PEDESTRIAN FLOW MODELING FOR PROTOTYPICAL MARYLAND CITIES

Prepared for

Maryland Department of Transportation
Division of Highway Safety Programs
Hanover, MD

November, 2004

TABLE OF CONTENTS

TABLE OF CONTENTS (cont'd)

LIST OF FIGURES

LIST OF FIGURES (cont'd)

LIST OF TABLES

INTRODUCTION

Pedestrian safety is emerging as a major area of concern for MPO's and planning agencies. Typically, pedestrian safety has been analyzed by either examining the absolute number of pedestrian crashes at a location, or computing an exposure rate from the number of crashes and the traffic volume. A more desirable measure would be an exposure rate based on the pedestrian volume, but it has not proven feasible to obtain pedestrian flow volumes on a wide-area basis to support this analysis. This report describes a pedestrian flow modeling process that was developed under the sponsorship of the Maryland DOT and the University of Maryland National Center for Smart Growth. The process provides micro-scale daily pedestrian flows on all sidewalks and crosswalks in a substantial coverage area. Two test cases were analyzed: an urban scenario comprising about 10 square miles of downtown Baltimore, and a suburban scenario comprising about 15 square miles of Langley Park in Prince Georges and Montgomery Counties.

The model structure is analogous to a standard four-step process, but with many nuances that reflect the specific features of pedestrian trip making. One important feature is its extensive use of accessibility to employment by type and multifamily housing in generation and distribution. The model is transferable, relying on generalized pedestrian travel characteristics that can be applied to standardized land use and network data. The model is founded on readily available, yet very detailed, input data. The model set includes custom software to process and integrate these various components:

- Census TIGER files that provide overall street network topology;
- Wide-area ortho-photography with 1 foot resolution that provides detailed information on streets, sidewalks, and land use without the need for extensive field observations; and
- Land use data derived from standard property tax record files. The model includes a custom geo-coding module that attaches each property to its correct TIGER line segment and accumulates block face totals for detailed land use categories;

Pedestrian data were derived from the New York home interview survey conducted by the New York Metropolitan Transportation Council and North Jersey Transportation Planning Authority. This survey tracked all household trips including non-motorized trips, and provides travel data for the full range of area types, ranging from the most intense CBD (Manhattan) to suburban and rural conditions.

A unique network builder and assignment process automatically generates sidewalks, intersection crosswalks, jay walk crosswalks and other features such as traffic signals and street widths that affect path choice and barrier effects. Methods are provided to manually override and refine this data. The model was validated against a large number of pedestrian counts and was shown to be a useful and acceptably accurate tool for pedestrian flow modeling. Other applications can also be supported, including analysis of land development patterns and the impact of urban design on travel behavior; connections between the built environment, physical activity, and public health outcomes; and further understanding of pedestrian risks.

This report presents the findings of this study in five sections:

LITERATURE REVIEW	Prior research has led to a number or approaches to pedestrian flow modeling. This

	section reviews prior work and establishes the context in which this work is performed.
TECHNICAL DESCRIPTION	Describes the structure of the model, data sources and their relationship to the model application, and calibration / validation findings.
USER GUIDE	Provides a complete description of how to use the model set. The application is a complex set of programs and data files using ArcGIS, TP+, and custom software developed for the application. The model chain is bound together and managed with the CENTRAL process controller.
CASE STUDIES	Describes the datasets and findings of the Baltimore and Langley Park examples, which produced pedestrian volume estimates for both study areas.
SAFETY ANALYSIS	Presents the findings of the comparison of pedestrian crash frequencies to estimated pedestrian volumes, to compute a pedestrian crash exposure rate.

In addition to this study report, a CD-ROM has been prepared that contains model setups and software as well as datasets for the two case studies.

LITERATURE REVIEW

Introduction

Transportation demand modeling has a long history and complex heritage in American traffic engineering and urban planning (Bates, 2000, Newell, 1980). The need to estimate the amount, type, and distribution of vehicular traffic in modern cities is clear, and traffic models have played an important role in the planning and governance of urban growth since the late 1950's (Hensher and Button, 2000; Ortúzar, 1994).

The need and ability to model pedestrian movement is a more recent development, however, resulting from increased interest in the public health, environmental, and social benefits of walking. Advances in computational power and understanding have made new modeling approaches possible, creating a newly emerging field of pedestrian modeling and simulation.

Pedestrian modeling has fundamental differences from vehicle modeling. These differences pose significant challenges to traditional traffic modeling approaches. Kerridge et al. (2001) point out that pedestrian trips are less homogenous than vehicle trips in terms of journey purpose. As a result, route choices are less well determined and subject to higher degrees of variability. Pedestrian trips are often smaller parts of larger trips or a tour of connected trips which use other modes, such as walking to or from a bus or subway stop. Pedestrian networks are also harder to define than vehicular networks because cities and buildings have numerous pathways available to pedestrians that are not available to vehicles. Pedestrians are also not limited to just crossing roads at intersections, for example, and movement through and within buildings are a feasible option as well.

To address these unique challenges, a large number of approaches have been proposed and tested in research communities around the world. This paper will review major developments in pedestrian volume modeling over the past three decades, with a special emphasis on the large number of innovative approaches which have surfaced in the last five to ten years. The review will focus both on the methodology and applied examples from each approach where possible, with an eye towards their relative applicability, accuracy, and usefulness for the planning practitioner.

This review will not concern itself with the large body of literature which addresses *why* people walk as opposed to taking other forms of transportation, commonly known as "mode choice" (Heggie, 1976; Ben – Akiva, 1985). Instead, it presents five major approaches taken by international researchers and practitioners, going into as much methodological detail as appropriate on each approach and their uses. This review will also examine the relative strengths and weaknesses of each approach and conclude with a discussion of future modeling needs and opportunities.

The Big Picture: Five Approaches to Pedestrian Modeling

Attempts to understand pedestrian movement dynamics date back nearly four decades. Early studies focused on the behavior of pedestrians in confined circumstances such as subways, airports, or building entrances (Carstens and Ring, 1970; Hankin and Wright, 1958; Navin and Wheeler, 1969), while others sought a broader understanding of pedestrians in central shopping districts (Pushkarev and Zupan, 1971; Behnam and Patel, 1977; Hoel, 1968; O'Flaherty and

Parkinson, 1972; Older, 1968). This set the stage for two parallel streams of pedestrian modeling research that persists to this day (Teknomo, 2002).

In recent years, measurement tools have become more powerful and sophisticated, resulting in more nuanced and complex models of pedestrian movement prediction. Many of these models have been developed for specific purposes, but all share the goal of helping planners and architects create efficient, comfortable, and safe operating environments in pedestrian facilities such as airports, shopping malls, multi-modal transfer points (Hoogendoorn, 2003). Helbing et al. (2001) classified the approach to these goals as that which those seeking to develop *level-of-service concepts* (Fruin, 1971; Moıri and Tsukaguchi, 1987; Polus et al., 1983), *design elements for pedestrian facilities* such as transport interchanges, public spaces, and fire escapes, (Pauls, 1984; Whyte, 1988), or *general planning guidelines* (Davis and Braaksma, 1988; TRB, 1985).

Whatever their ultimate application, a large number of simulation models have been proposed, including *queuing models* (Lövas, 1994; Yuhaski and Macgregor Smith, 1989), *stochastic models* (Modesti and Sciomachen, 1999; Ashford et al., 1976; Mayne, 1954), *route choice models* (Hoodgendoorn and Bovy, 2004; Helbing, 1997; Bovy and Stern, 1990; Hill, 1982; Borgers and Timmermans, 1986a; 1986b; Timmermans et al., 1992), *gas kinetic or fluid dynamics models*, (Henderson, 1971; 1974; Schadschneider, 2002), *configurational and graph-based models* (Hillier, 1996; Hillier et al., 1993, Raford and Ragland, 2004; Desyllas et al., 2003), *cellular automata and agent based models,* (Batty, 2003, 2001; Weifeng et al., 2003; Turner and Penn, 2002; Schadschneider et al., 2002; Kerridge et al., 2002; Blue and Adler, 2001, 19998; Kukla et al., 2001; Muramatsu et al., 1999; Gopal and Smith, 1990), and *direct estimation or sketch plan methods* (FHWA, 1999).

For ease of understanding, these models can be divided into five general categories of modeling. These are:

1. *Statistical physics models*, including particle dynamics, gas kinematics, and fluid flow approaches,
2. *Microsimulation models*, including agent based, artificial intelligence, and cellular automata approaches,
3. *Configurational models*, including space syntax and visibility graph analysis approaches,
4. *Sketch plan models,* including level of service approaches, aggregate demand estimation, and similar planning approaches,
5. *Origin – Destination / route choice models*; including discrete choice models, activity scheduling models, relative utility models, and stochastic models.

As with any categorization of a developing discipline, the lines between these approaches are not rigid. Several approaches draw from similar bodies of theory, such as microsimulation models that use discrete choice heuristics, for example. Nonetheless, these distinctions are useful guidelines which can help practitioners gain an understanding of the complicated and rapidly developing field of pedestrian modeling and every effort has been made to present a thorough and comprehensive overview of key papers and approaches.

Statistical Physics Models:

The statistical physics approach has been most recently summarized by Schadshneider (2002), one of the many European innovators in the field of pedestrian modeling. Schadschneider's approach is characteristic of the statistical physics school of pedestrian modeling, which draws heavily from the physical sciences for its inspiration and methods. The majority of works published in this school can be found in mathematical, physics, computer science, and statistical journals. They tend to share the goal of creating formalized mathematical models of individual pedestrians, crowd dynamics, and traffic flows.

The idea of traffic “flow” is central to statistical physics modeling. Schadshneider (2002) observes that there are several ways of distinguishing different physical science approaches to traffic modeling, creating a list which Teknomo (2002) elaborates upon. These approaches are:

1. Hydrodynamic models
2. Gas-kinetic models
3. Magnetic force models
4. Social force models

Hydrodynamic models were popularized by Henderson (1971) and his colleagues in the early 1970’s. This approach views traffic as a compressible fluid formed by the pedestrians or vehicles. The central variables in this approach are densities and flows, which are related through continuity equations that expresses the conservation location and momentum pedestrians and vehicles. Early approaches by Lighthill and Whitman (1955) assumed that flow was completely controlled by density, which allowed for modeling standing waves of congestion at bottlenecks and congestion. This can be visualized by imagining a smoothly flowing stream of people, which is suddenly interrupted at a given point. As the people at the front of the flow slow or stop, this causes a “ripple” effect backwards in the flow, resulting in a stably propagating wave of congestion. Real world traffic is much more complex, obviously, and later efforts attempted to deal with this by considering people as compressible fluids, which allowed for the modeling of unstable states.

Gas-kinetic models are an attempt to derive macroscopic behavior from microscopic equations (Prigonine and Herman, 1971). Traffic is considered as a gas of interacting particles with given vectors and positions. Pedestrians are seen as particles that bounce off each other and their surroundings, picking up momentum and direction. Congested situations result from many pedestrians in a compressed space with opposing vectors. Smooth flow occurs when pedestrian particles have similarly aligned vectors with little change. Gas kinetic models are useful in that they can usefully describe the emergence of stable states and dynamic phase changes.

Magnetic force models were originally developed in Japan by Okazaki (1979a, 1979b), with his colleagues Matsushita (1981, 1991), and Yamamoto (1981). This approach uses equations derived from Coulomb’s Law to calculate the relative attraction, repulsion, and motion of pedestrians and pedestrian environments. Pedestrians are represented with a positive pole, as are obstacles like walls, columns, and handrails. Destinations are weighted with negative poles and act as attractors to mobile pedestrians, who move to avoid collision and reach their destination. Like a modified gas-kinetic model, pedestrians rebound through space, avoiding each other and obstacles, seeking the path of least resistance to reach their goal.

All of these models produce interesting results, but ultimately prove to be less useful in predicting real world traffic flow than other approaches. Helbing et al. (2001) point out that realistic gas-kinetic or fluid-dynamic theory for pedestrians are inherently limited because they do not contain corrections due to their particular interactions (that is, avoidance and deceleration maneuvers) which of course do not conserve momentum and energy. The magnetic strength model shares the additional weakness of assuming arbitrary field strengths, which depending on their preliminary setting will result in arbitrary outcomes.

Helbing and Molnar (1995) attempt to ground statistical physics models in a more realistic framework by incorporating objectively measured variables into their approach. They do this by borrowing lessons from route choice theory, which will be elaborated on below. Instead of just physical forces operating on pedestrians, Helbing and Molnar (1995) suggest that social forces act strongly on individual pedestrians as well. They assume that individuals will always seek to reach a given destination as comfortably and efficiently as possible. Using this rule and operating like Newtonian particles, pedestrians in motion will assume a desired direction and a desired speed that will not change unless disturbed. If other barriers such as pedestrians or walls are present, pedestrians will be influence by a *repulsive effect.* Positive goals such as the destination will have an *attractive effect.* All of these effects are measured using physical science based approaches and due to the non-linear equations at their heart, tend to create the kind of dynamic, self-organizing patterns that can be seen in pedestrian crowds.

All such statistical physics approaches are able to produce realistic looking models of small scale pedestrian interaction. They are particularly successful at modeling pedestrian behavior at high densities and have been successfully used to model such situations in confined environments such as corridors and bottlenecks (Helbing and Molnar, 1995; Lovas, 1994, Timms, 1992; Timms and Cavalho, 1991), places free of automobile traffic such as subway and metro stations (Annesley et al., 1989; Daly et al., 1991; Harris, 1991), and for bridges and pedestrian walkways such as those used by pilgrims to Mecca (Selim and Al-Rubeh, 1991). As a result of these successes, statistical physics models are often at the root of popular evacuation modeling software packages and efforts (Helbing, 2003; Takimoto, 2003; Kirchner and Schadschneider, 2002; Chen et al. 2002; Thompson and Marchant, 1995a, 1995b; Watts, 1987).

The usefulness of such models is limited when applied to more complex, open-ended scenarios such as urban environments. They tend to represent pedestrians as continuous flows traveling in single directions, rather than as individuals with multiple goals and objectives. They are not able to account for individual variations in origins and destinations easily, nor can they represent more realistic pedestrian behavior such as trip-linking or decision-making heuristics such as utility maximization (Kurose et al., 2001). As a result, they have received little application or use in larger scale planning efforts that require pedestrian demand modeling on a city-wide or regional scale.

Microsimulation Models:

Microsimulation models came into widespread use in the 1990's when powerful computer processors become more widely available and less expensive (Kerridge et al., 2001). The availability of additional computing power gave researchers the opportunity to explore more advanced formulations of traditional statistical physics approaches to pedestrian dynamics, resulting in new and expanded methods for pedestrian behavior modeling.

Helbing et al. (2001) point out that although physics based models can provide interesting pictures of large scale crowd dynamics, the direct simulation of individual pedestrian motion is far more favorable for urban applications. Microsimulation offers a practical solution to this problem through the creation of thousands of individual, simulated pedestrians following predefined rules of behavior. By releasing these simulated people in simulated environments, researchers are able to analyze their behavior and create large scale predictions of travel demand and volumes at a given point.

Microsimulation takes several, related forms. The two most common are:

1) Cellular automata models (CA), and
2) Agent based models

Cellular automata (CA) and agent based approaches share a similar heritage. Both divide space into a uniform grid of cells. Cellular automata models assign values to each cell and have rules which govern the state of cells relative to the states of their neighbors. In this way it is possible to simulate the travel of individual pedestrians through the cell space by shifting the value of pedestrian presence from one cell to another. Hayes (1999) defines an agent as a "unit of computer code that is *autonomous* and *goal-directed."* Many agent-based approaches operated in cellular automata spaces, making the distinction between the two approaches blurry (Batty, 2003; Hoogendoorn, 2003; Schadschneider et al., 2002; Kerridge et al., 2001; Turner and Penn, 2002, Batty and Jiang, 1999). The key aspect about both CA and agent-based microsimulation approaches is that they both model individual actors operating independently, trying to achieve their goals in relation to their environment and the behavior of other agents. One effect of such a microscopic approach is the emergence of large scale complex phenomena such as traffic jams and congestion.

There are several different types of CA pedestrian models that will be explored in this review. Early approaches such as those outlined by Batty (2003) illustrate how a *random walk* approach could be combined with basic rules of attraction and repulsion for interesting results. A slightly more sophisticated approach is based upon the principle of *chemotaxis*, the phenomenon by which ants leave chemical trails for others to follow, resulting in coherent pathway formation and trail following behavior (Schadschneider et al., 2002; Helbing et al., 1997). Using a chemotactic approach, Schadschneider et al. (2002) were able to reproduce the type of self-organizing phenomenon often seen in high density pedestrian situations. This includes *jamming,* such as in bottleneck situations where many people simultaneously seek to enter or exit through a small door or when large volumes of people are walking in opposite directions, *lane formation*, such as when two groups of people spontaneously form continuous streams of traffic flowing in opposite directions; *oscillations*, such as when alternating spurts of people pass in opposite directions through a bottleneck; and *panic scenarios*, such as the “freezing-by-heating” effect caused when large numbers of people attempt to move in different directions at the same time, resulting in congestion and lack of movement. A more sophisticated chemotaxis approach that utilized attractors and realistically distributed crowd generators in a simulated urban environment was used by Batty et al. (2003) to simulate crowd dynamics in the Notting Hill Street Carnival in London, with good success.

Turner and Penn (2002) have successfully used agents with vision to replicate crowd movement in complex architectural medium-scale urban environments. Their approach, called the Extrasomatic Visual Agents System (EVAS), pre-computes visibility relationships amongst and urban space, then releases agents into the space that are told to walk towards areas with the greatest surface area. This straightforward approach is based on the early work of Benedikt (1979) and Gibson (1979), whose pioneering investigations into human visual processes and spatial cognition have found application in fields such as artificial intelligence, computer vision, pattern recognition, and software design (Cornwell et al., 2003; Borensten and Trettvik, 2001). The theoretical premise of the EVAS system is that humans use visual queues derived from their surroundings to help them understand and navigate space. This work also draws heavily from the theory behind configurational modeling, which will be explored in more detail below (Hillier, 1984, 1996). Conroy-Dalton (2001, 2003) provided evidence for vision-based approached when she examined how real humans and simulated agents explored different virtual environments. It was found that routes were chosen that minimized angular deflection in the direction of travel, and that the relationship between immediately visible spaces and global spatial structure played a strong role in way finding tests. Building on this connection between visibility, spatial cognition, and navigation, Turner (2003) demonstrated that sighted agents following these rules were able to replicate human movement in the City of London with up to with a statistically significant r-squared value of 0.67. Agents were also used to simulate the experience of art-viewers in the Tate Britain Gallery (Turner and Penn, 2002), and in enclosed retail shopping environments. These experiments make EVAS agents some of the most “perceptive” agents in the field of pedestrian modeling. Future work aims to incorporate specific destination and attraction relationships with agents, encoded as “taste” vectors, to test EVAS agents in more lifelike situations (Penn, 2003).

A third agent model, known as STREETS (Haklay et al., 2001), bears mention. This model, developed using the SWARM programming language pioneered at the Santa Fe Institute, is among the most complicated and nuanced models available. The STREETS approach uses a five step solver module to calculate the activity of each agent, based on their immediate visibility, medium term movement, and long term goals. The interaction between these modules allows agents to navigate the micro-environment of the virtual world (i.e., avoid walls, other pedestrians, etc.), the meso-environment (such as the length of the street and the geometry of the built environment), and the macro-environment (including their relative location to desired goals). The system also allows for the classification of different types of agents with different walking behavior, goals, and goal seeking strategies. Although STREETS has yet to be publicly tested in a real world scenario and its validity has yet to be tested, the system's flexibility and detail suggests it has significant potential.

Configurational Models:

Configurational models are those which emphasize the role that various aspects of the built environment have on influencing pedestrian movement dynamics. The most well researched and widely utilized application of configurational approaches is the space syntax approach, which originated at the University College London in the United Kingdom in the early 1980's (Hillier and Hanson, 1984; Hillier, 1996). Over 300 articles and four books have been published using space syntax, as well as a variety of reviews which focus on space syntax. The method has also been used successfully in a variety of applied planning and transportation studies in the United Kingdom and abroad (SSL, 2004).

The space syntax approach is based on measuring objective patterns of spatial relationships and linking these to patterns of movement within urban environments (Hillier et al., 1993). This is done through topological, graph-based analysis of pedestrian networks, which represents accessible spaces as nodes in a graph and processes their relationships using standard graph theory measures (Teklenberg et al., 1993). Early empirical studies found that these measures correlated well with observed pedestrian movement rates in many European cities, leading to further investigations and theoretical developments (Hillier et al., 1993; Penn and Dalton, 1994; Peponis, 1989; Hossain and Penn, 1999; Major et al., 1999, Desyllas and Duxburry, 2001).

Although not microsimulational in nature, space syntax studies regularly yield predictive accuracies of 60% to 80% with relatively little data requirements (Hillier et al., 1993; Penn and Dalton, 1994; Peponis, 1989; Hossain and Penn, 1999; Major et al., 1999, Desyllas and Duxburry, 2001). In circumstances where spatial configuration alone is less predictive, multivariate statistical regression models have been used to account for variations in pedestrian flow. Desyllas et al. (2003) utilized distance to transit, retail frontage, and sidewalk widths in central London to adjust their syntactic model using step-wise regression, and achieved an r-squared value of 0.74 when compared with observed counts. Stonor et al. (2002) combined distance to transit, land use composition, crossing design, and signal phase information in a multivariate regression model of south London. Raford and Ragland (2003) incorporated residential and housing densities derived from US Census data into their configurational model, yielding city-wide pedestrian volume predictions with an r-squared of 0.72 when compared to observed pedestrian traffic in the city of Oakland, California. This pedestrian volume model was then compared to pedestrian – vehicle crash data to create a pedestrian risk index for the city's first pedestrian Master Plan.
The key measure of space syntax research is *integration*, which measures the accessibility of a point in space relative to all others. Integration values are strongly influenced by the shape and accessibility relationships of urban spaces, which the space syntax literature refers to as spatial *configuration* (Hillier and Hanson, 1984; Hillier, 1996). The term "configurational analysis" thus derives from this approaches' emphasis on topological relational structures.

Penn (2003) suggests that the consistent correlation between configurational spatial properties and pedestrian movement could be explained by looking at the underlying mechanisms by which people perceive, understand, and then navigate their surroundings. Space syntax theory draws from neuro-scientific and cognitive science literature, which suggests that humans translate visual data of their immediate location into topological maps of the larger structure of their environment (Kitchin and Blades, 2002; Golledge, 1999). Configurational knowledge of spatial systems, as represented through topological relationships, has thus been proposed as the primary mechanism for linking human movement with topological representations of space (Hillier, 2003a, 2003b; O'Keefe, 1993; Smyth and Kennedy, 1982; Garling et al., 1983; Magliano et al., 1995).

Raford (2003) provides a discussion of the relative strengths and weaknesses of the space syntax approach through his application of the method in Oakland, California. He emphasizes that the simplicity and ease of availability of the necessary input data, combined with the relatively simple model construction and reasonably accurate results, makes the space syntax approach a useful model for urban planners and policy makers. He also points out several weaknesses of the approach, mainly its current de-emphasis on nuanced land use and attractor variables. Space syntax researchers argue that configurational analysis takes these variables into consideration, however, by considering the relationship of land use to spatial accessibility vis a vis the "multiplier effect". The multiplier effect asserts that different types of land uses are attracted to movement rich or movement sparse locations over time, which then attract additional movement and additional land uses in a multiplicative cycle (Hillier, 1996b).

Space syntax is unique among models discussed in that it is the only model which has given rise to a small industry of academic theory and scholarship, addressing issues of urban growth, sociology, anthropology, philosophy, linguistics, computer science, and artificial intelligence.

Sketch Plan Models:

Sketch plan methods comprise a group of approaches which are not simulation based in the way that statistical physics, cellular automata, and configurational approaches are, but instead attempt to approximate pedestrian demand based on simple planning guidelines and limited mathematics to produce "rules of thumb" (FHWA, 1999).

Many pedestrian sketch plan methods are generalized methods which attempt to predict pedestrian volumes through the use of pedestrian counts and regression analysis as a function of adjacent land uses (such as the number square feet of office or retail space) and/or indicators of transportation trip generation (parking capacity, transit volumes, traffic movements, etc.). Pushkarev and Zupan (1971) and Behnam and Patel (1977) were among the first researchers to attempt to forecast pedestrian volumes in central business districts using observed counts and land use measures, in Manhattan and Milwaukee, respectively. Demographic data on surrounding populations has also been combined with estimated trip generation and mode split rates to estimate levels of pedestrian traffic. Ercolano et al. (1997) used peak house vehicular counts and assumed mode split information to predict the peak pedestrian per hour in traffic analysis zones, then used land use data to distribute these trips to other zones.

The benefit of pedestrian sketch plans is that they require minimal data collection and no training in mathematical simulation or computer modeling. They are able to offer quick estimations of pedestrian volume, but are only effective at the aggregate level, if at all. Sketch plan models are not able to assign realistic pedestrian volumes to specific streets or intersections, they lack a model for dealing with congestion and traffic flow issues, and cannot account for the kind of dynamic goal-oriented behavior seen in CA and agent based systems. Sketch plans are often been applied in larger, regional and multi-zone urban environment plans where estimates of pedestrian volumes are desirable, but high accuracy or specific detail is not required for these reasons (1,000 Friends of Oregon, 1992 - 1997; Cambridge Systematics, 1994; Rossi et al. 1994; Chesapeake Bay Foundation et al., 1996).

Origin – Destination / Route Choice Models:

The final category of pedestrian modeling approaches are origin-destination / route choice models. These models resemble traditional vehicular travel demand models in many aspects (Hoogendoorn and Bovy, 2004; Ben-Akiva, 1985; McNally, 2000a, 2000b).

Models utilizing this approach are based on evidence that pedestrians use *utility maximization* heuristics when decided when to originate pedestrian trips and which routes and destinations to choose. Hill (1982) found that directness was an important factor in choosing pedestrian routes, with pedestrian frequently choosing the shortest perceived routes between given points (Senevarante and Morall, 1986). Additional factors such as habit, number of crossing and perceived pleasantness have also been found to be important to pedestrian route decision making (Bovy and Stern, 1990). McFadden (2001) found choice behavior to be highly dependant on individual psychological factors. Given a variety of route options, Hamacher and Tjandra (2001) hypothesized that pedestrians apply *subjective rational choices* during evacuation scenarios based on a variety of weighted variables such as those explored by Bovy and Stern (1990).

The basis of utility maximization and discrete choice theories is that all pedestrian actions are performed for a reason and therefore have utility relative to some goal set. Different types of utility maximization consider and route choice consider a finite number of routes (Gipps, 1986; Hamacher and Tjandra, 2001) or infinite routes (Hoogendoorn and Bovy, 2004), depending on their mathematical complexity.

Although technically distinct from other forms of microsimulation, many origin – destination / route choice models are utilized as underlying rule sets for more complex microsimulations. Hoogendoorn and Bovy (2004) utilize a microsimulation with an underlying route choice model composed of three levels:

1. *Strategic level,* which deals with departure time choice and activity pattern choice
2. *Tactical level,* which deals with activity scheduling, activity area choice, and route-choice to reach activity areas
3. *Operational level,* which deals with specific walking behaviors

Although their work addresses all three levels, it is the *tactical* level in which utility maximizing route choice heuristics have the most use. Given a predetermined schedule, their model supposes that pedestrians make decisions to minimize the cost of arrival at each scheduled destination by weighing the necessary activity pattern, different transit times utilizing alternate routes, and the velocity of each alternative. After a route has been decided, a microsimulation is used to release simulated pedestrians in a simulated environment with predetermined origins and destinations. Hoogendoorn and Bovy (2004) tested their heuristics on a small plaza in central Amsterdam, with very detailed and accurate results.

Kurose et al. (2001) introduce a route choice model which expands upon the concept of utility maximization by identifying three new types of cost-minimizing rule sets for pedestrian shopping behavior. These are local-distance-minimizing (LDM), total-distance-minimizing (TDM), and global-distance-minimizing (GDM) heuristics. The LDM choice heuristic states that a pedestrian would take the shortest route between successive stores on a trip whenever possible. In contrast, the TDM heuristic assumes that that pedestrians attempt to minimize the total distance of their overall route. Finally, GDM heuristics make the assumption that pedestrians do not always minimize the total distance traveled, but instead choose a global route choice decision based on a modified version of the optimal store choice route, but there are local deviations from a truly optimal route (Kurose et al., 2001). They go on to construct a complex decision making matrix applying these concepts, which they test against empirical data gathered in a Dutch central market. Within this enclosed environment, Kurose et al.'s (2001) model accounted for nearly 85% of total trips and route choices.

Each of these approaches use explicit rule based models based on utility maximization to assign pedestrian trips to available routes. In contrast to these approaches, which could be called deterministic due to their emphasis on pre-defined rules that leave little room for deviation, stochastic route assignment has been proposed (Urbitran Associates, 2004 for modeling pedestrian route choice dynamics. Stochastic assignment is used heavily in traditional vehicular travel models, and semi-randomly assigns vehicular trips to available routes based on a predetermined origin – destination trip matrix (Cantarella and Cascetta, 1998). Stochastic models are useful for dealing with conditions of uncertainty, as is often the case when generating trip distributions from population mode choice data or when estimating future demographic trends (Zhao and Kockelman, 2001).

A commonly used approach, the Stochastic User Equilibrium (SUE) was first introduced by Daganzo and Sheffi (1977), as an attempt to relax the unrealistic assumption of earlier traffic models that users had perfect knowledge of travel conditions and costs. Daganzo and Sheffi (1977) their predecessors introduced randomization into travel model variables as an effort to deal with errors in travelers' perception of travel costs (Hazelton, 1998).

Several pedestrian modeling approaches have used stochastic choice in their modeling assumptions. In the same study s the one discussed above, Hoogendoorn and Bovy (2004) introduced a stochastic differential equation for the directional vector of pedestrians in their microsimulation model of a pedestrian behavior in a Dutch shopping plaza. Modesti and Sciomachen (1998) analyze pedestrian trips as one mode in a multi-model trip model of the Italian city of Genoa. Their model seeks the optimal shortest distance pathway between origin and destination points, using any mode available. Each mode incurs varying costs and delays, resulting in a multi-dimensional matrix of route alternatives. Stochasticity is introduced into the route choice equations as per traditional vehicular transport modeling. The model produced a variety of travel times between abstracted traffic analysis zones (TAZ's), factoring in social class, costs of tickets, gasoline, parking, etc. The model was not validated against empirical data, however, so it is unclear as to how successful this approach was.

Discussion

This literature review has demonstrated five major approaches to pedestrian modeling. Each approach has strengths and weaknesses and each were designed to serve more or less specific purposes. Many are academic or theoretical in nature, but many have received some professional application, even if only experimentally.

Statistical physics models have a strong mathematical underpinning and have demonstrated effectiveness in predicting crowd behaviors in high density, confined areas with known origins and destinations. Academic models have been applied to airports, train stations, and queuing scenarios, while professional practitioners have begun to use statistical physics based simulations to model evacuation dynamics in complex buildings and offices. These models loose utility when applied to more open ended environments where pedestrians possess a variety of goals, as is the case in retail shopping environments or urban navigation models.

Cellular automata and agent based models arguably offer the most useful and potentially accurate approaches to pedestrian modeling, especially when combined with advanced route choice heuristics such as those explored above. CA and agent models are able to incorporate lessons learned from statistical physics models with regards to inter-pedestrian interaction and put them in the context of multiple goal-based environments, with different types of agents operating independently to pursue their own goals. Models such as the SWARM, STREETS, and EVAS offer the potential for multi-scale pedestrian analysis and may be able to shed light on unpredictable and emergent dynamic phenomena such as traffic jams and congestion.

The strength of CA and agent based models, their complexity and level of detail, may also be their most significant limitation. Most CA models require a high level of mathematical understanding and computer science knowledge to construct and operate. They also require a large amount of detailed data on environmental conditions and can require significant amounts of effort to prepare and calibrate. Finally, despite continuing increases in processing power, many CA and agent based models of the size necessary for practical planning purposes are still very computationally intense. It is likely that these factors - a steep learning curve requiring advanced training and knowledge combined with data and time intensive set up and operation - are the reason why so few of these models have been put to practical or professional planning use.

Despite (or perhaps because of) the relative simplicity of configurational approaches like space syntax, this modeling approach has been the most widespread and successfully applied urban pedestrian model of those reviewed. Literally hundreds of academic and professional applications of space syntax methodology have been conducted around the world to date. This is the case despite the fact that it remains relatively unknown in the American planning and research world. Owing to the fact that it is significantly less complex than CA and agent based approaches, it is also less computationally expensive, allowing for quick and easy modeling of a variety of urban scenarios and outcomes. Its lack of complex data requirements and easily grasped (if not fully understood) principles suggest that it may be one of the most useful models reviewed for planners and municipalities seeking to generate reasonably accurate models of pedestrian volume in their cities. To be fully accepted, however, it is likely that the space syntax approach will have to develop more consistent and sophisticated methods of incorporating traditional planning variables into its model, including population densities, level of service rankings, transit accessibility, and the effect large scale attractors a few examples.

Sketch plans remain well used in American planning circles. They are simple and easy to execute and provide quick answers to important questions. They do not provide very accurate or detailed results, however, and it is likely that they will be replaced over time with more valuable methods as they develop.

Like CA and agent based models, origin – destination / route choice models offer complex descriptions of the built environment and pedestrian behaviors within it. But these models suffer from the same weaknesses as CA and agent based approach do; namely that they require a large volume of data, a high level of specialized technical expertise and take time to set up, calibrate, and revise. These approaches have a long history of success in vehicular travel modeling however, and offer a significant amount of tailoring opportunities for highly detailed output. Land use variable such as jobs – housing balance are easily incorporated, as are the location of trip generators such as transit hubs and parking garages. Specific level of service issues such as sidewalk facilities, crossing signals, street width, and traffic volume can also be dealt with within the model, using either empirically observed behaviors for calibration or published findings from past studies. Any number of variables using route choice modeling and CA or agent based combinations can be created given enough data, time, money, and expertise. Given these conditions, it is possible that a combined approach using route choice heuristics may be the most likely type of models to produce nuanced and useful results for planning purposes.

Conclusion

The range of models and approaches reviewed in this literature review illustrate the wide variety of backgrounds and applications which pedestrian modeling has come into existence to serve. Clearly major issues remain which need to be addressed before the next generation of pedestrian models are to be created and deployed. It is likely that these models will begin to converge in their theory and applications as these models continue to develop in sophistication. Many CA models already use complex route choice heuristics, for example, and it is possible to image a joint space syntax / agent based model that would benefit from the simple effectiveness of the space syntax approach while becoming more sensitive to the types of land use and demographic variables included in most CA and route choice traffic models.

It is clear is that pedestrian modeling as a field is developing past the initial stages of research and development and is finding practical applications in industries around the world. Although no single solutions exist, practitioners are nearing the point where they will be able to select from a wide variety of modeling tools to suite any given problem.

In the future, hybrid models such as have been developed under this work are likely to develop with increased flexibility and power, thus allow for a more rapid and uniform approach to pedestrian modeling. As this occurs, the planning, engineering, and architecture professions will likely see increased benefits from pedestrian modeling and demand may grow for its application to a wide range of issues and challenges. If the modeling process becomes more available and less expensive, then it is possible that the true value of pedestrian simulations as a powerful decision support system and scenario planning tool for urban decision making will be realized.

Citations

1,000 Friends of Oregon. 1993, Making the Land Use Transportation Air Quality Connection: Volume 4A, The Pedestrian Environment. Portland, OR. Available at http://www.teleport.com/~friends/Lutraq2/Docs.htm

Annesley T, Dix M, Beswick A, Buchanan P, 1989, ``Development and application of pedestrian assignment models in London railway station studies'' Traffic Engineering and Control 30 345 - 352

Ashford N, O'Leary M, McGinity P D, 1976, ``Stochastic modelling of passenger and baggage flows through an airport terminal'' Traffic Engineering and Control 17 207 - 210

Bates J, 2000, "History of Demand Modelling" in Handbook of transport modeling, ed. Hensher D and Button K, Pergamon, New York

Batty M and Jiang B, 1999, "Multi-Agent Simulation: New Approaches to Exploring Space-Time Dynamic Within GIS", Working Paper 10, Centre for Advanced Spatial Analysis, University College London

Batty M, 2001, ``Agent-based pedestrian modeling'' Environment and Planning B: Planning and Design 28 321 - 326

Batty M, 2003, "Agent-Based Pedestrian Modelling", Working Paper 61, Centre for Advanced Spatial Analysis, University College London

Behnam J and Bharat P, 1977, "A Method for Estimating Pedestrian Volume in a Central Business District", Pedestrian Controls, Bicycle Facilities, Driver Research, and System Safety, Transportation Research Record 629, Washington, DC

Behnam, Jahanbakhsh and Bharat G. Patel, A Method for Estimating Pedestrian Volume in a Central Business District, Pedestrian Controls, Bicycle Facilities, Driver Research, and System Safety, Transportation Research Record 629, Washington, DC, 1977.

