, Regina and children, who took me in and held my hands along the way; Dr. Tim Strauss of University of Northern Iowa, Cedar Falls, who first gave me a scholarship from half a world away, cultivated and re-focused my academic curiosity; my cousin, Rev. Fr. Justin Podos, Nde Titus Kassem and my many friends, colleagues, roommates and fellow students whom I met along the way. You will all be remembered with fondness and gratitude.

To my sister, Gladys, who had always been the “being” of my pursuits
To dad, Augustine, who had always been the beacon
To mum, Paulina, who had always been the bedrock
Please take me in again
From out of the rain and dust
From out of the penury of spirit
From out of soulless oblivion
Into your embrace of love
Unconditional
Again
I Love You All

Name: LEONARD SITJI BOMBOM

Date of Degree: DECEMBER, 2013

Title of Study: DECIPHERING ACTIVITY PATTERNS USING THE TIME-GEOGRAPHY FRAMEWORK: A CASE STUDY OF OKLAHOMA STATE UNIVERSITY, STILLWATER CAMPUS

Major Field: GEOGRAPHY

Abstract: Human societies are organized around activities. Every individual participates in certain activities at all times, which are organized in both time and space. Therefore to understand how human societies are organized, it is important to understand how human activities are organized. Traditionally, methods of activity analysis have employed transportation planning, structural equation, simulation and other computational models. Most of these models use trips and trip making as the bases for activity analysis. Current practice however recognizes activities as the focus of activity analysis since trips are derived from the demand of people to participate in activities. This and other shortcomings of the traditional models have resulted in the search for new perspectives and tools to analyze activity patterns. Hagerstrand's time-geography presents an elegant framework to study and understand activity patterns through several important and clearly defined concepts such as stations, space-time paths, space-time prisms, and activity constraints. One of the most important attribute of this framework is its capacity to capture and represent the sequence of human activities in simple but effective ways. The space-time path is a three-dimensional (3D) trajectory that represents the locations of human activities in a two-dimensional (2D) plane and captures the time and sequence of activity participation through the third dimension - time. Activity constraints also provide an understanding of the necessary conditions needed for human activity to take place. Unfortunately, only a few studies have developed methods of activity analysis using this framework. This study adopts the time-geography framework and concepts to develop two new methods to decipher activity patterns. The daily activity schedule fragmentation index (DASFI) examines the propensity of individuals to organize their activities in chains or fragments. The daily activity intensity similarity index (DAISI) measures the degree of similarity between the activity profiles of people. Both indices can be used in cluster analysis to derive clusters which group individuals with similar characteristics in their activity patterns. A case study with the population at Oklahoma State University – Stillwater Campus proves useful in understanding how people organize their activities and could help in planning geographical space to meet the activity needs of people.

TABLE OF CONTENTS

Chapter Page

LIST OF TABLES

LIST OF FIGURES

# CHAPTER I

# INTRODUCTION

## 1.1 Introduction

Human societies revolve around different activities such as work, shopping, recreation, etc. These activities are usually distributed across space; therefore, participation in these activities requires movement from one activity point to another. The movement process to activity locations and participation in the activities both take some time (Pipkin, 1995). Consequently space and time are the basic components of activity participation.

Activities are engaged in at different locations, at different times and for different lengths of time (duration) depending on the type of activity, the purpose of activity, and environmental, social and economic constraints. These and other spatial and temporal constraints ultimately determine human activity behavior or pattern. For example, some shopping malls have set opening and closing times and any shopping activity at such locations can only be possible within the temporal window available for business activities. The operating times at these malls therefore constrain shoppers' behavior to either shop at the mall before the closing time or shop elsewhere. The continuous interaction among people in space therefore produces discernible patterns, which are important to understanding the complex spatiotemporal organization of society.

Conceptually, time and space are inextricably woven together in activity participation (Hagerstrand, 1970). For example, a mere movement from point A to point B (spatial component) consumes some time, irrespective of the distance between the two locations. A decision to partake in an activity also involves the considerations of the availability of an opportunity to participate in the activity, the time duration for activity engagement, and the temporal "distance" to the activity location. The significance of time is underscored by the fact that engaging in a particular activity engenders a forfeiture of participation in some other activities at the same time due to the nature of human indivisibility (Thrift, 1977a). If the window of temporal opportunity to engage in an activity is too tight for a given activity, there is a high probability that another activity that fits the allowable time period may be chosen, thus forfeiting a preferred activity alternative for temporal convenience. Yet despite the immense and acknowledged importance attached to time in activity selection, scheduling and participation, only recently has time been explicitly modeled in activity analysis (Shaw, 2006).

First, time was assumed to be an inevitable corollary of human activity engagement and therefore taken for granted. Second, it was much easier to represent the spatial component (location) of activities rather than their temporal elements (Langran, 1992; O'Sullivan, 1995; Yuan, 1996; Yu, 2006). This omission has prompted calls for a deeper examination of the role of time in human activity research (Hagerstrand, 1970; Shaw, 2006). Recent developments of geographic information systems (GIS) and science (GISc) open new and promising opportunities for analyzing activity patterns. Classical GIS is designed to accurately and effectively represent spatial data. However, its representation of time is only as a series of snapshots of spatial elements captured in a sequence of temporal representation. Such representation fails to capture the dynamic processes of change and motion that the time element embodies (Langran, 1992). The problem lies with both the retention of the historical cartographic framework by GIS (O'Sullivan, 1995) and the universality of time, which is not divisible in the same manner that space can be compartmentalized (Thrift, 1977a). These and other factors have added to the difficulty of modeling time.

Recent developments in time research, however, have improved the analysis and modeling of human behavior by explicitly incorporating the time dimension in human activity framework. One of these frameworks is *time geography*, proposed and developed by Hagerstrand (1970) and the Lund School of Geography in Sweden. Time geography examines the inter-relationship between activities in space and time and their constraints (Miller, 2004; Yu, 2006). The fundamental tenet of time geography is that all human activities have both spatial and temporal dimensions that cannot be meaningfully separated. This basic attribute presents opportunities and at the same time it poses constraints to human activity participation.

For quite some time, time geography provided a conceptual basis for activity pattern analysis with little operational analytical capacity (Miller, 2004; Yu, 2006). However, major developments in the field (e.g., Miller, 1991, 1999; Kwan, 2000b, 2003; Yu and Shaw, 2004; Yu, 2006) have ignited a lot of interest in time geography and have opened new frontiers for the application of the time-geography framework. In spite of its enormous potential for activity analysis, only limited work has been done to examine activity patterns using time-geography principles and concepts. This study explores this area of research.

## 1.2 Research Problem

Time geography possesses enormous potential for activity pattern analysis and modeling. The concepts of stations, space-time paths and space-time prisms hold a lot of promise for identifying and/or exposing individual and collective activity patterns, and the concept of activity constraints help to explain human activity behavior. Each space-time path captures the trajectory of an individual's activities by reconstructing the time duration and location of activities undertaken by the person. Identifying similar sequences of individual activities is important to unveil any hidden or underlying patterns, to understand the impact of activity locations and the social, economic, environmental constraints influencing human activity participation.

One of the main goals of studying human activity patterns is to improve the understanding and prediction of spatial interactions (Scott, 2006). For example, the extent to which the socio-economic and demographic backgrounds of people affect where, when and how often they engage in activities is important to planning and organizing human activity distribution as well as explaining human interactions in the society. Such an undertaking becomes more significant in the face of increasing advancements in technologies (e.g., virtual space) that have drastically influenced the perception of and participation in activities across space. Bhat and Lawton (2000) contend that some of the major research directions in activity analysis include the examination of differences in types of interactions in space and time for different activity types, effect of rigid temporal constraints on activity participation attributes, and the impact of space-time interactions on activity sequencing choices. These research foci are aimed towards unraveling activity patterns and they all fit neatly within the framework of time geography. However, very few studies have attempted to comprehensively model activity patterns within an explicitly time-geography framework (Scott, 2006).

The time-geography framework has been demonstrated to be suitable for analyzing human activity patterns. First, the approach focuses on individuals' use of their knowledge, objects and tools, and their social relations to participate in activities (Ellegard, 1999). Second, the approach provides an elegant representation for the trajectory of an individual's activity participations through a space-time path concept. A space-time path accurately captures the location of the activity, the time and duration of the activity, movement between activity locations and importantly, the sequence in which the activities are undertaken. Third, the framework accounts for constraints that influence activity participation by individuals. All these encompass the realm of activity analysis. The framework therefore possesses the requisite conditions necessary for effectively unraveling the human activity conundrum and enhancing our understanding of human activity patterns.

In spite of this fertile foundation, limited progress has been made in the past to develop practical computational models to operationalize the time-geography framework and concepts and apply them to facilitate activity analysis for a better understanding of human activity patterns in the real world. In response to the increasing research needs in this area, this study targets the development of time-geography-based analytical methods, which will put individual activities in their spatial and temporal contexts when analyzing and understanding human activity patterns, and applying them in practical cases.

The Oklahoma State University (OSU), Stillwater campus, provides an effective site for such a study. The university is ideal for studying activity patterns because it has a controlled environment with a more clearly demarcated activity schedule for different groups than is found in the general population. This provides a very good setting to acquire data, develop and test new indices because they could be easily validated by the known activity schedules of the population in the university.

## 1.3 Research Objectives

Specifically, the objectives of this research include:

1. Gain a better understanding of individuals' activity patterns by taking a time-geography approach to examining the types, locations, and linkages (i.e., movement characteristics between activity locations) of activities involved in the individuals' daily lives.
2. Develop new time-geography based indices which can be used to facilitate the exploration and identification of certain activity patterns presented in individuals' daily activities. The indices will be able to incorporate the spatial and temporal context information of the activities and reflect the spatio-temporal characteristics of the activity patterns with the measured values.
3. Apply the indices in a case study and test the effectiveness of using these indices to discern meaningful activity patterns in an individual-based activity database.

## 1.4 Research Questions

To achieve the objectives of this study, certain questions are posed for resolution:

1. What are the types of activities, levels of activity participation, and activity linkages involved in the daily activities of the population in OSU, Stillwater Campus?

To understand the activity patterns of individuals on campus, it is important to examine the types of activities being undertaken by the individuals, their levels or intensity of activity participation and the constraints of space and time that are imposed on their activities by their schedules. These provide the context within which the meaning and significance of an individual's real life schedules are determined, and which lays the foundation for a time-geography approach to activity analysis (Ellegard, 1999).

2. Taking the spatial and temporal contexts of activities into consideration, what is an effective method to examine, understand, and analyze activity patterns at the OSU, Stillwater campus?

The time-geography framework possesses a lot of potential for generating and developing human activity analysis indices. Traditionally, the propensity of people to chain their trips (trip-chaining) and the activity profiles of groups are important indices of activity pattern analysis. Unfortunately, no methods have been developed to measure these important indices of activity analysis in line with current practices in the field such as the time-geography framework. The activity-based approach requires new methods to incorporate these indices within the activity analysis methodology. This study proposes two indices, a fragmentation index and a similarity index, based on time-geography principles and concepts to measure activity-chaining and activity profiles, respectively.

(i) The fragmentation index utilizes the concepts of station (as anchor location) and space-time path. A station represents an activity location, which serves as the origin or destination of a trip. A space-time path, connecting the stations with their corresponding trips in a trajectory, chronicles the activity sequence of an individual. The fragmentation index focuses on examining how an individual organizes his/her activities in regards to their stations and analyzes the daily activity scheduling patterns of individuals.

(ii) The similarity index identifies groups of activity participants with similar activity intensity. It adopts a sequential approach that utilizes one of the most distinguishing features of the time-geography framework, which is the sequence of activities. These components are further developed in the operational framework.

Different people respond to or are influenced by constraints or stimuli differently based on several factors. For example, an undergraduate student with a large number of class credit units and attending classes at places distributed all over the campus may generate a different activity pattern from a graduate student who has fewer class credit units and is more laboratory-based. Generally, it may be expected that individuals with similar traits or characteristics may display similar activity patterns. This may allow for reclassifying people into more meaningful groups based on similar activity patterns.

## 1.5 Contributions of the Research

### 1.5.1 Theoretical Contributions

Thanks to recent advancements in geographic computational technology and availability of individual-based spatio-temporal activity data, there has been an increasing interest in employing the time-geography framework to explicitly model human activity patterns. This approach offers many advantages because it accounts for all aspects of human activity patterns. Several principles and concepts of time-geography have been applied in some studies to capture contexts (everyday, project, social, geographical) of human activities (Ellegard, 1999); explore space-time density of trajectories in 3D in space-time cubes (Demsar and Virrantaus, 2010); measure similarity between representative space-time trajectories (Wilson, 2008); measure dissimilarity between geographically dispersed space-time paths (Vanhulsel et al., 2011), etc. This study adds to the growing literature on the adoption and application of time geography to analyze human activity patterns through the employment of the concepts of space-time path and activity stations; and the use of the principle of activity sequencing that is embedded in the time-geography framework to develop new methods of analyzing human activity patterns.

The new methods developed in the study are expected to open up new frontiers of discussion and application in activity analysis and time geography study. For example, the fragmentation index adopts a graph theoretic approach that suits the elegant framework of time geography that is encapsulated in the space-time path but which is not commonly applied to activity analysis within the framework. This approach may serve as a bridge between activity analysis, time-geography and graph theory.

This study proposes a quantitative approach to the selection of places as anchor locations in the analysis of human activity patterns. Traditionally, an anchor location is determined by the functionality of a place and usually confined to a home or workplace location. Using the principles of frequency-of-visit and duration of time at locations, this study provides a data-driven method to determine the anchor locations involved in an individual’s daily activities. This method broadens the scope for the definition of an anchor location and can better accommodate untypical activity patterns. With a more flexible definition of an anchor location, this approach opens up new perspectives and offers new avenues to perceive, conceive and analyze human activity patterns.

### 1.5.2 Practical Contributions

The study can contribute to the development of suitable activity spaces to enhance the human activity participation experience. The fragmentation index employs the concept of anchor location (based on stations), which represents the organizing location of activities for individuals. Identifying groups of individuals with similar activity characteristics as well as identifying their important anchor locations is important to appreciating the activity patterns of both individuals and groups of persons; and, how individuals and groups complement one another and create effective activity spaces. This is important to the provision of more suitable and livable activity spaces based on characteristics of compositing groups, characteristics of activities and the level of significance of anchor locations to activity participation.

The two activity analysis indices developed in the study can not only help us increase our understanding of human activity patterns, but also add to the variety of human activity methods that may be employed to analyze and unravel human activity patterns. Activity chaining and activity profiles are two important components in understanding the organization of human activity patterns. However, few methods have been developed to effectively measure them. This study takes up the challenge and develops two indices to address the measurement of these two activity components respectively. These indices provide different perspectives to investigate human activity patterns and can promote further understanding of the mechanism of how individuals organize their daily activities in space and time.

# CHAPTER II

## REVIEW OF LITERATURE

### 2.1 Introduction

The study of human (activity) behavior has been in the forefront of academic research for a long time. Over the course of time, several theories, concepts and frameworks have been developed to enhance understanding and explanation of human behavior. Unwin (1992) identify two broad approaches for the study of human behavior. These are behavioral and humanistic approaches.

According to Unwin (1992), behavioral approaches seek to circumvent the shortcomings of the preceding rational and spatial models that assumed a perfect information environment and rational decision making behavior by people. It reaffirms the position that human beings operate in an environment of imperfect information on all opportunities and constraints, and do not necessarily have to process all available alternatives to arrive at decisions. In addition, it reinforces the fact that human beings do not always make decisions based on utility maximization (Axhausen and Garling, 1991). Humanistic approaches on the other hand are forged on a foundation based on human experience and knowledge of their environment.

Importantly, both approaches accept that human behavior consists of both spatial and temporal contexts. Generally, these approaches seek to identify, understand and explain behavior of large groups of people. As the attraction and desirability for disaggregated analysis of human behavior patterns gained currency, alternative frameworks and models were sought and developed. This resulted in a “Kuhnian” paradigm shift with emphasis placed on the individual rather than the aggregate or group level of human behavior (Pas, 1990). Consequently, aggregated patterns became the aggregation of individual (disaggregated) behavior whose decisions are influenced by different motivations. An explanation of aggregate patterns therefore became reliant on an understanding of the nature of individual behavior. In line with this, most studies of human activity-travel patterns have adopted two basic frameworks: trip-based and activity-based approaches. These frameworks provide approaches based on different perspectives on the nature and motivation of human activity patterns.

### 2.2 Activity-Pattern Analysis: Trip-Based Approach

Traditionally, activity analysis has focused on closely examining the patterns generated by movements between activity locations. Emphasis is placed on variables such as the number, frequency, rates, and lengths or distances of trips generated and distributed between origin and destination areas. The “trip” therefore became the central unit of activity analysis (Bhat and Lawton, 2000; Kulkarni and McNally, 2000; Miller and Shaw, 2004). This approach, referred to as the trip-based model, is typically represented by the Four-Step Transportation Planning Model (TPM), which consists of four separate but inter-related stages: trip generation, distribution, modal split and network assignment. To implement the model, traffic analysis zones (TAZ) are first demarcated. These are areas considered and demarcated to possess relatively homogeneous characteristics including socioeconomic and demographic as well as transportation attributes. The TAZs serve as the spatial units for the generation and distribution of trips (Taaffe et al., 1996). Several techniques and methods of analyzing each step of the process have been developed to include cross-classification and regression methods for trip generation; gravity-based models for trip distribution; and numerous algorithms to determine modal split and network assignments (Taaffe et al., 1996)

Classic examples of the application of the four-step transportation planning model include the Chicago Area Transportation Study (CATS) and the Detroit Metropolitan Area Traffic Study commissioned in the 1950s and 1960s, which were ambitious but were constrained by limited computing capabilities of the time (Hall, 1988; Weiner, 1992; Kitamura, 1996; Bhat and Lawton, 2000). Since then there have been phenomenal improvements in computation and software developments that have opened up the frontier of travel demand modeling. Despite these improvements, the four-step transportation model has been criticized for its inherent assumptions and limitations. Miller and Shaw (2004) criticize the model for its unrealistic assumptions on the sequence in which human activities are undertaken. They argue further that even when the model is solved in the sequence it recommends, there is no convergence to any particular solution because the expected feedbacks into the solution process are very subjective and therefore are liable to wide range of differences among modelers (e.g., Florian et al., 1975; Boyce et al., 1994; Hassan and Al-Gadhi, 1998).

Kitamura (1996) claims the model lacks a sound basis in actual behavior of trip makers. For example, people do not base traveling decisions on how many trips to make; instead they decide on where to go, how to get there and when. This last component, time, is another sour point for the model. The time of day, a critical component of trip decision making is not explicitly incorporated in the model. This is important because traffic congestion, which is one of the most important concerns of transportation planning and traffic forecasting, is heavily dependent on time-of-day. This deficiency makes it very difficult for the model to effectively analyze traffic peak spreading (spread of traffic congestion), assess impacts of congestion pricing (charge users at traffic peak periods higher prices for use of transport facilities) or predict destination of cold or hot starts (Environmental Protection Agency categories of transient modes based on their engine soak periods at a particular time. These criteria are used by researchers in travel surveys and traffic assignment procedures in traffic emission analysis).

Further criticism is aimed at the use of the trip as the fundamental unit of analysis. The model assumes trips are independent entities whereas in reality trips are linked together and the underlying decisions to make trips are also inter-related. Disjointing trips into individual units misses the whole essence of why the trip was made in the first place consequently the decisions underlying the revealed characteristics of trips are not taken into account (Bhat and Lawton, 2000; Krizek, 2003). For example, a person may decide to drive to work simply because he/she would like to go shopping after work.

Kitamura (1996) agrees that introducing new elements or modifying old elements may help overcome some of the limitations of the conventional transportation planning and travel demand models but submits that the atemporal structure of the trip-based models is difficult to eliminate. To include the important dimension of time, therefore, there was need for a framework that explicitly incorporated it. These and other related concerns led to the emergence of the activity-based models.

### 2.3 Activity-Pattern Analysis: Activity-Based Approach

Unlike the trip-based approaches that treat transportation as if it were desired for its own sake, activity-based approaches model transportation as a derived demand, i.e., people move because they have the need to undertake in activities such as work, shopping, recreation, etc. To achieve these important activities, transportation is necessary to overcome distance. Activities, not trips per se, become the organizing principle of revealed activity patterns (Pas, 1990; Misra and Bhat, 2000).

Human societies consist of many activities distributed in space over time. Movements are organized around these activity locations as the need to engage them arises. This results in spatial interaction among people, which produces discernible patterns that reveal human behavior. Misra and Bhat (2000) contend that the conceptual appeal of the activity-based approach is derived from the fact that it acknowledges that the desire and need to participate in activities is more fundamental than the travel that is generated by these activities. Travel is consequently a product of differences in lifestyles and participation in activities among the population and this is set within a richer and more holistic framework than was provided for by the earlier conventional models (Jones et al., 1990). McNally and Rindt (2008:5) define activity-travel patterns as “the revealed pattern of behavior represented by travel and activities (both home and non-home) over a specified time period (of a single day).” These activities, which serve as the organizing principle of revealed patterns, become the basic unit of analysis thus replacing trips. Participating in activity programs involve making decisions on activity schedules and travel plans, both of which produce the day's activity patterns. The emphasis is now on individuals and their motivations to make trips for the purposes of engaging in activities. This is an important shift in activity analysis framework (Pas, 1990).

According to Behrens (2000), early activity-based studies concentrated on understanding travel behavior, tackling new approaches to data collection and data analysis, while later studies were centered on development of activity-based travel forecasting models. McNally and Rindt (2008) concur but indicate that the works in both periods are still ongoing as activity-based models grapple with the availability of suitable data and methods to collect them. It was recognized early on that activity-based approaches will require more qualitative data than the conventional trip-based approaches and the traditional methods of survey, travel diaries, origin-destination diaries and trip generation surveys, though still important and useful may need to be replaced by more appropriate tools of data collection (Kenyon, 2006). Some of these methods of data collection are reviewed in the following sub-sections.

### 2.3.1 Activity-Based Diaries

Clarke et al. (1981) identified the pioneering work of researchers at the Transport Studies Unit (TSU) of Oxford University in the late 1970s and early 1980s in the application of activity-travel diaries as particularly useful in establishing activity-based methods. In a two-day activity-travel diary study of 196 households in Banbury, Oxfordshire, data was collected on how household members spent their time, the spatial and temporal environment within which activity schedules were made (e.g., activity location, opening and closing hours of potential activity locations), transportation supply and the roles that children play in determining household activity-travel behavior. The study found that children in households acted as agents that significantly transformed the activity paths and travel constraints of whole households. This is similar to results found by Goodwin (1983).

Stopher (1992) demonstrated using a one-day survey, that trips are more accurately measured by data collected from an activity-diary than the traditional travel diary and survey methods. Reported trip rates were higher basically because the activity-based diaries captured non-home trips better. The diary had pre-coded activity groups to guide respondents and ease data analysis.

Misra and Bhat (2000) analyzed the activity patterns of non-workers in the San Francisco Bay Area based on activity-travel diary data collected from the 1990 San Francisco Bay Area Study. The study found that households and individual characteristics have significant influence on out-of-home activity types and the propensity for activity chaining but that activity sequencing behavior depends on the types of activities to be pursued.

Doherty (2004) examined household activity scheduling in Toronto, Canada, using a computer-designed activity-based survey. Sixty different activity types were simplified into nine categories. Respondents were asked to record only activities that lasted 10 minutes or longer. To induce a favorable response, a $20 incentive was provided for respondents. Unfortunately, the response level was still very low (17%).

In summary, activity-based diaries have three basic advantages over traditional trip diaries and surveys. First, they account for both trips and the activities that generated them. Second, they incorporate multitasking, which are generally more common than are reported. Third, they allow for pre-coding, which simplifies data collection and analysis. Unfortunately, the intensive nature of the requirements of activity diaries renders it liable to a low response rate.

### 2.3.2 In-Depth Interviews

Behrens (2000) identified in-depth interview techniques as important data collection tools for analyzing activity-travel patterns. The Household Activity-Travel Simulator (HATS) developed by the TSU is an example of this method. The respondent plots their activities and inter-link travel activities for the previous day on a games board using different colored blocks. A range of hypothetical spatio-temporal constraints are then introduced and the respondent re-arranges activities to reflect how he or she would respond to the new changes. The revised individual activity-travel pattern is then quantified and analyzed.

Kenyon (2006) pointed out that it is not possible to collect all relevant data in a single survey, or by a single methodology. There is need for an integrated methodology that is both effective in collecting as much relevant information as possible and accurately, too. For example, location-aware technologies, such as the global positioning system (GPS), are very reliable tools of collecting spatial information (Zhou and Golledge, 1999). This technology has the added advantage of an option to incorporate some interview method.

2.3.3. Global Positioning System (GPS)

The use of GPS technology as a data collection tool is gaining prominence in activity research. Zhou and Golledge (1999) equipped the cars of 100 households with GPS tracking device to collect data for the study of travel behavior in Lexington, Kentucky. The study used statistical methods such as factor analysis, discriminant analysis and multivariate analysis of variance (MANOVA) techniques to analyze repetitive and cyclical patterns that may be hidden in the activity-travel patterns of households.

Rinner (2004) adopted the time-geography concepts of station, space-time path and space-time prism to represent activity data collected by GPS technology. The study simulated the travel of several individuals through the western downtown to the university building in Toronto, Canada and developed an interactive positioning of space-time stations in 3D scenes. The study illustrated the application of time-geography concepts to analyze activity-travel patterns.

Andrieko et al. (2007) employed GPS technology to track movement of cars and trucks including the duration of vehicles at locations. Stations with high temporal durations were designated as “significant places” (stop points or Points of Interests [POI]), which include homes, work places and places regularly visited such as shops. Similarly, trips were extracted and clustered based on time (duration) thresholds.

### 2.3.4 Classifying Activity Types

Typically, activity-based surveys involve the pre-coding of activities into a few categories. The main advantage of this categorization is to streamline potential activities and simplify data entry for respondents and researchers alike, and also to reduce the burden of post-data collection work and analysis. First, the range of activities at the disposal of individuals in a given time period may be astronomical. An open-ended activity type survey will result into an enormous amount of information that will consume a lot of time and resources to classify, and analyze. Second, similar activity types may be recorded differently by different people. It becomes difficult therefore to identify the different aliases or codes by which a given activity type may be labeled by different people and to standardize them. This may introduce inaccuracies in data analysis, misrepresentations of patterns and misspecifications in estimates. To guard against these, activity-travel diaries are pre-coded into a handful of recognizable activity types or groups based on criteria that are meaningful, understandable and acceptable within the context of the goals of the study.

Ellegard (1999) devised a hierarchical categorization method that reduced about 600 different activity types into five categories based on declining levels of details in activity diaries: sphere, category, class, sort and specification, which allows for joint household activities to be identified and easily compared. Kenyon and Lyons (2007) classified over 60 different activities into eight groups of activity types including internet use, telecommunication and activity-travel patterns of individuals.

These general coding of activity types have been employed by several other researchers, each with a slightly different group of classes to reflect its aims and objectives and based on the types of individual activities that are of interest in accomplishing stated goals (Golob, 1986; Zhou and Golledge, 1999; Misra and Bhat, 2000). Other studies adopt a more restricted number of groups of activity types. Reichman's (1976) classic classification of individual activity groups has become a standard categorization for many researchers: Subsistence activities (primary activities, e.g., work or business), Maintenance activities (activities necessary for sustenance of individual livelihoods, e.g., shopping), and Discretionary or Leisure activities (e.g., going to the movies) (Pas, 1982, 1984; Gould and Golob, 1997)

Garling et al. (1998) produced a trichotomy of activity types based on participation, and preparation: *Habitual activities* (routine activities that are repetitive and easy to discern, e.g., work); *Planned or Pre-arranged activities* (e.g., arrange to go shopping after work); and, *Impulsive or Spontaneous activities* (activities undertaken on the spur of the moment, e.g., an impromptu lunch break). Other classification typologies include *work* and *non-work* or *home* and *non-home* activity classes (e.g., Hanson, 1980; Misra and Bhat, 2000; Srinivasan and Rogers, 2005); and *flexible* and *fixed*/*inflexible* activities (Pipkin, 1995; Vilhelmson, 1999; Miller, 2004).

Krizek (2003:396) submits that aggregating activity types into more restricted number of groups allows for easier analysis in ways that are "more parsimonious than using eight activity types but more detailed than the simple work/non-work dichotomy."

## 2.4 Models of Activity-Pattern Analysis

This section reviews three models of activity patterns: the stochastic models, simulation/computational models and the time-geography framework. Most of the methods of activity analysis fall within one or more of these frameworks.

### 2.4.1 Stochastic Models

Many techniques based on mathematical probability theory have been applied to analyze activity-travel patterns. These include statistical methods such as regression and discrete choice models (e.g., least squares, tobit, logit, etc.), econometric, and structural equation models.

Misra and Bhat (2000) apply several varieties of the regression model, including binary and response ordered logit structures to model the influence of socioeconomic and demographic characteristics of respondents in the San Francisco Bay area on their number of stops for activity types, activity chaining and activity sequencing behavior. Lee-Gosselin and Miranda-Moreno (2009) and Krizek (2003) also employ regression methods to analyze socio-economic and demographic influences on the activity patterns of individuals. Zhou and Golledge (1999) use several statistical methods including factor analysis, discriminant analysis and K-group multivariate analysis of variance (MANOVA) to analyze household travel behavior in Lexington, Kentucky.

Golob and McNally (1997), Fuji and Kitamura (2000), Golob (2001), and Kuppam and Pendyala (2001) employ structural equation models to represent human interactions and activity allocations, and to model the relationship between activity duration and trip generation, activity frequency, and trip chaining. The models represented subsistence activities and duration at work, household maintenance, and discretional activities of male and female couples, which reflect a range of independent variables.

These types of models have become very common techniques in activity analysis (Kitamura et al., 1992; Golob et al., 1994). The main attraction of stochastic models lies in the fact that they provide easily quantifiable means of analyzing activity patterns.

### 2.4.2 Simulation/Computational Models

Generally, activity pattern forecasting is based on analyzing representative sample of a given population. However the availability of a representative sample is not always guaranteed. Simulation techniques offer an approach to update available sample or synthesize a representative sample based on aggregate data (Bhat and Lawton, 2000).

Several simulation models have been developed to analyze activity-travel patterns. Axhausen and Garling (1991) and McNally and Rindt (2008) claim that the precursor of activity-based simulation models may well be Lenntorp's (1976) PESASP (Program for Evaluating Sets of Alternative Sample Paths) model, which was developed on the basis of Hagerstrand's (1970) time geography. Nevertheless, one of the earliest of these simulation models, CARLA, developed by Jones et al. (1983) is a combinatorial algorithm that identifies and simulates feasible alternative activity schedules or paths (Kitamura, 1988; McNally and Rindt, 2008). Axhausen and Garling (1991), however, note that the constraints to activity scheduling identified by CARLA may not necessarily be the same as those that constrain activity scheduling regimes of people in real life.

Both PESASP and CARLA became the prototype for STARCHILD (McNally and Recker, 1986), a predictive model of complex travel patterns (Kitamura, 1988) developed to identify and choose between schedules of representative activities (Axhausen and Garling, 1991). Recker et al. (1986a, b) note that STARCHILD's basic assumption of utility-maximization as the driving force of individual's decision making is highly implausible. In real situations, decisions on activity schedules are limited and considered to be non-compensatory (Montgomery, 1990; Axhausen and Garling, 1991). McNally and Rindt (2008) opined that STARCHILD is more suited to theory and research work than general practical application.

Other simulation/computational process models developed for activity pattern analysis include SCHEDULER (Garling et al., 1994), SMASH (Ettema and Timmermans, 1995), AMOS (Kitamura, 1996) and ALBATROSS (Arentze and Timmermans, 2000). Most simulation and computational models avoid the reductionist approaches adopted by trip-based model systems for the holistic framework that are more suited for activity-based models. Despite their limitations, simulation and computational models provide promising directions for operational models and can be used to test alternative conceptual frameworks for activity behavior (McNally and Rindt, 2008).
It is instructive that the underlying trigger for the development of activity-based simulation/computational models is Hagerstrand's time-geography approach. Time geography has many qualities that make it an attractive framework for the study and analysis of activity patterns.

2.4.3 Time-Geography Approach

Time geography is a simple but profound approach to the study of human activities. Human activities are distributed across space. To engage in activities, movement is required between activity locations. This takes time. Additional time is also required to participate in the activities. Time geography takes the view that these two components, space and time, are resources that are intricately and inextricably woven together and should therefore be considered as equally important in understanding activity patterns. First proposed by Hagerstrand (1970), time geography presents an elegant framework to study human activities within a space-time context (Golledge and Stimson, 1997).

Additionally, time geography acknowledges the significant roles of various social, economic, and environmental limitations on human activity participation. These are referred to as constraints. No human activities can be completed without facing some constraints (Hagerstrand, 1970; Thrift, 1977). For example, to engage in an activity in another location, a person would need to overcome the friction of distance. Time geography provides basic concepts that capture both spatial and temporal components of human activities and their relationships to the constraints. The space-time path captures the location of activities, the time, and duration of activity participation in 3D with a 2D spatial plane and time as the third dimension. The space-time prism demarcates the range within which an individual can engage in activities given certain constraints. These types of information are important to understanding human activities.

Yu (2006) itemized three related appeals of the activity-based approach of time geography. First, the integrated space-time system of time geography explicitly represents time, which is important in activity participation. Second, the framework is well-suited to the derived nature of transportation in human activity engagement. Movement between location points are linked by tilted lines that indicate the relatedness of human activities, and the duration of time spent in an activity location is represented by a straight vertical line. Third, the issue of trip chaining with multiple stops/purposes is addressed by the sequential nature of the connecting lines between activity locations and movements between them. The connecting lines are not random rather they follow the sequence by which the activities are undertaken.  A fourth aspect may be added. The framework allows for analyzing the activity patterns of individuals as the space-time path deals with activity characteristics of individuals rather than whole groups (Peuquet, 1999; Shaw and Yu, 2009a). This is important in activity analysis because it allows for a more focused and deeper examination of the individual's purposes of trip-making activity, which are usually lost in aggregated data.

Several studies have adopted time geography as the framework for activity pattern analysis. Early studies by Lenntorp (1976, 1978), Carlstein et al. (1978), and Ellegard et al. (1977) set the tone for many future studies. Lenntorp's (1976, 1978) PESASP model of accessibility was developed to simulate and evaluate alternative sample paths of activities. The model seeks to identify all the possible ways, given activity locations and spatio-temporal constraints, which an activity program can be carried out. It plots out the number of possible alternatives for each travel mode against the number of origin points and examines the efficiency of movement between activity locations. The model has been criticized as theoretically appealing but practically unrealistic (Thrift, 1977a). Several recent developments in research have, however, attempted to address much of these criticisms by developing suitable methods to analyze activities.

Miller (1991) brought time geography to the forefront of academic research again after a hiatus in the late 1970s and the 1980s. His operationalization of the concept of space-time prism reignited interests in the framework and opened new frontiers for the application of time-geography concepts in several research fronts (Kwan, 1998; Miller, 1999; Lee and McNally, 2002; Yu and Shaw, 2004).

### 2.5 Time Geography and Activity Analysis

Hagerstrand (1970) first proposed time geography as an approach to the study of time, space and human activities (Golledge and Stimson, 1997). Time geography seeks to examine the relationships between human activities and their constraints within a space-time context (Pipkin, 1995; Golledge and Stimson, 1997).

Time and space are important resources, which should be treated as co-equals. They are also limited resources in the sense with which human beings use them. Hagerstrand (1975) identifies eight fundamental conditions of human reality: the indivisibility of a human being; a limited human lifespan; limited ability of human beings to participate in many activities at a time; the fact that human tasks have durations; movement between any two points takes time; limited packing capacity of space; limited outer size of terrestrial space; and, the fact that every situation is linked to past situations. These conditions, biophysical, ecological and locational, impose constraints on human use of space and time (Thrift, 1977a).

### 2.5.1 Activity Constraints

The basic underlying argument is that human activities are undertaken at some locations in space at particular time periods and within certain limiting factors (constraints). These factors impinge on individual freedom and constrain them to occupy certain space locations and time allocations, which ultimately shape individual's patterns of activities. When these constraints are effectively identified, it is possible to understand why an individual's participation in activities is patterned in certain ways rather than differently (Golledge and Stimson, 1997). Consequently, time geography explores human behavior by identifying the constraints imposed on individual action that limit possible behavior alternatives (McNally and Rindt, 2008).

Time geography identifies three broad types of constraints to human activity participation: capability, coupling and authority constraints (Hagerstrand, 1970). Capability constraints are barriers that limit participation in activities by demanding that large portions of time be allocated for physiological necessities such as sleeping, eating, etc (Golledge and Stimson, 1997). These constraints extend beyond biology to social and technological capabilities; for example, the capability to effectively operate a technological device to carry out a desired task.

Coupling constraints require a person to occupy a certain location for a fixed duration of time in order to conduct some activity and may involve joint activities with other individuals. Synchronization of time and/or space by all participating individuals or objects is therefore required for coupling activities to be undertaken. An example is engaging in a class group work. All participants will be required to congregate at the same time and in the same location, physically or by some proxy (e.g., teleconference).

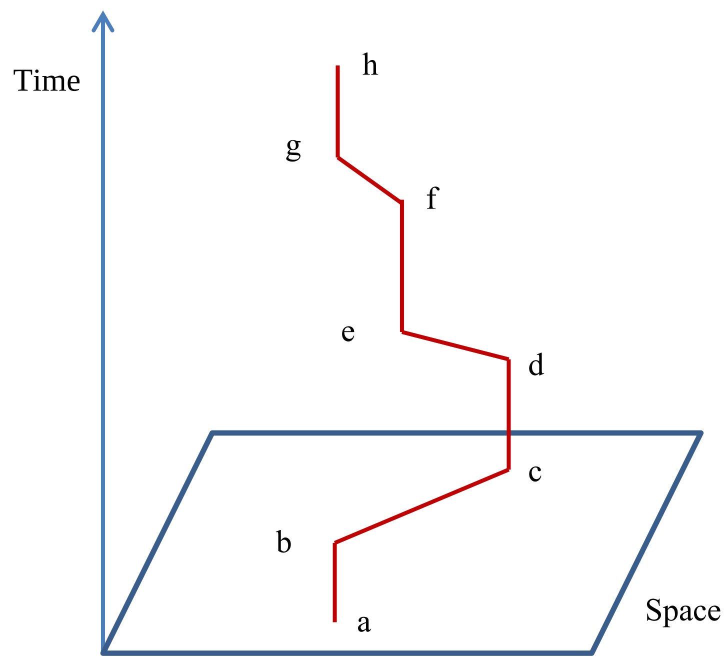
Authority constraints, on the other hand, are restrictions in space and time that can affect location and/or time activity. This involves laws, regulations, and institutional contexts within which activities are undertaken. For example, driving activities are generally constrained, by law, to roadways provided for the purpose.

These constraints are synonymous with the necessary, but not sufficient, conditions for the occurrence of an activity or interaction (Pred, 1977). Individuals must overcome constraints in order to participate in a single activity or to interact with other people. Time geography therefore concerns itself with what activities people are free to do, rather than what they actually do. In order words, time geography seeks to trace barriers that prevent certain activities from occurring and does not attempt to directly predict individual behavior. Instead, it examines the possibilities available for activity participation, also known as action space. Miller (2007) refers to this perspective of activity analysis as “constrained-based,” a view that is radically different from the traditional approaches that extract human behavior from observed human action (McNally and Rindt, 2008). Thrift (1977b) had argued that observed behavior does not tell us much about the constant change in the potential of each pocket of space-time, or the inputs and outputs which sustain it. By implication, analyzing human behavior through observation of human actions does not reflect the process by which the observed action has been produced. The representation of activity behavior in such manner will therefore be incomplete.

Time geography represents an individual’s activity pattern and activity space through the concepts of space-time path and space-time prism. These are important visualization tools that underscore the elegance of the time-geography framework.

## 2.5.2 Space-Time Path and Space-Time Prism

A space-time path is a trajectory of an individual's movement in physical space over time. It provides information on the spatial and temporal characteristics of an individual's activities, including time, location, temporal extents and the sequences of activities (Yu, 2006). The space-time path is a three-dimensional orthogonal coordinate system that represents time as the third dimension while the other two dimensions represent changes in activity locations in space (see Figure 2.1)




**Activities**:

**a-b**: 6:00-7:00am (Home)

**b-c**: 7:00-7:30 am (Travel: Bus)

**c-d**: 7:30-10:00am (Work)

**d-e**: 10:00-10:10am (Travel: Walk)

**e-f**: 10:10 – 1:00 pm (Shopping)

**f-g**: 1:00-1:10 pm (Travel: Bus)

**g-h**: 1:10 – 4:00 pm (Home)

Figure 2.1: A Space-time path and stations

Wang and Cheng (2001) define activity pattern simply as a sequence of staying at or moving between locations. Effectively, the space-time path represents human activity pattern by a set of continuous segments of vertical and non-vertical (tilted) line segments. Vertical line segments, as the third dimension, represent the duration of activity undertaken at a particular location. The non-vertical segment represents movement between activity locations. These line segments are linked together in the sequence in which the activities occurred. The velocity of movement between activity locations can be extracted from the space-time path by examining the distance between the two points and the time it took to arrive at the destination point from an origin point. A steeper slope for the non-vertical segment means that the movement between the activity locations requires more time to overcome (Miller, 2004).

A space-time prism (Figure 2.2) simply delimits the possible locations for the space-time path over a particular time span (Miller, 2004). It maps the limits of feasible human action space within a given time duration (Lenntorp, 1976). Since an individual can only be at one location at a time due to the indivisibility of human nature as a capability constraint, the time to act or be in the opposite direction is reduced.

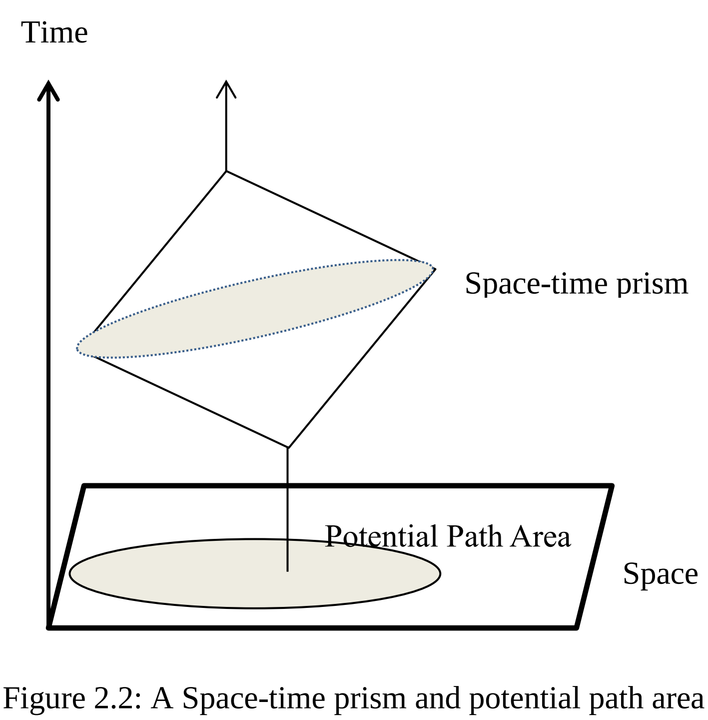



Figure 2.2: A Space-time prism and potential path area

This means that less time is available for the person to participate in activities that involved moving in the other directions or locations. McNally and Rindt (2008) explained that once an individual travels to a specific location, the potential action space for any subsequent activities will be reduced depending on the activity duration.

Consequently, the more time is taken to undertake a current activity, the less time there is to allocate to subsequent activities. The space-time prism therefore demarcates the sphere within which activities can occur for a particular time period and within certain constraints.

When the prism is projected onto a two-dimensional plane, it is referred to as potential path area (Figure 2.2). This establishes the physical area boundary that can be accessed by an individual within available time budget (Miller, 2004).

The seminal work of Hagerstrand and the accompanying work of his colleagues at Lund on time geography set the stage and tone for subsequent researchers. For example, several early studies of time geography concentrated on migration and population systems, planning and policy in Sweden (Hagerstrand, 1975). Nevertheless, for a period of time, time-geography was adopted mainly as an elegant conceptual, rather than an analytical, framework (Miller, 2004; Yu, 2007). Its simplicity, in theory and concept, has made it very attractive as a research tool. Recent advances in GIS technology, however, have improved the fortunes of time geography as an analytical tool. For example, Lin et al. (2009) demonstrate that the space-time path and space-time prisms can be used to geovisualize activity-travel patterns. The study is set in Temsui, Taiwan where map questionnaires were used to obtain trip activity data of 40 tourists. Using detailed Google maps for the area, the space-time paths of tourists are constructed in GIS as a means to study their spatial and temporal behavior, with the aim of identifying the most popular tourists' paths. Space-time prisms for tourists were also constructed from which potential path areas were projected to identify the areas and potential opportunities that were accessible to tourists during their travel time.

### 2.5.3 Extensions to Time-Geography Framework

Miller's (1991) classic work on operationalizing time-geography concepts reignited interest in time geography. He demonstrated that the space-time prism and potential path areas can actually be modeled in GIS and can be constructed on a network constrained with variable speeds that approximate the reality of human activity spaces. This work exhumed time geography from near irrelevance to a prime position in academic discourse. Following this lead, several extensions of time-geography framework have been implemented to realize the rich potentials of its concepts and to model developments in human activities.

Kwan (2000c) and Yu and Shaw (2004) have developed geovisualization tools for space-time paths using GIS. These studies demonstrate the construction of space-time paths of several individual human activity sequences in both space and time. Yu and Shaw's (2004) design allows for a temporal dynamic segmentation of space-time paths that makes it possible to identify locations of individuals at any particular time along the space-time path. Adams (2000), Kwan (2000a), Yu (2006, 2007), and Shaw and Yu (2009a) have explored several ways of extending time-geography concepts to represent both physical and virtual activities, and activity spaces in a hybrid or integrated space-time environment. The recent advances in Information and Communication Technologies

(ICT) have necessitated such an extension because the nature of human activity engagement has changed drastically. More time is now being spent by individuals on activities, e.g., shopping, working, etc. in virtual space, which have significantly influenced human activity patterns through substitution, generation and/or modification of physical activities (Mokhtarian, 1990, 1997). These works have enhanced the analytical strength of the time-geography framework and increased its appeal in human activity analysis.

Other studies have adopted the time-geography framework and have employed other approaches to investigate activity patterns. One of these approaches is similarity measures, which are methods that compare the activities of individuals to identify those that are identical and group them into clusters of similar activity patterns.

### 2.5.4 Similarity Measures

Shoval and Isaacson (2007a, b) employed a space-time similarity measure based on sequential alignment method to examine the activity patterns of tourists in Northern Israel. Tourists' activities and movements in Akko town were captured using GPS devices. A total of 246 visitors' tracks were obtained but only 139 tracks were complete and used in plotting space-time paths, 44 of which were ultimately analyzed. The space-time path aptly captures the essence of the tourists' spatial behavior summarized as "what" (the activity), "where" (activity location) and "when" (time of activity).

To simplify analysis, the study area was demarcated into 26 zones and each zone is represented by a single character for identification purposes (unique location ID). Because Akko is a very small town, the temporal resolution was high at one minute. The analysis was restricted to activity locations rather than actual activities in which the tourists participated. Consequently, the sequential alignment analysis was based on the order in which the tourists visited the sites in the town and the routes they took to reach them. The software, ClustalG, was employed in analyzing the data.

Sequence alignment measures the degree of differences between two sequences in terms of their element composition and sequence. It adopts a “biological” distance concept to compute the differences between two sequences of activities. The algorithm consists of three operations: insertion, deletion and substitution. One of these operations could be applied to one of the sequences to make the string (of sequences) identical to the other string. The more operations needed to make two strings identical, the longer the distance between the two sequences. And the longer the distance between the two sequences, the less identical they are. The biological distance of two sequences therefore indicates how identical or dissimilar they are in their element composition and sequencing.

The result is a taxonomic guide tree that revealed three well-defined groups of visitors distinguished by (a) the number of locations they visited, and (b) the order in which they visited the locations. The first group had 16 sequences, the second had 18, and the third had six sequences. The amount of time each group spent in each polygon (location) is also calculated by averaging the amount of time for all individuals in the group that was spent in the polygon. Relationships were determined between the three spatio-temporal patterns and the personal and travel information of the tourists. A contingency table analysis was used and statistical significance of relationships was sought between various categorical variables. The only variable that appeared to have had any significant influence on the patterns was the answer to the question whether individuals availed themselves of the "Information" at the Information Kiosk at the Visitors' Center. More than 80% of individuals in group 2 reported having used the information obtained at the Visitors' Center while only 50% and 33% of people in groups 1 and 3 used it respectively. For context, group 2 had the more intensive tour of the three groups. This bears out the literature on the importance of information as it affects how the cultural tourist resource is ultimately used.

Though several limitations of the study were identified (e.g., inability to model the effects of personal characteristics of tourists on activity patterns, assess reliability of the method, etc), the authors contended that the method provides researchers with an avenue to aggregate and create generalized space-time paths and develop typologies of space-time activities that are necessary for furthering developments in time geography. Sinha and Mark (2005) develop a distance function to measure the similarity of geospatial lifelines as a means to examine the exposure of individuals to health risks in the course of their residential locations over time. They argued that geospatial lifelines allow for better and direct comparisons of the space-time activities of individuals and health risk cases and also reveal locations of common risk factors that could identify causal effects. A geospatial lifeline is defined by Mark and Egenhofer (1998) as:

> "the continuous set of positions occupied by an object in geographic space over some time period. Geospatial lifelines consist of discrete space-time observations of a geospatial lifeline, describing an individual's location in geographic space at regular or irregular temporal intervals." (Sinha and Mark, 2005:117)

The resultant trajectories can be analyzed for similarities. A distance function that identifies similarities in individual residential histories is used to infer the potential cluster of cases that can be subjected to more rigorous cluster analysis.

The lifeline data model is inspired by time geography and is a representation of an individual's movement pattern in geographic space. The lifeline distance function is based on Minskowski metric for measuring the distance between two lifelines. The two principal parameters are proximity in space, and temporal duration of the proximity. The function is a "weighted average of successive separation distances between two residences, where the weights are the durations a particular separation distance was maintained before either one or both residences are changed." (Sinha and Mark, 2005:121)

The study adopted two different methods to simulate lifelines: (i) based on successive residential locations being completely independent and random within the study area (random positions), and (ii) based on a modified random walk model of an exponential model of 'migration move distances' observed in the US for migratory populations (exponential model). One major assumption is a deterministic relation between exposure and ill-health. Simulated individuals who live in risk areas for more than an arbitrary threshold number of years are labeled as cases, while others are considered the control.

The lifeline distance functions were evaluated for cohort of simulated subjects who were all born in the same year and lived to be 70 years old. For both simulation populations, histograms of distance functions are produced and statistics calculated for the databases including mean distances, standard deviation, and skewness. Using one sample Kolmogorov-Smirnov (KS) composite goodness-of-fit test the study tested the hypothesis that each empirical distribution is similar to a normal distribution with mean and standard deviation as estimated from the samples for an exposure of 10km radius.

Results indicated that the lifeline distance measure can distinguish between the distribution of cases and controls but the efficiency varies with exposure, size and duration. Results are better for shorter distance migrations than for purely random migration processes. As movements get more randomized, the spatial-temporal extent of risk clusters of cases become less compared to controls. The performance of the lifeline distance function therefore deteriorates. Two basic scenarios present themselves at which the lifeline distance functions for complete residential histories performs better than simple distances based on times of diagnosis: (i) when cases are clustered around an exposure early in life and spread out, and (ii) highly mobile populations make cases-control distributions similar but also make the use of lifeline distance functions more effective.

In a similar type of study based on migratory histories of people from the Great Plains region of the US, Yu et al. (2008) developed a measure of dissimilarity between individual space-time paths (lifelines) of migrants, which is a quantitative measure of how two lifelines are unlike one another. In this case, the space-time path represents the migratory record of persons, where they have lived (location), the amount of time spent at the place and the sequence of places they migrated to. The dissimilarity measure is calculated based on measuring the entire shape of space-time paths.

Three basic components of the space-time path are measured: the total movement distance and the number of movements, time-weighted mean center of residential locations and time-weighted standard deviation to the mean center. The dissimilarity measure is thus defined as the cumulated area of the gaps between the two paths representing the two sequences of activity attributes. A matrix of dissimilarity values reveals the "distance" between pairs of space-time paths based on their activity sequences. The higher the value the more the two sequences are dissimilar, in which case the cumulated area of the gap between the two sequences is larger. Three methods of hierarchical clustering analyses were then performed (simple or minimum method, complete or maximum method and average linkages) to identify inherent groups in the data. An algorithm is developed to calculate the dissimilarity measures and produce a graph of clusters of different groups in the datasets.

