A COMPUTATIONAL METHODOLOGY FOR ADDRESSING DIFFERENTIATED ACCESS OF VULNERABLE POPULATIONS DURING BIOLOGICAL EMERGENCIES

Martin Joseph O'Neill II

Dissertation Prepared for the Degree of

DOCTOR OF PHILOSOPHY

UNIVERSITY OF NORTH TEXAS

August 2014

APPROVED:

Armin R. Mikler, Major Professor
Bill Buckles, Committee Member
Yan Huang, Committee Member
Bruce Hunter, Committee Member
Tom Jacob, Committee Member
Chetan Tiwari, Committee Member
Barrett Bryant, Chair of the Department of Computer Science and Engineering
Costas Tsatsoulis, Dean of the College of Engineering
Mark Wardell, Dean of the Toulouse Graduate School

O'Neill II, Martin Joseph. A Computational Methodology for Addressing Differentiated Access of Vulnerable Populations during Biological Emergencies. Doctor of Philosophy (Computer Science and Engineering), August 2014, 115 pp., 6 tables, 68 illustrations, bibliography, 150 titles.

Mitigation response plans must be created to protect affected populations during biological emergencies resulting from the release of harmful biochemical substances. Medical countermeasures have been stockpiled by the federal government for such emergencies. However, it is the responsibility of local governments to maintain solid, functional plans to apply these countermeasures to the entire target population within short, mandated time frames. Further, vulnerabilities in the population may serve as barriers preventing certain individuals from participating in mitigation activities. Therefore, functional response plans must be capable of reaching vulnerable populations.

Transportation vulnerability results from lack of access to transportation. Transportation vulnerable populations located too far from mitigation resources are at-risk of not being able to participate in mitigation activities. Quantification of these populations requires the development of computational methods to integrate spatial demographic data and transportation resource data from disparate sources into the context of planned mitigation efforts. Research described in this dissertation focuses on quantifying transportation vulnerable populations and maximizing participation in response efforts. Algorithms developed as part of this research are integrated into a computational framework to promote a transition from research and development to deployment and use by biological emergency planners.

ACKNOWLEDGMENTS

I would like to thank my friends, my family, and God Almighty, without whom I would not have been able to make it this far. I would like to express my gratitude to Armin R. Mikler for his tireless dedication to my education. When I took my first undergraduate class from him, I had no idea he would later become my major professor, my mentor, and my friend. I am thankful to have such wonderful friends as Kyle Fleming, Billy O'Dell, and Shaun and Kate Szkolnik, who have always challenged and supported me.

I am grateful to my Mom and Dad for striving to provide every opportunity they could find for me to learn and explore throughout my childhood. This early education enabled me to appreciate math and science and has served as a solid foundation upon which I have built my academic career. I am grateful for my wife Tricia and our little dog Henry for making the bad days good and the good days even better. Their love and support have made this journey much more enjoyable.

I would like to recognize the other members of the Computational Response Planning research group at the University of North Texas Center for Computational Epidemiology and Response Analysis: Saratchandra Indrakanti, Tamara (Schneider) Jimenez, Armin R. Mikler, Susie Ramisetty-Mikler, Chetan Tiwari, and Josh Urbanovsky. Although the research at the focus of this dissertation is my own, the larger computational framework into which I integrated my new methodologies has been a team effort over the last five years. I would also like to recognize those who have collaborated with us on this project including Terry LaFon from the Texas Department of State Health Services as well as Stephen (Mark) Fulmer, Leslie Rodriguez, and Stewart DeJournett from the Tarrant County Public Health Department (TCPHD). Early funding for the Computational Response Planning project described in this dissertation was provided by TCPHD through the Cities Readiness Initiative. Research described in this dissertation was supported by Grant Number NIH 1R01LM011647-01 and Grant Number NIH 1R15LM010804-01 from the National Institutes of Health.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Mitigation response plans must be created to protect the population during biological emergencies resulting from the deliberate or accidental release of harmful biochemical substances. Although medical countermeasures (MCMs) have been stockpiled for such emergencies, the rapid application of these countermeasures to large, diverse populations presents a challenge to public health preparedness practitioners (PHPP). A computational framework to assist PHPP in creating, analyzing, and optimizing their response plans has been created. However, access disparities can result from vulnerabilities not considered during the planning process. Research presented in this document quantitatively identifies vulnerability in the population and explores new analysis and optimization techniques to minimize resulting access disparities using computational methodologies. Specifically, lack of access to mitigation resources arising from the unavailability of transportation resources is addressed spatially with respect to available demographic and public transit data.

1.1. Background

Terrorist incidents such as the March 1995 Tokyo subway attack using the nerve agent sarin prompted concerns regarding the ability of the United States to effectively respond to the release of chemical or biological agents. These concerns led to the creation of the National Pharmaceutical Stockpile (NPS) in 1997 [134]. In 2003, the NPS was renamed the Strategic National Stockpile (SNS). The SNS is currently overseen by the Department of Health and Human Services (DHHS) and the Centers for Disease Control and Prevention (CDC). It may be used to restock local emergency medical supplies in the case of a public health emergency [130].

A policy of rapid, aggressive treatment of the affected population in a biological emergency is supported by computational models [71][140], and delaying this treatment can increase the number of casualties [9] as well as the burden on hospitals [65]. For example, if the biological emergency resulted from the release of smallpox, the infection time line

in Figure 1.1 shows that treatment by vaccination is unlikely to be effective seven to ten days following initial viral infection. The time line illustrates that infected individuals may begin transmitting the disease up to four days before they become symptomatic. The last naturally occurring case of smallpox was in 1979, and earlier symptoms of smallpox resemble common diseases such as influenza or chickenpox [26]. This means that infected individuals may transmit smallpox to others before the outbreak is detected, leading to an increased number of casualties [100].

Over-the-counter drug sales have been monitored to detect public health events using the National Retail Data Monitor tool [63]. However, patient complaints are among the most used data sources in syndromic surveillance systems. [63]. Although many alternative strategies of disease surveillance have been explored in an effort to to achieve earlier outbreak detection [33][36][38][53][63][64][75][109], aggressive, mass-treatment campaigns must be implemented quickly in order to curtail the biological emergency.

The release of anthrax was identified as one of the new health threats of the 21st century by the World Health Organization (WHO) [148]. Anthrax vaccines may not be available in sufficiently large quantities [13] to treat entire populations immediately following a release of anthrax. A study of the cost-effectiveness of a pre-attack anthrax vaccination campaign found that that such a campaign is only cost-effective if the probability of attack is sufficiently high [51]. Otherwise, post-attack prophylactic treatment using a combination of antibiotics and vaccination are currently recommended, though new, non-antibiotic, non-vaccination treatment methods are currently being researched [93]. Even though these antibiotics have not been adequately tested or have some probability of significantly adverse reactions in children, pregnant women, or immuno-suppressed individuals, a working group including medical and public health experts from academic, research, and government institutions recommend their use when balancing risk of these antibiotics against that of exposure to an engineered, antibiotic-resistant strain of anthrax [67].

Continued federal funding to research, development, and stockpile MCMs underscores their importance in preparedness efforts. In fiscal year 2014, the Department of Health and

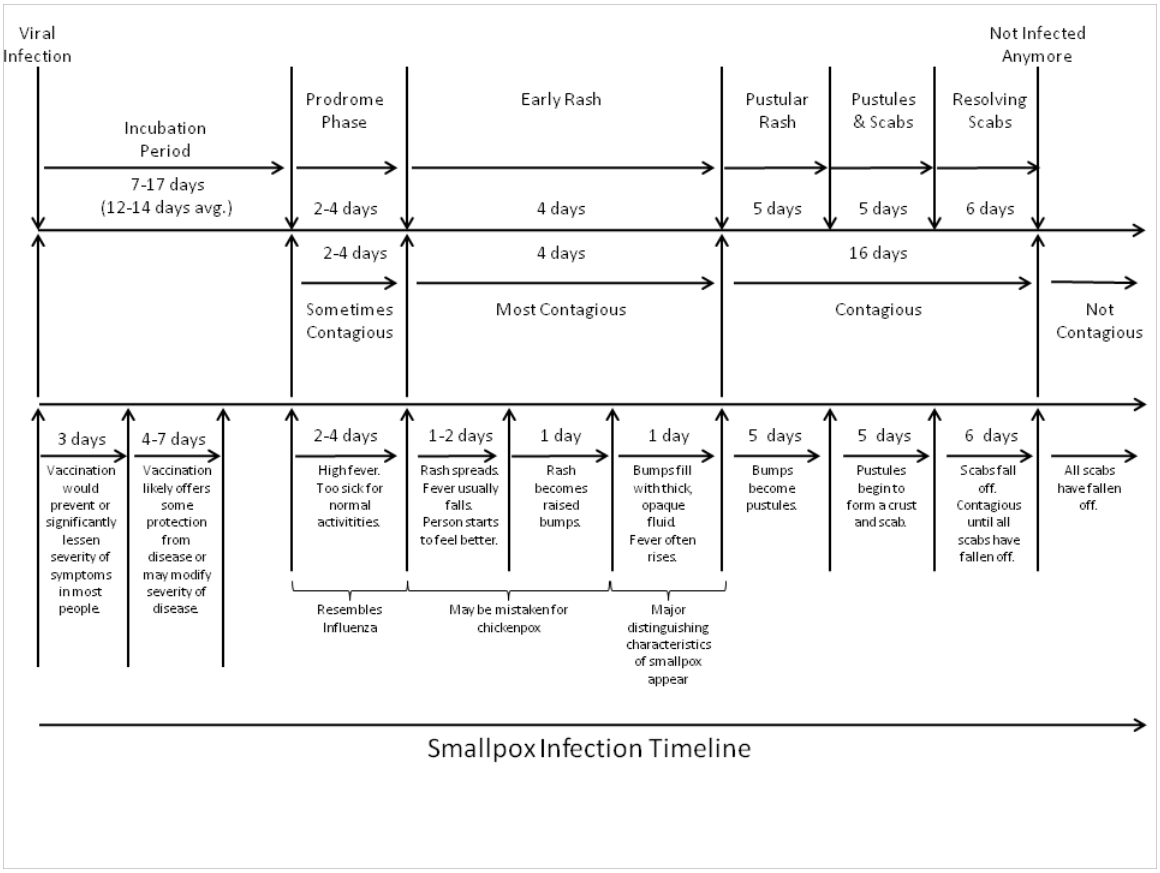


FIGURE 1.1. Smallpox infection time line using information from [25][26][34][147] to represent smallpox disease progression in an individual

Human Services has allocated $415 million to support advanced research and development of MCMs and an additional $250 million as the *first installment* of a multi-year commitment to acquire MCMs [115]. Although these MCMs include assets for response to chemical, biological, radiological, and nuclear (CBRN) events, the Department of Homeland Security has identified anthrax as the agent most likely to be used in bio-terrorist acts [70]. Further, in fiscal years 2004 - 2012, MCMs for response to the release of anthrax comprised 44% of all MCMs procured [70]. However, having MCMs is only the first step towards a successful mitigation campaign [43]. The Public Health Emergency Medical Countermeasures Enterprise (PHEMCE) recognizes the existence of diverse and unique vulnerabilities in the population and has expressed its commitment to address gaps in MCM application resulting from such vulnerabilities [116]. Therefore, participation in campaigns to distribute MCMs in a timely manner must be maximized to support the federal government's public health

security efforts.

Although the United States federal government has acquired many of the medical or pharmacological supplies needed to facilitate treatment of the population, the responsibility of actually dispensing medications to and treating individuals lies with local governments [27]. While this may seem like an elementary problem for smaller localities, the goal of reaching every individual [29] in significantly larger localities within short mandated time frames [28] presents a significant challenge. Vulnerabilities in the population may limit the participation of certain individuals, thus hindering the effectiveness of response plans and increasing access disparities. Therefore, localities must prepare themselves by planning for these contingencies.

The Cities Readiness Initiative (CRI) was instituted in 2004 by the DHHS to provide major metropolitan statistical areas (MSAs) with technical assistance for biological emergency mitigation planning [129]. These MSAs comprise over half of the entire population of the United States [30]. CRI plans focus on reaching the entire target population within only forty-eight hours [28][141]. The CDC sponsored a RAND Corporation evaluation of CRI which found that it had improved mass treatment preparedness [145].

1.2. Vulnerable Populations

When Hurricane Ivan threatened New Orleans, local residents were encouraged to evacuate the city, but, one noted, "They say evacuate, but they don't say how I'm supposed to do that... If I can't walk it or get there on the bus, I don't go. I don't got a car." [123] A year later, vulnerable residents during Hurricane Katrina were impacted not only by the storm, but also by the lack of feasible evacuation operations [42][124]. Although these emergency events were hurricanes rather than biological emergencies, the transportation problems observed remain relevant. If everyone in the population is instructed to visit a clinic for treatment, but some in the population are too far from and lack transportation to a clinic, these people will be unable to receive the mitigation resources which have been allocated and are waiting for them.

People and the social contexts in which they live are non-uniform. Nonetheless, they

play an important role in the success of public health interventions [87]. Individuals who fall into at-risk populations may have additional needs for communication, transportation, and medical care [8]. In July 2004, the *Individuals with Disabilities in Emergency Preparedness* Executive Order was signed to support the safety and security of disabled individuals during disasters [132]. The Pandemic and All-Hazards Preparedness Act of 2006 (PAHPA) specifically mandates that public health and medical needs of at-risk individuals must be taken into account during public health emergency mitigation [135]. A 2008 report by DHHS on the implementation of PAHPA-specified provisions specifically states that "No one should be left behind in emergency and disaster prevention (mitigation), preparedness, response, and recovery, whether an event is natural or man-made." [8] The report defines at-risk individuals as being those who have special needs which may interfere with their ability to receive mitigation resources. These needs are grouped into the following five functional areas: maintaining independence, communication, transportation, supervision, and medical care.

The CDC released a workbook detailing specific vulnerabilities of at-risk individuals and how they could be identified and addressed [29]. Nonetheless, Mastroianni asserts that "Little attention has been given to the impact of social vulnerabilities on the effectiveness of public health preparedness strategies, a shortcoming in both policy-making and in how we think about the ethics of public health." [87] Although many vulnerabilities exist which must be considered, the work presented in this document focuses primarily on vulnerabilities arising from the unavailability of transportation which may lead to differentiated access during biological emergency mitigation.

Quantifying vulnerability and identifying vulnerable populations in demographic data are nontrivial [2]. However, it may be possible to quantify a specific vulnerability stemming from a specific attribute or cause. Although the importance of vulnerability indicators to public health preparedness programs has been recognized, such indicators are rarely incorporated as integral components of preparedness [5]. Further, regional differences in culture may affect what variables are indicators of risk [72]. Local public health departments un-

derstand their communities in ways that may be difficult to capture in demographic data [87]. However, using only subjective perceptions of risk may lead to adverse consequences [72]. Therefore, quantitative tools are needed to assist local public health departments in identifying and including vulnerable populations in their mitigation plans.

There are many approaches in the literature to studying vulnerability, but *vulnerability* is generally defined as being risk of being adversely impacted by some hazard. Although a release of harmful biochemical substances is obviously hazardous, it is assumed (in the context of the research presented in this document) that the entire study population has already been exposed to this release. Mitigation resources to lessen or eliminate the impact of exposure to this release are needed by each individual of the population. Therefore, exposure to the release, combined with the inability to receive mitigation resources, is the the hazard used to establish the vulnerability of individuals in the population. However, since uniform exposure of the entire population is assumed, a barrier to receiving mitigation resources is a vulnerability, and the hazard is the non-receipt of mitigation resources within some emergency-specific time frame.

1.3. Problems to Be Addressed

Many disaster mitigation strategies involve the identification of points accessible to the target population and the implementation of mitigation activities at these points. This has given rise to two similar, yet distinct, approaches to mitigation, both known as PODs. In the field of emergency management, PODs (referred to hereafter in this document as PODsEM for the sake of clarification) are *points of distribution* where the public may acquire life sustaining supplies (e.g. food, water, or blankets) following a disaster [48]. Selection and design of PODsEM must consider traffic congestion concerns and walkability of surrounding areas [83]. In the field of public health preparedness, points of dispensing (PODs) are ad-hoc clinics implemented as part of biological emergency mitigation plans where medical countermeasures such as those provided by the SNS are dispensed to the public. Although PODs have been implemented as part of H1N1 vaccination campaigns [107], the lack of mandated time frames for treatment of entire populations make their concerns and constraints vastly

different from PODs discussed in this document. Depending upon the specific cause of the biological emergency, medical countermeasures could be in the form of antibiotic pills, vaccine injections, or courses of antivirals. The use of PODs during biological emergencies has been recommended by the CDC to alleviate patient load on hospitals, commercial pharmacies, and other health care institutions which are likely to, in such a situation, already be overwhelmed [27].

The design of mass-treatment dispensing campaign plans must rely on an analysis of quantitative data [21] and is a function of the size of the target population and the mandated time frame [27]. A particular response plan is considered to be *feasible* if it supports the application of mitigation resources to the entire target population within this specific, mandated time frame. However, the demographic heterogeneity and non-uniform spatial distribution of the target population complicate plan design by adding demographic and spatio-temporal concerns which are likely to impact plan feasibility. For example, the CDC has recognized that some people will be unable to travel to a POD to receive treatment [27]. In order to achieve feasibility, a plan must include these people by re-allocation mitigation resources or by identifying alternate forms of transportation. Therefore, analysis and development of feasible response plans are intractable without the use of computational tools for processing the large amounts of disparate, quantitative data.

The work presented in this document focuses on the development of algorithms to address transportation vulnerabilities in the context of biological emergency response plans. Available data are explored in an effort to geographically identify and quantify vulnerable populations. New computational methods are developed to determine the reach of existing response plans to include vulnerable populations and to quantify differentiated access resulting from these plans. Public transit system network and schedule data are included and examined in the context of population vulnerabilities and response mitigation plans to further enhance the analyses. Spatio-temporal public transit data are translated into a graph model to facilitate the application of methods developed in the domain of graph theory. Automatic optimization algorithms are developed to minimize differentiated access due to

transportation vulnerabilities by modifying POD placements, public transit stop locations, and public transit schedules. Implementations used in this work are incorporated into an existing response plan analysis framework to provide an easy transition from research and development to actual public health deployment and adoption.

1.4. Contributions of the Research

The RE-PLAN computational framework [101] was developed to facilitate data-driven, quantitative response planning by PHPP without the need for computer programming or GIS expertise. Other computational tools have been created to assist PHPP with response capabilities. However, these tools either focus on problems within each facility (e.g. POD procedures and/or interactions with the public) rather than on quantitative aspects of the overall response plans [24][91] or they distort spatial population data by assuming uniform population densities within the U.S. Census Bureau's defined geographic units [78]. The RE-PLAN framework has been adopted by the public health department of Tarrant County, Texas and has been used to create and modify the county's response plans. Although resulting plans are recognized as a significant step forward in biological emergency response planning, they do not consider the demographic composition of target populations. This may lead to access disparities in plans otherwise deemed feasible.

Work for this research includes the design of algorithms to identify, quantify, and analyze vulnerabilities within target populations using demographic and geographic data. Specifically, vulnerabilities arising from lack of access to transportation resources are addressed. Optimization algorithms to minimize access disparities resulting from specific vulnerabilities have been developed. These algorithms have been implemented and integrated into the RE-PLAN framework for deployment and use at public health departments. Therefore, contributions of this research include computational algorithms for spatial vulnerability analysis and access disparity minimization as well as computational tools for use in public health preparedness planning.

1.5. Overview of the Dissertation

This dissertation has been prepared for a degree in computer science. However, the interdisciplinary nature of the research documented herein requires, in many cases, translation of specific terminology between domains. The specific terms *at-risk* and *vulnerable* have proved to be particularly troublesome [2], necessitating some explanation regarding how these terms are used in this document. For clarification, in this dissertation, *vulnerable* individuals have specific attributes, referred to as *vulnerabilities*, which place them *at-risk* of not receiving mitigation resources. Therefore, the terms *vulnerable individuals* and *at-risk individuals* are used interchangeably in this document, though specific context may make one term preferred rather than the other.

Examples of spatial queries are included as algorithms in this dissertation due to their complexity and to the integral role they play in the computational methodology. However, in order to avoid the complexities of geographic projections and coordinate systems while presenting computational methodologies, spatial reprojections have been omitted from example queries. Nonetheless, these reprojections have been implemented into the actual tools. Specifically, spatial data is stored in the database with spatial reference identifier (SRID) 4323 using the World Geodetic System 1984 (WGS 84) [95] with units in decimal degrees. However, for spatial analysis, data is temporarily reprojected into the State Plane Coordinate System of 1983 [119] with SRID 32138 and units in meters. Examples of spatial analysis in this dissertation focus on data from the Dallas-Fort Worth Metroplex. The projection chosen for analysis is suitable since it is specifically for use with the region described as *Texas North Central*.

The structure of this dissertation is as follows: Chapter 2 includes a review of work related to that of this dissertation, as well as sources from which specific assumptions, parameters, and concepts used in this document were derived. Chapter 3 includes a rudimentary case study to be used as a proof of concept/proof of need. Chapter 4 details the vulnerability quantification process and the data needed for analysis. Chapter 5 explores reach maximization algorithms built on top of quantification methods from Chapter 4. Chapter 6

includes a detailed description of how public transit data is stored and how it can be used to created a directed graph model of the system. Chapter 7 concludes the dissertation with an exploration of the developed tools, their limitations, and their real-world impact on domains such as public health preparedness and emergency management.

CHAPTER 2

LITERATURE REVIEW

Strategies for the distribution of medical countermeasures among treatment facilities have been explored in [80], and a policy of delaying countermeasure distribution to facilities until a better estimate of need distribution can be observed is advocated by Arora, Raghu, and Vinze [7]. Routing and scheduling problems associated with the timely delivery of countermeasures to facilities have been examined in [62]. Nonetheless, countermeasure application to target populations remains a challenging problem [109], leading researchers and localities to drill, develop, evaluate, and improve new response plans [74][97][106][120].

Analysis and optimization methods for inside-clinic queuing and procedures during a biological emergency have been examined in the literature [54][78][79]. A model exploring advantages of specifically targeting travelers or commuters for prophylactic treatment is described in [84]. However, the constraints of this model were placed more on mitigation resources than time, thus making the model inapplicable to the biological emergency mitigation problem explored in this document. The REsponse PLan ANalyzer tool (RE-PLAN) for analyzing and optimizing facility placement with respect to spatially distributed target populations and the available road network infrastructure was developed at the University of North Texas Center for Computational Epidemiology and Response Analysis (CeCERA) [112][113][114]. This tool was implemented as an extensible framework for creating, analyzing, and optimizing response plans using a variety of disparate quantitative data [101]. However, new analysis and optimization methods are needed to address vulnerabilities which exist in the target population.

Lack of access to transportation capable of providing access to mitigation resources may be caused by many different factors. The RE-PLAN framework includes modules to analyze available road network infrastructure with respect to business as usual and/or response plan-generated traffic. However, cars may be insufficient or unavailable to some individuals. Large families or households may not have adequate seating capacity in their cars [42], and

cars in low-income households are generally older and may be unreliable [20]. Walking to a clinic to receive mitigation resources may be impossible due to long distances or difficult terrain. Public transit resources may be inadequate or unavailable to be used as a transportation option. Therefore, mitigation and transportation resources must be analyzed and allocated optimally to minimize access disparities arising from transportation vulnerabilities.

Guidelines for the development of computational models to facilitate quantitative response planning were recommended in [21]. Tools already created using the RE-PLAN framework are consistent with these guidelines. However, as more computationally difficult challenges are approached, approximation algorithms and heuristics must be developed to prevent the computational complexity of new tools from hindering their real-world usability.

2.1. POD Throughput

POD throughput (also known in public health preparedness literature as POD efficiency) is a rate (generally in population units per time unit) at which a POD can treat its assigned population [27]. POD throughput requirements can be determined using population size, mandated time frames, and mass treatment dynamics related to the cause of the biological emergency. These dynamics include the application method of mitigation resources to the population. For example, an emergency stemming from a release of anthrax would likely involve the dispensing of antibiotic pills [67] even though recent studies have found the vaccine to be safe [121]. An emergency stemming from a release of smallpox would likely involve a mass vaccination campaign [71]. In [74] a "head of household" strategy is described for dispensing pills where only a single member of each household would travel to the POD and would obtain enough medications for the entire household. This reduced the number of individuals who must travel to a POD. However, this strategy will not work when mitigation requires vaccination since each individual will have to be treated by a licensed health care professional [40].

2.2. Computational Tools for Biological Emergency Mitigation

The effect of queuing on POD throughput at the POD facility and at the facility's parking lot were examined with respect to non-uniform arrival rates by [10]. The model used

population data at the zip code level as well as different arrival rates drawn from evacuation and POD throughput studies from the literature. Queues in the parking lot resulting from queues at the POD were shown to lead to congestion on the streets near the parking lot. However, an analysis of roadway traffic resulting from this congestion was not conducted. In [66], a Monte Carlo model to study personnel resource needs resulting from non-uniform arrival rates at ad-hoc influenza vaccination clinics led to the conclusion that unstable arrival rates increase the quantity of personnel needed.

A model created using Visual Basic for Applications (VBA) within Microsoft Excel simulated how delaying mass prophylaxis or failing to meet CRI mandated time limits following a aerosolized release of anthrax would affect hospital surge [65]. A hospital triage model following a smallpox or anthrax bio-terrorist attack based on patient severity is described in [105]. A mathematical model of an anthrax bio-terror event presented in [140] examines different release levels, detection methods, and dispensing strategies in order to quantify burden on hospitals and resulting death tolls.

2.3. Addressing Transportation Vulnerabilities Using Public Transportation

Age may be used as an indicator of vulnerability in a population. Elderly individuals (generally over the age of 65) are regarded as being potentially vulnerable, not due to their age, but due to the correlation between advanced age and medical conditions which may limit their abilities to perform activities of daily living, thus causing them to be vulnerable [49]. In addition to the elderly population, the CDC recognizes children under the age of 18 and infants to be at-risk during an emergency [29].

Nutley suggests that automobile-based paratransit be employed to address remotely located transportation vulnerable populations [99]. A survey of eighty-nine transportation agencies revealed that less than 60% of responding agencies have plans for coordination with other agencies in the case of large disasters such as terrorist attacks [12]. Nonetheless, the need for disaster preparedness plans which coordinate with transit agencies may be further recognized through studies of how transit systems may be used in emergency scenarios.

In order for public transit to be an option for a traveler, the Transportation Research Board has identified the following requirements [126]:

- A transit stop must be within walking distance of the trip origin (Spatial availability).
- A transit stop must be within walking distance of the trip destination (Spatial availability).
- The schedule and routes of the transit system are known (Information availability).
- Service at or near the needed times are available (Temporal availability).
- Space is available on the transit vehicle at the desired time (Capacity availability).

The fixed routes and schedules of public transit systems may make certain areas inaccessible to individuals who lack access to a working automobile [137].

2.3.1. Walking Distance Estimations

From several studies, a maximum walking time of five minutes to bus or ten minutes to rail transit has been found. Assuming an average walking speed of 3 miles per hour, maximum walking distances of 0.25 miles to bus or 0.5 miles to rail transit are estimated. However, poor pedestrian environments which make walking less safe (lack of sidewalks) or more strenuous (areas with grades greater than 5%) diminish the distance covered during these times, thus also reducing the maximum walking distances. Further, population characteristics may affect the maximum walking distance of a region. For example, the elderly typically do not walk as far as young adults [126]. Therefore, different estimations of maximum walking distance based on regional or population characteristics should be investigated.

Estimations of maximum walking distance are separate for origin-to-transit and transit-to-destination distances [59][126]. Acceptable walking distance estimations may vary regionally due to factors such as terrain [94]. A model using estimations of walking effort along with actual walking distance was created to calculate equivalent walking distances which could be used to compare regions having different terrains [143]. Walker notes that the walking distance coverage area of a transit stop in a grid of streets should be computed using Man-

hattan distance or by drawing a ring around the stop using a radius which approximates the Manhattan distance coverage area [139]. Badland et al. [11] developed a software tool for urban design and public health planning based on an agent-based model to study the walkability of neighborhoods with respect to existing roadways. The tool includes parameters for variables such as maximum walking time, walking speed, maximum distance, and intersection wait time. Analysis of a survey of residents of twenty-five California housing projects within 0.5 miles of a transit stop suggests that residents seem to focus more on a walking time of five to eight minutes than on a specific walking distance [82]. Shortest path walking distance estimations were calculated using street and walking path data from several sources in [150]. A comparison of the results underscore the importance of using a complete data set and reveal possible advantages of using both freely available and commercial data sources. Characteristics of bus transit networks and stop locations which minimize pedestrian walking distance and promote safety are explored in [81].

2.4. Public Transit Models and Studies

The TRansportation ANalysis and SIMulation System (TRANSIMS) was developed at Los Alamos National Laboratory to provide simulated goal-oriented movements of the population across the transportation network in support of planning, traffic, and environmental research [118]. It was used by a disease outbreak model to simulate diseases in cities such as Portland, OR [14][46]. Direct ridership models of rail and bus use combine walking distance estimates, transit service features, and geographic population and transit data to predict the usage of transit resources [31][32]. A study presented in [82] found that the density of street lights had a greater impact than other factors such as street trees, furniture, or retail shops on whether individuals would walk to transit stations. The company *Urban Design 4 Health* developed tools for the San Diego Association of Governments using block group level data to create maps of health indicators related to transit network characteristics such as walkability and accessibility [127].

The access and accessibility of a public transit system are contrasted in [94] with access being the ability to use the service determined by such factors as acceptable walking

distances and the transit system's proximity to target populations and target destinations, and accessibility being a measure of how reachable a destination transit stop is from an origin transit stop. An extensive literature review regarding the concept of accessibility of public transit is provided in [69], which also describes a new index to quantify how public transit makes employment destinations available to the working population.

A survey investigating how transportation access affects the frequency of health care visits for chronic conditions found that, among rural populations, those with access to private vehicles had on average 1.92 times as many visits per year as those without access. Having a friend of family member who could provide transportation also increased the number of visits per year. However, access to public transit only increased the number of visits per year by an average of four [6]. Nonetheless, a study of cancer patients in New Mexico revealed impaired access to transportation as being a significant factor in non-receipt of cancer therapies [56].

A survey of research quantifying a region's transportation network with respect to the spatial distribution of desirable destinations (in this particular context, referred to as *accessibility*) is presented in [76]. Most relevant to this dissertation are the cumulative opportunity, gravity, and behavioral methods described in this survey. The cumulative opportunity method scores accessibility using the number of destinations within a specified measure of distance. The gravity method is similar to the cumulative opportunity method, but an impedance function is used to account for measured distances to each destination. In the behavioral method, a utility function is used to predict matches between individuals in a region and particular destinations.

An index to quantify accessibility to employment opportunities in the Los Angeles region using estimates of maximum travel time together with geographic population and employment site data was presented by Wachs and Kumagai in [137]. In [4],an agent-based disease spread model incorporating public transit was presented which was used to simulate the spread of influenza in the Greater Toronto Area. Techniques from the study of human-computer interaction were adapted by Schmehl et al. to identify barriers to the use of public transportation by special needs populations [110] . A new domain called Computational

Transportation Science is described as being centered around investigating how a variety of sensor data from different sources may be used to monitor, interpret, control, or manage traffic systems [146]. A parallelized algorithm described in [52] minimizes travel distance of taxi cabs by grouping individuals concurrently requesting service together to share a cab and by constructing an optimal pick-up and drop-off sequence.

2.5. Previous Work on Similar Problems

The set of location-allocation problems are generalized in [35] as being problems which must determine number, location, and capacity of sources given location of, requirements of, or shipping costs to each destination. In a location-allocation article [35] written in or before 1961, Cooper stated that, "The existence and use of digital computers barely makes a dent into some of the combinatorial problems arising when computation is undertaken." In 1963, about two years after this article was published, Gordon E. Moore published his predictions on the exponential growth (with respect to time) of components able to be "crammed" onto integrated circuits [92]. While location-allocation problems remain large and computationally intensive, fifty years of advances in digital computer technology have made more problems reasonably solvable using digital computers.

Methods for determining a spatial distribution of ambulance bases in order to maximize the proportion of the population the ambulances may serve within specified time constraints were explored in [1], and Harewood discusses how under-utilization of resources increases with proportion of population covered by availability of ambulance service [60]. A survey of location identification and analysis methods aimed at emergency medical and fire service officials stressed the importance of analytical approaches to ambulance and firetruck location selection to timely and safe emergency response [55]. Widener and Horner [144] developed a methodology built on the p-median [90] and capacitated-median [138] problems to find near optimal solutions which strategically place and assign mitigation resources near populations which need them following a natural disaster. Megiddo [88] proved the p-median problem to be NP-hard, and Brimberg et al. [22] studied heuristics towards solving the problem.

CHAPTER 3

CASE STUDY

As a case study, a county was chosen and a feasible response plan was created for it using previous methodology implemented as part of the REsponse PLan ANalyzer (RE-PLAN) framework. Using previous methodology, the details of this response plan and its feasibility are explored. Additional data are then integrated into the analysis which reveal plan shortcomings due to the disregard of vulnerabilities in the analysis process. These shortcomings serve to motivate the research presented in this dissertation.

3.1. Response Planning in Tarrant County, Texas as a Case Study

The computational response plan analysis team at the University of North Texas (UNT) Center for Computational Epidemiology and Response Analysis (CeCERA) has been working with Tarrant County Public Health to analyze and optimize their response plans since 2008. The REsponse PLan ANalyzer (RE-PLAN) software developed at CeCERA has been installed at Tarrant County Public Health and has been used to create and analyze response plans for the county. One noteworthy modification that has been made to the county's plans as a result of RE-PLAN analyses is the transition from walk-in Points of Dispensing (PODs) to drive-through PODs.

3.2. Tarrant County Response Plans Deemed Feasible by Previous Work

As a case study, a feasible response plan was created using 40 PODs with 25 drive-through lanes each (for a total of 1,000 drive-through lanes across the entire county). This response plan was created for the purposes of this case study. Other than the use of drive-through booths, any similarities between it and Tarrant County's actual response plans are entirely coincidental. Figure 3.1 shows a map of the response plan rendered through the RE-PLAN graphical interface as well as a legend detailing the symbols on this map. Although traffic congestion may be expected during certain times of the day, this is due more to base

traffic than POD traffic. Therefore, the traffic situation supports the conclusion that this plan is feasible.

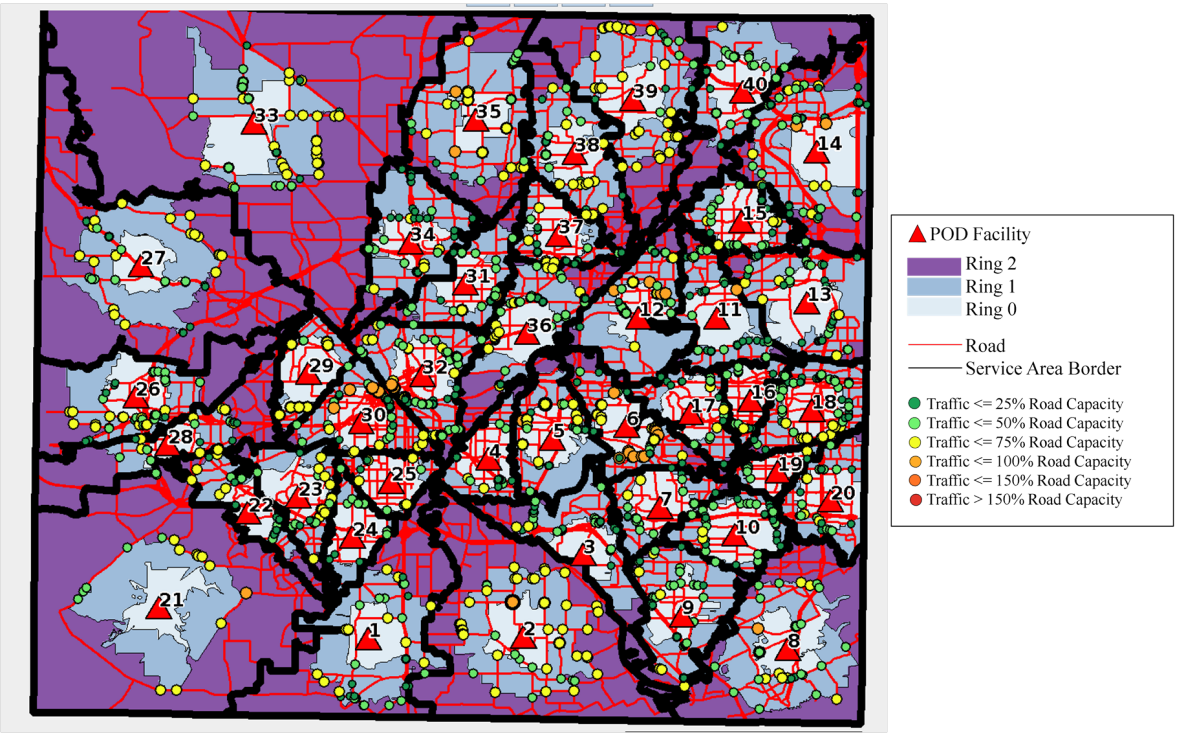


FIGURE 3.1. Map of a feasible response plan for Tarrant County, Texas created using the RE-PLAN tool

Figure 3.2 shows the population distribution across the PODs for this plan. Resource distribution decisions are influenced by population distribution estimations. Therefore, a uniform distribution of population across the PODs simplifies otherwise complex resource distribution decisions.