Ben-Akiva M, 1985, *Discrete choice analysis: theory and application to travel demand, MIT* Press, Cambridge, Massachusetts

Benedikt M, 1979, "To take hold of space: isovists and isovist fields", Environment and Planning B, Volume 6, pp. 47 – 65

Blue V and Adler J, 1998, "Emergent Fundamental Pedestrian Flows from Cellular Automata Microsimulation", Transportation Research Record 1644, pp 29-36

Blue V and Adler J, 2001, "Cellular automata microsimulation for modeling bi-directional pedestrian walkways", Transportation Research Part B, 35, p. 293 – 312

Borgers A and Timmermans H, 1986a, ``A model of pedestrian route choice and demand for retailfacilities within inner-city shopping areas'' Geographical Analysis 18 115 - 128

Borgers A and Timmermans H, 1986b, ``City centre entry points, store location patterns and pedestrian route choice behaviour: a microlevel simulation model'' Socio-Economic Planning Science 20, 25 - 31

Bovy P and Stern E, 1990, *Route Choice: Wayfinding in Transport Networks*, Kluwer Academic Publishers, Dordrecht

Cambridge Systematics, Inc., 1994, "Short-Term Travel Model Improvements, Travel Model Improvement Program", U.S. Department of Transportation; DOT-T-95-05, pp. 2-1 to 2-7

Cantarella G and Cascetta E, 1998, "Stochastic assignment to transportation networks: models and algorithms", in Marcotte P. and Nguyen S, eds., Equilibrium and Advanced Transportation Modelling, Kluwer Academic Publishers, Boston, pp. 87-108

Carstens R and Ring S , 1970, ``Pedestrian capacities of shelter entrances'' Traffic Engineering 41 38 - 43

Chen T, Song W, Fan W, Lu S, 2002, "Jamming transition of pedestrian evacuating flow in crossing exit", Process in Safety Science and Technology Part A, v 3, p 507-512

Chesapeake Bay Foundation, Environmental Defense Fund, et al.., 1996*, A Network of Livable Communities: Evaluating Travel Behavior Effects of Alternative Transportation and Community Designs for the National Capital Region*, Washington, DC

Conroy – Dalton R, 2003, "The Secret is to Follow Your Nose: Route Path Selection and Angularity", Environment and Behavior. January 2003 Special Issue on Spatial Cognition (forthcoming), available at http://undertow.arch.gatech.edu/pages/3sss/papers_pdf/47_conroy.pdf

Conroy-Dalton R, 2001, "Spatial navigation in immersive virtual environments", doctoral dissertation, University College London. Available at: http://undertow.arch.gatech.edu/pages/rdalton/thesis.htm

Cornwell J, O'Brien K, Silverman B, Toth J, 2003, "Affordance Theory for Improving the Rapid Generation, Composability, and Reusability of Synthetic Agents and Objects", Technical Report, Philadelphia, PA, Univ. of Penn/ACASA Technical Report. 2003. Available at http://www.seas.upenn.edu/ ~barryg/BRIMS2003_AffordanceTheory.doc

Daganzo C and Sheffi Y, 1977, "On stochastic models of travel assignment", Transportation Science, volume 11, pp. 253 – 274

Daly P N, McGrath F, Annesley T J, 1991, ``Pedestrian speed/flow relationships for underground stations'' Traffic Engineering and Control, 32, pp. 75 - 78

Davis D G, Braaksma J P, 1988, ``Adjusting for luggage-laden pedestrians in airport terminals'' Transportation Research A, volume 22, pp. 375 - 388

Desyllas J, Duxbury E, Ward J, Smith A, 2003, "Pedestrian Demand Modelling of Large Cities: An Applied Example from London", Working Paper 62, Centre for Advanced Spatial Analysis, University College London

Dial R, 1971, "A probabilistic Multipath Traffic Assignment Algorithm which obviates Path Enumeration", Transportation Research, number 5(2), pp. 81-111

Ercolano J, Olson J, Spring D, 1997, Sketch-Plan Method for Estimating Pedestrian Traffic for Central Business Districts and Suburban Growth Corridors, Transportation Research Record 1578, Washington, DC

Federal Highway Administration, 1999, “Guidebook on Methods to Estimate Non-Motorized Travel: Overview of Methods”, United States Department of Transportation, Publication No. FHWA-RD-98-165, Virginia.

Franklin S and Grassner A, 1997, “Is It an Agent, or Just a Program?: A Taxonomy for Autonomous Agents”, in Intelligent Agents III: Agent Theories, Architectures, and Languages, edited by Muller J, Woolridge M, and Jennings N, Springer – Verlag, Berlin, p. 21 – 35

Fruin J, 1971, ``Designing for pedestrians: a level-of-service concept'', in Highway Research Record Number 355: Pedestrians (Highway Research Board, Washington, DC) pp 1 - 15

Gibson J, 1979, The Ecological Approach to Visual Perception, Houghton Mifflin Company, New York

Gipps P, 1986, Simulation of Pedestrian Traffic in Buildings, Schriftenreihe des Instituts fuer Verkeswesen 35, University of Karlsruhe.

Golledge R, ed., 1999, Wayfinding behavior: Cognitive mapping and other spatial processes, Johns Hopkins University Press, Baltimore, Maryland

Gopal S, Smith T R, 1990, ``NAVIGATOR: an AI-based model of human way-finding in an urban environment'', in Spatial Choices and Processes Eds MMFischer, P Nijkamp,YYPapageorgiou (North-Holland, Amsterdam) pp 169 - 200

Haklay M, O’Sullivan D, Thurstain-Goodwin M, Schelhorn T, 2001, “``So go downtown'': simulating pedestrian movement in town centres”, Environment and Planning B: Planning and Design, volume 28, pp. 343 - 359

Hamacher H and Tjandra S, 2001, “Mathematical modeling of evacuation problems: a state of the art”, in Pedestrian and Evacuation Dynamics, Springer, Berlin, pp. 59 – 74

Hankin B and Wright R, 1958, ``Passenger flow in subways'' Operational Research Quarterly 9, pp. 81 - 88

Harris N G, 1991, ``Modelling walk link congestion and the prioritisation of congestion relief '' Traffic Engineering and Control, volume 32(2), pp. 78 - 80

Hayes C, 1999, ``Agents in a nutshell – a very brief introduction'', IEEE Transactions on Knowledge and Data Engineering, volume 11(1), pp. 127 - 132

Hazelton M, 1998, “Some Remarks on Stochastic User Equilibrium”, Transportation Research B, volume 32, number 2, pp. 101 – 108

Heggie I, ed., 1976, *Modal choice and the value of travel time,* Clarendon Press, Oxford

Helbing D, 1997, *Traffic Dynamics: New Physical Modeling Concepts*, Springer – Verlag, Berlin

Helbing D, Isobe M, Nagatani T, Takimoto K, 2003, “Lattice gas simulation of experimentally studied evacuation dynamics”, Physical Review E (Statistical, Nonlinear, and Soft Matter Physics), volume 67, number 6, pp. 67101-1-4

Helbing D, Molnar P, Farkas I J, Bolay K, 2001, “Self-organizing pedestrian movement”, Environment and Planning B: Planning and Design 2001, volume 28, pp. 361 - 383

Helbing D, Schweitzer F, Keltsch K, Molnar P, 1997, “Active walker model for the formation of human and animal trail systems”, Physics Review E, volume 56, pp. 2527 – 2539

Henderson L F, 1971, ``The statistics of crowd fluids'', Nature, volume 229, pp. 381 - 383

Henderson L F, 1974, ``On the fluid mechanics of human crowd motion'' Transportation Research 8 509 - 515

Hensher D and Button K, eds., 2000, *Handbook of transport modeling*, Pergamon, New York

Hill M, 1982, “Spatial Structure and Decision-Making of Pedestrian Route Selection Through an Urban Environment”, Ph.D. Thesis, University Microfilms International

Hillier B and Hanson J, 1984, *The Social Logic of Space,* Cambridge University Press, Cambridge, England

Hillier B, 1996, *Space is the Machine*, Cambridge University Press, Cambridge, England

Hillier B, 1996, “Cities as Movement Economies”, Urban Design International, volume 1 No 1, pp. 49 - 60

Hillier B, 1997, “Cities as movement economies” in ed Droege P, Intelligent Environments, Spatial Aspect of the Information Revolution, North Holland, p. 742

Hillier B, 2003a, “The human city beneath the social city”, Proceedings, 4th International Space Syntax Symposium, London, England

Hillier B, 2003b, “The architecture of seeing and going”, Proceedings, 4th International Space Syntax Symposium, London, England

Hillier B, Penn A, Hanson J, Grajewski T, Xu J, 1993, “Natural movement: or configuration and attraction in urban pedestrian movement”, Environment & Planning B: Planning & Design, volume 19, 29-66

Hoel L, 1968, ``Pedestrian travel rates in central business districts'' Traffic Engineering 38 10 - 13

Hoissan N and Penn A, 1999, “A Syntactic Approach to the Analysis of Spatial Patterns in Spontaneous Retail Development in Dhaka”, in Proceedings of the 2nd International Space Syntax Symposium, Brasilia, Brazil, March, 1999

Hoodgendoorn S and Bovy P, 2004, “Pedestrian route-choice and activity scheduling theory and models”, Transportation Research Part B, volume 38, p. 169 – 190

Hoogendoorn S, 2003, “Pedestrian Travel Behavior Modeling”, Proceedings of the 10th International Conference on Travel Behavior Research, Lucerne, 2003.

Jiang B, 1999, “Modelling Urban Environments with Open Spaces”, in Proceedings of the 2nd International Space Syntax Symposium, Brasilia, Brazil, March, 1999

Kerridge J, Hine J, Wigan M, 2001, “Agent-based modelling of pedestrian movements: the questions that need to be asked and answered”, Environment and Planning B: Planning and Design 2001, volume 28, pp. 327 - 341

Kerridge J, Hine J, Wigen M, 2001, ``Agent-based modelling of pedestrian movements: the questions that need to be asked and answered'' Environment and Planning B: Planning and Design 28 327 - 341

Kirchner A and Schadschneider A, 2002, “Simulation of evacuation processes using a bionics-inspired cellular automaton model for pedestrian dynamics”, Physica A: Statistical Mechanics and its Applications, v 312, n 1-2, p 260-276

Kitchin R and Blades M, 2002, *The Cognition of Geographic Space*, I.B, Tauris, London.

Kukla R, Kerridge J, Willis A, Hine J, 2001, “PEDFLOW: Development of an autonomous agent model of pedestrian flow”, Transportation Research Record 1774, pp 11 – 17

Kurose S, Borgers A, Timmermans H, 2001, “Classifying pedestrian shopping behaviour according to implied heuristic choice rules”, Environment and Planning B: Planning and Design 2001, volume 28, pp. 405 - 418

Lighthill M and Whitman G, 1955, Proceedings of the Royal Society of Physics, A, Volume 229, p. 281

Lovas G, 1994, ``Modelling and simulation of pedestrian traffic flow'' Transportation Research B 28 429 - 443

Matsushita S and Okazaki S, 1991, “A Study of a Simulation Model for Way Finding Behavior by Experiments in Mazes.” Journal of Architecture, Planning, and Environment Engineering, AIJ, No. 429, p. 51-59

Mayne A J, 1954, ``Some further results in the theory of pedestrians and road traffic'' Biometrika, volume 41, pp. 375 - 389

McFadden D, 1978, “Modeling the choice of residential location”, in Karlquit A, et al., eds., Spatial Interaction Theory and Residential Location, North Holland, Amsterdam, pp. 75 – 96

McNally M, 2000a, “The Four-step Model” in Hensher and Button, eds., Handbook of Transport Modelling, Pergamon, New York, New York

McNally M, 2000b, “The Activity Based Approach” in Hensher and Button, eds., Handbook of Transport Modelling, Pergamon, New York, New York

Modesti P, Sciomachen A, 1999, “A utility measure for finding multiobjective shortest paths in urban multimodal transportation networks” European Journal of Operational Research, 111, p. 495 – 508

Moıri M, Tsukaguchi H, 1987, ``A new method for evaluation of level of service in pedestrian facilities'' Transportation Research A 21 223 - 234

Navin P and Wheeler R, 1969, ``Pedestrian flow characteristics'' Traffic Engineering 39 31 - 36

Newell G, 1980, *Traffic flow on transportation networks,* MIT Press, Cambirdge, Massachusetts

O’Keefe J, 1993, “Kant and the sea-horse; an essay in the neurophilosophy of space”, in Eilan N, McCarthy R, and Brewer B, eds., Spatial Representation, Blackwell Publishers, Cambirdge, England

O'Flaherty C and Parkinson M, 1972, ``Movement on a city centre footway'' Traffic Engineering and Control 13 434 - 438

Okazaki S and Matsushita S, 1981, "A Study of Pedestrian Movement in Architectural Space Part 5: A Proubing walk and a guide walk by a guideboard.", Journal of Architecture, Planning, and Environment Engineering, AIJ, No. 302, p. 87 – 93

Okazaki S and Yamamoto C, 1981, "A Study of Pedestrian Movement in Architectural Space Part 4: Pedestrian Movement Represented in Perspective", Journal of Architecture, Planning, and Environment Engineering, AIJ, No. 299, p. 105 – 113

Okazaki S, 1979a, "A Study of Pedestrian Movement in Architectural Space Part 2: Concentrated Pedestrian Movement", Trans. Of A.I.J., No. 284, p. 101 – 110

Okazaki S, 1979b, "A Study of Pedestrian Movement in Architectural Space Part 3: Along the Shortest Path, Taking Fire, Congestion, and Unrecognized Space into Account", Journal of Architecture, Planning, and Environment Engineering, AIJ, No. 285, p. 137 – 147

Older S J, 1968, ``Movement of pedestrians on footways in shopping streets'' Traffic Engineering and Control 10 160 - 163

Ortúzar J, 1994, *Modelling transport,* Wiley, New York

Pauls J, 1984, ``The movement of people in buildings and design solutions for means of egress'' Fire Technology 20 27 - 47

Penn A, 2003, "Space Syntax and Spatial Cognition", Environment and Behavior B, volume 35, number 1, pp. 30 – 34

Penn A, 2003, presentation on advanced agent research, Bartlett School of Architecture, University College London

Penn A, Dalton N, 1994, ``The architecture of society: stochastic simulation of urban movement'', in Simulating Societies: The Computer Simulation of Social Phenomena Eds N Gilbert, J Doran (UCL Press, London) pp 85 - 125

Peponis J, 1989, "The spatial core of urban culture", Ekistics, Jan.-Apr., v.56, no.334-335, p.43-55

Peponis J, 1998, "On the generation of linear representations of spatial configuration", Environment and planning B, volume 25, n.4, p. 559-576

Polus A, Schofer J L, Ushpiz A, 1983, ``Pedestrian flowand level of service'' Journal of Transportation Engineering 109 46 - 56

Prigonine I and Herman R, 1971, *Kinetic Theory of Vehicular Traffic*, American Elsevier, New York.

Pushkarev, Boris and Jeffrey M. Zupan, 1971, "Pedestrian Travel Demand", Highway Research Record 355, Washington, D.C..

Raford N and Ragland D, 2004, "Space Syntax: An innovative tool for pedestrian volume modeling", forthcoming Transportation Research Record Paper No. 04-2977

Raford N, 2003, “Looking Both Ways: Space syntax for pedestrian exposure forecasting and collision risk analysis”, Proceedings, 4th International Space Syntax Symposium, London, England

Rossi T, Lawton T, Kim K. 1993, “Revision of Travel Demand Models to Enable Analysis of Atypical Land Use Patterns”, Cambridge Systematics, Inc. and Metropolitan Service District.

Sawhney A, 2003, “Spatial Cognition and Configurational Understanding”, Final Built Environment Report, Msc. Built Environment, Advanced Architectural Studies, University College London, London

Schadschneider A, 2002, “Traffic flow: a statistical physics point of view”, Physica A, 313, p. 153 – 187

Schadschneider A, Kirchner A, Nishinari K, 2002, “CA Approach to Collective Phenomena in Pedestrian Dynamics”, Proceedings of ACRI 2002, LNCS 2493, p 239 – 248

Selim S Z, Al-Rubeh A H, 1991, ``The modelling of pedestrian flow on the Jamarat Bridge'' Transportation Science 24 257 - 263

Senevarante P and Morall J, 1986, “Analysis of factors affecting the choice of route of pedestrians”, Transportation Planning and Technology, volume 10, p. 147 – 159

Space Syntax Limited, 2004, “Projects: Full Project List”, accessed via the world wide web on Friday, April 16th, 2004, http://www.spacesyntax.com/portfolio/Portfolio_Projects.htm

Stonor T, Campos M, Smith A, 2002, “Towards a Walkability Index”, presented at Walk 21: 3rd International Conference, Steps Towards Liveable Cities, San Sebastian, Spain

Takimoto K, 2003, “Spatio-temporal distribution of escape time in evacuation process”, Physica A, v 320, 15, p 611-21

Teklenburg, J A F, Timmermans H J P, and van Wagenberg, A, 1993, “Space Synax: Standardised Integration Measures and Some Simulations” Environment and Planning B 20(3) 347–357

Teknomo K, 2002, “Microscopic Pedestrian Flow Characteristics: Development of an Image Processing Data Collection and Simulation Model”, Ph.D. Dissertation, Department of Human Social Information Sciences, Graduate School of Information Sciences, Tohoku University, Japan.

Thompson P and Marchant E, 1995a, “A Computer Model Evaulation of Large Building Populations”, Fire Safety Journal, No. 24, p. 131 – 148

Thompson P and Marchant E, 1995b, “Testing and Application of the Computer Model ‘SIMULEX’”, Fire Safety Journal, No. 24, p. 149 – 166

Timms P, 1992, ``Putting pedestrians into network planning models'', paper presented at the 6th World Conference on Transport Research, Lyon.

Timms P, Cavalho S, 1991, ``Inclusion of pedestrians and cyclists in network planning models'', in Proceedings of the DRIVE Conference Advanced Telematics in Road Transport, Amsterdam

TRB, 1985, ``Pedestrians'', in Highway Capacity Manual special report 209 (Transportation Research Board, Washington, DC) chapter 13

Turner A and Penn A, 2002, “Encoding natural movement as an agent-based system: an investigation into human pedestrian behaviour in the built environment”, Environment and Planning B: Planning and Design 2002, volume 29, pp. 473 – 490

Turner A, 2003, “Analysing the visual dynamics of spatial morphology”, Environment and Planning B: Planning and Design, volume 30, pp. 657 - 676

Watts J, 1987, “Computer Models for Evacuation Analysis”, Fire Safety Journal, No. 12, p. 237 – 245

Weifeng F, Yang L, Fan W, 2003, “Simulation of bi-directional pedestrian movement using a cellular automata model”, Physica A, 321, p. 633 – 640

Yuhaski S J Jr, Macgregor Smith J M, 1989, ``Modelling circulation systems in buildings using state dependent queueing models'' Queueing Systems 4 319 - 338

Zhao Y and Kockleman M, 2001, “The Propagation Of Uncertainty Through Travel Demand Models: An Exploratory Analysis”, in Annals of Regional Science, available at http://www.ce.utexas.edu/prof/kockelman/public_html/ ARS01ErrorPropagation.pdf

TECHNICAL DESCRIPTION

Overview

The objective of this study is to develop and demonstrate a method for estimating pedestrian flows, and then to compare those flows to pedestrian accidents to identify high-priority locations.

This procedure is based on stochastic path-finding methods. It builds upon prior research and applications that Urbitran has explored in New Jersey, New York City, and Pennsylvania. While not an actual application of regional travel forecasting and traditional travel demand modeling methods, the proposed approach does use tools that are familiar to travel modelers. The method includes a land use allocation method, pedestrian travel generator, a distribution module that allocates pedestrian movements to their destination, and a stochastic assignment method that allocates walking pedestrians to a mix of paths that reflect pedestrian congestion, street-crossing barriers, and other factors. The software platform is a hybrid of ArcView GIS, CitiLabs' TP+ and VIPER travel modeling software, and custom software for special functions.

This method is designed to adapt to available street network, land use, and demographic data to simplify development of the model, while allowing opportunities to provide specific fine-grain data wherever it might be available. The model consists of the following key components:

Pedestrian Network Synthesis - Beginning with a street network database such as a Census TIGER file, GIS and specialized software are used to add descriptive data to the street network, then to convert the street file to sidewalks.

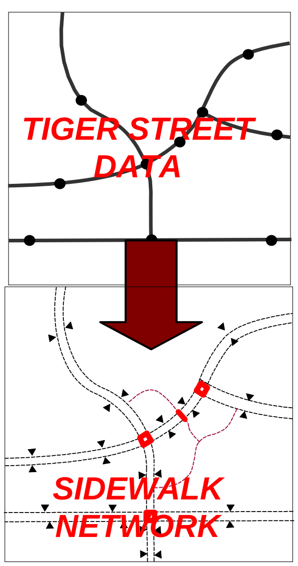


Figure 1
Street Network Conversion to Pedestrian Network

Street data includes such information as the functional class, number of lanes, cross sectional elements (median, shoulder, parking), traffic control devices (signals and stop signs), speed limits, and traffic volumes. This information is used by later modules to compute the barrier effects posed by the street system.

The pedestrian network is synthesized through a generalized method that creates sidewalks along all streets subject to the functional class (freeways have no sidewalks, for example). Then based on the control device – signal, stop sign – a crosswalk submodel creates crossings with appropriate characteristics, both at intersections and at mid-block locations.

Additional attributes of the pedestrian network are estimated, with the ability to override wherever observed data is available. Sidewalk quality can be estimated or input, providing a basis for affecting path finding. Load points for abutting land uses are added to each block face and tied to the sidewalk network.

This pedestrian network is created in a format that can be loaded into the VIPER network editor, which provides excellent tools for editing and displaying networks. Within VIPER the network can be further refined, adding sidewalks and pathways, changing default characteristics to better reflect actual conditions, etc. This tool is also useful for updating the pedestrian network to reflect the addition of facilities in the future.

Land Activity Synthesis – In Maryland data describing land uses and activity generators is fortuitously available for individual properties via the PropertyView system. This data was obtained for each of the case study areas, providing very detailed information as to specific parcel-by-parcel land uses. On a wider area basis, residential population and characteristics are available by Census block group, and that data was also obtained and used for the case studies.

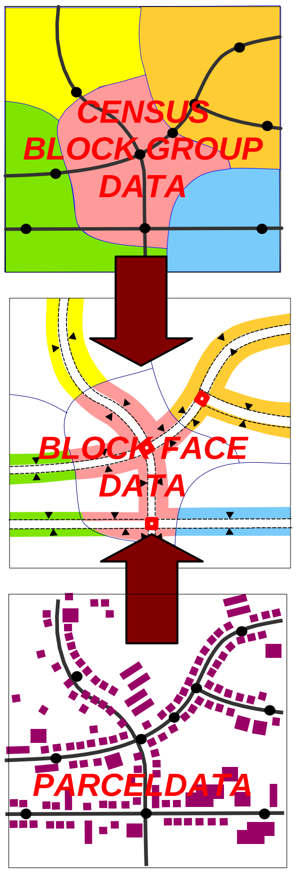


Figure 2
Land Use Data Conversion to Block Face Activity

Pedestrian activity generally occurs at a block-face level, reflecting the trip-making potential of the abutting land development. Therefore the model contains a land activity synthesizer geo-codes the parcel data to block faces, and allocate census characteristics such as income, and household size to the block face land uses.

Special transport facilities such as parking garages and transit stations that attract pedestrians can also be added to the database. Travel characteristics such as percentages of walk-to-transit and on-site vs. off-site parking can be specified either at the area level or for individual block faces.

Each element of block face activity is attached to the pedestrian network at load points, as illustrated in the image.

Pedestrian Travel Generator – The amount of pedestrian activity generated at each block face load point and its destinations is calculated. Analogous to a linked trip generation and distribution model, this pedestrian travel generator accounts for additional factors that encourage or stimulate walking, such as residential density, retail opportunities and the land use mix, street connectivity, and the adequacy of pedestrian facilities. Within a network-based modeling framework these and other important factors are readily computed using standard accessibility measures computed from network connectivity. (For example, one accessibility measure is the amount of retail reachable from an origin within a designated walk distance.) The number of pedestrian trips generated is then a function of the land activity and of accessibility to surrounding activities. Activity is stratified into an appropriate number of purposes, such as commutation, transport access (parking, transit stations), retail access, recreation, etc.

The computed pedestrian activity is distributed from each origin – a block face load point – to its destinations using a conventional distribution model adapted to the pedestrian context. In addition to distance and time, non-traditional factors such as sidewalk quality and availability, and street crossing times and exposure influence the destination choice. The result of this computation is a matrix of pedestrian movements from each block face load point to all other load points in the network. The pedestrian movements are stratified by time of day and purpose.

Pedestrian Path Allocation – This module assigns the pedestrian movements to appropriate paths through the network. When the movements from all origins to all destinations are accumulated, the result is the total pedestrian volume on each sidewalk and crosswalk segment.

The process uses conventional network assignment software adapted to the pedestrian context. Impedances are computed for each sidewalk and crosswalk segment on the basis of time, distance, quality and exposure, and then a stochastic assignment method allocates each origin-to-destination pedestrian movement to a set of paths. Some trip purposes such as commutation are highly sensitive to time and convenience and so closely follow the minimum-time path through the network. Other purposes are more discretionary and so they follow a variety of paths from origin to destination. Use of a stochastic assignment method allows these factors to be accounted for.

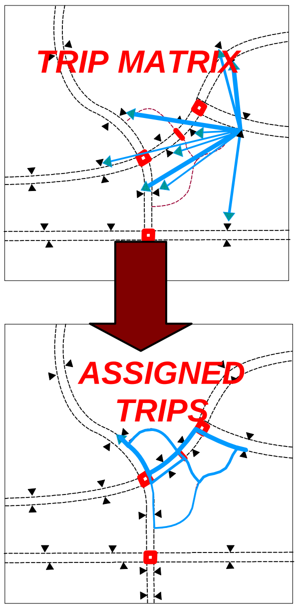


Figure 3
Stochastic Path Finding Through the Sidewalk Network

Total pedestrian volumes are accumulated and reported on each sidewalk segment and at crosswalks and intersections. This aggregate data is then reported back to the GIS for displays and to support spatial analysis of accidents and safety exposure.

Data Dependence and Calibration -All travel models are heavily dependent on descriptive data. This approach is designed to make maximum use of available data such as the Census and regional travel models for activity data (population, housing, employment, etc.) and TIGER or travel model network databases to derive the pedestrian network. The model provides tools to link these more generalized data resources to the fine-grain block face and sidewalk descriptions needed to support the model, and to the property parcel-level data that comprise the land use file. Clearly the derivation of detailed descriptions from more coarse data will be generalized, and refinement of land use and sidewalk data to reflect actual conditions improves the quality of the estimates produced by the model. This need for local detail needs to be balanced with the amount of effort needed to obtain it.

The model development requires the following general data items for the selected study area(s), as they may be available:

- Census TIGER files
- Census block group data (housing, population, characteristics)
- Specific building data (offices, retail, special generators, other) from MD PropertyView
- Recreational facilities
- Parking inventories (facilities, number of spaces, utilization)
- Transit stop / station locations and activity
- Traffic signal locations and characteristics
- Street physical inventory (functional class, width, cross section, traffic volume)
- Aerial ortho-photos

All data should be in ArcView GIS format, properly conflated to permit display and spatial analysis across all sources.

It is assumed that travel surveys and localized pedestrian travel data availability is typically minimal, although any such data should be obtained and can be used. Lacking local data, the trip generation / distribution model and other model components will generally rely upon data and characteristics obtained from the literature and comparable metropolitan areas.

The model has been calibrated to match observed local conditions. The most important such data were pedestrian counts at key intersection locations. Other observations such as sidewalk congestion, photos, and other qualitative information were used as well, as available.

The Pedestrian Travel Demand Model

The following documents the pedestrian demand model that was developed for this study. It describes the calibration process and results for the trip generation and trip distribution procedures.

Trip Generation

Typically the first step in a travel demand model is to estimate the number of trip ends that occur in each zone. For home-based trips, trip ends that are based at a residence are called *productions* and trip ends that are based elsewhere are called *attractions*. For non-home-based trips, the terms *origins* and *destinations* are used, with the normal convention being that origins = destinations (NHB trips into a zone = NHB trips out of a zone).

Purposes

The first step is to define the trip purposes that will be modeled. Because this model was going to be based on relatively detailed land use data, a decision was made to define the purposes fairly narrowly. Also, there was a feeling that certain purposes are more likely to be pedestrian trips than others. As noted above, the distinction between home-based and non-home-based purposes is critical. Thus, the following trip purposes were used:

- home-based work (HBWK): travel from home directly to work and back
- home-based eat meal (HBEM): travel from home to a restaurant and back
- home-based shop (HBSH): a trip to purchase goods
- home-based personal business (HBPB): a trip to obtain services or for a purpose not specified elsewhere
- home-based leisure (HBLS): a trip made in order to pursue recreational or social activities, or for which the trip itself is a leisure activity (e.g., jogging)
- home-based school (HBSC): travel from home directly to school and back
- non-home-based work (NHWK): travel between the workplace and some other destination (e.g., work-related business)
- non-home-based eat meal (NHEM): travel from someplace other than home (e.g., work) to a restaurant and back
- non-home-based shop (NHSH): a trip to purchase goods, not to/from home
- non-home-based personal business (NHPB): a trip to obtain services or for a purpose not specified elsewhere (e.g., lunchtime errands), not to/from home
- non-home-based leisure (NHLS): a trip made in order to pursue recreational or social activities, or for which the trip itself is a leisure activity (e.g., jogging), not to/from home
- non-home-based school (NHSC): travel between a non-home location and school

Many trips are made for more than one purpose, as part of a “tour” involving several stops. Modeling such tours is an extremely difficult process and is just now being undertaken by researchers and practitioners with million-dollar budgets and new home interview survey data. Given the nascent state of the art in tour modeling, we have decided to stick with the more conventional process of modeling individual trips. Splitting trips by these 12 purposes does a better job of identifying each component of a tour, however.

Survey

Most travel models are based on a home-interview survey of travel behavior. This typically gathers information on all trips made by the members of a household for a 24-hour period. In the past, these surveys have excluded pedestrian trips, since they have not historically been of much interest. Even though such trips are increasingly being included in these surveys, in most metropolitan areas there are not enough of them to use for model development. An exception is the New York metropolitan area, which performed an extensive survey of more than 11,000 households in 1996. Although New York is admittedly unique in many respects, this survey provides probably the richest database of pedestrian travel in the U.S. The survey covered 28 counties constituting the greater New York metropolitan area, including New Jersey and Connecticut. This area covers the full range of development density to be found in the U.S., from lower Manhattan to rural central New Jersey. With appropriate care, information from this survey can be transferred to other areas.

One particularly interesting observation was that the average pedestrian trip length from the NYMTC survey was 24.4 minutes. The 2002 *National Survey of Pedestrian & Bicyclist Attitudes and Behaviors* by USDOT/NHTSA found an average pedestrian trip length of 24.0 minutes. In addition, the two distributions of trips by travel time were very similar.

The principal survey database included 89,605 trip records representing 59.3 million daily trips. Of those, pedestrian trips were 12,274 records and over 9 million trips. These trip records were summarized by trip end (production vs. attraction), purpose, and traffic analysis zone (*traffic analysis zones*, or just *zones*, are neighborhood-sized geographic entities used to tabulate travel model data -- there are 3,586 such zones in the New York region).

Model Development

The trip survey database was augmented with other sources of information. One was the New York socioeconomic data file, which contained variables such as the number of persons, households (HH), and jobs, and average HH income by zone, for 1996. The other was data on *accessibility*.

Accessibility is a general concept in travel modeling that typically refers to the ability of people to reach various destinations. It measures both the degree of development activity and the travel time needed to get to those activities. It was theorized that accessibility is a primary influencing factor on the number of pedestrian trips that are made. Population and employment density are sometimes used to reflect the closeness of travel opportunities, but given the extremely small size of the traffic analysis zones to be used in this model (i.e., a single block face), it did not make a lot of sense to use density.

Accessibility is a zone-based measure and can be calculated from a matrix of zone-to-zone travel times and a vector of zonal “opportunities”. For the purposes of this study, a fairly conventional definition of accessibility was used:

$$Acc(i) = \sum [Opp(j) * F(i,j)] \qquad \text{(summed across all zones j)}$$

Where:
Acc(i) = accessibility of zone i
Opp(j) = opportunities in zone j – generally either employment or households
F(i,j) = an inverse function of travel time between zones i and j (as time increases, F becomes smaller); for this purpose, a gamma function is used:

$$F = t^{-1.5} * e^{-0.1t}$$

Where:

t = walk time between zones i and j, minutes (computed as the distance along the sidewalk at a speed of 3 mph)
e = base of natural logarithms (2.71828…)

This is similar to gamma functions used in other models and was examined carefully for mathematical and logical reasonableness. This accessibility function has the desired properties: as opportunities increase it goes up, but it goes up faster if the opportunities are nearby. Defined this way, "accessibility" really has no dimensions and no particular meaning as an absolute number. Its main usefulness is in describing the *relative* accessibility of one zone compared to another.

There can be many different definitions of accessibility, depending on which variable is used to define "opportunities". For example, accessibility to total employment might be related to the number of work trips while accessibility to retail employment might be related to the number of shopping trips. At the start, it was not known which definitions might be most useful, and so several versions of accessibility were calculated:

- total floor space (sq ft)
- non-residential floor space
- office floor space
- single-family dwelling units (SFDU)
- multi-family dwelling units (MFDU)
- total dwelling units
- total employment (jobs)
- total population
- retail employment

These different measures were then attached to the calibration file for each zone.

The original intent was to estimate a regression model of trip rates, at the traffic analysis zone level. However, the New York data proved to be too thin to accomplish that. Therefore the data was aggregated to a system of 123 districts. These were somewhat arbitrarily defined as aggregations of every 30 numerically contiguous zones, respecting county boundaries, and this proved to be a workable construct. Table 1 shows the total pedestrian trip rates by purpose.

Values shown are daily pedestrian trips per household or per 1,000 square feet of total floor space.

The next step was to create the dependent variable. Although trip generation models typically calculate both productions and attractions, usually the focus is on the production model because of the greater confidence placed in the allocation of HHs by zone (compared to employment). For any given study area, the sum of the productions must equal the sum of the attractions. Usually they don't and the most common practice is to assume that the sum of the productions is correct and that the attractions must be normalized so that their sum matches the production sum.

That same general logic was followed here. The primary focus was on calculating the trip productions per HH, for the HB purposes and trip productions per 1000 square feet of total floor space for the NHB purposes. All of the available zonal socioeconomic variables were tested to see how well they correlated with the pedestrian trip rate. The following independent variables were selected to go forward in the analysis:

- accessibility to MFDUs (ACCMFM)
- accessibility to total employment (ACCEMP)
- accessibility to retail employment (ACCRET)
- low income dummy (= 1 if the zonal average HH income < $41,000, else 0) (LOW)
- high income dummy (= 1 if the zonal average HH income >= $41,000, else 0) (HIGH)

Table 1
New York Trip Rates

	Productions per HH	Attractions per KSF
HB Work	0.062	0.031
HB Pers Bus	0.228	0.116
HB Eat Meal	0.148	0.075
HB Shop	0.144	0.073
HB Leisure	0.151	0.077
HB School	0.067	0.034
total	0.800	0.406

	Trips per KSF
NHB Work	0.034
NHB Pers Bus	0.037
NHB Eat Meal	0.046
NHB Shop	0.044
NHB Leisure	0.026
NHB School	0.0001
total	0.187

For each trip purpose, a model of the following type was estimated:

$$TR = ACCMFM^{A} * ACCEMP^{B} * ACCRET^{C} * (D * LOW + E * HIGH)$$

Where:

TR = trip rate (trips/HH for HB purposes, trips/KSF floor space for NHB purposes)
A, B, C, D, E = calibrated coefficients
(Note: for the NHB purposes, the “D” and “E” coefficients were set equal – there is no influence of income)

The models were calibrated using the method of least squares. For each district, the estimated trip rate was compared to the surveyed rate. The coefficients were adjusted so as to minimize the overall sum of the squared error. Table 2 shows the final model and the district-level coefficient of determination (r^2).

Walk accessibility to MFDUs consistently showed up as the most important variable, for almost all purposes. This makes sense, since apartments, condominiums, and townhouses represent the kind of high density development that is conducive to walking. Accessibility to employment was important for Work and Personal Business trips. Accessibility to retail employment was important for Eat Meal, Shop, and Leisure trips, which seems logical.

The influence of income was slight, but rational. Generally, lower income people tend to walk more. The low/high income breakpoint of $41,000 (HH income, 1990 $) represents approximately the lowest income 20% of HHs in the New York region.

Table 2
Trip Generation Production Model

Purpose	ACCMFM A	ACCEMP B	ACCRET C	Low Inc D	High Inc E	district r^2
Trip Rates per Household						
HB Work	0.0384	0.3655	0.0000	0.0148	0.0148	0.433
HB Pers Bus	0.2396	0.0223	0.0000	0.1578	0.1012	0.445
HB Eat	0.2039	0.0000	0.0212	0.1159	0.0740	0.312
HB Shop	0.3923	0.0000	0.0000	0.0735	0.0735	0.437
HB Leisure	0.2199	0.0000	0.0484	0.1097	0.1013	0.350
HB School	0.1430	0.0000	0.0000	0.0601	0.0347	0.201
Trip Rates per KSF of Total Floor Space						
NHB Work	0.0000	0.8050	0.0000	0.0004	0.0004	0.892
NHB Pers Bus	0.2363	0.3099	0.0000	0.0036	0.0036	0.643
NHB Eat	0.0000	0.0000	0.5948	0.0081	0.0081	0.741
NHB Shop	0.5315	0.0000	0.2370	0.0020	0.0020	0.620
NHB Leisure	0.2547	0.0000	0.2624	0.0055	0.0055	0.358
NHB School	0.3541	0.0000	0.0000	0.0076	0.0076	0.075

basic production equation = ACCMFM^A * ACCEMP^B * ACCRET^C * Inc Factor
Inc Factor = Low Inc (D) if avg HH inc < $41,000, else High Inc (E)
Low and High Inc Factors are set equal for NHB (no effect of income)

ACCMFM = walk accessibility to MFDU
ACCEMP = walk accessibility to total employment
ACCRET = walk accessibility to retail employment

As noted above, the model shown in Table 2 is used to estimate trip productions. These equations produce trip rates that average out to those shown in Table 1. The land use file must contain information on dwelling units (for this purpose, taken to be roughly equivalent to "households") and non-residential floor space for each traffic analysis zone. The HB trip rates per HH are multiplied by HHs. Total floor space is calculated as non-residential floor space + 1000 * MFDU + 2000 * SFDU. The NHB trip rates per KSF are multiplied by total floor space in thousands of square feet.