These studies aptly identify and showcase the potentials of time-geography concepts in directly modeling activity patterns. They also expose the wide range of research thrusts that time-geography framework encompasses.

### 2.6 Activity-Pattern Analysis on University Campuses

Most studies on activity patterns in universities in the US have concentrated on developing campus-specific transportation models (e.g., Poinsatte and Toor, 1999; Isler and Hoel, 2004) or trace analysis of wireless user patterns (e.g., McNett and Voelker, 2003; Balazinski and Castro, 2003; Henderson et al., 2004; Meng et al., 2004; Hsu and Helmy, 2005). The latter set of studies involved examination of wireless LAN users' behavior, including mobility behavior and traffic generated by the use of mobile equipment.

Hsu and Helmy (2005) undertake a comparative analysis of wireless local area network (WLAN) mobile user patterns using WLAN trace analysis in four university campuses. The data comprised 12,000 distinct users and over 13,000 access points (APs). They propose several metrics to measure behavior of mobile users, including associations to APs, mobility, repetitive, encounter, and friendship patterns. Two broad groups of analysis were adopted: (i) individual user behavior including activeness of users, long, and short term mobility of users; and, (ii) repetitive association pattern of users; and relationship between mobile modes.

An interesting feature of the research involved the development of a quantitative metric called Network Similarity Index (NSI) to measure repetitive pattern in user behaviors. Network similarity index is defined as the time gap in the average Location Similarity Index (LSI) for all users at particular time gap. The LSI, which is the "fraction of all such pairs where the user is associated with the same AP in both snapshots," (Hsu and Helmy, 2005:6) is measured for individual users. Snapshots of associated APs of users are taken every one minute. To glean the tendency for repetitive behavior of users at a certain time gap (e.g., 24 hours), all snapshot pairs that are separated by this time gap (of 24 hours) are considered. The LSI value indicates how likely a user re-appears at the same location after a chosen time gap.

The NSI is represented as a line graph, in which the "average value" of the curves reflects the fraction of users that always stay at the same location. Very high NSI values indicate a higher degree of repetitive behavior and stationarity in location, e.g., high degree of home-based (dormitory) activities. The results show that across the four campuses, the strongest repetitive pattern (the highest value) occurred at the time gap of one day and the second strongest pattern at the time gap of a week (second peak period). Kay et al. (2006) develop a similarity path algorithm to identify significant sequences of student group activities. The study aims to build tools that can identify interaction sequences that indicate problems as well as those patterns that reflect success in group activities of students.

The study adopts a data mining technique that preprocessed student group activity (software development teams) into suitable alphabets. The frequent sequential pattern mining algorithm is developed to account for the temporal nature of the data and cluster of group interactions. TRAC (an enhanced wiki and issue tracking system for software development projects) system data that traced students' actions (events) were collected and used to find frequent patterns that characterize some aspects of their team work. The general idea is to connect events if they are related to a specific task (e.g., group work) and track the frequency of event patterns.

The results of the study were inconclusive. However, two pertinent points can be raised for the study. First, the study correctly recognizes the significance of coupling constraints in a university community where joint activities are very common features of activity participation. Lectures, group work and discussions generally involve some synchronicity in time and/or location for participants. Second, the study recommends the adoption of contrast set mining algorithm (Bay and Pazzani, 2001), which considers only discrete and numerical data, as a potential candidate for transformation to formats that will recognize sequences of events that may be captured in other data forms. Such transformations will allow for developments of similarity measures.

Eom et al. (2009) examine daily activity patterns of students at North Carolina State University (NCSU). The study models student travel and activity behavior as a step towards a more comprehensive model of activity-based travel behavior. Data from 843 students were collected on their travel characteristics (trip rates, travel modes and trip types) and activity characteristics (activity types and location of activities, activity sequencing). Statistical analyses were used to find similarities in activity profiles of students based on their residential status (on- or off-campus), educational status (undergraduate or graduate) and gender status (male or female). The data for the study was collected using a student travel diary for one week-day only.

Using regression analysis to model trip rates, the study found that residential, educational and driver license statuses influence trip making behavior of students of NCSU. Gender status, however, does not. Student trip rates were also found to be much higher than the general local population's as measured by North Carolina Department of Transportation in 1999. The most frequent mode for primary trips (initial trips from home/dormitories to school) for on-campus students is walking (79.9%) while off-campus students use the automobile more often (68.9%). Only 3.5% of both student groups use the shuttle as primary modes. Secondary trips are associated with transfer from primary mode to another mode or to the final destination. Walking and the school shuttle system account for 95% of secondary trips. Walking is the only mode that showed up for tertiary trips. These are trips from the bus stop to the final destination.

A total of 4,883 individual activities were undertaken by 698 students. Most of the activities (45.5%) are university-work related, including class/school, study/research and work/volunteer activities. Meals and social/recreation activities also had very high frequencies. Average trip rate for university work-related activities was 1.61 per person per day, indicating that students attend classes, at least more than once a day. Meals have the second highest trip rate at 1.10 per person per day, showing that most students eat at least more than once a day. School, study and work activities have the largest allocation in students' time budgets (duration). Shopping activity frequency among students is higher than doctor/other professional activities but the latter take longer (larger time duration). Overall average trip time is 12 minutes, which signifies that most activities are centered in or around/near the NCSU campus.

Activity sequencing is examined using the activity transition matrix. The matrix represents the likelihood that a subsequent activity of a certain type will occur given an activity of a current type (Allison et al., 2005). The rows of the matrix indicate a current activity while the columns represent subsequent activities. The cell values are percentages of occurrence of a subsequent activity given a current activity type. The activity transition matrix has also been used by Misra and Bhat (2000) to study the activity-travel patterns of non-workers in the San Francisco Bay Area.

The most frequent subsequent activity following a current activity type is meals at 50% of the time. Schools/class and research activities are the most frequent subsequent activity type following meals. This may indicate that students usually have meals between classes, research and/or social/recreational activities. Among all possible combinations, however, the most frequent activity in the transition matrix is school/class activity, which occurs immediately after sleep (52.6%). This means that students probably attend classes before eating breakfast.

Activity profiles of student groups based on educational status (undergraduate or graduate) were compared using the F-statistics and represented as graphs. Activity profiles show the number of people engaged in each activity in each hour period during the day. This facilitates understanding of daily activity sequences of each group and helps embed the types of activities, their timing, and intensities (Alam and Goulias, 1999) into an activity-based model. For every hour of the day, the major activity type is identified and recorded to create a daily activity profile. No distinction is made between activity days (i.e., activities conducted on a Monday are treated in same way as activities reported for Friday). If more than one activity occurs within an hour time frame, the activity with the longest duration is recorded as the major activity. In the end all activities are aggregated into only five major activity types: home (all in-home activities); work (class and research-related); shopping; recreation (dining and leisure activities); and others.

Temporally, there are morning (10.00 am) and afternoon (2.00 pm) daily work activity peaks, which contrast with the usual urban peak period of 8.00 am and 5.00 pm. This probably results from the nature of class schedules. Undergraduate students were found to be more engaged in school work in the mornings and afternoons than graduate students (more engaged in the afternoons than morning, and more at night than undergraduate students). On-campus students also show a distinctive activity profile from off-campus students. On-campus students participate in more school activities and recreation than off-campus students. This may be attributed to the fact that more undergraduates live and make more on-campus trips than graduate students.

Overall, the activity profile (hourly activity participation and activity sequencing) does not seem to be significantly different among student groups, even though the proportions and frequency of activity types are slightly different. However, hourly activity type is different for student groups, which is indicative of a high correlation with time of the day.

## 2.7 Summary and Conclusion

From the literature review, several gaps are identified that need to be addressed. First, there is an observable gap in the employment of time-geography concepts, to develop framework to examine and understand activity patterns. Most studies have continued to employ the traditional stochastic models for activity analysis. Time-geography framework promises an alternative approach to the traditional models. Whereas the traditional approaches have been largely aggregative, time-geography provides the framework to develop disaggregated methods of analysis. Consequently, the development of a fragmentation index and a similarity index based on individual activity patterns and using concepts rooted in time geography would expand the frontiers of activity pattern research and further time-geography research.

Secondly, though many studies have adopted the activity-based framework to activity analysis, there is still much to be done to develop methods of analyzing and recognizing activity patterns. There is still much emphasis on theoretical developments, which need to be augmented by methods of analysis that are capable of providing new insights into the activities of people. Many components of activity patterns such as activity chains, activity transition, activity profiles, etc. are still being analyzed in aggregates even though there has been a refocus towards disaggregated methods of analysis. The need for new indices to address some of these problems is therefore overdue.

This study takes on the challenge to develop new methods that would contribute to addressing some of these gaps and concerns. The methods adopt time-geography principles and concepts, and a disaggregated approach to analyze some of the important measures of activity patterns.

# CHAPTER III

## OPERATIONAL FRAMEWORK

### 3.1 Introduction

Generally, studies on activity analysis have often adopted traditional methods such as those based on the four-step transportation planning process, stochastic and simulation approaches, time series analysis, and structural equation modeling. Each of these methods has made significant contributions to the study and understanding of human activity participation. However, they also have their weaknesses. For example, the methods are generally reductionist in conceptualization because the activity participation process is usually reduced to a series of its component parts (e.g., trips), which are then analyzed and returned as representative of the human activity experience. Second, traditional methods are usually aggregative and the individual activity experience is lost in the mesh of many others. This study adopts the more robust activity-based and individual-focused conceptual framework of time geography to address some of these inherent problems that afflict traditional methods of activity analysis.

Hagerstrand's time-geography framework possesses many capable components that endear themselves to the study and analysis of human activity patterns. Developed from studies of human migration, time-geography concepts lend themselves easily to activity analysis. The concepts offer elegant foundation to support the development of methods to analyze human activities. Therefore, this study develops an operational framework (Figure 3.1) to study activity patterns based on the principles and concepts of time-geography, such as space-time path, stations, activity constraints, and sequences of activity engagement. The activity analysis methods proposed in the study will glean activity patterns in respect to activity chaining propensity of individuals and their general activity profiles or their intensity of activity participation. The framework and indices are implemented in a case study using data collected from OSU campus, Stillwater.

Basically, the study (university) population can be categorized into two groups: employees and students. From these two groups, four potential activity participating groups are identified based on constraints inspired by their roles and their activity schedules on campus. These are the faculty, administrative and support staff (from the employee group), and graduate, and undergraduate students (from the student group). It is expected, based on differences in activity types and constraints that the activity patterns generated by these four groups would exhibit some but not exclusively, discernible differences. For example, among the student population, undergraduate students have a higher course credit load that involves taking and attending a larger number of classes in several more locations than graduate students. This schedule imposes restrictions on their ability to participate in other activities but also generates other activities associated with attending classes, e.g., homework assignments. The larger credit loads of undergraduate students require them to attend a greater number of class sessions of shorter duration, often strewn across different locations. Typically, this involves series of movements from one activity location to another that translates into activity chains.

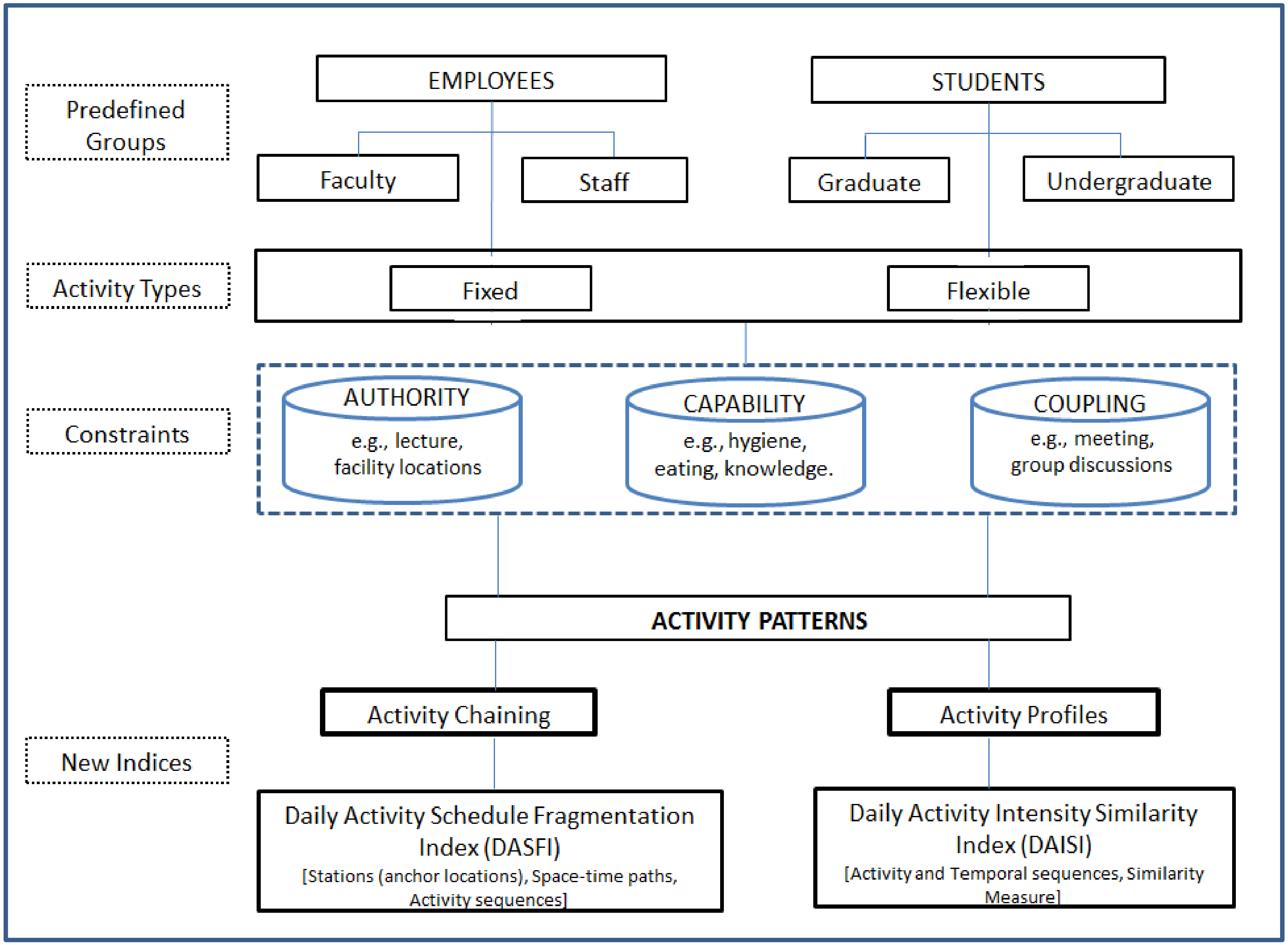



Figure 3.1: Operational framework for activity patterns analysis: a case study of Oklahoma State University

Strings of activity chains are therefore expected to characterize the activity patterns of the undergraduate student population. The graduate student, on the other hand, takes fewer classes of much greater intensity, in terms of time and associated activities. The rigors of graduate studentship require much longer periods of time for research work in laboratories, library or other locations. Longer spells of time may therefore be spent in fewer locations than may apply for undergraduate students. The activity pattern of graduate students is likely to exhibit close resemblance to the faculty's pattern because of the relative similarity between their activity schedules. Most graduate students have intensive research and teaching assistantship responsibilities not unlike the faculty (save for the class schedules). Their work schedules mimic those of the faculty and therefore their activity patterns may be similar as well.

The schedules for university employees are also different for faculty and staff. Faculty have the responsibility of teaching, researching and engaging in community development activities, which involve movement between more activity locations than is ordinarily obtained for most administrative and support staff. Faculty members are expected to teach at different classroom locations and conduct research work in offices, laboratories, or the library for long periods of time. Between these teaching responsibilities, the faculty could resort to a location on campus to wait for the next activity period or re-organize their remaining activity schedule. Consequently, it may be expected that the activity pattern of faculty may fall between the expected fragmented pattern of staff and the chained activity patterns of undergraduate students.

For administrative staff, most of them are generally ensconced in their office spaces where they are expected to discharge most of their duties. Activity movements outside of the offices, e.g., for meetings, lunch, or other responsibilities are usually followed by a return to their offices. Consequently, each foray away from the office fragments the activity pattern with the office largely serving as the anchor location. Most of the support staff also discharges their responsibilities at different locations. In most instances they are expected to return to its office location several times a day for new assignments or report on their assigned tasks. In this situation, it is easy to anticipate a fragmented activity pattern.

These expected patterns of activity schedules in the university are a function of several factors, which are adequately represented by the time-geography framework. These include the fact that the actual sequences in which activities are undertaken are captured and reproduced in the framework through space-time paths, space-time prisms, stations and activity constraints.

The space-time path (STP), a 3D trajectory that captures the entire human activity itinerary of an individual, is representative of the many concepts that time geography offers for the study of human activity. STP encompasses the activity location, the starting and ending times of each activity (and therefore activity duration) and the sequences in which the activities are conducted (Yu, 2006). But the framework extends beyond just the immediate activity undertakings to examine the potentials for activities that each person is allowed or capable of undertaking. The space-time prism measures the window of opportunity that individuals have within given constraints. For example, an engagement in a given activity demands the foregoing of other activity opportunities at other locations, especially within the given time frame. The space-time prism therefore measures the 3D sphere of access for activity participation for an individual. When projected to a 2D area, the space-time prism is known as "potential area" and measures the physical area on the earth's surface that an individual can have access to undertake activities, given the constraints. These constraints are one of the most important components of the time-geography framework and refer to the numerous conditions that need to be fulfilled or barriers to be overcome for activities to take place. Naturally, several of these constraints impose on the activity schedules of the different groups and are instrumental to the construction of expected group differences in activity chaining patterns.

## 3.2 Daily Activity Schedule Fragmentation Index (DASFI)

The first index presented in the study is the Daily Activity Schedule Fragmentation Index (DASFI), which measures the fragmentation of the activity schedules of individuals. DASFI is developed to investigate the activity chaining propensity of individuals. The index takes a disaggregated (i.e., individual-based) approach and employs space-time path as its basic construct for analysis. It also involves redefining the concept of anchor locations to include other potential locations other than home and workplaces only. Specifically, the study explicitly adopts the concepts of station, activity chains, space-time path and the sequences of human activity as the bedrock of the fragmentation index. These concepts and principles are employed and applied within the given constraints that condition or limit activity participation of individuals.

The location of human activities, referred to as station, is an important and integral part of the human activity experience. The station is the filter through which human activity paths pass through and are represented as “tubes” or “pillars.” They form the basis of interaction between individuals. When several paths flow through a given station, they form a “bundle.” The size of the bundle can be represented by the size of the pillar or tube (Thrift, 1977a).

The station is also the reference point for individual’s space-time paths and prisms. For example, the symmetry of a prism is determined by whether the station of activity origin of the individual is the same as the destination of activity. If the station of origin is same as destination then the space-time prism is considered to be symmetrical. Equally, the station is the activity location identifier in a space-time path and therefore is significant in considering the nature of individuals’ activity itinerary. Importantly, the station is recognized as the second of three models (along with Individual, and Time Supply and Demand models) of the societal operation (Thrift, 1977a). In essence, stations exert their own sets of constraints on the activity schedule of individuals, e.g., the opening hours and times of operations of a mall.

Stations can be used to analyze activity patterns at two levels: the individual and the collective levels. At the collective level, a station is the location of bundles (Thrift, 1977a) or where several activity paths pass through. The size of the bundle is indicative of the significance of the station as an activity location for the population. At the individual level, the station is the location at which most activities take place and therefore forms the basis for understanding individual activity participation and patterns.

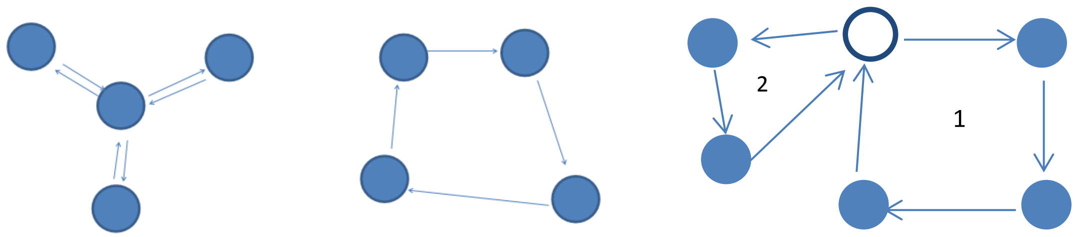
Generally, human activity participation confers some degree of pre-eminence on some activity locations than it does on others. These apparently more important locations are referred to as “anchor points or locations.” The concept of anchor points has metamorphosed and diversified through the years since it was first proposed by Lynch (1960) and hypothesized by Golledge and Spector (1978). The hypothesis "argued that key landmarks, nodes and areas individually and jointly 'anchor' subregions of space and link together hierarchically the items of information acquired about [that] space" (Couclelis et al., 1987:101). In essence, certain locations are "salient cues" in the environment around which individual activity experiences are anchored. Carrasco et al. (2008) invoked the concept of "social anchor points" in examining how the social network of individuals create their activity patterns, which is important in understanding the generation and spatial distribution of social activities and communication media behavior among individuals. Social anchor points are defined as "the main places where the individuals 'move around' when they interact with other network members" (Carrasco et al., 2008:5). These are usually key pivotal places that define the social activity space, e.g., homes, workplaces or even pubs and restaurants (Horton and Reynolds, 1971). Miller (2005) also acknowledged the home and workplace locations as important "space-time anchor points," which define the activity space of individuals.

There appears to be a consensus on the status of the home and workplace locations as activity anchor destinations (McGuckin and Nakamoto, 2004; Islam and Habib, 2012). However, locations other than the home and workplace can also function as key activity anchor points, depending on circumstances such as the perceptual or symbolic relevance of the location, relational-spatial (e.g., frequency of visits) and relational non-spatial (actual or potential significance) properties of the location (Couclelis et al., 1987). The frequency of visits and temporal duration at activity locations may be important defining principles for activity anchor points. For example, Ahas et al. (2010) employed temporal duration of two hours at an activity location to define and identify personal anchor points. Hanson and Huff's (1988) concept of "core stops" adopted repetitive activity-travel behavior (defined by four-attribute characteristics of activity, mode, arrival time and location) occurring at least three different days of the week to determine stability and regularity in individual activity patterns (Raux et al., 2011). Using repetitive behavior on activity locations, Raux et al. (2011) inferred a concentration of activity patterns around a few "anchor points or locations."

Though the importance of the home and workplace as key anchor points cannot be overemphasized, there are several merits to a more inclusive definition of an anchor point. For example, improved developments in information and communication technologies (ICT) have expanded the activity spaces of individuals and the confining regimes of the home and workplaces have been transcended (Kwan and Weber, 2003), which has encouraged fragmentation of activities and activity locations (Lenz and Nobis, 2007; Hubers et al., 2008; Alexander et al., 2011). Consequently, increasing number and types of activities are being undertaken away from the home and traditional workplace spaces. As weekend and leisure activities increase and social networking improves and develops locations other than the home and workplace may become more important environmental cues around which individuals organize their activity schedules. It would therefore be necessary, moving forward, to expand the scope of defining and identifying anchor locations. Temporal duration and frequency of visits to particular activity locations may be appropriate measures to adopt depending on the goals and scale of the study. For example, a long, unexpected and accidental delay at the airport may confer on the location (airport) an exaggerated significance in the schedule of activities for the day when a temporal duration principle is adopted. In similar vein, a frequent visit to the convenience room may have a similar effect for a frequency-of-visit principle.

Islam and Habib (2012) have indicated that it is traditional for an anchor point (usually the home or workplace) to be identified to classify trip chains. Understanding trip chains have been integral to transportation planning and modeling, as they reveal travel patterns. As demand for travel intensifies with increasing activity participation, there is an increased propensity to link various trips into a single journey called trip chains or trip tours. Activity-travel patterns therefore become more complex. Primerano et al. (2008) summarize several definitions of trip chains that have been postulated and used in the literature (e.g., Holzapfel, 1986; Thill and Thomas, 1987; McGuckin and Murakami, 1999). In general, a trip chain consists of trip segments between anchor locations. In some definitions, more than one anchor location, home and work, may be recognized (McGurkin and Murakami, 1995) while in some there is only a sole anchor location in a trip chain (Holzapfel, 1986). Holzapfel (1986) insists that a trip chain should consist of at least three trip segments or more.

Increasingly, the research focus has shifted from trip chains to activity chains with growing recognition of the derived demand for trip making. Trips are undertaken to fulfill the desire to participate in activities. Liu et al. (2008) argue that for individuals to attend more activities to get a higher utility, there is a need to chain activities one after another to decrease disutility caused by travel. They used the first diagram (a) in Figure 3.2 to show an activity pattern without activity chains, and the second figure (b) to illustrate an activity pattern with an activity chain. This definition of activity chains concurs with McNally's (2000). It is further argued here that a second string of activities after an initial return to base location constitutes a second activity chain and the location that binds these two activity chain episodes is the activity anchor location as illustrated in the third diagram (c) in Figure 3.2.




(a) (b) (c)

Figure 3.2: Activity patterns with or without activity chains and activity anchors

Hagerstrand's (1970) time geography presents an elegant frameowrk that bestows meaning on both activity chains and anchor locations. The space-time path traces the movement of an individual in space with respect to time (Miller, 2007). Activities are undertaken at space-time stations for limited time periods. An anchor location, based on time or frequency of visit, is invariably determined from the range of space-time stations at which the individual had participated in activities. One of the appeals of the use of space-time path in examining activity chains is the fact that it (space-time path) organizes the activity schedules of individuals in both spatial and temporal sequence in which the activities are undertaken. It is therefore a realistic and accurate depiction of the human activity participation experience. Identifying activity chains from a space-time path therefore may be more intuitive. As illustrated in Figure 3.2c, an activity chain is regarded as a series of activities undertaken at different locations between times spent at a sole anchor location. Once an anchor location is determined either by frequency of visits or duration of time, series of activities between visits to the location may be classified as activity chains and recognized as fragments of the space-time path. The anchor point is therefore at the epicenter of the disintegration of the space-time path into fragments that may or may not consist of activity chains. The resulting fragmentation index therefore allows for identification of the activity chaining propensity of individuals through a space-time path disintegration process that involves extracting movements to and from anchor locations to activity destinations.

The anchor location therefore can be viewed as an organizing point for the individual’s activity itinerary. The frequency of visit or the long temporal duration at a location suggests its preeminence in the schedule of the individual. It may be a location of convenience, comfort (zone) or it may be tied to an important work or social engagement. It may also be argued that the absence of the anchor location may have resulted in a substantially different activity pattern, both structurally and functionally. Consequently, it is proper to employ the anchor location as a central principle for organizing individual activity schedules. DASFI adopts this concept as the foundation upon which to fragment activity schedules and to measure the propensity of individuals to chain their activities. For each space-time path, the anchor location is determined on the basis of either being the most visited location or the location in which the largest amount of time is spent at.

The activity schedule of undergraduates is expected to be more intensive in regards to the number of activities they engage in. With larger number of classes, along with associated activities (e.g., group discussions, homework assignments, etc), it is envisaged that activity patterns will be less fragmented and therefore chained more often than not. For example, most of the undergraduate students do not have a permanent facility such as an office space to use as an organizing location. Consequently, they may employ several arbitrary locations while they wait for the next activity session (e.g., class lectures). The activity patterns will assume a continuous chain of activity locations with few returns to a sole location base that may be recognized as an anchor location. The faculty, staff and graduate students, however, have offices or laboratories to return to each time they undertake an activity outside of these locations. Their activity patterns would therefore be fragmented because there is a sole base location to return to and could easily be established as an anchor location.

Generally, participation in activities is constrained by physiological necessities (capability constraints, e.g., sleeping and eating), synchronizing activity times and/or places with other persons (coupling constraints, e.g., meetings and class lectures), and/or restrictions imposed by regulations (authority constraints, e.g., voting requirements). Consequently, for individuals to engage in the activities, they need to fulfill those conditions or overcome these barriers.

In a regulated system such as the university, authority constraints are probably the most important sets of constraints as exemplified by the university calendar and work and class schedules. In work environments, authority constraints are usually the most dominant constraint groups to influence activity patterns. When the university is in full session, the semester’s calendar sets the tone for most of the activities that are undertaken in the course of the semester. The work and class schedule also impose on the individual’s activity types, activity times (starting and ending times, therefore duration) and activity locations, which may or may not be desirable or convenient but which the individual participant is required to adhere to in order to partake in the activity. Work and class schedules are probably the most important authority constraints in a university campus. First, they dictate the times and locations of activities on campus. Duration of lectures and associated laboratory assignments, and administrative work are determined by already fixed work and class schedules as dictated by the university authority. These inflexible or fixed activities and locations usually form the fulcrum around which other activities are planned. By their fixed nature they lend themselves to being predictable and expected (habitual). Other activities such as entertainment, social events, leisure activities are examples of flexible activities whose times and locations are generally arranged to skirt around the fixed work and class schedules.

Second, authority constraints determine how much time is left for discretionary or flexible activities (activities for which rescheduling of time and location are easy). For example, if the work schedule of a professor in a given day requires extensive teaching assignments, then it would be expected that less time will be available for other non-work related or discretionary activities than at other days when teaching responsibilities are less intense. The daily activity pattern therefore is affected accordingly.

Third, the location of activity facilities on campus is also an important authority constraint, which is fixed and determines the dispersal of activity locations. Location of facilities influences decisions on the choice of mode to traverse the distance between activity locations, and the total duration of travel activities on campus. For example, a student with more wide-ranging locations for class lectures will naturally expend more time in moving between the dispersed activity locations than students with more closely-knit activity locations on campus.

Though not necessarily secondary to authority constraints, most activities with capability limitations in the university community are conducted around those that are influenced by the authority constraints. Similarly, activities that are subject to coupling constraints usually circumvent those with authority limitations imposed on them. These include the need for group discussions, homework assignments, seminars and workshops or conferences that are not included as part of the university's regular schedule of academic activities, even when they are important in the general scheme of things.

Because they are imposed by authority constraints, most of the fixed and inflexible activities become habitual or regular. Habitual activities are those that have become routine and are therefore repetitive, cyclical, or expected. These types of activities conform to activities with authority constraints. Pre-arranged activities are those for which location and time are planned before participation such as in activities with coupling constraints. These include all activities that require the synchronized participation of two or more people such as in group discussions. However, some coupling activities may also be spontaneous or impulsive, e.g., accepting an impromptu invitation to lunch. Spontaneous or impulsive activities are those which are undertaken on the spur of the moment and are neither habitual nor pre-arranged. An urge of bodily metabolism is hardly pre-arranged or planned even if it is routine in the sense that it is an expected part of human nature. These types of activities are usually individual-based and may be more appropriately thought of as being imposed by capability constraints.

It is imperative to point out that the boundaries between the influences of these constraints are usually fuzzy. For example, in a university, class schedules are imposed by authority constraints, however, attending a class involves being with other class mates and/or the professor, which is a coupling constraint.

Essentially, these constraints also set the tone for the intensity of activities conducted by different groups. Here, intensity is measured by the number of activities undertaken within an hourly period in any given day. Ihler and Smyth (2007) define human activity intensity as the "rate at which events occur." They maintained that these processes are typically inhomogeneous in time because they are the product of the aggregated behavior of individuals and therefore "exhibit a temporal dependence linked to the rhythms of the underlying human activity" (p.1). Such data, they suggest, is important for detecting unusual events and help in understanding behavior in the context of temporal patterns.

Basically, when certain responsibilities are fixed or inflexible they require that they be undertaken at certain locations and/or at certain time periods, e.g., attending classes. The duration and location of the class are set by the university calendar and schedule. Undergraduates with greater number of classes of relatively shorter duration to attend at different locations within a given day would, on average, undertake a greater number of activities in a day than a graduate student with fewer classes of longer duration. Consequently, it may be more common to have an undergraduate student complete a class session, move to another location and start a new class session all within a given hour period, while a graduate student may be attending only one class session for a three hour period. This data is usually captured in a component of activity patterns referred to as activity profiles. Activity profiles are typically analyzed using statistical tests such as Chi-Squared and ANOVA tests to investigate differences/variations in the proportions of groups performing a particular activity at any particular time period.

### 3.3 Daily Activity Intensity Similarity Index (DAISI)

The second index developed in the study is the Daily Activity Intensity Similarity Index (DAISI), which examines the similarity in the activity profiles or intensity (rate) of activity participation of individuals. The measure adopts the sequence of activity intensity of participants as the basic principle of time geography to build upon.

DAISI is developed based on sequential alignment of the number (or rate) of activities conducted by each individual within hourly time frames. The index (DAISI) uses a common time period for all individuals, synchronized from the first activity time to the last activity time such that all individuals have the same number of hours of activity time. Activity starting time for all individuals is therefore the same as the activity ending time. The period between the starting and ending times of activities are compartmentalized into hourly periods such that all individuals have equal number of hourly time frames, with same starting and ending times. This provides the common basis to sequentially compare the intensity of activities conducted by individuals.

Each activity conducted within the hourly period is counted and recorded for the hour. The number of activities in each time frame for an individual is compared to the corresponding number of activities in the time frames of another individual. Because all individuals have the same number of time frames with same starting and ending times, it allows for a sequential comparison of time frames in the activity profiles. The similarity index produces a matrix of values that measures the degree of similarity or dissimilarity between pairs of individuals in regards to their activity profiles. The values are then subjected to a cluster analysis such that groups of similar activity intensity are identified and their characteristics discerned. For example, most staff members have regular and stable work activity schedules. The regularity of their work tasks may also result in fewer records of activity types and number of activities. Undergraduate students, on the other hand, may display a higher frequency of activity engagement. They have more classes to attend, more homework assignments and examinations to write, more studying and research work to engage in, and more group discussions to attend to. Their activity profiles therefore may display greater numbers of activities per time frame, which would suggest a higher rate of activity engagement.

It is pertinent to point out that these expectations for groups do not always hold true for all individual members of the groups. Differences exist between individuals in each group and in some cases they may be significant. This explains why it is necessary to embrace disaggregated approaches to studying activity patterns. Both DASFI and DAISI appreciate this point and therefore measure fragmentation and activity profiles of individuals rather than of predefined groups of activity participants. The ensuing groups from the analysis are therefore more intuitive and form the basis for new classification of activity patterns. These patterns are then examined for their defining characteristics to understand and shed light on human interaction processes.

To accomplish this theoretical and analytical framework therefore appropriate methodology needs to be developed to collect the necessary data to fulfill the goals and objectives of the study.

CHAPTER IV

# METHODOLOGY

## 4.1 Introduction

This chapter discusses the methods and techniques of data collection, data presentation and data analysis. It also explains the choice of the study area and data collection methods.

## 4.2 Study Area

The study is restricted to the Stillwater campus of Oklahoma State University. The choice of the study area is informed by a few factors. First, the study intends to develop new indices to analyze activity patterns. This requires a known population to verify and validate the efficacy of the new indices. The population of the university, their responsibilities and schedules and general activity patterns and constraints are generally clear and known. This makes it easier to verify and make sense of the results that may be produced by the new indices.

Second, the nature of the data needed for the study is difficult to come by because it involves largely personal information. It is reasoned that it would be easier to collect the data from the university community because the researcher is part of the community.

Another advantage is the fact that members of the community are used to being involved with such surveys and may be more amenable to providing such information if the appropriate conditions (of anonymity and/or confidentiality) are met.

Finally, the university possesses the necessary structure and infrastructure to provide the necessary support information for a good research design. This includes a reliable sample frame of potential respondents that is provided by the office of Institutional Research and Information Management (IRIM) to researchers. These are some of the factors that informed the selection of the OSU Stillwater campus as a case study for the collection of data, verification, and validation of the indices developed in the study.

### 4.3 Sampling Design

Four different groups of participants have been identified as potential activity pattern groups in the Stillwater campus of OSU. These are undergraduate students, graduate students, staff, and faculty. The study collected data from a proportional number of these groups using a two-day activity-based survey.

First, the study aims to identify activity characteristics of respondents, and develop new indices to decipher and classify activity patterns. Consequently data are collected on the socio-demographic characteristics of respondents and their activity characteristics including location, types, duration, and nature of activities (habitual, pre-arranged or spontaneous). The data collected are used to develop and test the new methods of activity patterns that are developed in the study.

Second, the nature of the data and the data collection techniques make it difficult to aim for a large data set. Data needed include detailed movement of respondents, activity types and location, starting, and ending times of activities (duration). Privacy concerns are usually very difficult to overcome and only persons ready and willing to participate in the survey provided information for the study.

GPS technology has improved the accuracy and convenience of collecting both spatial and temporal activity data. Many studies have employed the technology to great benefit (e.g., Zhou and Golledge, 1999; Dykes and Mountain, 2003; Rinner, 2004; Andrieko et al., 2007) and most Departments of Corrections at the state levels in the USA use the technology to monitor parolees. Different types of GPS devices have been produced and can be worn with ease. Cell phones embedded with GPS also have proved to be useful tools in collecting data to measure activity and lifespaces of people in communities (Schenk et al., 2011).

Nevertheless, the application of GPS technology to collect spatial information is not very appealing to many people in the society. There is usually a strong distaste for techniques that appear to monitor and chronicle the activity and activity locations of persons. For activity analysis, participants are usually expected to wear a GPS tracking device that collects real-time information on all locations with great accuracy. After the data has been downloaded from the GPS device, participants are debriefed for detailed information on the activities they participated in at the different locations recorded by the GPS device, their modes of movement between activity locations, number of persons they may have undertaken the activity with, other activities being done simultaneously at locations and the purposes of the activities. In essence, a new questionnaire and/or interview are administered to each participant to make sense of the activity data collected using the GPS device. Such a rigorous procedure to insure that accurate and reliable data is collected may appear to be tedious and cumbersome to many potential respondents. Fortunately, the research problem of this study does not require detailed information on exact location of activities or even the particular routes taken between location points for analysis.

Of more importance are data on activity types, activity start and end times (activity duration), and activity sequencing. Consequently, the employment of GPS devices is not necessary for data collection and analysis. Moreover, preliminary investigation indicated that few respondents would be willing to provide activity information through the use of GPS technology. For the purposes of this study therefore an activity diary/questionnaire is adopted for data collection. This helped to circumvent these fundamental problems. Even at that only approximately 3% of the sample of potential respondents provided information on their activity characteristics. Though the response rate is low, the number of responses is sufficient to develop and test the new indices. Generally, activity-diary studies do not usually have large response rates (see Doherty, 2004; Frye et al., 2012).

Activity diaries have been employed in many studies of activity patterns including Clarke et al. (1981), Stopher (1992), Misra and Bhat (2000), Doherty (2004), Kenyon and Lyons (2007), and Eom et al. (2009). Some of these studies have incorporated an in-depth interview method to complement the activity diary.

To circumvent the problems associated with collecting data deemed by many potential respondents as personal information, the data collection was designed as an anonymous online survey. First the OSU Institutional Research Board (IRB) approved the data collection protocol after carefully scrutinizing it to ensure that it met the needed requirements and standards of strict confidentiality or anonymity that this type of research survey should adhere to. The initial application to conduct a human research on OSU Stillwater campus was submitted to the IRB in January, 2011. This included the application forms requesting for permission to conduct human research on campus, the activity diary/ questionnaire, the researcher's curriculum vitae and a recommendation letter from the advisor of the study. After a few corrections were made to the application at the instigation of the IRB, the protocol was finally approved on March 16, 2011 for a one-year period. The approval required the OSU Institutional Research and Information Management (IRIM) to provide the appropriate sample frame and generate the random sample of not more than 5,000 for the study. Being an online research survey, the researcher was advised to work in collaboration with the OSU Information Technology (IT) and OSU Communication Services to publish and send the email request letters to potential respondents.

Unfortunately, the number of responses fell short of expectations. By February 23, 2012, only 97 respondents, out of the 5,000 on the sample list, had filled out the activity dairy/questionnaire. Of this number, only 84 responses were complete enough to be used for analysis. On the advice of the dissertation committee, an application was made to the IRB for an extension of the IRB protocol to allow for collection of additional data. The extension was granted and the IRB protocol was extended to February 23, 2013. In the second session of data collection (February 2012 to February 2013), 66 new respondents were added with 63 responses fully complete to be used for analysis. In total

163 respondents filled out the activity diary/questionnaire and a total of 147 responses were complete and used for the study.

### 4.4 Sampling Method

The data were collected in a two-year period from April 2011 to February 2013. Most of the data were collected in the spring Semesters of 2011 (April), 2012 (January to April), and 2013 (January and February). This is typical of many activity studies, which collected data for one or two-day periods only because of the intensive nature of the data required (Eom et al., 2009; Jiang et al., 2012). The activity diary/questionnaire collected data for a two-day period within a week. This is typical of many studies that have used the activity diary/questionnaire largely because it is demanding and intensive (see Clarke et al., 1981; Stopher, 1992; Eom et al., 2009). Records from the OSU Institutional Research and Information Management (IRIM) department indicated that about 20,250 students were registered for the Spring 2011 semester at the Stillwater campus, and there were 1,500 faculty members and 4,450 staff. The sample frame from which samples could be generated was therefore 26,200.

However, the OSU Communication Services, which coordinates research activities on campus, along with the OSU Institutional Review Board (IRB) and the Institutional Research and Information Management (IRIM), has put a cap of 5,000 persons only per survey. Each semester, several permits are granted for surveys on campus and the risk of over-flooding people's emails with survey requests has become real. Consequently, the sample size of the study was 5,000 persons, which were randomly generated from the sample frame of 26,200 persons. To get a representative sample therefore, a proportional sampling technique was adopted for the four groups of potential respondents identified on campus (Faculty, Staff, Graduate students, and Undergraduate students). This breaks down as 3,000 undergraduate students, 750 graduate students, 375 faculty, and 875 staff.

The data were collected using an online survey. The activity diary/questionnaire was developed using SurveyMonkey, which is an online software that allows research surveys to be designed and administered. The results of the research survey were exported to an Excel database.

The researcher worked with the OSU Communication Services and then the OSU IT for several weeks to develop an appropriate strategy to send emails and the link to the research survey site (on SurveyMonkey) to all the 5,000 potential respondents. Since each email was to be personalized (i.e., addressed with the title and names of each potential respondent), permission was finally granted to the researcher to dispatch the emails through a Mail Merge service. There was to be an initial email sent to each person on the sample list, and only two email reminders to each person were allowed for the rest of the survey period. The researcher was advised to send not more than 2,000 emails in a day otherwise the OSU email system would recognize them (i.e., emails) as spam and shut down delivery from the initiating address probably for a week. The 2,000 emails were also to be spread over the entire day. Consequently, emails were sent to all 5,000 persons between Thursday, April 14 and Sunday, April 17, 2011. Five batches of 400 emails were sent at an interval of two hours each on Thursday, April 14 and Friday April 15; and, two batches of 500 emails were sent on Sunday, April 17, 2011. The emails contained a link to the research survey. Subsequently, the first set of email reminders were sent to the sample list between Thursday April 21 and Sunday, April 24, and the second set were sent between Thursday April 28 and Saturday, April 30. The email reminders appealed to respondents who had not filled out the questionnaire to kindly do so and thanked those who had provided information for the survey and asked them to ignore the reminder. By the second email, the list of those who had responded had been compiled and they were exempted from the reminder emails sent out.

The activity diary survey was uploaded and tested on Tuesday, April 12, 2011. Similar steps were taken after the extension to the IRB was granted in February, 2012 to enable the continuation of data collection till February 2013. An initial email soliciting cooperation from potential respondents was sent out to all persons in the sample frame. Another email reminding potential respondents was sent out two weeks after the initial email and a second reminder email was posted two weeks later. Each of these stages also involved several days as instructed by OSU IT. The data collection process was discontinued at the expiration of the IRB on February 23, 2013.

A lot of emphasis was put in making the survey anonymous in large part. No questions were asked on names and addresses or any personal information that may identify a respondent. An option to volunteer contact information for possible follow-up discussion was however incorporated to enable the researcher to conduct an in-depth interview with volunteer respondents, if necessary. The purpose of the in-depth interview was to glean information on the factors and constraints that people in the OSU community take into consideration in making decisions about what types of activities to engage in, the location of activities and when to participate in such activities. This follow-up option was eventually discarded in favor of complete anonymity of respondents, in the hope of improving the response rate.

## 4.5 Data Collection Design

Data on activities were collected through a comprehensive activity diary/questionnaire from faculty, staff, graduate, and undergraduate students on the OSU campus, Stillwater. Data included activity types, activity times (starting and ending times), activity locations, and activity scheduling process.

### 4.5.1 Activity Diary/Questionnaire Design

The activity diary/questionnaire consisted of two sections. The first section sought information on the socio-economic and demographic characteristics of participants, including gender, age, residential status (on-campus/off-campus), participant status (faculty, staff, graduate student, undergraduate student), residency status (Oklahoman, Out-of-State, International), and contact information (email) (optional). The second section was a table to collect information on activity type, activity location, activity starting and ending time (duration), travel modes between activity locations, and activity scheduling process (habitual activities such as lectures, sleeping, etc.; pre-planned activities, e.g., scheduled meetings; spontaneous activities, e.g., instantaneous activities). Codes for different activity types, travel modes and activity scheduling were provided to simplify data entry by respondents. A map of Stillwater showing census block groups and important and recognizable landmarks in each block group was provided to guide participants to enter information on activity locations outside the OSU campus. To boost anonymity of actual locations and encourage accurate completion of the diary, certain important generic locations (e.g., Wal-Mart East; Stillwater High School, Richmond Elementary School) were used as locations for activities outside the campus. Respondents were asked to select the generic location/landmark on the map that was closest to the actual address of their activity location.

### 4.5.2 Data Compilation and Processing

The SurveyMonkey, the online site which hosted the research survey, possesses several tools to process and store the data collected. First, it prepares an elaborate database into which the data are collected. The database can be uploaded as a Microsoft Excel file or a Microsoft Access file. This allows for individual researchers to customize the process of their data analysis. Second, each activity diary/questionnaire filled out by a respondent can be exported as is, which allows for easy verification of the data entered and the output of the information contained in the database. Third, summary of data and data entry are accessible to the researcher. The date of data entry by the respondent, IP address from which data is entered (optional, it can be turned off to maintain complete anonymity) and the general summary of each question on the activity diary/questionnaire are available. There are also options available to create and present summaries of both the characteristics of respondents and their responses in several forms, including tables and graphs. These are some of the features of the SurveyMonkey that endeared it to the researcher as a great data collection tool for this research survey. Since most of the work was done in Microsoft Excel, the data and summary were also downloaded in the Microsoft Excel option available.

### 4.6 Data Presentation and Analysis

Several techniques were employed to present and analyze the data. Socio-economic and demographic characteristics of respondents and activities, activity profiles, and activity transition are presented as graphs and tables.

The activity patterns of individuals are represented by space-time paths using Shaw and Yu's (2009a) custom extension of the Extended Time-Geographic Framework Tools in ArcGIS. The tool provides two options to create space-time paths: from point features and from line features. For point features, two fields are required from the database: a unique ID for every individual, and timestamp, which records the time when the information was collected at activity locations. The equivalent required fields for a line feature is the unique ID and the start time and end time of each activity.

For a point input feature class, the function simply connects the tracking points according to their temporal sequence. For a polyline input feature class, the function assumes that the end location of a previous trip is the start location of the next trip. This presupposes that the data set has complete records on trips with no missing gaps. To generate space-time paths from these trips, therefore, the function simply connects the start/end locations of trips according to their temporal sequence (Shaw and Yu, 2009).

The output in both situations is a graphical line trajectory that represents duration at activity locations as vertical line segments and movement between activity locations as tilted (non-vertical) line segments. The temporal dimension is represented as the third dimension. The study adopted the point feature option, which fitted better the format of the data collected.

The space-time path represents the individual pattern of activities, which is the crux of the new fragmentation index (Daily Activity Schedule Fragmentation Index - DASFI) that is developed in the study to analyze activity patterns. The fragmentation index measures the activity chaining propensity and/or the degree of disruption for a sequence of activity given the anchor location of an individual's activity schedule/itinerary. DASFI examines the propensity of individual for trip-chaining rather than undertaking solitary activity trips. Tour-based activities (activity-chaining) are not only more efficient but produce a different tapestry of activity patterns from those produced by solitary trip making. The fragmentation index (DASFI) also measures the significance of anchor points to the organization of individual activity schedules.

The similarity between and among the sequences of activity types by individuals are analyzed using the ClustalG software. Each activity type is coded by an alphabet or a number of alphabets, which are arranged in the sequence in which they are undertaken. The sequences are then compared using a similarity algorithm. A cluster of similar sequences allows for identification of factors that bind these patterns of activity sequences together and provides an understanding of how they influence human interaction at large.

The study develops a similarity index (Daily Activity Intensity Similarity Index - DAISI) that derives from the activity profiles of individuals, which is based on intensity of activities carried out within a time period. Activity profiles show the number of people that engage in each activity in each hourly time frame of the day (Eom et al., 2009). This provides useful insight into the significance of activity types, their timing and sequencing, and helps embed these important aspects of activity participation into the activity-based framework (Alam and Goulias, 1999).

The similarity index (DAISI) utilizes the proficiency in activity participation as a measure of activity patterns. DAISI measures the frequency (or rate) of activities within an hourly period by individuals and compares them for similarity. A cluster of similar activity frequency (or rate) sequences provides a new perspective to recognize and understand patterns of human activity, their intensity, timing, and sequencing.

The transition activity matrix is employed to examine activity sequencing by the different groups as well. The rows of the matrix represent current activities and the columns represent subsequent activities. The subsequent activity is the activity that succeeds a current activity. For a group of respondents, the activity sequencing pattern was established by identifying the succeeding activities with the highest frequency for each current activity. This identified patterns of activities that are important corollary of particular activities. Such information allows for detection of important activity types and therefore provides a platform to anticipate, plan and accommodate activities.

The next three chapters provide information on the characteristics of respondents and their activities, and on the development and implementation of the new indices using the data collected at the OSU campus, Stillwater.

# CHAPTER V

## DATA PRESENTATION

### 5.1 Introduction

The general characteristics of respondents and their activities are presented along with indices of activity pattern recognition such as activity profiles, activity transition, activity sequencing, etc. Graphs and tables are the primary methods of presenting data in the chapter. The chapter is intended to provide an overview of the socio-demographic characteristics of respondents, the types, frequency and duration of their activities and general activity characteristics of respondents at the OSU Stillwater campus.

### 5.2 Characteristics of Respondents

A total of 147 complete responses were received, which is only about 3% of the 5,000 persons sampled. This falls within the range of the expected percentage of respondents predicted in Chapter IV (Methodology). First, the nature and scope of the data sought was personal and time-consuming to provide, which may explain the low response rate. Secondly, the researcher did not have much control over the selection of potential respondents. Consequently, OSU university faculty, staff and students in Tulsa Campus were inadvertently included in the sample. This is apparent from the large volume of emails received by the researcher from Tulsa faculty, staff and students who indicated interest but could not participate as a result of their ineligibility because the study is based on the Stillwater Campus. Probably, the difference would not have been significant but it is difficult to quantify the number of probable respondents from Tulsa Campus that were, *ab initio*, rendered ineligible to respond as a result of this mix-up in sample selection. Table 5.1 presents characteristics of 147 respondents who correctly completed the activity diary/questionnaire.

Table 5.1: Characteristics of Respondents

| | **Characteristics** | | | |
|---|---|---|---|---|
| | **Respondents** | **Freq.** | **Survey %** | **OSU %** |
| **Role** | Faculty | 18 | 12.25 | 5.73 |
| | Staff | 14 | 9.52 | 16.98 |
| | Graduate | 44 | 29.93 | 15.00 |
| | Undergraduate | 71 | 48.3 | 62.29 |
| **Home Location** | On campus | 47 | 31.97 | NA |
| | Off Campus | 100 | 68.03 | NA |
| **Gender** | Female | 79 | 53.74 | 48.91 |
| | Male | 68 | 46.26 | 51.09 |
| **Residence** | Oklahoman | 81 | 55.1 | NA |
| | Out-of-State | 44 | 29.93 | NA |
| | International | 22 | 14.97 | 7.00 |
| **Age** | 18-24 | 75 | 51.02 | NA |
| | 25-29 | 16 | 10.88 | NA |
| | 30-34 | 12 | 8.16 | NA |
| | 35-39 | 12 | 8.16 | NA |
| | 40-44 | 8 | 5.44 | NA |
| | 45-49 | 7 | 4.76 | NA |
| | 50-54 | 5 | 3.4 | NA |
| | 55-59 | 7 | 4.76 | NA |
| | 60+ | 5 | 3.4 | NA |

OSU % calculated from Data obtained from IRIM Office
NA: Data Not Available

About 78% of respondents are students. Expectedly, most of the respondents also live off-campus; are Oklahoman residents; and are less than 30 years of age. There are more female respondents than male as well. This distribution of the broad groups of respondents (employee group: 21.77%; students group: 78.23%) reflects the general population distribution of the groups at the OSU Stillwater campus (employees: 22.72%, and students: 77.28%). However, the distributions within these broad groups are much different.

