

MODELING ACCESSIBILITY THROUGH GEOCROWDSOURCING

by

Han Qin
A Dissertation
Submitted to the
Graduate Faculty
of
George Mason University
in Partial Fulfillment of
The Requirements for the Degree
of
Doctor of Philosophy
Earth Systems and Geoinformation Sciences

Committee:

_____	Dr. Matthew T. Rice, Dissertation Director
_____	Dr. Chaowei Yang, Committee Member
_____	Dr. Dieter Pfoser, Committee Member
_____	Dr. Jyh-Ming Lien, Committee Member
_____	Dr. Anthony Stefanidis, Department Chairperson
_____	Dr. Donna M. Fox, Associate Dean, Office of Student Affairs & Special Programs, College of Science
_____	Dr. Peggy Agouris, Dean, College of Science
Date: _____	Spring Semester 2017 George Mason University Fairfax, VA

Modeling Accessibility through Geocrowdsourcing

A Dissertation submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy at George Mason University

by

Han Qin
Master of Science
Eastern Michigan University, 2012
Bachelor of Engineering
Wuhan University, 2009

Director: Matthew T. Rice, Associate Professor
Department of Geography and Geoinformation Science

Spring Semester 2017
George Mason University
Fairfax, VA



This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/).

DEDICATION

This dissertation is dedicated to my wife Manzhu Yu

And my parents – Xinxiang Qin, Weihong Han

ACKNOWLEDGEMENTS

First of all, I would like to express the deepest appreciation to my committee chair – Dr. Matt Rice for his endless support. Without his guidance and persistent help, I cannot finish my Ph.D. program and this dissertation in four years. Dr. Rice treats me as a family member and a close friend, and shares his research and teaching experiences with me as much as he can. He also helps my wife a lot on her teaching experience. I would also like to thank my committee members, Dr. Chaowei Yang, Dr. Dieter Pfoser, and Dr. Jyh-Ming Lien, who provide a lot of useful advices on my research works and presentations. In addition, I would like to thank Dr. Kevin Curtin, who helps me a lot on my publications and this dissertation. Finally, I would like to thank my family and friends at GGS department and Spatiotemporal Center. Thank you very much for all your support in the past four years at George Mason University.

TABLE OF CONTENTS

	Page
List of Tables	viii
List of Figures	ix
List of Equations	xi
List of Abbreviations	xii
Abstract	1
Chapter One: Introduction	3
The Emergence of Geocrowdsourcing	7
Introduction to the QA-CGD Research Project	10
Existing Data Quality Approaches	12
Chapter Two: Conceptual Framework	15
Geospatial Data Sources	16
Elements of Geospatial Data Quality Assessment	18
Quality Assessment of Geocrowdsourced Data	20
Scientific Problem Statement	21
Summary of Research Themes	22
Establishment of a Geocrowdsourcing Environment for Accessibility	22
Improvement of Geocrowdsourcing Testbed and Data Quality Assessment Framework with Public Participation	24
Assessment of the Impact and Use of Geocrowdsourced Data in a Pedestrian Network	25
Chapter Three: Geocrowdsourcing and Accessibility for Dynamic Environments	27
Abstract	27
Introduction	28
Geographical Setting and Context	29
Literature Review	33
Accessibility Mapping and Applications	33
Geocrowdsourcing and Volunteered Geographic Information	35

Quality Assessment of Data.....	36
Materials and Methods	37
Permanent and Transient Navigation Obstacles	38
The GMU Geocrowdsourcing Testbed	38
Pedestrian Network	41
Quality Assessment of Data.....	44
Results	45
Positional Accuracy of Obstacles	45
Pedestrian Network Coverage.....	47
Obstacles and their Impact on the Accessible Pedestrian Network	49
Routing and Accessibility Dynamics.....	51
Discussion	56
Summary	58
Future Work	59
Chapter Four: Obstacle Characterization in a Geocrowdsourced Accessibility System ..	61
Abstract	61
Introduction	61
The GMU Geocrowdsourcing Testbed	64
Moderation and Quality Assessment in the GMU Geocrowdsourcing Testbed...	65
Training and Obstacle Categorization in the GMU Geocrowdsourcing Testbed .	66
User Motivation in the GMU Geocrowdsourcing Testbed.....	73
Position Accuracy	74
Conclusion and Future Work	75
Chapter Five: Assessing the Importance of Pedestrian Network Segments in a Geocrowdsourced Accessibility System.....	77
Abstract	77
Introduction	78
Literature Review and Relevant Previous Work.....	81
Accessibility Systems	81
Formal Modeling of Pedestrian Networks	82
Geotechnology and Geocrowdsourcing.....	84
The George Mason University Geocrowdsourcing Testbed.....	85
Operations Research and Optimizing Maintenance.....	86

Statement of Purpose	87
Data and Methodology	87
Family of Models for Pedestrian Network Optimization	89
Model Notation and Formulations	89
Full Dataset from GMU-GcT.....	99
Results	102
Comparisons Among the Family of Models.....	109
Discussion	111
Summary	112
Chapter Six: Quality Assessment in Geosocial Media	114
“Visualizing Spatiotemporal Trajectories of Mobile Social Media Users Using Space– Time Cube”	116
A General Framework for Quality Analysis on Social Media Network	120
Chapter Seven: Conclusions and Future Research	125
Summary and Concluding Remarks.....	125
Future Research.....	131
References	134

LIST OF TABLES

Table	Page
Table 1 Quality Assessment Variables used in the GMU Geocrowdsourcing Testbed ...	44
Table 2 Results of different models	109

LIST OF FIGURES

Figure	Page
Figure 1 Hunter and Beard 1992, Classification of Error in GIS, from “Understanding Error in Spatial Databases”	14
Figure 2 A general workflow includes capture, store, filter, assess, analyze, and visualize geocrowdsourced data.....	16
Figure 3 Elements of geospatial data quality	20
Figure 4 Pedestrians crossing eight lanes of traffic in Tysons Corner, Virginia. (Photo by Ken Archer, (2014) used with permission).....	30
Figure 5 Temporary construction barricades, fencing, and debris blocking a pedestrian pathway	32
Figure 6 The GMU Geocrowdsourcing Testbed, geo.gmu.edu/vgi (accessed November 22, 2014)	40
Figure 7 GMU Geocrowdsourcing Testbed, Locator Icon	41
Figure 8 Google Maps pedestrian routing on the GMU Campus (September 2014)	42
Figure 9 A view of the pedestrian network (yellow) for the local area	43
Figure 10 Positional Accuracy for GMU Testbed Obstacles.....	47
Figure 11 Roads (in black) without adjacent sidewalks (at left, with all roads, and isolated, at right)	48
Figure 12 Pedestrian-inaccessible regions masked with polygons	48
Figure 13 Location of Stairs and Steep pathways (buffered for visibility), with pedestrian network (left) and isolated (right)	49
Figure 14 Location of Obstacles from the GMU Geocrowdsourcing Testbed.....	50
Figure 15 Location of pedestrian network segments impacted by obstacles, with pedestrian network (left) and isolated (right).....	51
Figure 16 Location of pedestrian network segments impacted by Stairs and Steep Paths (purple), and Obstacles (red), with pedestrian network (left) and isolated (right)	51
Figure 17 Routing Scenario 1: modest route length increase	54
Figure 18 Routing Scenario 2: substantial route length increase.....	55
Figure 19 Routing Scenario 3	56
Figure 20 Individual route preference (green) and the shortest cost path (blue), from Rice et al. 2014.....	58
Figure 21 Transient obstacles in the pedestrian corridors, Fairfax, Virginia.....	62
Figure 22 Training graphic with picture and obstacle type, obstacle duration, and obstacle priority characterization	67
Figure 23 GMU geocrowdsourcing training picture showing crowd/event	68

Figure 24 High agreement (95%) for categorization of Figure 23 as a crowd/event obstacle by 37 training participants	68
Figure 25 GMU geocrowdsourcing training picture showing sidewalk obstruction.....	69
Figure 26 Significant disagreement for categorization of Figure 25 by 37 training participants	69
Figure 27 GMU geocrowdsourcing training picture showing sidewalk obstruction.....	70
Figure 28 Significant disagreement for categorization of Figure 27 by 46 training participants	70
Figure 29 GMU geocrowdsourcing training picture.....	71
Figure 30 Moderate to low disagreement for categorization of Figure 29 by 46 training participants	71
Figure 31 Distribution of obstacle types in GMU-GcT, January 2014.....	72
Figure 32 Positional characteristics of an obstacle report	75
Figure 33 Transient obstacle impacting pedestrians, captured by the George Mason University Geocrowdsourcing Testbed (GMU-GcT)	80
Figure 34 Repairs of contiguous segments of sidewalk curbs, with traffic alteration	89
Figure 35 The segments in need of repair.....	92
Figure 36 The optimal subset of repairs for this problem instance.....	93
Figure 37 Concept of the neighborhood set: N_{ij}	95
Figure 38 Small dataset solution with contiguity enforced	96
Figure 39 Small dataset solution with contiguity encouraged	99
Figure 40 Obstacle Report with Details, from the GMU-GcT	101
Figure 41 Pedestrian network with the set of segments in need of repair	103
Figure 42 Result of Model 1	104
Figure 43 Result of model 2.....	106
Figure 44 Result of Model 3	108
Figure 45 Equivalent segments in these three models	110
Figure 46 The space–time trajectory of a Twitter user	118
Figure 47 The space–time trajectory of another Twitter user	119
Figure 48 Framework for Quality Analysis of Big Data on Social Media (Hajjar et al., 2015)	121
Figure 49 Point, polyline and polygon-based footprints for high volume users with user-ids.....	123
Figure 50 Unusual footprint: Gridded footprint.....	124

LIST OF EQUATIONS

Equation	Page
Equation 1 Objective Function Maximize Benefit	90
Equation 2 Constraint: overall budget for repair	90
Equation 3 Constraint: decision variable	90
Equation 4 Enforced contiguity constraint	95
Equation 5 Constraint: determine the values of a_{ij}	97
Equation 6 Modified objective function	97
Equation 7 Objective function for Model 3	98
Equation 8 Constraint: overall budget for repair	98
Equation 9 Constraint: decision variable	98
Equation 10 Encouraged contiguity constraint	98
Equation 11 Constraint: decision variable	99

LIST OF ABBREVIATIONS

Ambient Geospatial Information	AGI
American Disability Act	ADA
Crowdsourced Geospatial Data.....	CGD
École polytechnique fédérale de Lausanne.....	EPFL
geographic information systems	GIS
George Mason University Geocrowdsourcing Testbed.....	GMU-GcT
Human Dynamics in the Mobile Age	HDMA
International Cartographic Association's	ICA's
Kellar Institute for Human disAbilities.....	KIHD
National Center for Geographic Information and Analysis.....	NCGIA
National Map Accuracy Standard.....	NMAS
National Standard for Spatial Data Accuracy	NSSDA
OpenStreetMap	OSM
Principal Investigator	PI
Public Participation and Geographic Information Systems	PPGIS
Quality Assessment and Accessibility Applications of the Crowdsourced Geospatial Data	QA-CGD
Research and Development.....	R&D
scalable vector graphics	SVG
Visualizing Information Space in Ontological Networks.....	VISION
Volunteered Geographic Information	VGI
Volunteered Geographic Services.....	VGS

ABSTRACT

MODELING ACCESSIBILITY THROUGH GEOCROWDSOURCING

Han Qin, Ph.D.

George Mason University, 2017

Dissertation Director: Dr. Matthew T. Rice

Map-based crowdsourcing is one of the most significant contemporary trends in the geospatial sciences and has completely changed many data collection workflows, and added new sources of data. An important aspect of this emerging trend is the manner in which data quality is assessed, and how well these quality assessment processes match processes used in traditional map-based and geographic information systems-based quality assessment procedures. This dissertation studies the evolution of geographic data collection, and the methods of quality assessment, and builds a comprehensive quality assessment workflow for geocrowdsourced data. This workflow is based on many traditional formulations of quality, such as positional accuracy, temporal consistency, categorical accuracy, fitness-for-use, and lineage. These quality assessment workflows are studied through the George Mason University Geocrowdsourcing Testbed (GMU-GcT), which was designed to study dynamic aspects of map-based crowdsourcing. The GMU-GcT tests the implementation of techniques from the US National Map Accuracy

Standard (NMAAS) as well as the National Standard for Spatial Data Accuracy (NSSDA), as well as several new techniques, modified over time, that are shown to have value within the specific context of geocrowdsourcing conducted with the GMU-GcT. This research extends the quality assessment work with modeling of a pedestrian network and the accessibility characteristics associated with navigation obstacles, many of which have been crowdsourced with the GMU-GcT, and tests the feasibility of infrastructure maintenance using geocrowdsourced data and associated quality assessment parameters. The quality assessment techniques from traditional mapping domains are shown to have value in the domain of geocrowdsourcing, and the ability to model pedestrian network accessibility and maintenance optimization is demonstrated through this work. Extensions of this research into geosocial media is explored with mixed results, and future work in simplified, image-based geocrowdsourcing is explored to determine what quality assessment metrics can be derived from greatly simplified geocrowdsourcing methods. Additional modeling enhancements, based on alternative optimization strategies and weighting factors, is discussed as a future area for work. Summary of end-user and subject matter experts is discussed in context of future modifications to the GMU-GcT.

CHAPTER ONE: INTRODUCTION

There is an old Chinese saying “Three cobblers with their wits combined equal Zhuge Liang, the mastermind”, which means “Two heads are better than one”. This ancient view of crowdsourcing echoes the writing of Howe (Howe 2008), who discusses the power of the crowd, and coined the term crowdsourcing, which is described as a process where a company’s research and development (R&D) problems are solved by people not employed by that company. A decade prior in 1996, The National Center for Geographic Information and Analysis (NCGIA) designated the concept of Public Participation and Geographic Information Systems (PPGIS) as a critical research priority for GIS, and included it in their research objectives¹. With the Internet and geographical information technologies present during this time period, web-based PPGIS projects were used to create spatial data and make decision support from public contributions, which is closely related to the contemporary concept of crowdsourcing. Ten years later, with the advent Web 2.0 and related technologies, Goodchild (2007, 2009) coined the term volunteered geographic information (VGI) to indicate the special Web phenomenon of georeferenced user-generated content. Goodchild explains that VGI is created by motivated, yet untrained members of the public and can be an important alternative to traditional authoritative information from mapping agencies and corporations, a concept refined from an earlier publication emphasizing distributed online information sharing communities (2005). Meanwhile, Stefanidis et al. (2011) terms the social network

¹ <http://www.ncgia.ucsb.edu/research/initiatives.html>

information which contains locations or references to geographic entities in social media messages to be ambient geospatial information (AGI), and highlights that AGI focuses on passively contributed data. Later, Rice et al. (2012) also follows on VGI and emphasizes a form of geocrowdsourcing – crowdsourced geospatial data (CGD), which involves the participation of end-users. Many of them are untrained in the geospatial sciences but have a strong interest in geospatial technology. These representative concepts are also found in many popular web-based systems or social network services, such as OpenStreetMap, Wikimapia, Twitter, Facebook or Waze, which produce heterogeneous, complex, and dynamic geocrowdsourcing datasets. The power of the crowd has long been acknowledged in geospatial domain and geocrowdsourcing is the form of this power. With the conceptual theories and methodologies developed by the above-mentioned scholars, the concept of geocrowdsourcing and its applicability have been gradually refined and enhanced to provide more benefits to research and development.

Among the many concerns raised by Goodchild, Stefanidis, and other authors, data quality is generally one of the most important. The reliability and trustworthiness of a data source, especially, a publicly contributed data source, has been a common theme through many of the academic papers and discussions of the emerging phenomena of VGI and geocrowdsourcing, including some of the most highly cited works, such as Haklay (2010) and Girres and Touya (2010). As an important contributor to research on critical GIS and public participation, Sieber (2006) emphasizes the appropriateness of information, and enumerates several critical data issues - wrong format, incorrect resolution, completeness, and users' needs in PPGIS projects. She concludes that an

important challenge for PPGIS is accuracy - “the challenge in PPGIS is to understand the importance of accuracy and illuminate the assumptions underlying quantitative analysis” (Sieber 2006, p. 498). Furthermore, in his 2010 paper, Goodchild discusses the well-known crowdsourcing applications Wikipedia² and Wikimapia³. He asserts that the success of Wikipedia is that entries are reviewed by a hierarchy of volunteers who employ well-defined criteria that are appropriate to crowdsourcing, while Wikimapia attracts erroneous and sometimes malicious content making it difficult for crowdsourcing mechanisms to work to correct errors. Wikipedia has formal criteria on information validation, and in terms of the verifiability, Wikipedia requires contributors provide reliable sources and includes the citations to support their material directly. However, in Wikimapia system, users create features and categories in free format when they post geoinformation. The decision to preserve or delete the submitted category information, which is a key component to search relevant features, is based on category moderators’ subjective criteria, although those moderators could have been trained or have significant, relevant experience. In spite of the intervention of moderators to check some aspects of quality, designing a comprehensive quality assessment is still required to assess and improve the overall quality of information produced in a geocrowdsourcing project.

Assessing and managing the quality of geocrowdsourcing is a critical problem in our current crowdsourcing environments. As mentioned, Wikipedia has an extensive system for quality checking entries with a hierarchal system of editors and automated

² <https://www.wikipedia.org/>

³ <http://wikimapia.org/>

computer programs (bots) that search for inconsistencies and unusual editing patterns Rice et al. (2012). On the geospatial side, OpenStreetMap has developed a similar system of human and automated techniques for identifying errors and implementing quality assurance processes, discussed by Rice et al. (ibid.) and Paez (2014). In order to address the need for quality assessment in a moderated crowdsourcing system where positional, thematic, and temporal accuracy issues are critical, Goodchild and Li (2012) suggest that project researchers can field check and moderate crowdsourced data, which is a general quality assessment approach for geocrowdsourcing. This “social approach”, discussed by Goodchild and Li, and extended in Rice (2015a, 2015b) and Rice et al. (2016), consists of a “moderator consistency study framework” to evaluate the consistency of moderator quality assessments.

The research of this dissertation will focus on the “Quality Assessment and Accessibility Applications of the Crowdsourced Geospatial Data” (QA-CGD) project, and related work based on this project, such as a study of data quality with geosocial media data. This dissertation reviews the transitions of geospatial data formats, the design of a comprehensive data quality assessment method, and the representation of the values and applications of geocrowdsourcing in modern dynamic era. This dissertation is composed of three primary chapters (3, 4, and 5), which address fundamental concepts in the generation of a comprehensive quality assessment framework instituted in the GMU-GcT and applications facilitated by quality assessment in the GMU-GcT.

The present chapter introduces the concept and evolution of geocrowdsourcing, the context of QA-CGD project, and relevant work on methods for geospatial data

quality. Chapter two summarizes the fundamental concepts of this dissertation – the evolution of general geospatial data sources and elements of traditional geospatial data quality assessment. Chapter two also states the scientific research purposes of this dissertation, and reviews the three principle research chapters. Chapter three introduces the quality assessment framework instituted in the GMU-GcT and contains a preliminary accessibility analysis; chapter four addresses categorization, quality assessment revisions, and cognitive issues in the GMU-GcT; and chapter five addresses the use of the GMU-GcT and expert input to generate optimization models for pedestrian network maintenance. Chapter six discusses quality assessment in geosocial media, including strategies for archiving, processing, and assessing geosocial media as a source of data for GIS. Chapter seven concludes with a synthesis of the ideas in presented in this dissertation, and discusses future research in different domains and future development in different application areas.

The Emergence of Geocrowdsourcing

Before the late 19th century, science was primarily conducted by people who had time and additional resources for data acquisition and analysis (Bryson 2011, Goodchild 2009, Haklay 2013). However, citizen science has changed from only being available to a privileged few in the past to engaging wider part of society today (Silvertown 2009). Amateurs, usually nonscientists, who voluntarily participate in scientific research activities are called citizen scientists (ibid.). These individuals can offer significant contributions to define scientific problems, evaluate scientific arguments, provide people

power or resources, and/or collect, process, analyze, and/or disseminate data (Clark and Illman 2001, Cohn 2008, Gura 2013, Hand 2010, Silvertown 2009).

Compared with traditional scientific activities, citizen scientific activities provide a way to conduct research in areas that may have insufficient human and financial resources, and bridges the gap between citizens and scientists by offering partnership and collaboration through well-defined projects (Clark and Illman 2001). Goodchild (2007) outlines the emerging phenomenon of volunteered geographic information (VGI), tying the phenomena to developments in the Internet and World Wide Web. He indicates that humans can act as sensors within a Web 2.0 environment to create and publish VGI. The act of self-promotion is stated as a one of the clear motivations for contributing volunteered geographic information.

In addition, the value of VGI has been underscored by a few prominent projects associated with emergency response and geosocial media. Zook et al. (2010) focus on four web-based mapping services: CrisisCamp Haiti, OpenStreetMap, Ushahidi, and GeoCommons at the recovery of post-earthquake in Haiti, and demonstrate that geocrowdsourcing can play a key role in the logistics of disaster response. The Center of Human Dynamics in the Mobile Age (HDMA) at San Diego State University has developed a geo-targeted event observation viewer system (SMART Dashboard) to harvest geo-tagged tweets and visualize different issues from different domains by purposes (Tsou et al. 2015). For instance, the HDMA system is applied to wildfire emergency responses⁴ by recruiting over 1000 local volunteers in San Diego to contribute

⁴ <http://vision.sdsu.edu/hdma/wildfire/>

the real-time wildfire information, which includes textual information, location, images and more, through Twitter platform. Qin et al. (2015a, 2015b), Rice et al. (2015, 2014, 2013a, 2013b, 2012) and Rice et al. (2016) address the values and challenges of crowdsourced geospatial data (CGD) to provide transient obstacle information to blind, visually-impaired, and mobility-impaired individuals through crowdsourcing. Therefore, geocrowdsourcing is important in time-sensitive domains and where needs are not being met, such as the domain of accessibility mapping. Also, web-based applications and social media platforms are an important component of the geocrowdsourcing phenomena, and represent the expansion of modes of geospatial practice from tradition mapping to GIS to web-based methods, and more recently, to georeferenced social media (Croitoru et al. 2013).

Several researchers and practitioners have integrated traditional geographical information science (GIScience) practice with information technology. An example is Tsou et al. (2013), who integrates GIScience and web-search engines to track and analyze public web pages and their web contents with associated spatial relationships. Although volunteers can use social media platforms or self-developed applications to contribute information with special purposes and standards, Stefanidis et al. (2011) outline characteristics of Ambient Geographical Information (AGI), which is passively contributed geographical information from social media feeds and not required to specific topics, and emphasize that AGI can focuses on the human activities, such as event detection, emergency responses and disaster recovery.

Introduction to the QA-CGD Research Project

“Quality Assessment and Accessibility Applications of Crowdsourced Geospatial Data” is a four-year project (2011-2015) funded by US Army Corps of Engineers – Engineer Research and Development Center, with Dr. Matt Rice, as the Principal Investigator (PI), several GMU faculty members (Pfooser, Curtin, Fuhrmann) as Co-PIs, and the author (Han Qin) as a multi-year research assistant and project developer. This projects studied the technological and social elements of emerging phenomena of geocrowdsourcing, including the exploration, development, and assessment methods for geospatial quality assessment. The four phases in this project, conducted as a rate of one phase per year, are:

- **Phase 1:** Devise standard methods for identifying sources of user-contributed geographic information;
- **Phase 2:** Develop methods for comprehensive quality assessment of user-contributed geographic Information;
- **Phase 3:** Develop ways of combining volunteered geographic information with institutional data and geovisualizing the combined data, including quality measures and assessments;
- **Phase 4:** Develop strategies for incorporating user-contributed data in complex spatiotemporal environments.

While the funding agency for this project was primarily interested in technology to be used by the military, our application domain for our university research is the realm of assistive geotechnology, where web mapping technology and geographic information

systems are used to provide access to enabling information and services. This application domain contained similar dynamics as the domains of interest for the sponsor, especially the issue of rapidly changing environments and their representation in mapping systems. As is well-known, many public health service locations, public amenities, and other public resources cannot be accessed dependably without the use of crowdsourcing and other techniques for capturing dynamic, changing environmental conditions that impact the blind, visually-impaired, and mobility-impaired. Thus, web-based tools have been developed to address the progression of user-generated and crowdsourced content from static data contributions to dynamic place-based services, to provide transient obstacle information and to provide access for mobility impaired individuals through crowdsourcing. At the same time, the comprehensive data quality assessment approach developed in this research project is a key to satisfying end-users who rely on geocrowdsourced information.

The contributions of this dissertation for QA-CGD project include development of fully-functional web-based geocrowdsourcing testbed and the implementation of a unique geospatial data quality assessment workflow, based on the needs and requirements of this project. All the data contribution tools, web-based portals for quality assessment, and visualization environments were developed for the purpose of exploring the dynamics of geocrowdsourcing, consistent with the research sponsor's needs and the proposed goals of this dissertation.

Existing Data Quality Approaches

The National Map Accuracy Standard (NMAS) was developed in 1940s and published in 1947 by the US Geological Survey⁵. As reviewed in (Rice et al., 2012b), the NMAS state that for published maps, 90% of well-defined features sampled should have horizontal positional errors of less than 1/30th of an inch at publication scale (for maps published at 1: 20,000 or larger, the figure is reduced to 1/50th of an inch), and vertical errors of less than ½ the elevation contour interval for the same 90% sample. For example, according to NMAS, the standard US Geological Survey 7.5' topographic quadrangle map published at 1: 24,000, 90% of the well-defined features sampled for positional accuracy should be within 40 feet, or approximately 12 meters. The National Standard for Spatial Data Accuracy (NSSDA) was published in 1990s as an attempt to update the NMAS and develop an assessment methodology that reflected the nature of GIS data, and specifically the interaction between data specification and scale in digital data. Quality assessment methods for geocrowdsourcing have been modeled after NMAS and NSSDA; however, quality assessment of geocrowdsourcing is still a major concern due to the “open, lightly controlled process with few constraints, specifications, or quality assurance processes” (Rice et al. 2012b). To explore these quality assurance processes with geocrowdsourced data, Haklay (2010) and Girres et al. (2010) apply several traditional quality assessment metrics to OpenStreetMap: positional accuracy, attribute accuracy, logical consistency, completeness, semantic accuracy, and temporal quality. Goodchild and Li (2012) put forward three general quality assessment approaches

⁵ <http://nationalmap.gov/standards/nmas.html>

(crowdsourcing, social, and geographic) for geocrowdsourced data with different conditions. Rice et al. (2012b, 2013, 2014, 2015a, 2015b) and Qin et al. (2015b) discuss and implement a comprehensive quality assessment method to enhance the quality of CGD, based on the social moderation approach discussed in Goodchild and Li (2012).

Furthermore, Stefanidis et al. (2011) distinguish between AGI and CGD/VGI and emphasize that AGI, like passively contributed social media data and social media feeds, can be turned into knowledge about current events, geopolitics, political movements, and the relationship between groups of end-users. On the contrary, CGD/VGI is usually used to address a specific topic or purpose and applications are generally designed with a purpose in mind. While several studies address quality assessment in VGI and CGD and use methodology from NMAS, NSSDA and concepts discussed by Goodchild and Li (2012), very few studies have looked at the same quality measures and techniques applied to georeferenced social media data. A summary of efforts to implement quality assessment with social media data is contained in chapter six of this dissertation, with a review of other preliminary efforts in this domain, such as Han and Tsou (2016), who used Twitter as a source for detailed population estimates with a quality assessment and data validation process using LandScan and Census data.

In addition to traditional quality assessment methods, Foody et al. (2013) highlights the relationship between level of volunteer expertise, and its impact on the data quality. Dr. Alexis Comber, a colleague of Dr. Giles Foody, has worked extensively in geocrowdsourcing of land use and land cover data from remotely sensed imagery. He (Comber 2013) emphasizes that the quality of VGI can be controlled by the reliability of

VGI contributors, users, and producers. The contribution of individual contributors to quality problems is also highlighted in Hunter and Beard (1992)'s "errors" pyramid of GIS (Figure 1 Hunter and Beard 1992, Classification of Error in GIS, from "Understanding Error in Spatial Databases"). However, the original producers cannot foresee all potential uses and it has become crucial for a data producer to record important aspects about the data so that users may make informed judgments regarding fitness for use (Chrisman 2006). Thus, the above ideas have been considered in the implementation of comprehensive quality assessment procedures and metrics for geocrowdsourced data, presented in this dissertation.