Figure 3.3 is a screen capture from the RE-PLAN tool showing that all PODs are capable of serving their assigned populations within the forty-eight hour mandated time frame. Assuming only a single person in each car, a car would have to simultaneously enter and exit a POD approximately every five seconds. However, since there are twenty-five drive-through lanes, this means that each car must be served within approximately 125 seconds. A similar computation to compute the time under which each individual must be served yields a service time of 118 seconds. Therefore, this plan is feasible.

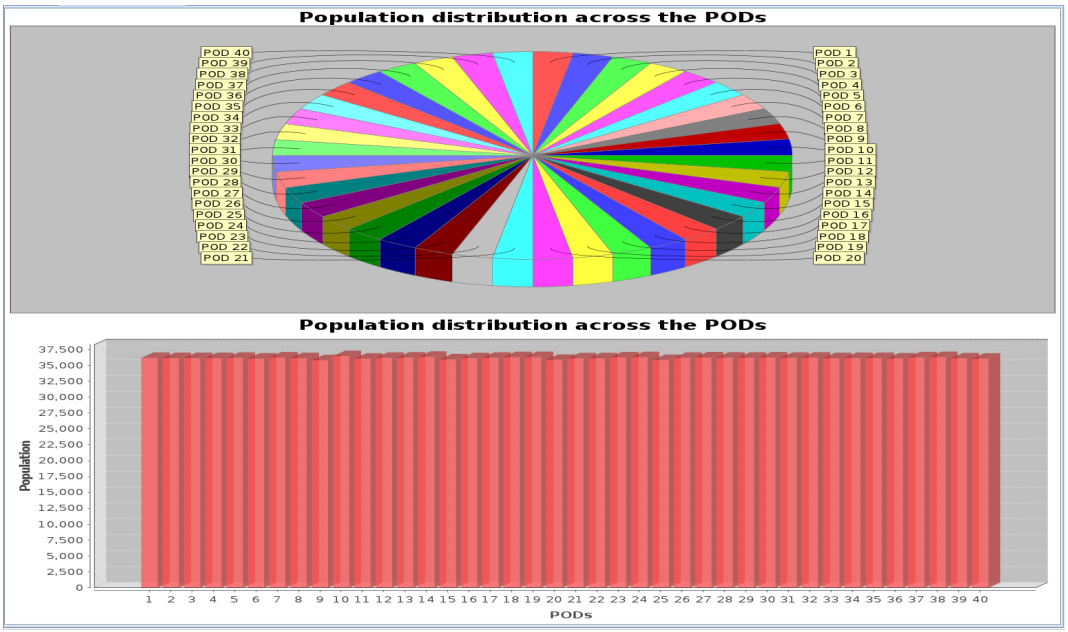


FIGURE 3.2. Population distribution of a feasible response plan for Tarrant County Texas created using the RE-PLAN tool

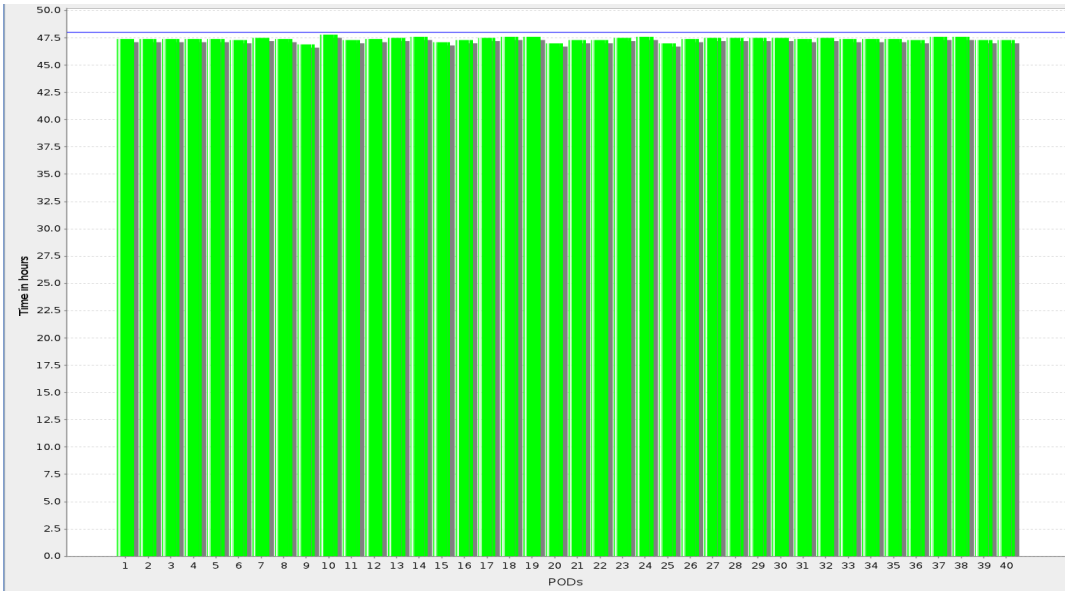


FIGURE 3.3. Time requirements for each POD to serve its population in a feasible response plan for Tarrant County, Texas created using the RE-PLAN tool

3.3. Data Used in Case Study

Although Tarrant County Public Health has been using decennial census population data at the census block level for their planning and analyses, the population data used in

this case study uses data from Five-Year Estimates (2006-2010) of the American Community Survey (ACS) from the U.S. Census Bureau at the block group level. This data source and lower level of spatial granularity were chosen for the additional demographic detail included in the ACS data. For vulnerability analyses below, additional ACS data at the census tract level were used as well as General Transit Feed (GTF) Specification data from Tarrant County's public transit authority *The T*. All road network and traffic data were provided by the North Central Texas Council of Governments (NCTCOG).

3.4. Study Regions of Tarrant County

Although plans from previous methodology were for the entire county, two study regions were selected to underscore the importance of sub-county regional analyses. Two approximately equal-area sections of Tarrant County were chosen for comparison and contrasting of vulnerability indicators. Figure 3.4 shows the population of Tarrant County, Texas block groups per square kilometer, and Figure 3.5 shows the two study regions chosen, their populations, and other data for comparisons. The northwest (NW) study region is labeled as being Non-Urban, and the southeast (SE) study region is labeled as being Urban.

3.5. Vulnerability Analyses of Feasible Response Plans from Previous Methodology

Two different vulnerabilities are studied in this section: vulnerability arising from language difficulties and vulnerability arising from lack of access to transportation. Although only the latter is addressed in the methodology of this dissertation, the former was also considered during preliminary studies and is included here to exemplify the variety of different vulnerabilities which may ultimately lead to access disparities.

3.5.1. Language Vulnerabilities

Language vulnerabilities arise from the inability to effectively receive directions from or communicate with public health personnel during a biological emergency. ACS data distinguishes between two different types of households where English is not the primary language: non-English and No-English. In non-English households, a language other than English is spoken as the primary language. However, there is at least one individual in the

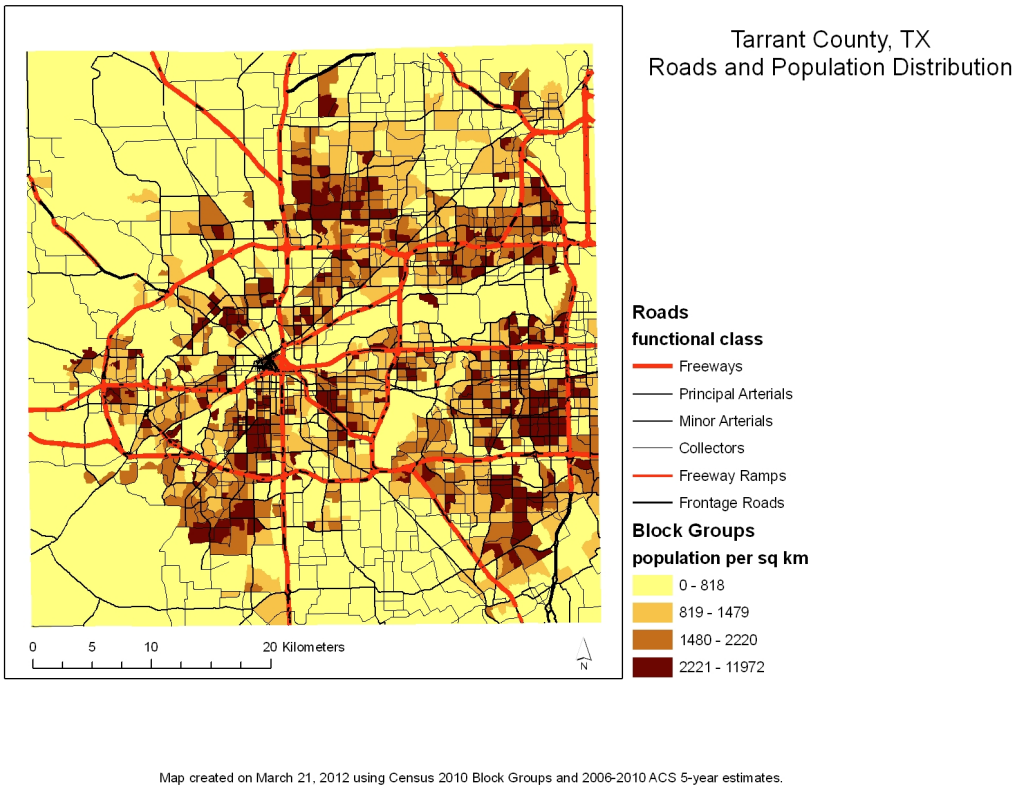


FIGURE 3.4. Road and population distribution across Tarrant County, Texas

household age 14 or over who speaks English well. In no-English households, English is not the primary language of the household, and there is no member of the household over age 14 who speaks English. For the purposes of this vulnerability analysis, no-English households are more vulnerable than non-English households since they have no adult English speaking members. Figure 3.6 shows a map of the NW and SE study regions and the numbers of non-English households in each POD's catchment area, and Figure 3.7 shows a chart of all non- and no-English households across the entire response plan. Lack of consideration for vulnerable populations such as those identified in these figures is likely to make response plans infeasible.

3.5.2. Transportation Vulnerabilities

Transportation vulnerabilities arise from a lack of ability to travel to a POD to receive mitigation resources. This travel may be by a variety of means including by foot, by car, or by public transit. ACS data linking household size to numbers of vehicles available were

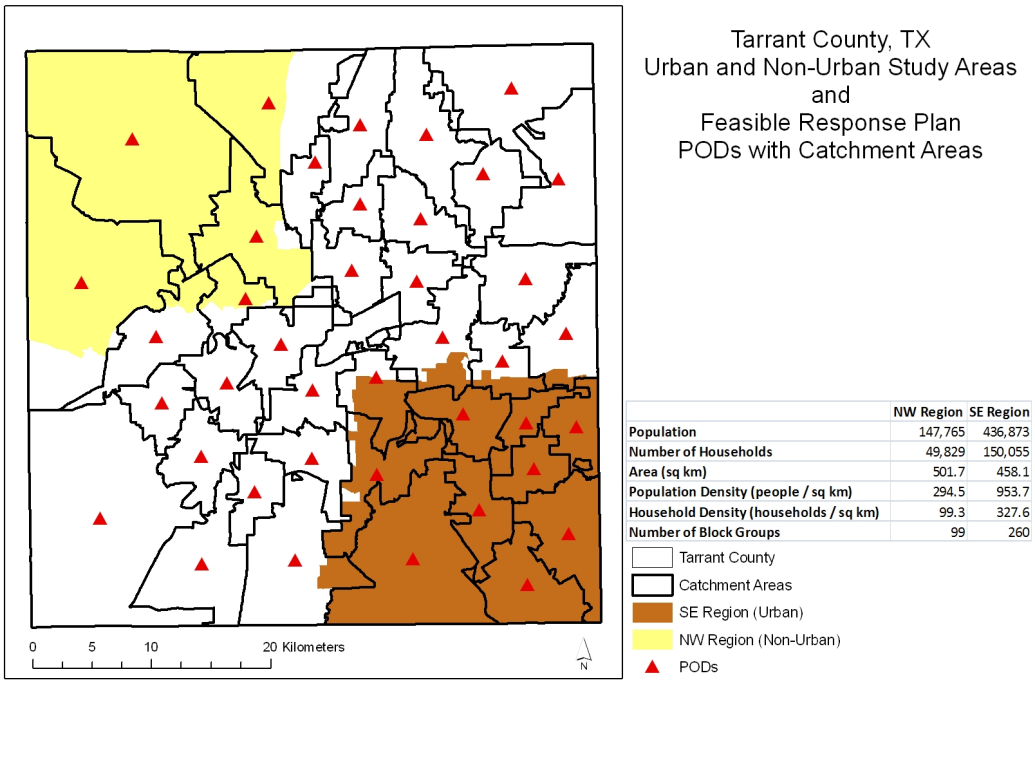
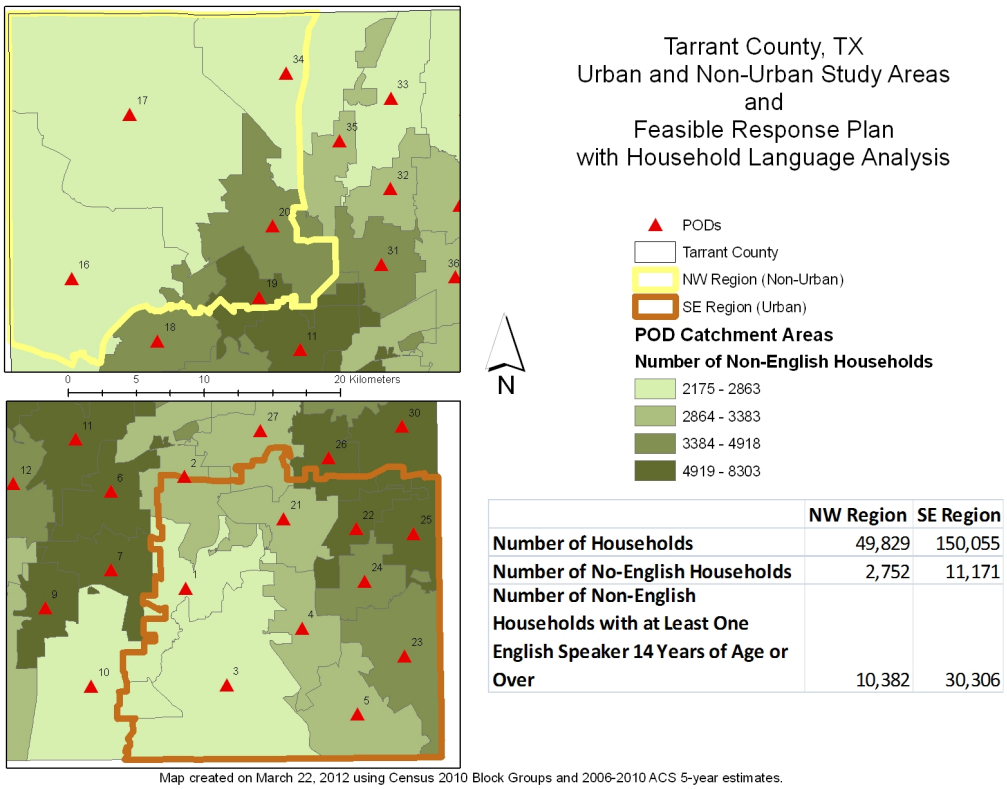


FIGURE 3.5. Tarrant County, Texas northwest and southeast regions chosen for vulnerability analyses displayed with feasible response plan for case study

used to approximate the numbers of individuals without access to a private vehicle across the county. Further, public transit route data were included in the analysis.

Those without access to a private vehicle in their household are ones who are initially considered vulnerable. However, many of these individuals are within walking distance of a POD or public transit which could take them to a POD. Therefore, the most vulnerable individuals are those who lack access to private vehicles, and are too far away from a POD or public transit to walk. For the purpose of this preliminary study, a walking distance of 1 km is assumed. Figure 3.8 shows the two study regions, the walkable area around public transit stops and PODs, and the population without access to private vehicles in each census tract.



	NW Region	SE Region
Number of Households	49,829	150,055
Number of No-English Households	2,752	11,171
Number of Non-English Households with at Least One English Speaker 14 Years of Age or Over	10,382	30,306

FIGURE 3.6. Numbers of non-English households of catchment areas as well as NW and SE study regions

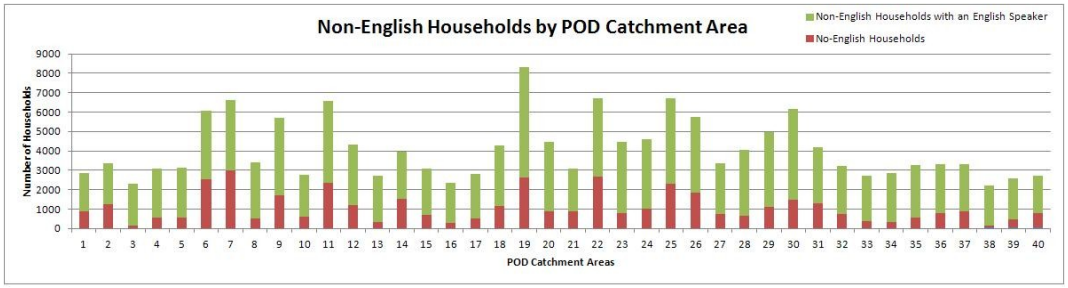
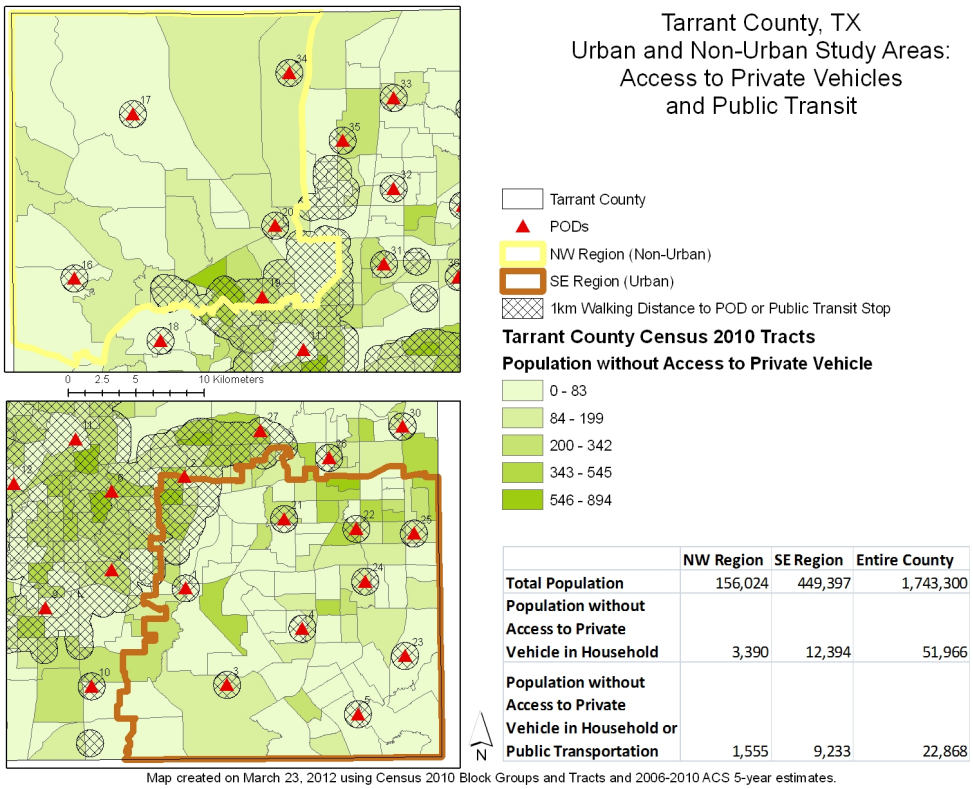


FIGURE 3.7. Numbers of no- and non-English households of catchment areas

3.6. Findings and Discussion

Vulnerable populations identified in the case study are likely to make infeasible the response plans deemed feasible by previous methodology. This motivates the development of new analysis and optimization methods to include vulnerable populations into the response planning process. Although both language and transportation vulnerabilities are explored in this case study, this dissertation focuses on identifying and addressing transportation



	NW Region	SE Region	Entire County
Total Population	156,024	449,397	1,743,300
Population without Access to Private Vehicle in Household	3,390	12,394	51,966
Population without Access to Private Vehicle in Household or Public Transportation	1,555	9,233	22,868

FIGURE 3.8. Transportation vulnerabilities in Tarrant County, Texas study regions vulnerabilities.

CHAPTER 4

MEASURING TRANSPORTATION VULNERABILITY

The Centers for Disease Control and Prevention (CDC) and Pandemic and All Hazards Preparedness Act (PAHPA) recognize that vulnerabilities in the population can serve as barriers placing certain individuals at-risk of not receiving critical medical resources [131]. Recent advances in response plan design and analysis assume ubiquitous access to private transportation and a population homogeneous in its ability to receive POD resources [78][79][101][113][114]. New methodologies must be developed to assess specific vulnerabilities that may lead to access disparities for otherwise feasible response plans.

This chapter focuses on the quantification of vulnerabilities arising from lack of access to transportation resources. The topic of measuring and analyzing transportation vulnerability are incrementally explored and serve as an introduction to Chapter 5 on response plan reach maximization. The "feasible" response plan for Tarrant County, Texas described in the case study presented in Chapter 3 and shown in Figure 3.1 is explored to demonstrate the implementation of the methodologies described. Data sources, example queries, and pseudocode are included to assist the reader in understanding and, if desired, reproducing the results.

4.1. Roadway Traffic Analysis

Traffic conditions must be analyzed in the context of response plans to ensure transportation resources are capable of delivering populations to their assigned PODs within time limitations. Congestion on the road network can lead to diminished arrival and/or departure rates at POD facilities. Lower arrival rates can lead to the under-utilization of POD resources. Lower departure rates can lead to crowds at the POD facilities which can, in turn, lead to further lowering of arrival rates. Therefore, the planning process must include a traffic analysis to assist planners in choosing suitable POD facility locations and distributing traffic control resources.

Road network data was obtained as an Environmental Systems Research Institute

TABLE 4.1. Road Network Shapefile Table Attributes

id	unique identifier of the road segment in the table
street	street name of the road segment
fucnl	functional class of the road segment
streeta	name of street on side "A" of this road segment
streetb	name of street on side "B" of this road segment
pklna	number of lanes in direction from street "A" to street "B"
pklnb	number of lanes in direction from street "B" to street "A"
amcap_ab	A.M capacity of road segment in direction from street "A" to street "B"
amcap_ba	A.M capacity of road segment in direction from street "B" to street "A"
pmcap_ab	P.M capacity of road segment in direction from street "A" to street "B"
pmcap_ba	P.M capacity of road segment in direction from street "B" to street "A"
opcap_ab	off-peak capacity of road segment in direction from street "A" to street "B"
opcap_ba	off-peak capacity of road segment in direction from street "B" to street "A"
areatype	type of area containing road segment

(ESRI) shapefile [45] from the North Central Texas Council of Governments (NCTCOG) for the following Texas counties: Collin, Dallas, Denton, Rockwall, and Tarrant. Data included in this shapefile are listed in Table 4.1. Maximum capacities (in numbers of vehicles per time) were needed to facilitate traffic analysis. However, capacities included were limited to one of three large portions of the day (A.M. period, P.M. period, and off-peak period). Additional information from the Dallas-Fort Worth Regional Transportation Model Description Document [98] was used to determine the fifteen-minute capacities for each road segment. These capacities were a function of the number of lanes of a road segment in each direction, the functional class of the road segment (e.g. freeways, collector, arterial), the relationship between opposite direction lanes of the road segment (e.g. divided or undivided), and the type of area in which the road segment was located (e.g. central business district, suburban residential, rural).

Traffic count data collected by the Thoroughfare Assessment Program (TAP), the City of Allen, the Texas Department of Transportation (TxDOT), and the TxDOT Saturation Count Program were obtained from the NCTCOG as a Microsoft Access database file. In this file, a *Counts* table provided the date, time, total count, and location id for each count record. Although equipment used to collect the traffic counts actually counted axles rather than vehicles, the counts in the dataset obtained were already scaled to represent

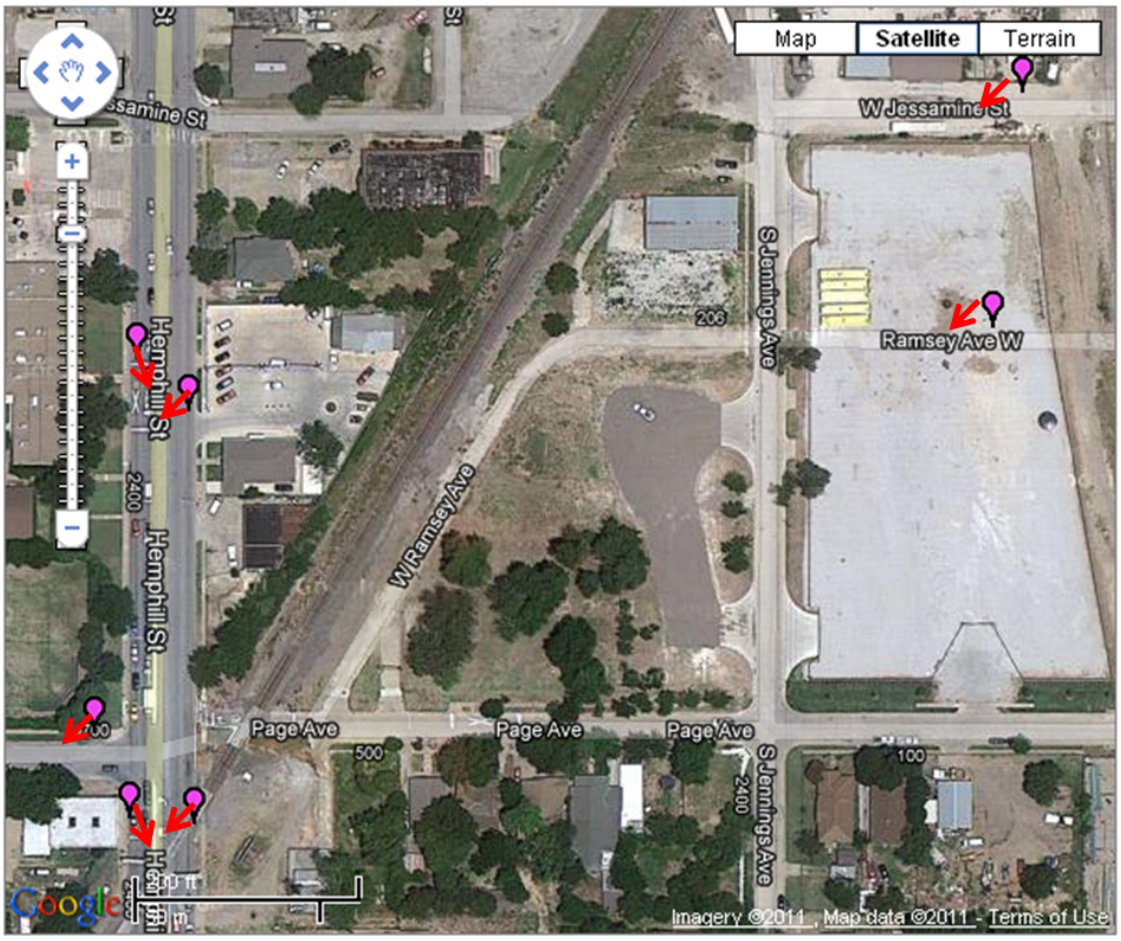


FIGURE 4.1. Example of spatially joining traffic count data points to the nearest road segment

numbers of vehicles by using average axles per vehicle estimations. Location ids were used to relate each traffic count record to a specific location in a *Locations* table also included in the database file. Each location record had an id, longitude, latitude, and roadway field.

Traffic counts obtained include both time and date attributes. Dates were used to obtain the day of the week each count represented, and time was represented as one of the ninety-six fifteen-minute intervals of a twenty-four hour day. Using the ESRI ArcMap software, each point location was spatially joined to the nearest road segment in the road network data as depicted in Figure 4.1. To validate the results of the spatial join, for each pair of joined records from the two tables, street names listed in the *street* field from the road network data were compared with the names listed in the *roadway* field of the traffic count data. Differences were investigated and were largely found to result from data entry

(e.g. "I820" versus ("Interstate 820") or roads which have more than one name. Therefore, this join method was deemed to be accurate.

Although 211,566 traffic count records were obtained from the NCTCOG, only those for which 15-minute intervals were available for the entire calendar day were extracted. From this subset of records, only counts within Tarrant County, Texas were selected for the case study, resulting in a set of 48,192 traffic count records at a total of 107 different locations across the county. This represents a sparse data problem since traffic count data is only available for 107 out of 7,417 road segments in the county. Therefore, traffic count data was categorized as being either *weekday* or *weekend*, and it is assumed that similar road segments experience similar traffic patterns. Specifically, for each fifteen-minute interval during either a weekday or weekend, it is assumed that all roads of the same functional class experience the same traffic per lane.

Traffic count data was used to estimate a business-as-usual traffic, hereafter referred to as *base traffic*. Base traffic estimations indicate the number of vehicles which crossed specific points on specific road segments at specific times. In order to represent traffic patterns over time, a different base traffic estimate is available for each fifteen-minute interval of the day for both weekdays and weekends, resulting in a total of 192 distinct base traffic estimations. Both the day of week (i.e. either weekday or weekend) and the time of day are needed to identify a specific base traffic estimate for analysis.

An algorithm to estimate the movement of individuals toward their assigned PODs within each catchment area was developed [112]. This algorithm identified specific road segments and calculated the number of individuals who must traverse each of them in order to participate in mitigation activities. The resulting information can be used to estimate traffic conditions resulting from implementation of POD plans, hereafter referred to as *POD traffic*. POD traffic is estimated using the population counts for each road segment, the number of people represented by each car, and the number of hours over which individuals have to travel to and from their assigned PODs.

Roadway traffic is derived from at least one of the two distinct sources described

above. Figure 4.2 summarizes the dependency of base traffic and POD on specific parameters. Either POD or base traffic can be analyzed separately, or both POD and base traffic can be combined to identify areas which are likely to become congested if plans are implemented on an otherwise normal day.

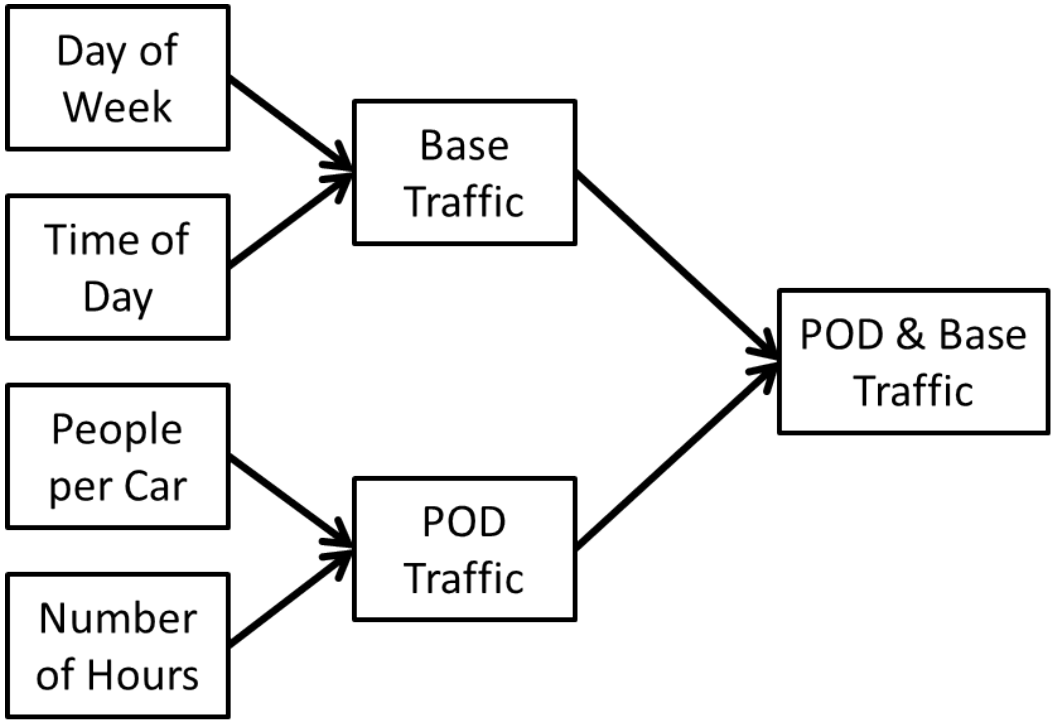


FIGURE 4.2. Traffic Calculation Dependency Diagram

Road segments in the road network data have properties which affect their maximum hourly throughput capacities. To this end, the number of cars per hour estimated for each road segment must be compared with the segment's maximum capacity. Estimated traffic on each road segment is then classified into one of the following six classes:

- estimated_traffic $< 25\%$ capacity
- $25\% \leq$ estimated_traffic $< 50\%$
- $50\% \leq$ estimated_traffic $< 75\%$
- $75\% \leq$ estimated_traffic $< 100\%$
- $100\% \leq$ estimated_traffic $< 150\%$
- estimated_traffic $\geq 150\%$

Colored dots referred to as *crossing points* are placed on the map to indicate estimated traffic conditions on specific road segments. Colors are indicative of the classification system specified above.

4.2. Identification and Quantification of Transportation Vulnerable Populations

Transportation vulnerability can be quantified by examining the spatial and demographic distribution of vulnerability indicators with respect to the locations of mitigation resources. Available public transit resources must be integrated into the calculations to examine what role they may play in connecting otherwise vulnerable populations with the resources they need. Data sources used to identify these indicators are shown in Table 4.2. Analysis results may be represented as risk maps [68] and used in the evaluation of existing response plans.

TABLE 4.2. Data Sources for the Quantification of Transportation Vulnerability

Vulnerability Indicators	**Data Source**
Transportation	Regional governments [98]
Public Transit	Local transit authority
Private car ownership per household	American Community Survey 5-year estimates
Elderly (65+)	American Community Survey 5-year estimates

4.2.1. Identifying Vulnerable Populations

Data concerning the availability of private vehicles in each household is available spatially at the 2010 decennial census tract level through the U.S. Census Bureau's American Community Survey table **B08201**. Although the universe of this data is the set of households, data is classified by the number of people in each household using the following scheme: *1-person*, *2-person*, *3-person*, and *4-or-more-person* households. The number of individuals without access to a private vehicle in each tract is estimated under the assumption that households of four or more individuals consist of exactly four individuals and by summing the number of individuals represented by each class in each tract. The tract geometry, total population, and estimate of individuals without access to a private vehicle in their household

is stored in database table *scenario_name_population_blocks* as shown in Table 4.3. Quantile classification is used to classify each tract in the region into one of five different classes representing the distribution of transportation vulnerability. A choropleth map using these classes is automatically generated and rendered in the RE-PLAN map interface. A screen capture of the map generated for Tarrant County, Texas is provided in Figure 4.4 using the color representations provided in 4.3.

