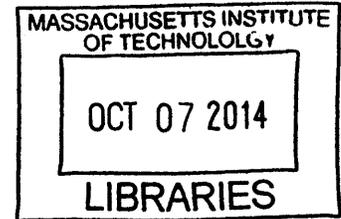


**Design and Deploy:  
Iterative Methods in Adapting Mobile Technologies for Data Acquisition,  
A Case Study in St. Louis, Missouri**

By

Kuan Butts

Bachelor of Science in Architecture  
Washington University in St. Louis  
St. Louis, Missouri (2011)



Submitted to the Department of Urban Studies and Planning  
in partial fulfillment of the requirements for the degree of

Master in City Planning

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September, 2014

© 2014 Massachusetts Institute of Technology. All Rights Reserved

The author hereby grants to MIT the permission to reproduce and to  
distribute publicly paper and electronic copies of the thesis document in  
whole or in part in any medium now known or hereafter created.

Signature redacted

Author

\_\_\_\_\_  
Signature redacted Department of Urban Studies and Planning  
(August 26, 2014)

Certified by

\_\_\_\_\_  
Signature redacted Professor Pericles Christopher Zegras  
Department of Urban Studies and Planning  
Thesis Supervisor

Signature redacted

Accepted by

\_\_\_\_\_  
Signature redacted Professor Dennis Frenchman  
Chair, MCP Committee  
Department of Urban Studies and Planning

# Design and Deploy: Iterative Methods in Adapting Mobile Technologies for Data Acquisition, A Case Study in St. Louis, Missouri

by

Kuan Butts

Submitted to the Department of Urban Studies and Planning  
on August 26, 2014 in Partial Fulfillment of the  
Requirements for the Degree of Master in City Planning

## Abstract

Advancements in mobile technology provide the opportunity to explore not only how data gathering (e.g., surveying) can be eased through digital input mechanisms, but also how such devices can bring new resolution to data gathered. This thesis covers the development history of an Android-based application, Flocktracker. Flocktracker incorporates techniques capitalizing on standard modern locational sensors on Android devices, demonstrating how data ranging from vehicle speeds to locations, directions, and on-board conditions can be relatively easily gathered. The research then deploys Flocktracker to explore the spatiotemporal dynamics of onboard security perception, as reported by users, along the 70 bus line in St. Louis.

Over a brief, three-day period in March, an on-board survey was implemented via Flocktracker. Based on this field work, the thesis presents aspects of the route data collected (origin-destination, ridership, speed, uploads activity by time of day), as well as a multivariate, ordered logit model of users' reported security perceptions, incorporating additional spatial data (e.g., on crime). Results from this model indicate the user-reported security perceptions relate significantly to highly localized aspects of a route, such as proximity to homicides, public disorder, property crimes, vacancies, vehicle speed, and relative location along the route.

Thesis Supervisor: Pericles Christopher Zegras  
Title: Associate Professor, Transportation and Urban Planning

## Foreword

The following thesis is written with the intent of taking the education I learned at MIT and researching abroad, from Bangladesh to Mexico, and bringing these two years' efforts back "home." By "home," I refer to St. Louis, Missouri, the city I adopted from 2007 to 2012.

The five years I spent in St. Louis were more formative and educational than those prior or since. While my time at MIT has allowed me to develop a wonderful array of skills and knowledge, the simple act of living in St. Louis provided an insight into what I believe is a microcosm of so many of the aspects that are both terribly wrong and terrifically right with the United States.

While the events of August 2014 regarding the shooting of Michael Brown, a Black North County young man, have brought forth many of the more ugly aspects of both St. Louis and American society at large, St. Louis has also an amazing, truly American history; one rich with industry, struggling to transform from to match the needs of a twenty-first century global economy. It is this transition that, while painful, is also a moment where the spatial nature of society becomes so critical.

Understanding the geographical dimensions of social complexity within a city is critical to fostering a future that is inclusive, economically prosperous, efficient, and, ultimately, functional. Failure to consider such broad complexities risks the tragic symptoms we observe today in many cities such as St. Louis: urban decay, vacancy, crime, segregation, and economic stagnation.

Taking the technology developed, lessons learned, and observations made from nearly two years of research in what would be considered the "Global South," I hope to adapt such efforts so as to produce an informed, albeit brief, analysis of the 70 Grand bus line, a transit service I believe represents one of the most symbolic tethers between two polar elements of St. Louis, both literally and physically: that of North and South St. Louis.

Kuan Butts

## **Acknowledgements**

Special thanks to Arturo Cadena, Daniel Palencia, and Danny Chiao; Flocktracker – none of this – would not have been possible without you. To that end, all those in Mexico City, especially the volunteers from the summer or 2013: thank you.

Endless gratitude, as well, to Chris Zegras for his continued support of the project and my aspirations in transportation throughout my time in DUSP. None of the experiences and opportunities that defined my time at MIT would have occurred had you not taken a chance on me that first semester.

To my family – my mother, father, and sister Tien – thank you for your continued support and patience. To Liz Kirby; thank you for being so understanding and encouraging, especially during the writing of this thesis.

Finally, I would like to acknowledge the Singapore-MIT Alliance for Research and Technology, whose Future Urban Mobility program provided funding to support this thesis.

# DESIGN AND DEPLOY

## Iterative Methods in Adapting Mobile Technologies for Data Acquisition, A Case Study in St. Louis, MO

Kuan Butts

### TABLE OF CONTENTS

<b>TERMINOLOGY</b> .....	7
<b>CHAPTER 1: INTRODUCTION</b> .....	9
THESIS OVERVIEW .....	9
ST. LOUIS CASE STUDY OVERVIEW .....	10
THESIS ROADMAP.....	13
<b>CHAPTER 2: SURVEY METHODS AND EMERGING TECHNOLOGIES</b> .....	14
METHODS INTRODUCTION.....	14
SMS SURVEYING: DEVELOPMENTS AND LIMITATIONS.....	16
COMMERCIAL ONLINE TOOLS .....	18
GOOGLE LABS DEVELOPMENTS: GOOGLE FORMS, FUSION TABLES, AND API .....	20
GOOGLE LABS DEVELOPMENTS: APP INVENTOR FOR ANDROID .....	21
EMERGING SOFTWARE SOLUTIONS FOR MOBILE SURVEY TECHNOLOGY .....	25
FLOCKTRACKER: EARLY DEVELOPMENT IN DHAKA.....	30
<b>CHAPTER 3: ST. LOUIS, MISSOURI CASE STUDY</b> .....	37
ST. LOUIS IN COMPARISON .....	37
INTRODUCTION TO THE CITY OF ST. LOUIS .....	39
ST. LOUIS METRO SYSTEM.....	45
URBAN CONTEXT OF 70 GRAND BUS LINE.....	48
<b>CHAPTER 4: METHODS AND TECHNICAL DESIGN</b> .....	58
INTRODUCTION TO FLOCKTRACKER'S DEVELOPMENT HISTORY .....	58
MAPPING DHAKA'S BUS NETWORK.....	59
PERCEPTION MAPPING IN MEXICO CITY.....	66
NATIVE REDEVELOPMENT OF FLOCKTRACKER .....	71
<b>CHAPTER 5: ST. LOUIS SURVEY DESIGN AND PRINCIPLES</b> .....	74
DESIGN OF ST. LOUIS OPERATION.....	74
ONBOARD SECURITY DISCUSSION .....	75
BEYOND SECURITY .....	77
<b>CHAPTER 6: SURVEY AND DATA ANALYSIS</b> .....	79
SURVEY INTRODUCTION.....	79
SURVEY METHOD .....	80
SURVEY RESULTS: INTRODUCTION .....	82
SPATIAL RESULTS: TRIP ORIGINS AND DESTINATIONS .....	92
SPATIAL RESULTS: ASPECTS OF FLOCKTRACKER DATA UPLOADS .....	98
SPATIAL RESULTS: ROUTE ONBOARD SECURITY PERCEPTION .....	104
<b>CHAPTER 7: STATISTICAL MODEL</b> .....	117
DESCRIPTIVE STATISTICS.....	117

COMMENTS ON DATA PREPARATION: DISTANCE MEASURES.....	120
MODEL SPECIFICATION AND ESTIMATION: MODEL VARIATIONS .....	122
DISCUSSION OF RESULTS: FINAL MODELS.....	125
DISCUSSION .....	129
LIMITATIONS.....	132
<b>CHAPTER 8: FUTURE POTENTIAL .....</b>	<b>134</b>
CONTEXT FOR FUTURE DEVELOPMENT .....	134
GOALS FOR THE FLOCKTRACKER APPLICATION .....	134
TECHNICAL END NOTES: FLOCKTRACKER REPOSITORY.....	138
<b>CHAPTER 9: CONCLUSIONS.....</b>	<b>140</b>
FINAL COMMENTS .....	140
<b>WORKS CITED .....</b>	<b>142</b>
<b>APPENDIX 1: SURVEY QUESTIONS .....</b>	<b>148</b>

## **Terminology**

**API: Application program interface; method for communicating data from one distinct application to another**

**Back End: server structure and management of information (non-visual components of software)**

**Background task: process designed to run without intrusion to experience of user (in terms of performance and visual aspects)**

**Call (Function Call): passes control of program activity to a predefined subroutine, which in part returns data or subsequent program instructions**

**Commit: changes made to a Git repository**

**Data Frame: two-dimensional array-like structure (table), where each column holds variable measurements, and each row values for a particular case**

**dBase: a database management system**

**DBF: dBase File's underlying file format**

**Django: open source, Python-language web framework**

**Freemium: a pricing strategy that allows base services for free, but charges for proprietary features of functionality**

**Function: a procedure or routine that represents a series of instructions; can return further instructions or some content**

**GIS: Geographic information systems; computer program designed to manage data spatially and present geographic data visually**

**Git:** Distributed revision control and source code management system

**GitHub:** web hosting service that supports Git and Git repositories

**GPS:** Global positioning system; provides specific location information

**Interval:** a period of time, to be elapsed, before an action occurs

**Java:** a computer programming language, the primary language used for  
developing applications for Android OS enabled hardware

**JavaScript:** computer programming language, most commonly used in websites

**JSON:** JavaScript Object Notation; a lightweight data-interchange format more  
easily accessible for non-technical users to write in and understand

**OS:** operating system, software designed to manage the computers hardware,  
facilitates the operation of software

**Repository:** data structure stored on a server that holds a series of files and  
directories, along with a history of changes that have been made to these its  
components

**SDK:** Android software development kit; allows for the creation of software, such  
as Android applications, but is not mandatory for creating them

**SMS:** Short message service; the mobile-to-mobile text messaging component of  
phone communications

**XML:** Extensible markup language, a format for data organization that is designed  
for both human comprehension and machine-readability; a format for holding  
data

## **CHAPTER 1: Introduction**

### **Thesis Overview**

This thesis is designed to accomplish three tasks. First, is to document, briefly, the development of survey software over time, with a particular focus on services dedicated to in-field data gathering, as well as data tethered to geospatial and temporal data. Within this development history is included a history of the developments and accomplishments of the Flocktracker application, as it has thus far progressed. The second task is the demonstration of the potential of the application, through implementation. This implementation involved a small-scale survey, which was performed in St. Louis, Missouri whose focus is an analysis of the 70 Grand line bus route. Within this analysis, both a multivariate model and a number of performance and specific, localized spatial data is presented, associated with specific spatiotemporal data. The purpose of this brief research initiative involving the application is to highlight a number of different analytical measurements possible through the use of the application, as well as to inspire new uses from in-field research where surveying is a component of data acquisition. Finally, the third component of the thesis will be to present a vision for what next steps ought to occur with Flocktracker, as well as how Flocktracker remains a valuable tool in light of other similar products' development.

## **St. Louis Case Study Overview**

In regards to the St. Louis case study, focus will be made on exploring perception of security at specific points in time along the 70 Grand bus line route. The purpose for using this measure is to explore how dynamics of the line affect or do not affect this aspect of the rider's experience. That is, how can the Flocktracker application add a layer of information that might provide greater insight to survey research than what was previously feasible? The concept of onboard security perception, in terms of fear of crime (as opposed to traffic safety opinions), was developed in an earlier research initiative involving Flocktracker. Thus, as it was seen as an effective and powerful measure of analysis in that research, it was deemed an effective measure for examining the potential of the application. Given St. Louis' documented history of crime and strained race relations, employing this as the variable for exploration seemed appropriate, within the short time frame of the research design and implementation.

Understanding perceptions of personal security is applicable to those who operate public transit in St. Louis, as well. Thus, the intent is not to demonstrate the technology's capability alone, but to explore it in a way that is constructive and representative of what could become valuable logistics, social science, or criminology research. To elaborate, crime is often a significant concern in urban public transit lines. Solutions are often geared - in dealing with this security issue - to addressing symptoms or immediate concerns where they occur. For example, a physical presence - uniformed officers - is the most commonly cited solution to dealing with

security issues along transit operations, according to a 1997 book published through the Transportation Research Board (Transit Cooperative Research Program, 1997, p. 1-4). While such directed measures do indeed address immediate concerns along a route, they are both cost intensive and temporary solutions. Additional research has also suggested that these solutions may also serve to “accentuate fear by increasing paranoia and distrust among people (Ellin, 1996, p. 153).

While this research does not seek to prescribe alternative measures to dealing with immediate security issues along a route, it does seek to understand more deeply the nature of security perception within a bus system. Specifically, it explores how environmental and location factors can influence perceptions, in relation to geospatial location and time of day.

For example, the connection between perceived security perception and urban environmental qualities is explored. In a 1988 article in *Transportation*, titled “The influence of personal security fears on women's travel patterns,” authors Lynch and Atkins report that graffiti and cleanliness were highlighted as significant factors contributing to perception of security by women. The report cited “desolation, disrepair, and untidiness” as correlating with the presence of graffiti and a general lack of cleanliness. This research harkens back to the broken windows theory, suggesting that people are influenced by their environments; if certain disorder is allowed in a neighborhood at an aesthetic level, then anything might be allowed. Thus, the area becomes insecure (Wilson and Kelling, 1982). Yavuz and Welch elaborate on this, citing Nasar and Fisher (1993) and Cozens et al. (2005), when

they write: “Transit environments that exhibit physical and social incivilities may stimulate feelings of fear and may invite criminals to perceive them as good places to commit crimes (Yavuz and Welch, 2010, p. 5).”

In addition to environmental factors, gender aspects can be examined, both for the individual respondent, and in regards to the gender balance present within the bus. These considerations build off of a concluding observation from the paper “Addressing Fear of Crime in Public Space,” in which the authors noted that “female passengers tend to be concerned about social incivilities in the transit environment, while they are less likely to be comforted by the presence of video cameras [than men] (Yavuz and Welch, 2010, p. 18).” In addition, results from a Chicago Transit Authority (CTA) Customer Satisfaction Survey (CSS), conducted in 2003 by the Northwest Research Group, revealed that both male and female respondents reported more safety problems related to other riders on trains, than in regard to crime in general (Yavuz and Welch, 2010, p. 12).

I offer an initial exploration into some of these security perception issues on a prominent bus line in St. Louis. My ultimate objective is to demonstrate the value of the Flocktracker technology, in providing potentially rich spatiotemporal data on a dynamic system. I aim to shed some light onto how on-board security perceptions of bus users might be influenced by gender, other individual characteristics, and spatial, temporal and contextual factors. The research is a preliminary demonstration, aimed at testing the technology and adding some knowledge to a complex social phenomenon.

## **Thesis Roadmap**

This thesis will first briefly provide an overview of development of digital technologies involved in surveying, in Chapter 2. Included here will be an overview of Flocktracker's development under Albert Ching, in 2011 and 2012. Chapter 3 will present the case study of St. Louis, Missouri. Included will be an introduction to the city and the bus line in particular that is researched in this thesis. Chapter 4 returns to the technological development of Flocktracker. In it, Flocktracker's developments and accomplishments over the past two years is outlined. The purpose of this section is to demonstrate the iterations of the applications and each of their deployment in various use cases. The chapter documents the development of Flocktracker up to its use in St. Louis, Missouri, in March of 2014.

The remainder of this thesis focuses on the specific case study in St. Louis and Flocktracker's future development. Chapter 6 presents the survey tool implemented in St. Louis, as well as a description of the results from the research. Chapter 7 builds on these results, presenting a statistical analysis of the site, including a model and discussion of the model's results. Chapter 8 returns to the Flocktracker application, presenting goals for continued development of the technology. Chapter 9 contextualizes work in the St. Louis study with technical goals outlined in Chapter 8, providing final comments and concluding remarks.

## **CHAPTER 2: Survey Methods and Emerging Technologies**

### **Methods Introduction**

From summer jobs canvassing street corners to multi-week research surveys in academia, numerous pen and paper in-person surveys are performed daily. Surveying is a key component of social science; vital for data acquisition in many academic fields, as well as other areas such as business and political operations. While performing surveys in an analog fashion (via pen and paper) is a tried and true method; new mobile technologies, combined with their ever-growing presence and affordability, make considering technological adaptations pertinent now more than ever. For example, in the United States, smartphone ownership reached 65.2% of the mobile market share, equating to 156 million handsets (comScore, 2014).

Limitations of a traditional analog survey method include the cost of data conversion into digital format; the errors that such conversions are prone to; the inability to monitor results as they are occurring and adjust and strategize based off those observations; the lack of quality assurance that the surveys are completed honestly; the difficulty in monitoring and controlling for environmental, spatial, and temporal inputs; and other physical constraints such as the difficulty of performing onboard surveys. It is in light of these issues that a brief synopsis of the state of technology in the field of surveying is pertinent.

In addition to ease of data entry, some surveys may benefit from features available on modern mobile devices, particularly smartphones and some tablet

models. Specifically, the ability to provide locational data and tie it to the point where each survey was performed, provides the opportunity to improve the resolution at which data is collected. For example, while, theoretically, monitoring the speed of the vehicle, the pinpoint location where the survey occurred, and the number of riders in the bus by gender is all possible on pen and paper, such a survey initiative would likely be overly ambitious. Furthermore, simply performing surveys with pen and paper onboard moving vehicles is often prohibitively difficult, particularly if answers involve more than circling a single option. That is, writing onboard a moving vehicle, particularly a bus, is inefficient, prone to legibility issues, and otherwise unnecessarily difficult. Surveys performed on smartphones help resolve these issues by enabling data entry to be easier, as well as more engaging, through responsive, digital interfaces.

Table 2.1: Comparison table of survey methods, traditional versus smartphone-based.

Traditional (“Pen and Paper”) Survey Method	Digitally Enhanced Survey Method (Smartphone)
Respondent data	Respondent data
Route Information	Route Information
Route Direction	Route Direction
	Respondent geospatial coordinates
	Vehicle speed
	In-vehicle ridership count, by gender
	Live data entry and visualization
	Surveyor in-field monitoring, safety

Table above demonstrates the comparative advantages that surveying on a smartphone or tablet can provide while in field.

This chapter will summarize developments within digital survey services in recent years, highlighting popular services and demonstrating the range of types of services available. First, SMS-based surveying will be discussed, as it is the least advanced survey technology presently available, in terms of phone operating system and capabilities requirements. Afterwards will be presented commercial browser-based tools, Google labs projects, smartphone based software, and finally Flocktracker's own past at Massachusetts Institute of Technology (MIT).

### **SMS Surveying: Developments and Limitations**

While creating custom hardware to respond to and optimize the observed shortcomings of traditional pen and paper surveying is capital intensive, mobile device operating systems have developed to the point where software capabilities are comparable to many modern desktop or laptop computers. Even in older model handsets, the limited technological capacity of the device is sufficient for some types of survey performance. In terms of survey deployment on pre-smartphone mobile device technology, mSurvey, a major mobile SMS-based survey company, is widely known and an example of a relatively successful and scaled operation. This company's service enables survey designers to build surveys based off of a repository of willing participants, sorted by a number of key characteristics. These attributes enable these survey designers to gain insights based off presorted statistical groups. Unfortunately, there are limitations to SMS-based survey methodology. First, in

order to gather additional data regarding the survey beyond the call, such as spatial location, agreements have to be made with telecom providers in order to enable the use of their services for such operations (Griffith, 2014). This bottleneck makes quickly deploying surveys in new environments and countries prohibitively difficult if implementers are not already in an arrangement with the survey provider, such as mSurvey. Furthermore, even when spatial location is possible, it is not at a high resolution. For example, in a recent survey performed in Trinidad and Tobago, a survey was performed in conjunction with the local telecommunications authority. The “resolution of data” achieved was only at the regional scale (as shown in Figure 2.1). For example, knowing what street the survey respondent was at was not possible. Finally, because SMS surveys are performed in a “call and response” method, wherein participants must wait for each question to be texted to them, quick and in-person interviews are not a reasonable use case for the tool.

Figure 2.1: mSurvey Trinidad and Tobago Digital Access Survey

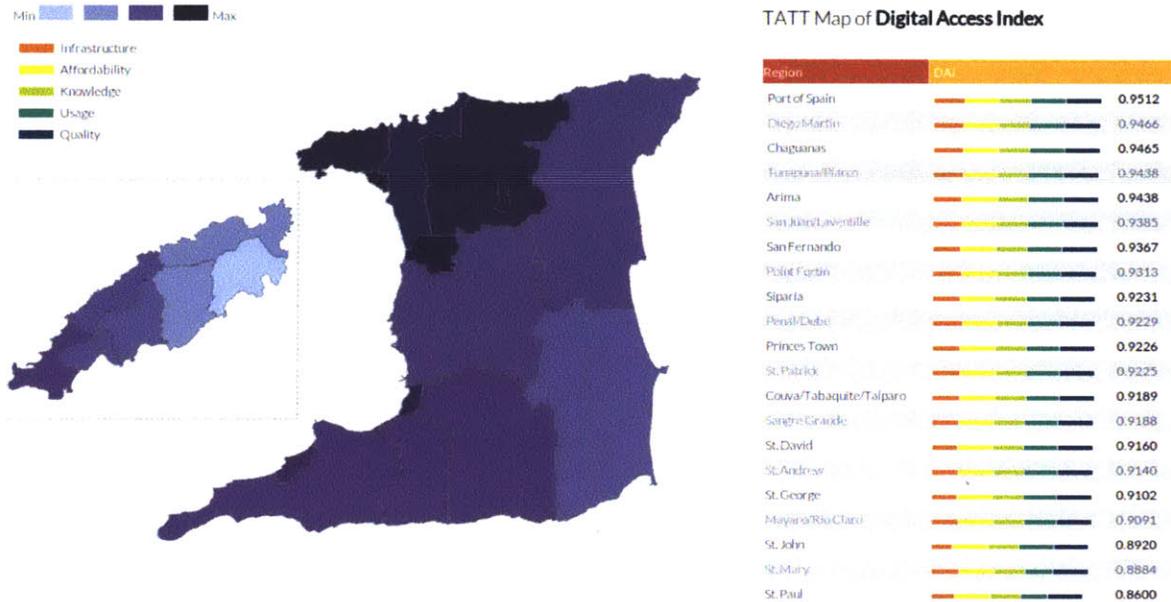


Image above is from mSurvey promotional materials demonstrating results from a survey performed for the Telecommunications Authority of Trinidad and Tobago.

### Commercial Online Tools

While the focus of this synopsis is specifically in-field survey solutions, a brief acknowledgement of static, browser-based surveys is necessary as a segue towards modern smartphone adaptations of the pen and paper survey. Most notable of these static services are Survey Monkey and Google Forms. Survey Monkey is but one of a myriad of software-as-a-service (SaaS) companies now doing business on the web in the field of digital surveying. Alternatives include competitors such as SurveyGizmo, Fluid Surveys, QuestionPro, Survey Analytics, Survs, and User Report. Survey Monkey, started in 1999, helped pioneer and popularize the field of

online survey prescription. Today, key package services that are offered bundled with commercial survey software include:

1. A graphic user interfaced (GUI) designed to visualize the survey creation process (enabling non-technical users to input queries intuitively into the survey architecture);
2. Curating of registered and willing survey participants by a number of characteristics; real-time visualizations of survey results and basic statistical analyses;
3. The ability to create custom data feeds to export live data results through an application programming interface (API);
4. And the presence of a robust security service that ensures data safety.

Unfortunately, all services ultimately come with limitations. Most browser-based survey software lacks the geospatial capabilities and in-field usability of a mobile-based (i.e, phone, tablet) application designed specifically for such uses. Operating the browser-based version of these applications requires a live internet connection and the downloading of content that would otherwise be locally stored (as in a standard smartphone “App”). As a result, these data-heavy interfaces are not optimized for the visual and energy-related limitations of a smartphone. That said, Survey Monkey and some of its competitors, such as QuestionPro and FluidSurveys, have made significant gains in the mobile surveying area and now provide a mobile phone version with the functionality of some elements of their desktop software. Ultimately, though, these services’ development have been

focusing on the more lucrative market of business and commercial analyses and solutions, as evidenced by the 2013 release of Survey Monkey's Enterprise platform. As a consequence, the private industry's development trajectory has not developed in the direction of survey operationalization in poorly connected environments, lean software, and geospatial specificity.

### **Google Labs Developments: Google Forms, Fusion Tables, and API**

Roughly concurrent to Survey Monkey's development has been the evolution of Google Docs and, specifically, Google Forms. Google Forms is essentially a visual interface for spreadsheet entry, placed over Google Docs' Google Spreadsheet software within the Google Docs productivity suite. In addition to a visualized data entry component, the recent Google Labs product development of Fusion Tables allows for graphics developed off the table data (e.g. pie charts, plots, graphs) to be visualized within a browser calling to Google's Fusion Tables API. While involving a greater learning curve than services such as Survey Monkey, the ability to visualize survey input data live is nonetheless possible.

The ability to enhance static spreadsheet information through embedded visualizations with Google's Fusion Tables API can allow for greater control over how data is visualized and curated than "freemium" sources, such as Survey Monkey. These "freemium" options restrict capability severely on the SaaS product, with the requirement that certain payments be made to "open up" additional

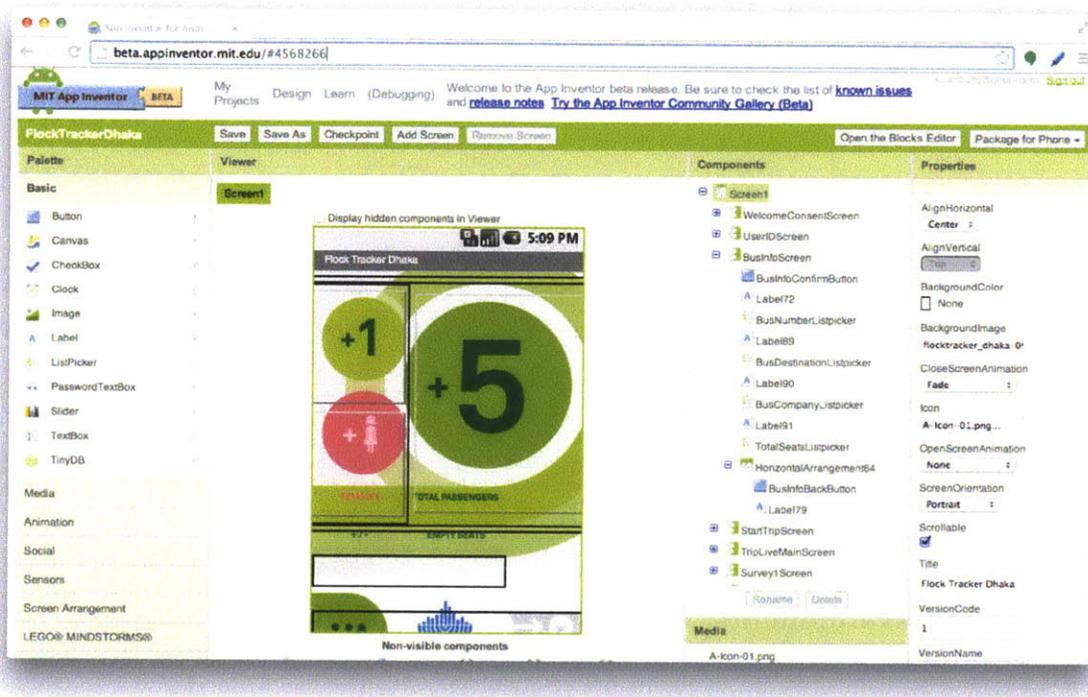
features. Google's service is, indeed, a "freemium" model as well, though the restrictions are comparatively flexible. In fact, they deal solely with the volume of information being transferred through Google's service, and not on the capabilities of the service itself. Thus, while Google's offerings are less tailored, their capabilities are broad so as to enable them to be a functional and highly customizable free replacement for some of the offerings that Survey Monkey and its competitors provide. Such services are valuable in that they unveil the potential for a more lightweight, basic solution to surveying than the more robust, yet more limited commercial SaaS survey offerings. That said, it does require a higher level of technical expertise.

### **Google Labs Developments: App Inventor for Android**

In addition to new offerings through Google Docs which enable the development and visualization of databases more easily through graphic interfaces; Google, jointly with MIT, has developed software that assists non-technical individuals in creating small applications. These have the ability to tap into Docs-based resources via devices running an Android operation system (OS). This product, App Inventor for Android, has an intertwined history between Google and MIT. Currently under the care of MIT, the project has been stewarded by both institutions, back and forth, since its conception in late 2010. App Inventor's primary purpose is to turn the act of programming into a simplified, visual interface where functions are created by

linking jigsaw puzzle-like pieces. Thus, only logical connections can occur because those pieces “fit.”

Figure 2.2: App Inventor for Android developer interface.



Shown is the in-browser user interface for App Inventor for Android. The version shown here is App Inventor Classic (AI1). A new edition of App Inventor (AI2) was released at the end of 2013.

App Inventor’s capabilities rely on a number of preprogrammed meta functions. With these functions, a user can string them together to define objects and functions. These results can create actions on screen or activity such as uploading data to a data repository like Fusion Tables. These also have the option of interacting with a back end database, via built-in support for Google’s Fusion Tables API. As a result, data gathering processes can be designed so that selections

on the application correlate with data entry into a specified Google Fusion Table. App Inventor for Android represents a valuable tool for early programmers, though its capabilities remain limited in terms of developing complex applications, handling multiple processes simultaneously.

Figure 2.3: App Inventor for Android Blocks Editor

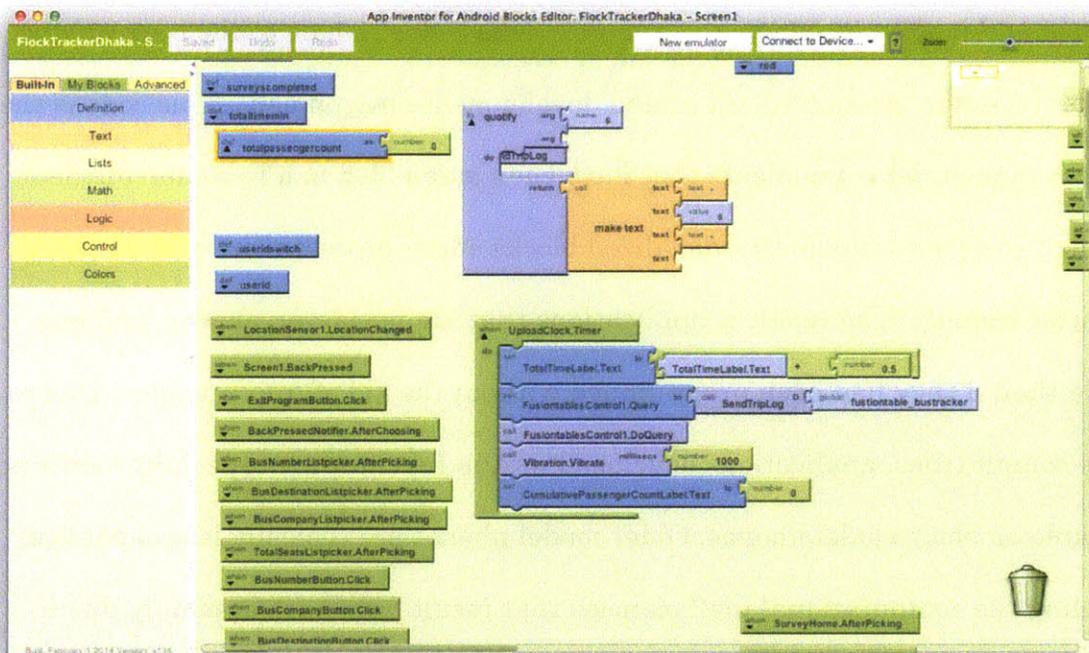


Image shows user interface for the “programming” component of App Inventor for Android. The interface is a downloadable program that operates simultaneous with the in browser screen. On this screen, the capabilities of physical items from the in-browser user interface are defined.

Paired with Google’s Fusion Table back end, App Inventor for Android can be a relatively simple solution to configure a mobile digital device to perform as a data-gathering tool. Unfortunately, App Inventor is still in its beta form and its present

limits can become burdensome when attempting to develop a more complex application. For example, data intensive actions, including frequent uploading and downloading, are significantly slowed on applications developed with App Inventor. Similarly, the built in GPS function has not been designed to be repeatedly called, especially not while other tasks are operating. Thus, designs that involve frequent and timed actions occurring automatically in the background, such as uploads of latitude and longitude, can be too complex. This is due in part to the code produced by App Inventor through its simulated development environment. This code is the product of modular components that have been assembled in a less than efficient manner, in order to make the visualized blocks editor process more standardized and user friendly. The result is applications that are far larger, slower, and less stable than if they had been programmed without the aid of App Inventor. Due to these complexities, applications developed in App Inventor are especially slow and unstable on older model phones. Older model phones are typically less capable of handling the computational inefficiencies that result from such assembly due to their limited memory and processing power.