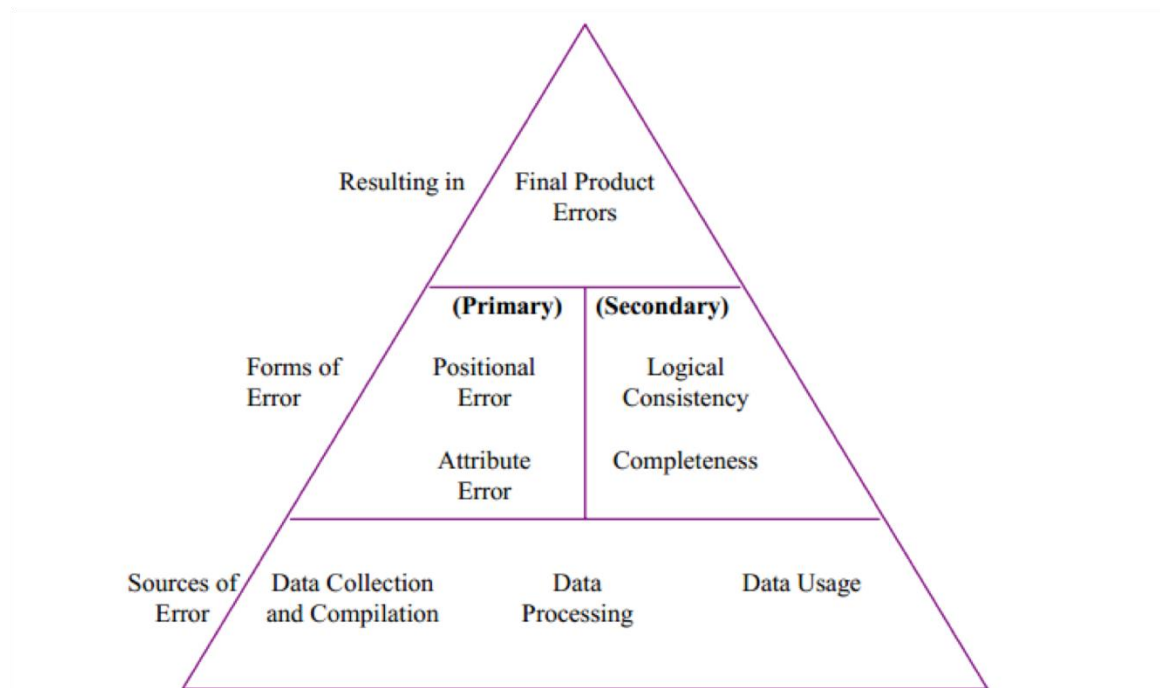


Figure 1 Hunter and Beard 1992, Classification of Error in GIS, from "Understanding Error in Spatial Databases"

CHAPTER TWO: CONCEPTUAL FRAMEWORK

During the past 4 decades, spatial-analytical tools have undergone a number of significant technical revolutions, documented by Monmonier (1985), Brunnett (2004), and others, who recognize the tremendous impact that computer systems, ontological networks, and the Internet have had on cartography, geographic information systems, spatial analysis, and the geospatial disciplines. Significant new influences in the geospatial domain include the participating web (Web 2.0), the semantic web (Web 3.0), the mobile web (Web 4.0), social media, and user-centered technologies, as well as the generation and use of very large, dynamic datasets. These areas of influence present many new opportunities and challenges, and this research focuses mainly on geocrowdsourcing, and the development of methods for quality assessment and visualization of geocrowdsourced data, building on the work of Goodchild (2007), Haklay (2010), Girres et al. (2010), Stefanidis et al. (2011), Rice et al. (2012a, 2012b, 2013a, 2013b, 2014) and others, who have helped define a new landscape of geocrowdsourced information and associated quality assessment strategies and techniques to make this information useful. Then a general workflow based on the idea of this research is illustrated in Figure 2 A general workflow

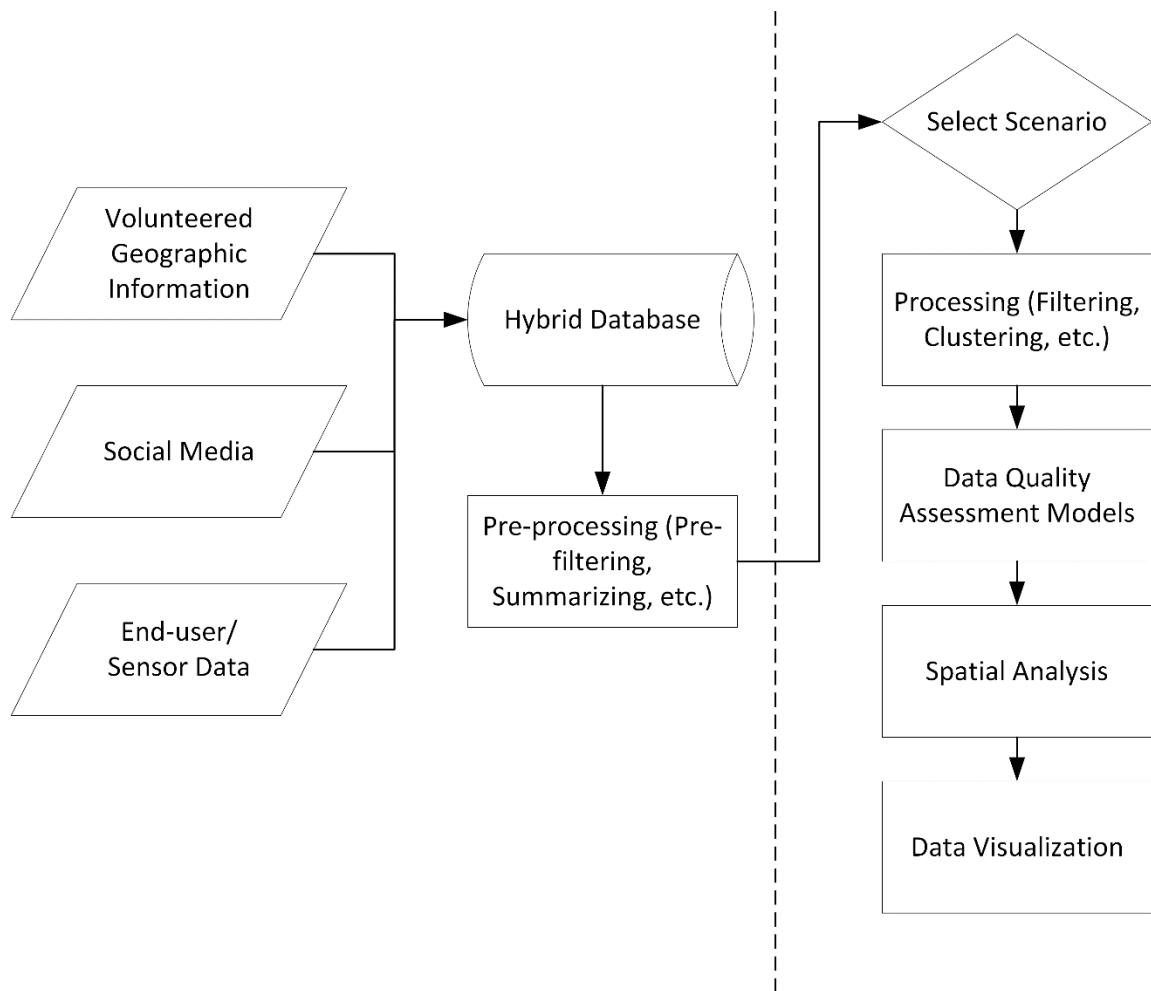


Figure 2 A general workflow includes capture, store, filter, assess, analyze, and visualize geocrowdsourced data

Geospatial Data Sources

Geospatial data, in its earliest form, came from paper maps or similar analog sources (Thrower 2008). Traditional data production was dominated by governments and military, who had the technical resources and expertise to initiate and manage large mapping projects. The US Census Bureau and the US Geological Survey are two major sources of traditional geospatial data during the 1950s, 1960s, 1970s, 1980s, and 1990s. In GIS, point, line, and area data from traditional sources has been augmented with

digital records, aerial photographs, and sensor data. Examples of new sensor-based data sources for GIS include embedded roadway sensors (Shekhar et.al 1997) and mobile sensor platforms, such as Valarm⁶. Text archives, including the content of newspapers and magazines such as National Geographic, are additional sources of data and can be georeferenced and used with geoparsing and data fusion techniques (Carroll 2006). Now, the largest source of new data in GIS is undoubtedly web-based user-contributed volunteered geographic information (VGI) and crowdsourced geospatial data (CGD), along with streaming social media data from sources such as Twitter. Goodchild (2007, 2009), Zook et al. (2010), and Croitoru et al. (2013) discussed these new sources of data and how they might be used to augment traditional mapping systems. Sui et al. (2013) describe the resulting new geoinformation landscape as a profound transformation in how geographic data, information and knowledge have been produced (p.1).

Data collection and storage strategies are critical for incorporating geocrowdsourcing information into geospatial workflows, due to the use of many different data sources, heterogeneous data formats and disparate data resolutions/scales. This raises a discussion about the need for universal standard for collecting and storing geocrowdsourcing. Tsou et al. (2013) developed the Visualizing Information Space in Ontological Networks (VISION) framework to analyze the contents of web pages and social media resources such as Twitter. Tsou et al. use the social media search engine (twitter API) and commercial web search engines (Yahoo and Bing API) to look for information that is geographically relevant. Using search engines is a common way to

⁶ <http://www.valarm.net/>

search for sources of geospatial data and extract the geolocations using IP geolocation and gazetteer-based geoparsing and georeferencing. A common approach in this research, following the identification of possible data sources and preliminary collection, is to store the information in SQL/NoSQL databases using numeric and text formats, and then convert them to traditional GIS data – digital maps, point, polyline, and polygon technically. However, VGI/CGD data is continuously provided by public participants, and streaming data (from sensor networks and from sources such as Twitter) are continuous and grow over time, but due to its nature, requires sorting and filtering as a preliminary quality assessment step to identify low quality, irrelevant, or useless data. Considering the many new sources of geospatial data discussed in this section and approaches used by other authors, a targeted geospatial data quality assessment approach is presented here. It is based on previous studies of geocrowdsourcing and existing traditional geospatial data quality methods, and has been revised extensively through interaction with end-users and subject matter experts.

Elements of Geospatial Data Quality Assessment

Positional accuracy perhaps is the best-known quality issue and has been explored by a number of mapping organizations and researchers for decades. US National Map Accuracy Standards (NMAS), published in 1947, was developed for the printed and fixed-scale maps to estimate positional accuracy at a given scale. Then National Standard for Spatial Data Accuracy (NSSDA) replaced the NMAS by using a statistical methodology for estimating the positional accuracy of maps and geospatial data in modern geospatial applications. The NSSDA is more appropriate for the data from GPS

and smartphones devices, because of its suitability for data specified and produced at any geographic scale.

In the book – “Elements of Spatial Data Quality”, Guphill and Morrison (1995) state that geospatial data includes positions, attributes, and relationships of features in space, and identify seven major elements of spatial data quality – lineage, positional accuracy, attribute accuracy, completeness, logical consistency, semantic accuracy, and temporal information. Meanwhile, Veregin (1999) indicates that “quality assessment and reporting is based on minimum quality standards (compliance testing or quality control), metadata standards (truth-in-labeling and fitness-for-use), or market standards (feedback from users)”, and components in data quality can be “assessed in space, time and theme (the three basic dimensions of geographical data)”. With the summaries of Guphill and Morrison (1995) and Veregin (1999), nine parameters of geospatial data quality assessment are cataloged in Figure 3 Elements of geospatial data quality. Combination of these nine quality assessment parameters are used various forms by authors such Girres et al. (2010), Haklay (2010), Ramm et al. (2011), Foody et al. (2013), and Rice (2013a, 2014, 2015a).

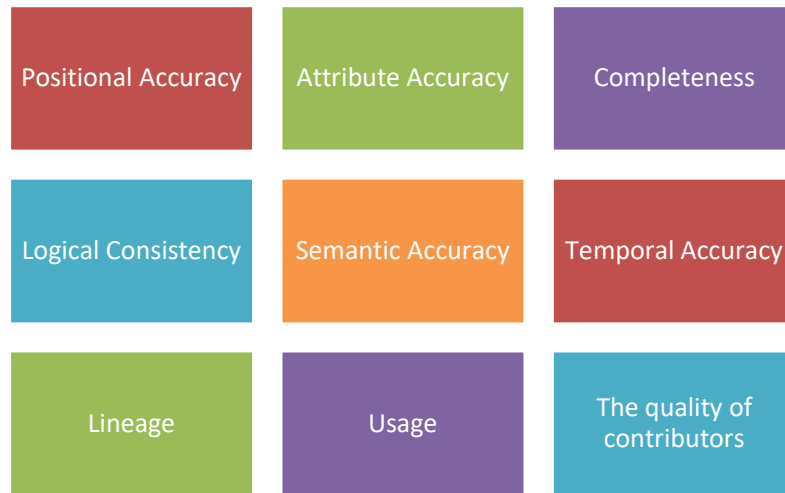


Figure 3 Elements of geospatial data quality

Quality Assessment of Geocrowdsourced Data

Many authors over the past decade have used the well-known geocrowdsourcing application, OpenStreetMap (OSM), as a vehicle for exploring quality assessment, reliability, and coverage dynamics of geocrowdsourcing. Haklay (2010) compares OSM data with authoritative Ordnance Survey data, and finds the positional accuracy of OSM roads data was within 6 meters, but also finds data coverage to be more complete in more affluent areas (Haklay et al. 2010). Girres and Touya (2010) have the similar findings with a quality assessment of French OSM datasets. The range of Euclidean distance between matching intersection points in the road networks is [31.58, 0.68] meters, and the average value is 6.65 meters. In geocrowdsourcing environments, crowdsourced datasets usually do not have enough ground-truth data for validation. Goodchild and Li (2012) suggest that crowdsourcing itself can be used to validate the quality of data. Based on Linus' Law, if enough people contribute, errors will be corrected (Haklay et al. 2010).

On the other hand, for small-scale project, the number of contributors could be limited, then Goodchild and Li (2012) also suggested a social moderation approach, where trained participants can provide a form of “ground truth” for quality assessment. A question based on this idea becomes relevant: How can geocrowdsourced data be moderated and validated and how does it compare to “authoritative” data? This research question has driven this dissertation as well as a number of theses produced by affiliated project researchers and students (Rice 2015b, Paez 2014, Pease 2014)

Scientific Problem Statement

The following three research questions form the scientific problem statement for this dissertation, and are addressed in the principle research chapters that follow:

- 1) Can transient navigation obstacles along a pedestrian network be identified, characterized, and mapped through geocrowdsourcing?
- 2) Can a comprehensive data quality metrics that derived from existing data quality methods, be used to assess and improve the quality of transient obstacles from geocrowdsourcing?
- 3) Can a spatial model be used to optimize the accessibility issue – network maintenance with a geocrowdsourcing testbed?

The first and second questions are addressed in detail in chapters three and four, with conclusions and results presented. The third question is addressed in chapter five, with complete results presented at the end of the chapter. An assessment of these questions and a synthesis of general findings and conclusions is presented in the final

chapter of this dissertation. A short preview of the major research themes (contained in chapters 3, 4, and 5) follows.

Summary of Research Themes

Based on the needs and requirements for establishing a geocrowdsourcing environment for disabled individuals, this dissertation has developed and implemented a unique geocrowdsourcing testbed and a comprehensive quality assessment framework. Iterative improvement of the geocrowdsourcing testbed and data quality assessment framework with public participation has been accomplished and documented in the three research themes that follow. As a final step, the validated geocrowdsourced data is used for measuring pedestrian network accessibility and optimization modeling of pedestrian network maintenance. The three research themes will be summarized in detail, below.

Establishment of a Geocrowdsourcing Environment for Accessibility

Print barriers and movement barriers are part of our built environment, and are partially responsible for the lack of participation in employment by the 60% of disabled individuals that are otherwise capable and willing (Golledge 2001). On the other hand, many channels for streaming information about physical obstacles and barriers are difficult or impossible for some blind and visually-impaired individuals to access. To provide transient obstacle information to disabled individuals (blind, visually-impaired and mobility-impaired) and assist the accessibility of local pedestrian networks under constraints presented by transient and permanent navigation obstacles, chapter three introduces the GMU Geocrowdsourcing Testbed (GMU-GcT), which is designed to

capture geocrowdsourced data in dynamic environmental conditions, and address data quality with a comprehensive geospatial quality assessment method.

The GMU-GcT consists of a web application and a separate mobile data collection tool developed using HTML5, JavaScript, JQuery, CSS, PHP and PostgreSQL/PostGIS. Several data quality assessment variables, including temporal consistency, location (X, Y), location text, image quality, obstacle type, duration, urgency, completeness, and moderator quality score, are used to determine the final quality score of captured data. In the results section of this work, obstacle information is validated and stored in the GMU-GcT dataset. Excluding three extreme outliers due to a technical error, the average positional accuracy of the validated obstacle reports is 3.0 meters, which is in the same general magnitude that reported by Haklay et al. (2010) and Girres and Touya (2010). In addition, the author (Han Qin) also analyzes origin-destination pairs for all 2,772 junctions in the pedestrian network for the study area, and calculates the value of average end-user route length under three different conditions – 1) no constraints 2) an end-user that wishes to avoid stairs and steep paths 3) an end-user that wishes to avoid stairs, steep paths, and validated transient obstacles. As the constraints increase from the first to the third condition, the average route length increases (from 1,834.7 meters to 2,215.5 meters), with three characteristic examples shown. The research concludes with a discussion of the result of this an analysis and the accessibility constraints in the local area, measured with the GMU-GcT system, and summarizes the significant benefit of geocrowdsourcing transient navigation obstacles in the pedestrian network.

Improvement of Geocrowdsourcing Testbed and Data Quality Assessment Framework with Public Participation

The ability of end-users to accurately identify, document, and characterize obstacles is a key to the data quality approach in the GMU-GcT. Project collaborator Paez (2014) designs training videos for the data contribution tools, and also ask potential contributors (including a wide range of students, faculty, staff, and local community members) to familiarize themselves with obstacle categorization used in the GMU-GcT through training exercises. The categories used in these exercises are sidewalk obstruction, construction detour, entrance/exit problem, poor surface condition, crowd/event. In a subsequent training step, participants are asked to identify the obstacle type from several simple unlabeled pictures, and the result shows certain high agreement examples, such as over 95% participants agree the categorization of a picture; and significant disagreement examples, such as six different suggested categorizations for a picture. This phenomenon suggests that a data quality assessment for geocrowdsourced needs to consider the impacts of contributor background and the impacts of miscategorization on assistive geotechnology. As discussed in Semple et al. (2013), feedback from several rounds of semi-formal evaluations can be used to update the prototype leading to an improved product. The author (Han Qin) then engages with end users and data contributors to study and refine the attribute categorizations, such as obstacle types, obstacle estimated duration, and obstacle urgency level. With the feedback from those potential contributors, several major changes were implemented in the GMU-GcT such as multiple selections for obstacle types and the use of broader and simpler duration categorizations. In this work, the author (Han Qin) uses an iterative,

interactive approach between researchers, contributors, and end-users to balance a shared understanding of categorization information, which is also conducive to the improvement of quality of information contributed to the GMU-GcT.

Assessment of the Impact and Use of Geocrowdsourced Data in a Pedestrian Network

Based upon previous geocrowdsourced data quality studies, this research presents a modeling methodology to identify high-value routing corridors in a dynamic geocrowdsourced accessibility system. In chapter five, data quality assessment metrics in (Qin et al., 2015b) are refined with the subject matter experts (SMEs) from GMU's Kellar Institute for Human disAbilities (KIHD), the Transportation Department in Fairfax City, and Smart Growth advocacy group in the City of Fairfax. In the case study presented in this chapter, the default cost and benefit of each segment have been identified by the material types of sidewalk (asphalt, concrete, brick, mortared brick), the usage frequency in complete pairwise routing analysis, and the priority assessment for individual segments throughout the entire network. A segment value optimization scenario is proposed as follows: An organization responsible for pedestrian facilities has a limited budget for sidewalk maintenance and improvement. They wish to decide where to allocate their resources for such maintenance in order to improve the facilities in a way which will provide the greatest overall benefit for the populace which they served, and the benefits of certain segments could be changed due to the contiguity of segments that need to repair. A family of spatial models, including a budget constraint model, a contiguity enforcement model, and a contiguity encouragement model, is developed to make the best decisions when a set of transient obstacles is encountered on the pedestrian

network. These optimization models can evaluate the dynamic environments repeatedly, and help governments and local experts to serve the public needs, such as the accessibility of disabled communities.

The following chapters contain full research presentations of these three themes, followed by a discussion of quality assessment in geosocial media and finally, conclusions, synthesis, and future work.

CHAPTER THREE: GEOCROWDSOURCING AND ACCESSIBILITY FOR DYNAMIC ENVIRONMENTS⁷

Abstract

A consequence of modern society's increasing reliance on digital communication is the concurrent multiplication and narrowing of information streams, with many channels of digital information, but channels which are difficult or impossible for some individuals to access. Blind and visually-impaired individuals are often left out of this communication, unless accommodations are carefully planned and made to present the information in a usable form. The context of this research is the realm of assistive geotechnology, where web mapping technology and geographic information systems are used to provide access to enabling information and services. This paper presents research on the development of tools to provide transient obstacle information to blind, visually-impaired, and mobility-impaired individuals through crowdsourcing, and research on the general accessibility of local pedestrian networks under constraints presented by transient and permanent navigation obstacles. We discuss the social and technological dynamics associated with the creation and use of our crowdsourcing system and present our effort for comprehensive quality assessment. We conclude that crowdsourcing is a crucial technique for successful deployment of assistive geotechnology, particularly those that

⁷ Published in GeoJournal on July 03, 2015. DOI: 10.1007/s10708-015-9659-x. Authors on manuscript: Han Qin, Rebecca M. Rice, Sven Fuhrmann, Matthew T. Rice, Kevin M. Curtin, and Eric Ong.

involve navigation, wayfinding, and travel in public space. We find that many public health services and other important resources cannot be accessed dependably without the use of crowdsourcing and other techniques for capturing dynamic, changing environmental conditions. More broadly, this paper concludes by addressing the progression of user-generated and crowdsourced content from static data contributions to dynamic place based services and the enabling role of assistive geotechnology in providing access and help to blind, visually-impaired, and mobility-impaired individuals.

Introduction

Professor Reginald Golledge, in a 2001 commencement speech at Simon Fraser University, emphasized two obstacles faced by blind and visually-impaired individuals seeking to participate fully in society: the print barrier, and the movement barrier. These two barriers are part of our built environment, and are partially responsible for the lack of participation in employment by the sixty percent of disabled individuals that are otherwise capable and willing (Golledge 2001). Research efforts using assistive geotechnology have been successful in confronting these barriers through tactile mapping (Miele 2011, Perkins 2001, Tatham 1991), GPS-based personal guidance systems (Loomis et al. 2001, 2005), multimodal mapping and spatial displays (Marston et al. 2006, Rice et al. 2005a), and accessible routing (Karimi et al. 2013, Kasemsuppakorn et al. 2009), yet serious problems remain. These problems are associated with transient navigation barriers in our built environment, and other changes to the environment that cannot be incorporated into standard mapping systems due to their transitory nature. This research paper presents an approach for geocrowdsourcing transient navigation obstacles,

and an analysis of the accessibility constraints that can be measured with this system. The results presented in this work underscore the difficulty in overcoming the movement barrier discussed by Golledge for blind, visually-impaired, and mobility-impaired members of society.

Geographical Setting and Context

Like many other cities Northern Virginia has experienced the phenomenon of urban sprawl, with its constant growth and desirable proximity to the District of Columbia. Urban sprawl gone amok has resulted in what Joel Garreau (1991, pp.3-4) has coined as an edge city, an urban center that contains all the functions a city has, though spread-out, absent of sidewalks, and connected not by “locomotives or subways, but by jetways, freeways, and rooftop satellite dishes”. Tysons Corner, Virginia serves as the quintessential edge city; what was once a rolling landscape abundant with orchards and farmland has now transformed into an exorbitant shopping mall and corporate headquarters center, a perpetual construction zone designed to fuel constant suburban growth. Requests to create pedestrian- and bike-friendly features in Tysons Corner have been denied by the Virginia Department of Transportation, due to the impact on vehicle speeds and traffic flow. As a result, Tysons Corner remains an urban center where automobiles and tall buildings dominate, and pedestrians cross eight lanes of traffic to walk from a new metro stop to nearby workplaces (Figure 4).



Figure 4 Pedestrians crossing eight lanes of traffic in Tysons Corner, Virginia. (Photo by Ken Archer, (2014) used with permission)

Twelve kilometers southwest of Tysons Corner and home to George Mason University, the City of Fairfax is situated in the suburban DC area and prides itself on its walkable downtown area. In most respects, it has little resemblance to Tysons Corner. In spite of its efforts to remain pedestrian-friendly, Fairfax faces several challenges in doing so, including consistent residential growth, a steadily expanding George Mason University student body, and relentless construction throughout the area. This has created an environment in which pedestrian routes can vary day-by-day, with serious consequences for disabled pedestrians (Avila 2014, Williams et al. 2013). George Mason University's Office of Disability Services, the entity responsible for providing support services for disabled students, counts approximately 300 students with some form of

functional disability at GMU. Seven out of eight of those students are mobility-impaired and use a wheelchair. Approximately 45 students are blind or visually-impaired.

Because of GMU's diverse student body, including those with mobility and visual impairments, achieving American Disability Act (ADA) compliance is a high priority. Standards for accessible design contained within the guidelines of the Americans with Disability Act, among many other requirements, require that pedestrian pathways not exceed a slope of 1:12, and paths must be wide enough for a wheelchair to pass through. Unfortunately, George Mason's campus is not ADA-compliant in its entirety. An effort to achieve compliance includes the release of an accessibility map on an annual basis, which documents entrances, exits, stairways, steep paths, and general construction perimeters. Although this map is useful, it suffers from the same standard pitfalls noted by Karimi et al (2014). It is static, and therefore cannot capture all obstacles pedestrians may encounter. Some obstacles are permanent—a narrow sidewalk, a steep and winding path, etc.—while some obstacles are only relevant for a certain period of time, whether it be a few weeks, a few days, or a few hours. These transient obstacles can be a closed sidewalk due to campus construction, or a poor surface condition due to rain or ice, which sometimes serve as a minor inconvenience but can also cause a major safety issue, especially those who are disabled (Figure 5).



Figure 5 Temporary construction barricades, fencing, and debris blocking a pedestrian pathway

Due to the abundant pedestrian traffic throughout the City of Fairfax, GMU campus, and surrounding areas, and the clear needs associated with blind, visually-impaired, and mobility-impaired residents, this article presents a system to aid pedestrians, particularly those with mobility or visual impairments. The system uses geocrowdsourcing to collect and add transient navigation obstacle information to an

accessibility map and allows users to create accessible-friendly routes through the area. The system also allows us to identify problems associated with transient obstacles that block the pedestrian walkways, and measure accessibility patterns in our local region.

Literature Review

For the problem under consideration, there is relevant research in areas related to accessibility mapping, geocrowdsourcing, and accuracy assessment in order to identify best practices and approaches that are most useful in designing a geocrowdsourcing application and conducting an analysis of the accessibility characteristics of the study area.

Accessibility Mapping and Applications

Accessibility mapping and accessible routing has been an ongoing project of the tactile mapping community, with many works originating in the International Cartographic Association's (ICA's) Commission on Maps and Graphics for Blind and Partially Sighted Persons, an ongoing commission of the ICA for thirty years. Commission participants Tatham (1991), Coulsen et al. (1991), Eriksson (2001), and Perkins (2001) evaluate the design, technologies, and use of tactile maps, while Miele (2011) and Miele et al. (2004, 2005, 2007) extend traditional tactile maps through automated tactile and audio enhancements built from geographic information systems (GIS) and scalable vector graphics (SVG). Coulsen et al.'s paper from the 1991 ICA proceedings is notable for predicting the use of GIS for accomplishing the tactile map automation, as presented later by Miele (2011), and presages many of the contemporary

accessibility devices and systems presented by Loomis et al. (2005), Beale et al. (2006), Karimi et al. (2014), and others.

Church and Marston explore accessibility mapping in their 2003 paper, and emphasize the variation in accessible route choice by individuals with vision and mobility impairments. Barbeau et al. present a notification system developed to communicate routing information to disabled individuals using public transit (2010). Kasemsuppakorn and Karimi (2009, 2013) developed a routing system for wheelchair users, taking into account slope, sidewalk conditions, traffic loads, etc., using impedance scores for sidewalk segments to determine optimal routes. A similar approach is used by Beale et al. (2006) to route individuals with mobility-impairments through the use of network analysis in GIS while accounting for slope, surface type, and presence of curb cuts. Laakso et al. (2011) review accessibility mapping services and develop formal data models for pedestrian networks to facilitate accessible routing. The concept of a personalized accessibility map is discussed by Karimi, et al. (2014), who developed a crowdsourced tool to aid disabled individuals in way-finding and navigation.

Notable open source accessibility routing applications can also be found in the pgRouting Gallery where applications take into account obstacles such as stairways or steep paths. The Portuguese Accessible Paths in Pinhel (2014), allows for the selection and display of paths suitable for wheelchair users as well as seniors and those with minor mobility impairments, taking slope and pathway material into account when routing. The École polytechnique fédérale de Lausanne (EPFL) also provides an open source routing tool that routes the end-user through buildings (if necessary) and notifies the end-user if

they will be required to change floors. Although these routing applications are a step in the right direction, end-user route preference can be diverse depending on the individual, regardless of impairments (Church and Marston 2003, Jacobson 1998, Golledge 1999, Williams et al. 2013, Avila 2014, Golledge et al. 2000).

Geocrowdsourcing and Volunteered Geographic Information

Goodchild's 2007 and 2009 papers outlining the emerging phenomenon of volunteered geographic information (VGI), and significant reviews of the phenomena by Elwood (2008), and Rice et al. (2012b) highlight the significant benefits associated with this emerging phenomena, which has been applied and studied in a variety of domains and settings (Elwood et al. 2012, 2013). A feature of accessibility efforts by Rice et al. (2011, 2012a, 2012b, 2013a, 2013b, 2014), and more recently Karimi et al. (2014) is the integration of geocrowdsourcing into applications for accessibility mapping, which allows additional information that cannot be captured in a traditional mapping process where the map base is produced in discrete (and often distant) time intervals. The work of Rice et al. and Karimi et al. was preceded by work using mobile communication devices to provide additional information to disabled individuals navigating through a dynamic environment (Nuernberger 2008, Barbeau 2010). The benefits of geocrowdsourcing are numerous, and for accessibility mapping, it appears to be the only effective way to capture the obstacles and navigation hazards that seriously limit the ability of blind, visually-impaired, and mobility-impaired individuals to overcome movement barriers. Cinnamon and Schuurman's work in data divides in public health (2013) underscores the

usefulness of geocrowdsourcing and VGI-based data collection, as a “pragmatic alternative” when official sources of data are missing (2013, 657).

Quality Assessment of Data

A notable weakness of geocrowdsourcing approaches is quality and reliability.

The dynamics between authoritative data (which benefits from technical expertise and training of its producers) and asserted data (produced by untrained end-users) is explored by Goodchild (2009), and later by Goodchild et al. (2012), where three general approaches for quality assessment are articulated. The social approach, articulated by Goodchild et al., includes the intervention and assessment by trained moderators, who fix errors and provide ‘ground truth’ for geocrowdsourced data. This quality assessment method is used in many geocrowdsourcing applications, including the one described in this work and presented initially by Rice et al. (2013a, 2013b, 2014). Elwood et al. (2012) notes that developing a framework for VGI applications is critical, including the development of an appropriate gazetteer and adequate maintenance of framework data, as well as quality assurance.

As reviewed by Ruitton-Allinieu (2011) and Rice et al. (2012b, 2014), Haklay’s 2010 study of OpenStreetMap (OSM) data in the United Kingdom found that the positional accuracy of OSM roads data, when compared to Ordnance Survey data, was within six meters. Girres et al. (2010) performed a comprehensive quality assessment of the French OSM datasets with similar findings. They used a comprehensive set of spatial quality measures articulated by Guptill and Morrison (1995) and Veregin (1999), including positional (geometric) accuracy, attribute accuracy, completeness, logical

consistency, semantic accuracy, temporal accuracy, lineage, and usage. The positional accuracy of features in their sample, determined through Euclidean distance measurements between matching intersection points in the road networks, averaged 6.65 meters, with a maximum of 31.58 meters and a minimum of 0.68 meters.

Assessing the quality of geospatial data can be difficult, with problems originating from many sources and attributable to many causes. The cumulative effects of positional error in thematic layers during GIS overlay operations is commonly cited as an example of the difficulty in tracing errors through source information to a final product. While Goodchild et al. (1989) and Hunter et al. (1992) present useful ways of conceptualizing this error in traditional GIS settings, geocrowdsourcing presents an additional challenge due to data being generated and edited by multiple untrained - and often unknown - participants. The approach used in this research borrows from Goodchild and Li (2012) in using a social approach for moderated quality assessment. This quality assessment process is crucial in being able to collect, analyze, and use transient obstacle data provided by end-users.

Materials and Methods

This effort characterizes the navigation obstacles and accessibility issues in the study area, and discusses the development of the GMU Geocrowdsourcing Testbed. The GMU Geocrowdsourcing Testbed uses a pedestrian network, similar to the work presented in Beale et al. (2006), Kasemsuppakorn et al. (2009), and Karimi et al. (2013, 2014), but also incorporates an extensive quality assessment process articulated here and in Rice et al. (2013a, 2014). The quality assessment process allows for the determination

of the relevancy and impacts of obstacles gathered by crowdsourcing, and the incorporation of those obstacles into the services provided by the system.