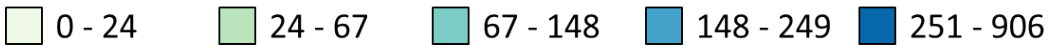


FIGURE 4.3. Color to value mappings used for all choropleth maps of vulnerability. Values are in units of individuals.

4.2.2. Preparing Public Transit Data

The effect of public transportation on access disparities is analyzed using local General Transit Feed (GTF) data, which consists of transit routes, schedules, and stop locations. The GTF specification was originally created by developers at Google and is used by such tools as Google Maps and Bing Maps to automatically provide public transit directions to mobile

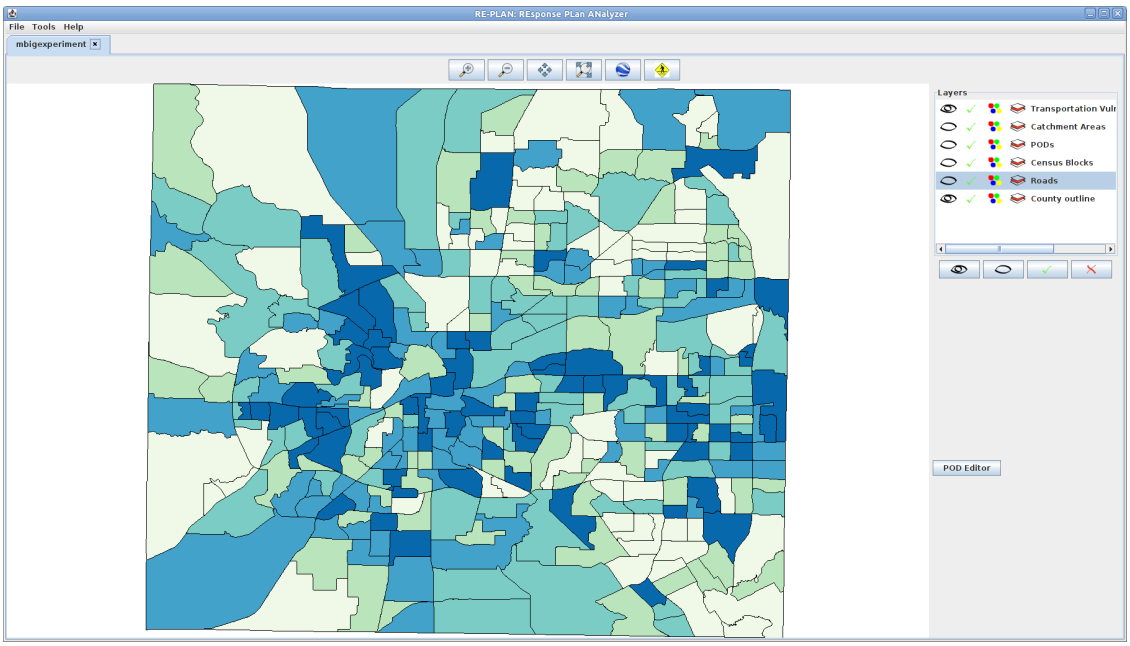


FIGURE 4.4. Tarrant County, Texas Vulnerable Population Quantile Classification

TABLE 4.3. Database Tables Used in PostgreSQL Queries and Columns Necessary for Computational Analysis

Table	Column	Column Description
scenario_name_population_blocks	the_geom	geometry of the tract polygon
	centroid	geometric centroid of the tract polygon
	logrecno	unique identifier of tract within the county
	p0010001	total population of the tract
	total_none	estimate of individuals in tract without access to a private vehicle
tarrant_gtfs_stops	the_geom	point location geometry of transit stop
	stop_name	name of the transit stop
scenario_name_pods	id	unique identifier of POD record
	location	geometry of POD's point location
	status	whether the POD is in use in current plans
	catchment_area	catchment area this POD is within
scenario_name_catchment	id	unique identifier of catchment area
	the_geom	geometry of catchment area polygon

and online users [57][89]. Transit data for Tarrant County was obtained from the Fort Worth Transportation Authority *The T*.

The GTF specification uses comma-separated values (CSV) files to store transit network data [57]. Figure 4.5 depicts the relationship between entities stored in the *stops*, *stop_times*, *trips*, and *routes* files. Each route is comprised of a set of trips. Each trip is comprised of a set of stops and a set of stop times. Each stop has a name (e.g. "Commerce & 5th - North Bound") and a geographic location (given as longitude and latitude coordinates). Stop times connect trips to stops at specific times. Although importing of non-spatial data to the PostgreSQL database is straightforward, conversion tools had to be developed to import spatial vector data from the GTF CSV files into the database.

As shown in the GTF specification relationships diagram, a single stop may be used by more than one route. The GTF specification includes several optional files which were

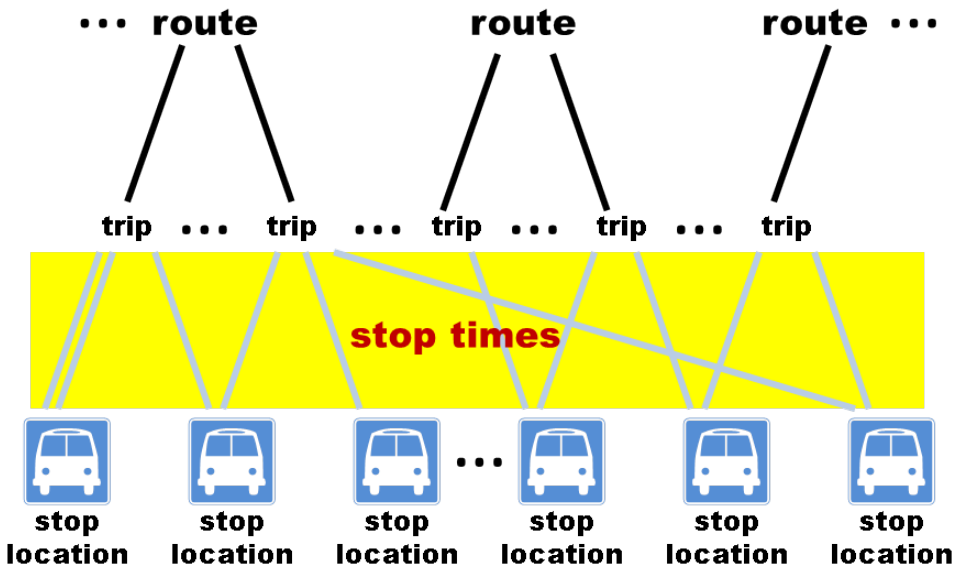


FIGURE 4.5. General Transit Feed Specification Entity Relationships

included in data obtained from *The T*. One of these files is the *transfers* file, which associates pairs of stop locations with a minimum transfer time estimate. In the case of data from *The T*, seventy-six transfer records are included, and all of them have a minimum transfer time of 0. Figure 4.6 shows an example of five stop locations which are all listed in the transfers file as being associated with each other. Further examination of the *transfers* file reveals that twenty-four of the seventy-six records are used to create pairwise associations between these five stops. The stop data describes all five of the stops as being part of the *East Side Transfer Center*. Using the Google Earth measurement tool, it can be observed that walking distance between pairs of stops at this location are less than 100 feet. Four routes are represented among these five stop locations. In order to simplify the analysis without reducing accuracy of the methodology, these five transit stops were collapsed into a single stop location record which is associated with all four transit routes. Converting the five point locations into a single *multipoint* feature was considered. However, a single point location central to the set of points was chosen in order to maintain consistency of the geometry attribute among records in the table.

FIGURE 4.6. Five transit stops of the East Side Transfer Center, *The T*, Tarrant County, TX

4.3. Evaluating Reach and Efficacy of Existing Response Plans

The spatial distribution of vulnerability resulting from lack of access to private transportation must be assessed with respect to the placement of each POD p in the set of PODs P of the existing response plans. Multiple definitions of maximum walking distance d_w can be used to represent limitations of different demographic groups [126]. In order to assess the impact of continued public transit service during biological emergency mitigation, algorithms must be developed to estimate plan reach with and without the availability of public transit

resources.

4.3.1. Evaluating Reach and Coverage Area of Existing Response Plans without Public Transit

The coverage area of existing response plans [139], given specific POD locations, is estimated using definitions of d_w [94][143]. A plan's reach of vulnerable populations can be calculated by overlaying the coverage area and the spatial distribution of vulnerable populations [82] as described in Figure 4.7. Although this figure specifically shows the computation of reach following directly from the computation of the coverage area, these two metrics were computed separately in the implementation of this particular component to facilitate the reuse of programming code and generalization of database queries.

An example database query to calculate the reach of vulnerable populations of a POD facility with $id = 6$ within its catchment area is provided in Algorithm 1. This query starts by selecting all population blocks within a specific catchment area (in this case, the catchment area with $id = 6$). From this set of population blocks, it selects only population blocks whose geometric centroids are within $d_w = 1000$ meters of the POD facility and sums their vulnerable populations. This sum represents the total reach of the POD.

The coverage area is calculated and provided with and without respect for population block boundaries. For simplicity, the coverage area which does not respect population block boundaries is hereafter referred to as the *coverage area buffer*, and the coverage area which does respect the population block boundaries is referred to as the *coverage area blocks*. The sequence of example queries in Algorithm 2 show how coverage areas are calculated without respect for population blocks. The first query creates a table *scenario_name_coverage_wo_trans* containing a walking distance buffer of $d_w = 1000$ meters around each active POD location (designated active by having $status =$'$true$'). The second query alters this table to add a new geometry field to hold the coverage areas of each POD clipped by the catchment area boundaries. The third query performs this clipping operation and updates the newly created field in the table.

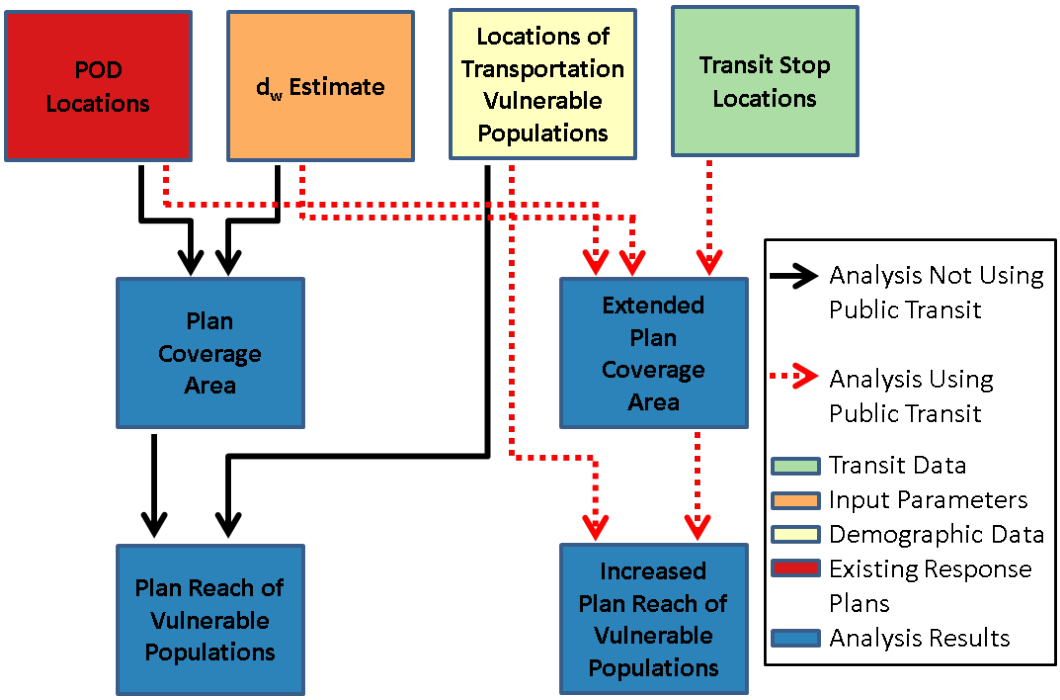


FIGURE 4.7. Procedure and data requirements for calculating POD reach of vulnerable populations

Input: An estimate of acceptable walking distance $d_w = 1000$ meters, POD $id = 6$, and tables of available facilities, catchment areas, and population blocks.

Output: The reach of POD with $id = 6$ given the current d_w estimate.

```
SELECT
    SUM(cen.total_none) AS reach
FROM
    (
        SELECT
            block.*
        FROM
            scenario_name_population_blocks AS block,
            scenario_name_catchment AS cat,
            scenario_name_pods AS pod
        WHERE
            pod.id=6 AND
            cat.id=pod.catchment_area AND
            ST_WITHIN(block.centroid,cat.the_geom)
    ) AS cen,
    scenario_name_pods AS pod
WHERE
    pod.id=6 AND
    ST_DISTANCE(cen.centroid,pod.location)<1000.0;
```

Algorithm 1: Example query to calculate reach of a POD facility with $id = 6$ within its catchment area with $d_w = 1000$ meters.

Input: An estimate of acceptable walking distance $d_w = 1000$ meters and tables of available facilities and catchment areas.

Output: The coverage area of all active PODs clipped by their catchment areas.

```
SELECT
   pods.catchment_area AS pod,
   ST_BUFFER(pods.location, 1000.0) AS coverage_wo_respect_ca
INTO
   scenario_name_coverage_wo_trans
FROM
   scenario_name_pods AS pods
WHERE
   pods.status='true';

ALTER TABLE
   scenario_name_coverage_wo_trans
ADD COLUMN
   clipped_coverage geometry;

UPDATE
   scenario_name_coverage_wo_trans AS cov
SET
   clipped_coverage = ST_INTERSECTION(
      cov.coverage_wo_respect_ca,
       cat.the_geom)
FROM
   scenario_name_catchment AS cat
WHERE
   cat.id=cov.pod;
```

Algorithm 2: Example sequence of queries to calculate coverage area of active PODs with $d_w = 1000$ meters without respect for population block boundaries.

The coverage area created using the queries in Algorithm 2 are used to select population blocks whose geometric centroids are within the coverage area buffer using the query in Algorithm 3. This results in a set of population blocks which represent the coverage area of the POD. Blocks representing the coverage area for each catchment area are dissolved into a single polygon for each catchment area using the aggregate version of a PostGIS union function.

Input: A table representing the coverage area buffer and a table of population blocks.

Output: Coverage area of PODs with respect to population block boundaries.

```
ALTER TABLE
   scenario_name_coverage_wo_trans
ADD COLUMN
   cen_cov geometry;

UPDATE
   scenario_name_coverage_wo_trans AS cov
SET
   cen_cov=temp.the_geom
FROM
   (
      SELECT
         pod,
         ST_MULTI(ST_UNION(the_geom)) AS the_geom
      FROM
         (
            SELECT
               cen.logrecno AS logrecno,
               cov.pod AS pod,
               blocks.the_geom AS the_geom
            FROM
               scenario_name_coverage_wo_trans AS cov,
               scenario_name_census_blocks AS blocks
            WHERE
               ST_WITHIN(blocks.centroid,cov.clipped_coverage)
         ) AS cen_cov
      GROUP BY
         pod
   ) AS temp
WHERE
   cov.pod=temp.pod;
```

Algorithm 3: Example sequence of two queries to convert the coverage area to respect population block boundaries and save the results.

A screen capture of the RE-PLAN interface showing the coverage area buffers of the response plan is provided in Figure 4.8, and one showing the coverage area blocks is provided in Figure 4.9. A map showing the vulnerable population outside of the reach of the plan is provided in the screen capture in Figure 4.10. This plan's reach of transportation vulnerable individuals (i.e. individuals without access to a vehicle in their households) without public

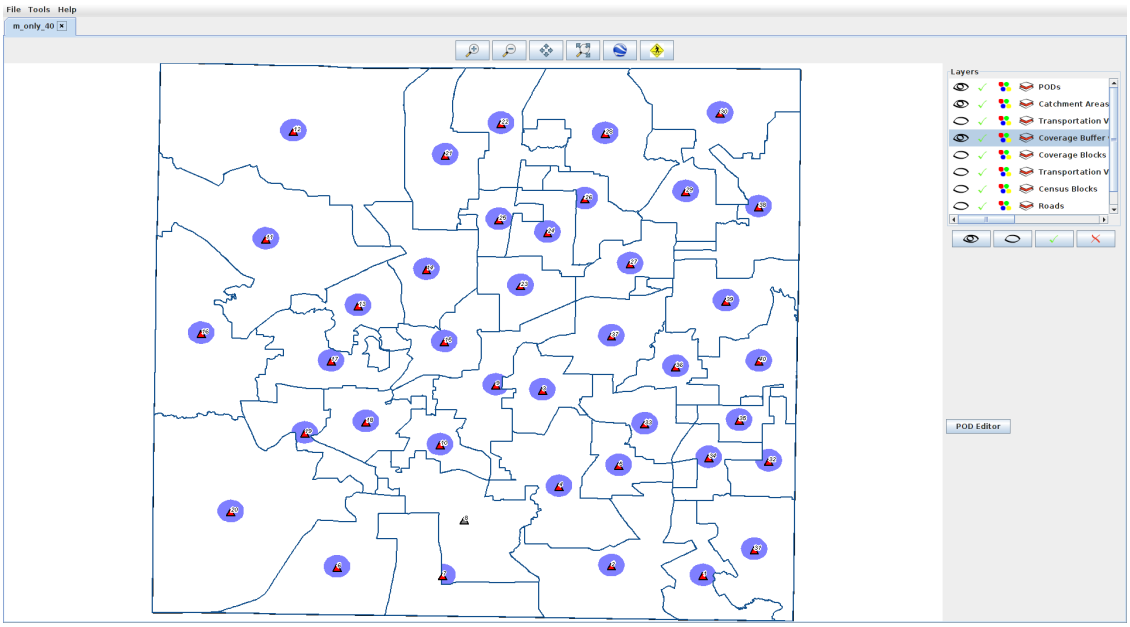


FIGURE 4.8. Screen capture of RE-PLAN with coverage area buffers for $d_w =$ 1 km around forty PODs in Tarrant County, Texas

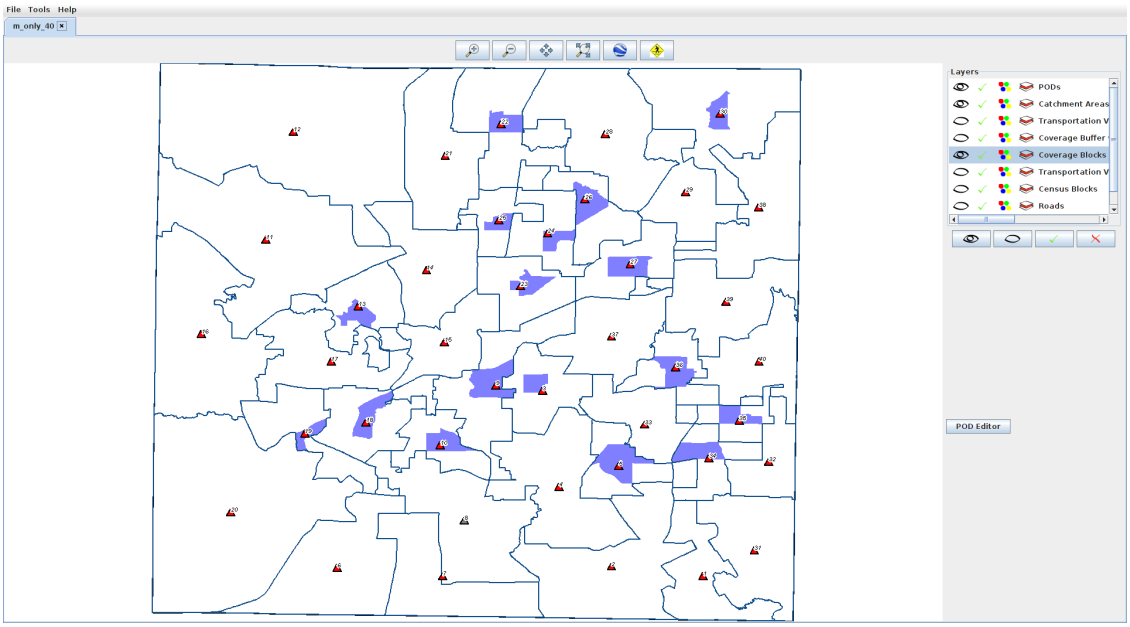


FIGURE 4.9. Screen capture of RE-PLAN with coverage area blocks for $d_w = 1$ km around forty PODs in Tarrant County, Texas

transit is calculated as 4,558, representing approximately 8.5% of the county's transportation vulnerable population.

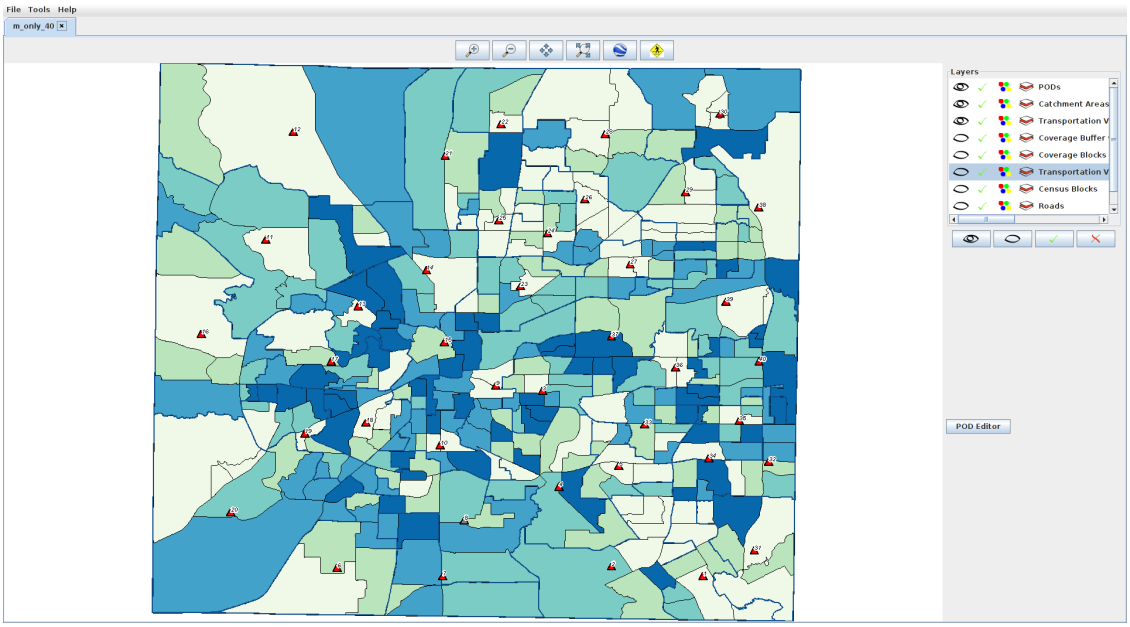


FIGURE 4.10. Transportation vulnerable population of Tarrant County, Texas at-risk of not being able to reach a POD in the scenario if public transit is unavailable

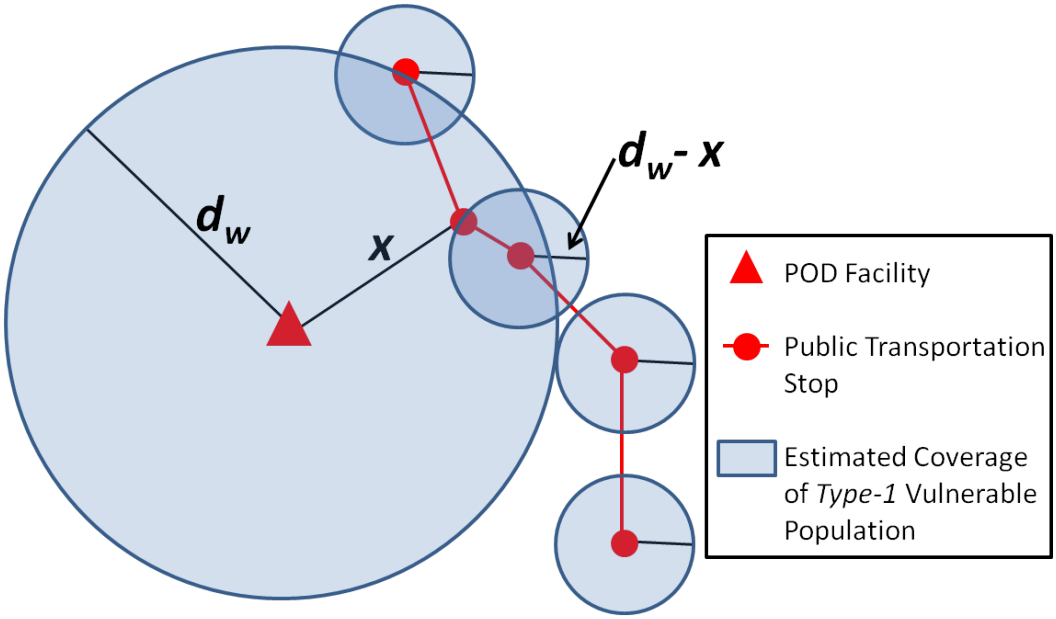


FIGURE 4.11. How walking distance coverage may be extended using public transit

4.3.2. Evaluating Reach and Coverage Area of Existing Response Plans with Public Transit

Public transit stop locations can be used in conjunction with existing POD locations and estimations of acceptable walking distances to increase the reach of PODs. Existing

response plans include a set of PODs P, and the public transit system includes a set of stops S. The distance between a POD p_i and a stop s_j is $d(p_i, s_j)$. For all $p_i \in P$ and $s_j \in S$, if $d(p_i, s_j) < d_w$, then all transit stops in the set $S - s_j$ may be used to extend the walking distance coverage of POD p_i. For each POD p_i, the transit stop s_j which has the minimum $d(p_i, s_j)$ value should be used in order to maximize the extended walking distance coverage area. However, only transit stops within a particular POD's catchment area should be considered in order to maintain the populations' assignments to catchment areas. Let x be the distance between a POD and the closest transit stop in its catchment area. Increased reach of vulnerable populations resulting from the extension of walking distance coverage can then be quantified. Although the walking distance coverage area C_{p_i} of a POD p_i may be calculated as $C_{p_i} \leq d_w^2\pi$ without use of public transit, public transit may be used to extend it to $C_{p_i} \leq d_w^2\pi + (d_w - x)^2(|S| - 1)\pi$ as illustrated in Figure 4.11. A summary of how a plan's reach of vulnerable populations are calculated from input data is depicted in Figure 4.7. Thus, deficiencies of existing response plans with respect to specific vulnerable populations can be identified.

The example query provided in Algorithm 4 calculates the distance x from a POD to the closest transit stop in its catchment area. This query is executed for each POD to determine the x of each POD. These values are compared with the d_w estimate to determine if public transit can be used to extend the coverage area of each POD.

Input: A POD $id = 5$, tables of PODs with their locations, catchment areas, and public transit stops.

Output: The distance between the POD with $id = 5$ and the closest transit stop in its catchment area. If not transit stops exist in the POD's catchment area, no records are returned.

```
SELECT
   stop.stop_id AS stop_id,
   cat.id AS catchment_area,
   ST_DISTANCE(pod.location,stop.the_geom) AS distance
FROM
   tarrant_gtfs_stops AS stop,
   scenario_name_catchment AS cat,
   scenario_name_pods AS pod
WHERE
   pod.id=5 AND
   ST_COVERS(cat.the_geom,stop.the_geom) AND
   cat.id=5
ORDER BY
   distance
LIMIT 1;
```

Algorithm 4: Example query to calculate the distance x from a POD with $id = 5$ to the closest transit stop in its catchment area.

The next step is creating the walking distance coverage area buffer and blocks for each POD. If for a particular POD $d_w - x \leq 0$, then the coverage area buffer, blocks, and reach of vulnerable populations are calculated using the methodology for analysis without public transit. Otherwise, the example query in Algorithm 5 employs the methodology depicted in Figure 4.11 to calculate the coverage area buffer, blocks, and reach of vulnerable populations.

A screen capture of the RE-PLAN interface showing the extended coverage area buffers of the response plan is provided in Figure 4.12, and one showing the extended coverage area blocks is provided in Figure 4.13. A map showing the vulnerable population outside of the extended reach of the plan is provided in the screen capture in Figure 4.14. This plan's reach of transportation vulnerable individuals extended by public transit is calculated as 15,867, representing approximately 29.7% of the county's transportation vulnerable population.

4.3.3. Estimations of Maximum Walking Distances and the Inclusion of Specific Populations

Methods described in this chapter rely on estimations of d_w to determine the reach of specific populations. An *inclusion property* can be observed which guarantees the coverage of specific populations under certain conditions. Increasing the maximum walking distance

Input: A POD $id = 1$, its $d_w - x = 972.1$ meters, tables of PODs with their locations, catchment areas, and public transit stops.

Output: The reach and walking distance coverage blocks of the POD.

```
SELECT
   sum(blocks.total_none) AS reach,
   1 AS catchment_area
FROM
   scenario_name_population_blocks AS blocks,
      (
         SELECT
            ST_INTERSECTION(buffer.single, cat.the_geom) AS coverage_area
         FROM
            scenario_name_catchment AS cat,
            (
               SELECT
                  ST_UNION(stop_buffer.single,pod_buffer.pod_buff) AS single,
                  stop_buffer.catchment_area AS catchment_area
               FROM
                  (
                     SELECT
                        stop_buffers.id AS catchment_area,
                        ST_MULTI(ST_UNION(stop_buffers.the_geom)) AS single
                     FROM
                        (
                           SELECT
                              cat.id AS id,
                              ST_BUFFER(stop.the_geom,972.1) AS the_geom
                           FROM
                              tarrant_gtfs_stops AS stop,
                              scenario_name_catchment AS cat
                           WHERE
                              ST_COVERS(cat.the_geom,stop.the_geom) AND
                              cat.id=1
                        ) AS stop_buffers
                     GROUP BY
                        stop_buffers.id
                  ) AS stop_buffer,
```

```
                    (
                        SELECT
                            ST_BUFFER(pod.location,1000.0) AS pod_buff
                        FROM
                            scenario_name_pods AS pod
                        WHERE pod.id= 1
                    ) AS pod_buffer
                ) AS buffer
            WHERE
                cat.id= 1
        ) AS clipped_buffer
    WHERE
        ST_COVERS(clipped_buffer.coverage_area,blocks.centroid)
    GROUP BY
        catchment_area ;
```

Algorithm 5: Example query to calculate the coverage blocks and vulnerable population reach of a POD with $id = 1$ given $d_w - x = 972.1$ meters.

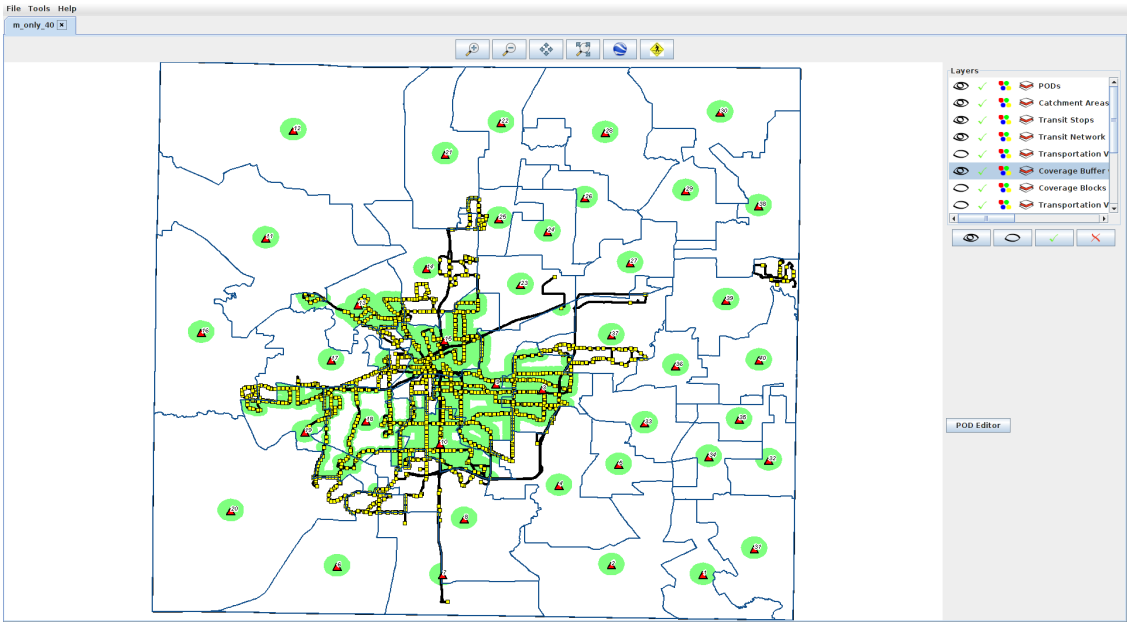


FIGURE 4.12. Screen capture of RE-PLAN with coverage area of forty PODs in Tarrant County, Texas extended using public transit and calculated using $d_w = 1$ km. Public transit routes and stops are also displayed on the map as a reference.

estimate can neither result in a decrease in the total covered population no the uncovering of a population already covered using existing distance estimates.