Trip attractions are those trip ends associated with the non-residential end of the trip. The attraction model was developed in a different fashion. This project's Baltimore land use file contains the number of square feet of floor space in each zone, for several different types of land uses, as shown in Table 3. It was desired to take advantage of this rich database, but since the New York socioeconomic database did not contain this level of detail, the New York data could not be used to develop the attraction model.

Table 3
Non-Residential Floor Space Categories

Variable	Description
HOTEL	Hotels, motels
AUTO_DLR	Auto dealers (new and used)
AUTO_SVCST	Auto service stations
AUTO_CONVN	Auto service stations with convenience stores
AUTO_OTHER	Other auto service establishments
REST_FAST	Fast-food restaurants
REST_OTHER	All other restaurants
STORE_DEPT	Department stores
STORE_OTHR	All other stores
OFFC_MED	Medical offices
OFFC_OTHER	All other offices
CARE_HOSP	Hospitals
CARE_DAYCR	Day care centers
CARE_OTHER	All other care-giving facilities (e.g., nursing homes)
BANK	Banks
WAREHOUSE	Warehousing
INDUSTRIAL	Industrial facilities
REC_PROPSF	Recreational property, general
REC_MOVIE	Movie theaters
REC_MUSEUM	Museums
REC_OTHER	All other recreational establishments
COM_POSTOF	Post offices
COM_CHURCH	Churches
COM_SCHOOL	Schools
COM_LIBR	Libraries
COM_OTHER	All other community facilities
SAFETY	Public safety facilities (e.g., fire station, police station)
PUB_MUNIC	Municipal public buildings
PUB_COUNTY	County public buildings
PUB_STATE	State public buildings
PUB_FED	Federal public buildings
UTILITIES	Public utility buildings

The basic premise of the attraction model was that most of the 12 trip purposes could be logically associated with some of the specific land uses listed above. Estimating the proper number of total attractions is not as important as estimating the productions, because the production estimate is what actually determines the total number of trips. The attraction estimates are used to allocate the non-residential trip ends to each zone. Therefore the attraction estimates should reflect the nature and intensity of the development in each zone.

The attraction model for each purpose is specified as a linear regression equation, based on the floor space by non-residential land use category, as shown in Table 4.

Table 4
Trip Attraction Equations

Purpose	Equation
HB Work	0.000094 * NONRES
HB Pers Bus	0.000349 * NONRES
HB Eat Meal	0.000226 * (REST_FAST + REST_OTHER)
HB Shop	0.000220 * (AUTO_DLR + STORE_DEPT + STORE_OTHR)
HB Leisure	0.000231 * (HOTEL + REC_PROPSF + REC_MOVIE + REC_MUSEUM + REC_OTHER)
HB School	0.000103 * COM_SCHOOL
NHB Work	0.000128 * NONRES
NHB Pers Bus	0.000164 * NONRES
NHB Eat Meal	0.007683 * (REST_FAST + REST_OTHER)
NHB Shop	0.001935 * (AUTO_DLR + STORE_DEPT + STORE_OTHR)
NHB Leisure	0.003854 * (HOTEL + REC_PROPSF + REC_MOVIE + REC_MUSEUM + REC_OTHER)
NHB School	0.000020 * COM_SCHOOL

Notes:
Variables are as defined in Table 3, plus the following:
NONRES = total non-residential floor space

In the final model application, the NHB origins are set equal to the average of the productions (Table 2) and attractions (Table 4) for each zone. The NHB destinations are set equal to the NHB origins.

Results

The Baltimore prototype study area includes the land use totals shown in Table 5. Application of the trip production and attraction models to these land use totals produced the weekday pedestrian trip totals shown in Table 6.

Table 5
Baltimore Study Area Land Use Totals

Variable	Total
POPULATION	90,840
EMPL_RETL	4,280
EMPL_NRETL	58,757
HH_TOTAL	42,471
HH_APT	4,569
HH_OTHER	37,902
HOTEL	490,386
AUTO_DLR	133,053
AUTO_SVCST	27,111
AUTO_CONVN	1,181
AUTO_OTHER	814,013
REST_FAST	85,448
REST_OTHER	1,027,283
STORE_DEPT	0
STORE_OTHR	4,975,871
OFFC_MED	625,169
OFFC_OTHER	6,320,600
CARE_HOSP	3,896,864
CARE_DAYCR	64,217
CARE_OTHER	192,151
BANK	268,640
WAREHOUSE	8,137,346
INDUSTRIAL	3,076,996
REC_PROPSF	419,314
REC_MOVIE	145,889
REC_MUSEUM	18,871
REC_OTHER	248,440
COM_POSTOF	259,449
COM_CHURCH	1,433,751
COM_SCHOOL	4,512,923
COM_LIBR	99,106
COM_OTHER	0
SAFETY	117,025
PUB_MUNIC	2,147,335
PUB_COUNTY	0
PUB_STATE	1,562,530
PUB_FED	0
UTILITIES	0

Note: values are shown in actual units for the first 6 variables; square feet of floor space for the other variables.

Table 6
Trip Totals by Purpose

Purpose	Trips
HB Work	2,944
HB Pers Bus	11,178
HB Eat Meal	7,046
HB Shop	7,573
HB Leisure	7,275
HB School	3,089
NHB Work	5,876
NHB Pers Bus	6,115
NHB Eat Meal	7,327
NHB Shop	7,125
NHB Leisure	4,246
NHB School	976
Total	70,770

Trip Distribution

The second step in most travel demand models is to allocate the trips between production zones and attraction zones, creating a matrix of zone-to-zone trip movements. This determines the pattern of trips.

The most common method of distributing trips is through a "gravity model". This model says that the number of trips between zone i and zone j is proportional to the number of trips produced in zone i, the number of trips attracted to zone j, and inversely proportional to the impedance separating the two zones:

$$T_{ij} = P_i * \frac{A_j F_{ij}}{\sum_j A_j F_{ij}}$$

Where:
T_{ij} = trips from zone i to zone j
P_i = trips produced in zone i
A_j = trips attracted to zone j
F_{ij} = impedance function, i to j

In Newton's original formulation of the gravity model, the impedance function is the square of the distance. This has been replaced with a gamma function, which is becoming the most commonly used equation for impedance functions:

$F = a * t^b * e^{gt}$

Where:
F = impedance
t = perceived walk time, minutes
a, b, g = calibrated coefficients
e = base of natural logarithms (2.71828…)

This is a minor variation of the same function that is used to compute the travel time function for accessibility, as described above. The travel time is taken as the perceived time, which is a combination of actual walking time and waiting time at crosswalks, weighted to reflect the fact that people don't like to wait.

Calibration of a gravity model mainly consists of these steps, performed separately for each trip purpose:

- hypothesize or borrow a starting set of gamma coefficients
- calculate the impedance function and apply the model
- from the resulting trip table, compute a trip time frequency distribution, showing the percentage of trips by each time increment and the average travel time
- compare this estimated distribution and average time to the observed data (usually obtained from a survey)
- adjust the gamma coefficients to improve the correspondence between the observed and estimated travel time distributions and average time
- repeat as necessary

The model is judged to be calibrated when the estimated and observed average travel times are within 1% of each other.

The model was calibrated using New York data. As noted above, this data exhibited a close correspondence to nationwide statistics, as shown in Table 7.

Table 7
Comparison of New York and National Walk Time Distributions

Time Increment	% of NY Trips	% of US Trips
< 5 min.	26.1%	26.9%
5 – 10 min.	21.2	19.6
10 – 20 min.	23.9	20.7
20 – 40 min.	18.4	18.0
> 40 min.	10.4	14.8

Thus, it was judged that it was reasonable to use the New York data to calibrate this model.

Table 8 shows the gamma coefficients for the impedance functions by purpose, along with the estimated and observed average walk time by purpose. Figure 4 displays the impedance function curves on a semi-log scale, with the purposes grouped according to similarity of their curves.

When this model is applied to the Baltimore dataset, the average trip times shown in Table 9 result. These seem reasonable.

Table 8
Distribution Model Coefficients and Results

Purpose	a	b	g	Obs. Avg. Tm.	Est. Avg. Tm.	% Error
HBWK	100,000	-0.2018	-0.0600	23.34	23.16	-0.8%
HBPB	100,000	-0.2018	-0.0600	22.57	22.70	0.6
HBEM	100,000	0.7259	-0.1476	18.77	18.81	0.2
HBSH	100,000	0.6249	-0.1417	18.19	18.04	-0.9
HBLS	100,000	2.5900	-0.2220	20.47	20.54	0.3
HBSC	100,000	2.5910	-0.2034	22.02	21.92	-0.5
NHWK	100,000	2.7580	-0.2225	19.92	19.89	-0.1
NHPB	100,000	2.7571	-0.2405	18.34	18.41	0.4
NHEM	100,000	2.7609	-0.1721	23.81	23.78	-0.1
NHSH	100,000	2.6936	-0.2659	16.05	16.10	0.3
NHLS	100,000	2.6947	-0.2492	18.09	18.05	-0.2
NHSC	100,000	2.7022	-0.1053	33.39	33.36	-0.1

Figure 4
Impedance Function Curves

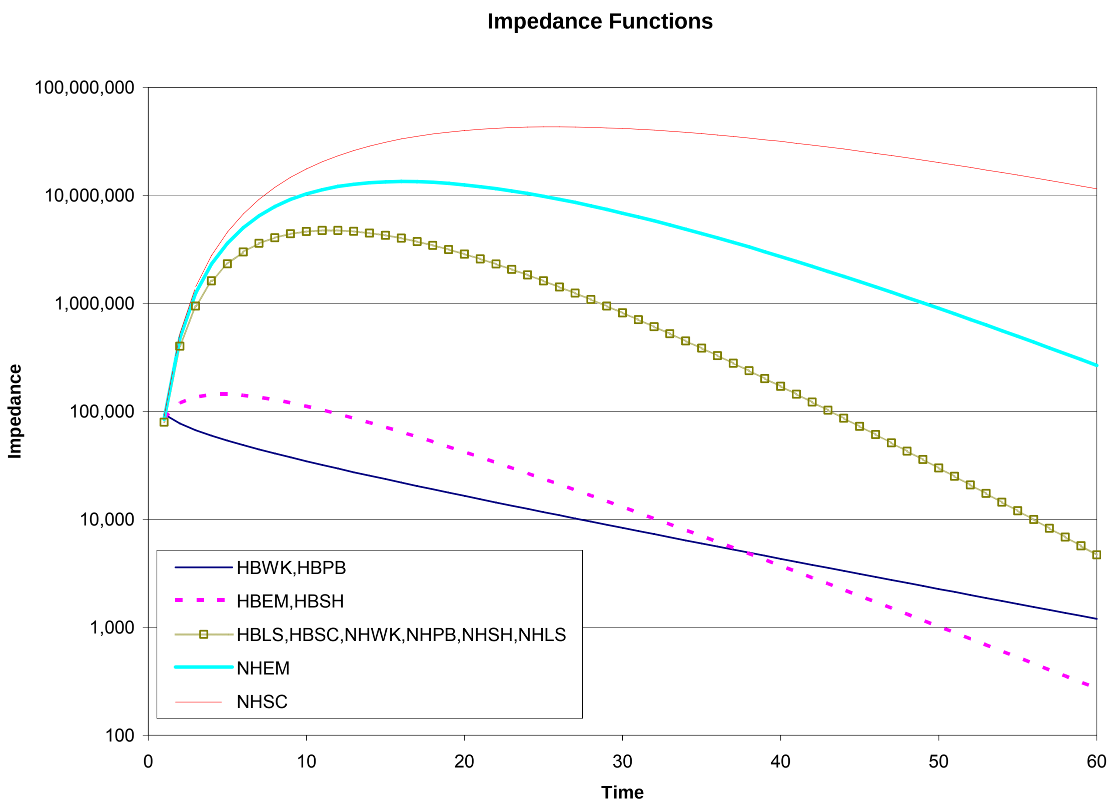


Table 9
Baltimore Average Trip Times

Purpose	Avg. Time (min.)
HB Work	23.17
HB Pers Bus	22.69
HB Eat Meal	18.77
HB Shop	18.05
HB Leisure	20.55
HB School	21.92
NHB Work	19.88
NHB Pers Bus	18.38
NHB Eat Meal	23.76
NHB Shop	16.11
NHB Leisure	18.03
NHB School	33.33
All Trips	21.39

Synthesis of the Pedestrian Network

The model is dependent on a pedestrian network representation that amply describes the path options available to all pedestrians in the study area. This network mainly consists of street sidewalks, but also includes intersection crosswalks, mid-block jay-walk locations, and even on-street walk paths where sidewalks are unavailable. Off-roadway pathways can be added to the network as well. This section describes the method by which this representation is developed.

Computational Framework

Due to its complexity the network building process spans several software platforms. It begins with manipulation of Census TIGER line files in ArcGIS, and then it uses several custom programs to perform the geometry, land use aggregation, and data manipulation calculations that are needed. Finally it builds a correctly formatted network file in Citilabs' TP+ to support the calculation of travel times and trip assignments. This complex process is bound together and run within the CENTRAL process controller, as is described in Section 4, the User Guide.

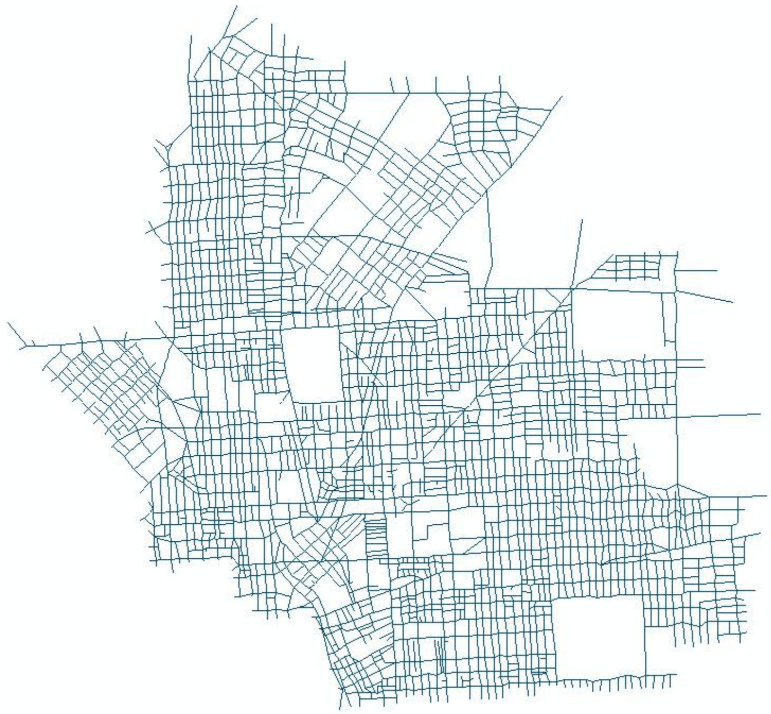
Pedestrian Network Topology and Data Development

The network synthesis relies on two primary data sources: First, an enhanced Census TIGER segment provides the basic street topology from which the pedestrian network is built. It is important that this enhanced file contain the Census Feature Code (CFCC) which provides the functional classification of the street. Other data items including the street functional class / facility type can be added as supplemental data.

Second, a parcel-specific land use database for the coverage area is obtained from the Maryland Property View system. Specific fields that are needed include property address, use of the structure (residential, office, etc.), legal descriptions such as acreage, and 2000 Census tract and block group numbers. These items are part of the standard Property View data record.

Once the data is in hand and an approximate study area boundary has been established, the specific cordon must be defined to include needed areas of concern, follow logical physical and neighborhood boundaries, and follow the rules outlined in Section 4 for selecting boundary segments. Within ArcGIS a process is used to tag and extract the selected TIGER line segments to a model coverage file. Care is needed at the boundary to cut individual segments that will create a logical external loading point. A typical extracted TIGER segment file is illustrated in Figure 5.

Figure 5
Typical Extracted TIGER Segment File



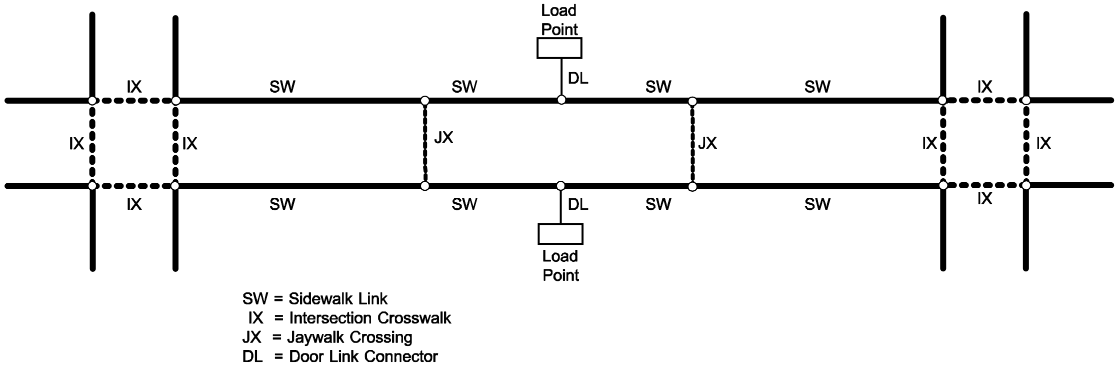
In addition to the data which comes to the street network from TIGER, supplemental data files can be developed in ArcGIS, using pre-defined fields to contain a variety of street and sidewalk data. The network can be overlaid on 1' resolution orthophotography to obtain much of this information without the need for substantial field inventory work, although field visits certainly are of value. External data that can be added includes:

- Street Facility Type to supplement the Census Feature Code
- Street width and cross section: Number of lanes, Medians and widths, Shoulders and parking
- Traffic signal locations and timing
- Sidewalk locations
- Traffic volumes from available sources, or estimated from functional class / facility type defaults

Once the data files have been extracted and populated, custom software converts each TIGER segment to a set of 14 sidewalk and crosswalk links. These are illustrated in Figure 6, and include conventional sidewalk link, crosswalk links at intersections, jaywalk links at mid-block, and doorway links that connect block-face load points to the sidewalk at the segment midpoint.

The software creates the basic geometry for the sidewalk system. In the above orthogonal example this is a fairly simple process, but real-world networks are far more complex with multi-leg intersections, non-orthogonal directionality, and TIGER file errors or anomalies. The software handles these peculiarities and produces correct geometry.

Figure 6
Sidewalk Link Topology



Starting with and retaining the TIGER segment and node numbering scheme, the software renumbers the network into a form that is assignable using TP+.

A typical portion from a pedestrian network built for Baltimore is illustrated in Figure 7. Note that no adjustments were made to the TIGER segment geometry, and yet the resulting fit with orthophotography is remarkably good.

Figure 7
Built Pedestrian Network Topology

Sidewalk and Cross-Walk Impedances

The accessibility calculations that underlie the previously described demand model, and the path finding that underlies the trip assignment model, all are based on travel times and impedances derived from the pedestrian network. Travel (walk) times are computed for each link in the network – sidewalks, intersection crosswalks and mid-block jay walks, doorways / load points, and other types. Then these times are weighted by a variety of factors to produce a set of impedances for each link that govern path-finding.

Basic sidewalk walk time is based on walking speed and distance. Average walk speed can be defaulted, or can be specified by the user. The default value for sidewalk walk speed is 3.5 mph.

Sidewalk quality factors are applied to modify the walk time to reflect perceived quality. For example, a high quality sidewalk would receive a quality factor of 1.0, whereas a poor-quality or non-existent sidewalk might receive a quality factor of 2.0. These factors can be set or overridden by the user. Default quality factors are as follows:

Table 10
Time Factors for Sidewalk Quality

Sidewalk Quality	Time Factor
High quality	1.0
Marginal quality	1.3
Poor quality	2.0
On-street walk	1.7
Other walkway types	1.0
Default Sidewalk Types for Street Facility Types:	
Freeway	None
Arterial	Marginal
Collector	High
Local	High
Alleyway	On-Street
Other	Marginal

At intersection crosswalks and mid-block jay walks, basic crosswalk times are based on walking speed (specified separately and typically faster than sidewalk walk speed), distance based on street width, and step-off conditions. Additional time is added to account for wait times for gaps in uninterrupted traffic (a function of the traffic volume), and wait times at signals (a function of signal timing and pedestrian phasing). Default crossing time parameters are shown in Table 11.

Table 11
Crosswalk Time Parameters

Parameter	Value
Crosswalk Walk Speed	4.5 mph
Reaction / Step-off Time	1.0 sec
Speed Risk Allowance	0.05 sec / mph
Crossing time factor if Pedestrian Phase at Signal	0.6
Crossing time factor if Pedestrian Actuation at Signal	0.8

Further adjustments are applied to increase walk time to account for crossing risk. Jay walks, for example, are riskier than intersection crossings. High traffic speeds are more risky than low speed streets. These risk factors and acceptable gap times are computed based on the facility type, speed, and volume. Default characteristics are shown in Table 12:

Table 12
Street Volume and Speed Defaults

Facility Type	Speed (mph)	Traffic Volume (Veh / hour / lane) Peak	Off-Peak
Freeway	60	1,200	850
Arterial	45	900	600
Collector	35	350	200
Local-1	25	150	80
Local -2	15	0	0
Local-3	15	0	0
Alleyway	15	0	0
Other	15	0	0

In the built pedestrian network the various components of travel time discussed above are preserved, and are combined to an overall peak and off-peak weighted time. Resulting weighted times for a typical section of the Baltimore network are shown below in Figure 8.

Figure 8
Weighted Network Walk Times

Assignment of Pedestrian Trips to the Network

As was described above, pedestrian trips from each block face to all other block faces are estimated by the pedestrian travel demand model. Paths are then found through the pedestrian network according to the above travel impedances, and the pedestrian trips are assigned to those paths. This section describes the method by which those paths are found, and the assignment method by which trips are allocated to those paths.

The assignment algorithm operates within the TP+ HWYLOAD software. It adapts standard network assignment methods to the needs of this specialized pedestrian trip assignment problem.

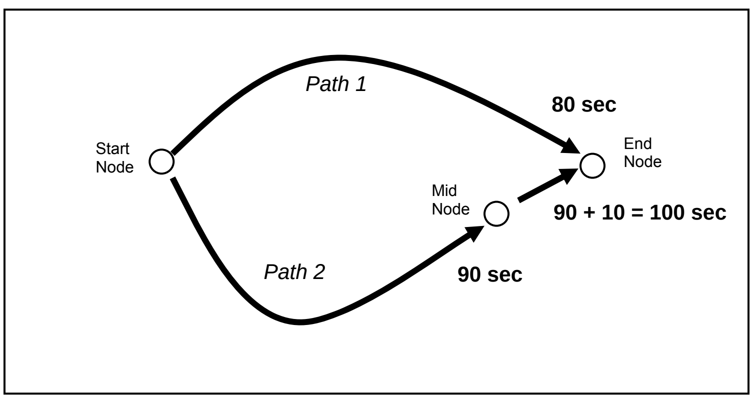
<u>The Stochastic Network Assignment Problem</u>

While moving from the same origin to the same destination, a group of pedestrians will use a variety of paths – some efficient with respect to time or impedance, some not so. To emulate this phenomenon the assignment method needs to find multiple paths from each origin to each destination and to proportionally load the trips along those paths.

Travel modeling packages such as TRANPLAN and MINUTP have provided a stochastic assignment algorithm devised by Robert Dial for this, but testing of the algorithm for this pedestrian problem revealed a fundamental flaw: The algorithm finds "efficient" paths by by comparing impedances at nodes and comparing impedances between the minimum time path and the candidate. To be considered an efficient path, the path must be getting further from the origin. Then a dispersion parameter Theta is used to calculate the proportion of trips using the two paths.

In the simplistic example illustrated in Figure 9, the minimum path is Path 1 and it can reach the End Node from the Start Node in 80 seconds. An alternate path being considered, Path 2, can reach the End Node in 100 seconds via the Mid Node.

Figure 9
Stochastic Algorithm: Finding an Efficient Path



However, the Path 2 time to the mid node is 90 seconds, which is already further from the origin than the End Node is by the minimum path, or 80 seconds. Therefore Path 2 is not considered an efficient path, and cannot be considered to receive a portion of the trips moving from Start to End.

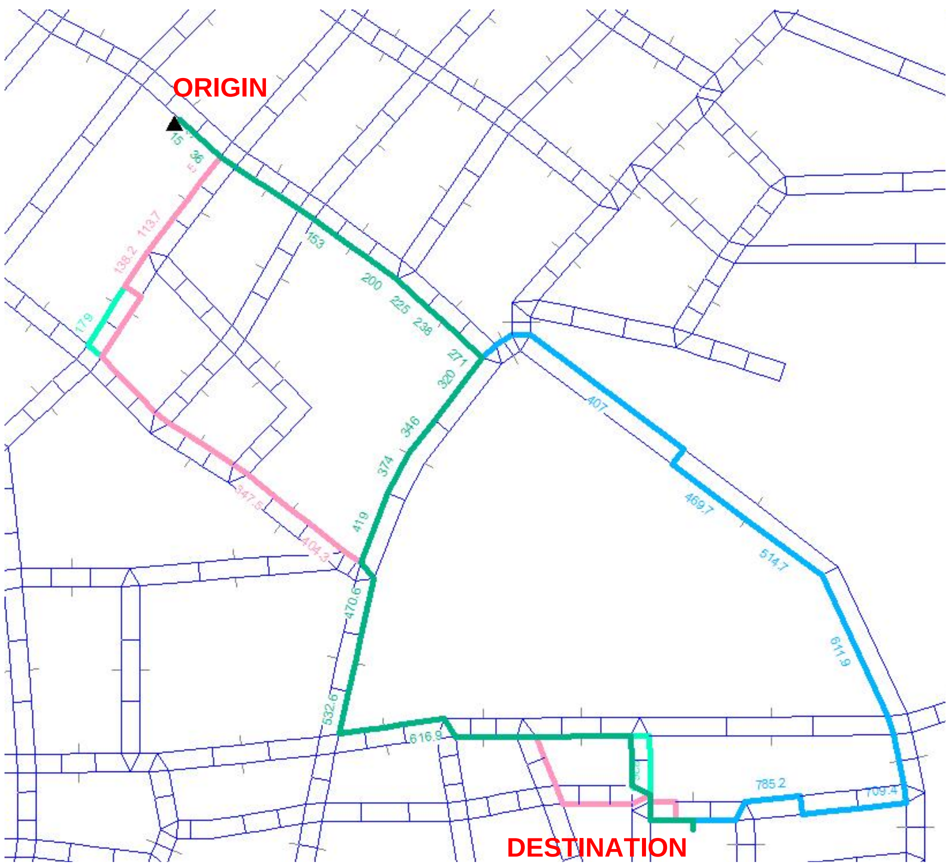
Normally in a typical highway assignment problem this definitional problem is not a significant issue, but the pedestrian network built by this model contains a multitude of short links that prevent alternative paths from being qualified and assigned. Therefore it was concluded that standard stochastic assignment methods cannot be used for this pedestrian model

The Pseudo-Stochastic Network Impedance Model

An alternative construct has recently been experimented with by Citilabs, William Allen, and Ohio DOT to deal with specialized traffic assignment issues such as truck multi-path assignment needs. This construct uses an iterative path-finding and assignment process, but randomly perturbates the link impedances before finding paths to emulate the random ways in which users perceive or react to actual impedances. After several iterations with these perturbated times, a family of paths is found to have been generated for each origin-to-destination movement that are a reasonable representation of multipath assignment.

A typical result is illustrated in Figure 10. Four paths were found from the origin in the top left of the diagram, to the destination in the bottom right. These four paths found some major variants (going around both sides of the large open block) as well as minor variants (using jay walks in stead of intersection crosswalks. It is precisely this performance that the assignment algorithm seeks to produce and estimate. (Refer to Figure 7 which shows the actual geography of the area on an ortho-photograph.)

Figure 10
Pseudo-Stochastic Paths



The implementation of this model in TP+ finds nine separate sets of perturbated paths for each origin-to-destination movement. These sets are developed as three random variants (A through C) of three levels of perturbation (1 through 3). Each trip purpose follows a perturbation level as shown in Table 13.

Table 13
Trip Purposes and Path Perturbation Levels

Trip Purpose	Perturbation Level	Perturbation Assignment Sets
HB Work	Minimum	1A,1B,1C
HB Personal Business	Medium	2A,2B,2C
HB Eat Meal	Maximum	3A,3B,3C
HB Shop	Maximum	3A,3B,3C
HB Leisure	Maximum	3A,3B,3C
HB School	Maximum	3A,3B,3C
NHB Work	Minimum	1A,1B,1C
NHB Personal Business	Medium	2A,2B,2C
NHB Eat Meal	Maximum	3A,3B,3C
NHB Shop	Maximum	3A,3B,3C
NHB Leisure	Maximum	3A,3B,3C
NHB School	Minimum	1A,1B,1C

A path set with minimum perturbation, used by such trip purposes as walking to work, is essentially the minimum path, and typically results in minor variations to jaywalk instead of using intersection crosswalks. A set with maximum perturbation, used by such trip purposes as leisure, will show a high level of variation and can typically result in going entirely around a block or finding another street to walk on.

The variations in travel impedances that comprise these perturbations are computed in one of two ways that can be selected by the user: Either the overall total impedance on a link can be perturbated, or the individual components of travel time (walk time, crossing time, crossing wait time, traffic speed penalties) can be perturbated. It appears that the individual component approach is more sensitive and delivers more appropriate paths, but further experimentation is needed in this regard.

The median values of each component, and of the total overall impedance, are computed using the defaults described above or user data if provided. Then for each of the nine impedance sets (1A through 3C in Table 13) the values are randomly varied, using a normal distribution with standard deviations that can be specified by the user. Default standard deviations built into the software, and suggested standard deviations that have been defined through practice, are as shown below in Table 14.

Table 14
Standard Deviations Used For Perturbation Levels

	Standard Deviation		
Travel Time Component	Minimum Perturbation (1A, 1B, 1C)	Medium Perturbation (2A, 2B, 2C)	Maximum Perturbation (3A, 3B, 3C0
Overall Impedance	0.1	0.2	0.3
Weighted Sidewalk Time	0.3	0.5	0.5
Sidewalk Quality	0.2	0.6	0.6
Street Crossing Time	0.4	0.8	0.8
Sidewalk Quality	0.2	0.5	0.8

<u>Pedestrian Network Assignment</u>

The matrix containing 24-hour pedestrian trips is assigned to the pedestrian network using the TP+ program HWYLOAD. One iteration of all-or-nothing assignment is used, with each trip purpose set assigned according to the three perturbated impedances comprising each set as shown in Table 14. Each set is then weighted with the following fractions. For any set (minimum, medium, or maximum), the fractions sum to 1.00.

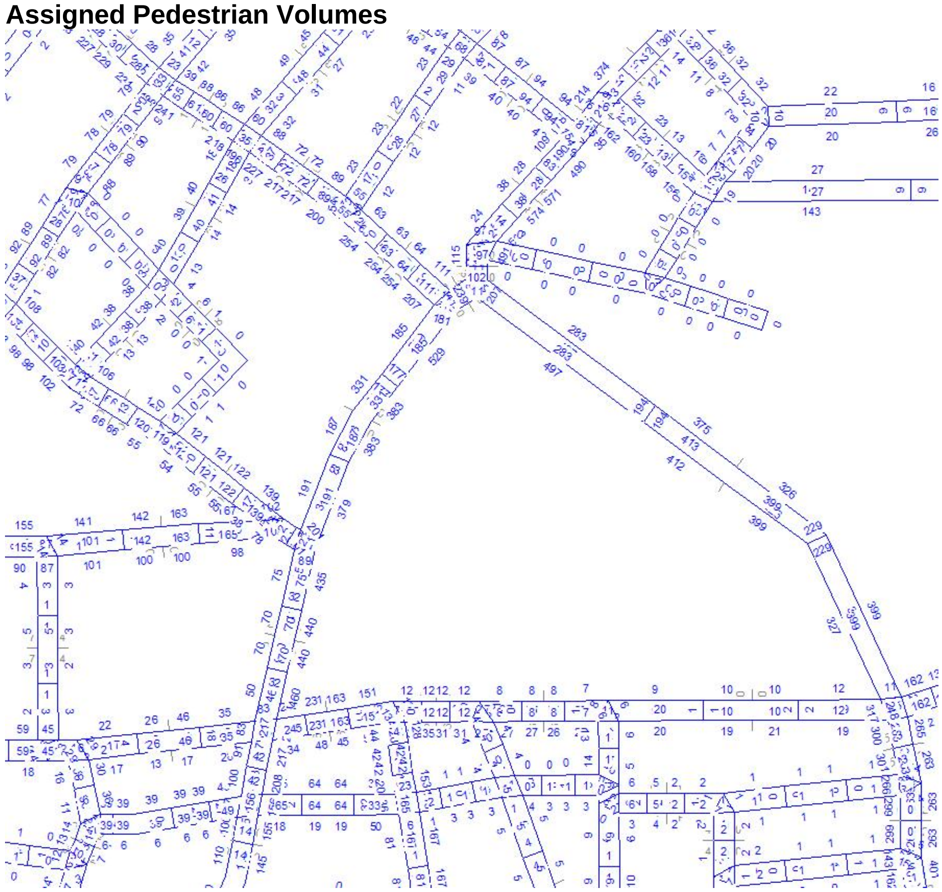
Table 15
Trip Assignment Set Weights

Perturbation Set (Purpose)	A	B	C
(1) Minimum (WK, SC)	0.40	0.30	030
(2) Medium (PB)	0.35	0.35	0.30
(3) Maximum (EM, SH, LS)	0.35	0.35	0.35

The product of this step is a loaded network containing estimated 24-hour pedestrian volumes on all links in the network: sidewalks, intersection crosswalks, jay walks, and door links / load points. A typical portion of the Baltimore network is illustrated in Figure 11.

Assigned pedestrian volumes were compared with observed pedestrian counts, where available, to validate the model. The results of this validation are presented in Section 5, Case Studies.

Figure 11
Assigned Pedestrian Volumes



The PEDCONTEXT Model Job Stream

The set of software, data files, and methods that comprise the pedestrian flow model have been named PEDCONTEXT, reflecting the ability of the model to estimate pedestrian flows in a way that is sensitive to the surrounding land use and geographical context.

The programs that are used to run the model and their basic functions are shown in Table 16.

Table 16
PEDCONTEXT Model Program Functions

Environment	Program	Program Source	Function
ARCGIS	ARCGIS	ESRI	Prepare study area TIGER network data files
	ARCGIS	ESRI	Prepare study area street network supplemental link and node data files
	ARCGIS	ESRI	Prepare study area Property View data file
CENTRAL	STREETNAME	Custom	Break down property file address street names to components
	GIS2PREP	Custom	Standardize address names, prepare street network and property data files for processing
	GIS2WALK	Custom	Compute the pedestrian network, geo-code property addresses to block faces
	NODE_RECODE	Custom	Build and renumber pedestrian network topology
	HWYNET	Citilabs TP+	Build sample TP+ network file
	SIDEWALK_CODE	Custom	Attach sidewalk / crosswalk attributes
	HWYNET	Citilabs TP+	Build final TP+ sidewalk network
	HWYLOAD	Citilabs TP+	Skim sidewalk network for distance
	LAND_USE_BUILD	Custom	Build block-face land use data file from property data
	HWYLOAD	Citilabs TP+	Skim sidewalk network for distance and weighted times
	MATRIX	Citilabs TP+	Calculate accessibility
	TRIPGEN	Citilabs TP+	Calculate block-face pedestrian trip generation
	MATRIX	Citilabs TP+	Generate F-factors
	TRIPDIST	Citilabs TP+	Trip distribution, all purposes
	MATRIX	Citilabs TP+	Format trip time frequency distributions
	MATRIX	Citilabs TP+	Convert production-attraction format to origin-destination format
	MATRIX	Citilabs TP+	Compress trip tables by purpose to assignable trip tables

Table 16 (Continued)
PEDCONTEXT Model Program Functions

Environment	Program	Program Source	Function
	HWYLOAD	Citilabs TP+	Assign pedestrian trips to network
	HWYNET	Citilabs TP+	Strip unneeded fields from loaded network
	HWYNET	Citilabs TP+	Round and unload volumes from loaded network
	VIPER2GIS	Custom	Compress sidewalk topology back to TIGER GIS topology
	HWYNET	Citilabs TP+	Rebuild TIGER-topology TP+ loaded network
	VIPER	Citilabs TP+	Export loaded network (TIGER topology) shape file for GIS
ARCGIS	ARCGIS	ESRI	Display / analyze data

USER GUIDE

The PEDCONTEXT model consists of a series of data preparation steps, GIS methods, custom programs, and TP+ demand model runs that together create the needed land use and network databases, run the model, and compile output data.

This documentation describes the steps by which a new application for a study area can be created. It assumes that the reader has at least a rudimentary knowledge of data handling methods using Visual FoxPro, of the ARC-GIS geographic information system software, and of the TP+ and CENTRAL packages that are used for travel demand modeling.

The document follows the steps involved in setting up a new application for Langley Park, Maryland. Directory names, file names and the like reflect that geography. However, some illustrations were assembled while the initial case study for Baltimore was being conducted and have not been updated to or replaced by the Langley Park application. For purposes of this User Guide this disconnect should not be problematic.

ORGANIZATION OF THE PEDCONTEXT DIRECTORIES

The PEDCONTEXT system consists of program libraries, data files, and run directories. The software is set up to run according to a very specific arrangement of files and directories and, although there is flexibility for a knowledgeable user to modify the structure, it is strongly suggested that the basic scheme not be modified.