Permanent and Transient Navigation Obstacles

In order to assess the accessibility of an area, one needs to determine the characteristics of the pedestrian pathways, and capture information pertaining to obstacles on any given path. This process can be time consuming, depending on the design of the system and specifications of the data models used. Beale et al. (2006), Laakso (2013), Rice (2013a, 2013b, 2014) and the Pinhel Accessibility Platform (2014) offer useful examples of formal approaches for modeling pedestrian infrastructure for accessibility. There are two types of obstacles that are collected by the GMU Geocrowdsourcing Testbed: permanent obstacles and transient obstacles. Permanent obstacles include stairways or steps, paths with a steep slope, narrow sidewalks, crosswalks, curb cuts, pathway surface conditions (such as gravel, cobble stones, etc.), protrusions, and misplaced urban furniture (such as a trash can, bench, or table blocking a pathway). These permanent obstacles are incorporated into the GMU Geocrowdsourcing Testbed. Transient obstacles include items that appear in the walkways but are transient. These items include vehicles on walkways, construction detours, construction barricades, temporary fencing, and barriers related to crowds or special events. These transient obstacles are crowdsourced, as described below.

The GMU Geocrowdsourcing Testbed

The GMU Geocrowdsourcing Testbed was developed from 2010 work presented by Aburizaiza et al. (2011) and Rice et al. (2011), and fully developed over the next two years (Rice 2012a, 2012b, 2013a, 2013b, 2014). The system consists of a web application (Figure 6) developed using HTML5, JavaScript, JQuery, CSS, and PHP, and a separate mobile data collection tool built with JQuery from the same codebase. The data is collected through forms using JavaScript, passed to the server-side using Ajax, and stored in PostgreSQL tables using PHP. The visualization capabilities in the GMU Geocrowdsourcing Testbed are built using PHP, JavaScript, and the Google Maps API. For routing functionality, we use ArcGIS Server to publish the network analysis server and map server, along with a PostgreSQL gazetteer database of building names, street names, and landmarks, along with address points. Transient obstacle data collected through crowdsourcing in our system is converted to barriers in the network using JavaScript and the Esri JavaScript API.

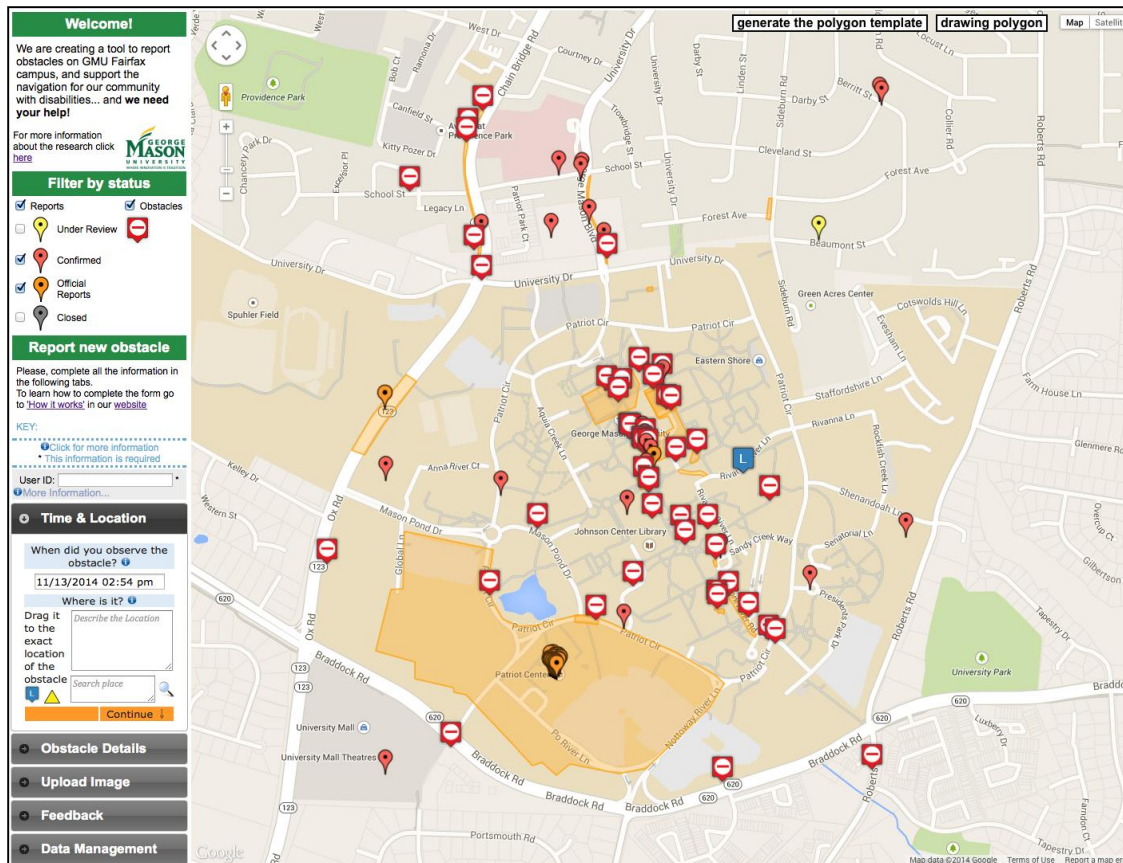


Figure 6 The GMU Geocrowdsourcing Testbed, geo.gmu.edu/vgi (accessed November 22, 2014)

The GMU Geocrowdsourcing Testbed, profiled in Rice 2013a and 2014, relies on a network of contributors who submit reports using the web application. The data contributor drags a location icon to the report location (Figure 7), and supplies the time, date, and report attributes, which include estimated duration, urgency, obstacle type, a text-based location description, and an obstacle description. The reports are then moderated, quality assessed, and converted into obstacles that are displayed through simple point symbols and info-windows in a web application developed with the JavaScript and the Google Maps API. The current transient obstacle data in our system includes 88 obstacles distributed over a study area of 25 square kilometers.

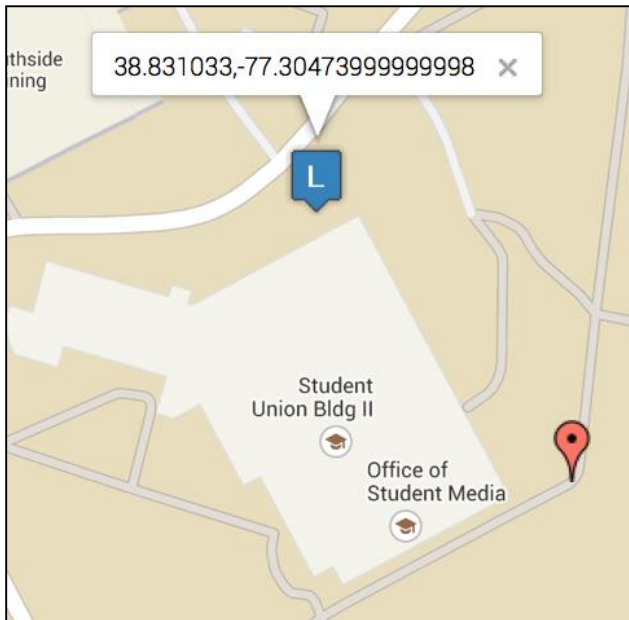


Figure 7 GMU Geocrowdsourcing Testbed, Locator Icon

Pedestrian Network

An analysis of pedestrian routing across several large online mapping websites indicates that nearly every routing service for pedestrians actually uses roadways. Figure 8 shows a typical example for routing between two proximate buildings on the GMU university campus, where pedestrian routes are given on roadways, and written directions refer to travel on the roadways. The actual pedestrian route indicated in Figure 8 is a steep area without ADA compliant walkways, ramps, and crossing points. There are no significant public domain data sources for pedestrian infrastructure in the area, especially for coverage of large areas. As in other similar applications by other researchers, it was necessary to develop an application specific pedestrian network and supporting datasets.

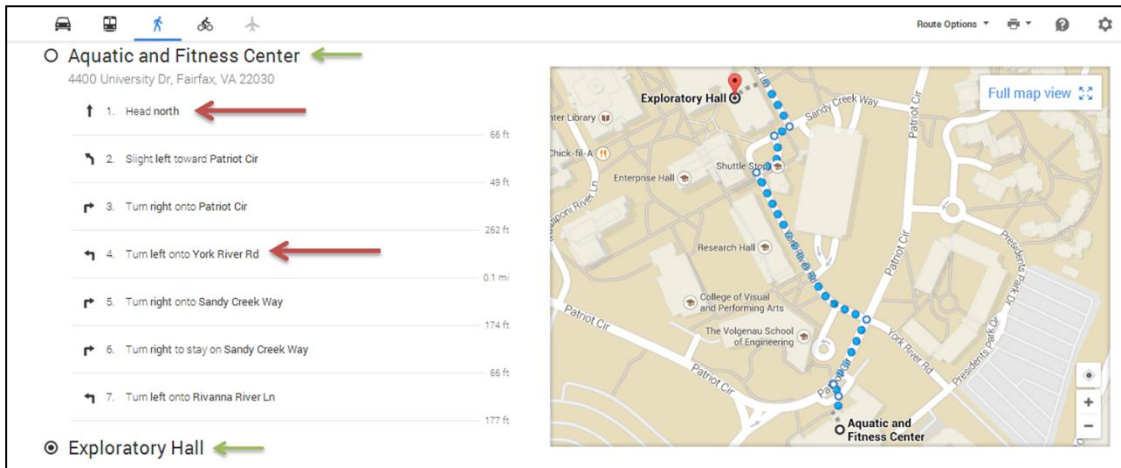


Figure 8 Google Maps pedestrian routing on the GMU Campus (September 2014)

The pedestrian network (Figure 9) was developed through digitizing from ortho-imagery and a sidewalk edge dataset with partial coverage of the study area, which encompasses the entire George Mason University campus, a portion of Fairfax County near the south side of the GMU campus, and the central portion of the City of Fairfax adjacent to campus on the west, north, and east sides. The study area has a 5 km by 5 km extent, for a total area of 25 square kilometers. The study area was chosen to include regions identified by local planning and transportation agencies as being origin and destination points for students, faculty, staff, and residents walking, riding shuttles, or commuting to campus, including students who are blind, visually-impaired, and mobility-impaired.

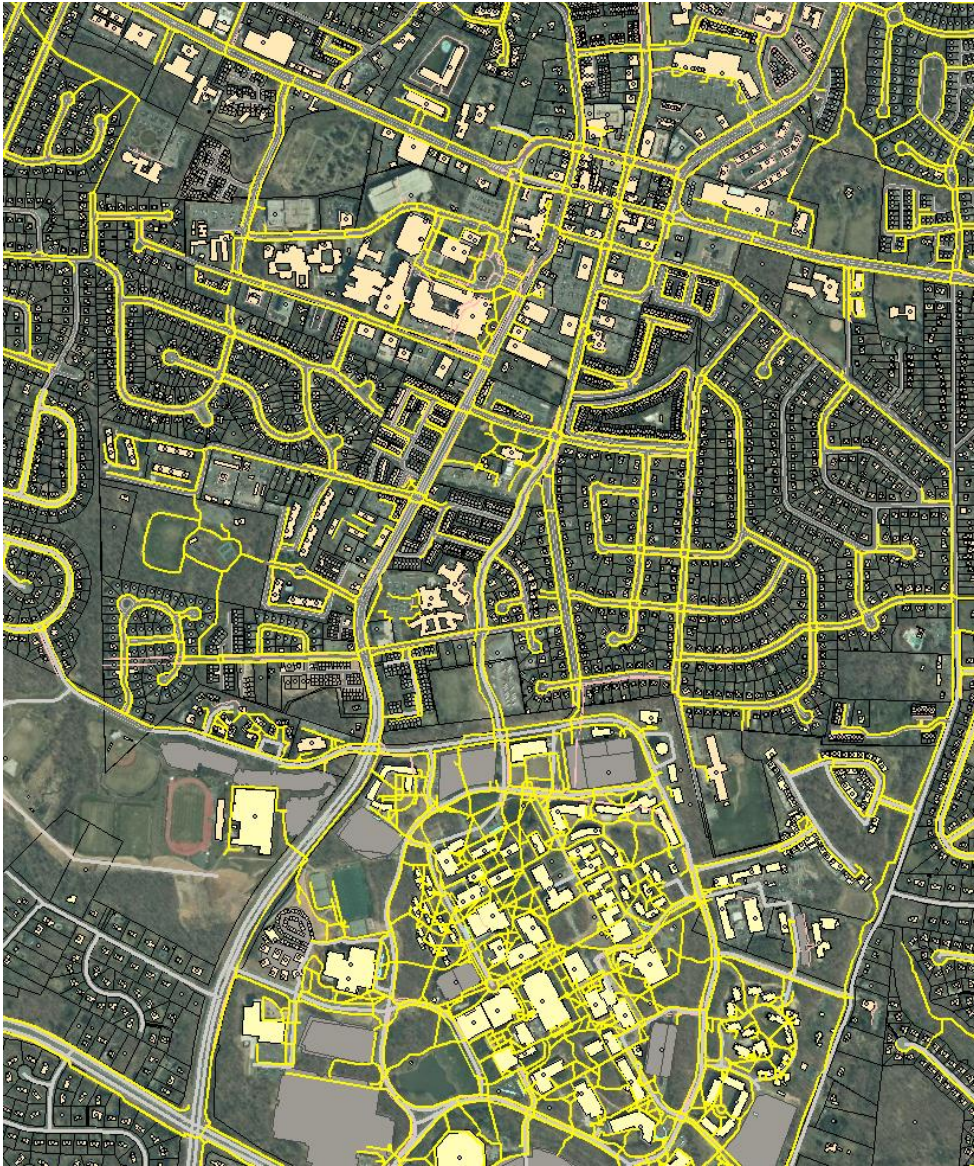


Figure 9 A view of the pedestrian network (yellow) for the local area

The pedestrian network and supporting data includes sidewalk centerlines, stairs, steps, steep paths (with greater than 1:12 slope), crosswalks, and curb cuts. The pedestrian network dataset used for the routing and accessibility studies contains

approximately 2700 network segments. Editing and refinement of this network is ongoing and updates are made on a weekly basis.

Quality Assessment of Data

Goodchild and Li (2012) identify three general approaches for the quality assurance of geocrowdsourced data: the crowdsourcing approach, the social approach, and the geographic approach. The system developed here relies on the social approach for quality assessment, using a small team of experienced moderators to check, validate, and provide ground truth for incoming reports. Once reports are moderated, a quality assessment algorithm is applied to produce a quality assessment score, based on the moderation of the report. The variables used in developing a weighted quality assessment score, as presented in Table 1, include temporal consistency, positional accuracy, location description, image quality, obstacle type, duration, urgency, report completeness, and moderator quality score. These quality assessment parameters are based on spatial data quality metrics and best practices mentioned in the literature review, namely the works of Haklay (2010), Gupta and Morrison (1995), Veregin (1999) and Girres and Touya (2010).

Table 1 Quality Assessment Variables used in the GMU Geocrowdsourcing Testbed

Quality Assessment Variables	Values	Ranks	Weight (%)
QA: Temporal Consistency	0,1	7	6
QA: Location (X,Y)	Max = 1, Min = 0	2	17
QA: Location text	0,1	8	5
QA: Image Quality	0,1,2,3	3	15
QA: Obstacle type	0,1,2	5	11
QA: Duration	0,1,2	6	10
QA: Urgency	0,1,2	4	12

QA: Completeness	0-100	9	4
QA: Moderator Quality Score	1-5	1	20
			100

The quality assessment metrics allow for the display of obstacle quality metrics and the identification of patterns in contributions that are meaningful, and areas that require better spatial or temporal coverage.

Results

The development of the GMU Geocrowdsourcing Testbed offers a unique opportunity to study the dynamics of geocrowdsourcing and to analyze the accessibility patterns in the local area and specifically, the impacts of obstacles on the pedestrian network. The research goal for this project, stated earlier, is to present an approach for geocrowdsourcing transient navigation obstacles, and to conduct an analysis of the accessibility constraints in the local area, measured with this system. Results associated with the collection and quality assessment of transient navigation obstacles are presented below, followed by a study of the accessibility characteristics of the local area through routing.

Positional Accuracy of Obstacles

A detailed summary of positional accuracy characteristics of obstacles in the GMU Geocrowdsourcing Testbed is contained in Rice et al. (2014). Contributors to our Geocrowdsourcing Testbed using mobile devices determine obstacle position through device-based GPS, with a relatively high level of positional accuracy. The majority of obstacles currently in our system have been contributed by users of our desktop data

contribution application, which involves the dragging of a location icon to a report's location on a map. This process, described in Paez (2014) and Pease (2014), can, in some cases, result in relatively large locational errors due to problems with precisely moving the location icon. Figure 10 and Figure 11 show the distribution of positional errors, with Figure 11 having four inordinately large outliers removed. The average positional accuracy of 3.0 meters per obstacle (Figure 11) is an accurate representation of the quality of the positioning for the usual data contributor. Note that there are many obstacles with a very low positional error (many with positional error of 0) indicating that the user has positioned their report well enough that the moderators, via field check, have accepted the position as correct. Figure 10 shows an average positional accuracy of 22.7 meters for the full obstacle dataset. A revised dataset under current use removes four obstacles with unusually large positional errors, reducing the average positional accuracy to 3.0 meters. These figures (3.0 meters to 22.7 meters) are in the same magnitude as the positional errors for crowdsourced feature locations from OSM datasets reported by Haklay (2010) and Girres et al. (2010).

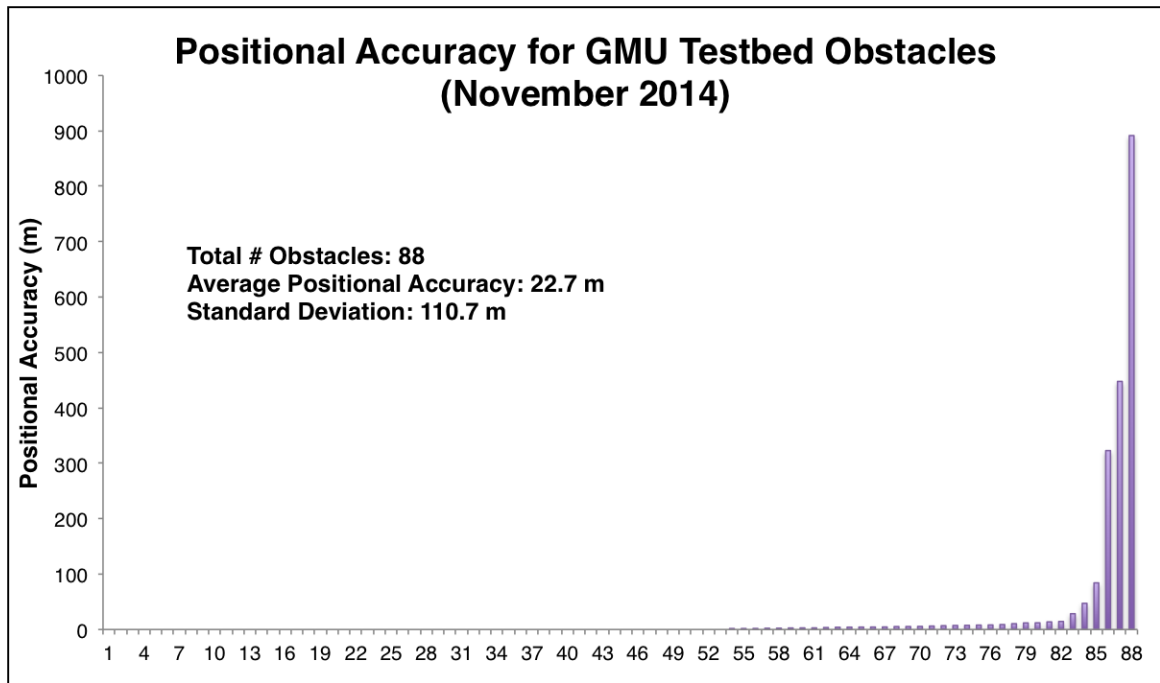


Figure 10 Positional Accuracy for GMU Testbed Obstacles

Pedestrian Network Coverage

Note that one useful result of the generation of a large pedestrian network is the immediate ability to assess the general accessibility characteristics for individuals that require pedestrian pathways for transit to and from the university and local establishments. Figure 11 shows the roadways in the study areas without adjacent sidewalks, and Figure 12 shows an alternative representation with pedestrian-inaccessible areas of the region masked with polygons. The total length of roadways in the study area is 187.2 kilometers and the total length of the pedestrian network (Figure 9, with both sides of the roadways included) is 230.0 kilometers. The total length of the pedestrian-inaccessible roadways (Figure 11) is 58.8 kilometers, or 31.4% of the total roadway

length. This number appears consistent with the observed conditions in the field study area.



Figure 11 Roads (in black) without adjacent sidewalks (at left, with all roads, and isolated, at right)

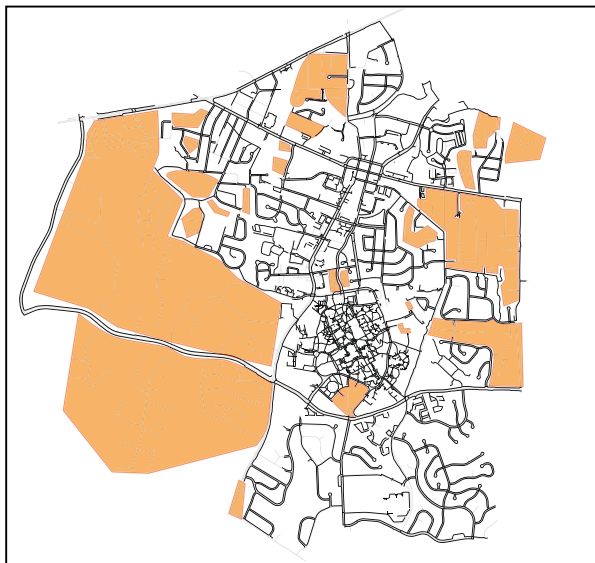


Figure 12 Pedestrian-inaccessible regions masked with polygons

Obstacles and their Impact on the Accessible Pedestrian Network

Figure 13, Figure 14 show the distribution of stairs, steep paths, and obstacles in the GMU Geocrowdsourcing Testbed. The stairs and steep paths are more heavily concentrated on the GMU campus (where the density of buildings and terrain variability is higher), with one notable section in the downtown City of Fairfax. The transient obstacles contributed to our system (Figure 14) are concentrated on a central transportation corridor from the downtown City of Fairfax through the center of the GMU campus.



Figure 13 Location of Stairs and Steep pathways (buffered for visibility), with pedestrian network (left) and isolated (right)



Figure 14 Location of Obstacles from the GMU Geocrowdsourcing Testbed

Figure 15Figure 16 extend the analysis to the underlying pedestrian network. The segments of the pedestrian network impacted by the presence of stairs, steep paths, and transient obstacles are shown in Figure 15Figure 16, and include many lengthy segments in the central transportation axis from the downtown City of Fairfax through the center of the George Mason University campus. The total length of pedestrian network segments currently impacted by stairs and steep paths, and transient obstacles is 14.1 kilometers, or 6.1% of the total length of the pedestrian network. While this proportion seems rather small, the distribution of impacted segments can, and does, have a significant impact on the accessible routes, as seen in the next section.

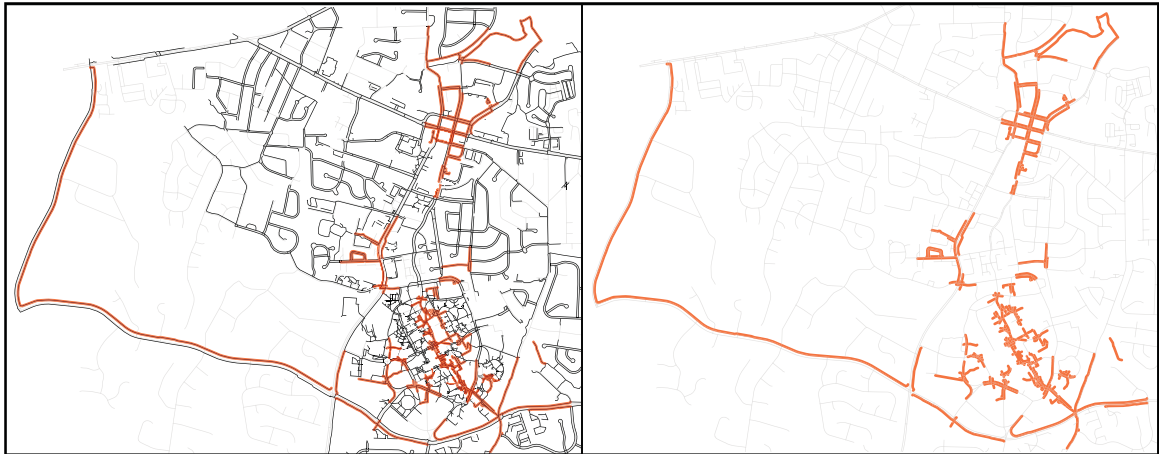


Figure 15 Location of pedestrian network segments impacted by obstacles, with pedestrian network (left) and isolated (right)



Figure 16 Location of pedestrian network segments impacted by Stairs and Steep Paths (purple), and Obstacles (red), with pedestrian network (left) and isolated (right)

Routing and Accessibility Dynamics

The large pedestrian network created for the GMU Geocrowdsourcing Testbed (Figure 9) contains 2700 individual junctions that can be used as origins and destinations for network routing. Examples of accessible routing (Beale 2006, Laakso 2013, Pinhel Accessibility Platform 2014) include many of the same items used in this study (steep

paths, stairs, steps, obstacles) but often lack the element of crowdsourcing, which is a practical way of capturing transient obstacles that are otherwise difficult to represent in a mapping system. Other geoaccessibility projects (Loomis et al. 2005, Marston et al. 2006, Miele 2007) have been successful, but have not used elements of crowdsourcing, which is one of the only practical ways to capture this type of information. Rice et al. (2011) and more recently, Karimi et al. (2014) have adopted this approach with success.

The underlying pedestrian network (Figure 9) contains 2,772 junctions, which generate 7,681,212 origin-destination pairs. Of these pairs, 6,683,094 produce routing results under normal conditions, while 998,118 origin-destination pairs failed to route, which is indicative of the normal topological disruptions of a real-world pedestrian network, where segments of the network are topologically stranded from larger sections.

In order to study the general accessibility characteristics of the study area, routing was performed under three general conditions. First, a study of routing across the entire pedestrian network was conducted using every possible pairwise combination of network junctions. This represents routing across the entire pedestrian network without any accessibility constraints imposed, and resulted in 6,683,094 successful routes and an average route length of 1834.7 meters. This is our initial unrestricted condition.

Second, we repeated the same routing study, but imposed a condition restricting routing across network segments containing stairs and steep paths. This resulted in 4,381,686 successful routes and an average route length of 2121.23 meters, or an increase of 15.62% from the unrestricted condition.

Third, we repeated the routing study again, but imposed the most restrictive condition: routing across the entire network but restricted network segments that contain stairs, steep paths, and transient obstacles. This resulted in 3868459 successful routes and an average route length of 2215.5 meters, or an increase of 20.76% above the unrestricted condition.

The second condition (routing with stairs and steep paths restricted) represents the most common accessibility routing approach. The third condition (routing with stairs, steep paths, and transient obstacles restricted), while the most restrictive condition, is the closest to reality for many of the end-users interviewed for our research (Rice et al. 2013a, 2014). When these end-users encounter stairs, steep paths, or transient obstacles, it forces them to return and reroute themselves along a different pedestrian pathway. The locations of permanent features (steep paths and stairs) can be learned, but the transient obstacles are difficult to predict and therefore, nearly impossible to route around in advance.

The three routing result scenarios presented next in Figure 17, Figure 18, Figure 19, represent three characteristic results. They include routing with modest (Figure 17), substantial (Figure 18), and enormous (Figure 19) increases in path length. Each figure shows the network routing under the three conditions articulated above, with progressively restrictive accessibility conditions. Figure 17 shows an initial routing scenario from an origin to a destination 1,982 meters away. This route takes the user directly through the center of the GMU campus. When stairs and steep paths are restricted, the shortest cost path increases by 4.4% to 2070 meters, which represents a

small reroute or deviation from the shortest cost path. The more restrictive condition, where transient obstacles are used to eliminate additional network segments, increases the length of the path to 2115 meters, a 6.7% increase over the least restrictive condition. This represents many routing scenarios encountered in our system, where routing around stairs, steep paths, and obstacles requires a modest increase in path length.

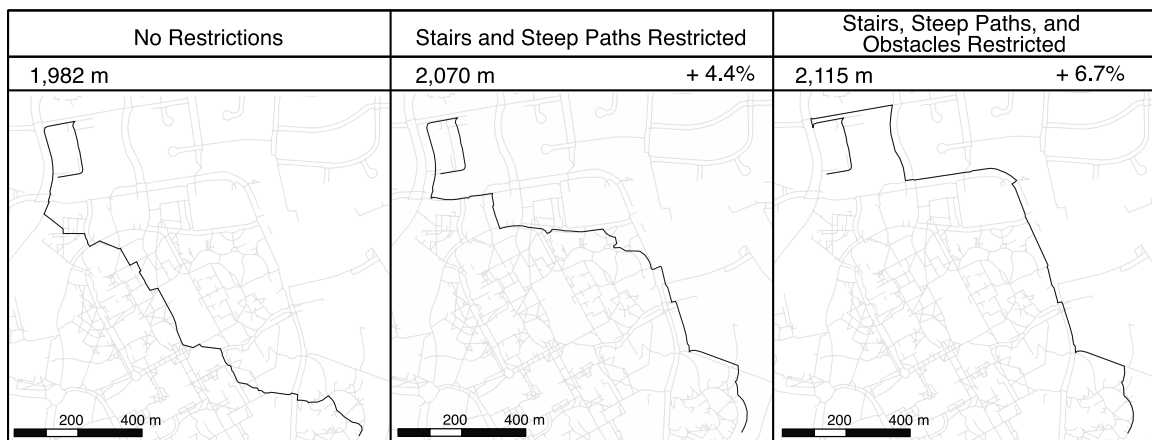


Figure 17 Routing Scenario 1: modest route length increase

Figure 18 shows a route from the study with a substantial path increase under restrictive accessibility conditions. The first condition (no restrictions) results in a path length between origin and destination or 1,319 meters. Under the second condition (stairs and steep paths restricted), the path length changes modestly, increasing to 1,432 meters or 8.6%. Under the third condition (stairs, steep paths, and transient obstacles restricted) the path length increases substantially to 4,065 meters, for an increase of 208.2%. Closer inspection of this routing scenario yields some clues. Transient obstacles along a main road with sidewalks along only one side of the street have resulted in a long route around

an entire neighborhood, to arrive at a destination close to the south side of the study area. This substantial increase in the length of an accessible path is less common than Figure 17, but happens when critical areas with poor sidewalk coverage are impacted by transient obstacles.