LEMMA 4.1. *Assuming stationery populations, facilities, catchment areas, and transit stops,*

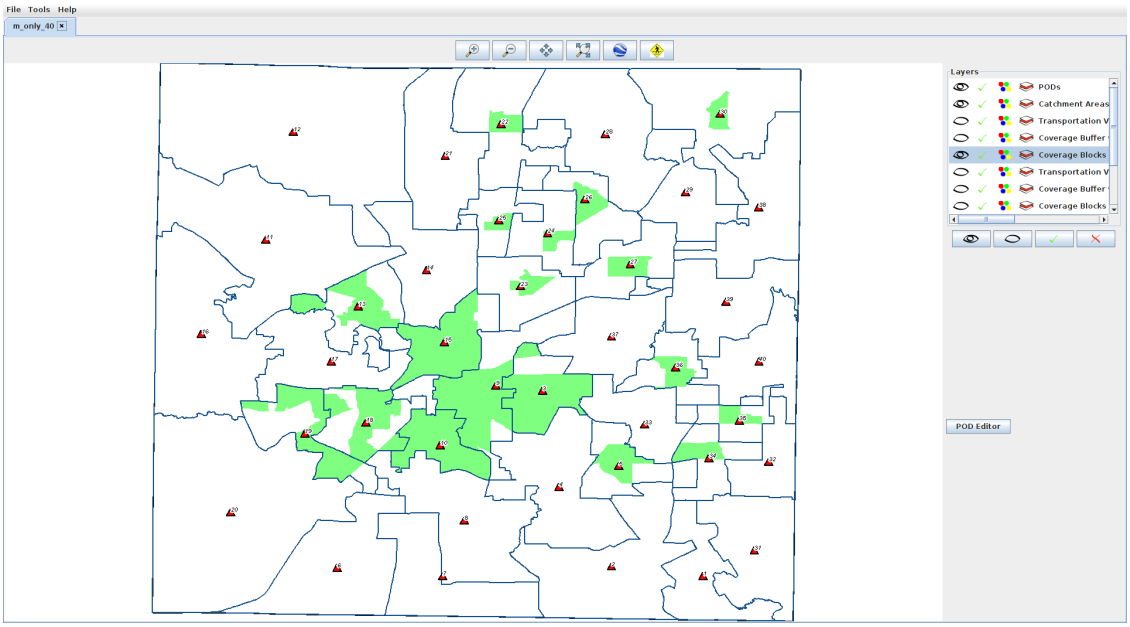


FIGURE 4.13. Screen capture of RE-PLAN with coverage area blocks around forty PODs in Tarrant County, Texas extended using public transit and calculated using $d_w = 1$ km

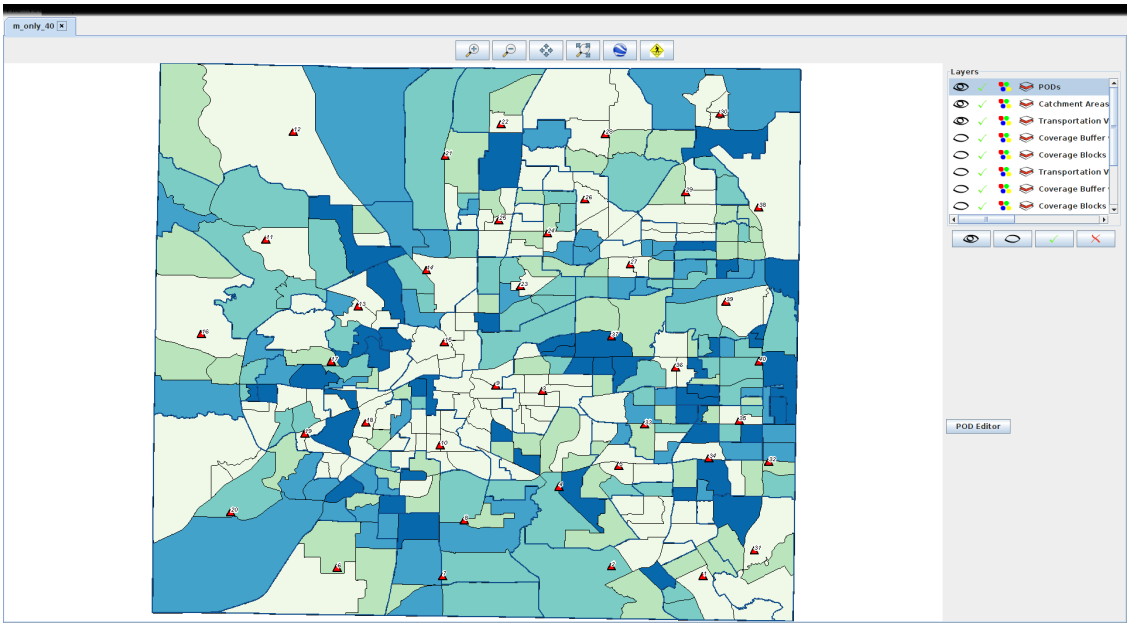


FIGURE 4.14. Transportation vulnerable population of Tarrant County, Texas at-risk of not being able to reach a POD in the scenario if public transit is available

if a population pop_i *is covered by a facility location* p_j *given a maximum walking distance* d_{w_1}*, this implies that* pop_i *will be covered by* p_j *given a maximum walking distance* d_{w_2} *if* $d_{w_1} \leq d_{w_2}$.

PROOF. Suppose a population pop_i is included in the reach of its assigned POD using a walking distance estimate of d_{w_1}. Also suppose there is a second walking distance estimate $d_{w_2} \geq d_{w_1}$.

Without the use of public transit, superimposing walking distance coverage areas of the two estimates on the same map reveals a pair of concentric circles as depicted in Figure 4.15. The smaller circle corresponds to d_{w_1} and is completely covered by the larger circle which corresponds to d_{w_2}. Therefore, any population covered by d_{w_1} must also be covered by d_{w_2}.

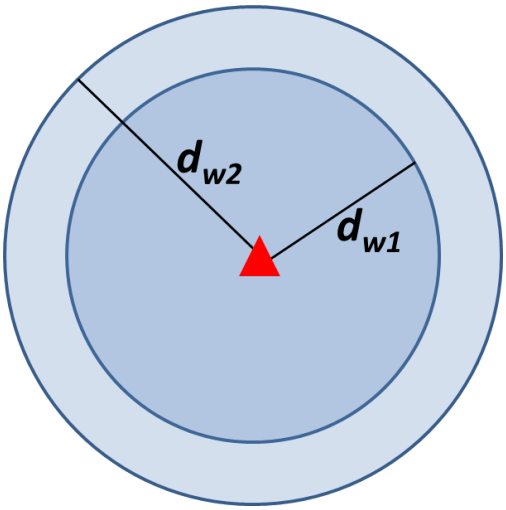


FIGURE 4.15. Diagram showing that, given two estimate of maximum walking distance d_{w_1} and d_{w_2}, if $d_{w_1} < d_{w_2}$, then the walking distance coverage area or d_{w_2} completely covers the walking distance coverage area of d_{w_1}.

If public transit is included, concentric circles may be observed around transit stops as well as the facility. Concentric circles around transit stops follow the same pattern as those around the facility. Therefore, Lemma 4.1 must also hold when considering public transit.

□

COROLLARY 4.2. *A facility's reach of vulnerable populations is non-decreasing with increasing estimates of* d_w.

A realistic estimate of acceptable walking distance must be based on population characteristics and the *walkability* of pedestrian environments [126]. The methods described

in this chapter can be used to identify areas where participation can be most greatly expanded by increasing an area's acceptable walking distance estimates. An increased d_w can be accomplished through the improvement of an area's *walkability*.

CHAPTER 5

MAXIMIZING REACH OF TRANSPORTATION VULNERABLE POPULATIONS

The signing into law of the Pandemic and All-Hazards Preparedness Reauthorization Act (PAHPA) on March 13, 2013 serves as a directive to public health and emergency planners that addressing the vulnerabilities of populations remains a priority . New methodologies must be developed to address specific vulnerabilities in order to minimize access disparities which could hamper the efficiency of otherwise feasible response plans. To facilitate their real-world impact, these new methodologies must be integrated into a framework which is accessible to and usable by public health and emergency planners.

Methodology described in Chapter 4 identified and analyzed the transportation vulnerable population of specific regions, and data representing the existing public transit system have been integrated into the analysis of transportation access disparities. Existing response plans which do not consider transportation vulnerabilities in the population must be modified to adjust and optimize POD placement and public transit infrastructure to maximize response plan reach of transportation vulnerable populations [3][41][86][102]. The optimization methods described in this chapter are applied incrementally to an existing response plan in an effort to maintain the plan's key characteristics such as the number of PODs, the geographic boundaries of catchment areas, and the population assigned to each catchment area.

The response plan described in the Case Study and used in Chapter 4 was deemed feasible by previous methodology and is used as a basis for the coverage maximization examples. This response plan consists of 40 PODs, each with 25 drive-though lanes. The RE-PLAN equal population partitioning algorithm was used to create an additional 120 POD locations in a new RE-PLAN scenario. The status of these locations were toggled to *off*, exported as a Comma Separated Values (CSV) file, and imported into the scenario for the "feasible" response plan. This resulted in three to six potential POD locations within each catchment area. A screen capture from RE-PLAN scenario with these POD locations

is provided in Figure 5.1.

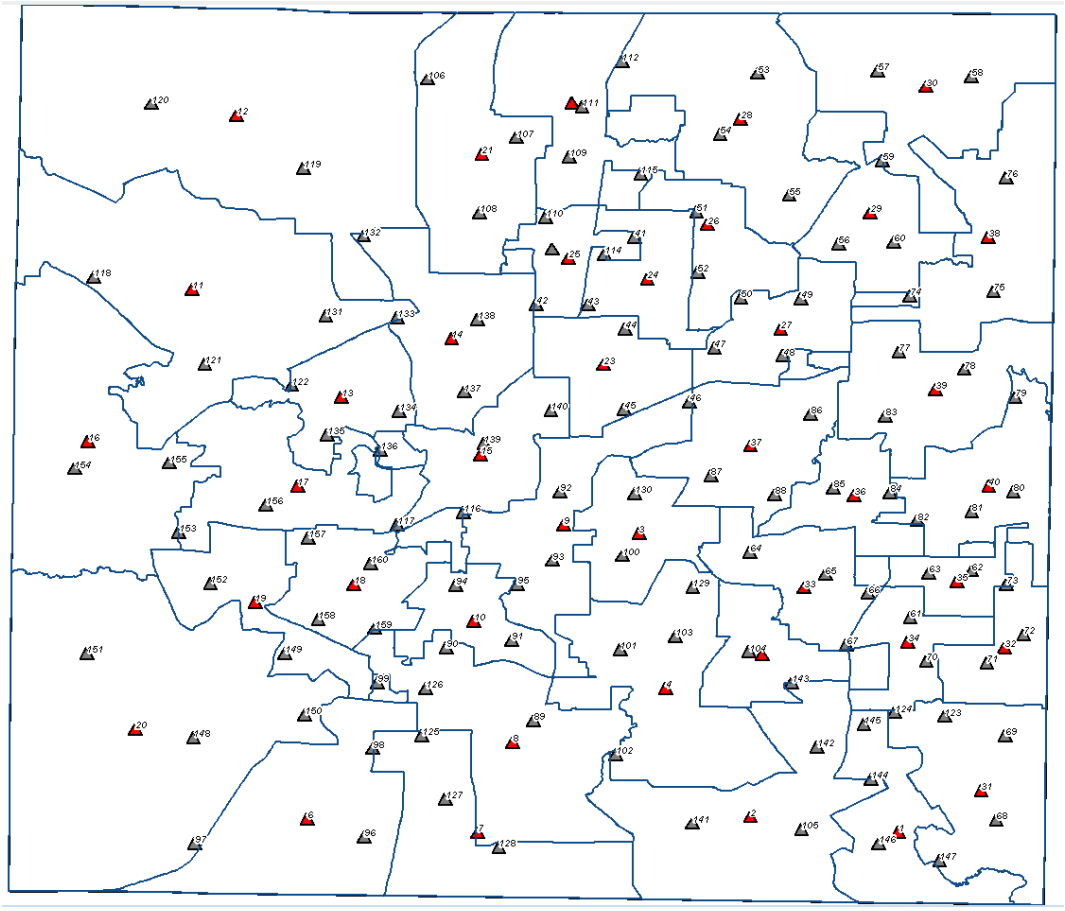


FIGURE 5.1. "Feasible" response plan from case study with 120 extra POD locations added. The red triangles represent locations of the forty original PODs, and the grey triangles represent locations of the 120 new PODs.

5.1. Maximizing Response Plan Reach while Preserving Existing Catchment Areas

Although a large set of facilities may exist in a region which are capable of hosting PODs during a biological emergency event, a subset of these facilities are generally chosen by planners to become PODs. If more than one suitable facility exists in a catchment area, tools are needed to assist planners in identifying which facility maximizes reach of vulnerable populations in this catchment area. For the purposes of the research described in this dissertation, optimality is defined in terms of reach of transportation vulnerable populations. Therefore, the optimal POD facility in a catchment area is the one which has the greatest reach of transportation vulnerable populations.

Selecting the POD location in each catchment area which yields the greatest reach of vulnerable populations will maximize the overall reach of the plan. Therefore, the maximization effort requires that the reach of vulnerable populations by each of the potential POD locations of a catchment area be computed. Methods for maximizing reach of existing response plans are detailed, and examples of these methods are presented.

5.1.1. Maximizing Response Plan Reach without Public Transit

Methodology to maximize response plan reach without public transit builds upon analysis techniques described in Chapter 4. As depicted in Figure 5.5, the coverage area of each POD location is computed, and the POD is re-assigned to the location with the coverage area that reaches the greatest vulnerable population [16]. A detailed method for performing these tasks computationally is provided in Algorithm 6. The facility with the greatest reach in each catchment area is chosen to host the POD for that catchment area. This is accomplished in the RE-PLAN Framework by setting the 'status' of chosen facilities to 'true' and unchosen ones to 'false'.

Input: An estimate of acceptable walking distance d_w, tables of available facilities, catchment areas, and transportation vulnerability population blocks. Exactly one facility in each catchment area has status = 'true'.

Output: POD locations chosen from facility locations maximize coverage of transportation vulnerable populations while maintaining catchment area (CA) boundaries. Exactly one facility in each catchment area has status = 'true'.

```
foreach catchment area i in the region do
    max_reach = 0
    max_facility = 0
    foreach facility j in catchment area i do
        if j.status == 'true' then
            original_facility = j
        end
        reach = results of query in Algorithm 1
        if reach > max_reach then
            max_reach = reach
            max_facility = j
        end
    end
    if max_reach > 0 AND original_facility != max_facility then
        original_facility.status = 'off'
        max_facility.status = 'on'
    end
end
```

Algorithm 6: Maximizing coverage of transportation vulnerable populations without public transit while maintaining catchment area boundaries

Once optimal facility locations are chosen, the coverage area buffers and coverage area blocks are calculated by using the queries in Algorithms 2 and 3, respectively. A screen capture of the RE-PLAN interface showing the coverage area buffers of the optimized response plan is provided in Figure 5.2, and one showing the coverage area blocks is provided in Figure 5.3. A map showing the vulnerable population outside of the reach of the plan is provided in Figure 5.4. This optimized plan's reach of vulnerable individuals without public transit is calculated as 10,569, representing approximately 19.8% of the county's transportation vulnerable population.

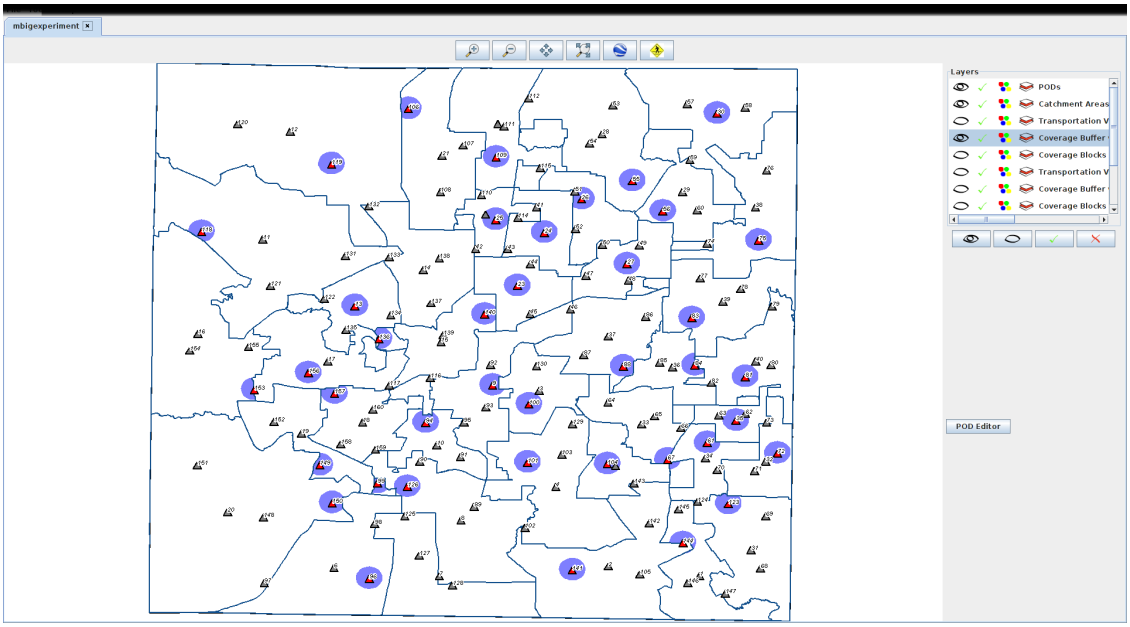


FIGURE 5.2. Screen capture of RE-PLAN with coverage area buffers for $d_w =$ 1 km around forty optimal PODs in Tarrant County, Texas

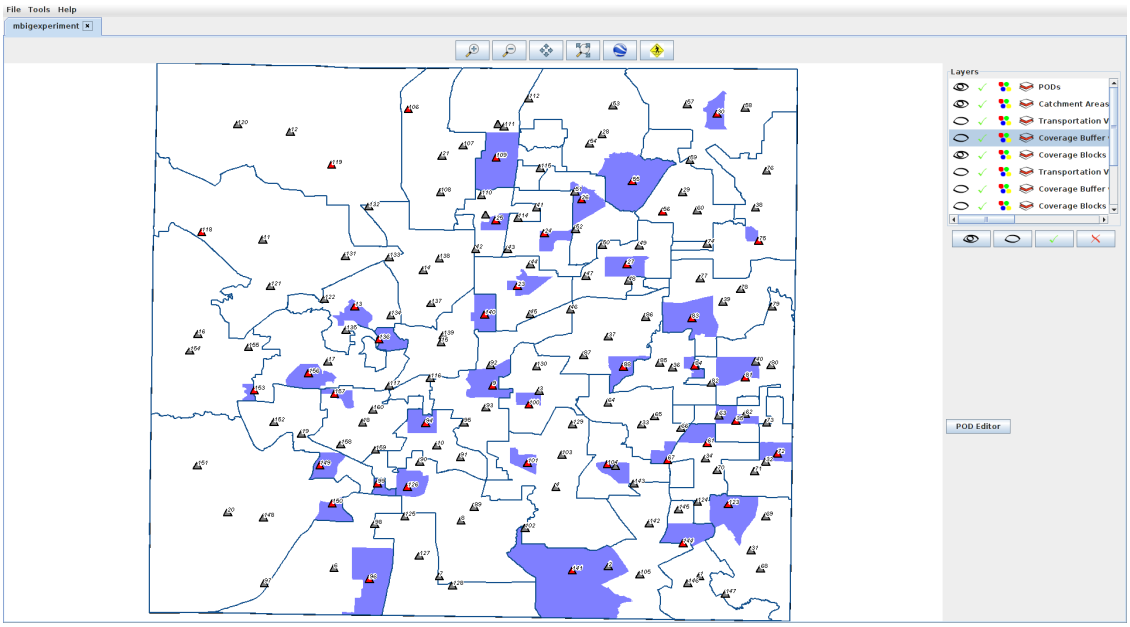


FIGURE 5.3. Screen capture of RE-PLAN with coverage area blocks for $d_w = 1$ km around forty optimal PODs in Tarrant County, Texas

5.1.2. Unconstrained Facility Location Selection to Maximize the Reach of Transportation Vulnerable Populations

If the choice of facility locations in a catchment area is not constrained by the availability of a set of specific facilities, a method to choose a facility location which maximizes

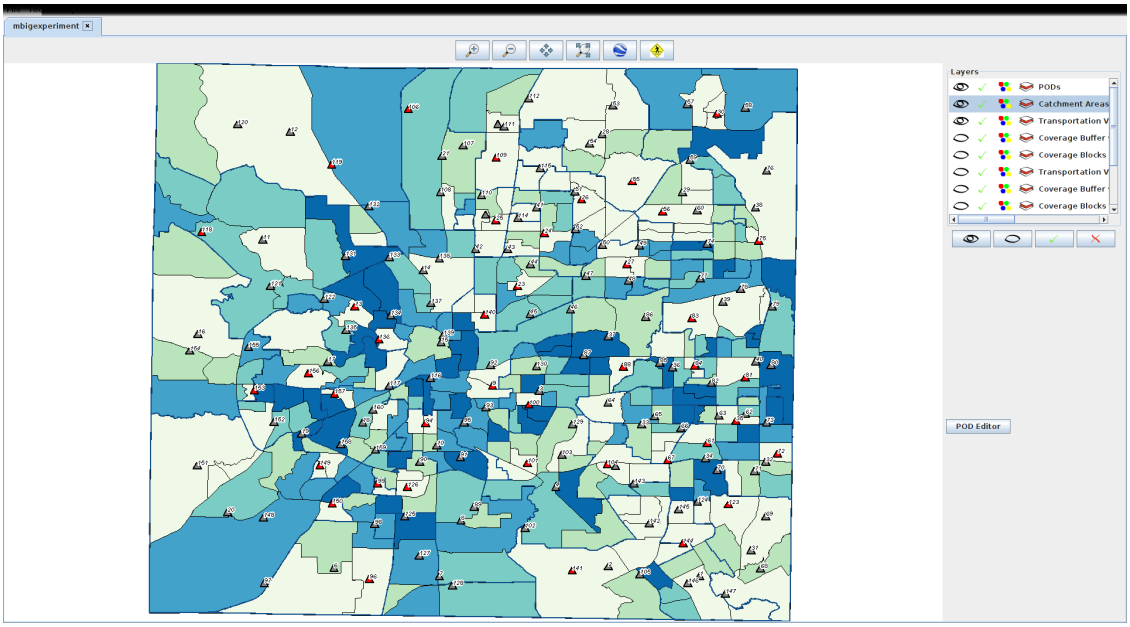


FIGURE 5.4. Transportation vulnerable population of Tarrant County, Texas at-risk of not being able to reach a POD in the scenario after optimization without public transit using the color to value mappings in Figure 4.3

.

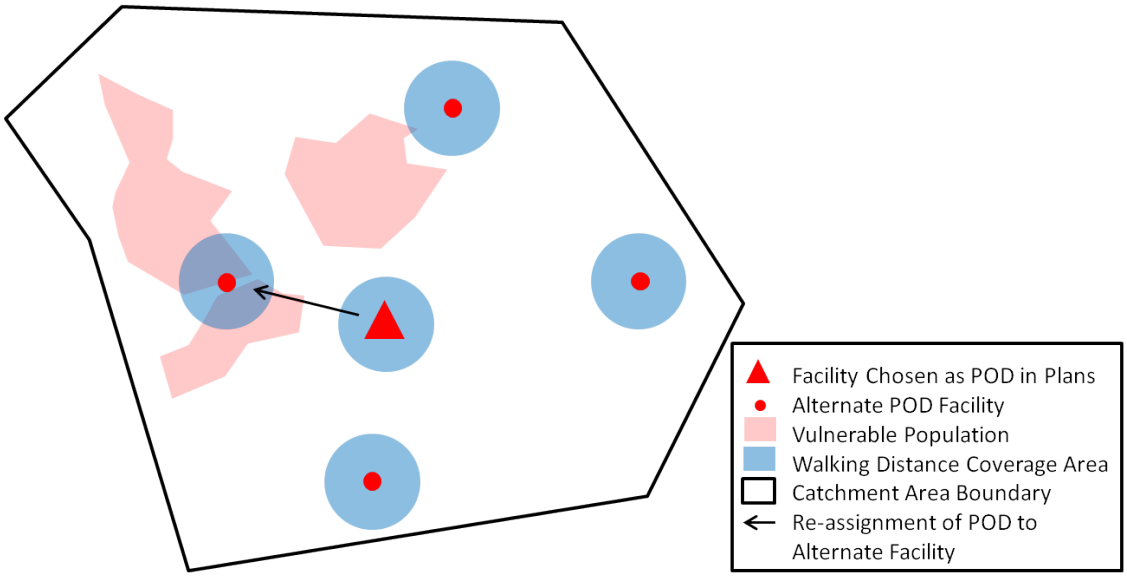


FIGURE 5.5. Method for maximizing reach of transportation vulnerable populations without the use of public transit by reassigning the POD to an alternate facility.

.

the reach of transportation vulnerable populations should be employed. An estimate of acceptable walking distance d_w can be used with vulnerable population locations and sizes to

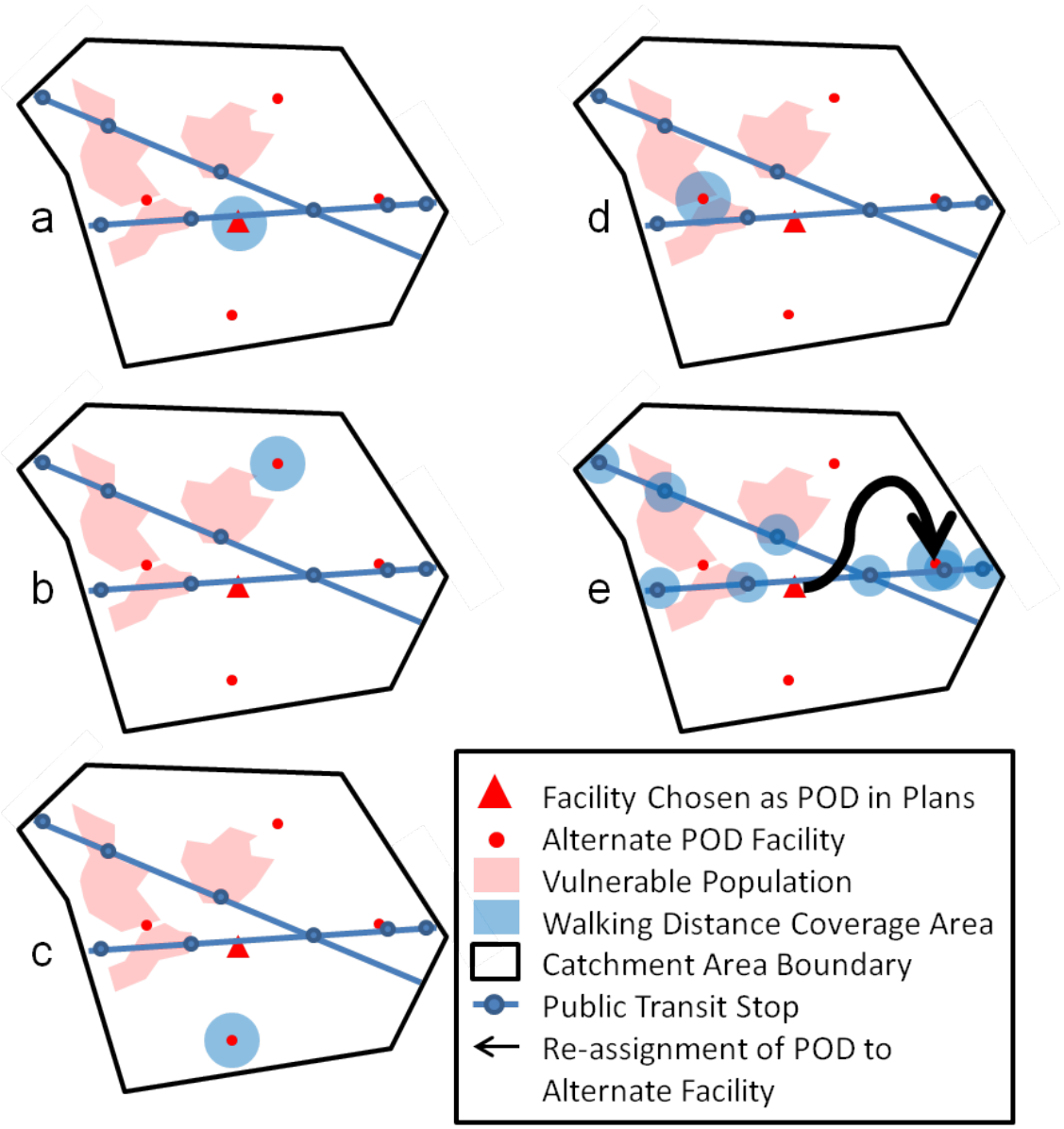


FIGURE 5.6. Five maps illustrating the method for maximizing reach of transportation vulnerable populations using public transit by reassigning the POD to an alternate facility.

calculate optimal locations for facility placement. To simplify the algorithm, populations are represented spatially as point locations. Two approaches are presented: a naive approach and a graph-based approach.

Besides simply choosing the geometric centroid of the catchment area, the naive approach is one of the most obvious approaches to unconstrained facility location selection.

It uses a center of gravity model [85] to minimize the average distance-to-POD over the entire vulnerable population of the area. However, depending on the geographic distribution of the vulnerable population, this is likely to lead to the selection of a POD location with less than maximal reach. This concept is demonstrated in Figure 5.7. A set of vulnerable populations and their sizes are shown in a single catchment area. A center of gravity model is used to select the POD location shown. However, when drawing a buffer around the POD of radius d_w, it is clear that the reach of the POD is only 20 vulnerable individuals. This POD location is clearly not maximal as there are two vulnerable populations which each exceed this reach. Simply placing the POD at either of these two locations would yield greater reach.

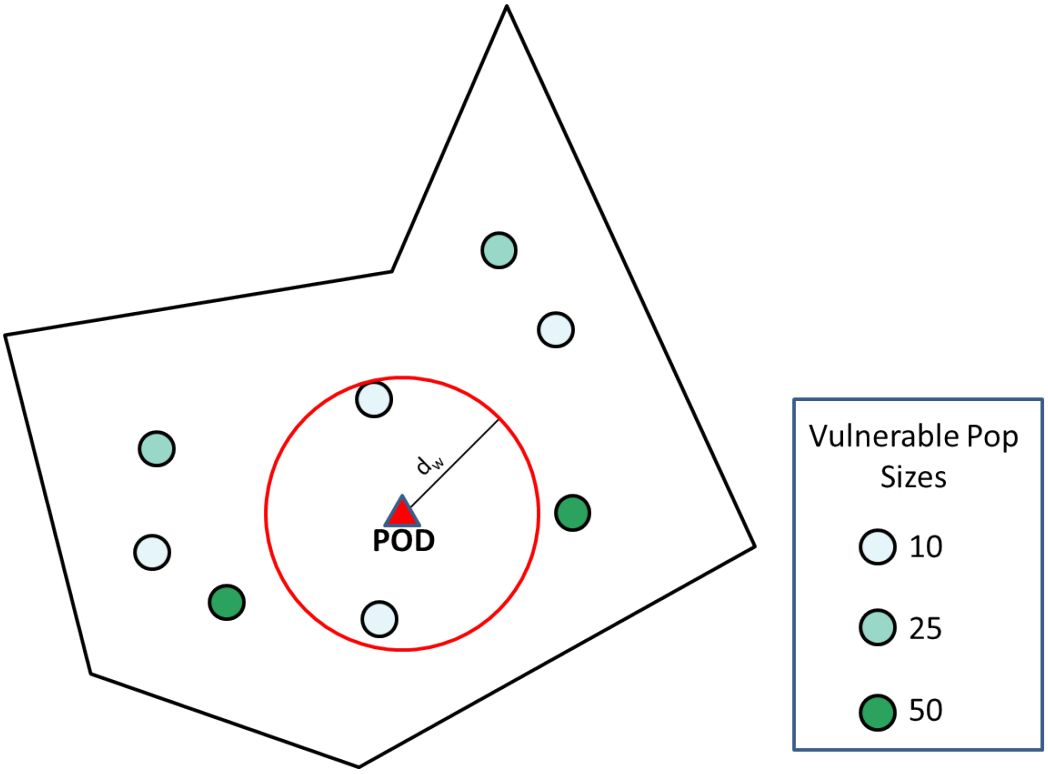


FIGURE 5.7. Example of naive approach to unconstrained facility location selection for a single catchment area with eight vulnerable populations. A buffer of radius d_w illustrates the low reach of this POD despite its central location.

The graph-based approach uses the same input data as the naive approach, but the POD location selected provides maximal coverage for the catchment area. As illustrated in

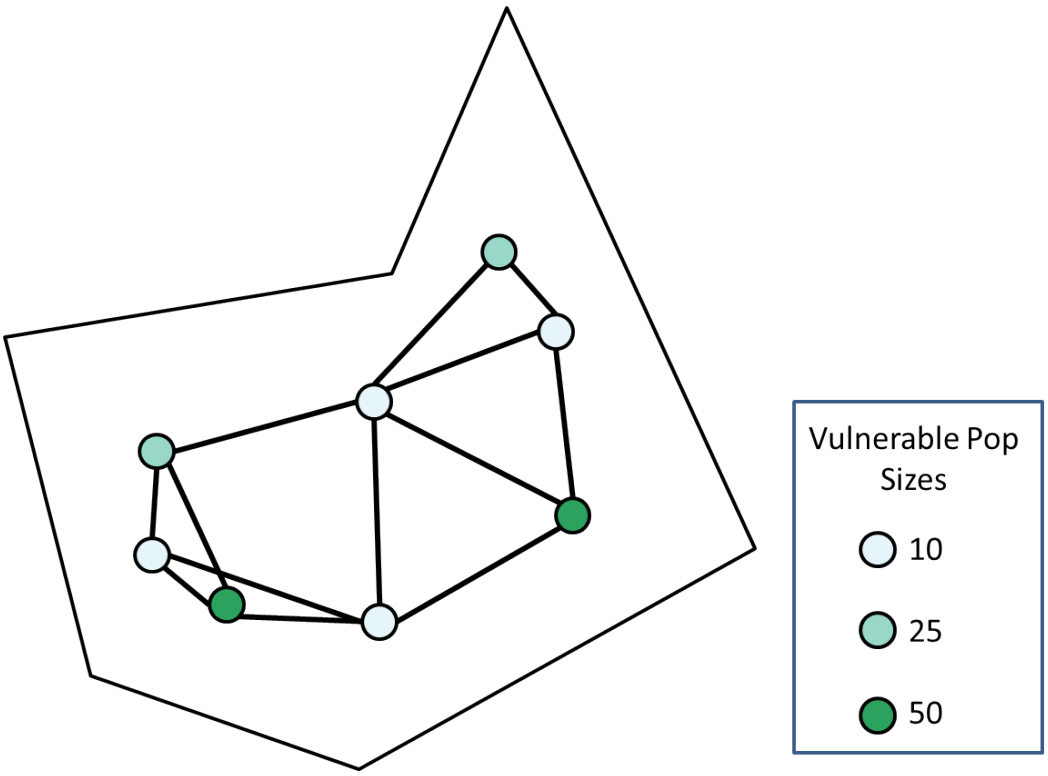


FIGURE 5.8. Example of graph approach to unconstrained facility location selection for a single catchment area with eight vulnerable populations.

Figure 5.8, a mathematical graph $G = (V, E)$ (consisting of a set of vertices V and a set of edges E) is created by representing each population as a vertex $v \in V$ and connecting each pair of populations within $2d_w$ distance of each other (shown as being the greatest distance between two populations both covered by the same facility location in Figure 5.9) by an edge $e \in E$. This graph can be used to determine a region in which to locate a facility with maximal reach.

A graph clique is a complete subgraph (i.e. a graph with an edge between each pair of vertices) [58]. A clique is defined to be a *maximal clique* if no additional vertices in the graph can be added such that the resulting subgraph is still a clique [125]. Let the set of maximum vulnerable population maximal cliques $MVPMC$ be the set of maximal cliques whose vertices represent the largest total vulnerable population (not to be confused with *maximum cliques*). The set $MVPMC$ can be used with d_w to identify specific locations which maximize coverage of vulnerable populations. This process involves the task of maximal

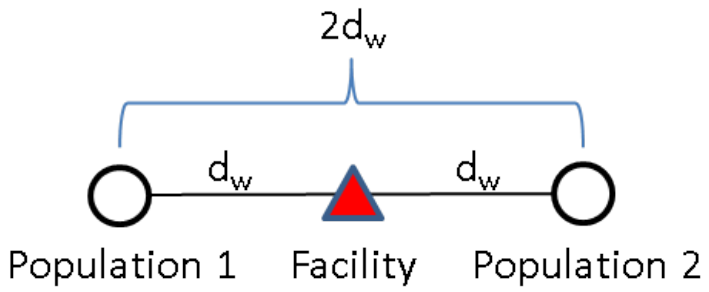


FIGURE 5.9. Illustration of how the maximum distance between two populations covered by the same facility using the same d_w walking distance estimate is $2d_w$.

clique enumeration, for which, although known to be an NP-hard problem (i.e. with respect to the size of the problem, a nondeterministic polynomial time problem which is *verifiable* in polynomial time [37]), parallel algorithms have been developed [111]. As depicted in Figure 5.10, a buffer of radius d_w is drawn around each population in $MVPMC$, and the POD must be located within the area where all of these buffers intersect as described in Theorem 5.1. Lemma 5.2 serves as a basis upon which a theoretical framework to explore how the graph representation of populations within $2d_w$ of each other can be exploited to solve the unconstrained facility location selection problem. Theorem 5.3 uses the number of maximal cliques of a graph to determine how many facility locations are required to reach the entire vulnerable population of a region.