App Inventor is ultimately a tool limited by the extents to which it has been programmed. As it is still in beta status, many functions that are easily included when programming have not yet been built in App Inventor, or have yet to be completed and refined sufficiently. As an introductory tool, App Inventor is valuable to help individuals become familiar with the logic structures necessary to program. That said, in order to develop a stable application capable of running numerous

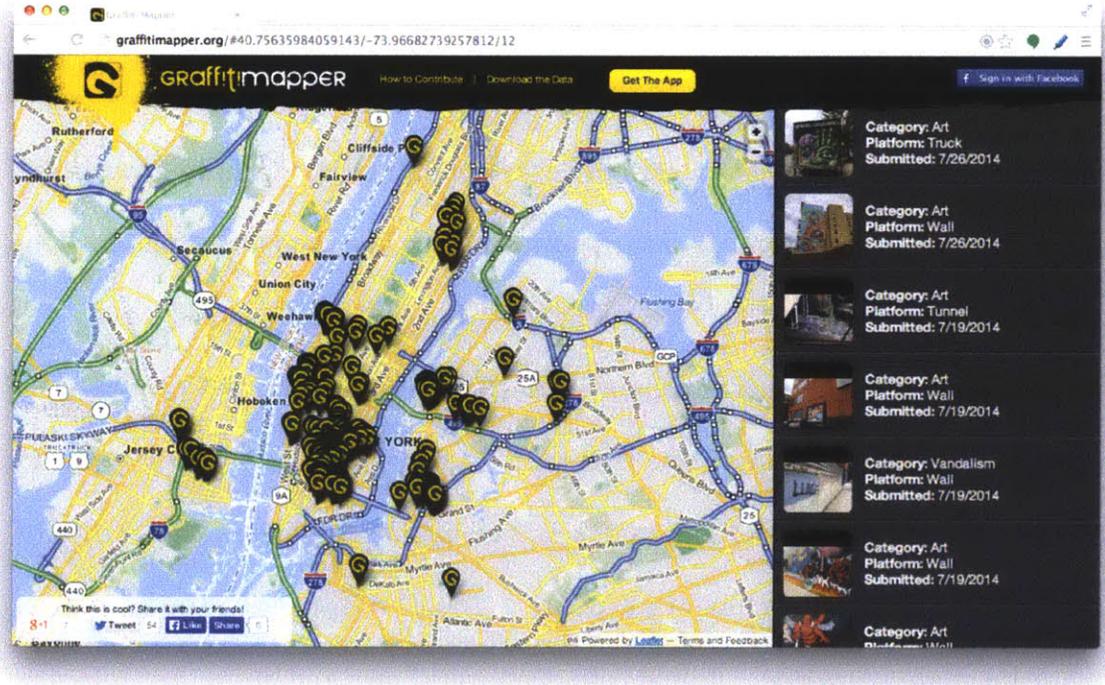
automatic, background tasks, without interfering with the performance of, for example, an interactive survey, an application needs to be programmed in Java, rather than via visual constructionist learning programming environments, such as App Inventor for Android.

### **Emerging Software Solutions for Mobile Survey Technology**

In October of 2011, software developer Spatial Network, Inc. from Clearwater, Florida released a product called Fulcrum. While Fulcrum does not appear to be documented as a tool used in any academic studies, it has maintained a strong user base and has been employed by users in dozens of industries; including utilities, public works, transportation, environmental, engineering, and construction. This figure is based on brief correspondence between Coleman McCormick, Vice President of Spatial Networks, Inc, and myself in early August. McCormick explained the project had developed “years ago” to support the company’s own “data collection projects, building GIS datasets and using mobile devices [and] GPS for field verification [and] enrichment purposes. ... [Spatial Networks, Inc.] decided to take our intimate knowledge of field collection workflows and build that into a product, which became Fulcrum (McCormick, 2014).” Fulcrum has been used successfully as the base platform for a number of social, crowd-sourced mapping initiatives, including Graffiti Mapper, a service that retains and presents uploaded photos of graffiti that are tagged to locations on a world map. Unfortunately,

Fulcrum lacks some features that are necessary to performing in-vehicle monitoring, such as the ability to maintain counts correlated to the same trip identification, at regular intervals. An example of this would be measuring ridership in a vehicle by gender, with the ability to add or subtract to the rider count at any time and have the new figures be updated and tagged to a specific trip, point in space, and time.

Figure 2.4: Graffiti Mapper

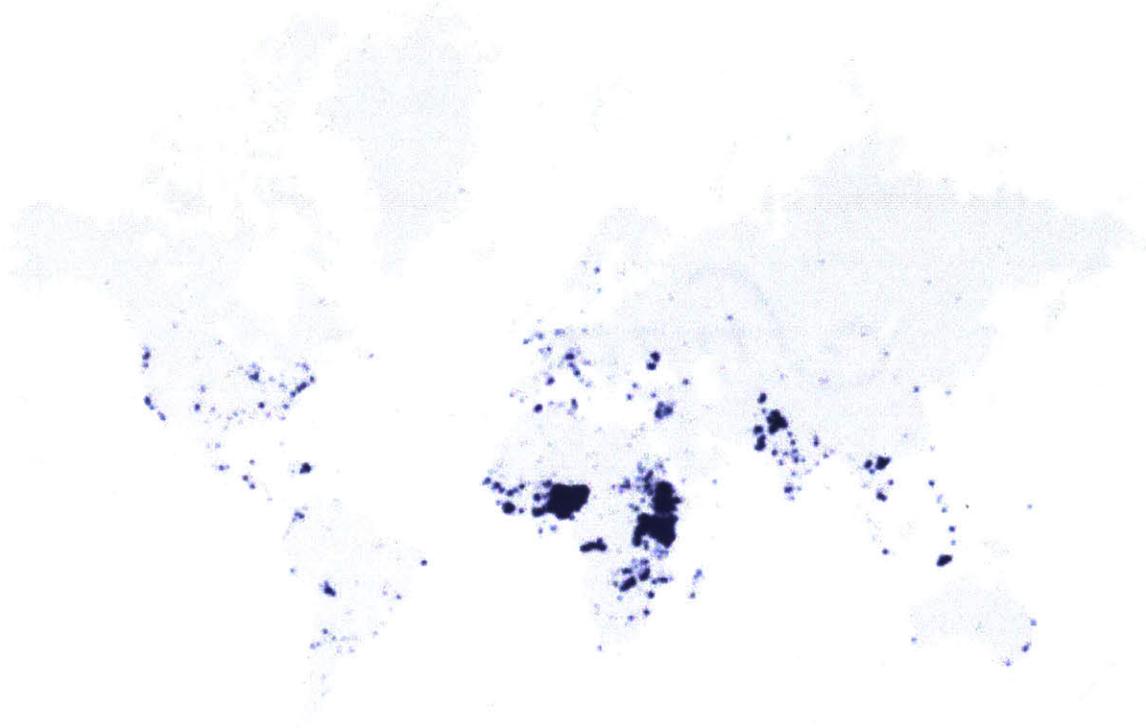


Above image shows screen capture of Graffiti Mapper web interface. Website accessed August 12, 2014.

Numerous comparable solutions to Fulcrum have been developed, with varying levels of quality. One such product that has been particularly successful is an open

source application, Formhub, free to users. In 2012, the Modi Research Group and the Earth Institute at Columbia University in New York was working with government and non-governmental organizations in Africa to gather data on a number of projects. In order to optimize the process, an internal technical solution was developed to perform the task of improving survey methods, as well as ensuring quality of data. As work continued, there was a desire to convert the technology to a more scalable solution. A first build of Formhub was completed and released later that year. Formhub includes a data and project management interface, including the ability to observe results in real time on color-coded charts and maps. Additional applications have been built on top of Formhub in the past few years, including Enketo Smart Paper (<https://enketo.org/>). This program essentially takes the core structure of Formhub and adds an enhanced user interface. Similar repackaging of the core architecture is found also in KoboForm within KoBo Toolbox (<http://www.kobotoolbox.org/>), from the Harvard Humanitarian Initiative.

Figure 2.5: ModiLabs' first million survey submissions.



Modified image from Modi Labs MapBox June 2013 visualization of its first one million Formlab submissions (Formhub, “One Million Surveys Visualized”). Image retrieved from Modi Labs’ Formhub blog.

While Formhub does employ Django (a high-level Python web framework) as a backend solution for its data management; it is predicated on a relatively old standard for the process of data conversion. The technology Formhub and other similar offerings employ to handle form conversion into surveys is known as Open Data Kit (ODK). ODK (<http://opendatakit.org/>) was originally developed by Google.org in the spring of 2008 by Gaetano Borriello. Future development occurred jointly with the University of Washington’s Department of Computer Science and Engineering. The premise of the system was to enable form conversion from non-

technical formats into an XML readable by hardware running an Android operating system. ODK is maintained at the University of Washington's Department of Computer Science and Engineering and members of a multidisciplinary university group known as Change (<http://change.washington.edu/>). This project is maintained financially through the hosting organization, presently the University of Washington, and continued funding from a Google Focused Research Award, as well as other donations.

Formhub's reliance on ODK has created some unfortunate complexities, resulting in an obtuse data structure. Presently, the Django backend actually needs to compile XML forms into a 1x1 data frame and send that out to ODK to parse and return. This reliance on ODK is overly complex, and a critical bottleneck in a service that could, in theory, exist independent of ODK. Formlab is currently experimenting with shifts away from its current backend framework; a GitHub account has been established under Columbia's Sustainable Engineering Lab for "formhub.R" (<https://github.com/prabhasp/formhub.R>). It has an interest in developing a relational database solution to replace the current Django and ODK-based framework.

Additional complexities with the current database include the necessity of pulling all data from a data set in order to perform calculations. This makes the live visualizations of survey statistics far more convoluted than they need be. Such methods are computationally intensive and put the platform at risk when operating at scale, as the application is presently.

Although such limitations do occur at the scale of millions of surveys, the key contribution Formhub has made to digital surveying is its ability to provide latitudinal and longitudinal data, as well as a timestamp, on data uploads. Furthermore, it accomplishes this, complete with private accounts for data management, for free to users. This key innovation enables even small-scale surveys to tap into existing data sets and correlate to other spatial information. The application's geospatial tagging is particularly valuable when importing and correlating other spatial data, such as those associated with census data or neighborhood crime statistics, into a generated dataset from Formhub surveys.

### **Flocktracker: Early Development in Dhaka**

Roughly concurrent to Formhub's development at Modi Labs, a project was initiated as part of the Singapore-MIT Alliance for Research and Technology – Future Mobility (SMART-FM) to identify mobile-phone-based paths to sustainable mobility in the global south. Albert Ching, a Masters student within the Department of Urban Studies and Planning (DUSP), began initial fieldwork in the summer of 2011. A report released just prior to this fieldwork, in May of 2011, provided initial identification of real-time information-based mobility innovations arising in para-transit services in other parts of South East Asia (Indonesia, Thailand, Malaysia). The goal of this was to understand the technology adaptation

process among such service providers and ultimately determine how innovations developed by SMART-FM might be appropriately transferred to such contexts.

After arriving at Dhaka as the city of focus, Ching's research spent the Fall of 2011 on beginning to understand App Inventor. Part of this process involved developing prototypes of software. The idea for mobile-based bus information (as a replacement AVL) came out of design charrettes (with residents of Dhaka, primarily from Kewkradong, a local bicyclist and urbanist activist group) and rapid prototyping done in Dhaka in January 2012. Another Masters student from DUSP, Stephen Kennedy, participated as well. An initial mapping exercise to operationalize the technology in-field was performed in the Spring of 2012.

One outcome of the research is the concept of "Flocksourcing" (Ching, 2012). This concept is described as an "experimental data collection technique where users become the sensors to generate a large amount of information that improves a public service like public transport (Ching, 2012, p. 3)." By taking users and using them as "sensors" through the incorporation of smartphone technology; the data collection process can be reimaged. While the concept of outfitting humans so that they become sensors has been discussed prior, its implementation has been limited. Thomas Erickson, a designer and researcher in the Social Computing Group at the IBM Watson Research Center in New York, acknowledged this potential as a development that had yet to fully form. Erickson commented in a 2010 paper that "tapping human intelligence to augment the intelligence of cities" in an effort to make cities "smarter" will, in time, become more legitimate (Erickson, 2010, p. 1-2).

Perception data – independent voices and opinions – can gain additional value when enhanced through a digital lens. For example, Flocktracker’s efforts demonstrated that a survey can now include the point and time at which it occurred, similar to the approach of the Formhub development team, in regards to improving data acquisition. In gathering large amounts of data for analysis, it is valuable to digitize the method of acquisition not only to create easier and more automated methods of data entry, but also to tap into the capabilities of modern mobile devices and capitalize on the increased amount of information one can generate from the same survey effort.

Figure 2.6: Flocksourcing Diagram

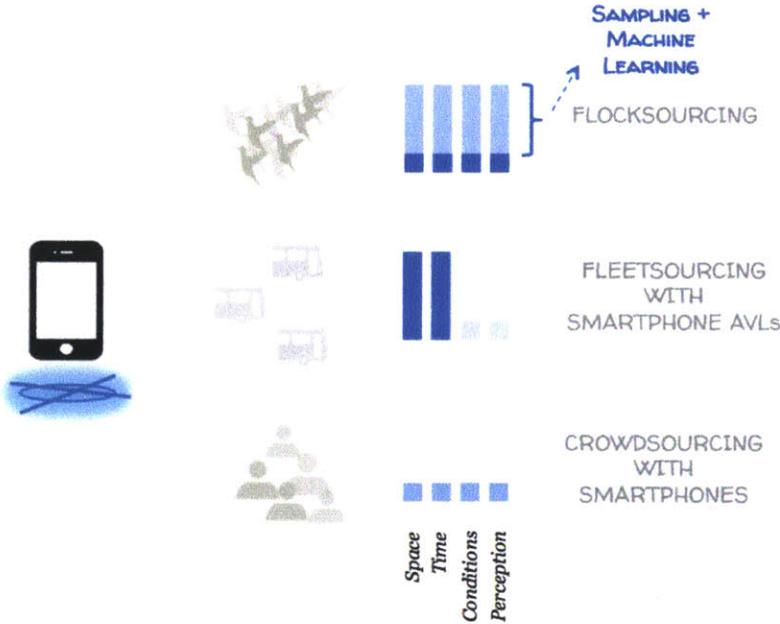
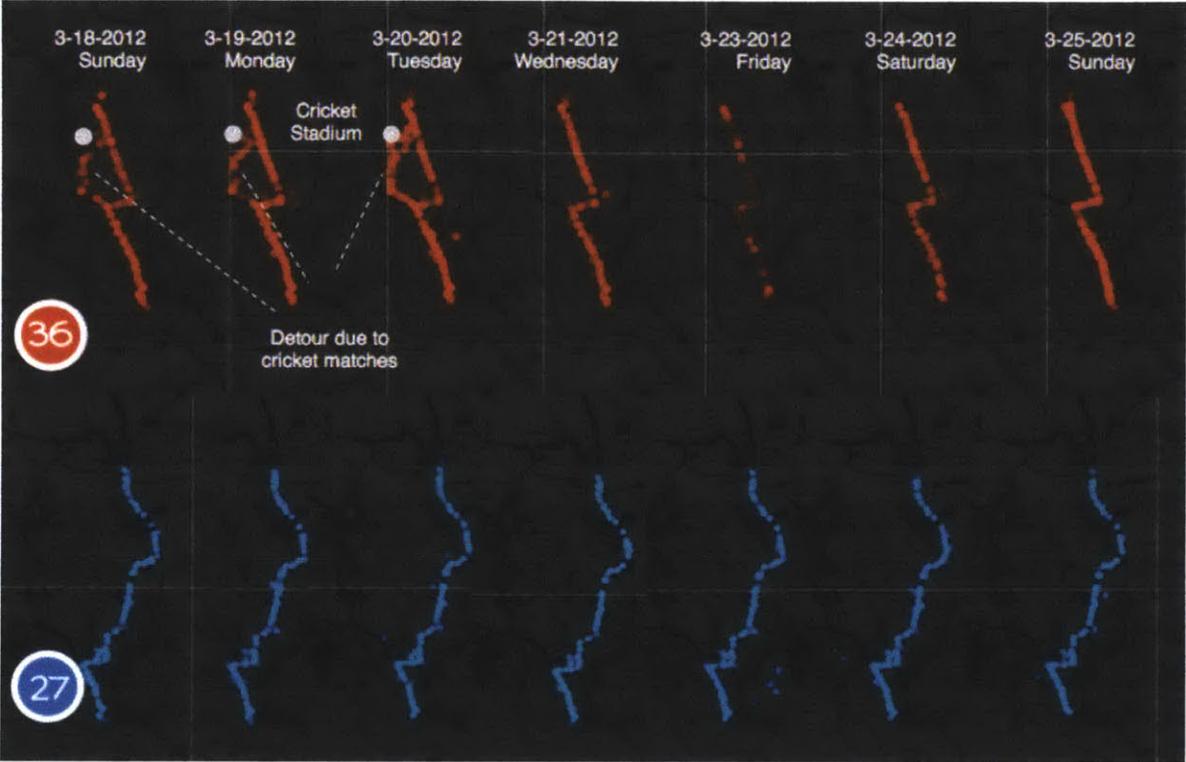


Diagram from Albert Ching’s thesis demonstrating the potential of smartphones as monitors of both fleet operations for bus systems and user surveys (Ching, p. 42).

What separates the Flocksourcing research from Formhub's solution is its focus on in-vehicle surveying. Flocksourcing was designed with a specific situation in mind: public transportation monitoring. Although it is completely functional in out-of-vehicle survey situations, the inclusion of what is often referred to as "bread crumbing" allows the application to update two live datasets, one for the surveys and one for the exact location of the application. "Bread crumbing" is the act of creating spatial data traces at regular time intervals. For example, if the latitude and longitude of the applications location are reliably uploaded every thirty seconds, the resulting data set can tell you not only the route the vehicle took, but also its speed, thus assist in tasks such as identifying congestion points along the route. One data repository contains information related to survey responses while the other contains information related to the trip. The two databases are then capable of being joined by a custom, random Trip ID that is generated at the beginning of each trip and included in each upload to either of the datasets. Data points can be arranged correctly after joining by Trip ID and sorting the resulting points by time. The "bread crumbing" strategy was initially selected to address the expensive nature of automated vehicle location (AVL) devices as used in public transportation services, which can cost upwards of \$15,000.00 (Ching, 2012, p. 36). Flocktracker's design seeks to build off its potential as a smartphone-based AVL replacement, enabling such services to be integrated with other capabilities, in particular onboard surveying and passenger counting, and providing the capability

for such data to be linked to spatiotemporal characteristics derivable from other sources (e.g., censuses).

Figure 2.7: Bus trace visualization.



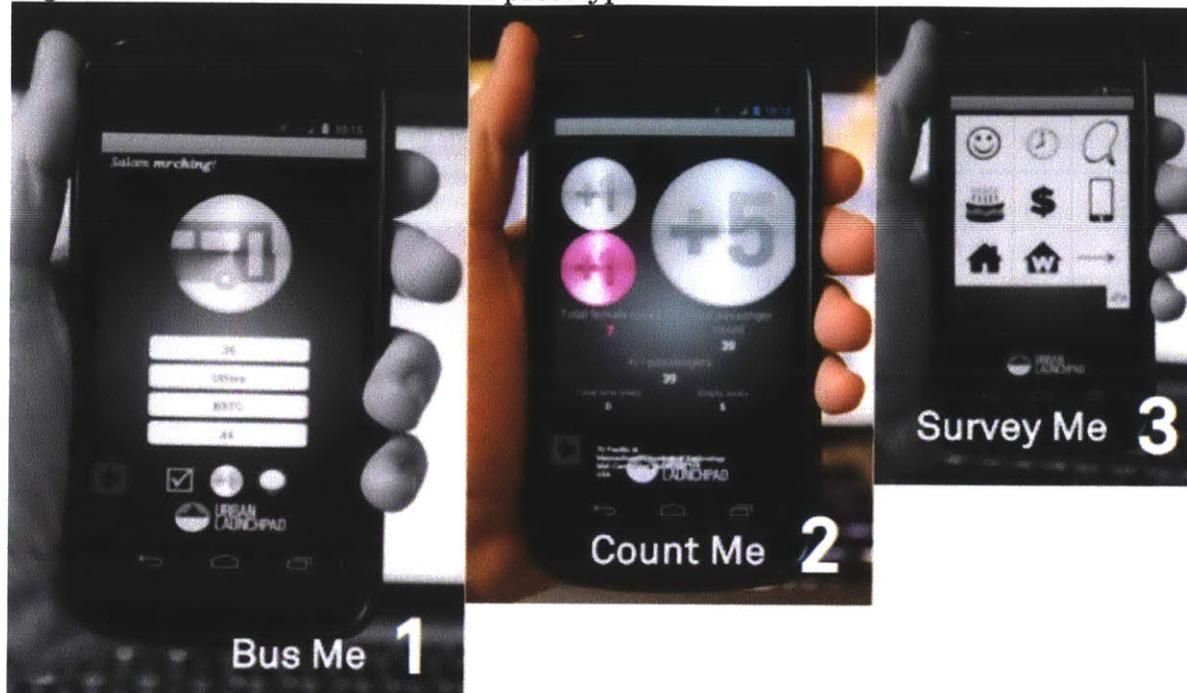
Traces from locational uploads over a one week-long in-field test operation in Dhaka, Bangladesh (Ching, p. 64).

In Albert Ching’s thesis, a general architecture for the application was outlined, with a backend that relies on Google’s Fusion Tables. The build outlined in Ching’s thesis is based on App Inventor for Android and employs the predefined functions made available through App Inventor for Android’s Blocks Editor software. The key concept of this build is that the back end of the software would involve two Google Fusion Tables. The first table’s purpose would be to receive and retain uploaded

survey responses. The second table's purpose would be a repository to keep track of the device's latitude and longitude. This information would be uploaded at a predetermined set interval (e.g., every thirty seconds), and the function call (in addition to its actions) would be controlled by the application itself. This system of two tables allows for information, such as bus crowding and speed to be observed over space and time, as well as correlated with survey results.

The original intent of the prototype run was modest, covering two bus corridors in Dhaka: "Identify the deployment scale required to develop an accurate portrait of the two routes' primary operating characteristics, including speeds, crowding, adherence to routes and their variation across times and days. In addition, we wanted to test the viability of using the technology for carrying out onboard surveys. The plan was to cover the routes for a week, aiming to achieve a minimum of 120 one-way vehicle rides, in total" (Ching, 2012, p. 61).

Figure 2.8: User interface for 2012 prototype.



Initial design of the application contained three primary “screens:” an initialization screen that contains details regarding trip; a main hub counter monitors the number of riders in the vehicle; and a survey hub.

Results from the initial survey piqued interest in the application’s potential: “The members of the flock ultimately carried out 270 rides, recording over 10,000 passenger counts and bus location points. In addition, they used the devices to survey, onboard, over 1,000 users. Three Flocksourcers rode the bus for over 2,000 minutes in that period, equivalent to over 33 hours” (Ching, 2012, p. 61). The pilot applications, although effectively demonstrating the potential of the approach, revealed limitations in the original app, and App Inventor in general. These, and their subsequent remedies, will be discussed in greater detail in Chapter 4.

## **CHAPTER 3: St. Louis, Missouri Case Study**

### **St. Louis in Comparison**

The study and selection of St. Louis is spurred by interest regarding a very particular urban condition. This condition is that of the shrinking, post-industrial, mid-sized Midwestern city. While some cities have seen gains in recent years, particularly at the scale of the metropolitan statistical area, the urban core has tended to experience either stagnant growth or a shrinking population, with these characteristics exacerbated most severe in the city center.

Shrinking Rust Belt cities that have experienced a population loss greater than one-third in the second half of the twentieth century, and of a population greater than 250,000 today are: Baltimore, Buffalo, Cincinnati, Cleveland, Detroit, Newark, Pittsburgh, Rochester, St. Louis (US Census, 2011). Within this group, those that fall under the consideration of Midwestern would primarily be Cincinnati, Cleveland, Detroit, Pittsburgh, and St. Louis. Buffalo and Rochester, of upstate New York, arguably bear great similarity to Midwestern urban conditions, rather than to the Eastern seaboard's cities. That said, the urban condition of the cities of upstate New York may likely be considered as a case particular to themselves, much in the way that Baltimore's condition is unique, given its proximity to robust economies such as that of Washington, D.C. (Gordon, 2008).

Table 3.1: Metropolitan Employment figures.

Metro Area	1998			2006			% Δ		
	Tot. Jobs (35m radius of dt)	Share of Jobs Within 3 Miles	Share of Jobs 3 to 10 Miles	Share of Jobs Beyond 10 Miles	Tot. Jobs (35m radius of dt)	Share of Jobs Within 3 Miles		Share of Jobs 3 to 10 Miles	Share of Jobs Beyond 10 Miles
Cincinnati	876,497	20.4%	31.4%	48.2%	917,480	17.1%	30.6%	52.3%	4.1%
Cleveland	1,000,374	18.0%	38.9%	43.1%	959,388	16.2%	38.1%	45.7%	2.5%
Detroit	1,787,063	7.3%	18.5%	74.2%	1,679,362	7.0%	15.7%	77.4%	3.1%
Pittsburgh	973,150	27.0%	29.7%	43.4%	1,008,801	25.9%	29.1%	45.1%	1.7%
St. Louis	1,125,324	15.3%	27.0%	57.7%	1,178,589	14.1%	25.0%	60.9%	3.3%

Adapted from “Job Sprawl Revisited: The Changing Geography of Metropolitan Employment (Brookings, 2009, p. 19).”

Colin Gordon, in his book “Mapping Decline,” touches briefly on this concept of the particular case of the shrinking, post-industrial, mid-sized midwestern city (Gordon, 2008, p. 222-225). While little research has been performed to categorize these cities into a unique grouping, they tend to emerge clustered in various analyses on American cities. Thus, their commonalities yield, for example, similar spatial behavior in terms of job location as reported by Elizabeth Kneebone in the Brookings Metropolitan Policy Program’s Metro Economy report on job sprawl (Kneebone, 2009, p. 12).

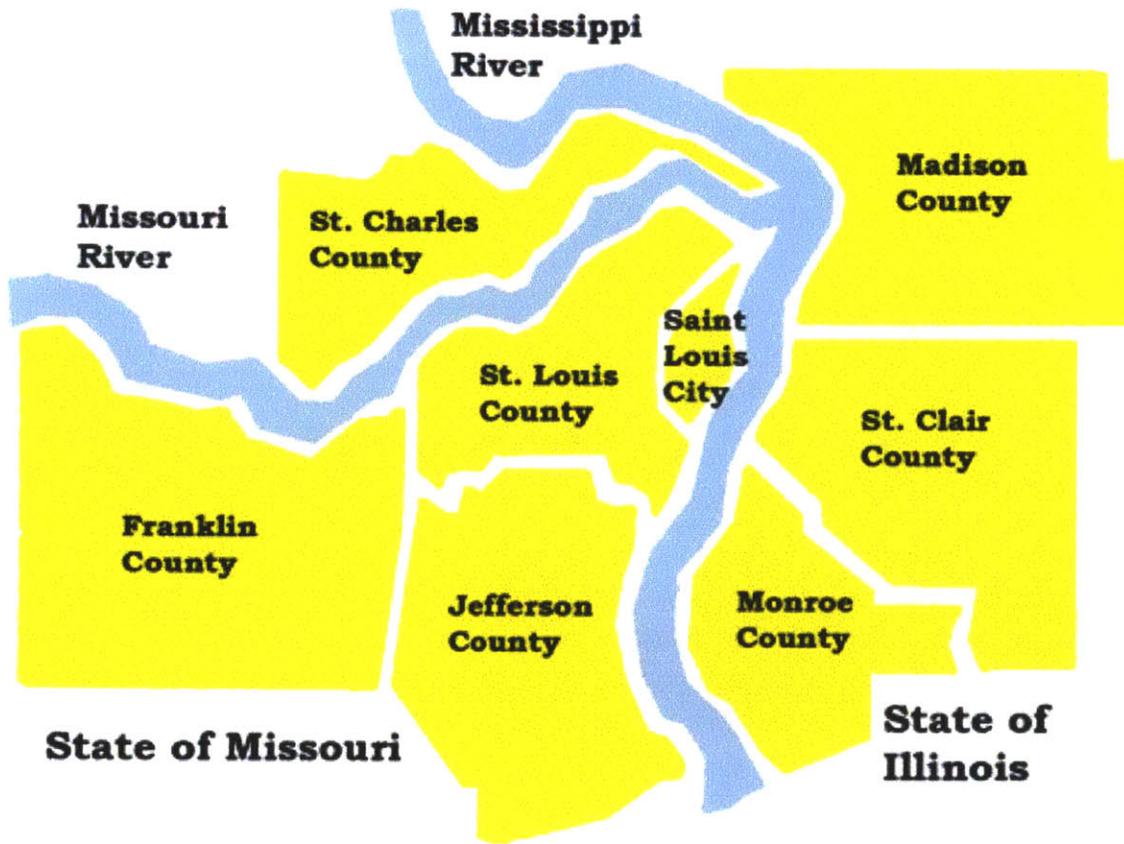
## **Introduction to the City of St. Louis**

St. Louis is a mid-sized city in the Midwestern USA, located along the eastern border of Missouri, at the shore of Mississippi River. It was primarily established as a fur trading post, and grew thanks to its strategic position just fifteen miles south of the confluence of the Missouri and Mississippi Rivers. This location served to fuel substantial growth in the late 19th and early 20th century, when river access, as well as substantial rail infrastructure investment served to fuel the city's fortune. While one of the key commercial hubs within a metropolitan statistical area (St. Louis, MO-IL) of nearly 2.8 million, today St. Louis itself only boasts, as of 2012, a population of 318,172 (St. Louis, 2014). This population size is in part due to the 1876 establishment of St. Louis City's boundaries via a severance from the County, as well as the establishment of a home rule charter.

Figure 3.1A: St. Louis' Location within the United States

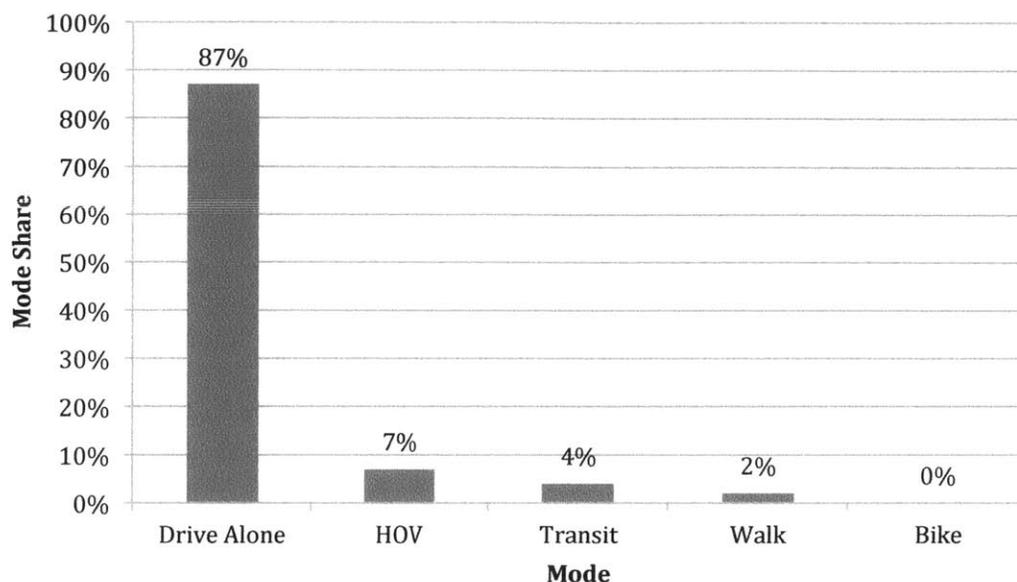


Figure 3.1B: St. Louis Area Counties, within State of Missouri and State of Illinois (Florida, 2014)



In addition to the historical separation of the City from the County, St. Louis has suffered from systemic population loss, due in large part to well-documented racial stigmatization. Chapter 2 of Gordon Levit's "Mapping Decline," titled "The Steel Ring: Race and Realty in Greater St. Louis," covers this topic in great depth. Systemic rezoning and restrictive deed covenants dominated housing sales and developments in the first half of the 20<sup>th</sup> century, cementing what is today a heavily segregated city and county (Gordon, 2008, p. 68-111). Segregation in St. Louis has been exacerbated by a stagnant economy and limited growth in recent years. Recent trends have been seen as hopeful signs of economic revitalization. For example, St. Louis' Cortex life sciences corridor was recently highlighted in a May 2014 Brookings report (Katz, 2014, p. 3). In spite of this good news, the city did lose another 10% of its population over the 2000-2010 period, a trend that has persisted since the mid-twentieth century (US Census 2004).

