

Display and Analysis of Crash Data

By

Dr. Andrew Graettinger, Dr. Jay K. Lindly and Mr. Gautam J. Mistry
Department of Civil and Environmental Engineering
The University of Alabama
Tuscaloosa, Alabama

Prepared by

UTCA

University Transportation Center for Alabama

The University of Alabama, The University of Alabama in Birmingham
and The University of Alabama at Huntsville

UTCA Report Number 03102
April 1, 2005

Technical Report Documentation Page

1. Report No UTCA 03102		2. Government Accession No.		3. Recipient Catalog No.	
4. Title and Subtitle Display and Analysis of Crash Data			5. Report Date January 31, 2005		
			6. Performing Organization Code		
7. Authors Dr. Andrew Graettinger, Dr. Jay K. Lindly, and Mr. Gautam J. Mistry			8. Performing Organization Report No. UTCA Final Report 03102		
9. Performing Organization Name and Address The University Transportation Center for Alabama The University of Alabama Box 870205 Tuscaloosa, AL 35487-0205			10. Work Unit No.		
			11. Contract or Grant No.		
12. Sponsoring Agency Name and Address UTCA			13. Type of Report and Period Covered. Final Report: April 1, 2003 – December 31, 2004		
			14. Sponsoring Agency Code		
15. Supplementary Notes					
16. Abstract This report describes the development and testing of a new Geographic Information System (GIS) application for the display and analysis of crash data stored in the Critical Analysis Reporting Environment (CARE) system. The primary objective of this study was to map crashes and correlate them with existing roadway features like bridges, crossroads, railroad grade crossings, etc. This application was developed to enable CARE data to be presented graphically on maps, which is not currently possible in CARE alone. The application utilizes the advantages of both CARE and GIS software. Reports can be created in CARE for detailed study, while representative maps can be generated in GIS. Crashes can be selected according to different parameters such as time of day, driving under the influence (DUI), pavement conditions, number of lanes, etc. This selection can occur in either CARE or GIS. Spatial analyses are performed on crash data that are mapped using the GIS application to identify “hot spots” at intersections, road segments, and mileposts. The average daily traffic (ADT) data for different highway sections can be represented graphically along with the crashes. The application also identifies the route with the least impedance, in the form of crashes, between origin and destination of a travel path using the highway network analysis feature of GIS. This application also identifies errors in the intersection node numbers used to identify crash locations. This error results in mapping the same crash to multiple intersections. Errors like this are difficult to detect without displaying the data graphically.					
17. Key Words crash data, geographic information system, GIS, spatial analysis, crash analysis			18. Distribution Statement		
19. Security Class (of this report) Unclassified		20. Security Class (of this page) Unclassified		21. No. of Pages 69	22. Price

Contents

Contents	iii
Tables	vi
Figures.....	vii
Executive Summary	viii
1.0 Introduction.....	1
1.1 Intoduction.....	1
1.2 Problem Statement.....	1
1.3 Scope of Study.....	2
1.4 Report organization.....	3
2.0 Review of the Literature	4
2.1 Introduction	4
2.2 Project Description	4
2.3 Analysis and prediction of Road Accidents Using GIS-Based Integrated Model.....	5
2.4 Prototype Traffic Safety GIS for Honolulu, Hawaii.....	6
2.5 Accident Data Management for Mexican Federal Roads Using GIS.....	7
2.6 Maryland Spatial Analysis of Crashes (MSCA) Project	8
2.7 Iowa State project	9
2.8 GIS-Based Tracking of Highway Crashes.....	9
2.9 CODES Data Linked to GIS.....	10
2.10 GIS Used to Enhance Pedestrian Safety in Clark County, Nevada.....	11
2.11 Other Related Studies	11
2.11.1 Graphical Representation of Crash Data.....	11
2.11.2 Crash Patterns	13
2.11.3 Crash Type	13
2.11.4 Crash Rates	14
2.11.5 Traffic Safety	14
2.11.6 Crash Data Linked to GIS.....	15
2.11.7 Network Analysis.....	16
2.12 Summary.....	17
3.0 Methodology.....	18
3.1 Introduction	18
3.2 Crashes Occurring at Intersections	18
3.2.1 Intersection Crash Recorded in DPS Database	18
3.2.2 Intersection in GIS	18

3.2.3	Connecting Node Shapefile in GIS to Crash Attribute Table for Intersections	19
3.3	Crashes Occurring on Road Segements and State Highways in Urban Areas ...	20
3.3.1	Road Segment Crash Recorded in DPS	21
3.3.2	Road Segments in GIS	21
3.3.3	Connecting Road Shapefile to DPS Database.....	21
3.4	Crashes Occurring on Interstates and State Highways in Rural Areas.....	22
3.4.1	Interstate and State Highway Crashes (Rural Area) Recorded in DPS	22
3.4.2	Interstate and State Highways in GIS	23
3.4.3	Connecting Highway Route File to DPS Database for Interstate and Rural State Highways	23
3.5	Summary.....	24
4.0	Spatial Analysis and Display of Crash Data.....	25
4.1	Introduction.....	25
4.2	Analyses and Display of Intersection Crashes.....	25
4.2.1	Total Number of Intersection Crashes	25
4.2.2	Crash Severity.....	27
4.3	Analyses and Display of Segmental Crashes.....	28
4.3.1	Total Number of Segmental Crashes	28
4.4	Analysis and Display of Highway Crashes.....	30
4.4.1	Total Number of Crashes	30
4.4.2	Crash Severity.....	32
4.4.3	Crashes on I-20/59 Normalized by ADT Data	33
4.5	Summary.....	35
5.0	Network Analysis.....	36
5.1	Introduction	36
5.2	Methodology.....	36
5.3	Network Analysis	37
5.4	Summary.....	40
6.0	Duplicate Node Numbers.....	41
6.1	Introduction and Problem Statement	41
6.2	Matching Paper Maps, Digital Maps, and CARE Data.....	41
6.3	Summary.....	44
7.0	Conclusions and Recommendations	45
7.1	Introduction.....	45
7.2	Conclusions.....	45
7.3	Recommendations.....	46
8.0	References.....	47
9.0	Abbreviations.....	49

Appendix A. Procedures for Creating Shapefiles for Intersections, Road Segements and Highways	50
Appendix B. Procedure for Connecting DPS Data with GIS Maps	54
Appendix C. Procedure for Creating and Using Networks.....	57

Tables

Number	Page
5-1. Summary of safe route maps	40

Figures

Number	Page
1-1. “Hot Spots” on I-20/59 represented in CARE.....	2
2-1. Fatal crash spots on 400-level highways in 1992.....	6
2-2. Injuries in central Honolulu, 1990.....	7
2-3. Road segments (500 m) thematically mapped in Queretaro.....	8
2-4. State highway crashes in Brazos County, Texas.....	10
2-5. High crash locations in Ames, Iowa for 1997.....	12
2-6. Crashes occurring on weekdays between 10 am and noon, Charlottesville.....	15
2-7. Map showing intersection crashes.....	16
2-8. Map showing safe route selected from origin to school.....	17
3-1. Attribute table showing crash records in CARE.....	19
3-2. Map showing nodes digitized in the Tuscaloosa area.....	20
3-3. Table showing crashes on Interstate I-359 with milepost.....	22
3-4. Milepost mapped along highways using linear referencing.....	24
4-1. Thematic map of crashes in Tuscaloosa area.....	26
4-2. Percentage of total crashes occurring at intersections during evening rush hour.....	27
4-3. Thematic map for crash severity at intersections.....	28
4-4. Crashes on road segments in Tuscaloosa area on Fridays.....	29
4-5. Crashes on road segments within 0.25 mile buffer around highways.....	30
4-6. Crashes on Interstate and state highways according to mileposts.....	31
4-7. Crashes on highways due to driving under influence during 1996-2000.....	32
4-8. Crash severity on highways of Tuscaloosa county during 1996-2000.....	33
4-9. Crashes on I-20/59 with ADT data for 2001.....	34
4-10. Crashes on I-20/59 in 2001 normalized by ADT.....	35
5-1. Map showing path selected based on length of route as weight.....	38
5-2. Map showing path selected based on number of crashes as weight.....	39
5-3. Map showing path selected based on a combination of length of route and number of crashes as weight.....	39
6-1. Map showing intersections with same node numbers.....	42
6-2. Thematic map for categories of errors.....	43
A-1. Editor session open for “county_roads” shapefile.....	51

Executive Summary

This report describes the development and testing of a new geographic information system (GIS) application for the display and analysis of Department of Public Safety (DPS) crash data stored in the Critical Analysis Reporting Environment (CARE) system. The primary objective of this study was to map crashes and correlate them with existing roadway features like bridges, crossroads, railroad grade crossings, etc. This application was developed to enable crash data to be presented graphically on maps, which is not currently possible in CARE alone.