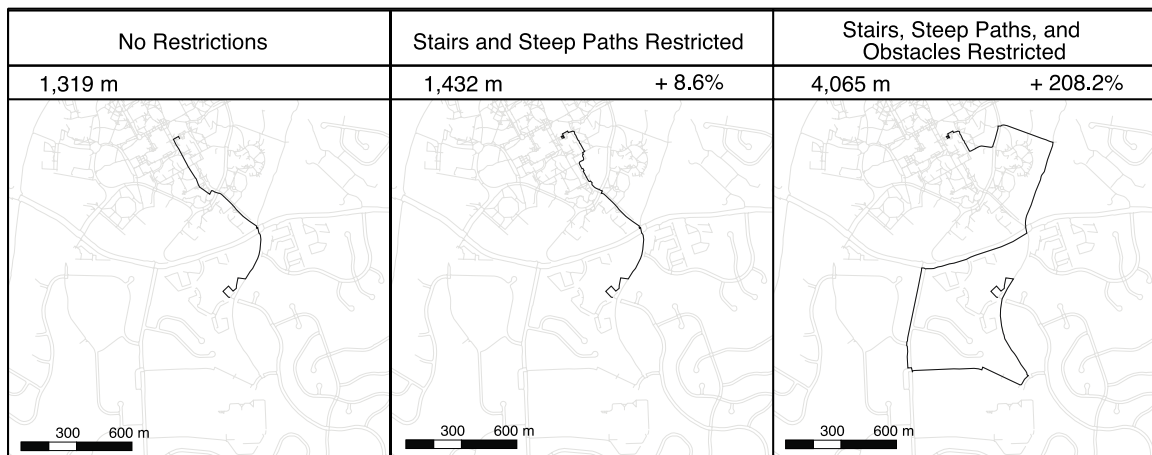


Figure 18 Routing Scenario 2: substantial route length increase

Figure 19 shows the progression from normal conditions to restrictive accessibility conditions in one of the most unusual examples covered by this study. The figure shows a short route from an origin and destination 69 meters apart, under no restrictions. Imposing the second condition (stairs and steep paths restricted) yields an enormous increase in length, to 1,052 meters or an increase of 1421%. The route is increased even further by the imposition of the third condition (stairs, steep paths, and transient obstacles restricted), with a final route of 1,294 meters, or 1915% above the initial route length. Inspection of this scenario allows us to conclude that an unusual combination of stairs, steep paths, obstacles, and natural features (buildings) causes a

reroute of unusual length. This scenario is rare, but not unheard of, and represents one of the worst cases of route length increase under restrictive accessibility conditions.

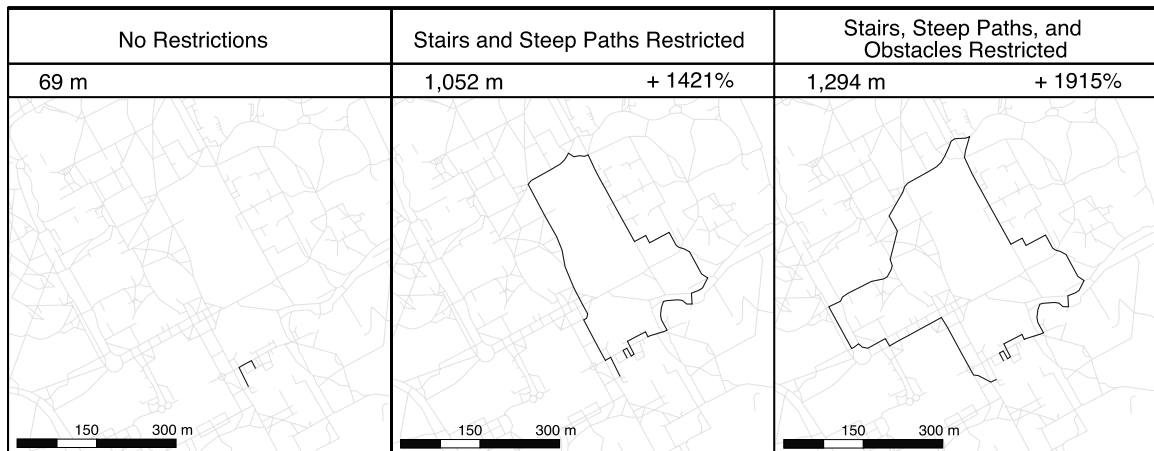


Figure 19 Routing Scenario 3

Discussion

Routing in the GMU campus region is difficult due to the overwhelming emphasis on vehicular transportation as the preferred mode of transit. While portions of the local area are urbanized and walkable (notably the City of Fairfax), other portions of the study area remain largely inaccessible, due to the lack of sidewalks along roadways, lack of curb cuts and connecting crossing points, and the presence of transient obstacles. These transient obstacles have a very large impact on the accessibility of the local area. Figure 17Figure 18Figure 19 demonstrate some of the characteristic results of imposing accessibility constraints on routing. These examples demonstrate that restricting routing to stairs and steep paths increasing the route length and forces rerouting. For the second condition (stairs and steep paths restricted) the average path length increase for the study

was 15.62% above normal conditions. For the third condition (stairs, steep paths, and obstacles restricted) the average path length increased 20.76% above normal conditions. Figure 12 already shows that large portions of the study area are inaccessible due to the lack of sidewalks and other infrastructure such as curb cuts and crosswalks. Figure 16 shows that large portions of the study region with sidewalks are in fact inaccessible for many individuals because of the presence of stairs, steep paths, and transient obstacles that block access to the pedestrian network. The impact is felt directly by blind, visually-impaired, and mobility-impaired individuals who require accessible sidewalks, but it is also felt by senior citizens, families with strollers, and individuals with minor and temporary mobility conditions.

Church and Marston (2003) note that route choice is highly individual, and that traditional accessibility measures do not take into account the significant physical and mobility differences of individuals. Their 2003 study demonstrated these individual differences in 4 individually-selected routing results between the same origin and destination with lengths varying from 182 meters to 610 meters. This 335% increase, due to individual preferences, was reported to represent common variation between individuals. Interviews with end-users of the GMU Geocrowdsourcing testbed included anecdotal information confirming the large individual variation in routing preferences, due to conditions such as tree cover, sunlight, ambient temperature, and minor variations in cross slope (Rice et al. 2014, Figure 20). This individual variation in routing preference makes a single accessible routing portal difficult to imagine, unless extensive routing customization is allowed for each user. In this regard, the recent work of Karimi

et al. (2014) is important and innovative. They follow best practices for accessible routing from the work of other researchers, and build individual preferences into the routing process.



Figure 20 Individual route preference (green) and the shortest cost path (blue), from Rice et al. 2014

Summary

Recognizing the large impact that transient obstacles have on accessibility for the blind, visually-impaired, and mobility-impaired, this research effort presents an approach for geocrowdsourcing transient navigation obstacles, and analyzes the accessibility constraints in the local area. The GMU Geocrowdsourcing Testbed uses the best practices from geocrowdsourcing and quality assessment metrics developed and presented by authors such as Haklay (2010), Goodchild et al. (2012), and Girres et al. (2010). The GMU Geocrowdsourcing Testbed allows us to conduct routing under normal, moderately

restrictive, and restrictive conditions associated with the requirements of blind, visually-impaired, and mobility-impaired subjects, and provides insights on the accessibility of the local area.

Future Work

The intention is to expand our study area beyond the 25 square kilometer region surrounding the George Mason University campus, to larger areas with an interest in sustainable transportation and accessibility, notably, Washington, D.C. and Portland, Oregon. These efforts will require redesign and revision of our GMU Geocrowdsourcing Testbed, and the collaboration of other researchers who have shown success in their efforts to implement novel accessibility research programs, with Karimi et al. (2014) and Laakso (2013) as aspirant collaborators. These efforts, to some extent, may be directed or coordinated through the International Cartographic Association's Commission on Maps and Graphics for Blind and Partially Sighted Persons, which has a long and distinguished history of innovation and progressive thought in this area.

Ideally researchers could implement a crowdsourced routing paradigm, modeled after Kulyukin et al. (2008), which may be useful in allowing for stable route generation through difficult areas, most of which are frequently transited by end-users of the GMU Geocrowdsourcing Testbed. The intelligence gathered through crowdsourcing routes will not only be practical to system users but will allow us to study the selection of alternative routes and measure their characteristics.

Another avenue for future research involves the identification of elements of the pedestrian network that are most influenced by barriers. That is, the methods described

above can be altered to identify the number of times a particular segment of the network is unavailable for travel between origins and destinations. This is a measure of the seriousness of the obstacle, or alternatively a measure of the benefit that could be achieved if the obstacle were removed. When all network segments have such an associated measure, a determination can be made regarding how to optimally allocate resources for mitigating obstacles. This would represent an extension to the integration of GIScience and optimal facilities location science that has been progressing over the past decade or more (Church 2002, Curtin et al. 2005).

CHAPTER FOUR: OBSTACLE CHARACTERIZATION IN A GEOCROWDSOURCED ACCESSIBILITY SYSTEM⁸

Abstract

Transitory obstacles – random, short-lived and unpredictable objects – are difficult to capture in any traditional mapping system, yet they have significant negative impacts on the accessibility of mobility- and visually-impaired individuals. These transitory obstacles include sidewalk obstructions, construction detours, and poor surface conditions. To identify these obstacles and assist the navigation of mobility- and visually-impaired individuals, crowdsourced mapping applications have been developed to harvest and analyze the volunteered obstacles reports from local students, faculty, staff, and residents. In this paper, we introduce a training program designed and implemented for recruiting and motivating contributors to participate in our geocrowdsourced accessibility system, and explore the quality of geocrowdsourced data with a comparative analysis methodology.

Introduction

Mapping dynamic geographic phenomena is often difficult, due to the requirements for frequent updates and changes that occur over time. In urban areas, pedestrian corridors and transportation infrastructure are the most in-demand, critical

⁸ Published in ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences on August 2015. Authors on manuscript: Han Qin, Ahmad O. Aburizaiza, Rebecca M. Rice, Fabiana Paez, Matthew T. Rice.

features to map, yet they are frequently impacted by sidewalk obstructions, construction detours, and changing surface conditions (Figure 21).



Figure 21 Transient obstacles in the pedestrian corridors, Fairfax, Virginia

For individuals with vision or mobility impairments, changes to the pedestrian corridors (even temporary ones) are very difficult, due to the necessity for rerouting, inconvenience, and the increased risks associated with safety hazards. Mapping these areas with high-frequency coverage is essential. Where remote sensing and automated feature extraction is not a realistic approach due to cost or required frequency of updates, a different approach is needed. Rein (2009) provides evidence of the great cost and difficulty associated with improving and maintaining a functional, accessibly multimodal transportation network. Beale et al. (2006), Laakso et al. (2011, 2013), Karimi et al. (2013), and Kasemsuppakorn et al. (2009) provide useful approaches for pedestrian network modeling and applications for routing disabled individuals. Nuernberger (2008) and Barbeau et al. (2010) show how communication devices can be used to enhance

navigation and improve communication with disabled individuals. The primary missing element in these approaches is the ability to capture and map transient events, which is difficult due to their dynamic nature.

Rice et al. (2011, 2012a) present a useful approach for mapping transient obstacle data using open source software, gazetteer-based geoparsing, and geocrowdsourcing. Subsequently, Rice et al. (2012b, 2013a, 2013b, 2014) extend this approach with quality assessment approaches, routing tools, and visualization. Paez (2014) presents a study of training and semantic understanding in geocrowdsourcing systems, and documents approaches for teaching contributors to identify and characterize transient obstacles in a geocrowdsourcing system. Through comparisons to other training systems embedded in OpenStreetMap, Google Map Maker, and the USGS's The National Map, Paez identifies key strategies and approaches for training and summarizes those approaches. Rice et al. (2014) review the work of Paez (2014) and provide additional insight into the use of training systems and quality assessment of geocrowdsourced data.

Goodchild (2007, 2009) first introduced the domain of geocrowdsourcing or 'volunteered geographic information' (VGI) in order to describe the phenomena of non-professionals creating and utilizing geographic data. The first generation of VGI applications aimed to collect georeferenced data and observations to be stored in a database. A newer concept of VGI, proposed by Thatcher (2013), is known as 'volunteered geographic services', or VGS. VGS differs from traditional VGI in the sense that while VGI is collected once and remains static, VGS permits users to add, edit, or delete the entries. VGS can be considered as a dynamic version of VGI. It is focused

more on actions between users who offer or use services. One good example of VGS is SeeClickFix, which permits users to report neighborhood issues to authoritative agencies such as tree falling, power outage, traffic light problems, offended graffiti, and potholes. SeeClickFix permits users to vote on problems to elevate the visibility of the report. The initiator can also close a case once the problem is resolved. Other examples of dynamic VGI or VGS applications are Carma Carpooling (formerly known as Avego) and Waze. The work presented here uses data contributed by the public (VGI) but also enhances the active use of this data through active map-based routing and other services (VGS). A critical aspect of the provision of services is an active quality assessment system, which depends on an understanding of the obstacle characterization abilities of the data contributors. The following sections of this paper present an overview of the GMU Geocrowdsourcing Testbed and its moderation and quality assessment program, the training and obstacle characterization studies, a short discussion of user motivations, a summary of positional accuracy characterization, and finally, conclusions and future work.

The GMU Geocrowdsourcing Testbed

The GMU Geocrowdsourcing Testbed (GMU-GcT), presented in Rice et al. (2014), was developed as an experimental approach for mapping transient navigation obstacles in the region surrounding the George Mason University Campus in Fairfax, Virginia. The GMU-GcT uses crowdsourced data contributions from members of the public, who identify, document, position, and describe obstacles through a map-based crowdsourcing approach. This system is built on the work of Paez (2014) and other best

practices for map-based geocrowdsourcing. The system uses a menu-driven, semi-structured reporting process where end-users provide location information, temporal tags, images, and attribute data for obstacles in the Fairfax, Virginia region (Figure 21). The GMU-GcT has both desktop and mobile contribution tools to facilitate obstacle reporting by the public and provides services such as obstacle-avoidance routing (Qin et al., 2015b).

Moderation and Quality Assessment in the GMU Geocrowdsourcing Testbed

Many authors have explored data quality issues in geocrowdsourcing, including Haklay (2010) and Girres et al. (2010) who explore the accuracy of crowdsourced features in OpenStreetMap through comparisons with authoritative Ordnance Survey and French IGN data. The assessments in Girres et al. (2010) are noteworthy for the thoroughness of the evaluation in the context of well-known map accuracy and GIS data accuracy characterization methods, such as those published by Gupthill et al. (1995), Hunter et al. (1992), and Veregin (1999). Goodchild and Li (2012) suggest different methods for quality assessing geocrowdsourced data. Because of the unofficial nature of the data, the GMU-GcT uses Goodchild and Li's social approach, where a team of trained moderators checks data contributions for errors. This quality assessment system is discussed in Rice et al. (2014) and Paez (2014). A key to the data quality approach in the GMU-GcT is the ability of end-users to accurately identify, document, and characterize obstacles. While some elements of data contribution, such as map-based obstacle positioning, are objective and easy to analyze for errors, other elements are more difficult. Obstacle characterization in the GMU-GcT requires a shared understanding of

obstacle categories between the system administrators and system end-users. This categorical characterization process, as noted by Camponovo et al. (2014) in the context of the Haiti Earthquake response, is problematic. They report that over fifty percent of the messages to the Ushahidi web platform were mischaracterized with regard to emergency need. Foody et al. (2013) develop approaches for assessing the relative accuracies of end-user categorization of land cover. Galloway et al. (2013) tested the ability of young contributors to identify tree species, finding the contributors over-reported rarer classes. Quality assessment for crowdsourced geospatial data is a challenge, as noted by many of the fore-mentioned authors. The approach for quality assessment used in the GMU-GcT is built on the best practices for these approaches, with Goodchild and Li's moderator-based social quality assessment as the model.


Training and Obstacle Categorization in the GMU Geocrowdsourcing Testbed

To improve the quality of information contributed to the GMU-GcT, Paez (2014) conducted a thorough review of training strategies in map-based social applications and geocrowdsourcing and found the most effective methods of training to be those that were embedded within the data contribution tools, such as those embedded within Google Map Maker and OpenStreetMap. Based on Paez's work, training videos are embedded within the GMU-GcT, through the "How it Works" link. While effectiveness or success of the training methods can be difficult to assess, Paez (2014) provides a summary and assessment of best practices.

For the GMU-GcT, the primary means of characterizing transient navigation obstacles is through the placement of a map-based location icon, the categorization of the

obstacle in question, and categorical assessments of the obstacles expected duration and urgency. While the positioning of obstacles is of high interest (Rice et al., 2015a), a critical aspect is the contributor's obstacle categorization. Potential contributors to the GMU-GcT (which includes a wide range of students, faculty, staff, and local community members) were asked to participate in a training exercise to learn obstacle categorization, and were shown a series of pictures with authoritative obstacle characterizations shown (Figure 22). The general obstacle types shown in Figure 22 and used in the GMU-GcT (sidewalk obstruction, construction detour, entrance/exit problem, poor surface condition, crowd/event) were derived from an end user study conducted in 2013 and reported in Rice et al. (2013a).

Examples (cont'd)



Obstacle Type	Duration
Sidewalk obstruction	Low (<1 day)
Construction detour	Medium (1-7 days)
Entrance/Exit problems	Long (> 7 days)
Poor surface conditions	
Crowd/Event	

Priority
Low
Medium
High



Where Innovation Is Tradition

Figure 22 Training graphic with picture and obstacle type, obstacle duration, and obstacle priority characterization

A subsequent training step asked potential contributors to identify the obstacle type from a simple unlabeled picture (Figure 23) using the obstacle types presented in the previous step. While some obstacle pictures were relatively easy to characterize (Figure 23, Figure 24) others were much more complex (Figure 25, Figure 26).



Figure 23 GMU geocrowdsourcing training picture showing crowd/event

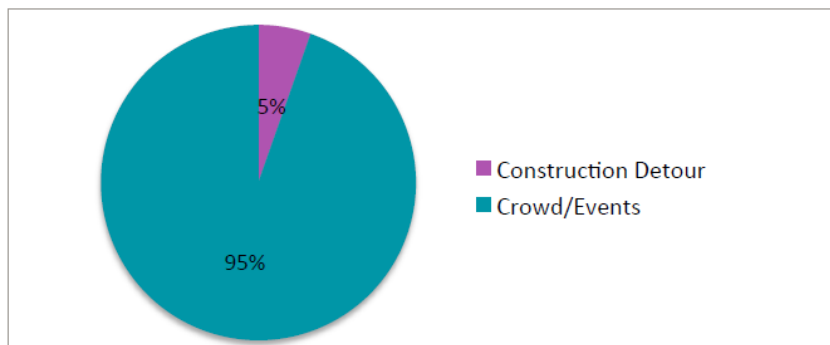


Figure 24 High agreement (95%) for categorization of Figure 23 as a crowd/event obstacle by 37 training participants

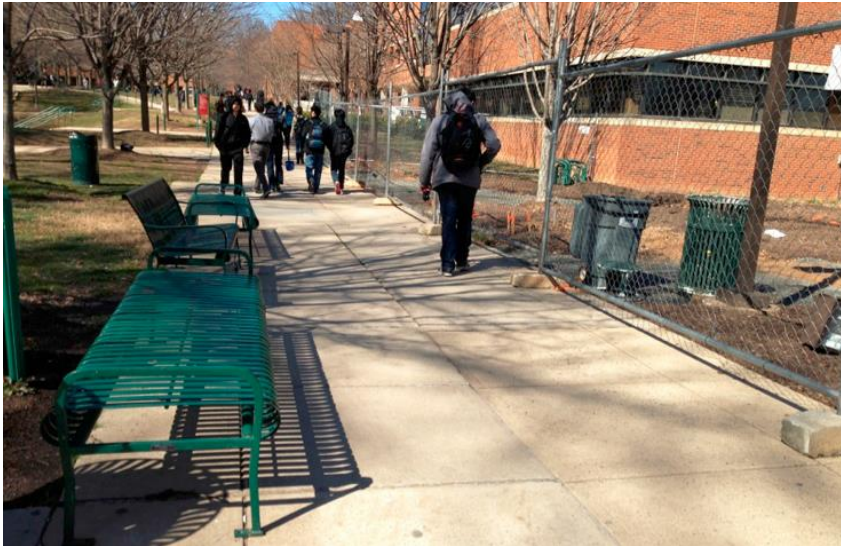


Figure 25 GMU geocrowdsourcing training picture showing sidewalk obstruction

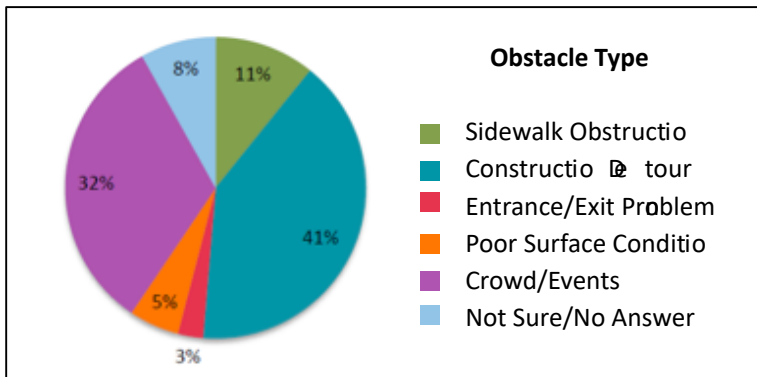


Figure 26 Significant disagreement for categorization of Figure 25 by 37 training participants

Figure 27Figure 28, similar to Figure 25Figure 26, show relatively high levels of disagreement among the forty-six members of a second training cohort in early 2014. In Figure 27, the water hose draped across the sidewalk is interpreted as an obstacle, as are various construction barricades and construction vehicles visible off-screen.

Figure 29Figure 30 show relatively low levels of disagreement for the category of the object shown in Figure 29, which may be attributable to the less complex scene. A

majority of participants categorized the obstacle in this image as a poor surface condition.

Figure 31 shows the frequency of all obstacle category tags in the GMU-GcT as of January 2014.



Figure 27 GMU geocrowdsourcing training picture showing sidewalk obstruction

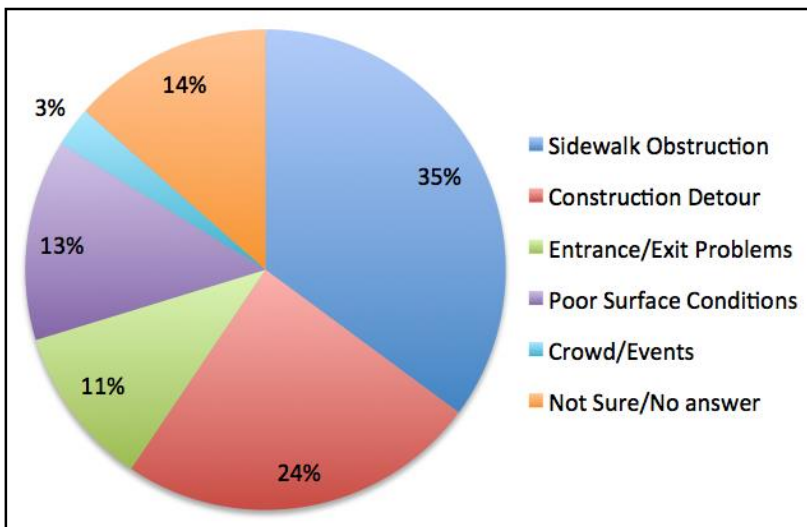


Figure 28 Significant disagreement for categorization of Figure 27 by 46 training participants



Figure 29 GMU geocrowdsourcing training picture

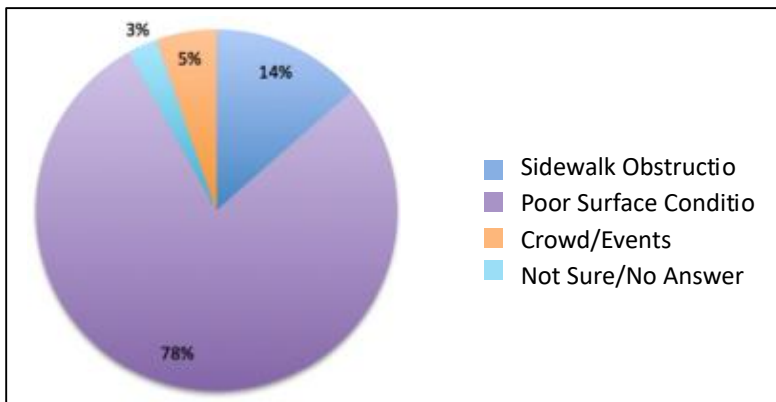


Figure 30 Moderate to low disagreement for categorization of Figure 29 by 46 training participants

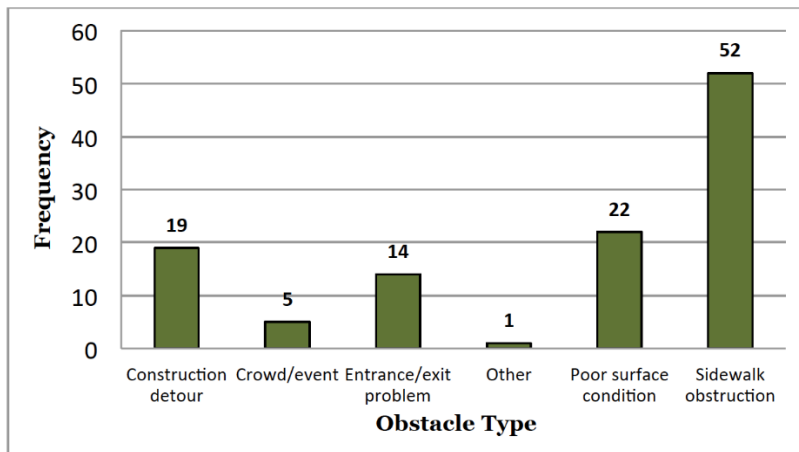


Figure 31 Distribution of obstacle types in GMU-GcT, January 2014

During the course of Paez’s study (Paez, 2014, Rice et al., 2014), 150 potential system contributors were trained in obstacle characterization. A few important results emerged from the training. First, participants recorded some disagreement in obstacle characterization for complex scenes (Figure 25, Figure 28) where more than one potential obstacle can be seen. Initial training subjects requested the ability to declare more than one category of obstacle for each training image. A resulting change in the GMU-GcT was the allowance for multiple obstacle types for each object. Another change precipitated by feedback from training subjects was the use of broader and simpler duration categorizations, which are now instituted as shown in Figure 22, with durations being low (less than 1 day), medium (1-7 days), and long (greater than 7 days).

Current efforts in the training of participants for the GMU-GcT is the modularization of training materials into short videos, and the embedding of training material directly into the contribution tools, as recommended by end-users and noted in Paez (2014) and Rice et al. (2014).

User Motivation in the GMU Geocrowdsourcing Testbed

The success of many geocrowdsourcing projects depends on the motivations and willingness of participants. For OpenStreetMap, an initial motivation for participation is described by Coast (2006) as resentment over the data pricing and data licensing practices of the Ordnance Survey. Furthermore, Coleman et al. (2009) note that differences in user needs and motivations are based on the different contexts in which they contribute. A study by Pease (2014) underscores the potential problems associated with user motivation in geocrowdsourcing. While typical positional accuracy levels for the GMU-GcT are in the same range as feature position accuracy positions noted by Haklay (2010) and Girres et al. (2010) for OSM data (Rice et al., 2014), Pease noted lower levels positional accuracy, and a higher level of disagreement for obstacle categorizations. In the study by Pease (2014), students were asked to characterize the position and attributes of obstacles along a path. The obstacles had previously been thoroughly characterized by trained moderators and project staff. Pease noted an average positional error of 13.07 meters for obstacle reports, and greater level of disagreement for obstacle categorization than for similar reports collected by the same GMU-GcT from the same area. The differences, according to project staff (Paez, 2014, Rice et al., 2014) were due to the motivation of the study participants. While in general, participation in the GMU-GcT is based on altruism and a desire to correct accessibility problems, the study participants in Pease (2014) were recruited from a general subject pool where extra credit was offered, which was suggested as a reason for their performance.

Position Accuracy

The GMU-GcT uses the positioning of reported obstacles in order to facilitate obstacle avoidance routing on a pedestrian network, which is motivated by the accessibility mapping and accessibility wayfinding work of Golledge (1999), Golledge et al. (2000, 2005, 2006), Church et al. (2003), Jacobson (1998), Pingel (2010) and Rice et al. (2005). Implemented in the system are four methods of evaluating positional accuracy: (1) human-georeferenced geographic coordinates via computer, (2) mobile GPS coordinates from user's current position, (3) embedded geo-tags in an image of reported obstacle, and (4) convex hulls created from geoparsed text descriptions of the obstacle's location, based on a comprehensive gazetteer. This work, characterized in Rice et al. (2015a), involves the development of metric georeferenced footprints from geoparsed location descriptions (red outline in Figure 32, based on Rice et al., 2011), geo-tagged positions embedded within contributed obstacle images (blue marker in Figure 32), obstacle positioning from a movable location icon (yellow marker in Figure 32) and positioning determined through a social moderation of the report (green marker, Figure 32). This work continues, with recent efforts to use electronic compass data embedded in image headers to validate the positions of obstacles (Rice et al., 2015a).

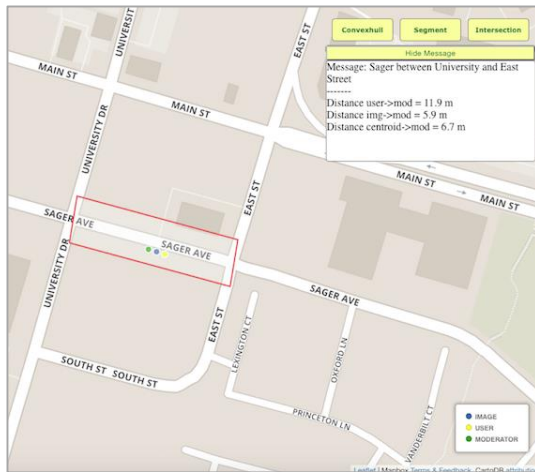


Figure 32 Positional characteristics of an obstacle report

Conclusion and Future Work

The GMU-GcT has been built to gather information about transient navigation obstacles in the local environment. Gathering these obstacles using a geocrowdsourcing system provides many benefits, including the ability for blind, visually-impaired, and mobility-impaired individuals to be informed in advance about unexpected changes to a pedestrian pathway. It will also allow these individuals to re-route, and avoid risks and significant delays. The data quality assessment is done through a social moderation process and based on best practices. The ability of data contributors to characterize and categorize obstacles is important to the system, and has been studied and improved through an iterative process and feedback from end-users and contributors (Paez, 2014, Rice et al., 2014). The GMU-GcT allows for obstacle avoidance routing, and to do so requires obstacle data whose positional characteristics and attributes are well known. Project researchers will continue to improve the system and maintain high quality geographic information. Critical future work will refine the routing elements and active-

use elements of the system, including mobile data validation tools and mobile moderation and quality assessment tools. A recent publication by Karimi et al. (2014) offers insights and strategies in a similar project, which will provide opportunities to adopt best practices and workflows from a different geographic setting.