Figure 3.2: Mode share in the St. Louis Region



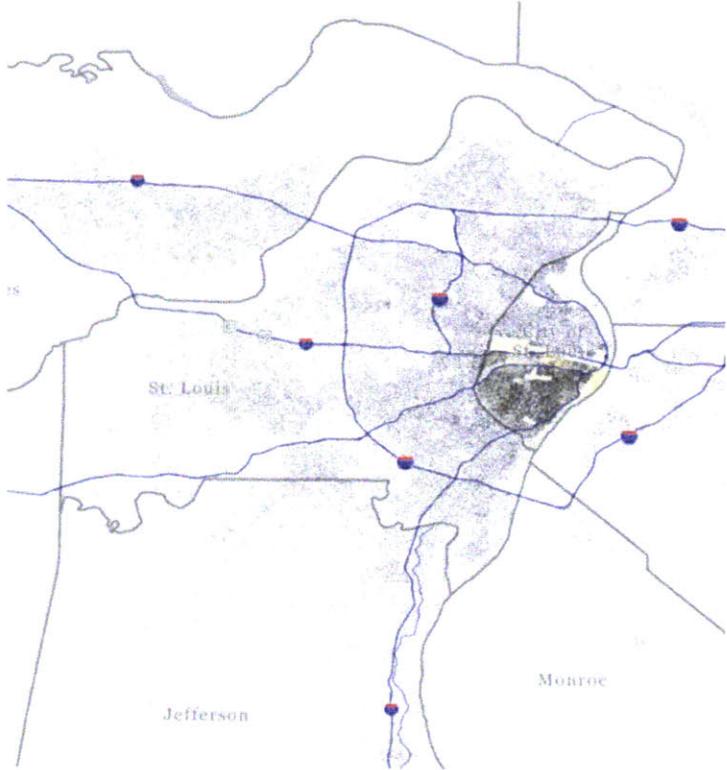
Regionally, St. Louis' public transit mode share sits slightly below the national average for the fifty largest metropolitan statistical areas (US Census, 2009, p. 8). That said, results from St. Louis' 2002 Household Travel Survey provided higher estimates on transit use for city residents specifically. When counting travel mode share for travel by location of home, in terms of county, results indicated that 6.4% of St. Louis City residents used transit. This contrasts with the County, a separate entity due to a city-county split, that reported only 1.2% of travelers using transit (East-West Gateway Coordinating Council, 2003, p. 8). Metropolitan areas exhibiting similar mode share include Sacramento, New Orleans, Cincinnati, and Louisville. Particularly in terms of Cincinnati and Louisville, St. Louis bears much in relation. Specifically, their built form exhibits a predominant legacy architecture, retained from a heavily industrial past, as well as severe cases of disinvestment and

urban abandonment, combined with significant suburban growth in the latter half of the twentieth century.

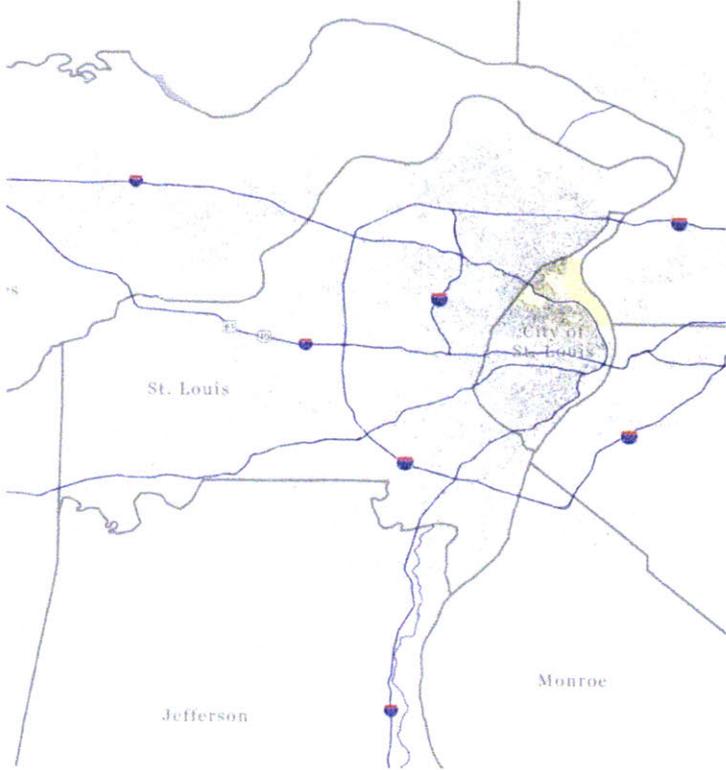
While the pull of commuters for work trips to both South and North St. Louis extend well beyond the limits of St. Louis City, intra-city commuters also exist, as shown the following Figure 3.3. That is to say, many living in North St. Louis travel to South St. Louis and relatively smaller numbers do so in the other direction, as well.

Figure 3.3: Residential location of workers in St. Louis.

**Residential Location of Low & Moderate Income Workers Who Commute Into South (Saint Louis, MO)**



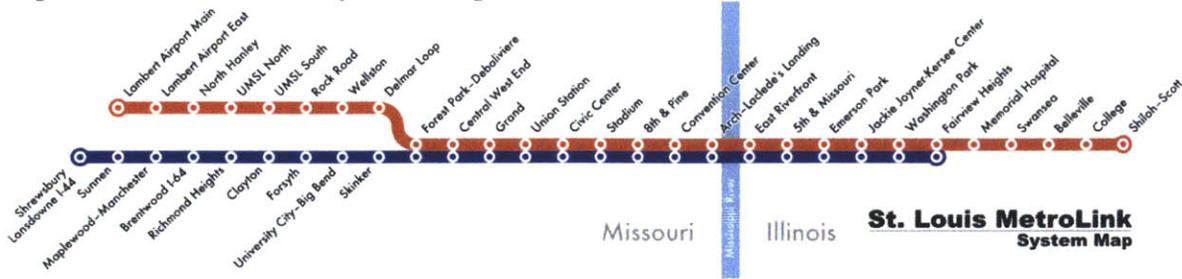
**Residential Location of Low & Moderate Income Workers Who Commute Into North (Saint Louis, MO)**



## **St. Louis Metro System**

Transit within the St. Louis metropolitan region is managed by the Bi-State Development Agency (BSDA). This agency is the product of a coordinated effort to service the St. Louis region, which spans across multiple political entities in the states of Illinois and Missouri. BSDA is funded through sales taxes from the City of St. Louis, St. Louis County, and the St. Clair County Illinois Transit District, as well as through federal grants and fare-paying customers. Through this funding, the agency operates MetroBus, special event MetroBus routes, MetroLink (the regions light rail service), and Call-A-Ride (point-to-point service designed for the elderly and handicapped). In addition, the agency is also taxed with the operation of the downtown airport, surrounding industrial parks throughout the city, tram operations inside the St. Louis Gateway Arch, and a number of other smaller operations.

Figure 3.4: MetroLink system map.



Formed in 1949, the BSDA was originally tasked with driving regional development. This broad task has led the BSDA to perform a wide range of tasks over its 60-plus year history, including a pollution risk study of the Missouri River, the formation of the Metropolitan Sewer District to handle regional waste infrastructure, the study to develop a modern port terminal at the Granite City Dock, and public transit operator. The role of public transit provider was not assumed by the BSDA until 1963 (Mueller, 2011). During the early 1960’s fifteen private operators operated transit within St. Louis City. The system had become chaotic and, as a result, the fifteen private systems were purchased and consolidated into a single system (Metro Transit) under the BSDA’s authority. Today, the Metro Transit system includes 46 miles of the MetroLink light rail transit line on which 87 vehicles operate; a MetroBus fleet of 387 vehicles operating on 75 routes, and a fleet of 120 para-transit Call-A-Ride vans.

Table 3.2: System Boarding

Boardings	MetroBus, Fixed Route	MetroLink	Metro System Total
	FY13	FY13	FY13
System	29,274,347	17,054,484	47,054,481
Missouri	26,255,224	13,732,038	40,600,563
Illinois	3,019,123	3,322,446	6,453,918

Table shows system boarding counts, by state, for fiscal year 2013.

Missouri stations make up a majority of all Metro transit system boardings. In 2013, Illinois made up one-fifth of all MetroLink boardings and roughly one-tenth of all MetroBus boardings. Buses make up the greatest portion of system-wide boardings; in 2013, MetroBus accounted for roughly 62% of all boardings, while MetroLink accounted for 36%.

Table 3.3: Fixed Route Boarding

	FY 2013	%Δ over '12	Avg. Day	Weekday	Saturday	Sunday
70 Grand	2,697,898	-3.52%	7,392	8,567	5,725	3,801
11 Chippewa	1,437,580	1.86%	3,939	4,485	3,286	2,162
95 Kingshighway	1,393,386	-1.87%	3,817	4,543	2,743	1,640
90 Hampton	1,186,149	-0.09%	3,250	3,747	2,553	1,723
74 Florissant	1,109,608	0.02%	3,040	3,510	2,517	1,476

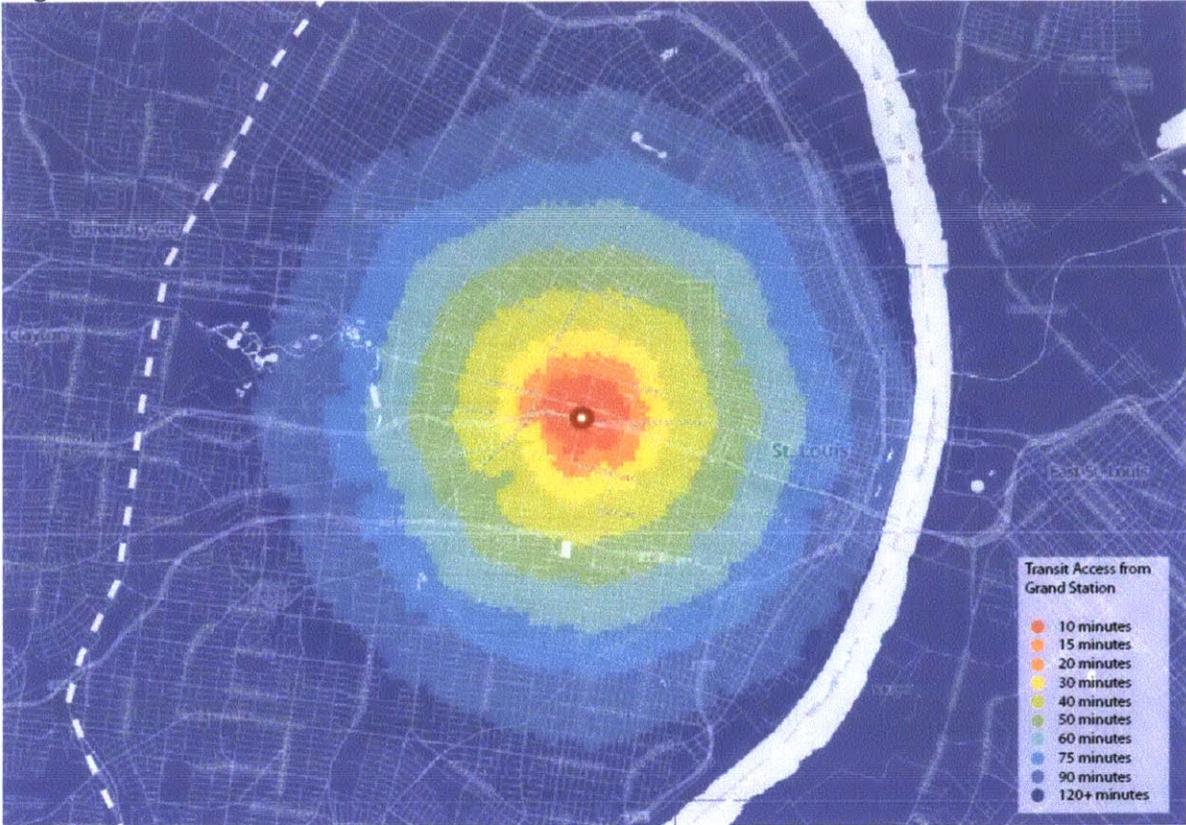
Boarding data for fixed routes shown for the five highest volume MetroBus routes in fiscal year 2013.

Of the routes in the bus system, roughly one-quarter (26.7%) of all boardings are made on five lines. Of these five highest volume lines, the north-south 70 Grand route is noteworthy as the highest volume route, with over 80% more boardings than the second highest volume route, 11 Chippewa in South City. Of the other four lines, two are similar to the 70 Grand in that they transect the city in a north to south fashion, while the other two straddle east to west their respective portions of the city; north side for the 74 Florissant and south side for the 11 Chippewa.

### **Urban Context of 70 Grand Bus Line**

The 70 Grand, one of the primary bus routes in the Metro Transit system, has recently been subject to a number of improvements, including the \$22 million reconstruction of the Grand Boulevard Bridge. This bridge, which re-opened in August of 2012, features enhanced transit facilities, including bus turnouts and an additional \$7 million dollar transit plaza, connecting the bus facilities to the MetroLink light rail station beneath it, as well as a 58-space park-and-ride area. More recently, the city has invested in a fleet of 15 new articulated buses to be deployed along the 70 route to handle increasing ridership and peak hour congestion. The first of these 60-foot articulated buses was placed into operation in early June. The remainder are planned to be operational by the end of the year (Hibbard, 2014).

Figure 3.5: Grand Center Transit Access



Transit travel time map demonstrating how a significant portion of the city is within one hour of public transit from the Grand Station MetroLink.

The 70 line’s significance is exhibited not just by recent capital improvement projects. For example, it inherited its numerical title from a streetcar that had previously operated along the route since September of 1859 (Hibbard, 2014). The streetcar’s legacy lives on both through the bus line and the significant institutional and cultural assets located along the Grand Boulevard corridor, where the 70 runs. In terms of ridership and frequency, the 70 line also stands out. Of the north-south Metro routes, the 70 Grand is the sole high frequency connector between the northern and southern portions of the city. The 90 Hampton runs at 25 minute

headways throughout the day, while the 95 Kingshighway route operates slightly more frequently, at 20 minute headways. By comparison, the 70 Grand runs at 12 minute intervals for much of the day.

Table 3.4: Top MetroLink stations, by boarding counts.

MetroLink Station	Boardings
Central West End	1,703,917
Forest Park	1,428,949
North Hanley	1,164,489
Civic Center	1,044,587
Grand Boulevard	752,836

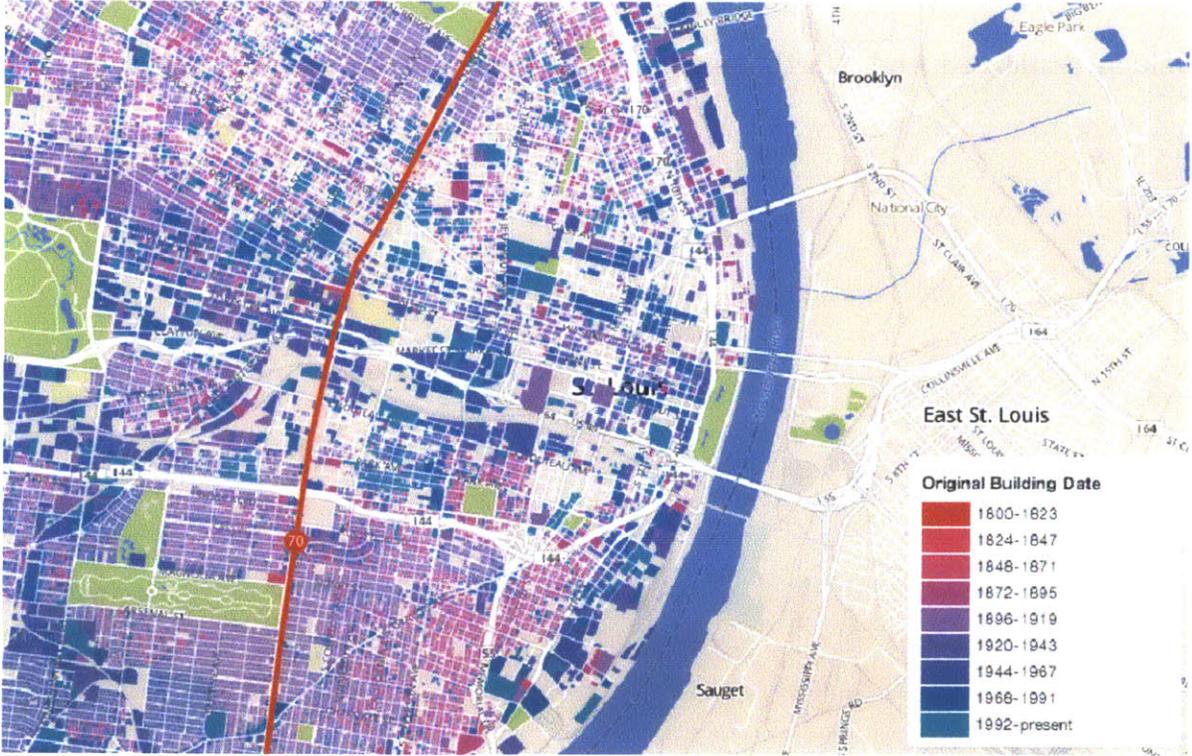
The cultural core of Grand Boulevard, Grand Center, is located in the central portion of the city, as seen in Figure 3.5. Grand Center boasts the The Grand Center District, an art and entertainment area featuring 12,000 theater seats, 12 museums, galleries, and 1,500 cultural events annually. This four-block area also sits adjacent to St. Louis University and Harris-Stowe State University. An estimated 1.5 million people travel to this area yearly for its arts, culture, dining, and other entertainment options. In addition, Grand Center rests just north of the Grand MetroLink station, one of the five busiest stations in the MetroLink system. Grand Center’s central location and numerous transit lines gives it a high degree of accessibility to much of the core of the city, as well as neighborhoods to the north and south, via public transit.

Figure 3.6: Grand MetroLink Station



Grand MetroLink station access from Grand Bridge. Construction on the \$7 million Scott Avenue Transit Plaza on the north end of the bridge was completed in August of 2012.

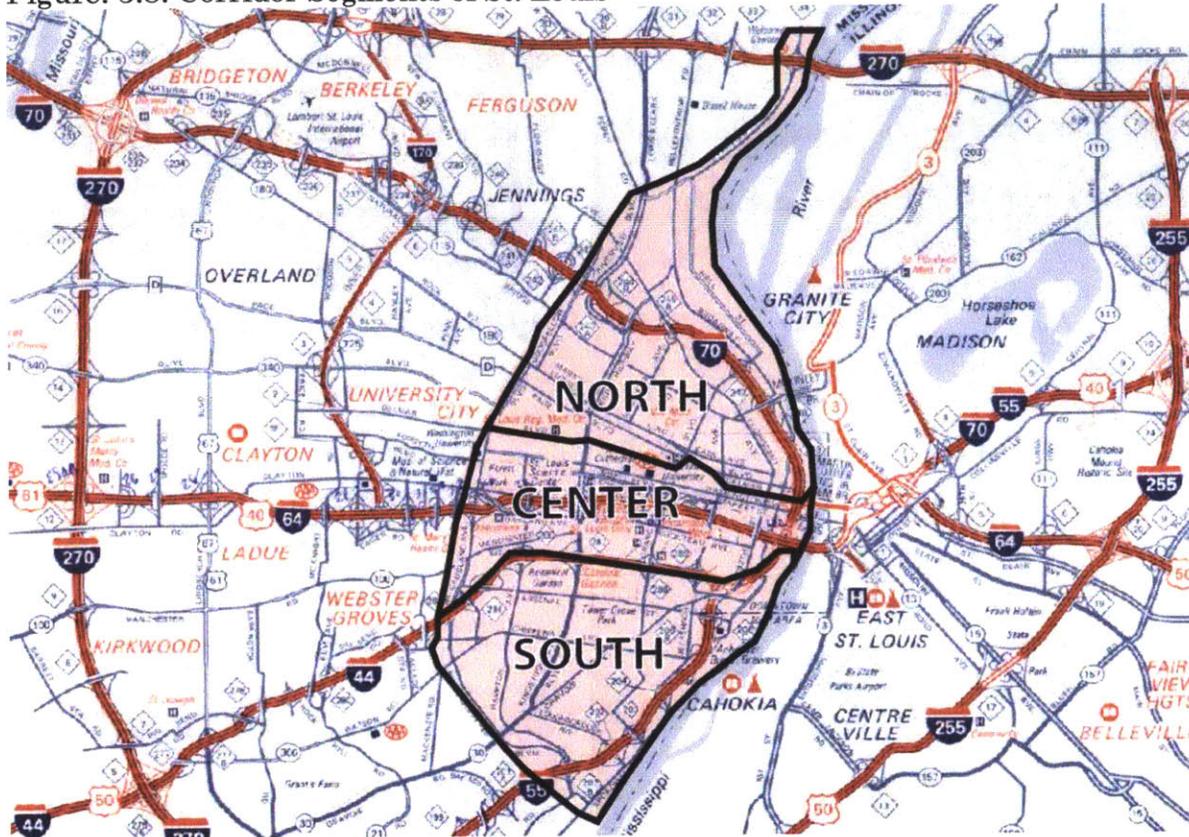
Figure 3.7: Building Age in St. Louis



70 line route highlighted in red. Route moves through some of oldest neighborhoods in city.

St. Louis housing stock is significantly more decayed to the north, whereas it has been relatively preserved in the south, particularly south of Interstate 44. The urban fabric is significantly deteriorated north of the Grand Center District, the prior discussed arts and culture area anchoring the central corridor of the city along the Grand Boulevard axis. The neighborhoods that the 70 Grand bus line directly passes through or abuts boast larger populations to the south than the north. To maintain consistency in description of where in St. Louis north, center, and south are; dividing lines will be established based off of those used by the St. Louis Planning Department. “South” is considered as all neighborhoods south of Interstate 44, “Center” as those in the corridor between Interstate 44 and south of Delmar Boulevard, and “North” as all remaining neighborhoods to the north.

Figure: 3.8: Corridor Segments of St. Louis

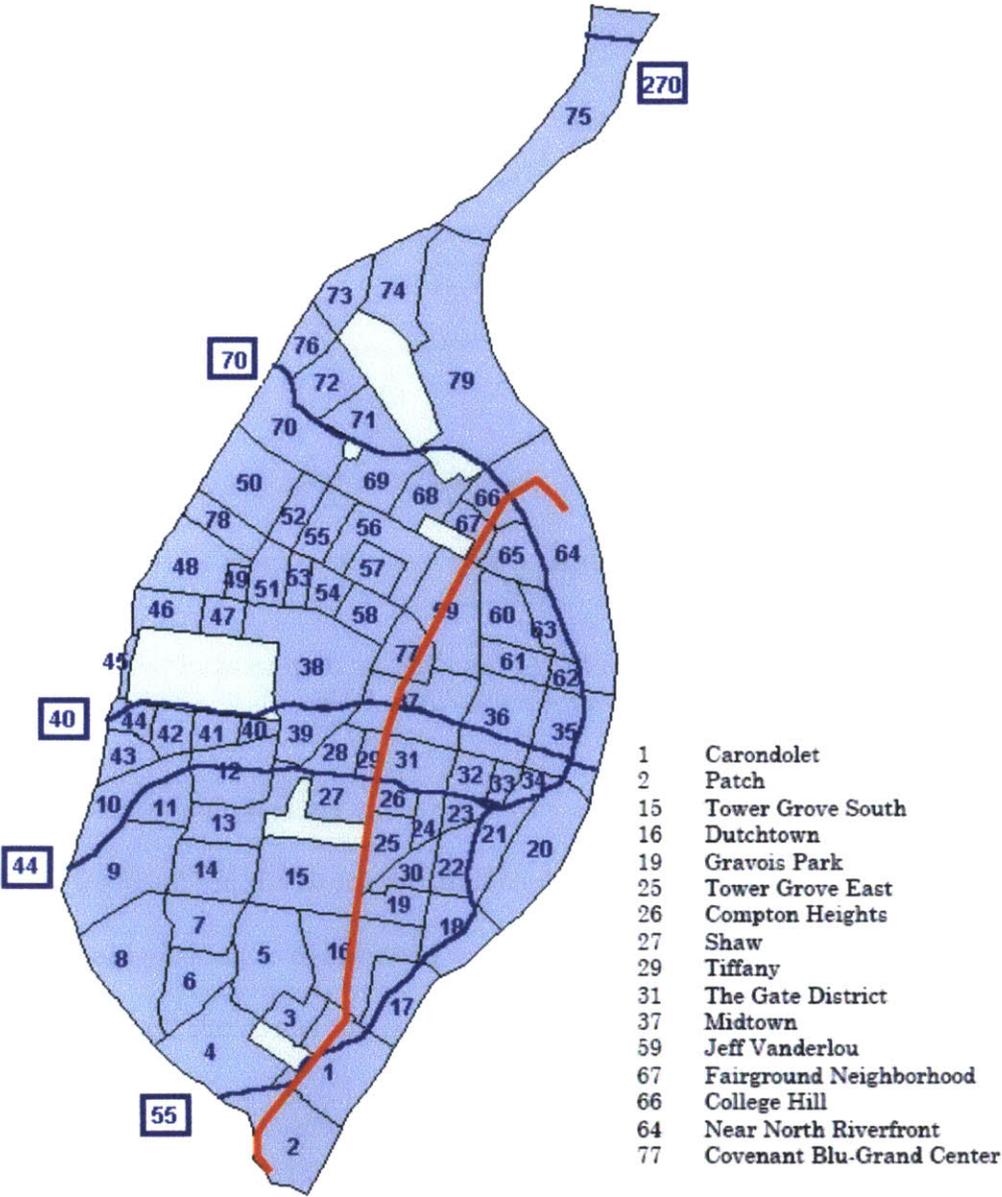


North, center, and south portions of St. Louis City, as demarcated by City of St. Louis Department of Planning. Base map via AAA, 1996.

Regarding households that own no cars (“Zero Car Households”), most are located in the south. This is due to the fact that the south simply boasts a larger population than the north, though. In terms of percentage of households that are “Zero Car,” greater proportions of neighborhoods in the North are “Zero Car” than in the south. The difference in levels of transit dependency by neighborhoods in the south compared to those in the north is observable in Table 3.5. Although 64% of the zero car households along the Grand 70 bus route are in these southern neighborhoods, they account for less than 20% of the total households on the south

side. In comparison, the percentage of zero car households in the more vacant north side is more than double that of the south side; but this represents only 18% of the total number of carless households in all the neighborhoods along the 70 line.

Figure 3.9: St. Louis Neighborhoods



Neighborhoods listed are those that abut the 70 Grand line, highlighted in red. Neighborhood map image via the City of St. Louis (Neighborhood Stabilization Team, 2014).

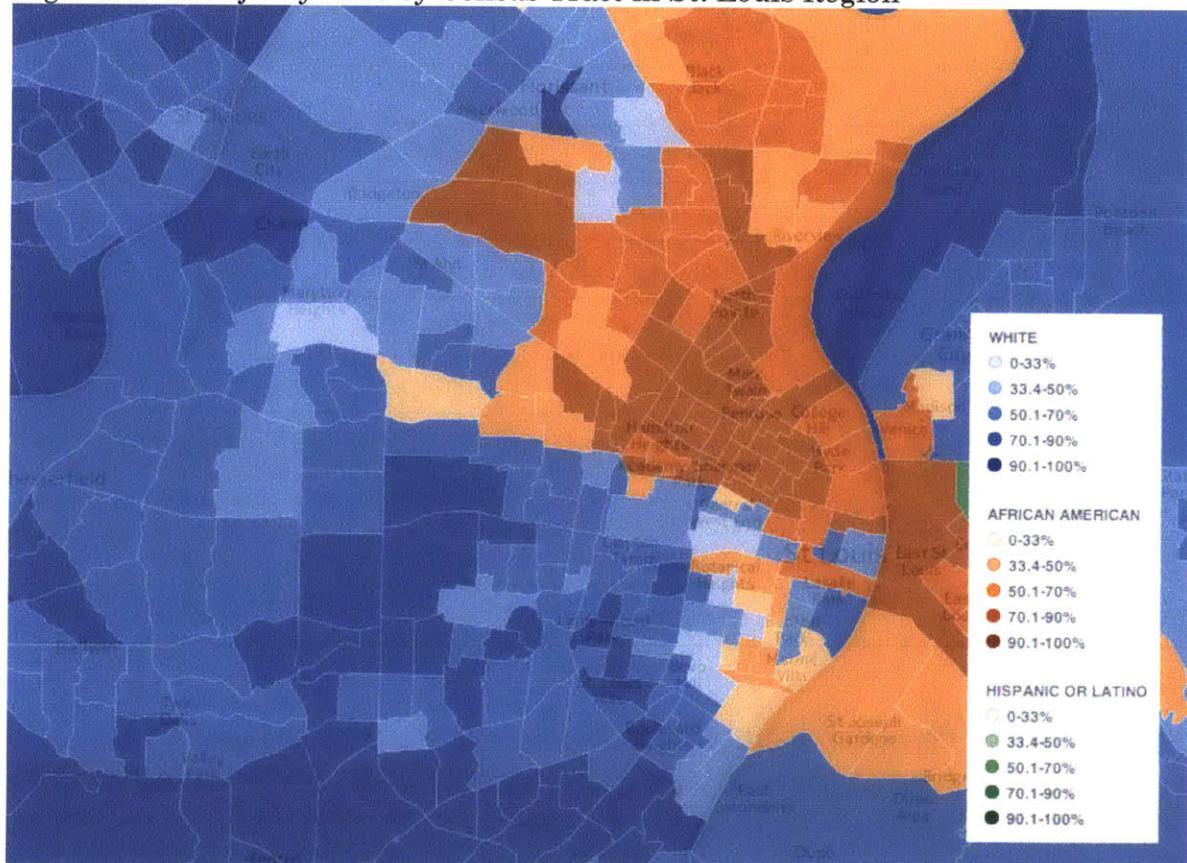
Table 3.5: Neighborhoods along 70 line, car ownership characteristics.

Neighborhood	Area	Households (HH)	Zero Car HH	%0 Car HH
Near North Riverfront	North	91	26	28.6%
College Hill	North	703	216	30.7%
Fairground	North	620	288	46.5%
Jeff Vanderlou	North	1,943	881	45.3%
Midtown	Center	1,590	672	42.3%
BC - Grand Center	Center	1,166	398	34.1%
The Gate District	Center	1,579	357	22.6%
Tiffany	South	425	98	23.1%
Compton Heights	South	567	35	6.2%
Shaw	South	2,988	390	13.1%
Patch	South	1,141	252	22.1%
Tower Grove South	South	6,055	1,095	18.1%
Tower Grove East	South	2,785	416	14.9%
Dutchtown	South	6,441	1,695	26.3%
Gravois Park	South	2,021	612	30.3%
Carondelet	South	3,698	496	13.4%
<b>Cumulative Results</b>	<b>Area</b>	<b>Households (HH)</b>	<b>Zero Car HH</b>	<b>%0 Car HH</b>
Cumulative Northside	North	3,357	1,411	42.032%
Cumulative Center	Center	4,335	1,427	32.918%
Cumulative Southside	South	26,121	5,089	19.482%

These transit dependency differences between the northern and southern neighborhoods also correlate with differences in the racial makeup of the neighborhoods. While the north is predominantly Black/African American, the south is comparatively mixed. The most heterogeneity is in the Center region, likely

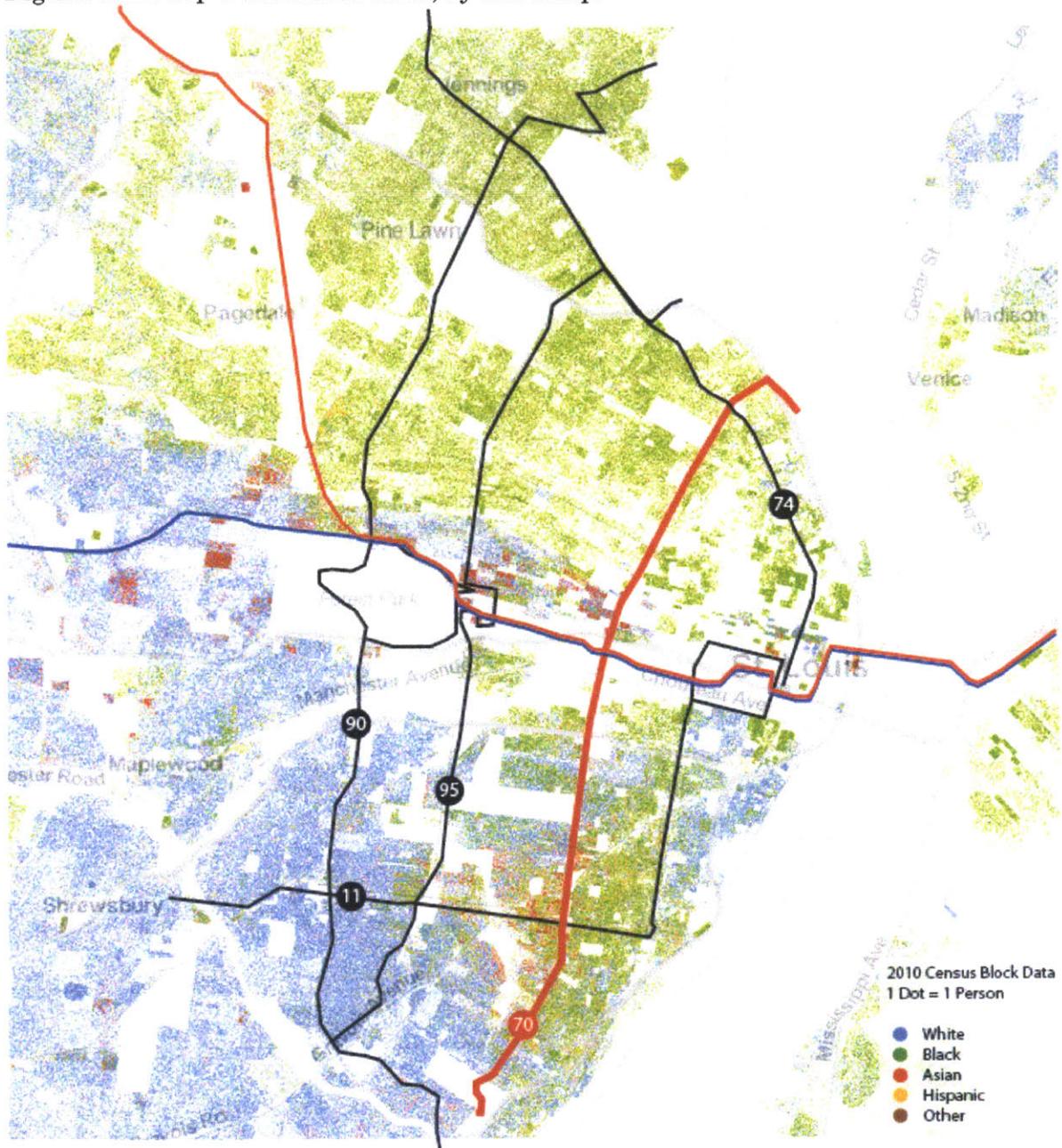
due to the presence of St. Louis University, as well as the Washington University Medical campus to the west in Central West End. The immediate areas surrounding these institutions host primarily Asian and white populations, with greater mix increasing with distance. The white population remains predominant in census tracts to the south and west, continuing its trend out into the suburbs of St. Louis County, where proportions of zero car households are also substantially lower. These spatial trends regarding race are visible in Figure 3.10.