The application utilizes the advantages of both CARE and GIS software. Reports can be created in CARE for detailed study, while representative maps can be generated in GIS. Crashes can be selected according to different parameters such as time of day, driving under the influence (DUI), pavement conditions, number of lanes, etc. This selection can occur in either CARE or GIS.

Spatial analyses are performed on crash data that are mapped using the GIS application to identify “hot spots” at intersections, road segments and highway mileposts. The average daily traffic (ADT) data for different highway sections can also be represented graphically along with the crashes. The application also identifies the route with the least impedance, in the form of crashes, between origin and destination of a travel path using the network analysis feature of GIS. This application identifies errors in the intersections node numbers used in crash data. This error results in mapping the same crash to multiple intersections. Errors like this are difficult to detect without displaying the data graphically.

Section 1.0 Introduction

1.1 Introduction

This project developed a new Geographic Information System (GIS) application for the display and analysis of crash data. Data about vehicle crashes were obtained from the Critical Analysis and Reporting Environment (CARE) software, which is a data analysis tool maintained by the CARE Research and Development Laboratory of the University of Alabama. Multi-year crash data from Tuscaloosa County, Alabama were mapped on a commercially available ETAK base map, and those crash locations were compared with existing roadway features.

The primary objective of this study was to map vehicle crashes. Other objectives of the project were to perform spatial analysis and “hot spot” identification, to identify errors in intersection numbering, and to a perform network analysis to identify the route with least crashes.

1.2 Problem Statement

The Alabama Department of Public Safety (DPS) crash database contains detailed information about vehicle crashes that occurred in Alabama. Crash information of interest can be extracted from the database by “filtering” the data to obtain a subset which displays only those crashes meeting a specific criterion. The data about these crashes are typically displayed in a tabular form with associated statistics. The crash information includes the location of each crash along with the municipal area where the crash occurred. The crash locations are described as nodes (intersections), links (road segments) or route-milepost (distance along a state highway or Interstate).

CARE can graphically display crash location frequency on a state highway or Interstate as shown in Figure 1-1. The figure shows a section of typical Interstate highway and the associated crashes mapped according to milepost. The numbers on the horizontal line represent the mileposts along the route, while the vertically stacked dots represent the number of crashes that occurred at that location. This graphical representation of crash data does not correlate crash location with roadway features such as bridges, cross roads, curves, rail-road crossing, etc. The GIS application developed in this study enables a better representation of multi-year crash data on maps, which correlate existing roadway features and crash data.

The locations of crashes stored in CARE are verbally described in the form of nodes, links or route-milepost. This verbal description is not in a format that can be easily used in a GIS, which requires a seamless coordinate system. The GIS application developed in this project has the advantage of having both a verbal description for mapping crash data as well as a coordinate system that will accept Global Position System (GPS) data that may be used in the future.

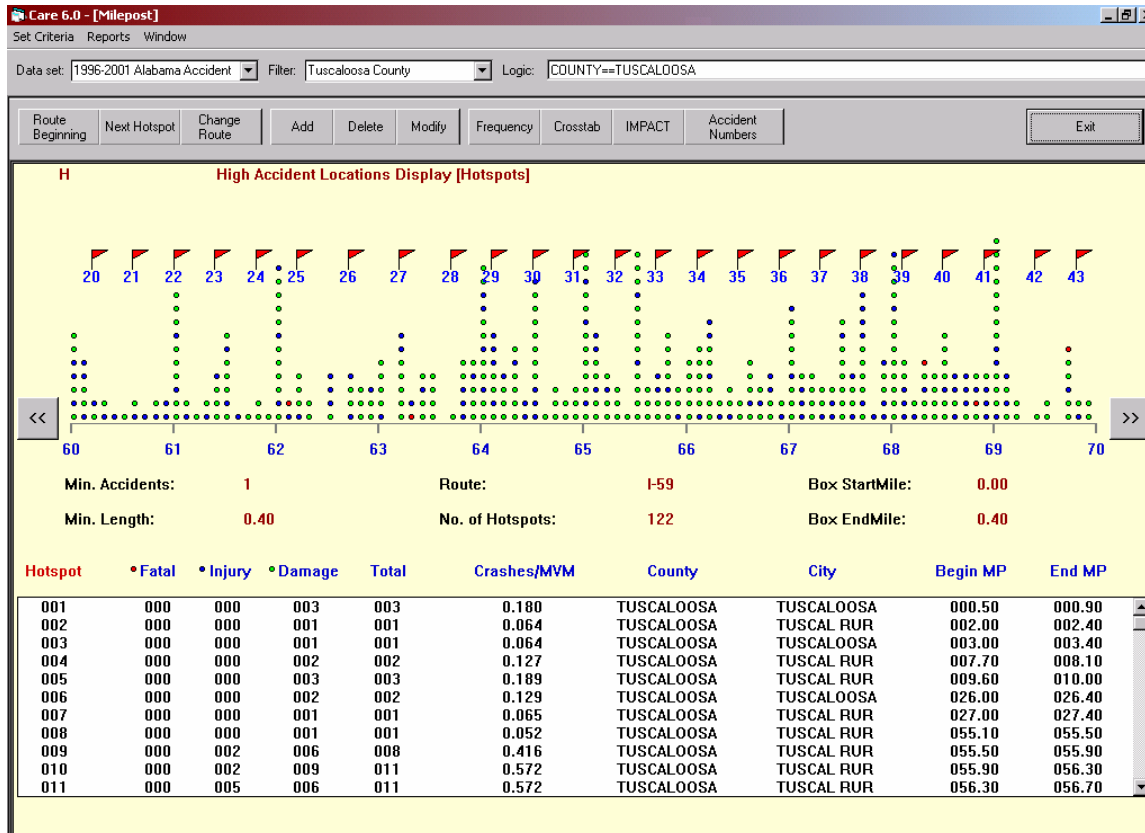


Figure 1-1. Typical Interstate “Hot Spots” represented in CARE

1.3 Scope of Study

The base map used for this project was commercially purchased and therefore did not have the ALDOT node numbers, link numbers, and municipal codes. This information was manually inserted on the base map by matching intersections and road segments on paper maps that were provided by ALDOT with the same intersections and road segments on the digital base map. The process of coding the base map with descriptive node-link information was very time consuming.

Considering the amount of time and resources available, the project was limited to Tuscaloosa County. Tuscaloosa County has both urban and rural areas, an Interstate highway, state highways, county roads, and multiple municipalities, which cover most of the variations that will be encountered if this GIS application is implemented statewide. This project focuses on a GIS

vehicle crash application for Tuscaloosa County that can act as a pilot study for future statewide implementation.

1.4 Report Organization

This report is organized in seven sections. Section one introduces the research project topic and objectives and also defines the scope of the CARE-GIS study. Section two is a literature review of GIS applications and explains the similarities/differences between the literature and the current project. Section three describes the methodology adopted to connect the CARE database with the GIS for three different descriptive crash location types: intersections, road segments, and route-mileposts. Section four illustrates the technique used to spatially analyze CARE data that is in the GIS. Section five describes an example of a network analysis carried out in GIS using CARE data to identify the safest route between an origin and destination selected by the user. Section six describes the detection of errors in assigning node numbers to intersections and displaying the errors on maps to help identify locations where the errors occurred. Finally, section seven presents the conclusions of this study and discusses future work.

Section 2.0

Review of the Literature

2.1 Introduction

A comprehensive literature search was performed to gather information related to crash data displayed and analyzed with GIS. GIS combines graphics depicting the location of features with the data describing the features. There are many GIS applications related to transportation engineering such as: transit programs, infrastructure management, airport management, and pavement management, to name a few. However, this literature review was restricted to applications that assist in displaying and analyzing crash data spatially. The primary sources of literature were FHWA publications, Transportation Research Records, ITE publications, and the worldwide web.

2.2 Project Description

The primary purpose of this project was to present DPS crash data accompanying CARE on maps so that the crashes could be related to existing roadway features like bridges, crossroads, railroad crossings, etc. The maps allow crashes to be represented in such a way that locations with high numbers of crashes can be identified at a glance. These high crash locations are called “hotspots,” which occur at intersections, road segments, or mileposts along a highway. ArcGIS software was used to map the crashes. Various spatial analyses were performed on crash data with respect to crash severity, number of crashes, and factors such as weather conditions and light conditions.