As noted in the two previous chapters and the following chapter, pedestrian infrastructure is a critical part of the urban landscape. Maintenance and condition of pedestrian walkways are important in quality of life measurements and real estate marketing, and can facilitate opportunities for business and commerce (Leinberger and Alonzo 2012). The previous two chapters have addressed a geocrowdsourcing testbed and the reporting of obstacles with the testbed, as well as the comprehensive, iterative, development of a quality assessment workflow. Geocrowdsourced data contributions are useful by themselves, as reports of obstacles, but can also be used in routing analyses and accessibility analysis, such as that presented in chapter three, where accessible routes between an origin and destination get dramatically longer when stairs, steep paths, and obstacles are avoided. The next chapter looks at the use of geocrowdsourcing reports as a primary input for maintenance operations. These reports indicate problems that need to be resolved, and can be trusted in the context of maintenance operations due to the existence of a significant quality assessment workflow. The reports form a basis for repair operations, along with other data gathered from subject matter experts. As such the following chapter is one of many extensions of the previous chapters and the general capabilities of the GMU-GcT.

CHAPTER FIVE: ASSESSING THE IMPORTANCE OF PEDESTRIAN NETWORK SEGMENTS IN A GEOCROWDSOURCED ACCESSIBILITY SYSTEM⁹

Abstract

Pedestrian infrastructure is an essential part of the urban fabric. Typically, it is carefully planned and maintained by governments and local experts, who recognize the benefits to health, well-being, and even economics associated with a walkable environment.

Pedestrian walkway characteristics, including running slope, cross slope, curb cuts, cross walks, sidewalk widths, and signalization are a part of the comprehensive design elements used by most municipalities. However, barriers or obstacles, including temporary obstructions, construction detours, and surface irregularities make this infrastructure difficult for mobility and vision impaired individuals to use.

Crowdsourcing can assist these individuals by providing information about transient and permanent navigation obstacles, through an accessibility mapping system. Accessibility mapping systems, several examples of which are discussed in this paper, provide routing functions to make navigation easier for vision impaired and mobility impaired individuals. A geocrowdsourced accessibility system can also identify deficiencies in a pedestrian network dynamically, and can provision routing and obstacle avoidance functions in real-time, with data about transient events provided by the public. This paper

⁹ Under review in GeoJournal. Authors on manuscript: Han Qin, Kevin M. Curtin, and Matthew T. Rice.

is based upon previous geocrowdsourced data quality studies, and presents a modeling methodology to identify high-value routing corridors in a dynamic geocrowdsourced accessibility system. The corridor measurement can help civic employees (City Public Works and Transportation departments) prioritize maintenance of a pedestrian infrastructure, including the rectification of obstacles identified through crowdsourcing. In this paper, we augment geocrowdsourcing data quality metrics with input from subject matter experts trained in orientation and mobility services, and discuss the accessibility elements that could directly influence the usability of the pedestrian infrastructure. We also present a cost optimization model to measure the value of a pedestrian network segment. Lastly, this paper analyzes how the value of a network segment in a geocrowdsourced accessibility system changes with network conditions and how this relates to prioritization of maintenance tasks through optimization criteria.

Introduction

In a 2001 commencement address, Dr. Reginald Golledge (2001) described two critical barriers – the print barrier and the movement barrier – that limit the participation in employment by the 60% of disabled individuals that are otherwise capable and willing. The use of digital technology within modern communication systems raises the print and movement barriers even higher, by making critical information less accessible and more difficult for the disabled to use. Many different tools, programs, and policies have been created, tested, and modified to help bridge this digital divide and overcome these barriers. Accessibility mapping systems, such as Miele et al.’s Tactile Map Automated Production system (TMAP, 2006) or Loomis et al.’s Personal Guidance System (PGS,

2001) are able to assist blind and visually-impaired individuals in route planning, navigation, and spatial awareness by presenting geographic information in tactile and auditory forms. These systems, and related work for improving the accessibility of web pages, e.g. Thatcher et al. (2002), will help the disabled with these difficult print and movement barriers. A premise of these systems is that geographic information is a key for overcoming the movement barrier, and essential for providing information necessary to function in our modern world.

Access to pedestrian infrastructure and public space is a critical feature for the blind, visually-impaired, and mobility-impaired. In many urban areas, the design of public pedestrian infrastructure has improved, through an awareness of important accessible design standards associated with the Americans with Disabilities Act (ADA) of 1990, which specifies limitations for running slope, cross slope, signalization, and other parameters. Useful reviews of ADA design standards for pedestrian infrastructure have been contributed by Dixon (1996), Kockelman et al. (2001, 2002), and Rodgers (2016).

A significant challenge presented in even the best-designed and most accessible environments are transient obstacles and pedestrian network problems that are not easily captured in any traditional mapping workflow. These transient or temporary problems limit accessibility significantly, re-introducing movement barriers into the built environment, and are difficult to capture in any traditional mapping workflow (Figure 33).



Figure 33 Transient obstacle impacting pedestrians, captured by the George Mason University Geocrowdsourcing Testbed (GMU-GcT)

An over-arching goal of our research work has been to develop crowdsourcing methods to quickly identify, characterize, assess, and map these transient obstacles and then disseminate them to the public, where they can be used in web-mapping applications or in public maintenance workflows. This paper extends our previous work with geographically-explicit models that use geocrowdsourcing inputs, cost functions, and selection criteria to optimize the maintenance of a pedestrian network. The starting point for identifying problems is geocrowdsourcing, which has become a key input for municipalities. The growing popularity of mobile tools such as Fix311 and SeeClickFix are evidence of the public's interest in providing information on maintenance and condition problems with public infrastructure. Rice et al. (2012b, 2015a) provides a review of popular geocrowdsourcing tools.

The following sections of this paper review the formulation and evolution of a geocrowdsourcing testbed, and provide a rationale for geocrowdsourcing as a tool for urban planning and modeling, and specifically the use of geographically-explicit optimization models for pedestrian network maintenance based on geocrowdsourcing inputs. A conceptual formulation of a family of models is followed with an implementation using a local pedestrian network and geocrowdsourcing data contributed by the public, leading to a synthesis of findings and future research directions.

Literature Review and Relevant Previous Work

We review several relevant research themes and contributions that form a conceptual framework for our work. These previous projects, publications, and contributions motivate our work. Notable motivations and themes include accessibility systems, formal modeling of pedestrian networks, geocrowdsourcing and the GMU Geocrowdsourcing Testbed (GMU-GcT), and optimization procedures and maintenance.

Accessibility Systems

In order to provide navigation and wayfinding assistance after losing vision as an adult, Reginald Golledge and research collaborators constructed and tested a personal guidance system, built with GPS, geographic information systems, auditory cues, and tactile feedback through a handheld pointer. This system was tested extensively and improved over a 20-year period, and documented in more than forty publications (Loomis et al. 2001). A weakness of this system, noted in Rice (2013a) and Qin (2015a) is the inability to provide real-time information about infrastructure changes and restrictions, which are common in dynamic urban environments. Recognizing the

importance of building a real-time information layer on top of accessibility, Nuernberger (2008) tested and reported on communication protocols via cell phone, that could be added on top of the PGS and similar systems. Barbeau (2010) and others have developed similar techniques for real-time geographically sensitive communication used by the disabled. The addition of real-time updates and alerts is now an essential part of navigation systems, with systems such as Waze combining crowdsourced incident reporting tools with navigation features. In similar fashion, the GMU-GcT adds an important layer of data to accessibility systems, by supplying real-time transitory obstacle data provided by the public through the GMU-GcT. This crowdsourced information could be added to any accessibility system with a network connection, such as Miele's Tactile Map Automated Production system, or other accessibility systems such as Beale et al. (2006) and Karimi et al. (2014).

Formal Modeling of Pedestrian Networks

Pedestrian networks are the most fundamental feature of urban systems for the disabled, as they are the primary means of access to public places. The elements of a pedestrian network are broader than just sidewalks. They include the actual walkway surface, as well as curbs and curb cuts, crossing signals, stairs, steps, ramps, railings, and crosswalks. Efforts by several researchers to carefully model the connections between the elements of a pedestrian network are noteworthy, and have been useful in illustrating the needs of blind, visually-impaired, and mobility-impaired end-users. First, Beale et al. (2006) effectively illustrates the difference in needs between able-bodied pedestrians and wheelchair users, and provide an example of the great detail and attention required to

provide meaningful routing along a pedestrian network. These issues are echoed by the work of Avila (2014) who notes the frequent request for micro-geographic data such as transitions in pedestrian walkway surface composition, which provides navigation cues to blind pedestrians in urban settings. These surface texture changes denote, for instance, the entry point into a roadway crossing. Laakso et al. (2011, 2013) create detailed pedestrian infrastructure models that allow for different levels of accessible accommodation based on user needs. Karimi et al. (2014) used detailed infrastructure data for outdoor and indoor pedestrian corridors to provide customized accessibility maps. In each case, the elements and structure of the pedestrian infrastructure are modeled in detailed form, and used to facilitate mobility. In the United States, and in the European Union (with exemplars such the Pinhel Accessibility Platform (reviewed in Rice et al. 2014), legal compliance and regulation is a driver, with design standards for accessibility published in summaries such as Russ (2009) and Steiner et al. (2012). Pedestrian infrastructure is commonly assessed for accessibility through analysis of items such as running slope, cross slope, width, curb cuts, and signalized crosswalks. Many of these items, such as running slope, are easily derived from elevation datasets, and are routinely included in accessible mapping protocols. Other items, such as cross slope, are much more difficult to measure (Rodgers 2015). Beale et al. (2006), Chen et al. (2015) and the Pinhel Accessibility Platform are notable for the degree to which they provide meaningful accessibility information and routing through careful modeling of pedestrian infrastructure.

Geotechnology and Geocrowdsourcing

The use of technology to improve quality of life in cities and for disabled citizens is not a new idea. Perkins (2002), notably, discusses the relevance of technological approaches for accessibility systems, advocating for user-centered, bottom-up, and socially-aware approaches, rather than techno-centric approaches that ignore wider social issues and views of impairment. In the context of map-based systems, Perkins suggests that “researchers should focus more on the social context of map use, and let that drive design decisions, instead of spending large research grants on often inappropriate technological solutions.” (2002, 526). Geotechnical innovations for engaging Golledge’s print and movement barriers should be contextualized with social issues, user feedback, and the needs of end-users. We keep this advice in mind, and engage end-users, subject matter experts, and the public in our activities, which is a natural extension of the crowdsourcing process.

Sui et al. (2014) describe the phenomena of geocrowdsourcing as a “profound transformation on how geographic data, information, and knowledge are produced and circulated”. Goodchild’s related concept of volunteered geographic information (2007, 2009) highlights key benefits of this approach, namely, local geographic expertise, reduced costs, rapid or continuous data updates, and improved interoperability. This important phenomenon was presaged by Goodchild et al. (2005) and others, who recognized the rapidly evolving nature of geographic information networks and the influence of social media. Elwood et al. (2012), Rice et al. (2012b), and Sui et al. (2014) provide broad reviews of the geocrowdsourcing phenomenon along with benefits, drawbacks, research considerations, and quality assessment strategies. With this context

in mind, the GMU-GcT was created to assist disabled individuals in identifying transitory obstacles in the local pedestrian network. This purpose and the goals of this system will be introduced, along with the primary research themes of this study.

The George Mason University Geocrowdsourcing Testbed

Identifying the benefits of geocrowdsourcing for enhancing accessibility, Rice et al. (2013b, 2014) presented the design of a geocrowdsourcing system for identifying, characterizing, and disseminating transient obstacle data. Their system, the GMU-GcT contains a map-based obstacle contribution and mapping system with an extensive quality assessment process (Qin et al. 2015b, Rice et al. 2016) and a gazetteer used for automatically generating spatial footprints from obstacle location descriptions (Rice et al. 2011, Aburizaiza and Rice 2016). This system is based on two of the primary benefits articulated by Goodchild (2007): local geographic expertise, and rapid updates. No other GIS or map-based data collection strategy has been identified for collecting small, but significant changes to the local pedestrian infrastructure. The formal modeling of pedestrian infrastructure, i.e., Beale et al. (2006), Chen et al. (2014), Karimi et al. (2013, 2014), and the information infrastructure associated with pedestrian networks (Laakso et al. 2011, 2013), are also critical components of our work. With these other research efforts in mind, we have developed the GMU-GcT, and modified it several times based on feedback and input. The system functions as a collection of web-based desktop and mobile applications that enable identification, characterization, reporting, moderation, tracking, visualization, and avoidance of obstacles along pedestrian networks. The system borrows concepts from the UCSB Personal Guidance System (Loomis et al. 2001),

Waze¹⁰, Nuernberger's work on mobility and communication (2008), previously cited work on pedestrian network modeling, and research from Jacobson (1998) addressing cognitive issues in wayfinding and navigation.

Operations Research and Optimizing Maintenance

Within the discipline of urban planning, Keirstead and Shah suggest that the ultimate goal of urban planning models is to ensure social, economic and environmental performance of their cities, through the “deliberate positioning of activity and transportation facilities” (2013, 175). They enumerate the past and present applications of optimization in urban planning and specify that the scale, model fidelity and meaning of optimization are major challenges in urban planning. Laakso et al. (2013), Beale et al. (2006), Rodgers (2015), and Avila (2014) note the significant geographic, topological, and structural detail required for successful pedestrian modeling. Karimi (2014) and Rice et al. (2014) demonstrate that analytical tools and geocrowdsourcing techniques are useful in developing effective accessibility systems built on top of detailed pedestrian networks. The addition of maintenance optimization, within a context of an accessibility system, is an important next step, because it recognizes that the pedestrian network is dynamic and in a state of constant change. The ability to reflect change is not easily incorporated into pedestrian network models, and is a major motivation for geocrowdsourcing approaches, such as those advocated by Rice et al. (2013a).

In this research work, geocrowdsourcing and optimization techniques are used to identify parts of the pedestrian network to monitor, repair, and maintain, with a goal of

¹⁰ See <https://www.waze.com/livemap> (accessed Nov. 17, 2016)

maximizing benefits while recognizing the inherent limits on financial, human, and technical resources used to address problems.

Statement of Purpose

Through geocrowdsourcing techniques, the GMU-GcT supplies validated information about areas of the local pedestrian network that require attention, and in many cases, repair or replacement. Local government authorities have a goal of maintaining and improving the pedestrian network, recognizing the general economic benefits of well-maintained pedestrian infrastructure (Leinberger and Alonzo 2012).

The research questions addressed here is, given a limited budget for maintenance and repair of the pedestrian network, can a spatial optimization approach be used to identify the most critical segments of the pedestrian network to repair? Is geocrowdsourcing (including surveys from subject matter experts and the public) a useful approach for identifying areas in need of repair? In order to explore these research questions, we have collected geospatial data, surveyed subject matter experts and members of the public, and built an optimization system for the GMU-GcT to explore the process of optimizing areas in the local pedestrian network that need repair. The data and methodology are described below.

Data and Methodology

Typical planning processes for sidewalk and road maintenance are ad hoc and there is no known system for determining the set of segments to repair that would provide the greatest overall value to the population being served. With crowdsourcing information about segments needing repair, obtained from the GMU-GcT, an optimal set

of segments can be determined based on their value. This section demonstrates how that can be achieved.

Consider the following scenario: An organization responsible for pedestrian facilities (e.g. a university campus facilities department, or a small city government) has a limited budget for sidewalk maintenance and improvement. They wish to decide where to allocate their resources for such maintenance in order to improve the facilities in a way which will provide the greatest overall benefit for the populace which they serve. In addition to the entire pedestrian network geometry data, this organization has determined the benefit and cost of repairing each section of the pedestrian network in need of maintenance. The benefit value is a composite measure that represents the general improvement if a segment is repaired, the usage frequency in complete pairwise routing analysis, and the priority assessment for individual segments throughout the entire network. The cost value represents the real cost of repairing a segment, which is based on material type (asphalt, concrete, brick) and segment length.

For some repair operations it is likely that greater efficiency – and probably lower overall cost – can be achieved if the segments chosen for repair are spatially associated in some way. This could mean that they are simply clustered together, or that they are contiguous to one another. The latter example is true of operations such as a lengthy sidewalk (or paving) project, where simply continuing to the next segment is more easily accomplished in contrast to moving around equipment to disconnected segments and restarting the process (Figure 34). Thus, contiguity is used as a spatial element in the family of models. However, to enforce contiguity or to encourage contiguity is a question

to be resolved during the planning phase of repair operations. In the following sections, a family of models is presented with their formulations and discussed with a comparison of their results.



Figure 34 Repairs of contiguous segments of sidewalk curbs, with traffic alteration

Family of Models for Pedestrian Network Optimization

The following family of models and formulations is presented, along with parameters and values used for optimizing sidewalk maintenance operations. The optimization models are then used with a simple pedestrian network to demonstrate their use, followed by their implementation with a full data set and obstacle information from the GMU-GcT.

Model Notation and Formulations

Consider the following notation:

i, j = the indices of junctions in the sidewalk database that need repair

B = Overall budget for sidewalk maintenance

b_{ij} = Benefit received from repairing segment ij (the segment going from junction i to junction j)

$x_{ij} = 1$ if segment ij is chosen for repair, and 0 otherwise

c_{ij} = Cost of repairing segment ij

N_{ij} = neighboring segments of segments

1) Model 1: basic model with only a budget constraint

Equation 1 Objective Function Maximize Benefit

$$\text{Max } Z = \sum_{ij} b_{ij} x_{ij}$$

Equation 2 Constraint: overall budget for repair

$$\sum_{ij} (c_{ij} x_{ij}) \leq B$$

Equation 3 Constraint: decision variable

$$x_{ij} = 0, 1$$

The objective function in Equation 1 examines the benefit for each segment multiplied by the decision variable that indicates whether or not a segment is to be repaired. That benefit will be maximized. Equation 2 considers the overall budget for repair, and constrains the segments to be repaired given their respective costs. In this

family of models, the decision variables are required to have a value of zero or one, no fractional repairs are permitted (Equation 3).

Consider the small network in Figure 35. It contains 30 segments and 23 vertices. Each segment has an associated benefit value that is the number of times that segment appears in a shortest path between two junctions. A subset of the segments has been identified as having an obstacle or being otherwise in need of repair or attention from the sidewalk authority. The segments in need of repair are highlighted in Figure 35. There is a cost associated with each potential segment repair, and the total budget for repair is known. Model 1 above can be used to identify the segment repairs that offer the greatest benefit while remaining within the overall budget. The optimal subset of repairs for this problem instance is shown in Figure 36.

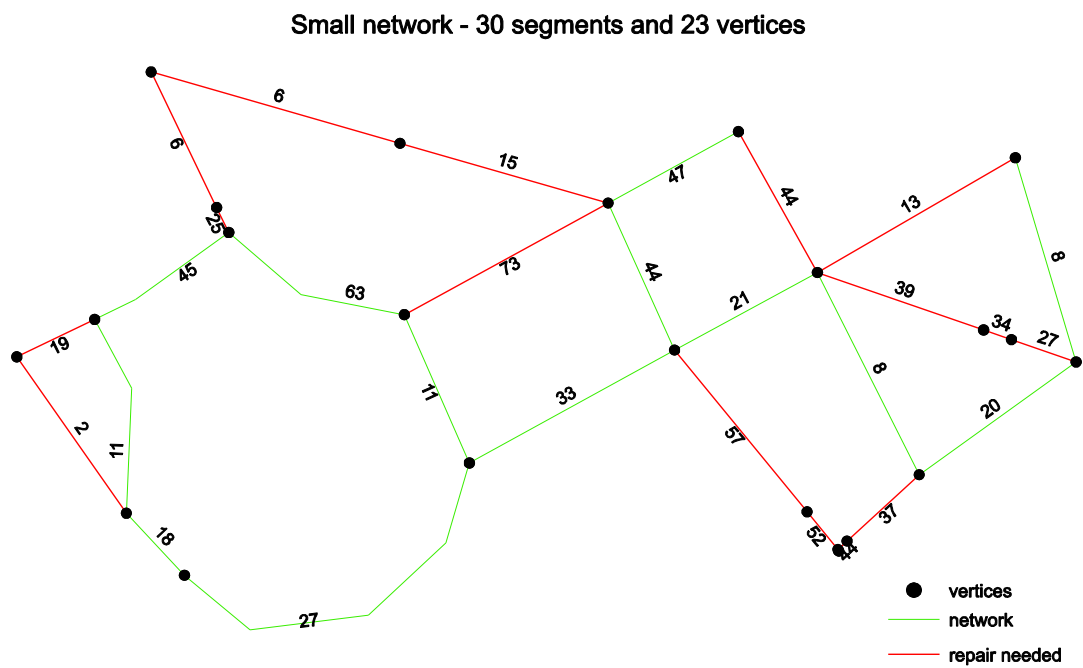


Figure 35 The segments in need of repair

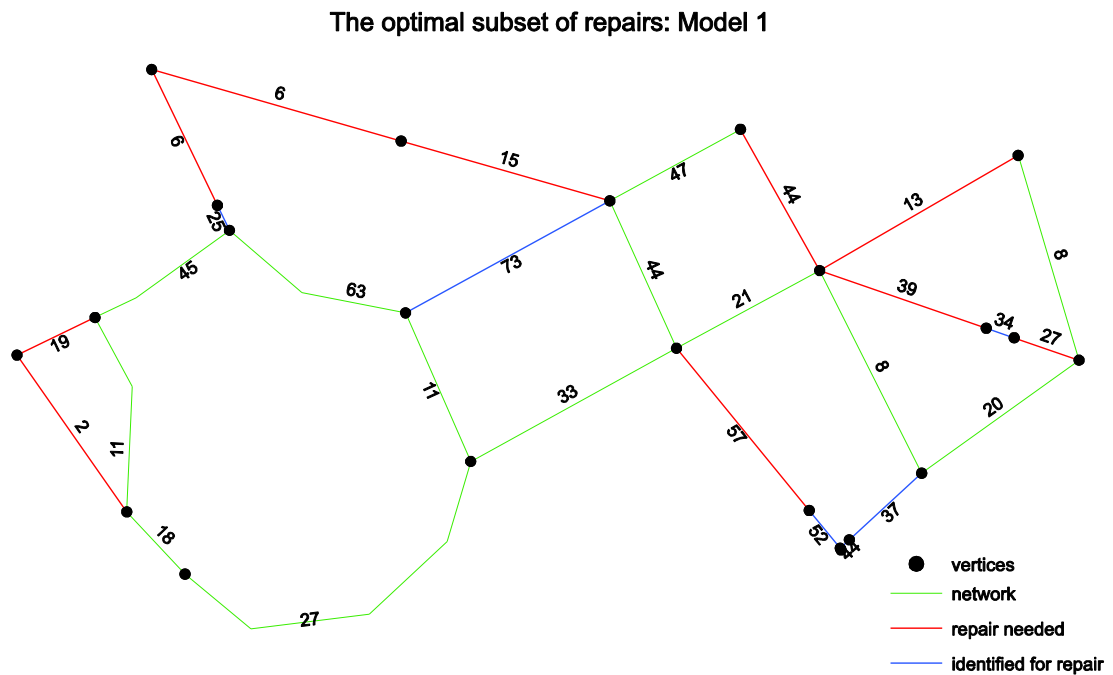


Figure 36 The optimal subset of repairs for this problem instance

In the absence of any further constraints the sidewalk authority could use this information to plan for and execute their repairs. With a problem instance of this size, the formulation can be coded in Microsoft Excel, and solved optimally with the off-the-shelf optimal Data Solver add-in.

In some cases, repair operations may be more efficient, and done at lower cost, if the segments selected are spatially associated. This might be the case with small segments that are clustered near each other, or segments that are contiguous. For segments that are contiguous, efficiency is increased by repairing them at the same time, with equipment and materials already deployed to the same location. Repairing single segments and then relocating equipment to repair a disconnected or distant segment can

be inefficient. Public works employees report that cost savings can be achieved when materials (such as concrete) are ordered in larger amounts and deployed to a single location, which is possible when selected segments are contiguous. The models below expand on Model 1 to encourage spatial association.

2) Model 2: enforce contiguity constraints in addition to a budget

Contiguity can either be enforced with constraints, or it can be encouraged with variable costs. To explore the idea of requiring contiguity we introduce the notion of the neighborhood set of each segment. The neighborhood set of any segment i,j , is comprised of all segments that are connected to segment i,j . For example, in Figure 37, the neighborhood set of the segment extending from node 11 to node 19 (11,19) consists of the segments (21,19), (22,19), (15,19), (8,11), (16,11) and (6,11). Given that neighborhood relationship, we introduce Equation 4 which enforces contiguity.

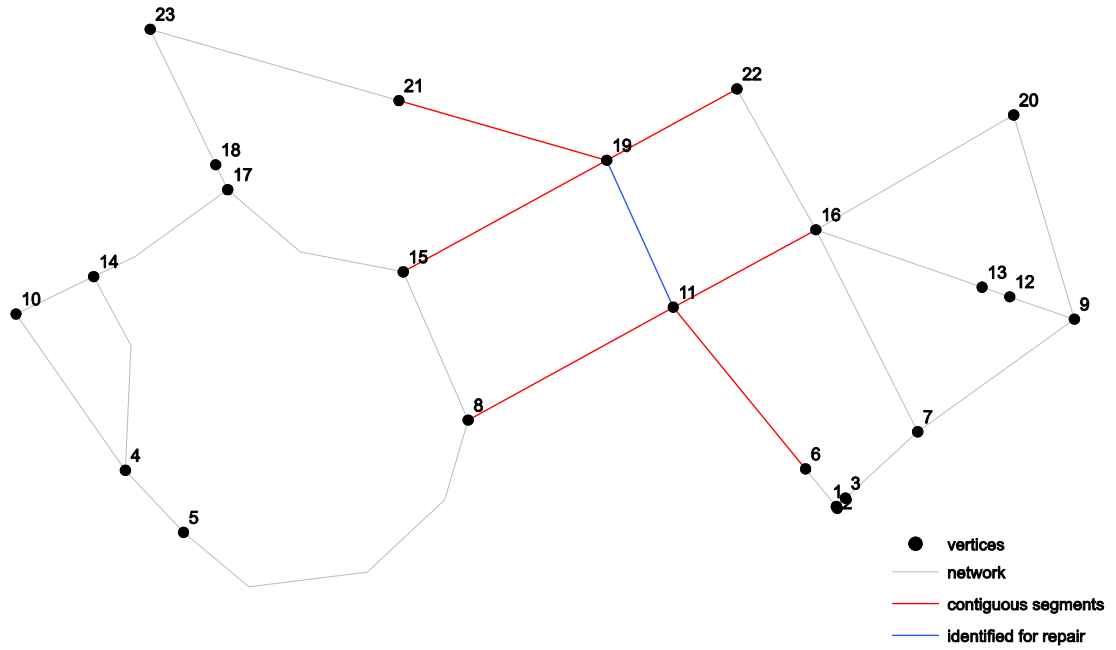


Figure 37 Concept of the neighborhood set: N_{ij}

Equation 4 Enforced contiguity constraint

$$x_{ij} \leq \sum_{N_{ij}} x_{ij}$$

With this constraint, a segment can only be identified for repair ($x_{ij} = 1$) if one of the segments in its neighborhood set is also identified for repair. That is, no single unconnected segment can be included in the optimal repair solution. Figure 38 shows the optimal solution on the small test dataset when contiguity is enforced. The overall benefit of the solution decreases from 265 to 251 (5.3%), which is the sum of the usage frequency of the segments being repaired. However, the contiguity can bring extra benefit for the road maintenance in the real life. For Model 2, the real benefit could be larger

than 251, while there are more potential parameters need to be considered with specific situations. For instance, additional savings will be captured in time, effort, and materials transportation costs, which are not included in this model, but are benefits of spatial association noted by local public works employees. At the present stage, Model 2 illustrates a different pattern by enforcing a spatial element.

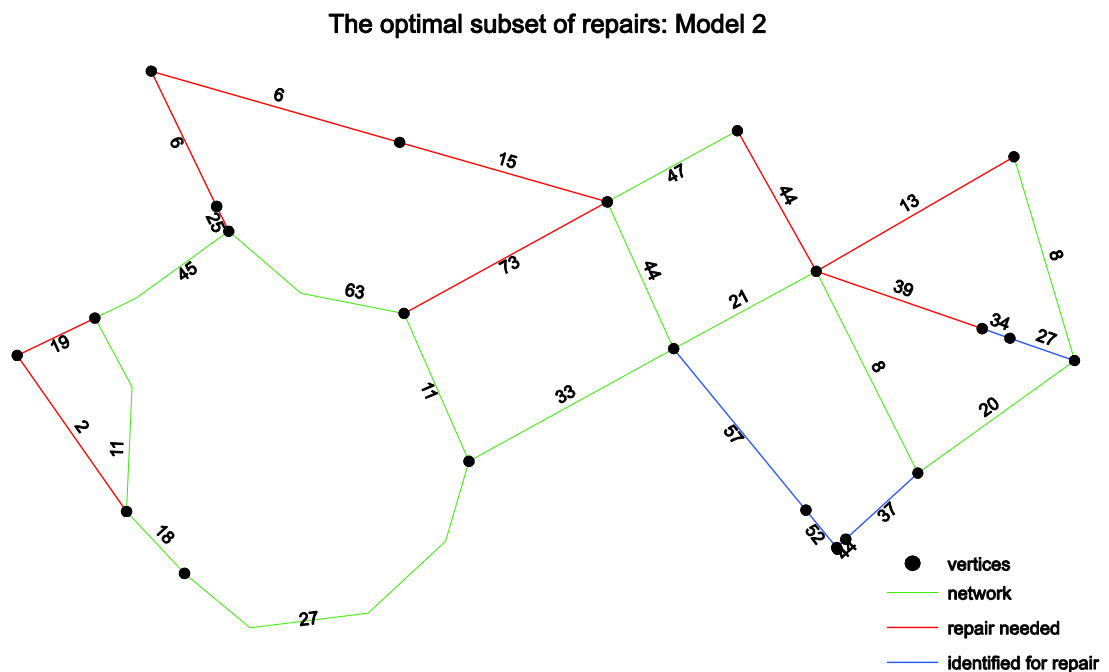


Figure 38 Small dataset solution with contiguity enforced

It may well be that a compromise solution (compromise between reduced benefit and reduced cost) can be obtained if contiguity is encouraged but not enforced. Consider Model 3.

- 3) Model 3: encourage contiguity in addition to a budget and extra benefit is offered to encourage contiguity.