THEOREM 5.1. *The reach of a facility placed within the intersection of d_w radius buffers around populations in a maximum vulnerable population maximal clique provides the maximum reach of vulnerable populations such that its reach cannot be dominated by the selection of a different facility location.*

PROOF. Assume there exists a location l outside of the intersection of d_w radius buffers around populations in a maximum vulnerable population maximal clique such that the sum of all vulnerable populations within its reach is greater than the reach of the facility within the intersection. This would imply that the vulnerable populations covered by a facility at location l are all within $2d_w$ of each other and form a maximal clique representing a

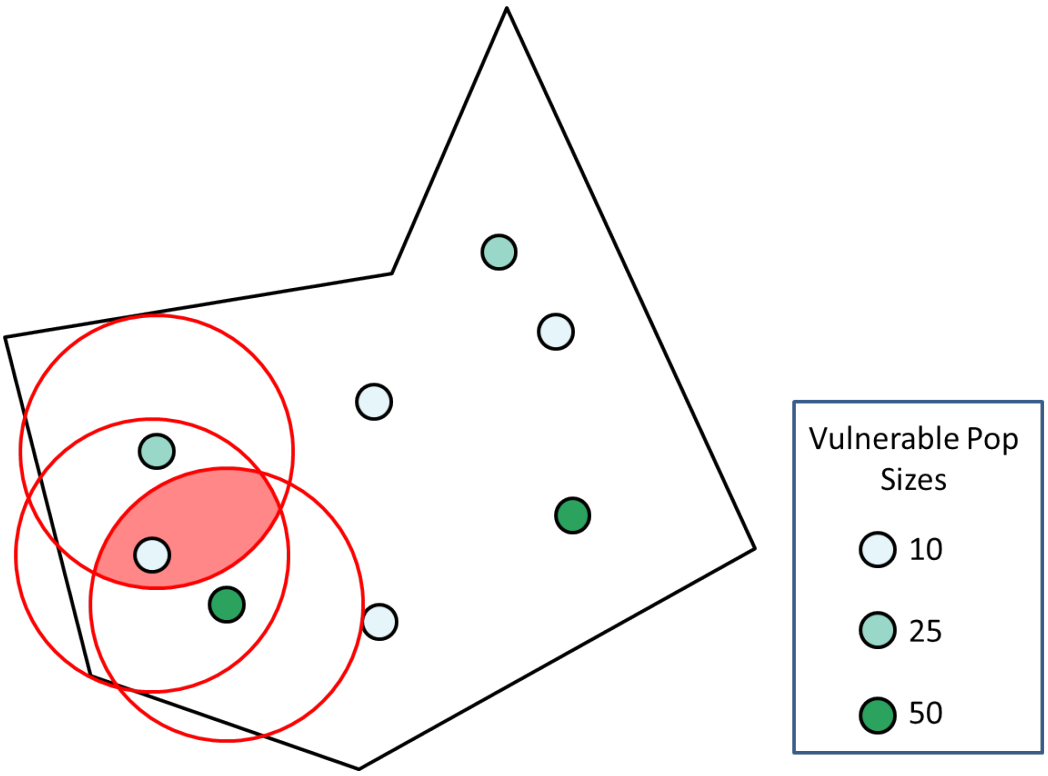


FIGURE 5.10. The intersection of the buffers drawn around the points representing the largest population clique is shaded, showing the area of maximal reach for POD facility location.

larger vulnerable population than the one initially chosen. However, this contradicts the method under which the graph was constructed and the criteria for selection of the original facility. Therefore, the original facility must provide the maximum reach of vulnerable populations. □

LEMMA 5.2. *The set of all maximal cliques of a graph represent a clique coverage of the graph.*

PROOF. Let $G = (V, E)$ where V is a set of vertices and E is a set of edges which connect pairs of vertices in V. Let MC_G represent the set of all maximal cliques mc_i of graph G. Assume $\exists v_i \in V \mid \forall mc_i \in MC_G, v_i \notin mc_i$. In other words, assume there exists a vertex $v_i \in V$ which is not included in any maximal clique in MC_G. However, the vertex v_i by itself is a clique. This clique must either be a maximal clique by itself or be a part of a larger clique which is a maximal clique. In either case, this vertex would be a member of a clique included

in MC_G, thus contradicting the initial assumption that $\forall mc_i \in MC_G$, $v_i \notin mc_i$. Therefore, the set of all maximal cliques must include all vertices in the graph, and, by definition [58], represent the clique coverage of the graph. □

THEOREM 5.3. *Let F be a minimum cardinality set of facility locations required to reach all vulnerable populations of a region. Let d_w be an estimate of maximum walking distance for transportation vulnerable individuals. Let $G = (V, E)$ be a graph such that V represents the set of populations, and E represents a set of edges between populations in V which are within $2d_w$ distance of each other. Let NV_V be the set of vertices in V which represent populations containing no vulnerable individuals and NV_E be the set of all edges in E incident upon any vertex in NV_V. Let $G_2 = (V - NV_V, E - NV_E)$. Let MC_{G_2} be the set of all maximal cliques in the graph G_2 . Then, $|F| \leq |MC_{G_2}|$.*

PROOF. Since consideration of reach is only explored for vulnerable populations, providing reach to all populations in G_2 is equivalent to providing reach to all vulnerable populations of G. The set MC_{G_2} of maximal cliques also represents the *clique cover* of the graph G. This implies that all vertices in G_2 are included in at least one clique in MC_{G_2}.

Assume there exists a population in $V - NV$ which cannot be covered through the use of $|MC_{G_2}|$ facility locations, resulting in the need for an additional facility location and causing $|F| > |MC_{G_2}|$. Since all vulnerable populations are included in at least one clique in MC_{G_2}, then one of the cliques must not be coverable by a single facility location. However, this contradicts the initial construction of the graph such that the only edges included in E (and, by extension, $E - NV_E$) are between populations within $2d_w$ of each other. Therefore, $|F| \leq |MC_{G_2}|$ holds.

□

Although $|MC_{G_2}|$ serves as an upper bound for $|F|$, it cannot be used to guarantee an exact value for $|F|$. Figure 5.11 shows an example graph representing populations within $2d_w$ distance of each other such that $|MC_{G_2}| = 3$ and $|F| = 2$. The three maximal cliques are $\{A, B\}$, $\{B, C\}$, and $\{C, D\}$. However, the entire vulnerable population can be covered

with only two facility locations: one which covers populations A and B, and another which covers populations C and D.

FIGURE 5.11. Example of graph such that $|F| < |MC_{G_2}|$.

From a GIS perspective, the same result can be achieved by creating buffers of radius d_w around each vulnerable population. These buffers represent the area each population can walk given the assumed d_w. Each buffer must include an attribute to represent the number of vulnerable individuals it represents. Using these buffers, the reach of a particular point can be calculated by summing the number of vulnerable individuals represented by each buffer which covers it. A union operation should be performed on the buffers to create a single polygon layer where each polygon represents an area of uniform reach. After calculating the reach of each polygon, the maximal reach polygon(s) can be determined by selecting the polygon(s) of maximum reach.

5.1.3. Examining Walking Distance Burden of Transportation Vulnerable Populations

A tool was developed to examine the walking distance burden of transportation vulnerable populations. The geometric centroid of each vulnerable population is used to represent the population's location. The distance from each population to its assigned POD facility is measured and plotted on a cumulative chart. The tool is used to examine the effect of plan maximization on the walking distance burden of transportation vulnerable populations.

Acceptable walking distances d_w of 1, 2, 3 and 4 kilometers were used with the reach maximization algorithm. Charts of distances of transportation vulnerable populations to their assigned PODs for the entire region are given in Figures 5.12,5.13,5.14,5.15, and 5.16, and for a selected catchment area in Figures 5.17, 5.18, and 5.19. Although the total reach of transportation vulnerable individuals is nondecreasing with increasing estimates of d_w, using

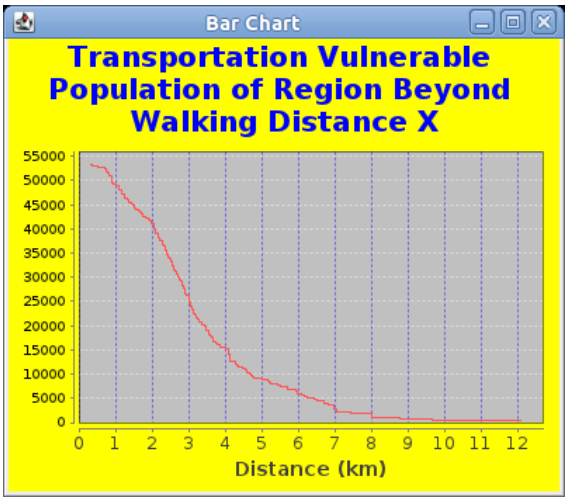


FIGURE 5.12. Region-wide vulnerable population distance-to-assigned-POD before maximization algorithms are applied

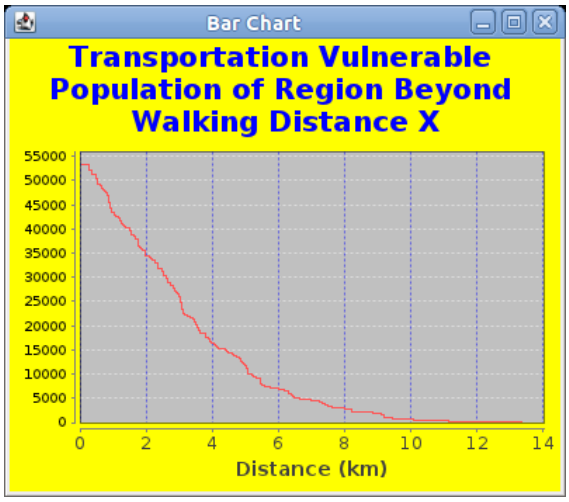


FIGURE 5.13. Region-wide vulnerable population distance-to-assigned-POD after maximization with $d_w = 1$ km

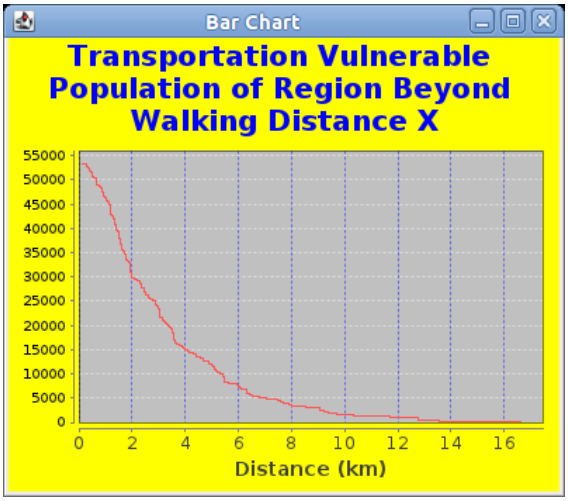


FIGURE 5.14. Region-wide vulnerable population distance-to-assigned-POD after maximization with $d_w = 2$ km

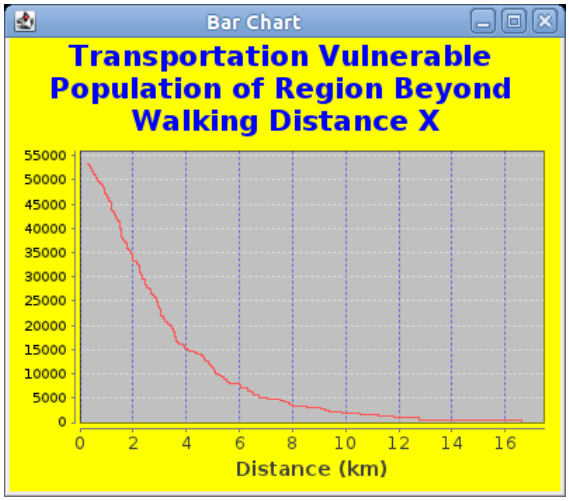


FIGURE 5.15. Region-wide vulnerable population distance-to-assigned-POD after maximization with $d_w = 3$ km

a larger estimate of d_w in the maximization procedure is likely to yield a greater average walking distance for vulnerable populations than a smaller estimate.

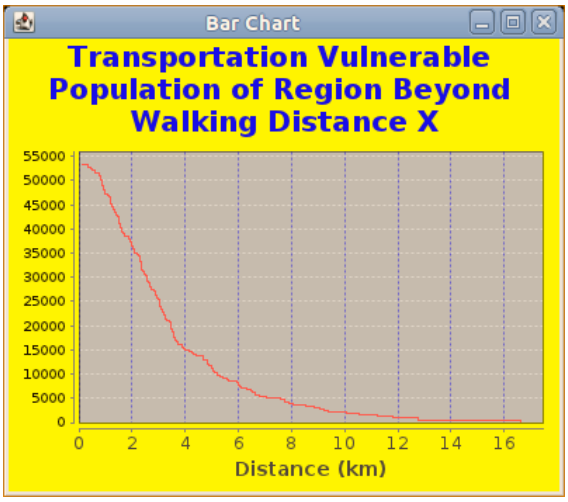


FIGURE 5.16. Region-wide vulnerable population distance-to-assigned-POD before maximization and after maximization with $d_w = 4$ km

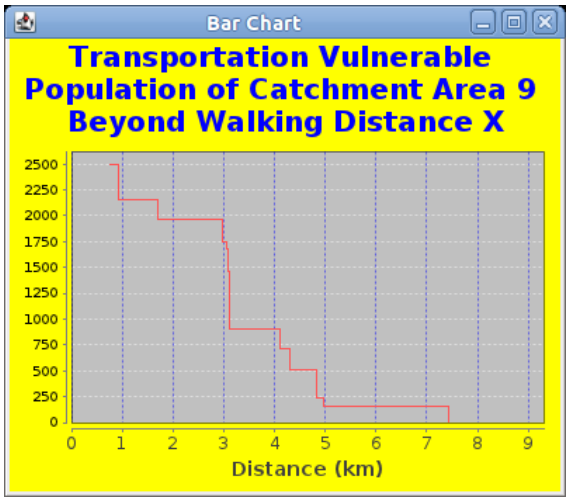


FIGURE 5.17. Catchment area 9 vulnerable population distance-to-assigned-POD before maximization and after maximization with $d_w = 1$ km

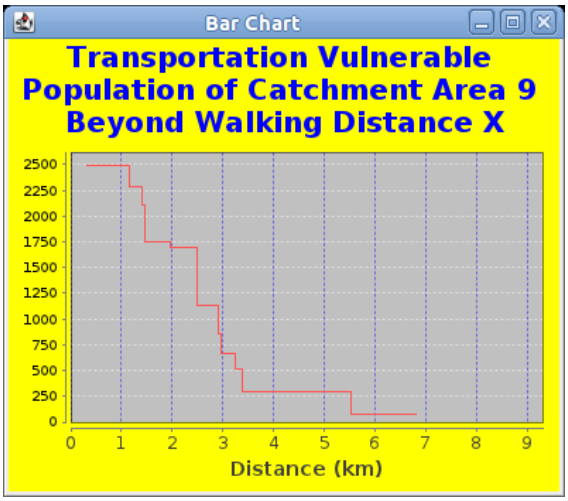


FIGURE 5.18. Catchment area 9 vulnerable population distance-to-assigned-POD after maximization with $d_w = 2$ km

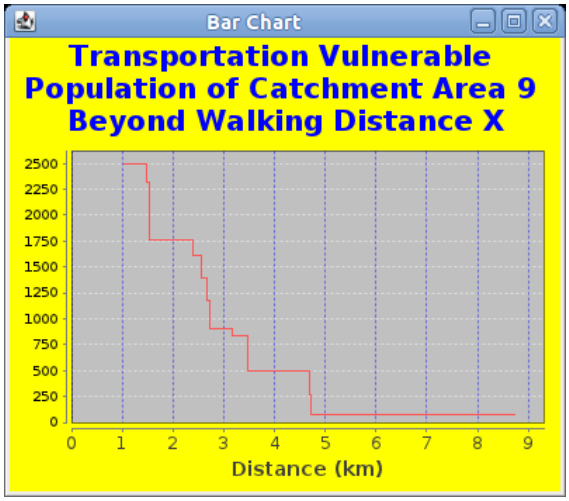


FIGURE 5.19. Catchment area 9 vulnerable population distance-to-assigned-POD after maximization with $d_w = 3$ km and $d_w = 4$ km

5.1.4. Maximizing Response Plan Reach Using Public Transit

Extending the walking distance coverage area of a POD location using public transit may increase its reach of vulnerable populations. The increased reach of each facility in

TABLE 5.1. Experimental Results for Regional Vulnerable Population Coverage Maximization

Experiment	MIN (km)	MAX (km)	AVG (km)	STDDEV (km)	Num PODs Changed
Before	0.324	12.076	3.269	1.900	0
$d_w = 1.0$ km	0.271	13.342	3.698	2.536	28
$d_w = 2.0$ km	0.116	16.604	3.903	2.859	34
$d_w = 3.0$ km	0.271	16.604	4.017	3.042	34
$d_w = 4.0$ km	0.116	16.604	4.065	3.038	35

TABLE 5.2. Experimental Results for Catchment Area 9 Vulnerable Population Coverage Maximization

Experiment	MIN (km)	MAX (km)	AVG (km)	STDDEV (km)	POD Changed
Before	0.756	7.427	3.436	1.871	false
$d_w = 1.0$ km	0.756	7.427	3.436	1.871	false
$d_w = 2.0$ km	0.310	6.803	2.801	1.846	true
$d_w = 3.0$ km	1.043	8.734	3.267	2.071	true
$d_w = 4.0$ km	1.043	8.734	3.267	2.071	true

a catchment area must be calculated, and the POD must be assigned to the facility with the greatest reach. Figure 5.6 shows an example of the methodology to choose the optimal facility in a catchment area to host a POD. The coverage areas of facility locations explored in maps **a**, **b**, and **c** do not intersect vulnerable populations. The coverage area of the facility location in map **d** provides some coverage (and therefore, reach) of nearby vulnerable populations. However, the close proximity of the facility location in map **e** to a public transit stop allows the coverage area of the facility location to be extended to intersect with vulnerable populations near three different transit stops, thus greatly increasing the facility's reach of vulnerable populations. For simplicity, in this example the vulnerable population is assumed to be uniformly distributed across the areas of vulnerability. The facility location closest to the transit stop has the maximal coverage area (and in this case reach) and is selected to host the POD. Nonetheless, the methodology presented considers non-uniform population densities of vulnerable populations by focusing on plan reach calculated using the spatial distribution of vulnerability and the estimated walking distance coverage area described in Chapter 4.

Algorithm 7 details the methodology employed to maximize a plan's reach of vulnerable populations using public transit. It builds upon methodology from Chapter 4 to analyze reach of vulnerable populations in the context of public transit. Facilities chosen to host PODs in plans following execution of Algorithm 7 maximize reach of vulnerable populations using available public transit infrastructure while maintaining existing catchment areas.

Input: An estimate of acceptable walking distance d_w, tables of available facilities, transit stops, catchment areas, and transportation vulnerability population blocks. Exactly one facility in each catchment area has status = 'true'.

Output: POD locations chosen from facility locations maximize coverage of transportation vulnerable populations using public transit while maintaining catchment area (CA) boundaries. Exactly one facility in each catchment area has status = 'true'.

```
foreach catchment area i in the region do
    max_reach = 0
    max_facility = 0
    foreach facility j in catchment area i do
        if j.status == 'true' then
            original_facility = j
        end
        x = results of query in Algorithm 4
        dw_minus_x = d_w − x
        reach = results of query in Algorithm 5 using dw_minus_x and d_w
        if reach > max_reach then
            max_reach = reach
            max_facility = j
        end
    end
    if max_reach > 0 AND original_facility != max_facility then
        original_facility.status = 'off'
        max_facility.status = 'on'
    end
end
```

Algorithm 7: Maximizing coverage of transportation vulnerable populations with public transit while maintaining catchment area boundaries

Figure 5.21 shows the coverage area buffers around the forty PODs chosen to maximize the response plan in Tarrant County, Texas with $d_w = 1$ km. The coverage area blocks are shown in Figure 5.21, illustrating where vulnerable populations are within reach of the optimized response plan. The transportation vulnerable population at-risk of not being able to reach a POD by either walking or using public transit is depicted in Figure 5.22. Following

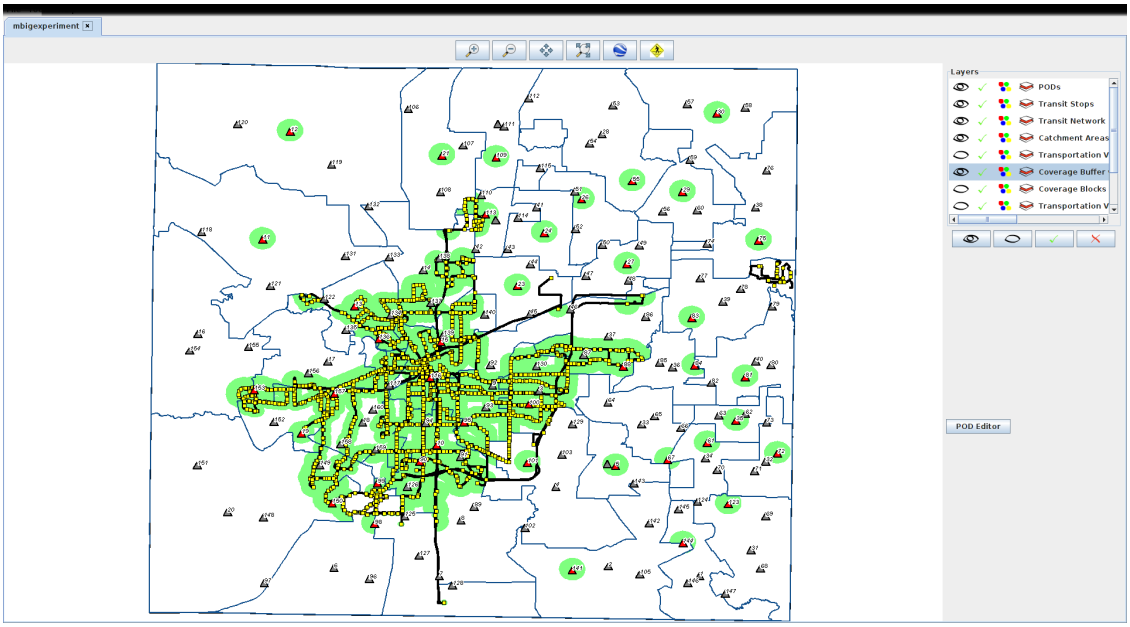


FIGURE 5.20. Screen capture of RE-PLAN with coverage area buffers for $d_w = 1$ km using public transit around forty optimal PODs in Tarrant County, Texas. Public transit stops and routes are also displayed on the map as a reference.

maximization using public transit with $d_w = 1$ km, the plan has a reach of vulnerable populations of 27,493, representing approximately 51.5% of the transportation vulnerable population of the county.

5.1.5. Maximizing Response Plan Reach by Adding Public Transit Stops Near PODs

Benefits of the algorithm to maximize plan reach using public transit rely not only on the existence of public transit stops in a region, but also on transit stops being located close to POD locations. This can lead to situations where public transit infrastructure is too far from PODs in a catchment area to be used to extend the walking distance coverage area. To overcome this problem, a new stop can be added to the public transit network. For example, Figure 5.24 shows a view of Netwon Rayzor Elementary School in Denton, TX using Google Earth. The closest large road to this school is Highway 380 (shown near the top of the map). A bus stop added to Highway 380 near the school would be about 150 m from the school.

Let y be the distance from the new stop to a POD location. If $y > x$, the only

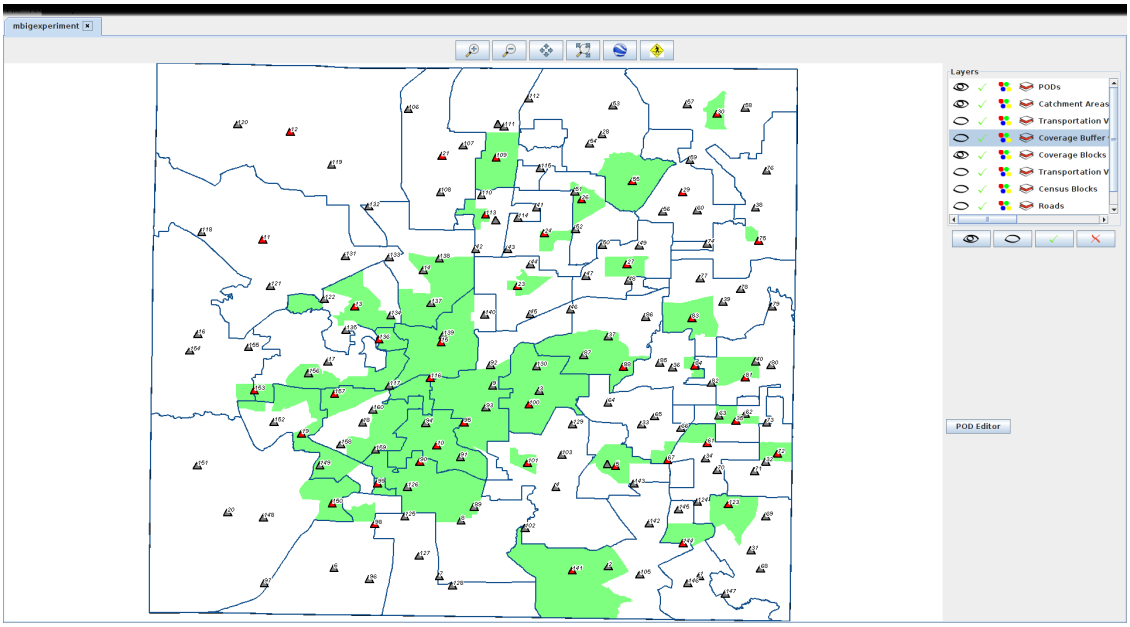


FIGURE 5.21. Screen capture of RE-PLAN with coverage area blocks for $d_w =$ 1 km around forty optimal PODs in Tarrant County, Texas

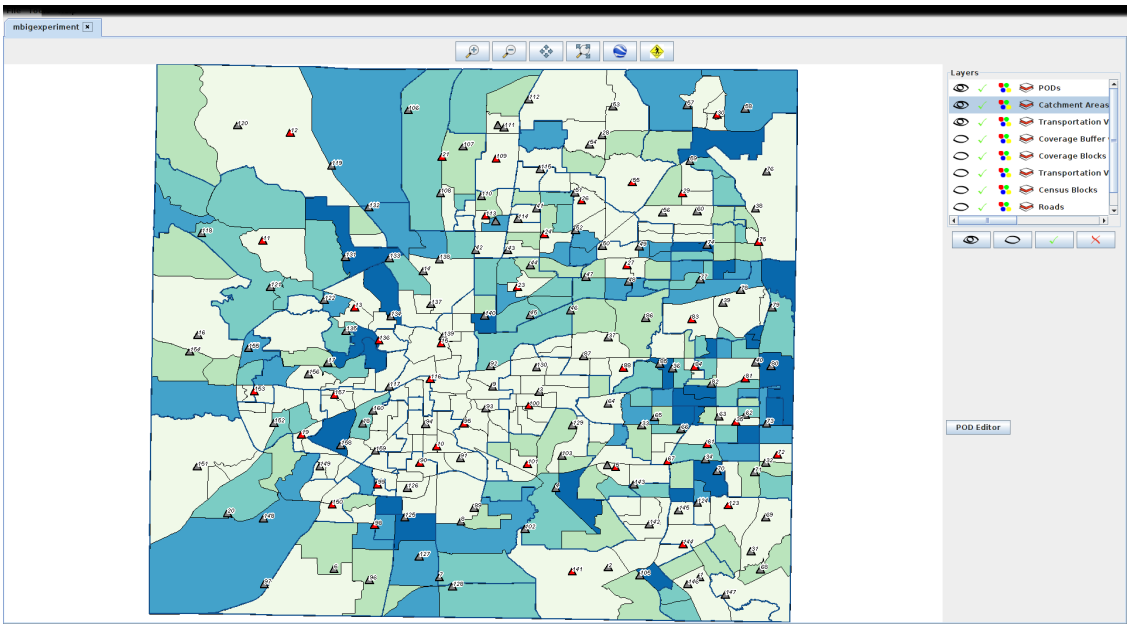


FIGURE 5.22. Transportation vulnerable population of Tarrant County, Texas at-risk of not being able to reach a POD in the scenario after optimization with public transit using the color to value mappings in Figure 4.3

.

increase in reach will be of populations within $d_w - x$ of the new transit stop. However, as depicted in Figure 5.23, if $y < x$, the radius of coverage around each transit stop in the set of transit stops in this catchment area S will be increased from $d_w - x$ to $d_w - y$: a difference

of $x - y$. This results in an increase of walking distance coverage area C_{p_i} for each POD p_i of $\Delta C_{p_i} \leq \pi(|S| - 1)(y^2 - x^2)$.

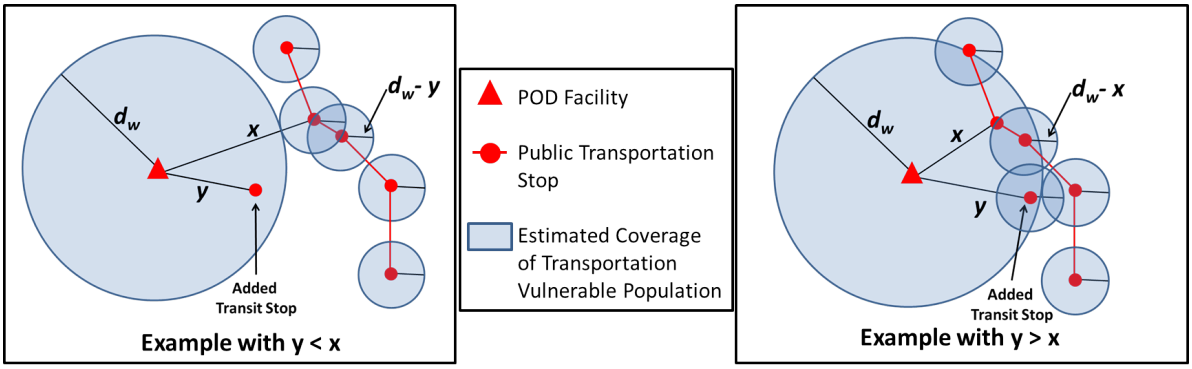


FIGURE 5.23. Effects of adding a transit stop within a distance y of a POD facility.

.

FIGURE 5.24. Example of adding a public transit stop 150 m from a facility which may be used as a POD.

.

Algorithm 8 details the procedure used to explore plan maximization resulting from increases in reach from the addition of transit stops near each POD facility. It chooses the

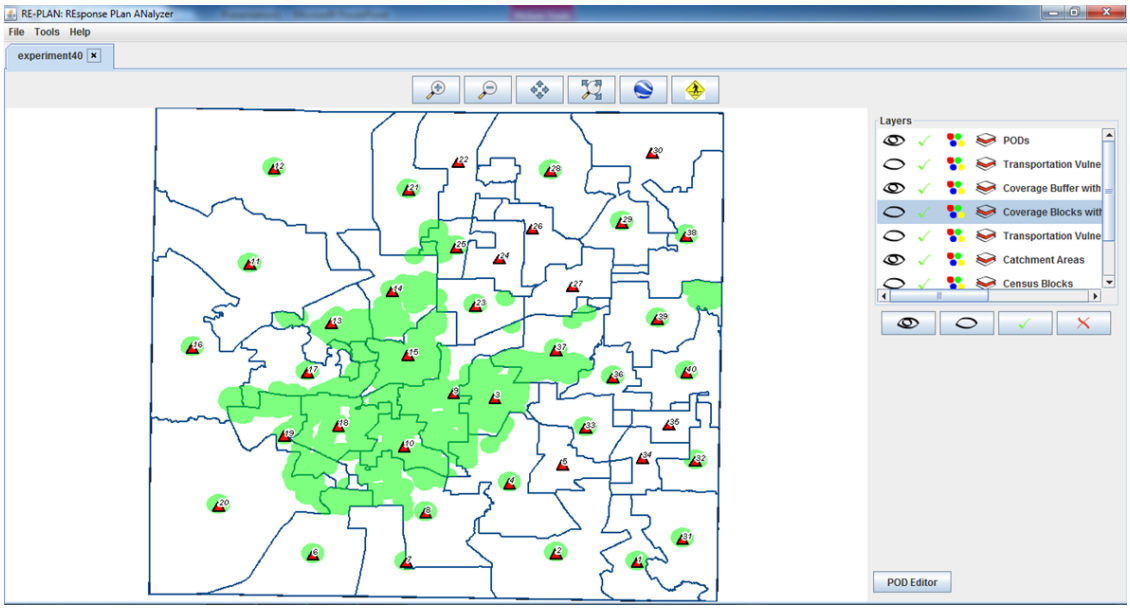


FIGURE 5.25. Screen capture of RE-PLAN with coverage area buffers for $d_w = 1$ km after adding a transit stop within 150 m of each of the original 40 PODs in Tarrant County, Texas.

.

facility location in each catchment area which maximizes plan reach using the public transit network and added transit stops such that $y < x$. Since calculations of reach are only affected by the addition of the transit stop if $y < x$, the precise location of the transit stop does not need to be determined. Rather, any location on a circle of radius y from the facility location can be chosen. If for all PODs in a CA, $y \geq x$, Algorithm 8 would choose the same POD locations and calculate the same reach as Algorithm 7.

The increased coverage area buffers resulting from adding a transit stop within 150 m of each of the original 40 PODs such that $y = 150$ is shown in Figure 5.25, and the corresponding coverage area blocks are shown in Figure 5.26. A map showing the remaining vulnerable population still at-risk of not being able to participate in mitigation activities is shown in Figure 5.27. The total reach of vulnerable populations resulting from adding transit stops 150 m from each POD is 27,328.

If the maximization algorithm is used to select the POD with greatest reach for each catchment area, a map showing the resulting coverage area buffers is provided in Figure 5.28, and a map showing the corresponding coverage area blocks is provided in Figure 5.29. A

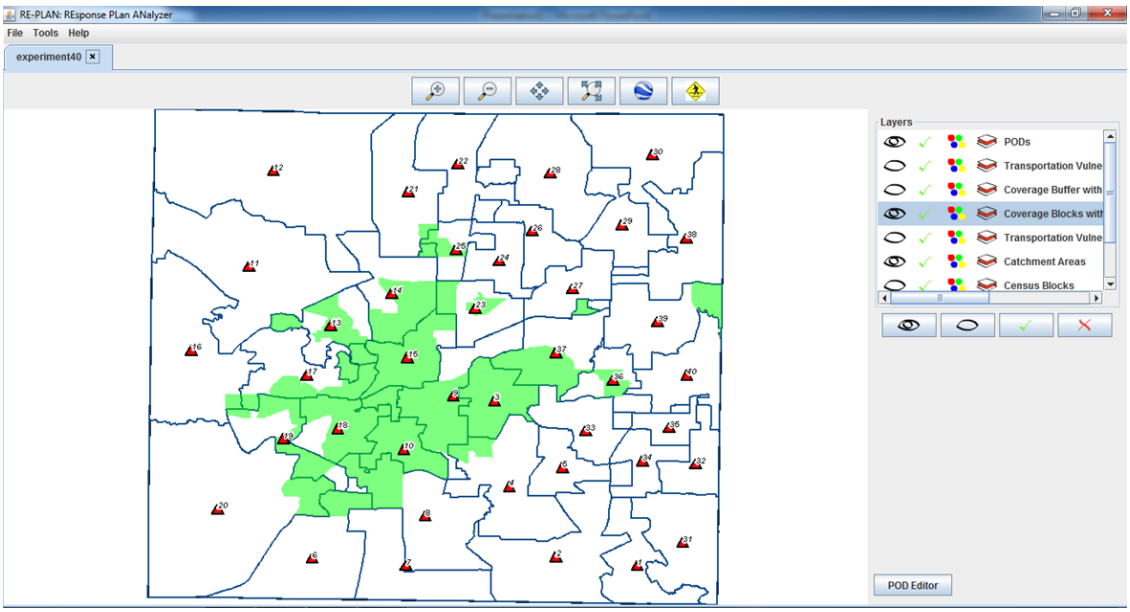


FIGURE 5.26. Screen capture of RE-PLAN with coverage area blocks for $d_w =$ 1 km after adding a transit stop within 150 m of each of the original 40 PODs in Tarrant County, Texas.

.