One of the more advanced features of this project was the network analysis, which can find the safest route between an origin and a destination. The route selected is the route that exhibited the minimum number of crashes between the origin and destination. Along with an origin and destination a user selects a time frame for the analysis, which can be any number of years from 1993 to the most recent year for which crash data is available.

The remainder of this literature review describes work done by other states and agencies related to mapping crashes and performing spatial and network analysis. Some work by other agencies is limited to certain types of crashes such as pedestrian crashes, bicycle crashes, truck crashes, crashes along a curve, etc. This project deals with crashes of all types and can be restricted to deal with a specific type of crash if required.

2.3 Analysis and Prediction of Road Accidents using GIS-Based Integrated Model

Frank et al. (2001) from the University of Waterloo carried out a research project that explored the effective functionality of GIS in data management, spatial referencing, and graphic visualization of data to provide a platform for analysis of crashes. The project involved analysis and prediction of road accidents using a GIS-based integrated model. The model integrated relevant information from different databases in a seamless and automated manner, and also displayed the results of the analyses on accident prediction maps, fatal crash maps, etc.

The Waterloo project developed a model to predict and analyze the number of crashes that would occur on a specific portion of highway, along a corridor, or at a particular intersection for a given time duration. The model required road geometry, traffic characteristics, crash data, weather conditions, etc. This information was stored in separate databases, which were integrated into a GIS for analysis and display of the results. A list of the Ontario Ministry of Transportation databases that were used follows:

- Digital cartographic reference base, which contains information related to the road network, rail network, lakes, rivers, parks etc.,
- Accident database system, which contains detailed information of crashes,
- Highway inventory management system, which contains geometric information of each road section, and
- Traffic volume inventory system, which contains existing and future traffic volumes.

The databases were integrated using a linear highway reference system (LHRS). LHRS is used in each database to distinctively identify a particular length of highway with comparable traffic and geometric characteristics. Figure 2-1 is a map displaying fatal crash spots on 400-level highways (network of controlled-access freeways in Ontario, Canada) for the year 1992.

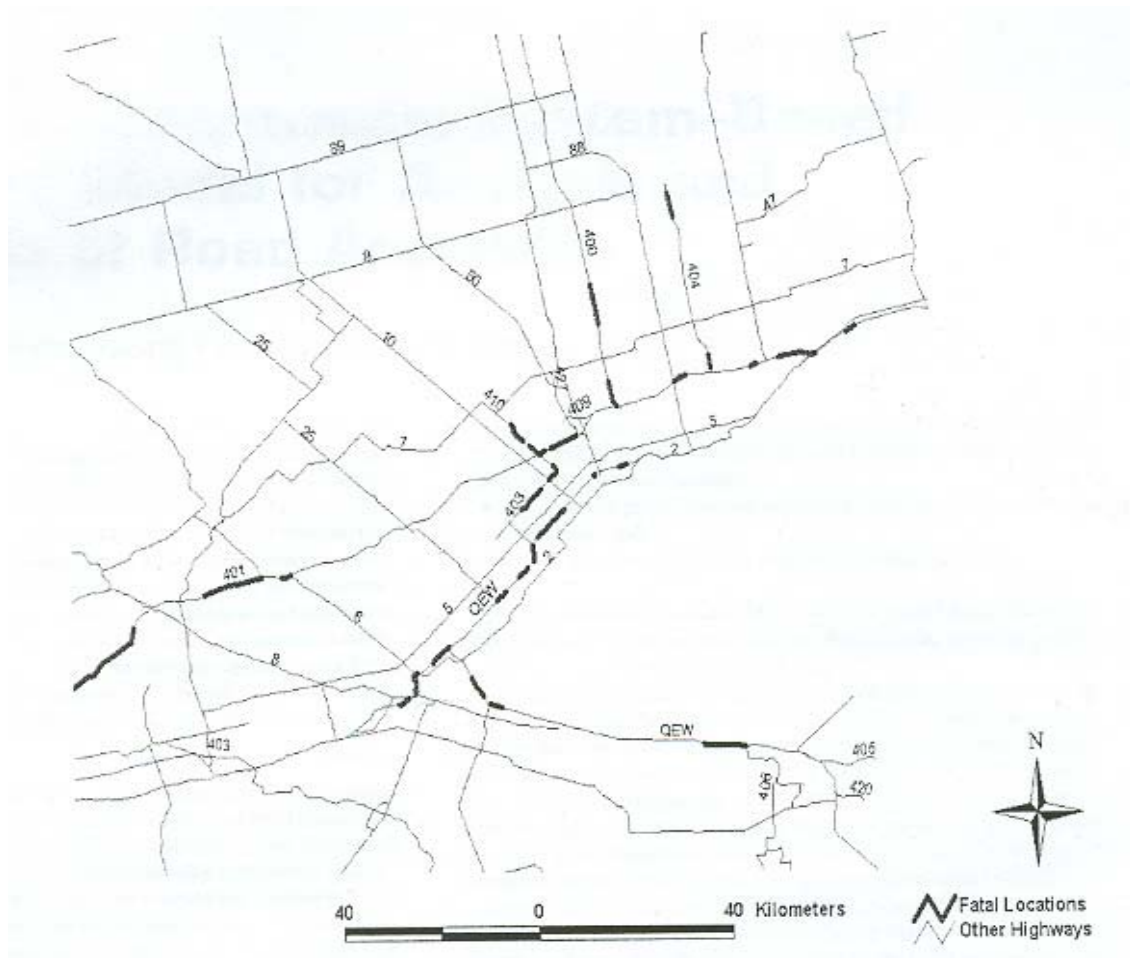


Figure 2-1. Fatal crash spots on 400-level highways in 1992

2.4 Prototype Traffic Safety GIS for Honolulu, Hawaii

Kim et al. (1995) conducted research at the University of Hawaii at Manoa, Honolulu, to develop a GIS spatial analysis system for vehicular crashes. The research produced a prototype GIS which would help in spatially analyzing crash data in combination with a crash outcome data evaluation system (CODES). CODES is a database containing detailed information about vehicular crashes that can be used to perform analyses on various traffic safety interventions such as motorcycle helmets, seatbelt effectiveness, etc.

A combination of software was used to create Hawaii's GIS application. A workstation environment was used to store and manage attribute data for crashes. A software engine called SAS carried out statistical analysis and database management. A desktop vector GIS using ATLAS*GIS and MAPINFO were used for displaying and manipulating crash data spatially on maps. Cross-platform and cross-software transfers allowed users to take advantage of the strengths of various software packages and to evade some of their weaknesses.

The Hawaii project successfully mapped crashes that occurred at intersections and on road segments. However, crashes that occurred on freeways were not mapped because milepost locations were not available in digital form. Figure 2-2 shows crash severity at intersections in Honolulu. This GIS application was used for studying collision patterns, formulating meaningful safety plans, identifying areas with high levels of crashes, identifying general patterns of crashes such as fatalities, alcohol related crashes, etc.