Rather than requiring that all edges that are fixed be connected to another edge being fixed (essentially requiring a series of connected repair paths), it is possible to encourage contiguity by adjusting the benefits for selecting contiguous segments. In order to do so, we introduce two new pieces of notation:

e = a contiguity efficiency parameter

a_{ij} = a contiguity indicator where $a_{ij} = 1$ if x_{ij} has a contiguous neighbor being repaired and 0 otherwise

We define Equation 5 to determine the values of a_{ij} :

Equation 5 Constraint: determine the values of a_{ij}

$$a_{ij} \leq \sum_{N_{ij}} x_{ij} \text{ for all } ij$$

and we modify the objective function to consider both the benefit from fixing a given edge ij , and an additional benefit if ij has a neighboring edge also being fixed:

Equation 6 Modified objective function

$$\text{Max } Z = \sum_{ij} b_{ij} x_{ij} + \sum_{ij} a_{ij} b_{ij}(e)$$

It is presumed that the contiguity efficiency parameter would – for example – be a value between 0 and 1 that would allow a fraction of the benefit to be included in the objective function value. For example, imagine that the benefit (b_{ij}) for a particular edge ij was 100, and the contiguity efficiency parameter (e) was set to 0.2. If both ij and an edge contiguous to ij were selected to be fixed, then both b_{ij} (100) and $b_{ij} \cdot e$ (20) would be added to the objective function. Thus, contiguous solutions will be encouraged when they are available, while still allowing standalone segments to be fixed up to the limit of the budget constraint. With this additional information, the full Model 3 becomes:

Equation 7 Objective function for Model 3

$$\text{Max } Z = \sum_{ij} b_{ij} x_{ij} + \sum_{ij} a_{ij} b_{ij}(e)$$

Equation 8 Constraint: overall budget for repair

$$\sum_{ij} (c_{ij} x_{ij}) \leq B$$

Equation 9 Constraint: decision variable

$$x_{ij} = 0,1$$

Equation 10 Encouraged contiguity constraint

$$a_{ij} \leq \sum_{N_{ij}} x_{ij}$$

Equation 11 Constraint: decision variable
 $a_{ij} = 0,1$

Figure 39 shows the optimal solution on the small test dataset when contiguity is encouraged.

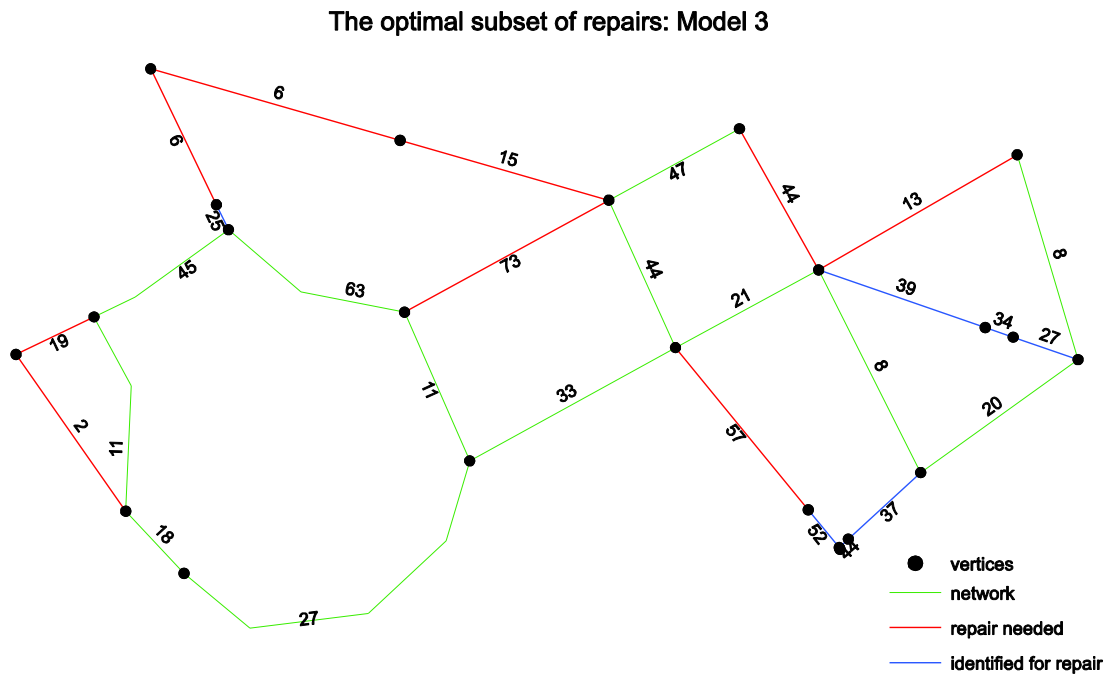


Figure 39 Small dataset solution with contiguity encouraged

Full Dataset from GMU-GcT

The full datasets in this study include the entire pedestrian network of Fairfax City and the GMU campus dataset used in GMU-GcT; and priority, usage frequency, and material type of each segment. The priority dataset is generated through the combined input of a wide group of subject matter experts from the Fairfax City government, Smart

Growth advocates, GMU Facilities and Maintenance Department, and students from GMU. Usage frequency is calculated with a complete pairwise routing across the entire pedestrian network, with each segment receiving counts based on the number of times they are used in a shortest-cost path between an OD pair. The most commonly used pedestrian segments therefore have the highest usage frequencies.

The pedestrian network for the Fairfax City and GMU campus area has 3592 segments, with 2878 junctions with 155 segments in need of repair for various reasons (e.g. sidewalk obstruction, construction detour, or poor surface condition). An example of a report identifying the segments in need of repair, gathered through the GMU-GcT crowdsourcing portal, is shown in Figure 40. While processing the model, the computational requirements exceeded the capacity of the Data Solver add-in in Excel software, and therefore more robust optimization software must be employed. While there are a range of options for formulating and solving linear programs, in this case an open-source option called OpenSolver, was employed.

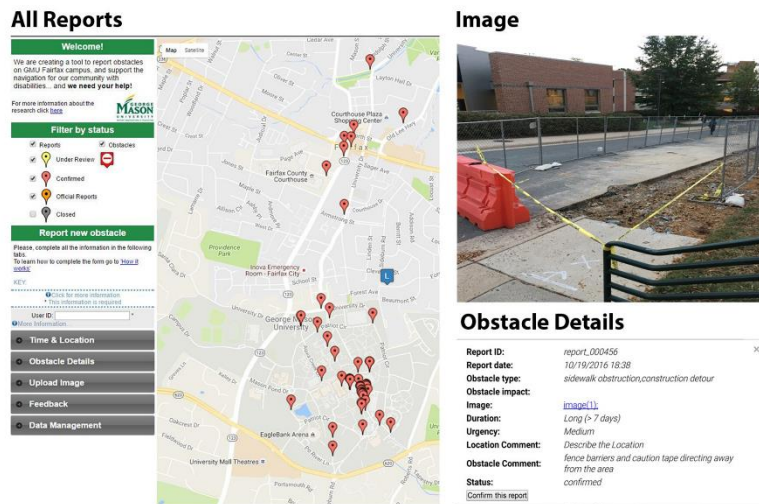


Figure 40 Obstacle Report with Details, from the GMU-GcT

According to the estimated cost and budget information from municipal public works and transportation engineers, the repair cost of each pedestrian segment in the study area is dependent on the materials and composition. As a guide, the following material price list was used, after conversion from square feet to square meters: brick sidewalk, \$322.92 per square meter; concrete sidewalk, \$161.46 per square meter; asphalt sidewalk, \$107.64 per square meter. In a particular scenario, the budget for repair or replacement of sidewalk surfaces in this study area is \$287,750.48, which is based on a standard allocation of \$1 per square foot per year or equivalently, \$10.76 per square meter per year. An aerial survey of the pedestrian network from high-resolution imagery resulted in our use of an average width of 1.83 meters for sidewalks in our study area, which allows the generation of a cost value for each segment, which is calculated by material price*width*length of segment. The benefit value used in this study is generated with a linear equation combining three variables: 1) priority assessments for each

segment from subject matter experts, 2) usage frequency determined from segment use counts during an exhaustive pairwise routing across the network, and 3) an urgency value, derived from geocrowdsourcing contributors, who provide an estimate of the seriousness of needed repair and maintenance. These three benefit values recognize the many different contributions to the value of a segment and the associated benefit of its repair.

Results

First, a preliminary result is obtained by identifying the segments in need of repair without considering contiguity (Model 1), which serves as a baseline for other comparisons (Table 2, Figure 42). Figure 41 and Figure 42 show the entire pedestrian network with the set of segments in need of repair, and the segments identified as providing the optimal total benefit if repaired under a budget constraint. This solution identifies 77 of 155 segments for repair, and demonstrates that under specific budget constraints, a small municipality can use geocrowdsourced data, its own GIS database, and open source solution software to prioritize their sidewalk repair operations.

GMU-GcT dataset: inaccessible segments

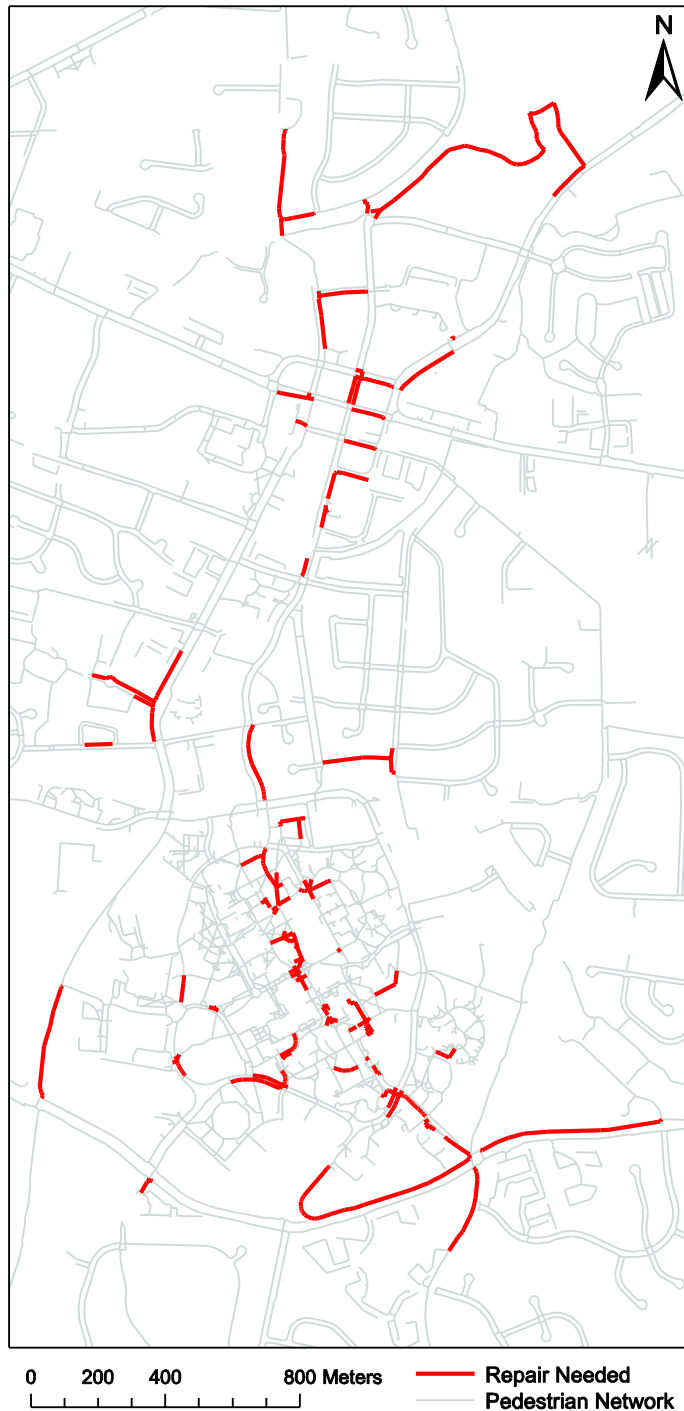


Figure 41 Pedestrian network with the set of segments in need of repair

Optimal set to repair under budget constraint

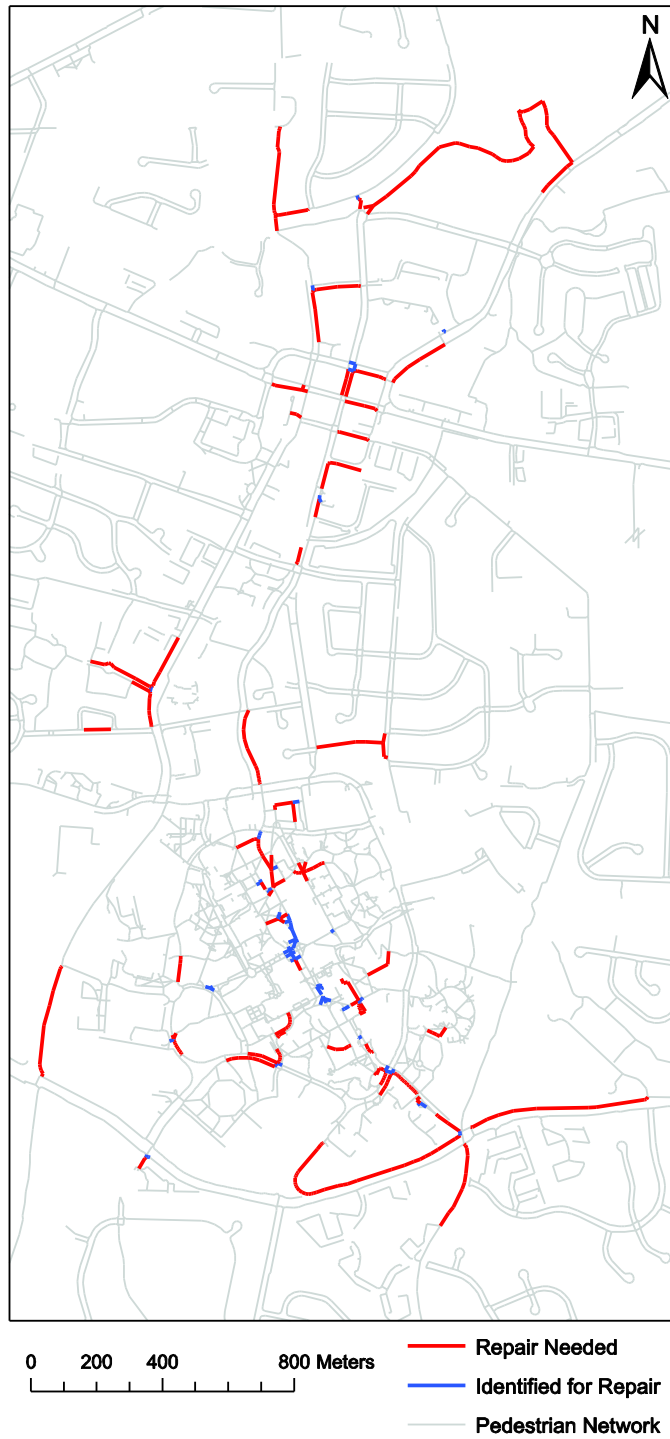


Figure 42 Result of Model 1

Next, additional solutions are generated by integrating a spatial element – contiguity. The contiguity constraint (Model 2) can be applied to the same dataset from the GMU-GcT. The result is presented in Figure 43, and summarized in Table 1. In this case there is a decrease in the number of segments repaired (from 77 to 74), and the resulting benefit achieved – the overall benefit of the solution decreases from 1366.01 to 1306.53 (4.35%). The complexity of the sidewalk network is such that not all the segments with identified obstacles or other needed repairs are explicitly contiguous. In this case, the initial budget is limited to fix all the contiguous segments, thus, the overall benefits of Model 1 and Model 2 have a slight difference as well as the total cost amount. However, there exist equivalent patterns from these two solutions, which is the key point of this family of models.

Contiguity is enforced with constraints

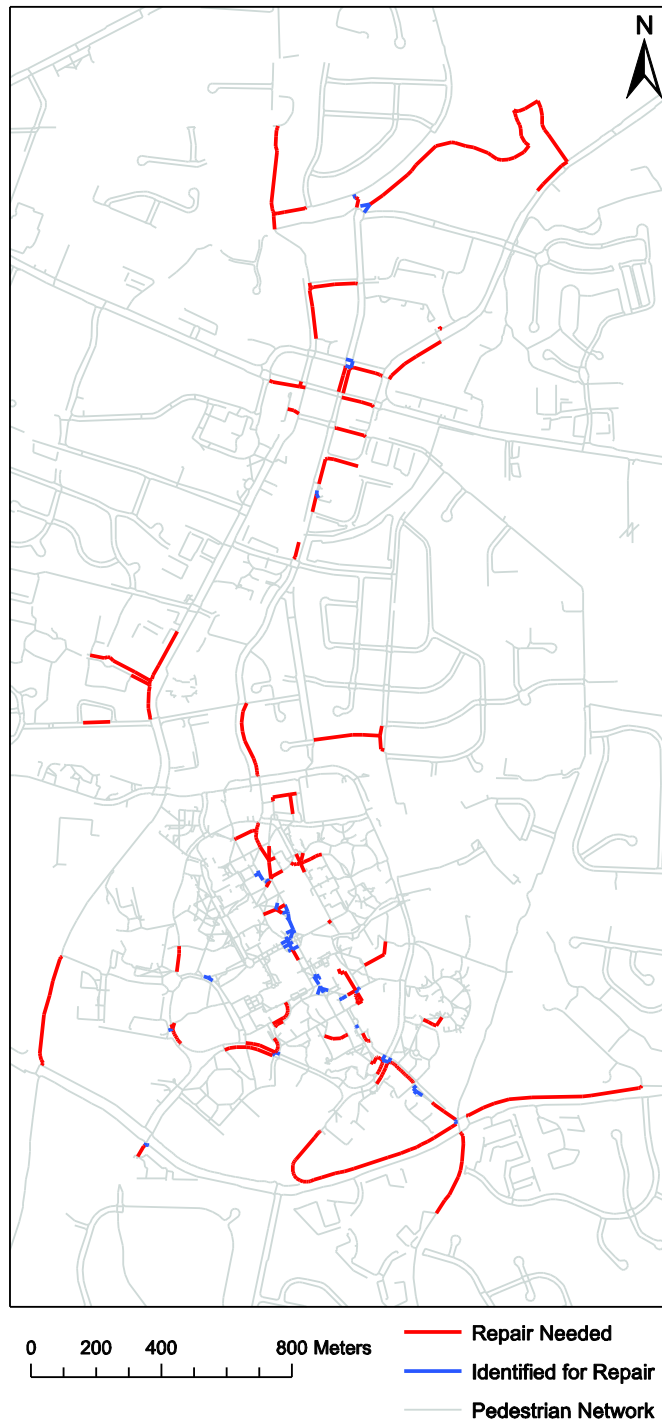


Figure 43 Result of model 2

Finally, contiguous encouraged solution (Model 3) is applied to the same full dataset from GMU-GcT, while still allowing standalone segments to be fixed up to the limit of the budget constraint, and the final result is shown in Figure 44 and Table 2. The total number of identified segments is 75 of 155 total segments.

Contiguity is encouraged with constraints

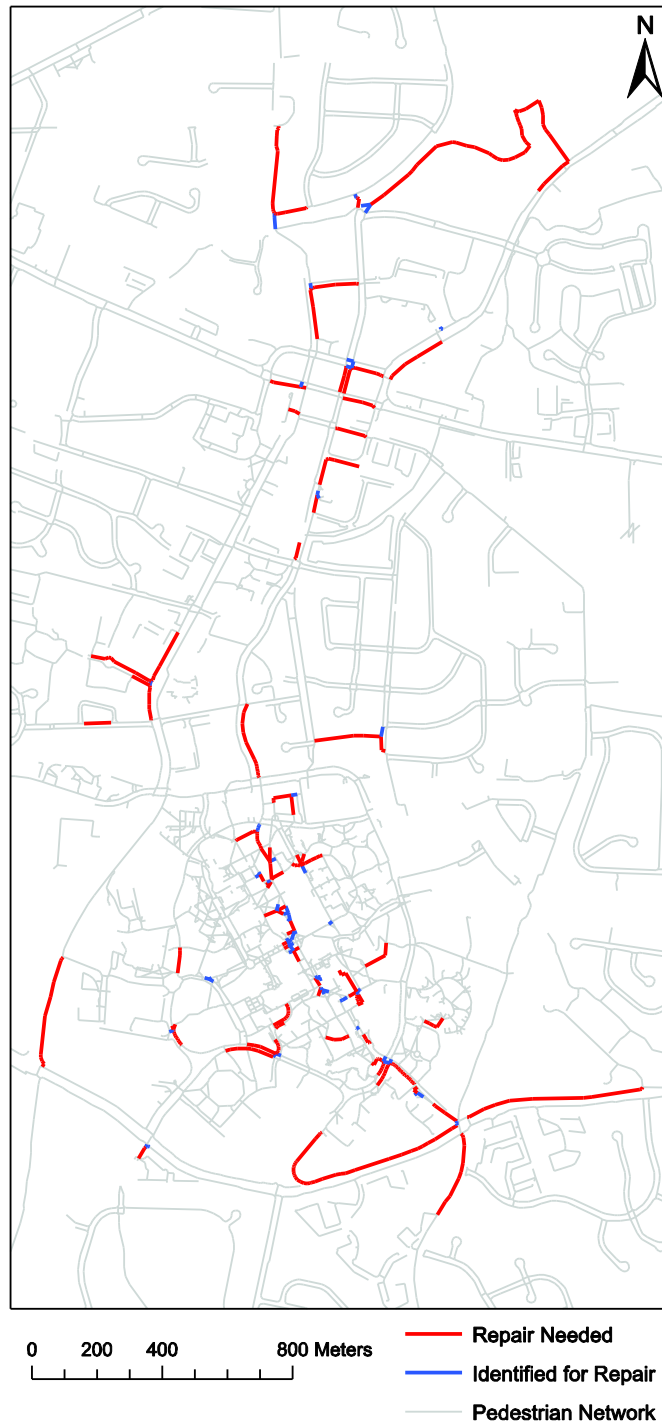


Figure 44 Result of Model 3

Comparisons Among the Family of Models

Table 2 Results of different models summarizes the results of different solutions generated with the family of models. There are 62 equivalent segments in these three models (Figure 45), indicating general similarity in results. While the numbers among these results are quite close, there are significant differences in the spatial arrangement and distribution of segments to be repaired under the solutions, as seen in Figures 41-45.

Table 2 Results of different models

	Model 1	Model 2	Model 3	Equivalent segments
Total number	77	74	75	62
Total length (meter)	1040.77	1029.07	1040.91	729.27
Total cost (USD)	287648.96	287568.71	287556.36	199044.54

Same identified segments with different models

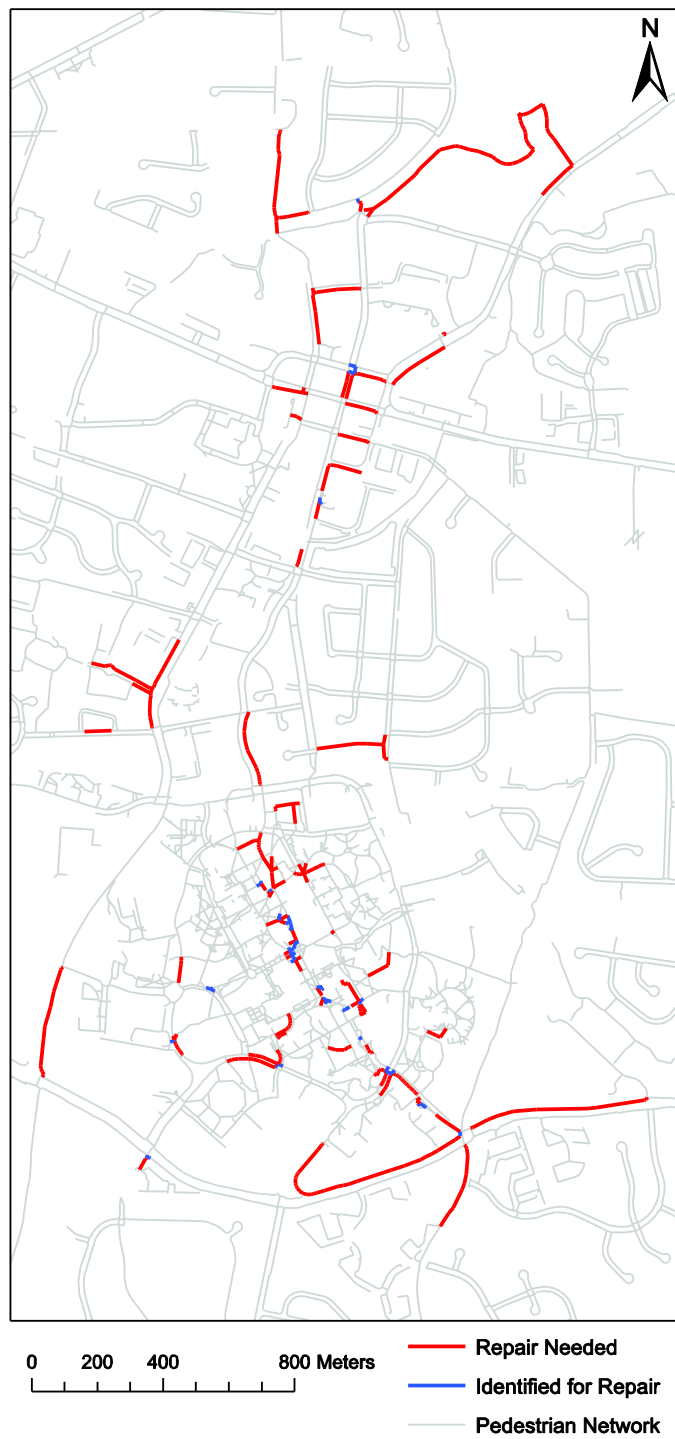


Figure 45 Equivalent segments in these three models

Discussion

There are several important findings of the work presented above. First, it has been demonstrated that a family of models can be generated for determining an optimal allocation of repair resources in order to maximize the benefit to the pedestrian population from those repairs. The family of models is predicated on the input of geocrowdsourcing data that identifies areas needing repair, as well as certain prioritization information gathered from subject matter experts and from the public. This work presented models that 1) would maximize benefit with only a budget constraint, 2) would maximize benefit and enforce contiguity constraints in addition to a budget, and 3) would maximize benefits while encouraging – but not enforcing – contiguity. It was also demonstrated that both models 1 and 2 could be run consecutively when the enforcement constraint dominated the budget constraint. While this family of models could clearly be expanded further, these example models show that real-world problem instances could be solved using optimization techniques in order to most efficiently use limited repair resources. In a larger sense this shows a benefit of using crowdsourcing techniques. The data collected through crowdsourcing informs a technical process that can lead to real-world decision-making.

One benefit of developing a family of models, is that these models – as with any scientific model – are a simplified expression of a real-world problem. With variant models one can solve similar, but not identical problems, with different assumptions/constraints and interpret any differences or similarities in the results. In the work presented above, it was noted that a significant number of the segments in need of repair were identified using consecutive enforcement and selection models, and the

contiguity encouragement model. This suggests that robust solutions could be generated by running multiple models and choosing segments that appear in the solutions to many or most of those different models.

While there are many possible ways in which this family of models could be expanded, some approaches come quickly to mind. It could be that particular paths should be targeted (rather than segments). Therefore, a model that defined (a priori) important paths, and assigned a benefit to selecting segments along those paths could produce a greater overall benefit to the pedestrians using those paths. Subject matter experts and members of the local transportation and smart growth communities advocated for such an approach, which would lead to a focus on collections of segments with a similar function, i.e., segments connecting the local bus stops to the public library.

Summary

Geocrowdsourcing is a valuable tool for mapping, and represents a major transformation in the way geospatial data is collected. A particularly important advantage in geocrowdsourcing is rapid and continuous update of geospatial data that otherwise would be mapped in a series of slower, discrete updates cycles (Goodchild 2007). In the GMU-GcT, the identification, characterization, assessment, and display of transient obstacle data has helped improve accessibility for members of the public, including the disabled, who are adversely impacted when a preferred route is inaccessible. The generation of quality assessment measures and data collection strategies within the GMU-GcT, and the general scientific implications have been addressed in Qin et al. (2015b) and Rice et al. (2014, 2015a). This paper extends this previous work by

introducing a family of optimization models that are able to quickly identify areas of a pedestrian network to repair. These models begin with a detailed pedestrian network, and with geocrowdsourcing inputs, that is, a set of publicly-generated reports about sidewalk segments that need attention and repair. The models also use prioritization information and usage frequency, which determine the benefit associated with the repair of a specific pedestrian network segment, with segments having high intrinsic value (due to priority and/or usage) generating higher benefit if repaired. The models also use cost factors, related to sidewalk surface material, and a fixed budget, which constrains the number of segments that can be repaired. The parameters of the models and the underlying data are all based on real geographic data and information volunteered by local planning, transportation, and public works supervisors. Within our analysis, we generated a set of solutions using these models that identify segments for repair based on a fixed budget. In certain cases, a contiguity requirement can be useful, and in other cases an encouragement for contiguity can be used to select areas for repair. This approach, based on real circumstances and data, shows the feasibility and value of including geocrowdsourcing as a data input for prioritizing maintenance and repair, and a family of optimization models that can focus the maintenance and repair in a way that maximizes benefits within a budgetary constraint. Future work will address the scalability and computational complexity of the optimization models, as they move from a study area of this size to a much larger city or geographic region, and methods for differential weighting geocrowdsourcing inputs based on contributor reliability.

CHAPTER SIX: QUALITY ASSESSMENT IN GEOSOCIAL MEDIA

Over the past ten years, social media data has been used extensively to enrich and augment traditional sources of geographic information used for decision-making. Li and Goodchild (2010) state social networks can play important roles in effective emergency management - “information generated and disseminated over social networks is incredibly valuable for disaster response” and “the study of the relationships, behaviors, and interactions in social networks may provide important insights for gathering information, planning evacuations and sheltering, and other rescue efforts”. Stefanidis et al. (2011) suggest “as the popularity of social media is growing exponentially we are presented with unique opportunities to identify and understand information dissemination mechanisms and patterns of activity in both the geographical and social dimensions, allowing us to optimize responses to specific events, while the identification of hotspot emergence helps us allocate resources to meet forthcoming needs”. Xu et al. (2015) emphasize “the social media approach for studying human mobility is far more data rich than traditional methods”. Han and Tsou (2016) compare the tweets to Landsat population estimates by testing validity of Twitter data for predicting the number of population. Social media data are also known for their extremely large volume, high velocity, wide variety, but easy access. Like most social media platforms, Twitter provides open application program interfaces (APIs) to allow people to archive streaming

or historical tweets, with many different characteristics, depending on the search methodology. There are an estimated 500 million tweets per day¹¹, but a small percentage of them is georeferenced in some way. In the study of Stefanidis et al. (2011), there is approximately 16% of the collected twitters with coordinates, while another 45% of those tweets had locational information at coarser granularity (e.g. city level). In this dissertation, the author archives millions of steaming tweets with geolocation restraint from January 2014 to June 2015, these tweets are generated by a large number of twitter users, with a wide variety of backgrounds, interests, and styles. For example, there are 5,974,723 geotagged tweets from Feb 07 to Feb 28 with 281891 different user IDs in a selected area (Bounding box: -77.33984726850588, 38.80382953366614, -77.28294152204592, 38.86360187983721). These studies can show the significant benefits from social media data with the advancement of web technologies and wide popularity of mobile devices in our daily life, such as easy access, frequent update, cost savings, and accessible location information (e.g., as GPS coordinates). However, social media data is contributed by the public (most of whom are not experts in geographic domains), and several quality problems emerge while data quality assessment is the largest challenge to be addressed in making social media data fit users' needs.