The following is the general arrangement of directories:

```
d:\MDOT_PED
        \COMMON.100
        \LIBRARY.100
        \PROGRAMS.100
        \BALT_N_02
                \CENSUS
                \COUNTS
                \INPUTS
                \OUTPUTS
        \LANGLEY_01
                \CENSUS
                \COUNTS
                \INPUTS
                \OUTPUTS
```

where:

d:	is the drive on which the system is installed (currently delivered on D:)
MDOT_PED	is the main directory containing all PEDCONTEXT components
COMMON.100	is a directory containing execution scripts and data files that are used by all applications and do not change
LIBRARY.100	is a directory containing all executable programs and components (Note ARC-GIS, TP+, and CENTRAL are

	installed in their own locations)
PROGRAMS.100	is a directory containing custom program source code and development libraries. Executables have been copied to the LIBRARY.100 directory
BALT_N_02	is the directory containing the Baltimore case study run. In addition to the standard subdirectories that follow, other subdirectories can and should be added as needed to organize the assembly of input data. When the model is run intermediate data files are accumulated here.
\INPUTS	contains input files assembled from various sources as described below
\OUTPUTS	contains key output files and reports
\PRINTOUT	contains reports from model runs
\COUNTS	contains traffic and pedestrian counts
\CENSUS	contains Census data
LANGLEY_01	is the directory containing the Langley Park case study

STEP 1: ASSEMBLE DATA

Data files must be obtained and organized for input to the process. The following are needed:

1. TIGER Line Files — These must be enhanced TIGER line files (as opposed to the files typically distributed with ARG-GIS or TransCAD or downloadable from various sources) that contain the Census Feature Code field (CFCC). Required fields in the raw files are:

TLID	tiger line id number	Numeric
FEDIRP	feature direction	Character
FETYPE	feature type	Character
FENAME	feature name	Character
CFCC	feature class code	Character
FRADDL	from address left side	Numeric
FRADDR	from address right side	Numeric
TOADDL	to address left side	Numeric
TOADDR	to address right side	Numeric

2. PROPERYVIEW Files — These are obtained from the State of Maryland PropertyView system and contain property attributes for each parcel in the study area. Required fields in the raw files are:

ACCTID	Parcel account id number	Numeric
DIGXCORD	X-coordinate	Numeric
DIGYCORD	Y-coordinate	Numeric
CT2000	Census tract number	Numeric
BG2000	Census block group number	Numeric
ADDRESS	Full street address	Character
STRTNUM	Street number	Character
STRTDIR	Street prefix direction	Character
STRTNAM	Street address name	Character
STRTTYP	Street address type	Character
STRTSFX	Street address suffix	Character
STRTUNT	Street address units	Character
CITY	City name	Character
ZIPCODE	Zipcode	Character

LEGAL1	Legal description, Line 1	Character
LEGAL2	Legal description, Line 2	Character
LEGAL3	Legal description, Line 3	Character
TOWNCODE	Town code number	Character
DESCTOWN	County / Town description	Character
ZONING	Zoning code	Character
MZI	Multiple zoning indicator	Character
MFI	Multiple family indicator	Character
LU	Land use code	Character
DESCLU	Land use descriptor	Character
ACRES	Parcel land area (acres)	Numeric
STRUCODE	Code for type of structure	Character
STRUSTRY	Number of stories	Character
DESCSTRY	Stories description	Character
STRUDWEL	Type of dwelling	Character
DESCDWEL	Dwelling type description	Character
SQFTSTRC	Foundation square footage	Character
CIUSE	Commercial / industrial use code	Character
DESCIUSE	Commercial use description	Character
DWLL_TOTAL	Number of dwellings	Numeric
APT_UNITS	Number of apartments	Numeric
SEQNUMB	Database record number	Numeric

Note that STRTNUM, STRTDIR, STRTNAM, STRTTYP, STRTSFX, and/or STRTUNT may not be available in a particular county file. Where this data is missing, PEDCONTEXT software parses these components from the full ADDRESS field information.

STEP 2: PREPARE THE TIGER LINE FILE

The street segments comprising the study area must be selected from the raw wider area (county) TIGER line file, and additional data fields added as described below.

A. Be sure the file as shipped has the correct coordinate system declared for it (GCS_NAD_1927_Definition_1976) for TIGER line files). Start ArcCatalog. In the catalog panel find the subject TIGER file. Right click on it, select PROPERTIES | SHAPE=GEOMETRY | SPATIAL REFERENCE click Pulldown | SELECT | GEOGRAPHIC COORDINATE SYSTEMS | NORTH AMERICA | NAD 1927 (Definition 1976).prj.

   Now change the coordinate system from GCS_NAD_1927_Definition_1976 (which is long/lat) to MD (which is state plane feet by exporting to an Access database while changing coordinates, then re-import. In the file panel right click on the above TIGER line file, EXPORT | SHAPEFILE TO GEODATABASE | enter JUNK in Output Geodatabase box (becomes c:\junk.mdb). Click on 'ENTER THE NAME…' box, then click CHANGE SETTINGS | CHANGE | SELECT | PROJECTED COORDINATE SYSTEMS | STATE PLANE | NAD 1983 (FEET) | NAD 1983 StatePlane Maryland FIPS 1900 (Feet).prj. Back out, click OK to export.

   In the catalog panel double-click on C:\JUNK.MDB to reveal component file names in the Contents panel. In the Contents panel right click on the exported file | EXPORT | GEODATABASE TO SHAPEFILE | Select shapefile location and name. This will be a new shapefile for the TIGER line file with stateplane coordinates (feet).

Exit ArcCatalog.

B. Make sure the appropriate columns, column headings, and column values are included in the data being used, this should be done before any further work is started. Required columns, column headings, and values can be found in bold in the data dictionary located above in Step 1.

C. Call up the TIGER line shape (.shp) file. Create a new (empty) map. In the catalog panel right click LAYERS | ADD DATA | select file (tgr24031lka.shp).

D. If the study area spans two or more counties, call up both TIGER files and merge them (TOOLS | GEOPROCESSING WIZARD | MERGE LAYERS TOGETHER). Load / select the merged TIGER file.

E. Attach stateplane x-y coordinates to the TIGER link records. In ArcMap, open the TIGER line shape file that was converted to state plane above. Right click on the layer.| OPEN ATTRIBUTE TABLE | OPTIONS | ADD FIELD. Specify datatype = DOUBLE, precision = 15, and scale = 3. Do this once for each of the following fields:

START_X
START_Y
END_X
END_Y
MID_X
MID_Y
FIRSTJW_X
FIRSTJW_Y
LASTJW_X
LASTJW_Y

Right click each of the above field headings in the attribute table | CALCULATE VALUES | YES | LOAD a unique script for each field as follows:

START_X	*scriptlib*\POLYLINE_GET_X_FROMPOINT.CAL
START_Y	*scriptlib*\POLYLINE_GET_Y_FROMPOINT.CAL
END_X	*scriptlib*\POLYLINE_GET_X_TOPOINT.CAL
END_Y	*scriptlib*\POLYLINE_GET_Y_TOPOINT.CAL
MID_X	*scriptlib*\POLYLINE_GET_X_MIDDLEPOINT.CAL
MID_Y	*scriptlib*\POLYLINE_GET_Y_MIDDLEPOINT.CAL
FIRSTJW_X	*scriptlib*\POLYLINE_GET_X_ONETHIRD_POINT.CAL
FIRSTJW_Y	*scriptlib*\POLYLINE_GET_Y_ONETHIRD_POINT.CAL
LASTJW_X	*scriptlib*\POLYLINE_GET_X_TWOTHIRD_POINT.CAL
LASTJW_Y	*scriptlib*\POLYLINE_GET_Y_TWOTHIRD_POINT.CAL

where: *scriptlib* = d:\MDOT_PED\PROGRAMS.100\CALCULATE_TIGER_XYCOORDS

Click OK to calculate.

F. Determine the area to be analyzed. The study area should not encompass more than 4,500 Tiger Links after the selection process has been made. If there are more than 4,500 links within the selected study area, the area will have to be modified accordingly.

Figure 12 shows an example of the Tiger Road Network for Baltimore City that is to be used in the analysis. The following steps need to be followed to properly select and cut the study area street network:

Figure 12
Baltimore TIGER Network

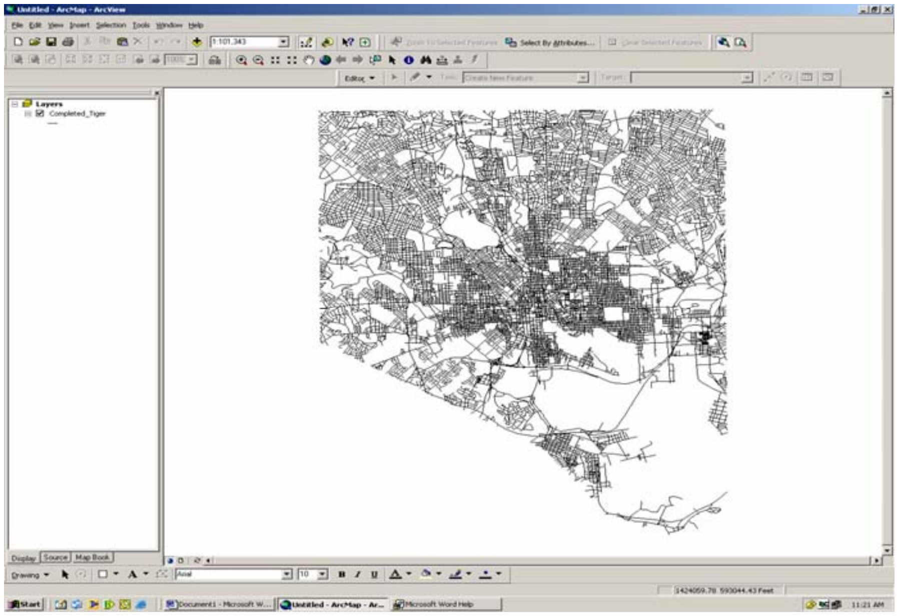


Figure 13 shows the base Tiger GIS network being used, for Baltimore City along with a selected area of study (links in light green). From this Tiger GIS network the GIS operator will select the Tiger links to be used in study. The selected study area links must be saved as a new GIS file (right-click source TIGER file on file panel **|** DATA **|** EXPORT-DATA).

Figure 13
Study Area TIGER Coverage

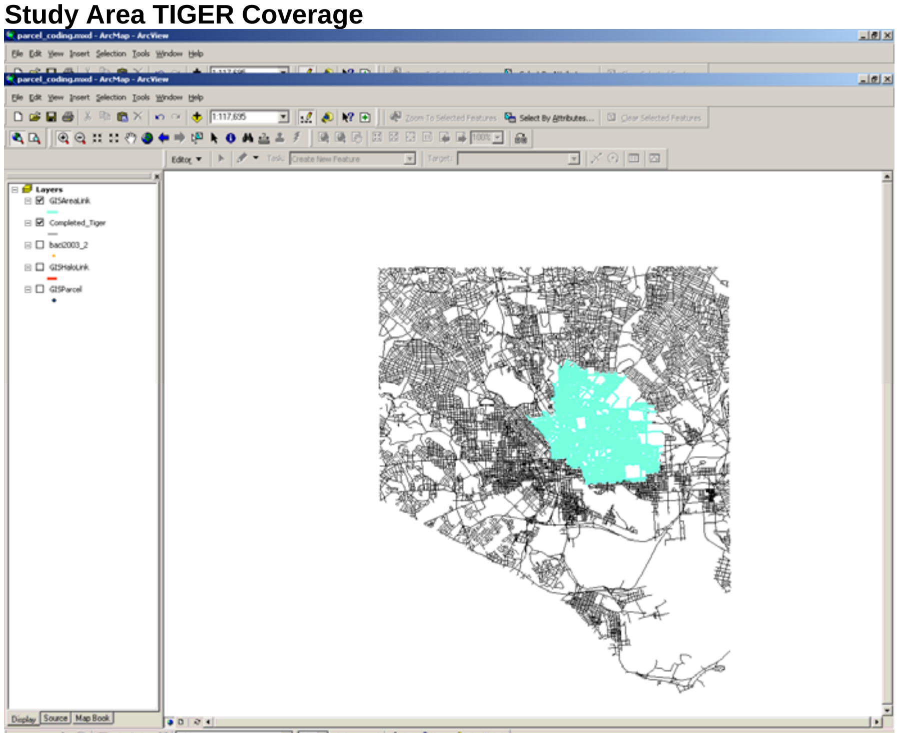


Figure 14 shows the study area in greater detail. Selecting the proper area can be done using one of many of ESRI's selection processes. This figure shows the links that fall into the study area that is to be analyzed

The beige area shows the boundary of the original Tiger road network selected. When selecting the study area the Tiger roads being selected must end on a single link. Figure 15 shows an example of how the study area links should be selected from the larger Tiger network. The Tiger roads within the red area can't be used as the end of study area because you have more than one link ending at the same point. The edge of the selected area must have a single link like those inside the blue circles. Figure 16 is a close-up of Figure 15.

Figure 14
Study Area Detail of TIGER network

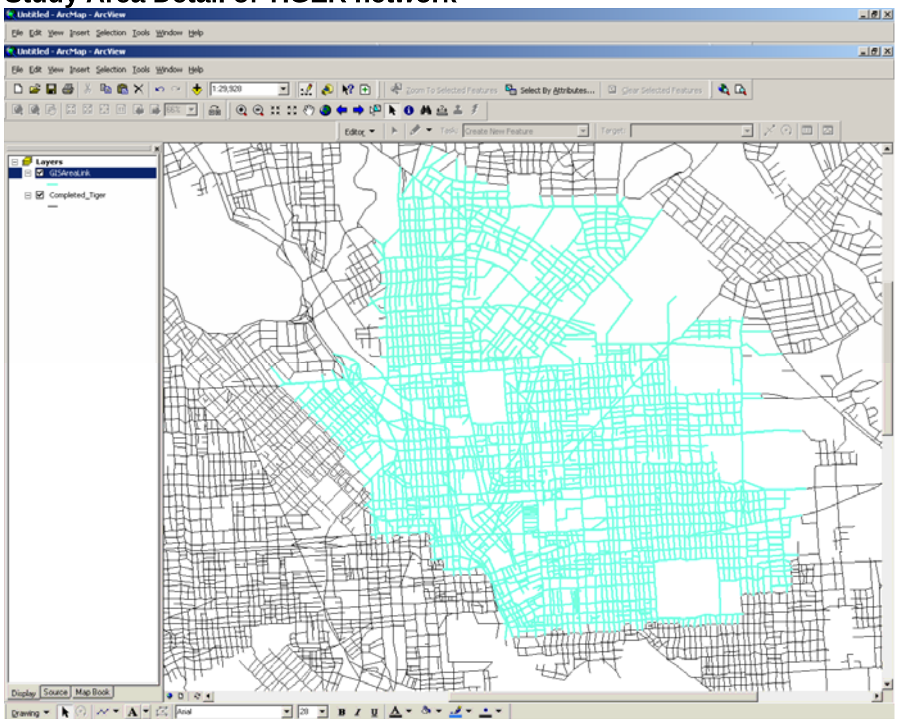


Figure 15
Study Area Edge Condition

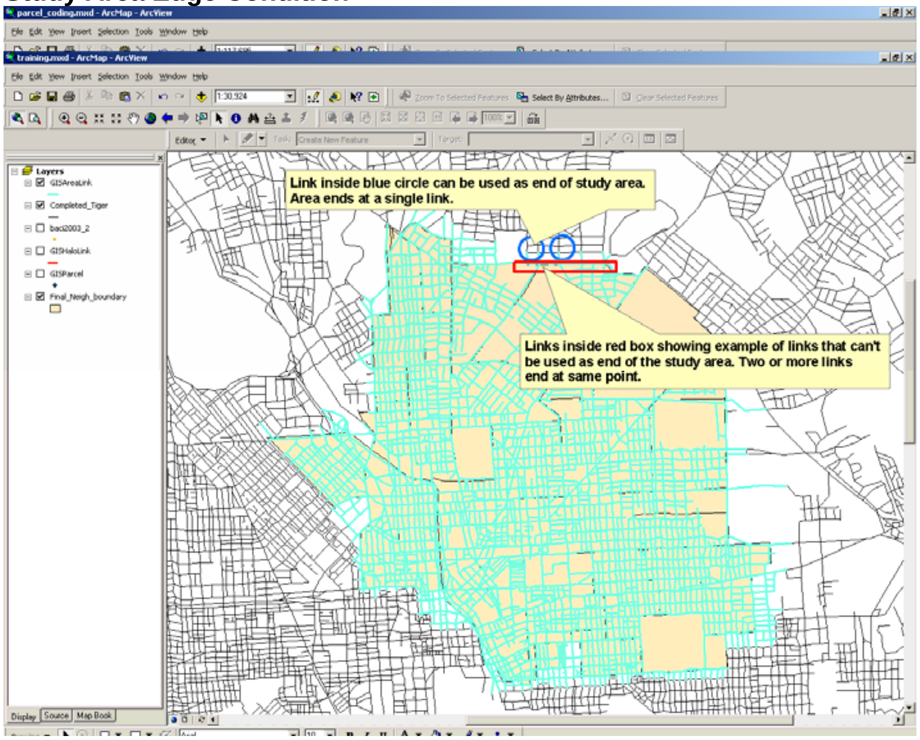


Figure 16
Detail of Study Area Edge Condition

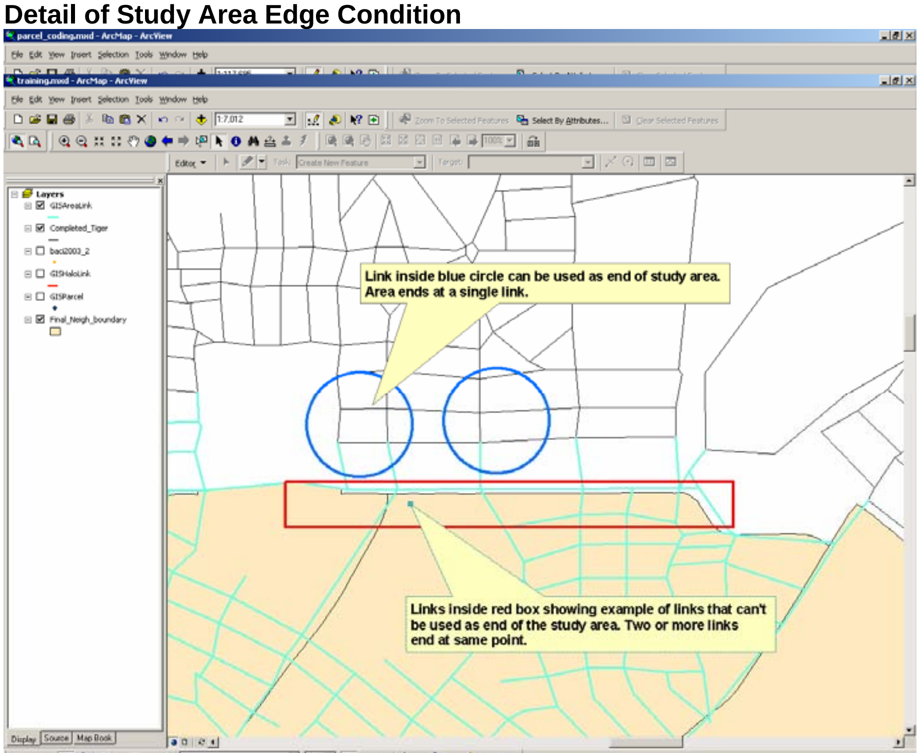


Figure 16 shows a part of the study area where ending in a single link can become difficult. The study area may have to be enlarged like it was here in order to get the proper end links.

The beige area represents the original network boundary. The study area has been extended in order to satisfy data requirements. The end of the study area must end like those inside the blue circle. Once you have properly adjusted the study area save it using the file name mentioned before.

Figure 17
Difficult Edge Condition

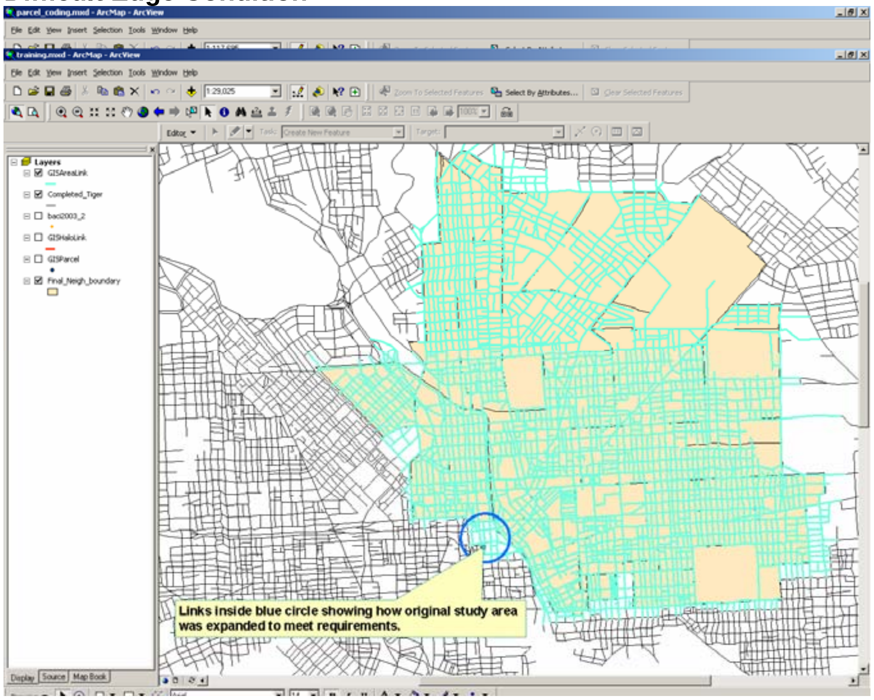


Figure 18
Detail of Difficult Edge Condition

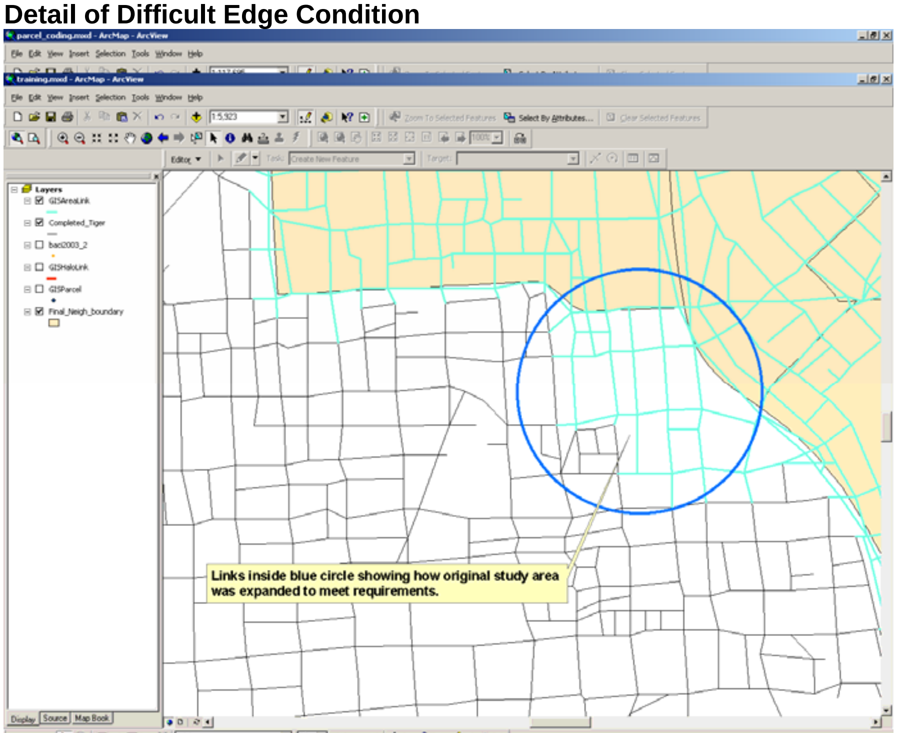


Figure 18 is a close-up of Figure 17; this shows how study area has been extended to accommodate data needs.

G. Boundary links just outside the study area must be selected and output. These links define the “End of the World”. The halo links which are represented as the red links below are to be included so that the pedestrian network building program can distinguish between dead ends within the study area and where the study area ends.

Figure 19
Boundary "Halo" Links

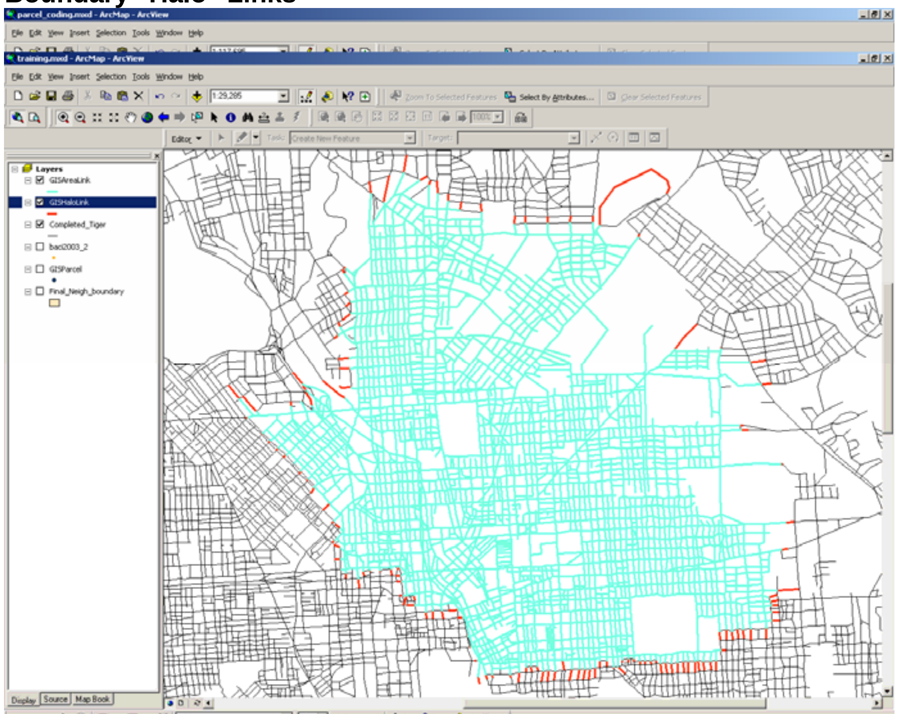
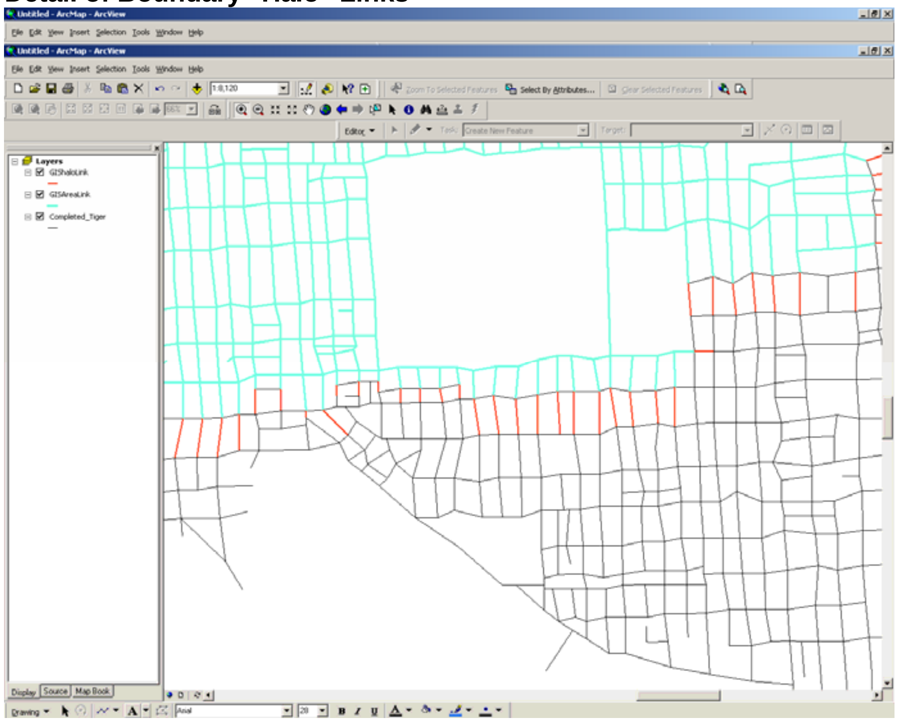


Figure 20
Detail of Boundary "Halo" Links



The light green Tiger links represent the study area that has been selected. The red links represent the halo links for the study area. Halo links are a single link extension around the study area. Only one halo link can be associated to a study area link. You can not use a single halo link to end two or more study area links.

STEP 3: PREPARE THE PROPERTY VIEW PARCEL DATA FILE

This file contains data regarding building types, i.e. vacant lot, office building, residential for each property in the study area. This file provides address locations and building square footage that is accumulated by the PEDCONTEXT software to land use totals for each block face.

Once you have selected your study area TIGER links (Step 2) you must then select the proper parcel data. This can be easily done by just using the select tool in ESRI to draw a box around the area links. Just make sure you draw the box big enough to capture all parcels that relate to selected area links. It is acceptable to have a larger coverage in this file than the area covered by the network, thereby having extra properties (See Figure 22). Save the selected parcel data as **GISParcel**.

Figure 21
Property View Parcels, City of Baltimore

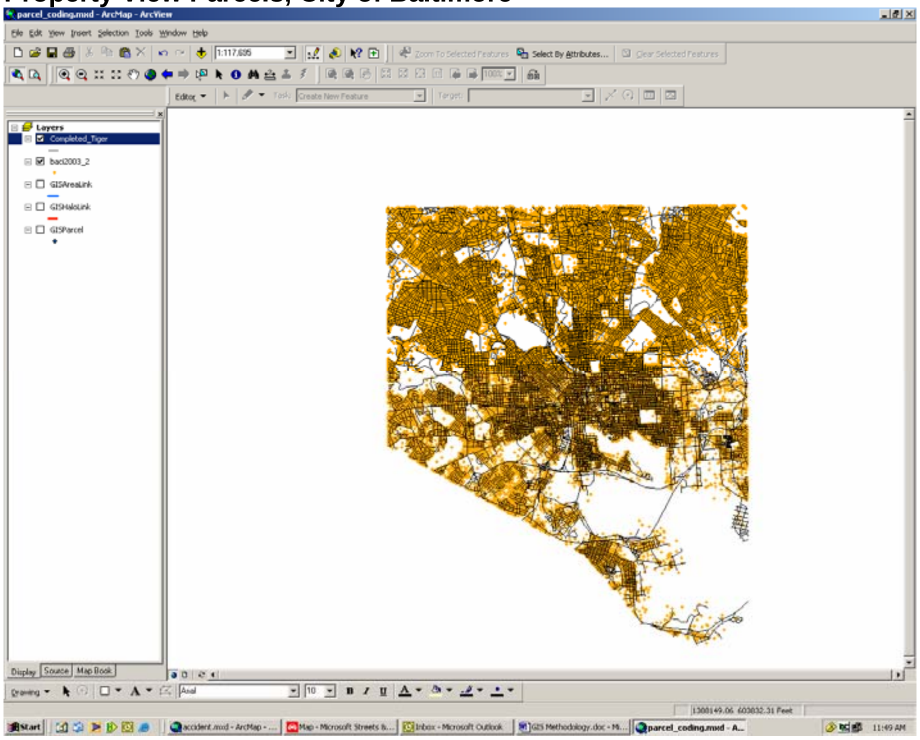
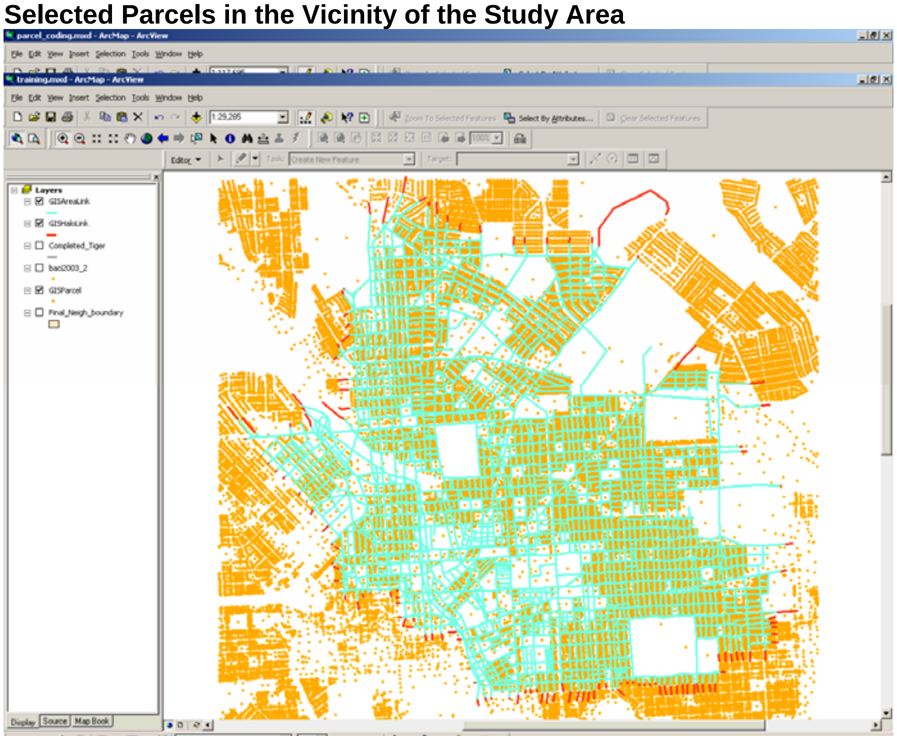


Figure 22
Selected Parcels in the Vicinity of the Study Area



STEP 4: RUN THE INITIAL DATA BUILD

Run the initial model steps to prepare network data, build the sidewalk network, and renumber sidewalk nodes to create a VIPER-compatible network.

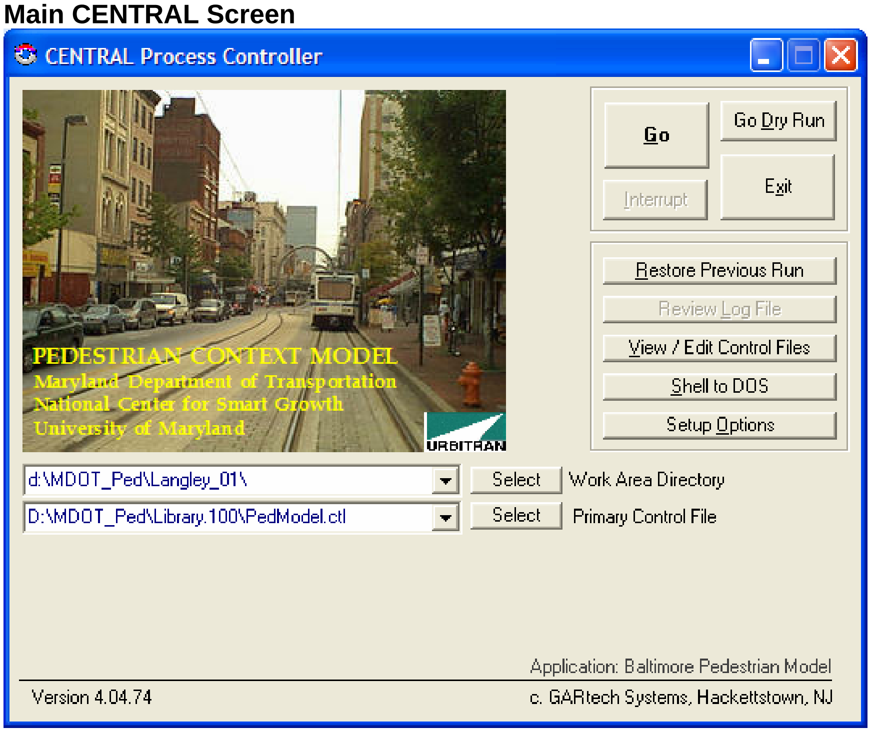
A. Copy the necessary network and property files to the *d:scenario*/INPUTS directory:

Selected Study Area TIGER links
Selected Halo Area TIGER links
Selected Property file records

In addition, copy from a prior scenario INPUTS directory, or from the COMMON.100 directory, a PEDMODEL.SET file. This file contains settings from prior dialogs that can be adapted to the current run. Lacking this file, it is necessary to override default settings provided by the system.

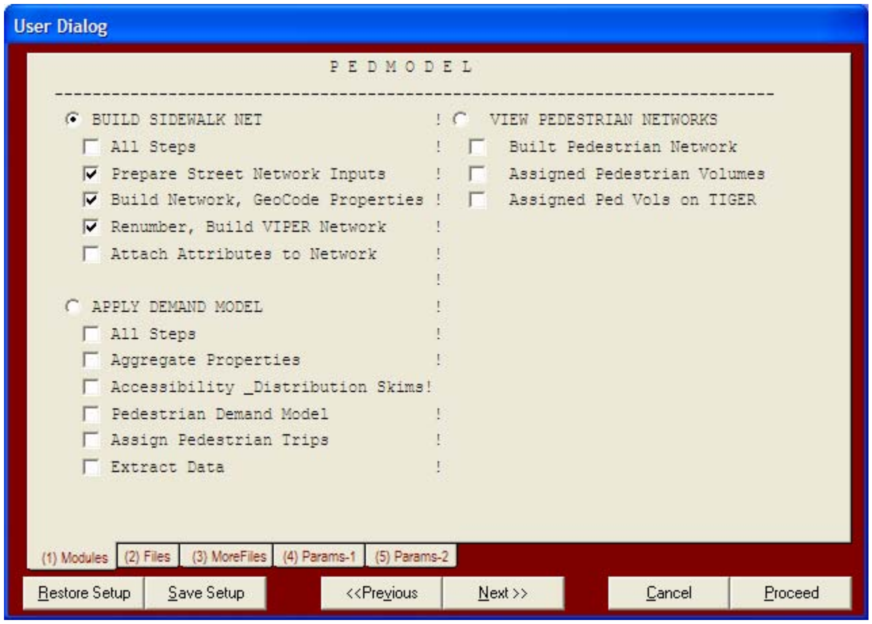
B. Start CENTRAL. In the "Work Area Directory" box select the directory in which the intermediate run data will be created (i.e. d:\MDOT_PED|LANGLEY_01). Be sure the "Primary Control File" is set to d:\MDOT_PED\LIBRARY.100\PEDMODEL.CTL.