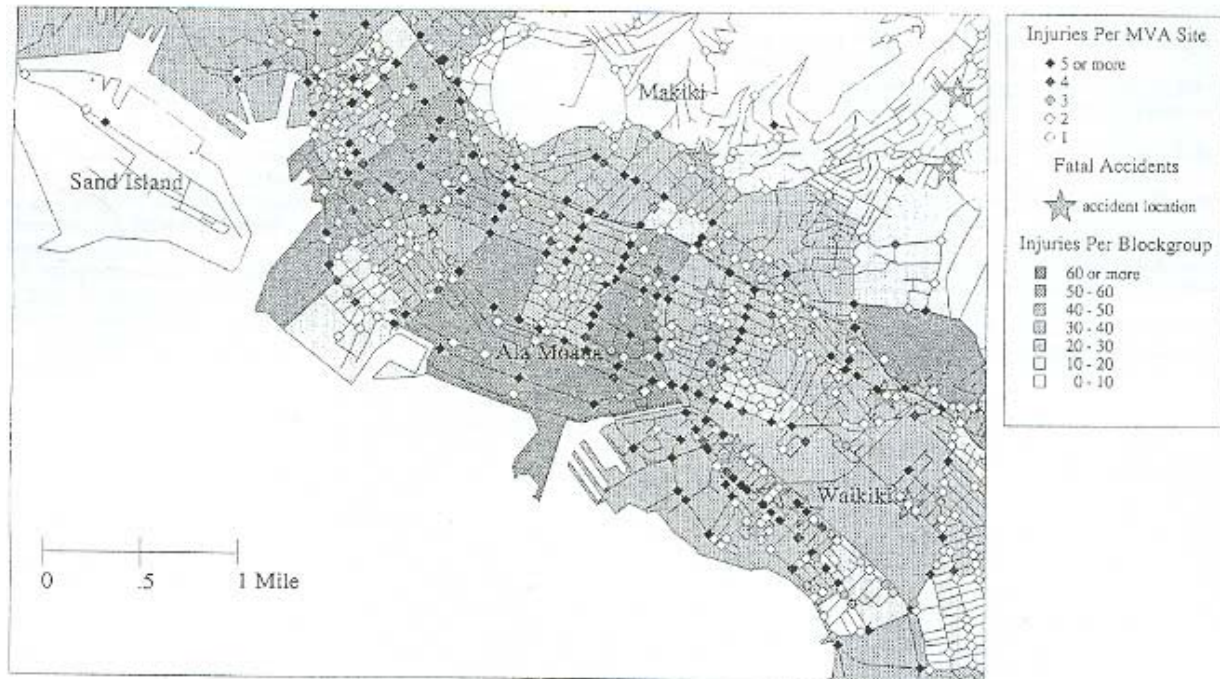


Figure 2-2. Injuries in central Honolulu, 1990

2.5 Accident Data Management for Mexican Federal Roads using GIS

Mendoza et al. (2000 & 2001) from the Mexican Transportation Institute developed a GIS-based accident data management system for federal roads in Mexico. Information about accidents that occurred in 1997 were retrieved and stored in electronic files by the Federal Highway Police. The record for each crash was divided into two sub-bases, which represented accident data and participant data.

Each crash record is represented on a map as a reference theme having a geocoded index. Geocoding creates a point feature for each crash record in the database. These point features are positioned at the midlength of the road segment on which they occurred. To obtain better geocoding accuracy, the road links were broken into 500 m segments to generate smaller georeferenced segments. Subsequently the two databases were geocoded, and queries were performed on the mapped data such as identification of the site with the most severe crashes, locating the nearest emergency medical service center for a particular location, etc. Figure 2-3

displays a thematic map of 500 m road segments displaying the number of accidents recorded in 1997 in the state of Queretaro.

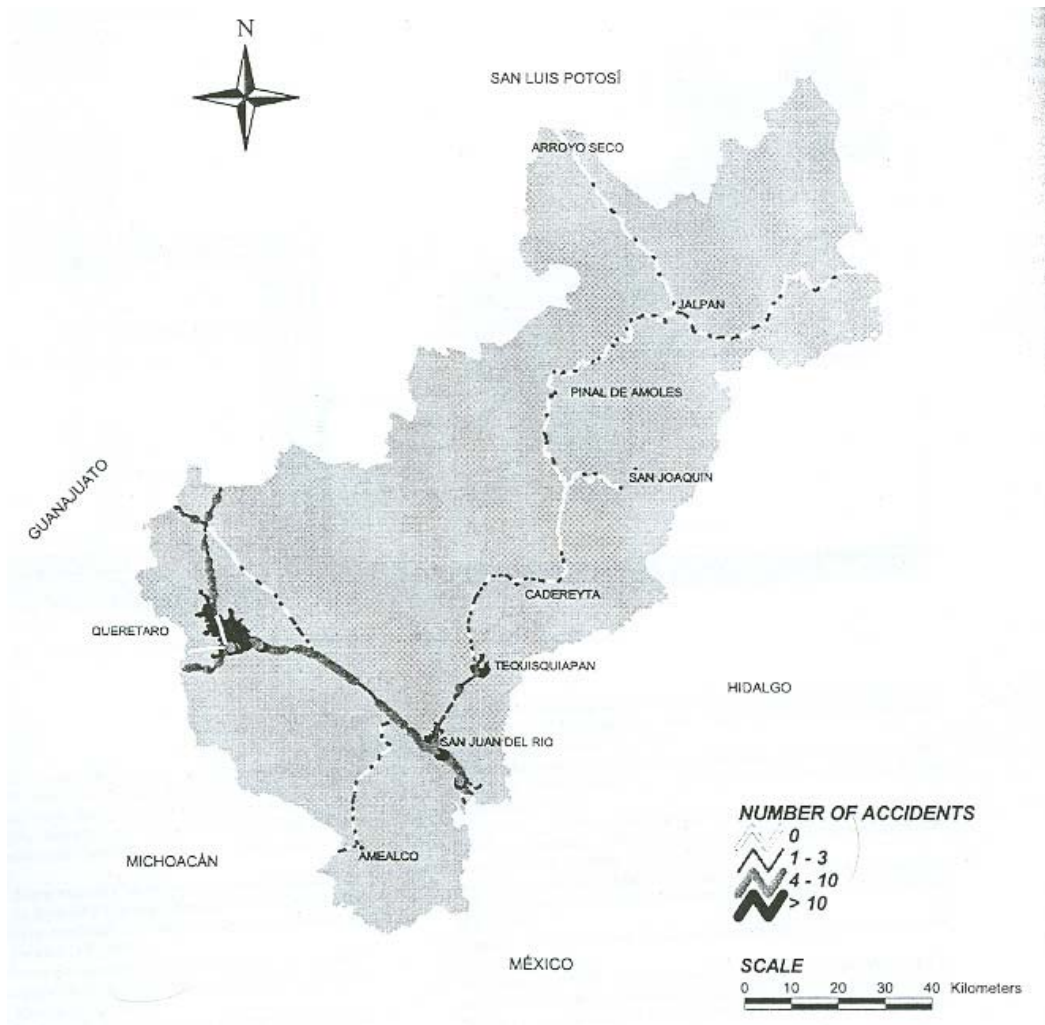


Figure 2-3 Road segments (500m) thematically mapped in Queretaro

The Mexican project used specific geographically referenced addresses of crashes. In order to successfully geocode these points, a base map is required with the same system of addresses. The roads were broken down into small segments and a point feature with an address was created at the midlength of each segment.

2.6 Maryland Spatial Analysis of Crashes Project

Bapna et al. (2003) carried out a research project at Morgan State University to design a GIS application for the Maryland Department of Transportation. The project explored GIS capabilities to provide online information about crashes and their statistics for commercial vehicles. The system showed graphical representation of crash information such as locations,

crash severity, time of crash, causal factors, etc. The system was deployed over the Internet, which makes it very accessible; however, the project did not perform network analysis using crash data.

2.7 Iowa State Project

Souleyrette et al. (2001) from the Center for Transportation Research and Education (CTRE, Iowa State University), developed GIS tools and procedures to systematically identify high crash locations due to hazardous roadway locations, design, etc. Two databases were integrated with existing digital imagery to present the data in the desired manner. The first database was a Geographic Information Management Systems (GIMS), which stored information about roadway characteristics such as lane width, surface type, traffic volume, shoulder type, etc. The other database was a Geographic Information System Accident Analysis and Location System (GIS-ALAS), which stored 10-years of information for crashes occurring on public roads.

The GIS tool developed produced maps used to identify high crash frequency locations. GIMS and GIS-ALAS databases were used in combination to relate crash rates with roadway design characteristics that contributed to the increase in crash rates. The highlight of this project was developing a statistical model for the relationship between geometric features and crash rates.

2.8 GIS-Based Tracking of Highway Crashes

The Texas Transportation Institute teamed up with Mallick and Sung (2001) of the Department of Statistics at Texas A & M University, to develop a GIS-based model to visually track crashes occurring on state highways for a project sponsored by the Bureau of Transportation Statistics, U.S. Department of Transportation. Crash data in most Texas Counties are manually entered in different databases, which are difficult to access and understand. Crash data for required counties were merged from the existing databases, and the locations of crashes on highways in those counties were represented on maps using GIS-based technology. These interactive visualizations provided the user with information such as date of crash, drunk driving, number of occupants, speeding, etc. The “hot spots” identified on maps assisted police departments in deploying officers to areas with high crash rates related to speed or drunk driving.

Figure 2-4 shows a map of Brazos County with crashes plotted along the state highway system. Crashes that were non-incapacitating, incapacitating, or fatal were mapped with different colors to visually indicate the problem areas along Brazos County highways.