The faculty and graduate groups are overrepresented in the survey respondents, while the staff and undergraduate groups are underrepresented. It may be speculated that the nature of the survey appears to appeal more to the faculty and graduate groups, especially since they have either undertaken surveys of similar nature of their own or would be embarking on one soon. The difficulty of conducting a research survey may not therefore be lost on them and the desire to contribute to its success may be higher. They probably are more driven to response by their own personal experiences or the experiences of their colleagues, students and/or mentors/advisors. Most staff and undergraduates do not undertake in many research surveys and though they may respond to solicitations to provide information, they may not be as motivated as those (faculty and graduates) who are engaged in similar research work and either feel the desire to help or may be hoping that their own research surveys may be favorably responded to as well.

The international group also is overrepresented in the survey respondents. This may, however, be more cultural. The research survey requires some personal information from respondents that many Americans may consider very private. The researcher received many emails from potential respondents who indicated they would have liked to participate but feel the information being required of them is private and personal. Many internationals, especially from countries other than western, developed countries, may not find the information too private to share. This may have contributed to the overrepresentation of this group of respondents.

### 5.3 Activity Characteristics of Respondents

Table 5.2 presents the activity characteristics of the four groups of respondents. In descending order of cumulative duration, sleeping/resting/idle, researching/studying, work/teaching, and attending classes are the most significant. For graduate students, meals take the place of attending classes in the order of duration. For staff and faculty, work/teaching is the most prominent activity, followed by sleeping. For the staff group household activities are third in line, while meals are the third most prominent activities for the faculty group.

Expectedly, both faculty and staff did not attend any classes and therefore did not spend time on the activity. Faculty did not report engaging in either financial/banking or religious (religious) activities; neither did the staff report spending time on researching/studying activities. Generally, the amount of time expended on these activities fall within expected ranges. For example, on an average day sleeping may be the activity with the longest continuous duration for most participants. It is an important activity and a capability constraint, which all members of the four groups engage in. Consequently, it is expected to have the longest cumulative duration. It also has the longest average duration for all activities.

Table 5.2: Activity Frequencies, Duration and Average Durations

| | GRADUATES | | | UNDERGRADUATES | | | STAFF | | | FACULTY | | | ALL | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Activity | Freq. | Duration | Av. Duration | Freq. | Duration | Av. Duration | Freq. | Duration | Av. Duration | Freq. | Duration | Av. Duration | Freq. | Duration | Av. Duration |
| Attending Classes | 60 | 3271 | 54.52 | 265 | 22049 | 83.20 | 0 | 0 | 0.00 | 0 | 0 | 0.00 | 325 | 25320 | 77.91 |
| Shopping | 18 | 730 | 40.56 | 11 | 352 | 32.00 | 2 | 115 | 57.50 | 3 | 95 | 31.67 | 34 | 1292 | 38.00 |
| Communication | 52 | 2351 | 45.21 | 58 | 2945 | 50.78 | 8 | 344 | 43.00 | 12 | 725 | 60.42 | 130 | 6365 | 48.96 |
| Day Care/Medical | 5 | 100 | 20.00 | 6 | 36 | 6.00 | 6 | 320 | 53.33 | 2 | 10 | 5.00 | 19 | 466 | 24.53 |
| Exercise/Sports | 23 | 1665 | 72.39 | 45 | 3134 | 69.64 | 7 | 570 | 81.43 | 14 | 786 | 56.14 | 89 | 6155 | 69.16 |
| Financial/Banking | 2 | 20 | 10.00 | 1 | 5 | 5.00 | 1 | 20 | 20.00 | 0 | 0 | 0.00 | 4 | 45 | 11.25 |
| Social event/meeting | 32 | 1781 | 55.66 | 26 | 2504 | 96.31 | 6 | 495 | 82.50 | 7 | 530 | 75.71 | 71 | 5310 | 74.79 |
| Household Activities | 93 | 4910 | 52.80 | 42 | 2276 | 54.19 | 25 | 1956 | 78.24 | 30 | 1992 | 66.40 | 190 | 11134 | 58.60 |
| Leisure/Entertainment | 49 | 4740 | 96.73 | 94 | 10904 | 116.00 | 10 | 1560 | 156.00 | 22 | 2448 | 111.27 | 175 | 19652 | 112.30 |
| Meals | 148 | 5643 | 38.13 | 228 | 9395 | 41.21 | 45 | 1927 | 42.82 | 55 | 2543 | 46.24 | 476 | 19508 | 40.98 |
| Hygiene | 111 | 4200 | 37.84 | 169 | 7075 | 41.86 | 40 | 1600 | 40.00 | 37 | 1239 | 33.49 | 357 | 14114 | 39.54 |
| Researching/Studying | 179 | 22617 | 126.35 | 186 | 28332 | 152.32 | 0 | 0 | 0.00 | 33 | 3112 | 94.30 | 398 | 54061 | 135.83 |
| Sleeping/Resting/Idle | 103 | 25428 | 246.87 | 168 | 43741 | 260.36 | 36 | 7144 | 198.44 | 43 | 7521 | 174.91 | 350 | 83834 | 239.53 |
| Travel | 385 | 5164 | 13.41 | 777 | 12110 | 15.59 | 114 | 2471 | 21.68 | 139 | 2006 | 14.43 | 1415 | 21751 | 15.37 |
| Religious | 5 | 306 | 61.20 | 10 | 655 | 65.50 | 2 | 250 | 125.00 | 0 | 0 | 0.00 | 17 | 1211 | 71.24 |
| Work/Teaching | 68 | 7523 | 110.63 | 42 | 7510 | 178.81 | 55 | 10895 | 198.09 | 66 | 8250 | 125.00 | 231 | 34178 | 147.96 |
| TOTAL | 1333 | 90449 | 67.85 | 2128 | 153023 | 71.91 | 357 | 29667 | 83.10 | 463 | 31257 | 67.51 | 4281 | 304396 | 71.10 |

In descending order of significance, sleeping/resting/idle, work/teaching, researching/studying, and attending classes have the largest average durations. The case for attending classes is important for the fact that both staff and faculty groups did not report engaging in the activity at all. Nevertheless, attending classes is among the most frequent activities and have one of the largest cumulative and average activity durations (time span). Similar observation is made of researching/studying, which the staff group did not report participating in, yet it is also one of the most frequented, and ranks among the activities with most duration, both cumulatively and on average. These underscore the significance of these activities, which in themselves may not be surprising since the study area is an academic community. These observations, however, lend credence to much of the data and speak to its reliability.

The most frequently reported activities are attending classes for undergraduates, researching/studying for graduates and work/teaching for both staff and faculty. Sleeping/resting/idle have the highest average duration for all groups except undergraduates for whom researching/studying activity has the most average time. This may have resulted from the fact that undergraduate students have a larger number of classes and consequently larger number of homework assignments, studying periods for tests, peer group discussions and other necessary study-related activities to deal with.

The data presented in table 5.3 represents the activity diary of 147 respondents covering 265 days' worth of activities among them, and approximately 4,281 activity episodes (excluding travel events) for an average of 16.15 activity episodes a day per person.

Table 5.3: Summary of Activity Characteristics of Respondents

| | Freq. | Number of Activity days | Number of Activities | Average No. of Activities/Person | Average No. of Activities/day |
|---|---|---|---|---|---|
| Undergraduate | 71 | 122 | 2128 | 29.97 | 17.44 |
| Graduate | 44 | 84 | 1333 | 30.30 | 15.87 |
| Faculty | 18 | 32 | 463 | 25.72 | 14.47 |
| Staff | 14 | 27 | 357 | 25.50 | 13.22 |
| Total | 147 | 265 | 4281 | 29.12 | 16.15 |

As expected, undergraduate students have the highest number of activities per day. Graduate and undergraduate students have nearly identical average number of activities per person as do the staff and faculty. The large difference in average number of activities may be accounted for by the “attending class” activity, which both the staff and faculty did not report undertaking. This activity has increased the number of activities for both the graduate and undergraduate students, and increased their average number of activities per person. As expected, the average number of activities slightly increased, progressively, from staff to faculty to graduate and then undergraduate students in that order.

### 5.4 General Activity Pattern Analyses

Conventionally, some of the important variables to consider in activity pattern analysis include the activity profiles, activity transition and sequencing of activity types. These are important to understanding the nature of activities, activity scheduling processes and actual activity participation.

5.4.1 Activity Profiles

Activity profiles show the number of people that engage in each activity in each hourly time frame of the day (Eom et al., 2009). It is also employed to define the number or proportion of activities undertaken by groups of individuals within specified time frames. This provides useful insight into the significance of activity types, their timing and sequencing, and helps embed these important aspects of activity participation into the activity-based framework (Alam and Goulias, 1999).

Figure 5.1 shows the combined activity profiles of the four groups of respondents.

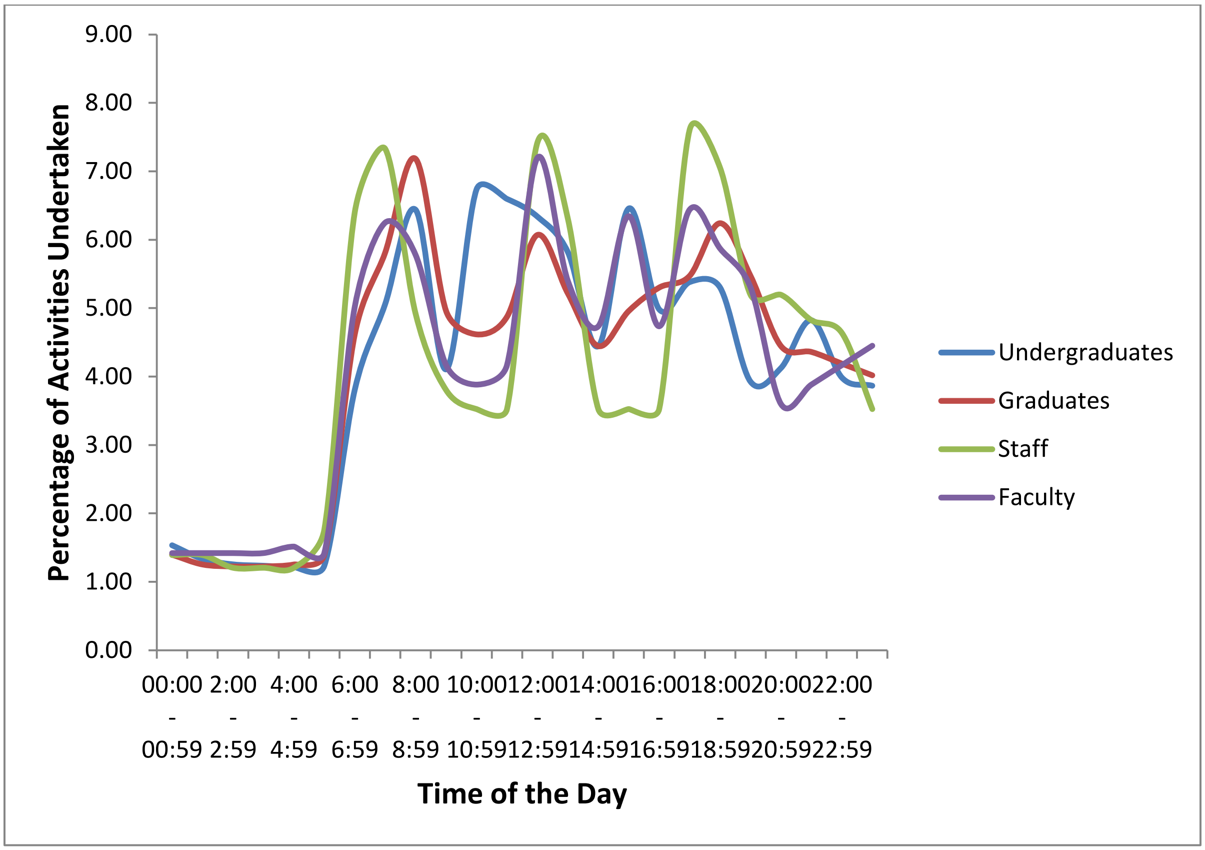



Figure 5.1: Activity profiles of groups

The activity profile is defined simply as the number of activities undertaken by an individual or group of people within given time periods, usually hourly time periods. It therefore measures the intensity of participation in activities. The general trend of the activity profiles suggests that the groups have similar trends of activity intensity. The groups have similar periods of activity peaks and apparent lull, with few differences or exceptions. Until 7:00 am there was hardly any difference in the activity profiles of the four groups. It may be assumed that most of the respondents may have been engaged in sleeping/resting/idle activity, which is suggested by the uniform level of activity reported in the early hours of the day. Between 7:00 am and 8:00 am, there was a spike in activities with graduate students and staff having slightly higher percentages (between 6.50% and 7.50%) of activities than faculty and undergraduate students. Between 8:00 am and 10:00 am the percentage of activity participation fell to around 4%. For undergraduates, this seeming lull period is shorter, with an immediate rise in activities from 9:00 am to a peak at between 10:00 am and 11:00 am. There is then a gradual decrease until 2:00 pm. This period (late morning to early evening) is the most active for undergraduates.

Faculty recorded its activity peak at between 12:00 noon and 1:00 pm, while the staff group has a triple peak between 7:00 am and 8:00 am, 12:00 noon and 1:00 pm, and between 5:00 pm and 6:00 pm. This interesting pattern points to a flurry of activities for the staff on the hours just before work, lunch time and after work, respectively. Between these periods, the staff has relatively low and stable activity intensity. This trend is expected and understandable. The staff group is the only group of respondents with fixed work starting and ending times with more routine activities. While the starting and ending times for teaching, attending classes and research/studying activities are usually staggered, the staffs have to report to work each day at 8:00 am and close at 5:00 pm, sometimes well before the end of school activities for other groups. At most days, the staff also performs routine activities, which may explain the relatively more stable and low levels of activities between the start of work, lunch and the end of work periods. Once the rhythm of the day's activities has been established, the staff has fewer other interruptions by other activities.

Another interesting pattern in activity profiles is between 1:00 pm and 4:00 pm. After the lunch period (during which activities for all groups are at low), while the activity intensity of other groups, especially undergraduates and faculty rise steeply (graduate students record a gentler rise in activity intensity after the lunch break), the staff group has a sharply contrasting low and stable activity register. All groups appear to wind down on activities from 6:00 pm, with staff having a slightly higher degree of activity intensity in the evening period, which declines more steeply as midnight approaches.

Generally, the activity profiles of all groups are similar, with peaks and dips recorded at similar time periods. Subtle differences in activity profile trends, however, exist that show the faculty and graduate students have similar activity profile patterns in the morning periods and the faculty and undergraduate students have similar patterns in the afternoon. Staff and faculty exhibit similar patterns (low activity intensity) in the late morning period as well. Undergraduates and faculty have four clearly demarcated peak periods each, while the staffs have three conspicuous peaks. The graduate students also have three peaks, with a slight hump of rising activity intensity between 3:00 pm and

4:00 pm that culminated in the third peak. This differentiates it from the lower and more stable activity levels of the staff at the same period.

The general trend of activity intensities for faculty, graduates and undergraduates is similar, which suggests that the academic component of their activities may be responsible. The trend supports the anecdotal evidence gathered about the relationship between the OSU class and lecture schedules and activity intensity trends displayed in Figure 5.1. In the morning periods, undergraduate classes are largely taught by graduate assistants. This may explain the slightly higher level of activity rates for these groups than for the faculty. In the afternoon period and evening periods, however, the graduate students attend their classes, which are invariable taught by the faculty. The activity rates of the groups therefore increase. Probably some of the higher level undergraduate courses are taught in the early afternoons. Most of these are taken by undergraduates (high rates) and taught by faculty (high rates). The graduate students' rates dip after their morning sessions. As the activity rates of the undergraduates fall towards the evening periods, the activity rates of the graduates and faculty pick up. This bears the mark of authority constraints, represented by the university class and work schedules, as discussed in Chapter III.

Overall, however, the marked difference in intensity of activities among the groups is the afternoon peak for other groups with the staff displaying a stable low activity regime at the same period. Other than this, the trends appear similar even though the degree of intensities suggests significant variations at different time frames. This similarity in trends may be due to the fact that the groups operate in the same environment and are influenced by similar sets of constraints especially authority constraints such as academic and work schedules that regulate campus activities of the groups for most of the day.

Using the percentages of activity intensity for hourly periods, the activity profiles of the four groups are presented in Table 5.4.

Table 5.4: Activity Profiles as Percentage of Number of Activities within Hourly Periods

| Time | Undergraduates | Graduates | Staff | Faculty |
|---|---|---|---|---|
| 00:00 - 00:59 | 1.53 | 1.40 | 1.39 | 1.42 |
| 1:00 - 1:59 | 1.34 | 1.25 | 1.39 | 1.42 |
| 2:00 - 2:59 | 1.26 | 1.23 | 1.21 | 1.42 |
| 3:00 - 3:59 | 1.24 | 1.23 | 1.21 | 1.42 |
| 4:00 - 4:59 | 1.22 | 1.25 | 1.21 | 1.52 |
| 5:00 - 5:59 | 1.24 | 1.40 | 1.76 | 1.42 |
| 6:00 - 6:59 | 3.81 | 4.62 | 6.40 | 5.02 |
| 7:00 - 7:59 | 5.06 | 5.81 | 7.33 | 6.25 |
| 8:00 - 8:59 | 6.44 | 7.18 | 4.92 | 5.78 |
| 9:00 - 9:59 | 4.11 | 4.96 | 3.80 | 4.17 |
| 10:00 - 10:59 | 6.74 | 4.62 | 3.53 | 3.88 |
| 11:00 - 11:59 | 6.60 | 4.87 | 3.53 | 4.17 |
| 12:00 - 12:59 | 6.34 | 6.07 | 7.42 | 7.20 |
| 13:00 - 13:59 | 5.82 | 5.22 | 6.31 | 5.40 |
| 14:00 - 14:59 | 4.44 | 4.45 | 3.53 | 4.73 |
| 15:00 - 15:59 | 6.46 | 4.96 | 3.53 | 6.34 |
| 16:00 - 16:59 | 4.98 | 5.30 | 3.53 | 4.73 |
| 17:00 - 17:59 | 5.38 | 5.47 | 7.61 | 6.44 |
| 18:00 - 18:59 | 5.30 | 6.24 | 7.05 | 5.87 |
| 19:00 - 19:59 | 3.93 | 5.47 | 5.19 | 5.30 |
| 20:00 - 20:59 | 4.13 | 4.45 | 5.19 | 3.60 |
| 21:00 - 21:59 | 4.82 | 4.36 | 4.82 | 3.88 |
| 22:00 - 22:59 | 3.99 | 4.19 | 4.64 | 4.17 |
| 23:00 - 23:59 | 3.87 | 4.02 | 3.53 | 4.45 |
| | 100 | 100 | 100 | 100 |

The table reveals that the largest percentage of the activities of undergraduates is in the late morning and early afternoon, while the staff group undertakes more activities in the early morning and evening hours.

The graduate students have their highest activity percentage in the morning, mid-afternoon and early evening periods. For the faculty group, activity rates are higher in the afternoon and evening periods. These distributions make sense.

There are many undergraduate class sessions in the morning periods that are taught by graduate teaching assistants. The graduate students in turn attend their class sessions in the latter parts of the day, which are taught by the faculty. Consequently, both undergraduates and graduate students have high activity percentages in the early morning period. The faculty group becomes more active in the afternoon and evening periods. These distributions support the suggestion made in the operational framework of the large influence that authority constraints exert in the university environment. The class and work/teaching schedules are part of the institutional regulatory mechanism of the university, and therefore, are part of the authority constraints.

### 5.4.2 Activity Sequencing Analysis

An activity sequence is the chain of consecutive activity episodes that makes up the activity schedule of an individual. Generally, an individual's activity pattern is a function of the roles he or she plays in the different socio-cultural and economic settings within which he or she operates (Harvey and Wilson, 2001). Two methods of activity sequencing analysis are employed: an activity transition matrix and an activity sequence alignment method using the ClustalG software package. A third method, daily activity intensity similarity index (DAISI) is proposed, developed and introduced in Chapter VII of this study to examine activity profiles using a sequential alignment method.

#### 5.4.2.1 Activity Transition Matrix

The activity transition matrix provides a platform to study and understand the activity sequencing behavior of groups (Eom et al., 2009; Lockwood et al., 2004). A matrix is designed to reveal the transition from one activity to another and it represents the likelihood that a subsequent activity of a certain type will occur given an episode of a current activity type (Lockwood et al., 2004).

Table 5.5 presents the activity matrix for all groups. All activities are allocated by their current and subsequent activities. Current activities, represented by the rows, are the "activities of the moment", the reference upon which the succeeding or subsequent activities, represented by the columns, are based. The value entry in each cell indicates the percentage of occurrence of a subsequent activity type after a particular current activity type.

Meals activity is the most frequent subsequent activity (appearing six times) following financial/banking (40%), work/teaching (33.03%), household (28.26%), researching/studying (22.83%), hygiene (19.54%) and religious (17.65%) activities. It (Meals) also ties with researching/studying and sleeping/resting/idle activities as the second most frequent activity type. Researching/studying activity is the second most frequent subsequent activity following day care/medical (30%), meals (27.96%), communication (19.83%), and meetings (14.29%) activities.

Table 5.5: Activity Transition Matrix (Percentage)

| | Subsequent Activities | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Current Activities** | Attending Classes | Shopping | Commu nication | Day Care/ Medical | Exercise/ Sports | Financial/ Banking | Social event/ meeting | Household Activities | Leisure/ Entertain | Meals | Hygiene | Research/ Study | Sleep/ Rest/Idle | Religious | Work/ Teaching |
| Attending Classes | 28.00 | 0.92 | 4.31 | 0.00 | 1.54 | 0.31 | 2.15 | 4.31 | 5.23 | 24.00 | 2.15 | 16.62 | 6.77 | 0.00 | 3.69 |
| Shopping | 0.00 | 16.22 | 2.70 | 0.00 | 2.70 | 0.00 | 8.11 | 21.62 | 5.41 | 16.22 | 2.70 | 16.22 | 5.41 | 0.00 | 2.70 |
| Communication | 9.09 | 1.65 | 1.65 | 0.00 | 3.31 | 0.00 | 2.48 | 4.13 | 7.44 | 8.26 | 9.92 | 19.83 | 16.53 | 0.83 | 14.88 |
| Day Care/Medical | 0.00 | 5.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 20.00 | 0.00 | 15.00 | 0.00 | 30.00 | 0.00 | 0.00 | 30.00 |
| Exercise/Sports | 0.00 | 1.16 | 3.49 | 0.00 | 0.00 | 0.00 | 1.16 | 4.65 | 10.47 | 26.74 | 41.86 | 5.81 | 2.33 | 0.00 | 2.33 |
| Financial/Banking | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 40.00 | 0.00 | 20.00 | 20.00 | 0.00 | 20.00 |
| Social event/meeting | 1.43 | 4.29 | 12.86 | 0.00 | 1.43 | 0.00 | 14.29 | 5.71 | 5.71 | 12.86 | 1.43 | 14.29 | 12.86 | 0.00 | 12.86 |
| Household Activities | 4.35 | 1.09 | 8.15 | 3.26 | 3.26 | 0.00 | 2.17 | 1.63 | 7.61 | 28.26 | 8.15 | 17.93 | 7.61 | 0.00 | 6.52 |
| Leisure/Entertain | 2.41 | 1.81 | 6.02 | 1.20 | 4.22 | 0.60 | 3.01 | 2.41 | 2.41 | 10.24 | 10.24 | 16.27 | 34.34 | 1.20 | 3.61 |
| Meals | 15.48 | 1.08 | 5.81 | 0.22 | 2.58 | 0.00 | 2.58 | 6.24 | 11.61 | 0.43 | 3.44 | 27.96 | 9.25 | 0.65 | 12.69 |
| Hygiene | 12.93 | 0.29 | 5.17 | 0.86 | 3.45 | 0.29 | 1.44 | 10.92 | 4.89 | 19.54 | 1.15 | 11.49 | 15.52 | 2.30 | 9.77 |
| Researching/Study | 15.49 | 0.52 | 3.94 | 1.05 | 2.89 | 0.26 | 2.89 | 6.56 | 8.92 | 22.83 | 8.92 | 3.15 | 14.96 | 0.52 | 7.09 |
| Sleep/Rest/Idle | 9.58 | 1.25 | 0.83 | 0.00 | 4.58 | 0.42 | 1.67 | 7.50 | 2.92 | 9.17 | 42.08 | 9.17 | 6.67 | 0.00 | 4.17 |
| Religious | 11.76 | 5.88 | 11.76 | 0.00 | 0.00 | 0.00 | 0.00 | 5.88 | 5.88 | 17.65 | 11.76 | 5.88 | 17.65 | 0.00 | 5.88 |
| Work/Teaching | 4.52 | 0.90 | 3.62 | 1.36 | 2.26 | 0.45 | 3.62 | 10.86 | 4.07 | 33.03 | 2.71 | 11.76 | 7.24 | 0.90 | 12.67 |

Meals are important capability constraint activities, which are necessary for sustenance and for engagement in other activities. It stands to reason that involvement in meals activities would be ubiquitous and would follow participation in other activities. Also, researching/studying activities, along with attending classes and work/teaching are the basic blocks of activities on a university campus. They (attending classes, researching/studying, and work/teaching) are controlled by strict university schedules and calendar, which impose authority constraint on undertaking them. Researching/studying is the only activity out of the three that possesses more leverage in terms of authority constraint and is undertaken to by three of the four groups in any large measure. Except for staff, other groups are invested in undertaking researching/studying activities with great frequency and for long periods of time. It makes sense therefore that it is one of the main activities (along with meals) that is most frequently undertaken.

Attending classes and meetings, e.g., social events, and group discussions, are the two activities which are the most frequent subsequent activities to themselves. This means that a class session or a meeting is often followed by another class session or meeting respectively. Shopping and work/teaching are the second most frequent subsequent activity after themselves. However, because shopping events are infrequent not much weight may be placed on its significance in this discussion.

It may be noted that students could have chains of class activities succeeding one another on busy school days or be engaged in a succession of group discussions. It is therefore common to have students rush out of one class to attend another class with no break in-between. This pattern therefore is expected for these activity types and reinforces the expectation (as described in Chapter III) that their activities may be more chained than other groups. Moreover, Table 5.2 shows that undergraduate students have the largest average frequency of attending classes (approximately four class sessions per person) against only approximately two class sessions per graduate student) (also as pointed out in Chapter III). Similar explanation can be extended for the meeting activities, a substantial number of which may have resulted from the large number of classes that the students undertake. Many students engage in class group discussions, collaborative homework assignments, etc and other such activities that result from their class schedules. Consequently, when a student takes many classes, the probability that they may participate in more joint class activities with other students is high.

These patterns were expected and explained in the operational framework (Chapter III) of the study. They are also borne out by the results of the new indices in the later chapters (VI and VII)

The use of conventional activity transition matrix and activity profiles as methods to analyze activity patterns of groups in any spatio-temporal settings has several limitations. First, both the activity transition matrix and activity profile are generalized and aggregated representations of activity participation characteristics. Though the percentages of likelihood can be calculated for both, the differences between individuals cannot be measured. Individual activity behaviors are therefore subsumed under a broader context that masks any significant patterns of activity outliers or extreme activity episodes. Second, even when statistically significant differences or variations between activity profiles or transition sequences can be measured among groups, these groups are usually pre-determined and the results are also aggregative. Activity sequential alignment methods generally account for these limitations and are more commonly being applied to activity analysis. The study applies the principle of sequential alignment to examine and analyze activity profiles of individuals through a similarity index.

### 5.4.2.2 Activity Sequence Alignment Method

Sequential alignment analysis defines measures of similarity between two or more character sequences that represent sequences of events (Wilson, 2001) with a view to identifying normative behavior and how such behaviors influence individual lives (Harvey and Wilson, 2001). Similarity between pairwise sequences of events are defined as the maximum matching score (or minimum conversion score) to convert one sequence into another through the processes of substitution, insertions and/or deletions. Figure 5.2 shows the graphical unit interface of the ClustalG package.

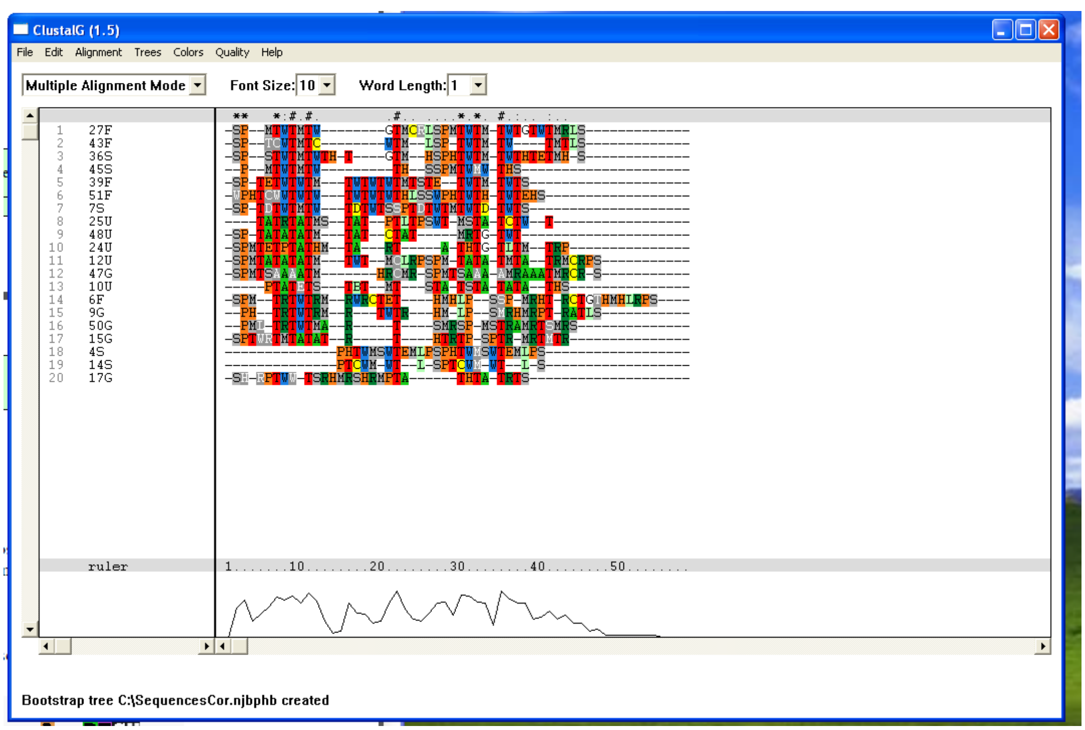



Figure 5.2: Interface of ClustalG

Generally, each activity type is represented by a single letter or a combination of letters. A sequence of activities is therefore represented by a chain of letters, each representing an activity type, in the sequence in which they occur. This allows for two or more sequences to be aligned one on top of the other and compared for similarity using penalty scores for insertions and deletions (referred to as Indels) (Wilson, 2001) from which a similarity score is calculated. Figure 5.2 presents the graphical unit interface (GUI) of ClustalG used in analyzing activity patterns by activity sequence alignment method. Table 5.6 presents the activity codes used in ClustalG to analyze the activity type sequences of individual respondents.

Table 5.6: Activity Codes for Sequential Alignment

| **Activity** | **Code** |
|---|---|
| Work/Teaching | W |
| Attending Classes | A |
| Shopping | B |
| Meals | M |
| Sleeping/Resting/Idle | S |
| Researching/Studying | R |
| Household Activities | H |
| Hygiene | P |
| Day care/Medical | D |
| Social Event/Meeting | G |
| Entertainment/Leisure | L |
| Exercises/Sports | E |
| Financial/Banking | F |
| Religious | V |
| Communication | C |
| Travel | T |

A sequence SHPMTA means: Sleep, H/hold, Hygiene, Meals, Travel, Attend classes in that sequence

The use of sequential alignment methods in social activity analysis is relatively new compared to the biological sciences where the method has been in use since the 1970s (e.g., Wilson, 1998a, b; Wilson, 1999; Harvey and Wilson, 2001; Wilson, 2001; Wilson et al., 2005; Shoval and Isaacson, 2007b). Wilson (2001) contends that the contribution of alignment methods to activity pattern analysis is the similarity measure generated as well as the patterns that are implied by a true multiple sequence algorithm. First, groups of activity participants are identified by the alignment of their activity sequences. They are therefore not pre-defined or pre-determined. The advantage of this is that any hidden patterns or underlying influences to the patterns could be discerned by examining the characteristics of individuals that make up the clusters. Second, the alignment method allows for comparison between any pairs of individuals on large scale bases. This allows for easy detection of abnormal patterns of behavior.

Using the ClustalG package the sequential alignment of activities of the respondents was analyzed. The result was augmented by the employment of TreeView software package (Page, 1996) to produce graphical representation of clusters of similarly aligned activity sequences (Figure 5.3).

The sequence alignment method identified four broad clusters based on activity sequencing patterns of respondents. Two of these groups are populated mostly by undergraduates; the third group has an overwhelming majority of faculty and staff involved; and, the fourth cluster has the largest number of graduate students of any group.

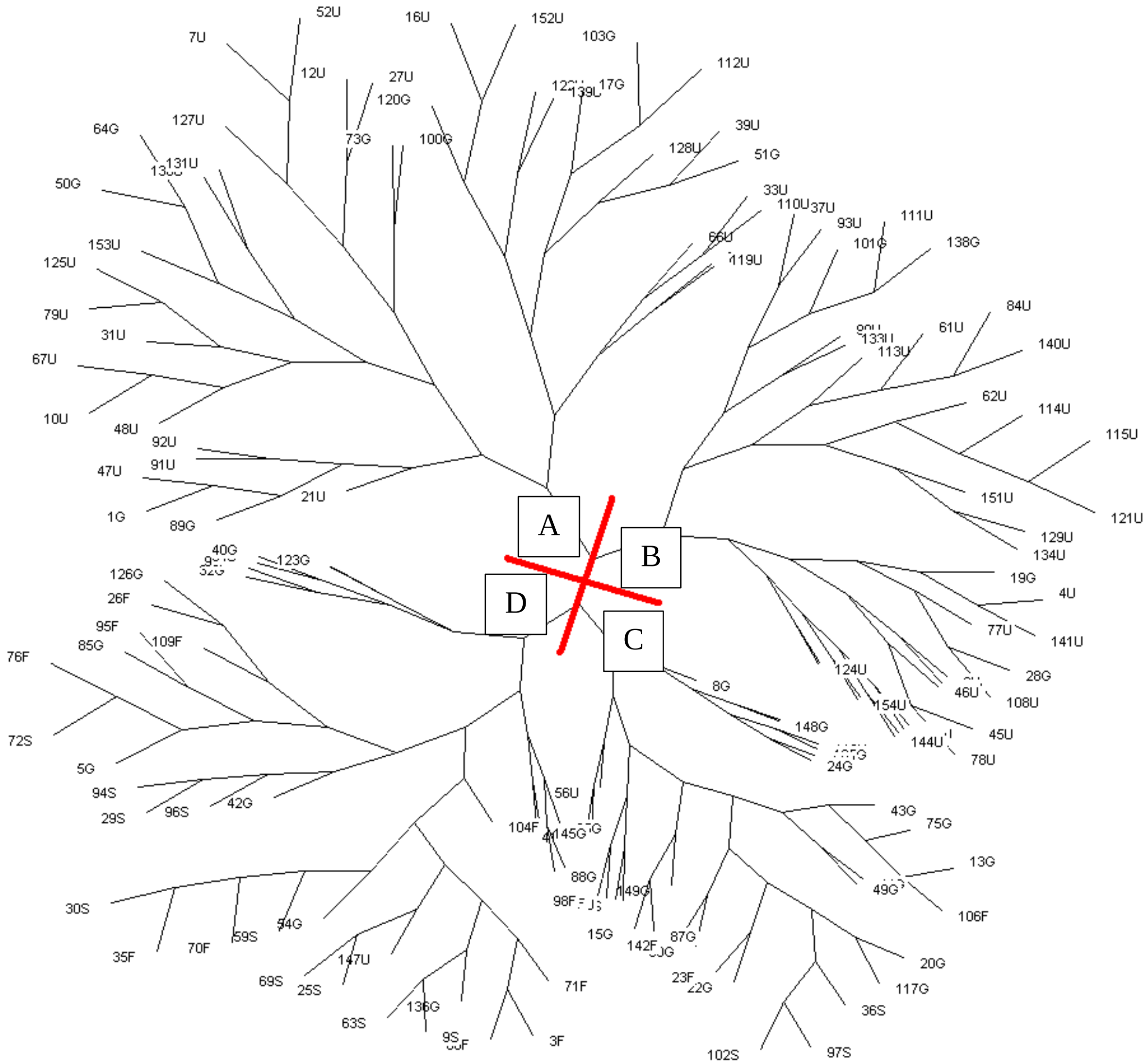


Figure 5.3: Radial tree illustration of similarity in activity patterns of respondents

*Undergraduate students (U), graduate students (G), Faculty (F), and Staff (S).

Table 5.7 presents a summary of the percentage of each group in the four pattern groups. The table indicates that groups 1 and 2 are Patterns A and B respectively, and they are both dominated by undergraduate students, while groups 3 and 4 are presented as

Patterns C and D and are dominated by graduate students, and the university employee (faculty and staff) groups respectively.

Table 5.7: Summary of Percentage of Respondent Groups in Patterns

| | Pattern A | Pattern B | Pattern C | Pattern D | Total |
|---|---|---|---|---|---|
| Faculty | 0 (0%) | 0 (0%) | 4 (22.22%) | 14 (77.78%) | 18 (100%) |
| Staff | 0 (0%) | 0 (0%) | 4 (28.57%) | 10 (71.43%) | 14 (100%) |
| Graduate | 10 (22.72%) | 4 (0.09%) | 19 (43.18%) | 11 (25%) | 44 (100%) |
| Undergraduate | 30 (42.25%) | 33 (46.48%) | 5 (7.04%) | 3 (4.23%) | 71 (100%) |
| Total | 40 (27.21%) | 37 (25.17%) | 32 (21.76%) | 38 (25.85%) | 147(100%) |

The distinguishing characteristics of the four patterns are examined and presented in the following subsequent sections.

5.4.2.2.1 Pattern A: Undergraduate Students Group

The most distinguishing features of this group are attending classes and studying/research activities. This group is dominated by the undergraduate students (75% of the group) who reported an average of four class sessions per person. The other 25% is made up of graduate students. Table 5.8 presents the activity matrix of this group.

There are no faculty or staff groups included in this cluster. The group has the highest average frequency and average duration time spent on attending classes (four class sessions; 304 minutes per person) and researching/studying (four sessions; 616 minutes per person) activities. This makes sense considering the high number of classes that undergraduates take and the large workload including examinations and homework assignments that come with each class.

Table 5.8 Activity Matrix for Pattern A (Percentages)

| | Subsequent Activities | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Current Activities** | Attending Classes | Shopping | Commu nication | Day Care/ Medical | Exercise/ Sports | Financial/ Banking | Social event/ meeting | Household Activities | Leisure/ Entertain | Meals | Hygiene | Research/ Study | Sleep/ Rest/Idle | Religious | Work/ Teaching |
| Attending Classes | 31.08 | 0.68 | 4.73 | 0.00 | 0.68 | 0.68 | 2.03 | 2.03 | 0.68 | 27.03 | 2.03 | 22.30 | 3.38 | 0.00 | 2.70 |
| Shopping | 0.00 | 0.00 | 14.29 | 0.00 | 0.00 | 0.00 | 14.29 | 57.14 | 0.00 | 14.29 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Communication | 11.11 | 0.00 | 3.70 | 0.00 | 0.00 | 0.00 | 0.00 | 7.41 | 14.81 | 7.41 | 0.00 | 18.52 | 22.22 | 3.70 | 11.11 |
| Day Care/Medical | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 33.33 | 0.00 | 16.67 | 0.00 | 50.00 | 0.00 | 0.00 | 0.00 |
| Exercise/Sports | 0.00 | 4.35 | 4.35 | 0.00 | 0.00 | 0.00 | 4.35 | 0.00 | 0.00 | 30.43 | 56.52 | 0.00 | 0.00 | 0.00 | 0.00 |
| Financial/Banking | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 100.00 | 0.00 | 0.00 | 0.00 |
| Social event/meeting | 0.00 | 0.00 | 11.11 | 0.00 | 0.00 | 0.00 | 11.11 | 0.00 | 5.56 | 16.67 | 0.00 | 22.22 | 22.22 | 0.00 | 11.11 |
| Household Activities | 13.51 | 0.00 | 10.81 | 5.41 | 0.00 | 0.00 | 2.70 | 2.70 | 0.00 | 21.62 | 13.51 | 27.03 | 0.00 | 0.00 | 2.70 |
| Leisure/Entertain | 0.00 | 8.33 | 16.67 | 0.00 | 4.17 | 0.00 | 0.00 | 4.17 | 0.00 | 4.17 | 16.67 | 16.67 | 29.17 | 0.00 | 0.00 |
| Meals | 26.05 | 0.00 | 1.68 | 0.00 | 1.68 | 0.00 | 2.52 | 3.36 | 5.04 | 1.68 | 1.68 | 40.34 | 14.29 | 0.84 | 0.84 |
| Hygiene | 21.84 | 0.00 | 4.60 | 1.15 | 0.00 | 0.00 | 0.00 | 11.49 | 5.75 | 16.09 | 0.00 | 16.09 | 14.94 | 4.60 | 3.45 |
| Researching/Study | 21.48 | 1.48 | 2.22 | 2.22 | 5.93 | 0.00 | 3.70 | 4.44 | 4.44 | 18.52 | 8.15 | 4.44 | 20.00 | 0.74 | 2.22 |
| Sleep/Rest/Idle | 12.68 | 0.00 | 0.00 | 0.00 | 9.86 | 0.00 | 4.23 | 7.04 | 4.23 | 7.04 | 40.85 | 8.45 | 2.82 | 0.00 | 2.82 |
| Religious | 28.57 | 14.29 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 28.57 | 0.00 | 0.00 | 28.57 | 0.00 | 0.00 |
| Work/Teaching | 6.25 | 0.00 | 0.00 | 0.00 | 12.50 | 0.00 | 0.00 | 0.00 | 0.00 | 50.00 | 0.00 | 18.75 | 12.50 | 0.00 | 0.00 |

There are no faculty or staff groups included in this cluster. The group has the highest average frequency and average duration time spent on attending classes (four class sessions; 304 minutes per person) and researching/studying (four sessions; 616 minutes per person) activities. This makes sense considering the high number of classes that undergraduates take and the large workload including examinations and homework assignments that come with each class.

Attending classes, researching/studying and meals are the most frequent current activities, with a frequency of 148, 135 and 119 recorded cases respectively. Research/studying activities are the most frequent subsequent activity (five times) in the group. This is followed by attending classes (four times) and Sleeping/Resting/Idle (four times) activities.

The group is made up of young students aged between 18 and 24. Only four respondents are above the age of 30 (all graduate students). There are 27 females (69%) in the group, which is 34.62% of female respondents, the highest among the groups. Pattern A is similar to Pattern B in many respects.

5.4.2.2.2 Pattern B: Undergraduate Students Group

This group is also dominated by undergraduate students (33 students) with four graduate students and no faculty or staff. Its most distinguishing feature is the high frequency of attending classes with an average of four class sessions per person and an average duration of 285 minutes per person. Research/studying activities returned only an average two sessions per person and an average duration of 337 minutes per person. This group has a considerable amount of time spent on working/teaching activities (164 minutes per person) than those in Pattern A (85 minutes). So it is the working population of the undergraduate student respondents, which is its major distinguishing difference with the Pattern A. Table 5.9 presents the activity matrix of the group

Table 5.9: Activity Matrix for Pattern B (Percentages)

| | Subsequent Activities | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Current Activities** | Attending Classes | Shopping | Commu nication | Day Care/ Medical | Exercise/ Sports | Financial/ Banking | Social event/ meeting | Household Activities | Leisure/ Entertain | Meals | Hygiene | Research/ Study | Sleep/ Rest/Idle | Religious | Work/ Teaching |
| Attending Classes | 29.10 | 0.75 | 3.73 | 0.00 | 4.48 | 0.00 | 1.49 | 3.73 | 8.96 | 23.13 | 0.75 | 10.45 | 8.21 | 0.00 | 5.22 |
| Shopping | 0.00 | 16.67 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 16.67 | 16.67 | 33.33 | 0.00 | 16.67 | 0.00 | 0.00 | 0.00 |
| Communication | 13.51 | 5.41 | 0.00 | 0.00 | 0.00 | 0.00 | 2.70 | 2.70 | 5.41 | 8.11 | 13.51 | 24.32 | 21.62 | 0.00 | 2.70 |
| Day Care/Medical | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Exercise/Sports | 0.00 | 0.00 | 4.17 | 0.00 | 0.00 | 0.00 | 4.17 | 0.00 | 12.50 | 16.67 | 45.83 | 12.50 | 4.17 | 0.00 | 0.00 |
| Financial/Banking | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Social event/meeting | 7.69 | 0.00 | 23.08 | 0.00 | 7.69 | 0.00 | 15.38 | 0.00 | 7.69 | 7.69 | 0.00 | 0.00 | 23.08 | 0.00 | 7.69 |
| Household Activities | 11.11 | 0.00 | 11.11 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 22.22 | 22.22 | 11.11 | 16.67 | 5.56 | 0.00 | 0.00 |
| Leisure/Entertain | 9.52 | 0.00 | 9.52 | 0.00 | 1.59 | 0.00 | 6.35 | 1.59 | 0.00 | 14.29 | 3.17 | 20.63 | 26.98 | 1.59 | 4.76 |
| Meals | 20.34 | 0.00 | 7.63 | 0.00 | 5.08 | 0.00 | 3.39 | 0.00 | 16.10 | 4.24 | 5.93 | 22.88 | 7.63 | 0.00 | 6.78 |
| Hygiene | 24.18 | 0.00 | 5.49 | 0.00 | 5.49 | 0.00 | 2.20 | 3.30 | 4.40 | 20.88 | 5.49 | 7.69 | 13.19 | 1.10 | 6.59 |
| Researching/Study | 20.00 | 0.00 | 5.88 | 0.00 | 0.00 | 0.00 | 1.18 | 2.35 | 18.82 | 23.53 | 15.29 | 0.00 | 12.94 | 0.00 | 0.00 |
| Sleep/Rest/Idle | 15.15 | 3.03 | 3.03 | 0.00 | 3.03 | 0.00 | 0.00 | 4.55 | 4.55 | 12.12 | 37.88 | 10.61 | 3.03 | 0.00 | 3.03 |
| Religious | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 50.00 | 0.00 | 0.00 | 0.00 | 50.00 | 0.00 | 0.00 |
| Work/Teaching | 28.57 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 17.86 | 35.71 | 0.00 | 10.71 | 7.14 | 0.00 | 0.00 |

Despite the predominance of research/studying activities in the group, attending classes is the most frequent current activity (134 cases), followed by meals (118 cases). The most frequent subsequent activity, however is researching/studying activities (four times), followed by sleeping/resting/idle activities (three times). The average age group of this cluster is also 18-24, with only 4 persons above the age of 30. It also consists of 22 females (59.46% of this group), which is 28.21% of the female respondents in the study.

5.4.2.2.3 Pattern C: Graduate Group

Researching/studying (average of three cases, 360 minutes per person), meals (average of three cases, 110 minutes per person), and hygiene (average of two cases, 76 minutes per person) activities are the most frequent activities. Work/teaching activities (average of only one case per person) are the most dominant in terms of duration other than sleeping/resting/idle activities. It has an average of 196 minutes per person. This is followed by entertainment/leisure activities (average case of one) with average time duration of 127 minutes per person.

The dominance of research/studying and work/teaching activities in this group is expected since the group is made up largely of graduate students. Research/studying activities are as much an integral part of their studies. They have fewer class credit loads and sessions, which are complemented by intense research/studying activities than undergraduate students. They may also be involved in teaching activities, which may be part of their assistantship or work on research projects for professors and departments as graduate research assistants. This explains the dominating feature of the group. Table 5.10 presents the activity matrix for Pattern C.

Table 5.10: Activity Matrix for Pattern C (Percentages)

| | Subsequent Activities | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Current Activities** | Attending Classes | Shopping | Commu nication | Day Care/ Medical | Exercise/ Sports | Financial/ Banking | Social event/ meeting | Household Activities | Leisure/ Entertain | Meals | Hygiene | Research/ Study | Sleep/ Rest/Idle | Religious | Work/ Teaching |
| Attending Classes | 0.00 | 0.00 | 4.35 | 0.00 | 0.00 | 0.00 | 4.35 | 17.39 | 8.70 | 17.39 | 4.35 | 34.78 | 8.70 | 0.00 | 0.00 |
| Shopping | 0.00 | 16.67 | 0.00 | 0.00 | 0.00 | 0.00 | 16.67 | 25.00 | 0.00 | 16.67 | 0.00 | 25.00 | 0.00 | 0.00 | 0.00 |
| Communication | 8.33 | 0.00 | 0.00 | 0.00 | 8.33 | 0.00 | 8.33 | 4.17 | 4.17 | 8.33 | 20.83 | 25.00 | 8.33 | 0.00 | 4.17 |
| Day Care/Medical | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 100.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Exercise/Sports | 0.00 | 0.00 | 6.67 | 0.00 | 0.00 | 0.00 | 0.00 | 20.00 | 13.33 | 26.67 | 26.67 | 0.00 | 6.67 | 0.00 | 0.00 |
| Financial/Banking | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 50.00 | 0.00 | 50.00 | 0.00 | 0.00 | 0.00 |
| Social event/meeting | 0.00 | 18.75 | 6.25 | 0.00 | 0.00 | 0.00 | 18.75 | 18.75 | 6.25 | 0.00 | 0.00 | 25.00 | 0.00 | 0.00 | 6.25 |
| Household Activities | 4.69 | 3.13 | 1.56 | 0.00 | 1.56 | 0.00 | 0.00 | 0.00 | 4.69 | 53.13 | 9.38 | 12.50 | 3.13 | 0.00 | 6.25 |
| Leisure/Entertain | 2.33 | 2.33 | 0.00 | 0.00 | 2.33 | 2.33 | 0.00 | 2.33 | 2.33 | 6.98 | 25.58 | 13.95 | 37.21 | 0.00 | 2.33 |
| Meals | 6.14 | 1.75 | 7.02 | 0.88 | 1.75 | 0.00 | 2.63 | 8.77 | 15.79 | 0.00 | 4.39 | 33.33 | 8.77 | 0.88 | 7.89 |
| Hygiene | 4.82 | 1.20 | 4.82 | 0.00 | 1.20 | 0.00 | 1.20 | 10.84 | 7.23 | 20.48 | 0.00 | 15.66 | 21.69 | 2.41 | 8.43 |
| Researching/Study | 5.61 | 0.93 | 5.61 | 0.00 | 2.80 | 0.93 | 1.87 | 11.21 | 8.41 | 26.17 | 8.41 | 5.61 | 11.21 | 0.93 | 10.28 |
| Sleep/Rest/Idle | 2.13 | 0.00 | 0.00 | 0.00 | 2.13 | 0.00 | 2.13 | 19.15 | 0.00 | 14.89 | 42.55 | 4.26 | 8.51 | 0.00 | 4.26 |
| Religious | 0.00 | 0.00 | 0.00 | 20.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 40.00 | 20.00 | 0.00 | 0.00 | 20.00 |
| Work/Teaching | 0.00 | 0.00 | 2.44 | 0.00 | 7.32 | 0.00 | 2.44 | 12.20 | 0.00 | 24.39 | 2.44 | 26.83 | 7.32 | 2.44 | 12.20 |

Meals (114 cases) and researching/studying (107 cases) activities are the most frequent current activities, while researching/studying is the most important subsequent activity (seven times), followed by meals (four times). This underscores the influence of researching/studying activities in this group.

The group has a slightly more male population (17, which is 53.13% of the group and 24.64% of the total number of male respondents in the study). Expectedly, the group has a slightly older population than Patterns A and B, with 17 (53.13%) of the group being 30 years and older.

5.4.2.2.4 Pattern D: Faculty and Staff Group

This group is dominated by work/teaching activities. It is by far the most important activity in the pattern. This group reported the highest frequency, total duration and average duration for work/teaching activities reported by any group. There is an average of four cases per person, with an average duration of 656 minutes per person. It actually has more duration than sleeping/resting/idle activities (average of two cases per person and average of 598 minutes per person), which has the greater number of cases and average duration per person in other groups. The next closest activity of significance is researching/studying (average of two cases per person and average duration of 200 minutes per person). Expectedly, therefore, work/teaching activities are the most frequent current activities (141 cases) and also the most frequent subsequent activity (seven times). This is followed by meals as current activity (105 cases) and as subsequent activity (three times). Table 5.11 presents the activity matrix for the group.