Figure 23
Main CENTRAL Screen



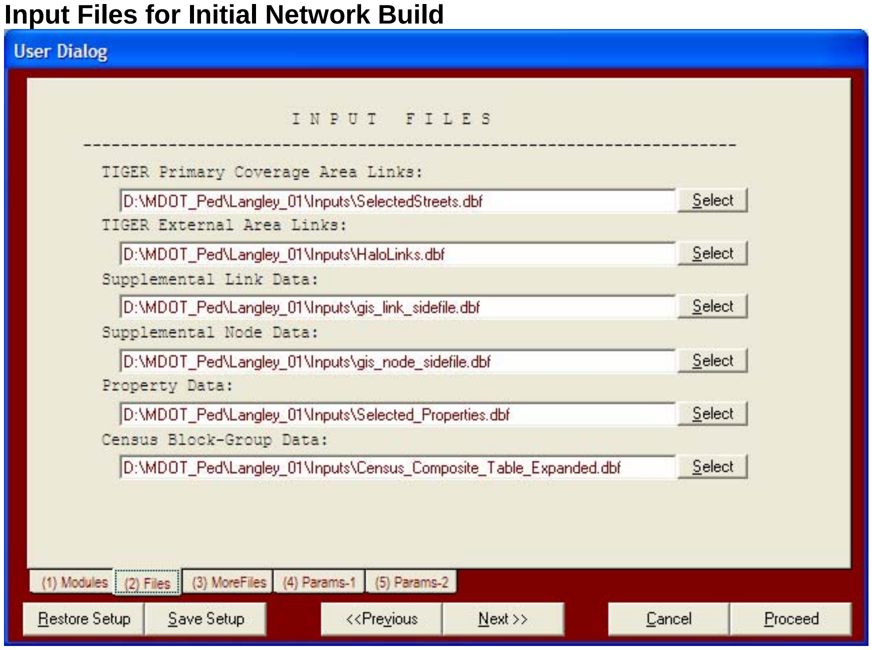
C. Click GO. On the MODULES screen select to BUILD SIDEWALK NET, and only the steps to Prepare the Street Network, to Build the Sidewalk Network and Geocode Properties, and to Renumber the Sidewalk Network and Build a VIPER Network (see Figure 24).

Figure 24
Module Selection Screen: Initial Network Build



D. On the FILES screen specify input files. For this initial build, only the selected study area TIGER links, the TIGER external area links, and the property data files are required (see Figure 25).

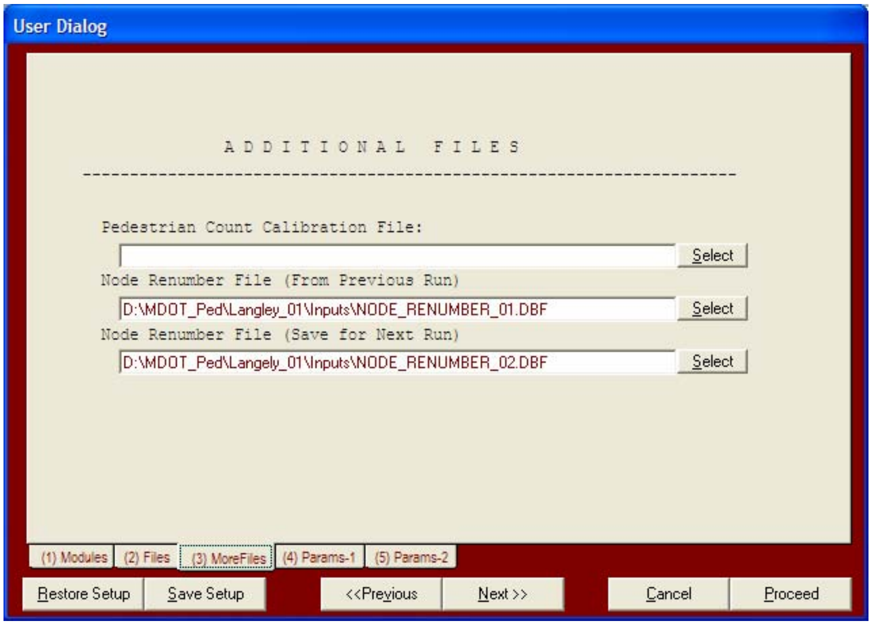
Figure 25
Input Files for Initial Network Build



E. The network nodes are renumbered each time the network is built, in sequence according to link number. Adding or removing a link will change the stream of numbers, so node numbers can change from build to build. If this matters (i.e. outside data such as counts are referenced by node number, then it is desirable to control the renumbering with a RENUMBER file.

If this network build has been run before, you probably will want to use the same renumbering scheme so that intersection / node numbers are consistent wherever possible. Each time the NODE_RECODE program is run it produces a RENUMBER file which records the renumbering scheme. This RENUMBER file can be saved then restored during the next run, using the file declarations on the "MORE FILES" screen shown in Figure 26:

Figure 26
Additional Files Screen



F. Click PROCEED to run the model.

G. A "green" screen indicates a successful model run. Review run reports by clicking on "Review Print Files" and select a file to view.

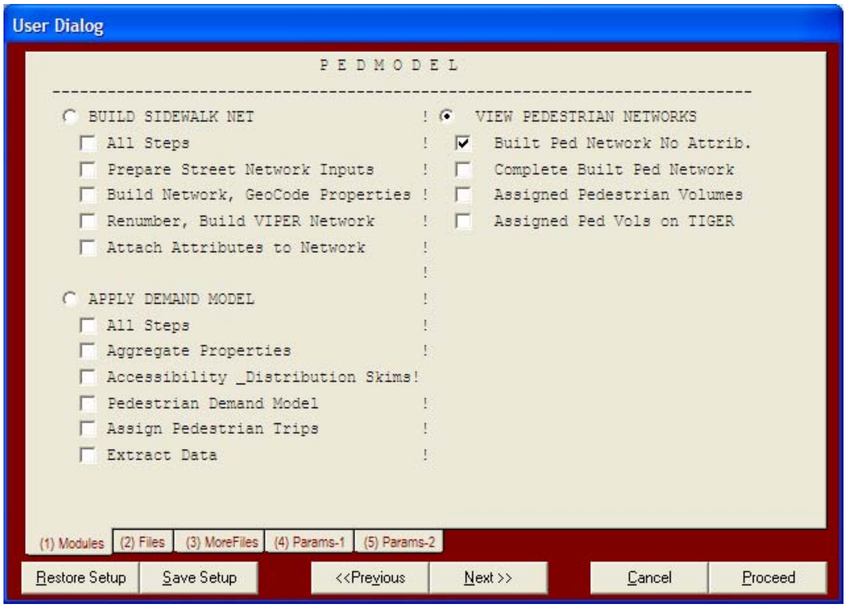
H.. Start the VIPER network editor program and load the SIDEWALK_NOATTRIBUTES.NET file that was just created. This can be done using the module selection shown in Figure 27.

Examine the sidewalk network topology that was built. If the layout looks normal, then most likely the sidewalk network was built without errors. If there is an area near an external boundary that has long disconnected lines (this will be obvious to the eye) then network construction errors occurred, most likely during the process of extracting the study area and halo links from the source TIGER file.

In VIPER, build and trace paths from a series of representative zones distributed across the study area to other representative zones. Look for disconnects and lack of access. It may be that there are node numbering disconnects in the FNODE and TNODE fields on the source TIGER file. If disconnects are found, edit the selected street TIGER file in ARCGIS to make the FNODE and TNODE values the same at intersection points. Then rerun the network builder and test paths again until there are no disconnects.

THIS PATH CHECKING IS PARTICULARLY IMPORTANT IF THE NETWORK HAS BEEN ASSEMBLED FROM TWO COUNTIES' TIGER FILES, AS WAS THE LANGLEY PARK CASE STUDY

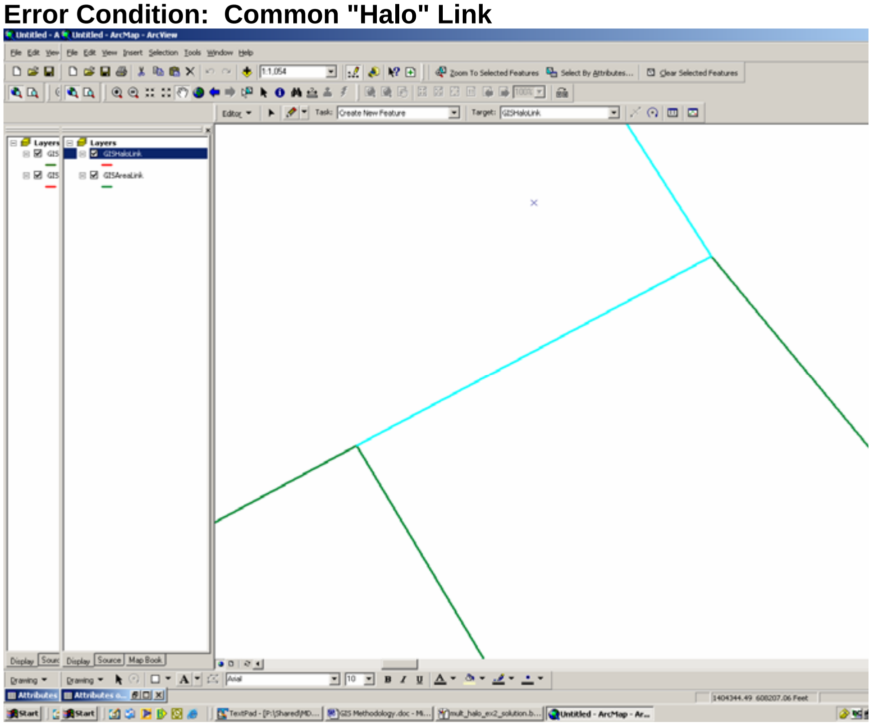
Figure 27
Module Selection Screen: View Built Pedestrian Network



Step 5: CORRECT ERRORS IN THE SELECTION PROCESS

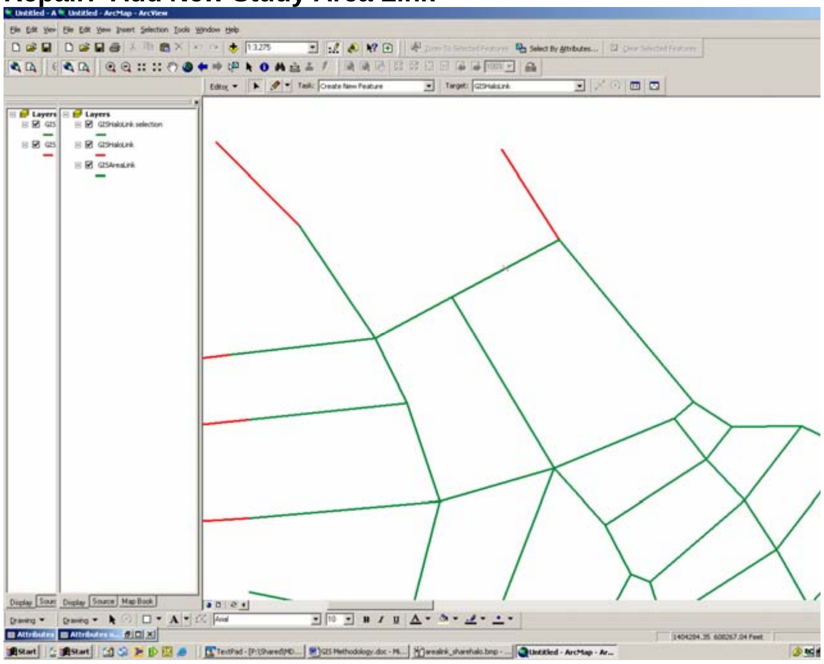
After the model is run using the DBF's from the GISAreaLink file and the GISHaloLink file the program will let you know if there were any errors in the selection process. The following figures are examples of some error messages one might get.

Figure 28
Error Condition: Common "Halo" Link



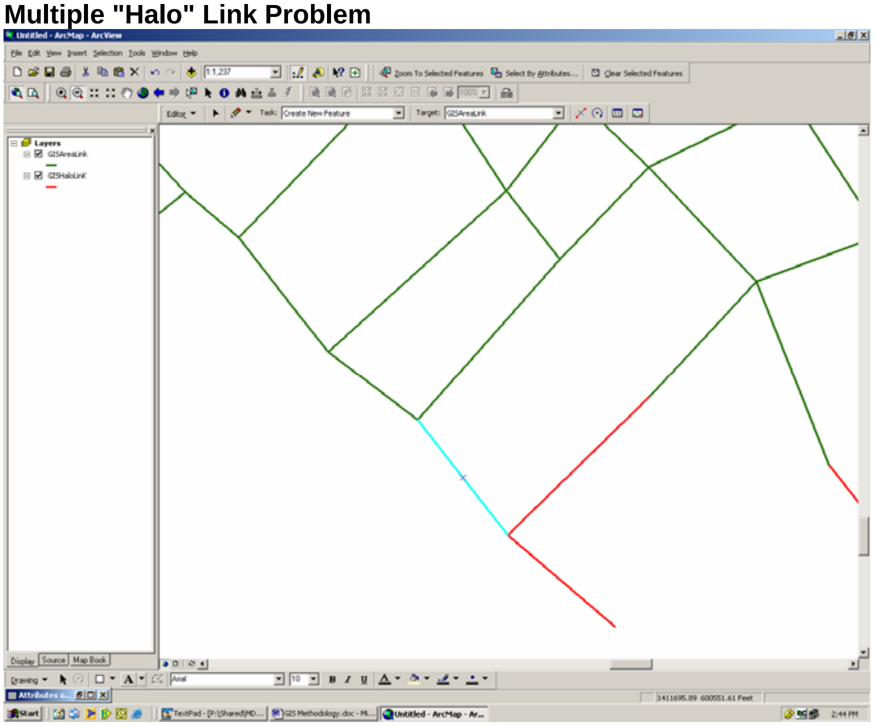
If you get an error message saying "Area links have common Halo link" the above figure shows what the error is. In this case a halo link is being shared by two area links. You can only have one halo link per area link. In order to fix this problem a new area link was added so that halo links are not being shared

Figure 29
Repair: Add New Study Area Link



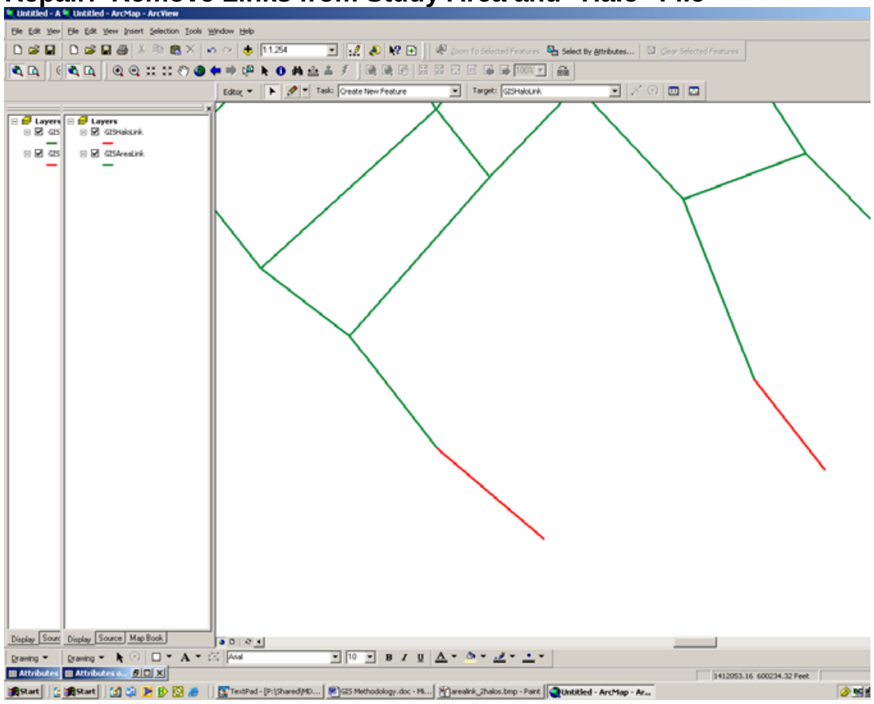
.The figure below shows another example of the multiple halo problems like figure 12

Figure 30
Multiple "Halo" Link Problem



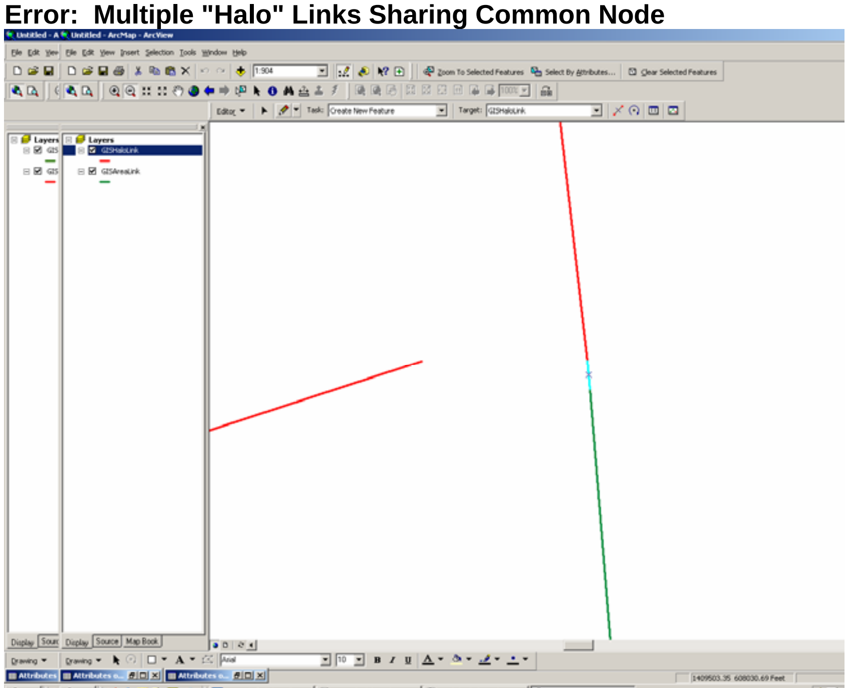
The figure below shows that some links have been removed from both the area link file and the halo link file. In this case, this was the easiest way to fix in order to fix the halo link sharing problem.

Figure 31
Repair: Remove Links from Study Area and "Halo" File



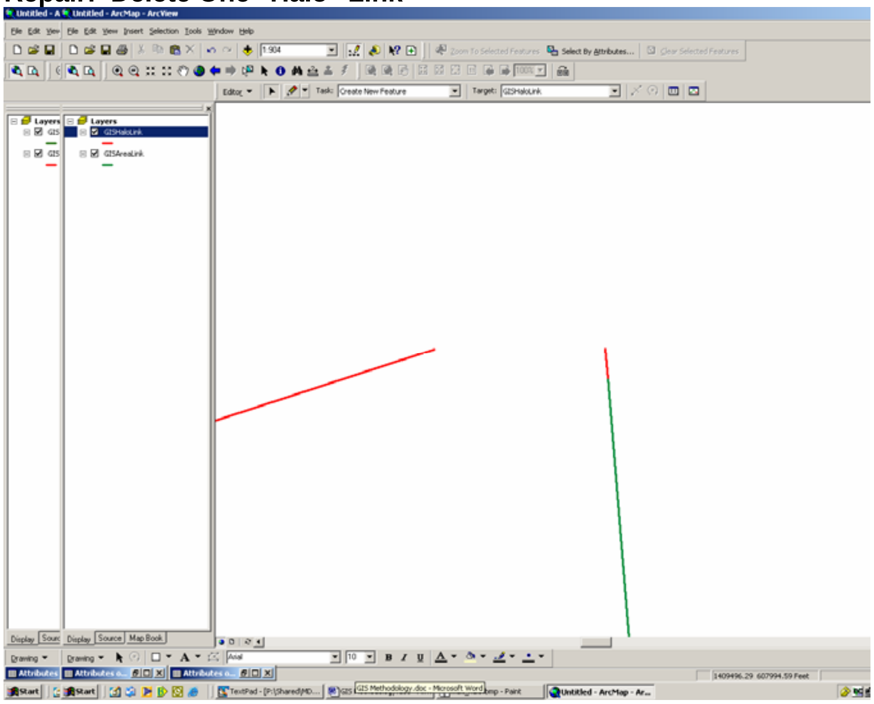
The figure below shows error message "Multiple halo links sharing common node". This happens when you have multiple halo links sharing the same fnode or tnode.

Figure 32
Error: Multiple "Halo" Links Sharing Common Node



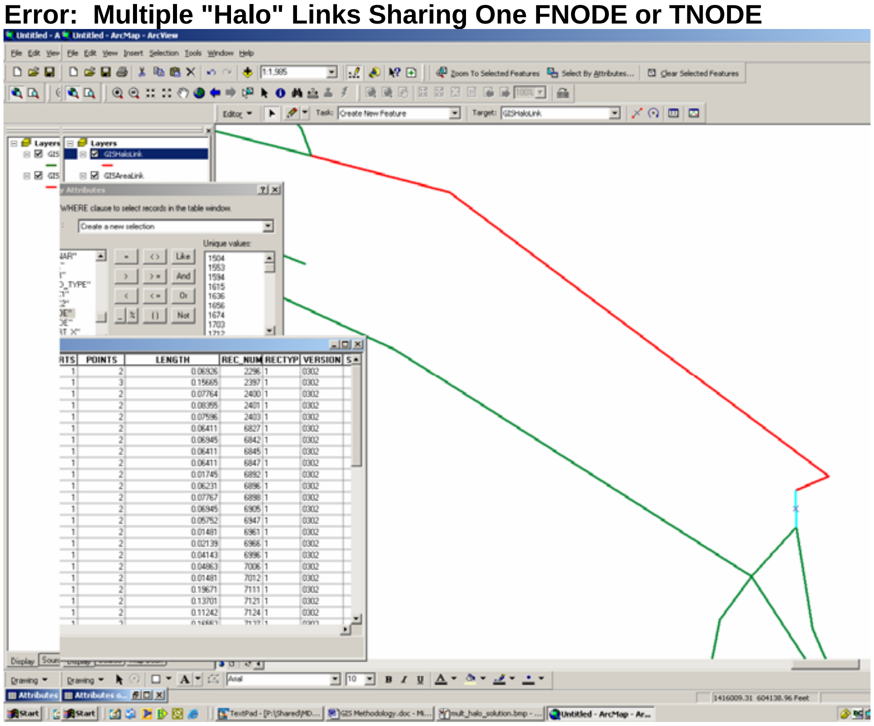
The solution for this problem is to delete one of the halo links.

Figure 33
Repair: Delete One "Halo" Link



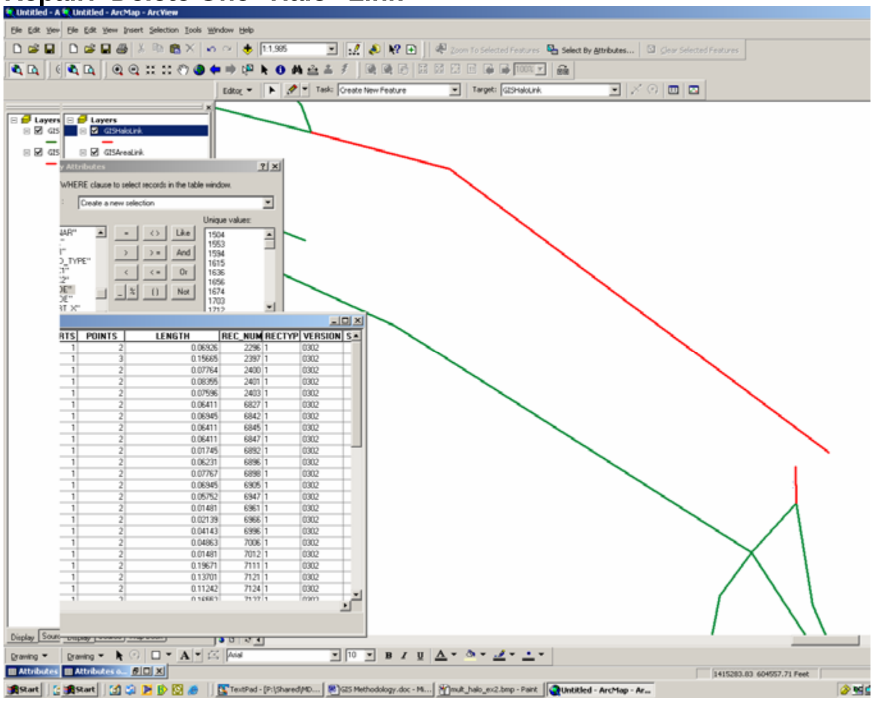
The figure below shows error message"". This happens when you have multiple halo links sharing the same fnode or tnode.

Figure 34
Error: Multiple "Halo" Links Sharing One FNODE or TNODE



The solution for this problem is to delete one of the halo links.

Figure 35
Repair: Delete One "Halo" Link



If you get the error message "Link not connected" this means that one of the halo links is not properly connected to the end point of an area link.

Figure 36
Error: Link Not Connected

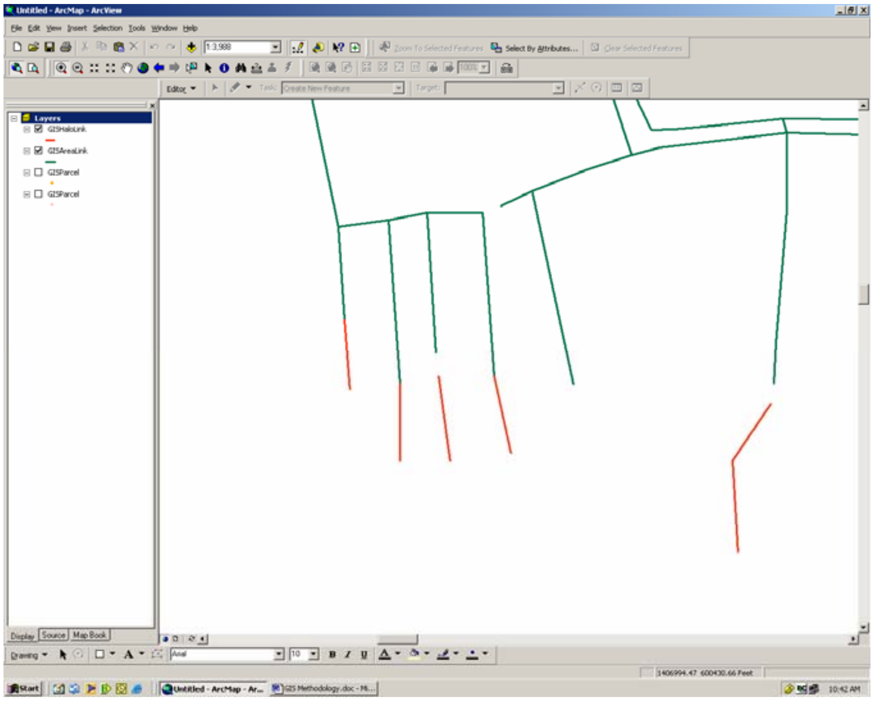
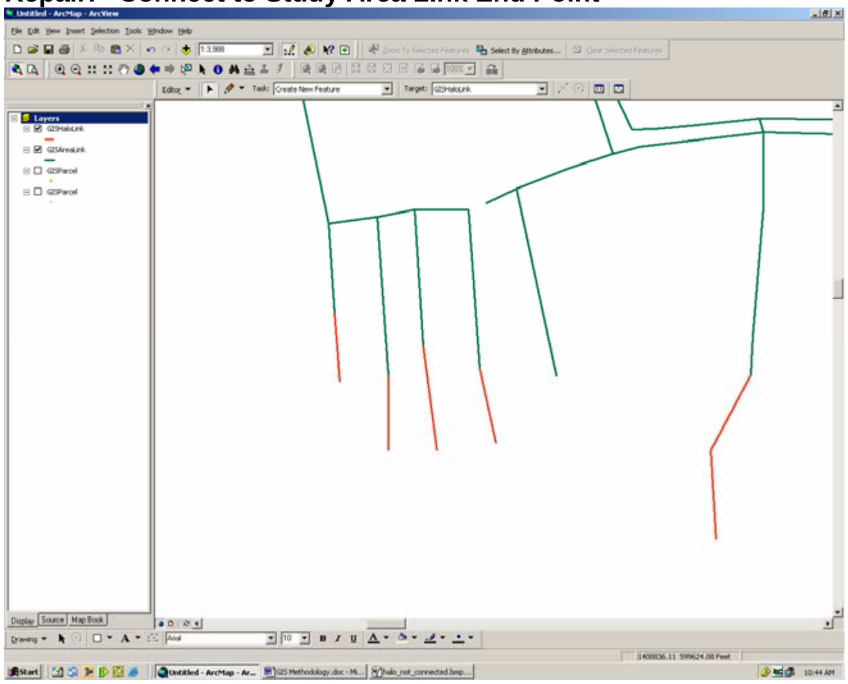


Figure 37
Repair: Connect to Study Area Link End Point



STEP 6: PREPARE THE SUPPLEMENTAL LINK AND NODE DATA FILES

Supplemental link and node data files contain additional street and sidewalk data beyond that available in the TIGER line files. These files must be created and edited through the GIS.

A. Copy the GIS_LINK_SIDEFILE.DBF and GIS_NODE_SIDEFILE.DBF files from a previous scenario.

B. Attach the files to the TIGER GIS and create a record for each TIGER link in the selected streets layer.

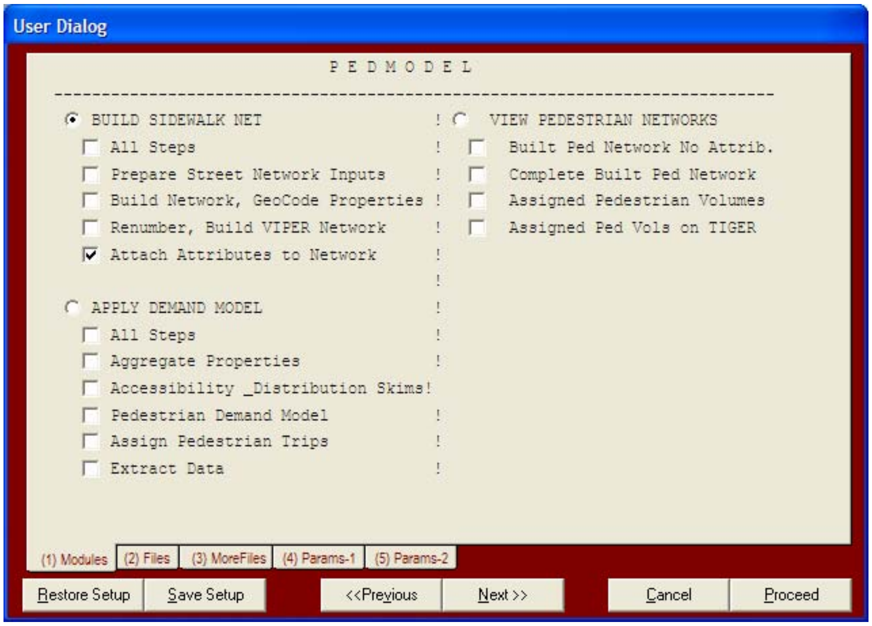
C. Add street and sidewalk attribute data as appropriate.

STEP 7: RUN THE STEP TO ATTACH SIDEWALK ATTRIBUTES TO THE NETWORK

A. In CENTRAL, restart the model by clicking GO

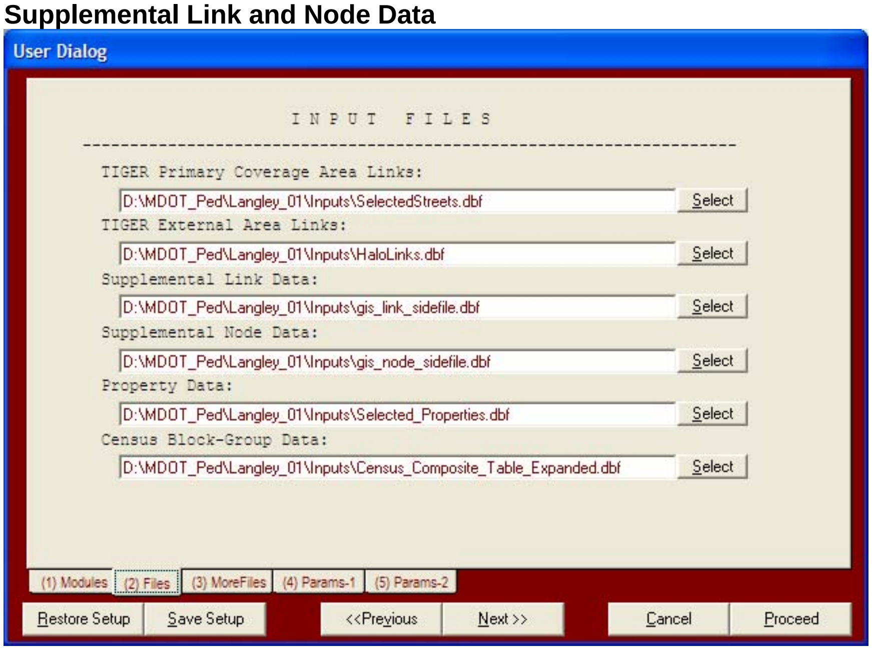
B. On the Modules screen, select ONLY the options to BUILD SIDEWALK NET and ATTACH SIDEWALK ATTRIBUTES TO NETWORK, as shown in Figure 38.

Figure 38
Module Selection Screen: Attach Attributes



On the Files screen, specify the SUPPLEMENTAL LINK DATA and SUPPLEMENTAL NODE DATA files.

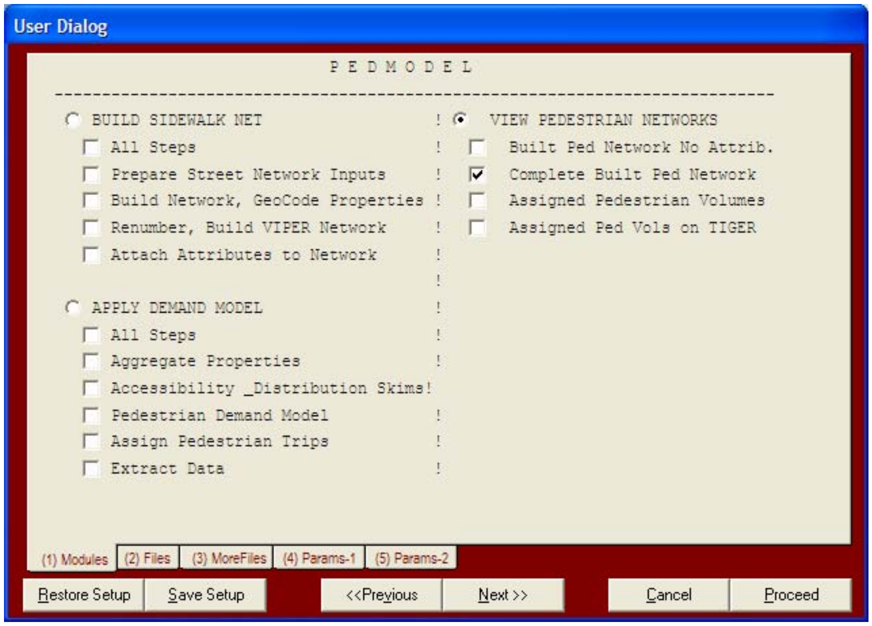
Figure 39
Supplemental Link and Node Data



C. Click PROCEED to begin the run. Review the print files to be sure execution completed without errors.

D. Load the complete built sidewalk network file into VIPER to review and validate topology and data coding. Use the Modules screen to select VIEW PEDESTRIAN NETWORKS | COMPLETE BUILT PED NETWORK:

Figure 40
Module Selection Screen: View Build Network



STEP 8 DOWNLOAD AND PREPARE CENSUS DATA

Population and housing characteristics are obtained at the block group level and applied to all households within the block group. To support this, block group data must be downloaded and prepared, using the CENSUS_COMPOSITE_TABLE_EXPANDED.XLS Excel workbook. Use the following procedure:

A. Go to the Census web site:
   http://factfinder.census.gov/servlet/DatasetMainPageServlet?_program=DEC&_lang=en&_ts=

B. Select Census 2000 Summary File 3 (SF-3)

C. Click “LIST ALL TABLES”

D. For each of the following tables:

DP-4	Profile of Selected Housing Characteristics
P-1	Population
P-16	Own Children <18years by Family Type by Age
P-52	Household Income
P-53	Median Household Income
QT-P23	Journey to Work

1. Click on the table name
2. Select a geography type = BLOCK GROUP
3. Select State = Maryland
4. Select County = Baltimore City
5. Select a Census Tract from the above list, select All Block Groups, click Add. Repeat for all tracts in the study area.
6. Click SHOW RESULT

7. Click PRINT/DOWNLOAD, then DOWNLOAD
8. Select MICROSOFT EXCEL, click OK
9. Locate the file destination and name appropriately.
10. Do for only the 1st file set – it actually contains data for all block groups.