Figure 3.10: Majority Race by Census Tract in St. Louis Region



Majority racial groups by census tract based off of 2012 American Community Survey Census data.

Figure 3.11: Top 5 MetroBus lines, by ridership.



Base map via the Racial Dot Map Project. The five major bus routes are highlighted in black with the 70 Grand highlighted in thick red. MetroLink routes, running east-west through the central corridor are also highlighted (Cable, 2013).

## CHAPTER 4: Methods and Technical Design

### Introduction to Flocktracker's Development History

As outlined in Chapter 2, Flocktracker was initially designed, developed and demonstrated through an early application in Dhaka, Bangladesh. After the initial pilot's promising results, as documented in Ching (2012) and Kennedy (2012), I joined the research project in October of 2012. The purpose of this chapter is to outline the developments of the technological approach from October 2012 to July 2014. For the purposes of creating a standardized method of referring to each iteration of the application that has been built over the past few years, the naming convention in Table 4.1 will be used to refer to each version of the application.

Table 4.1: Version Names and Details

Version	Programmed In	Details
1	App Inventor for Android	Prototype used under Albert Ching's thesis in Dhaka, Bangladesh in Spring 2012
2	App Inventor for Android	Modification developed in October – December 2012 for Dhaka's bus system mapping, January 2013
3	App Inventor for Android	Survey redevelopment and expansion; Mexico City, June – August 2013
4	Eclipse Java IDE	Complete rebuild, native coding in Java, prototyped in January 2014
5	Eclipse Java IDE	Modified version, from January test runs, used in St. Louis; February – May 2014
6	Android Studio / IntelliJ IDE	Complete refactor of Version 6 code; June – August 2014

Versions will be elaborated on later, referring to version names as identified here.

## **Mapping Dhaka's Bus Network**

While Ching (2012) and Kennedy (2012) demonstrated the viability of using the original Flocktracker software design to map bus routes and perform onboard counts; they also formed an organization “Urban Launchpad” to advance such technology-oriented urban innovations. In the wake of this work, Urban Launchpad, Kewkradong (a local social advocacy/mobilization organization), and MIT-based researchers set out to utilize the prototype software to develop a map of Dhaka's bus network. At the time of this effort, no complete, publicly available bus map of the system existed, and the limited data the city itself had on bus operator licenses was outdated and inaccurate. Operating in this data void, the first goal during October 2012 was to establish a new, more stable and user-friendly version of the application.

The piloted version of the application from January of 2012 (Version 1) had been built within the App Inventor for Android environment, which proved limiting given the field requirements. For example, for the bus route mapping effort in Dhaka in January 2013, (Version 2) the team wanted to utilize the phone's camera function. The purpose was to enable the user to take a photo of the bus before s/he boarded the vehicle so that reference to the line could include a visual component. Due to the inconsistent naming conventions of bus lines in Dhaka, this would assist in later identifying particular routes. Unfortunately, it was a feature shelved in the final moments because of the applications poor performance in terms of handling multiple, concurrent function calls. These inefficiencies were heightened due to the

technical limitations of the Samsung Galaxy Y. The Samsung Galaxy Y was the only Android-outfitted smartphone at the time available in Dhaka, Bangladesh and was essentially a repackage of older model Samsung offerings, tailored to the Southeast Asian market. As a result of the age of the internal technology used, particularly processing speed and memory, the device's performance was not stable when running the App Inventor-based programs.

Other issues present with Version 2 of the application included the inability to suspend the sleep function on the phone, and the inability to optimize the locational calls. The inability to suspend the phone's sleep function had two consequences. First, if the phone fell asleep, the application would either crash or simply miss a number of uploads, detrimental to data quality and route tracing accuracy. The second consequence was that the operator of the software would have to manually keep the phone "awake" by touching the screen constantly. This kept the screen turned on, but, because the screen is one of the greatest consumers of energy on a mobile device, it drastically reduced the battery life. Also, if the phone went to sleep, the calls to upload the location of the phone could not be performed.

App Inventor for Android's built-in location finder was not designed for repeated calls and the effort was both battery-intensive and otherwise inefficient. This means that the built-in function for geo-location on the application was slow to pinpoint the location of the device. When the function to pinpoint the location of the device was called again (this was designed to occur every 30 seconds), it had to completely repeat the process. A common and current solution to this problem is to

cache the previous location of the device in order to more quickly pinpoint the new location (something that the present version of the application, Version 6, does). Because the built-in function in App Inventor did not include this functionality, its locational process was significantly slower. If it was unable to find the location before the interval had ended and the call was made again, it would result in the application crashing, providing an incorrect location, or simply providing no location. The inability to cache the prior location and use that data in location updates, as well as optimize application performance by moving certain functions to background tasks that would not interfere with primary data uploads (those of greatest importance, such as survey responses), were significant limitations to the 1<sup>st</sup>, 2<sup>nd</sup>, and 3<sup>rd</sup> versions of the application.

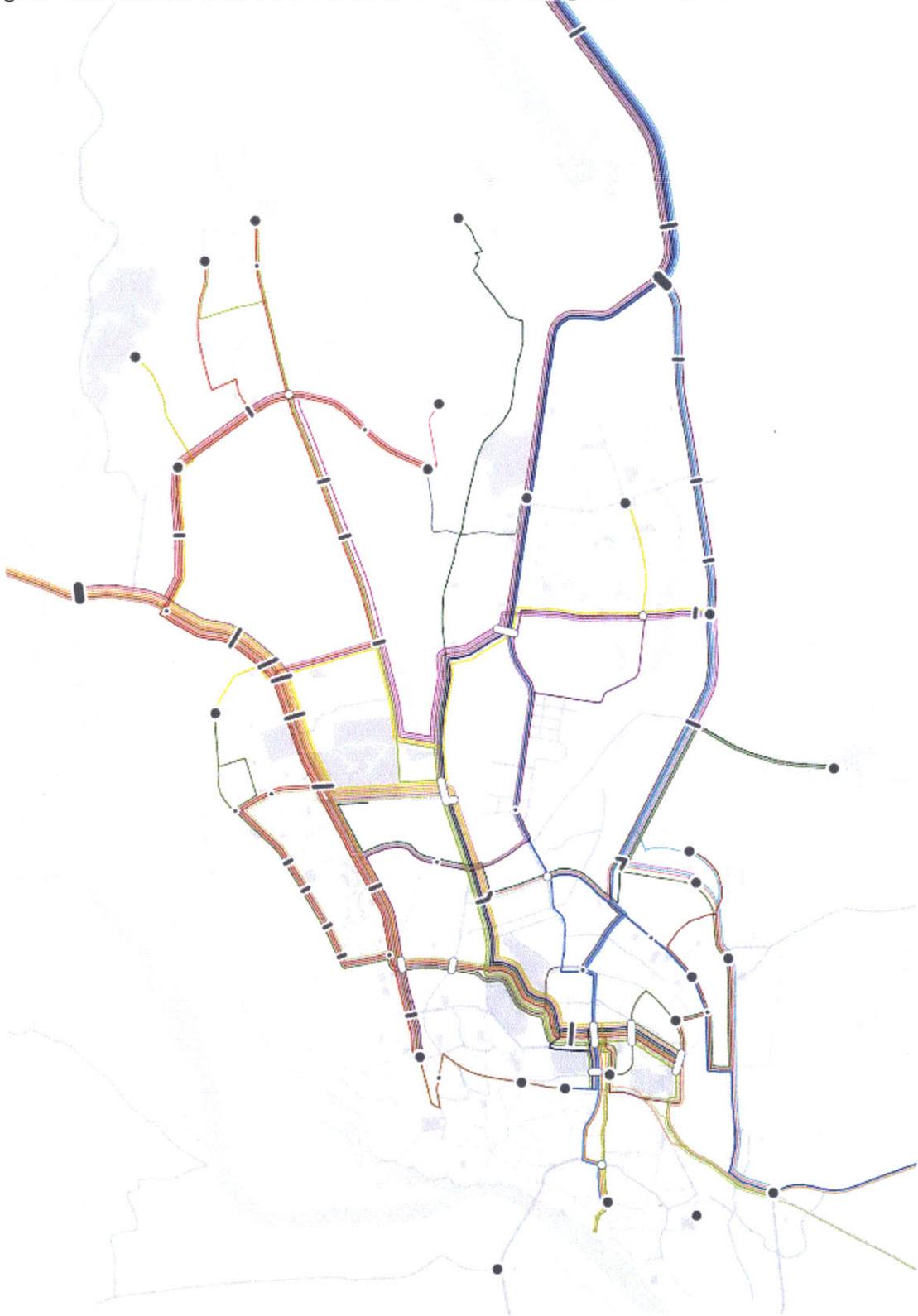
Figure 4.1: Version 2 Screens



The three primary screens from the rebuild of the application for the January 2013 bus mapping effort in Dhaka, Bangladesh. From left to right: counter screen (by gender), main hub, and survey screen.

Although Version 2 did still struggle with these technical limitations, a build was completed for use in January of 2013. The build was implemented as a tool in the process of mapping the bus system of Dhaka, Bangladesh. There, it successfully operated as a “spot checker,” meaning it was used in specific instances to double check or discover portions of routes that were unclear or unknown, rather than cover whole stretches of a bus line from start to end. Thus, by only employing the software for an hour or a fraction of an hour at a time (versus multiple hours) issues such as limited battery life and locational inaccuracies were mitigated. Most of the bus routes during the January effort in Dhaka could be determined and confirmed through multiple interviews, follow ups, and cross referencing between operators. Some routes had discrepancies in travel patterns, though, and Version 2 of the application (which had been coined “Bus Tracker Dhaka”) served as an effective and simple method of clarification. A volunteer would be deployed to the portion of the route in question and would ride it multiple times with the tracking portion of the application turned on. Targeting smaller portions of track and repeating the route a number of times both ensured better accuracy in determining the specific route and made cataloguing the specific path the bus took feasible. This method helped resolve inconsistencies in reports for the portions of the routes in question.

Figure 4.2: Plotted KML lines of all bus routes mapped in Dhaka.



The application's limited locational accuracy (in Version 2) was sufficient for brief periods of use, enough to develop the Dhaka Bus Map. While the application was developed to upload locational data at thirty second intervals, the instability of the application would often result in a locational call not being completed before the interval of time had run its course and the function was called again. Such instances, which were frequent, could cause the application to crash or upload faulty data. Using additional capabilities, such as the survey screen, would only serve to exacerbate this, as would length of use. Thus, short-term use cases were reasonable in that, if the application crashed, the user could dedicate the time and attention to monitoring it and restarting it, as the period of mapping was only an hour or fraction thereof.

One consequence of the limited usability of the application was the inability to develop metrics, such as bus speed, with a high degree of accuracy. Because point location and consistency of upload interval were not reliable, calculations to determine the distance between two upload points and time between them could not be utilized.

Figure 4.3: Dhaka Bus Map (March 2013 Release)



Previous page: Bus map shown is from April 2013 release. Updates and changes to the map have been made since. The most recent iteration is available for download online at [urbanlaunchpad.org](http://urbanlaunchpad.org).

## **Perception Mapping in Mexico City**

Despite its limitations, the Version 2 of the application was still successfully used, in combination with a large field effort, to develop and disseminate a map of Dhaka's existing bus system (Zegras et al, 2014 (forthcoming)). With the objective of further field testing the app, examining its functionality in a different urban context and towards a different analysis, another mega-city of the global south was turned to: Mexico City. The project was a collaboration with Urban Travel Logistics (UTL), a company under Grupo Prodi in Mexico City that specializes in developing and assessing large scale transit facilities, especially intermodal transfer stations. These transfer stations are known as CETRAMs or Centers for Modal Transfers, in Spanish *Centros de Transferencia Modal* (Gobierno del Distrito Federal, 2014). The research team (myself working with students from MIT and UNAM, the Universidad Nacional Autónoma de México), examined various dimensions of CETRAM performance, focusing on two UTL projects in the north of the Federal District (see Zegras et al, 2014 (forthcoming) for additional details).

Figure 4.4: In-Field Operations Monitoring and Application Synchronizing



In contrast to Dhaka, which focused on developing the information necessary to generate a bus map of the city, the Mexico case aimed at exploring the viability of the tool as a survey platform (Version 3). The interest derived, in part, from prior projects UTL had where the validity of survey data was uncertain (partially due to the difficulty in monitoring and auditing survey implementation). Towards these ends, we used the Flocktracker application as a platform for surveying bus system users. Specifically, Version 3 was modified to include a complete survey interface, customized for the purposes of analyzing semi-formal transit services that provide services at two CETRAMs. In developing the survey component, a dashboard was also designed. The dashboard was a web page that allowed viewers to monitor data

results in real time, as well as to visualize the location and movement of volunteers surveying in the field on an interactive map. While developed and used in a limited capacity in Dhaka, the dashboard became a critical component of the Mexico City research, allowing the research team to monitor, in real-time, and manage survey implementation on routes emanating from two CETRAMs in the northwest and northeast of the metropolitan area, simultaneously.

Figure 4.5: Mexico City Dashboard Screen Capture

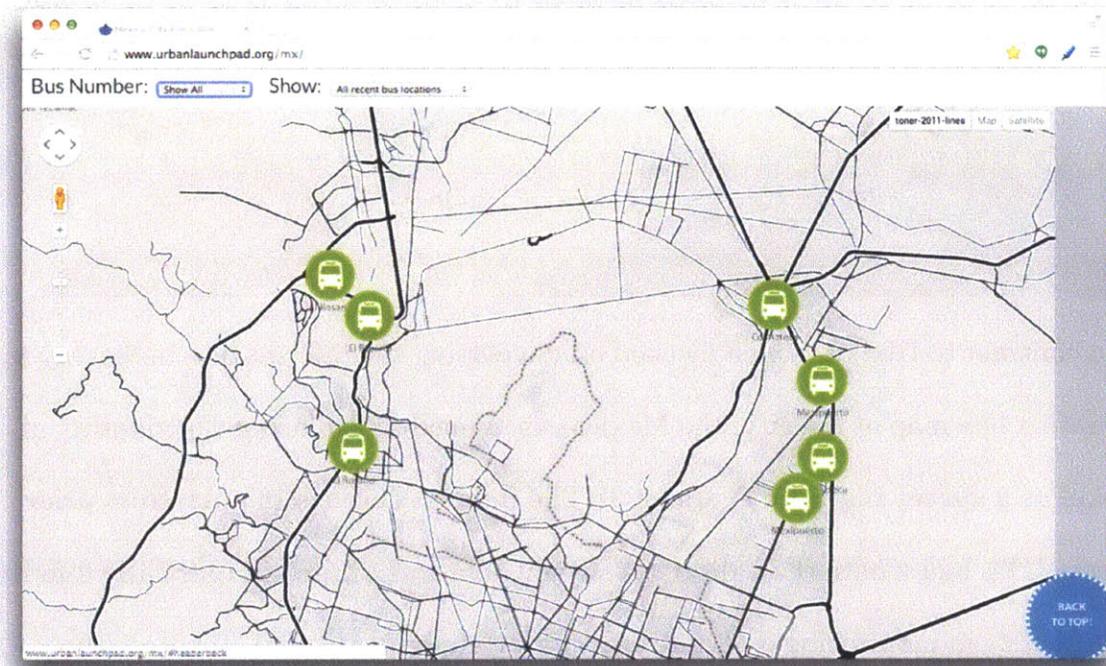
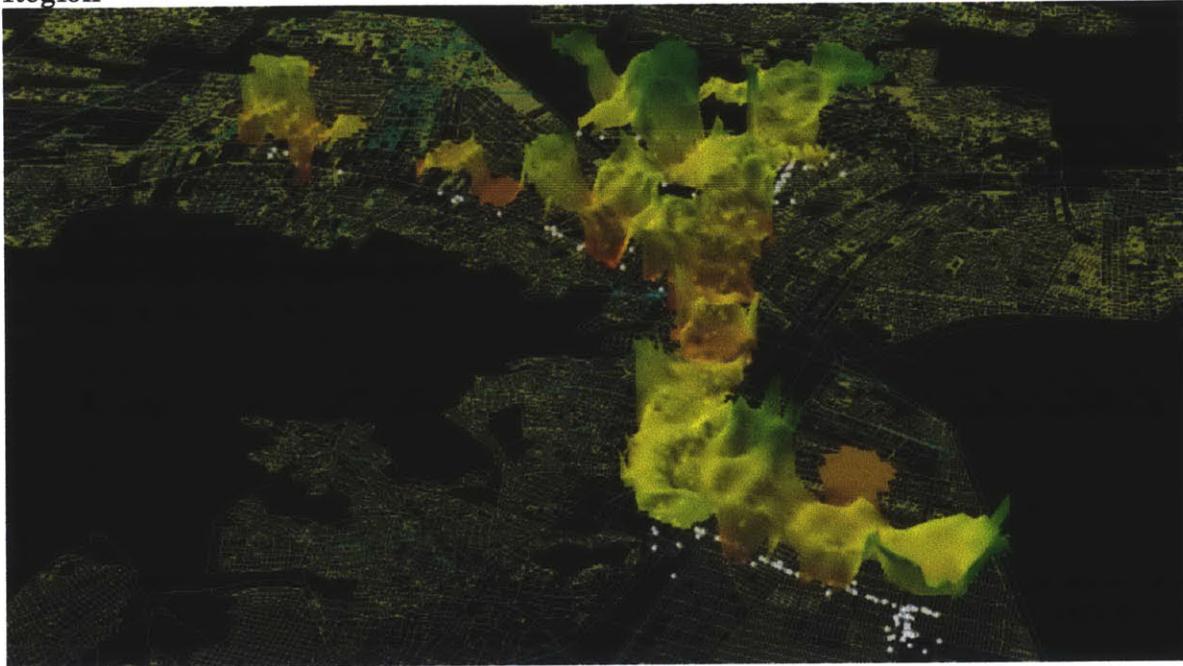


Figure 4.6: Topography of Personal Security Perception in Northeast of Mexico City Region

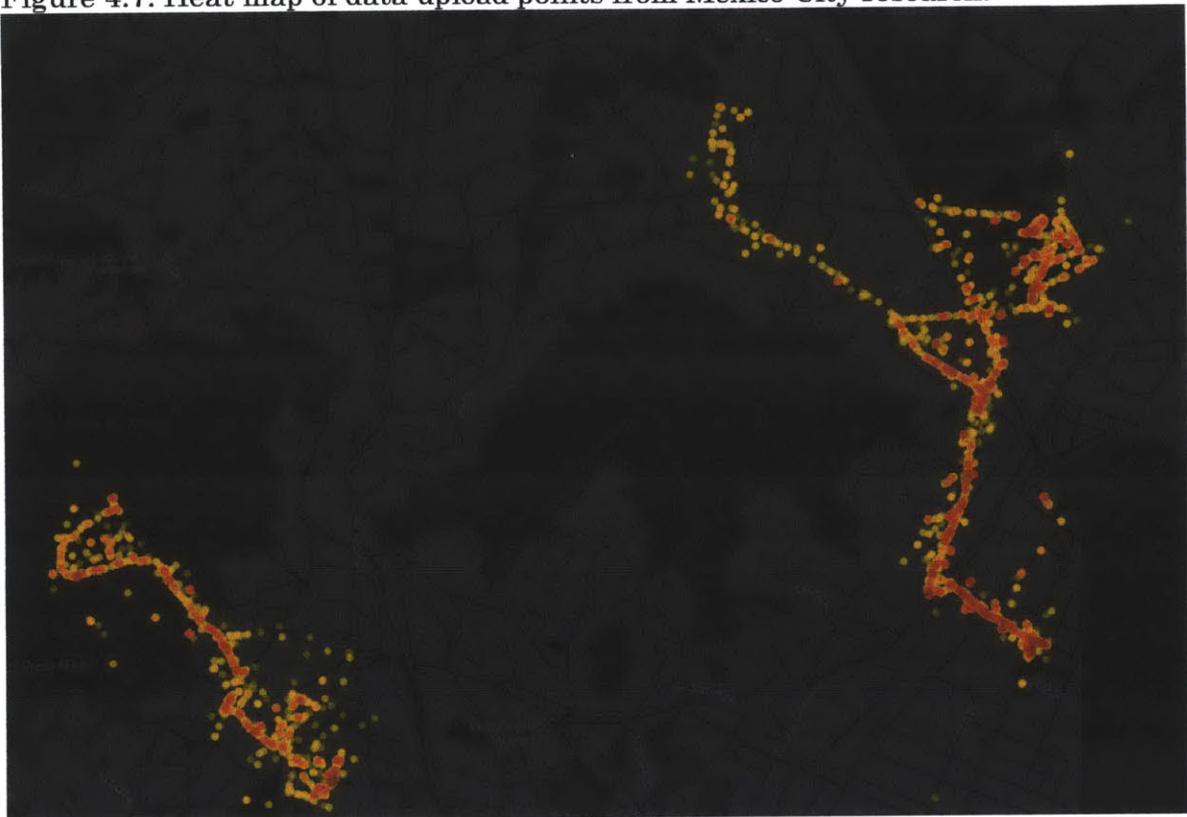


Topographic representation of security perception at specific locations throughout the northeastern region of Mexico City, along “peseros” routes developed by Liqun Chen and Kuan Butts (Zegras et al, 2014 (forthcoming)).

In addition to demonstrating the capacity of the dashboard to ease the process of running surveys in two disparate locations simultaneously, the Mexico City pilot (Version 3) served to showcase the potential of the Flocktracker application to develop data sets that more effectively matched each data point with place and time stamps. Should additional geospatial data exist, as is increasingly the case in a large number of urban areas, globally, those additional spatial data can be linked, locationally, to the survey data for additional analysis. This characteristic facilitates joining the data collected with, for example, context-specific variables from other sources such as socio-demographic data. That said, Version 3’s locational accuracy

was still limited, and interval consistency for uploads had not yet been resolved. The result was an improved but imperfect product.

Figure 4.7: Heat map of data upload points from Mexico City research.



The limitations of Version 3's use in Mexico City were exacerbated by the scale of the operations. Because dozens of users were operating the survey simultaneously, maintenance was constantly necessary to correct and update upload errors from live users. This process required constant attention by the surveyors (application operators) to make sure the application was functioning

correctly during survey implementation. In part to prepare the surveyors for these challenges, a day-long training session was carried out.

### **Native Redevelopment of Flocktracker**

In light of the limitations experienced with Version 3 of the Flocktracker application, a complete native development of the application was initiated in September 2013 (Version 4). A team of three helmed this native coding: Kuan Butts (MIT), and Danny Chiao (MIT), Daniel Palencia (UNAM). Another, Arturo Cadena (UNAM) provided additional support. The redevelopment focused on three main goals. First, the application needed to be stabilized. Specifically, background tasks, such as completed survey uploads, needed to be programmed so as to not interrupt survey and ridership count activities. Additionally, geospatial location tracking - the act of “bread-crumbling” data points over a route at specified intervals - needed to be secured and optimized in order for locational data to be accessed efficiently, so as to conserve battery power and improve accuracy. Secondly, the application’s workflow needed to condense the steps needed to process surveys, thus reducing the complexity of moving around survey parts during implementation. Part of this goal implied developing a method by which a researcher could jump between questions in the survey dynamically, without impacting completed responses. The final goal of the application was to enable the survey component to be self-configuring. That is, the application needed to be able to take in a file (JSON, a lightweight data-

interchange format, was selected as the standard), read it, and reconfigure itself to the selected inputs. In such a way, the application would have the potential to be reprogrammable by nontechnical individuals, able to perform an infinite number of different surveys.

Figure 4.8A: Spatial accuracy in Mexico City with Version 3, Summer 2013.

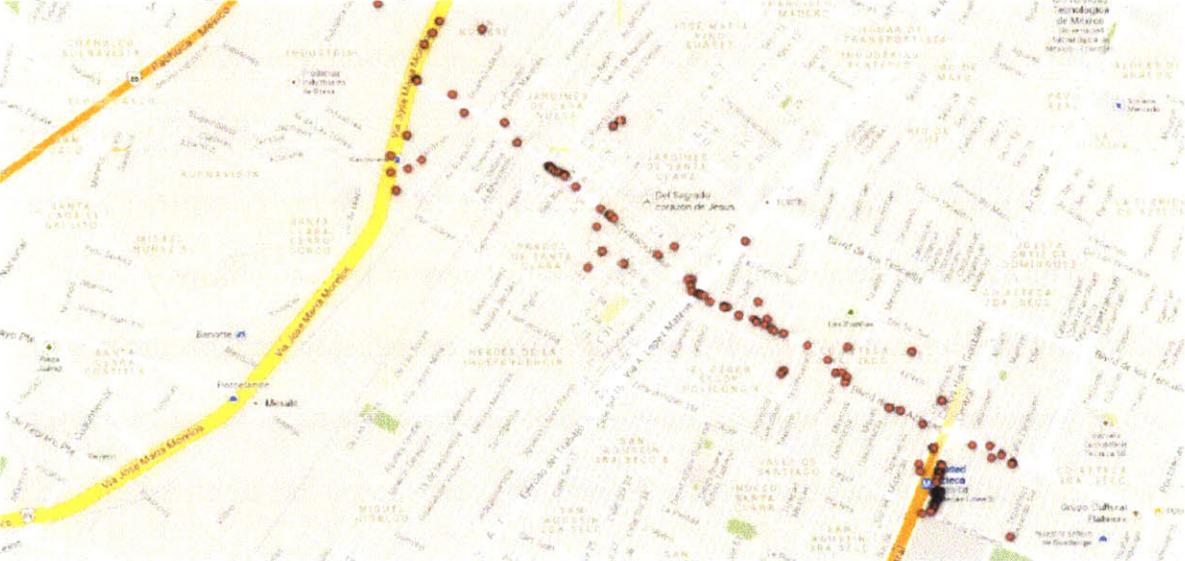


Figure 4.8B: Spatial accuracy in Mexico City with Version 4, Summer 2014.



Version 4 was tested in Mexico City in January of 2014. The results from this run demonstrated the significantly improved accuracy of spatial locations when performing location calls (Butts, 2014, 9). Figure 4.8 compares results from Version 3's use in the summer of 2013 with that of Version 4 in January of 2014. Both instances involved making locational calls every 30 seconds. The improved quality of data is clear in 4.8B, which used the new, App Inventor-free, natively coded application (Version 4). Through this complete rebuild, accuracy was improved from around 200 meters in Version 3 to 5 meters or less in Version 4 (Butts, 2014, 8). Modifications to Version 4 were made to correct for issues found while operating the application in January of 2014. The resulting application, Version 5, was used in St. Louis, Missouri, for this thesis' research.

## **CHAPTER 5: St. Louis Survey Design and Principles**

### **Design of St. Louis Operation**

With the significantly improved capabilities of Version 5 of the Flocktracker application, I designed an experiment to demonstrate some of the application's key strengths:

1. The ease of constructing a survey and implementing it,
2. The speed with which the survey could be installed and prepared for a surveyor on a tablet or smartphone,
3. And the potential for the results to be spatially linked, with high precision to a number of additional environmental and route-related factors that would not have been easily gathered with a traditional, pen and paper survey.

The prior pilots, especially Version 4 in Mexico City, had led to a robust and reliable application. This allowed me to focus the St. Louis experiment on gathering a team, designing a short survey that leveraged a number of the application's capabilities, and examining a topic of real potential concern to a public transit agency in a city such as St. Louis, Missouri. This enabled me to fully deploy the application as a data-collection tool – to be used in-field to perform surveys and live analyses of bus route's dynamics – rather than a technological prototype.

Unfortunately, due to unforeseen complexities in securing permissions from the transit agency (BSDA) to perform the onboard survey during the last week of March, I had to put the project on hold. As a result, I lost the team of volunteers

who had initially agreed to assist in implementing the onboard survey. I was ultimately able to secure permission to perform the onboard survey, halfway through the week during which survey implementation had originally been scheduled. Without a survey team, I was still able to install the survey on my own tablet and begin surveying within 30 minutes of receiving permission to do so from BSDA. The results of this intensive three-day effort will be discussed in detail in the following Chapter. The demonstration of the application as a method for quickly deploying a survey and generating live results was successful, showing how a single researcher could generate a sizeable amount of high resolution data in a relatively short time.

### **Onboard Security Discussion**

Similar to the Mexico City research, I chose to focus the experiment on “onboard security perception.” This security concept is intended to represent crime and fear of crime, not traffic safety. Bannister and Fyfe (2001) elaborate on this concept extensively, documenting the emergence of fear as a component of lifestyle that, ultimately, begins to intervene and establish itself in urban space. They build from Hale’s 1996 “Fear of Crime,” which suggests that observations and measurements of fear of crime are not merely measurements of this fear. Rather, the fear itself is a proxy element, indicative of “some other attribute, which might be better characterized as ‘insecurity with modern living’, ‘quality of life’, ‘perception of disorder,’ or ‘urban unease (Hale, 1996, p. 84).” Nilay and Welch, in their 2010

paper, attempt to elaborate further on a definition of “fear of crime,” acknowledging that there is no, universal, agreed-upon definition, by citing a 1995 paper, “Fear of Crim: Interpreting Victimization Risk.” In this paper, the author portrays fear of crime as “an emotional response of dread or anxiety to crime or symbols that a person associates with crime (Ferraro, 1995, p. 4).”

The complex range of emotional, social, and personal influences that are thus embodied in this concept of personal security are inherently unique to each individual. That is, elements of one’s own personal experience serve to influence one’s personal perception. Prior instances in which one observed crime or heard others describe being victim to a crime affects his/her perceptions of crime in transit (Clark, 2003). That said, fear of crime is often a primary component in an individual’s decision to utilize public transportation or not. For example, studies have shown that fear of crime is one of the most significant reasons for individuals to not choose to use public transportation (Lynch et al, 1988). While research often centers on females’ fear of crime and the limits this places on their ability to engage in the public environment, security concerns have also been demonstrated to be prevalent amongst men as well, restricting their activity in public spaces (Day et al, 2003).

In addition to restricting the freedom of potential riders to feel safe in public, reduced perceptions of personal security serve to undermine the public’s image of the government as a public protector (Hanmer and Stanko, 1985). Thus, in analyzing onboard transit security perceptions, research can assess primary

concerns for both male and female transit customers. As was suggested at the conclusion of the paper “Addressing Fear of Crime in Public Space,” by Yavuz and Welch, “in order to identify safety strategies that could effectively address fear of crime, it is first necessary to pay more attention to the factors that make individuals feel more vulnerable to crime in public space (Yavuz and Welch, 2010, p. 18).”

### **Beyond Security**

In light of these findings, adding a high resolution spatial and temporal overlay on top of survey data geared at understanding perception of security can become a powerful way of understanding the dynamics of this variable along the length of the route being studied. Such variables that explore how far north the respondent is when answering, their proximity to recent homicides or other crimes, vacant and abandoned buildings, and the speed of the vehicle can all be incorporated to enhance understanding of personal security perception.

Clearly, numerous factors play a role in influencing perception of security for a rider. These influences might be thought of within 3 categories:

1. **Characteristics of the rider:** Examples include race, age, gender, trip purpose, and origin and destination
2. **Characteristics of the trip:** Examples include speed, northbound or southbound, location, time of day, and rider count by gender

3. **Characteristics of the surroundings: Examples include crimes, abandonment rates, vacancy rates, distance north, and area auto ownership levels**

With the new version of Flocktracker and its improved location determination capabilities, the application can now enable all the security perceptions gathered during onboard surveys to be closely linked to other variables and data sources. This has provided an unprecedentedly rich set of potential explanatory variables.