In this chapter, quality assessment of geosocial media data is discussed through a review of the most widely used quality assessment variables (Guptill and Morrison 1995, Veregin 1999, Girres and Tuoya 2010), and an analysis of a large sample of tweets. Many of the quality assessment parameters discussed in previous chapters of this dissertation

¹¹ <http://www.internetlivestats.com/twitter-statistics/>

(3, 4, 5) need to be significantly modified or changed to have relevance in the quality assessment of social media data. For instance, positional accuracy can be addressed in a number of ways, including the accuracy of the georeferencing of an individual tweet or twitter-user's location, as well as the spatial mean center of a cluster or tweets with a common space-time-attribute connection. This chapter includes work published by the author in CaGIS (Xu et al., 2015), and presented as a part of an NSF-sponsored I/UCRC research project titled "Visualizing Data Quality in a Crowdsourcing Environment". The examples, figures, and results of these two research efforts are summarized as they related to the quality assessment parameters discussed in previous chapters, which are based on the items shown in Figure 3.

"Visualizing Spatiotemporal Trajectories of Mobile Social Media Users Using Space-Time Cube"¹²

The use of social media in GIS domain enables users to actively report their activities with geographical locations. In this paper, the co-author (Han Qin) uses a visualization tool, named CommonGIS¹³, to illustrate active users' mobility by using space-time cube in this study. The individual in Figure 46 represents two major geographical locations with several sporadic activities at other locations. The second individual in Figure 47 posted tweets at one geographical area intensively. This paper provides several initial visualization results, and also points out an interesting topic – the fundamental differences between traditional top-down approach and bottom-up approach. Traditional top-down approach refers to the use of contextual information in visual

¹² Authors: Chen Xu, Han Qin, Manzhu Yu; <http://dx.doi.org/10.1080/15230406.2015.1059253>

¹³ <http://geoanalytics.net/eda/software.html>

perception, and the bottom-up approach is well-known as the data-driven processing. This paper focuses on the individuals, and the backgrounds of different individuals affect the spatial patterns of their footprints directly. Although the challenges of assessing the dataset are not discussed in this paper, the study with more quantifiable depictions of spatiotemporal patterns is required to depict georeferenced social media data. In conclusion, while standard procedures don't work well for geosocial media data, the relevant data quality assessment methods can be developed by considering *usage*, *lineage* and *the quality of contributors*, which refer to Figure 3. In the next section, a framework for social media big data quality analysis (Hajjar et al., 2015) is discussed and several approaches for improving the data quality of Twitter, which is a typical geosocial media data, are presented.

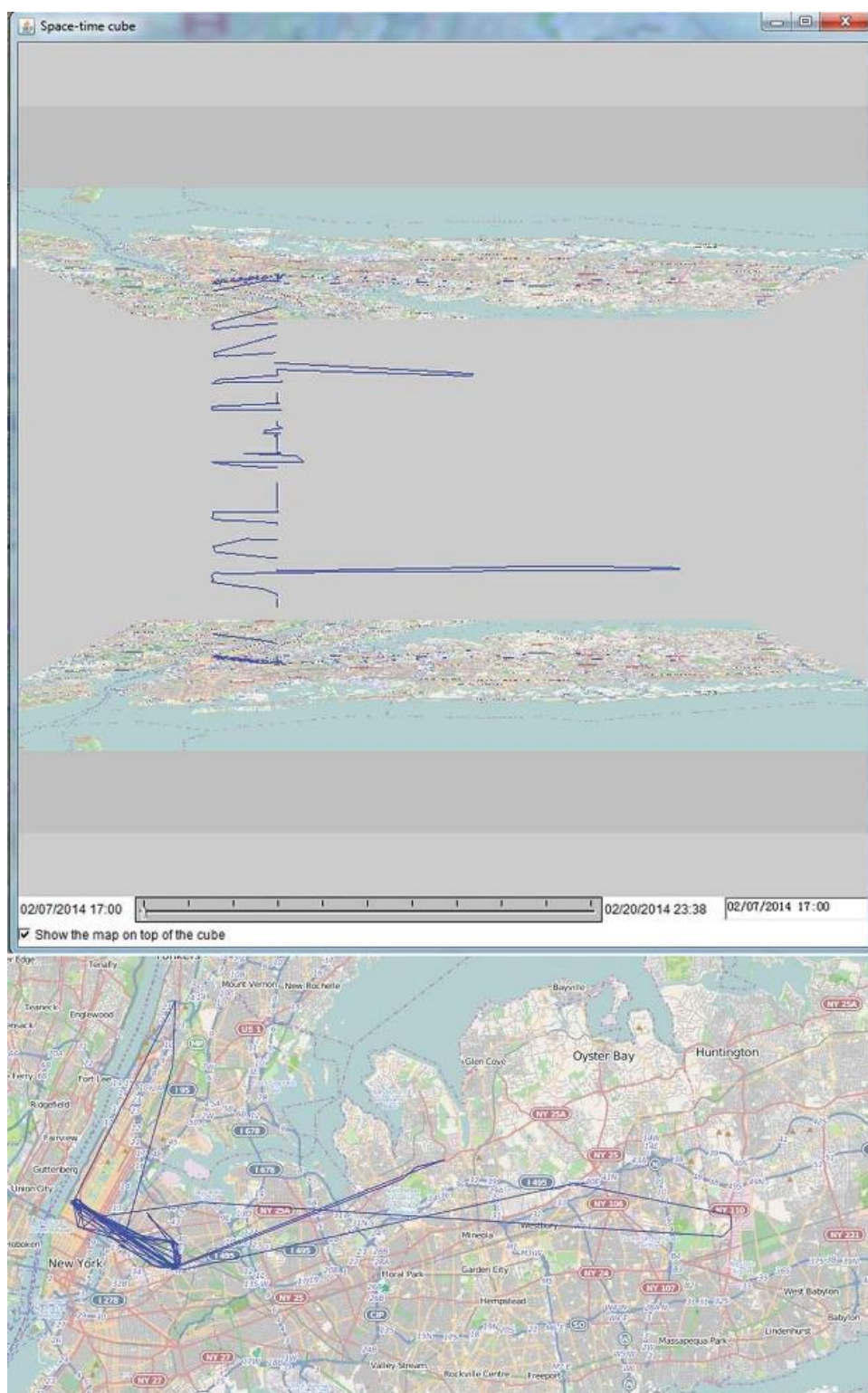


Figure 46 The space-time trajectory of a Twitter user



Figure 47 The space–time trajectory of another Twitter user

A General Framework for Quality Analysis on Social Media Network

Accuracy is an important aspect of data quality, but fitness for use is also a necessary aspect for users. Thus, many social media networks have considered about different factors of data quality which could improve the degree of satisfaction for end users. Filtering or pre-processing data is common method to find the related information from massive social media data. For instance, the real-time push function is a typical data processing method to deliver ‘useful’ information to end users, based on user behaviors, hot events, or other filter conditions. In Figure 1, we know that ‘errors’ or ‘misleading results’ could exist during data processing, and they may affect decision making. Hajjar et al. (2015) discuss data information from popular social media sites, such as Facebook, Twitter, Flickr, LinkedIn, and summarizes quality issues of several technologies that are used to capture and process those data (Figure 48). This paper summarizes “an overview of quality issues of techniques used in the capture, analysis and processing of big data on social media” (Hajjar et al. 2015), and discusses the suitability of using different techniques to satisfy different quality factors. In addition, this paper suggests that when all quality factors in the Figure 48 are satisfied, data quality challenges can be better controlled or addressed by these social media providing companies.

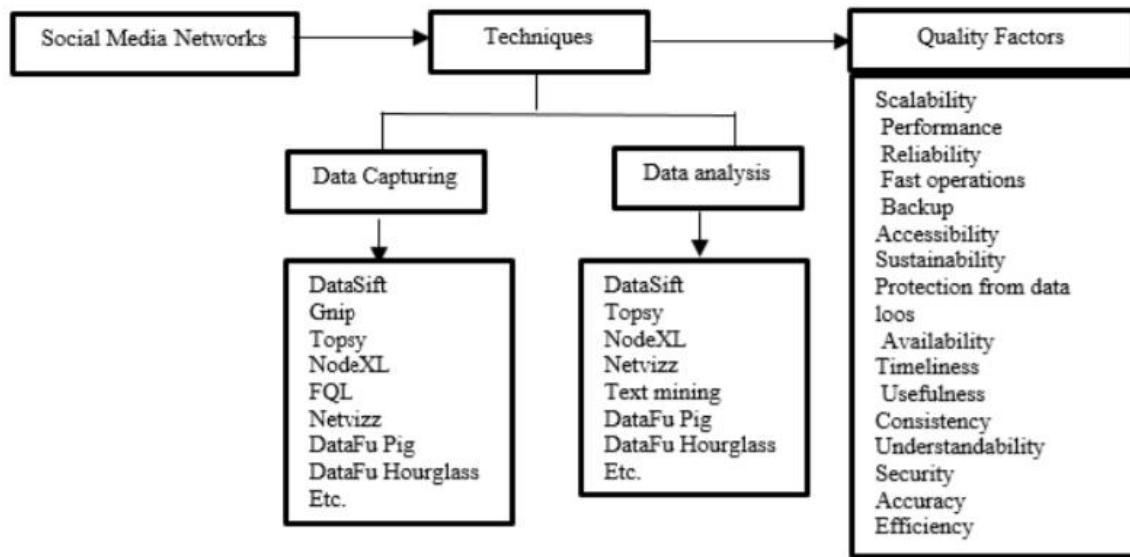


Figure 48 Framework for Quality Analysis of Big Data on Social Media (Hajjar et al., 2015)

To insight into geosocial media data, the author (Han Qin) uses Python scripts to download the streaming geotagged tweets in New York Area and detects different footprints of individuals with coordinates information in the archived data. Removing incorrect, inaccurate or irrelevant parts of the data is another way to detect the high quality parts of the data. Thus, two novel approaches – distinguishing active or inactive users within a specified geographic activity space and detecting unusual patterns are used to filter and improve the quality of raw dataset. Figure 49 includes point patterns (all activities are at a single location), polyline patterns (activities are at two locations) and polygon patterns (activities are at three or more locations).

In New York City area, users with a point activity pattern and users with polyline activity patterns are analyzed as a preliminary filtering step in the process of determining usefulness. Due to current dynamics associated with mobile devices and twitter use, a

primary indicator of a less useful or spam-related social media user is a spatial footprint that does not move at all, or is inconsistent with the topical and content-based analysis of their tweets. Research contributors analyzing event-based tweets in New York City often note that high volume twitter accounts with unusual point locations are frequently associated with computer bots and other less useful sources, where frequent tweets about the Occupy Wall Street movement include a geographic location on the ocean (Wayant et al., 2012). Author observes that many user accounts with simple point-based spatial footprint close after a short time period of activity. In contrast, long-term, high volume twitter users with a simple point footprint are sometimes tweeting in an official capacity from a workplace, and are therefore useful sources of information. For long-term, high-volume twitter users with a simple point or polyline-based spatial footprint, additional content analysis is required to determine whether inconsistencies in location are present, which requires a gazetteer-based geoparsing approach. Users with polygon activity patterns are numerous, and filtering by polygon areas is a good preliminary filtering step to identify high-value user accounts. Users with large polygon spatial footprints seem more likely to be high value, active social media contributors. Thus, an active user can be assumed that the area of spatial pattern is larger than a certain value.

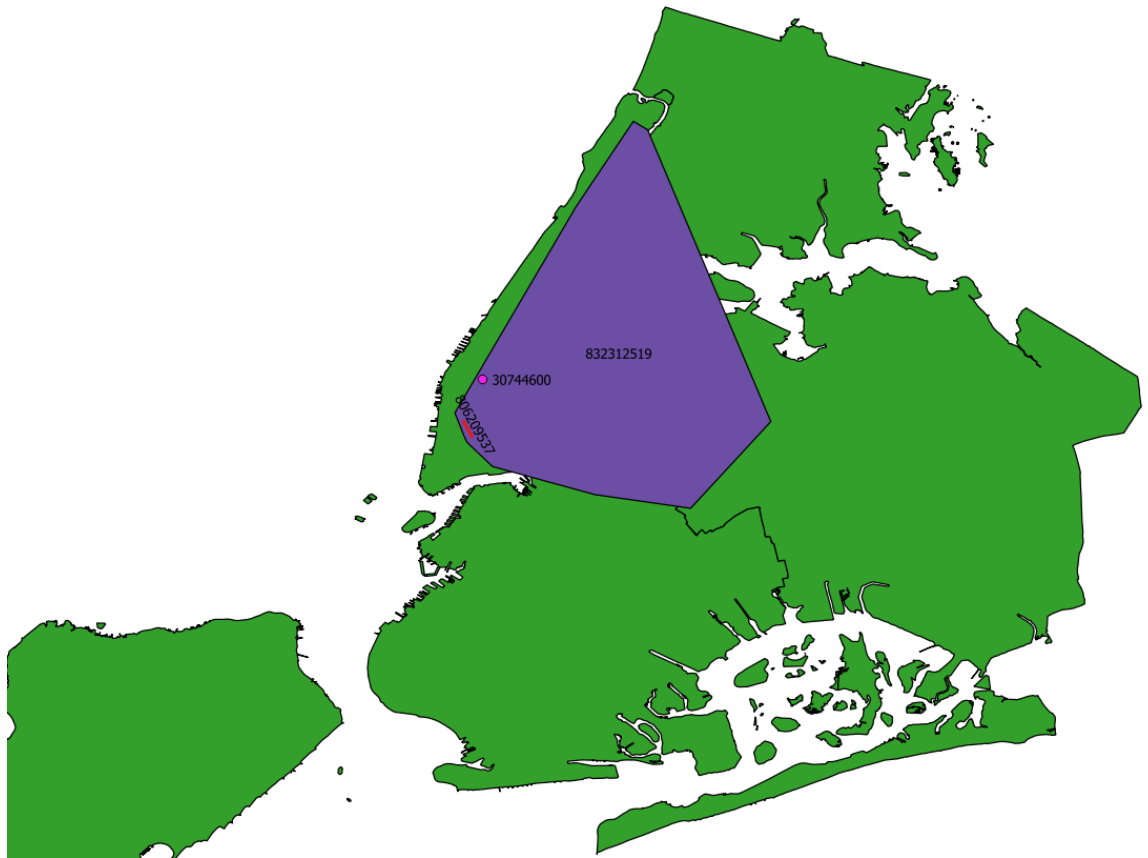


Figure 49 Point, polyline and polygon-based footprints for high volume users with user-ids

In terms of unusual individual footprints, Figure 50 represents a gridded/regular footprint. This user account is reporting vehicle identification number. Locations are constrained on a regularly spaced lattice and clearly the result of some type of artificial aggregation process, similar to other ‘geo-privacy’ masking done with crime reports. This approach which looks at the identification of unusual point patterns in georeferenced tweets provides an indicator of unusual activity and possible quality problems.

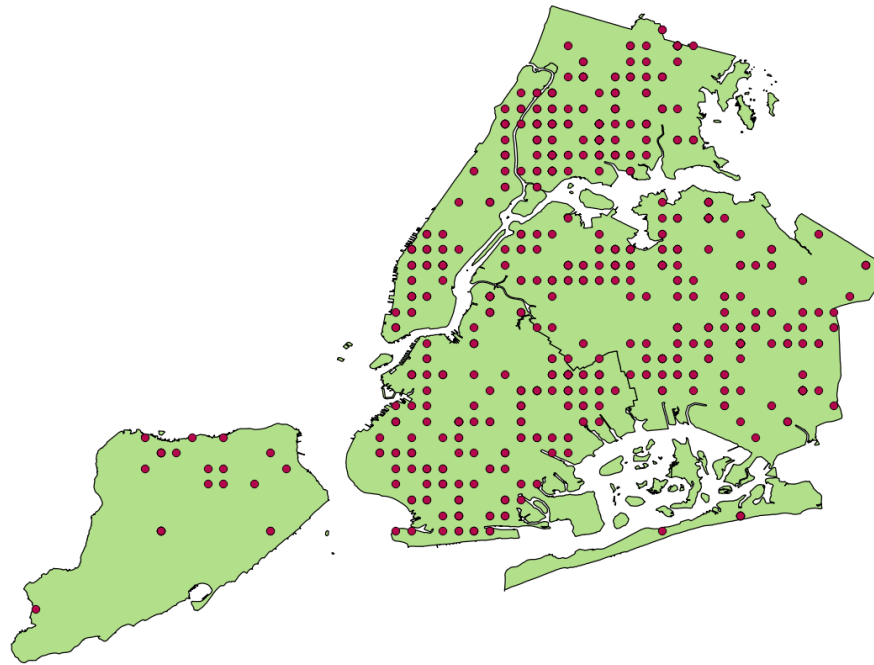


Figure 50 Unusual footprint: Gridded footprint

The chapter provides some guidance on quality assessment methods for geosocial media data. The use of gazetteers for geoparsing unstructured social media is a fruitful area of work. The issue of scale within gazetteers and the applicability of local POI names and are relevant topics that would be useful in advancing the use of geosocial media data. Detailed gazetteers could be used, for instance, in monitoring the status and circumstances of public facilities. Exemplar projects and computational strategies should be explored to create useful, real-time data visualizations of geosocial media, including elements of content and sentiment analysis, which are important in identifying high-value and low-value elements in the social media corpus.

CHAPTER SEVEN: CONCLUSIONS AND FUTURE RESEARCH

To identify geocrowdsourcing information, the author (Han Qin) developed a testbed to archive and map geocrowdsourced obstacle reports, investigated traditional data quality approaches, including the National Standard for Spatial Data Accuracy (NSSDA) and National Map Accuracy Standard (NMAS), and built on key geocrowdsourcing work from Haklay (2010), Girres et al. (2010), Foody et al. (2013), and Goodchild and Li (2012) to develop comprehensive data quality metrics for geocrowdsourcing. These quality metrics include positional accuracy, attribute accuracy, completeness, logical consistency, semantic accuracy, temporal accuracy, lineage, usage, and the quality of contributors.

The unique contribution of this dissertation is the development, implementation, and iterative improvement of a comprehensive quality assessment framework for geocrowdsourced data, which is then used for measuring pedestrian network accessibility and optimization modeling of pedestrian network maintenance. This chapter provides a summary of the entire dissertation and concluding remarks. It also discusses several directions for future research.

Summary and Concluding Remarks

This dissertation is composed of three primary chapters (3, 4, and 5), which present the creation, iterative improvement, and use of a comprehensive data quality

assessment framework for conducting accessibility analysis of a pedestrian network, and the optimization modeling of pedestrian network maintenance.

In review, the application setting for this research is the geocrowdsourcing of transitory obstacles in a pedestrian network, and the development of methods for quality assessment and visualization of geocrowdsourced data, built on the previous work of Goodchild (2007), Haklay (2010), Girres et al. (2010), and others, who have made important research contributions, defining a new landscape of volunteered geographic information (VGI) and associated quality assessment strategies and techniques. While geocrowdsourcing is not limited to VGI, geocrowdsourced information is gathered from end-users, crowdsourcing applications, distributed sensors, and most importantly, social media applications such as Twitter. In this dissertation, the author (Han Qin) has developed a map-based reporting tool in web-based and mobile versions that have been tested and used by a number of contributors, including students and faculties from GMU, and research staff from the United States Army Corps of Engineers, Engineer Research and Development Center) to report obstacles on the GMU campus. Those tools have implemented spatial data quality assessment variables (Table 1) and generated preliminary quality scores for each crowdsourcing report. The initial spatial data quality metrics (Table 1) were developed by the QA-CGD research team, led by Dr. Matt Rice, and including Han Qin, Rebecca Rice, Fabiana I. Paez, Eric W. Ong and others. After the moderator-based quality assessment process is complete, an obstacle information record is produced by related high-quality reports from contributors and the reports from moderators (Qin 2015b, Rice 2015b).

Rather than focusing only on the geometry of geospatial data, this dissertation also focuses on attribute information to validate geometry with a social moderation approach (Goodchild and Li 2012, Rice 2015b), which relies on a team of well-trained moderators to conduct field check and moderate crowdsourced data for ground truth. Based on the validated obstacle dataset, the author (Han Qin) then used a full origin-destination matrix and shortest-path algorithm to calculate the value of average route length for all 2,772 junctions in the pedestrian network of the study area. This calculation method has been used under three different conditions – 1) no constraints 2) an end-user that wishes to avoid stairs and steep paths 3) an end-user that wishes to avoid stairs, steep paths, and validated transient obstacles. The third condition is the most restrictive condition and closest to reality for many of the mobility-impaired and vision-impaired end-users interviewed as a part of this research. The result under this most restrictive condition shows that the average route length increases from 1,834.7 meters to 2,215.5 meters or an increase of 20.76% above the unrestricted condition, which reflects the significant impact of obstacles for disabled communities. Interesting proposed modifications and extensions of this research, including new weighting formulations are discussed in the Future Research section of this chapter.

In chapter four, the author presents work on improving the quality and utility of obstacle information from geocrowdsourcing. This research studied the attribute categorizations and cognitive assessments of data contributors. Building on earlier thesis research of Paez (2014), the author (Han Qin) engaged with end users and data contributors to study and refine the attribute categorizations, such as obstacle types,

obstacle estimated duration, and obstacle urgency level. In chapter five, the author conducted surveys and assessments of five groups of local subject matter experts (SMEs) to collect priority information of the pedestrian network in study area. A family of spatial optimization models was developed to identify the most critical segments of the pedestrian network with geocrowdsourcing. The optimization models incorporated the SME data, segment usage counts from an exhaustive routing analysis, and geocrowdsourced obstacle reports from GMU-GcT platform, and used this information for optimizing sidewalk maintenance operations. This research (chapter five) underscores the benefit of using geocrowdsourcing as an input to a family of optimization models used in real-world decision-making.

Finally, this dissertation discusses quality assessment in geosocial media. The usefulness of NMAS, NSSDA, and traditional data quality assessment approaches in Girres et al. (2010) and Haklay (2010) begins to crumble when faced with unstructured geosocial media. Thus, an approach that considers fitness-for-use, lineage and the quality of contributors is discussed in the chapter six. An additional data quality dynamic that emerged, due to the massive volume of geosocial media data, is the necessity of using filters, selection criteria, and other fitness-for-use constraints to address basic quality assessment issues in geosocial media datasets. For instance, to validate the quality of contributors, the author (Han Qin) presents work on activity patterns and user dynamics in geosocial media. For millions of individual social media records, users with a large geographic activity patterns showed a high probability of being an active user.

Conversely, unusual activity patterns, such as single points or gridded points indicated a potentially artificial account.

In conclusion, this dissertation contributes to both GIScience and general scientific benefit in several ways. The primary conclusions of this research are summarized below:

- 1) A customized geocrowdsourcing platform with public input and feedback can lead to improvement in the quality of geocrowdsourced data. A map-based prototype – GMU-GcT was built to capture, store, filter, summarize, assess, analyze, and visualize the geocrowdsourced information from the public, with a goal of sharing transient obstacle information with the disabled communities. By using input, feedback and detailed quality assessment information, the author refined the attribute categorizations used in GMU-GcT, the geocrowdsourcing data contributors can provide more accurate information through a shared understanding of categories.
- 2) This dissertation reviews and discusses general theories and formal standards of geospatial data quality, such as NMAS and NSSDA. These standards address the most commonly explored component of accuracy (horizontal positional accuracy), through the work of Girres et al. (2010) and Haklay (2010), which compare OSM to authoritative government geospatial datasets. Average horizontal positional error of OSM data from these studies is 6 meters, which is slightly higher than the average error from the GMU-GcT, summarized in chapter three and in Rice et al. (2016). Other works about spatial data quality assessment, such as Foody et al. (2013) and

Hajjar et al. (2015) contributed to the formulation of our accuracy assessment methodology.

- 3) Methodologically, this dissertation proposes a comprehensive data quality assessment method by identifying the high quality geocrowdsourced reports with several quality assessment variables, and then improves the quality of the report process through a social moderation approach and input from local SMEs. Routing analysis with the pedestrian network and obstacle data from GMU-GcT is used to measure the impact of geocrowdsourced obstacle information on disabled individuals. A family of spatial optimization models is developed to assess the importance of pedestrian network segments with validated geocrowdsourced data and identify areas of a pedestrian network to repair. These studies demonstrate the usefulness of crowdsourced geospatial data and the possible uses of the GMU-GcT.
- 4) Technologically, open source is used widely in this dissertation. Especially, to address a great amount of computation, open source and cloud computing resources are leveraged. Compared with ArcGIS server, PostgreSQL/PostGIS is easier to access and free to use. Millions of calculations in the network analysis portion of this research were completed in a short time by using Python and PostGIS running on multiple virtual machines, provisioned by the OpenStack cloud platform. In contrast, earlier iterations of the same routing analysis took several months of processing time when using ArcPy and ArcGIS Server with same numbers of physical machines.
- 5) Traditional quality assessment approaches cannot be easily applied to evaluate unstructured social media data. To assess the quality of geosocial media data, this

dissertation presented a study of activity patterns of individuals and unusual patterns in large geosocial media dataset. Two general quality assessment factors – fitness for use and contributor quality, were deemed to be most relevant.

Overall, this dissertation implements a comprehensive GIS-centric quality assessment approach with semi-structured geocrowdsourced data, using a purpose-built geocrowdsourcing testbed, the GMU-GcT. An additional purpose of this research (assisting disabled persons with geocrowdsourced obstacle data) demonstrates the general importance and benefits of geocrowdsourcing. This dissertation addresses significant challenges, such as improving GMU-GcT with end-users' feedback, the recruitment of contributors, and the validation of geocrowdsourced data a moderator-based quality assessment system.

The author concludes that it is difficult to define a universal standard for assessing the quality of geocrowdsourcing, and that the data quality workflow should be different based on different purposes and geographical scales. For the GMU-GcT and the semi-structured obstacle data collected as a part of this research, traditional data quality approaches based on NMAS and NSSDA have value, with outcomes similar to other studies. The author hopes the research and results in this dissertation add value and assist future GIScience studies and related work from other domains.

Future Research

The GMU-GcT continues to develop with the assistance of data contributors, subject matter experts, and end-users. The geocrowdsourcing testbed combines a traditional map-based geocrowdsourcing tool with routing-based accessibility analysis

and optimization modeling, as presented in chapters three, four, and five. With the feedback from testbed end-users and subject matter experts, several further improvements will be implemented:

- 1) Instead of using a form-style interface to submit reports with a desktop computer, the GMU-GcT will incorporate an Instagram-style image contribution tool emphasizing streamlined mobile-based contributions. A quicker, easier, and fun contributing process could improve the user experiences and attract more contributors.
- 2) The author has collected detailed sloped information, and have discussed better slope-based routing to avoid the steep paths with a user-defined slope parameter. Currently, the GMU-GcT uses a standard Dijkstra shortest-cost path algorithm (chapter three). The weighting function only includes length of each segment as the cost variable. The cost variable could include two or more factors, such as priority and price, and different weighting formulations based on end-user preference. Other work by research team members to develop weighting models based on tree canopy coverage are underway.
- 3) Although the author leverages cloud resources and multi-thread programming to calculate a solution for the entire O-D matrix, much higher efficiency can be achieved by simplifying the computation through eliminating a large number of O-D pairs. Several minimum threshold values will be used to eliminate O-D pairs where the sum of the distance between pairs is less than a certain value, or where an O-D pair has minimal connectivity or impact on the rest of the network.

- 4) Analysis of network accessibility based on isochrones will be conducted to complement existing methods showing accessibility changes using path length increase. Isochrone maps, such as those developed by Sir Francis Galton (Thrower 2008, 152) are a more useful, functional way of depicting and mapping network accessibility and could be added to GMU-GcT with a few minor changes.
- 5) Chapter Five of this dissertation presents a benefit function comprised of segment routing usage, priority values assigned by SMEs, and a geocrowdsourced urgency parameter, which assesses the impact of an obstacle on an impacted segment. Future network analysis will combine these three factors with sensitivity analysis, identifying the network edges and junctions with the maximum impact on routing, and the most detrimental impact during disruption.
- 6) More complex models will be added to the family of optimization models for decision making. In the future, temporal issues will be introduced to the models with analysis of the order of repairs. For instance, prioritizing the sequence of segment repair to reduce the traffic issues could also impact the final decision making. Another important model parameter to explore will be the cost savings (in terms of materials) for contiguous repairs, which can be done more cheaply with larger batches of repair materials (e.g., concrete) delivered at one time.
- 7) Pareto optimization, and other algorithms used for multi-objective optimization will be explored with more complex optimization models, to explore the dynamics of optimization and the Pareto efficiency associated with changes to allocated resources.

REFERENCES

- Alpert, D. (2014, July 31). Why is Tysons walkability and bikeability so bad? Retrieved October 12, 2014, from <http://greatergreaterwashington.org/post/23802/why-is-tysons-walkability-and-bikeability-so-bad/>
- Antoniou, V., Morley, J., & Haklay, M. (2010). Web 2.0 geotagged photos: Assessing the spatial dimension of the phenomenon. *Geomatica*, 64(1), 99–110.
- Association, A. P., Steiner, F. R., & Butler, K. (2012). *Planning and Urban Design Standards*. John Wiley & Sons.
- Avila, K. (2014). The experiences of pedestrians with visual impairments in a metropolitan setting: An ethnographic inquiry. In *Proceedings, Biannual International Conference of the Association for Education and Rehabilitation for the Blind and Visually Impaired*. San Antonio, TX.
- Barbeau, S. J., Winters, P. L., Georggi, N. L., Labrador, M. A., & Perez, R. (2010). Travel assistance device: utilising global positioning system-enabled mobile phones to aid transit riders with special needs. *Intelligent Transport Systems, IET*, 4(1), 12–23.
- Barron, C., Neis, P., & Zipf, A. (2014). A Comprehensive Framework for Intrinsic OpenStreetMap Quality Analysis. *Transactions in GIS*, 18(6), 877–895. <https://doi.org/10.1111/tgis.12073>
- Batty, M., Hudson-Smith, A., Milton, R., & Crooks, A. (2010). Map mashups, Web 2.0 and the GIS revolution. *Annals of GIS*, 16(1), 1–13. <https://doi.org/10.1080/19475681003700831>
- Beale, L., Field, K., Briggs, D., Picton, P., & Matthews, H. (2006). Mapping for wheelchair users: Route navigation in urban spaces. *The Cartographic Journal*, 43(1), 68–81.
- Biba, S., Curtin, K. M., & Manca, G. (2014). Who Can Access Transit? Reviewing Methods for Determining Population Access to Bus Transit. *Cityscape: A Journal of Policy Development and Research*, 16(2), 193–204.