E. Extract each Excel workbook from its zip file. Rename DT_DEC_2000_SF3_U_DATA1.XLS to:
   DP4_HOUSING-CHARACTERISTICS
   P1_POPULATION.XLS
   P16_CHILDREN-BY-AGE.XLS
   P52_HOUSEHOLD-INCOME.XLS
   P53_MEDIAN-INCOME.XLS
   QT-P23_JOURNEY TO WORK

F. Copy to the current working directory the CENSUS_COMPOSITE_TABLE_EXPANDED.XLS spreadsheet file (refer to it as the CENSUS file). Extract the QT_DEC_2000_SF3_U_GEO.XLS file from the one of the downloaded zip files (refer to it as the GEO file). Copy and paste the data he GEO_ID column from the GEO file into the CENSUS file. Be sure to clear all extra data in the column, or to copy the logic to additional rows if needed. Notice that Tract and BlkGrp are computed from the GEO_ID. Verify them. You may need to adjust the parsing formulas for tract (column B) and block-group (column C) to correct them from the GEO_ID field.

G. From each of the renamed data workbooks, copy and paste the data into the appropriate tabs of the CENSUS file. Check carefully that all tracts/blocks align. Re-download and fill in missing values as necessary.

H. When the SUMMARY worksheet is complete to your satisfaction, save the workbook. Then delete the first row (long titles) of the SUMMARY worksheet. Save the worksheet in dBaseIV .dbf format.

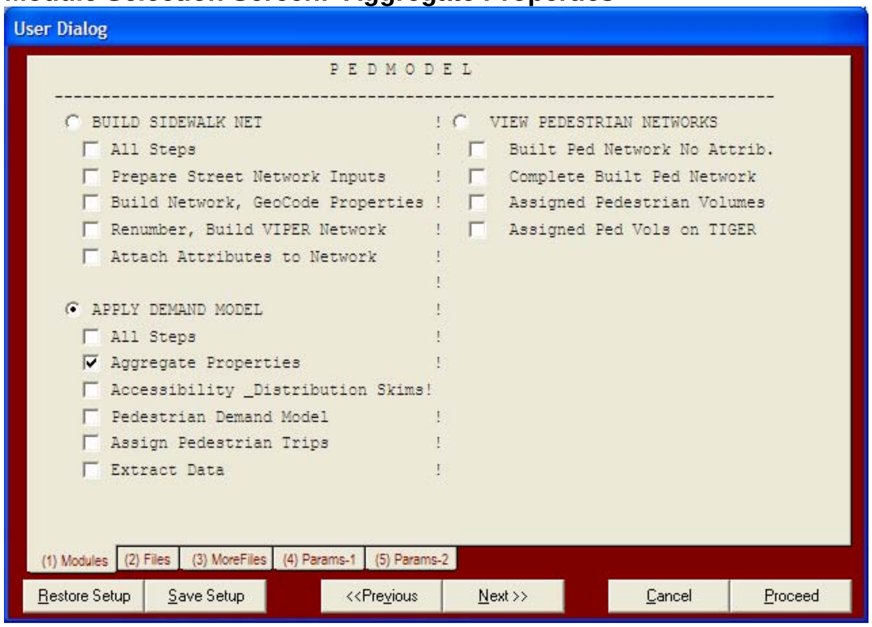
STEP 9: RUN THE LAND USE AGGREGATION PROGRAM

Land use data is aggregated to block face totals and census data must be attached to each block face.

A. In CENTRAL, restart the model by clicking GO

B. On the Modules screen, select ONLY the option to AGGREGATE PROPERTIES TO LAND USE FILE. On the Files screen specify the census block-group data file. Click PROCEED to run the model

Figure 41
Module Selection Screen: Aggregate Properties



C. After the run, check the print file. In particular at the bottom check the list of tracts and block groups not in the census block group data file. Download the needed census block group data and rebuild the data file as described above.

D. Rerun the model if data corrections were made.

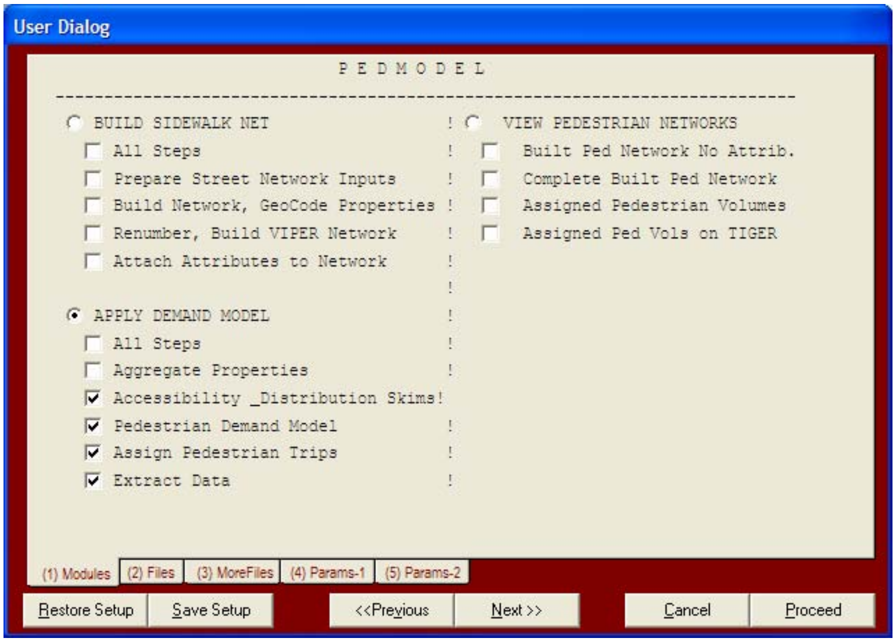
STEP 10: RUN THE TRAVEL DEMAND PORTION OF THE MODEL

The accessibility computations, trip generation, distribution, and assignment portions of the model are normally run together.

A. In CENTRAL, restart the model by clicking GO

B. Select APPLY DEMAND MODEL, and ACCESSIBILITY / DISTRIBUTION SKIMS, PEDESTRIAN DEMAND MODEL, and ASSIGN PEDESTRIAN TRIPS as shown in Figure 42. Also select EXTRACT DATA if it is desired to unload pedestrian volume estimates to summary VIPER network files and to shape files that can be loaded into ArcGIS.

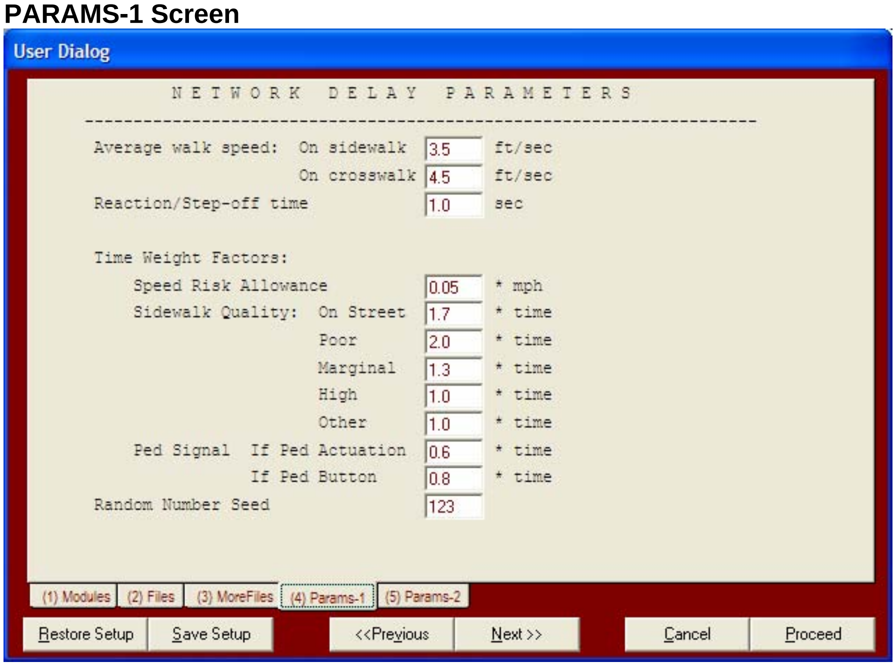
Figure 42
Module Selection Screen: Run the Demand Model



C. Specify assignment model parameters, using the PARAMS-1 and PARAMS-2 screens shown in Figures 43 and 44.

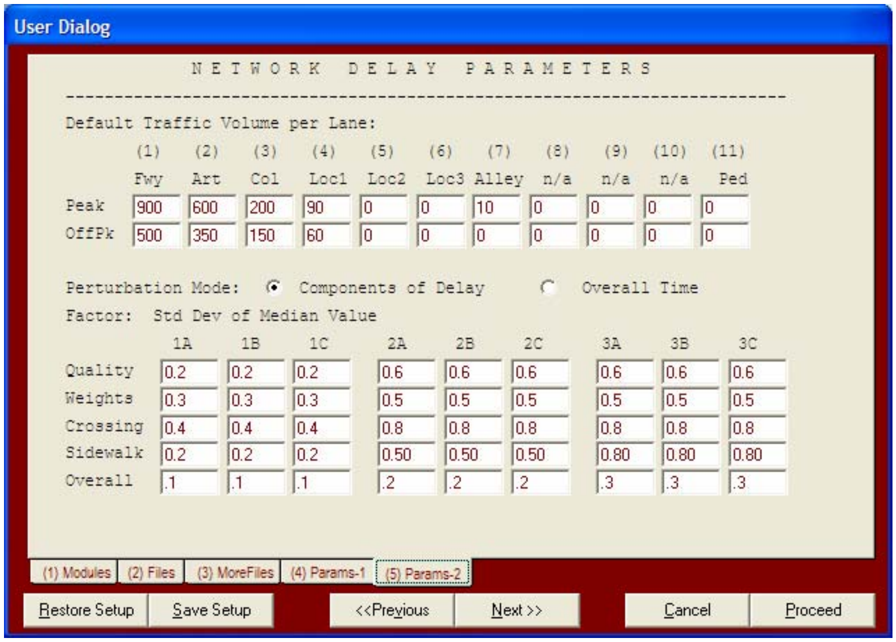
Values on PARAMS-1 (Figure 43) specify walk speeds and delay factors. In addition, a random number seed can be specified which, if changed from run to run, will produce a different distribution for the pseudo-stochastic assignment method.

Figure 43
PARAMS-1 Screen



Values on the PARAMS-2 screen (Figure 44) specify the default roadway traffic volumes during peak and off-peak hours, for the various facility types.

Figure 44
PARAMS-2 Screen



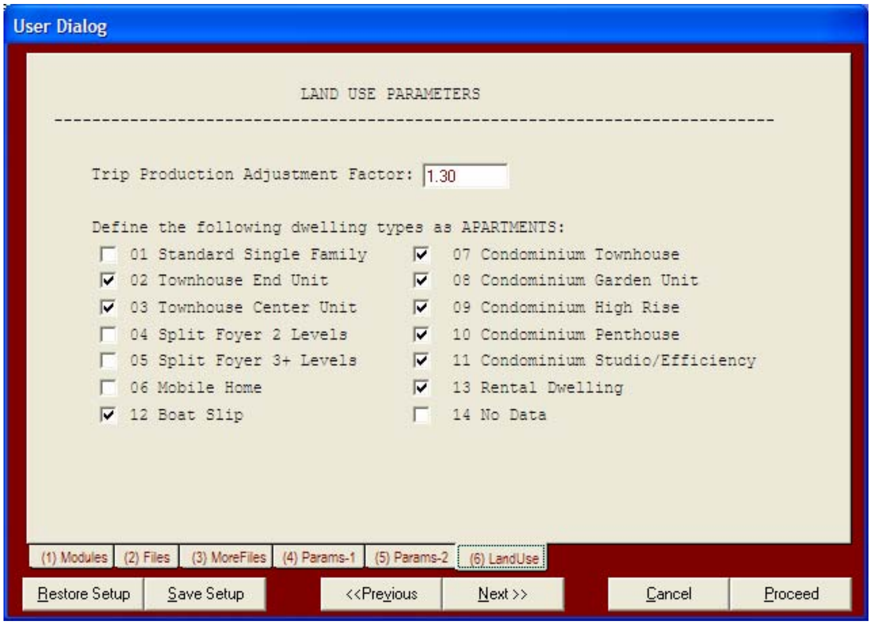
Also specified on the PARAMS-2 screen are the factors which set the standard deviation for each walk time distribution. Nine columns are provided, for each of the nine separate assignment sets. For each assignment set there are five factors that are used to compute a standard deviation for that time element. For example, time penalties added due to sidewalk quality are randomly estimated from a distribution with a median value as specified on PARAMS-1, and standard deviation computed by this factor times the median value. Increasing the factor, then, increases the variability.

Time perturbations can either be computed individually for the four components of delay (Quailty, Weight, Crossing time, and Sidewalk type) if the perturbation mode is wet to "Components of Delay", or simply on the Overall time if the perturbation mode is set to "Overall Time".

On the LANDUSE screen (Figure 45) settings relating to land use and trip generation can be specified. First, the Trip Generation Adjustment Factor allows an across-the-board adjustment of trip making levels. This factor is normally developed at the time of calibration, but can also be used for other analysis purposes such as accounting for growth or daily / seasonal variation. The default is 1.00, and reflects the calibration of the Baltimore case study to pedestrian counts. No adjustment is needed (factor = 1.00) for the default condition.

Also on this screen is specification of those dwelling types found in the parcel database STRUCODE field, that should be accounted for as apartments in trip generation. The apartment land use is indicative of lower income occupants with smaller household sizes. Townhouses within Baltimore, for example, are likely to be row-homes and of a character similar to apartments, whereas in Langley Park townhouses are more likely the typical suburban town home.

Figure 45
LANDUSE Screen



D. Click "GO" to run the model. At the completion of the run, review print files to insure the run has been performed correctly.

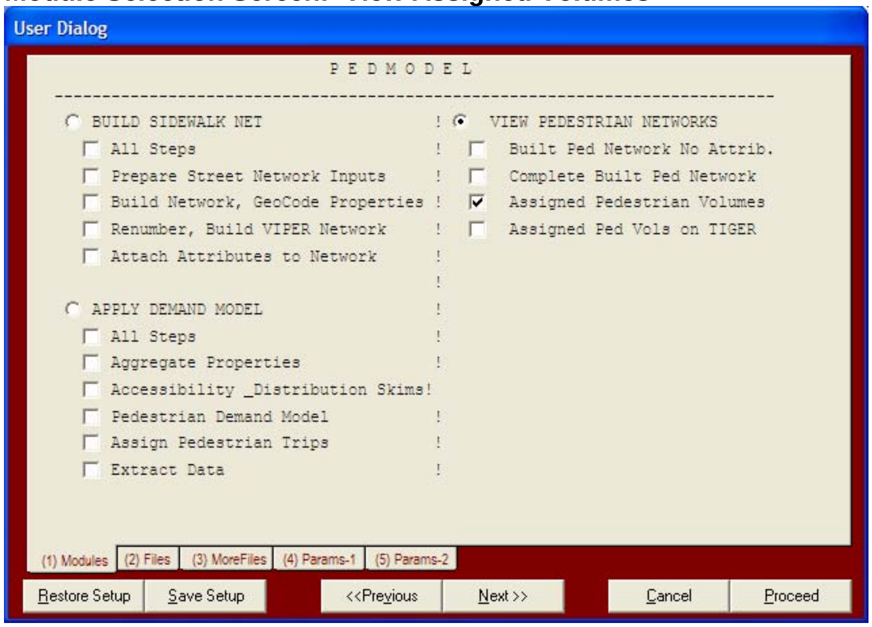
STEP 11: VIEW THE ASSIGNED PEDESTRIAN VOLUMES WITH VIPER

Using the Citilabs VIPER program, the loaded pedestrian network volumes and other attributes can be viewed.

A. In CENTRAL, restart the model by clicking GO

B. Select VIEW PEDESTRIAN LOADS and ASSIGNED PEDESTRIAN VOLUMES to start VIPER and view the assignment results. Click Proceed.

Figure 46
Module Selection Screen: View Assigned Volumes



C. Use VIPER to display the assigned 24-hour pedestrian volumes in field PEDVOL24. Zoom and post to the area of interest.

Figure 47
VIPER View of Pedestrian Network

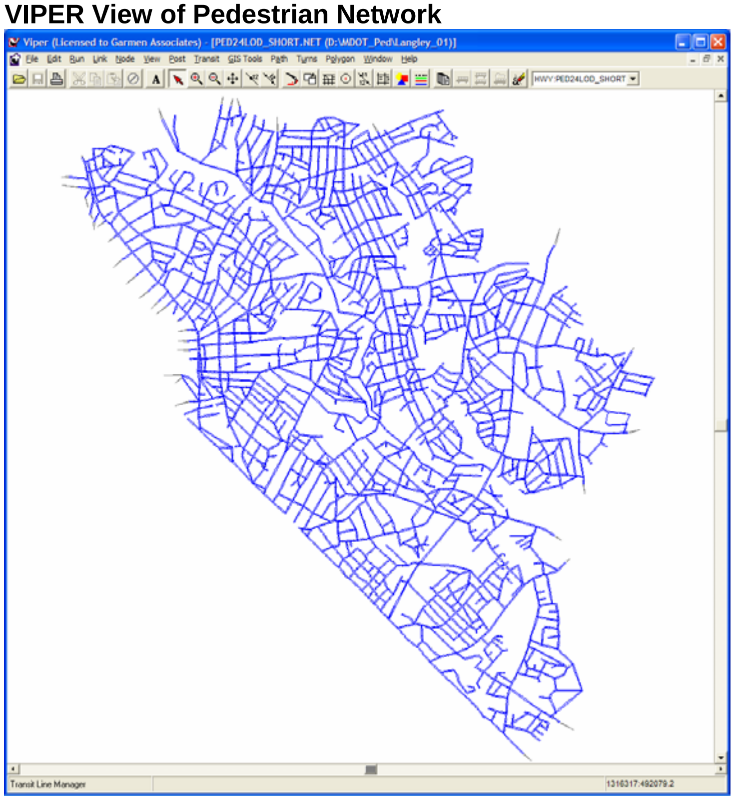
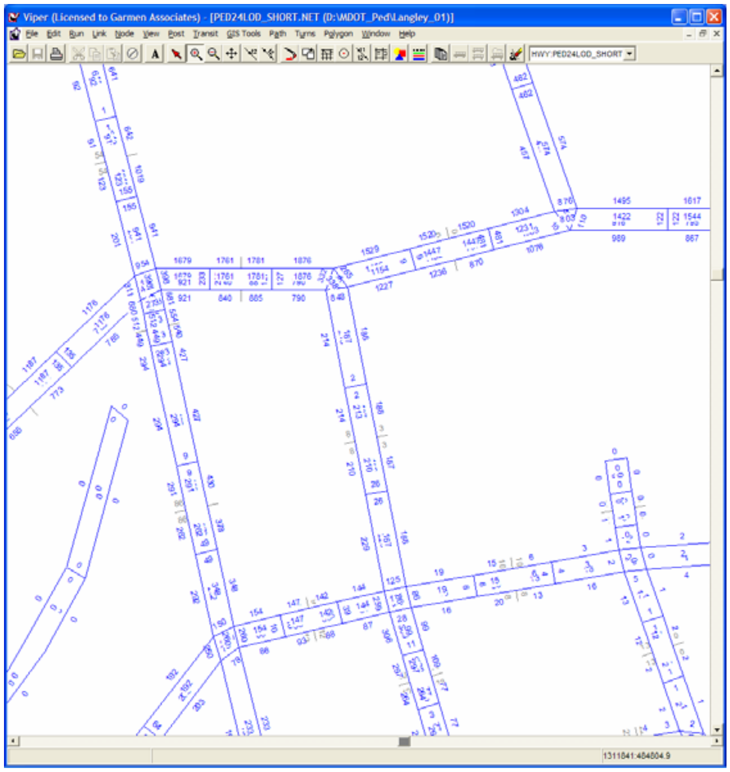


Figure 48
VIPER View of Pedestrian Volumes



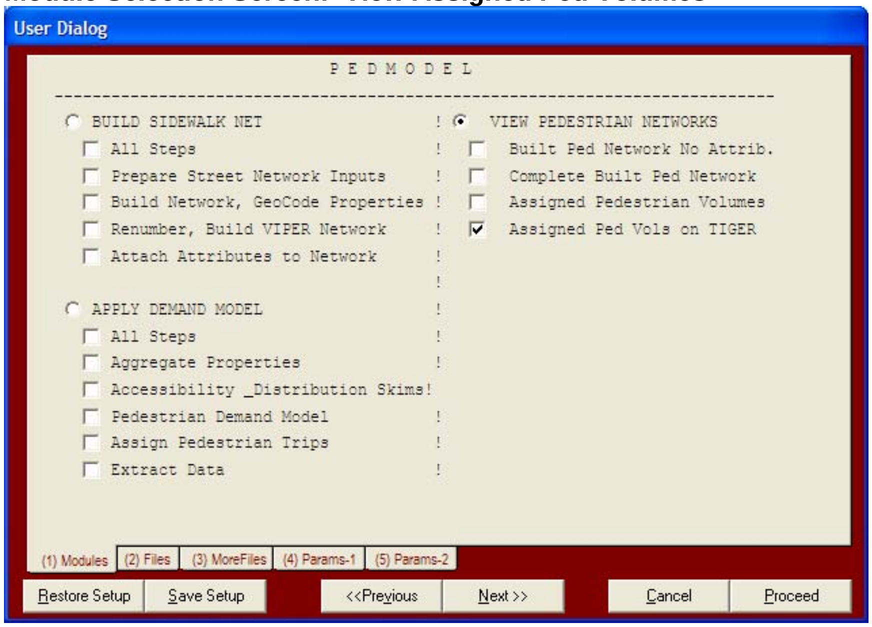
STEP 12: UNLOAD SHAPE FILES FROM VIPER

If the EXTRACT option has been selected then the detailed sidewalk network (PED24LOAD.NET) has been compressed to the less detailed TIGER link detail in GISLOADTIGER.NET, which is still a VIPER format file. Using VIPER, this network can then be converted to link and node shape files that can be viewed and analyzed in ArcGIS.

A. In CENTRAL, restart the model by clicking GO

B. Select VIEW PEDESTRIAN LOADS and ASSIGNED PED VOLS ON TIGER to start VIPER and view the compressed loaded network (see Figure 49). Click Proceed.

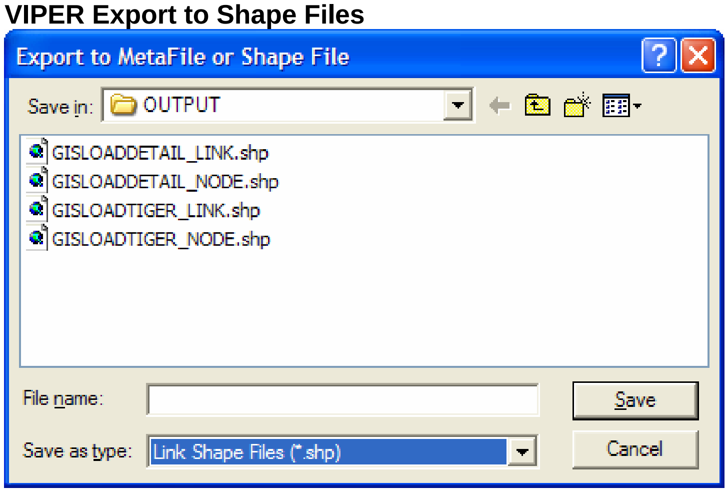
Figure 49
Module Selection Screen: View Assigned Ped Volumes



C. Check the data and topology to insure the compression was performed correctly.

D.. Export the shape files from VIPER. Click FILES then EXPORT. In SAVE AS TYPE, select LINK SHAPE FILES. Specify a name, click save, and a shape file for link data is produced.

Figure 50
VIPER Export to Shape Files



The following fields are in the LINK TIGER file:

TLINKID	Tiger line id number
FACTYPE	Street facility type
DISTANCE	Link length (miles)
PEDCNT24	Observed pedestrian count (24 hrs)
PEDCNTAM	Observed pedestrian count (am period)
PEDCNTMD	Observed pedestrian count (midday period)

PEDCNTPM	Observed pedestrian count (pm period)
SWVOL24	Assigned pedestrian volume on sidewalks (24 hrs)
JAYVOL24	Assigned pedestrian volume on jaywalks (24 hrs)
PRODS24	Pedestrian trips produced by land uses on this link (24 hrs)
ATTRS24	Pedestrian trips attracted to land uses on this link (24 hrs)

Repeat for NODE SHAPE FILES.

The following fields are in the NODE TIGER file:

TNODEID	Tiger node id number
NODETYP	Node Type: 6 = Intersection 7 = Link Dead-End
XWALKVOL24	Sum of assigned pedestrian volumes on all crosswalks within the node

KEY OUTPUT FILES

The PEDCONTEXT model produces a large number of intermediate and final output files. Key files that would be of use or interest to the user are located in the *scenario*\OUTPUT directory and include the following:

PED24LOAD.NET	Viper-format loaded network containing all assigned pedestrian volumes for the fully detailed network (sidewalks, crosswalks, etc.)
GISLOADTIGER.NET	Viper-format loaded network compressed to TIGER segment topology, with selected loaded volume and pedestrian count fields
GISLOADTIGER_LINK.xxx	Group of shape and data files exported from Viper, containing TIGER link load data. These files are manually exported and named in the last step as discussed above.
GISLOADTIGER_NODE.xxx	Group of shape and data files exported from Viper, containing TIGER node load data. These files are manually exported and named in the last step as discussed above.

CASE STUDIES

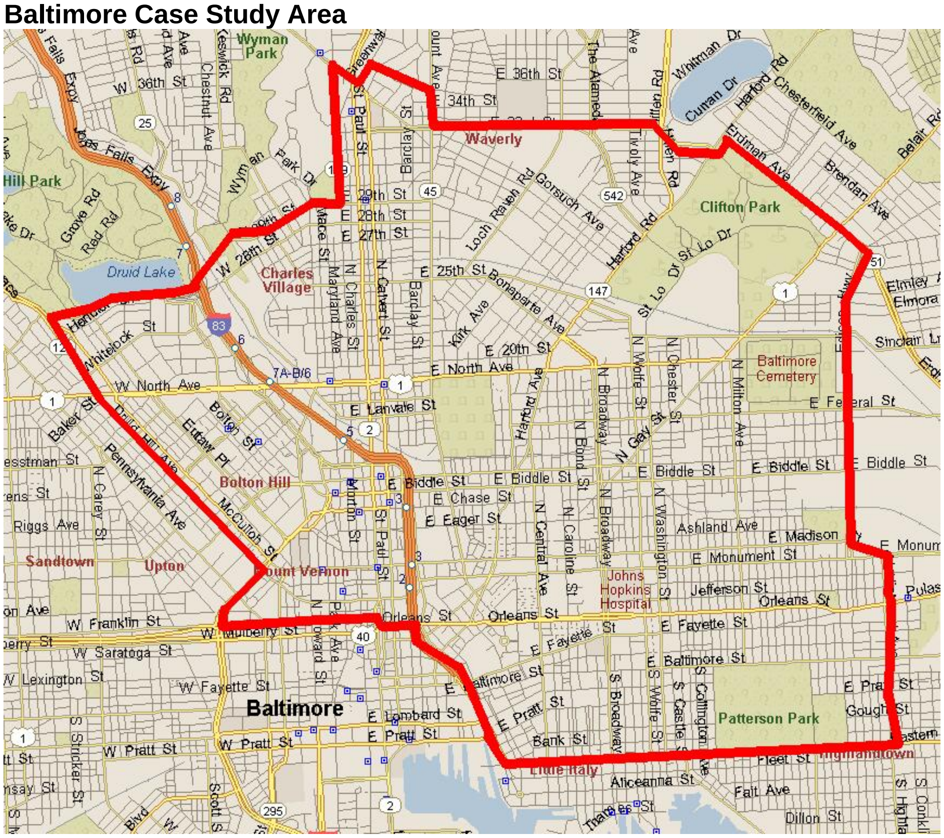
Two case study locations were selected as prototypes for development of the pedestrian flow model. These locations served both as test beds to test model functions and as demonstration areas to apply the model results to pedestrian safety problems.

The two locations were in the City of Baltimore, and the vicinity of Langley Park in Montgomery and Prince Georges Counties.

The Baltimore Case Study

The Baltimore case study encompassed a large portion of the City of Baltimore, as is shown in Figure 51. The study area included sections of the Mount Vernon, Bolton Hill, Charles Village, Patterson Park, and Johns Hopkins Hospital neighborhoods.

Figure 51
Baltimore Case Study Area



Altogether the study area encompasses about 10 square miles. The area has a population of about 97,000 peope, about 63,000 jobs, and contains about 46,000 individual properties. There is a great diversity of street types, ranging from freeways (I-83) down to local alleyways. Many streets are multi-lane with, in many cases, large park-like medians.

Using the tools of the PEDCONTEXT model the sidewalk network was built from TIGER street segments. A considerable amount of supplemental data was obtained and used to refine the default networks, including use of high resolution ortho-photos to obtain street geometrics, signal locations, and sidewalk / pathway details. A limited amount of field observation was also performed.

The resulting sidewalk network is illustrated in Figure 52.

Figure 52
Baltimore Sidewalk Network

The Baltimore portion of the Maryland Property View database was obtained, and 45,622 parcel records were processed. The resulting land use totals for the sturdy area are summarized in Table 17. The properties were aggregated to 6,050 block faces. Additional data was obtained from other sources including, from Census, block-group population, income, and household size data; and transit station locations and routes.

Figures 53, 54, and 55 show the locations of residential dwelling units, office floor space, and retail employment in the study area.

From this data it was estimated that the study area population is 97,323 persons. Total employment is 63,053 persons, of which 4,284 are retail employees and 58,769 are non-retail.

Table 17
Baltimore Study Area Land Use Activity

Land Use	Properties	Activity	
Residential:			
Apartments	40,377	44,826	du
Other	837	257	du
Subtotal	41,214	45,083	du
Commercial:			
Hotel	16	337,282	sf
Auto_Dlr	20	133,053	sf
Auto_PkLot	651	12,816	sf
Auto_Garag	36	7,597	sf
Auto_SvcSt	21	27,111	sf
Auto_Convn	2	1,181	sf
Auto_Other	146	2,387	sf
Rest_Fast	21	85,448	sf
Rest_Other	220	1,027,283	sf
Store_Dept	0	0	sf
Store_Othr	951	4,975,870	sf
Offc_Med	320	625,174	sf
Offc_Other	530	6,320,605	sf
Bank	29	268,640	sf
Warehouse	525	8,137,345	sf
Industrial	63	3,076,995	sf
Subtotal	3,551	25,038,785	sf
Recreation:			
Rec_PropSF	14	419,314	sf
Rec_PropAc	14	124	ac
Rec_LandAc	4	272	ac
Rec_Movie	8	145,890	sf
Rec_Museum	3	18,871	sf
Rec_Other	26	248,441	sf
Subtotal	69	832,912	sf
Community:			
Care_Hosp	34	3,896,865	sf
Care_DayCr	8	64,217	sf
Care_Other	6	192,151	sf
Com_PostOf	4	259,449	sf
Com_Church	217	1,433,750	sf
Com_School	52	21,366,040	sf
Com_Libr	2	99,106	sf
Subtotal	323	27,311,578	sf
Public:			
Safety	8	117,025	sf
Pub_Munic	114	2,147,330	sf
Pub_County	0	0	sf
Pub_State	16	1,562,530	sf
Pub_Fed	0	0	sf
Utilities	13	0	sf
Subtotal	151	3,826,885	sf

Figure 53
Residential Dwelling Unit Locations

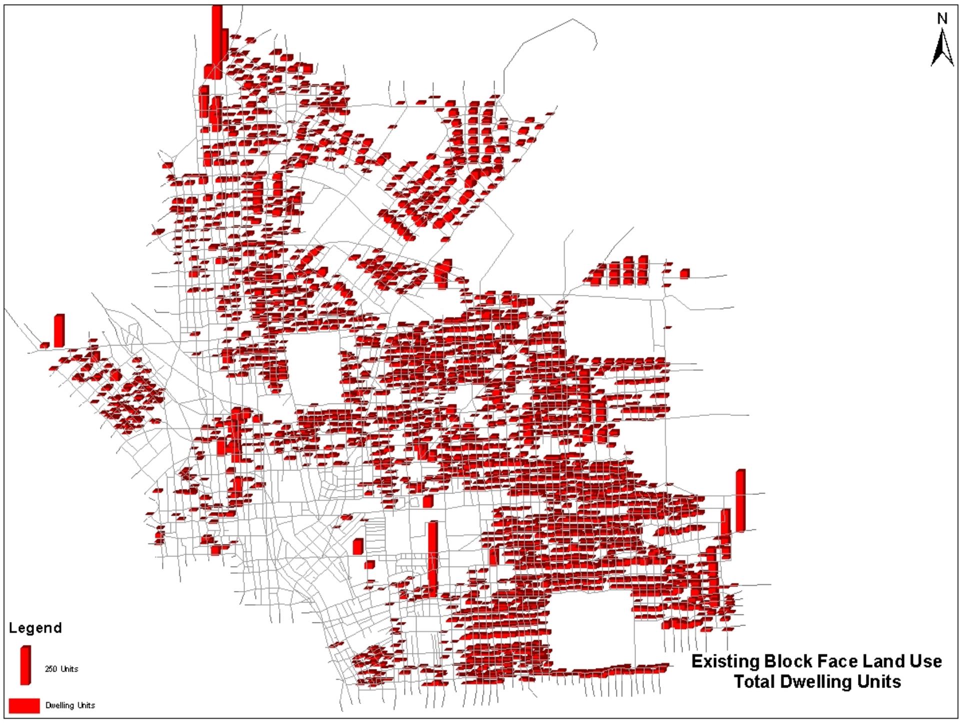


Figure 54
Office Floor Space Locations

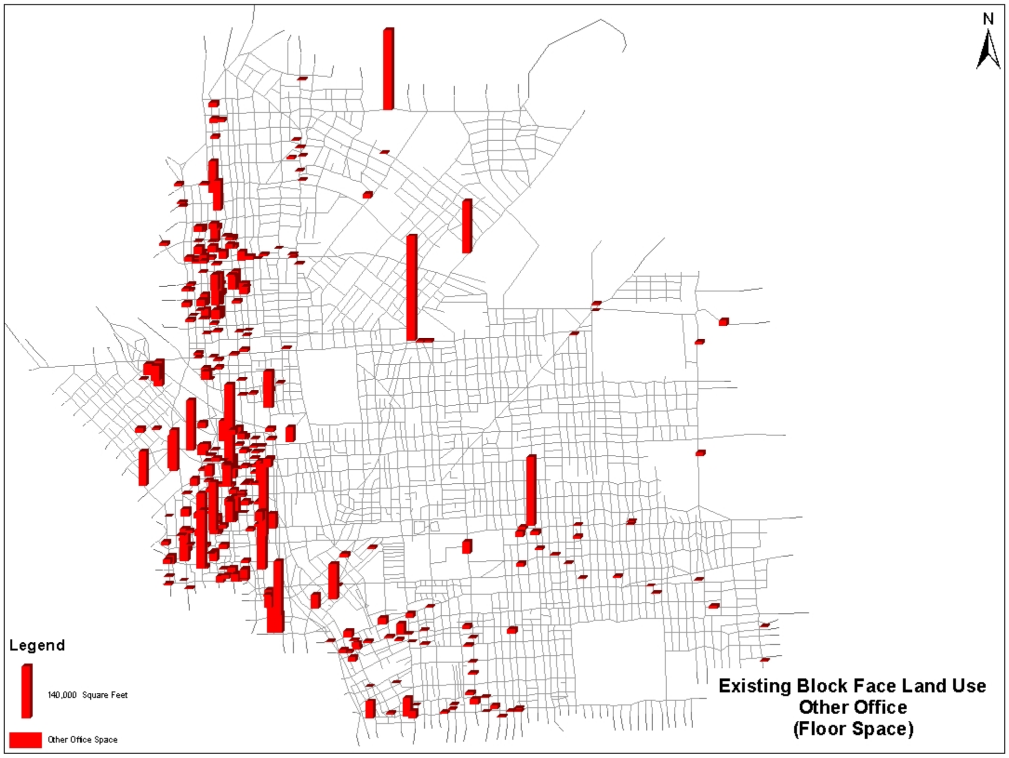
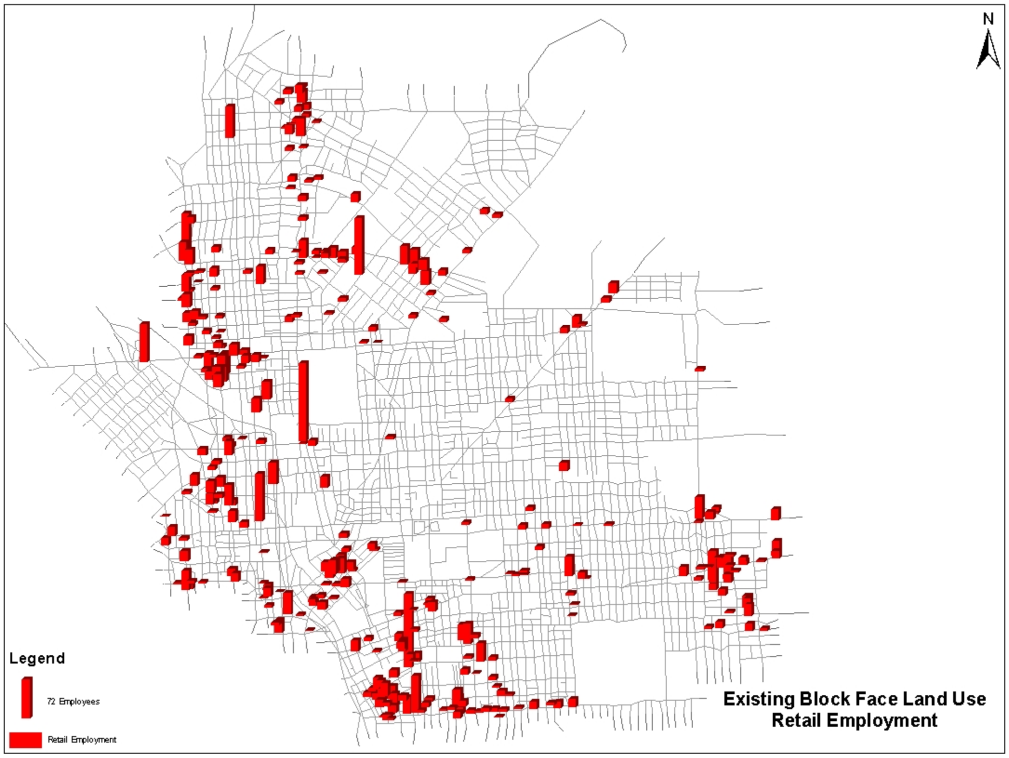


Figure 55
Retail Employment Locations



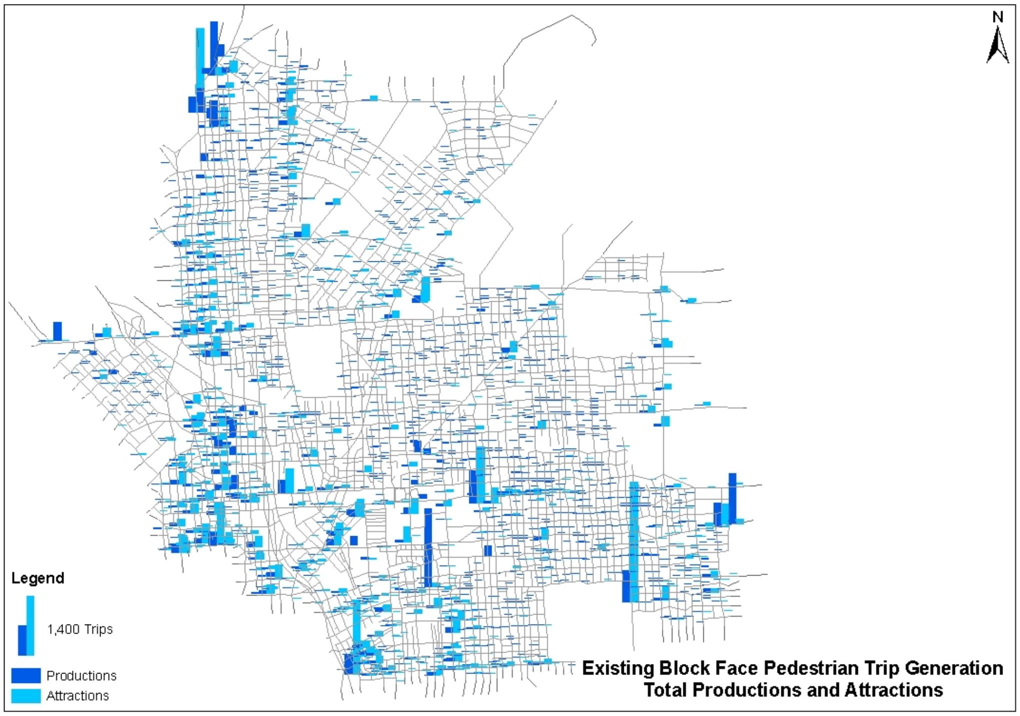
The trip generation model estimated that the above study area activities generate about 117,554 walk trips per day, as is shown in Table 18. Summing residential population with employment to give a rough indication of total study area activity (some 160,376 people), there are an estimated 0.73 daily walk trips per person in the study area.