## CHAPTER 6: Survey and Data Analysis

### Survey Introduction

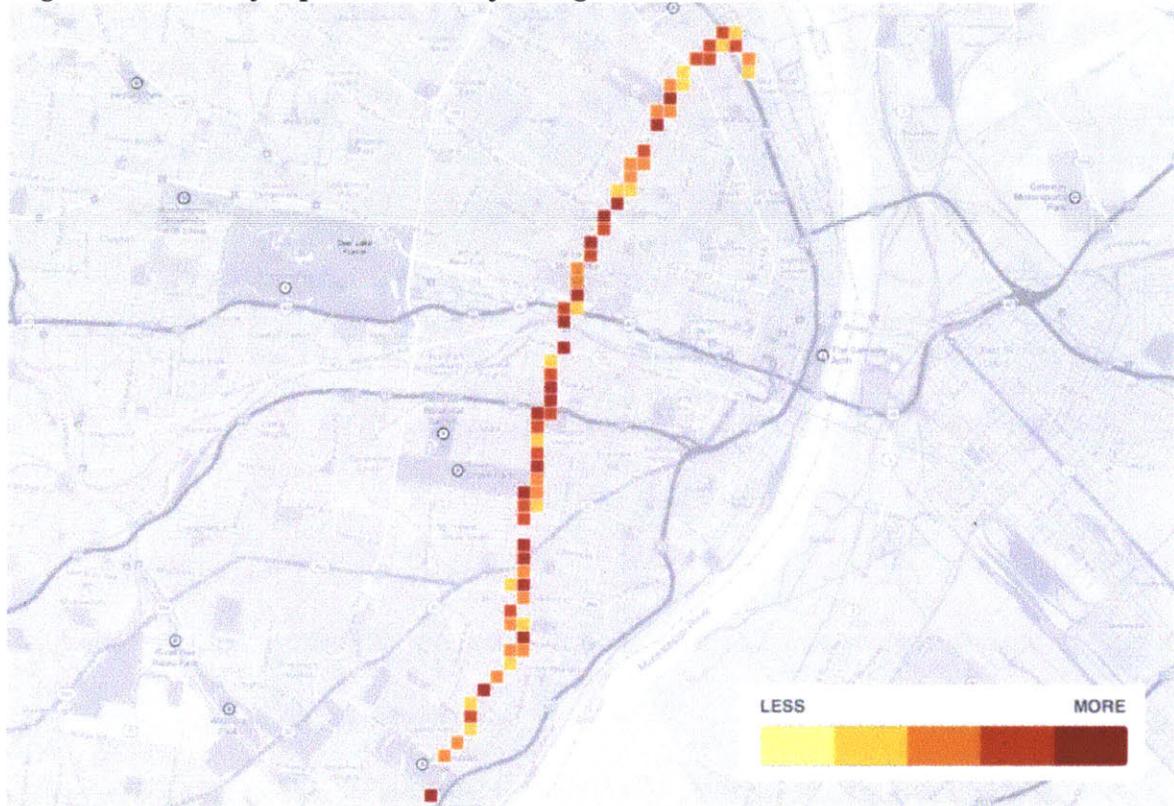
I designed and implemented the onboard survey used in this thesis. I was trained in survey implementation and ethics via COUHES,<sup>1</sup> and had substantial prior experiences with using various versions of Flocktracker in the field (as I have led the development of the application since November 2012).

For the St. Louis implementation, I chose the Grand Avenue 70 bus line, due to its significance in terms of both ridership, and its role as the sole high frequency north-to-south connector within the core of the city. To carry out onboard surveys on this line, I rode its length from its southern termination point south of Carondelet Park at Loughborough Commons to its northern terminus just above Interstate 70 on the industrial zone of the north shore of St. Louis, near the Mississippi River. Each morning, I boarded the 70 Grand at the intersection of Shenandoah and Grand Boulevard. This location sits roughly in the middle of the route. Additionally, I would alternate the direction I initially headed in so as to balance morning observations for the earliest portion of the day.

---

<sup>1</sup> COUHES. (2014). "Training: Research Involving Human Subjects."

Figure 6.1: Survey Uploads Density along Line 70 Route.

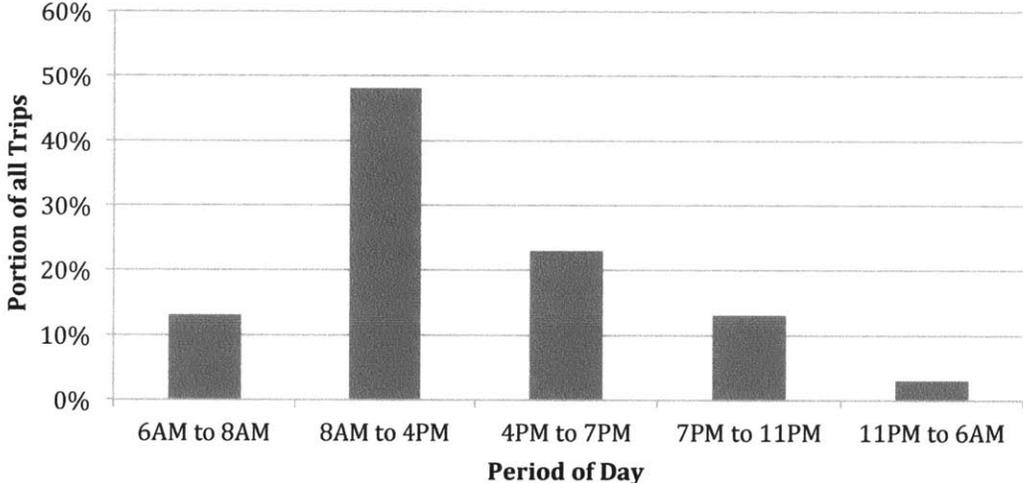


## Survey Method

Appendix 1 contains the full survey questions. Surveys were performed in St. Louis on March 26, 27, and 28. Survey operations began at 2:00 PM on March 26, and 8:00 AM on March 27 and 28. Survey operations completed by 8:00 PM each night, at the latest. Riding was not nonstop – I took midday breaks and, on Friday, an extended detour to head downtown to present work on the project to BSDA (Metro). Thirty hours, roughly, were dedicated to performing onboard surveys over this three-day period. Within these 30 hours, an estimated 22+ hours were spent

inside a bus while it was operating. I targeted the 8 AM to 7 PM time window, during which an estimated 71.1% of all trips in St. Louis are performed<sup>2</sup>. While the intent was to survey until 7:00 PM, trips could not be timed so precisely. Data and surveys were conducted until complete round trips were finished, which often extended beyond 7:00 PM. Thus, data are reported for as late as 8:00 PM.

Figure 6.2: Times of Travel for All Trips in St. Louis Region



Data gathered from 2003 St. Louis Region Household Travel Survey Report (East-West Gateway Coordinating Council, 2013, p. 12). Data of 46,909 unlinked trips representing 9,457,294 total trips in the region.

Respondents were selected by methodically moving up and down the bus, approaching individuals who had not yet participated or declined to participate. I did not record rejection rates and therefore cannot provide an estimated rate of rejection. I did observe that riders were noticeably less receptive to taking a survey in the early morning or when the bus was extremely full (standing room only).

<sup>2</sup> NuStats' 2003 St. Louis Region Household Travel Survey Report

Additionally, when the bus was very full, it was harder for me to maneuver around the vehicle, thus limiting my ability to approach new potential participants. Thus, these time periods, when the bus was very full and during the early morning, are likely underrepresented in the survey responses collected.

Morning trips alternated between northbound and southbound trips. On Wednesday, when I began surveying at 2:00 PM, I took my first bus headed northbound from the intersection of Shenandoah and Grand Boulevard. On Thursday at 8:00 AM, I started with a northbound trip from the same intersection. Finally, on Friday, I began my morning from the same intersection, but caught a southbound morning. In this way, I sought to balance my morning observation counts with both northbound and southbound initial directions.

### **Survey Results: Introduction**

In total, the survey included 248 riders on the 70 Grand bus line. Of those 248, 236 observations were kept after cleaning the dataset. The eight surveys removed were due to incomplete form entry. In addition, 1,754 data points were collected. Each data point represents a “bread crumb,” uploaded at 30-second intervals, containing current trip information, including:

1. Trip ID
2. User ID
3. Geospatial coordinates

4. Exact time
5. Route direction (northbound or southbound)
6. Rider count (separated by gender)

These results were enabled by a substantially more stabilized application, with significant improvements over the previous research endeavors using prior Flocktracker builds. For example, in our Mexico City research project on CETRAMs, 650 volunteer hours produced 10,977 data points while conducting 1,528 surveys. A “buddy system” was employed in Mexico which meant that, jointly, each duo conducted 4.7 surveys and produced 33.76 data points per hour, on average (Zegras et al, 2014 (forthcoming), p. 17). Furthermore, in order to combat the unstable nature of the application (Version 4) used in Mexico City, both students in each “buddy pair” were operating the application on their phones at the same time. Thus, each phone was only successfully uploading, on average, 16.88 “bread crumbs” per hour, when it should have been achieved roughly 80-100 (accounting for downtime, failure to turn on the tracker, or other user errors).

The St. Louis operation achieved greater efficiency than the Mexico City one, averaging 11.26 surveys per hour (versus 4.7 in Mexico City per team of two) and 79.73 data points per hour (versus 33.76 in Mexico City per team of two). Acknowledging that I performed user errors, such as failing to turn on the tracker function while using the application on some runs, I did nearly achieve the desired goal of 80-100 upload points per hour. The app was successfully optimized to upload location at 30 second intervals, and thus the quality of the data was sufficient to

determine speed by calculating the distance from one point to the next to determine the rate of travel over that 30 second period.

While a multitude of factors make the results, in terms of survey completion rates, not comparable between St. Louis and Mexico, upload rates are comparable. Upload rates are comparable because, in both operations, the application was designed to upload at 30 second intervals. The consistency of data point uploads, intended to be two per minute, was achieved in St. Louis with the new build of the application, thanks to significant improvements in the application architecture and its ability to operate background tasks smoothly and efficiently.

Figure 6.3: Age Distribution of Survey Respondents

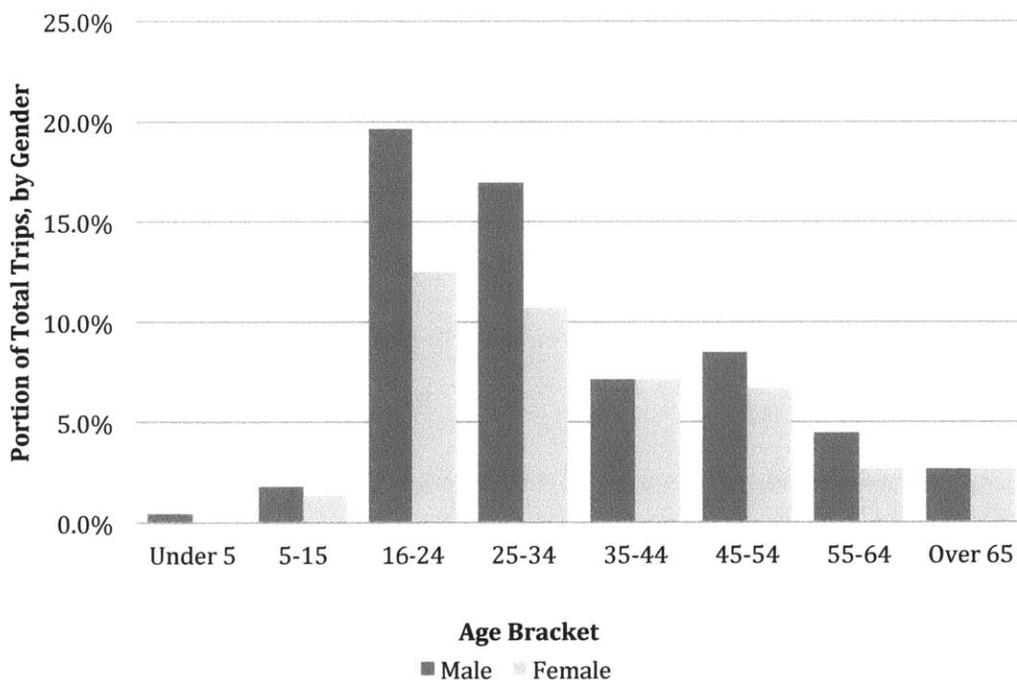


Figure 6.4: Results Comparison against 2002 HHTS

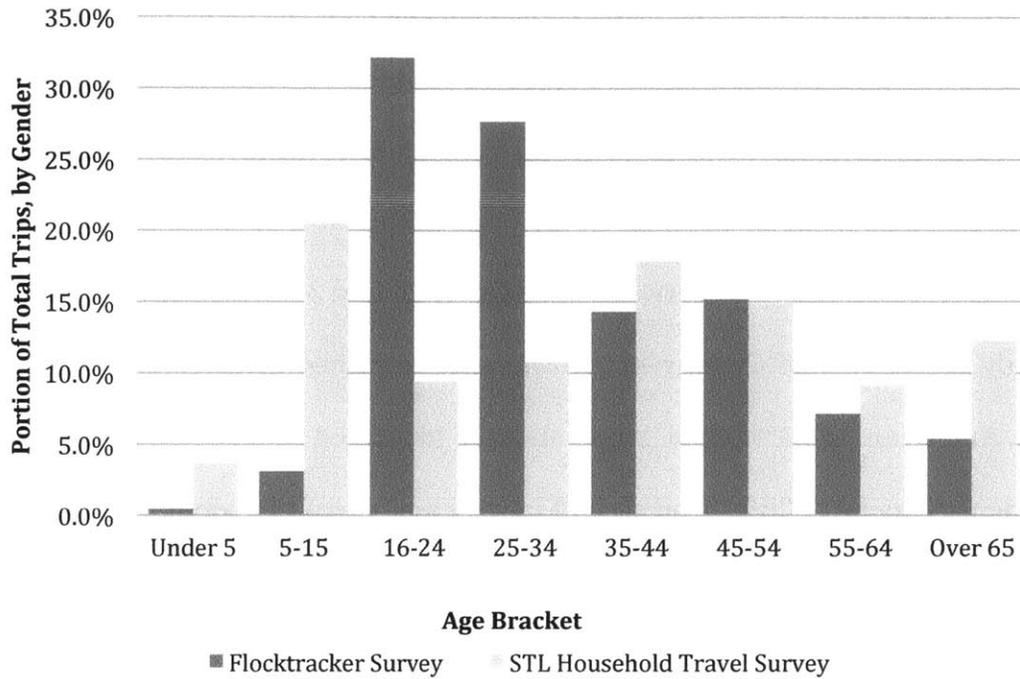
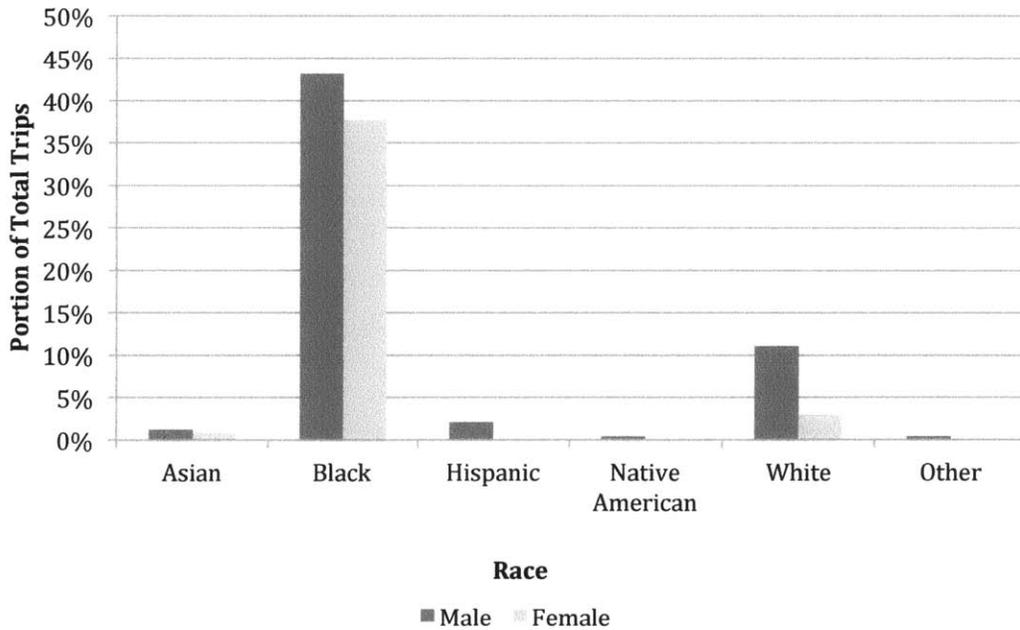


Figure 6.5: Survey Respondent Ethnicities



The average age of a survey respondent was 34 years old. Most (58.5%) respondents were male. Thus, male respondents were over-represented, assuming that the gender balance of riders on the 70 Grand was similar to that of the general population in St. Louis, where males made up roughly 48% of the total population. Half of all respondents were between the ages of 19 and 34. Respondents were primarily black; over 80% of survey participants described themselves as “Black / African American.” 53.4% of black respondents identified themselves as male. Of those identifying as “White” (the second largest portion of the population, at 13.98%), 78.7% identified as male.

Figure 6.6: Image of respondent filling out questions on tablet, by Kuan Butts.



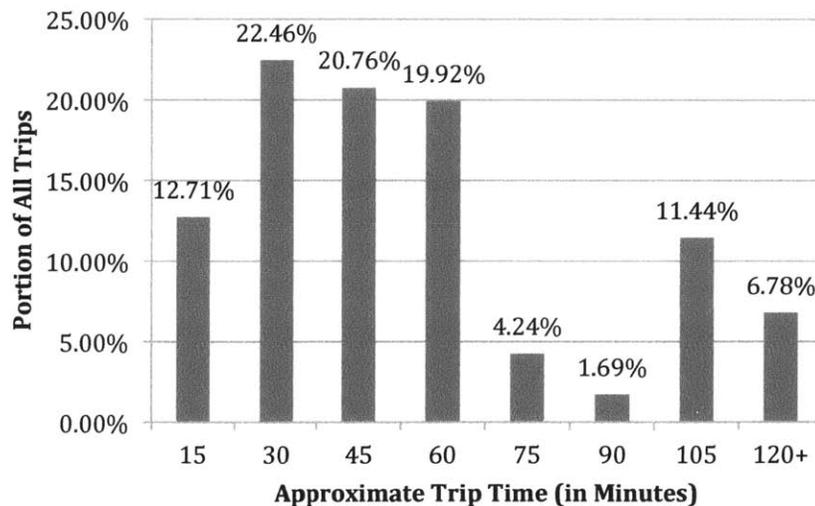
Comparing the results to the 2002 St. Louis Household Travel Survey (HHTS) reveals various possible biases. That said, the demographic data from the HHTS is for the metropolitan region as a whole, versus specific to riders using the MetroBus on route 70, or even those in the City itself. The HHTS also represents a glimpse of travelers in St. Louis more than 10 years before my survey. Thus, it is unclear how representative the HHTS demographics are of individuals specifically riding the 70 line during my experiment.

In light of these comparability challenges, my survey reveals an apparent bias towards younger individuals, particularly in the age bracket of 16-24 (East-West Gateway Coordinating Council, 2003). This could be caused by the survey's use of tablets and potential generational bias towards technology in younger audiences versus older ones. That is, older riders may have been more prone to decide not to participate in the survey, because it was performed on a tablet. In a 2014 paper co-authored by Todd Hennessy, the manager of market research for BSDA (Metro), 50% or more of every age group reported owning a smartphone and 85% or more of all respondents were found to have reported owning any type of cellular phone (Windmiller et al, 2014, p. 9). The 19 to 35 age group reported the highest levels (all over 95%) of cellular phone ownership (and over 78% owning smartphones), with those between the ages of 25 and 30 reaching nearly 100% (80% smartphone ownership). Ownership for those ages 51 to 64 was lowest, at 85%, with only 50% owning a smartphone (Windmiller et al, 2014, p. 9). While these numbers indicate high levels of mobile device adoption across age groups, younger groups' higher

smartphone ownership may reveal a higher propensity for participating in the tablet-based onboard survey.

The number of respondents that were young children was notably below the HHTS numbers. While, again, the representative nature of the HHTS figures to the 70 line is unknown, the limited number of child respondents may be due to the specific time period of the survey: portions of the day when children were likely to be in school or after-school programming. The lower share of female respondents in my onboard survey may also reflect lower transit usage by women, as some research has suggested that low income women, especially mothers, have a disproportionate need to use a car (Schulz et al, 1996, p. 553) and are either isolated or limited in their ability to utilize public transit. In the end, the representativeness of my sample, vis-à-vis the actual users of the 70 or the broader metropolitan area, and the reasons for any biases, warrants further investigation.

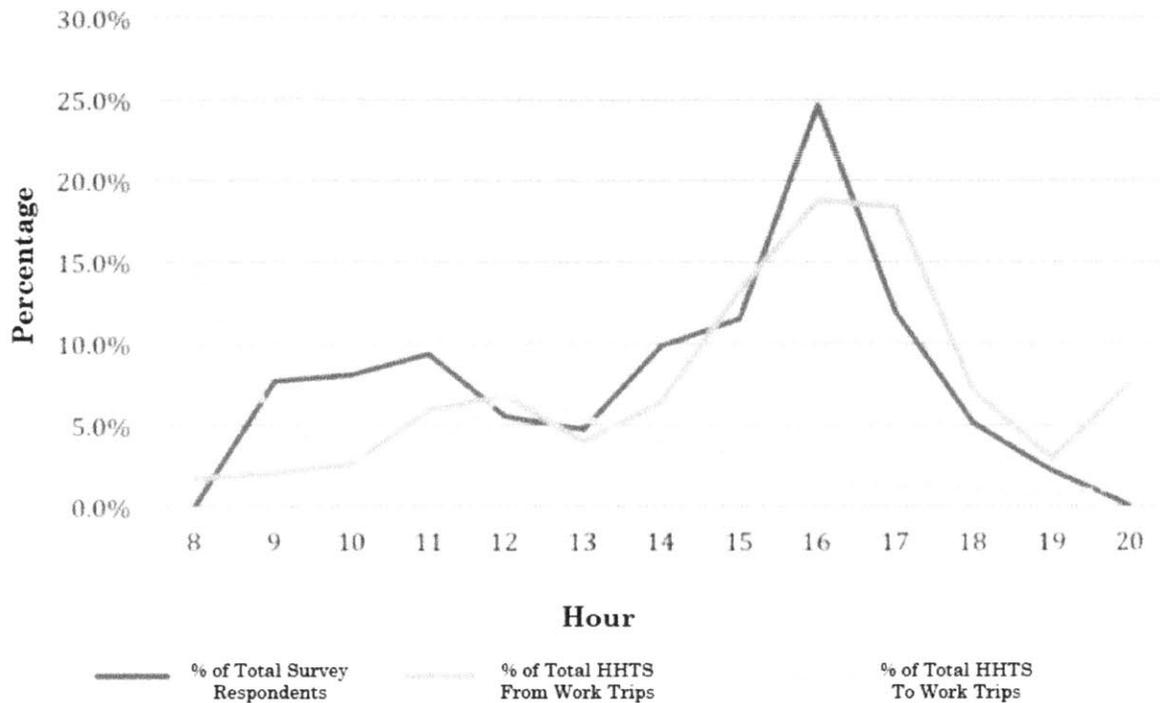
Figure 6.7: Travel Time Estimate Distribution



To estimate approximate travel times for the trips being undertaken by respondents, I used their stated origins and destinations (by neighborhood). For each neighborhood I established a centroid latitude and longitudes and used the subsequent matrix of origins and destinations to feed into Google's Distance Matrix API. Results were rounded up to the nearest 15-minute interval to determine approximate trip times. Problems with this indicator include the inconsistent sizes of neighborhoods in St. Louis and the potential for riders to have misreported and/or interpreted differently what neighborhood they were coming from or going to. Also, for rides into St. Louis County or East St. Louis, only one centroid was used for each. As a result, locations for trips to or from these areas are very crude.

From these estimates, nearly two-thirds of riders were travelling between 30 minutes and an hour from their origin to their destination. According to HHTS, in the City of St. Louis, trips taking 30 – 59 minutes represented 9% of all trips on all modes (East-West Gateway Coordinating Council, 2003, p. 19). Only 226 of 46,909 reported trips in the HHTS dataset were listed as having occurred on a public bus, with an average trip duration of 12.36 minutes. Of the 128 trips that listed their home county as St. Louis City, the average travel time was a mere 11.30 minutes.

Figure 6.8: Average Portion of Surveys Performed by Hour and To/From Work Trip Start Times Distribution



The number of onboard surveys performed, on average, per hour of the day, peaked at roughly 4:00 PM. This time corresponds with the peak period for return from work trips in the city, according to the HHTS. The less-than-receptive attitude towards surveying I encountered in the morning hours may have influenced the smaller portion of total survey results obtained in the morning versus the afternoon peak. Overall, results in survey counts from northbound and southbound trips were balanced; there were 120 surveys performed on northbound trips and 116 on southbound trips.

## Spatial Results: Trip Origins and Destinations

Surveyed riders' trips originated from locations in St. Louis City, St. Louis County, and East St. Louis. Using the aggregate areas of the corridors, presented in Chapter 3 ("North" defined as corridor north of Delmar Boulevard, "South" as corridor south of Interstate 44, and "Center" as linear portion of city from downtown to Forest Park). Nearly two thirds of trips originated from either the north or the south areas of St. Louis City.

Table 6.1: North and South Origin-Destination Trip Combinations

Trip Type	Trip Count	% of Total
North - North	18	7.63%
North - South	17	7.20%
South - North	28	11.86%
South - South	42	17.80%
Combined	105	44.49%

In discussion with BSDA (Metro) transit operators, the intention of the north-south bus routes, including the Grand 70 line, is to act as a feeder system into the MetroLink, which runs east-west through the central corridor of the city. From my results, the Grand 70 bus serves a substantial number of individuals using the bus as primarily a north-south system, as well. Specifically, 105 respondents, or 44.5% of the respondents, indicated that they were traveling in such a manner (see Table 6.1). Furthermore, 37.2% of trips (88) were from a neighborhood intersecting or adjacent to Grand Boulevard to a neighborhood also intersecting or adjacent.

Figure 6.9A: Origin Counts

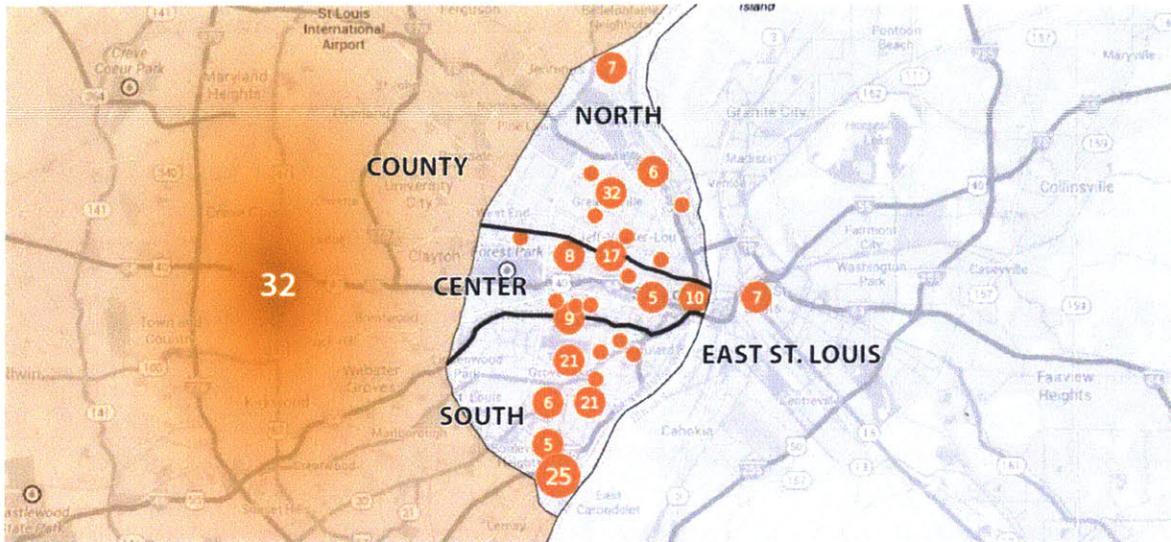
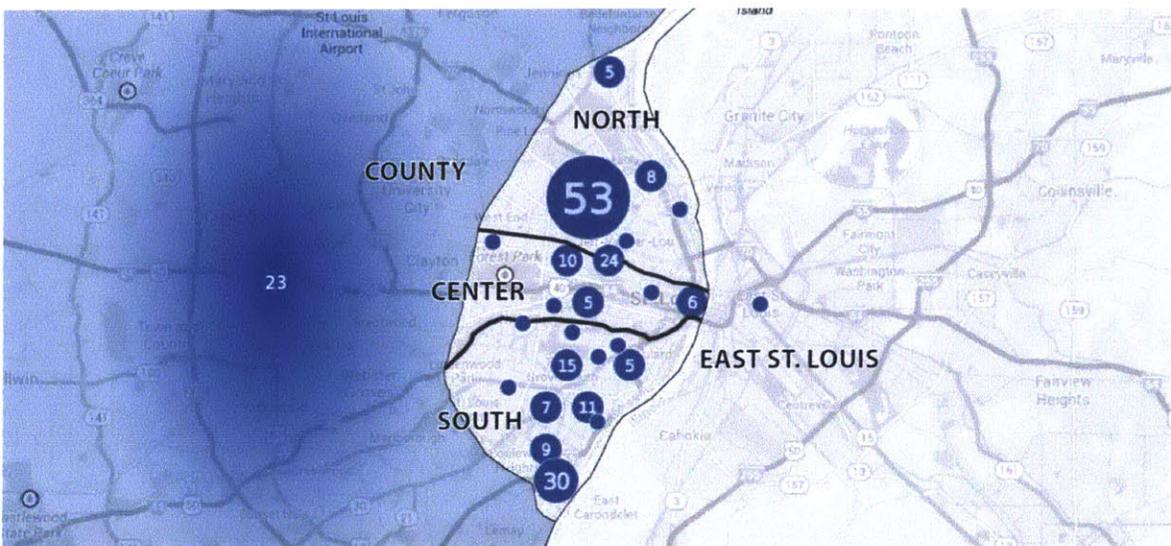


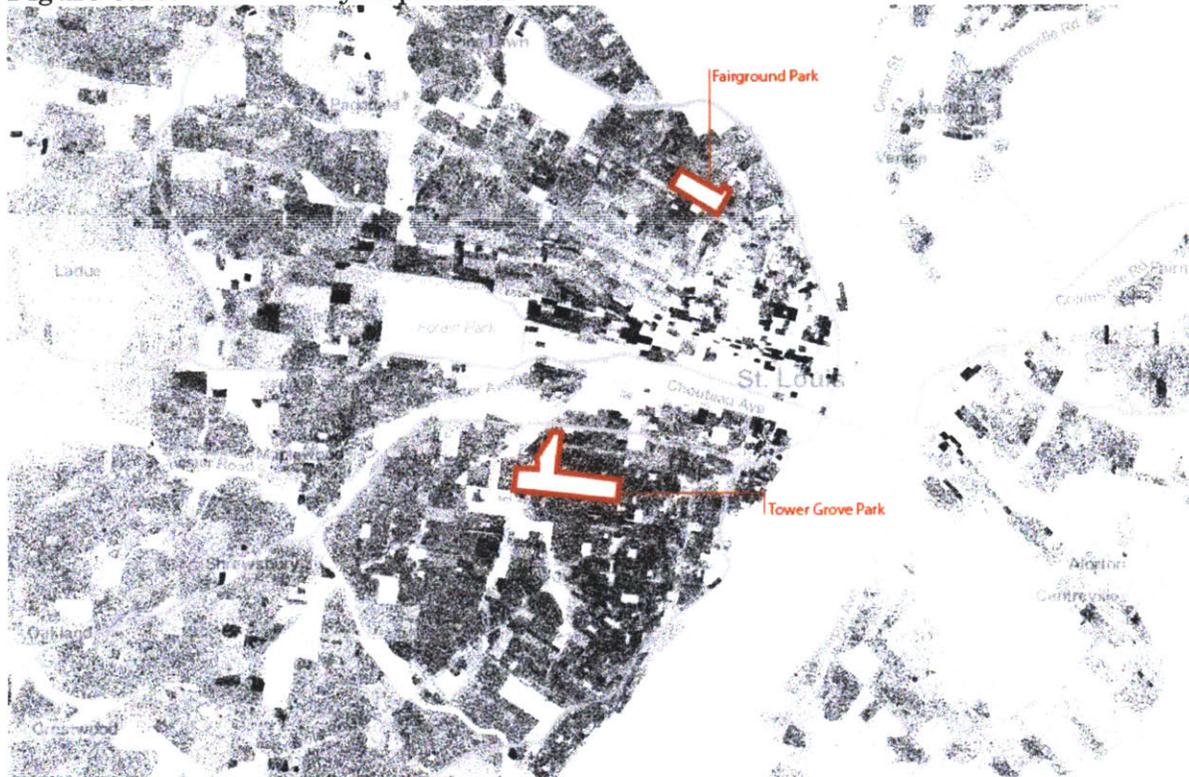
Figure 6.9B: Destination Counts



Transparent pochéd region indicates aggregated St. Louis County origins or destinations.

Looking more specifically at the reported neighborhood origins and destinations, the evidence suggests that destinations in the south were comparatively more dispersed than those in the north (see Figures 6.9A and 6.9B). Additionally, the number of responses that stated “Fairground” as their destination was particularly high, relative to other stated destinations in the north. There are multiple possible interpretations as to why this is the case. For one, the area around Fairground Park, specifically to the south and the west, contains housing stock that is fairly intact, relative to much of the north side’s level of decay. This is exhibited through higher densities, as seen in Figure 6.10. Thus, these areas may indeed represent a greater portion of all the north side’s trip destinations. Technically, however, they are not the Fairground neighborhood; they are Greater Ville and O’Fallon. It is possible that respondents mistakenly chose Fairgrounds because it is the major public space near these neighborhoods.

Figure 6.10: Dot Density Population



Dot density population map of St. Louis, Missouri. Fairground Park highlighted in red (Cable, 2013).