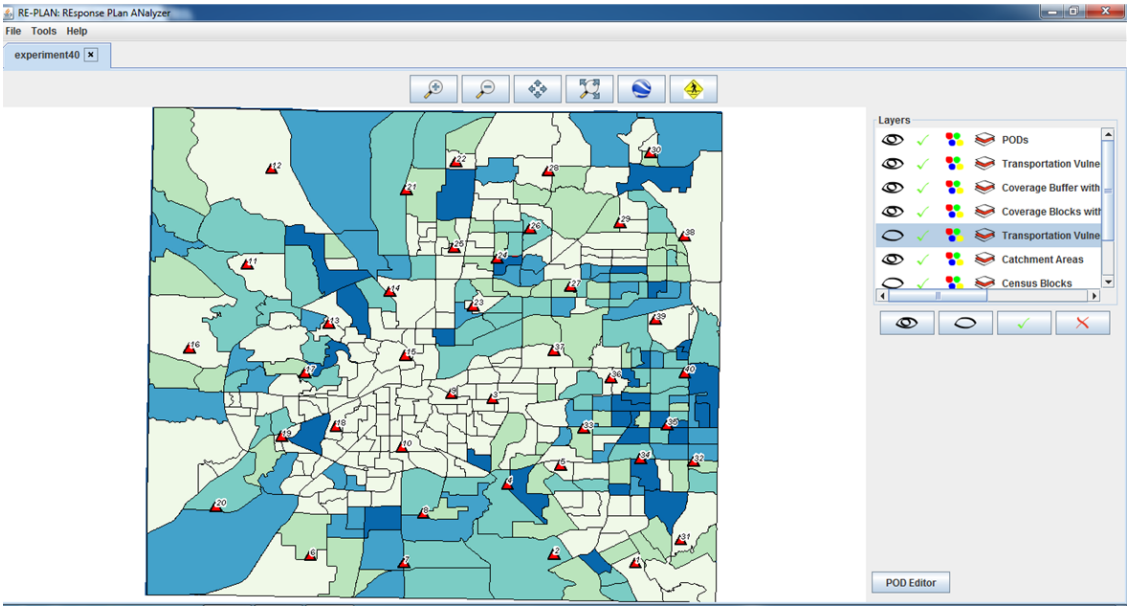


FIGURE 5.27. Screen capture of RE-PLAN showing vulnerable population still at risk of not receiving mitigation resources after adding a transit stop within 150 m of each of the original 40 PODs in Tarrant County, Texas using $d_w = 1$ km.

.

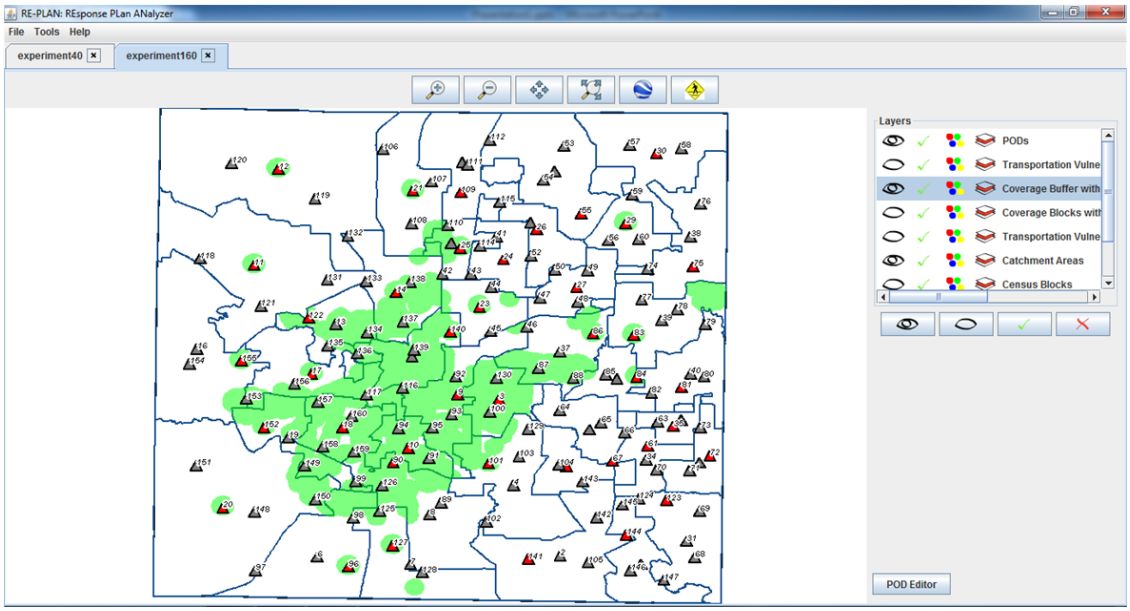


FIGURE 5.28. Screen capture of RE-PLAN showing the best 40 of 160 PODs with coverage area buffers for $d_w = 1$ km after adding a transit stop within 150 m of each POD in Tarrant County, Texas.

.

map showing the remaining vulnerable population still at-risk of not receiving mitigation resources is shown in Figure 5.29. The total reach of vulnerable populations resulting from choosing the best 40 POD locations considering an additional transit stop added at 150 m from each location is 30,597.

If a new transit stop is added at each POD such that $y = 0$, Figure 5.31 shows a map of the coverage area buffers of the original 40 PODs, and Figure 5.32 shows the corresponding coverage area blocks. Figure 5.33 shows the remaining vulnerable population still at-risk of not receiving mitigation resources. The total reach of vulnerable populations using the original 40 POD locations with transit stops added to each POD is 28,848.

Choosing the POD location with the greatest reach in each catchment area assuming a transit stop is added at each POD location results in the coverage area buffers map provided in Figure 5.34. The corresponding coverage area blocks map is provided in Figure 5.35, and a map showing the remaining vulnerable populations still at risk of not being able to participate in mitigation activities is provided in Figure 5.36. This method results in a total reach of 31,771 vulnerable individuals, representing a reach of nearly 59.5% of the

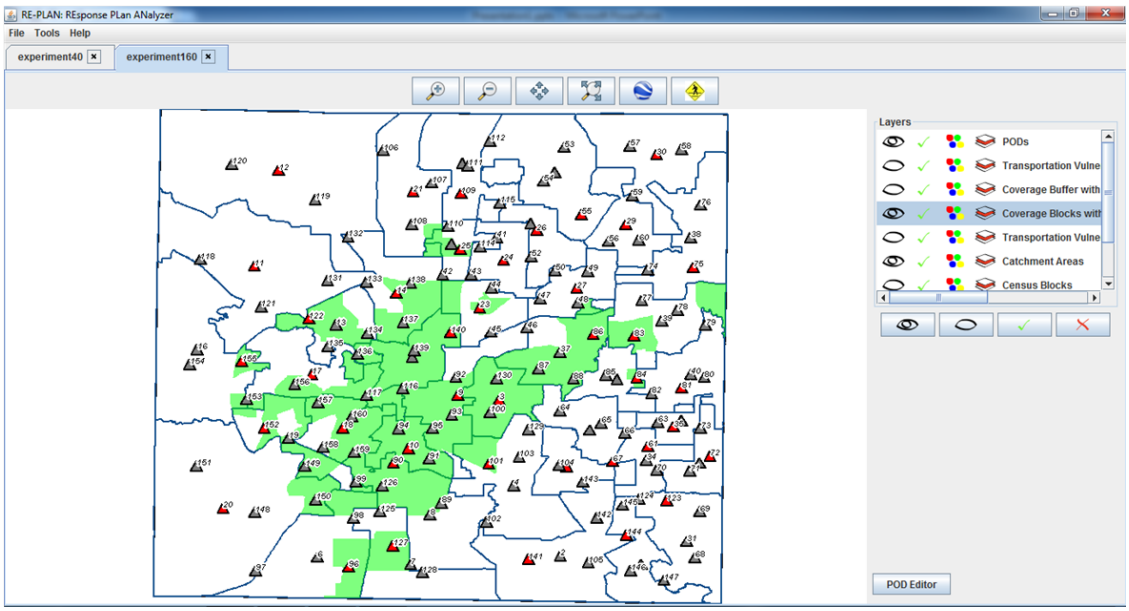


FIGURE 5.29. Screen capture of RE-PLAN showing the best 40 of 160 PODs with coverage area blocks for $d_w = 1$ km after adding a transit stop within 150 m of each POD in Tarrant County, Texas.

.

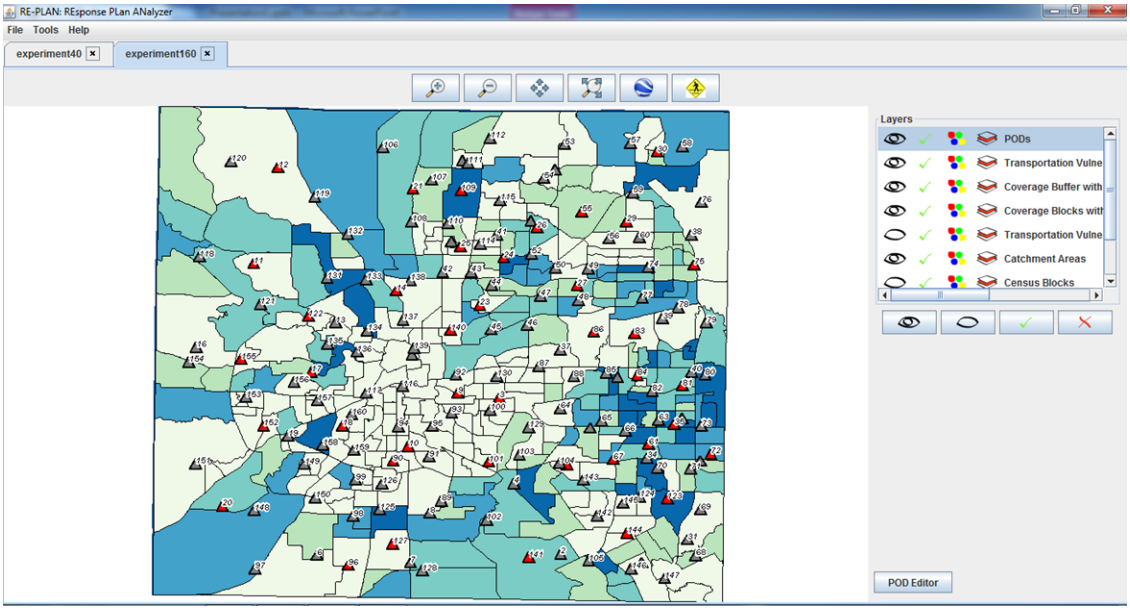


FIGURE 5.30. Screen capture of RE-PLAN showing the vulnerable population still at risk of not receiving mitigation resources after choosing the best 40 of 160 PODs and adding a transit stop within 150 m of each POD in Tarrant County, Texas using $d_w = 1$ km.

.

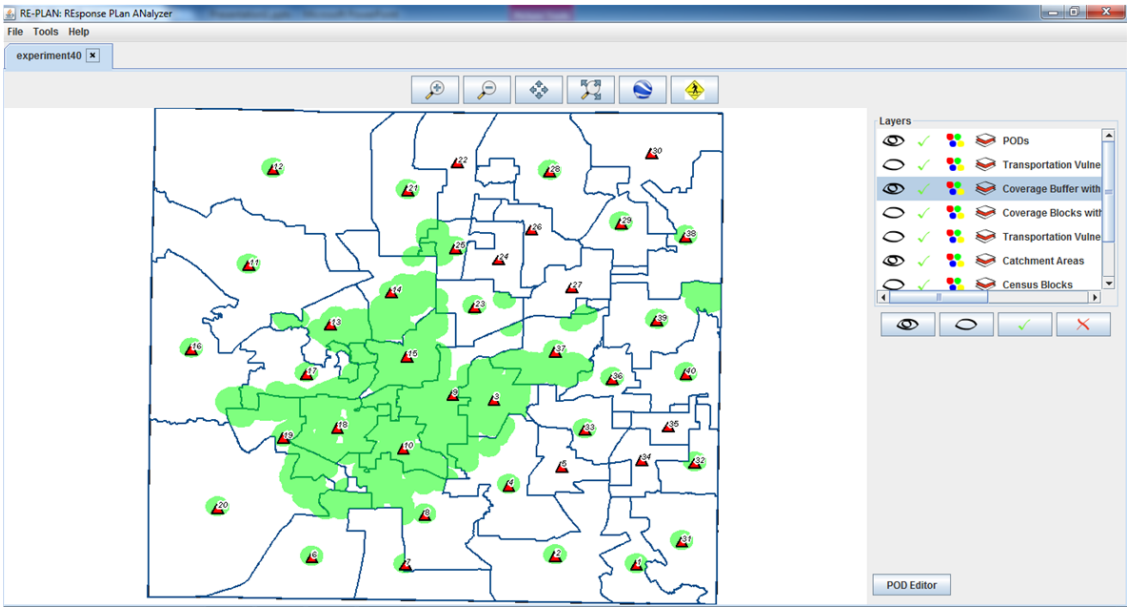


FIGURE 5.31. Screen capture of RE-PLAN with coverage area buffers for $d_w = 1$ km after adding a transit stop at each of the original 40 PODs in Tarrant County, Texas.

.

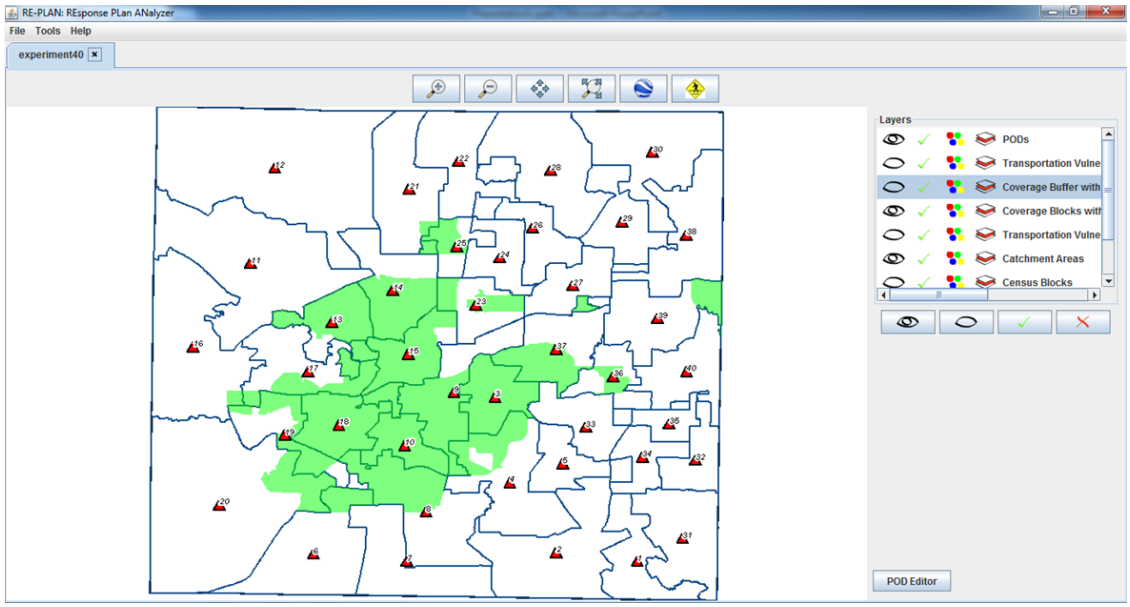


FIGURE 5.32. Screen capture of RE-PLAN with coverage area blocks for $d_w = 1$ km after adding a transit at each of the original 40 PODs in Tarrant County, Texas.

.

transportation vulnerable population.

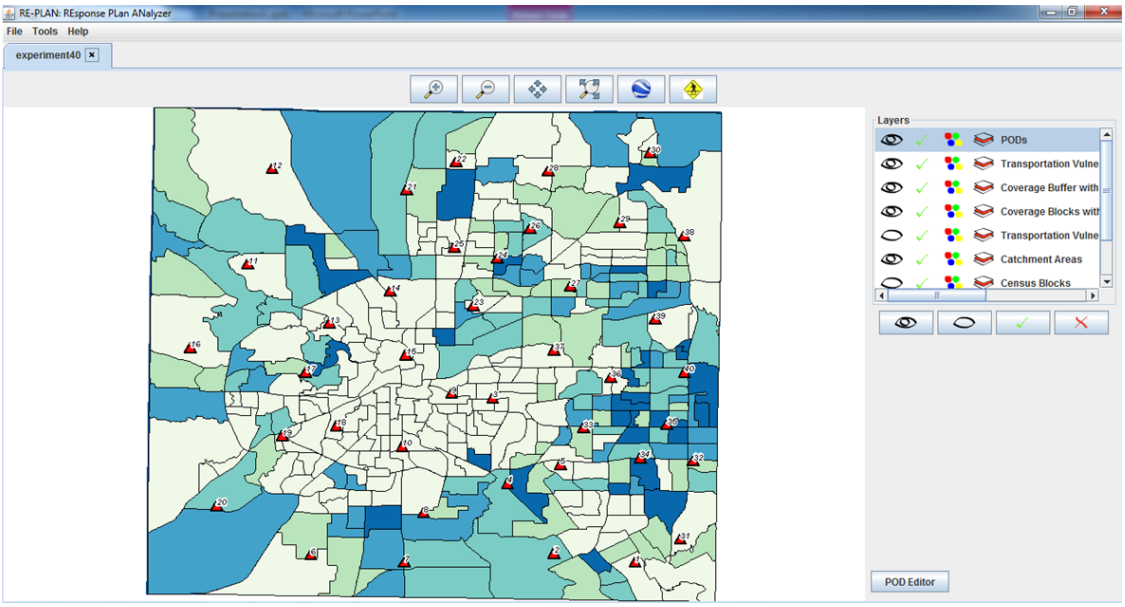


FIGURE 5.33. Screen capture of RE-PLAN showing vulnerable population still at risk of not receiving mitigation resources after adding a transit stop at each of the original 40 PODs in Tarrant County, Texas using $d_w = 1$ km.

.

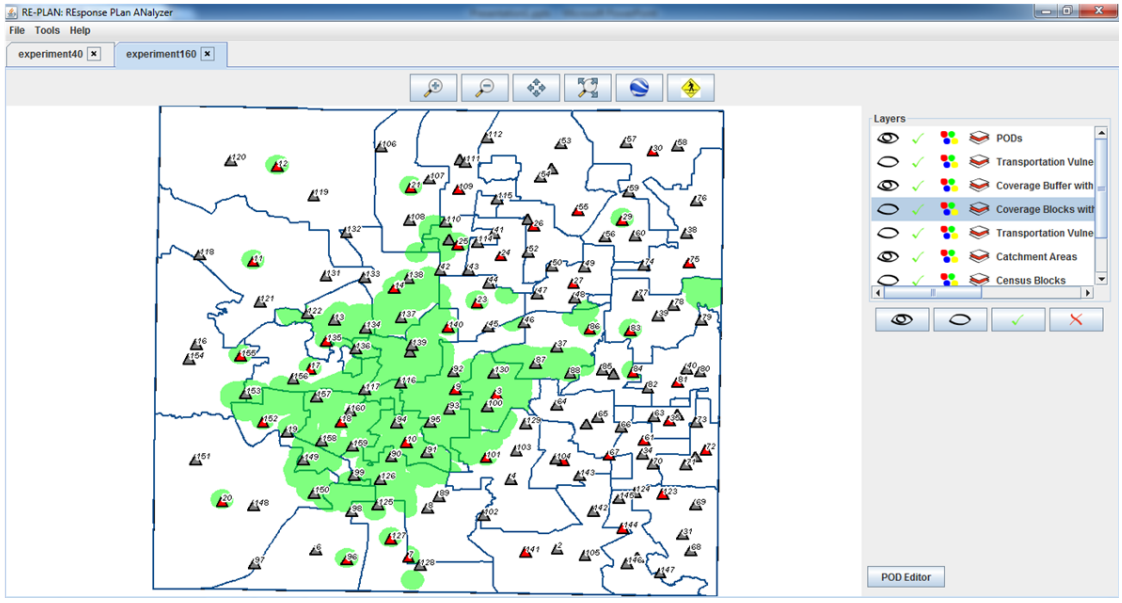


FIGURE 5.34. Screen capture of RE-PLAN showing the best 40 of 160 PODs with coverage area buffers for $d_w = 1$ km after adding a transit stop at each POD in Tarrant County, Texas.

.

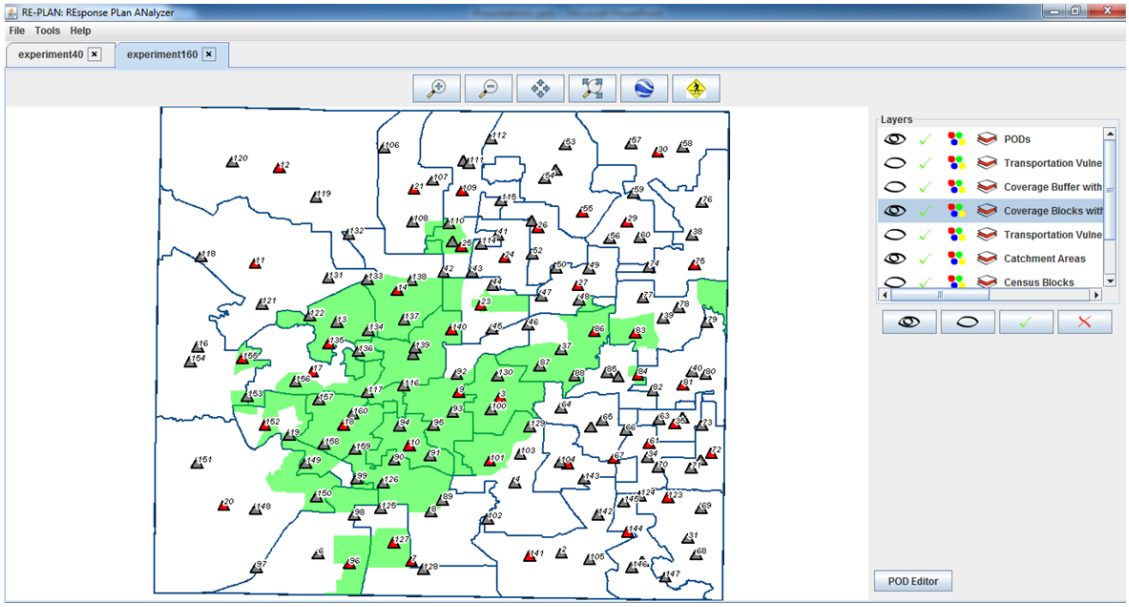


FIGURE 5.35. Screen capture of RE-PLAN showing the best 40 of 160 PODs with coverage area blocks for $d_w = 1$ km after adding a transit stop at each POD in Tarrant County, Texas.

.

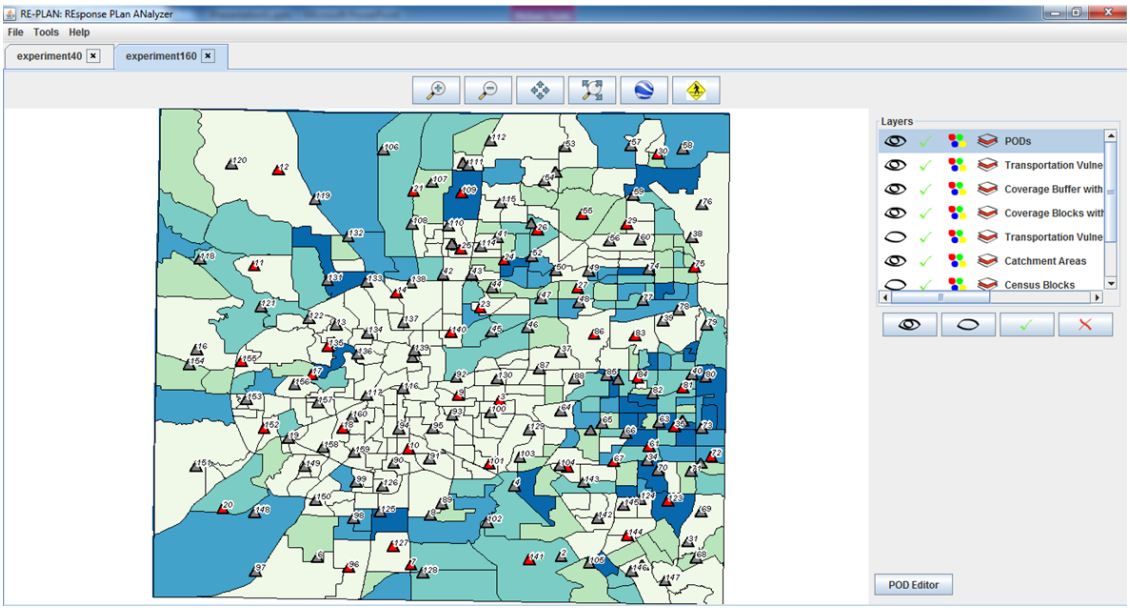


FIGURE 5.36. Screen capture of RE-PLAN showing the vulnerable population still at risk of not receiving mitigation resources after choosing the best 40 of 160 PODs and adding a transit stop at each POD in Tarrant County, Texas using $d_w = 1$ km.

.

Input: An estimate of acceptable walking distance d_w, a distance y to specify how far the added transit stops will be from each POD, tables of available facilities, transit stops, catchment areas, and transportation vulnerability population blocks. Exactly one facility in each catchment area has status = 'true'.

Output: POD locations chosen from facility locations maximize coverage of transportation vulnerable populations using public transit (including the added transit stop) while maintaining catchment area (CA) boundaries. Exactly one facility in each catchment area has status = 'true'.

```
foreach catchment area i in the region do
    max_reach = 0
    max_facility = 0
    foreach facility j in catchment area i do
        if j.status == 'true' then
            original_facility = j
        end
        x = results of query in Algorithm 4
        if y < x then
            dw_minus_x = d_w − y
        end
        else
            dw_minus_x = d_w − x
        end
        reach = results of query in Algorithm 5 using facility j, dw_minus_x, and d_w
        if reach > max_reach then
            max_reach = reach
            max_facility = j
        end
    end
    if max_reach > 0 AND original_facility != max_facility then
        original_facility.status = 'off'
        max_facility.status = 'on'
    end
end
```

Algorithm 8: Maximizing coverage of transportation vulnerable populations by adding a transit stop within a distance of y of each POD while maintaining catchment area boundaries and choosing the POD which maximizes reach of vulnerable populations within each catchment area.

5.1.6. Re-assigning Populations to Different Catchment Areas to Maximize Response Plan Reach

Re-assigning populations to different catchment areas may increase the reach of vulnerable populations estimated using methodology detailed in Chapter 4. A situation in which several vulnerable populations are beyond walking distance of their assigned PODs is depicted in Figure 5.38. However, POD 3's close proximity to a transit stop facilitates the use of public transit to extend its walking distance coverage area beyond its normal catchment area. Transportation vulnerable populations covered by POD 3, but not covered by their assigned PODs, can be reassigned to POD 3. This action changes the number of individuals assigned to each POD. Therefore, proportional reallocation of resources based on this reassignment must be conducted and/or catchment areas must be recalculated.

Populations outside of the reach of their assigned PODs in the case study were examined to determine if they could be reassigned to a different POD. Figure 5.37 shows three reassignment options for the population of a specific, out-of-reach census tract. Additional POD locations which can serve as candidates for reassignment of this population are listed in Table 5.3.

Population reassignment options are dependent upon the distance from each POD to its closest transit stop and the distance from each out-of-reach population to its closest transit stop. If, for a particular population and a particular POD, the sum of these two distances is less than d_w, then the population may be reassigned to this POD. Specifically, let pop_i represent a specific out-of-reach population. Let $stop_{pop_i}$ be the transit stop of minimum distance from pop_i. Let P be the set of all POD locations p_j, and $stop_{p_j}$ be the transit stop of minimum distance from p_j. Then, pop_i may be reassigned to any POD location $p_j \in P \mid dist(p_j, stop_{p_j}) \leq d_w - dist(pop_i, stop_{pop_i})$. Therefore, if for a particular POD p_j, $dist(p_j, stop_{p_j}) > d_w$, out-of-reach populations cannot be reassigned to this POD. Further, if for an out-of-reach population $dist(pop_i, stop_{pop_i}) > d_w$, then this population cannot be reassigned to another POD.

In practice, limitations on transit capacities will influence the movement of popula-

TABLE 5.3. Alternate POD Facilities to which Planners May Choose to Re-Assign the Vulnerable Population of Tract 439100300 While Respecting the Maximum Walking Distance Estimate $d_w = 2000$ Meters.

POD id	Distance from POD to closes transit stop	Total distance vulnerable population would have to traverse
116	17 m	996 m
88	21 m	999 m
153	47 m	1,026 m
90	48 m	1,027 m
150	68 m	1,046 m
136	96 m	1,074 m
10	141 m	1,119 m
100	183 m	1,162 m
95	218 m	1,196 m
15	288 m	1,266 m
13	290 m	1,268 m
113	344 m	1,322 m
157	350 m	1,329 m
19	423 m	1,401 m
99	737 m	1,716 m
98	833 m	1,811 m

tions. Graph theoretical methodologies must be employed to fully assess these limitations. However, reassignment decisions must also consider that the POD location in a plan with $min(dist(p_j, stop_{p_j}))$ will minimize the total walking distance for all populations eligible for POD reassignment. If the sum of reassigned populations is too large, this POD facility (and the transit network leading to it) may be overwhelmed. Therefore, a strategy which starts by reassigning populations to candidate POD locations which maximize total walking distance should be explored to take advantage of as many candidate POD locations as possible.

Populations which are neither within walking distance of a POD nor a public transit stop must be addressed by either adding additional POD(s) or strategically modifying the public transit system to reach these populations (i.e. paratransit). The demand-capacity ratio resulting from implementation of response plans must be analyzed for each link in the public transit system to ensure sufficient transit infrastructure is provided. Effectively eliminating the distance between PODs and transit stops by implementing shuttles between them would maximize the coverage areas of PODs via the public transit system.

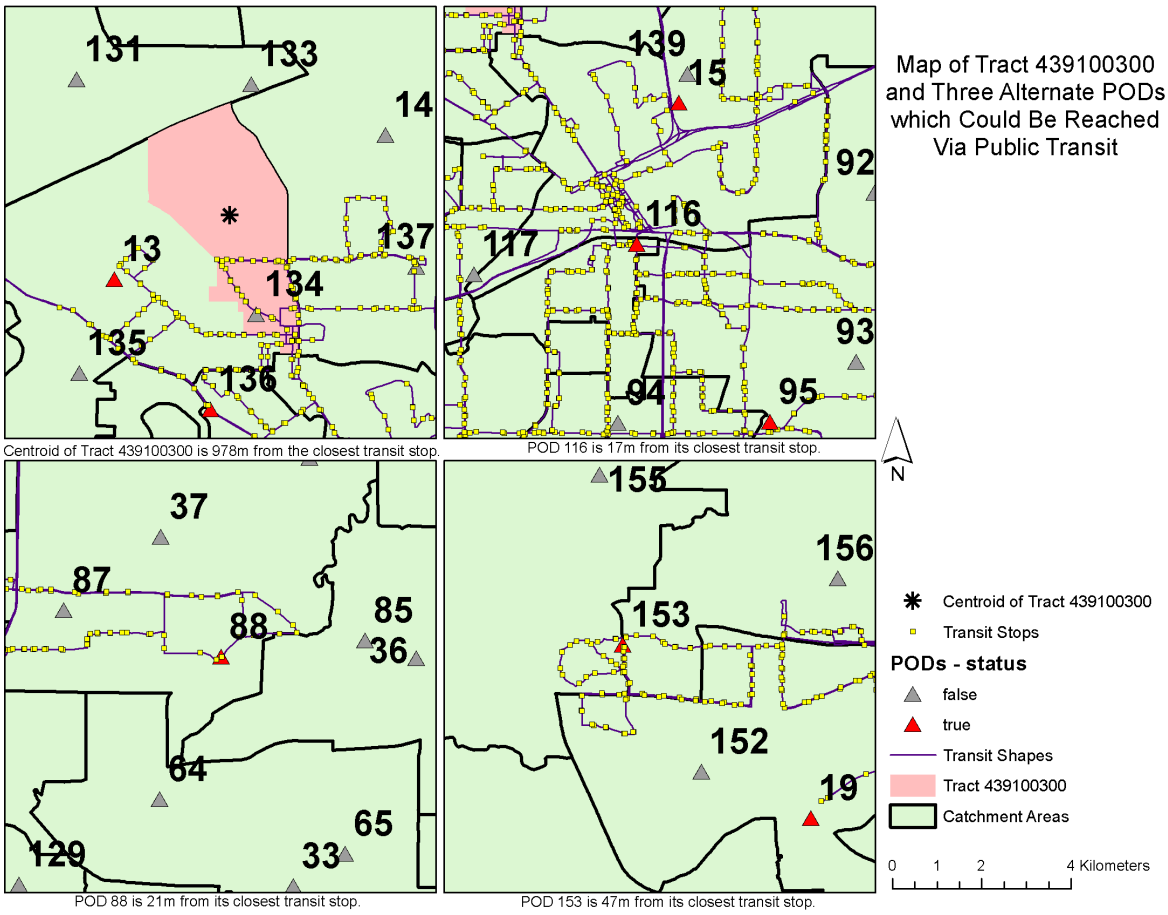


FIGURE 5.37. Census tract 439100300 and the three PODs to which this tract's population could be reassigning while minimizing total walking distance

5.2. Examining Sensitivity of Efforts towards Reach Maximization

The inclusion property described in Lemma 4.1 is not guaranteed by maximization methods described in this chapter. Increasing estimates of d_w can result in the selection of a different facility location by reach maximization algorithms. This different facility location is likely to cover a different set of populations, thus violating the inclusion property. Nonetheless, specific properties exist regarding the sensitivity of maximization algorithms described in this chapter.

Lemma 5.4 states that, given the maximization methods described in Algorithms 6 and 7, the reach of a vulnerable populations is nondecreasing with increasing estimates of d_w. Lemma 5.5 states that using Algorithm 8, the reach of vulnerable populations is nondecreasing with decreasing values of y. Corollary 5.6 follows from the proof of Lemma 5.5

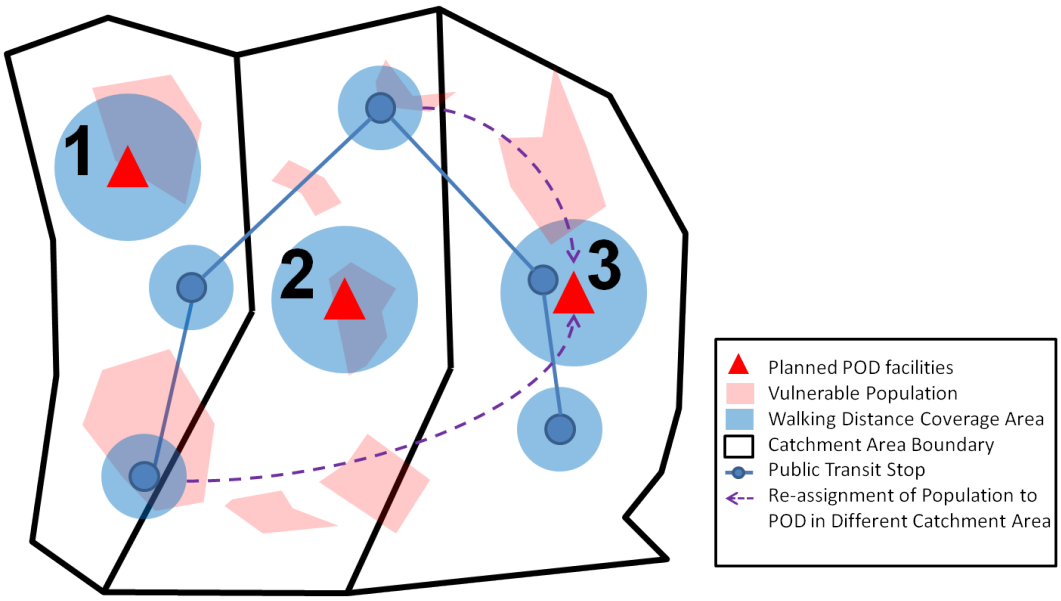


FIGURE 5.38. Method for maximizing reach of transportation vulnerable populations by reassigning populations to PODs in different catchment areas.

and defines an inclusion property for Algorithm 8 which exists only for populations within $d_w - y$ of transit stops when the value of y is less than the minimum distance between all transit stops and candidate facility locations in a particular catchment area.