Table 5.11: Activity Matrix for Pattern D (Percentages)

| Group 4 | Subsequent Activities | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Current Activities** | Attending Classes | Shopping | Commu nication | Day Care/ Medical | Exercise/ Sports | Financial/ Banking | Social event/ meeting | Household Activities | Leisure/ Entertain | Meals | Hygiene | Research/ Study | Sleep/ Rest/Idle | Religious | Work/ Teaching |
| Attending Classes | 22.22 | 11.11 | 0.00 | 0.00 | 11.11 | 0.00 | 11.11 | 22.22 | 0.00 | 11.11 | 0.00 | 11.11 | 0.00 | 0.00 | 0.00 |
| Shopping | 0.00 | 11.11 | 0.00 | 0.00 | 11.11 | 0.00 | 0.00 | 0.00 | 11.11 | 11.11 | 11.11 | 11.11 | 22.22 | 0.00 | 11.11 |
| Communication | 0.00 | 2.94 | 0.00 | 0.00 | 8.82 | 0.00 | 2.94 | 2.94 | 2.94 | 8.82 | 8.82 | 8.82 | 14.71 | 0.00 | 38.24 |
| Day Care/Medical | 0.00 | 8.33 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 16.67 | 0.00 | 8.33 | 0.00 | 16.67 | 0.00 | 0.00 | 50.00 |
| Exercise/Sports | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 3.85 | 15.38 | 30.77 | 30.77 | 7.69 | 3.85 | 0.00 | 7.69 |
| Financial/Banking | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 100.00 |
| Social event/meeting | 0.00 | 0.00 | 9.52 | 0.00 | 0.00 | 0.00 | 14.29 | 4.76 | 4.76 | 14.29 | 4.76 | 9.52 | 4.76 | 0.00 | 33.33 |
| Household Activities | 0.00 | 0.00 | 11.11 | 6.35 | 7.94 | 0.00 | 3.17 | 0.00 | 12.70 | 14.29 | 4.76 | 11.11 | 19.05 | 0.00 | 9.52 |
| Leisure/Entertain | 0.00 | 0.00 | 3.57 | 7.14 | 10.71 | 0.00 | 3.57 | 7.14 | 7.14 | 3.57 | 3.57 | 14.29 | 25.00 | 3.57 | 10.71 |
| Meals | 0.95 | 2.86 | 7.62 | 0.00 | 1.90 | 0.00 | 1.90 | 13.33 | 6.67 | 0.00 | 2.86 | 13.33 | 5.71 | 0.95 | 41.90 |
| Hygiene | 1.22 | 0.00 | 9.76 | 2.44 | 6.10 | 1.22 | 2.44 | 15.85 | 2.44 | 15.85 | 0.00 | 7.32 | 15.85 | 0.00 | 19.51 |
| Researching/Study | 8.93 | 0.00 | 3.57 | 1.79 | 0.00 | 0.00 | 5.36 | 10.71 | 14.29 | 10.71 | 3.57 | 1.79 | 10.71 | 0.00 | 28.57 |
| Sleep/Rest/Idle | 0.00 | 0.00 | 0.00 | 0.00 | 3.39 | 0.00 | 0.00 | 3.39 | 0.00 | 6.78 | 57.63 | 5.08 | 11.86 | 0.00 | 11.86 |
| Religious | 0.00 | 0.00 | 33.33 | 0.00 | 0.00 | 0.00 | 0.00 | 33.33 | 0.00 | 33.33 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Work/Teaching | 0.71 | 1.42 | 4.26 | 2.13 | 1.42 | 0.00 | 5.67 | 12.06 | 3.55 | 36.88 | 4.26 | 6.38 | 6.38 | 0.71 | 14.18 |

The context to this pattern is that it is dominated by the employee group (faculty and staff) and graduate students. There are 14 faculty (36.84% of the group; and, 77.78% of the faculty population in the study) and 10 staff (26.32% of group and 71.43% of the staff population in study). There are also 11 graduate students (28.95% of group; and, 25% of the graduate respondents). Only three undergraduate students are included in this group. The staff did not report any academic activities (attending classes or researching/studying). They are mostly engaged in office work. Faculty also are primarily engaged in teaching and research, while the graduate students have working/teaching and researching/studying activities, along with attending classes to undertake. The preponderance of recorded work/teaching activities therefore correctly overshadows any other activities.

The composition of the group is indicative of an older population, with 29 persons (76.32% of the group) being 30 years old and above. There are also more males (24 persons, which is 63.16% of the group and 40.58% of the male respondents).

These are the four activity patterns identified through the activity sequencing method. This approach places the premium on the sequences in which individual activities are carried out by people. This is an important consideration in analyzing human activity interaction.

The sequential alignment technique employed by the ClustalG algorithm is only one of many methods that employ activity codes to calculate a similarity score for activity sequence patterns. However, Wilson (2001) has identified several problems that may influence the results. First, the pairwise scores depend on length of sequences and the sizes of the alphabets or letters that are used to represent activity types. When alphabets of disparate sizes are used, sequences would not align perfectly and the algorithm is unable to correctly identify corresponding sequences of letter characters. A Calibri font type is recommended to mitigate this problem. Second, when a number of alphabets are used to incorporate temporal duration into the alignment measure, scores of activity sequences are inflated at larger temporal resolution. Third, there is hardly any rigorous justification for setting the insertion, deletion and substitution scores for the alignment algorithm.

Though sequences of activities are important, other measures could be employed to identify activity patterns. This study adopts an activity intensity method that engages the sequences in which frequency of activities is undertaken, rather than sequences of activity types. The method (the Daily Activity Intensity Similarity Index - DAISI) also adopts a uniform starting and ending period for all activities, which addresses some of the concerns, especially that of disparate activity sequencing lengths, identified by Wilson (2001). It uses a form of activity profile that summarizes activities by frequency rather than by types to calculate an index score of similarity between pairs of individuals. By using real number of activity frequency rather than alphabets to represent them or activity types as in other methods, the daily activity intensity similarity index circumvents the problem faced by differences in font types and sizes of alphabets as described by Wilson (2001).

The Daily Activity Schedule Fragmentation Index also adopts an activity sequencing approach to calculate the propensity to chain activities or the degree of disruption in an individual’s activity schedule given an anchor location for organizing the itinerary. The space-time path of time geography is employed as the structure from which the index is calculated. It may be important to note that one of the most striking components of the time-geography framework is the capacity to view the sequences in which activities are undertaken, which puts it apart from traditional time study approaches. The space-time path is one major instrument that captures this sequence elegantly, and therefore chains activities in the order in which they occurred. This makes it possible for the fragmentation index to be calibrated as described in the next chapter.

CHAPTER VI

DAILY ACTIVITY SCHEDULE FRAGMENTATION INDEX (DASFI): A SPACE-TIME PATH-BASED FRAGMENTATION INDEX FOR EXPLORING ACTIVITY CHAINS

## 6.1 Introduction

This chapter deals with the basis, development, formulation and testing of the Daily Activity Schedule Fragmentation Index (DASFI). This index measures the propensity of an individual to chain or fragment their activities. The index is based on identifying an anchor location in the space-time path of an individual, which is generated from their activity data.

## 6.2 The Basis for a Daily Activity Schedule Fragmentation Index

Conventionally, travel demand models consider only trips to a single destination for the purpose of carrying out a single activity (such as a trip to work or shop, etc.) (Spiess, 1996). However, people typically undertake more multiple-stop trips than single trips and in the process create a chain of trips (Kitamura, 1988; Axhausen and Hertz, 1989). Trip chains reveal how people organize and arrange their trips to fulfill their activity participation needs. Though several models to analyze trip chains based on sophisticated statistical and mathematical algorithms have been proposed (e.g., Golob, 1997; Fellendorf et al., 2000; Abdelghany and Mahmassani, 2003), there has been a shift to models based on activity chains, which account for both temporal and spatial constraints (Carpenter and Jones, 1983). Instead of linking trip chains, an individual's activities are connected into chains because trips are typically derived from the needs by people to engage in activities at different locations. According to Ellegard and Svedin (2012), the basic principles of activity-based travel demand models and analysis are rooted in Hagerstrand's (1970) time-geography framework.

Hagerstrand's (1970) time geography provides an elegant framework to represent spatio-temporal characteristics of individual activities with an integrated three-dimensional (3D) space-time system. One of the key concepts in the framework is a space-time path, which is a trajectory of an individual's movement in physical space over time. The space-time path provides information on the spatial and temporal characteristics of an individual's activities, including time, location, temporal extents and the sequences of activities (Yu, 2006). It effectively represents human activity patterns by a set of continuous segments of vertical (representing time spent at particular locations for certain activities) and non-vertical (tilted) line segments (representing movements between activity locations) in a 3D space-time system. These line segments are linked together in the sequence in which the activities took place, forming chains of activities. Activity chaining is a phenomenon that exists but is rarely investigated, probably due to the difficulty in extracting the information from diary surveys, among others (Primerano et al., 2008). This is a challenge that this section seeks to address.

Measuring how individuals organize their activity chains in their daily lives presents an interesting yet challenging research question. Different people are affected by different constraints and therefore organize their daily activity chains different from others. For example, some individuals may fragment their daily activity schedules into many shorter activity chains while others organize their activities into fewer longer chains. Among all activity locations involved in a person's daily life, certain locations are of greater importance because they function as anchor locations where people can make a stop, reorganize, and get ready for their future activities. Generally, the home and workplace locations are often regarded as anchor locations (McGuckin and Nakamoto, 2004). However, other places where people stop by frequently or stay for longer duration can also function as important destination points. These most frequented locations or activity points where a person spends the most time become the individual's anchor location, which breaks down a person's daily schedule into fragments. Finding the spatial and temporal distribution of activities associated with these anchor locations plays a key role in defining activity chains and assessing the magnitude of fragmentation in a person's daily activity schedule.

Time geography's space-time path concept allows effective depiction of the organization of individual activities as process and integrated examination of both spatial and temporal characteristics of human activity behavior. Therefore, in this study, a space-time path-based fragmentation index is proposed to measure the magnitude of how individuals organize their daily activities into activity chains based on anchor location.

### 6.3 The Concept of Activity Fragmentation

The concept of fragmentation is used in many disciplines and several measures of fragmentation have been developed. These include the study of temporal fragmentation of leisure activities in relation to time pressure in Sociology; the development of specialized zones of infrastructural fragmentation, and relocation of activities in Economic Geography and Spatial Planning; and hard disk fragmentation in Computer Sciences, etc. (Hubers et al., 2008). Other applications include land fragmentation in agricultural related studies (Todorova and Lulcheva, 2005), fragmentation of political parties (Caulier, 2010), and fragmentation of raw muscles in animals (Calkins and Davis, 1980). More recently, there has been an increasing interest in how the concept of fragmentation applies to human activities in space and time. Currently, most of the studies have focused on how Information and Communication Technology (ICT) influences or encourages the fragmentation of human activities both temporally and spatially (Couclelis, 2000; Lenz and Nobis, 2007; Alexander et al., 2010).

Studies have suggested that ICT has blurred the spatial and temporal boundaries between workplaces and other spaces such that many activities that were traditionally tied down to particular locations can now be fragmented and undertaken in non-traditional spaces, which opens up new opportunities for individuals to multitask (Couclelis, 2000; Kwan and Weber, 2003; Lenz and Nobis, 2007; Hubers et al., 2008). Though first articulated by Couclelis (2000, 2004), one of the earliest empirical support for ICT-induced activity fragmentation was the study by Lenz and Nobis (2007). Since then several other studies have developed measures of activity fragmentation to examine and explain the influence of ICT on human activity patterns (Hubers et al., 2008; Alexander et al., 2011). Irrespective of discipline and definitional particularisms associated with them, there is an underlying understanding of fragmentation as a disintegration process (Couclelis, 2004).

The shift from the traditional trip-based models to activity-based models implies that activity chains rather than trip chains are now the new frame of reference in activity analysis (Axhausen and Herz, 1989). A (activity) chain is the combination of all (activities) performed by an individual starting from a given base (usually a home or workplace) until the considered individual returns to this base (McNally, 2000; Hammadou et al., 2004). This base (home or workplace) is traditionally referred to as the anchor location. Hagerstrand's space-time path (STP) provides a framework that weaves activities and activity locations in a seamless sequence at which they occurred, and allows for identification of anchor locations and activity chains within the daily activity schedules of individuals. Figure 6.1 presents the structure of the space-time path

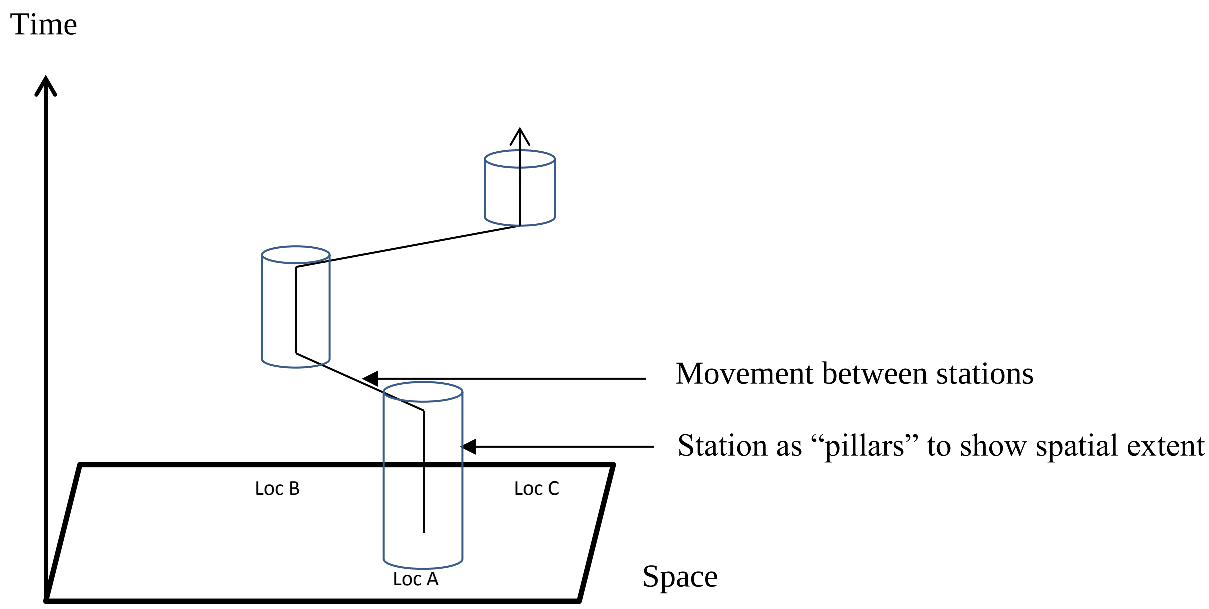



Figure 6.1: Space-time path

This study utilizes this strength to identify anchor locations, and derive and measure activity chains by developing a fragmentation index of activity schedules based on the space-time path.

### 6.4 Development of a Space-Time Path-Based Fragmentation Index

An anchor location is central to the integration of the space-time path. Because the space-time path is a trajectory connecting activity locations, removing movement segments to the anchor location introduces a disruption in the continuity of the path, thereby creating fragments in the space-time path. The degree of connectivity of activity locations among the resultant fragments from the disintegrated space-time path becomes central to measuring activity chains. The daily activity schedule fragmentation index (DASFI) applies a formula that relates the fragmented component of the space-time path to the average degree of connectivity among the fragmented sections from the space-time path.

DASFI is calculated as (where v > 1):

$$DASFI = \frac{\mathrm{y} - \frac{\mathrm{d}}{\mathrm{vx}}}{e}$$

Where: y = number of movements used to connect to the anchor location

d = number of movements not connected to the anchor location

v = number of vertical segments in the space-time path (representing activity locations)

x = number of fragments;

e = number of movements between activity locations in the space-time path.

There are three basic components to the formula. The first component, y, consists of the number of movements used to connect activity locations to the anchor location. These segments hold the structure of the space-time path into a coherent whole and are therefore the key to the disintegration of the space-time path. The movement segments in the space-time path are therefore of great significance in the index because they are responsible for the connectivity of the vertical segments (space-time stations or activity locations) into one coherent whole framework or structure.

The natural form of connectivity in the space-time path is expressed as: v = e + 1. Movement is a derived demand of the desire to engage in activities at different locations. Consequently, each movement segment has an origin and a destination, both of which are particular activity locations. To retain this natural form of connectivity among activity chains, therefore, the space-time path is fragmented by extracting movement segments to and from anchor locations. Removing the activity anchor locations themselves would result in “unnaturally” dangling movement segments where the movement segment has neither an origin nor a destination location.

The second component of the formula deals with the degree of connectivity of the fragments resulting from the disintegration of the space-time path. The measure of connectivity is derived from beta (β) index of network connectivity, which relates the number of edges (movement segments) to the number of vertices (activity location) for each fragment.

Connectivity of one fragment $=\left(\frac{d_j}{v_j}\right)$ (1)

where $d_j$ = the number of movement segments in each space-time path fragment, j;

$v_j$ = the number of vertical segments (activity locations) in each space-time path fragment or activity chain such that $d_j = v_j - 1$ or $v_j = d_j + 1$.

Consequently, $\frac{d_j}{v_j}$ can also be expressed as: $\frac{(v-1)-y}{v} = 1 - \frac{y+1}{v}$. As the activity locations become increasingly chained together, the value of d tends towards zero (0).

A space-time path may be disaggregated into several fragments depending on its structure and the position of the anchor location. The larger the number of locations in the space-time path that constitute the anchor location, the more fragmented the space-time path becomes when disintegrated. A connectivity index for all fragments therefore is the cumulative of all movement segments against the aggregate of all vertical segments (activity locations) thus:

$$\text{Connectivity for all fragments} = \left(\frac{\sum d_j}{\sum v_j}\right) \qquad (2)$$

The average connectivity for all the fragments is therefore calculated as follows:

$$\left(\frac{1}{x}\right)\left(\frac{\sum d_j}{\sum v_j}\right) = \left(\frac{d}{xv}\right) \qquad (3)$$

Where $\left(\frac{1}{x}\right)$ is an averaging factor, with x = number of fragments.

The third component, e, is the number of movement segments in the entire space-time path, consisting of both movement segments directly linked to the anchor location (y) and those that are not linked to the anchor location (d) such that e = y + d. This is used as a normalizing factor to maintain the index values between 0 and 1.

The fragmentation index has values between 0 and 1, which can be expressed as percentages. The higher the values, the more fragmented the space-time path. Lower index values indicate minimum fragmentation. An index value of 1 (100%) indicates maximum fragmentation, a situation in which all other activity locations are directly linked to the anchor location. A value of zero (0) will indicate complete chaining of activity schedules with no fragmentation. Generally, this index (0) value is not common as daily activity schedules would involve two or more activity locations. However, a situation where an individual spends the entire activity space within a single location may be envisaged, e.g., a bed-ridden invalid's activity schedule may consist of strings of activities undertaken at, technically, a single location. In such scenarios the fragmentation index equals zero by default.

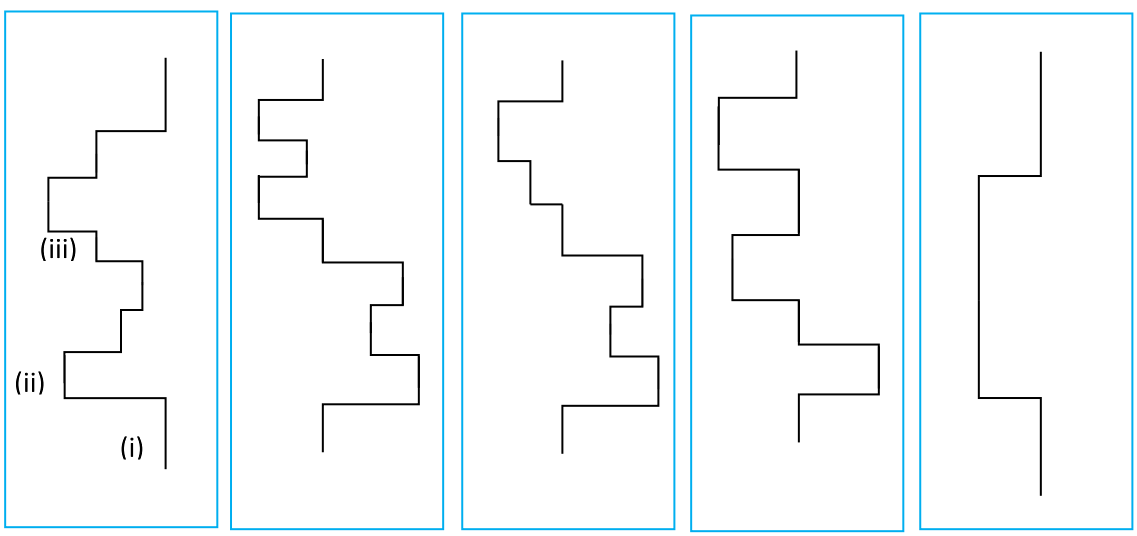
The zero-fragmentation index scenario also raises the specter of granularity in data collection, and temporal and spatial measurements of the index. There are several ways to collect activity data. This includes location-based technologies and activity diaries. The data collected could be open-ended or closed-ended (with activity choices provided). The degree to which data is aggregated or disaggregated, before or after data collection, may determine the groupings and representations of activities in a space-time path.

Data granularity, however, has few direct effects on the fragmentation index, since it is measured only after the data has been transformed into space-time paths. The temporal and spatial granularities also do not present much of a problem if consistency is maintained. For example, the index could be derived for space-time paths that represent decades of migration across countries or cities or for minutes of agitated pacing on a balcony. Generally, granularity affects the process of generating space-time paths therefore it may influence the results of the index only indirectly as it relates to alternative choices made in the construction of the space-time path. Consequently, the fragmentation index of daily activity schedules would reflect the final product of the decisions made in selection of the data, and temporal and spatial resolutions that produced the space-time paths in the first place, as it would affect any discussions and analysis attached to the paths as well.

### 6.5 Measuring the Daily Activity Schedule Fragmentation Index (DASFI)

For convenience and demonstration purposes, this section adopts the frequency-of-visits principle therefore the activity location most frequented by the individual during the study period is assumed as the anchor location. Realistically a space-time path may consist of more than one competing anchor location. In such cases, the location with the largest temporal duration (accumulated time spent at location) among the competing locations is recognized as the anchor location. Similarly, a tie between two locations with same activity duration may be broken by assuming the locations most frequented as the anchor location.

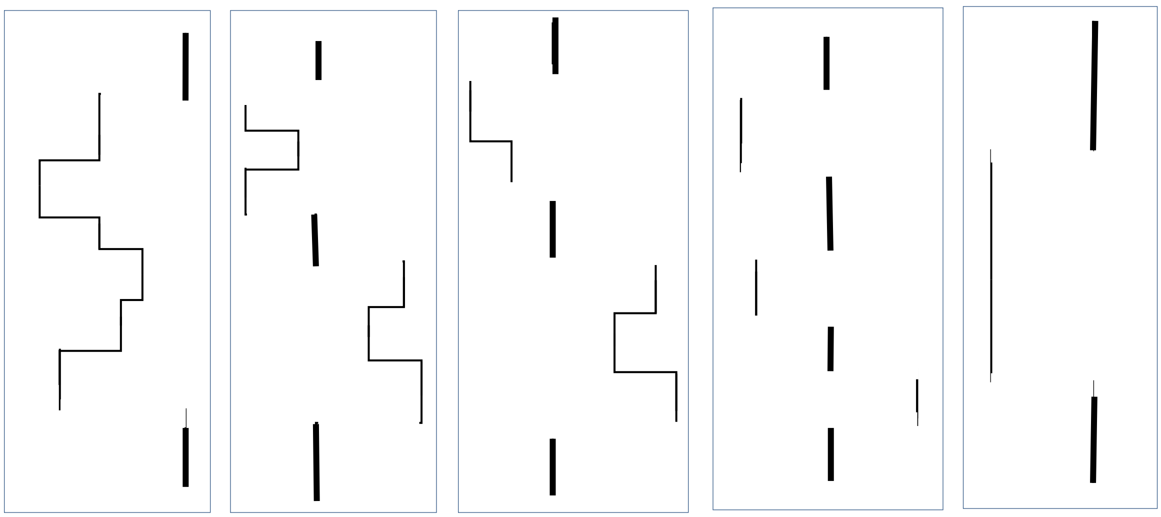
Figures 6.2a, b, c, d and e are examples of space-time paths representing the activity patterns of five individuals. The anchor location of each space-time path is determined by the number of visits the individual paid to the location. The anchor locations for Figures 6.2b, c, d and e are easily discernible because there is only one location in each space-time path that qualifies. For Figure 6.2a, however, three locations were visited twice each. In this situation the location with the largest cumulative duration is selected as the anchor location, which in this case is location (i).




a b c d e

Figure 6.2: Space-time paths showing activity patterns of five individuals

When the movement segments connecting other locations to the anchor location (y) are extracted, a fragmented space-time path is created. Figure 6.3a, b, c, d and e show the fragmented space-time paths from Figures 6.2a, b, c, d and e, respectively and their anchor locations (represented as thick vertical lines).



a b c d e

Figure 6.3: Fragmented space-time paths showing activity chains

Figures 6.3a, b, c, d and e have 3, 5, 5, 7 and 3 fragments, respectively, from the decomposition of their space-time paths. Figure 6.3a has an activity chain in the second fragment. Fragments 2 and 4 of both Figures 6.3b and c have activity chains at different levels of connectivity. Figures 6.3d and e do not have any activity chains because no vertical segment (activity location) is connected to another by a movement segment. These two (Figures 6.3d and e) represent maximal fragmentation, a situation where the individual apparently returns to the anchor location after undertaking activities at any other location.

The fragments for Figures 6.3a, b and c are at different levels of connectivity. For instance, Figure 6.3a has three fragments from the decomposition of its space-time path. The second fragment is a continuous connection of six intermediate activity locations between the vertical segments representing the anchor location. The individual therefore projects a propensity towards activity chaining. As indicated earlier, as the number of locations chained together increases, the value of the index tends towards zero. The daily activity schedule fragmentation index of Figure 6.3a is therefore expected to be a low value, tending towards zero.

Both Figures 6.3b and c have a total number of 5 fragments each, with 3 of the fragments constituting the anchor location. However, they differ in the size of fragment number 4 (starting from the bottom). In Figure 6.3b fragment 4 (counting starts from the bottom) has three vertical segments and two movement segments, while fragment 4 in Figure 6.3c has only two vertical segments and one movement segment between them. Since all other aspects of the fragments are similar, the degree of connections in fragment 4 should determine the difference in activity chains of these individuals. Since Figure 6.3b is more elaborately chained than Figure 6.3c, it is expected that the value of its fragmentation index would be lower, tending towards greater activity chaining (see Figure 6.4b and c, respectively).

Figures 6.4a, b. c, d and e present the fragmentation index values of each space-time path using the formula:

$$DASFI = \frac{\mathrm{y} - \frac{\mathrm{d}}{\mathrm{vx}}}{e}$$

| | | | | |
|---|---|---|---|---|
| $FI = \frac{2 - \frac{5}{8 * 3}}{7}$ = 0.2559 | $FI = \frac{4 - \frac{4}{9 * 5}}{8}$ = 0.4889 | $FI = \frac{4 - \frac{3}{8 * 5}}{7}$ = 0.5607 | $FI = \frac{6 - \frac{0}{7 * 7}}{6}$ = 1.0000 | $FI = \frac{2 - \frac{0}{3 * 3}}{2}$ = 1.0000 |
| a | b | c | d | e |

Figure 6.4: Fragmentation index values for activity chains from space-time paths.

The DASFI values are consistent with the expectations expressed as described in Section 6.5. A real life example of the fragmentation index is calculated for data collected on the Oklahoma State University, Stillwater, for illustrative purposes.

### 6.5.1 Practical Example with Data Collected on OSU Campus, Stillwater

Using a two-day online activity diary/questionnaire, data were collected from 147 respondents consisting of 18 faculty, 14 staff, 44 graduate students, and 71 undergraduate students at the Oklahoma State University, Stillwater (see Table 5.1). Data gathered include role of respondent at the university (faculty, staff, undergraduate student, or graduate student), age, residential status, gender, and whether the respondent lived on-campus or off-campus. Activity data collected include types of activities, the starting and ending times for each activity, sequence in which activities were undertaken, activity location, mode of travel to activity location and whether the activity was habitual, spontaneous or planned.

Using the data collected a space-time path for each individual is constructed using a modified version of Shaw and Yu’s (2009) Extended Time-Geography Framework Tools

Extension for ArcGIS 9.3, generated in ArcScene 10.0. Figure 6.5 shows all the space-time paths generated from the data collected.

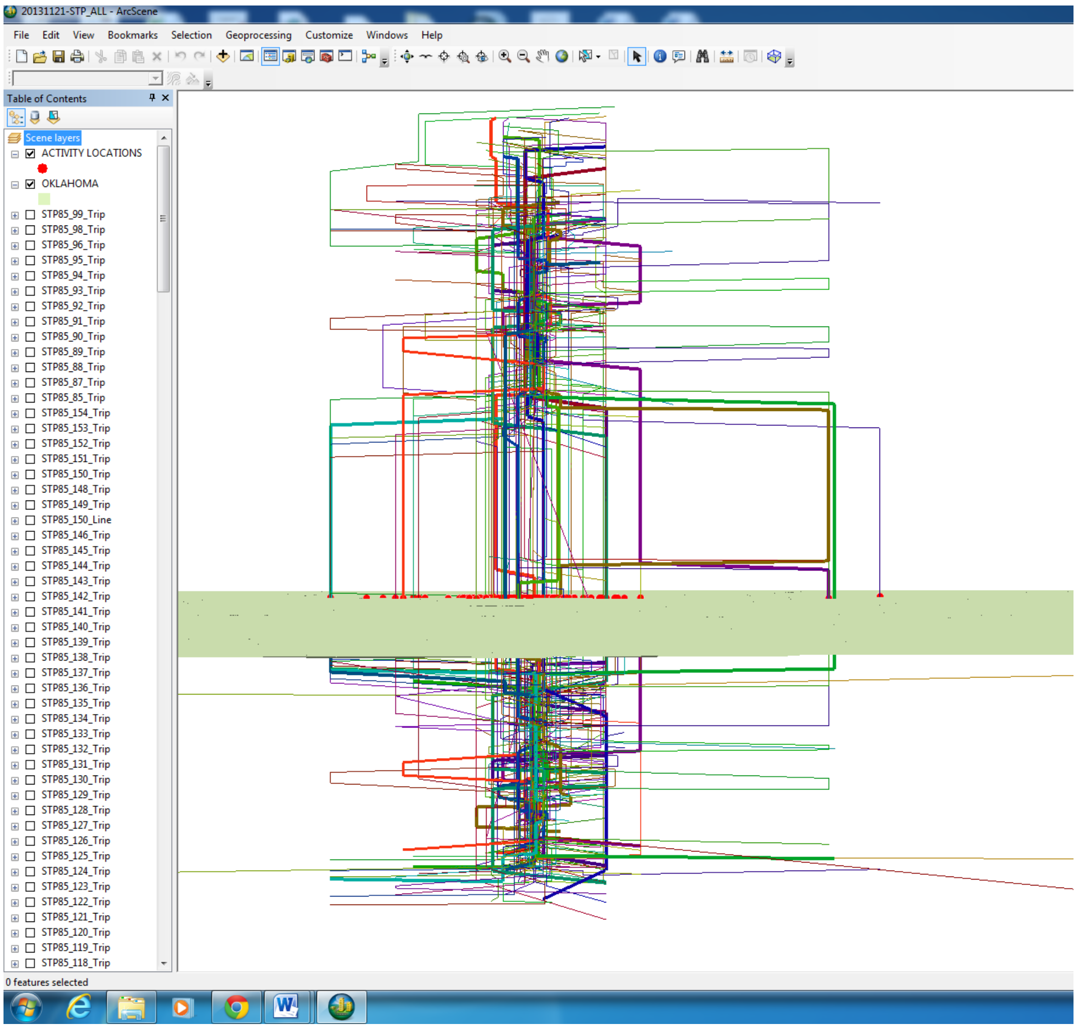



Figure 6.5: Space-time paths of all respondents in and around OSU campus, Stillwater

The space-time paths are then disaggregated along the determined anchor locations as explained and demonstrated in Figures 6.2 and 6.3, and the daily activity schedule fragmentation index (DASFI) formula is applied to measure the activity chaining patterns of respondents as in Figure 6.4.

The fragmentation indices are subjected to a cluster analysis to allow for a comparative analysis of homogeneous (cluster) groups following Lenz and Nobis (2007). First, a hierarchical cluster analysis was performed (Ward-Linkage) to determine the optimal number of cluster groups to use in the subsequent stage. Second, a k-means clustering is then performed using the number of clusters determined by the hierarchical clustering method. The analysis of the fragmentation index is based on the k-means clustering analysis

### 6.5.2 Results of Daily Activity Schedule Fragmentation Index (DASFI)

Figure 6.6 presents a summary of the fragmentation index of the 147 respondents. The index values are plotted using a comparative histogram chart. The index ranges from 0 to 1, where 1 represents maximal fragmentation, a situation where an individual returns to the anchor (base) location after undertaking an activity at any other location; and, values closer to 0 represent high activity chaining tendencies where individuals return to the anchor location probably only after having participated in activities in a string of other locations.

A cursory examination of Figure 6.6 suggests that undergraduate students are better represented at the medium and lower values than staff and faculty. A closer examination indicates that below the 0.450 index values, no staff and faculty members are represented, yet they are conspicuous in the high index values, with a few in the medium value ranges.

#### 6.5.2.1 Fragmentation Index Clusters

The hierarchical clustering method identified 5 groups of fragmentation indices. This set the tone for k-means cluster analysis, using k=5. Below is a synopsis of the character of the clusters created from the k-means cluster analysis:

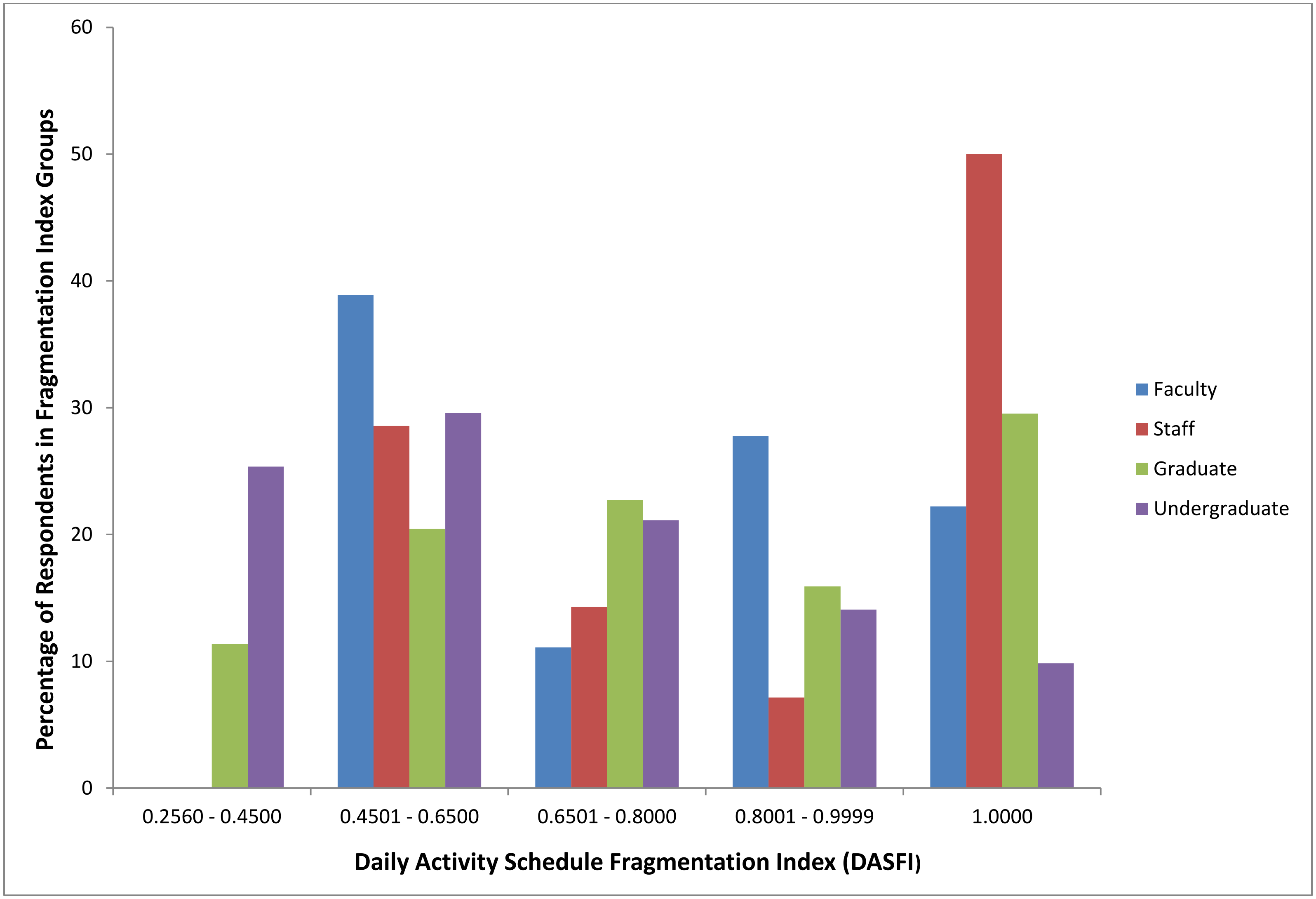



Figure 6.6: DASFI values for respondents based on frequency of visit principle

Cluster 1: High Fragmenters (Index values: 0.908 – 1.000) (*n* = 36)

Cluster 2: Fairly High Fragmenters (Index Values: 0.725 – 0.888) (*n* = 34)

Cluster 3: Moderate Fragmenters (Index Values: 0.561 – 0.713) (*n* = 34)

Cluster 4: Moderate Fragmenters (Index values: 0.433 – 0.540) (*n* = 25)

Cluster 5: Activity Chainers (Index values: 0.256 – 0.399) (*n* = 18)

These five clusters could be summarized into three general groups of high fragmenters, medium fragmenters and activity chainers.

6.5.2.1.1 High Activity Fragmenters (DASFI: 0.725-1.000, n=70)

Two of the five cluster groups (Clusters 1 and 2) consist of 70 individuals (47.6% of total respondents) with a high propensity to return to their anchor locations immediately after an activity and before proceeding to the next activity location. Figure 6.7 presents an example of a high fragmenter's activity pattern.

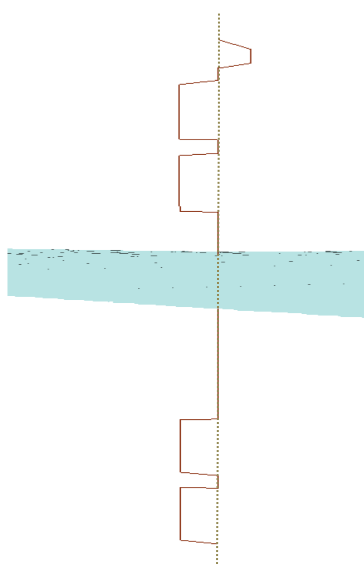


Figure 6.7: High fragmenter's activity pattern (FI = 1.000)

This group of individuals places a high premium on the anchor location as the organizing location of their activity itinerary. Cluster 1 consists of 5 faculty (27.78% of all faculty), 16 graduate students (36.36%), 7 staff (50%) and 8 undergraduates (11.27%). Cluster 2 has 6 faculty (33.33%), 9 graduate students (20.45%), 3 staff (21.43%) and 16 undergraduate students (22.54%). Collectively, these two cluster groups account for 71.43% of staff, 61.11% of faculty, 56.81% of graduates and only 33.81% of undergraduates. This indicates that a larger percentage of staff, faculty and graduate students are high activity fragmenters. Generally, the activity schedules of faculty and graduate students have a similar trajectory. Both have research and teaching (many graduate students do) responsibilities, while the graduate student attends classes (depends on his/her stage of academic development) as well. These activities are conducted in a few select places (offices, labs or classrooms depending on the number of class sections taught).

In terms of time, these three groups (i.e., faculty, staff, and graduate students) have relatively high degree of flexibility. Outside of teaching class schedules, the faculty group has large windows to choose when they will engage in research and any other activities they need to accomplish. Graduate students, likewise, enjoy relatively larger levels of academic flexibility than undergraduates. They take fewer credit hours and may therefore be able to arrange other activities around these few classes, their research and/or teaching responsibilities. The greater degree of flexibility allows this group to choose their activity locations and times with more deliberation and therefore create a fragmented activity schedule pattern. With apparently large windows of spare time between activities, it is easier to select a location to return to for other activities or to wait for the next activity. The staff group has fixed office hours (8:00 am to 5:00 pm) that do not accord them great flexibility, but their activity spaces on campus are also narrow because most of them are centered around their offices, a facility (office) they, along with faculty and graduate students possess, from which they may choose to operate and coordinate their activities. For staff, most activities take place in the offices but any other activity outside of the office is much likely to be succeeded by a return to the office location. This pattern of activity schedule is likely to result into a fragmented structure, and the activity behavior inherent in it is quite predictable in the sense that a high fragmenter is likely to return to the anchor location more frequently than expected of others. This pattern fits the expectations of activity chaining patterns for faculty, staff and graduate students as described in the operational framework.

6.5.2.1.2 Moderate Activity Fragmenters (DASFI: 0.561-0.713, n = 59)

Clusters 3 and 4 consist of 59 persons with moderate fragmentation indices, which suggest that they have several degrees of activity locations strung together away from the anchor location. Figure 6.8 presents an example of a moderate fragmenter's activity pattern

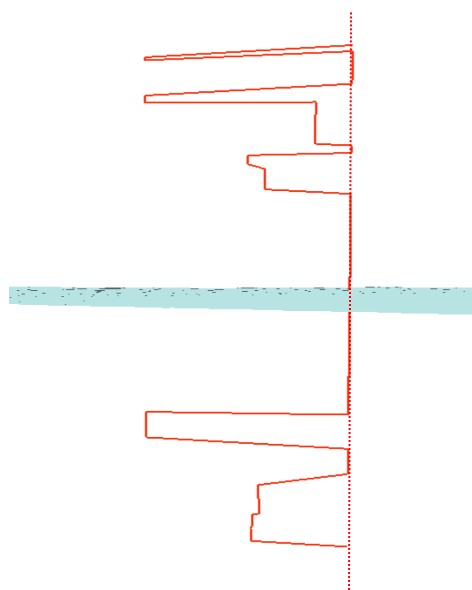


Figure 6.8: Moderate fragmenter's activity pattern (FI = 0.665)

Fragmentation indices for Cluster 3 indicate a more extensive string of activity chaining (less fragmentation) than for Cluster 4. Cluster 3 consists of three faculty (16.67%), 11 graduates (25.00%), three staff (21.43%) and 17 undergraduates (23.94%). Cluster 4 has four faculty (22.22%), four graduate students (9.09%), one staff (7.14%) and 16 undergraduates (22.54%). Together the clusters of moderate fragmenters account for 47.48% of the undergraduate population, which is the largest among the groups. It may be argued that the large number of classes, which undergraduate students take tailor them towards stringing several activities together. Table 5.2 indicates that attending classes are the most frequent subsequent activity to attending classes. This means that one class session is likely to be preceded or succeeded by another class session. Moreover only graduate and undergraduate students reported attending classes as an activity with undergraduates having the higher average number of class sessions. This point may be buttressed by the fact that Cluster 5 is exclusively composed of graduate and undergraduate students only, the two groups that take classes.

6.5.2.1.3 Activity Chainers (DASFI: 0.256-0.540, n = 18)

Cluster 5 is the group of activity chainers. Figure 6.9 presents an example of activity schedule pattern of an activity chainer.

The index values suggest a high degree of activity linkages with little return to anchor locations in between them. Their activity behaviors are also predictable in the sense that they may probably not be returning to their base after an initial activity very frequently. The group is composed of only four graduate and 14 undergraduate students, which comprise 9.09% and 19.72%, respectively.

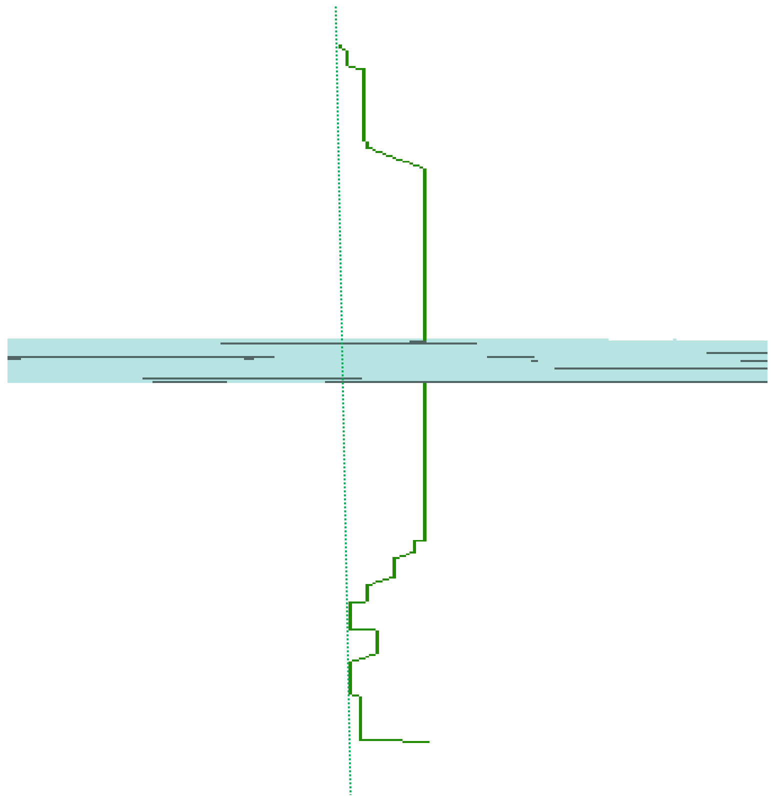


Figure 6.9: Activity chainer's activity pattern (FI = 0.298)

It may be explained that the diversity in the activities and spread of activity locations of graduate and undergraduate students, i.e., classrooms, laboratories, etc., which are scattered across the campus, the sheer number of activities and the limited time to engage them present challenges that only a linking of activities could alleviate.

Table 6.1 shows that the role of the respondent (faculty, staff, graduate and undergraduate), and whether the individual lived on or off-campus were important determinants of the cluster group of the respondent.

Table 6.1: Results of the Pearson Chi-Squared Test

| | Valid Cases | | Pearson Chi-Squared | | |
|---|---|---|---|---|---|
| **Characteristics of Respondents** | **N** | **%** | **Value** | **df** | **Sig.** |
| Respondent's Status | 147 | 100 | 24.306[a] | 12 | .018 |
| Campus | 147 | 100 | 11.359[a] | 4 | .023 |
| Residence | 147 | 100 | 2.813[a] | 8 | .946 |
| Age | 147 | 100 | 36.488[a] | 32 | .268 |

The results presented in Table 6.1 align with the characteristics of members of the respondent groups. All but one staff and faculty each reported living off-campus, a large proportion of graduate students (68.18%) and undergraduate students (56.34%) as well. Consequently, most faculty, staff and graduate students are high fragmenters while most undergraduate students are activity chainers.

## 6.6 DASFI Based on Temporal Duration Principle

The study suggests two principles on which the anchor location could be identified and applied to determine the daily activity schedule fragmentation index. The first approach is the frequency of visit principle as discussed in the preceding sections of this chapter. The second approach is the temporal duration principle, based on the amount of time spent at a given location. The duration-based approach identifies the anchor location as the location where the individual spent the most amount of time over the course of a day or the study period. In the case of a tie between locations, the more frequented among the competing locations is selected as the anchor location.

### 6.6.1 Fragmentation Index Clusters

Using k=5 for the k-means cluster analysis, the five groups are presented as:

Cluster 1: High Fragmenters (Index values: 0.908 – 1.000) (*n* = 30)

Cluster 2: Fairly High Fragmenters (Index Values: 0.713 – 0.888) (*n* = 32)

Cluster 3: Moderate Fragmenters (Index Values: 0.440 – 0.690) (*n* = 31)

Cluster 4: Moderate Fragmenters (Index values: 0.385 – 0.497) (*n* = 33)

Cluster 5: Activity Chainers (Index values: 0.189 – 0.367) (*n* = 21)

These five clusters were also grouped into three: high fragmenters, medium fragmenters and activity chainers as in the frequency-based method. This allows for comparison of the composition of the groups.

#### 6.6.1.1 High Fragmenters (DASFI: 0.713-1.000, n = 62)

This group consists of 62 respondents, composed of 8 faculty (44.4%), 8 staff (57.1%), 22 graduate students (50%), and 24 undergraduate students (33.8%). Together 50% of the employee group (faculty and staff) and 43.4% of the student group (undergraduate and graduate students) are high fragmenters. Generally, this group possesses similar characteristics as the high fragmenters in the frequency-of-visit principle.

#### 6.6.1.2 Medium Fragmenters (DASFI: 0.433-0.690, n = 64)

There are 64 individuals in the medium fragmenters' group. Cluster 3 has 33 respondents and Cluster 4 has 31 respondents. These groups comprise of 8 faculty (44.44%), 5 staff (35.72%), 14 graduate students (31.82%), and 37 undergraduate students (52.11%). Together 13 respondents (40.62%) are of the employee group, while the student group (graduate and undergraduate students) has 51 individuals (44.35%).

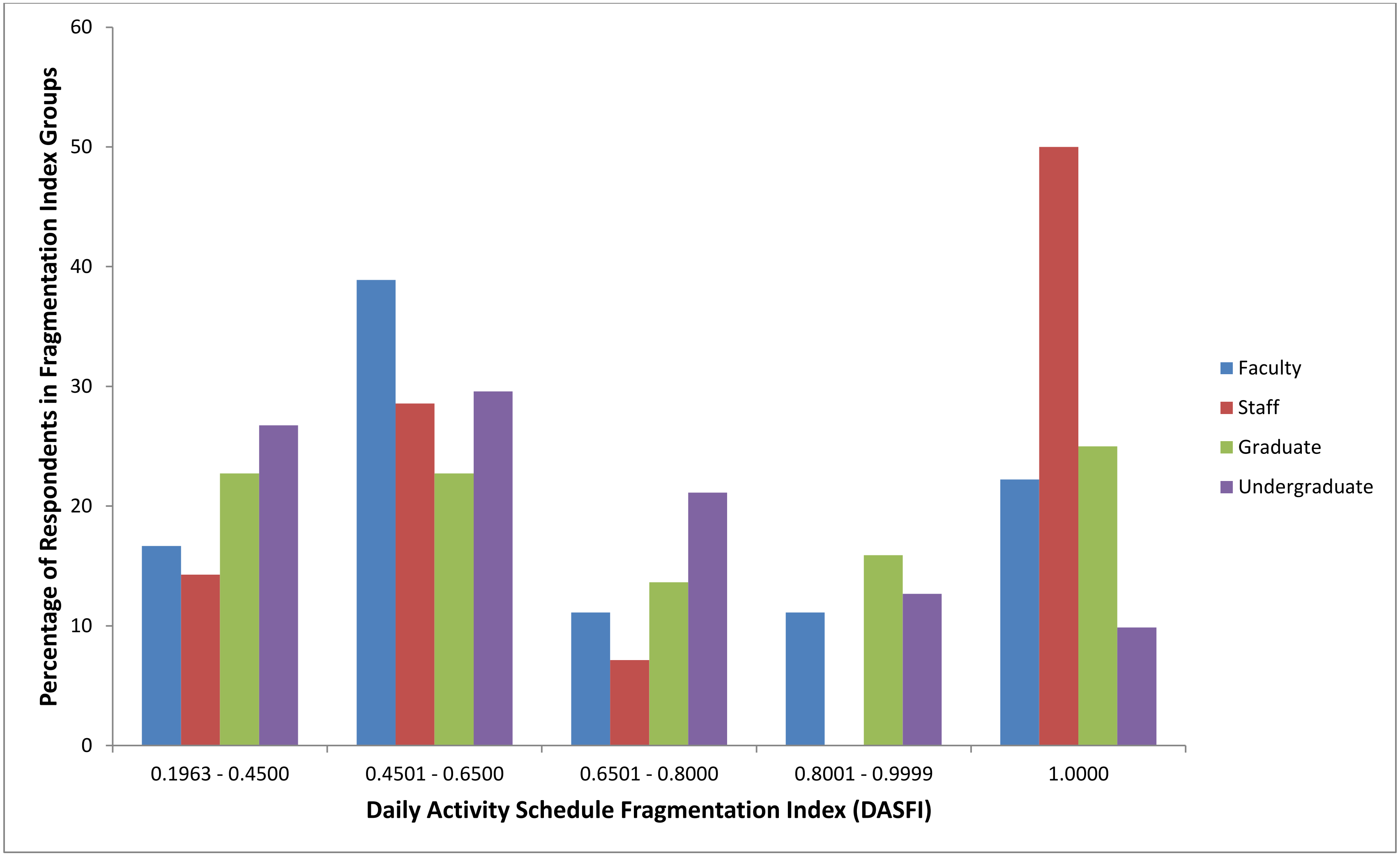



Figure 6.10: DASFI values for respondents based on temporal duration principle

6.6.1.3 Activity Chainers (DASFI: 0.189-0.367, n = 21)

There are 2 faculty (11.11% of faculty), 1 staff (7.14% of staff), 8 graduate students (18.18% of graduate students) and 10 undergraduate students (14.09% of undergraduate students). In total there are only 21 respondents in the activity chainers group, made up of 3 members of the employee group (9.38% of the group) and 18 from the students group (15.65% of the group).