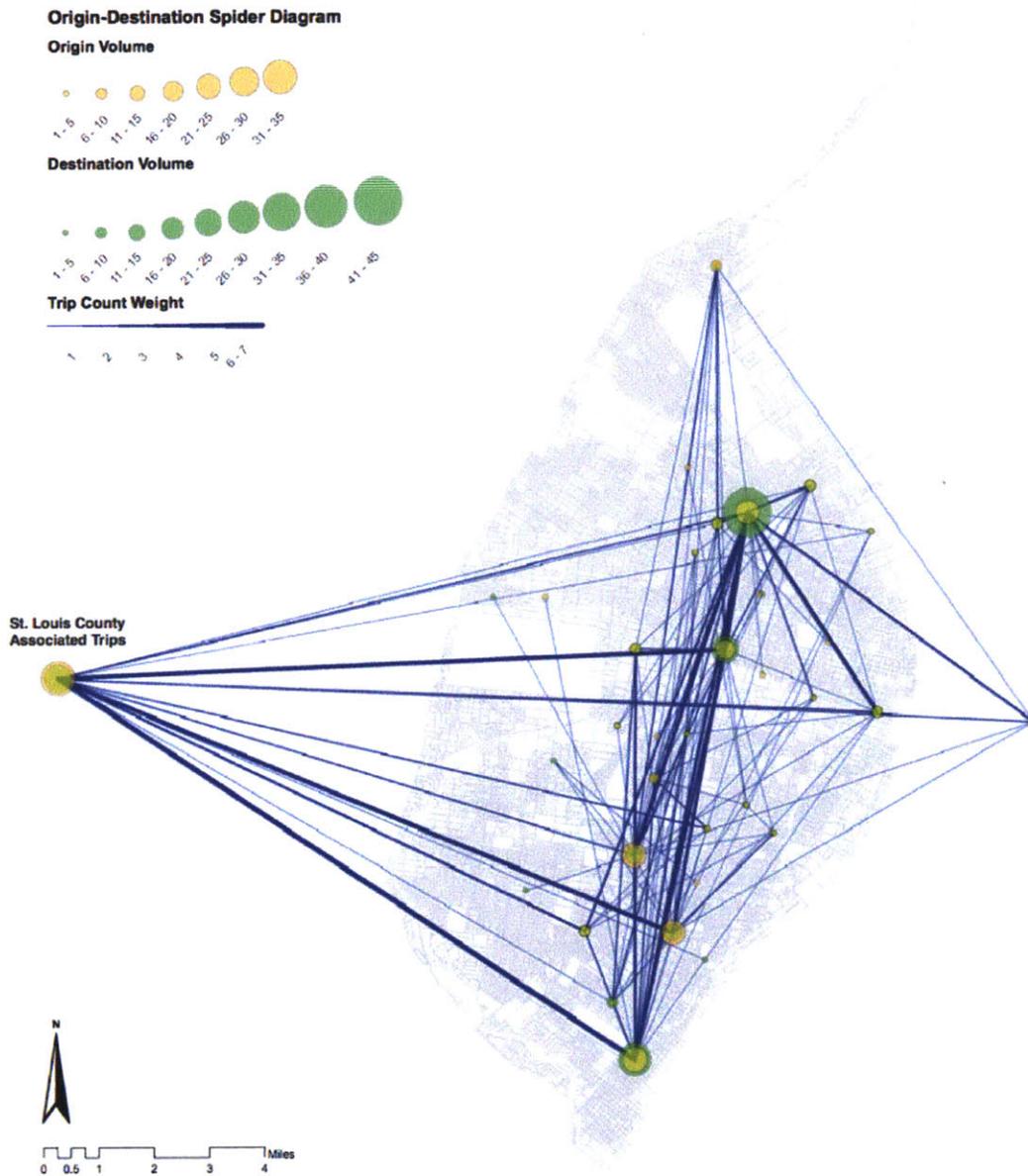
Another possible explanation is that individuals were heading to neighborhoods around Fairground, other than Greater Ville and O'Fallon, and simply identified Fairground as their destination instead. This may be due to major landmarks (such as the Fairgrounds Park, as distinguished from the neighborhood Fairground) having a stronger and greater sphere of influence in neighborhoods where there is a diminished quality of consistence in urban housing stock (Manzi et al, 2010).

Although the reason for the high number of Fairground-bound trips is unclear, it is peculiar in light of the fact that the Fairground neighborhood itself has a

relatively low level of zero car (and thus transit dependent) households (5.8%). In comparison, two major neighborhoods along Grand Boulevard in the south, Carondelet and Tower Grove South feature 13% and 18% zero-car households, respectively.

Another unexpected result from the report was the low number of trips to and from major employment centers in the urban core, including the Central West End, the location of Washington University Medical Center, a 230-acre, 12-city block institutional anchor (Katz, 2014). For scale, Barnes-Hospital, just one of the major employers in the medical hub, employs nearly 10,000 persons. Similarly, the downtown core region also represented a limited portion of all trips. The relatively low share of reported trips to these areas may be due to the fact that most people traveling to these parts of the city do not travel by the 70 Grand; or, if they do, they may have been less inclined to participate in the on-board survey.

Figure 6.11: Spider line map of origin-destination volumes.

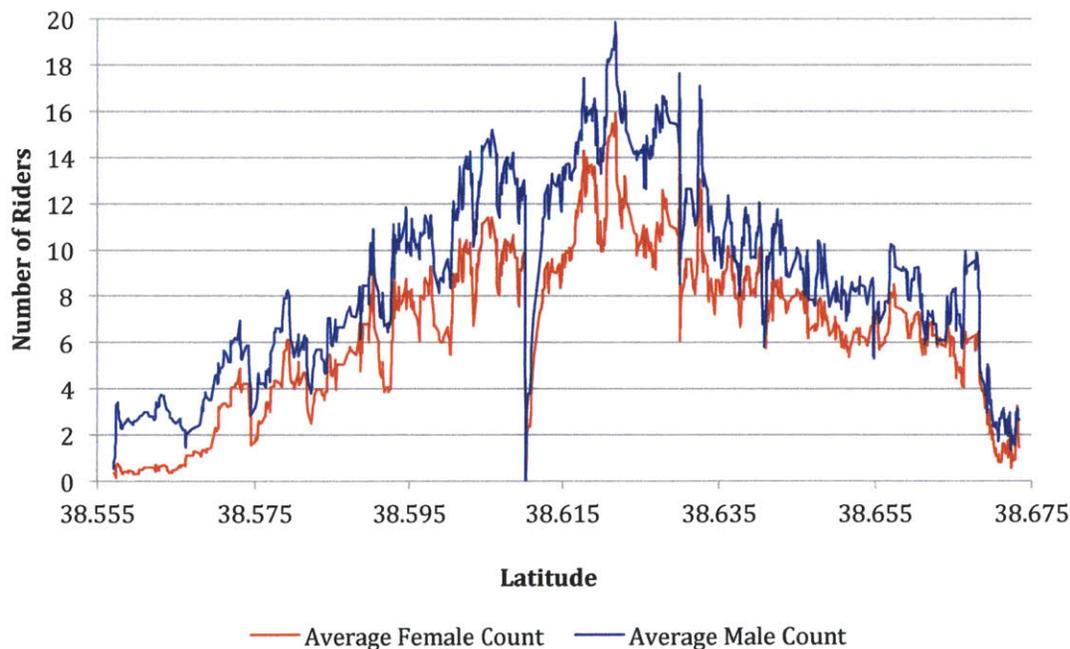


Some primary conclusions can be made regarding the gathered origin-destination data. First, while trips to St. Louis County, and out along the MetroLink corridor to Central West End, Downtown, Downtown West, and Illinois

do represent a significant portion of trips reported, trips along the Grand Boulevard central artery also make up a very important portion of total trips. While the survey failed to specify whether the respondent was going to or coming from their stated trip purpose (i.e. coming home from work versus going to it), it still sheds light on the 70 Grand line's role in relation to other routes, including MetroBus and MetroLink: 19.5% of recorded trips involved multiple buses, while 27% involved a MetroLink (light rail) transfer.

### Spatial Results: Aspects of Flocktracker Data Uploads

Figure 6.12: Ridership along 70 line.



All ridership upload points, from trip monitoring data, along latitude from south (left) to north (right).

As discussed in Chapter 4, a powerful functionality of Flocktracker is the ability to visualize survey responses and other variables, with temporal and/or spatial precision. In the case of Figure 6.12, ridership by gender is plotted over latitude. Thanks to the roughly linear path of the bus line, plotting results across latitude from south (left) to north (right) effectively charts the results. Figure 6.12 also reveals an aspect of the trip caused by my starting at the same location every morning. At latitude 38.6095, a sudden drop in average riders to zero can be seen. This is the result of me starting the bus every morning at that point (the intersection of Shenandoah and Grand) with zero riders. Overall, results show peak ridership occurring between Grand's intersection with Park Avenue and the Grand MetroLink.

Figure 6.13: Line 70 average bus speed.

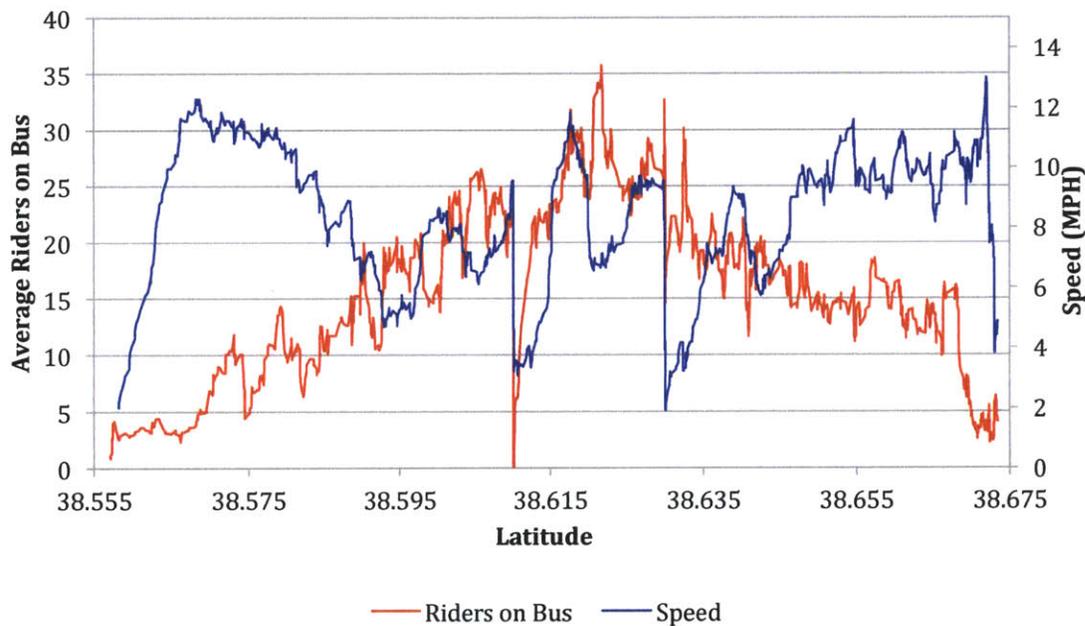


Heatmap indicating speed of 70 bus along route: lower speeds (red) observed near the Gravois intersection in South City, Grand Station by I-64, and to the north.

The comparison of variables over latitude enables comparisons ranging from purely performance based, to those that incorporate environmental aspects and perception levels. In Figure 6.13 speed is shown along the course of the route. Using the tracking data created from operating the application, speed is shown through a

heat map, highlighting where clustering of “bread crumbed” data points is densest. Speed can be calculated by determining distance covered between each point (uploaded at thirty second intervals). This result, when averaged, can be plotted with ridership, for example, as is shown in Figure 6.14.

Figure 6.14: Ridership and



Here, higher speeds at the northern and southern portions of the route are paired with lower average ridership. A brief spike in speed occurs from roughly latitude 38.615 to 38.622, which is the stretch of Grand from Compton Hill Reservoir Park to Chouteau Avenue, as can be seen in Figure 6.15. This stretch of Grand is an area with limited stops and dispersed institutional facilities related to St. Louis University’s Medical School.

Figure 6.15: Google Street View looking south down Grand from Chouteau Avenue



By segmenting the ridership and speed data by hour, aspects of peak and off-peak periods can be compared. Taking the ridership and speed comparison, we can observe that ridership in the south and north side is higher in the evening period (5:00 PM – 8:00 PM) than during the day. The center of the route, around Grand MetroLink, remains the period of highest ridership, save for in the evening when it peaks farther south, around the Arsenal and Grand intersection.

Figure 6.16A - B: Line 70 Ridership Segmented by Time of Day; A (8:00 AM – 10:59 AM), B (11:00 AM – 1:59 PM)

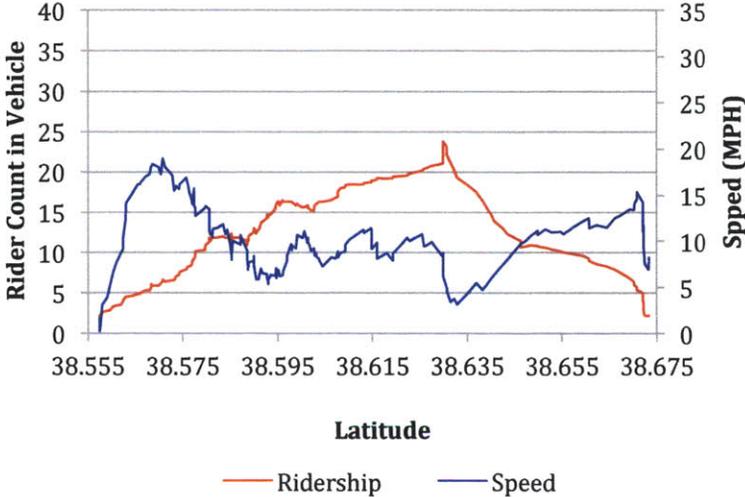
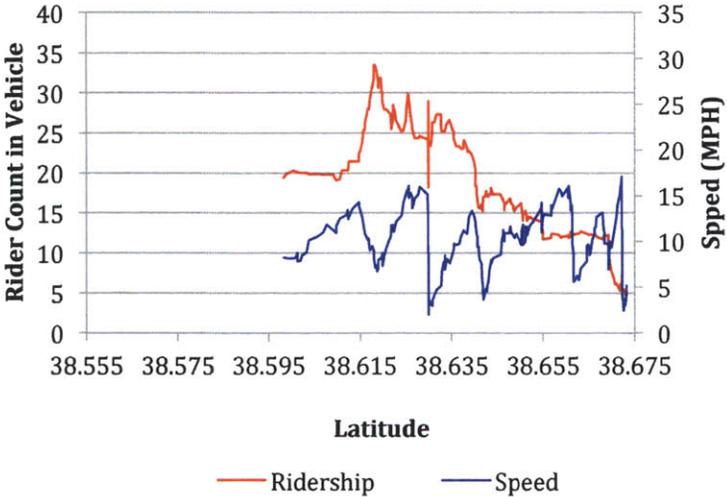
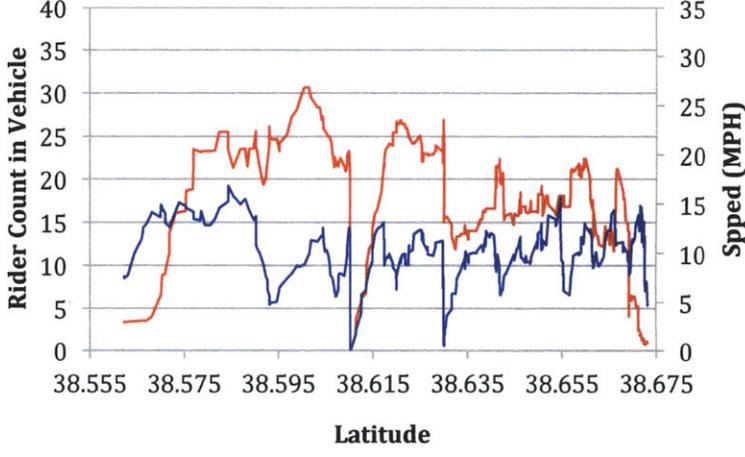
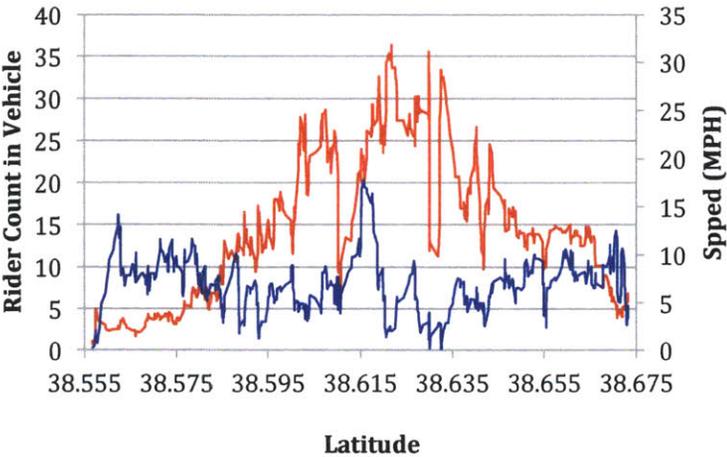


Figure 6.16C - D: Line 70 Ridership Segmented by Time of Day; C (2:00 PM – 4:59 PM), D (5:00 PM – 7:59 PM)

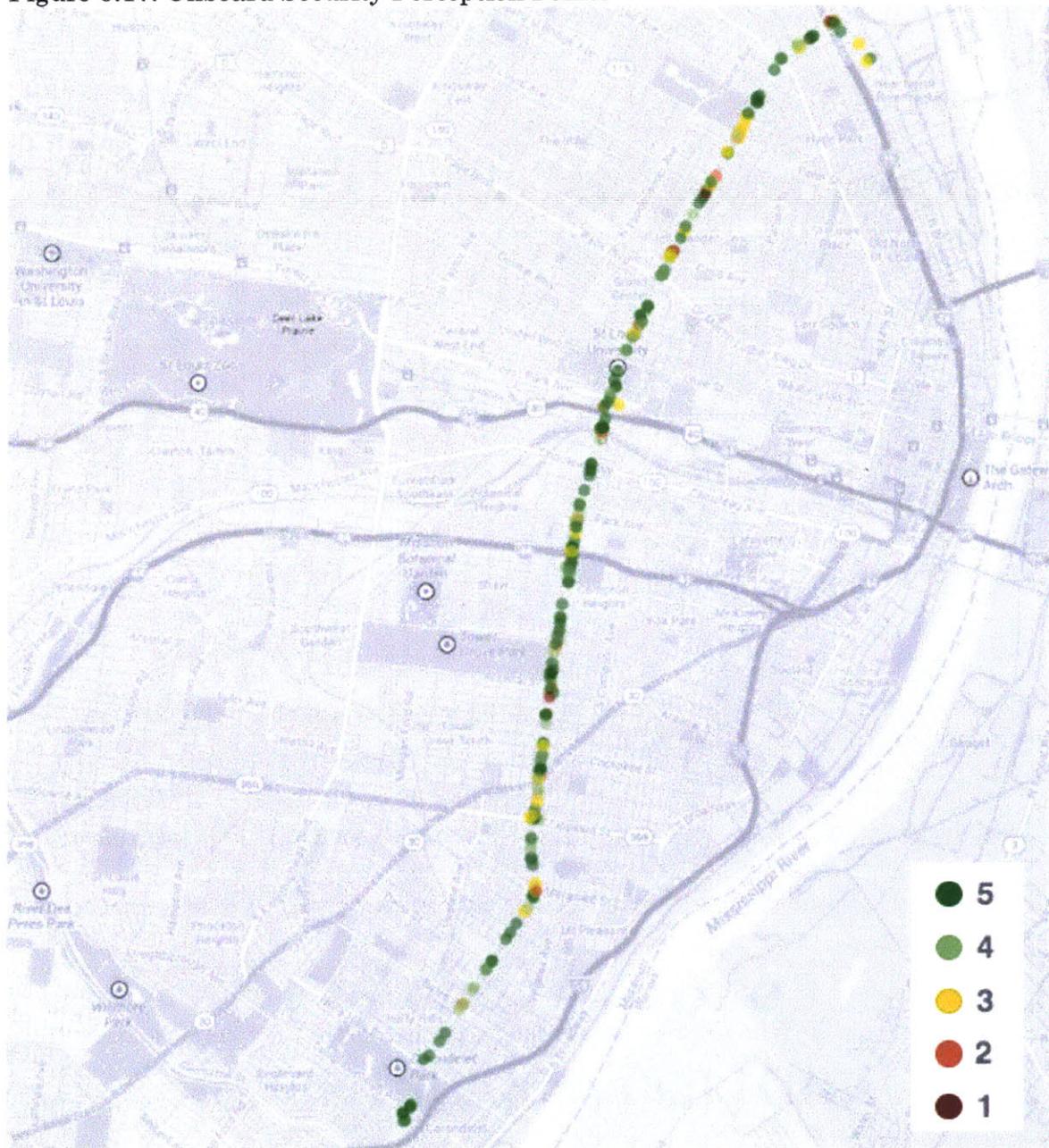


Another aspect of the Grand 70 observable from this segmentation is the consistency of north side ridership. While the south side features substantial fluctuations in ridership throughout the day; the north side's ridership, while lower on average, remains comparatively steady. AM peaks are not as easily comparable, due to the limited the "bread crumbed" data on buses operating in the south during this time period.

### **Spatial Results: Route Onboard Security Perception**

As outlined earlier, a core question of the survey how personally secure the respondent felt at that moment in time and space. The results are shown spatially in Figure 6.17.

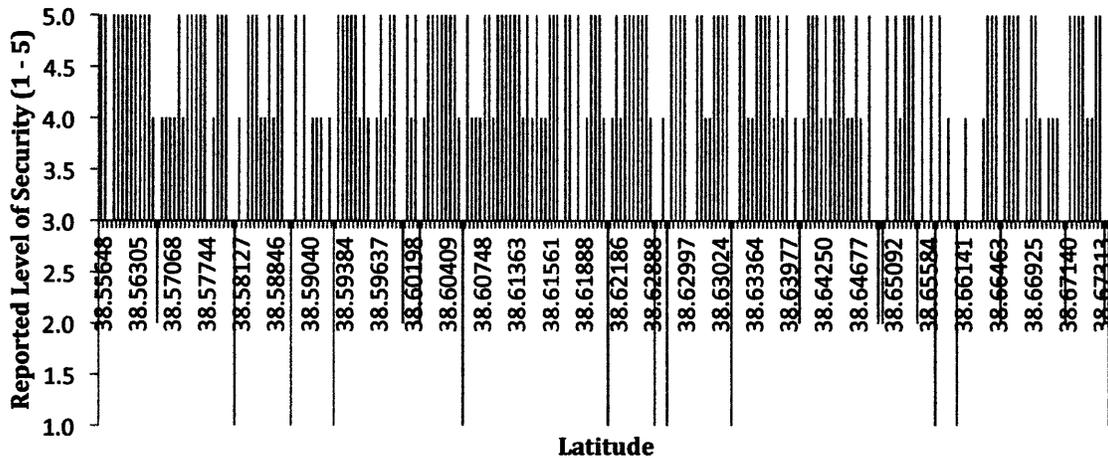
Figure 6.17: Onboard Security Perception Points



While a rating of “5”, or “Very Secure,” was the most common response amongst riders (52.1% of all respondents), there is a somewhat observable trend in onboard security perception as one moves farther north. Towards the south, ratings of 5 and

4 tend to be slightly more common than in the north. This tendency is visualized in Figure 6.18.

Figure 6.18: Point Perception



Plot of security perception by latitudinal points (south being on the left side of the graph and north being to the right).

Figure 6.19 visualizes ridership and average perceived level of onboard security along latitude. Notable along the plotted security rankings is the relative clustering of low security scores. These clusters exist in pockets along the corridor and, while more severe (that is, featuring lower averages) in the north, they nonetheless occur throughout the route. This can be seen in the moving average trend line generated in Figure 6.19, which also shows ridership levels, demonstrating how spatial security perception information can be correlated with other environmental factors to reveal possible relationships.

Figure 6.19: Average Point Perception Against Ridership

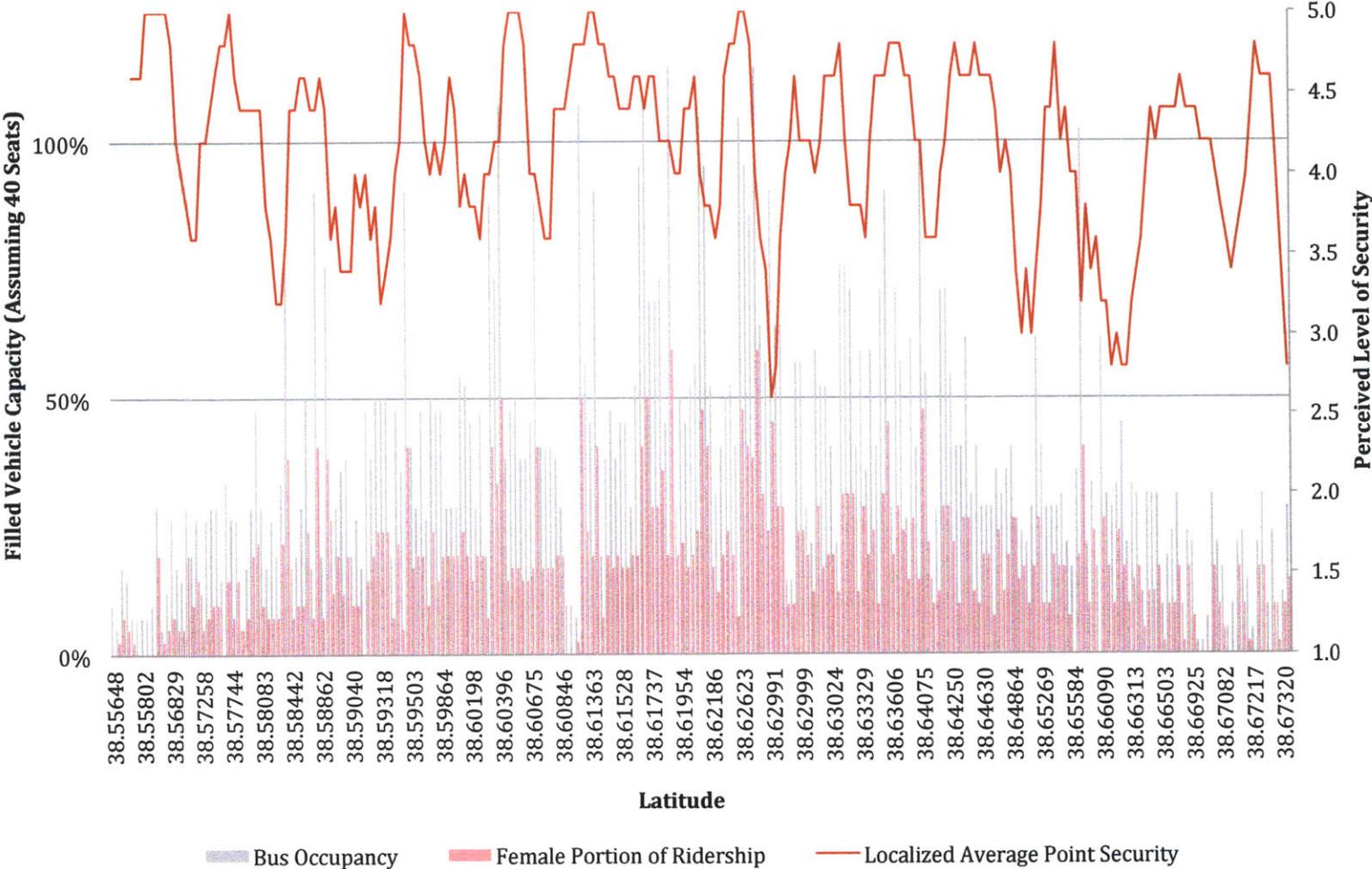


Figure 6.20: Crime Incidents Related to Onboard Security

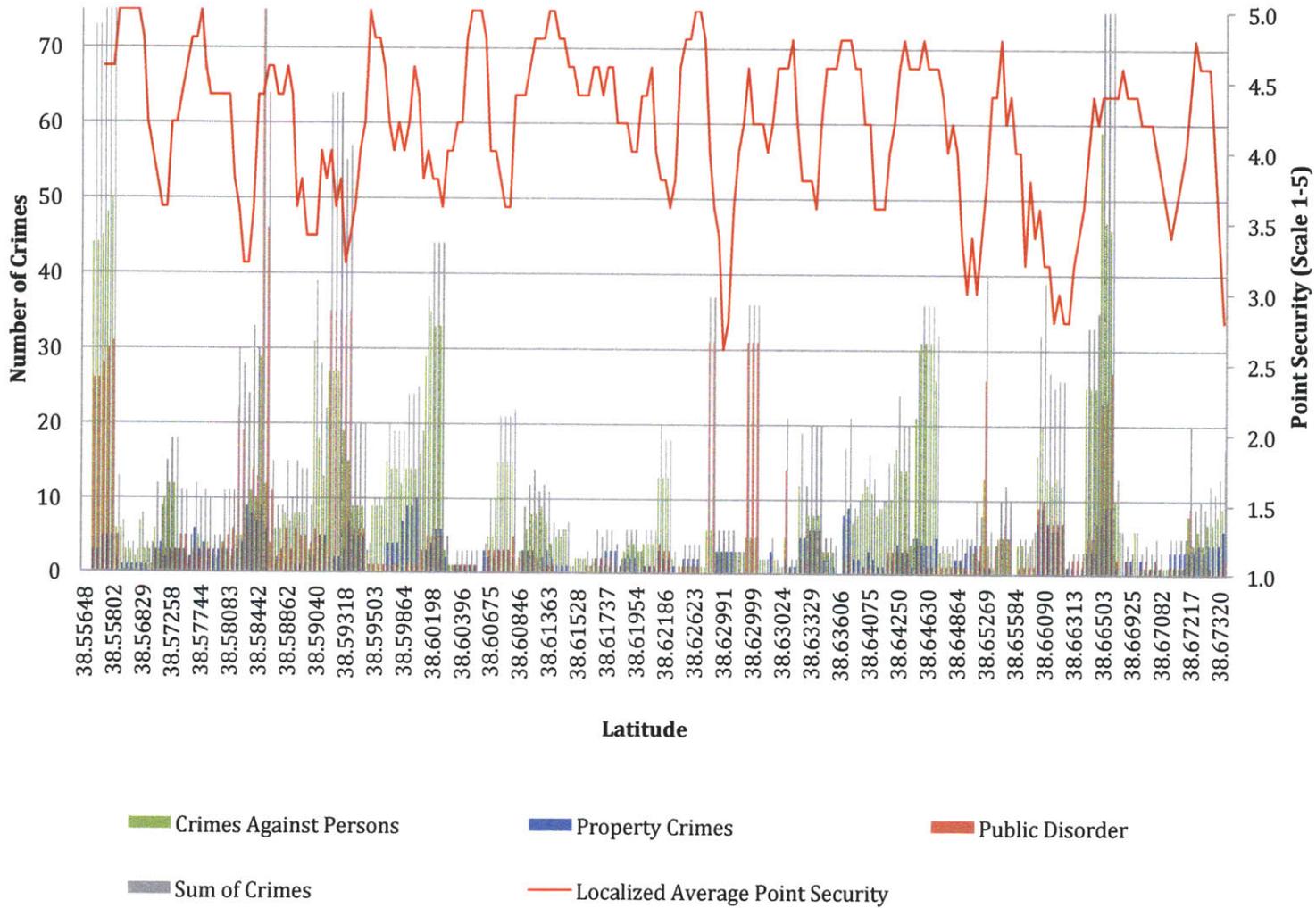


Figure 6.21: Crime In St. Louis City with 70 Line Route

### Crime Heatmap

Incident Rate (April 2013 to March 2014)