LEMMA 5.4. *Assuming stationery populations, facilities, catchment areas, and transit stops, let* $reach(f_i, d_w)$ *be the reach (in numbers of individuals) of a facility* f_i *with maximum walking distance* d_w*. Then, using Algorithms 6 or 7,* $d_{w_1} < d_{w_2} \implies reach(f_i, d_{w_1}) \leq reach(f_i, d_{w_2})$ *for catchment area* i*.*

PROOF. Let d_{w_1} and d_{w_2} be estimates of maximum walking distance such that $d_{w_1} < d_{w_2}$. If the facility chosen by Algorithms 6 or 7 remains the same for both d_{w_1} and d_{w_2}, then Corollary 4.2 applies, and the inclusion property presented in Lemma 4.1 holds as a more rigorous form of Lemma 5.4. Otherwise, if the facility chosen by Algorithms 6 or 7 changes from f_1 to f_2 with the change from d_{w_1} to d_{w_2}, further exploration is needed.

Let facility f_1 be chosen for d_{w_1} and f_2 be chosen for d_{w_2}. This implies that $reach(f_1, d_{w_1}) \leq reach(f_2, d_{w_2})$ since, by Lemma 4.1, for a particular facility f_i, $reach(f_i, d_{w_1}) \leq reach(f_i, d_{w_2})$. Therefore, $reach(f_2, d_{w_1}) \leq reach(f_1, d_{w_1}) \leq reach(f_1, d_{w_2}) \leq reach(f_2, d_{w_2})$, and the Lemma holds. □

LEMMA 5.5. *Assuming stationery populations, facilities, catchment areas, and transit stops, let* $min(x_i)$ *be the minimum distance among all candidate facilities in a catchment area* i *to their closest transit stops. Let* y *be a distance which a new transit stop will be added to the chosen facility location. Let* $reach(c_i, y)$ *be the reach (in numbers of individuals) of the facility location selected by Algorithm 8. Let* y_1 *and* y_2 *be two different distances at which a transit stop may be added near the chosen facility location. Then, for catchment area* i*,* $min(x_i) \geq y_1 \geq y_2 \implies reach(c_i, y_1) \leq reach(c_i, y_2)$.

PROOF. Adding a new transit stop at a distance y closer to the facility location than the closest existing transit stop will have no effect on the walking distance coverage area drawn around the facility itself. However, if $y < d_w$, the walking distance coverage area drawn around the transit stops will increase through the addition of the new transit stop.

Assume that $min(x_i) \geq y_1 \geq y_2$ and that $reach(c_i, y_1) > reach(c_i, y_2)$. By decreasing the distance from the facility to the new transit stop from y_1 to y_2, the radius of the coverage area around each transit stop is increased by $(y_1 - y_2)$ as depicted in Figure 5.39. The choice of particular facility location has no effect on coverage areas around transit stops since a new transit stop must be added at the specified distance of the chosen facility. The only difference between the coverage areas of different facility locations within a catchment area are the areas within d_w of the facility itself. Therefore, since the coverage area (and associated reach) around the facility itself does not change with changes in y, and the coverage area (and associated reach) around transit stops is nondecreasing as y is increased, then $reach(c_i, y_1) \not> reach(c_i, y_2)$. Therefore, Lemma 5.5 holds since decreasing y leads to an increase in coverage area and nondecreasing reach.

□

COROLLARY 5.6. *Assuming stationery populations, facilities, catchment areas, and transit stops, let* $min(x_i)$ *be minimum distance among all candidate facilities in a catchment area* i *to their closest transit stops. Let* y_1 *and* y_2 *be distances at which a new transit stop will be added to the chosen facility location such that* $min(x_i) \geq y_1 \geq y_2$*. Using Algorithm 8, if a*

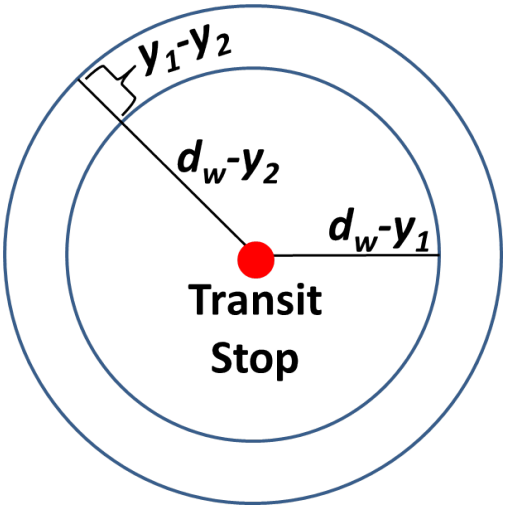


FIGURE 5.39. Diagram showing that, given two distances y_1 and y_2 at which a new transit stop is added near the facility such that the new transit stop is the closest transit stop to the facility and $y_1 > y_2$, the walking distance coverage area increases with the decrease in distance from the new transit stop to the facility.

population pop_i *is covered by the chosen facility location* f_i *given a specific maximum walking distance* d_w *with a transit stop added* y_1 *distance from* f_i, pop_i *will also be covered by* f_i *given the same maximum walking distance* d_w *if the transit stop is added* y_2 *distance from* f_i.

CHAPTER 6

EXPLORING THE APPLICATION OF GRAPH ALGORITHMS TO PUBLIC TRANSIT DATA TO ANALYZE AND MAXIMIZE PLAN REACH

Although public transit network data has been used in the methodology of Chapters 4 and 5 to estimate coverage areas of facility locations, a more detailed computational representation which respects the network's spatio-temporal nature is needed [108]. Many public transit authorities currently make their complete schedule and route information available online using a General Transit Feed (GTF). A computational model which represents the movement of transit vehicles in both time and space (i.e. geographic location) is needed to facilitate further analysis and optimization of public transit resources.

A description of GTF data is provided in 4.2.2, and the entity relationships within the GTF specification are depicted in Figure 4.5 [57]. Although each trip corresponds to a single route, an individual stop location can be used by many different trips of many different routes. This property allows passengers to transfer from one route to another. However, it also complicates the data structure required to build an accurate computational model of the transit system.

One approach to modeling GTF data is through the use of graphs. Data from a system's GTF can be translated into a directed graph which represents both spatial and temporal aspects of the transit system's routes and schedule. Once the directed graph model of a GTF has been constructed, it may be analyzed and modified using graph theoretical algorithms.

6.1. Constructing the Graph Data Structure from GTF Data

A graph vertex must be created for each unique [location, time] pair represented in the GTF. Three types of edges must be created in the directed graph in order to accurately model the actual transit system:

- Moving edges

- Waiting-with-transit-vehicle edges
- Waiting-without-transit-vehicle edges

Moving edges represent the movement of transit vehicles in geographic space (i.e. from one stop to another). Waiting-with-transit-vehicle edges represent the transit vehicle remaining at a particular stop location for a specified duration of time. Waiting-without-transit-vehicle edges represent the ability of transit riders to simply wait for a transit vehicle at a transit stop for some duration of time. The set of vertices and all three types of edges can be identified using spatio-temporal transit schedule data located in the *stop_times.txt* file of a transit authority's GTF.

Each record in the *stop_times.txt* file has fields for *trip_id*, *arrival_time*, *departure_time*, *stop_id*, and *stop_sequence*. The *stop_id* field relates each record in this file to a specific record in the *stops.txt* file which specifies a geographic location. For each record in *stop_times.txt* for which *departure_time* is different than *arrival_time*, two vertices must be created in the directed graph with an edge leading from the *arrival_time* vertex to the *departure_time* vertex. This directed edge is a *waiting-with-transit-vehicle* edge which represents a duration of time the transit vehicle is waiting at the stop. Otherwise, if arrival and departure times are equal, only a single vertex is required to represent both arrival and departure. The *stop_sequence* field represents the order in which stop locations are visited on each trip and must be used to define the *moving* edges in the graph. Finally, subsequent stop times at each stop location must be examined to create the set of *waiting-without-transit-vehicle* edges. Figure 6.1a depicts a series of vertices at the same location with edges between them. Figure 6.1b depicts an example graph with all three types of edges representing a GTF.

Formally, let graph $G = (V, E)$ be a directed graph such that V is a set of vertices and E is a set of directed edges between vertices in V. The set V represents a set of times at specific geographic locations drawn from the *stop_times.txt* file of a GTF. Each vertex in V is labeled using a string built from the concatenation of the *stop_id*, the underscore (_), and the time of day (either *arrival_time* or *departure_time*) represented by the vertex.

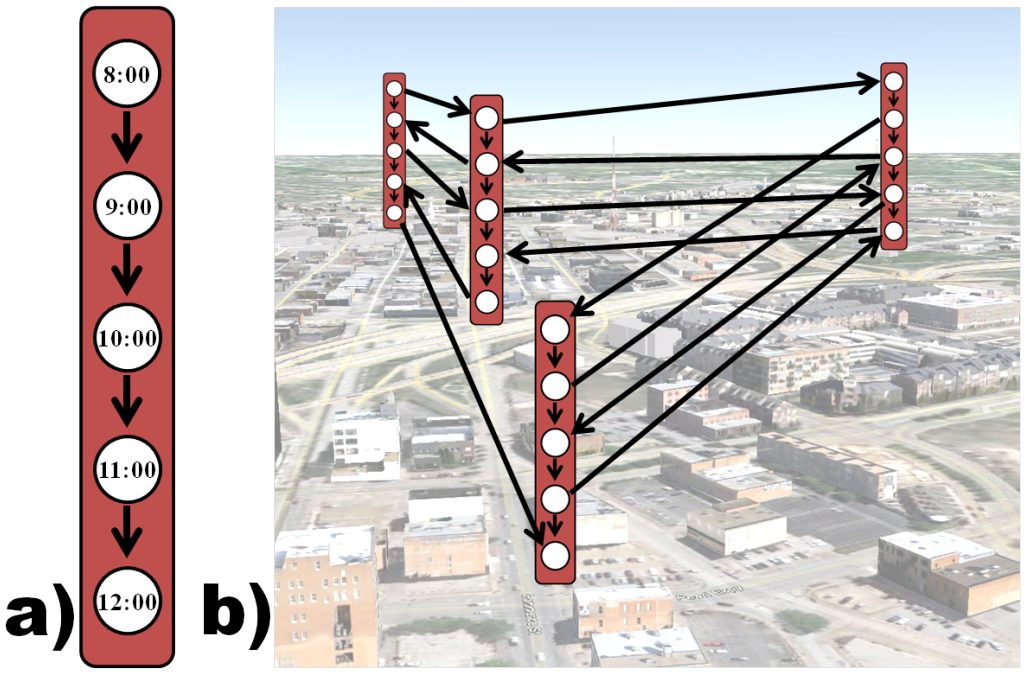


FIGURE 6.1. Example directed graph representation of GTF data

Input: A table *stop_times* representing the *stop_times.txt* GTF file and a specific *trip_id* in the table: 377214.

Output: All records in table *stop_times* with *trip_id*=377214, ordered by ascending *stop_sequence*.

```
SELECT
    *
FROM
    stop_times
WHERE
    trip_id=377214
ORDER BY
    stop_sequence ASC;
```

Algorithm 9: Example query to return records from *stop_times* for a specific *trip_id* ordered by *stop_sequence*.

Input: A table *stop_times* representing the *stop_times.txt* GTF file and a specific *stop_id* in the table: 1468.

Output: All times represented in table *stop_times* with *trip_id*=1468, ordered by ascending time.

```
SELECT
   time
FROM
   (
      SELECT
         departure_time AS time,
         stop_id,
         'departure' AS type
      FROM
         stop_times
      UNION ALL
      SELECT
         arrival_time AS time,
         stop_id,
         'arrival' AS type
      FROM
         stop_times
   )   AS unioned
WHERE
   stop_id=1468
ORDER BY
   time ASC
```

Algorithm 10: Example query to return all times (both arrival and departure)represented in records from *stop_times* for a specific *stop_id* ordered by time of day.

The database query provided in Algorithm 9 selects all records in the database table *stop_times* (representing the *stop_times.txt* GTF file) which have a specific *trip_id*, returning results ordered by the *stop_sequence* field. The query provided in Algorithm 10 returns an ordered list of all times (whether arrival or departure) specified in the *stop_times* table for a specific *stop_id*. These queries are used in the pseudocode provided in Algorithm 11 to build the sets V and E of the directed graph.

```
Input: A database table stop_times representing the stop_times.txt GTF file

Output: Graph object representing transit network specified in GTF file
        stop_times.txt

Graph graph = new Graph()
List distinct_trip_ids = list of trip ids
foreach trip in distinct_trip_ids do
    List stop_times = Object representing results of query in Algorithm 9
    previous_departure_vertex = null
    foreach s in stop_times do
        graph.createVertex(s.stop_id + "_" + s.arrival_time)
        if stop_time.arrival_time < stop_time.departure_time then
            graph.createVertex(s.stop_id + "_" + s.departure_time)
            graph.createEdge(s.stop_id + "_" + s.arrival_time, s.stop_id + "_" +
            s.departure_time)
        end
        if previous_departure_vertex ≠ null then
            graph.createEdge(previous_departure_vertex, s.stop_id + "_" +
            s.arrival_time)
        end
        previous_departure_vertex = s.stop_id + "_" + s.departure_time
    end
end
foreach stop_id in stops do
    List times=Object representing results of query in Algorithm 10
    Time last_time = times[0]
    for int i=1; i<times.size; i++ do
        graph.createEdge(stop_id + "_" + last_time, stop_id + "_" + times[i])
        last_time = times[i]
    end
end
```

Algorithm 11: Pseudocode to create a graph model of a public transit network using GTF data

The sizes of $|V|$ and $|E|$ for a graph are bounded by specific properties of the GTF it was created to represent. In the largest-case scenario, the arrival and departure times for each record in the *stop_times.txt* file are different. Let T be the set of all records in the *stop_times.txt* file. Then, $|V| \leq 2 * |T|$ since, for each record in T, at most two new vertices must be added to G. Further, if two records in T specify the same location and time, a new vertex does not need to be created.

To determine a bound for $|E|$, each of the three types of edges must be examined separately. Let E_m represent the set of *moving edges* in G and Υ represent the set of all trips

in the *trips.txt* file. Then, $|E_m| = |T| - |\Upsilon|$ since, for each trip, the number of vertices is one more than the number of edges. Let E_w represent the set of all *waiting-with-transit-vehicle edges*. Then, in the largest case, each record in T requires the creation of an edge in E_w, and $|E_w| \leq |T|$. Let E_{wo} represent the set of all *waiting-without-transit-vehicle edges*. Let L represent the set of records in the *stops.txt* file. Then, assuming the stop location in each record of L corresponds to at least one record in T, $|E_{wo}| \leq |T| - (|L| + 1)$ since, for each additional stop used, one more record in T must represent an edge in E_m. As an example, in the GTF data obtained for Tarrant County, Texas, $|T| = 115,276$, $|L| = 1,996$, $|\Upsilon| = 2,959$, and $|V| = 86,364$.

For each edge in G, there must be a corresponding duration of time. This time either represents a waiting time at a transit stop or the time it takes for a transit vehicle to move from one stop to another. A cycle in the graph would imply that it is possible to move backwards in time. Therefore, the graph is acyclic [117] and falls into the class of directed acyclic graphs (DAGs). DAGs consist of a set of vertices with directed edges placed between them such that, when traversing the graph along the links, it is impossible to visit the same vertex twice [37]. They have been well-studied in mathematics and computer science [39][58][77][103][122][125][136]. Existing DAG algorithms from the literature can be adapted and applied to the biological emergency public transit problem [17][19][23].

6.2. Analyzing Transit Resources Using the GTF Graph

Feasible response plans must allow populations to make a round trip to their assigned POD facility location. Public transit schedules which are limited to certain times of the day may create situations where travel to the POD is possible, but where a return trip is unavailable. Further, the first trips of the day originating at transit stops near each POD location do not contribute to the transit system's capability of providing round-trips to the POD. Using the DAG, these can be approached through the application of graph theoretical algorithms [15][47][50][61][142][149].

Whether a specific POD is reachable by a population assigned to it using the public

transit network must be considered. Formally stated, let $l_m \in L$ be the location of a population and $l_n \in L$ be the location of the POD to which it is assigned. Let $V_{l_m} \subset V$ be a set of vertices representing different times at location l_m and $V_{l_n} \subset V$ be a set of vertices representing different times at location l_n such that $V_{l_m} \bigcap V_{l_n} = \emptyset$. The reachability of a POD by the population can be determined using G by calculating the reachability of each vertex in V_{l_m} from each vertex in V_{l_n}. The reachability of G can be studied using techniques described by King in [73] which employ Las-Vegas type randomized algorithms to construct and maintain a tree representing the transitive closure of the graph. These algorithms are dynamic and will thus support modifications of the graph which are likely to result from plan participation maximization efforts.

An all-pairs shortest paths algorithm for directed graphs is also presented by King [73]. Given two vertices $v_i, v_j \in V$, this algorithm will determine the shortest path in G from v_i to v_j. When modeling the flow of populations through the transit network, this algorithm will facilitate the choice of specific paths populations can take through the graph.

Feasible time intervals for travel from specific locations to and from each POD may be limited. In order to examine these intervals, POD service time estimates (observed during real-life POD drills [74][120] recommended by the Centers for Disease Control and Prevention [27]) must be considered. For the population at location l_m to participate in the mitigation plans, it must be able to make a round trip to its assigned POD at location l_n. Participation of this population requires the existence of paths in G from some vertex $v_p \in V_{l_m}$ to some vertex $v_q \in V_{l_n}$ and from some vertex $v_r \in V_{l_n}$ to some vertex $v_s \in V_{l_m}$ such that $time(v_q) < time(v_r)$. Further, let τ represent the time an individual must spend at a POD in order to be served. As illustrated in Figure 6.2, this service time limits the number of different paths an individual may take to return home from the POD by requiring that $\big(time(v_q) + \tau\big) < time(v_r)$. Both reachability and all-pairs shortest paths algorithms must be employed to calculate feasible time intervals for travel with respect to POD service time estimates.

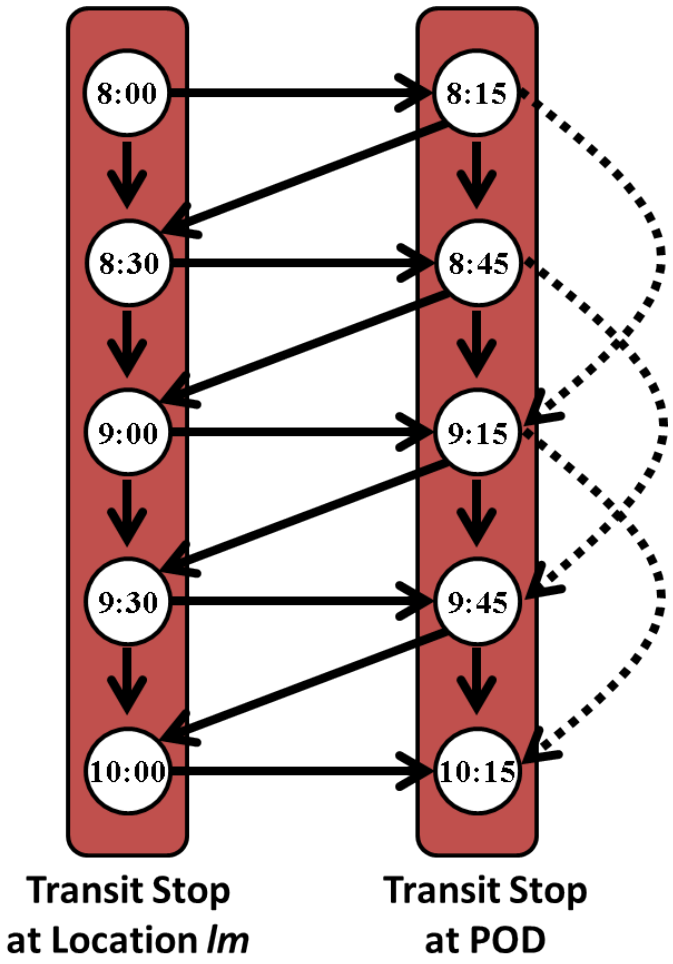


FIGURE 6.2. How POD service time affects transit network reachability and capacity. Solid arrows represent directed edges in the graph of the GTF. Dotted arrows indicate earliest possible departure time from POD for each arrival time given a POD service time (τ) of one hour.

CHAPTER 7

CONCLUSION

Response plans must be created to protect the population during biological emergencies resulting from the deliberate or accidental release of harmful biochemical substances. The Centers for Disease Control and Prevention (CDC) has recommended that mitigation plans include the creation of a set of ad-hoc clinics known as Points of Dispensing (PODs) designed to quickly provide medical countermeasures (MCMs) to the entire population of a region [27].

The importance of MCMs in emergency response efforts is evidenced by continued federal funding to develop and procure them. The Department of Health and Human Services has allocated $415 million in fiscal year 2014 to support advanced research and development of MCMs. Further, an additional $250 million has been allocated in fiscal year 2014 as the *first installment* of a multi-year commitment to acquire MCMs through Project BioShield [115]. Although these MCMs include assets for response to chemical, biological, radiological, and nuclear (CBRN) events, the Department of Homeland Security has identified anthrax as the agent most likely to be used in bio-terrorist acts [70]. Further, in fiscal years 2004 - 2012, MCMs for response to the release of anthrax comprised 44% of all MCMs procured [70]. However, having MCMs is only the first step towards a successful mitigation campaign [43]. The Public Health Emergency Medical Countermeasures Enterprise (PHEMCE) recognizes the existence of diverse and unique vulnerabilities in the population and has expressed its commitment to address gaps in MCM application resulting from such vulnerabilities [116]. Therefore, research described in this dissertation on maximizing participation in efforts to distribute MCMs in a timely manner is of utmost importance to the federal government's public health security efforts.

Individuals in the population must travel to their assigned PODs in order to participate in mitigation activities. However, certain spatial and/or demographic attributes can preclude travel to a POD by portions of the population [131]. In this context, these

attributes represent *transportation vulnerabilities* which must be addressed in order to maximize participation in response efforts. To this end, methods to identify and quantify specific vulnerabilities were described, and a variety of strategies to maximize participation in response efforts were explored. These methods and strategies have been incorporated into an existing computational framework in order to facilitate their adoption by real-world practitioners and policymakers.

Algorithms to quantify and analyze the reach of transportation vulnerable populations resulting from specific response plans were described in Chapter 4. Given a set of POD facility locations with corresponding service areas and an estimate of maximum walking distance (d_w), these algorithms determine whether each vulnerable population in a region is capable of participating in response efforts. Separate algorithms were presented for plans including operational and non-operational public transit systems. An *inclusion property* was formally defined in Lemma 4.1 to guarantee inclusion of specific populations under increasing estimates of d_w.

The analysis and quantification algorithms presented in Chapter 4 serve as a foundation upon which reach maximization algorithms described in Chapter 5 were constructed. The following three types of reach maximization strategies were explored:

- Unconstrained facility location selection
- Constrained population-to-POD assignment preserving
- Constrained population-to-POD reassignment

Although the *inclusion property* described in Lemma 4.1 does not hold for algorithms based on these maximization strategies, each of these algorithms are shown to have non-decreasing total reach with increasing estimates of d_w.

The unconstrained facility location selection strategy described in Section 5.1.2 uses the location and sizes of vulnerable populations in each service area together with the d_w estimate to identify an optimal set of POD locations to maximize plan reach within each service area. This same strategy was described in two ways: once using methods common in the field of geographic information systems (GIS), and once using graph theoretical

methodologies. The maximum number of facility locations needed to reach 100% coverage of vulnerable populations was defined using clique enumeration and coverage techniques from graph theory. The number of facility locations required to achieve complete coverage can be further reduced by busing individuals from identified locations directly to and from their assigned PODs.

The constrained strategies described in Algorithms 6 and 7 preserve population-to-POD assignments. These assignments were calculated to maintain uniform population service areas created using previously developed methodologies [112]. One approach uses a set of candidate facilities in each service area and an estimate of d_w to choose the facility location in the service area which maximizes reach of vulnerable individuals. A second approach adds a new public transit stop within a specific distance of the POD facility in order to increase the number of vulnerable populations who can use public transit to travel to and from their assigned PODs. A third approach combines these two methods by selecting the location with the greatest reach assuming a single transit stop is added to each service area within a specific distance of the chosen POD location.

The constrained population-to-POD reassignment strategy identifies vulnerable populations which are outside of the reach of their assigned POD and determines reassignment options to send them to different POD locations using public transit. This method can be used to reach otherwise excluded vulnerable populations which are within d_w distance of their closest transit stop. However, the strategy which adds transit stops within a distance y of each POD facility can be employed to include these same populations.

The addition of transit stops can be implemented by busing individuals between a POD and a transit stop. Further, let the endpoint of this new bus route closest to the POD be y_{POD} distance from the POD, and the endpoint closest the transit stop be y_{stop} distance from the closest transit stop. If busing individuals all the way to the POD or to the transit stop is not feasible, Algorithm 8 still applies as long as $y = y_{POD} + y_{stop}$. Lemma 5.4 states that reach of vulnerable populations is nondecreasing with increasing estimates of d_w. Lemma 5.5 states that reach is nondecreasing with decreasing values of y, and Corollary 5.6

defines a special version of the *inclusion property* for populations surrounding public transit stops.

Maximization algorithms described in this dissertation choose the optimal facility location out of a set of locations on a *per catchment area* basis. In order to preserve benefits of previous work on uniform population partitioning, catchment area boundaries are not modified during reach maximization algorithms. Although the maximization algorithms presented in this dissertation do calculate the best possible solution, they are not greedy.

A comparison of results from the different reach and analysis methods developed which respect catchment area boundaries is provided in Figure 7.1. The maximization method which added a transit stop at each POD location is calculated to reach over 59% of Tarrant County's 53,411 vulnerable individuals. This represents a significant improvement over the approximately 8.5% reach of the original plan without the use of public transit.

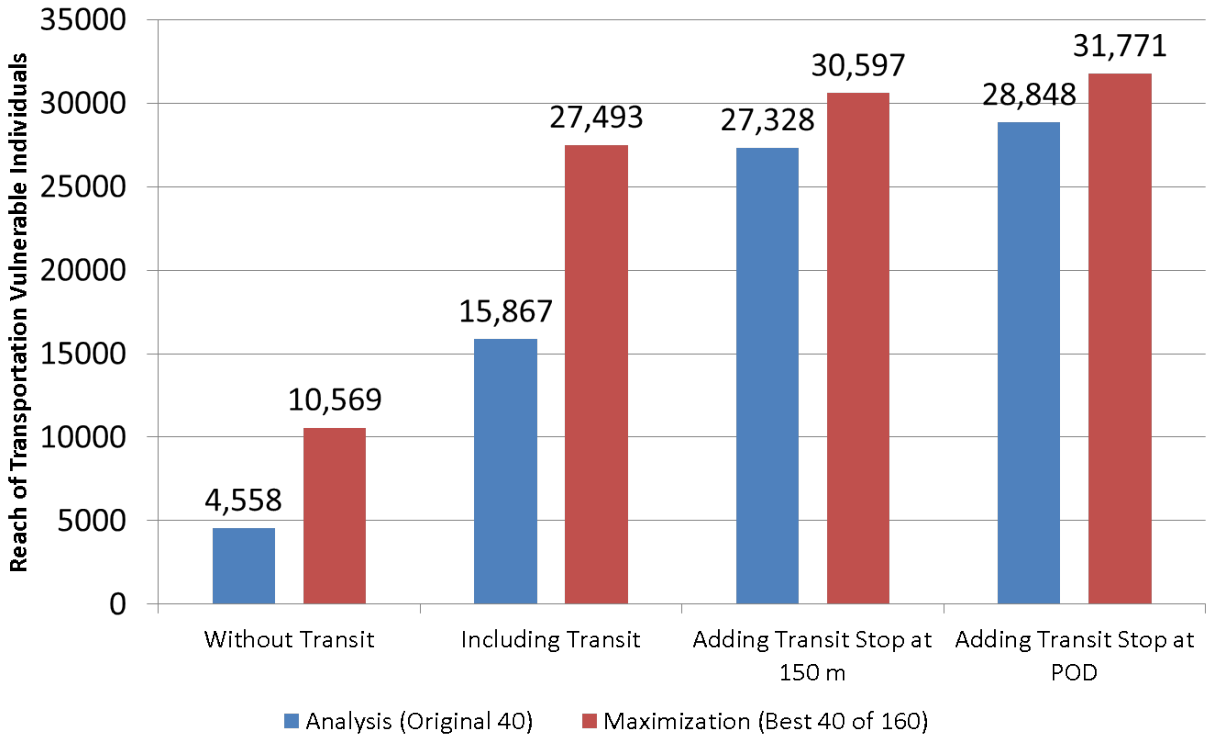


FIGURE 7.1. Comparison of vulnerable population reach of different analysis and maximization methods using plans and algorithms described in Chapters 3, 4, and 5 chapters.

If additional facility locations are available, reach maximization algorithms can dra-

matically improve a plan's reach of transportation vulnerable populations. However, the availability of public transit is the single factor which most greatly increase a plan's reach. In the examples provided, without adding transit stops, optimization approximately doubles a plan's reach of vulnerable populations while the availability of public transit approximately triples it. In the analysis presented, a combination of optimization and public transit availability increase the plan's reach sixfold.

An approach using a directed graph to more accurately model the spatio-temporal aspects of a transit system is described in Chapter 6. Algorithm 10 specifies the steps needed to construct the mathematical directed graph model. Vertices of the resulting graph represent geographic locations at specific points in times. Therefore, barring the invention of backward time travel technology, this graph falls into the class of *directed acyclic graphs*. The graph model of the public transit data will facilitate the application of algorithms from graph theory in order to further explore the properties of the transit network with respect to given response plans.

Analysis and maximization algorithms implemented during research described in this dissertation were built on top of the RE-PLAN Framework [101]. The RE-PLAN Framework manages database tables used in parameterized queries, making newly implemented algorithms region-independent. It also provides a graphical interface for constructing plans, executing analysis and optimization algorithms, and exploring resulting on an interactive map. Traffic analysis methodology described in Chapter 4.1 represents early work on the RE-PLAN Framework. Implementing these algorithms using the RE-PLAN Framework provides a direct path from research and development in academia to deployment and use by public health preparedness practitioners, thus enhancing the real-world impact of algorithms described in this dissertation.

Early methodologies which led to the RE-PLAN Framework grew out of a collaboration which began in 2008 between Tarrant County Public Health (TCPH) and the University of North Texas (UNT) Center for Computational Epidemiology and Response Analysis (CeCERA). Since 2010, software based on the these methodologies has been an integral part

of the biological emergency response planning process at TCPH. Software user training sessions have been conducted for TCPH planners, and county stakeholders have been trained regarding plans created using the software.

Methodologies developed for response plan design and analysis led to the implementation of the RE-PLAN Framework during the project "A Computational Framework for Assessing the Feasibility of Bio-emergency Response" (funded by NIH 1R15LM010804-01). The region independence and extensibility of this framework were designed to facilitate the real-world impact of newly developed analysis and design methods. Methodology described in this dissertation to identify and address vulnerabilities was performed during the project "Minimizing Access Disparities in Bio Emergency Response Planning" (funded by NIH 1R01LM011647-01).

7.1. Limitations

Limitations of methods described in this dissertation largely stem from the availability, accuracy, and granularity of data they require. Low granularity estimations of otherwise unavailable data can lead to inaccuracies. For example, traffic count observation records obtained and used for traffic analysis do not accurately reflect the situation on the road network. As depicted in Figure 7.2, a particular traffic count could result from few cars on the road (i.e. low density) or from many cars on the road (i.e. high density) moving slowly. A speed component is needed to distinguish between these two situations. Although equipment exists to collect traffic speed data in addition to traffic counts, only traffic count data was available when methods described in this dissertation were implemented. Therefore, low traffic density is assumed for all traffic count records in the *business as usual* traffic estimation method described in Chapter 4.1. This assumption can lead to underestimation of roadway congestion for periods of high density during which the counts were initially observed and collected.

Further, population data used by methods described in this dissertation cause problems related to availability, accuracy, and granularity. Spatial data linking specific populations to specific geographic locations are limited by the accuracy and currency of data

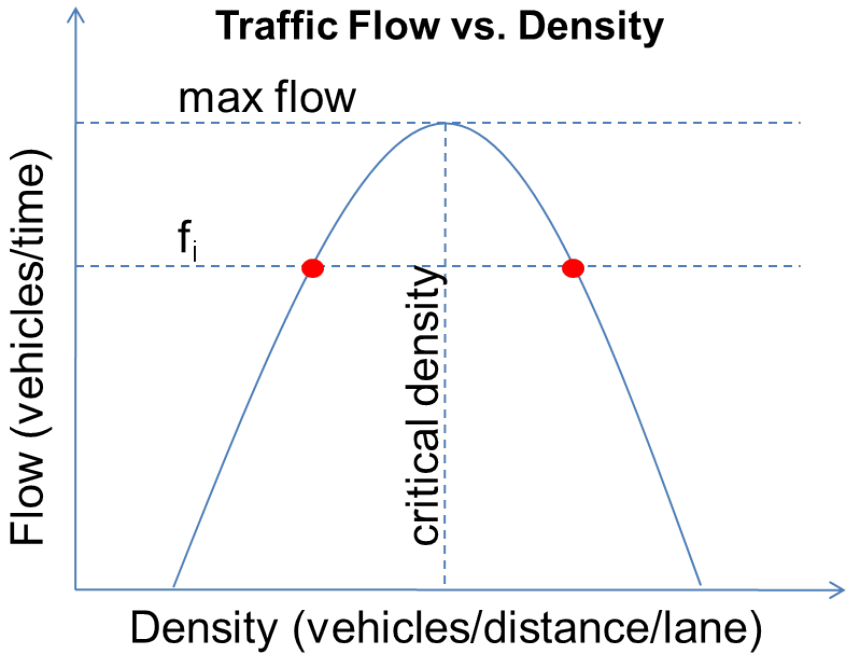


FIGURE 7.2. Given only an observation of a specific flow of traffic f_i, it is impossible to know whether this flow is the result of high speed and low vehicle density, or of low speed and high vehicle density [133].

collection methods. Examples presented in this dissertation utilized data from the U.S. Census Bureau at the spatial granularity of counties, tracts, block groups, and blocks depicted in the hierarchy in Figure 7.3. These data include trade-offs between spatial granularity, currency, and accuracy. For example, decennial census population data represent actual counts of the entire population at geographic units as fine as the census block. However, counts are only updated every ten years. Population data from the American Community Survey are derived from samples collected using a survey rather than a complete count of the population, and these counts are only available for geographic units as fine as block groups or census tracts. Nevertheless, these counts represent more current estimates of the population. Further, many different demographic attributes (e.g. private vehicle ownership by household size) are available in the American Community Survey which are not available in the decennial census. After considering trade-offs between the availability, accuracy, and granularity of population data, the American Community Survey was selected as the source of data for examples presented in this dissertation. Nonetheless, the methods described were

designed to use more accurate, higher granularity data if it becomes available.