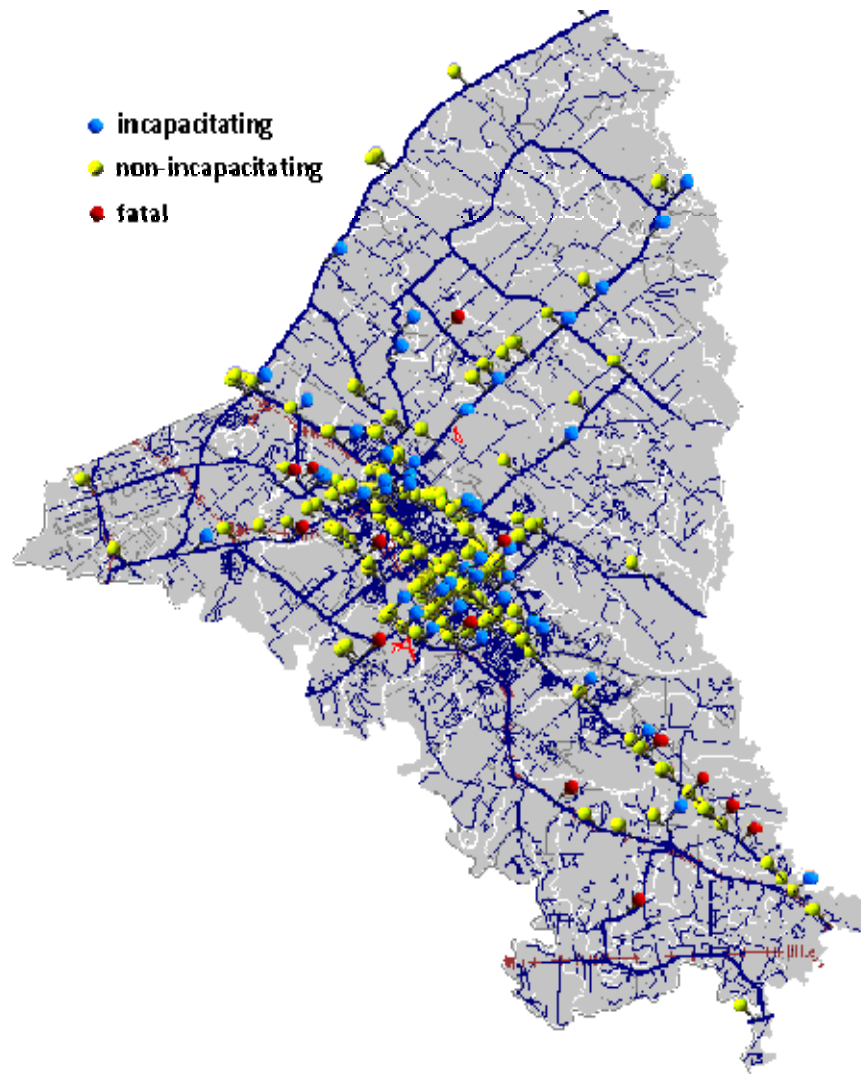


Figure 2-4. State highway crashes in Brazos County, Texas

2.9 CODES Data Linked to GIS

Four Crash Outcome Data Evaluation System (CODES) states, Hawaii, Maryland, New Hampshire, and South Carolina participated in a collaborative effort to link the CODES database and GIS to map crashes, which facilitates spatial analysis and identification of crash patterns. Kim and Kerns (2001) along with a team of researchers and professionals carried out this project. In addition to information such as police crash data, emergency medical services transport data, hospital data, and insurance claims, CODES also provides location information such as police beat, census tract, county, street name, distance from a milepost or nearest intersection, etc.

In addition to linking CODES to GIS, each state had individual research goals. The Hawaii GIS project developed different procedures for geocoding, spatial analysis, and mapping. The Maryland CODES project calculated incidence rates and spatially represented crash sites using

GIS. It was also used to reinforce a pedestrian safety campaign in Ocean City, MD. The New Hampshire GIS facilitated graphical representation of crash data such as frequencies and trends, while the South Carolina GIS mapped crashes, described CODES in spatial terms, and helped in community development programs.

2.10 GIS used to Enhance Pedestrian Safety in Clark County, Nevada

Srinivas et al. (2002) from the University of Nevada-Las Vegas Transportation Research Center developed a GIS application to identify problem areas for pedestrians in Clark County, Nevada. It enhanced pedestrian safety through study of temporal and spatial patterns of crashes and developed a methodology to integrate crashes with other important information such as land use information, demographic information, and socio-economic data. Crash data maintained by the Nevada Department of Transportation was geocoded with the help of three referencing approaches: offsets from intersections, street addresses, and route numbers and milepost locations. The primary focus of the project was to identify zones with high crash rates for pedestrians and to obtain results from geospatial analysis using GIS.

2.11 Other Related Studies

Various articles provide a clear idea about how GIS can be instrumental in improving the interpretation and representation of crash data. Based on this literature review, most studies related to GIS and crash data can be delineated into categories according to fundamental issues such as: graphical representation of crash data, crash patterns, crash types, crash rates, traffic safety, and crash data linked to GIS and network analysis. All of the research discussed in this section includes one or more of these GIS/crash analysis topics.

2.11.1 Graphical representation of crash data

One of the most important benefits of GIS is its ability to graphically represent a large amount of data on a single map. This type of representation is very helpful in understanding multiple data types and values at a glance according to the location. Roche (2000) emphasized various possible macroscopic and microscopic applications of GIS in crash data display and analyses. The macroscopic analysis included the identification of locations on large area maps where a high number of crashes occurred at intersections. Tabular crash data was connected to a GIS, allowing a visual representation of the data to assist in the analysis at a larger scale. Microscopic analysis involved scrutinizing subsets of crash data at a single intersection. This type of analysis helped to identify a pattern or patterns of crashes occurring at a particular intersection. Pattern identification indicated the necessity for emergency response improvements and new enforcement measures to improve the situation. Figure 2-5 shows a map that identifies locations with high crash frequencies in Ames, Iowa.