Overall, there are both subtle and obvious differences in the composition of the three groups between the results for the frequency-based and duration-based approaches. Table 6.2 presents a summary of the results.

Table 6.2: Summary of DASFI for Frequency-Based and Duration-Based Approaches

| | **Frequency-Based Index** | | | | **Duration-Based Index** | | | |
|---|---|---|---|---|---|---|---|---|
| Groups | Faculty | Staff | Graduates | Under-graduates | Faculty | Staff | Graduates | Under-graduates |
| High Fragmenters | 11 (61.11%) | 10 (71.43%) | 25 (56.82%) | 24 (33.80%) | 8 (44.4%) | 8 (57.14%) | 22 (50%) | 24 (33.80%) |
| Moderate Fragmenters | 7 (38.89%) | 4 (28.57%) | 15 (24.09%) | 33 (46.48%) | 8 (44.44%) | 5 (35.72%) | 14 (31.82%) | 37 (52.11%) |
| Activity Chainers | 0 (0%) | 0 (0%) | 4 (9.09%) | 14 (19.72% | 2 (11.11%) | 1 (7.14%) | 8 (18.18%) | 10 (14.09%) |
| Total | 18 (100%) | 14 (100%) | 44 (100%) | 71 (100%) | 18 (100%) | 14 (100%) | 44 (100%) | 71 (100%) |

Though the difference in the number of persons in the activity chainers group of both the frequency-based and duration-based approaches is only three, there are more substantive changes underlying this simple difference. There are 11 activity chainers common to both frequency-based and duration-based approaches. However, seven individuals identified as activity chainers using the frequency-based approach were replaced, and four more persons added to the group using the duration-based approach. Between the two approaches, 120 individuals remained within the clusters they were originally classified in the frequency-based approach. These comprise of 11 from the activity chainers, 50 from medium fragmenters and 59 from high fragmenters. A total of 27 (18.37%) individuals therefore switched clusters between the two approaches. Though this may not appear to be a significant number, there are a few significant individual changes in groups.

Most of the changes are between juxtaposed groups, which mean that an individual index changes from one group to another next to it. For example, there were 13 changes between the activity chainer and the medium fragmenter groups, and ten changes between the medium fragmenter and high fragmenter groups. Seven indices classified as activity chainers by the frequency-based approach were grouped under the medium fragmenter group by the duration based approach. All but one (graduate student) are undergraduate students.  On the other hand, six indices from the medium fragmenters cluster are regrouped as activity chainers. These consist of two undergraduate students, two graduate students, one staff and one faculty. From the medium fragmenters in the frequency-based approach, three indices (two graduates and one undergraduate) are reclassified as high fragmenters in the duration-based approach. In the opposite direction, seven high fragmenters (one undergraduate student, two graduate students, two staff and two faculty) are regrouped as medium fragmenters. The significant leaps are between the activity chains and the high fragmenters, the two groups at the end poles of the range of the index. All four changes (four graduate students and one faculty) involved changes from high fragmentation to activity chaining in the frequency-based approach to the duration-based approach. For clarity, the space-time paths of the four individuals are diagrammatized by Figures 6.11 to 6.18.

Respondent 1 (R1) (Graduate, Out-of-State US resident, 40-44 years old, Male, Off-Campus) had Math Sciences building as the anchor location for the frequency-based approach while Richmond Elementary school was the anchor location for the duration-based approach. Figures 6.12a, b and c show the difference between the two approaches for Respondent 1.

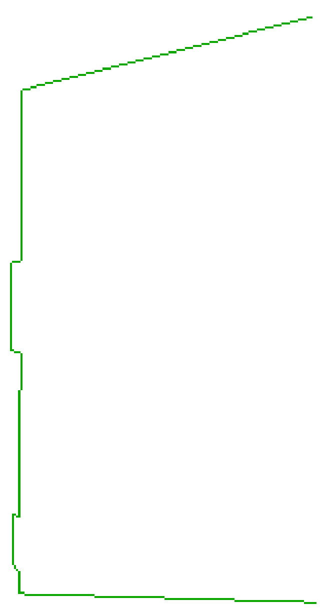
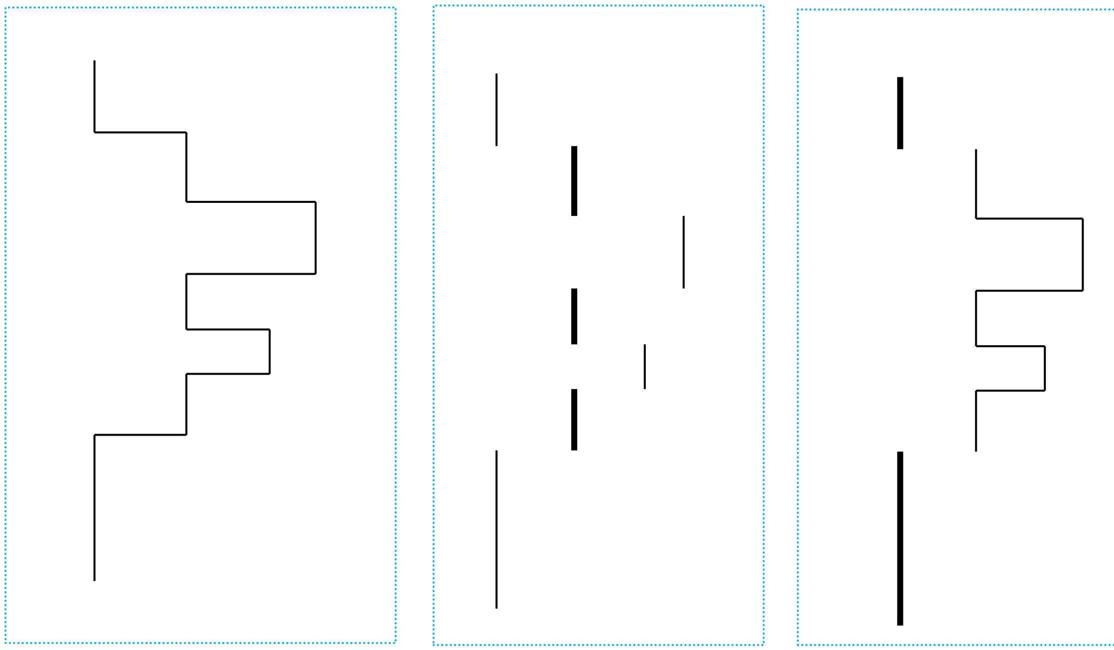


Figure 6.11 Space-time path of respondent 1 (R1)



(b) (c)

$$DASFI = \frac{6 - \frac{0}{7 * 7}}{6}$$

$= 1.0000$

$$DASFI = \frac{2 - \frac{4}{7 * 3}}{6}$$

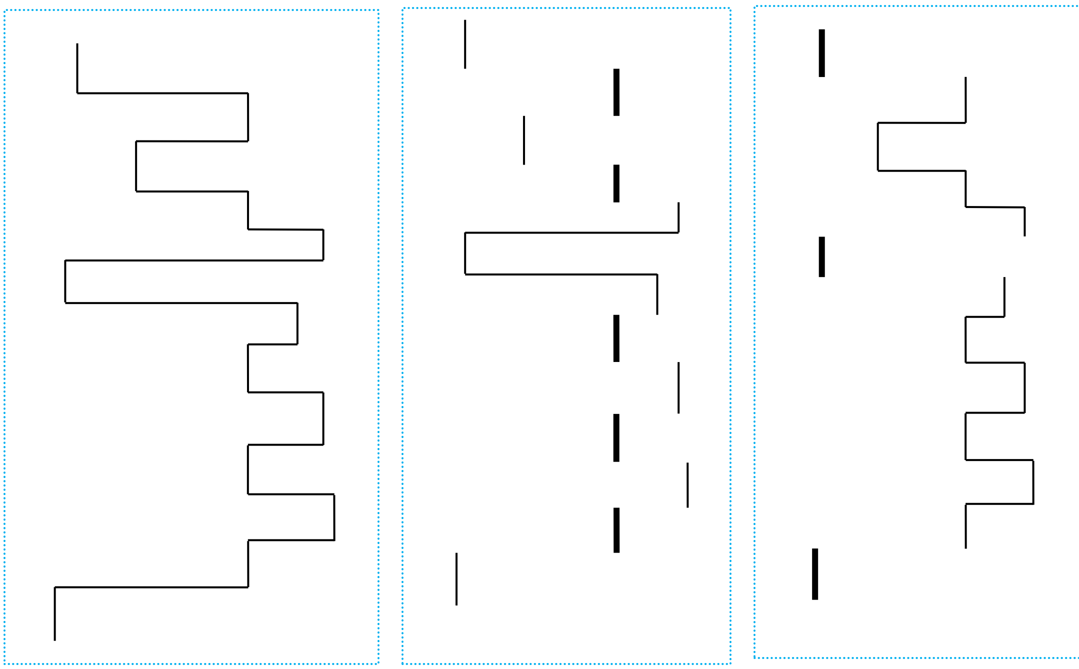
$= 0.30159$

Figures 6.12: (a) Diagrammatized space-time path of R1 (b) Frequency-based (c) Duration-based

Respondent 1 visited the Math Sciences building 3 times in the course of the period and visited Richmond Elementary two times. However, the cumulative duration spent at the Math Sciences building was only 7 hours 29 minutes (449 minutes) while duration of time spent at Richmond Elementary school location was 13 hours 10 minutes (790 minutes). Based on frequency of visit principle, the space-time path disintegrates completely. However, based on duration, there is evidence of an extensive chaining of activities in the space-time path. This accounts for the large differences in index values between the two approaches. It may be pertinent to note that much of the 13 hours spent at Richmond location included sleeping/resting/idle activity (11 hours). If this activity type were not to count, Math Sciences building would be the anchor location for both frequency-based and duration-based approaches.

Respondent 24 (R24) (Graduate, Oklahoman, 35-39 years old, Male, Off-Campus) had Morrill Hall as the anchor location for frequency-based approach and Factories area location as the anchor location for the duration-based approach.

Figure 6.13 The Space-time path of respondent 24 (R24)



(a) (b) (c)

$$DASFI = \frac{10 - \frac{2}{13 * 11}}{12}$$

$= 0.8322$

$$DASFI = \frac{4 - \frac{8}{13 * 5}}{12}$$

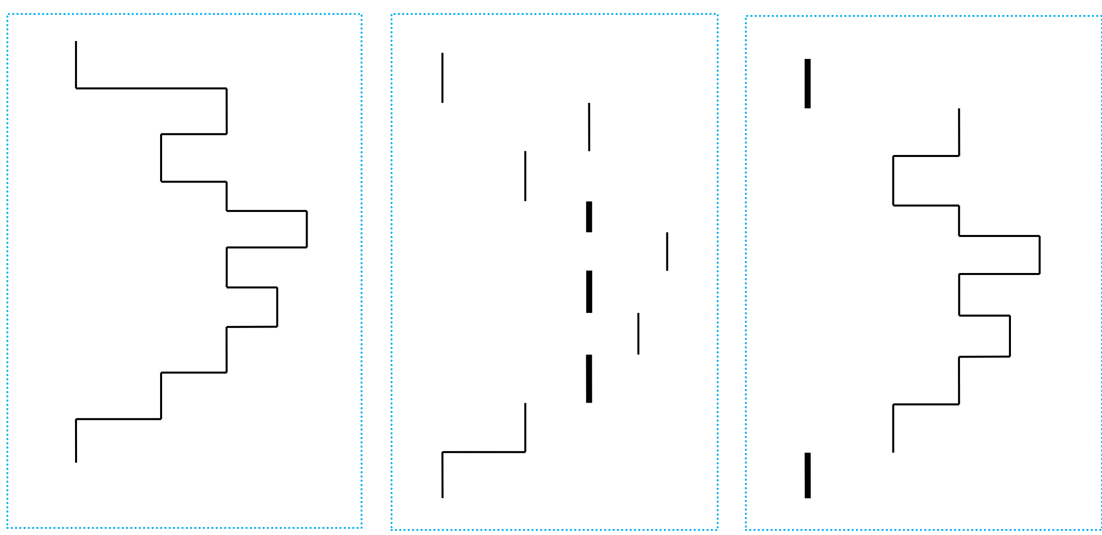
$= 0.3231$

Figures 6.14: (a) Space-time path of R24 (b) Frequency-based (c) Duration-based

R24 visited Morrill Hall five times and had a cumulative duration of 11 hours 45 minutes (705 minutes), while the factories area location was visited three times for an accumulated duration of 19 hours 29 minutes (1169 minutes). About 8 hours 49 minutes (529 minutes) was spent on sleeping/resting/idle activity type.

Respondent 106 (R106) (Faculty, Out-of-State US resident, 30-34 years old, Male, Off-Campus) had Murray Hall as the anchor location for the frequency-based approach and Lakeside Memorial Golf Club location as anchor location for duration-based approach.

Figure 6.15 The Space-time path of respondent 106 (R106)



(a) (b) (c)

$$DASFI = \frac{8 - \frac{1}{10 * 9}}{9}$$

= 0.8877

$$DASFI = \frac{2 - \frac{7}{10 * 3}}{9}$$

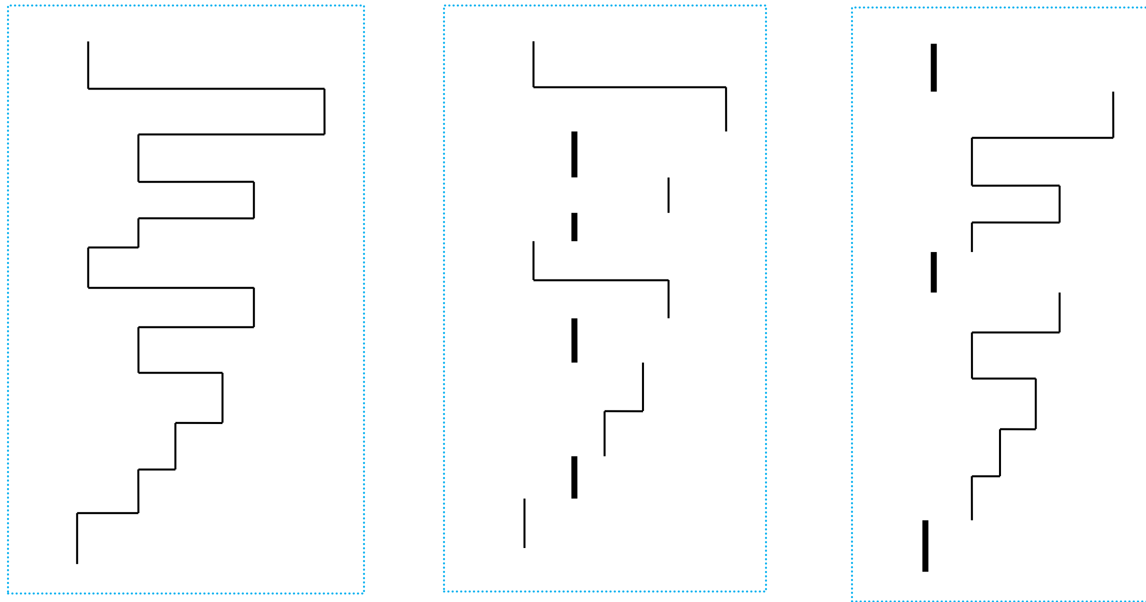
= 0.1963

Figures 6.16: (a) Space-time path of R106 (b) Frequency-based (c) Duration-based

Murray Hall was visited four times, with an accumulated time of 16 hours 15 minutes (975 minutes) spent at the location of, while the Lakeside Memorial Golf Club location was visited only twice with cumulative time duration of 18 hours 44 minutes (1124 minutes), 10 hours 59 minutes (659 minutes) of which were contributed by sleeping/resting/idle activity.

Respondent 126 (R126) (Graduate, Out-of-State US resident, 40-44 years old, Male, Off-Campus) had a more elaborate space-time path.

Figure 6.17 Space-time path of respondent 126 (R126)



(a) (b) (c)

$$DASFI = \frac{8 - \frac{3}{12 * 9}}{11}$$

= 0.7248

$$DASFI = \frac{4 - \frac{7}{12 * 5}}{11}$$

= 0.3530

Figures 6.18: (a) Space-time path of R126 (b) Frequency-based (c) Duration-based

For Respondent 126, the Classroom Building on campus was the anchor location for the frequency-based approach. The location was visited four times within the period, with time duration of 15 hours 36 minutes (936 minutes). Based on this, the space-time path is more fragmented and the index indicates so. For the duration-based approach, the anchor location is the Boomer Lake area. The space-time path also shows evidence of chaining activities. The Boomer Lake location was visited three times, with cumulative time duration of 18 hours 9 minutes (1089 minutes). Sleeping/resting/idle activities contributed 8 hours 44 minutes (524 minutes) of the duration at Boomer Lake area. The balance is less than the time spent at the classroom building.

The four cases presented have several characteristics in common: all of them are 30 years and above, which is higher than the average age of all respondents; they are all male and stay off-campus. Probably the most important of these is the fact that they all live off-campus. All the affected respondents in these cases have their frequency-based anchor locations on campus but their duration-based anchor locations are off-campus (home locations). This fits the expectation that those who live off-campus may have anchor locations on campus, which may be locations where they engage in more activities. The home locations are more unlikely to be anchor locations because of constraints on movement. This is especially true of those who live farther away from campus like in these four cases (Richmond Elementary, Boomer Lake, Lakeside Memorial Golf Club and Factories area of Stillwater).

Interestingly, however, for most respondents a single anchor location qualifies for both the frequency-based and duration-based approaches. In the four cases presented, a single location would have sufficed for both approaches if activities other than sleeping/resting/idle were considered. For example, in cases where the study period extends only between the morning and early evening periods, sleeping/resting/idle activity type would be less significant and pronounced. In such situations it is possible to have both approaches having same anchor locations for individuals.

### 6.7 Characteristics of Anchor Locations

There is a large diversity of activity anchor locations for faculty, staff, and students at Oklahoma State University, Stillwater. These locations are therefore classified into three basic groups of home, office/class and others. Home locations include homes, hostels and dormitories, which are either on-campus or off-campus. Figure 6.19 shows the classes of anchor locations of respondents.

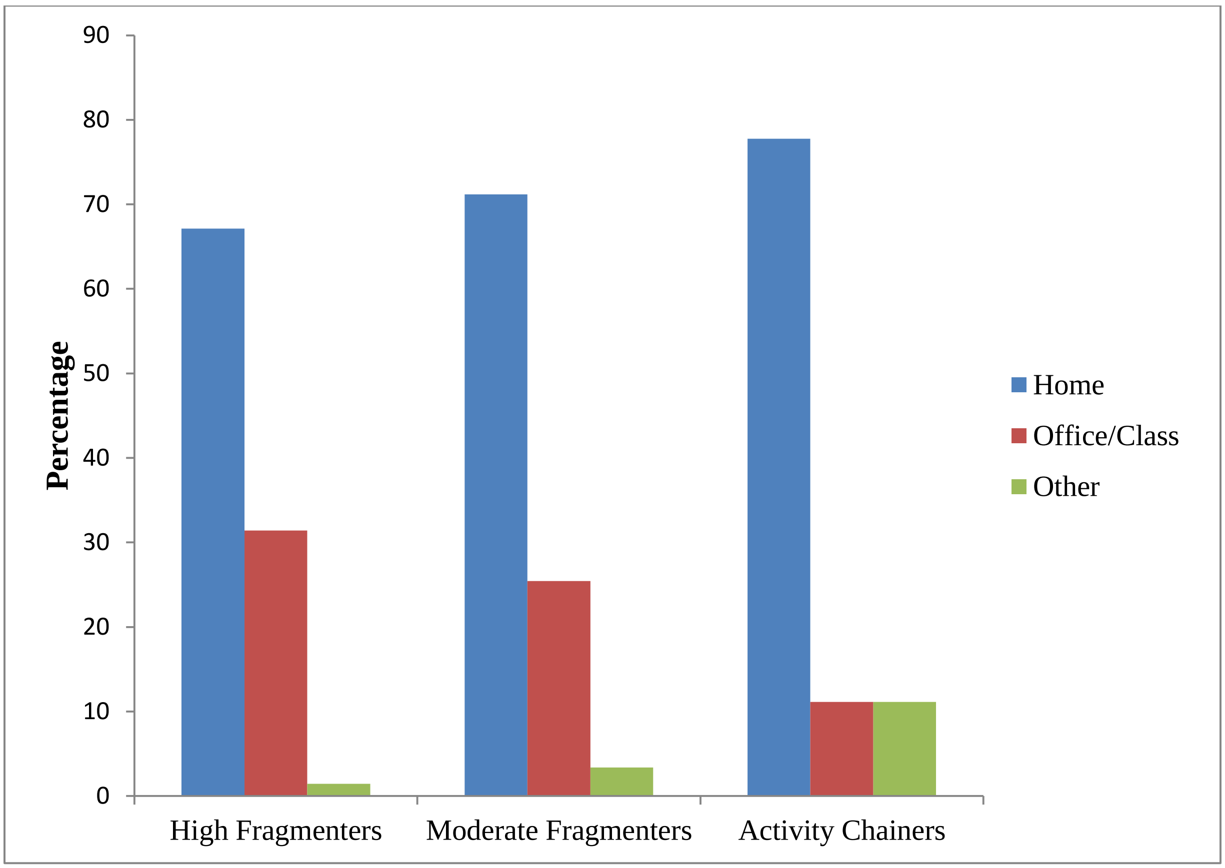



Figure 6.19: Anchor locations of respondents

About 70% of respondents have their homes as the anchor locations. All but one faculty and staff each reported living off-campus while most students live either on-campus or in close proximity to it. For students who live on-campus therefore, their dormitory accommodations are close enough to anchor their activity itinerary. Other anchor locations are offices for staff, faculty and graduate students and classrooms largely for undergraduate students. Library and social gathering locations such as the Students Union building are classified as “other.” Several students reported these as anchor locations. Though the home and workplaces feature highly as activity anchor locations for high and moderate fragmenters, other locations were equally as important as workplace locations for activity chainers. This justifies the redefinition of an anchor location to accommodate other locations other than the traditional home and workplaces alone. In situations where neither home nor workplace information are available, it is easy to appreciate the significance of the redefinition of an anchor location. There may also be many other circumstances where the home and workplace locations would not be as important in the scheme of things as other locations.

Most of the anchor locations also are on-campus. This includes the on-campus housing units for most of the students, the offices for faculty, staff and graduate students, the classroom buildings, library and Student Union buildings. Figure 6.20 provides information on location of anchor locations in respect to whether they are on-campus or off-campus.

This is interesting in the fact that most respondents (68.03%) live off-campus, yet most of the anchor locations for the three groups are on-campus locations.  This speaks to the influence of authority constraints imposed on the respondents by the university calendar and schedules, which ensure that most of their activities for a typical day are conducted on-campus.

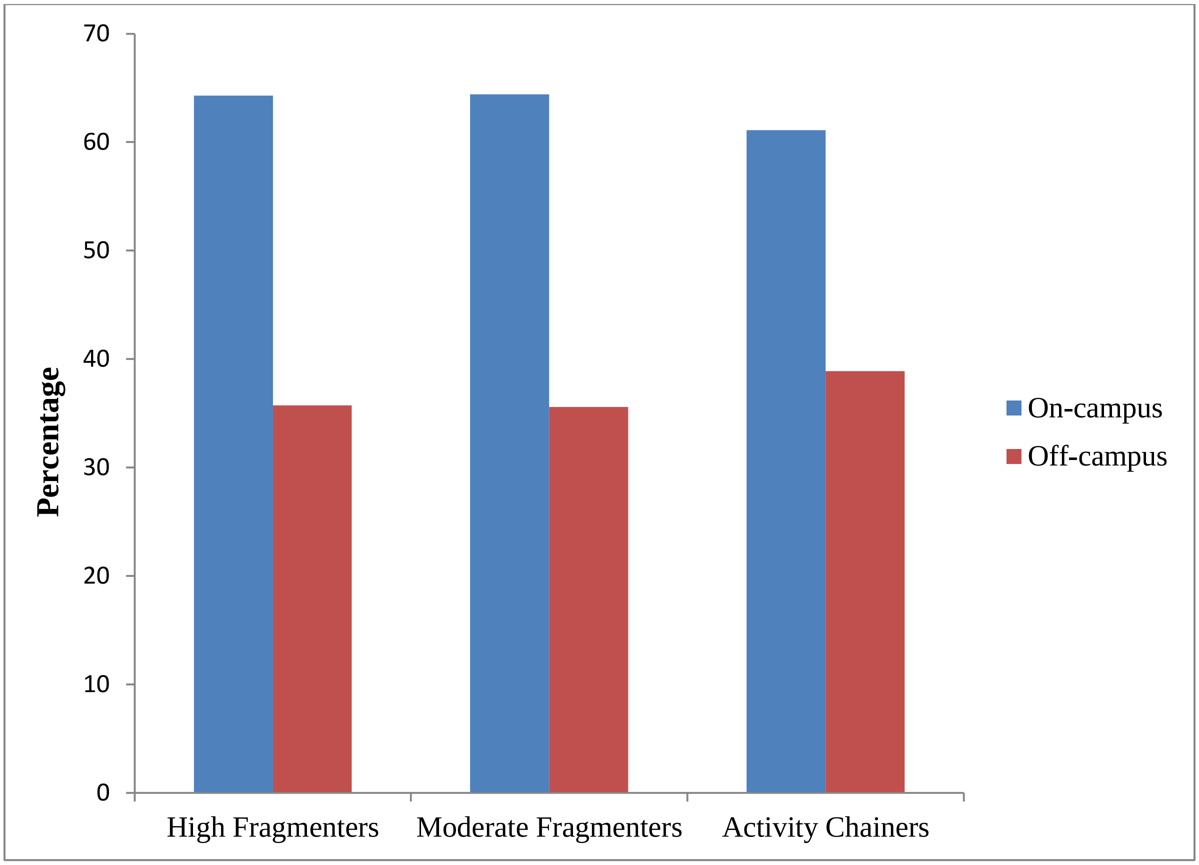



Figure 6.20: On-campus and off-campus locations of anchor points.

It also affirms the theoretical expectation that respondents who live off-campus would require on-campus anchor locations to organize their activities.

A large percentage of activity chainers (about 38%) has their anchor locations off-campus. These may be the group of students with off-campus accommodation and no easy recourse to their living quarters as organizing locations. This may explain the activity chaining propensity of many of them.

### 6.8 Significance and Implication of the Daily Activity Schedule Fragmentation Index (DASFI)

The study has developed and demonstrated a daily activity schedule fragmentation index based on space-time paths to investigate the propensity of individuals to chain their activities. This provides for better understanding of the activity patterns of people. It also allows for better appreciation of the significance of certain locations around which other activities are organized. Identifying the basic characteristics of anchor locations of a group of people may be important to planning appropriate transportation or activity participation facilities and infrastructure. These are potential uses to which the fragmentation index could be applied

The index allows for planning effective activity-conducive environments. There are two principal groups in the index: the activity fragmenters and the activity chainers. The fragmenters are those who have a clear activity organizing location. The anchor location is most often more important to their activity scheduling and pattern than for the activity chainers. Consequently, the primary anchor location (defined as the one most used by people) may be developed for higher efficiency, e.g., comfort, convenience, etc. Determining the characteristics of anchor locations and of the people who use them, and the time periods within which they are most intensely used would be important to plan improvements to the activity spaces for better activity participation experience.

For activity chainers, the anchor location is less important compared to their needs for efficient transition from one activity location to another. To plan effective movement between activity locations, there is need to identify two things: (i) time periods when travel between locations are more intense (most people are transiting from one activity location to another), and (ii) directions of movement between activity locations.

These will allow for the creation of an appropriate environment for better transition between locations. For instance, in a campus, blue (vehicle-free) zones, pedestrian-friendly or bike-friendly zones between locations during transition-intense time periods may be as important as the provision of transit buses between important locations that are farther apart. Appropriate plans would be needed to improve movement between locations.

It is recommended that the index be further developed and expanded into a similarity index to allow for direct pair-wise comparison between activity patterns of individuals with respect to their chaining characteristics. Developing a GIS prototype that include the construction of a space-time path, the capacity to identify an anchor location based on selectable options (e.g., frequency of visit to a location or temporal duration at a location), the ability to deconstruct the space-time path into fragments, and calculate the fragmentation index, organize the report into types of anchor locations, etc. would be an important contribution to exploring the concept of activity fragmentation and activity chains.

CHAPTER VII

DAILY ACTIVITY INTENSITY SIMILARITY INDEX (DAISI): A SEQUENTIAL ALIGNMENT-BASED MEASURE OF ACTIVITY PROFILE

## 7.1 Introduction

This chapter presents the basis, development, formulation and testing of the daily activity intensity similarity index (DAISI). DAISI measures the similarity between and among the activity profiles of individuals. The index is based on the sequence of activity frequency performed by individuals as generated from an activity diary/questionnaire.

## 7.2 The Basis for the Daily Activity Intensity Similarity Index (DAISI)

Traditionally, an activity profile is one of the major indicators of activity patterns and is used to measure the intensity of activity participation of a group of individuals. Kulkarni and McNally (2000) define an activity profile as the proportion of the members of a group that are participating in each specified activity type (home, work, maintenance, and discretionary) at any particular time. In essence, the activity profile shows the number of people engaged in each activity in each hourly time frame during a day (Eom et al., 2009). Determining activity profiles has its advantages. Generally, it provides a good snapshot of a representative activity pattern from which the given activity pattern can be described and visualized (Kulkarni and McNally, 2000). It also facilitates understanding of the daily activity sequence of each group, the type of activities, their starting and ending times, and the sequences in which the activities are undertaken (Alam and Goulias, 1999; Eom et al., 2009). The activity types are then aggregated for predefined groups to examine their patterns of activity participation.

Basically, each activity type is recorded for every hour to create an activity profile. When more than one activity type is undertaken at a given hour frame (multiple activities), the most important activity type, based on temporal duration, is adopted and recorded for the time frame (Eom et al., 2009). The activity types are then aggregated for predefined groups to examine their patterns of activity participation. Several methods of analyzing activity profiles have been used in different studies. These include simple descriptions of the graphs constructed from activity profile data to simulation measures. For example, Kulkarni and McNally (2000) compared simple statistical measures such as mean error (ME), mean absolute error (MAE) and root mean squared error (RMSE) of each group of representative activity pattern (RAP) they had identified against the activity profile simulated for all individuals in the study. They explained that the mean error (ME) allows for determination of any bias in the simulated patterns towards a particular activity type while the mean absolute error (MAE) and root mean squared error (RMSE) provide insight into the accuracy of the patterns. Eom et al. (2009) used the F-Statistic to determine if any variations exist in the activity profiles of groups of students at a university. These include graduate versus undergraduate students, male versus female students, and on-campus versus off-campus students at North Carolina State University (NCSU). Jiang et al. (2012) employed a “star” diagram of clusters of activity patterns and used simple but elegant and detailed graphs to display activity profiles of members of each group. The data was processed into bits of 5-minute intervals instead of the traditional one hourly time frame. In analyzing the activity patterns inherent in the data collected, Jiang et al. (2012) employed a k-means clustering method via principal component analysis (PCA). The PCA had been used to compute eigenvalues that identified important activities (referred to as “eigenactivities”) in the daily activity structure of individuals. The daily activity patterns discerned are then clustered using the k-means method. They validated their clustering effectiveness through the Dunn and Silhouette Indices, which examine the compactness of cluster membership.

Generally, activity profiles reflect the activity intensity of groups. This means that the groups are either predetermined, i.e., generated before their activity profiles are constructed. In many cases, the groups are homogeneous. The descriptions are therefore based on aggregates of individuals within the groups and comparisons are made between or among groups to determine variations in proportions of members undertaking an activity type within a given time frame, usually an hourly time period. The socioeconomic and demographic characteristics of members of the group are then examined to try to explain the patterns. The daily activity intensity similarity index (DAISI) adopts the general principle of the activity profile with several twists of its own.

The activity profile is an established and well-used method of analyzing activity patterns. It provides a glimpse into the degree of involvement of members of a group in certain activity types within given time periods. Its adoption as the basis for the daily activity intensity similarity index is therefore well-founded. DAISI however adopts a more time-geography approach in several ways. First, the activity profiles are constructed for individuals and not groups. In essence, the groups eventually generated are not predetermined.  One of the attractions of activity-based approaches like the time-geography framework is that individuals are the basis for analyzing activities (Thrift, 1977a; Pas, 1990; Peuquet, 1999; Shaw and Yu, 2009). Eventually, groups are created from clustering persons of similar activity participation tendencies. Second, the intensity of activity participation is measured on an hourly time frame as conventional but irrespective of activity type. Consequently the number of all activities undertaken within each hourly frame is recorded for each individual. Third, the sequences of the activities are taken into consideration in computing similarity scores between and among individuals. This is important because it is one of the distinguishing features that set the time-geography framework apart from the traditional time studies approach.

DAISI utilizes the time-geography principles already identified to measure the similarity between the activity intensities of a pair of individuals. A matrix of similarity values are generated and clustered into groups of similar activity profiles.

## 7.3 Requirements for Developing a Similarity Measure

There have been scores of similarity measures developed for different purposes. These range from measures to find similarity between sentences or segments of text (e.g., Metzler et al., 2007; Achananuparp et al., 2008) to similarity in trajectories in spatial networks (e.g., Tiakas et al., 2009), graph or shape matching (Bai and Latecki, 2008), and of geographically dispersed space-time paths (Vanhulsel et al., 2011). Different approaches also have been adopted and applied to determine similarity. Koutra et al. (2011) summarized several algorithms used for graph and subgraph similarity and matching to include edit distance/graph isomorphism, feature extraction, and iterative approaches to substructure index-based approximate graph alignment, tensor analysis, and convex relaxation.

Basically, isomorphism identifies a bijection (a mathematical function that gives an exact pairing of the element of two sets) between the nodes of two graphs, which preserves adjacency. It includes computing edit distance (addition/deletion of nodes and edges in the graph) between the graphs and determining the maximum and/or minimum common subgraph. Iterative methods, on the other hand, define two graphs as similar if their neighborhoods are similar (Zager and Verghese, 2008).

Gunopulos and Das (2000) presented several methods of time series similarity measures to include Euclidean similarity measure, normalization of sequences, dynamic time warping, multidimensional scaling, and indexing techniques such as the variants of the R-trees, kd-trees, vp-trees, and sequential scan. These methods range from the simple to sophisticated algorithms. In a review of 20 similarity indices, Hayek (1994) identified a common notation that underlies several requirements of a similarity index:

$a$ = the number of entries that are common to both lists.

$b$ = the number of entries in the first list that are not in the second.

$c$ = the number of entries in the second list that are not in the first.

$n$ = the maximum number of entries that could occur in either list.

$d$ = the number of entries (of the maximum $n$) that do not appear in either list.

DAISI accounts for two of these lists ($a$ and $n$) explicitly, and adapts the remaining components. Unlike the factors in the indices reviewed by Hayek (1994), activity rates (time frames in this study) are either identical or not identical rather than whether they are present or absent in a list. Consequently, notations *b* and *c* are not available options in DAISI. However, they are accommodated by calculating the differences between activity rates for pair of activity profile lists. Component *d* is based on component *n*. Component *a*, is defined by DAISI as the activity rates of activity profile time frames that are identical. Initially however, time frames are composed of activity frequencies, not activity rates. Each time frame contains the number of activities undertaken by an individual in a time frame. A maximum number of activities (*n*) that can be undertaken and therefore recorded in a time frame is assumed for time frames. Using this maximum number of activities (*n*), the activity frequency is converted into activity rates by dividing the number of activities undertaken by an individual within an hourly frame by the maximum number of activities (n) that could be engaged in within an hourly frame. Between a pair of activity profile lists, the differences in activity rates will be contributed by those time frames that are not identical. Component *d* is therefore measured as the average value of the differences in activity rates between the two activity profile lists. DAISI, therefore, accounts for these notations as much as they can be accommodated within the framework of the type of data being used.

Generally, similarity measures are expected to fulfill certain broad conditions, which are referred to as requirements. Tulloss (1997:125-126) summarizes these requirements into eight and are reproduced as follows:

"REQUIREMENT 1: *A similarity index shall be sensitive to the relative size of the two lists to be compared; and great difference in size shall be interpreted to reduce the value of the similarity index."*

“REQUIREMENT 2: *A similarity index shall be sensitive to the size of the sublist shared by a pair of lists; and an increase in difference in size between the smaller of the two lists and the sublist of common entries shall be interpreted to reduce the value of the similarity index*.”

“REQUIREMENT 3: *A similarity index shall be sensitive to the percentage of entries in the larger list that are in common between the lists and to the percentage of entries in the smaller list that are in common between the two lists and shall increase as these two percentages increase. For logical completeness, we add the following definition*:

“DEFINITION 1: When two lists to be compared by means of a similarity index are of the same size (cardinality), one shall arbitrarily be selected to be called “the larger.” The remaining list shall be “the smaller.”)”

“REQUIREMENT 4: *A similarity index shall yield values having fixed upper and lower bounds.”*

“REQUIREMENT 5: *A similarity index shall have the property that when two lists are identical, the similarity index for the two lists shall be equal to the upper bound of the index.”*

“REQUIREMENT 6: *A similarity index shall have the property that when two lists have no entries in common, the similarity index for the lists shall be equal to the lower bound of the index.”*

“RECOMMENDATION 1: The upper bound of a similarity index should be one; the lower bound of a similarity index should be zero.”

“REQUIREMENT 7: *Distribution of values of a similarity index between zero and one shall be such that (a) if the size of two input lists is fixed, then the output shall vary roughly directly as the number of entries shared between the lists; and (b) if the smaller list is a subset of the larger* *list, then the value of the similarity index shall vary roughly inversely as the size of the larger list.”*

REQUIREMENT 7 part (a) is a variation of the definition of “linearity” of Hayek (1994). Experience with other similarity indices shows that an additional requirement must be added to the list. It relates to convenience in using a program that implements a similarity index.

“REQUIREMENT 8: *A similarity index program shall check its input data to verify that the following relationships hold:*

$a + b > 0$

$a + c > 0$”

Tulloss (1997) further reviewed several similarity indices in the light of these requirements and found many of them in violation of one or several of them. These indices include the Ochiai/Otsuka, Dice, Jaccard, Kulczynski, Mountford, Sokal and Sneath coefficients, etc. A few of these are reviewed to provide perspectives on these requirements for a similarity index. The Jaccard, Dice and Bray-Curtis indices are measures of similarity used in plant and biological sciences but which have gained currency in several other fields including the social sciences. The Jaccard coefficient defines similarity between sample sets as the size of the intersection of the sample sets divided by the sum of the sample sets.

$$Sj = \frac{a}{a+b+c} \tag{1}$$

Where $a$ = number of specimens in both sample sets

$b$ = number of specimens in the second sample set only

$c$ = number of specimens in the first sample set only

The simplicity of the measure has been a great attraction.

The Dice index is similar to the Jaccard coefficient. It also uses a binary function of either presence or absence of specimens in a sample unit but with an averaging factor for the components that are not identical. The basic formula is presented thus:

$$Sj = \frac{a}{d + \frac{b+c}{2}} \qquad (2)$$

According to Tulloss (1997) the two coefficients (Jaccard and Dice) encounter similar problems. The normalizing factors (i.e., *b* and *c* variables of the index) are insensitive to the difference in size of the two lists, which contravenes *requirement 1*. The values of the Jaccard index may also be understated since it lacks the averaging factor of the Dice index. The index, according to Hayek (1994), is also not linear. This charge of non-linearity in indices has also been leveled against the Sokal and Sneath, Mountford and Kulczynski coefficients. This means that progressive differences in values of factors do not yield commensurate progressive index values. Other problems include a “division by zero” problem in situations where there are no items in other lists (e.g., *b* and *c*). This violates *requirements 4 and 5* and is common in indices where the normalizing factors could result in a value that equals zero (0). An example is the Mountford coefficient.

One of the most popular methods of measuring similarity is the employment of a sequence analysis. Halpin (2007) summarized some of these methods. Hamming distance measures similarity between sequences of equal length by comparing them element by element. However, Hamming distance does not do indels (insertion, deletion, and substitution of elements) and consequently it can only recognize similarity at the same location. Degenne’s method defined sequence similarity as a "function of the vectors of cumulated duration in each state, measured at each time point" (Halpin, 2007:14). Basically, the method measures the angle between the vectors at each time point. Although the sequences in which things are done may be different between two sequences, if they have the same cumulated duration at the end, they will necessarily end at the same point.

Wilson et al. (1999) developed the ClustalG algorithm to analyze sequential events based on insertion, deletion and substitutions (indels) of elements between two sets of events. The basic idea is to compute the distance between the two events, which involves indels operations necessary to make the two strings of events identical. When the distance score (that is the number of operations necessary to make the two strings identical) of the strings is low, it indicates a higher degree of similarity between them.

DAISI is a measure of similarity between strings of activity intensities based on simple principles of sequencing. Like the Hamming index, DAISI does not use "indels" in the strict sense of the principle. However, it calculates and uses the average rate of activity needed per time frame to make the two activity profiles identical as a means to determining similarity between them.

## 7.4 Daily Activity Intensity Similarity Index (DAISI) and the Similarity Index Requirements

The daily activity intensity similarity index (DAISI) requires the lengths of the activity profiles of individuals being compared to be equal. The number of time frames should be the same and the starting and ending time for each time frame should be the same for all individuals. This accounts for *requirement 1*, which needs the index to be sensitive to differences in sizes of the two lists to be compared. By ensuring the equality of the two lists, DAISI complies with this requirement as well as circumvents it in some ways. The reporting of activities using an activity diary presents a problem in which people use arbitrarily different times to report the start and end of their activities. This could complicate computations as identified by Wilson (2001). Equal length of lists mitigates the problem.

By maintaining equality in the activity profile lists of the individuals, *requirement 2* is also accommodated. Since no differences exist between the lengths of the lists, any sublists would have similar trait. Moreover, in the activity profile list, there are no sublists to consider. DAISI applies the proportion of entries of activity frequencies between two individuals as a determinant of their similarity. As the number of common entries between the individuals increases, their similarity also increases, which complies with *requirement 3*. DAISI also applies the average of the differences in activity rates between corresponding hourly time frames of the individuals as a measure of "activity distance" between them. When the activity distance between the activity rates entries of two individuals is small, the similarity between them is larger. These are the two principal components of DAISI.

*Requirement 4* dictates that an index should be bounded by an upper and a lower limit. The values of the index therefore fall within the range of these limits. DAISI has a range between 0 and 1, where zero (0) is the lower limit denoting lack of similarity, and 1 is the upper limit signifying complete similarity. These features of DAISI also satisfy requirements 5 and 6. *Requirement 5* requires that when two lists are completely identical, their similarity index should equal the upper limit; and, for *requirement 6*, when there are no entries in common between the two lists, the value of the similarity index should equal the lower limit. DAISI fulfills all these three requirements.

DAISI has equal lengths (or sizes) for the time frames of individuals' activity profiles as assumed by *requirement 7a*. In such situations, the output is expected to display some linearity, in which case it should vary roughly directly as the number of entries shared between the lists. DAISI attempts to fulfill this requirement to the extent that a progressively larger number of shared entries of activity rates results into a progressively greater level of similarity between individual activity profiles. DAISI, however, does not need to fulfill *requirement 7b*, because there are no subsets involved. All activity profiles are required to be of equal lengths.

There are no zero (0) activity entries in any time frame of the activity profiles as explained in Section 7.5. Because of this any combinations of the components will yield values greater than zero (0). Consequently, the conditions $a + b > 0$ and $b + c > 0$ are met in all situations. This meets *requirement 8*.

DAISI has attempted to fulfill all the requirements for a similarity index as submitted by Tulloss (1997). The proceeding section presents the development of DAISI in line with these requirements.

### 7.5 Developing the Daily Activity Intensity Similarity Index (DAISI)

The daily activity intensity similarity index (DAISI) is composed of two components, which are encapsulated in the simple formula:

$$\text{DAISI} = 1 - (rd * dd) \qquad (3)$$

Where *rd* = ratio of dissimilarity between a pair of individuals

*dd* = ratio of dissimilarity (activity) distance between the pair of individuals

Each of these components (ratios of dissimilarity and dissimilarity distance) involves several steps, which are presented shortly. Before then, the activity profile data need to be processed into a format that would allow the two components to be computed.

The activity profile is basically the number of activities that an individual undertakes within given hourly time frames of the day. DAISI measures the similarity between the activity profiles of two individuals at a time. For several individuals, DAISI measures the similarity between the activity profiles of each individual against all other individuals, one at a time. The final product is therefore a matrix of similarity values, which indicates the degree to which the activity profiles of two individuals are similar.

First, the length of time of the activity profiles is demarcated into time frames such that each activity profile has uniform hourly time frames (e.g., 8:00 am – 8:59 am; 9:00 – 9:59; 10:00 – 11:00). The length of hourly episodes are kept same (the beginning and end hourly time frames for all individuals are identical). All time frames should start at the same time and end at the same time for all persons (e.g., from 8:00am – 12:00 noon; 6:00 am – 12:00 midnight) and therefore ensure equal length of activity profile lists. This addresses some of the concerns expressed by Wilson (2001) where unequal length of sequences poses some problems in deriving accurate results of similarity between pairs of sequences. This also meets *requirement 1* as listed in Section 7.3.

Second, count and enter the number of activities each individual undertook within each hourly time frame. This provides the profile of activity intensity for individuals. The profile represents the activity participation of individuals (not predefined groups), irrespective of activity type. In this sense DAISI measures the intensity of human activity participation and does not measure tendency to particular activity types. Each activity conducted within the hourly period is counted for the particular hour. For example, when a student completes a class session at location A at 9:20 am, walks to location B and starts the next class session at 9:30 am, three activities are recognized for the hour period (2 class sessions and a movement activity). When an activity extends beyond the hour mark demarcated, the activity is recorded for the two hourly periods it traverses. For example, when a class session takes place between 9:30 am and 10:20 am, it would be counted as two different activities, each on one side of the 10:00 am hourly time line and recorded for both the 9:00 am to 9:59 am and 10:00 am to 10:59 am hourly time frames. This solves the problem of “zero” (0) activity within any hourly period, which is rather impractical in human activity experience. The study examined the option of counting an activity for only the hourly period within which it was initiated but discarded it because of the “zero” activity problem. In a three hour lecture, which is common for graduate students and also for faculty, two of those hourly periods may register zero activity if the activity is counted only for the hour within which the class started or even just once for any of the time frames. Same can be said of the work of many staff, which stretches across several hours at a time. Though it would not affect the mathematical formulation of the index, it undermines the practical reality of the human activity experience. To maintain a more real human activity scenario, the index accommodates the redundancy inherent in counting a single activity twice across an arbitrary temporal divide. Since no zero (0) is accommodated requirement 8 is met.

Third, the number of activities recorded in each time frame is normalized by a factor of 60 to convert the activity frequencies into activity rates. This section accommodates notations *b*, *c*, and *d* by Hayek (1994). In a given day, there are 24 hours within which individuals participate in any number of activities. In any given hour, there may be only one or many different activities to undertake. Hagerstrand (1975) lists eight conditions necessary to measure the degree of reality present in a geographic model or theoretical framework. Two of these conditions are important to this step of the process.

The first is the realization that human beings are limited in the number of tasks they can perform at a time; and, secondly that every task has a duration (time limit). Taking these conditions into consideration, it is assumed that individuals can perform only a limited number of activities in any given day and each activity consumes limited time span or duration within the given day. The number of activities an individual can undertake in a given time period is therefore not infinity. According to Wilson (1998a) a typical activity diary contains between 10 and 40 separate activity episodes in a day, where an episode is the period of activity from the beginning to an ending time. An episode is therefore the "coarsest recording interval for activities" and people generally recognize the episode as the building block of daily activities. Sequences of activities can therefore be generated from the episodes, taking into account the beginning and ending periods of each activity episode.

Wilson (1998a:1027) further contends that in a given day, the 1-minute activity episode interval is "probably the shortest practicable sequence element and generates a very fine activity record, 1440 elements per day, while remaining manageable." Allowing therefore for such wide latitude in the number and time duration of activities, it is inconceivable to imagine an individual conducting more than one activity every minute for any appreciable length of time. Consequently, every individual is assumed a capacity to undertake not more than one activity for every minute of the day, which is 1440 activities in a day (Wilson, 1998a; Vanhulsel et al., 2011). This breaks down to 60 activity episodes every hour, which is the maximum number of activities that can be recorded in any time frame. This accounts for *n* in Hayek's (1994) notation.

The number of activities actually conducted by an individual, and recorded in the time frame, is therefore measured against the maximum allowable number of activities that could be conducted in each time frame. This converts the activity frequency reported by respondents to activity rates. This averaging factor (maximum number of activities per time frame) normalizes all the entries into a range of values between 0 and 1, where values tending towards 1 indicate large number of activities in the time frame and values tending towards zero (0) indicate fewer numbers of activities. This normalization is crucial to maintaining the index values within the specified upper bound of 1 and lower bound of zero (0) and helps fulfill *requirements 4, 5 and 6*.

A table of activity profiles represented as activity rates for individuals at different time frames forms the basis for the computation of the index. The procedures to compute the components of the index are presented shortly and a worked example is presented in Appendix II. The calculation of the index (DAISI) is accomplished in three stages.

First, a ratio of dissimilarity (*rd*) is first calculated. This is determined as the number of time frames with non-identical activity participation rate divided by the number of time frames for the study period. Secondly, the absolute differences between the activity rates of corresponding respondents are determined at all the time frame levels. This is referred to as the dissimilarity distance between any pair of activity rates. These are then summed up for the respondents accordingly and the average is computed. At the third stage, the product of the first (*rd*) and second (*dd*) stages is subtracted from 1, to invert the index values from dissimilarity (where values tending towards 1 indicate lesser similarity) into measures of similarity (where values tending towards 1 indicate greater similarity).

7.5.1 Stage 1: The *rd* Component

The *rd* is basically the ratio of dissimilarity between the activity profiles, represented as hourly activity rates, of a pair of individuals. The number of time frames in the study period is counted and represented as *y*. For example, if the study period is between 6:00 am and 12:00 midnight, there are 18 hourly time frames, which means $y = 18$. The notation *y* is therefore the length of the activity profile list. In the example provided (see Appendix II), this is captured in Section A.

Section B (Appendix II) describes the determination of the variable *a*, which is the number of hourly time frames between the activity profiles of two individuals that have different values of activity rates. In Hayek (1994), the notation *a,* is the number of entries which is common to both lists. DAISI however takes on a reversed role for the notation *a*, for an obvious reason. Take three individuals (A, B and C) with no common activity rates between any pair of them. If *rd* is based on common entry of values among them, then the value returned will be zero for each pair (AB, AC, BC). Irrespective of how close the activity rate of one individual (A) is to another (B) compared to the third (C), the eventual similarity index will be zero (0). This would suggest that the three individuals (A, B, C) have same or similar activity rates. For example, it may be argued that a string of values (A) [0.2, 0.3, 0.2] may be more similar to a second string (B) [0.3, 0.4, 0.3] than a third string (C) [0.8, 0.9, 0.7], if they all represent same variables of real numbers as is the case with activity rates. Since there are no common identical entries between them, all three strings would return an index score of zero (0). Technically, this may be correct; however, it renders the second component (*dd*) of the index unnecessary because the value will still be zero, irrespective of any value that may be multiplied by the *rd* component. It is therefore necessary to ameliorate this situation and imbue the index with the capacity to differentiate, even if only slightly, between activity profiles whose differences are large from those whose differences may be smaller even when they do not have any common entries between them. By inverting the role of notation *a*, only individuals with total similarity between them will return a value of zero, which signify exact similarity. This solves the problem and still meets the conditions set out in requirement 1, albeit in a different way. Consequently, *rd* is calculated as the number of dissimilar activity rate entries divided by the number of time frames. For example, if 6 out of the 18 time frames (*y*) have different values of activity rates, then *a* = 6.