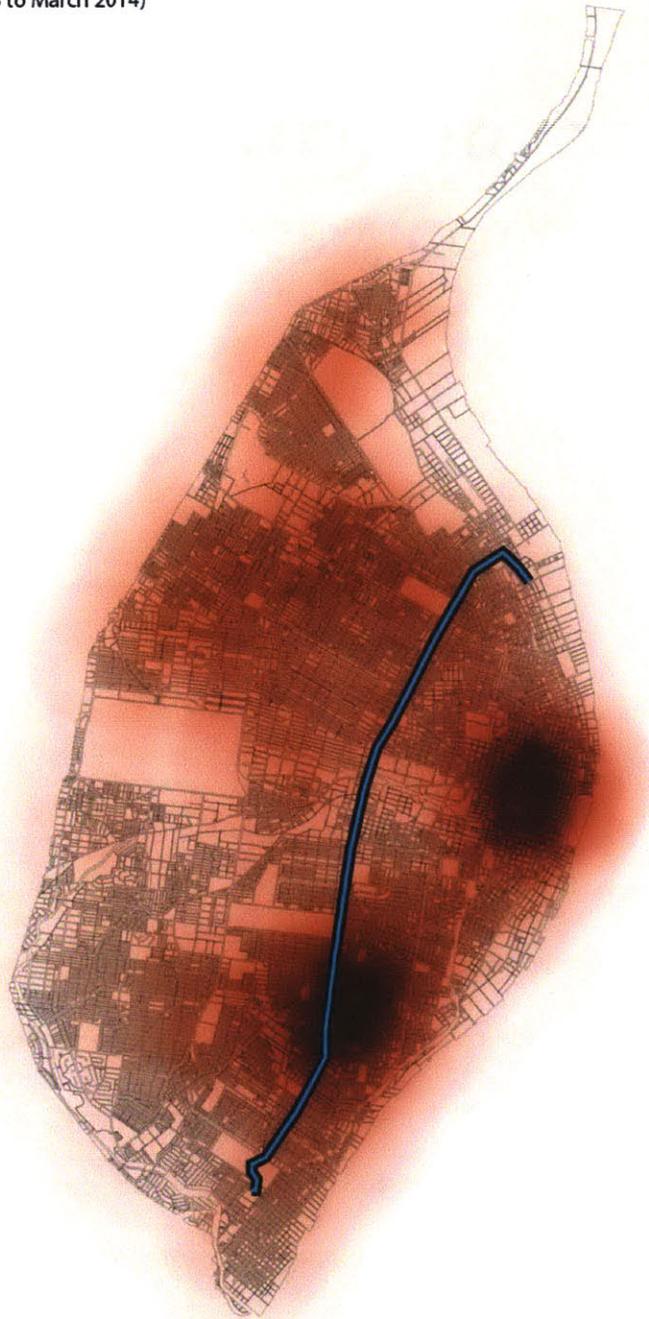
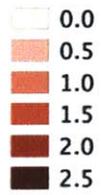
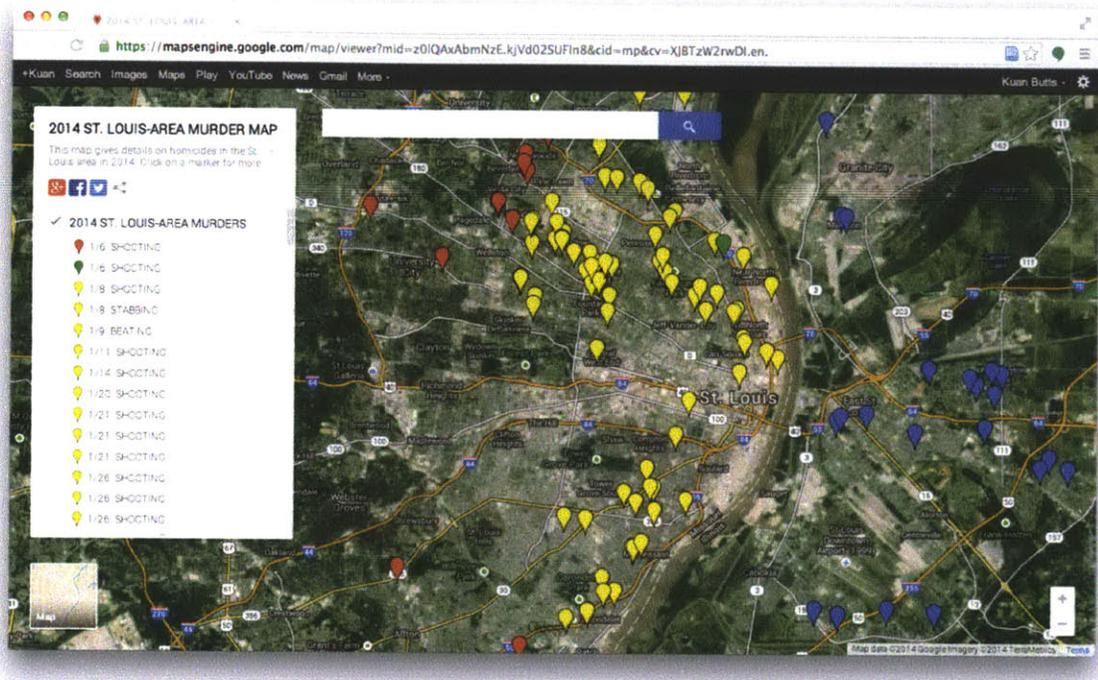


Figure 6.22: St. Louis Metropolitan Police Department Homicide Feed

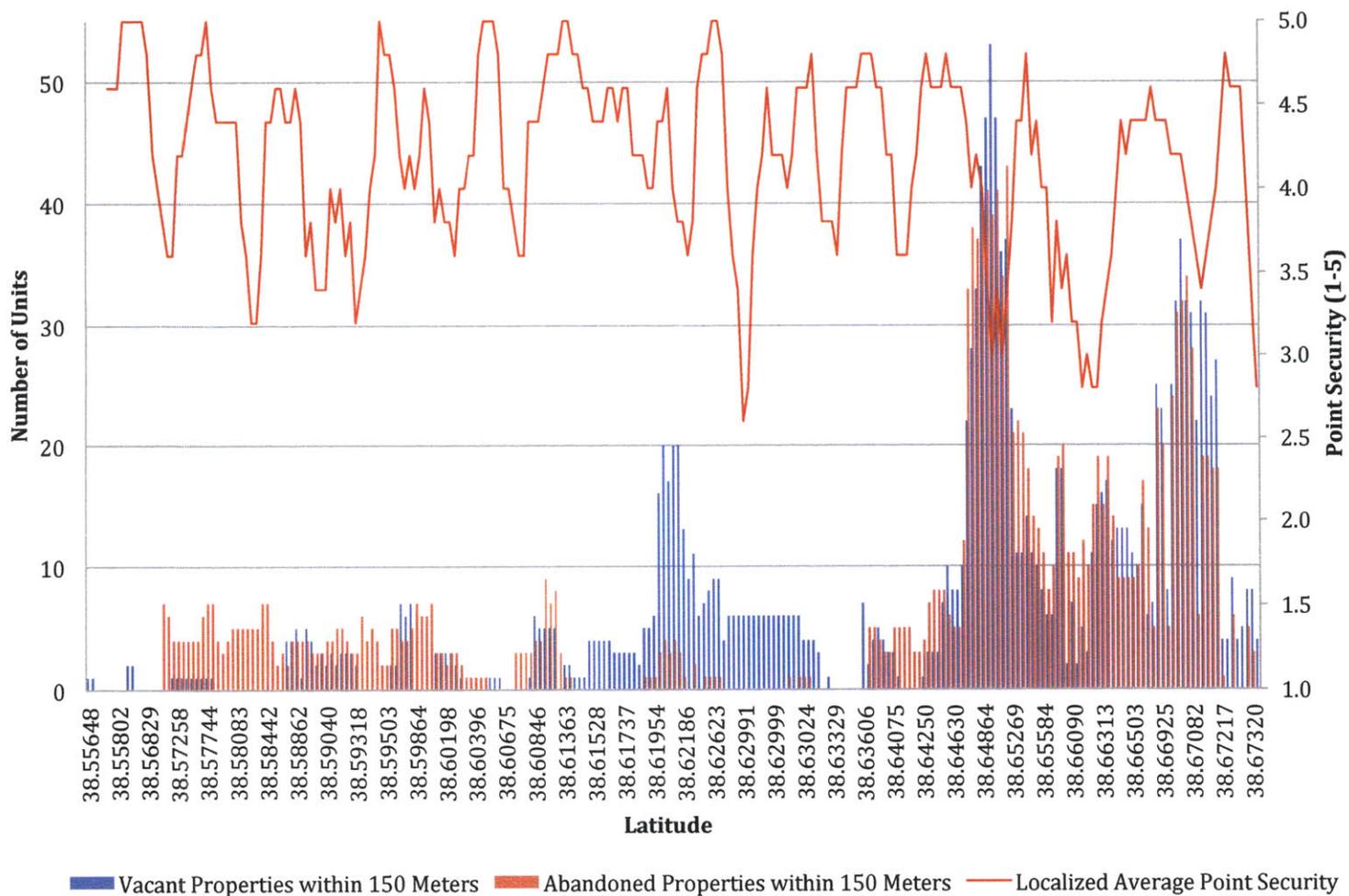


Interactive map of homicides recorded by the St. Louis Metropolitan Police Department in 2013, maintained by the St. Louis Post-Dispatch. Yellow markers indicate city homicides.

Another characteristic that might correlate with onboard security perception is criminal activity in the surroundings. The St. Louis Metropolitan Police Department (SLMPD) provides information on crimes through their online Crime Reports portal. Figure 6.22 shows a crime heat map for the city during the period April, 2013 – March 2013, created from SLMPD’s monthly datasets. To compare reported criminal activities to security perceptions, I aggregated all crimes within a 100-meter radius around each survey data point and sorted them by type. Crimes such as fraud and domestic abuse were not included; instead I focused on public crimes of three categories: crimes against persons (isolating homicide from others),

crimes against property, and crimes of public disorder. Counts of criminal activity, when plotted against average onboard security at latitudinal points spread along the 70 line bus route, do not exhibit a strong visual correlation with poor security perception. Additionally, points of high security perception do appear to correlate with lower overall crime levels. This relationship will be explored more completely in the statistical models in the following chapter.

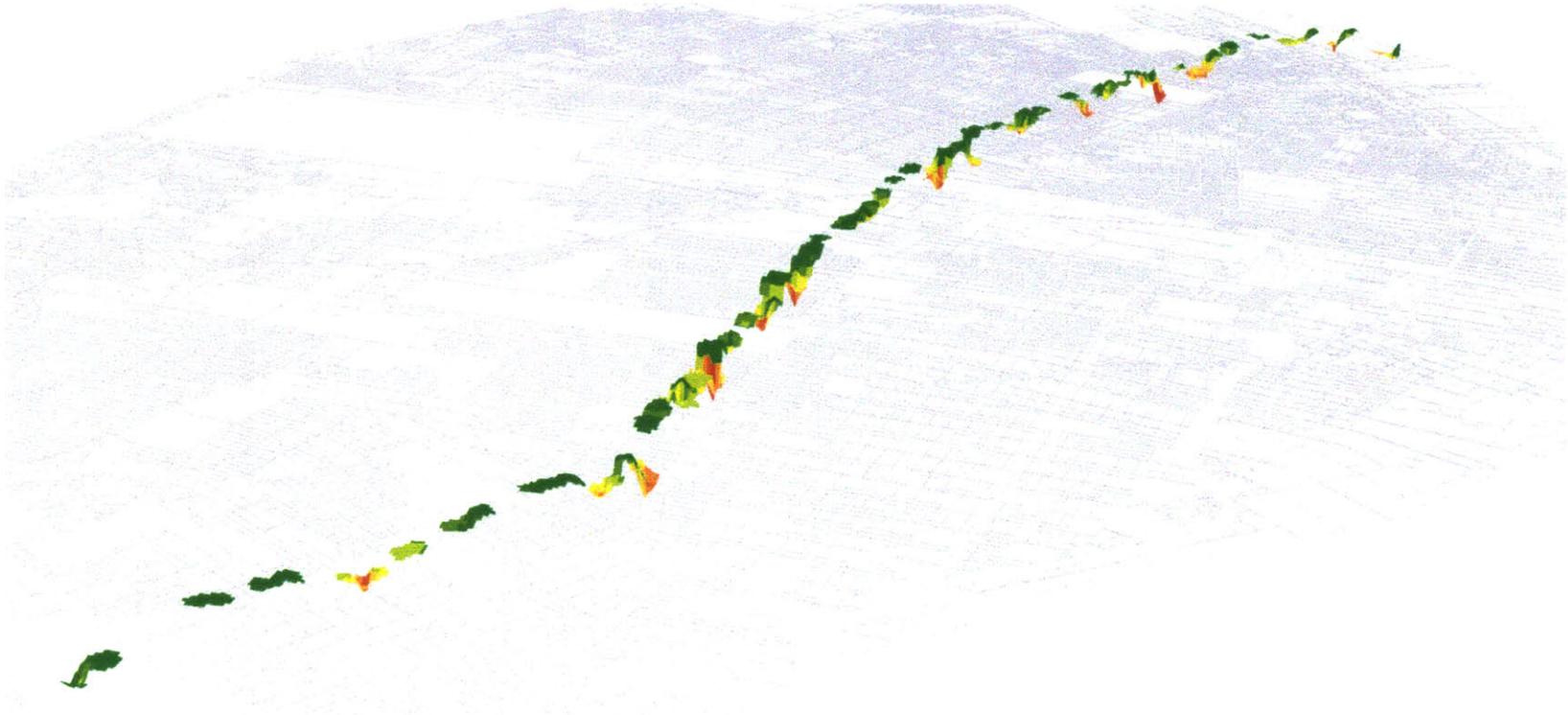
Figure 6.23: Vacancy Levels Related to Average Onboard Security



A similar analysis to the relationship between onboard security perception and instances of crime can be performed with regard to vacant properties. A 150-meter buffer was created around each point of reported security perception to tally the number of vacant or abandoned properties in the area, using data from St. Louis' tax records in their Access Assessor Dataset, from August of 2014. Vacant property and abandoned property are not included together, as sums, due to the fact that variable counts overlap in some cases. That is, some properties are both vacant and abandoned, whereas others are just one of the two. Plotting these data reveals a fairly unsurprising trend; the number of properties listed as either vacant or abandoned increases as one moves north, especially above latitude 38.64, which corresponds approximately with Delmar Boulevard, used by the city and this research as the “dividing line” between the central corridor and the north side. This region also does visually correspond with a slightly lower average onboard security perception rating. This relationship will be explored more completely in the statistical models in the following chapter.

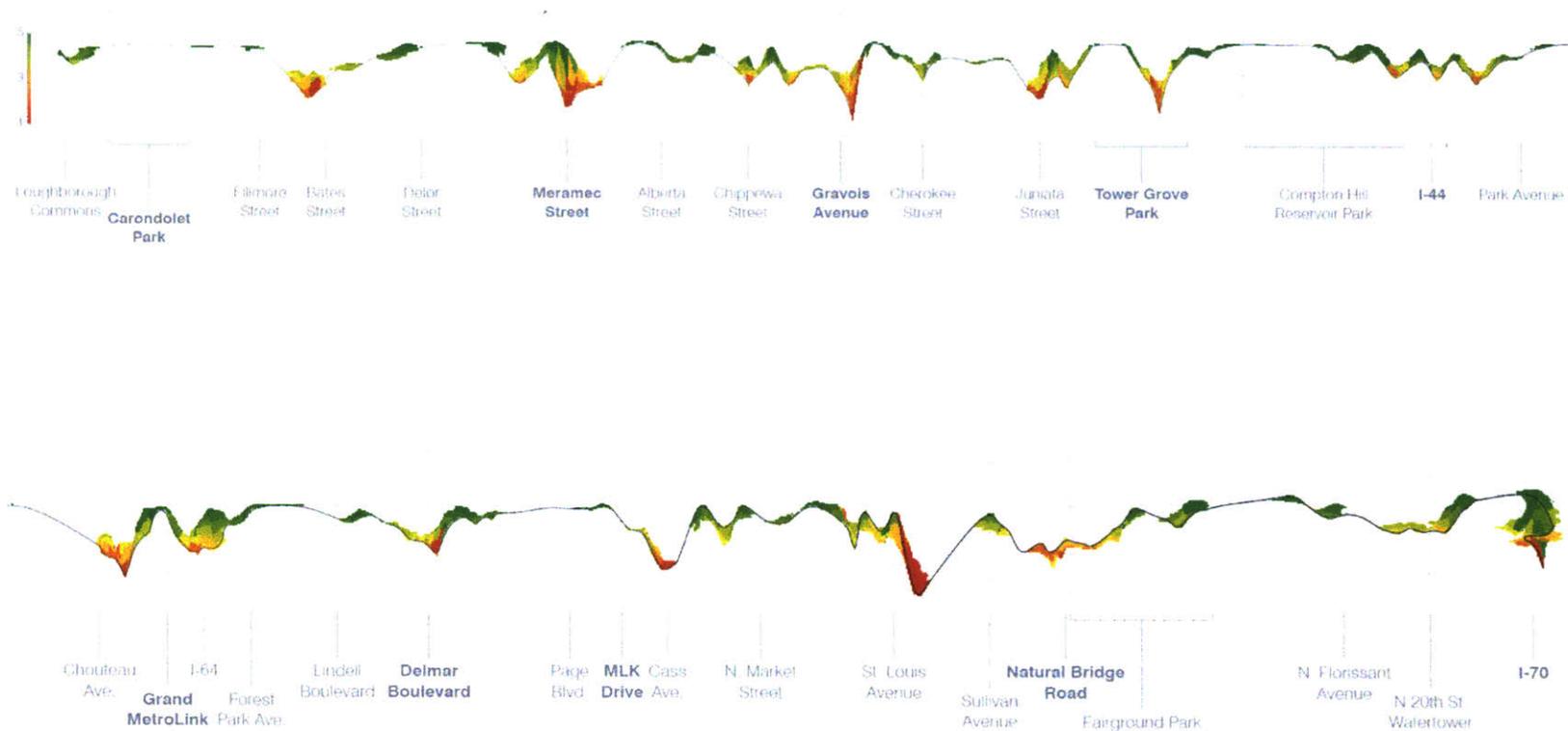
Finally, mapping, in three dimensions, average security perception – a topography – provides a potent visualization of users perceived security along the extent of the route (Figure 6.24 and 6.25). Low points (red) indicate lower perceptions of security and higher areas (green) higher.

Figure 6.24: Security Perception Topography



Average onboard security perception, looking northwest. Generated using the interpolated distance weighted function in ArcGIS to calculate weighted average raster point values along the route.

Figure 6.25: Elevation View of Security Perception Topography



To conclude, these graphs should serve as both an exploratory and an inspirational exercise. These initial examples are meant to explore the potential insights possible from such pairings. Particularly in pairing metrics against one another, such as speed and ridership, I intend to visualize more performance-based aspects of the route. In pairing environmental aspects to average reported security perception, I hope to have demonstrated and explored the potential for novel insight. In a more comprehensive survey, additional factors could be incorporated. Thus, numerous possibilities and variable pairings exist to further explore the route in such a manner in the future.

## CHAPTER 7: Statistical Model

This chapter demonstrates the power of the data collected in the St. Louis case, through the estimation of multivariate models of transit user perceptions of onboard security. The modeling shows the potential value of incorporating spatial and temporal resolution as enabled by the Flocktracker application.

### Descriptive Statistics

Selected variables are included in the table below, from variables discussed in the prior chapter.

Table 7.1: Descriptive Statistics Table

	Tot.	% Tot.	Blk.	Male	% Male	Fem.	% Fem.	Min	Max	Avg,	SD
<b>Total Dataset</b>											
<b>N</b>	236	100.00 %	191	138	58.47%	98	41.53%	1	5	4.15	1.12
<b>Age Brackets</b>											
<b>Under 5</b>	1	0.42%	1	1	0.42%	0	0.00%	4	4	4.00	0.00
<b>5-15</b>	7	2.97%	7	4	1.69%	3	1.27%	3	5	4.00	0.82
<b>16-24</b>	72	30.51%	61	44	18.64%	28	11.86%	1	5	4.08	1.20
<b>25-34</b>	62	26.27%	50	38	16.10%	24	10.17%	2	5	4.44	0.76
<b>35-44</b>	32	13.56%	25	16	6.78%	16	6.78%	1	5	4.06	1.22
<b>45-54</b>	34	14.41%	26	19	8.05%	15	6.36%	1	5	4.09	1.29
<b>55-64</b>	16	6.78%	12	10	4.24%	6	2.54%	1	5	3.50	1.46
<b>65+</b>	12	5.08%	9	6	2.54%	6	2.54%	2	5	4.50	0.90
<b>Gender</b>											
<b>Male</b>	138	58.47%	102	138	58.47%	0	0.00%	1	5	4.33	1.02

<b>Female</b>	98	41.53%	89	0	0.00%	98	41.53%	1	5	3.91	1.22
<i><b>Ethnicity</b></i>											
<b>Asian American</b>	5	2.12%	0	3	1.27%	2	0.85%	3	5	4.40	0.89
<b>Black / African American</b>	191	80.93%	191	102	43.22%	89	37.71%	1	5	4.10	1.15
<b>Hispanic</b>	5	2.12%	0	5	2.12%	0	0.00%	4	5	4.80	0.45
<b>Native American</b>	1	0.42%	0	1	0.42%	0	0.00%	5	5	5.00	0.00
<b>White</b>	33	13.98%	0	26	11.02%	7	2.97%	1	5	4.27	1.07
<b>Other</b>	1	0.42%	0	1	0.42%	0	0.00%	5	5	5.00	0.00
<i><b>Response Hour</b></i>											
<b>8 - 9 AM</b>	0	0.00%	0	0	0.00%	0	0.00%	0	0	0.00	0.00
<b>9 - 10 AM</b>	18	7.63%	17	0	0.00%	0	0.00%	1	5	3.94	1.06
<b>10 - 11 AM</b>	19	8.05%	19	0	0.00%	0	0.00%	2	5	4.00	0.94
<b>11 - 12 PM</b>	22	9.32%	16	0	0.00%	0	0.00%	3	5	4.32	0.72
<b>12 - 1 PM</b>	13	5.51%	9	0	0.00%	0	0.00%	1	5	4.15	1.28
<b>1 - 2 PM</b>	11	4.66%	8	0	0.00%	0	0.00%	3	5	4.45	0.82
<b>2 - 3 PM</b>	23	9.75%	20	0	0.00%	0	0.00%	1	5	4.17	1.37
<b>3 - 4 PM</b>	27	11.44%	24	0	0.00%	0	0.00%	1	5	4.04	1.16
<b>4 - 5 PM</b>	58	24.58%	42	0	0.00%	0	0.00%	1	5	4.21	1.10
<b>5 - 6 PM</b>	28	11.86%	27	0	0.00%	0	0.00%	1	5	3.86	1.53
<b>6 - 7 PM</b>	12	5.08%	8	0	0.00%	0	0.00%	3	5	4.58	0.67
<b>7 - 8 PM</b>	5	2.12%	1	0	0.00%	0	0.00%	4	5	4.60	0.55
<b>8 - 9 PM</b>	0	0.00%	0	0	0.00%	0	0.00%	0	0	0.00	0.00
<i><b>Origin Perception</b></i>											
<b>1</b>	16	6.78%	14	7	2.97%	9	3.81%	1	5	2.31	1.89
<b>2</b>	11	4.66%	11	7	2.97%	4	1.69%	2	4	2.91	0.70
<b>3</b>	30	12.71%	27	13	5.51%	17	7.20%	1	5	3.53	1.01
<b>4</b>	51	21.61%	38	30	12.71%	21	8.90%	1	5	3.98	0.81
<b>5</b>	128	54.24%	101	81	34.32%	47	19.92%	2	5	4.70	0.64

<b>Destination Perception</b>											
<b>1</b>	14	5.93%	13	7	2.97%	8	3.39%	1	5	2.14	1.56
<b>2</b>	9	3.81%	8	7	2.97%	5	2.12%	1	5	3.00	1.50
<b>3</b>	25	10.59%	22	13	5.51%	15	6.36%	2	5	3.16	0.75
<b>4</b>	56	23.73%	47	30	12.71%	26	11.02%	1	5	4.02	0.92
<b>5</b>	132	55.93%	101	81	34.32%	44	18.64%	2	5	4.69	0.64
<b>Primary Decision Aspect</b>											
<b>Comfort</b>	26	11.02%	15	17	7.20%	9	3.81%	1	5	4.08	1.20
<b>Convenience</b>	44	18.64%	31	28	11.86%	16	6.78%	3	5	4.45	0.73
<b>Cost</b>	34	14.41%	30	17	7.20%	17	7.20%	1	5	4.32	0.98
<b>Personal Security</b>	73	30.93%	67	39	16.53%	34	14.41%	1	5	3.85	1.39
<b>Timeliness</b>	59	25.00%	48	37	15.68%	22	9.32%	1	5	4.24	0.97
<b>Trip Purpose</b>											
<b>Caring for others</b>	7	2.97%	6	3	1.27%	4	1.69%	3	5	3.86	0.90
<b>Dining</b>	1	0.42%	0	1	0.42%	0	0.00%	5	5	5.00	0.00
<b>School</b>	29	12.29%	23	19	8.05%	10	4.24%	1	5	4.07	1.22
<b>Shopping / Errands</b>	47	19.92%	37	17	7.20%	30	12.71%	1	5	4.13	1.15
<b>Social</b>	47	19.92%	40	33	13.98%	14	5.93%	1	5	3.91	1.30
<b>Work</b>	105	44.49%	85	65	27.54%	40	16.95%	1	5	4.30	1.00
<b>Most Secure Mode</b>											
<b>Bike</b>	2	0.85%	0	2	0.85%	0	0.00%	5	5	5.00	0.00
<b>Car</b>	83	35.17%	65	45	19.07%	38	16.10%	1	5	4.07	1.16
<b>MetroBus</b>	88	37.29%	80	44	18.64%	44	18.64%	1	5	4.16	1.12
<b>MetroLink</b>	35	14.83%	24	25	10.59%	10	4.24%	1	5	4.17	1.22
<b>Private Shuttle</b>	5	2.12%	5	3	1.27%	2	0.85%	2	5	3.40	1.14
<b>Walk</b>	23	9.75%	17	19	8.05%	4	1.69%	3	5	4.48	0.79
<b>Bus Direction</b>											

<b>Grand 70 Northbound</b>	120	50.85%	95	78	33.05%	42	17.80%	1	5	4.16	1.12
<b>Grand 70 Southbound</b>	116	49.15%	96	60	25.42%	56	23.73%	1	5	4.15	1.13

### Comments on Data Preparation: Distance Measures

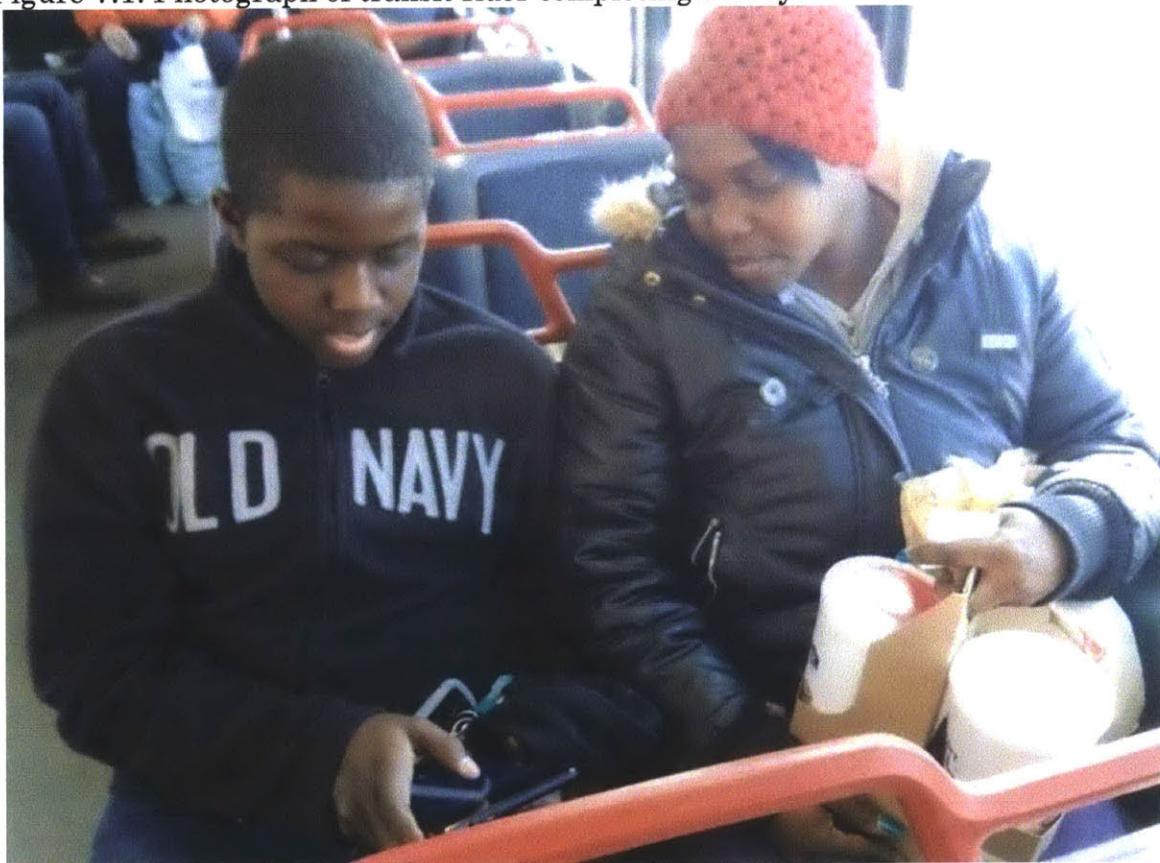
In developing the model, three metrics were produced in regards to represent the concept of “distance.” The initial purpose for developing such a concept was to understand if onboard security perception was related to how far north the survey took place. As discussed in Chapter 6, transit users’ security perceptions may well vary in space in St. Louis, reflecting the socioeconomic and racial disparities in the city, epitomized by the “north-south divide.” Reasonably representing this concept is a challenge, however. I initially considered employing latitude, a variable easily acquired in the data acquisition process, and naturally representing degrees north. The issue with latitude was that the relational aspect of it implied that distance was being measured not from anywhere within St. Louis, but rather from the equator. Conceptually, this was deemed somewhat inappropriate, given that riders would not consider their location in such a manner.

Table 7.2: Descriptive Statistics Comparing Distance (Miles) Alternative to Latitude

Method from Loughborough Commons	Minimum	Maximum	Average	Std. Dev.
Euclidean Distance	0.02708	8.58526	4.61513	2.35378
Route Distance Matrix Output	0.06214	9.38270	4.87968	2.43181

I also derived and tested two additional variables to represent the potential north-south influence. The first, Euclidean distance, measured absolute distance, as the crow flies, from the southernmost point of the route, the Loughborough Commons stop on route 70 (latitude 38.557582, longitude -90.262618) to the point of the specific observation. This approach only accounts for distance from the southernmost point, ignoring “north-ness,” which in itself has no meaning in this case, but represents simply straight-line distance.

Figure 7.1: Photograph of transit rider completing survey.



I also developed a third variable, “route distance,” which also measures distance from the Southernmost point on the route, but along the actual bus route. I developed this variable using the specific observation’s latitude-longitude and the route distance from that point to the Loughborough Commons stop.<sup>3</sup> The resulting values from this method of measurement are similar to those resulting from the Euclidean distance. In initial models, I tested each of the distance-related variables. In the end, I present the models including the route distance approach, which was both conceptually most meaningful and statistically most appropriate.

### **Model Specification and Estimation: Model Variations**

This section presents notes on development of the final mode. I estimated an ordered logit model, due to the ordered nature of the response variable (perceived security). The model estimations use clustered standard errors to account for surveyed riders in the same vehicle. Table 7.2 presents the final model results.

Ultimately, among the numerous models tested, results underscore the complexity of dealing with human subjects and their stated security perceptions on a bus. Understanding how these perceptions relate to a bus route across time and space highlights the myriad potential variables of relevance and the limitations in

---

<sup>3</sup> I used Google’s Distance Matrix API service, using the WALKING travelMode getDistanceMatrix input. This approach was possible since the route ran along Grand Boulevard in a linear path, thus making all walking routes match identically with the bus routes (in both cases Grand Boulevard is the sole road used in determining shortest paths).

both considering and capturing these myriad aspects in an operational model.

Nonetheless, their presence merits discussion.

All response variables from the survey were explored in testing various model specifications. From the prior discussed “bread crumbing” method, speed and route distance from the southernmost point of the 70 line, Loughborough Commons, were calculated. Age was segmented into 6 categories (under 16, 16-18, 19-24, 25-34, 35-49, 59-64, and 65+), to match those of the 2002 St. Louis Household Travel Survey. Not included in the final model was respondents’ ethnicity. This variable was found to be insignificant. This result may be partly influenced by the fact the great majority of respondents (80%) identified themselves as “Black / African American.”

Decision variables, the respondent’s primary concern when travelling (comfort, convenience, cost, security, or timeliness), were all variables included in the survey that the author (myself) initially assumed to be a significant component of onboard security perception. Timeliness, for example, seeks to explore the role of service reliability, defined as “how often service is provided when promised” (Transit Cooperative Research Program, 2003). Service reliability has been linked with perceived safety (Yavuz and Welch, 6). Studies indicate that the three attributes of service reliability (frequency, timeliness, and wait times) negatively associate with perceived personal security (Nsour, 1999). That is, lower frequencies, poorer timeliness, and longer wait times reduce a passenger’s perceived level of security (Reed et al., 2000).

Trip purpose options (work, school, caring for others, social, shopping/errands, dining, and other) were based off of those provided in the Trip Data File Dictionary, accompanying the 2002 St. Louis Household Travel Survey (East-West Gateway Coordinating Council, 2003). These options loosely correspond to those provided for personal travel in the National Household Survey (US DOT, 2010).

Spatially derived variables used while testing variations of the model included crime, property vacancy and abandonment, and neighborhoods' percentage of zero car households. Crime, vacancy, and abandonment counts were tallied through a 100-meter buffer drawn around each survey upload point. In addition to the aforementioned segmentation of crime by type, interaction variables between these crime types and gender are also explored. Crimes such as domestic violence and fraud are not included. The first type of crime is crimes against persons, including all forms of assault, save for homicide, which was placed in a category of its own. Property crimes represent crimes of a destructive nature involving another's property. Crimes of public disorder include behavior such as loitering, stalking, and disturbing the peace.

Vacancy and abandoned properties data developed for inclusion in the model was sourced from the St. Louis City's "Assessors Office Real Estate Records."<sup>4</sup> Prior research in Los Angeles (Loukaitou-Sideris, 2002) finds no statistically significant correlation between crime rates and number of vacant units. This research seeks to

---

<sup>4</sup> Due to the overlapping nature of many vacant and abandoned properties, I included the variables separately, in different models. The results were similar in either case.

explore if, regardless of the environment's influence on actual crime statistics, there is a correlation with fear of crime; that is, perception of security.

Zero car household data was sourced from the 2011 American Community Survey. This data, which is attributed to neighborhoods, was conjoined with survey results, based off respondents' stated origin and destination. Four segments of zero car households were explored: 0-10%, 10-15%, 15-20%, and 20+%.