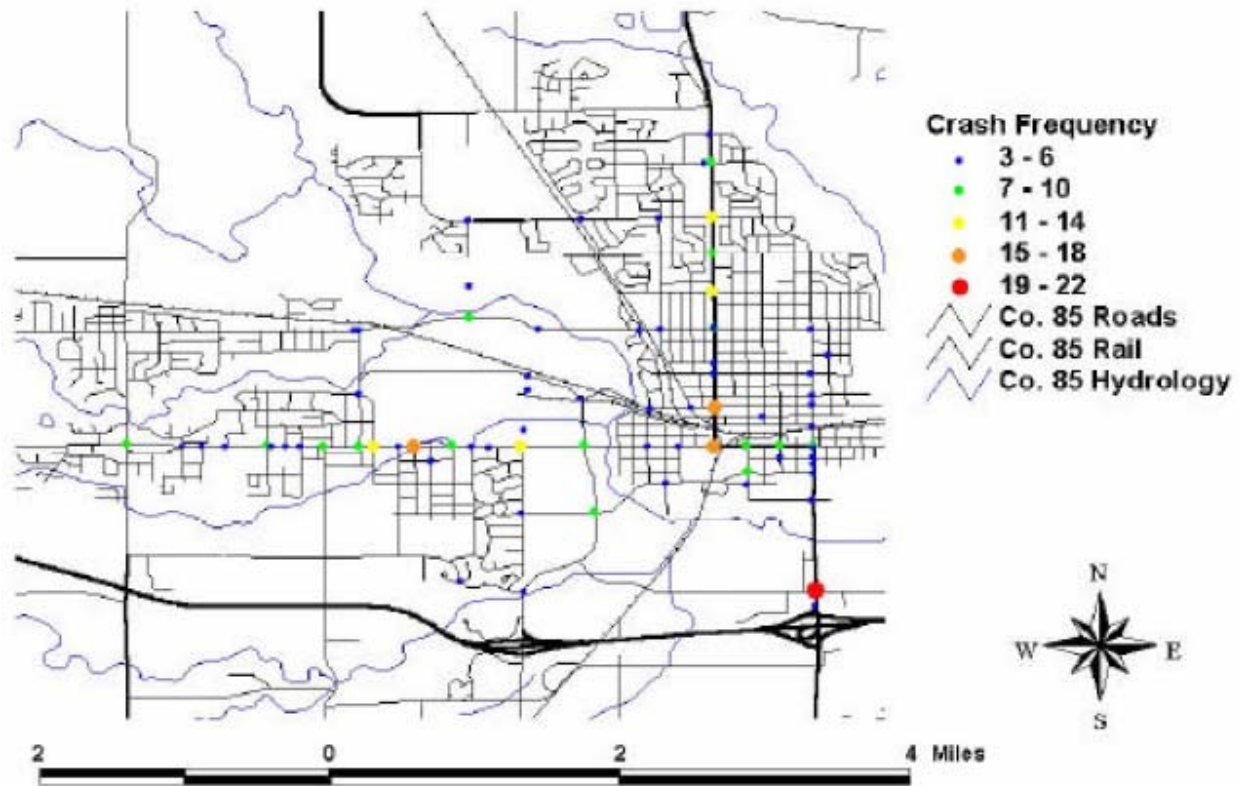


Figure 2-5. High crash locations in Ames, Iowa for 1997

A select group of states including California, Illinois, Maine, Michigan, Minnesota, North Carolina, Utah, and Washington, worked together to develop a multi-state crash analysis system which integrates various information including crash data, roadway inventory and traffic operations data. This crash analysis system was integrated with a GIS within the Highway Safety Information System by Harkey (1999) to display crashes graphically, to identify high-truck-crash locations, and to take advantage of the spatial analysis capabilities within the GIS. High truck crash locations along primary designated truck routes were identified using the sliding scale program. Small segments, of specified lengths, of a route were selected progressively to check for crash rates higher than or equal to a threshold crash rate set by the user. At each intersection the program traversed three miles in each direction of the intersection to check for truck crash rates, and then proceeded along a designated route. To identify truck crash locations along undesignated truck routes, a program called the “corridor analysis program” was used. This helped in identifying routes which were not designated for trucks, but were still used by trucks due to the shorter routes to terminals.

Thnay et al. (2000) emphasized the use of a digital database and a GIS to visually represent crash data and draw collision diagrams that made analysis of high crash locations more efficient. Large amounts of data can be analyzed to identify high crash locations, and to allow geocoding as a pin-map using GIS. Collision diagrams are spatially plotted on base maps to identify crash patterns and relate those patterns to existing roadway objects with the help of orthophotos and GIS-based right-of-way maps.

2.11.2 Crash Patterns

One of the other important abilities of GIS is to analyze various types of data in combination with crash data to identify patterns of crashes with respect to time, demographic composition, etc. Arthur et al. (1997) used GIS to develop a monitoring system for the City of Calgary to make city roads and intersections safer. The authors used GIS to identify locations with a high number of crashes and high crash severity in an attempt to establish a correlation between speed and crashes. The project not only involved identifying accident locations, but also involved identifying the time period when the accidents occurred.

Chakraborty et al. (1996) explained the use of GIS to identify the demographic composition of a community affected by the release of toxic substances resulting from truck accidents by adapting a Geographic Plume Analysis (GPA) approach. A toxic dispersal “footprint” produced by a chemical dispersal model was superimposed on a GIS map. Demographic data, such as racial and economic information of the effected population were determined. Areas with high populations of low-income and minorities were identified as locations most likely to be exposed to hazards commonly transported by trucks.

Miller (2001) published a paper that explored the analytical capabilities of GIS to evaluate crash data. Some of the GIS techniques emphasized in this paper are: identifying crash location patterns, producing collision diagrams, identifying contrasting sources of crash data, using spatially based statistical applications, and providing appropriate routes for hazardous material carriers, bicyclists, and pedestrians.

2.11.3 Crash Type

GIS can analyze crash data for a specific type of crash that a group of officials may be interested in, such as truck crashes, crashes involving alcohol, animal-vehicle crashes, etc. Schutle and Aultman (1999) conducted a study to inspect the characteristics of alcohol-related crashes in the state of Kentucky, with varying levels of alcohol prohibitions. The counties were categorized as “wet” or “dry.” The GIS analysis included the use of highway milepost and home zip code of drivers’ residences to estimate the distances from the drivers’ homes to the crash sites.

Harkey et al. (1999, 2001) at the Highway Safety Research Center of the University of North Carolina worked with the North Carolina Center for Geographic Information and Analysis to map truck crashes to pinpoint problem locations throughout the state. Identification of such locations with the help of GIS assisted in assigning their limited resources to locations to most reduce crashes and fatalities. The system also identified routes with high truck crash rates that occurred due to drivers bypassing weigh stations or using small roads to reduce longer truck routes.

Norris et al. (2003) demonstrated GIS as a useful tool to identify high wildlife-vehicle collision locations along State Route 260 in central Arizona. Limited Rocky Mountain elk in the area were instrumented with GPS receiver collars and satellite data transmission capabilities, which

were integrated with a GIS to identify elk highway crossing locations and areas with high populations of elk near highways. Information received from these collars was used to determine the locations requiring fencing to prevent elk from crossing the highway. The system was also used to evaluate the fence effectiveness based on before and after collision numbers.

2.11.4 Crash Rates

Crash rate/frequency is an important measure to identify locations that should be given priority for traffic safety improvement. A report by Schneider (2001) illustrated identification of factors associated with pedestrian crashes on the University of North Carolina at Chapel Hill Campus through a GIS spatial analysis. Two distributions were compared spatially: 1) perceived risk distribution and 2) police-reported crash distribution. This spatial comparison helped to identify locations that were perceived to be more crash prone, and to show how they were related to actual crash frequency.

Aty and Saidi (1999) conducted a similar GIS-based study to determine a relationship between driver characteristics and traffic crashes. GIS was used to identify regions with high crash rates. Ticketed drivers from regions with high crash rates were provided with training programs to educate them about traffic safety problems and to improve their driving behavior.

2.11.5 Traffic Safety

Traffic safety is one aspect of transportation engineering that has benefited from the use of GIS. Austin et al. (1997) applied GIS technology to develop methods that emphasize road safety improvements. One of the applications developed identified high crash rates in school zones during peak hours of travel. This type of zonal identification helped in picking suitable areas for road safety campaigns. Another application identified routes taken by children to walk to schools and analyzed those routes for safety.

Jerome et al. (2000) emphasized the application of GIS to improve highway safety by providing a better visualization of highway safety conditions. Macro level tools were developed, such as a spatial grid-based analysis tool and a network analysis tool to identify high crash activity areas without getting into detailed crash analysis. Grid-based analysis involved overlaying a grid of tiles of specified dimensions on a spatial crash database to identify high crash areas by thematic representation of tiles. Network analysis involved identifying crash locations within a specified distance from a selected corridor. Micro level tools involved identification of locations based on individual crashes and theme-based crash groups. Theme-based grouping typically restricts the study area to a particular corridor, intersection, interchanges or type of roadway.

2.11.6 Crash Data Linked to GIS

Most states already have an extensive and well-maintained database of crashes. Information from these databases may not be compatible with a GIS. There are many studies which required integration of a database with a GIS to perform an analysis. Brewer et al. (2000) discussed the benefits of integrating GIS with crash data and collision reports. A methodology was developed using GIS for the identification and evaluation of locations based on crashes. Identified locations were prioritized for safety improvements. The study focused on demonstrating GIS as a part of a systematic approach to enhance roadway safety.

Miller (1999) from the Virginia Transportation Research Council integrated a software package called Micro Traffic Records System, used by law enforcement agencies to record crashes, with a GIS to map crashes occurring at intersections or on streets. The study demonstrated the use of GIS as a tool to assist crash analysis at both a macroscopic and microscopic level. For example, Figure 2-6 illustrates a map showing crashes that occurred over a large area on weekdays between 10 a.m. and noon. This map informed the user about problem locations in the downtown area, along Route 29 and Route 250 Bypass and helped identify locations where more enforcement was required.