*rd* is then calculated as:

$$rd = \frac{a}{y} \tag{4}$$

Where *y* = number of activity time frames

Using the values provided in the example, *rd* is calculated as:

$$rd = \frac{6}{18} = 0.3333 \ (or\ 33.33\%) \tag{5}$$

For an individual's time frame measured against itself, *rd* = 0 because all the time frames will be identical (i.e., if there are 18 time frames, then *y* = 18. If all the time frames have same value entries, then *a* = 0. The value of *rd* will therefore be 0 out of 18, which is 0.) In *a* situation where no time frames have any identical common entries between a pair of activity profiles, *rd* = 1 (i.e., *y* = 18; *a* = 18; therefore *rd* = 18/18 = 1). This fulfills requirement 3 in a similar way as requirement 1 is fulfilled. Consequently, on a scale of 0 to 1, pairs of individuals with *rd* values tending towards zero are more similar than those with *rd* values tending towards 1. Since similarity index expects higher values to represent similarity (Hayek, 1994), the *rd* component actually measures ratio of dissimilarity between individuals.

The next stage is the *dd* component.

### 7.5.2 Stage 2: The *dd* Component

The *dd* component represents the measure of average dissimilarity distance between two individual activity profiles. The *dd* is calculated as:

$$dd = \quad \Sigma \text{ abs}(r_i - r_j) / y \qquad (6)$$

where: *dd* is the measure of dissimilarity, calculated as the average of differences between the activity rates of a pair of individuals

$r_i$ and $r_j$ = activity rates for corresponding time frames for individuals i and j, respectively.

In many sequential alignment methods, the principle of "indels" (insertion, deletion, and substitution) is employed to measure the "distance" between two sets to be compared. Wilson et al. (1999) and Isaacson and Shoval (2007) define this "distance" as the number of operations (which may include insertion, deletion, and/or substitutions) needed to make the two sets identical. DAISI defines "distance" as the average value of the absolute difference between the two activity profiles that is needed per time frame to make the two profiles identical (*dd*). This becomes the equivalent of the "indels" operation for DAISI. Averaging the difference between the activity rates of a pair of individuals has several advantages, including (i) maintaining an index range of 0 to 1 in keeping with requirements 4, 5 and 6; and, (ii) accounting for any subtle differences between activity rates of individuals, which is not accounted for by *rd*.

To calculate *dd*, therefore, absolute differences between corresponding values of time frames are determined. The activity rate in time frame for 8:00 – 8:59 am for individual A is subtracted from the activity rate in time frame for 8:00 - 8:59 am for individual B. Only absolute values are returned. The sequence of the differences is maintained to ensure that corresponding time frames are being compared for the individuals. The average of these differences then becomes the measure of dissimilarity between the activity profiles of the individuals. It represents the average value necessary to make the activity rates of the pair of time frames identical. This is calculated in Section C of Appendix II.

It may be pertinent to point out that the range of values for both *rd* and *dd* are the same (from 0 to 1) and are interpreted similarly. For *rd* values tending towards 1 are indicative of dissimilarity between a pair of individuals. When 4 out of 5 entries in time frames in a pair of activity profiles are identical, the level of dissimilarity is 0.2, which signifies that the pair is more similar than not. For the same range for *dd* (0 to 1), values tending towards 1 are indicative of dissimilarity between a pair of individuals. When the average of the difference in activity rates is 0.2222, it indicates that the activity profiles are more similar than when the value is 0.5555.

At this stage, all the components of the formula for DAISI have been computed. Since DAISI is meant to be a similarity measure where higher values should indicate greater degree of similarity and lower values should signify lesser degree of similarity between a pair of individuals' activity profiles (in keeping with requirements 4, 5 and 6), the product of *rd* and *dd* is subtracted from 1. The range of values is 0 and 1. Since a value of 1 indicates total dissimilarity, subtracting 1 from itself, would invert the range such that values of 0 signify total dissimilarity and value of 1 indicates complete similarity. The similarity in activity profiles (DAISI) of a pair of individuals therefore is calculated as:

$$DAISI = 1 – (rd * dd) = 1 – (0.3333 * 0.7778) = 1 - 0.2592 = 0.7408 \text{ or } 74.08\% \qquad (7)$$

This example indicates that the pair of activity profiles has a 74.08% similarity between them. It is pertinent to point out again that the similarity index is calculated for all pairs of individuals such that each individual, compared against itself returns a perfect score of 1 for complete similarity. A value of zero (0) indicates complete dissimilarity between the pair of activity profiles for the respective individuals. The final product therefore is a symmetrical matrix of similarity scores between pairs of individuals, which is then subjected to a clustering analysis to identify patterns of similarity in activity profiles.

Using the data collected from Oklahoma State University, Stillwater campus, the matrix is first subjected to a hierarchical clustering method to determine the optimal number of clusters. The number of clusters generated, broadly, is employed as an input for k-means. This follows Lenz and Nobis (2007) (see Appendix II).

### 7.5.3 Stage 3: Clustering the Similarity Score

The similarity indices for all individuals are collected into a symmetrical matrix (see Section E of Appendix II), which is then subjected to a clustering method to produce groups of individuals with similar activity profiles. For this study, the k-means clustering method is adopted where k = 5. One of the most important components is the choice of clustering method and probably, the number of clusters.

Hierarchical clustering techniques do not require the choice of the number of clusters. The algorithm chooses the number of clusters intuitively using gaps identified in the data. K-means clustering however requires the choice and input of a number of cluster groups to be created. Several methods have been developed to ensure the selection of an optimal number of cluster groups. These include the use of measures of compactness such as Silhouette and Dunn's indices, and also the employment of the number of clusters generated by a hierarchical clustering method as the optimal number of groups for k-means clustering (Lorenz and Nobis, 2007). The cluster groups are then mined for any variables that may provide insight into the patterns of activity profiles generated from the index.

### 7.6 Practical Example with Data Collected on OSU Campus, Stillwater

Data collected with the activity dairy/questionnaire was processed and used as a proof of concept for identifying activity profile patterns using the daily activity intensity similarity index. First, the respondents reported disparate starting and ending times for their activity schedules. For instance, 26 respondents reported only one day of activities, while 47 reports did not start at 6:00 am on the first day. Additionally, 40 respondents did not complete records of activities for the second day. This means that comparison between individual schedules will have to contend with disparate lengths of activity time frames, which may affect accuracy of results (Wilson, 2001). DAISI requires identical lengths for all schedules; consequently the data is trimmed to accommodate as many responses as possible that could produce identical lengths of schedules. To achieve this, the time frames are set for between 9:00 am and 12:00 midnight of the first day only to maximize the number of respondents that can be included in the dataset. This has led to

143 respondents (Faculty: 17; Staff: 14; Graduate: 44; Undergraduate: 68) being part of this case study.

Between 9:00 am and 12:00 midnight, there are 15 hourly time frames. The first stage of the analysis required that the number of activities conducted by each person within each hour of the 15 hourly time frames be recorded. Each activity conducted within the hourly period is counted for the hour as described in Section 7.4 of this chapter.

The number of activities conducted within each hourly period for each respondent is divided by 60 to produce an activity rate. The activity rate of each respondent is compared to every other respondent's including itself. The end result is a symmetrical matrix of similarity index values that shows how each respondent compares to every other respondent with regards to similarity in activity profiles. A matrix of all similarity indices is then subjected to a hierarchical clustering using IBM SPSS 21 to determine optimal number of clusters. Five broad clusters are determined and applied in k-means (k = 5) clustering to classify individuals of similar activity profiles into five groups. The five clusters of activity profiles are presented in Figure 7.1.

Incidentally, two of the five hardly qualify as clusters. They are largely outliers from the rest of the activity profiles. Clusters 4 and 5 consist of only two and one activity profiles, respectively.

The characteristics of the five clusters are examined and discussed as follows:

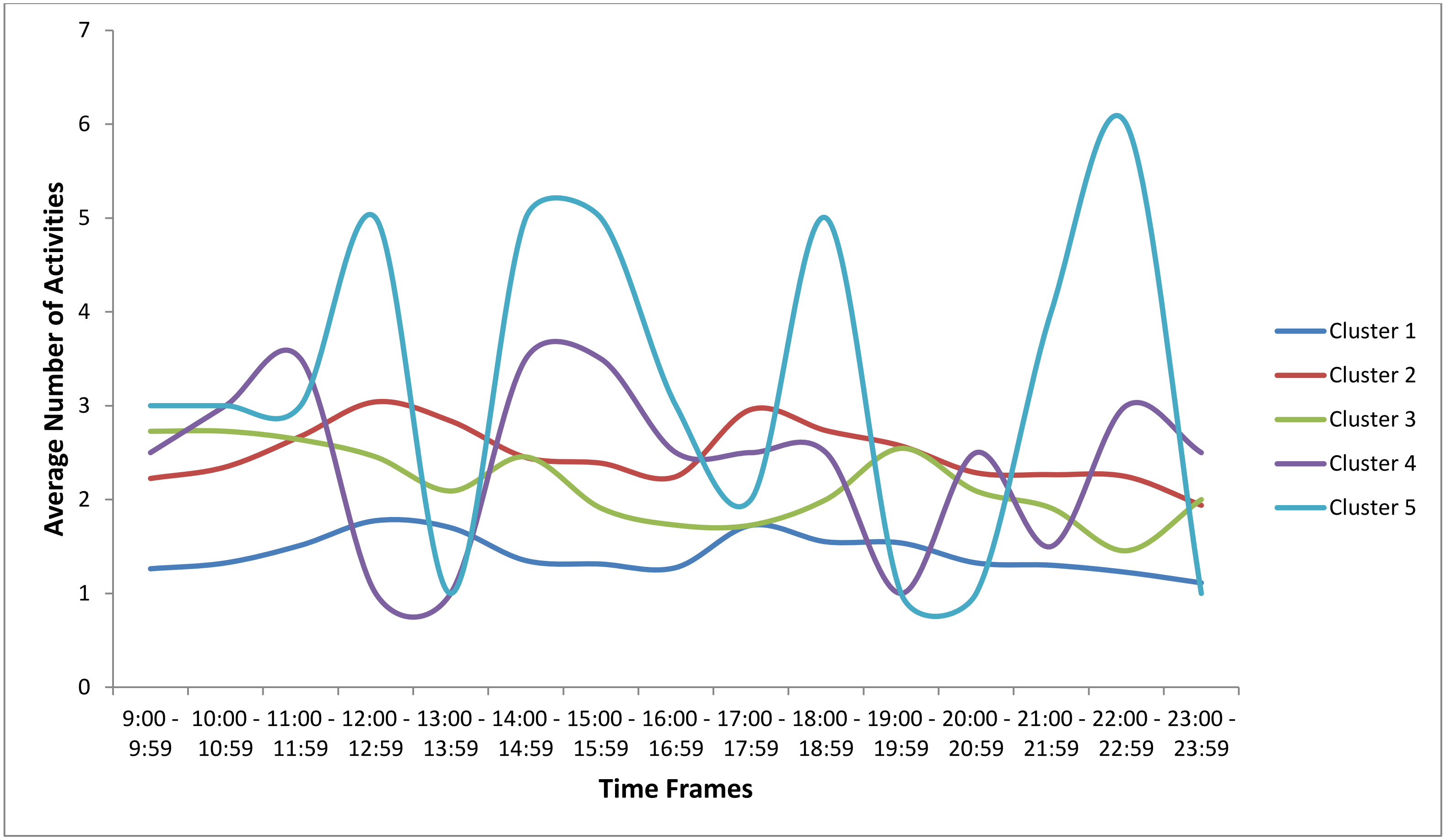



Figure 7.1: Clusters of daily activity intensity similarity index (DAISI)

7.6.1 Cluster 1: (F=11; S=14; G=25; U=30; n=80)

The members of this group have the least average number of activities in a day. On average, only about 1.5 activities are undertaken within every hour of the day. There are only two discernible peaks in the afternoon and evening time periods each. The afternoon peak is between 12:00 noon and 2:00 pm and the evening peak period is between 5:00 pm and 6:00 pm after which the intensity of activities dips into the evening. The group is dominated by graduate students (56.82%), staff (100%) and faculty (64.71%). Only 44.12% of the undergraduate population belongs to this group.

About 62.12% of the male population and 50.65% of the female population belong in this group. Of those who live on-campus, 48.89 % fall into this cluster, while 59.18% of non-campus dwellers are in the group. Interestingly, but expectedly, all respondents between the ages 45 and 59 (100%) and 80% of those 60 years and above are found in this cluster. About 71.43% of those 30 years and above and 45.97% of those under 30 years old are in this group. This seems appropriate since a large portion of the cluster is populated by faculty, staff and graduate students, who generally are older than undergraduate students. The group also consists of International (56.52%), Oklahoma (56.41%) and 54.76% of respondents from other states other than Oklahoma.

7.6.2 Cluster 2: (F=5; S=0; G=15; U=29; n=49)

This group shares an uncannily similar trend of activity profiles as Cluster 1. The peak periods are between 12:00 noon and 2:00 pm for the afternoon period, and 5:00 and 6:00 pm for the evening period, after which the activity rates dip, progressively, into the evening as well. The major difference is that the group has a higher average of activity rate of 2.5 per hour. Unlike Cluster 1, the relative percentage of undergraduates is higher (42.65%) as compared with zero (0%) for staff, 29.41% for faculty and 34.09% for graduate students. About 36.36% of the female population, and 31.82% of the male respondents; 37.78% of campus dwellers and 32.65% of those who live off-campus are found in this group. There is a higher percentage of those from other states other than Oklahoma (40.48%) than there are Oklahoma (32.05%) and International (30.43%) persons. About 41.38% of those who are less than 30 years of age and 23.21% of those who are 30 years and older are in this group. This is a much younger group than Cluster 1.

### 7.6.3 Cluster 3: (F=1; S=0; G=3; U=7; n=11)

This cluster has an activity profile that falls between Clusters 1 and 2. The average activity rate is about 2 activities per hour. The morning period has a higher activity rate than the afternoon and evening periods. The other two peak periods are between 2:00 and 3:00 pm and between 7:00 and 8:00 pm. The peak periods for Clusters 1 and 2 coincide, generally, with the periods of lower activity rates for Cluster 3. Between 3:00 pm and 6:00 pm and 9:00 pm – 11:30 pm the activity rates fall below the 1 activity mark and rise again around 11:30 pm when the activity rates of Clusters 1 and 2 are at their lowest.

About 5.88% of faculty, 6.82% of graduates and 10.29% of undergraduate populations make up the cluster group. About 7.69% of the Oklahoman, 13.04% of the Internationals and 4.76% of those from other states (other than Oklahoma) are also in the group. The remaining characteristic compositions of the group include 4.55% and

10.39% of the male and female populations, respectively; 11.11% and 6.12% of campus and off-campus residents, respectively; and, 81.82% of the group is less than 30 years old.

7.6.4 Cluster 4: (F=0; S=0; G=1; U=1; n=2)

Cluster 4 is one of the two outlier activity profiles. It consists of only two individuals, with high but very erratic activity profile behavior. Periods of high activity rates are punctuated by periods of very low activity rates. Peak periods are between 10:00 am and 11:00 am, 2:00 pm and 3:00 pm and 10:00 pm and 11:00 pm. The periods of very low activities are between 12:30 pm and 2:00 pm, 6:30 pm and 8:00 pm and 10:00 pm and 11:00 pm.

Both members of this group are Oklahoman, female, and live off-campus. One of them is between 18 and 24 years of age and undergraduate, and the other is between 40 and 44 years old and a graduate.

7.6.5 Cluster 5: (F=0; S=0; G=0; U=1; n=1)

Cluster 5 has only one activity profile, which is similar in trend to cluster 4 but of much higher intensity. The peaks are higher and the dips are as low compared to cluster 4. All the peak periods at 12:00 noon, 2:00 pm and 3:00 pm, 6:00 pm and 10:00 pm recorded 5 activities or more. The low rates are at 1:00 pm, 5:00 pm and 10:00 pm. Generally, each peak is followed immediately by a very low rate, producing an activity rate trend of very erratic character.

The only member of this group is undergraduate, Oklahoman, male, lives on campus and is between 18 and 24 years old.

It is instructive that clusters with older population tend to have relatively lower and more stable activity profiles than those with younger population. Clusters 1 and 2 basically have identical trends in activity profiles except that Cluster 2 has a younger population and higher activity rate. Most of the older population is found in Cluster 1, which explains the low and more stable activity rate. The composition of the group offers another compelling explanation to the trend of activity profile in the cluster. All staff and most of the faculty and graduate students fall within this group. As predicted in the operational framework, these groups in the university population are expected to have more stable work schedules. Unless other activities factor in, most staff has and would report only routine activities that do not vary much over time.

Faculty and graduate students have similar activity schedules, which may be few in number but intense in time consumption. Since DAISI measures intensity in terms of number/variety of activities being conducted, it is understandable that the activity rates for these groups is relatively low, compared with undergraduate students, who take large number of classes over relatively shorter periods of time. These groups of respondents are more likely to report lower activity rates than undergraduates. Incidentally, these groups also fall largely within the high fragmenters' group as well. As explained earlier for the daily activity schedule fragmentation index (DASFI), these groups usually have established anchor locations. For the staff and faculty, their offices are most probably the most obvious locations to return to after activities in other locations. Graduate students also have office or laboratory spaces to relocate to after engaging in activities elsewhere.

The younger population and higher proportion of undergraduate students in Cluster 2 may explain why the activity rates are higher than for Cluster 1 even though the trends are similar. The group also does not have any staff within its ranks. It may be safe to say that the staff group probably reports the least number of activities of any group in the university community because of the tendency for them to regard their daily schedule as routine. Since they are not involved in the group, the general activity rates do not appear to suffer any loss from averaging rates reported by other groups.

The activity rate of Cluster 3 falls between Clusters 1 and 2. The groups' composition is similar to Cluster 2. The population is younger, and largely undergraduate. This also lends credit to the proposition that cluster with a younger population has a more active activity profile.

Few concrete statements could be made about the activity profile patterns in Clusters 4 and 5 since these are outlier activity rates that do not fit into any of the first three cluster groups. Also there are hardly any effective statements that could be made on the time periods of the day that may or may not have higher rates of activity. The presence of the outliers distorts the picture and even when based only on the three "regular" clusters, there are no significant stand-out periods of the day for unusually high or low activity rates.

### 7.7 Significance and Implication of the Daily Activity Intensity Similarity Index (DAISI)

The daily activity intensity similarity index (DAISI) is significant for two basic reasons: first, it provides a method to statistically measure activity profiles, and second, it enhances understanding of how the activity profiles of individuals are similar or dissimilar, and how activity patterns of groups are created.

Conventionally, activity profiles have adopted simple statistical methods to test for variations between predefined groups. The results usually indicate whether any variations exist in the intensity of activities between groups. Individuals are subsumed under the groups, which are predefined and therefore generally presumed to be homogenous. DAISI adopts an approach that does not assume or pre-define groups. It also deals with individual activity profiles and compares one against the others to determine which individuals' activity profiles are most similar. In this way, the approach captures more the essence of the contributions of individuals to understanding activity participation rather than examine groups, which may have disparate activity profiles that may not necessarily be identified.

DAISI has individuals as the central components in understanding activity patterns. It therefore builds groups from sets of activity participants whose activity profiles are similar. The groups produced are therefore generally more inclined towards similar activity participation rate. The index therefore is able to expose any patterns that may be hidden in the data with regards to activity participation of individuals than may have been possible through predefined groups.

The use of a similarity measure to investigate patterns of activity participation is also an advantage of the index. Through a rigorous process, the approach provides another method through which activity profiles may be studied and patterns discerned. Traditionally, the process has been more descriptive in nature, with only simple statistical tests to examine any differences that may exist between groups. In this approach, a more rigorous procedure and process is brought to bear on investigating and understanding activity profiles. This does not only add a layer of method to the methodology of studying activity patterns but also lends it a veneer of rigor that enhances an extraction of patterns from the data rather than merely describe it.

Understanding these patterns provides the possibility to create enabling environment for better activity engagement. For example, understanding important time frames for high level of activity participation within a given time period may help in planning strategies to improve the activity participation experience. This may involve identifying types of activities, locations of activities and the nature of constraints that need to be overcome. DAISI provides the background through which such activity enhancement could be made possible.

DAISI employs already existing software packages to compute. The index and matrix can easily be calculated in Microsoft Excel Spreadsheet. The clusters can be generated by any statistical software package with clustering capabilities. This underscores the simplicity of the index, its accessibility and compatibility with already existing packages.

The index has some limitations as well. By insisting on same length of time frame for all individuals, the index is restrictive and rigid. It is not suitable for comparing differences across time scales and for disparate lengths of time. The index expects synchronized time frames for all individuals. This is a prerequisite for a meaningful computation of the daily activity intensity similarity index.

Secondly, the index is biased towards activity profiles with common entries between them. The *rd* component carries a lot of weight such that when two strings of activity profiles do not have any identical common entries but are numerically very close, the index returns a similarity score that is close to zero. Sequences with more common activity rates between them but large differences at other time frames would return a higher index value.

Third, it may be acknowledged that most the index values are stacked towards larger values. Invariably, the distribution may hardly be normal in most circumstances. This may not be unconnected with the fact that compared to the maximum number of 60 activities per hourly time frame used to normalize activity frequencies, most actual human activity numbers are very low. For the data collected and used as example in the study, the average activity frequency per hourly time frame is 1.8 activities. The highest number of activities reported at any given time frame is 6. The low values of actual activities, relative to the maximum number that may be allowed therefore may have skewed the values in favor of large index values because since most people would be reporting significantly smaller values, their similarity potentials are enhanced and probably exaggerated.

Finally, there is the issue of linearity that is still largely unexplored in the index. Preliminary evidence suggests progressive increase in index values with progressive similarity between activity profiles. The depth of this linearity however is currently untested. Though it may not be a limitation yet, its inclusion here is meant to acknowledge the possibility of its limitation.

CHAPTER VIII

SUMMARY AND RECOMMENDATION

## 8.1 Introduction

This chapter summarizes the study and makes some recommendations for future studies. The summary presents the basic research question and a survey of the time-geography framework, the bases for the indices and the results of the case study of the indices. The recommendations examine other related issues that could enhance the application of the indices and the time-geography framework in activity analysis.

## 8.2 Summary of the Dissertation

Human activity patterns are important to the understanding of human societies and interactions. Over the years, there have been several changes in the study of human activity patterns to reflect the reality of the human experience. Traditional methods of trip analysis as a basis for understanding activity patterns have given way to more realistic activity analysis. The underlying principle is that participating in activities across both space and time is the main reason behind trip making. Trips are therefore derived from the demand for activities. Consequently, activities are the basis of activity analysis.

Several frameworks have been developed and applied to activity analysis. This includes statistical and mathematical modeling, simulation methods, time studies and analysis, etc. Recently, attention is being paid to time geography whose framework possesses the necessary ingredients to effectively study and analyze human activity patterns. Several concepts of the framework, including stations, space-time paths and space-time prisms, activity constraints (capability, coupling and authority), and on activity sequencing are very appealing for the study of activity patterns. Importantly, it provides a mechanism to incorporate the time dimension into the study that hitherto had been limited to mostly studying human activities in space.

Time geography views time and space as inseparable and co-equal elements in human activity because every activity requires a location (space) but it equally is situated within a time frame. Both space and time are influential to participating in activities and therefore should be equally essential in studying them. Unfortunately, the elegant concepts embodied in the framework were for a large part inoperable until the 1990s and an improvement in computing technology opened the doors for engaging the framework in new and exciting ways. However, the fact that only recently has the framework been effectively engaged also means that there is still much work to be done. New methods of analyzing human activities need to be developed within the framework. This is the challenge that this dissertation sets out to accomplish.

Using data collected from faculty, staff, graduate students, and undergraduate students at Oklahoma State University, Stillwater Campus, this study develops two indices that enhance the understanding of activity patterns within a time-geography framework. The first index is the daily activity schedule fragmentation index (DASFI), and the second is the daily activity intensity similarity index (DAISI). Both indices utilize the concepts in the time-geography framework and components of activities participation.

The daily activity schedule fragmentation index (DASFI) uses the concepts of station to redefine anchor location as a central theme to fragmentation of an activity schedule. It also adopts the space-time path with its many appealing features, including maintaining the sequences of activity participation, as the primary construct to identify anchor locations for individual schedules, and for computing the index of fragmentation through a disintegration process of the path.

The station is basically the location in which activities take place. Traditionally the home and workplace locations are the designated anchor locations because they supposedly are the most important stations in the schedules. However, this dissertation argues that other locations other than these two may equally be considered. It therefore redefines an anchor location as the station that is most visited by the activity participant or the one with the largest share of time spent. Identifying this anchor location sets the tone for the disintegration of the space-time path to reveal the fragments of activity chains (or lack thereof) that the index measures. With a range between 0 and 1, the index identifies values closer to 1 as highly fragmented patterns and values closer to zero (0) as chained activity patterns. These values are then clustered, using a clustering technique (in this study, the k-means is used) to find groups of activity patterns based on fragmentation of activity schedules of individuals.

The fragmentation index tackles an important component of activity patterns. This is the fact that human beings try to maximize utility in participating in activities. Consequently, individuals tend to string their activities into chains rather than split them into bits and pieces, especially if they can help the situation. No methods have been devised to effectively tackle this problem and the daily activity schedule fragmentation index is presented as a means to investigate this trait in human activity patterns. Added to this is the fact that the index could identify important anchor locations that could be improved to enhance activity participation. It may therefore be an important tool in activity planning.

Using DASFI, three broad classes of activity patterns were discerned for both frequency-based and duration-based approaches. These are the high fragmenters, moderate fragmenters and activity chainers. High fragmenters are a group of individuals who resort to the anchor locations very frequently in the course of their activities. For the frequency-based approach, this group consisted of large numbers of faculty and staff along with graduate students. The moderate fragmenters group is composed of those whose return to the anchor locations is not as frequent as the high fragmenters', yet do not have long chains of activities in their schedule. This consisted largely of graduate and undergraduate students with a sprinkle of staff and faculty. The activity chainers group consisted only of undergraduate and a few graduate students. The duration-based approach had similar numbers of individuals in the group with similar compositions. The activity chainers, however, contained a few staff and faculty as well. The principal difference between the two approaches was the contribution of the sleeping/resting/idle activity type, which in the duration-based approach resulted into a change in the classification of at least 27 individuals. It is argued that were the sleeping/resting/idle activity type excluded, most of the individuals would have same anchor location for both frequency-based and duration-based approaches.

The daily activity intensity similarity index (DAISI) adopts the concept of activity sequences to measure activity profiles of individuals. Conventionally, activity profiles are an important measure of activity participation for groups of persons. Statistical tests are applied to determine any variations in the proportion of groups that may be engaged in a particular activity at particular time frames.

The DAISI provides a new method by which to analyze and understand patterns of activity profiles. Instead of comparing the activity profiles of predefined groups, the index measures and compares the activity profiles of individuals, which is more aligned with the time-geography framework. It is also based on a similarity measure rather than finding just variations between groups. In this way it seeks to find individuals that have similar activity profiles and group them together through a clustering technique (k-means is used in this study). This yields more intuitive groups of individuals with similar characteristics of activity participation than predefined groups of participants.

The number of activities that each participant undertook within hourly time frames is counted for the duration of the study. The sequences in which these numbers of activities are undertaken are kept and compared to other individuals of same time frames to derive a ratio of similarity between individual participants. The sum of the (absolute) differences between corresponding time frames is measured against the sum total for the activities in the corresponding time frames for the individuals being compared. The result is multiplied by the ratio of similarity to produce an index of similarity for activity profiles between two individuals. These similarity index values are organized into a symmetrical matrix that is subjected to a k-means clustering technique. The resulting groups are the quizzed for any characteristics that may throw more light into the organization of activity profiles by individuals.

Five clusters of activity profiles were recognized. The first cluster consisted of individuals who were very active in the morning but as the day wore on became less involved in many activities. Most of the members were young undergraduates. The second cluster had individuals who were involved in many activities throughout the day. Most members of this group were young and female. The third cluster consisted of older population, with most of the staff and some of the faculty involved. The activity rate is generally low and stable. The fourth cluster members were most active in the afternoon and late evening. It is dominated by graduate and undergraduate students. Most of them live off-campus and are young. The fifth cluster has average activity rates, which is stable throughout the day. It has a fair share of faculty and staff but is largely dominated by undergraduates. More than a quarter of the group is 35 years and older. Generally, clusters with larger proportion of older individuals indicated lower and more stable levels of activity rates.

The index of similarity for activity profiles provides insight into the characteristics of individuals that may be engaged in different levels of activity rates and also the time frames in which high or low levels of activity rates are common. This may be important in organizing activities and enhancing the human activity experience by providing necessary facilities at appropriate times and to the appropriate persons.

The two indices developed in this study therefore possess potentials for revealing hidden patterns of activity participation in a dataset. They also have the capacity to help better the understanding of activity patterns and in the process to enhance the planning of and provision of the necessary facilities to promote activity participation. For example, DASFI identifies two principal groups: the fragmenters and the chainers.

The fragmenters are generally individuals who have a clear activity organizing location, which they return to very often or spend large amount of time at in the course of carrying out daily activities. This anchor location is an important component of their activity schedule. Anchor locations identified based on the frequency of visit principle suggest a stream of human traffic both to and from the anchor location, especially when the location serves as anchor location for many individuals, e.g., a building on campus. Access to such a location is therefore important to the many individuals whose activity anchor needs are served by the building. It may be appropriate to restrict vehicular traffic around the location to allow for easy access to the location.

When duration at a location is the basis for identifying anchor locations, enhancing convenience and comfort at the locations may be appropriate strategy to adopt. Depending on the socio-economic and demographic characteristics of the people and the physical characteristics of the location appropriate facilities to foster interaction and generate synergy.

For activity chainers, the anchor location is usually not as important compared to their needs for efficient transition from one activity location to another. To plan effective movement between activity locations, there is need to identify two things: (i) time periods when travel between locations are more intense (when most people are transiting from one activity location to another), and (ii) directions of movement between activity locations. These will allow for the creation of an appropriate transit response to provide seamless transition between important activity locations. For instance, in a university campus such as OSU, Stillwater, blue (vehicle-free) zones or pedestrian-friendly or bike-friendly zones between locations may be as important as the provision of transit buses between important locations at transition-intense time periods.

DAISI may be important in identifying activity intense periods of the day and the characteristics of the individuals involved. Activity characteristics gleaned from the clusters generated may provide the basis on which to plan effective strategies to ease transition between important activity locations and provide facilities and amenities at important anchor locations and during important transition times.

It is important to indicate that this planning framework could be extended to urban and regional planning. Appropriate urban and regional centers that are identified as anchor locations for both fragmenters and chainers could receive the appropriate attention and response that may serve the needs to multiple populations and provide a more vibrant activity engagement process.

### 8.3 Recommendations for Further Studies

Time geography is an extensive framework whose appeal for activity analysis is tremendous. There are many concepts, precepts and principles that have been employed to analyze activities. There are still many more that can be applied to further the study of activity patterns. This study makes a few recommendations on improving the methods developed in the study and other aspects of the time-geography framework that could be applied to further activity analysis.

The space-time prism has become a staple for studying human accessibility and activity space. Both the prism and potential path areas are important concepts that shed light on the nature of human interactions and capability. It is recommended that these concepts be further extended to develop individual methods to measure actual activity participation within the activity space. Though much work has been done on the theoretical component of the time-geography framework, few actual methods have been derived to implement these elegant frameworks.

The data collected for this study were based on a two weekdays recall activity diary/questionnaire. One of the major criticisms of this technique is the notorious problem of recall lapse. When people have to struggle to recall every bit of activity, its location, the approximate starting and ending times, and other characteristics, it tends to increase the time needed to fill out the diary. This discourages people who may be willing to respond to the research survey. It is easy to envisage this problem as one of the contributors to the low response rates associated with this type of survey. The employment of location-aware technologies such as the GPS has the potentials to enhance the accuracy and response rates of activity data collection because it collects real-time activity data, when used appropriately. Though it has its problems such as its high costs and perceived intrusiveness in the private lives of people especially that it needs to be carried along, the GPS technology possesses the capacity to collect larger quantity and more reliable data as it measures exact locations. An integrated approach involving the use of GPS and in-depth interview to obtain the qualitative data on activity characteristics and scheduling processes is recommended for future data collection in activity studies.

The daily activity schedule fragmentation index (DASFI) and the daily activity intensity similarity index (DAISI) could both be developed into GIS prototypes to enhance easier and more accurate processing of data and analysis. This may be especially true for the fragmentation index, which relies on an extended framework of temporal GIS extension to construct the space-time path necessary for computing the index. Aligning the index extension with the temporal GIS extension it relied on to construct the space-time paths would make for easier software bundling and computing convenience.

## REFERENCES


Abdelghany, A.F. and Mahmassani, H.S. (2003) Temporal-spatial micro-assignment and sequencing of travel demand with activity/trip chains. *Transportation Research Record*, 1831:89-97.

Achananuparp, P., Hu, X. and Shen, X. (2008) The evaluation of sentence similarity measures. Proceedings of the 10th International Conference on Data Warehousing and Knowledge Discovery, 305-316. Berlin, Germany: Springer-Verlag.

Adams, P. C. (2000) Application of a CAD-based accessibility model. In D. G. Janelle and Hodge, D. C. (Eds.). *Information, place and cyberspace: Issues in accessibility*. Berlin, Germany: Springer-Verlag.

Ahas, R., Aasa, A., Silm, S. and Tiru, M. (2010) Daily rhythms of suburban commuter's movements in the Tallinn metropolitan area: Case study with mobile positioning data. *Transportation Research Part C*, 18:45-54.

Alam, S. B. and Goulias, K. G. (1999) Dynamic emergency evacuation management system using spatio-temporal models of behavior. *Transportation Research Record,* 1660:92-99.

Alexander, B., Ettema, D. and Dijst, M. (2010) Fragmentation of work activity as a multi-dimensional construct and its association with ICT, employment and sociodemographic characteristics. *Journal of Transport Geography,* 18:55-64.

Alexander, B., Hubers, C., Schwanen, T., Dijst, M. and Ettema, D. (2011) Anything, anywhere, anytime? Developing indicators to assess the spatial and temporal fragmentation of activities. *Environment and Planning B: Planning and Design,* 38(4):678-705.

Allison, M. L., Sivaramakrishnan, S. and Chandra, R. B. (2005) Exploratory analysis of weekend activity patterns in the San Francisco Bay Area, California. *Transportation Research Record,* 1926:70-78.

Andrieko, G., Andrieko, N. and Wrobel, S. (2007) Visual analytics tools for analysis of movement data. ACM SIGKDD Explorations Newsletter (Special Issue), 9(2).

Arentze, T. and Timmermans, H. (2000) ALBATROSS: A learning-based transportation oriented simulation system. EIRASS, Technische Universiteit Eindhoven.

Axhausen, K. and Garling, T. (1991) Activity-based approaches to travel analysis: Conceptual frameworks, models, and research problems. *Transport Reviews*, 12:324-341.

Axhausen, K. W. and Herz, R. (1989) Simulating activity chains: German approach. *Journal of Transportation Engineering*, 115(3):316-325.

Bai, X. and Latecki, L. J. (2008) Path similarity skeleton graph matching. *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)*, 30(7):1-11.

Bay, S. D. and Pazzani, M. J. (2001) Detecting group differences: Mining contrast sets. *Data Mining and Knowledge Discovery*, 5(3):213-246.

Balazinska, M. and Castro, P. (2003) Characterizing mobility and network usage in a corporate wireless local-area network. Proceedings of MobiSys 2003:303-316, May.

Behrens, R. (2000) Activity-travel analysis: A review of theoretical origins, recent developments and local application. South African Transport Conference *'Action in Transport for the New Millennium'* Organized by Conference Planners South Africa, 17-20 July.

Bhat, C. R. and Lawton, K. (2000) Passenger travel demand forecasting. *Transportation in the new millennium*. Transportation Research Board, Washington, D.C.

Boyce, D. E., Zhang, Y-F. and Lupa, M. R. (1994) Introducing feedback into four-step travel forecasting procedure versus equilibrium solution of combined model. *Transportation Research Record*, 1443:65-74.

Calkins, C. R. and Davis, G. W. (1980) Fragmentation index of raw muscle as a tenderness predictor of steaks from US good and US standard steer and bullock carcasses. Faculty Papers and Publications in Animal Science, University of Nebraska-Lincoln.

Carlstein, T., Parkes, D. and Thrift, N. (1978) *Human activity and time-geography*. New York, USA: John Wiley and Sons.

Carpenter, S. and Jones, P. (1983) *Recent advances in travel demand modeling*. Aldershot, UK: Gower.

Carrasco, J. A., Hogan, B., Wellman, B. and Miller, E. J. (2008) Collecting social network data to study social activity-travel behaviour: An egocentric approach. *Environment and Planning B: Planning and Design*, 35(6):961-980.

Caulier, J-F. (2010) On the measurement of fragmentation. Cahier du CEREC 2010/4. http://centres.fusl.ac.be/CEREC/document/2010/april10_frag.pdf (accessed on June 15, 2011).

Clarke, M., Dix, M., Jones, P. and Heggie, I. (1981) Some recent developments in activity-travel analysis and modeling. *Transportation Research Record*, No. 794

Couclelis, H., Golledge, R.G., Gale, N. and Tobler, W. (1987) Exploring the anchor-point hypothesis of spatial cognition. *Journal of Environmental Psychology*, 7:99-122.

Couclelis, H. (2000) From sustainable transportation to sustainable accessibility: Can we avoid a new tragedy of commons? In Janelle, D.G. and Hodge, D.C. (Eds.). *Information, place and cyberspace*. Berlin, Germany: Springer-Verlag.

Couclelis, H. (2004) Pizza over the Internet: E-commerce, the fragmentation of activity, and the tyranny of the region. *Entrepreneurship & Regional Development*, 16(1):41-54.

Demsar, U. and Virrantaus, K. (2010) Space-time density of trajectories: Exploring spatiotemporal patterns in movement data. *International Journal of Geographical Information Science*, 24(10):1527-1542.

Doherty, S. T. (2004) Rules for assessing activity scheduling survey respondents' data quality. *Journal of the Transportation Research Board*, 1870:109-115.

Dykes, J. A. and Mountain, D. M. (2003) Seeking structure in records of spatio-temporal behaviour: Visualization issues, efforts and applications. *Computational Statistics & Data Analysis*, 43:581-603.

Ellegard, K. (1999) A time-geographical approach to the study of everyday life of individuals – a challenge of complexity. *GeoJournal*, 48(3):167-175.

Ellegard, K., Hagerstrand, T. and Lenntorp, B. (1977) Activity organization and the generation of daily travel: Two future alternatives. *Economic Geography*, 53(2):126-152, April.

Ellegard, K. and Svedin, U. (2012) Torsten Hagerstrand's time-geography as the cradle of the activity approach in transport geography. *Journal of Transport Geography*, 23:17-25.

Eom, J. K, Stone, J. R. and Ghosh, S. K. (2009) Daily activity patterns of university students. *Journal of Urban Planning and Development*, ASCE / December:141-149.

Ettema, D. and  Timmermans, H. (1995) Theories and models of activity patterns. In D. Ettema and Timmermans, H. (Eds.). *Activity-based approaches to travel analysis*. Kidlington, UK: Elsevier Press.

Fellendorf, M., Naokel, K. and Handke, N. (2000) VISUM-online -traffic management for the EXPO 2000 based on a traffic model. 7th World Congress on Intelligent Transport Systems, Turin, Italy, November.

Florian, M., Nguyen, S. and Ferland, J. (1975) On the combined distribution-assignment of traffic. *Transportation Science*, 9:43-53.

Frye, L., Berger, P. D. and Weinberg, B. D. (2012) An analysis of college-student travel-destinations and its marketing implications. *International Journal of Humanities and Social Science*, 2(11):45-55.

Fuji, S. and Kitamura, R. (2000) Evaluation of trip-reducing effects of new freeways using a structural equations model system of commuters' time use and travel. *Transportation Research Part B: Methodological*, 34:339-354.

Garling, T., Gillholm, R. and Garling, A. (1998) Reintroducing attitude theory in travel behavior research: The validity of an interactive interview procedure to predict car use. *Transportation,* 25:129-146.

Garling, T., Kwan, M-P. and R. Golledge (1994) Computational process modeling of household activity scheduling. *Transportation Research B*, 28:355-364.

Golledge, R. G. and Spector, A. (1978) Comprehending the urban environment: Theory and practice. *Geographical Analysis*, 10:403-426.

Golledge, R. G. and Stimson, R. J. (1997) *Spatial behavior: A geographic perspective*. New York, USA: The Guildford Press.

Golob, T. F. (1986) A non-linear canonical correlation analysis of weekly trip chaining behavior in the Netherlands. *Transportation Research A*, 20A:385-389.

Golob, T. F. (1997) A model of household demand for activity participation and mobility. Working paper UCTC No. 335. The University of California Transportation Center, University of California, Berkeley.

Golob, T. F. (2001) TravelBehavior.com: Activity approaches to modeling the effects of information technology on personal travel behavior. In D.A. Hensher (Ed.). *Travel behaviour research: The leading edge.* Oxford, UK: Pergamon Press.

Golob, T. F., Kitamura, R. and Lula, C. (1994) Modeling the effects of commuting time on activity duration and non-work travel. Paper presented at Annual Meeting of Transportation Research Board, Washington, DC, January.

Golob, T. F. and McNally, M. G. (1997) A model of activity participation and travel interactions between household heads. *Transportation Research B*, 31(3):177-194.

Goodwin, P. (1983) Some problems in activity approaches to travel demand. In S. Carpenter and Jones, P. (Eds.). *Recent advances in travel demand analysis*. Aldershot, UK: Gower.

Gould, J. and Golob, T. F. (1997) Shopping without travel or travel without shopping: An investigation of electronic home shopping. *Transport Reviews,* 17:355-376.

Gunopulos, D. and Das, G. (2000) Time series similarity measures (tutorial PM-2). Proceeding KDD '00 Tutorial notes of the sixth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Boston, MA, August.

Hagerstrand, T. (1970) What about People in Regional Science? *Papers of the Regional Science Association,* 24:7-21.

Hagerstrand, T. (1975) Space, time and human conditions. In A. Karlqvist, Lundqvist, L. and Snickars, F. (Eds.). *Dynamic allocation of urban space*. Massachusetts, USA: Lexington Books.

Hall, P. (1988) *Cities of tomorrow*. Oxford, UK: Blackwell Publishers.

Halpin, B. (2007) Sequence analysis. Essex Summer School University of Limerick, July 13-15. http://teaching.sociology.ul.ie/seqanal/essex/labs.pdf (accessed, May 20, 2013).

Hammadou, H., Thomas, I., van Hofstraeten, D. and Verhetsel, A. (2004) Distance decay in activity chains analysis: A Belgian case study. Paper presented at the ERSA 2003 Congress, Jyvaskyla, Finland, September.

Hanson, S. (1980) Spatial diversification and multipurpose travel: Implications for choice theory. *Geographical Analysis*, 12:245-257.

Hanson, S. and Huff, J. (1988) Systematic variability in repetitious travel. *Transportation*, 15:111-135.

Hasan, M. and Al-Gadhi, S. (1998) Application of simultaneous and sequential transportation network equilibrium models to Riyadh, Saudi Arabia. *Transportation Research Record,* No. 1645:127-132.

Harvey, A. and Wilson, C. (2001) Evolution of daily activity patterns from 1971 to 1981: A study of Halifax activity panel survey. *Canadian Studies in Population* (Special Issue on Longitudinal Methodology), 28(2):459-489.

Hayek, L. A. C. (1994) Analysis of amphibian biodiversity data. In W. R. Heyer, Donnely, M. A. and McDiarmid, R. W. (Eds.). *Measuring and monitoring miological diversity: Standard methods for amphibians*. Washington D. C.: Smithsonian Institution.

Henderson, T., Kotz, D. and Abyzov, I. (2004) The changing usage of a mature campus-wide wireless network. In Proceedings of ACM MobiCom, Philadelphia, September.

Holzapfel, H. (1986) *Trip relationships in urban areas*. Aldershot, UK: Gower.

Horton, F. E. and Reynolds, D. R. (1971) Action space differentials in cities. In H. McConnell and Yaseen, D. (Eds.). *Perspectives in geography: Models of spatial interaction.* DeKalb, USA: Northern Illinois University Press.

Hsu, W-J. and Helmy, A. (2005) *IMPACT*: *I*nvestigation of *m*obile-user *p*atterns *a*cross university *c*ampuses using WLAN *t*race analysis. http://arxiv.org/PS_cache/cs/pdf/0508/0508009v1.pdf (accessed Sept. 4, 2010).

Hubers, C., Schwanen, T. and Dijst, M. (2008) ICT and temporal fragmentation of activities: An analytical framework and initial empirical findings. *Journal of Economic and Social Geography (TESG),* 99(5):528-546.

Ihler, A. T. and Smyth, P. J. (2007) Learning time-intensity profiles of human activity using non parametric Bayesian models. *Neural Information Processing Systems,* 19:625-632.

Islam, M. T. and Habib, K. M. N. (2012) Unravelling the relationship between trip chaining and mode choice: Evidence from a multi-week travel diary. *Transportation Planning and Technology,* 35(4):409-426.

Isler, E. and Hoel, L. A. (2004) The effect of land use planning on university transportation systems. Center for Transportation Studies at the Univ. of Virginia. Prepared for the Mid-Atlantic Universities Center for Transportation, Research Rep. No. UVACTS-5-14-67.

Jiang, S., Ferreira, J. and Gonzalez, M. C. (2012) Clustering daily patterns of human activities in the city. *Data Mining and Knowledge Discovery,* 25:478-510.

Jones, P. M., Dix, M. C., Clarke, M. I. and Heggie, I. G. (1983) *Understanding travel behavior.* Aldershot, UK: Gower.

Jones, P., Koppelman, F. and Orfeuil, J-P. (1990) Activity analysis: State-of-the-art and future directions. In P. Jones (Ed.). *Developments in dynamic and activity-based approaches to travel analysis*. Aldershot, UK: Gower.

Kay, J., Maisonneuve, N., Yacef, K. and Zaiane, O. R. (2006) Mining patterns of events in students' teamwork data. Proceedings of Educational Data Mining workshop, Jhongli, Taiwan, July.

Kenyon, S. (2006) Accessibility diaries: A new methodological approach to understand the impact of Internet use upon personal travel. *Journal of Transport Geography*, 14(2):123-134.

Kenyon, S. and Lyons, G. (2007) Introducing multitasking to the study of travel and ICT: Examining its extent and assessing its potential importance. *Transportation Research A*, 41:161-175.

Kitamura, R. (1988) An evaluation of activity-based travel analysis. *Transportation*, 15(1-2):9-34.

Kitamura, R. (1996) Applications of models of activity behavior for activity based demand forecasting. Proceedings of "Activity-based Travel Forecasting" Conference US Department of Transportation, Washington, DC, Report DOT-T-97-17.

Kitamura, R., Robinson, J. P., Golob, T. F., Bradley, M., Leonard, J. and van der Hoorn, T. (1992) A comparative analysis of time use data in the Netherlands and California. Institute of Transportation Studies, University of California, Davis, Research Report UCD-ITS-RR-92-09.

Koutra, D., Parikh, A., Ramdas, A. and Xiang, J. (2011) Algorithms for graph similarity and subgraph matching. http://www.cs.cmu.edu/~jingx/docs/DBreport.pdf (accessed May 22, 2013).

Krizek, K. J. (2003) Neighborhood services, trip purpose, and tour-based travel. *Transportation*, 30:387-410.

Kulkarni, A. and McNally, M. G. (2000) An activity-based microsimulation model for generating synthetic activity-travel patterns: Initial results. UC Irvine: Center for Activity Systems Analysis. http://www.escholarship.org/uc/item/1596c2wb (accessed, June 15, 2010).

Kuppam, A. R. and Pendyala, R. M. (2001) A structural equations analysis of commuter activity and travel patterns. *Transportation*, 28(1):33-54.

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

Kwan, M-P. (2000a) Human extensibility and individual hybrid-accessibility in space–time: A multi-scale representation using GIS. In D.J. Janelle and Hodge, D. C. (Eds.). *Information, place, and cyberspace: Issues in accessibility*. Berlin, Germany: Springer-Verlag.

Kwan, M-P. (2000b) Interactive geovisualization of activity-travel patterns using three-dimensional geographical information systems: A methodological exploration with a large data set. *Transportation Research C: Emerging Technologies*, 8:185-203.

Kwan, M-P. (2000c) Analysis of human spatial behavior in a GIS environment: Recent developments and future prospects. *Journal of Geographical Systems*, 2(1):85-90.

Kwan, M-P. (2003) Geovisualization of activity-travel patterns using 3D geographical information systems. Proceedings of the Tenth International Conference on Travel Behaviour Research, Lucerne, Switzerland, August.

Kwan, M-P. and Weber, J. (2003) Individual accessibility revisited: Implications for geographical analysis in the twenty-first century. *Geographical Analysis*, 35(4):341-353.

Langran, G. (1992) *Time in geographic information systems*, London: Taylor and Francis.

Lee, M.S. and McNally, M.G. (2002) Measuring physical accessibility with space-time prisms in a GIS: A case study of access to health-care facilities. http://www.its.uci.edu/~casa/casa-pubs.html (accessed February 6, 2013).

Lee-Gosselin, M. and Miranda-Moreno, L. F. (2009) What is different about urban activities of those with access to ICTs? Some early evidence from Québec, Canada. *Journal of Transport Geography,* 17:104-114.

Lenz, B. and Nobis, C. (2007) The changing allocation of time activities in space and time by the use of ICT-fragmentation as a new concept and empirical results. *Transportation Research Part A*, 41:190-204.

Lenntorp, B. (1976) Paths in space-time environments: A time geographic study of movement possibilities of individuals. Lund Studies in Geography Number 44, Royal University of Lund, Sweden.

Lenntorp, B. (1978) A time geographic simulation model of individual activity programmes. In T. Carlstein, Parkes, D. and Thrift, N. (Eds.). *Timing space and spacing time, Vol. 2: Human activity and time-geography*. London, UK: Longman.

Lin, M. L., Chu, C. M., Tsai, C. H., Chen, C. C. and Chen, C. Y. (2009) Geovisualization of tourist activity travel patterns using 3D GIS: An empirical study in Tamsui, Taiwan. *World Academy of Science, Engineering and Technology,* 36:454-458.

Liu, M., Mao, B., Gao, F., Guo, J. and Gao, L. (2008) Analysis on commuter's activity chain choice behavior. Proceedings of the Traffic and Transportation Studies Congress Conference, ASCE, Nanning, China, August.

Lockwood, A. M., Srinivasan, S. and Bhat, C. R. (2004) An exploratory analysis of weekend activity patterns in the San Francisco Bay Area. Technical Report No. SWUTC/o4/167820-1, Southwest Region University Transportation Center, Texas Transportation Institute, Texas A & M University System.

Lynch, K. (1960) *The image of the city*. Massachusetts, USA: Cambridge Press.

Mark, D. M. and Egenhofer, M. J. (1998) Geospatial lifelines. In O. Guenther, Sellis, T. and Theodoulidis, B. (Eds.). *Integrating spatial and temporal database*, Dagstuhl Seminar Report No. 228.

McGuckin, N. and Murakami, E. (1999) Examining trip-chaining behaviour – a comparison of travel by men and women. *Transportation Research Record*, 1693.

McGuckin, N. and Nakamoto, Y. (2004) Trips, chains, and tours – using an operational definition. NHTS Conference. http://www.trb.org/conferences/nhts/McGuckin.pdf.(accessed March 21, 2011).

McNally, M. G. (2000) The activity-based approach. In D. A. Hensher and Button, K. J. (Eds.). *Handbook of transport modeling*. New York, USA: Pergamon Press.

McNally, M. G. and Rindt, C. (2008) *The activity-based approach*. Center for Activity Systems Analysis, Institute of Transportation Studies, UC Irvine http://escholarship.org/uc/item/86h7f5v0 (accessed, June 15, 2010).

McNally, M. G. and Recker, W. W. (1986) *On the formation of household travel/activity patterns*. Institute of Transportation Studies, University of California, Irvine, CA. USDOT Final Report.

McNett, M. and Voelker, G. (2003) Access and mobility of wireless PDA users, *ACM SIGMOBILE Mobile Computing and Communications Review*, 7(4), October.

Meng, X., Wong, S., Yuan, Y. and Lu, S. (2004) Characterizing flows in large wireless data networks. Proceedings of ACM MobiCom, Philadelphia, September.

Metzler, D, Dumais, S. and Meek, C. (2007) Similarity measures for short segments of text. *Advances in Information Retrieval Lecture Notes in Computer Science*, 4425:16-27.

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

Miller, H. J. (1999) Measuring space-time accessibility benefits within transportation networks: Basic theory and computational methods. *Geographical Analysis*, 31:187-212.

Miller, H. J. (2004) Activities in space and time. In P. Stopher, Button, K., Haynes, K. and Hensher, D. (Eds.). *Handbook of transport 5: Transport geography and spatial systems*. Pergamon/Elsevier Science.