## Discussion of Results: Final Models

Table 7.2: Model

Variable	M1	Coef.	z	P	Sig.
Pseudo R2	0.1194				
Prob > chi2	0				
Std. Err. adjusted for clusters in TripID	29				
Route Distance * Female		-0.1666107	-2.81	0.005	***
Vehicle Speed		0.0361758	2.24	0.025	**
Age 16-18		-1.483961	2.84	0.004	*
Age 50-64		-1.123633	-1.83	0.067	***
Age 16-18 * Female		3.503427	-2.68	0.007	***
Origin Zero Car HH 0-10%		0.680024	1.42	0.255	•
Origin Zero Car HH 10-15%		0.7754367	2.31	0.021	**
Origin Zero Car HH 15-20%		0.5241391	1.36	0.172	•
Dest. Zero Car HH 10-15%		-0.6060422	-2.04	0.041	**
Trip Purpose Work		0.6839805	2.26	0.024	*
Safest Mode: Bike		13.39583	16.13	0	***
Decision: Security		-0.4391163	-1.41	0.16	•
Vacant Properties Count		-0.0231859	-1.74	0.083	*

Public Disorder Crimes	-0.0374545	-2.62	0.009	***
Homicides * Female	-1.040172	-1.94	0.053	*
Property Crimes * Female	-0.1183076	-2.23	0.026	**
Public Disorder Crimes * Female	0.0288087	1.59	0.111	•

\*\*\* < 0.01, \*\* = <0.05, \* = < 0.1, • = < 0.2

The final model developed is shown in Table 7.2. Route distance from the southern terminus (Loughborough Commons), is significant only when interacting with gender. This suggests that there females' sense of security declines with distance northward. Vehicle speed is another significant factor; higher speeds correlate with higher levels of perceived onboard security. Higher speeds are present in areas with fewer stops (such as the more suburban regions to the far south on the 70 line).

Some age categories were found to be significant in the model. Specifically, those age 16-18 and 50-64 tended to feel less secure than others. On the other hand, teenage females age 16-18 correlated with much higher levels of average perceived security; this interesting outcome suggesting young males feel less secure than young women warrants further research. Previous research along the Blue Line in Los Angeles linked younger populations with increased crime rates (Loukaitou-Sideris, 2002, p. 142). Specifically, having a higher portion of the population younger than 24 years of age correlated with higher rates (crimes per 100 riders on the Blue Line), while increases in the portion of the population over 65 years correlated with decreases in the same rates (Loukaitou-Sideris, 2002, p. 142).

Lower percentages of zero car households at the respondents' stated origins correlated with higher perceived onboard security. Why this is the case is unclear. As mentioned earlier, the survey failed to determine if individuals were headed to or from their stated trip purpose (i.e. going to or from work). Similarly, insight is limited as to the significance of proportion of zero car households at the stated destination. According to the final model, lower levels of zero car households (10-15%) correlate with lower perceived levels of onboard security. Combined, these coefficients suggest that individuals, when heading to areas where car ownership is lower, tend to feel less secure. On the other hand, in areas where they have already been to and left safely (presumably), they feel more confident. Thus, it is the uncertainty of arriving at a destination with lower auto ownership that correlates with reduced security perception, whereas that uncertainty is no longer applicable to the location from which the trip originated.

Trip purpose was initially included as a variable in order to explore if more "standard" trips, such as work trips, correlated with higher levels of perceived point security. The hypothesis is that such trips would have become a standard experience to the respondent and thus likely something they had become used to and comfortable with. The results support this hypothesis – work trips do correlate with higher levels of personal security perception.

In terms of the perceived most secure mode of travel, bikes correlated strongly with high levels of personal security perception. One possible reason for this is that those who feel biking is the most secure form of transportation likely feel more

confident biking in urban environments. Thus, their confidence in personal security is likely greater than those who do not feel secure biking around a city. That is, these individuals may just feel more secure in public spaces.

Of the decision variables respondents indicated were most important to them when travelling, security alone was significant. Riders indicating security was their primary concern when making travel decisions felt, on average, less secure than others. That those who are most concerned about their personal security feel least confident about it makes sense. These concerns lead to the final portion of the model, which deals with crimes. As previously discussed, crimes have been broken down into four categories deemed applicable to the research.

Public disorder crime counts correlated with lower perceived onboard security. Public disorder crimes might be considered an indicator of a greater presence of the behavior that earns these types of infractions in an environment. For example, loitering crimes might be indicative of an area where loitering of an unsavory kind might occur – and frequently. Thus, such attributes may indeed contribute to an inhospitable climate, one that makes riders feel less secure than particular instances of crimes against a person or property.

Interaction variables created between female gender and homicides, property crimes, and public disorder were all significant. Interaction variables between female gender and homicides and female gender and property crimes correlated with lower levels of perceived onboard security. On the other hand, female gender and public disorder crime counts correlated with higher reported levels of perceived

onboard security. While public disorder crimes, generally, are associated with lower levels of perceived security; this effect is less so for females. Why females feel more secure than males in areas where public disorder crimes are reported may be due to the fact that policing is occurring in these areas. Thus loitering and other unwelcoming behavior is patrolled, whereas other areas may exhibit the same environment and persons that would have been arrested, but were not due to lack of policing. This might be attributed to a sensitivity to poorly lit and understaffed environments that is higher amongst females in particular (Trench et al., 1992). Males, on the other hand, may feel at greater risk, due to the nature of the public disorder and/or the potential police presence in response to it.

## **Discussion**

While preliminary and indicative, the model results underscore the complex nature of a perception-based metric, such as personal security. Yavuz and Welch (2010) elaborate on this phenomenon, citing Koskela (1997), who comments that fear of crime does not come from “mathematical functions of actual risk but are rather highly complex products of each individual’s experiences, memories, and relations to space (Koskela, 1997, p. 304).” Thus, the measures of security perception I derive from users and the statistical models attempting to explain these perceptions can serve as just one aspect of an analysis of the 70 Grand bus line. As the 70 Grand line remains the most used bus line in the city of St. Louis

(despite a recent drop in ridership of roughly 3%), identifying vulnerable points along the routes, such as areas that suffer most from congestion, poor security perception, or overloaded buses provide information that can be employed in improving the network.

While I examined riders' perceived on-board security for the pilot research demonstration, Flocktracker's versatility could assist in diagnosing other dimensions of system performance. For example, crowding on the 70 Grand has been enough of a concern to prompt the BSDA to procure 15 new 60 foot-long articulated buses. All 15 of these buses will be operational on the route by the end of 2014. As stated by BSDA, these articulated buses are designed to reduce crowding on the 70 Grand bus line by accommodating roughly 25% more passengers than the standard 40 foot-long buses currently running the route, increasing capacity from 40 seats to 54. At a cost of \$430,000 per bus, these vehicles were purchased refurbished for roughly half of what the cost of a new bus would have been. Combined, the purchase price for all 15 buses was roughly \$6.4 million (MetroSTL, 2014).

The articulated buses have been purchased to address an acute and specific problem along the route, peak-hour crowding. As shown in Figures 6.19A-D, the south side appears to exhibit greater fluctuations in ridership over the course of a day. The north side, on the other hand, has lower average ridership, but maintains that level fairly consistently throughout the day. So, just as an example of the value of spatiotemporal data collected through the Flocktracker application; and given the

investment and attention by the BSDA on improving 70 line service, future research could explore the comparative impacts of investment in increased frequency versus increased bus capacity. The interaction of frequency with personal security is also important; research suggests that passengers' concerns regarding their personal security might be reduced through more frequent and on time bus performance; particularly in public spaces where riders are unable to control their environment and thus tend to feel more vulnerable (Yavuz and Welch, 2010, p. 6; Farrall et al., 1997; Johnston, 2001; Crime Concern, 2002). So, while additional research is clearly needed, my Flocktracker demonstration on the 70 Grand shows how the application can quickly be adapted and deployed for highly nuanced analysis of a range of performance dimensions.

Using Flocktracker to understand complex variables, such as security perception, offers the opportunity to correlate perception and rider behavior with environmental and temporal aspects. Issues of specific locational security have been addressed by actions such as placing officers at high impact locations along the route (the most oft cited solution to dealing with problem security points; Transit Cooperative Research Program, 1997), as well as by investing in some physical urban amenities improvements along the route.

“Many of the current ways of tackling fear in the city seem to rely on intensified forms of surveillance of potentially unruly people and places,” write Bannister and Fyfe in “Introduction: Fear and the City (Bannister and Fyfe, 2001, p. 811). While placing officers does address symptoms of a public security issue, Flocktracker can

assist in teasing out more nuance in the nature of its cause. Only by seeking to understand, more holistically, underlying causes of aspects of a trip, such as perception of personal security, will BSDA be better positioned to improve service in the long run.

### **Limitations**

Ultimately, the application, the surveying effort, and the resulting models and analysis serve to demonstrate the insights possible from a survey effort that lasted less than a single work week, with only one individual carrying out the field work. As such, this should be considered first and foremost as a demonstration of the potential of the application and an example of what might be the initial efforts of a more complete, application-based, urban diagnosis. The validity and reliability of the data in this particular application are limited by a number of factors, such as:

1. The survey was performed by myself, a male in his mid-20s who might have been unconsciously selective in who he chose to interview.
2. In that same fashion, respondents might have been more or less likely to respond to me, given my race and age.
3. In addition, responses may have been influenced, given my age, gender, dress, or any number of other unintended social cues.
4. The time period was very brief and could have represented unusual system performance, ridership, or user characteristics.

This effort nonetheless shows what can be achieved in a remarkably brief amount of time. For example, I was able to pause from riding the bus at a midpoint on Friday and, while connecting to the MetroLink downtown to Metro's office to present my research, take all the data which had been automatically input into Fusion Tables, and create ridership plots by latitude as well as descriptive statistics of the overall survey results (completed just a few moments before). Such rapid deployment and generation of data, in the field by a single individual, would almost certainly be impossible using traditional data collection techniques.

## **CHAPTER 8: Future Potential**

### **Context for Future Development**

Geospatially tagged surveying is still an emerging field. Private actors, such as mSurvey and Fulcrum exist, albeit with unclear financial viability. Meanwhile, non-profit initiatives such as Modi labs' Formhub have achieved success at scale, but suffer from limited institutional funding, necessary to support and maintain such a service. Flocktracker, to date, has been a research tool, developed in an academic institution (MIT). Whether and in what form it survives into the future remains to be seen.

### **Goals for the Flocktracker Application**

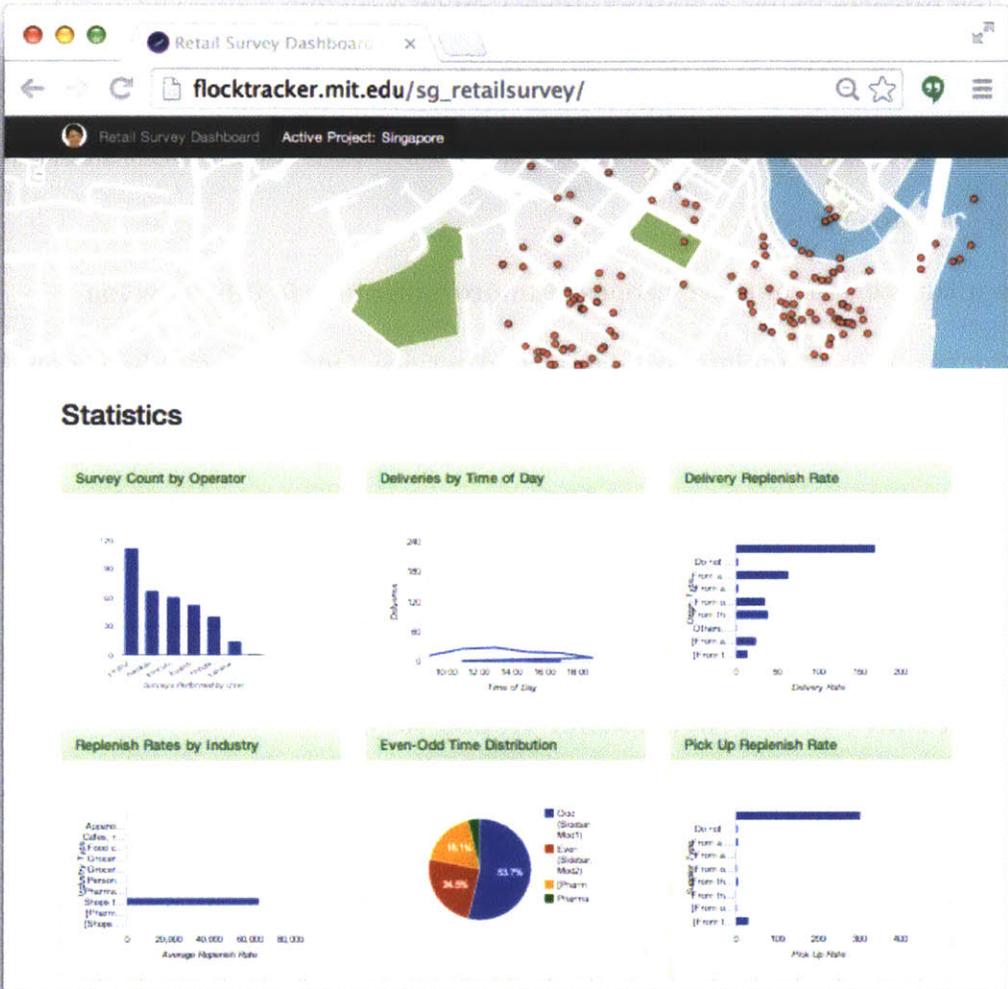
Ultimately, for Flocktracker to become a successful and scalable service, a custom back end must be developed. By weaning Flocktracker from its reliance on Google Fusion Tables and developing a custom database framework, creators of projects will eventually be able to develop surveys, manage data, and visualize results themselves. Currently, users must be manually added, due to limitations with Google's Fusion Tables. Furthermore, updates and adjustments to projects must also be handled manually. This severely limits the potential for others to easily begin to use the application and is presently the most severe roadblock to a stable and public-ready version of the application.

Other advantages from moving off of Google include the ability to avoid various limitations imposed on Google's Fusion Tables API. For example, the application is presently tied to a single account, in order to access Google's Fusion Tables API capabilities. Among the downsides to this approach: a single Google Docs account is essentially distributed within the copy of the application to all users; and, call limitations are strict and limit large numbers of users from accessing the application. For example, per table query limits are presently capped at 30 write requests per minute. What this means is that, when multiple users are submitting data to a single project table, the application must effectively check to make sure the data were accepted. If too many users upload data simultaneously, the data equivalent of a traffic jam could occur, causing data to be significantly backed up while waiting to access Fusion Tables.

What exacerbates the problem is that the current iteration (Version 6) of the application breaks up long surveys into smaller portions. This design element is necessary to allow a queue to form that holds both survey responses and automatically uploaded locational data. By separating the survey results into smaller portions, large blocks of data do not sit in the queue, a risk that requires the application to cache these large files while continuing to operate all other tasks, including developing new survey response uploads. This means that, instead of a survey response equating to a single API call, it can now become 5, 10, 100, or more API calls. With long surveys, this can lead to data quality upload issues as a single survey. For example, in a recent pilot application (a survey of retailers in Singapore

in July 2014) showed that a long survey (over 60 questions) could lead to Fusion Tables temporarily blocking data uploads, including rejecting only portions of the survey once it had been broken down into a series of uploads. Because each instance of the application on a smartphone is not aware of others uploading to the same table, simultaneously, a queuing strategy or some form of cross-device queuing is not presently possible. The consequence is a chaotic upload queue, exacerbated by increasing number of simultaneously operating instances of the application that could become unstable and unreliable when larger numbers of phones are running simultaneously.

Figure 8.1: Example of a Personal Dashboard (Prototype)



Dashboard visualizations taken from prototype developed for summer 2014 research in Singapore.

Finally, with a custom database, a public interface to peruse other data created with the application is possible. Creating such a searchable repository could create a digital environment to foster social exchange and access of sourced data. Such a

system would enable users planning a survey initiative to explore ongoing or completed projects under a similar subject. Such a system would serve to not only better host data gathered with the Flocktracker application, but it could also become a social repository for urban data in the same fashion that a service such as GitHub operates for code.

The ultimate goal of these proposed improvements is to fully develop Flocktracker into a complete, self-contained service. Once this goal has been achieved, only light technical maintenance will be necessary for continued facilitation of the service.

### **Technical End Notes: Flocktracker Repository**

As of mid-August 2014, a completely refactored version of the application has been pushed onto the GitHub repository.<sup>5</sup> The refactored application architecture is based on a class system with the intent of organizing the application and making it highly readable to interested authors. The next step for the application with this repository, completely open to the public, is to foster a developer community to assist the improvement of various application functions, particularly once the application has been migrated from Fusion Tables to an AWS or similar backend platform. Currently, the application has also been forked to Urban Launchpad ([www.urbanlaunchpad.org](http://www.urbanlaunchpad.org)) and a modified form of the application is being explored for use as a garment factory quality-monitoring device to connect consumers to

---

<sup>5</sup> This repository can be found online at [github.com/UX-Mexico/surveyappmaker](https://github.com/UX-Mexico/surveyappmaker).

factory employees. This device will play a role as a component of New Market Goods' (<http://newmarketgoods.com/>) product line (a spinoff of Urban Launchpad). This branch can be found online at [github.com/urbanlaunchpad/surveyappmaker](https://github.com/urbanlaunchpad/surveyappmaker).

## **CHAPTER 9: Conclusions**

### **Final Comments**

Mobile technology holds great promise to improve field data collection efforts. Flocktracker, as it has been developed and enhanced over the past years, has demonstrated the value of digitally enhanced surveys and mapping. From gathering the data to produce a bus map of Dhaka, Bangladesh in less than two weeks to deploying and managing team of 30 volunteers to simultaneously survey four *peseros* routes in opposite corners of Mexico City for over a combined 650 hours; Flocktracker has demonstrated itself as a capable field data-collection method, enhancing the resolution of data by adding geospatial and temporal aspects to its results.

Flocktracker's role in this thesis' research in St. Louis has enabled me to explore the relationship between surveyed riders' perception of their personal security and a variety of spatial and vehicle related factors. Through the geospatial specificity of the survey uploads, correlations between performance (speed, route distance), environmental aspects (crime, vacancy), and perception (personal security) can be observed. Such variables would previously have been prohibitively difficult to produce in such a survey. While ultimately limited by the scale of the survey operations, the St. Louis research serves as both a preliminary exploration into bus route dynamics along the 70 line and inspiration for how Flocktracker can be applied in future research.

One concern found in the conclusion of Albert Ching's thesis (Ching, 2012) was the question of who would continue data collection efforts in Dhaka, Bangladesh. With Flocktracker's complete redevelopment, a more stable and universal application is within reach. Such a technology, complete with a web-based portal through which to upload and manage projects, could significantly reduce barriers to establishing Flocktracker-based research projects around the world. With greater ease of deployment, more open approaches to warehousing the data, and simple access for others in the same field of interest, Flocktracker may empower anyone from motivated citizens to public transit organizations. The ultimate goal is to make Flocktracker valuable at a range of scales; from just a single concerned citizen interested in crosschecking technocratic decisions, to specialists looking to understand more precisely the spatiotemporal dimensions of social phenomena.

## Works Cited

- Bannister, Jon and Fyfe, Nick. (2001). "Introduction: Fear and the City." *Urban Studies* (Sage Publications, Ltd.). Volume 38, Issue 5-6, Pages 807-813.
- Florida, Jennfier A. (2014). "St. Louis Area Counties, State of Missouri, and State of Illinois." St. Louis City Recorder of Deeds and Vital Records Registrar. St. Louis, Missouri.
- Bi-State Development Agency (Metro St. Louis). (2014). "Metro Passenger Counts." Dataset, including: "Assignment By Line," "MetroBus," "MetroLink," "System." Procured April 2.
- Butts, Kuan. (2014). "State of the Application." <http://www.scribd.com/doc/206591546/State-of-the-Application>. Flocktracker (flocktracker.mit.edu). Accessed August 3.
- Cable, Dustin. (2013). "The Racial Dot Map." Demographics Research Group. Weldon Cooper Center for Public Service, University of Virginia. <http://www.coopercenter.org/demographics/Racial-Dot-Map>. Charlottesville, VA
- Ching, Albert. (2012). "A User-Flocksourced Bus Intelligence System in Dhaka." MIT Department of Urban Studies and Planning, Master of City Planning Thesis.
- Clark, J. (2003). "Fear in fear-of-crime." *Psychiatry, Psychology and Law*. Volume 10. Pages 267–282.
- COUHES. (2014). "Training: Research Involving Human Subjects." MIT.
- comScore. (2014). "MobiLens" and "Mobile Metrix" data. Key trends from US Smartphone Industry. December 2013 data.
- Cozens, P., Saville, G. and Hillier, D. (2005). "Crime prevention through environmental design." *CPTED: A review and modern bibliography*. *Journal of Property Management*. Volume 23. Pages 328–356.
- Crime Concern (1997). "Perceptions of safety from crime on public transport." *Crime Concern and Transport and Travel Research*. London.
- Crime Concern (2002). "People's perceptions of personal security and their concerns about crime on public transport: the literature review." Department of Transport. London.

- Day, K., Stump, C. and Carreon, D. (2003). "Confrontation and loss of control: Masculinity and men's fear in public space." *Journal of Environmental Psychology*. Volume 23, Pages 79-86.
- East-West Gateway Coordinating Council. (2003). "Household Travel Survey: Final Report of Survey Results." *St. Louis Region Household Travel Survey, NuStats*. St. Louis, Missouri.
- East-West Gateway Coordinating Council. (2003). "Household Travel Survey: St. Louis Final Methods Report." *St. Louis Region Household Travel Survey, NuStats*. St. Louis, Missouri.
- Ellin, N. (1997). "Shelter from the storm or form follows fear and vice versa." *Architecture of Fear*. Pages. 15 – 45. New York: Princeton Architectural Press.
- Erickson, T (2010). "Geocentric Crowdsourcing and Smarter Cities: Enabling Urban Intelligence in Cities and Regions." 1st Ubiquitous Crowdsourcing Workshop at UbiComp.
- Farrall, S., Bannister, J., Ditton, J. and Gilchrist, E. (1997). "Questioning the measurement of the 'fear of crime'; findings from a major methodological study." *The British Journal of Criminology*. Volume 37. Pages 658–675.
- Ferraro, K. F. (1995). "Fear of Crime: Interpreting Victimization Risk." New York. State University of New York Press.
- Ferraro, K. F. and LaGrange, R. (1987). "The measurement of fear of crime." *Sociological Review*. Volume 57. Pages 70–101.
- Flusty, S. (1997). "Building paranoia." *Architecture of Fear*. Pages 47–59. New York. Princeton Architectural Press.
- Formhub Blog. (2013). "One Million Surveys Visualized." Modi Labs, Columbia University. <http://blog.formhub.org/2013/06/27/1m-submissions-vizualized/>. Accessed July 26.
- Gobierno del Distrito Federal. (2014). "Acerca de CETRAM." CETRAM OM. <http://www.cetram.df.gob.mx/>. Accessed August 16.
- Gordon, Colin. (2008). "Mapping Decline: St. Louis and the Fate of the American City." *Politics and Culture in Modern America Series*. University of Pennsylvania Press.
- Griffith, Kenfield. (2014). "Changing the Discourse of Surveys: Research in Emerging Economies." mSurvey Corporate Presentation, July 10, 2014. MIT. Cambridge, Massachusetts.

- Hale, C. (1996). "Fear of crime: a review of the literature." *International Review of Victimology*. Volume 4, Pages 79–150.
- Hamilton, K. and Jenkins, L. (2000). "A gender audit for public transport: a new policy tool in the tackling of social exclusion." *Urban Studies*. Volume 37. Pages 1793–1800.
- Hanmer, J. and Stanko, E. A. (1985). "Stripping away the rhetoric of protection: violence to women, law and the state in Britain and the USA." *International Journal of the Sociology of Law*. Volume 13. Pages 357–374.
- Hibbard, Matthew. (2014). "Take A Ride Back In History on the #70 Grand Line." Next Stop, Metro Transit – St. Louis. <http://www.nextstopstl.org/12048/take-a-ride-back-in-history-on-the-70-grand-line/>. Accessed June 10.
- Johnston, L. (2001). "Crime, fear and civil policing." *Urban Studies* (Sage Publications, Ltd.). Issue 38, Pages 959–976.
- Katz, Bruce and Wagner, Julie. (2014). "The Rise of Innovation Districts: A New Geography of Innovation in America." Metropolitan Policy Program at Brookings. Washington, D.C.
- Kneebone, Elizabeth. (2009). "Job Sprawl Revisited: The Changing Geography of Metropolitan Employment." Metro Economies Series for the Metropolitan Policy Program. Brookings.
- Koskela, H. (1997). "Bold Walk and Breakings: Women's Spatial Confidence Versus Fear of Violence." *Gender, Place, and Culture*. Volume 4. Pages 301–319.
- Loewen, L. J., Steel, G. D. and Suedfeld, P. (1993). "Perceived safety from crime in the urban environment." *Journal of Environmental Psychology*. Volume 13. Pages 323–331.
- Loukaitou-Sideris, A. and Fink, C. (2009). "Addressing women's fear of victimization in transportation settings: a survey of U.S. transit agencies." *Urban Affairs Review*. Volume 44. Pages 554–587.
- Loukaitou-Sideris, Anastasia, et al. (2002). "The Geography of Transit Crime: Documentation and Evaluation of Crime Incidence on and around the Green Line Stations in Los Angeles." *Journal of Planning Education and Research*. Volume 22, Issue 2, Pages 135–150.
- Lynch, G. and Atkins, S. (1988). "The influence of personal security fears on women's travel patterns." *Transportation*. Volume 15, Issue 3. Pages 257–277.

- Mallach, Alan and Brachman, Lavea. (2010). "Ohio's Cities at a Turning Point: Finding the Way Forward." Metropolitan Policy Program. Brookings Governmental Studies. Brookings.
- Mason, Susan and Fredericksen, Elizabeth. (2009). "Fostering neighborhood viscosity: does design matter?" Community Development Journal. Volume 46, Issue 1, Pages 7–26.
- Manzi, Tony; Lucas, Karen; Jones, Tony Lloyd; and Allend, Judith. (2010). "Social Sustainability in Urban Areas: Communities, Connectivity and the Urban Fabric." Routledge.
- McCormick, Coleman. (2014). Vice President, Spatial Networks, Inc. Email Correspondence. August 9–12, 2014.
- MetroSTL. (2011). "Fiscal Year 2011 Comprehensive Annual Financial Report." [www.metrostlouis.org/Libraries/Annual\\_Financial\\_Reports/FY\\_2011\\_Comprehensive\\_Annual\\_Financial\\_Report.pdf](http://www.metrostlouis.org/Libraries/Annual_Financial_Reports/FY_2011_Comprehensive_Annual_Financial_Report.pdf). Accessed August 10, 2014.
- MetroSTL. (2014). "Articulated MetroBus Fact Sheet." Operations Committee Open Session: 4/16/14. [www.metrostlouis.org](http://www.metrostlouis.org). Accessed August 8.
- MetroSTL. (2014). "Metro Call-A-Ride." [www.metrostlouis.org/PlanYourTrip/MapsSchedules/CallARide.aspx](http://www.metrostlouis.org/PlanYourTrip/MapsSchedules/CallARide.aspx). Accessed August 9.
- MetroSTL. (2014). "Talk trains and buses with Metro." [live.stltoday.com/Event/Talk\\_trains\\_and\\_buses\\_with\\_Metro\\_3](http://live.stltoday.com/Event/Talk_trains_and_buses_with_Metro_3). Accessed August 8.
- Mueller, Courtney. (2011). "The Story of the Bi-State Development Agency." Next Stop, Metro Transit – St. Louis. <http://www.nextstopstl.org/>. Accessed August 19.
- Nasar, J. L. and Fisher, B. (1993). "'Hot spots' of fear and crime: a multi-method investigation." Journal of Environmental Psychology. Volume 13. Pages 187–206.
- Neighborhood Stabilization Team. (2014). "Neighborhoods of the City of St. Louis." The City of St. Louis. <https://www.stlouis-mo.gov/neighborhoods/>. Accessed July 29.
- Nivola, Pietro S. (1998). "Fat City." Brookings Review.
- Nsour, S. A. (1999). "Safety and security of transit services: enhancing transportation safety in the 21st century." ITE International Conference Proceedings. Institute of Transportation Engineers. Washington, DC.

- Paddison, Ronan. (2001). "Handbook of Urban Studies." SAGE Publications Limited. London, UK.
- Pain, R. H. (1993). "Women's fear of sexual violence: explaining the spatial paradox." *Crime and the Urban Environment*. (H. Jones) Pages 46–55.
- Pain, R. H. (1995). "Elderly women and violent crime: the least likely victims?" *British Journal of Criminology*. Volume 35. Pages 584–598.
- Pain, R. H. (1997). "Social geographies of women's fear of crime." *Transactions of the Institute of British Geographers*. Volume 22. Pages 231–244.
- Schulz, Dorothy and Gilbert, Susan. (1996). "Women and Transit Security: A New Look at an Old Issue." Office of Highway Policy Administration. Federal Highway Administration, US Department of Transportation. Proceedings from the Women's Travel Issues Second National Conference. Baltimore, MD. Chapter 30, Pages 551-562.
- St. Louis, City of. (2014). "Census Results." <http://dynamic.stlouis-mo.gov/census/corridor.cfm>. Accessed August 12.
- St. Louis, City of. (2014). "History of St. Louis." Planning and Urban Design Agency, Research Division <https://stlouis-mo.gov/visit-play/stlouis-history.cfm>. Accessed July 14.
- Sweet, Elizabeth L. and Escalante, Sara Ortiz. (2014). "Bringing bodies into planning: Visceral methods, fear and gender violence." *Urban Studies* (Sage Publications, Ltd.). Online publishing, July 11.
- Transit Cooperative Research Program. (1997). "TCRP Synthesis 21 Improving Transit Security: A Synthesis of Transit Practice." Transportation Research Board. National Research Council. Washington, D.C.
- Transit Cooperative Research Program. (2003). "Transit capacity and quality of service manual." 2nd Edition. Transportation Research Board of the National Academies. Washington, D.C.
- Trench, S.; Oc, T.; and Tiesdell, S. (1992). "Safer cities for women: perceived risk and planning measures." *Town Planning Review*. Volume 63. Pages 279–293.
- US DOT. (2010). "2009 National Household Travel Survey: Summary of Travel Trends." Federal Highway Administration.
- US Census. (2009). "Commuting in the United States: 2009." <http://www.census.gov/prod/2011pubs/acs-15.pdf>. American Community Survey Reports. Accessed July 28.

- US Census. (2011). "Population Estimates: City and Town Totals: 2011."  
<http://www.census.gov/popest/data/cities/totals/2011/>. Accessed August 25.
- Windmiller, Henessy, Watkins. (2014). "Accessibility of Communication Technology and the Rider Experience: A Case Study of St. Louis Metro." Internally distributed copy. Also: TRB 93rd Annual Meeting Compendium of Papers. Washington, D.C.
- Yavuz, Nilay and Welch, Eric W. (2010). "Addressing Fear of Crime in Public Space: Gender Differences in Reaction to Safety Measures in Train Transit." *Urban Studies* (Sage Publications, Ltd.). Volume 47, Issue 12, Pages 2491–2515.
- Zegras, P. Christopher, Butts, Kuan, Cadena, Arturo, Palencia, Daniel. (2014, forthcoming). "Spatiotemporal Dynamics in Public Transport Personal Security Perceptions: Digital Evidence from Mexico City's Periphery." *ICT for Transport: Opportunities and Threats* (P. Rietveld, M. Givoni, N. Thomopoulos, Eds.). Edward Elgar Publishing.

## **APPENDIX 1: Survey Questions**

The survey was short; I designed it to be performed in under three minutes.

1. What is your gender?
  - a. Male
  - b. Female
  - c. Other
2. What is your age?
  - a. Numerical input
3. Please state your race.
  - a. Black/African American, White, Hispanic, Asian, Native American, Other
4. What was the origin of this trip you are currently on? Where did you start your trip today?
  - a. Complete neighborhood list of St. Louis City, plus County and East St. Louis.
5. What will be the final destination of this trip you are currently on? Where are you heading?
  - a. Complete neighborhood list of St. Louis City, plus County and East St. Louis.
6. Does trip contain another transfer?
  - a. Yes or no
7. For this trip, what was your initial mode? How did you start your trip, did you...
  - a. Walk, Bike, MetroBus, MetroLink, Shuttle (Call-A-Ride), Car, Other
8. What will be your final mode for this trip?
  - a. Walk, Bike, MetroBus, MetroLink, Shuttle (Call-A-Ride), Car, Other
9. What is the purpose of your trip today?
  - a. Work, School, Caring for Others, Social, Shopping, Dining, Other
10. What do you feel is most secure mode (type) of transport in St. Louis, if you could select any?
  - a. Walk, Bike, MetroBus, MetroLink, Shuttle (Call-A-Ride), Car
11. What do you feel is the most important consideration when making travel plans?
  - a. Timeliness, Convenience, Cost, Comfort, Personal Security
12. Please state your perception of personal security at this time. That is, how safe do you feel at this moment in time, in this area?
  - a. Value from 1 to 5
13. Please state your perception of personal security at your origin. That is, how safe did you feel at the point where you started this trip?
  - a. Value from 1 to 5
14. Please state your perception of personal security at your destination. That is, how safe do you think you will feel at your point of destination, where you are headed?
  - a. Value from 1 to 5