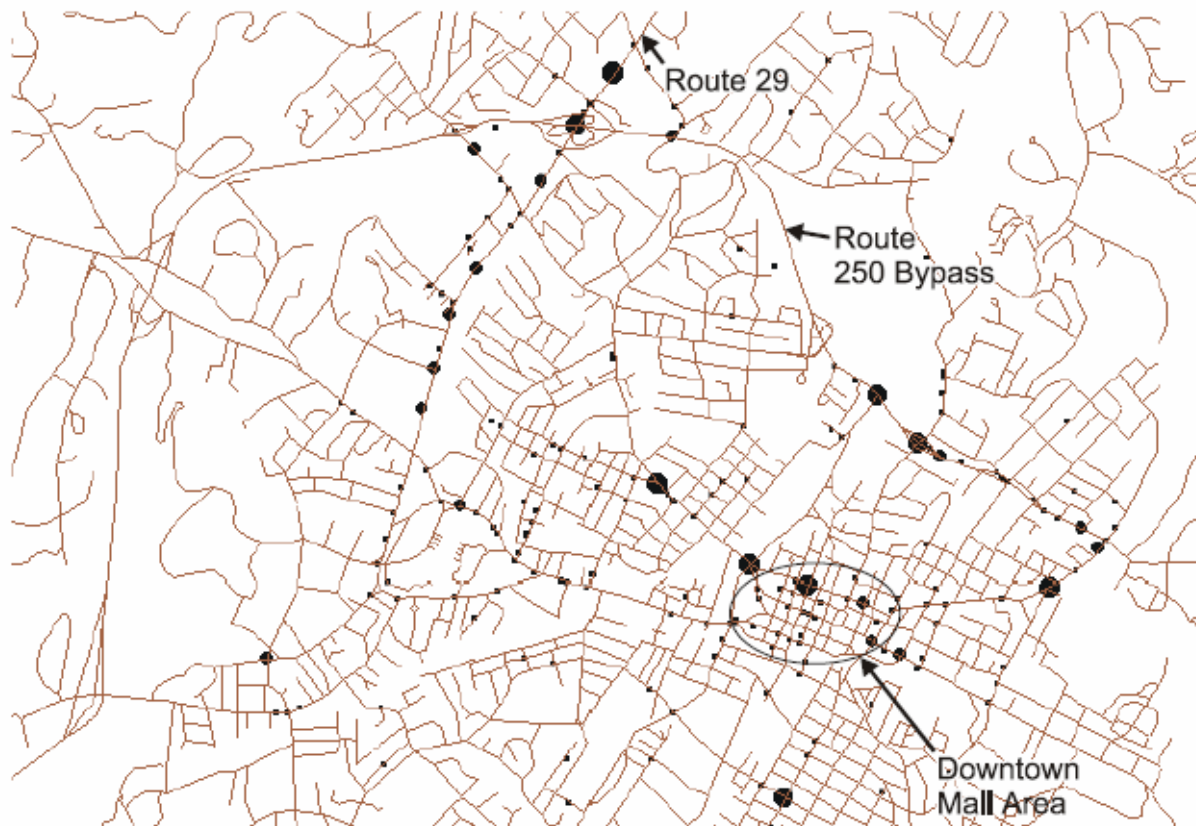


Figure 2-6. Crashes occurring on weekdays between 10 a.m. and noon, Charlottesville (Virginia)

Harkey (1999) developed a crash referencing and analysis system within a GIS for Wake County, North Carolina. Information associated with both crashes and roadway parameters were incorporated into the GIS. Other data bases included in the system were the traffic signal inventory, railroad grade crossing inventory, average daily traffic file, and pavement condition inventory. Various types of analysis were conducted using different programs integrated into the GIS such as a spot analysis program, strip analysis program, cluster analysis program, sliding scale analysis program and corridor analysis program.

2.11.7 Network Analysis

Network analysis, a relatively new feature in GIS, can help select routes. Smith et al. (2000) produced a report discussing the combination of GIS with traditional safety data. This report served as an educational document for GIS professionals and traffic safety engineers. Maine and Washington were selected for case studies in an attempt to understand GIS-based safety analysis tools. The report described GIS-based analysis such as intersection analysis, strip analysis, cluster analysis, corridor analysis, safe pedestrian route selection, bicycle compatibility of roads, and high pedestrian crash zones.

Figure 2-7 is a map with intersections selected according to the number of crashes, fatalities, injuries, etc. Figure 2-8 is a map displaying a safe route selected from an origin to school based on traffic and roadway safety. The report also discussed the identification of high truck crash road segments and high pedestrian crash zones.

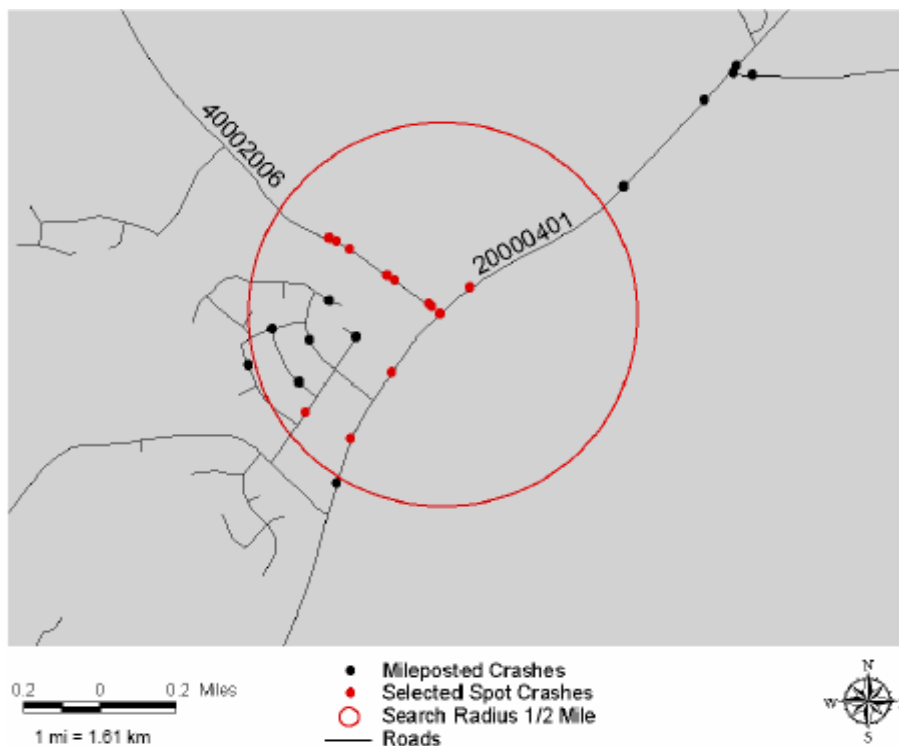


Figure 2-7. Map showing intersection crashes

Austin et al. (1997) applied GIS technology to develop methods that emphasize road safety improvements for children. One of the applications identified routes taken by children that walk to school, and analyzed those routes for safety. Harkey (1999) developed a separate program called the corridor analysis program that helped identify routes which were not designated for trucks, but which were used by trucks due to shorter routes to terminals.