Table 18
Daily Walk Trip Productions – Baltimore Study Area

	Daily Walk Trip Productions			
Trip Purpose	Home Base	Non Home Base	Total	Percent
Work	3,684	4,478	8,162	7%
Personal Business	25,646	5,548	31,194	27%
Eat Meal	14,778	5,576	20,354	17%
Shop	24,372	6,560	30,932	26%
Leisure	15,985	3,811	19,796	17%
School	5,448	1,668	7,116	6%
TOTAL	89,913	27,641	117,554	100%

The resulting daily walk trip productions and attractions are shown in Figure 56 for individual block faces. The intensity of walk trip generation varies considerably across the study area: In the residential areas to the east and north the number of trips generated on each block face is uniformly low, whereas to the south and west, where commercial and office uses predominate, walk trip levels are considerably higher.

Figure 56
Daily Walk Trip Productions and Attractions (By Block Face)



The matrix of daily pedestrian trips was assigned to the pedestrian network using the methods described in Section 3. The resulting sidewalk and intersection crosswalk volumes are illustrated in Figure 57 for the full study area, and in Figures 58 through 61 for more detailed quadrants of the study area.

Figure 57
Assigned Daily Pedestrian Volumes - Baltimore

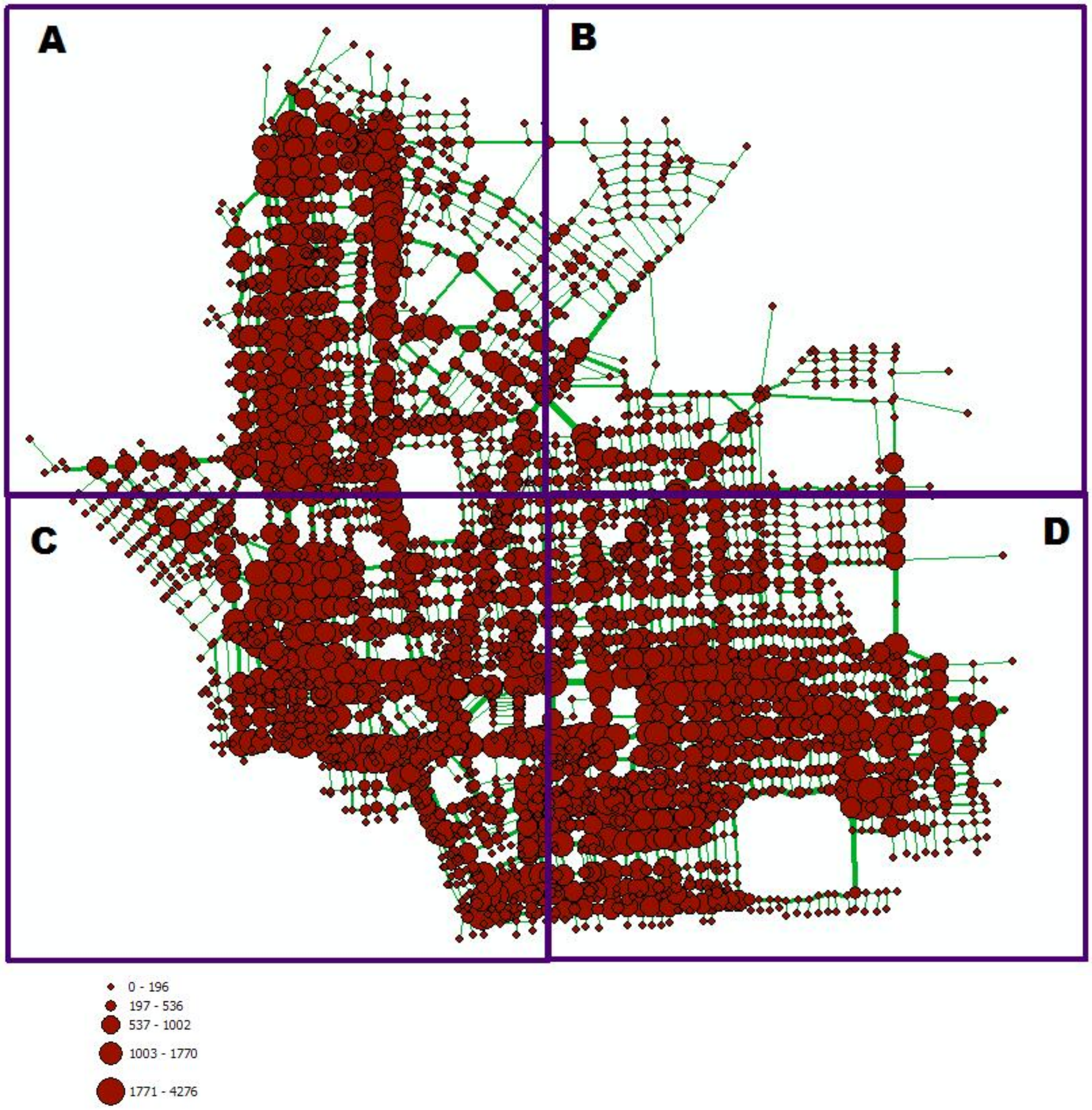


Figure 58
Assigned Daily Pedestrian Volumes - Baltimore
PANEL A -- NORTHWEST

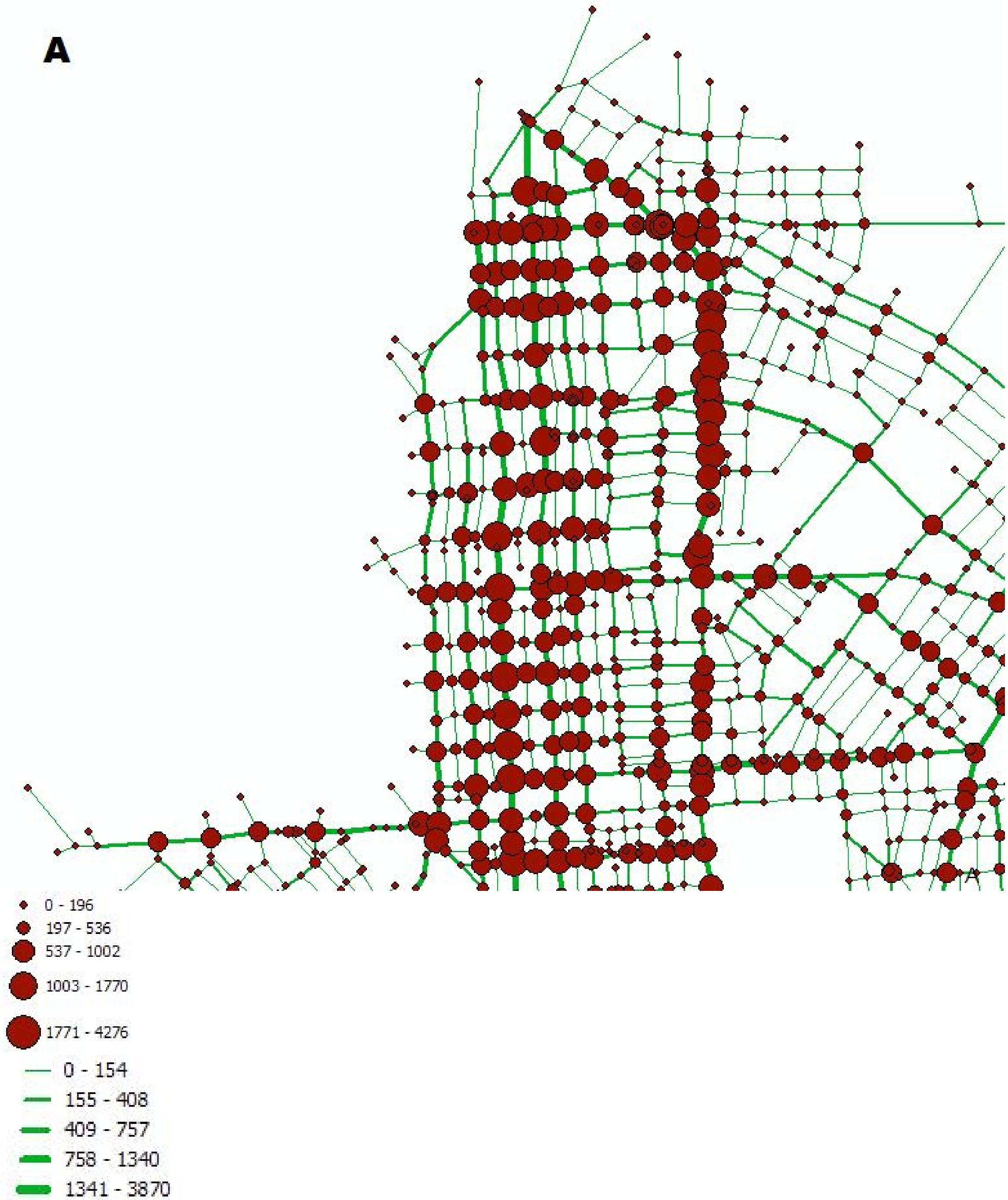


Figure 59
Assigned Daily Pedestrian Volumes - Baltimore
PANEL B -- NORTHEAST

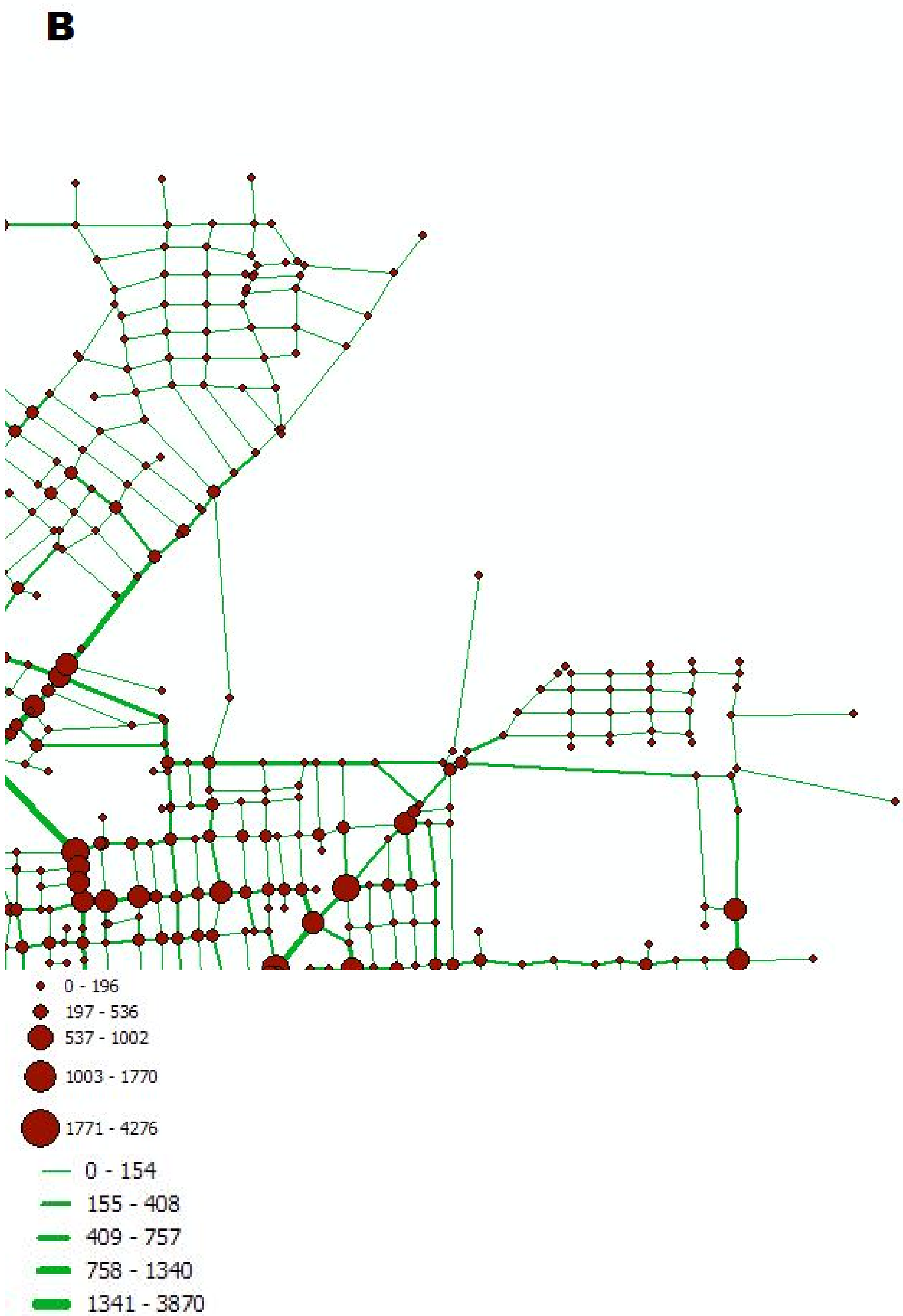


Figure 60
Assigned Daily Pedestrian Volumes - Baltimore
PANEL C -- SOUTHWEST

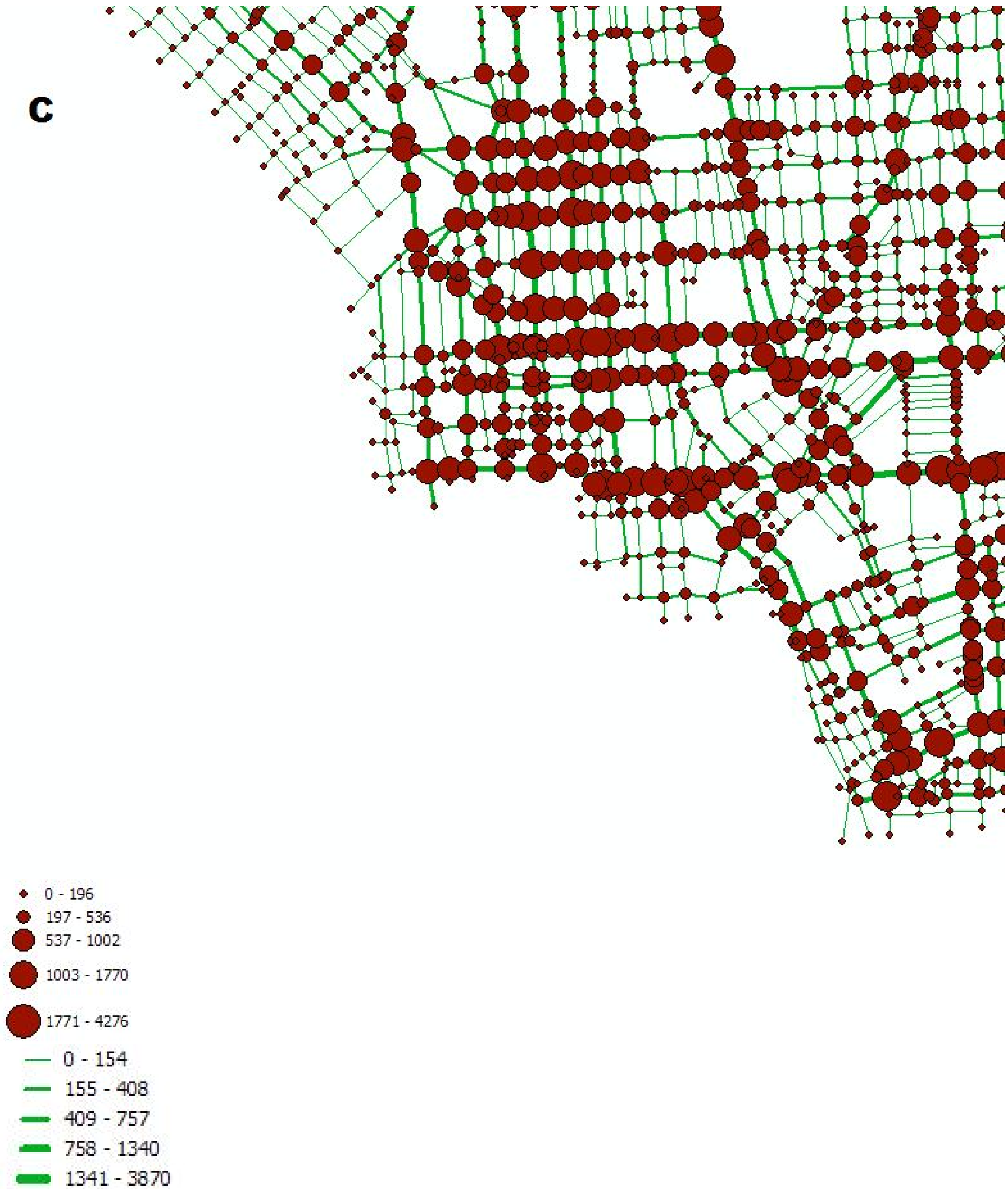


Figure 61
Assigned Daily Pedestrian Volumes - Baltimore
PANEL D -- SOUTHEAST

The pedestrian volumes that the model estimated were compared with volumes counted at various locations in the study area to obtain a sense of how well the overall model performs. It must be noted, however, that such comparisons are difficult to make and may be misleading. Obviously the model has inaccuracies, since it is but a representation of actual conditions. Pedestrian counts also have substantial inaccuracies, though, and they must be used carefully.

In the instance of this Baltimore study area, the pedestrian counts were provided for morning, midday, and evening peak periods. All counts were conducted at intersection crosswalks; there were no sidewalk counts provided. Analysis of the counts indicated a number of locations where the count quality was suspect: Directional patterns in one period were not paralleled by similar (or reverse) patterns in another period, for example. Nonetheless the counts were retained and used for the following comparison.

In addition, it is important to note that the counts were made only for specific periods, not for the whole 24-hour day. The model, by contrast, estimates 24-hour pedestrian flows. The counts were expanded to 24-hour totals using factors derived from the NYMTC survey data, but this is acknowledged to be an inaccurate process at best.

Nonetheless the following statistics were compiled:

The ratio of estimated to observed pedestrian volumes was 104%. In other words, the overall pedestrian volume estimated by the model at counted locations was only 4% higher than the sum of the counts at the same locations. This was taken to be a strong indication of the model's accuracy.

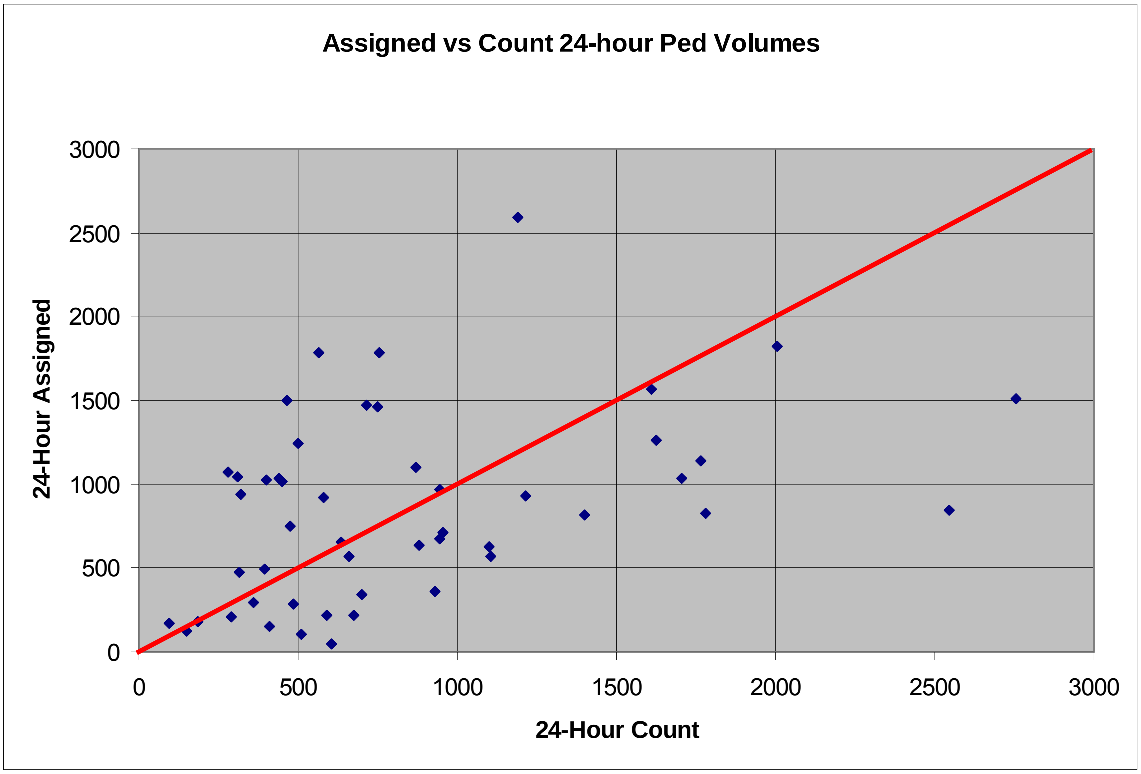
The Percent Root Mean Square Error was 96%. This indicates the average error and is substantially higher than one would wish. But, allowing for the above problems with the count database, it was decided that this is an acceptable showing.

Finally the counted vs. assigned pedestrian volumes were graphed, with the result shown in Figure 62. This graph indicates a reasonable comparison between the model estimate and observed pedestrian volumes.

This evaluation of the Baltimore case study led to several conclusions with respect to the usefulness of the PEDCONTEXT modeling framework:

- The PEDCONTEXT model construct is viable. It is based on extensions of traditional travel demand estimation techniques, it uses available software and hardware platforms, it is founded on data that is obtainable at reasonable effort, and it produces reasonable results.
- The detailed data needed for the model can be assembled from available sources in Maryland with reasonable effort. Similar data sources are available in other states, so the model can be transferred to other jurisdictions with reasonable efforts.
- The Model is sensitive to real-world factors that do affect pedestrian travel, such as land use, physical sidewalk network connectivity and quality, and the barrier effects of street crossings.
- Output from the model can be used for a variety of planning and operations functions.

Figure 62
Comparison of Assigned vs Counted Pedestrian Volumes

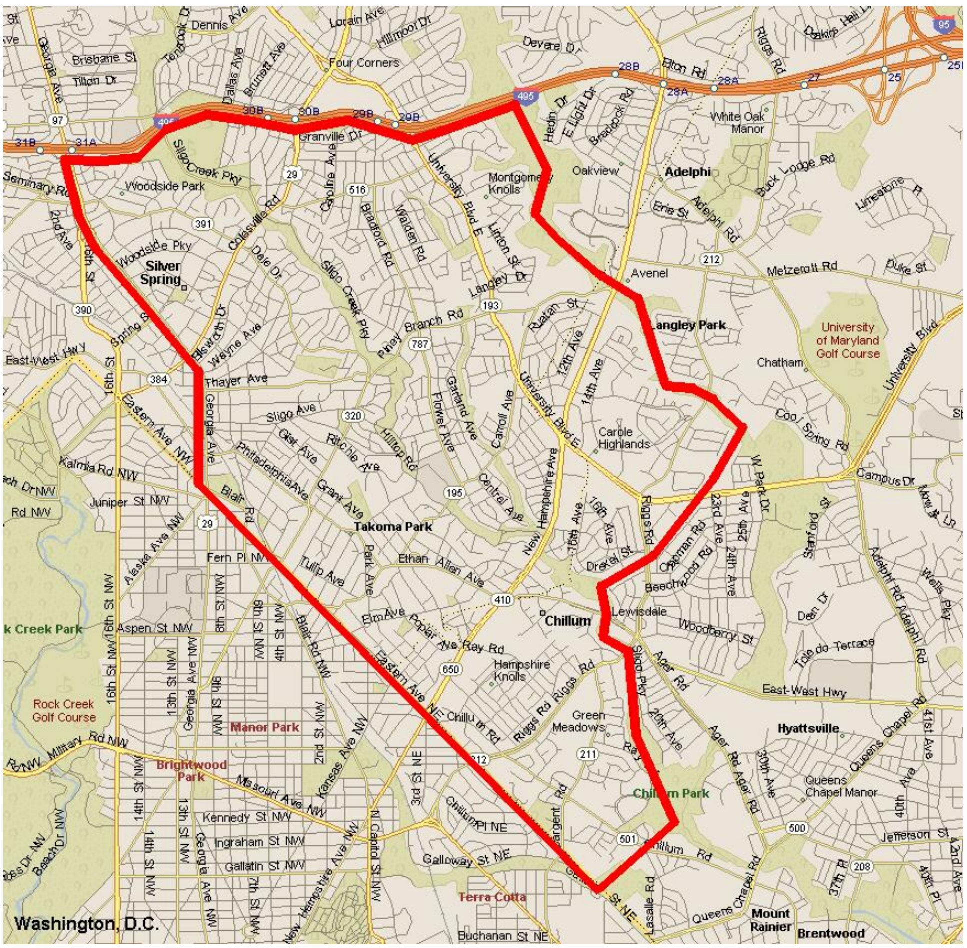


The Langley Park Case Study

A second case study was undertaken, covering the vicinity of Langley Park in Montgomery and Prince Georges Counties. Unlike the Baltimore case study which provided the test bed for developing the PEDCONTEXT model, and therefore was thoroughly vetted with respect to data quality, the objective of the Langley Park case study was to determine if reasonable results could be obtained with minimal effort by simply transferring the model and running it with new local data.

The study area is shown in Figure 63. It includes portions of Silver Spring, Takoma Park, Chillum, and Langley Park. The coverage of this study area is slightly larger than the Baltimore study area: about 13 square miles, as compare to 10 square miles for Baltimore. The area is considerably less intensely developed, however. It contains a residential population of only about 66,000 persons, about 17,000 employees, and about 16,100 properties.

Figure 63
Langley Park Case Study Area



Using the tools of PEDCONTEXT the sidewalk network was built from TIGER street segments. In contrast to the Baltimore case study, supplemental data was not available. High resolution ortho photography could not be obtained, and field investigations were not conducted. Instead, this case study demonstrates that reasonable results can be obtained from available off-the-shelf data sources.

The resulting sidewalk network is illustrated in Figure 64.

The Montgomery and Prince Georges County portions of the Maryland Property View database was obtained and merged, and 16,142 parcel records were processed. The resulting land use totals for the study area are summarized in Table 19. The properties were aggregated to 3,550 block faces.

Figure 64
Langley Park Sidewalk Network

It should be noted that the property data for Langley Park does not appear to be as complete as was the Baltimore data. Public and community properties such as parks, churches, post offices, libraries, and municipal / county offices appear to be under-represented in the data file. This is understandable since the primary use of the Property View files is for property taxation. However for this pedestrian modeling purpose, these public uses attract leisure trips and, lacking them, leisure trips will be under estimated.

From this data it was estimated that the study area population is 66,125 persons. Total employment is 17,150 persons, of which 995 are retail employees and 16,155 are non-retail.

Table 19
Langley Park Study Area Land Use Activity

Land Use	Properties	Activity	
Residential:			
Apartments	756	16,917	du
Other	14,822	14,440	du
Subtotal	15,578	31,357	du
Commercial:			
Hotel	2	238,266	af
Auto_Dlr	2	19,325	sf
Auto_PkLot	24	1,262	sf
Auto_Garag	1	658	sf
Auto_SvcSt	8	21,887	sf
Auto_Convn	1	1,343	sf
Auto_Other	44	4,543	sf
Rest_Fast	6	18,269	sf
Rest_Other	30	96,750	sf
Store_Dept	1	99,200	sf
Store_Othr	214	2,669,701	sf
Offc_Med	30	21,239	sf
Offc_Other	153	3,611,482	sf
Bank	8	38,578	sf
Warehouse	21	323,118	sf
Industrial	4	136,580	sf
Subtotal	549	7,302,201	sf
Recreation:			
Rec_PropSF	0	0	sf
Rec_PropAc	0	0	ac
Rec_LandAc	0	0	ac
Rec_Movie	0	0	sf
Rec_Museum	0	0	sf
Rec_Other	2	17,176	sf
Subtotal	2	17,176	sf
Community:			
Care_Hosp	1	314,266	sf
Care_DayCr	1	14,270	sf
Care_Other	3	132,901	sf
Com_PostOf	1	2,520	sf
Com_Church	0	0	sf
Com_School	1	7,210	sf
Com_Libr	0	0	sf
Subtotal	7	471,167	sf
Public:			
Safety	1	6,839	sf
Pub_Munic	0	0	sf
Pub_County	0	0	sf
Pub_State	0	0	sf
Pub_Fed	0	0	sf
Utilities	5	0	sf
Subtotal	6	6,839	sf

The trip generation model estimated that the above study area activities generate about 51,400 walk trips per day, as is shown in Table 20. Summing residential population with employment to produce a rough estimate of total study area activity yields 83,275 persons. The estimated daily walk trip generation of 51,400 trips is equivalent to 0.62 daily walk trips per person. This is a significantly lower rate than the 0.73 trips per person estimated for Baltimore.

Table 20
Daily Walk Trip Productions – Langley Park Study Area

	Daily Walk Trip Productions			
Trip Purpose	Home Base	Non Home Base	Total	Percent
Work	1,367	802	2,169	4%
Personal Business	11,697	1,274	12,971	25%
Eat Meal	6,785	670	7,455	14%
Shop	13,445	3,163	16,608	32%
Leisure	7,958	843	8,801	18%
School	2.624	772	3,396	7%
TOTAL	43,876	7,524	51,400	100%

Other differences between Langley Park and Baltimore are apparent from this table (refer to Table 18). The number of walking work trips is substantially lower in Langley Park than Baltimore – 2,169 vs. 8,162. The number of walk trips for personal business are also lower in Langley Park – 12,971 in Langley Park vs. 31,194 in Baltimore. There are also significantly fewer Shopping walk trips in Langley Park than Baltimore – 16,608 vs. 30,932. The less dense housing, higher incomes, and greater walking distances to shopping and other attractions are responsible for these differences.

The matrix of daily pedestrian trips was assigned to the pedestrian network using the methods described in Section 3. The resulting sidewalk and intersection crosswalk volumes are illustrated in Figure 65 for the full study area, and in Figures 66 and 67 for more detailed quadrants of the study area.

Figure 65
Assigned Daily Pedestrian Volumes – Langley Park

Figure 66
Assigned Daily Pedestrian Volumes – Langley Park
PANEL A - NORTH

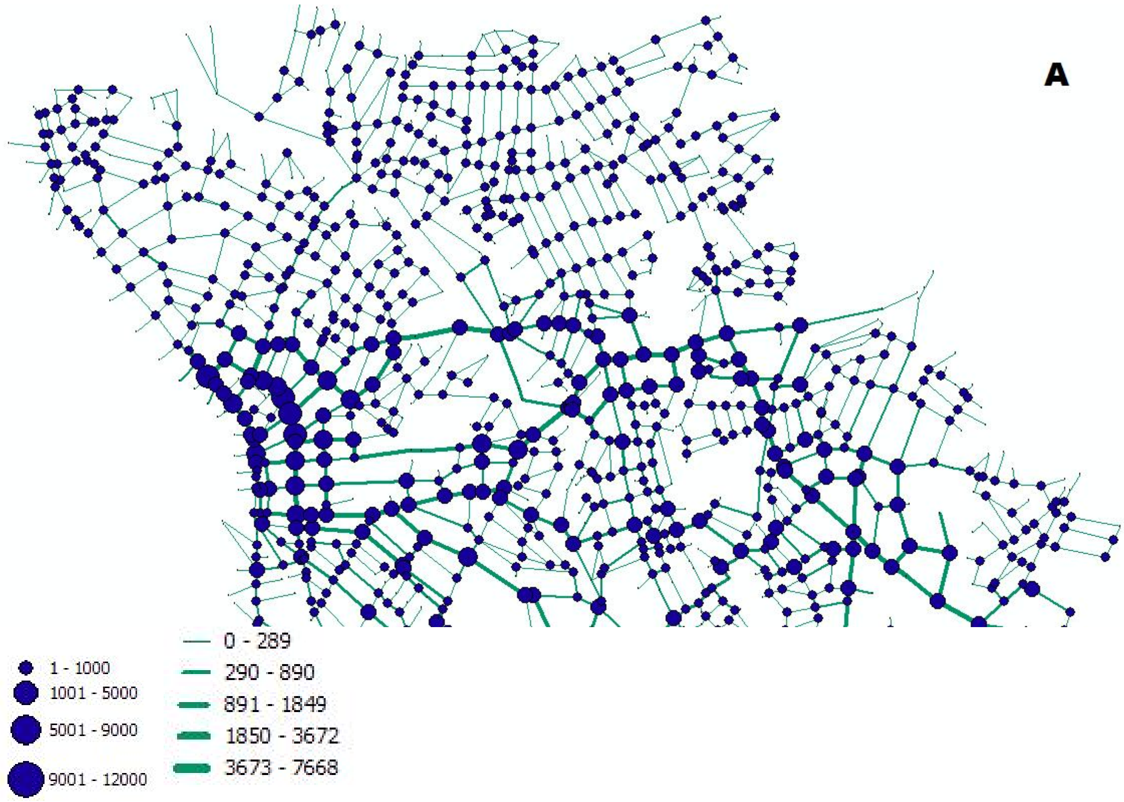
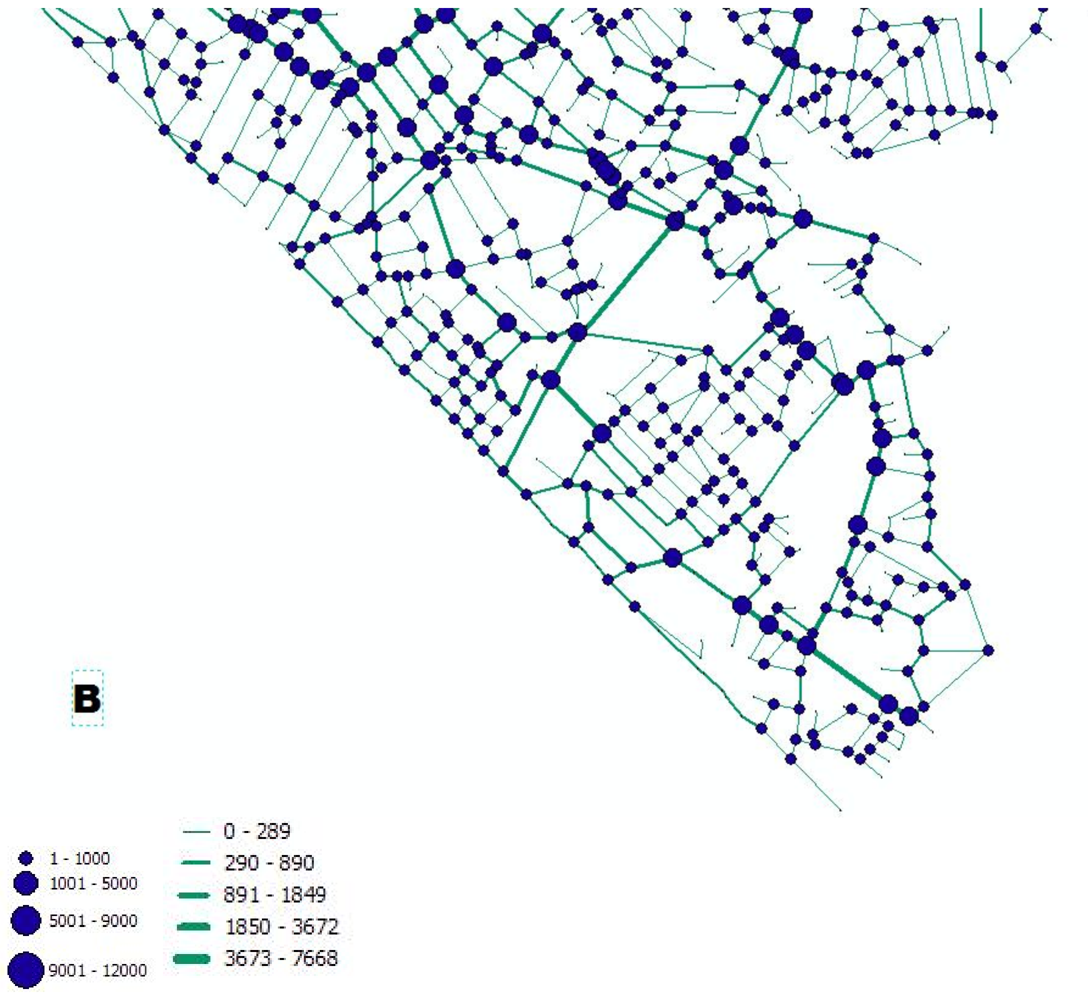


Figure 67
Assigned Daily Pedestrian Volumes – Langley Park
PANEL B - SOUTH



SAFETY ANALYSIS

Pedestrian safety is emerging as a major area of concern for MPO's and planning agencies. Typically, pedestrian safety has been analyzed by either examining the absolute number of pedestrian crashes at a location, or computing an exposure rate from the number of crashes and the traffic volume. A more desirable measure would be an exposure rate based on the pedestrian volume, but it has not proven feasible to obtain pedestrian flow volumes on a wide-area basis to support this type of analysis.