Miller, H. J. (2005) Necessary space-time conditions for human interaction. *Environment and Planning B: Planning and Design*, 32:381-401.

Miller, H. J. (2007) Social exclusion in space and time. In K.W. Axhausen (Ed.). *Moving through nets: The social and physical aspects of travel*. Oxford, UK: Elsevier Press.

Miller, H. J. and Shaw, S-L. (2004) *Geographic information systems for transportation: Principles and applications*. New York, USA: Oxford University Press.

Misra, R. and Bhat, C. R. (2000) Activity travel patterns of non-workers in the San Francisco Bay Area: Exploratory analysis. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 1718, TRB, National Research Council, Washington, D.C..

Mokhtarian, P. L. (1990) A typology of relationships between telecommunications and transportation. *Transportation Research,* 24A(3):231-242.

Mokhtarian, P. L. (1997) Now that travel can be virtual will congestion virtually disappear? *Scientific American,* 277:93.

Montgomery, H. (1990) From cognition to action: The search for dominance in decision making. In H. Montgomerya and Svenson, O (Eds.). *Process and structure in human decision making.* New York, USA: John Wiley and Sons.

O'Sullivan, D. (1995) Geographical information science: Time changes everything. *Progress in Human Geography,* 29(6):749-756.

Page, R. (1996) TREEVIEW: An application to display phylogenetic trees on personal computers. *Computer Applications in the Biosciences,* (12):357-358.

Pas, E. I. (1982) Analytically derived classifications of daily travel-activity behavior: Description, evaluation and interpretation. *Transportation Research Record,* No. 879.

Pas, E. I. (1984) The effect of selected sociodemographic characteristics on daily travel-activity behavior. *Environment and Planning A*, 16:571-81.

Pas, E. I. (1990) Is travel demand analysis and modelling in the doldrums? In P. Jones (Ed.). *Developments in dynamic and activity-based approaches to travel analysis*. Oxford Studies in Transport. Aldershot, UK: Gower.

Peuquet, D. J. (1999) Time in GIS and geographical databases. In D. J. Maguire, Goodchild, M., Rhind, D. W. and Longley, P. A. (Eds.). *Geographical information systems: Principles and applications*. London, UK: Longman Press.

Pipkin, J. (1995) Disaggregate models of travel behaviour. In Hanson, S. (Ed.). *The geography of urban transportation*. New York, USA: The Guilford Press.

Poinsatte, F. and Toor, W. (1999) Finding a new way: Campus transportation in the twenty-first century. Final Report, Univ. of Colorado Environmental Center, Boulder, Colorado.

Pred, A. (1977) *City systems in advanced economies: Past growth, present processes, and future development options*. London, UK: Hutchinson.

Primerano, F., Taylor, M. A. P., Pitaksringkarn, L. and Tisato, P. (2008) Defining and understanding trip chaining behavior. *Transportation*, 35 (1):55-72.

Raux, C., Ma, T-Y and Cornelis, E. (2011) Variability versus stability in daily travel and activity behavior: The case of a one week travel diary. http://halshs.archives-ouvertes.fr/halshs-00612610/ (accessed April, 2012).

Recker, W. W., McNally, M. G. and Root, G. S. (1986a) A model of complex travel behavior -part 1: Theoretical development. *Transportation Research A*, 20A:307-318.

Recker, W. W., McNally, M. G. and Root, G. S. (1986b) A model of complex travel behavior -part 2: An operational model, *Transportation Research A*, 20A:319-330.

Reichman S (1976) Travel adjustments and life styles: A behavioral approach. In P. Stopher and Meyburg, A. (Eds.). *Behavioral travel demand models.* Massachusetts, USA: Lexington Books.

Rinner, C. (2004) Three-dimensional visualization of activity-travel patterns. Proceedings of the Munster GI Days: 231-237.

Schenk, A. K., Witbrodt, B. C., Hoarty, C. A., Carlson, R. H., Goulding, E. H., Potter, J. F., and Bonasera, S. J. (2011) Cellular telephones measure activity and lifespace in community-dwelling adults: Proof of principle. *The American Geriatrics Society*, 59:345-352.

Scott, D. M. (2006) Constrained destination choice set generation: a comparison of GIS-based approaches. 85th Annual Meeting of the Transportation Research Board: Compendium of Papers CD-ROM, Washington, D.C., January.

Shaw, S-L. (2006) What about 'time' in transportation geography? *Journal of Transport Geography*, 14(3):237-240.

Shaw, S-L, and Yu, H. (2009a) A GIS-based time-geography approach of studying individual activities and interactions in a hybrid physical-virtual space. *Journal of Transport Geography*, 17:141-149.

Shaw, S-L and Yu, H. (2009b) Extended time-geography framework tools extension for ArcGIS 9.3 http://web.utk.edu/~sshaw/NSF-Project-Website/download.htm (accessed May 24, 2009).

Shoval, N. and Isaacson, M. (2007a) Tracking tourists in the digital age. *Annals of Tourism Research*, 34(1):141-159.

Shoval, N. and Isaacson, M. (2007b) Sequence alignment as a method for human activity analysis in space and time. *Annals of the Association of American Geographers*, 97(2): 281-296.

Sinha, G. and Mark, D. M. (2005) Measuring similarity between geospatial lifelines in studies of environmental health. *Journal of Geographical Systems*, **7**:115-136.

Spiess, H. (1996) Computing activity chain based trip distribution models. http://emme2.spiess.ch/archive/postscript/tchain.pdf (accessed June 2, 2010).

Srinivasan, S. and Rogers, P. (2005) Travel behavior of low-income residents: Studying two contrasting locations in the city of Chennai, India. *Journal of Transport Geography*, 13(3):265-274.

Stopher, P. R. (1992) Use of an activity-based diary to collect household travel data. *Transportation*, 19:159-176.

Taaffe, E., Gauthier, H. and O'Kelly, M. (1996) *Geography of transportation*. Englewood Cliffs, USA: Prentice-Hall.

Thill, J-C. and Thomas, I. (1987) Toward conceptualizing trip-chaining behavior: A review. *Geographical Analysis*, 19(1):1-17.

Thrift, N. (1977a) An introduction to time-geography. Concepts and Techniques in Modern Geography No 13, Geo Abstracts, Norwich.

Thrift, N. (1977b) Time and theory in geography 2. *Progress in Human Geography*, (October) 1: 413-457.

Tiakas, E., Papadopoulos, A. N., Nanopoulos, A. Manolopoulos, Y., Stojancivic, D. and Djordjevic-Kajan, S. (2009) Searching for similar trajectories in spatial networks. *Journal of Systems and Software*, 82:772-788

Todorova, S.A. and Lulcheva, D. (2005) Economic and social effects of land fragmentation on Bulgarian agriculture. *Journal of Central European Agriculture*, 6 (4):555-562.

Tulloss, R. (1997) Assessment of similarity indices for undesirable properties and a new tripartite similarity index based on cost functions. In M. E. Palm and Chapela, I. H. (Eds.). *Mycology in sustainable development: Expanding concepts, vanishing borders.* Boone, USA: Parkway Publishers.

Unwin, T. (1992) *The place of geography*. Harlow, UK: Longman.

Vanhulsel, M., Beckx, C., Janssens, D., Vanhoof, K. and Wets, G. (2011) Measuring dissimilarity of geographically dispersed space-time paths. *Transportation*, 38:65-79.

Vilhelmson, B. (1999) Daily mobility and the use of time for different activities: The case of Sweden. *GeoJournal*, 48:177-185.

Wang, D. and Cheng, T. (2001) A spatio-temporal data model for activity-based transport demand modeling. *International Journal of Geographical Information Science*, 15(6):561-585.

Weiner, E. (1992) History of urban transportation planning. In G. E. Gray and Hoel, L. A. (Eds.). *Public Transportation.* Englewood Cliffs, USA: Prentice-Hall.

Wilson, C. (1998a) Activity pattern analysis by means of sequence alignment methods. *Environment and Planning A*, 30:1017-1038.

Wilson, C. (1998b) Analysis of travel behavior using sequence alignment methods. *Transportation Research Record*, No. 164.

Wilson, C. (1999) Sequence alignment analysis of daily activities. In J. Mertz and M. Ehling (Eds.). *Time use – research, data and policy*. Baden-Baden, Germany: Nomos.

Wilson, C., Harvey, A. and Thompson, J. (1999) ClustalG: Software for analysis of activities and sequential events. Workshop on Sequence Alignment Methods, Research Development Initiatives Program, Social Science and Humanities Research Council of Canada, Halifax, October.

Wilson, C. (2001) Activity pattern of Canadian women: Application of ClustalG sequence alignment software. *Transportation Research Record* 1777, No. 01-2292.

Wilson, C. (2008) Activity patterns in space and time: Calculating representative Hagerstrand trajectories. *Transportation*, 35:485-499.

Yu, H. (2006) Spatio-temporal GIS design of an exploring interactions of human activities. *Cartography and Geographic Information Science*, 33(1):3-19.

Yu, H. (2007) Visualizing and analyzing activities in an integrated space–time environment: Temporal GIS design and implementation. *Transportation Research Record: Journal of the Transportation Research Board*, 2024:54-62.

Yu, H. and Shaw, S-L. (2004) Representing and visualizing travel diary data: a spatio-temporal GIS approach. Proceedings of the 2004 ESRI International User Conference.

Yu, H., Shaw, S-L and Bombom, L. S. (2008) Exploring patterns in individual-based activity datasets. Presented at the 104th Annual Meeting of The Association of American Geographers, Boston, MA, April 15-19.

Yuan, M. (1996) Modeling semantics, temporal, and spatial information in geographic information systems. In M. Craglia and Couclelis, H. (Eds.). *Geographic Information Research: Bridging the Atlantic*, London: Taylor and Francis.

Zager, L. A. and Verghese, G. C. (2008) Graph similarity scoring and matching. *Applied Mathematics Letters*, 21:86-94.

Zhou, J. and Golledge, R. (1999) A GPS-based analysis of household travel behavior. Paper Presented at WRSA Annual Meeting, Kauai, Hawaii.

# APPENDICES

## APPENDIX I

**Step-by-step Computation of the Daily Activity Schedule Fragmentation Index (DASFI)** (Frequency-Based)

| S/N | RESPONDENT_ID | v | e | y | d | x | vx | d/vx | y-(d/vx) | DASFI | Cluster |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 152UO18MN | 8 | 7 | 2 | 5 | 3 | 24 | 0.2083 | 1.7917 | 0.2560 | 5 |
| 2 | 39UO55FN | 8 | 7 | 2 | 5 | 3 | 24 | 0.2083 | 1.7917 | 0.2560 | 5 |
| 3 | 89GO45FN | 8 | 7 | 2 | 5 | 3 | 24 | 0.2083 | 1.7917 | 0.2560 | 5 |
| 4 | 122UI18FY | 14 | 13 | 4 | 9 | 5 | 70 | 0.1286 | 3.8714 | 0.2978 | 5 |
| 5 | 84UN18FN | 7 | 6 | 2 | 4 | 3 | 21 | 0.1905 | 1.8095 | 0.3016 | 5 |
| 6 | 133UO18MY | 10 | 9 | 3 | 6 | 4 | 40 | 0.1500 | 2.8500 | 0.3167 | 5 |
| 7 | 140UI18FN | 13 | 12 | 4 | 8 | 5 | 65 | 0.1231 | 3.8769 | 0.3231 | 5 |
| 8 | 120GO18FN | 12 | 11 | 4 | 7 | 5 | 60 | 0.1167 | 3.8833 | 0.3530 | 5 |
| 9 | 151UN18MY | 12 | 11 | 4 | 7 | 5 | 60 | 0.1167 | 3.8833 | 0.3530 | 5 |
| 10 | 134UN18FY | 6 | 5 | 2 | 3 | 3 | 18 | 0.1667 | 1.8333 | 0.3667 | 5 |
| 11 | 15GI35FY | 6 | 5 | 2 | 3 | 3 | 18 | 0.1667 | 1.8333 | 0.3667 | 5 |
| 12 | 81UO18MN | 6 | 5 | 2 | 3 | 3 | 18 | 0.1667 | 1.8333 | 0.3667 | 5 |
| 13 | 121UO18FY | 19 | 18 | 7 | 11 | 8 | 152 | 0.0724 | 6.9276 | 0.3849 | 5 |
| 14 | 148GN30MN | 11 | 10 | 4 | 6 | 5 | 55 | 0.1091 | 3.8909 | 0.3891 | 5 |
| 15 | 16UO18FN | 11 | 10 | 4 | 6 | 5 | 55 | 0.1091 | 3.8909 | 0.3891 | 5 |
| 16 | 130UO18FY | 11 | 10 | 4 | 5 | 5 | 55 | 0.0909 | 3.9091 | 0.3909 | 5 |
| 17 | 153UO18MN | 16 | 15 | 6 | 9 | 7 | 112 | 0.0804 | 5.9196 | 0.3946 | 5 |
| 18 | 47UO18MY | 31 | 30 | 12 | 18 | 13 | 403 | 0.0447 | 11.9553 | 0.3985 | 5 |
| 19 | 112UO40MN | 10 | 9 | 4 | 5 | 5 | 50 | 0.1000 | 3.9000 | 0.4333 | 4 |
| 20 | 116UN18MY | 10 | 9 | 4 | 5 | 5 | 50 | 0.1000 | 3.9000 | 0.4333 | 4 |
| 21 | 40GN30MN | 10 | 9 | 4 | 5 | 5 | 50 | 0.1000 | 3.9000 | 0.4333 | 4 |
| 22 | 131UN18FY | 19 | 18 | 8 | 10 | 9 | 171 | 0.0585 | 7.9415 | 0.4412 | 4 |
| 23 | 129UO18FN | 12 | 11 | 5 | 6 | 6 | 72 | 0.0833 | 4.9167 | 0.4470 | 4 |
| 24 | 124UO45MN | 5 | 4 | 2 | 2 | 3 | 15 | 0.1333 | 1.8667 | 0.4667 | 4 |
| 25 | 26FO45MN | 5 | 4 | 2 | 2 | 3 | 15 | 0.1333 | 1.8667 | 0.4667 | 4 |

DASFI: Frequency-Based

| S/N | RESPONDENT_ID | v | e | y | d | x | vx | d/vx | y-(d/vx) | DASFI | Cluster |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 26 | 34UO18MN | 5 | 4 | 2 | 2 | 3 | 15 | 0.1333 | 1.8667 | 0.4667 | 4 |
| 27 | 7UO18FN | 5 | 4 | 2 | 2 | 3 | 15 | 0.1333 | 1.8667 | 0.4667 | 4 |
| 28 | 98FO35FN | 5 | 4 | 2 | 2 | 3 | 15 | 0.1333 | 1.8667 | 0.4667 | 4 |
| 29 | 104FO50MN | 9 | 8 | 4 | 4 | 5 | 45 | 0.0889 | 3.9111 | 0.4889 | 4 |
| 30 | 139UO18MN | 9 | 8 | 4 | 4 | 5 | 45 | 0.0889 | 3.9111 | 0.4889 | 4 |
| 31 | 141UN18MY | 9 | 8 | 4 | 4 | 5 | 45 | 0.0889 | 3.9111 | 0.4889 | 4 |
| 32 | 144UO25MN | 9 | 8 | 4 | 4 | 5 | 45 | 0.0889 | 3.9111 | 0.4889 | 4 |
| 33 | 58FI35MN | 9 | 8 | 4 | 4 | 5 | 45 | 0.0889 | 3.9111 | 0.4889 | 4 |
| 34 | 27UO18FY | 11 | 10 | 5 | 5 | 6 | 66 | 0.0758 | 4.9242 | 0.4924 | 4 |
| 35 | 74UN18FY | 11 | 10 | 5 | 5 | 6 | 66 | 0.0758 | 4.9242 | 0.4924 | 4 |
| 36 | 123GO18FN | 13 | 12 | 6 | 6 | 7 | 91 | 0.0659 | 5.9341 | 0.4945 | 4 |
| 37 | 61UO18MY | 13 | 12 | 6 | 6 | 7 | 91 | 0.0659 | 5.9341 | 0.4945 | 4 |
| 38 | 149GN25MN | 17 | 16 | 8 | 8 | 9 | 153 | 0.0523 | 7.9477 | 0.4967 | 4 |
| 39 | 110UO18FN | 19 | 18 | 9 | 9 | 10 | 190 | 0.0474 | 8.9526 | 0.4974 | 4 |
| 40 | 17GI25FY | 16 | 15 | 8 | 7 | 9 | 144 | 0.0486 | 7.9514 | 0.5301 | 4 |
| 41 | 62UN18FN | 12 | 11 | 6 | 5 | 7 | 84 | 0.0595 | 5.9405 | 0.5400 | 4 |
| 42 | 79UN18FY | 12 | 11 | 6 | 5 | 7 | 84 | 0.0595 | 5.9405 | 0.5400 | 4 |
| 43 | 96SO55FN | 12 | 11 | 6 | 5 | 7 | 84 | 0.0595 | 5.9405 | 0.5400 | 4 |
| 44 | 103GN55MY | 8 | 7 | 4 | 3 | 5 | 40 | 0.0750 | 3.9250 | 0.5607 | 3 |
| 45 | 147UO18MN | 8 | 7 | 4 | 3 | 5 | 40 | 0.0750 | 3.9250 | 0.5607 | 3 |
| 46 | 28GI25FY | 8 | 7 | 4 | 3 | 5 | 40 | 0.0750 | 3.9250 | 0.5607 | 3 |
| 47 | 44GO35FN | 8 | 7 | 4 | 3 | 5 | 40 | 0.0750 | 3.9250 | 0.5607 | 3 |
| 48 | 60GN35FN | 8 | 7 | 4 | 3 | 5 | 40 | 0.0750 | 3.9250 | 0.5607 | 3 |
| 49 | 97SO55FN | 8 | 7 | 4 | 3 | 5 | 40 | 0.0750 | 3.9250 | 0.5607 | 3 |
| 50 | 2GI25FY | 15 | 14 | 8 | 6 | 9 | 135 | 0.0444 | 7.9556 | 0.5683 | 3 |

DASFI: Frequency-Based

| S/N | RESPONDENT_ID | v | e | y | d | x | vx | d/vx | y-(d/vx) | DASFI | Cluster |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 51 | 46UN18FN | 15 | 14 | 8 | 6 | 9 | 135 | 0.0444 | 7.9556 | 0.5683 | 3 |
| 52 | 91UO18MN | 15 | 14 | 8 | 6 | 9 | 135 | 0.0444 | 7.9556 | 0.5683 | 3 |
| 53 | 93UO18FN | 13 | 12 | 7 | 5 | 8 | 104 | 0.0481 | 6.9519 | 0.5793 | 3 |
| 54 | 108UO18FY | 11 | 10 | 6 | 4 | 7 | 77 | 0.0519 | 5.9481 | 0.5948 | 3 |
| 55 | 99FN30MN | 11 | 10 | 6 | 4 | 7 | 77 | 0.0519 | 5.9481 | 0.5948 | 3 |
| 56 | 127UN18MN | 14 | 13 | 8 | 5 | 9 | 126 | 0.0397 | 7.9603 | 0.6123 | 3 |
| 57 | 59SO40FN | 14 | 13 | 8 | 5 | 9 | 126 | 0.0397 | 7.9603 | 0.6123 | 3 |
| 58 | 70FN30MN | 14 | 13 | 8 | 5 | 9 | 126 | 0.0397 | 7.9603 | 0.6123 | 3 |
| 59 | 83UN18FY | 14 | 13 | 8 | 5 | 9 | 126 | 0.0397 | 7.9603 | 0.6123 | 3 |
| 60 | 92UO18MN | 14 | 13 | 8 | 5 | 9 | 126 | 0.0397 | 7.9603 | 0.6123 | 3 |
| 61 | 31UO18MY | 17 | 16 | 10 | 6 | 11 | 187 | 0.0321 | 9.9679 | 0.6230 | 3 |
| 62 | 29SI30MN | 4 | 3 | 2 | 1 | 3 | 12 | 0.0833 | 1.9167 | 0.6389 | 3 |
| 63 | 41FO60MN | 4 | 3 | 2 | 1 | 3 | 12 | 0.0833 | 1.9167 | 0.6389 | 3 |
| 64 | 88GI45FN | 4 | 3 | 2 | 1 | 3 | 12 | 0.0833 | 1.9167 | 0.6389 | 3 |
| 65 | 115UO25MN | 7 | 6 | 4 | 2 | 5 | 35 | 0.0571 | 3.9429 | 0.6571 | 3 |
| 66 | 20GN18FN | 7 | 6 | 4 | 2 | 5 | 35 | 0.0571 | 3.9429 | 0.6571 | 3 |
| 67 | 75GO55N | 7 | 6 | 4 | 2 | 5 | 35 | 0.0571 | 3.9429 | 0.6571 | 3 |
| 68 | 78UN40MN | 7 | 6 | 4 | 2 | 5 | 35 | 0.0571 | 3.9429 | 0.6571 | 3 |
| 69 | 51GI30FN | 10 | 9 | 6 | 3 | 7 | 70 | 0.0429 | 5.9571 | 0.6619 | 3 |
| 70 | 55UO18FY | 10 | 9 | 6 | 3 | 7 | 70 | 0.0429 | 5.9571 | 0.6619 | 3 |
| 71 | 85GN18FN | 13 | 12 | 8 | 5 | 9 | 117 | 0.0427 | 7.9573 | 0.6631 | 3 |
| 72 | 128UO18MN | 13 | 12 | 8 | 4 | 9 | 117 | 0.0342 | 7.9658 | 0.6638 | 3 |
| 73 | 90UO18MN | 13 | 12 | 8 | 4 | 9 | 117 | 0.0342 | 7.9658 | 0.6638 | 3 |
| 74 | 21UO18FN | 16 | 15 | 10 | 5 | 11 | 176 | 0.0284 | 9.9716 | 0.6648 | 3 |
| 75 | 32GN30MN | 16 | 15 | 10 | 5 | 11 | 176 | 0.0284 | 9.9716 | 0.6648 | 3 |

DASFI: Frequency-Based

| S/N | RESPONDENT_ID | v | e | y | d | x | vx | d/vx | y-(d/vx) | DASFI | Cluster |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 76 | 45UO18FN | 14 | 13 | 9 | 4 | 10 | 140 | 0.0286 | 8.9714 | 0.6901 | 3 |
| 77 | 125UO18MY | 15 | 14 | 10 | 4 | 11 | 165 | 0.0242 | 9.9758 | 0.7126 | 3 |
| 78 | 10UO18FY | 12 | 11 | 8 | 3 | 9 | 108 | 0.0278 | 7.9722 | 0.7247 | 2 |
| 79 | 126GN40MN | 12 | 11 | 8 | 3 | 9 | 108 | 0.0278 | 7.9722 | 0.7247 | 2 |
| 80 | 49GI25FY | 12 | 11 | 8 | 3 | 9 | 108 | 0.0278 | 7.9722 | 0.7247 | 2 |
| 81 | 66UO18FN | 12 | 11 | 8 | 3 | 9 | 108 | 0.0278 | 7.9722 | 0.7247 | 2 |
| 82 | 94SN55FN | 9 | 8 | 6 | 2 | 7 | 63 | 0.0317 | 5.9683 | 0.7460 | 2 |
| 83 | 109FI40MY | 13 | 12 | 9 | 3 | 10 | 130 | 0.0231 | 8.9769 | 0.7481 | 2 |
| 84 | 80UN18MY | 17 | 16 | 12 | 4 | 13 | 221 | 0.0181 | 11.9819 | 0.7489 | 2 |
| 85 | 56UO18FN | 19 | 18 | 14 | 4 | 15 | 285 | 0.0140 | 13.9860 | 0.7770 | 2 |
| 86 | 12UN18FY | 19 | 18 | 14 | 3 | 15 | 285 | 0.0105 | 13.9895 | 0.7772 | 2 |
| 87 | 50GI25FY | 15 | 14 | 11 | 3 | 12 | 180 | 0.0167 | 10.9833 | 0.7845 | 2 |
| 88 | 64GI25FY | 15 | 14 | 11 | 3 | 12 | 180 | 0.0167 | 10.9833 | 0.7845 | 2 |
| 89 | 145GO35MN | 6 | 5 | 4 | 1 | 5 | 30 | 0.0333 | 3.9667 | 0.7933 | 2 |
| 90 | 30SO55FY | 11 | 10 | 8 | 2 | 9 | 99 | 0.0202 | 7.9798 | 0.7980 | 2 |
| 91 | 3FO50MN | 11 | 10 | 8 | 2 | 9 | 99 | 0.0202 | 7.9798 | 0.7980 | 2 |
| 92 | 4UO18FY | 11 | 10 | 8 | 2 | 9 | 99 | 0.0202 | 7.9798 | 0.7980 | 2 |
| 93 | 48UN18FY | 16 | 15 | 12 | 3 | 13 | 208 | 0.0144 | 11.9856 | 0.7990 | 2 |
| 94 | 52UO18FY | 17 | 16 | 13 | 3 | 14 | 238 | 0.0126 | 12.9874 | 0.8117 | 2 |
| 95 | 33UO18FY | 7 | 6 | 5 | 1 | 6 | 42 | 0.0238 | 4.9762 | 0.8294 | 2 |
| 96 | 111UN18FN | 13 | 12 | 10 | 2 | 11 | 143 | 0.0140 | 9.9860 | 0.8322 | 2 |
| 97 | 24GO35MN | 13 | 12 | 10 | 2 | 11 | 143 | 0.0140 | 9.9860 | 0.8322 | 2 |
| 98 | 67UO18FN | 13 | 12 | 10 | 2 | 11 | 143 | 0.0140 | 9.9860 | 0.8322 | 2 |
| 99 | 119UO18FY | 14 | 13 | 11 | 2 | 12 | 168 | 0.0119 | 10.9881 | 0.8452 | 2 |
| 100 | 114UO18MN | 8 | 7 | 6 | 1 | 7 | 56 | 0.0179 | 5.9821 | 0.8546 | 2 |

DASFI: Frequency-Based

| S/N | RESPONDENT_ID | v | e | y | d | x | vx | d/vx | y-(d/vx) | DASFI | Cluster |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 101 | 11UN18FY | 8 | 7 | 6 | 1 | 7 | 56 | 0.0179 | 5.9821 | 0.8546 | 2 |
| 102 | 136GO50MN | 8 | 7 | 6 | 1 | 7 | 56 | 0.0179 | 5.9821 | 0.8546 | 2 |
| 103 | 18FO60FN | 8 | 7 | 6 | 1 | 7 | 56 | 0.0179 | 5.9821 | 0.8546 | 2 |
| 104 | 102SO35MN | 10 | 9 | 8 | 1 | 9 | 90 | 0.0111 | 7.9889 | 0.8877 | 2 |
| 105 | 106FN30MN | 10 | 9 | 8 | 1 | 9 | 90 | 0.0111 | 7.9889 | 0.8877 | 2 |
| 106 | 113UN18FN | 10 | 9 | 8 | 1 | 9 | 90 | 0.0111 | 7.9889 | 0.8877 | 2 |
| 107 | 19GO18FY | 10 | 9 | 8 | 1 | 9 | 90 | 0.0111 | 7.9889 | 0.8877 | 2 |
| 108 | 68FO50MN | 10 | 9 | 8 | 1 | 9 | 90 | 0.0111 | 7.9889 | 0.8877 | 2 |
| 109 | 6UO18MY | 10 | 9 | 8 | 1 | 9 | 90 | 0.0111 | 7.9889 | 0.8877 | 2 |
| 110 | 71FO50MN | 10 | 9 | 8 | 1 | 9 | 90 | 0.0111 | 7.9889 | 0.8877 | 2 |
| 111 | 8GI18FY | 10 | 9 | 8 | 1 | 9 | 90 | 0.0111 | 7.9889 | 0.8877 | 2 |
| 112 | 118UO18MN | 12 | 11 | 10 | 1 | 11 | 132 | 0.0076 | 9.9924 | 0.9084 | 1 |
| 113 | 42GI25MY | 12 | 11 | 10 | 1 | 11 | 132 | 0.0076 | 9.9924 | 0.9084 | 1 |
| 114 | 76FO35FN | 12 | 11 | 10 | 1 | 11 | 132 | 0.0076 | 9.9924 | 0.9084 | 1 |
| 115 | 143GI35MN | 15 | 14 | 13 | 1 | 14 | 210 | 0.0048 | 12.9952 | 0.9282 | 1 |
| 116 | 54GN25MN | 15 | 14 | 13 | 1 | 14 | 210 | 0.0048 | 12.9952 | 0.9282 | 1 |
| 117 | 100GN25FN | 7 | 6 | 6 | 0 | 7 | 49 | 0.0000 | 6.0000 | 1.0000 | 1 |
| 118 | 101GI30FY | 9 | 8 | 8 | 0 | 9 | 81 | 0.0000 | 8.0000 | 1.0000 | 1 |
| 119 | 117GI18MN | 5 | 4 | 4 | 0 | 5 | 25 | 0.0000 | 4.0000 | 1.0000 | 1 |
| 120 | 135GN18MY | 5 | 4 | 4 | 0 | 5 | 25 | 0.0000 | 4.0000 | 1.0000 | 1 |
| 121 | 137UI18FY | 12 | 11 | 11 | 0 | 12 | 144 | 0.0000 | 11.0000 | 1.0000 | 1 |
| 122 | 138GN60MY | 4 | 3 | 3 | 0 | 4 | 16 | 0.0000 | 3.0000 | 1.0000 | 1 |
| 123 | 13GI25MN | 4 | 3 | 3 | 0 | 4 | 16 | 0.0000 | 3.0000 | 1.0000 | 1 |
| 124 | 142FO35FN | 5 | 4 | 4 | 0 | 5 | 25 | 0.0000 | 4.0000 | 1.0000 | 1 |
| 125 | 14UO18MN | 3 | 2 | 2 | 0 | 3 | 9 | 0.0000 | 2.0000 | 1.0000 | 1 |

DASFI: Frequency-Based

| S/N | RESPONDENT_ID | v | e | y | d | x | vx | d/vx | y-(d/vx) | DASFI | Cluster |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 126 | 154UO18MN | 9 | 8 | 8 | 0 | 9 | 81 | 0.0000 | 8.0000 | 1.0000 | 1 |
| 127 | 1GN40MN | 7 | 6 | 6 | 0 | 7 | 49 | 0.0000 | 6.0000 | 1.0000 | 1 |
| 128 | 22GN25FN | 9 | 8 | 8 | 0 | 9 | 81 | 0.0000 | 8.0000 | 1.0000 | 1 |
| 129 | 23FO40MN | 3 | 2 | 2 | 0 | 3 | 9 | 0.0000 | 2.0000 | 1.0000 | 1 |
| 130 | 25SO25FN | 13 | 12 | 12 | 0 | 13 | 169 | 0.0000 | 12.0000 | 1.0000 | 1 |
| 131 | 35FN35MN | 5 | 4 | 4 | 0 | 5 | 25 | 0.0000 | 4.0000 | 1.0000 | 1 |
| 132 | 36SO60MN | 5 | 4 | 4 | 0 | 5 | 25 | 0.0000 | 4.0000 | 1.0000 | 1 |
| 133 | 37UO18FN | 4 | 3 | 3 | 0 | 4 | 16 | 0.0000 | 3.0000 | 1.0000 | 1 |
| 134 | 38UN18FY | 13 | 12 | 12 | 0 | 13 | 169 | 0.0000 | 12.0000 | 1.0000 | 1 |
| 135 | 43GN30MN | 9 | 8 | 8 | 0 | 9 | 81 | 0.0000 | 8.0000 | 1.0000 | 1 |
| 136 | 57SO45FN | 5 | 4 | 4 | 0 | 5 | 25 | 0.0000 | 4.0000 | 1.0000 | 1 |
| 137 | 5GO30MN | 9 | 8 | 8 | 0 | 9 | 81 | 0.0000 | 8.0000 | 1.0000 | 1 |
| 138 | 63SO45FN | 5 | 4 | 4 | 0 | 5 | 25 | 0.0000 | 4.0000 | 1.0000 | 1 |
| 139 | 65GI30FN | 7 | 6 | 6 | 0 | 7 | 49 | 0.0000 | 6.0000 | 1.0000 | 1 |
| 140 | 69SO25FN | 13 | 12 | 12 | 0 | 13 | 169 | 0.0000 | 12.0000 | 1.0000 | 1 |
| 141 | 72SN60MN | 5 | 4 | 4 | 0 | 5 | 25 | 0.0000 | 4.0000 | 1.0000 | 1 |
| 142 | 73GN18FN | 5 | 4 | 4 | 0 | 5 | 25 | 0.0000 | 4.0000 | 1.0000 | 1 |
| 143 | 77UO18FN | 14 | 13 | 13 | 0 | 14 | 196 | 0.0000 | 13.0000 | 1.0000 | 1 |
| 144 | 82UN18FN | 3 | 2 | 2 | 0 | 3 | 9 | 0.0000 | 2.0000 | 1.0000 | 1 |
| 145 | 87GO18FN | 3 | 2 | 2 | 0 | 3 | 9 | 0.0000 | 2.0000 | 1.0000 | 1 |
| 146 | 95FO40MN | 5 | 4 | 4 | 0 | 5 | 25 | 0.0000 | 4.0000 | 1.0000 | 1 |
| 147 | 9SO45FN | 5 | 4 | 4 | 0 | 5 | 25 | 0.0000 | 4.0000 | 1.0000 | 1 |

DASFI: Duration-Based

| **S/N** | RESPONDENT_ID | **v** | **e** | **y** | **d** | **x** | **vx** | **d/vx** | **y-(d/vx)** | **DASFI** | **Cluster** |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 106FN30MN | 10 | 9 | 2 | 7 | 3 | 30 | 0.2333 | 1.7667 | 0.1963 | 5 |
| 2 | 152UO18MN | 8 | 7 | 2 | 5 | 3 | 24 | 0.2083 | 1.7917 | 0.2560 | 5 |
| 3 | 39UO55FN | 8 | 7 | 2 | 5 | 3 | 24 | 0.2083 | 1.7917 | 0.2560 | 5 |
| 4 | 89GO45FN | 8 | 7 | 2 | 5 | 3 | 24 | 0.2083 | 1.7917 | 0.2560 | 5 |
| 5 | 17GI25FY | 16 | 15 | 4 | 11 | 5 | 80 | 0.1375 | 3.8625 | 0.2575 | 5 |
| 6 | 122UI18FY | 14 | 13 | 4 | 9 | 5 | 70 | 0.1286 | 3.8714 | 0.2978 | 5 |
| 7 | 59SO40FN | 14 | 13 | 4 | 9 | 5 | 70 | 0.1286 | 3.8714 | 0.2978 | 5 |
| 8 | 70FN30MN | 14 | 13 | 4 | 9 | 5 | 70 | 0.1286 | 3.8714 | 0.2978 | 5 |
| 9 | 92UO18MN | 14 | 13 | 4 | 9 | 5 | 70 | 0.1286 | 3.8714 | 0.2978 | 5 |
| 10 | 1GN40MN | 7 | 6 | 2 | 4 | 3 | 21 | 0.1905 | 1.8095 | 0.3016 | 5 |
| 11 | 133UO18MY | 10 | 9 | 3 | 6 | 4 | 40 | 0.1500 | 2.8500 | 0.3167 | 5 |
| 12 | 128UO18MN | 13 | 12 | 4 | 8 | 5 | 65 | 0.1231 | 3.8769 | 0.3231 | 5 |
| 13 | 140UI18FN | 13 | 12 | 4 | 8 | 5 | 65 | 0.1231 | 3.8769 | 0.3231 | 5 |
| 14 | 24GO35MN | 13 | 12 | 4 | 8 | 5 | 65 | 0.1231 | 3.8769 | 0.3231 | 5 |
| 15 | 85GN18FN | 13 | 12 | 4 | 8 | 5 | 65 | 0.1231 | 3.8769 | 0.3231 | 5 |
| 16 | 120GO18FN | 12 | 11 | 4 | 7 | 5 | 60 | 0.1167 | 3.8833 | 0.3530 | 5 |
| 17 | 126GN40MN | 12 | 11 | 4 | 7 | 5 | 60 | 0.1167 | 3.8833 | 0.3530 | 5 |
| 18 | 151UN18MY | 12 | 11 | 4 | 7 | 5 | 60 | 0.1167 | 3.8833 | 0.3530 | 5 |
| 19 | 134UN18FY | 6 | 5 | 2 | 3 | 3 | 18 | 0.1667 | 1.8333 | 0.3667 | 5 |
| 20 | 15GI35FY | 6 | 5 | 2 | 3 | 3 | 18 | 0.1667 | 1.8333 | 0.3667 | 5 |
| 21 | 81UO18MN | 6 | 5 | 2 | 3 | 3 | 18 | 0.1667 | 1.8333 | 0.3667 | 5 |
| 22 | 121UO18FY | 19 | 18 | 7 | 11 | 8 | 152 | 0.0724 | 6.9276 | 0.3849 | 4 |
| 23 | 130UO18FY | 11 | 10 | 4 | 6 | 5 | 55 | 0.1091 | 3.8909 | 0.3891 | 4 |
| 24 | 148GN30MN | 11 | 10 | 4 | 6 | 5 | 55 | 0.1091 | 3.8909 | 0.3891 | 4 |
| 25 | 16UO18FN | 11 | 10 | 4 | 6 | 5 | 55 | 0.1091 | 3.8909 | 0.3891 | 4 |

DASFI: Duration-Based

| S/N | RESPONDENT_ID | v | e | y | d | x | vx | d/vx | y-(d/vx) | DASFI | Cluster |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 26 | 153UO18MN | 16 | 15 | 6 | 9 | 7 | 112 | 0.0804 | 5.9196 | 0.3946 | 4 |
| 27 | 47UO18MY | 31 | 30 | 12 | 18 | 13 | 403 | 0.0447 | 11.9553 | 0.3985 | 4 |
| 28 | 102SO35MN | 10 | 9 | 4 | 5 | 5 | 50 | 0.1000 | 3.9000 | 0.4333 | 4 |
| 29 | 112UO40MN | 10 | 9 | 4 | 5 | 5 | 50 | 0.1000 | 3.9000 | 0.4333 | 4 |
| 30 | 116UN18MY | 10 | 9 | 4 | 5 | 5 | 50 | 0.1000 | 3.9000 | 0.4333 | 4 |
| 31 | 40GN30MN | 10 | 9 | 4 | 5 | 5 | 50 | 0.1000 | 3.9000 | 0.4333 | 4 |
| 32 | 71FO50MN | 10 | 9 | 4 | 5 | 5 | 50 | 0.1000 | 3.9000 | 0.4333 | 4 |
| 33 | 131UN18FY | 19 | 18 | 8 | 10 | 9 | 171 | 0.0585 | 7.9415 | 0.4412 | 4 |
| 34 | 129UO18FN | 12 | 11 | 5 | 6 | 6 | 72 | 0.0833 | 4.9167 | 0.4470 | 4 |
| 35 | 32GN30MN | 16 | 15 | 7 | 7 | 8 | 128 | 0.0547 | 6.9453 | 0.4630 | 4 |
| 36 | 124UO45MN | 5 | 4 | 2 | 2 | 3 | 15 | 0.1333 | 1.8667 | 0.4667 | 4 |
| 37 | 26FO45MN | 5 | 4 | 2 | 2 | 3 | 15 | 0.1333 | 1.8667 | 0.4667 | 4 |
| 38 | 34UO18MN | 5 | 4 | 2 | 2 | 3 | 15 | 0.1333 | 1.8667 | 0.4667 | 4 |
| 39 | 7UO18FN | 5 | 4 | 2 | 2 | 3 | 15 | 0.1333 | 1.8667 | 0.4667 | 4 |
| 40 | 98FO35FN | 5 | 4 | 2 | 2 | 3 | 15 | 0.1333 | 1.8667 | 0.4667 | 4 |
| 41 | 104FO50MN | 9 | 8 | 4 | 4 | 5 | 45 | 0.0889 | 3.9111 | 0.4889 | 4 |
| 42 | 139UO18MN | 9 | 8 | 4 | 4 | 5 | 45 | 0.0889 | 3.9111 | 0.4889 | 4 |
| 43 | 141UN18MY | 9 | 8 | 4 | 4 | 5 | 45 | 0.0889 | 3.9111 | 0.4889 | 4 |
| 44 | 144UO25MN | 9 | 8 | 4 | 4 | 5 | 45 | 0.0889 | 3.9111 | 0.4889 | 4 |
| 45 | 43GN30MN | 9 | 8 | 4 | 4 | 5 | 45 | 0.0889 | 3.9111 | 0.4889 | 4 |
| 46 | 58FI35MN | 9 | 8 | 4 | 4 | 5 | 45 | 0.0889 | 3.9111 | 0.4889 | 4 |
| 47 | 5GO30MN | 9 | 8 | 4 | 4 | 5 | 45 | 0.0889 | 3.9111 | 0.4889 | 4 |
| 48 | 27UO18FY | 11 | 10 | 5 | 5 | 6 | 66 | 0.0758 | 4.9242 | 0.4924 | 4 |
| 49 | 74UN18FY | 11 | 10 | 5 | 5 | 6 | 66 | 0.0758 | 4.9242 | 0.4924 | 4 |
| 50 | 123GO18FN | 13 | 12 | 6 | 6 | 7 | 91 | 0.0659 | 5.9341 | 0.4945 | 4 |

DASFI: Duration-Based

| S/N | RESPONDENT_ID | v | e | y | d | x | vx | d/vx | y-(d/vx) | DASFI | Cluster |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 51 | 61UO18MY | 13 | 12 | 6 | 6 | 7 | 91 | 0.0659 | 5.9341 | 0.4945 | 4 |
| 52 | 149GN25MN | 17 | 16 | 8 | 8 | 9 | 153 | 0.0523 | 7.9477 | 0.4967 | 4 |
| 53 | 31UO18MY | 17 | 16 | 8 | 8 | 9 | 153 | 0.0523 | 7.9477 | 0.4967 | 4 |
| 54 | 110UO18FN | 19 | 18 | 9 | 9 | 10 | 190 | 0.0474 | 8.9526 | 0.4974 | 4 |
| 55 | 118UO18MN | 12 | 11 | 6 | 5 | 7 | 84 | 0.0595 | 5.9405 | 0.5400 | 3 |
| 56 | 62UN18FN | 12 | 11 | 6 | 5 | 7 | 84 | 0.0595 | 5.9405 | 0.5400 | 3 |
| 57 | 76FO35FN | 12 | 11 | 6 | 5 | 7 | 84 | 0.0595 | 5.9405 | 0.5400 | 3 |
| 58 | 79UN18FY | 12 | 11 | 6 | 5 | 7 | 84 | 0.0595 | 5.9405 | 0.5400 | 3 |
| 59 | 96SO55FN | 12 | 11 | 6 | 5 | 7 | 84 | 0.0595 | 5.9405 | 0.5400 | 3 |
| 60 | 103GN55MY | 8 | 7 | 4 | 3 | 5 | 40 | 0.0750 | 3.9250 | 0.5607 | 3 |
| 61 | 147UO18MN | 8 | 7 | 4 | 3 | 5 | 40 | 0.0750 | 3.9250 | 0.5607 | 3 |
| 62 | 44GO35FN | 8 | 7 | 4 | 3 | 5 | 40 | 0.0750 | 3.9250 | 0.5607 | 3 |
| 63 | 60GN35FN | 8 | 7 | 4 | 3 | 5 | 40 | 0.0750 | 3.9250 | 0.5607 | 3 |
| 64 | 97SO55FN | 8 | 7 | 4 | 3 | 5 | 40 | 0.0750 | 3.9250 | 0.5607 | 3 |
| 65 | 2GI25FY | 15 | 14 | 8 | 6 | 9 | 135 | 0.0444 | 7.9556 | 0.5683 | 3 |
| 66 | 46UN18FN | 15 | 14 | 8 | 6 | 9 | 135 | 0.0444 | 7.9556 | 0.5683 | 3 |
| 67 | 91UO18MN | 15 | 14 | 8 | 6 | 9 | 135 | 0.0444 | 7.9556 | 0.5683 | 3 |
| 68 | 93UO18FN | 13 | 12 | 7 | 5 | 8 | 104 | 0.0481 | 6.9519 | 0.5793 | 3 |
| 69 | 108UO18FY | 11 | 10 | 6 | 4 | 7 | 77 | 0.0519 | 5.9481 | 0.5948 | 3 |
| 70 | 30SO55FY | 11 | 10 | 6 | 4 | 7 | 77 | 0.0519 | 5.9481 | 0.5948 | 3 |
| 71 | 99FN30MN | 11 | 10 | 6 | 4 | 7 | 77 | 0.0519 | 5.9481 | 0.5948 | 3 |
| 72 | 127UN18MN | 14 | 13 | 8 | 5 | 9 | 126 | 0.0397 | 7.9603 | 0.6123 | 3 |
| 73 | 83UN18FY | 14 | 13 | 8 | 5 | 9 | 126 | 0.0397 | 7.9603 | 0.6123 | 3 |
| 74 | 29SI30MN | 4 | 3 | 2 | 1 | 3 | 12 | 0.0833 | 1.9167 | 0.6389 | 3 |
| 75 | 41FO60MN | 4 | 3 | 2 | 1 | 3 | 12 | 0.0833 | 1.9167 | 0.6389 | 3 |

DASFI: Duration-Based

| S/N | RESPONDENT_ID | v | e | y | d | x | vx | d/vx | y-(d/vx) | DASFI | Cluster |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 76 | 88GI45FN | 4 | 3 | 2 | 1 | 3 | 12 | 0.0833 | 1.9167 | 0.6389 | 3 |
| 77 | 115UO25MN | 7 | 6 | 4 | 2 | 5 | 35 | 0.0571 | 3.9429 | 0.6571 | 3 |
| 78 | 20GN18FN | 7 | 6 | 4 | 2 | 5 | 35 | 0.0571 | 3.9429 | 0.6571 | 3 |
| 79 | 78UN40MN | 7 | 6 | 4 | 2 | 5 | 35 | 0.0571 | 3.9429 | 0.6571 | 3 |
| 80 | 84UN18FN | 7 | 6 | 4 | 2 | 5 | 35 | 0.0571 | 3.9429 | 0.6571 | 3 |
| 81 | 51GI30FN | 10 | 9 | 6 | 3 | 7 | 70 | 0.0429 | 5.9571 | 0.6619 | 3 |
| 82 | 55UO18FY | 10 | 9 | 6 | 3 | 7 | 70 | 0.0429 | 5.9571 | 0.6619 | 3 |
| 83 | 90UO18MN | 13 | 12 | 8 | 4 | 9 | 117 | 0.0342 | 7.9658 | 0.6638 | 3 |
| 84 | 21UO18FN | 16 | 15 | 10 | 5 | 11 | 176 | 0.0284 | 9.9716 | 0.6648 | 3 |
| 85 | 45UO18FN | 14 | 13 | 9 | 4 | 10 | 140 | 0.0286 | 8.9714 | 0.6901 | 3 |
| 86 | 125UO18MY | 15 | 14 | 10 | 4 | 11 | 165 | 0.0242 | 9.9758 | 0.7126 | 2 |
| 87 | 10UO18FY | 12 | 11 | 8 | 3 | 9 | 108 | 0.0278 | 7.9722 | 0.7247 | 2 |
| 88 | 49GI25FY | 12 | 11 | 8 | 3 | 9 | 108 | 0.0278 | 7.9722 | 0.7247 | 2 |
| 89 | 66UO18FN | 12 | 11 | 8 | 3 | 9 | 108 | 0.0278 | 7.9722 | 0.7247 | 2 |
| 90 | 94SN55FN | 9 | 8 | 6 | 2 | 7 | 63 | 0.0317 | 5.9683 | 0.7460 | 2 |
| 91 | 109FI40MY | 13 | 12 | 9 | 3 | 10 | 130 | 0.0231 | 8.9769 | 0.7481 | 2 |
| 92 | 80UN18MY | 17 | 16 | 12 | 4 | 13 | 221 | 0.0181 | 11.9819 | 0.7489 | 2 |
| 93 | 56UO18FN | 19 | 18 | 14 | 4 | 15 | 285 | 0.0140 | 13.9860 | 0.7770 | 2 |
| 94 | 12UN18FY | 19 | 18 | 14 | 3 | 15 | 285 | 0.0105 | 13.9895 | 0.7772 | 2 |
| 95 | 50GI25FY | 15 | 14 | 11 | 3 | 12 | 180 | 0.0167 | 10.9833 | 0.7845 | 2 |
| 96 | 64GI25FY | 15 | 14 | 11 | 3 | 12 | 180 | 0.0167 | 10.9833 | 0.7845 | 2 |
| 97 | 145GO35MN | 6 | 5 | 4 | 1 | 5 | 30 | 0.0333 | 3.9667 | 0.7933 | 2 |
| 98 | 3FO50MN | 11 | 10 | 8 | 2 | 9 | 99 | 0.0202 | 7.9798 | 0.7980 | 2 |
| 99 | 4UO18FY | 11 | 10 | 8 | 2 | 9 | 99 | 0.0202 | 7.9798 | 0.7980 | 2 |
| 100 | 48UN18FY | 16 | 15 | 12 | 3 | 13 | 208 | 0.0144 | 11.9856 | 0.7990 | 2 |

DASFI: Duration-Based

| S/N | RESPONDENT_ID | v | e | y | d | x | vx | d/vx | y-(d/vx) | DASFI | Cluster |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 101 | 52UO18FY | 17 | 16 | 13 | 3 | 14 | 238 | 0.0126 | 12.9874 | 0.8117 | 2 |
| 102 | 33UO18FY | 7 | 6 | 5 | 1 | 6 | 42 | 0.0238 | 4.9762 | 0.8294 | 2 |
| 103 | 111UN18FN | 13 | 12 | 10 | 2 | 11 | 143 | 0.0140 | 9.9860 | 0.8322 | 2 |
| 104 | 67UO18FN | 13 | 12 | 10 | 2 | 11 | 143 | 0.0140 | 9.9860 | 0.8322 | 2 |
| 105 | 119UO18FY | 14 | 13 | 11 | 2 | 12 | 168 | 0.0119 | 10.9881 | 0.8452 | 2 |
| 106 | 114UO18MN | 8 | 7 | 6 | 1 | 7 | 56 | 0.0179 | 5.9821 | 0.8546 | 2 |
| 107 | 11UN18FY | 8 | 7 | 6 | 1 | 7 | 56 | 0.0179 | 5.9821 | 0.8546 | 2 |
| 108 | 136GO50MN | 8 | 7 | 6 | 1 | 7 | 56 | 0.0179 | 5.9821 | 0.8546 | 2 |
| 109 | 18FO60FN | 8 | 7 | 6 | 1 | 7 | 56 | 0.0179 | 5.9821 | 0.8546 | 2 |
| 110 | 28GI25FY | 8 | 7 | 6 | 1 | 7 | 56 | 0.0179 | 5.9821 | 0.8546 | 2 |
| 111 | 113UN18FN | 10 | 9 | 8 | 1 | 9 | 90 | 0.0111 | 7.9889 | 0.8877 | 2 |
| 112 | 19GO18FY | 10 | 9 | 8 | 1 | 9 | 90 | 0.0111 | 7.9889 | 0.8877 | 2 |
| 113 | 68FO50MN | 10 | 9 | 8 | 1 | 9 | 90 | 0.0111 | 7.9889 | 0.8877 | 2 |
| 114 | 6UO18MY | 10 | 9 | 8 | 1 | 9 | 90 | 0.0111 | 7.9889 | 0.8877 | 2 |
| 115 | 8GI18FY | 10 | 9 | 8 | 1 | 9 | 90 | 0.0111 | 7.9889 | 0.8877 | 2 |
| 116 | 42GI25MY | 12 | 11 | 10 | 1 | 11 | 132 | 0.0076 | 9.9924 | 0.9084 | 1 |
| 117 | 143GI35MN | 15 | 14 | 13 | 1 | 14 | 210 | 0.0048 | 12.9952 | 0.9282 | 1 |
| 118 | 54GN25MN | 15 | 14 | 13 | 1 | 14 | 210 | 0.0048 | 12.9952 | 0.9282 | 1 |
| 119 | 100GN25FN | 7 | 6 | 6 | 0 | 7 | 49 | 0.0000 | 6.0000 | 1.0000 | 1 |
| 120 | 101GI30FY | 9 | 8 | 8 | 0 | 9 | 81 | 0.0000 | 8.0000 | 1.0000 | 1 |
| 121 | 117GI18MN | 5 | 4 | 4 | 0 | 5 | 25 | 0.0000 | 4.0000 | 1.0000 | 1 |
| 122 | 135GN18MY | 5 | 4 | 4 | 0 | 5 | 25 | 0.0000 | 4.0000 | 1.0000 | 1 |
| 123 | 137UI18FY | 12 | 11 | 11 | 0 | 12 | 144 | 0.0000 | 11.0000 | 1.0000 | 1 |
| 124 | 138GN60MY | 4 | 3 | 3 | 0 | 4 | 16 | 0.0000 | 3.0000 | 1.0000 | 1 |
| 125 | 13GI25MN | 4 | 3 | 3 | 0 | 4 | 16 | 0.0000 | 3.0000 | 1.0000 | 1 |