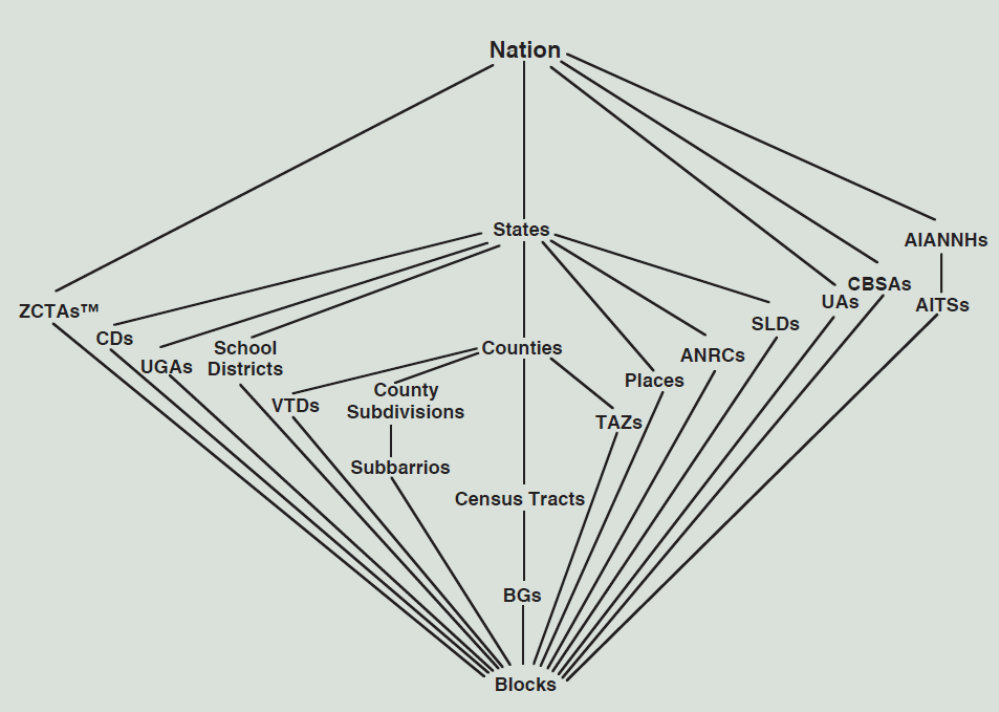


FIGURE 7.3. Geographic Hierarchy of Census Entities [128]

In analysis and optimization methods described in Chapters 4 and 5, public transit is assumed to be unlimited in capacity, and temporal aspects of the system are not considered. If POD operations are assumed to continue around the clock during the emergency mitigation period, schedule constraints are of less concern than capacity constraints. Therefore, these methods focus on the Transportation Research Board's *spatial availability* [126] requirements of public transit listed in Chapter 2.3.

The joining of data using spatial attributes can be achieved using a variety of methods. For examples presented in this dissertation, joins between polygons from different tables were implemented by using the polygon attributes from one table and calculated geometric centroids of the polygon attributes from the second table. Specifically, when spatially joining walking distance coverage area buffers to population block polygons, a population block was assumed to be completely within a buffer if its geometric centroid was within this buffer. This can lead to over- or under- estimations of the vulnerable population actually within walking distance coverage areas. An alternate method for performing spatial joins involves joining

walking distance coverage areas to all population blocks they overlap. This would guarantee an over-estimate of the vulnerable population within the walking distance coverage area. Alternatively, the proportion of each population block covered could be used to estimate the proportion of each block's vulnerable population which is covered. Nonetheless, due to the non-uniform population distribution within each population block, this method can also lead to over- or under-estimations of the vulnerable population actually within walking distance coverage areas.

7.2. Future Work

Many opportunities remain for further development, exploration, and expansion of concepts described in this dissertation. The use of Global Positioning System (GPS) directions by individuals traveling to their assigned PODs could be modeled using routing algorithms on a graph representation of the road network. Results from this graph routing method should be compared with results from the method currently implemented as part of the RE-PLAN Framework [112]. This may lead to a better estimate of business as usual traffic.

The public transit graph model described in Chapter 6 should be implemented and explored. This model would consider all public transit requirements (listed in Chapter 2.3) from the Transportation Research Board except for *information availability*. Walking distances between populations' home locations and beginning transit stops, as well as the walking distance from the ending transit stops to the PODs, can be integrated using methodology similar to that described in Chapter 4 to consider *spatial availability* requirements. The *temporal availability* will be considered as a result of the spatio-temporal nature of the directed graph. If capacity data is available, *capacity availability* analyses can be performed since the model will allow specific populations to be routed across specific edges in the directed graph which represent trips (specific transit vehicles) at specific times and locations.

Many different graph libraries can be used to implement the road network and public transit graph models. However, the RE-PLAN Framework is currently implemented using a Java client and a database running PostgreSQL. Therefore, the Java graph library

JGraphT [96] and the PostgreSQL database routing library *pgRouting* [104] should be explored. Nonetheless, custom written graph methods may also represent desirable options.

Once the public transit graph model has been implemented, it is important to distinguish between what is theoretically possible and what is realistic. Maximum flow algorithms may be used to calculate the maximum performance of the transit network with respect to vulnerable populations and their assigned PODs. However, if individuals prefer taking the shortest path through the network, this maximum performance estimate may be unrealistic. Therefore, theoretical maximum capacity estimates may be calculated using shortest path algorithms on each vulnerable population and its assigned POD. Further, if ridership data is available for a transit feed, it can be used to validate the calculated maximum capacity estimates for this feed.

Methods described in this dissertation utilize population data linked to spatial polygon (vector) data. New population data sources such as the LandScan project under development at Oak Ridge National Laboratory (ORNL) utilize raster data formats. LandScan uses a variety of data including census data and satellite imagery to estimate population counts during different times of day at 3 arc-seconds resolution (equivalent to a grid cell on the Earth's surface $\frac{1}{1200}$ degrees longitude by $\frac{1}{1200}$ degrees latitude in size [44]) [18]. Methods described in this dissertation may be modified to use raster data in analysis and maximization algorithms as an alternative to vector data.

Many algorithms presented in this dissertation can be considered *embarrassingly parallel*. Specifically, those algorithms (e.g. Algorithms 7 or 8) which perform tasks on a per-catchment-area basis can be separated into a number of parallel tasks with very little effort. For purposes of scalability, these algorithms may be parallelized to enhance their performance on larger data sets and over increasingly large numbers of catchment areas.

7.3. Broader Impacts

Research described in this dissertation focuses specifically on the development and analysis of methods to maximize participation in biological emergency response plans. Spatial population and demographic data are used with general transit feed data to identify

transportation vulnerable populations, and methods to maximize a plan's reach of these vulnerable populations are presented. Methodologies employed to address these problems were built upon previous work in topics including transportation science, location allocation modeling, graph theory, and public health preparedness. Nonetheless, methods described are applicable to many different domains.

Acceptable walking distances *to* and *from* transit stops are important in the fields of transit research and planning [126]. In the transit research literature, these two distances are estimated separately. For example, Guerra, Cervero, and Tischler [59] note that evidence supports a 0.25 mile distance around places of employment and 0.5 mile distance around places of residence. These distance estimates are used to calculate transit catchment areas, plan land use, and model expected ridership [59]. The methods described in this dissertation contribute to the fields of transit research and planning by providing catchment areas for specific locations using public transit, thus linking the *to* and *from* distances into a single distance estimate.

The ability to estimate catchment areas for specific points using public transit also provides a contribution to the fields of retail geography and public health. Specific populations may be targeted by calculating their catchment areas. These catchment areas may be used to determine optimal retail or health services locations (respectively) for the target populations.

Transportation vulnerability analysis methods described in this dissertation also contribute to the field of emergency planning. During large scale evacuations prior to foreseeable hazardous events (e.g. hurricanes or floods), these methods can be used to identify locations of populations who lack transportation resources. Once identified, walking distance catchment areas can be calculated for these populations. The placement of emergency transit resources may be planned according to intersections of populations' catchment areas. This would facilitate the use of emergency transit resources to evacuate transportation vulnerable populations.

BIBLIOGRAPHY

[1] B. Adenso-Diaz and F. Rodriguez, *A Simple Search Heuristic for the MCLP : Application to the Location of Ambulance Bases in a Rural Region*, Omega, Int. J. Mgmgt Sci. 25 (1997), no. 2, 181–187.

[2] W. Neil Adger, *Vulnerability*, Global Environmental Change 16 (2006), no. 3, 268–281.

[3] Pankaj K. Agarwal and Micha Sharir, *Efficient Algorithms for Geometric Optimization*, ACM Computing Surveys 30 (1998), no. 4, 412–458.

[4] Dionne M Aleman, Theodorus G Wibisono, and Brian Schwartz, *ACCOUNTING FOR INDIVIDUAL BEHAVIORS IN A PANDEMIC DISEASE SPREAD MODEL*, Proceedings of the 2009 Winter Simulation Conference (Austin, TX), no. Newman 2002, 2009, pp. 1977–1985.

[5] Dennis P Andrulis, Nadia J Siddiqui, and Jenna L Gantner, *Preparing racially and ethnically diverse communities for public health emergencies.*, Health affairs 26 (2007), no. 5, 1269–79.

[6] Thomas A. Arcury, John S. Reisser, Wilbert M. Gesler, and James M. Powers, *Access to Transportation and Health Care Utilization in a Rural Region*, The Journal of Rural Health 21 (2006), no. 1, 31–38.

[7] Hina Arora, T.S. Raghu, and Ajay Vinze, *Resource allocation for demand surge mitigation during disaster response*, Decision Support Systems 50 (2010), no. 1, 304–315.

[8] Assistant Secretary for Preparedness & Response, *Pandemic and All-Hazards Preparedness Act Progress Report on the Implementation of Provisions Addressing At-Risk Individuals*, Tech. Report August, Department of Health & Human Services, 2008.

[9] Prasith Baccam and Michael Boechler, *Public health response to an anthrax attack: an evaluation of vaccination policy options.*, Biosecurity and bioterrorism : biodefense strategy, practice, and science 5 (2007), no. 1, 26–34.

[10] Prasith Baccam, David Willauer, Justin Krometis, Yongchang Ma, Atri Sen, and Michael Boechler, *Mass Prophylaxis Dispensing Concerns: Traffic and Public Access*

to PODs, Biosecurity and bioterrorism : biodefense strategy, practice, and science 9 (2011), no. 2, 139–151.

[11] Hannah Badland, Marcus White, Gus Macaulay, Serryn Eagleson, Suzanne Mavoa, Christopher Pettit, and Billie Giles-Corti, *Using simple agent-based modeling to inform and enhance neighborhood walkability*, International Journal of Health Geographics 12 (2013), 58.

[12] John N. Balog, *Public Transportation Security: Public transportation emergency mobilization and emergency operations guide*, Tech. report, Transportation Research Board, Washington, D.C., 2005.

[13] Salvatore Barese, Frank Acevedo, and Lawrence Herman, *Anthrax: Medical Response to a Biologic Terrorist Attack*, Physician Assistant 25 (2001), no. 12.

[14] Chris L Barrett, Stephen G Eubank, and James P Smith, *If smallpox strikes Portland....*, Scientific American 292 (2005), no. 3, 54–61.

[15] Richard Bellman, *On a Routing Problem*, Quart. Appl. Math. 16 (1958), 87–90.

[16] Oded Berman, Jorg Kalcsics, Dmitry Krass, and Stefan Nickel, *The ordered gradual covering location problem on a network*, Discrete Applied Mathematics 157 (2009), no. 18.

[17] Oded Berman and Dmitry Krass, *The generalized maximal covering location problem*, Computers & Operations Research 29 (2002), no. 6, 563–581.

[18] Budhendra Bhaduri, Edward Bright, Phillip Coleman, and Marie L. Urban, *LandScan USA: a high-resolution geospatial and temporal modeling approach for population distribution and dynamics*, GeoJournal 69 (2007), no. 1-2, 103–117.

[19] Randeep Bhatia, Sudipto Guha, Samir Khuller, and Yoram J Sussmann, *Facility Location with Dynamic Distance Functions*, Journal of Combinatorial Optimization 2 (1998), no. 3, 199–217.

[20] Evelyn Blumenberg and Margy Waller, *The Long Journey to Work: A Federal Transportation Policy for Working Families*, The Brookings Institution Series on Transportation Reform (2003).

[21] Margaret L Brandeau, Jessica H McCoy, Nathaniel Hupert, Jon-Erik Holty, and Dena M Bravata, *Recommendations for modeling disaster responses in public health and medicine: a position paper of the society for medical decision making.*, Medical decision making : an international journal of the Society for Medical Decision Making 29 (2009), no. 4, 438–60.

[22] Jack Brimberg, Pierre Hansen, Nenad Mladenović, and Eric D Taillard, *IMPROVEMENTS AND COMPARISON OF HEURISTICS FOR SOLVING THE UNCAPACITATED MULTISOURCE WEBER PROBLEM*, Operations Research 48 (2000), no. 3, 444–460.

[23] C. Caruso, a. Colorni, and L. Aloi, *Dominant, an algorithm for the p-center problem*, European Journal of Operational Research 149 (2003), no. 1, 53–64.

[24] Center for the Advancement of Distance Education at the School of Public Health of the University of Illinois at Chicago, *The POD Game*.

[25] Centers for Disease Control and Prevention, *Smallpox Disease and Its Clinical Management*, 2003.

[26] ———, *Smallpox Disease Overview*, 2004.

[27] ———, *Receiving, Distributing, and Dispensing Strategic National Stockpile Assets: A Guide for Preparedness, Version 10.02*, Tech. Report August, 2006.

[28] ———, *Emergency Preparedness and Response: Key Facts about the Cities Readiness Initiative (CRI)*, 2008.

[29] ———, *Public Health Workbook to Define, Locate, and Reach Special, Vulnerable, and At-risk Populations in an Emergency*, 2010.

[30] ———, *Cities Readiness Initiative*, 2011.

[31] Robert Cervero, *Alternative Approaches to Modeling the Travel-Demand Impacts of Smart Growth*, Journal of the American Planning Association 72 (2006), no. 3, 285–295.

[32] Robert Cervero, Jin Murakami, and Mark Miller, *Direct Ridership Model of Bus Rapid Transit in Los Angeles County*, 2009.

[33] Hsinchun Chen, Daniel Zeng, and Ping Yan, *Syndromic Surveillance Data Sources and Collection Strategies*, Infectious Disease Informatics 21 (2010), no. 1, 33–48.

[34] James Chin, *Control of Communicable Diseases Manual*, 17 ed., American Public Health Association, 2000.

[35] Leon Cooper, *LOCATION-ALLOCATION PROBLEMS*, Operations Research 11 (1963), no. 3, 331–344.

[36] Courtney D Corley, Armin R Mikler, Karan P Singh, Diane J Cook, Fort Worth, and Computer Science, *Monitoring Influenza Trends through Mining Social Media*, Proceedings of the 2009 International Conference on Bioinformatics and Bioengineering (BIOCOMP09) (Las Vegas, NV), 2009.

[37] Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein, *Introduction to Algorithms*, 2 ed., The Massachusetts Institute of Technology, Cambridge, Massachusetts, 2001.

[38] Monica Crubézy, Martin O Connor, Zachary Pincus, and Mark A Musen, *Ontology-Centered Syndromic Surveillance for Bioterrorism*, IEEE Intelligent Systems 20 (2005), no. 5, 26–35.

[39] Narsingh Deo, *Graph Theory with Applications to Engineering and Computer Science*, Prentice Hall, 2004.

[40] Division of Bioterrorism Preparedness and Response, NCPDCID, and Office of Infectious Diseases, *Emergency Preparedness and Response: Questions and Answers about Smallpox Vaccine*, 2009.

[41] Zvi Drezner and Horst W. Hamacher, *Facility Location: Application and Theory*, Springer, New York, 2001.

[42] David P Eisenman, Kristina M Cordasco, Steve Asch, Joya F Golden, and Deborah Glik, *Disaster planning and risk communication with vulnerable communities: lessons from Hurricane Katrina.*, American journal of public health 97 Suppl 1 (2007), no. Supplement 1, S109–15.

[43] Christina England, *The United States Medical Countermeasure Enterprise: A Broken*

Link in US Biopreparedness, Journal of Homeland Security and Emergency Management 11 (2014), no. 1, 117–130.

[44] Environmental Systems Research Institute, *ArcUser Online: Measure in Arc-Seconds.*

[45] ———, *ESRI Shapefile Technical Description*, Tech. Report July, 1998.

[46] Stephen Eubank, Hasan Guclu, V S Anil Kumar, Madhav V Marathe, Aravind Srinivasan, Zoltan Toroczkai, and Nan Wang, *Modelling disease outbreaks in realistic urban social networks*, Nature 429 (2004), no. May, 180–184.

[47] Shimon Even and Yossi Shiloach, *An On-Line Edge-Deletion Problem*, Journal of the ACM 28 (1981), no. 1, 1–4.

[48] Federal Emergency Management Agency and US Army Corps of Engineers, *IS-26 Guide to Points of Distribution (PODs)*, Tech. Report December, 2008.

[49] Lauren S Fernandez, Deana Byard, Chien-Chih Lin, Samuel Benson, and Joseph A Barbera, *Frail elderly as disaster victims: emergency management strategies.*, Prehospital and disaster medicine 17 (2002), no. 2, 67–74.

[50] Robert W Floyd, *Algorithm 97: Shortest path*, Communications of the ACM 5 (1962), no. 6, 345.

[51] Robert A Fowler, Gillian D Sanders, Dena M Bravata, Bahman Nouri, Jason M Gastwirth, Dane Peterson, Allison G Broker, Alan M Garber, and Douglas K Owens, *Cost-Effectiveness of Defending against Bioterrorism : A Comparison of Vaccination and Antibiotic Prophylaxis against Anthrax*, Annals of Internal Medicine 142 (2005), no. 8, 601 – 610.

[52] Gyozo Gidofalvi, Tore Risch, Torben Bach Pedersen, and Erik Zeitler, *Highly Scalable Trip Grouping for Large Scale Collective Transportation Systems*, EDBT'08 (Nantes, France), 2008, pp. 678–689.

[53] Jeremy Ginsberg, Matthew H Mohebbi, Rajan S Patel, Lynnette Brammer, Mark S Smolinski, and Larry Brilliant, *Detecting influenza epidemics using search engine query data.*, Nature 457 (2009), 1012–4.

[54] M. Giovachino, T. Calhoun, N. Carey, B. Coleman, G. Gonzalez, B. Hardeman, and

B. McCue, *Optimizing a District of Columbia Strategic National Stockpile Dispensing Center*, J Public Health Manag Pract 11 (2005), no. 4, 282–90.

[55] Jeffrey B Goldberg, *Operations Research Models for the Deployment of Emergency Services Vehicles*, EMS Management Journal 1 (2004), no. 1.

[56] J S Goodwin, W C Hunt, and J M Samet, *Determinants of cancer therapy in elderly patients.*, Cancer 72 (1993), no. 2, 594–601.

[57] Google Developers, *GTFS and GTFS-realtime*.

[58] Jonathan L. Gross and Jay Yellen, *Handbook of Graph Theory*, CRC Press, Boca Raton, Florida, 2003.

[59] Erick Guerra, Robert Cervero, and Daniel Tischler, *Half-Mile Circle*, Transportation Research Record: Journal of the Transportation Research Board (2012), no. 2276, 101–109.

[60] S I Harewood, *Emergency ambulance deployment in Barbados : a multi-objective approach*, Journal of the Operational Research Society 53 (2002), 185–192.

[61] Monika R. Henzinger and Valerie King, *Randomized fully dynamic graph algorithms with polylogarithmic time per operation*, Journal of the ACM 46 (1999), no. 4, 502–516.

[62] Jeffrey W Herrmann, Sara Lu, and Kristen Schalliol, *A Routing and Scheduling Approach for Planning Medication Distribution*, Proceedings of the 2009 Industrial Engineering Research Conference, 2009.

[63] Hsinchun Chen, Daniel Zeng and Ping Yan, *Infectious Disease Informatics: Syndromic Surveillance for Public Health and Bio-Defense*, Springer, New York, 2010.

[64] Anette Hulth, Gustaf Rydevik, and Annika Linde, *Web queries as a source for syndromic surveillance.*, PloS one 4 (2009), no. 2.

[65] Nathaniel Hupert, Daniel Wattson, Jason Cuomo, Eric Hollingsworth, Kristof Neukermans, and Wei Xiong, *Predicting hospital surge after a large-scale anthrax attack: a model-based analysis of CDC's cities readiness initiative prophylaxis recommendations.*, Med Decis Making 29 (2009), no. 4, 424–37.

[66] Nathaniel Hupert, Wei Xiong, Kathleen King, Michelle Castorena, Caitlin Hawkins,

Cindie Wu, and John A Muckstadt, *Uncertainty and Operational Considerations in Mass Prophlyaxis Workforce Planning*, Disaster Medicine and Public Health Preparedness 3 (2009), no. SUPPL. 2, S121–S131.

[67] Thomas V Inglesby, Donald A Henderson, John G Bartlett, Michael S Ascher, Edward Eitzen, Arthur M Friedlander, Jerome Hauer, Joseph Mcdade, Michael T Osterholm, Tara O Toole, Gerald Parker, Trish M Perl, Philip K Russell, and Kevin Tonat, *Anthrax as a Biological Weapon*, Journal of the American Medical Association 281 (1999), no. 18, 1735–1745.

[68] Tamara Jimenez, Chetan Tiwari, Armin R. Mikler, and Marty O'Neill II, *Maps, Rates, and Fuzzy Mountains: Generating Meaningful Risk Maps*, IEEE International Conference on Bioinformatics and Biomedicine (Philadelphia, PA), 2012.

[69] Mairi Joyce and Roger Dunn, *A Methodology for Measuring Public Transport Accessibility to Employment a case study , Auckland CBD , NZ*, ITE 2010 Annual Meeting and Exhibit (Vancouver, Canada), 2010, pp. 1–14.

[70] Robert Kadlec, *Policy Brief: Renewing the Project BioShield Act*, Tech. Report January, Center for a New American Security, 2013.

[71] Edward H Kaplan, David L Craft, and Lawrence M Wein, *Emergency response to a smallpox attack: the case for mass vaccination.*, Proceedings of the National Academy of Sciences of the United States of America 99 (2002), no. 16, 10935–40.

[72] Roger E. Kasperson, Ortwin Renn, Paul Slovic, Halina S. Brown, Jacque Emel, Robert Goble, Jeanne X. Kasperson, and Samuel Ratick, *The Social Amplification of Risk: A Conceptual Framework*, Risk Analysis 8 (1988), no. 2, 177–187.

[73] Valerie King, *Fully Dynamic Algorithms for Maintaining All-Pairs Shortest Paths and Transitive Closure in Digraphs*, 40th Annual Symposium on Foundations of Computer Science (New York City, NY), 1999, pp. 81–89.

[74] Howard K. Koh, Lori J. Elqura, Christine M. Judge, John P Jacob, Amy E Williams, M Suzanne Crowther, Richard A Serino, and John M Auerbach, *Implementing the*

Cities Readiness Initiative: Lessons Learned from Boston, Disaster Medicine and Public Health Preparedness 2 (2008), no. 1, 40–49.

[75] Edward Krenzelok, Erma MacPherson, and Rita Mrvos, *Disease surveillance and non-prescription medication sales can predict increases in poison exposure.*, Journal of medical toxicology 4 (2008), no. 1, 7–10.

[76] Kevin J Krizek, *CHAPTER 6 PERSPECTIVES ON ACCESSIBILITY AND*, Access to Destinatons (David Levinson and Kevin J. Krizek, eds.), Elsevier, London, UK, 2005, pp. 109–130.

[77] Steffen L. Lauritzen, *Graphical Models*, Oxford University Press, Oxford, 1996.

[78] E. K. Lee, C.-H. Chen, F. Pietz, and B. Benecke, *Modeling and Optimizing the Public-Health Infrastructure for Emergency Response*, Interfaces 39 (2009), no. 5, 476–490.

[79] Eva K. Lee, Siddhartha Maheshwary, Jacquelyn Mason, and William Glisson, *Large-Scale Dispensing for Emergency Response to Bioterrorism and Infectious-Disease Outbreak*, Interfaces 36 (2006), no. 6, 591–607.

[80] Young M. Lee, *Analyzing Dispensing Plan for Emergency Medical Supplies in the Event of Bioterrorism*, Proceedings of the 2008 Winter Simulation Conference, 2008, pp. 2600–2608.

[81] Herbert S. Levinson, Samuel Zimmerman, Jennifer Clinger, James Gast, Scott Rutherford, and Eric Bruhn, *Bus Rapid Transit Volume 2: Implementation Guidelines*, Tech. report, Transportation Research Board, Washington, D.C., 2003.

[82] Hollie M Lund, Robert Cervero, and Richard W. Wilson, *Travel Characteristics of Transit-Oriented Development in California*, Tech. Report January, California Department of Transportation, Sacramento, CA, 2004.

[83] Ted Maddry, *Planning and Executing a POD Mission*.

[84] Liang Mao and Ling Bian, *A Dynamic Network with Individual Mobility for Designing Vaccination Strategies*, Transactions in GIS 14 (2010), no. 4, 533–545.

[85] F. E. Maranzana, *On the Location of Supply Points to Minimize Transport Coststst*, Operations Research Quarterly 15 (1964), 261–270.

[86] Vladimir Marianov and Daniel Serra, *Location - Allocation of Multiple-Server Service Centers with Constrained Queues or Waiting Times*, Annals of Operations Research 111 (2002), no. 1, 35–50.

[87] Anna C Mastroianni, *Slipping through the net: social vulnerability in pandemic planning.*, The Hastings Center report 39 (2009), no. 5, 11–2.

[88] Nimrod Megiddo and Kenneth J. Supowit, *On the complexity of some common geometric location problems*, SIAM J. Comput. 13 (1984), no. 1.

[89] Microsoft, *Bing Transit - Add your transit data to Bing Maps*, 2012.

[90] Nenad Mladenović, Jack Brimberg, Pierre Hansen, and José a. Moreno-Pérez, *The p-median problem: A survey of metaheuristic approaches*, European Journal of Operational Research 179 (2007), no. 3, 927–939.

[91] Colleen Monahan and John A. Franco, *From Research to Practice : Creating a Virtual POD Using a Collaborative Virtual Environment*, 2011.

[92] Gordon E Moore, *Cramming More Components onto Integrated Circuits*, Electronics 38 (1965), no. 8, 114–117.

[93] M Mourez, R S Kane, J Mogridge, S Metallo, P Deschatelets, B R Sellman, G M Whitesides, and R J Collier, *Designing a polyvalent inhibitor of anthrax toxin.*, Nature biotechnology 19 (2001), no. 10, 958–61.

[94] Alan T Murray, Rex Davis, Robert J Stimson, and Luis Ferreira, *PUBLIC TRANSPORTATION ACCESS*, Transpn. Res.-D 3 (1998), no. 5, 319–328.

[95] National Imagery Mapping Agency, *Department of Defense World Geodetic System 1984*, Tech. report, 2000.

[96] Barak Naveh, *JGraphT*.

[97] Christopher Nelson, Edward W. Chan, Anita Chandra, Paul Sorensen, Henry H. Willis, Kathering Comanor, Hayoung Park, Karen A. Ricci, Leah B. Caldarone, Molly Shea, John A. Zambrano, and Lydia Hansell, *Recommended Infrastructure Standards for Mass Antibiotic Dispensing*, Tech. report, RAND Corporation, Santa Monica, CA, 2008.

[98] North Central Texas Council of Governments, *Dallas-Fort Worth Regional Travel Model (DFWRTM): Model Description*, Tech. report, Arlington, TX, 2009.

[99] Stephen D Nutley, *Rural transport problems and non-car populations in the USA: A UK perspective*, Journal of Tra 4 (1996), no. 2, 93–106.

[100] Oak Ridge Institute for Science and Education, *Biological Attacks: Communication Challenges*.

[101] Martin O'Neill II, Armin R. Mikler, and Tamara Schneider, *An Extensible Software Architecture to Facilitate Disaster Response Planning*, BIOCOMP'11 (Las Vegas, NV), 2011.

[102] B. Pelegrin and L. Canovas, *A new assignment rule to improve seed points algorithms for the continuous k-center problem*, European Journal Of Operational Research 104 (1998), 366–374.

[103] Sriram Pemmaraju and Steven Skiena, *Computational Discrete Mathematics*, Cambridge University Press, New York, 2003.

[104] pgRouting Community, *pgRouting Project*.

[105] Jomon Aliyas Pual and Govind Hariharan, *Hospital Capacity Planning for Efficient Disaster Mitigation during a Bioterrorist Attack*, Proceedings of the 2007 Winter Simulation Conference (Washington, D.C.), no. 1976, 2007, pp. 1139–1147.

[106] Anke Richter and Sinan Khan, *Pilot Model: Judging Alternate Modes of Dispensing Prophylaxis in Los Angeles County*, Interfaces 39 (2009), no. 3, 228–240.

[107] Anne Rinchiuso-Hasselmann, Ryan L. McKay, Christopher A. Williams, David T. Starr, Beth Maldin Morgenthau, Jane R. Zucker, and Marisa Raphael, *Protecting the Public from H1N1 through Points of Dispensing (PODs)*, Biosecurity and bioterrorism : biodefense strategy, practice, and science 9 (2011), 13–21.

[108] Kenneth H. Rosen, *Discrete Mathematics and Its Applications*, 4 ed., WCB McGraw-Hill, Boston, 1999.

[109] Lisa D Rotz and James M Hughes, *Advances in detecting and responding to threats*

from bioterrorism and emerging infectious disease., Nature medicine 10 (2004), no. 12 Suppl, S130–6.

[110] Susanne Schmehl, Stephanie Deutsch, Johann Schrammel, Lucas Paletta, and Manfred Tscheligi, *Directed Cultural Probes: Detecting Barriers in the Usage of Public Transportation*, INTERACT 2011 (Lisbon, Portugal), 2011, pp. 404–411.

[111] Matthew C. Schmidt, Nagiza F. Samatova, Kevin Thomas, and Byung-Hoon Park, *A scalable, parallel algorithm for maximal clique enumeration*, Journal of Parallel and Distributed Computing 69 (2009), no. 4, 417–428.

[112] Tamara Schneider, *A Framework for Analyzing and Optimizing Regional Bio-Emergency Response Plans*, Dissertation, University of North Texas, 2010.

[113] Tamara Schneider and Armin R. Mikler, *RE-PLAN: A Computational Tool for Response Plan Analysis*, International Journal of Functional Informatics and Personalised Medicine 3 (2010), no. 2, 103–121.

[114] Tamara Schneider, Armin R. Mikler, and Martin O'Neill II, *Computational Tools for Evaluating Bioemergency Contingency Plans*, Proceedings of the 2009 International Conference on Disaster Management, 2009.

[115] Kathleen Sebelius, *Statement on the President's Fiscal Year 2014 Budget*, 2013.

[116] U.S. Department of Health Services and Human, *2012 Public Health Emergency Medical Countermeasures Enterprise (PHEMCE) Strategy*, Tech. report, 2012.

[117] Matthew R Silver and Olivier L De Weck, *Time-Expanded Decision Networks : A Framework for Designing Evolvable Complex Systems*, Systems Engineering 10 (2007), no. 2, 167–186.

[118] Laron Smith, Richard Beckman, Keith Baggerly, Doug Anson, and Michael Williams, *TRANSIMS: Transportation Analysis and Simulation System*, Tech. Report 26295, Los Alamos National Laboratory, 1995.

[119] James E. Stem, *State Plane Coordinate System of 1983*, Tech. report, NOAA, Rockville, MD, 1990.

[120] Andy Stergachis, Catherine M Wetmore, Michelle Pennylegion, Randal D Beaton,

Bryant T Karras, Dean Webb, Diane Young, and Michael Loehr, *Evaluation of a mass dispensing exercise in a Cities Readiness Initiative setting.*, Am J Health-Syst Pharm 64 (2007), no. 3, 285–93.

[121] Brock Stewart, Charles E Rose, Jerome I Tokars, Stacey W Martin, Wendy a Keitel, Harry L Keyserling, Janiine Babcock, Scott D Parker, Robert M Jacobson, Gregory A Poland, and Michael M McNeil, *Health-related quality of life in the CDC Anthrax Vaccine Adsorbed Human Clinical Trial*, Vaccine 30 (2012), no. 40, 5875–9.

[122] Kozo Sugiyama, *Graph Drawing and Applications for Software and Knowledge Engineers*, World Scientific, River Edge, NJ, 2002.

[123] The Associated Press, *New Orleans Empties Ahead of Ivan*, 2004.

[124] The White House, *The Federal Response to Hurricane Katrina*, Tech. Report February, Washington, D.C., 2006.

[125] K. Thulasiraman and M. N. S. Swamy, *Graphs: Theory and Algorithms*, John Wiley & Sons, Inc., New York, 1992.

[126] Transit Research Board, *Transit Capacity and Quality of Service Manual*, 2 ed., Washington, D.C., 2003.

[127] Urban Design 4 Health, *San Diego Region Health Impact and Forecasting Assessment.*

[128] U.S. Census Bureau, *Strength in numbers: Your Guide to Census 2010 Redistricting Data from the U.S. Census Bureau*, Tech. Report 6172, February 2010.

[129] U.S. Department of Health & Human Services, *FY 2005 Peformance and Accountability Report*, Tech. report, 2005.

[130] ———, *Strategic National Stockpile (SNS) - Radiation Emergency Medical Management*, 2011.

[131] U.S. Department of Health and Human Services, *Strategic National Stockpile (SNS)*, 2012.

[132] U.S. Department of Homeland Security, *Individuals with Disabilities in Emergency Preparedness: Executive Order 13347*, Tech. Report July, 2005.

[133] U.S. Department of Transportation Federal Highway Administration, *The 1985 High-*

way Capacity Manual - A Summary, Office of Traffic Operations, Washington, D.C., 1986.

[134] U.S. General Accounting Office, *COMBATING TERRORISM: Chemical and Biological Medical Supplies Are Poorly Managed*, 1999.

[135] U.S. Government Printing Office, *Pandemic and All-Hazards Preparedness Act, Public Law No. 109-417*, 2006.

[136] Tyler John VanderWeele, *Contributions to the theory of causal directed acyclic graphs*, Ph.D. thesis, Harvard University, 2006.

[137] Martin Wachs and T. Gordon Kumagai, *Physical Accessibility as a Social Indicator*, Socio-Econ. Plan. Sci. 7 (1973), 437–456.

[138] Harvey M. Wagner, *ON A CLASS OF CAPACITATED TRANSPORTATION PROBLEMS*, Management Science 5 (1959), no. 3, 304–318.

[139] Jarrett Walker, *Human Transit*, 2011.

[140] Lawrence M Wein, David L Craft, and Edward H Kaplan, *Emergency response to an anthrax attack.*, Proceedings of the National Academy of Sciences of the United States of America 100 (2003), no. 7, 4346–51.

[141] White House, *Homeland Security Presidential Directive/HSPD-21: Public Health and Medical Preparedness*, 2007.

[142] P.D. Whiting and J.A. Hillier, *A Method for Finding the Shortest Route through a Road Network*, Operational Research 11 (1960), no. 1, 37–40.

[143] Sony Sulaksono Wibowo and Piotr Olszewski, *MODELING WALKING ACCESSIBILITY TO PUBLIC TRANSPORT TERMINALS : CASE STUDY OF SINGAPORE MASS RAPID TRANSIT*, Journal of the Eastern Asia Society for Transportation Studies 6 (2005), 147–156.

[144] Michael J. Widener and Mark W. Horner, *A hierarchical approach to modeling hurricane disaster relief goods distribution*, Journal of Transport Geography 19 (2011), no. 4, 821–828.

[145] Henry H. Willis, Christopher Nelson, Shoshana R. Shelton, Andrew M. Parker, John A.

Zambrano, Edward W. Chan, Jeffrey Wasserman, and Brian A. Jackson, *Initial Evaluation of the Cities Readiness Initiative*, Tech. report, RAND Health, Santa Monica, CA, 2009.

[146] Stephan Winter, Monika Sester, Glenn Geers, and Ouri Wolfson, *Towards a Computational Transportation Science*, SIGMOD Record 39 (2010), no. 3, 27–32.

[147] World Health Organization, *Smallpox*, 2001.

[148] ———, *The World Health Report 2007: A Safer Future - Global Public Health in the 21st Century*, Tech. report, Geneva, Switzerland, 2007.

[149] F Benjamin Zhan and Charles E Noon, *Shortest Path Algorithms : An Evaluation using Real Road Networks*, Transportation Science 32 (1998), no. 1, 65–73.

[150] Dennis Zielstra and Hartwig Hochmair, *Comparison of Shortest-Path Lengths for Pedestrian Routing in Street Networks Using Free and Proprietary Data*, Transportation Research Board 91st Annual Meeting, no. 954, 2012.