The implementation of this PEDCONTEXT pedestrian flow model now permits the pedestrian crash exposure rate to calculated, based on the pedestrian volume. This Section reports the result of a safety analysis for the Baltimore and Langley Park case study areas, using the daily pedestrian volumes estimated by the model as described in Section 5.

Geocoded pedestrian crash data was provided for the case study areas by the Maryland Department of Transportation, Division of Highway Safety Programs. The data covered three years – 2000, 2001, and 2002.

Baltimore Case Study Area

In the Baltimore case study area a total of 876 crashes involving pedestrians were recorded during the three years 2000 to 2002. These crashes occurred at 493 distinct locations. It should be noted that the geo-coding of all crashes was such that they appear to have occurred exclusively at intersections. This is likely not the case in actuality, since inevitably some pedestrians are likely to have been struck at mid-block locations when jay walking, or at other locations. Given this geo-coding convention, all crash rates were computed on the basis of total intersection crosswalk pedestrian volumes.

Figure 68 shows the location of the 876 pedestrian crashes, with the size of the dot proportional to the number of crashes during the three-year period. Table 21 lists the locations with three or more crashes in the three-year period, ranked by number of crashes consistent with the data presented in Figure 68.

A typical practice is to weight crashes by their severity, so that those with a higher extent of injuries or fatalities essentially count for more. A code describing the injury-severity of each crash was provided in the data, with a range from 1 (less severe) to 5 (more sever). Each crash was weighted by the severity value, and the severity-weighted number of crashes was accumulated at each location. Figure 69 shows the result of this, displaying the number of severity-weighted crashes at each intersection. Interestingly the severity weighting does not appear to significantly affect the ranking of crash locations.

Table 22 lists the crash locations with three or more crashes, ranked according to their severity, consistent with the data presented in Figure 69.

Figure 68
Pedestrian Crash Locations – Baltimore Case Study Area
(By Number of Crashes)

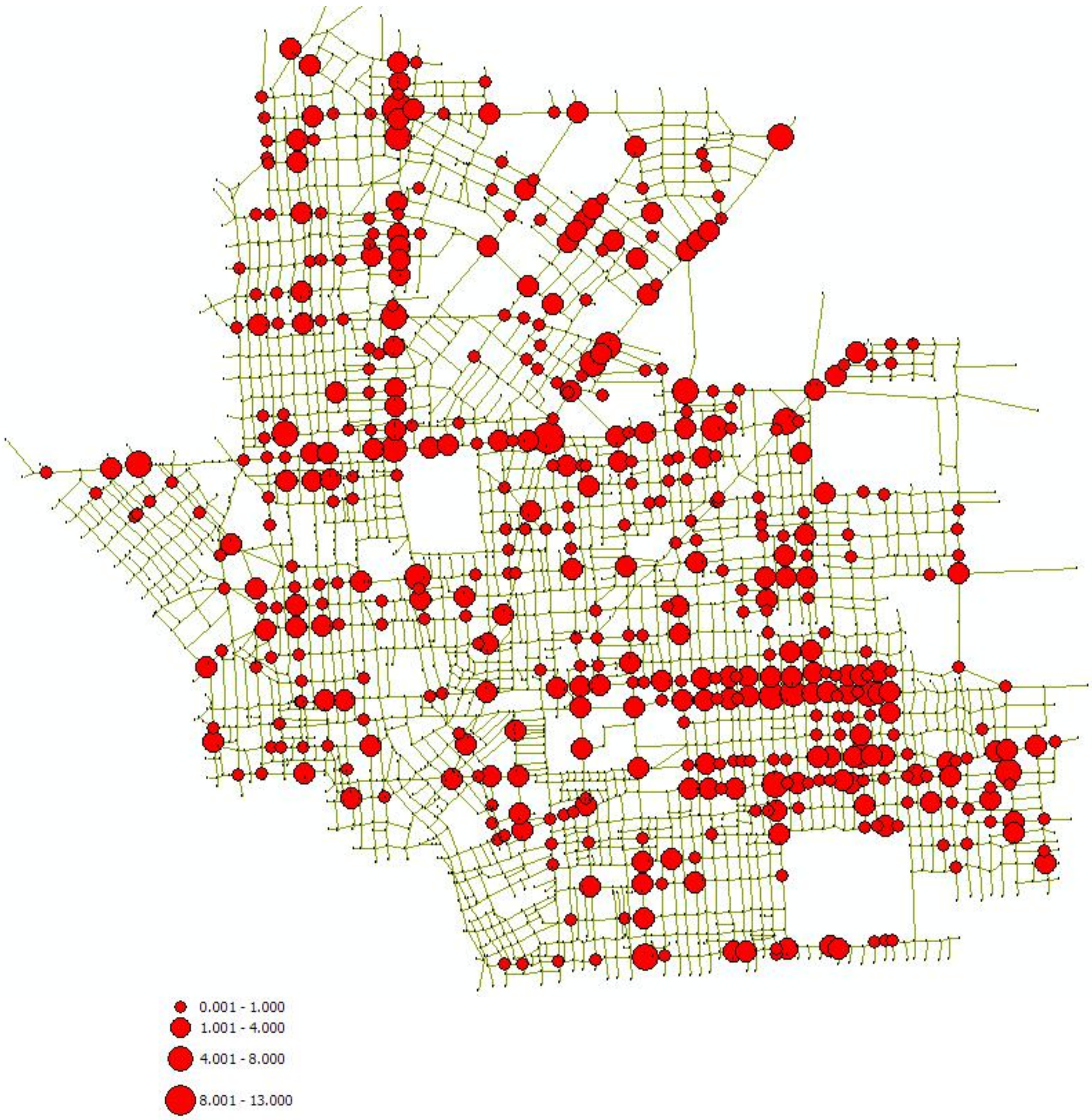


TABLE 21: RANKED BY COUNT

Figure 69
Pedestrian Crash Locations – Baltimore Case Study Area
(By Severity-Weighted Number of Crashes)

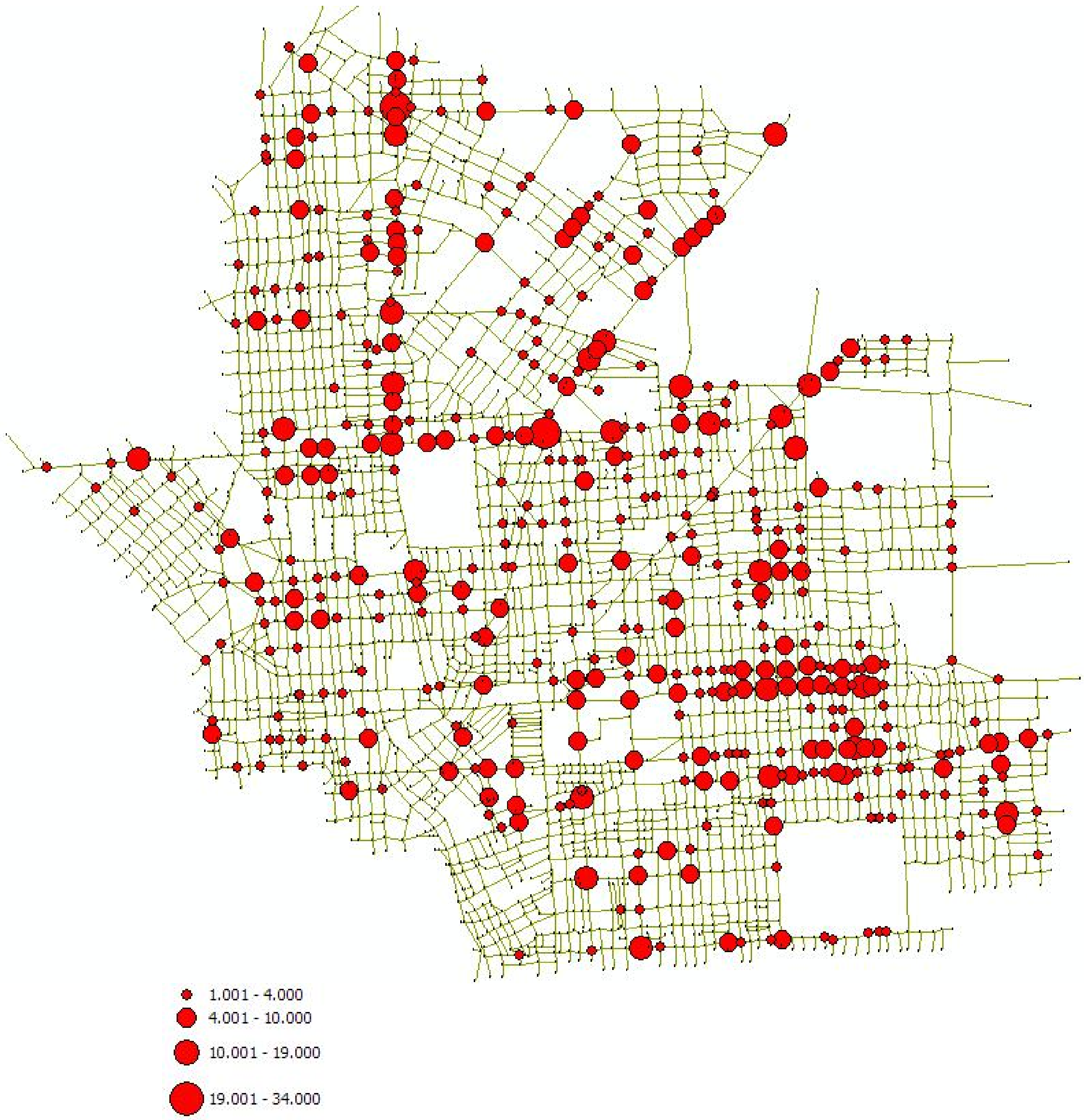


TABLE 22: RANKED BY SEVERITY WEIGHTED COUNT

The PEDCONTEXT model provides daily pedestrian volumes on sidewalks and intersection crosswalks. Since the pedestrian crashes were geo-coded to intersections, the number of daily pedestrians using all the crosswalks within an intersection were summed to produce a daily pedestrian total. These data were displayed graphically in Figures 57 through 61 for the Baltimore case study area.

Since the crash data represents three years of data, daily (weekday) pedestrian volumes were expanded by multiplying by 365 x 3. (The model represents a typical weekday. Lacking available pedestrian count data for weekends on a sufficiently large sample, it is assumed that weekends have approximately the same pedestrian activity as weekdays. Therefore multiplying by 365 is assumed to convert the daily estimate to an annual estimate.)

Figure 70 and Table 23 show the priority locations, ranked by exposure rate. The exposure rate is expressed in terms of pedestrian crashes per million pedestrians at the subject location.

Figure 71 and Table 24 show the same priority locations ranked by exposure rate, but with the rate calculated on the basis of severity-weighted crashes.

Figure 70
Pedestrian Crash Locations – Baltimore Case Study Area
(By Exposure Rate)

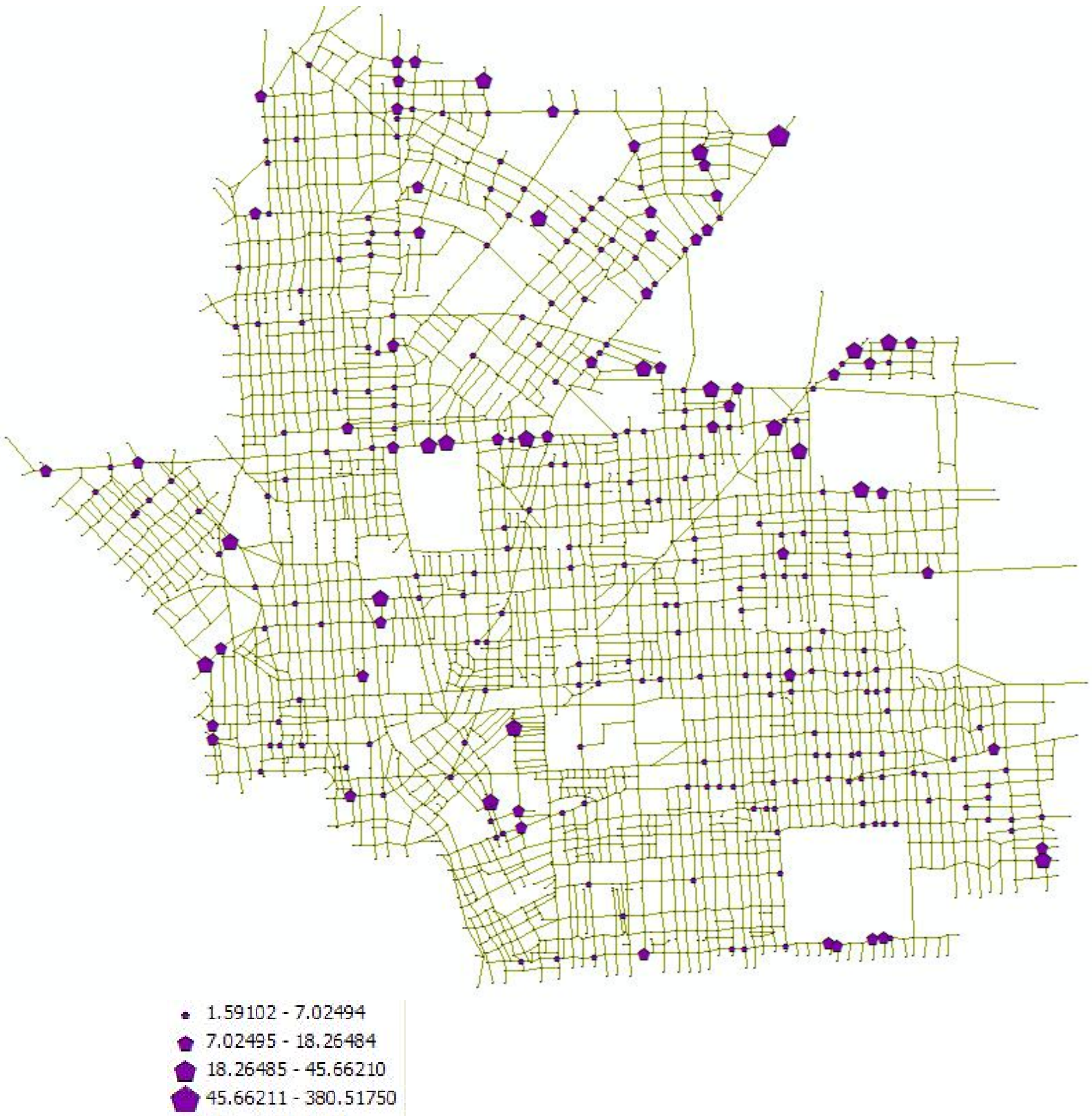


Table 23: RANKED BY RATE

Figure 71
Pedestrian Crash Locations – Baltimore Case Study Area
(By Severity-Weighted Exposure Rate)

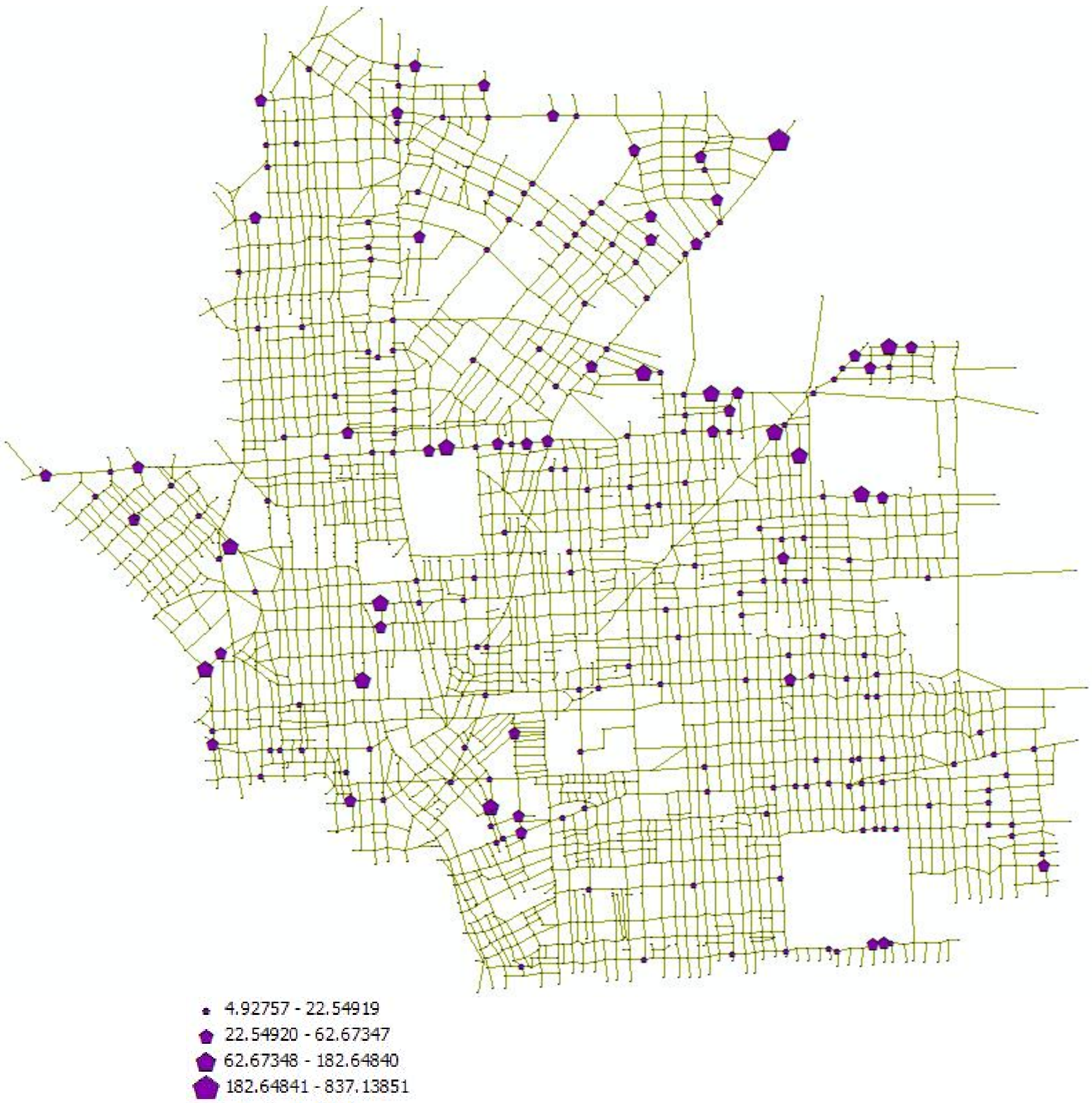


Table 24: RANKED BY SEVERITY-WEIGHTED RATE

Langley Park Case Study Area

A similar crash exposure analysis was conducted for the Langley Park case study area, using crash data for the three years 2000 to 2002, and the daily pedestrian volumes estimated by the PEDCONTEXT model.

Within the Langley Park case study area there were 161 pedestrian crashes at 90 locations.

Figure 72 and Table 25 show the 90 crash locations in the study area ranked according to the number of crashes that occurred during the three-year observation period.

Figure 73 and Table 26 show the same information, but ranked according the number of severity-weighted crashes.

Using pedestrian volumes estimated by the PEDCONTEXT model for the Langley Park case study area (refer to Figures 65 through 67), a pedestrian crash exposure rate was computed for each of the 90 crash locations. This exposure rate is expressed in terms of crashes per million pedestrians crossing at the subject intersection. The results are shown in Figure 74 and Table 27 for a rate based on the unweighted count of crashes.

Crash rates based on severity-weighted crash counts are shown in Figure 75 and Table 28.

Figure 72
Pedestrian Crash Locations – Langley Park Case Study Area
(By Number of Crashes)

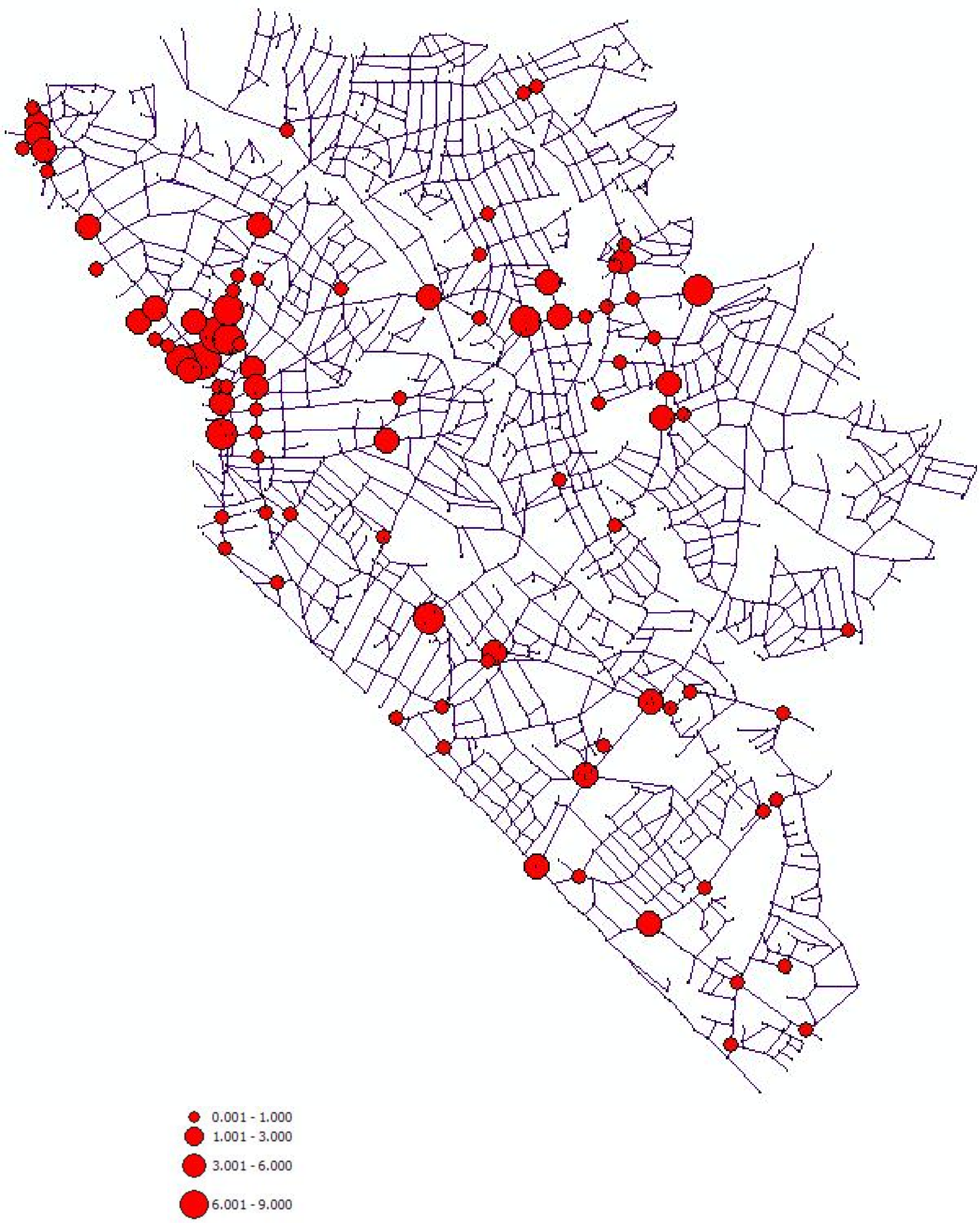


Table 25 - RANKED BY COUNT

Figure 73
Pedestrian Crash Locations – Langley Park Case Study Area
(By Severity-Weighted Number of Crashes)

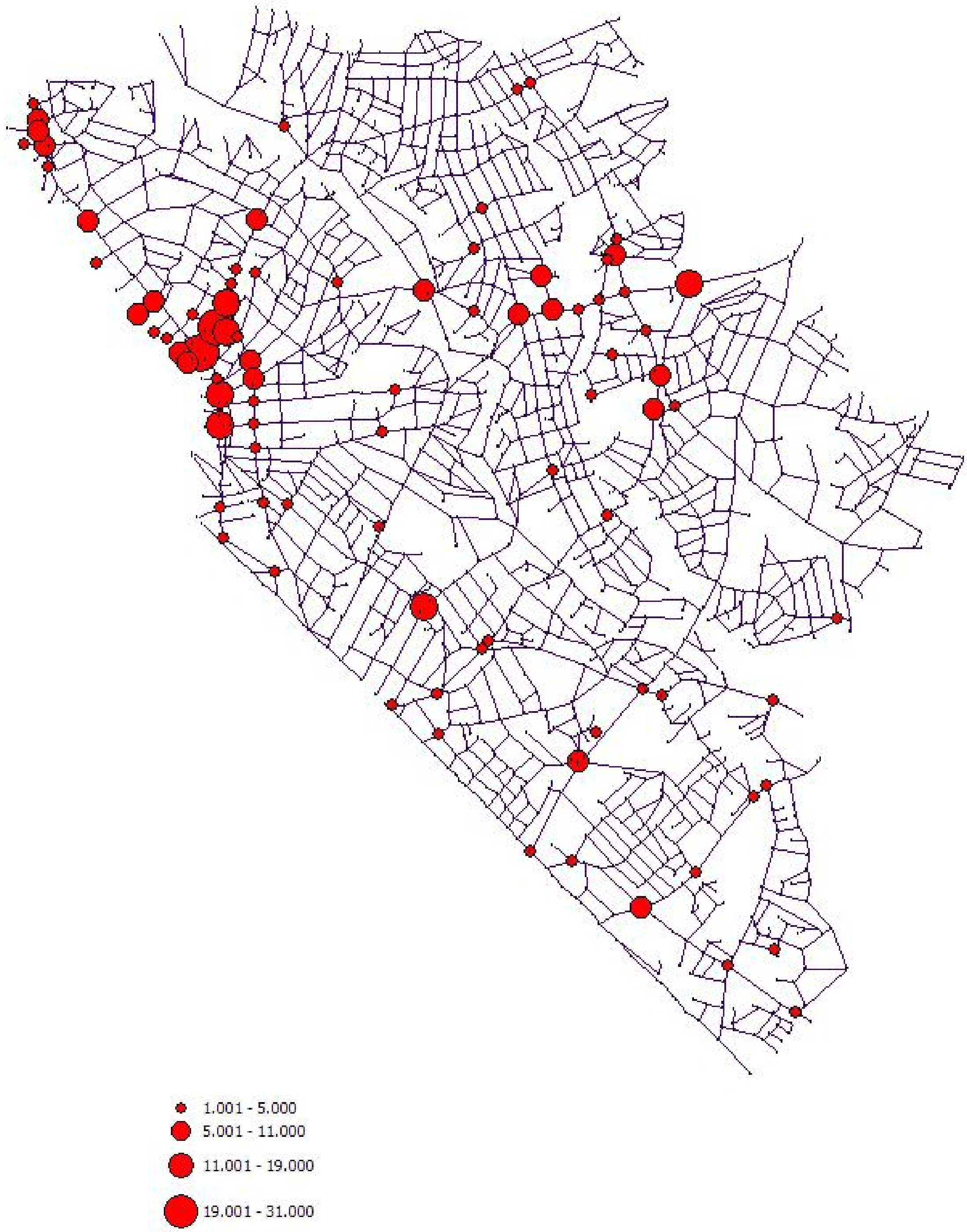


Table 26 - RANKED BY SEVERITY WEIGHTED COUNT

Figure 74
Pedestrian Crash Locations – Langley Park Case Study Area
(By Exposure Rate)

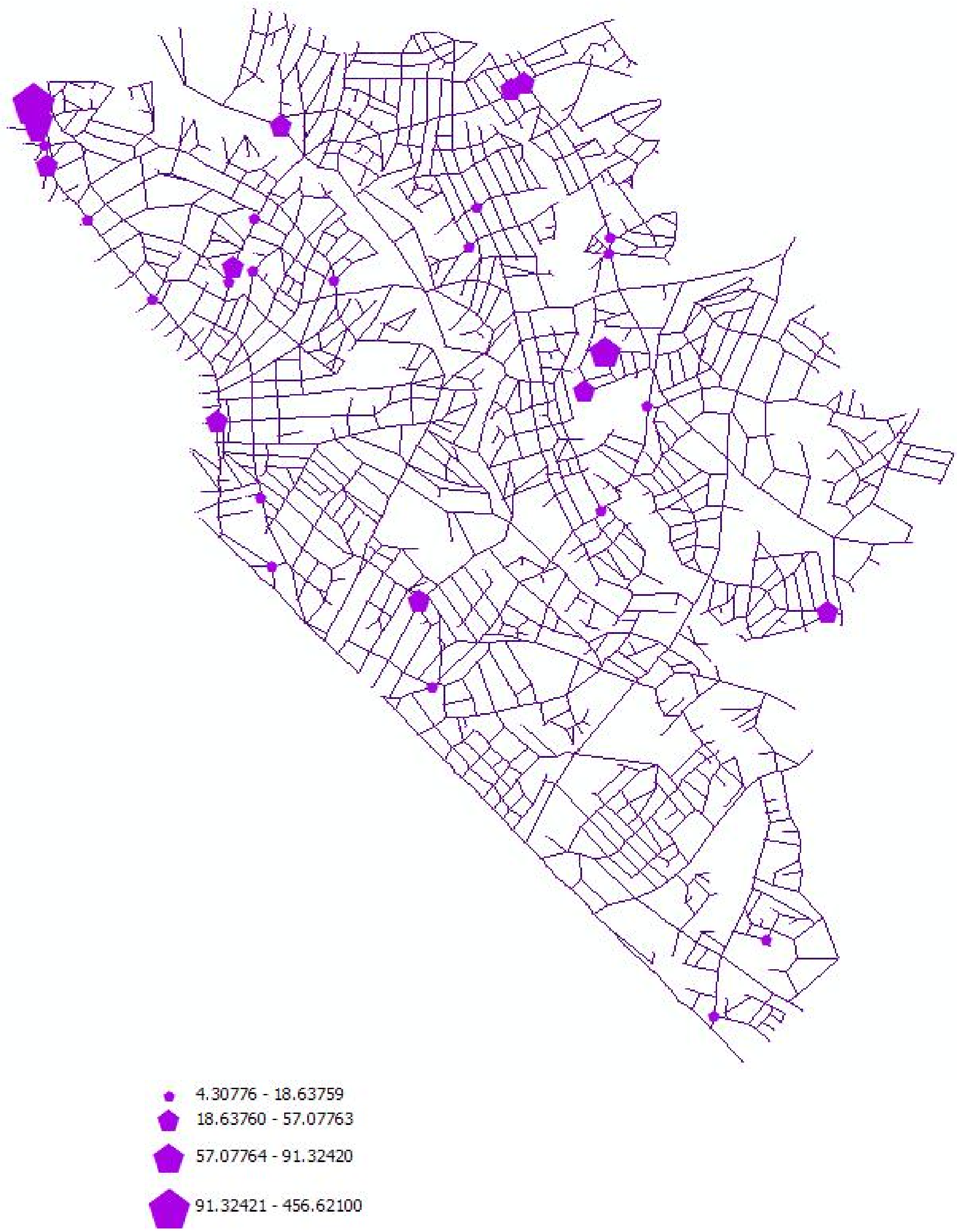


Table 27 -- RANKED BY RATE

Figure 75
Pedestrian Crash Locations – Langley Park Case Study Area
(By Severity-Weighted Exposure Rate)

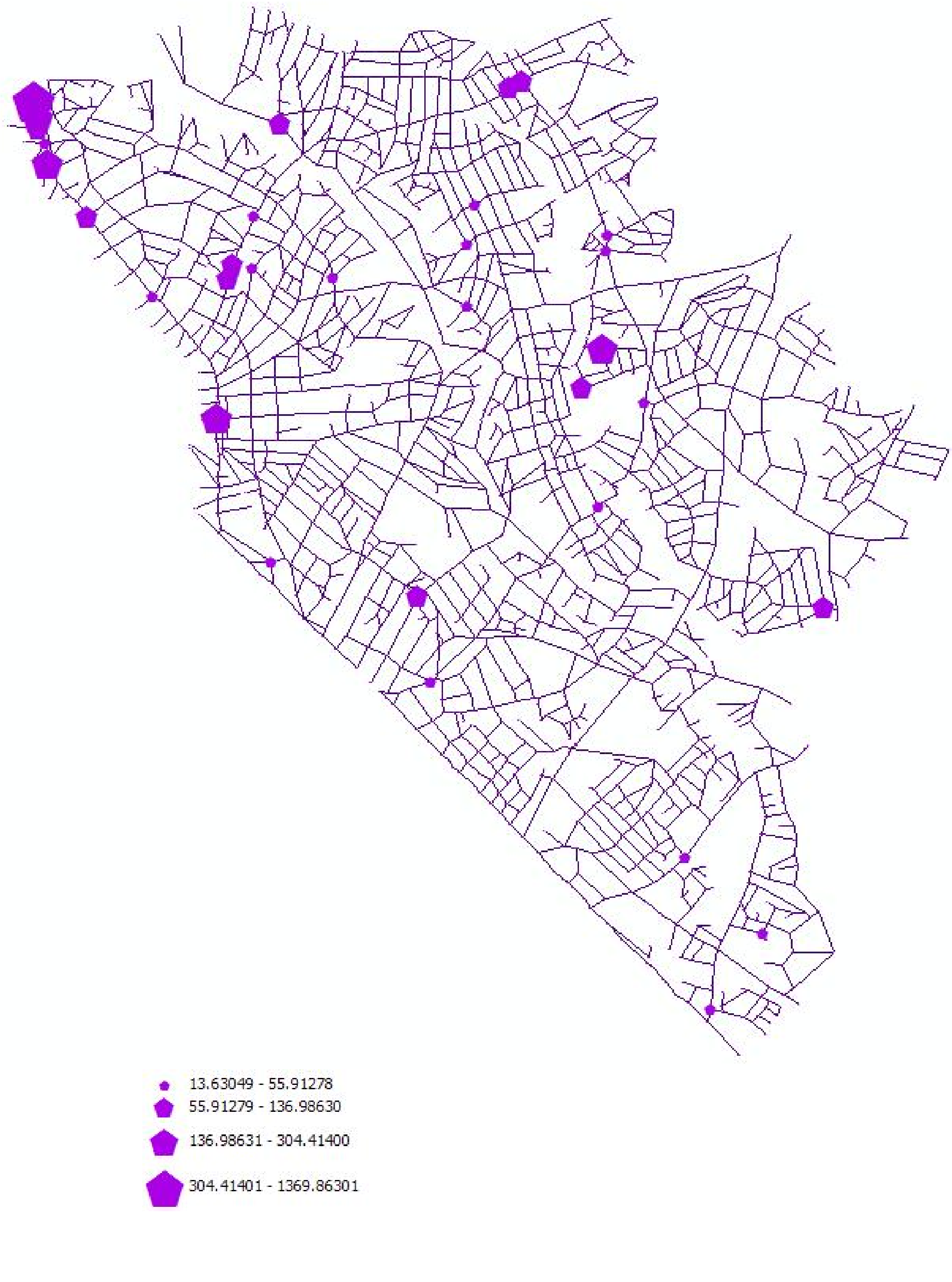


Table 28 -- RANKED BY SEVERITY-WEIGHED RATE