- Boulos, M. N. K., Resch, B., Crowley, D., Breslin, J., Sohn, G., & Burtner, R. (2011). Crowdsourcing, citizen sensing and sensor web technologies for public and environmental health surveillance and crisis management: trends, OGC standards and application examples. *International Journal of Health Geographics*, 10(1), 67–95. <https://doi.org/10.1186/1476-072X-10-67>
- Boulos, M. N. K., Warren, J., Jianya Gong, & Peng Yue. (2010). Web GIS in practice VIII: HTML5 and the canvas element for interactive online mapping. *International Journal of Health Geographics*, 9, 14–26. <https://doi.org/10.1186/1476-072X-9-14>
- Brunnett, G. (2004). *Geometric Modeling for Scientific Visualization*. Springer Science & Business Media.
- Bryson, B. (2011). *At Home: A Short History of Private Life* (Reprint edition). New York: Anchor.
- Butler, R. E. (1994). Geography and Vision-Impaired and Blind Populations. *Transactions of the Institute of British Geographers*, 19(3), 366. <https://doi.org/10.2307/622329>
- Camponovo, M. E., & Freundschuh, S. M. (2014). Assessing uncertainty in VGI for emergency response. *Cartography and Geographic Information Science*, 41(5), 440–455. <https://doi.org/10.1080/15230406.2014.950332>
- Carroll, A. (2006). A Case-Study-In-Progress: How a Media Organization Tackles the Georeferencing Challenge/Opportunity. Presented at the International Workshop on Digital Gazetteer Research & Practice. Retrieved from <http://ncgia.ucsb.edu/projects/nga/docs/carroll-paper.pdf>
- Cartwright, W. (2012). Neocartography: Opportunities, Issues and Prospects. *South African Journal of Geomatics*, 1(1).
- Chen, M., Lin, H., Liu, D., Zhang, H., & Yue, S. (2015). An object-oriented data model built for blind navigation in outdoor space. *Applied Geography*, 60, 84–94. <https://doi.org/10.1016/j.apgeog.2015.03.004>
- Chen, M., Lin, H., Lu, G., He, L., & Wen, Y. (2011). A spatial-temporal framework for historical and cultural research on China. *Applied Geography*, 31(3), 1059–1074. <https://doi.org/10.1016/j.apgeog.2011.01.022>
- Chrisman, N. (2006). Development in the Treatment of Spatial Data Quality. In R. Devillers & R. Jeansoulin (Eds.), *Fundamentals of Spatial Data Quality* (pp. 21–30). ISTE. Retrieved from <http://onlinelibrary.wiley.com/doi/10.1002/9780470612156.ch1/summary>

Church, R. L., & Marston, J. R. (2003). Measuring Accessibility for People with a Disability. *Geographical Analysis*, 35(1), 83–96. <https://doi.org/10.1111/j.1538-4632.2003.tb01102.x>

Cinnamon, J., Rinner, C., Cusimano, M. D., Marshall, S., Bakele, T., Hernandez, T., ... Chipman, M. L. (2009). Evaluating web-based static, animated and interactive maps for injury prevention. *Geospatial Health*, 4(1), 3–16.

Cinnamon, J., & Schuurman, N. (2010). Injury surveillance in low-resource settings using Geospatial and Social Web technologies. *International Journal of Health Geographics*, 9(1), 25. <https://doi.org/10.1186/1476-072X-9-25>

Cinnamon, J., & Schuurman, N. (2013). Confronting the data-divide in a time of spatial turns and volunteered geographic information. *GeoJournal*, 78(4), 657–674. <https://doi.org/10.1007/s10708-012-9458-6>

Clark, A. (2014). Where 2.0 Australia's Environment? Crowdsourcing, Volunteered Geographic Information, and Citizens Acting as Sensors for Environmental Sustainability. *ISPRS International Journal of Geo-Information*, 3(3), 1058–1076. <https://doi.org/10.3390/ijgi3031058>

Clark, F., & Illman, D. L. (2001). Dimensions of Civic Science Introductory Essay. *Science Communication*, 23(1), 5–27. <https://doi.org/10.1177/1075547001023001002>
Cohn, J. P. (2008). Citizen Science: Can Volunteers Do Real Research? *BioScience*, 58(3), 192–197. <https://doi.org/10.1641/B580303>

Coleman, D. J. (2013). Potential Contributions and Challenges of VGI for Conventional Topographic Base-Mapping Programs. In D. Sui, S. Elwood, & M. Goodchild (Eds.), *Crowdsourcing Geographic Knowledge* (pp. 245–263). Springer Netherlands. Retrieved from http://link.springer.com/chapter/10.1007/978-94-007-4587-2_14

Coulsen, M., Riger, M., & Wheate, R. (1991). Progress in creating tactile maps from Geographic Information Systems output. In *Mapping the Nations, 15th Conference and 9th General Assembly* (Vol. 1). Bournemouth, UK.

Croitoru, A., Crooks, A., Radzikowski, J., & Stefanidis, A. (2013). Geosocial gauge: a system prototype for knowledge discovery from social media. *International Journal of Geographical Information Science*, 27(12), 2483–2508. <https://doi.org/10.1080/13658816.2013.825724>

Crooks, A. T., & Wise, S. (2013). GIS and agent-based models for humanitarian assistance. *Computers, Environment and Urban Systems*, 41, 100–111. <https://doi.org/10.1016/j.compenvurbsys.2013.05.003>

- Dixon, L. (1996). Bicycle and Pedestrian Level-of-Service Performance Measures and Standards for Congestion Management Systems. *Transportation Research Record: Journal of the Transportation Research Board*, 1538, 1–9. <https://doi.org/10.3141/1538-01>
- Dodge, M., & Kitchin, R. (2013). Crowdsourced cartography: mapping experience and knowledge. *Environment and Planning A*, 45(1), 19 – 36. <https://doi.org/10.1068/a44484>
- Dunn, C. E. (2007). Participatory GIS — a people’s GIS? *Progress in Human Geography*, 31(5), 616–637. <https://doi.org/10.1177/0309132507081493>
- Egenhofer, M. J. (1996). Spatial-Query-by-Sketch. Presented at the IEEE Symposium on Visual Languages, Boulder, CO: IEEE Press.
- Elwood, S. (2006). Negotiating Knowledge Production: The Everyday Inclusions, Exclusions, and Contradictions of Participatory GIS Research. *Professional Geographer*, 58(2), 197–208. <https://doi.org/10.1111/j.1467-9272.2006.00526.x>
- Elwood, S. (2010). Geographic Information Science: Emerging Research on the Societal Implications of the Geospatial Web. *Progress in Human Geography*, 34(3), 349–357.
- Elwood, S., Goodchild, M. F., & Sui, D. Z. (2012). Researching Volunteered Geographic Information: Spatial Data, Geographic Research, and New Social Practice. *Annals of the Association of American Geographers*, 102(3), 571–590. <https://doi.org/10.1080/00045608.2011.595657>
- Engelbrektsson, P., Karlsson, I. C. M., Gallagher, B., Hunter, H., Petrie, H., & O’Neill, A.-M. (2004). Developing a navigation aid for the frail and visually impaired. *Universal Access in the Information Society*, 3(3-4), 194–201. <https://doi.org/10.1007/s10209-003-0088-0>
- Footy, G. M., See, L., Fritz, S., Van der Velde, M., Perger, C., Schill, C., & Boyd, D. S. (2013). Assessing the Accuracy of Volunteered Geographic Information arising from Multiple Contributors to an Internet Based Collaborative Project. *Transactions in GIS*, 17(6), 847–860. <https://doi.org/10.1111/tgis.12033>
- Forbus, K. D., Usher, J., & Chapman, V. (2003). Qualitative Spatial Reasoning about Sketch Maps. Proceedings of the Fifteenth Annual Conference on Innovative Applications of Artificial Intelligence.
- Foth, M., Bajracharya, B., Brown, R., & Hearn, G. (2009). The Second Life of urban planning? Using NeoGeography tools for community engagement. *Journal of Location Based Services*, 3(2), 97–117. <https://doi.org/10.1080/17489720903150016>

- Geertman, S. (2002). Participatory planning and GIS: a PSS to bridge the gap. *Environment and Planning B: Planning and Design*, 29(1), 21 – 35. <https://doi.org/10.1068/b2760>
- Geertman, S. C. M., & Ritsema Van Eck, J. R. (1995). GIS and models of accessibility potential: an application in planning. *International Journal of Geographical Information Systems*, 9(1), 67–80. <https://doi.org/10.1080/02693799508902025>
- Girres, J.-F., & Touya, G. (2010). Quality Assessment of the French OpenStreetMap Dataset. *Transactions in GIS*, 14(4), 435–459. <https://doi.org/10.1111/j.1467-9671.2010.01203.x>
- Golledge, R. G. (1999). *Wayfinding behavior: Cognitive mapping and other spatial processes*. JHU Press.
- Golledge, R. G., Jacobson, R. D., Kitchin, R., & Blades, M. (2000). Cognitive maps, spatial abilities, and human wayfinding. *Geographical Review of Japan Series B*, 73(2), 93–104.
- Golledge, R. G., Rice, M., & Jacobson, D. (2005). A Commentary on the Use of Touch for Accessing On-Screen Spatial Representations: The Process of Experiencing Haptic Maps and Graphics. *The Professional Geographer*, 57(3), 339–349. <https://doi.org/10.1111/j.0033-0124.2005.00482.x>
- Golledge, R. G., Rice, M. T., & Jacobson, R. D. (2006). Multimodal interfaces for representing and accessing geospatial information. In *Frontiers of Geographic Information Technology* (pp. 181–208). Springer.
- Goodchild, M. (2009). NeoGeography and the nature of geographic expertise. *Journal of Location Based Services*, 3(2), 82–96. <https://doi.org/10.1080/17489720902950374>
- Goodchild, M. F. (2007). Citizens as Sensors: The world of volunteered geography. Keynote Address, Session 1 presented at the NCGIA Workshop on Volunteered Geographic Information, December 13-17, 2007, UCSB, Santa Barbara, CA, December 13-17, 2007. Retrieved from http://ncgia.ucsb.edu/projects/vgi/docs/present/Goodchild_intro.pdf
- Goodchild, M. F. (2009). NeoGeography and the nature of geographic expertise. *Journal of Location Based Services*, 3(2), 82–96. <https://doi.org/10.1080/17489720902950374>
- Goodchild, M. F., Kyriakidis, P., Rice, M., & Schneider, P. (2005). *Report of the NCGIA Specialist Meeting on Spatial Webs*. Citeseer. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.432.3011&rep=rep1&type=pdf>

Goodchild, M. F., & Li, L. (2012). Assuring the quality of volunteered geographic information. *Spatial Statistics*, 1, 110–120. <https://doi.org/10.1016/j.spasta.2012.03.002>

Guptill, S. C., & Morrision, J.L. (1995). *Elements of Spatial Data Quality* (Vol. 202). Elsevier Science Limited.

Gura, T. (2013). Citizen science: Amateur experts. *Nature*, 496(7444), 259–261. <https://doi.org/10.1038/nj7444-259a>

Haklay, M. (2010). How good is volunteered geographical information? A comparative study of OpenStreetMap and Ordnance Survey datasets. *Environment and Planning B: Planning and Design*, 37(4), 682–703. <https://doi.org/10.1068/b35097>

Haklay, M. (2013). Citizen Science and Volunteered Geographic Information: Overview and Typology of Participation. In D. Sui, S. Elwood, & M. Goodchild (Eds.), *Crowdsourcing Geographic Knowledge* (pp. 105–122). Springer Netherlands. Retrieved from http://link.springer.com/chapter/10.1007/978-94-007-4587-2_7

Haklay, M. (2013). Neogeography and the delusion of democratisation. *Environment and Planning A*, 45(1), 55 – 69. <https://doi.org/10.1068/a45184>

Haklay, M. (Muki), Basiouka, S., Antoniou, V., & Ather, A. (2010). How Many Volunteers Does it Take to Map an Area Well? The Validity of Linus' Law to Volunteered Geographic Information. *The Cartographic Journal*, 47(4), 315–322. <https://doi.org/10.1179/000870410X12911304958827>

Hand, E. (2010). Citizen science: People power. *Nature News*, 466(7307), 685–687. <https://doi.org/10.1038/466685a>

Han, S. Y., & Tsou, M.-H. (2016). Dynamic Population Mapping Using Social Media Data (Twitter). In *AutoCarto Proceedings Papers 2016*.

Howe, J. (2008). *Crowdsourcing: Why the Power of the Crowd Is Driving the Future of Business* (1st ed.). New York, NY, USA: Crown Publishing Group.

Hudson-Smith, A., Batty, M., Crooks, A., & Milton, R. (2009). Mapping for the Masses Accessing Web 2.0 Through Crowdsourcing. *Social Science Computer Review*, 27(4), 524–538. <https://doi.org/10.1177/0894439309332299>

Hunter, G. J., & Beard, K. (1992). Understanding error in spatial databases. *Australian Surveyor*, 37(2), 108–119. <https://doi.org/10.1080/00050326.1992.10438784>

Jacobson, R. D. (1998). Cognitive mapping without sight: Four preliminary studies of spatial learning. *Journal of Environmental Psychology*, 18(3), 289–305.

Karimi, H. A., & Kasemsuppakorn, P. (2013). Pedestrian network map generation approaches and recommendation. *International Journal of Geographical Information Science*, 27(5), 947–962.

Karimi, H. A., Zhang, L., & Benner, J. G. (2014). Personalized accessibility map (PAM): a novel assisted wayfinding approach for people with disabilities. *Annals of GIS*, 20(2), 99–108. <https://doi.org/10.1080/19475683.2014.904438>

Kasemsuppakorn, P., & Karimi, H. A. (2009). Personalised routing for wheelchair navigation. *Journal of Location Based Services*, 3(1), 24–54. <https://doi.org/10.1080/17489720902837936>

Keirstead, J., & Shah, N. (2013). The Changing Role of Optimization in Urban Planning. In A. Chinchuluun, P. M. Pardalos, R. Enkhbat, & E. N. Pistikopoulos (Eds.), *Optimization, Simulation, and Control* (pp. 175–193). Springer New York. Retrieved from http://link.springer.com/chapter/10.1007/978-1-4614-5131-0_11

Kockelman, K., Heard, L., Kweon, Y.-J., & Rioux, T. (2002). Sidewalk Cross-Slope Design: Analysis of Accessibility for Persons with Disabilities. *Transportation Research Record: Journal of the Transportation Research Board*, 1818, 108–118. <https://doi.org/10.3141/1818-17>

Kockelman, K., Zhao, Y., & Blanchard-Zimmerman, C. (2001). Meeting the intent of ADA in sidewalk cross-slope design. *Journal of Rehabilitation Research and Development*, 38(1), 101–110.

Kraak, M.-J. (2011). Is There a Need for Neo-Cartography? *Cartography and Geographic Information Science*, 38(2), 73–78. <https://doi.org/10.1559/1523040638273>

Kulyukin, V. A., Nicholson, J., Ross, D. A., Marston, J. R., & Gaunet, F. (2008). The Blind Leading the Blind: Toward Collaborative Online Route Information Management by Individuals with Visual Impairments. In *AAAI Spring Symposium: Social Information Processing* (pp. 54–59). Retrieved from <http://www.aaai.org/Papers/Symposia/Spring/2008/SS-08-06/SS08-06-011.pdf>

Laakso, M., Sarjakoski, T., & Sarjakoski, L. T. (2011). Improving accessibility information in pedestrian maps and databases. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 46(2), 101–108.

- Li, L., & Goodchild, M. F. (2010). The Role of Social Networks in Emergency Management. *International Journal of Information Systems for Crisis Response and Management*, 2(4), 48–58. <https://doi.org/10.4018/jiscrm.2010100104>
- Lockwood, K., Forbus, K., Halstead, D. T., & Usher, J. (n.d.). Automatic Categorization of Spatial Propositions. Northwestern University.
- Loomis, J. M., Golledge, R. G., & Klatzky, R. L. (2001). GPS-based navigation systems for the visually impaired. In *Fundamentals of wearable computers and augmented reality* (pp. 429–446). Mahwah, NJ: Lawrence Erlbaum Associates.
- Loomis, J. M., Marston, J. R., Golledge, R. G., & Klatzky, R. L. (2005). Personal guidance system for people with visual impairment: A comparison of spatial displays for route guidance. *Journal of Visual Impairment & Blindness*, 99(4), 219.
- Marston, J. R. (2002). Towards an Accessible City: Empirical Measurement and Modeling of Access to Urban Opportunities for those with Vision Impairments, Using Remote Infrared Audible Signage. *University of California Transportation Center*. Retrieved from <http://escholarship.org/uc/item/9df5b5ts>
- Marston, J. R., Loomis, J. M., Klatzky, R. L., Golledge, R. G., & Smith, E. L. (2006). Evaluation of Spatial Displays for Navigation Without Sight. *ACM Trans. Appl. Percept.*, 3(2), 110–124. <https://doi.org/10.1145/1141897.1141900>
- Michael F. Goodchild, J. A. G. (2010). Crowdsourcing geographic information for disaster response: a research frontier. *Int. J. Digital Earth*, 3(3), 231–241. <https://doi.org/10.1080/17538941003759255>
- Miele, J. A. (2007). Automated Creation and Detailed Annotation of Audio/Tactile Maps Using Scalable Vector Graphics (SVG). Presented at the International Cartographic Conference, Moscow, Russia.
- Miele, J. A. (2011). Tactile Map Automated Production (TMAP): Using GIS Data to Automatically Generate Raised-Line Street Maps. Retrieved from <http://www.ski.org/Rehab/TMAP/>
- Miele, J. A., & Gilden, D. B. (2004). Tactile Map Automated Production (TMAP): Using GIS Data to Generate Braille Maps. In *Proceedings of the CSUN International Conference on Technology and Persons with Disabilities*. Los Angeles, CA.
- Miele, J. A., & Marston, J. R. (2005). Tactile Map Automated Production (TMAP): Project Update and Research Summary. In *Proceedings of the CSUN 20th Annual International Conference*. Los Angeles, CA.

Monmonier, M. S. (1985). *Technological Transition in Cartography*. Mark Monmonier. Nedkov, S., & Zlatanova, S. (n.d.). Enabling Obstacle Avoidance for Google Maps' Navigation Service.

Nuernberger, A. (2008). *Presenting Accessibility to Mobility-Impaired Travelers* (UCTC Dissertation). University of California Transportation Center, Santa Barbara, CA.

Orsi, F., & Geneletti, D. (2013). Using geotagged photographs and GIS analysis to estimate visitor flows in natural areas. *Journal for Nature Conservation*, 21(5), 359–368. <https://doi.org/10.1016/j.jnc.2013.03.001>

Paez, F. I. (2014, May). *Recruitment, Training, and Social Dynamics in Geo-Crowdsourcing for Accessibility* (Master of Science). George Mason University, Fairfax, VA.

Palazzi, C. E., Teodori, L., & Rocchetti, M. (2010). Path 2.0: A participatory system for the generation of accessible routes. In *2010 IEEE International Conference on Multimedia and Expo (ICME)* (pp. 1707–1711). <https://doi.org/10.1109/ICME.2010.5583240>

Pease, P. A. (2014, July). *The Influence of Training on Position and Attribute Accuracy in VGI* (Master of Science). George Mason University, Fairfax, VA.

Perkins, C. (2001). Tactile campus mapping: Evaluation designs and production technologies. In *Proceedings, The 20th International Cartographic Conference* (Vol. 5, pp. 2906–15). Beijing, China.

Perkins, C. (2002). Cartography: progress in tactile mapping. *Progress in Human Geography*, 26(4), 521–530. <https://doi.org/10.1191/0309132502ph383pr>

Pingel, T. J. (2010). Modeling slope as a contributor to route selection in mountainous areas. *Cartography and Geographic Information Science*, 37(2), 137–148.

Pingel, T. J. (2010). *Strategic elements of route choice*. University of California, Santa Barbara.

Qin, H., Aburizaiza, A. O., Rice, R. M., Paez, F., & Rice, M. T. (2015a). Obstacle characterization in a geocrowdsourced accessibility system. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 1, 179–185.

Qin, H., Rice, R. M., Fuhrmann, S., Rice, M. T., Curtin, K. M., & Ong, E. (2015b). Geocrowdsourcing and accessibility for dynamic environments. *GeoJournal*, 1–18. <https://doi.org/10.1007/s10708-015-9659-x>

Ramm, F., Topf, J., & Chilton, S. (2011). *OpenStreetMap: using and enhancing the free map of the world*. UIT Cambridge Cambridge. Retrieved from <http://library.wur.nl/WebQuery/clc/1958758>

Rice, M. T., Aburizaiza, A. O., Jacobson, R. D., Shore, B. M., & Paez, F. I. (2012a). Supporting Accessibility for Blind and Vision-impaired People With a Localized Gazetteer and Open Source Geotechnology. *Transactions in GIS*, 16(2), 177–190. <https://doi.org/10.1111/j.1467-9671.2012.01318.x>

Rice, M. T., Curtin, K. M., Paez, F. I., Seitz, C. R., & Qin, H. (2013a). Crowdsourcing to Support Navigation for the Disabled: A Report on the Motivations, Design, Creation and Assessment of a Testbed Environment for Accessibility.

Rice, M. T., Curtin, K. M., Pfoser, D., Rice, R. M., Fuhrmann, S., Qin, H., ... others. (2015a). Social Moderation and Dynamic Elements in Crowdsourced Geospatial Data: A Report on Quality Assessment, Dynamic Extensions and Mobile Device Engagement in the George Mason University Geocrowdsourcing Testbed. George Mason University Fairfax United States. Retrieved from <http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=AD1001943>

Rice, M. T., Jacobson, R. D., Caldwell, D. R., McDermott, S. D., Paez, F. I., Aburizaiza, A. O., ... Qin, H. (2013b). Crowdsourcing techniques for augmenting traditional accessibility maps with transitory obstacle information. *Cartography and Geographic Information Science*, 40(3), 210–219.

Rice, M. T., Paez, F. I., Mulhollen, A. P., Shore, B. M., & Caldwell, D. R. (2012b). Crowdsourced Geospatial Data: A Report on the Emerging Phenomena of Crowdsourced and User-Generated Geospatial Data.

Rice, M. T., Paez, F. I., Rice, R. M., Ong, E. W., Qin, H., Seitz, C. R., ... others. (2014). *Quality Assessment and Accessibility Applications of Crowdsourced Geospatial Data: A Report on the Development and Extension of the George Mason University Geocrowdsourcing Testbed*. GEORGE MASON UNIV FAIRFAX VA. Retrieved from <http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA615952>

Rice, R. M. (2015b). Validating VGI Data Quality in Local Crowdsourced Accessibility Mapping Applications: a George Mason University Case Study. George Mason University.

Rice, R. M., Aburizaiza, A. O., Rice, M. T., & Qin, H. (2016). Position Validation in Crowdsourced Accessibility Mapping. *Cartographica: The International Journal for*

Geographic Information and Geovisualization, 51(2), 55–66.
<https://doi.org/10.3138/cart.51.2.3143>

Rodgers, R. E. (2016, June 30). *A Statistical Comparison of Sidewalk Slopes Derived from Multi-Resolution Digital Elevation Models in Support of Accessibility* (Thesis). Retrieved from <http://mars.gmu.edu/handle/1920/10295>

Rodríguez, A., Yebes, J. J., Alcantarilla, P. F., Bergasa, L. M., Almazán, J., & Cela, A. (2012). Assisting the Visually Impaired: Obstacle Detection and Warning System by Acoustic Feedback. *Sensors*, 12(12), 17476–17496. <https://doi.org/10.3390/s121217476>

Russ, T. (2009). *Site Planning and Design Handbook, Second Edition*. McGraw Hill Professional.

Semple, H., Qin, H., & Sasson, C. (2013). Development of a Web GIS Application for Visualizing and Analyzing Community Out of Hospital Cardiac Arrest Patterns. *Online Journal of Public Health Informatics*, 5(2), 212. <https://doi.org/10.5210/ojphi.v5i2.4587>

Shekhar, S., Coyle, M., Goyal, B., Liu, D.-R., & Sarkar, S. (1997). Data Models in Geographic Information Systems. *Commun. ACM*, 40(4), 103–111.
<https://doi.org/10.1145/248448.248465>

Sieber, R. (2006). Public participation geographic information systems: A literature review and framework. *Annals of the Association of American Geographers*, 96(3), 491–507.

Silvertown, J. (2009). A new dawn for citizen science. *Trends in Ecology & Evolution*, 24(9), 467–471. <https://doi.org/10.1016/j.tree.2009.03.017>

Stefanidis, A., Crooks, A., & Radzikowski, J. (2011). Harvesting ambient geospatial information from social media feeds. *GeoJournal*, 78(2), 319–338.
<https://doi.org/10.1007/s10708-011-9438-2>

Sui, D., Elwood, S., & Goodchild, M. F. (Eds.). (2013). *Crowdsourcing Geographic Knowledge Volunteered Geographic Information (VGI) in Theory and Practice*. New York, NY: Springer.

Tatham, A. F. (1991). The design of tactile maps: Theoretical and practical considerations. In *Mapping the Nations: Proceedings, 15th Conference and 9th General Assembly of the International Cartographic Association* (Vol. 1, pp. 157–166). Bournemouth, UK.

Tellex, S., & Roy, D. (2009). Grounding Spatial Prepositions for Video Search. MIT.

- Thatcher, J. (2013). From Volunteered Geographic Information to Volunteered Geographic Services. In D. Sui, S. Elwood, & M. F. Goodchild (Eds.), *Crowdsourcing Geographic Knowledge* (pp. 161–173). Dordrecht: Springer Netherlands. Retrieved from http://www.springerlink.com/index/10.1007/978-94-007-4587-2_10
- Thatcher, J., Waddell, C., Henry, S., Swierenga, S., Urban, M., Burks, M., & Bohman, P. (2003). *Constructing Accessible Web Sites* (Softcover reprint of the original 1st ed. edition). Birmingham, U.K. : New York: Apress.
- Thies, J., & Smith, V. (2011). Transitions in Digital Map Production: An Industry Perspective. *Cartography and Geographic Information Science*, 38(3), 310–312. <https://doi.org/10.1559/15230406382310>
- Thrower, N. J. W. (2008). *Maps and Civilization: Cartography in Culture and Society, Third Edition*. University of Chicago Press.
- Tsou, M.-H., Jung, C.-T., Allen, C., Yang, J.-A., Gawron, J.-M., Spitzberg, B. H., & Han, S. (2015). Social Media Analytics and Research Test-bed (SMART Dashboard). In *Proceedings of the 2015 International Conference on Social Media & Society* (pp. 2:1–2:7). New York, NY, USA: ACM. <https://doi.org/10.1145/2789187.2789196>
- Tsou, M.-H., Yang, J.-A., Lusher, D., Han, S., Spitzberg, B., Gawron, J. M., ... An, L. (2013). Mapping social activities and concepts with social media (Twitter) and web search engines (Yahoo and Bing): a case study in 2012 US Presidential Election. *Cartography and Geographic Information Science*, 40(4), 337–348. <https://doi.org/10.1080/15230406.2013.799738>
- Tulloch, D. (2014). Crowdsourcing geographic knowledge: volunteered geographic information (VGI) in theory and practice. *International Journal of Geographical Information Science*, 28(4), 847–849. <https://doi.org/10.1080/13658816.2013.873440>
- Tulloch, D. L. (2008). Is VGI participation? From vernal pools to video games. *GeoJournal*, 72(3/4), 161–171.
- Veregin, H. (1999). Data quality parameters. *Geographical Information Systems*, 1, 177–189.
- Vorapatratorn, S., & Nambunmee, K. (2014). iSonar: an obstacle warning device for the totally blind. *Journal of Assistive, Rehabilitative & Therapeutic Technologies*, 2(0). <https://doi.org/10.3402/jartt.v2.23114>
- Wang, F. (2000). Modeling Commuting Patterns in Chicago in a GIS Environment: A Job Accessibility Perspective. *The Professional Geographer*, 52(1), 120–133. <https://doi.org/10.1111/0033-0124.00210>

- Wayant, N., Crooks, A., Stefanidis, A., Croitoru, A., Radzikowski, J., Stahl, J., & Shine, J. (2012). Spatiotemporal clustering of twitter feeds for activity summarization. *GIScience (short Paper)*. Retrieved from http://www.academia.edu/download/30811448/giscience2012_paper_78.pdf
- Williams, M. A., Hurst, A., & Kane, S. K. (2013). “Pray before you step out”: describing personal and situational blind navigation behaviors. In *Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility* (pp. 1–8). Bellevue, Washington: ACM.
- Xu, C., Qin, H., & Yu, M. (2015). Visualizing spatiotemporal trajectories of mobile social media users using space–time cube. *Cartography and Geographic Information Science*, 42(sup1), 75–83. <https://doi.org/10.1080/15230406.2015.1059253>

BIOGRAPHY

Han Qin received his Bachelor of Engineering from Wuhan University in 2009 in China. Then he came to Eastern Michigan University through International Joint Program with Wuhan University and received Master of Science in Geographic Information Systems in 2012. As a master's student at Eastern Michigan and a doctoral student at George Mason University, Han Qin published the following peer-reviewed journal papers as a part of his research, as well as several of technical reports that can be viewed at <https://scholar.google.com/citations?user=brZ3eZIAAAAJ&hl=en>.

- 1) Qin, H., Aburizaiza, A. O., Rice, R. M., Paez, F., & Rice, M. T. (2015). Obstacle Characterization in a Geocrowdsourced Accessibility System. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 1, 179–185.
- 2) Qin, H., Rice, R. M., Fuhrmann, S., Rice, M. T., Curtin, K. M., & Ong, E. (2015). Geocrowdsourcing and accessibility for dynamic environments. *GeoJournal*, 1–18. <https://doi.org/10.1007/s10708-015-9659-x>
- 3) Rice, M. T., Jacobson, R. D., Caldwell, D. R., McDermott, S. D., Paez, F. I., Aburizaiza, A. O., ... Qin, H. (2013). Crowdsourcing techniques for augmenting traditional accessibility maps with transitory obstacle information. *Cartography and Geographic Information Science*, 40(3), 210–219. <https://doi.org/10.1080/15230406.2013.799737>
- 4) Rice, R. M., Aburizaiza, A. O., Rice, M. T., & Qin, H. (2016). Position Validation in Crowdsourced Accessibility Mapping. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 51(2), 55–66. <https://doi.org/10.3138/cart.51.2.3143>
- 5) Semple, H., Qin, H., & Sasson, C. (2013). Development of a Web GIS Application for Visualizing and Analyzing Community Out of Hospital Cardiac Arrest Patterns. *Online Journal of Public Health Informatics*, 5(2), 212. <https://doi.org/10.5210/ojphi.v5i2.4587>
- 6) Xu, C., Qin, H., & Yu, M. (2015). Visualizing spatiotemporal trajectories of mobile social media users using space–time cube. *Cartography and Geographic Information Science*, 42(sup1), 75–83. <https://doi.org/10.1080/15230406.2015.1059253>