DASFI: Duration-Based

| S/N | RESPONDENT_ID | v | e | y | d | x | vx | d/vx | y-(d/vx) | DASFI | Cluster |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 126 | 142FO35FN | 5 | 4 | 4 | 0 | 5 | 25 | 0.0000 | 4.0000 | 1.0000 | 1 |
| 127 | 14UO18MN | 3 | 2 | 2 | 0 | 3 | 9 | 0.0000 | 2.0000 | 1.0000 | 1 |
| 128 | 154UO18MN | 9 | 8 | 8 | 0 | 9 | 81 | 0.0000 | 8.0000 | 1.0000 | 1 |
| 129 | 22GN25FN | 9 | 8 | 8 | 0 | 9 | 81 | 0.0000 | 8.0000 | 1.0000 | 1 |
| 130 | 23FO40MN | 3 | 2 | 2 | 0 | 3 | 9 | 0.0000 | 2.0000 | 1.0000 | 1 |
| 131 | 25SO25FN | 13 | 12 | 12 | 0 | 13 | 169 | 0.0000 | 12.0000 | 1.0000 | 1 |
| 132 | 35FN35MN | 5 | 4 | 4 | 0 | 5 | 25 | 0.0000 | 4.0000 | 1.0000 | 1 |
| 133 | 36SO60MN | 5 | 4 | 4 | 0 | 5 | 25 | 0.0000 | 4.0000 | 1.0000 | 1 |
| 134 | 37UO18FN | 4 | 3 | 3 | 0 | 4 | 16 | 0.0000 | 3.0000 | 1.0000 | 1 |
| 135 | 38UN18FY | 13 | 12 | 12 | 0 | 13 | 169 | 0.0000 | 12.0000 | 1.0000 | 1 |
| 136 | 57SO45FN | 5 | 4 | 4 | 0 | 5 | 25 | 0.0000 | 4.0000 | 1.0000 | 1 |
| 137 | 63SO45FN | 5 | 4 | 4 | 0 | 5 | 25 | 0.0000 | 4.0000 | 1.0000 | 1 |
| 138 | 65GI30FN | 7 | 6 | 6 | 0 | 7 | 49 | 0.0000 | 6.0000 | 1.0000 | 1 |
| 139 | 69SO25FN | 13 | 12 | 12 | 0 | 13 | 169 | 0.0000 | 12.0000 | 1.0000 | 1 |
| 140 | 72SN60MN | 5 | 4 | 4 | 0 | 5 | 25 | 0.0000 | 4.0000 | 1.0000 | 1 |
| 141 | 73GN18FN | 5 | 4 | 4 | 0 | 5 | 25 | 0.0000 | 4.0000 | 1.0000 | 1 |
| 142 | 75GO55N | 7 | 6 | 6 | 0 | 7 | 49 | 0.0000 | 6.0000 | 1.0000 | 1 |
| 143 | 77UO18FN | 14 | 13 | 13 | 0 | 14 | 196 | 0.0000 | 13.0000 | 1.0000 | 1 |
| 144 | 82UN18FN | 3 | 2 | 2 | 0 | 3 | 9 | 0.0000 | 2.0000 | 1.0000 | 1 |
| 145 | 87GO18FN | 3 | 2 | 2 | 0 | 3 | 9 | 0.0000 | 2.0000 | 1.0000 | 1 |
| 146 | 95FO40MN | 5 | 4 | 4 | 0 | 5 | 25 | 0.0000 | 4.0000 | 1.0000 | 1 |
| 147 | 9SO45FN | 5 | 4 | 4 | 0 | 5 | 25 | 0.0000 | 4.0000 | 1.0000 | 1 |

# APPENDIX II

**Step-by-step Computation of the Daily Activity Intensity Similarity Index (DAISI)**

SECTION A: Activity Frequencies and Activity rates

| Activity Frequency | 8:00 - 8:59 | 9:00 - 9:59 | 10:00 - 10:59 | 11:00 - 11:59 | 12:00 - 12:59 |
|---|---|---|---|---|---|
| A | 1 | 1 | 5 | 1 | 1 |
| B | 1 | 1 | 5 | 1 | 1 |
| C | 3 | 3 | 8 | 3 | 3 |
| D | 3 | 3 | 8 | 3 | 3 |
| E | 5 | 3 | 6 | 4 | 2 |
| F | 9 | 9 | 8 | 9 | 9 |
| Activity Rates | 8:00 - 8:59 | 9:00 - 9:59 | 10:00 - 10:59 | 11:00 - 11:59 | 12:00 - 12:59 |
| A | 0.0167 | 0.0167 | 0.0833 | 0.0167 | 0.0167 |
| B | 0.0167 | 0.0167 | 0.0833 | 0.0167 | 0.0167 |
| C | 0.0500 | 0.0500 | 0.1333 | 0.0500 | 0.0500 |
| D | 0.0500 | 0.0500 | 0.1333 | 0.0500 | 0.0500 |
| E | 0.0833 | 0.0500 | 0.1000 | 0.0667 | 0.0333 |
| F | 0.1500 | 0.1500 | 0.1333 | 0.1500 | 0.1500 |

SECTION B: Calculating *rd*

| *Indv A* | 8:00 - 8:59 | 9:00 - 9:59 | 10:00 - 10:59 | 11:00 - 11:59 | 12:00 - 12:59 | SUM | AVG (*rd*) |
|---|---|---|---|---|---|---|---|
| A | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| B | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| C | 1 | 1 | 1 | 1 | 1 | 5 | 1 |
| D | 1 | 1 | 1 | 1 | 1 | 5 | 1 |
| E | 1 | 1 | 1 | 1 | 1 | 5 | 1 |
| F | 1 | 1 | 1 | 1 | 1 | 5 | 1 |
| *Indv B* | 8:00 - 8:59 | 9:00 - 9:59 | 10:00 - 10:59 | 11:00 - 11:59 | 12:00 - 12:59 | SUM | AVG (*rd*) |
| A | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| B | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| C | 1 | 1 | 1 | 1 | 1 | 5 | 1 |
| D | 1 | 1 | 1 | 1 | 1 | 5 | 1 |
| E | 1 | 1 | 1 | 1 | 1 | 5 | 1 |
| F | 1 | 1 | 1 | 1 | 1 | 5 | 1 |
| *Indv C* | 8:00 - 8:59 | 9:00 - 9:59 | 10:00 - 10:59 | 11:00 - 11:59 | 12:00 - 12:59 | SUM | AVG (*rd*) |
| A | 1 | 1 | 1 | 1 | 1 | 5 | 1 |
| B | 1 | 1 | 1 | 1 | 1 | 5 | 1 |
| C | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| D | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| E | 1 | 0 | 1 | 1 | 1 | 4 | 0.8 |
| F | 1 | 1 | 0 | 1 | 1 | 4 | 0.8 |

| *Indv D* | 8:00 - 8:59 | 9:00 - 9:59 | 10:00 - 10:59 | 11:00 - 11:59 | 12:00 - 12:59 | SUM | AVG (*rd*) |
|---|---|---|---|---|---|---|---|
| A | 1 | 1 | 1 | 1 | 1 | 5 | 1 |
| B | 1 | 1 | 1 | 1 | 1 | 5 | 1 |
| C | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| D | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| E | 1 | 0 | 1 | 1 | 1 | 4 | 0.8 |
| F | 1 | 1 | 0 | 1 | 1 | 4 | 0.8 |
| *Ind E* | 8:00 - 8:59 | 9:00 - 9:59 | 10:00 - 10:59 | 11:00 - 11:59 | 12:00 - 12:59 | SUM | AVG (*rd*) |
| A | 1 | 1 | 1 | 1 | 1 | 5 | 1 |
| B | 1 | 1 | 1 | 1 | 1 | 5 | 1 |
| C | 1 | 0 | 1 | 1 | 1 | 4 | 0.8 |
| D | 1 | 0 | 1 | 1 | 1 | 4 | 0.8 |
| E | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| F | 1 | 1 | 1 | 1 | 1 | 5 | 1 |
| *Ind F* | 8:00 - 8:59 | 9:00 - 9:59 | 10:00 - 10:59 | 11:00 - 11:59 | 12:00 - 12:59 | SUM | AVG (*rd*) |
| A | 1 | 1 | 1 | 1 | 1 | 5 | 1 |
| B | 1 | 1 | 1 | 1 | 1 | 5 | 1 |
| C | 1 | 1 | 0 | 1 | 1 | 4 | 0.8 |
| D | 1 | 1 | 0 | 1 | 1 | 4 | 0.8 |
| E | 1 | 1 | 1 | 1 | 1 | 5 | 1 |
| F | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

SECTION C: Calculating *dd*

| *Indv A* | 8:00 - 8:59 | 9:00 - 9:59 | 10:00 - 10:59 | 11:00 - 11:59 | 12:00 - 12:59 | SUM | AVG (*dd*) |
|---|---|---|---|---|---|---|---|
| A | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| B | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| C | 0.0333 | 0.0333 | 0.0500 | 0.0333 | 0.0333 | 0.1833 | 0.0367 |
| D | 0.0333 | 0.0333 | 0.0500 | 0.0333 | 0.0333 | 0.1833 | 0.0367 |
| E | 0.0667 | 0.0333 | 0.0167 | 0.0500 | 0.0167 | 0.1833 | 0.0367 |
| F | 0.1333 | 0.1333 | 0.0500 | 0.1333 | 0.1333 | 0.5833 | 0.1167 |
| *Indv B* | 8:00 - 8:59 | 9:00 - 9:59 | 10:00 - 10:59 | 11:00 - 11:59 | 12:00 - 12:59 | SUM | AVG (*dd*) |
| A | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| B | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| C | 0.0333 | 0.0333 | 0.0500 | 0.0333 | 0.0333 | 0.1833 | 0.0367 |
| D | 0.0333 | 0.0333 | 0.0500 | 0.0333 | 0.0333 | 0.1833 | 0.0367 |
| E | 0.0667 | 0.0333 | 0.0167 | 0.0500 | 0.0167 | 0.1833 | 0.0367 |
| F | 0.1333 | 0.1333 | 0.0500 | 0.1333 | 0.1333 | 0.5833 | 0.1167 |
| *Indv C* | 8:00 - 8:59 | 9:00 - 9:59 | 10:00 - 10:59 | 11:00 - 11:59 | 12:00 - 12:59 | SUM | AVG (*dd*) |
| A | 0.0333 | 0.0333 | 0.0500 | 0.0333 | 0.0333 | 0.1833 | 0.0367 |
| B | 0.0333 | 0.0333 | 0.0500 | 0.0333 | 0.0333 | 0.1833 | 0.0367 |
| C | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| D | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| E | 0.0333 | 0.0000 | 0.0333 | 0.0167 | 0.0167 | 0.1000 | 0.0200 |
| F | 0.1000 | 0.1000 | 0.0000 | 0.1000 | 0.1000 | 0.4000 | 0.0800 |

| *Indv D* | 8:00 - 8:59 | 9:00 - 9:59 | 10:00 - 10:59 | 11:00 - 11:59 | 12:00 - 12:59 | SUM | AVG (*dd*) |
|---|---|---|---|---|---|---|---|
| A | 0.0333 | 0.0333 | 0.0500 | 0.0333 | 0.0333 | 0.1833 | 0.0367 |
| B | 0.0333 | 0.0333 | 0.0500 | 0.0333 | 0.0333 | 0.1833 | 0.0367 |
| C | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| D | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| E | 0.0333 | 0.0000 | 0.0333 | 0.0167 | 0.0167 | 0.1000 | 0.0200 |
| F | 0.1000 | 0.1000 | 0.0000 | 0.1000 | 0.1000 | 0.4000 | 0.0800 |
| *Ind E* | 8:00 - 8:59 | 9:00 - 9:59 | 10:00 - 10:59 | 11:00 - 11:59 | 12:00 - 12:59 | SUM | AVG (*dd*) |
| A | 0.0667 | 0.0333 | 0.0167 | 0.0500 | 0.0167 | 0.1833 | 0.0367 |
| B | 0.0667 | 0.0333 | 0.0167 | 0.0500 | 0.0167 | 0.1833 | 0.0367 |
| C | 0.0333 | 0.0000 | 0.0333 | 0.0167 | 0.0167 | 0.1000 | 0.0200 |
| D | 0.0333 | 0.0000 | 0.0333 | 0.0167 | 0.0167 | 0.1000 | 0.0200 |
| E | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| F | 0.0667 | 0.1000 | 0.0333 | 0.0833 | 0.1167 | 0.4000 | 0.0800 |
| *Ind F* | 8:00 - 8:59 | 9:00 - 9:59 | 10:00 - 10:59 | 11:00 - 11:59 | 12:00 - 12:59 | SUM | AVG (*dd*) |
| A | 0.1333 | 0.1333 | 0.0500 | 0.1333 | 0.1333 | 0.5833 | 0.1167 |
| B | 0.1333 | 0.1333 | 0.0500 | 0.1333 | 0.1333 | 0.5833 | 0.1167 |
| C | 0.1000 | 0.1000 | 0.0000 | 0.1000 | 0.1000 | 0.4000 | 0.0800 |
| D | 0.1000 | 0.1000 | 0.0000 | 0.1000 | 0.1000 | 0.4000 | 0.0800 |
| E | 0.0667 | 0.1000 | 0.0333 | 0.0833 | 0.1167 | 0.4000 | 0.0800 |
| F | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |

SECTION D: Matrix of Similarity Indices [DAISI = 1 – (*rd* * *dd*)]

| | A | B | C | D | E | F |
|---|---|---|---|---|---|---|
| A | 1.0000 | 1.0000 | 0.9633 | 0.9633 | 0.9633 | 0.8833 |
| B | 1.0000 | 1.0000 | 0.9633 | 0.9633 | 0.9633 | 0.8833 |
| C | 0.9633 | 0.9633 | 1.0000 | 1.0000 | 0.9840 | 0.9360 |
| D | 0.9633 | 0.9633 | 1.0000 | 1.0000 | 0.9840 | 0.9360 |
| E | 0.9633 | 0.9633 | 0.9840 | 0.9840 | 1.0000 | 0.9200 |
| F | 0.8833 | 0.8833 | 0.9360 | 0.9360 | 0.9200 | 1.0000 |

APPENDIX III

**Daily Activity Intensity Similarity Index Clusters of 143 Respondents and k-means Clusters, where k = 5**

DAISI: Cluster Groups

| No. | ID | Cluster | Status | Residency | Age | Gender | Campus |
|---|---|---|---|---|---|---|---|
| 1 | 73GN18FN | 1 | G | N | 18 | F | N |
| 2 | 62UN18FN | 1 | U | N | 18 | F | N |
| 3 | 82UN18FN | 1 | U | N | 18 | F | N |
| 4 | 84UN18FN | 1 | U | N | 18 | F | N |
| 5 | 113UN18FN | 1 | U | N | 18 | F | N |
| 6 | 7UO18FN | 1 | U | O | 18 | F | N |
| 7 | 45UO18FN | 1 | U | O | 18 | F | N |
| 8 | 129UO18FN | 1 | U | O | 18 | F | N |
| 9 | 117GI18MN | 1 | G | I | 18 | M | N |
| 10 | 127UN18MN | 1 | U | N | 18 | M | N |
| 11 | 14UO18MN | 1 | U | O | 18 | M | N |
| 12 | 34UO18MN | 1 | U | O | 18 | M | N |
| 13 | 81UO18MN | 1 | U | O | 18 | M | N |
| 14 | 114UO18MN | 1 | U | O | 18 | M | N |
| 15 | 118UO18MN | 1 | U | O | 18 | M | N |
| 16 | 147UO18MN | 1 | U | O | 18 | M | N |

| No. | ID | Cluster | Status | Residency | Age | Gender | Campus |
|---|---|---|---|---|---|---|---|
| 17 | 154UO18MN | 1 | U | O | 18 | M | N |
| 18 | 137UI18FY | 1 | U | I | 18 | F | Y |
| 19 | 38UN18FY | 1 | U | N | 18 | F | Y |
| 20 | 79UN18FY | 1 | U | N | 18 | F | Y |
| 21 | 83UN18FY | 1 | U | N | 18 | F | Y |
| 22 | 134UN18FY | 1 | U | N | 18 | F | Y |
| 23 | 19GO18FY | 1 | G | O | 18 | F | Y |
| 24 | 27UO18FY | 1 | U | O | 18 | F | Y |
| 25 | 52UO18FY | 1 | U | O | 18 | F | Y |
| 26 | 108UO18FY | 1 | U | O | 18 | F | Y |
| 27 | 116UN18MY | 1 | U | N | 18 | M | Y |
| 28 | 31UO18MY | 1 | U | O | 18 | M | Y |
| 29 | 61UO18MY | 1 | U | O | 18 | M | Y |
| 30 | 22GN25FN | 1 | G | N | 25 | F | N |
| 31 | 100GN25FN | 1 | G | N | 25 | F | N |
| 32 | 25SO25FN | 1 | S | O | 25 | F | N |

| No. | ID | Cluster | Status | Residency | Age | Gender | Campus |
|---|---|---|---|---|---|---|---|
| 33 | 69SO25FN | 1 | S | O | 25 | F | N |
| 34 | 13GI25MN | 1 | G | I | 25 | M | N |
| 35 | 54GN25MN | 1 | G | N | 25 | M | N |
| 36 | 144UO25MN | 1 | U | O | 25 | M | N |
| 37 | 28GI25FY | 1 | G | I | 25 | F | Y |
| 38 | 50GI25FY | 1 | G | I | 25 | F | Y |
| 39 | 64GI25FY | 1 | G | I | 25 | F | Y |
| 40 | 42GI25MY | 1 | G | I | 25 | M | Y |
| 41 | 29SI30MN | 1 | S | I | 30 | M | N |
| 42 | 99FN30MN | 1 | F | N | 30 | M | N |
| 43 | 40GN30MN | 1 | G | N | 30 | M | N |
| 44 | 148GN30MN | 1 | G | N | 30 | M | N |
| 45 | 5GO30MN | 1 | G | O | 30 | M | N |
| 46 | 101GI30FY | 1 | G | I | 30 | F | Y |
| 47 | 53GI33MY | 1 | G | I | 33 | M | Y |
| 48 | 98FO35FN | 1 | F | O | 35 | F | N |

| No. | ID | Cluster | Status | Residency | Age | Gender | Campus |
|---|---|---|---|---|---|---|---|
| 49 | 142FO35FN | 1 | F | O | 35 | F | N |
| 50 | 143GI35MN | 1 | G | I | 35 | M | N |
| 51 | 145GO35MN | 1 | G | O | 35 | M | N |
| 52 | 102SO35MN | 1 | S | O | 35 | M | N |
| 53 | 15GI35FY | 1 | G | I | 35 | F | Y |
| 54 | 59SO40FN | 1 | S | O | 40 | F | N |
| 55 | 1GN40MN | 1 | G | N | 40 | M | N |
| 56 | 78UN40MN | 1 | U | N | 40 | M | N |
| 57 | 23FO40MN | 1 | F | O | 40 | M | N |
| 58 | 95FO40MN | 1 | F | O | 40 | M | N |
| 59 | 88GI45FN | 1 | G | I | 45 | F | N |
| 60 | 9SO45FN | 1 | S | O | 45 | F | N |
| 61 | 57SO45FN | 1 | S | O | 45 | F | N |
| 62 | 63SO45FN | 1 | S | O | 45 | F | N |
| 63 | 26FO45MN | 1 | F | O | 45 | M | N |
| 64 | 124UO45MN | 1 | U | O | 45 | M | N |

| No. | ID | Cluster | Status | Residency | Age | Gender | Campus |
|---|---|---|---|---|---|---|---|
| 65 | 3FO50MN | 1 | F | O | 50 | M | N |
| 66 | 68FO50MN | 1 | F | O | 50 | M | N |
| 67 | 71FO50MN | 1 | F | O | 50 | M | N |
| 68 | 104FO50MN | 1 | F | O | 50 | M | N |
| 69 | 136GO50MN | 1 | G | O | 50 | M | N |
| 70 | 94SN55FN | 1 | S | N | 55 | F | N |
| 71 | 96SO55FN | 1 | S | O | 55 | F | N |
| 72 | 97SO55FN | 1 | S | O | 55 | F | N |
| 73 | 39UO55FN | 1 | U | O | 55 | F | N |
| 74 | 75GO55N | 1 | G | O | 55 | M | N |
| 75 | 30SO55FY | 1 | S | O | 55 | F | Y |
| 76 | 103GN55MY | 1 | G | N | 55 | M | Y |
| 77 | 18FO60FN | 1 | F | O | 60 | F | N |
| 78 | 72SN60MN | 1 | S | N | 60 | M | N |
| 79 | 36SO60MN | 1 | S | O | 60 | M | N |
| 80 | 138GN60MY | 1 | G | N | 60 | M | Y |

| No. | ID | Cluster | Status | Residency | Age | Gender | Campus |
|---|---|---|---|---|---|---|---|
| 1 | 20GN18FN | 2 | G | N | 18 | F | N |
| 2 | 85GN18FN | 2 | G | N | 18 | F | N |
| 3 | 46UN18FN | 2 | U | N | 18 | F | N |
| 4 | 111UN18FN | 2 | U | N | 18 | F | N |
| 5 | 87GO18FN | 2 | G | O | 18 | F | N |
| 6 | 120GO18FN | 2 | G | O | 18 | F | N |
| 7 | 16UO18FN | 2 | U | O | 18 | F | N |
| 8 | 21UO18FN | 2 | U | O | 18 | F | N |
| 9 | 67UO18FN | 2 | U | O | 18 | F | N |
| 10 | 77UO18FN | 2 | U | O | 18 | F | N |
| 11 | 93UO18FN | 2 | U | O | 18 | F | N |
| 12 | 90UO18MN | 2 | U | O | 18 | M | N |
| 13 | 91UO18MN | 2 | U | O | 18 | M | N |
| 14 | 92UO18MN | 2 | U | O | 18 | M | N |
| 15 | 128UO18MN | 2 | U | O | 18 | M | N |
| 16 | 139UO18MN | 2 | U | O | 18 | M | N |

| No. | ID | Cluster | Status | Residency | Age | Gender | Campus |
|---|---|---|---|---|---|---|---|
| 17 | 152UO18MN | 2 | U | O | 18 | M | N |
| 18 | 153UO18MN | 2 | U | O | 18 | M | N |
| 19 | 8GI18FY | 2 | G | I | 18 | F | Y |
| 20 | 122UI18FY | 2 | U | I | 18 | F | Y |
| 21 | 11UN18FY | 2 | U | N | 18 | F | Y |
| 22 | 12UN18FY | 2 | U | N | 18 | F | Y |
| 23 | 48UN18FY | 2 | U | N | 18 | F | Y |
| 24 | 74UN18FY | 2 | U | N | 18 | F | Y |
| 25 | 131UN18FY | 2 | U | N | 18 | F | Y |
| 26 | 4UO18FY | 2 | U | O | 18 | F | Y |
| 27 | 10UO18FY | 2 | U | O | 18 | F | Y |
| 28 | 55UO18FY | 2 | U | O | 18 | F | Y |
| 29 | 141UN18MY | 2 | U | N | 18 | M | Y |
| 30 | 151UN18MY | 2 | U | N | 18 | M | Y |
| 31 | 125UO18MY | 2 | U | O | 18 | M | Y |
| 32 | 133UO18MY | 2 | U | O | 18 | M | Y |
| 33 | 149GN25MN | 2 | G | N | 25 | M | N |

| No. | ID | Cluster | Status | Residency | Age | Gender | Campus |
|---|---|---|---|---|---|---|---|
| 34 | 115UO25MN | 2 | U | O | 25 | M | N |
| 35 | 2GI25FY | 2 | G | I | 25 | F | Y |
| 36 | 49GI25FY | 2 | G | I | 25 | F | Y |
| 37 | 51GI30FN | 2 | G | I | 30 | F | N |
| 38 | 65GI30FN | 2 | G | I | 30 | F | N |
| 39 | 70FN30MN | 2 | F | N | 30 | M | N |
| 40 | 106FN30MN | 2 | F | N | 30 | M | N |
| 41 | 43GN30MN | 2 | G | N | 30 | M | N |
| 42 | 60GN35FN | 2 | G | N | 35 | F | N |
| 43 | 76FO35FN | 2 | F | O | 35 | F | N |
| 44 | 44GO35FN | 2 | G | O | 35 | F | N |
| 45 | 24GO35MN | 2 | G | O | 35 | M | N |
| 46 | 126GN40MN | 2 | G | N | 40 | M | N |
| 47 | 112UO40MN | 2 | U | O | 40 | M | N |
| 48 | 109FI40MY | 2 | F | I | 40 | M | Y |
| 49 | 41FO60MN | 2 | F | O | 60 | M | N |

| No. | ID | Cluster | Status | Residency | Age | Gender | Campus |
|---|---|---|---|---|---|---|---|
| 1 | 58FI35MN | 3 | F | I | 35 | M | N |
| 2 | 17GI25FY | 3 | G | I | 25 | F | Y |
| 3 | 140UI18FN | 3 | U | I | 18 | F | N |
| 4 | 32GN30MN | 3 | G | N | 30 | M | N |
| 5 | 80UN18MY | 3 | U | N | 18 | M | Y |
| 6 | 123GO18FN | 3 | G | O | 18 | F | N |
| 7 | 66UO18FN | 3 | U | O | 18 | F | N |
| 8 | 110UO18FN | 3 | U | O | 18 | F | N |
| 9 | 119UO18FY | 3 | U | O | 18 | F | Y |
| 10 | 121UO18FY | 3 | U | O | 18 | F | Y |
| 11 | 130UO18FY | 3 | U | O | 18 | F | Y |

| No. | ID | Cluster | Status | Residency | Age | Gender | Campus |
|---|---|---|---|---|---|---|---|
| 1 | 56UO18FN | 4 | U | O | 18 | F | N |
| 2 | 89GO45FN | 4 | G | O | 45 | F | N |

| No. | ID | Cluster | Status | Residency | Age | Gender | Campus |
|---|---|---|---|---|---|---|---|
| 1 | 47UO18MY | 5 | U | O | 18 | M | Y |

APPENDIX IV

## Oklahoma State University Institutional Review Board (IRB) and Questionnaire

**Oklahoma State University Institutional Review Board**

| | |
|---|---|
| Date: | Wednesday, March 16, 2011 |
| IRB Application No | AS1132 |
| Proposal Title: | Analyzing Activity Patterns in a University Campus: A Time Geography Approach |
| Reviewed and Processed as: | Exempt |

**Status Recommended by Reviewer(s): Approved Protocol Expires: 3/15/2012**

Principal Investigator(s):

Leonard Sit Ji Bombom
90 S. Univ. Place
Stillwater, OK 74075

Hongbo Yu
337 S. Murray
Stillwater, OK 74078

The IRB application referenced above has been approved. It is the judgment of the reviewers that the rights and welfare of individuals who may be asked to participate in this study will be respected, and that the research will be conducted in a manner consistent with the IRB requirements as outlined in section 45 CFR 46.

☒ The final versions of any printed recruitment, consent and assent documents bearing the IRB approval stamp are attached to this letter. These are the versions that must be used during the study.

As Principal Investigator, it is your responsibility to do the following:

1. Conduct this study exactly as it has been approved. Any modifications to the research protocol must be submitted with the appropriate signatures for IRB approval.
2. Submit a request for continuation if the study extends beyond the approval period of one calendar year. This continuation must receive IRB review and approval before the research can continue.
3. Report any adverse events to the IRB Chair promptly. Adverse events are those which are unanticipated and impact the subjects during the course of this research; and
4. Notify the IRB office in writing when your research project is complete.

Please note that approved protocols are subject to monitoring by the IRB and that the IRB office has the authority to inspect research records associated with this protocol at any time. If you have questions about the IRB procedures or need any assistance from the Board, please contact Beth McTernan in 219 Cordell North (phone: 405-744-5700, beth.mcternan@okstate.edu).

Sincerely,

Shelia M. Kennison

Shelia Kennison, Chair
Institutional Review Board

**OSU, STILLWATER CAMPUS ACTIVITY PATTERN RESEARCH SURVEY, 2011**

Dear [OSU Faculty/Staff/Student's name]

My name is Leonard Sitji Bombom, a PhD student in the Department of Geography, Oklahoma State University and I am conducting a research survey for my dissertation.

I would like to invite you to participate in a research survey of activity patterns on OSU campus. Such a study may have significant implications for designing and locating campus-wide facilities for different types of activity spaces, etc.

Your participation in this research survey would consist of completing a **15-20-minute online questionnaire or a hard copy** (pls, contact me or my advisor by email or phone and a drop-by can be arranged), which include basic questions **on your activities** (types and locations of activities, starting and ending time of activities, and travel mode) for **Wednesday and Thursday** of a week. You may fill in the questionnaire at the end of the activity day or the next day for a previous activity day.

No names, addresses or contact information are required unless voluntarily provided. **Some results of the study will be posted online** (see link below). Each questionnaire will have a unique ID that will be known only to you and **which you may use to access and view the image of your activity pattern** and compare them to the general activity patterns on campus.

http://www2.geog.okstate.edu/users/bombom/Index.html

Participation in this research survey is **voluntary** and you are free to withdraw from the study at any time while completing the questionnaire. The research survey is also **completely anonymous**. Information collected will be kept in strict confidence, for a period not exceeding one year only; in secure locations in the offices of the researcher and his advisor at 361 and 342 Murray Hall, respectively. Your name or identity may not be known, and if volunteered, will NOT be used in the analysis or research findings or dissertation report in any way.

If you have questions or concerns about your involvement in this study, please do not hesitate to contact me at (405) 744 2901 or by email: leonard.s.bombom@okstate.edu or the Chair of my dissertation committee, Dr. Hongbo Yu at (405) 744 9167 or by email: hongbo.yu@okstate.edu. If you have questions about your rights as a research volunteer, you may contact the Oklahoma State University Institutional Review Board (IRB) Chair, Dr. Shelia Kennison, 219 Cordell North, Stillwater, OK 74078, 405-744-3377 or irb@okstate.edu.

Thank you,

Sincerely
Leonard Sitji Bombom
PhD Student

Okla. State Univ.
IRB
Approved 3/16/11
Expires 3/15/12
IRB# AS-11-32

## Oklahoma State University Institutional Review Board

| | | | |
|---|---|---|---|
| Date: | Friday, February 24, 2012 | Protocol Expires: | 2/23/2013 |
| IRB Application No: | AS1132 | | |
| Proposal Title: | Analyzing Activity Patterns in a University Campus: A Time Geography Approach | | |

Reviewed and Processed as: **Modification/Continuation**

Status Recommended by Reviewer(s) **Approved**

Principal Investigator(s) :

| | |
|---|---|
| Leonard Sit Ji Bombom<br>90 S. Univ. Place<br>Stillwater, OK 74075 | Hongbo Yu<br>337 S. Murray<br>Stillwater, OK 74078 |

---

Approvals are valid for one calendar year, after which time a request for continuation must be submitted. Any modifications to the research project approved by the IRB must be submitted for approval with the advisor's signature. The IRB office MUST be notified in writing when a project is complete. Approved projects are subject to monitoring by the IRB.

☒ The final versions of any printed recruitment, consent and assent documents bearing the IRB approval stamp are attached to this letter. These are the versions that must be used during the study.

Signature :

Shelia M. Kennison

Shelia Kennison, Chair, Institutional Review Board

Friday, February 24, 2012

Date

If you have questions or concerns about your involvement in this study, please do not hesitate to contact me at (405) 744-2901 or by email: leonard.s.bombom@okstate.edu or the Chair of my dissertation committee, Dr. Hongbo Yu at (405) 744-9167 or by email: hongbo.yu@okstate.edu. If you have questions about your rights as a research volunteer, you may contact the Oklahoma State University Institutional Review Board (IRB) Chair, Dr. Shelia Kennison, 219 Cordell North, Stillwater, OK 74078, 405-744-3377 or irb@okstate.edu .

Thank you,

Sincerely
Leonard Sitji Bombom
PhD Student



[redacted]

RESEARCH SURVEY ON ACTIVITY PATTERNS ON OSU CAMPUS, STILLWATER

Summary of Project

Human societies revolve around activities. Each day, everybody is involved in a number of activities, which take place in different locations. It takes time to engage in activities and also to move between activity locations. Consequently, participation in activities requires both space (location) and time. Using lines, movement between activity locations (tilted lines) and duration of participation in activity locations (vertical lines), can be plotted for the activity itinerary (pattern) of individuals. This line is referred to as a space-time path. When many space-time paths (of individuals) are put together, you have a mesh of lines that do not appear to make much sense (see diagram in questions below).

This study hopes to bring some order into this mesh by developing measures that will allow for space-time paths (lines of individual activity itinerary) to be categorized into groups of similarities based on activity characteristics. This will then allow for better understanding of the factors that may have resulted into groups of individuals having similar activity patterns (itinerary).

The university campus presents an environment within which to collect data and develop the methods needed to achieve this goal.

This is the summary of the study. More at: http://www2.geog.okstate.edu/users/bombom/Research.html

**1. What is your Status (role) in OSU?**

- What is your Status (role) in OSU? Undergraduate Student
- Graduate Student
- Faculty
- Researcher
- University Staff
- Others

**2. Where do you come from?**

- ○ Where do you come from?  Oklahoma Resident
- ○ Out-of-State Resident
- ○ International

**3. What is your age range?**

- ○ What is your age range?  18 - 24
- ○ 25 - 29
- ○ 30 - 34
- ○ 35 - 39
- ○ 40 - 44
- ○ 45 - 49
- ○ 50 - 54
- ○ 55 - 59
- ○ 60 +

**4. What is your gender?**

- ○ What is your gender?  Male
- ○ Female

**5. Do you live on campus?**

- ○ Yes
- ○ No

**INSTRUCTIONS AND EXAMPLE**

1. Please record your full activities for the day. This includes travel activities between places and the approximate time they took.

2. If your home is off-campus, pls, select or use the nearest activity locaton (on- or off-campus) from the list provided (see map for help, too) that may be nearest to your home location as starting location. On-campus housing is also generically stated for privacy concerns

3. If you undertook an activity outside of the OSU campus, pls select one of the locations provided for places (on- or off campus) that may be closest to your activity location. This is to safeguard your privacy regarding the actual location of your activity/home. List of OFF-CAMPUS locations are below the On-Campus locations (in alphabetical order, too)

4. Ex: If you woke up at 6.00 am and prepared breakfast, then activity 1 is “H/hold Activities”, and location 1 is your “home address” and activity schedule is “Habitual”;
if you then had your bath after that, then activity 2 is "hygiene,"your activity location is home, state start and end times and activity schedule may be "Habitual;"
if you then left home at 8.15 am to library, then activity 3 is “Travel,” no activity location, travel mode may be "Walk" and activity schedule may be "Habitual.".
If you arrived library at 8:30 for a scheduled group discussion and left at 9:20 am, then activity 4 is “Studying/Researching”, activity location is “Edmon Low Library”, and activity scheduling is “Planned”.
If you accepted a friend’s impromptu invitation (not previously planned) to go to “Student Union”, then activity schedule is “Spontaneous.” Start time and End time depend on when you stand/end each activity.
If you attended two different sessions of classes/lectures in the same location consecutively, you report them as two different activities, with different start/end times.
If you visited a friend's apartment near Boomer Lake, please select Boomer Lake as activity location in the drop-down box...

5. The Start time for a new activity should coincide with the End time of the previous activity

(Pls, see example below)

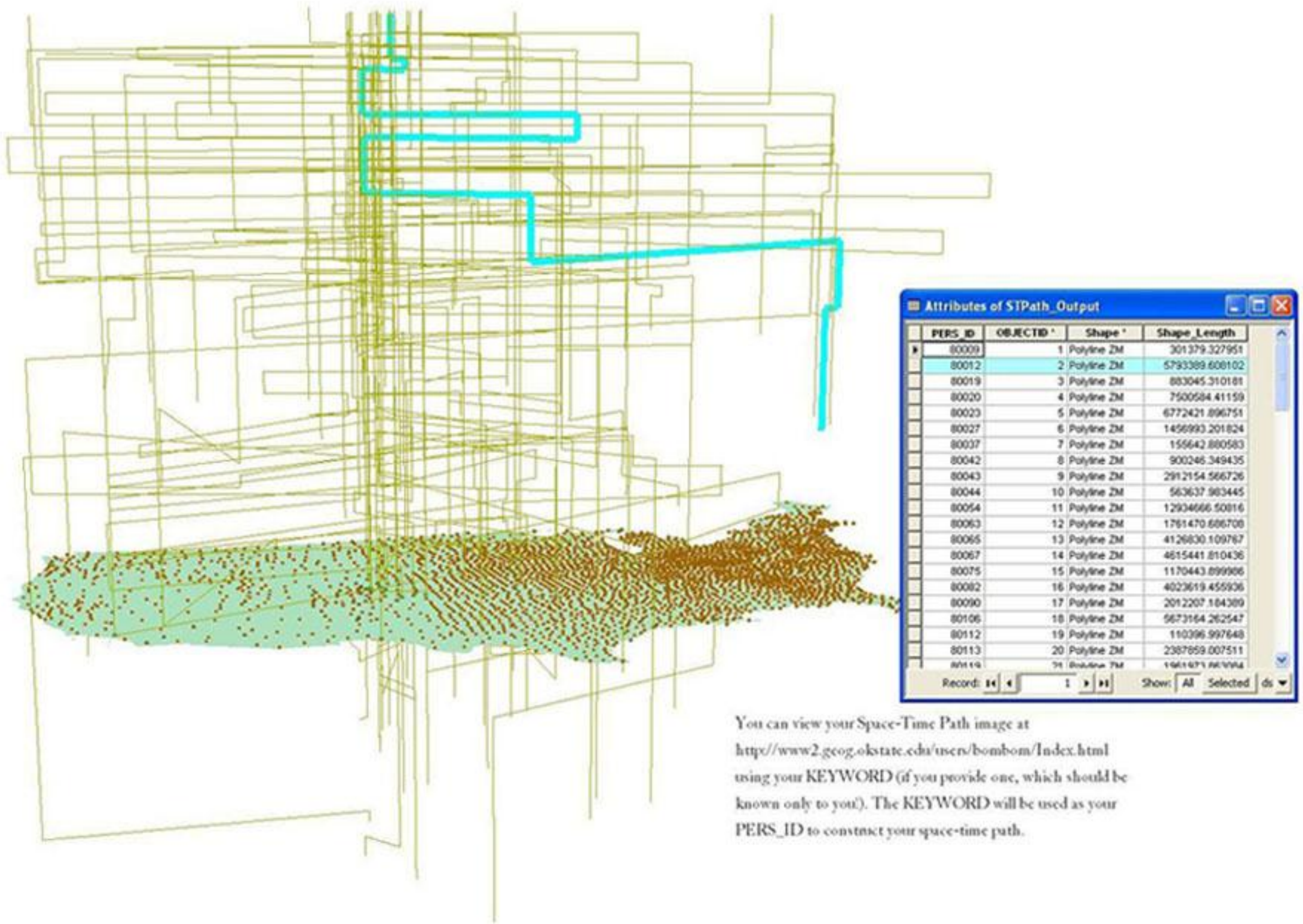
**6. FILL IN ACTIVITY INFORMATION FOR WEDNESDAY**

**A. [Activity Scheduling: Habitual (Fixed activities, e.g., Lectures); Planned (e.g., Dept. Meeting); Spontaneous (e.g., impromptu meeting)]**

**B. Please enter information for "Primary" activity, e.g., if you are texting while working, "work" is your primary activity**

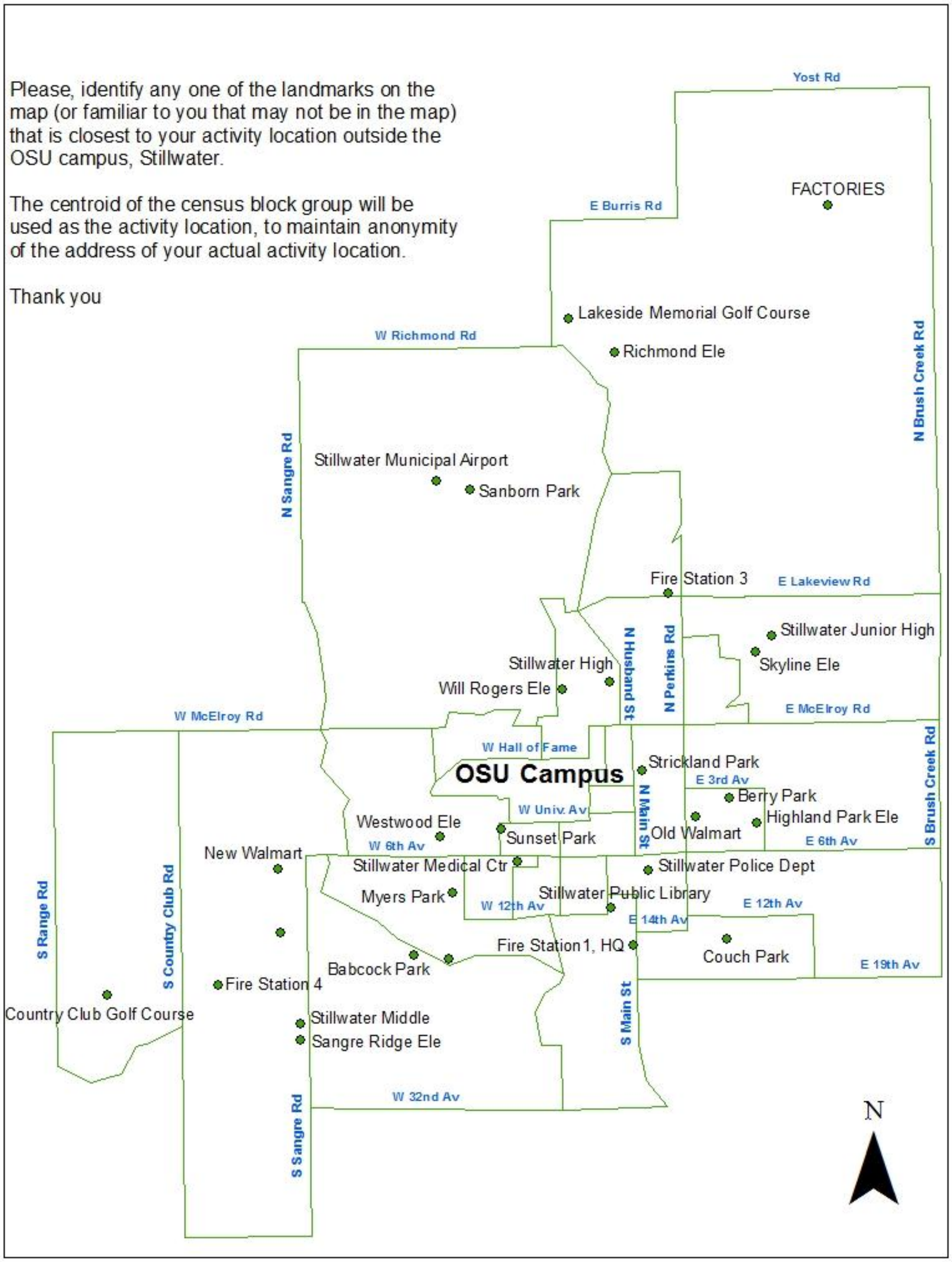
| | Activity Type | Start Time (Hours) | Start Time (Mins) | End Time (Hours) | End Time (Mins) | Activity Location (If Activity is NOT Travel) | Travel Mode (If activity is travel) | Activity Scheduling |
|---|---|---|---|---|---|---|---|---|
| Activity 1 | Hygiene (Bathing, dressing, etc) | 6.00 am | 45 | 7.00 am | 20 | STEVENS APTS S90 – S92 | | Habitual |
| Activity 2 | Household Activities (Cooking, cleaning, child care, etc) | 7.00 am | 20 | 8.00 am | | STEVENS APTS S90 – S92 | | Habitual |
| Activity 3 | Travel (Between activity locations) | 8.00 am | | 8.00 am | 25 | | Walk | Habitual |
| Activity 4 | Communication (Cellphone/message/textg/email/internet,etc) | 8.00 am | 25 | 9.00 am | 40 | MURRAY HALL NORTH/SOUTH | | Habitual |
| Activity 5 | Work/Teaching | 9.00 am | 40 | 11.00 am | | MURRAY HALL NORTH/SOUTH | | Habitual |
| Activity 6 | Studying/Researching | 11.00 am | | 12 Noon | | MURRAY HALL NORTH/SOUTH | | Planned |
| Activity 7 | Travel (Between activity locations) | 12 Noon | | 12 Noon | 10 | | Walk | Habitual |
| Activity 8 | Eating | 12 Noon | 10 | 1.00 pm | | STUDENT UNION/ PAUL MILLER JOURN & BR /ATHERTON | | Spontaneous |
| Activity 9 | Travel (Between activity locations) | 1.00 pm | | 1.00 pm | 15 | | Walk | Habitual |
| Activity 10 | Communication (Cellphone/message/textg/email/internet,etc) | 1.00 pm | 15 | 1.00 pm | 45 | MURRAY HALL NORTH/SOUTH | | Spontaneous |
| Activity 11 | Travel (Between activity locations) | 1.00 pm | 45 | 2.00 pm | | | Walk | Habitual |
| Activity 12 | Attend Classes | 2.00 pm | | 3.00 pm | 30 | CLASSROOM BUILDING/ BOOKSTORE (TEMPORARY) | | Habitual |
| Activity 13 | Travel (Between activity locations) | 3.00 pm | 30 | 4.00 pm | | | Bus (Campus Transit) | Spontaneous |
| Activity 14 | Sleeping/Resting/Idle | 4.00 pm | | 6.00 pm | 30 | STEVENS APTS S90 – S92 | | Planned |
| Activity 15 | Travel (Between activity locations) | 6.00 pm | 30 | 6.00 pm | 50 | | Bicycle | Planned |

Migration patterns of selected people across the US. The highlighted space-time path shows the migration pattern of one person:



OSU Campus Map (to help with identifying activity locations). The Campus map is saved on the geography department's server and is linked to the survey. If it does not show, pls, exit the survey, click on the survey link again and click 'No' to the prompt. There is NO spam, pls.

Choose Locations Closest to Actual Activity Locations Outside of Campus



6. FILL IN ACTIVITY INFORMATION FOR WEDNESDAY

A. [Activity Scheduling: Habitual (Fixed activities, e.g., Lectures); Planned (e.g., Dept. Meeting); Spontaneous (e.g., impromptu meeting)]

B. Please enter information for "Primary" activity, e.g., if you are texting while working, "work" is your primary activity

| | Activity Type | Start Time (Hours) | Start Time (Mins) | End Time (Hours) | End Time (Mins) | Activity Location (If Activity is NOT Travel) | Travel Mode (If activity is travel) | Activity Scheduling |
|---|---|---|---|---|---|---|---|---|
| Activity 1 | | | | | | | | |
| Activity 2 | | | | | | | | |
| Activity 3 | | | | | | | | |
| Activity 4 | | | | | | | | |
| Activity 5 | | | | | | | | |
| Activity 6 | | | | | | | | |
| Activity 7 | | | | | | | | |
| Activity 8 | | | | | | | | |
| Activity 9 | | | | | | | | |
| Activity 10 | | | | | | | | |
| Activity 11 | | | | | | | | |
| Activity 12 | | | | | | | | |
| Activity 13 | | | | | | | | |
| Activity 14 | | | | | | | | |
| Activity 15 | | | | | | | | |

7. ACTIVITY INFORMATION FOR WEDNESDAY (Contd)

A. [Activity Scheduling: Habitual (Fixed activities, e.g., Lectures); Planned (e.g., Dept. Meeting); Spontaneous (e.g., impromptu meeting)]

B. Please enter information for "Primary" activity, e.g., if you are texting while working, "work" is your primary activity

| | Activity Type | Start Time (Hours) | Start Time (mins) | End Time (Hours) | End Time (Mins) | Activity Location (If Activity is NOT Travel) | Travel Mode (If activity is Travel) | Activity Scheduling |
|---|---|---|---|---|---|---|---|---|
| Activity 16 | | | | | | | | |
| Activity 17 | | | | | | | | |
| Activity 18 | | | | | | | | |
| Activity 19 | | | | | | | | |
| Activity 20 | | | | | | | | |
| Activity 21 | | | | | | | | |
| Activity 22 | | | | | | | | |
| Activity 23 | | | | | | | | |
| Activity 24 | | | | | | | | |
| Activity 25 | | | | | | | | |
| Activity 26 | | | | | | | | |
| Activity 27 | | | | | | | | |
| Activity 28 | | | | | | | | |
| Activity 29 | | | | | | | | |
| Activity 30 | | | | | | | | |

**8. ACTIVITY INFORMATION FOR THURSDAY (same instructions as above)**

**A. [Activity Scheduling: Habitual (Fixed activities, e.g., Lectures); Planned (e.g., Dept. Meeting); Spontaneous (e.g., impromptu meeting)]**

**B. Please enter information for "Primary" activity, e.g., if you are texting while working, "work" is your primary activity**

| | Activity Type | Start Time (Hrs) | Start Time (Mins) | End Time (Hrs) | End Time (Mins) | Activity Location (If Activity is NOT Travel) | Travel Mode (If activity is Travel) | Activity Scheduling |
|---|---|---|---|---|---|---|---|---|
| Activity 1 | | | | | | | | |
| Activity 2 | | | | | | | | |
| Activity 3 | | | | | | | | |
| Activity 4 | | | | | | | | |
| Activity 5 | | | | | | | | |
| Activity 6 | | | | | | | | |
| Activity 7 | | | | | | | | |
| Activity 8 | | | | | | | | |
| Activity 9 | | | | | | | | |
| Activity 10 | | | | | | | | |
| Activity 11 | | | | | | | | |
| Activity 12 | | | | | | | | |
| Activity 13 | | | | | | | | |
| Activity 14 | | | | | | | | |
| Activity 15 | | | | | | | | |

9. ACTIVITY INFORMATION FOR THURSDAY (Contd) (same instructions as above)

A. [Activity Scheduling: Habitual (Fixed activities, e.g., Lectures); Planned (e.g., Dept. Meeting); Spontaneous (e.g., impromptu meeting)]

B. Please enter information for "Primary" activity, e.g., if you are texting while working, "work" is your primary activity

| | Activity Type | Start Time (Hrs) | Start Time (Mins) | End Time (Hrs) | End Time (Mins) | Activity Location (If Activity is NOT Travel) | Travel Mode (If activity is Travel) | Activity Scheduling |
|---|---|---|---|---|---|---|---|---|
| Activity 16 | | | | | | | | |
| Activity 17 | | | | | | | | |
| Activity 18 | | | | | | | | |
| Activity 19 | | | | | | | | |
| Activity 20 | | | | | | | | |
| Activity 21 | | | | | | | | |
| Activity 22 | | | | | | | | |
| Activity 23 | | | | | | | | |
| Activity 24 | | | | | | | | |
| Activity 25 | | | | | | | | |
| Activity 26 | | | | | | | | |
| Activity 27 | | | | | | | | |
| Activity 28 | | | | | | | | |
| Activity 29 | | | | | | | | |
| Activity 30 | | | | | | | | |

10. OPTIONAL

1. If you would let us meet with you, if we have further questions, please provide contact information, e.g., email address. If not, disregard
2. Please, write down a KEYWORD of your choice as your unique ID which you may use to retrieve your activity pattern image when the preliminary results of the study are posted online

OPTIONAL: If you would let us meet with you, if we have further questions, please provide contact information, e.g., email address. If not, disregard 2. Please, write down a KEYWORD of your choice as your unique ID which you may use to retrieve your activity pattern image when the preliminary results of the study are posted online

VITA

Leonard Sitji Bombom

Candidate for the Degree of

Doctor of Philosophy

Thesis: DECIPHERING ACTIVITY PATTERN USING TIME-GEOGRAPHY FRAMEWORK: A CASE STUDY OF OKLAHOMA STATE UNIVERSITY, STILLWATER CAMPUS

Major Field: Geography

Biographical:

Education:

Completed the requirements for the Doctor of Philosophy in Geography at Oklahoma State University, Stillwater, Oklahoma in December, 2013.

Completed the requirements for the Certificate in GIS at Oklahoma State University, Stillwater, Oklahoma in December, 2008.

Completed the requirements for the Master of Arts in Geography at University of Northern Iowa, Cedar Falls, Iowa/USA in December, 2005.

Completed the requirements for the Master of Science in Environmental and Resources Planning at University of Jos, Jos/Nigeria in 1999.

Completed the requirements for the Bachelor of Science in Geography and Planning at University of Jos, Jos/Nigeria in 1994.

Experience:

Professional Memberships:

Association of American Geographers (AAG)
Association of Nigerian Geographers (ANG)
Gamma Theta Upsilon (GTU), International Geographic Honors Society