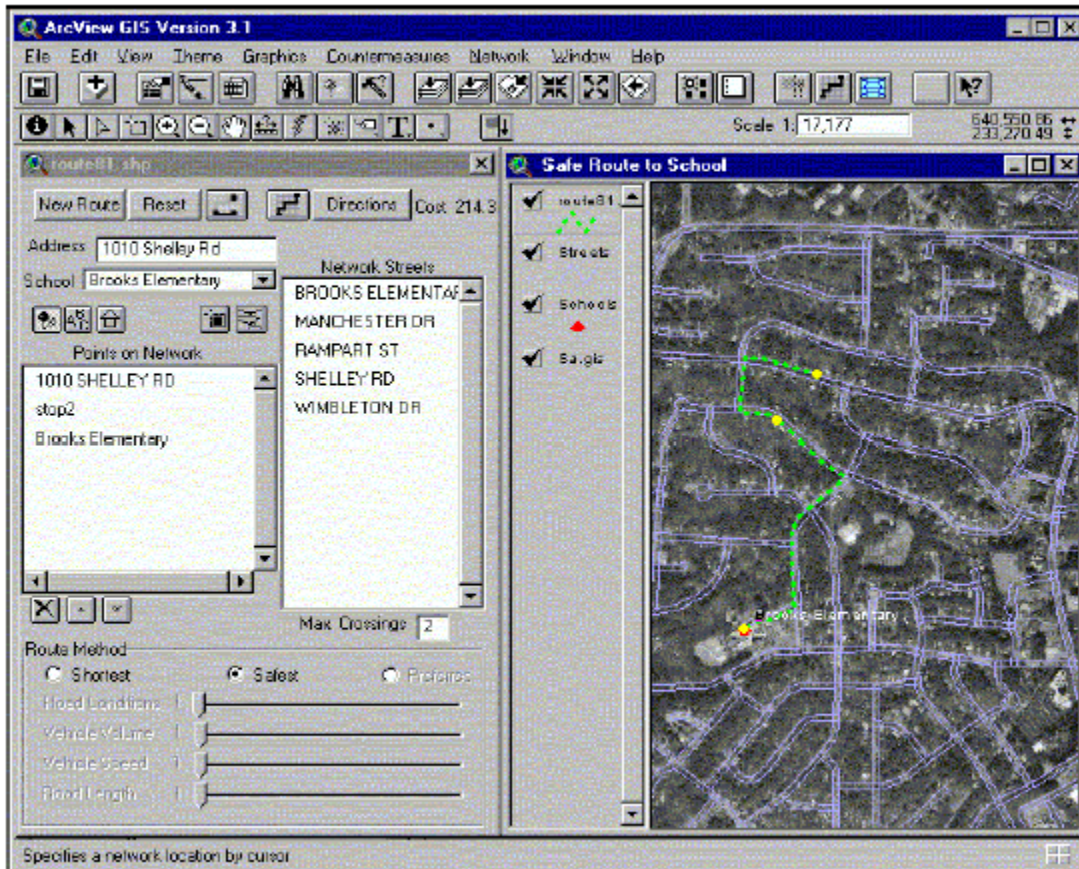


Figure 2-8. Map showing safe route selected from origin to school

2.12 Summary

A large amount of research has been conducted on combining GIS and crash databases for the purpose of displaying and analyzing crash data spatially. All of the reviewed research projects focused on improving community and traffic safety, and they were presented according to fundamental issues such as graphical representation of crash data, crash patterns, crash types, crash rates, traffic safety, crash data linked to GIS, and network analysis.

Section 3.0 Methodology

3.1 Introduction

This section describes the methodology adopted for mapping crash data in ArcGIS. The techniques used to connect DPS crash information with an ArcGIS basemap depended on the crash location information stored in the database. Different techniques were developed to map intersection, road segment, or route-milepost location data.

In the DPS database, crashes are stored in one of three ways depending on the location at which a crash occurred: nodes (intersections), links (road segments with a start node and end node), or route-milepost (a distance along a state highway or an Interstate). This information, along with the municipality where a crash occurred, is adequate to describe a crash location in words. Unfortunately, the location information in words cannot be easily displayed in a GIS, which requires a seamless coordinate system, such as latitude and longitude or the state plane coordinate system (northing and easting). Consequently, an alternative procedure was developed to establish a connection between DPS data and GIS location.

3.2 Crashes occurring at Intersections

3.2.1 Intersection Crash Recorded in DPS Database

Each intersection and segment crash record in Alabama has three fields describing the location at which the crash occurred. These fields contain information about the road segment (Link_ID) and may contain two intersections (Node_1 and Node_2). For an intersection crash, the value stored for the Node_2 field is zero, which denotes that the crash occurred at an intersection (Node_1) while the vehicle was traveling on a road segment (Link_ID). Figure 3-1 shows a portion of the attribute table of crash information extracted from the DPS database and opened in GIS. Records highlighted in Figure 3-1 indicate crashes that occurred at an intersection.

3.2.2 Intersection Mapping in GIS

For intersection crash mapping, a separate point shapefile or layer was created. For each intersection, nodes were screen digitized as point features, and node numbers were manually inserted in the attribute table in a field called "Node_ID". Paper node-link maps provided by ALDOT were used to obtain the node numbers for each intersection. All intersections in Tuscaloosa County, Alabama that had experienced a crash were digitized as described in Appendix A. After geocoding all the nodes in the county, the intersection shapefile was linked

with the DPS database. Figure 3-2 shows a map of a portion of the Tuscaloosa downtown area with all the nodes digitized. The entire process consumed approximately 220-person hours of work.

INJURIES_1	ROAD_DEFE	LIGHT_COND	WEATHER_C	POLICE_ARR	AIM_INTERS	LINK_ID	NODE_1	NODE_2	ConnectID
NO SECOND	NO SECOND	DAYLIGHT	CLEAR	46-60	INTERSECT	359	9141	0	9141
1	NONE	DARKNESS	CLEAR	11-20 MIN	INTERSECT	6016	71	0	71
2	NONE	DARKNESS	CLOUDY	1-10 MINU	INTERSECT	6	269	0	269
NO SECOND	NO SECOND	DAYLIGHT	SNOW	1-10 MINU	SEGMENTAL	5188	327	328	5188_327_328
NO SECOND	NO SECOND	DARKNESS	CLOUDY	1-10 MINU	SEGMENTAL	1995	8997	9039	1995_8997_9039
1	NONE	DAYLIGHT	CLOUDY	1-10 MINU	SEGMENTAL	5970	39	40	5970_39_40
NO SECOND	NO SECOND	DAYLIGHT	CLOUDY	0	INTERSECT	5168	276	0	276
NO SECOND	NO SECOND	DARKNESS	SNOW	11-20 MIN	INTERSECT	5185	372	0	372
NO SECOND	NO SECOND	DARKNESS	SNOW	21-45 MIN	SEGMENTAL	215	227	228	215_227_228
1	NONE	DARKNESS	SNOW	11-20 MIN	INTERSECT	6	16	0	16
NO SECOND	NO SECOND	DARKNESS	SNOW	21-45 MIN	SEGMENTAL	5840	3450	3457	5840_3450_3457
NO SECOND	NO SECOND	DARKNESS	SNOW	1-10 MINU	INTERSECT	5185	372	0	372
1	NONE	DAYLIGHT	CLEAR	1-10 MINU	SEGMENTAL	5172	155	348	5172_155_348
1	NONE	DAYLIGHT	CLOUDY	1-10 MINU	SEGMENTAL	215	165	166	215_165_166
1	NONE	DAYLIGHT	CLEAR	1-10 MINU	SEGMENTAL	215	166	167	215_166_167
1	NONE	DAYLIGHT	CLEAR	1-10 MINU	INTERSECT	5869	567	0	567
1	NONE	DAYLIGHT	CLEAR	OVER 180	INTERSECT	5168	265	0	265
1	NONE	DAYLIGHT	CLOUDY	1-10 MINU	SEGMENTAL	6	269	2025	6_269_2025
1	NONE	DAYLIGHT	CLEAR	1-10 MINU	INTERSECT	5449	1040	0	1040
NO SECOND	NO SECOND	DAYLIGHT	CLEAR	1-10 MINU	INTERSECT	216	694	0	694
NO SECOND	NO SECOND	DAYLIGHT	CLEAR	21-45 MIN	INTERSECT	5558	542	0	542
1	NONE	DAYLIGHT	CLEAR	1-10 MINU	SEGMENTAL	7	16	20	7_16_20
1	0	DARKNESS	SNOW	1-10 MINU	SEGMENTAL	1766	3501	3502	1766_3501_3502
NO SECOND	NO SECOND	DAYLIGHT	SNOW	1-10 MINU	SEGMENTAL	5449	1021	1025	5449_1021_1025
NO SECOND	NO SECOND	DARKNESS	SNOW	1-10 MINU	INTERSECT	13	372	0	372
1	NONE	DAYLIGHT	CLEAR	1-10 MINU	INTERSECT	69	11469	0	11469
1	NONE	DAYLIGHT	CLEAR	11-20 MIN	INTERSECT	6	65	0	65
1	NONE	DAYLIGHT	CLEAR	1-10 MINU	SEGMENTAL	69	1248	7349	69_1248_7349
1	NONE	DAYLIGHT	CLEAR	1-10 MINU	INTERSECT	5558	283	704	5558_283_704
1	NONE	DAYLIGHT	CLEAR	11-20 MIN	INTERSECT	7	92	0	92
2	NONE	DAYLIGHT	CLEAR	1-10 MINU	INTERSECT	5449	1039	0	1039

Figure 3-1. Attribute table showing crash records in CARE

3.2.3 Connecting Node Shapefile in GIS to Crash Attribute Table for Intersections

To connect the intersection node shapefile in GIS to the DPS data, intersection crashes first were separated from road segment crashes in the database. As each crash record had three different crash location fields, it was difficult to associate three fields in the database to one field in GIS. To overcome this problem, a new field named “ConnectID” was created in the attribute table for crashes. The procedure to create and populate the “ConnectID” column is described in Appendix B. All records that had a Node_2 value of “0”, which indicates an intersection crash, were identified, and the value of Node_1 was placed in the “ConnectID” field. This task was accomplished by writing a Visual Basic (VB) program in GIS.

The “ConnectID” field was then linked to the Node_ID field of the shapefile using the “Join” command in ArcGIS. Once the join is established, the data is ready for spatial analysis. For

other types of non-intersection crashes, additional strategies were developed as discussed in the following sections.

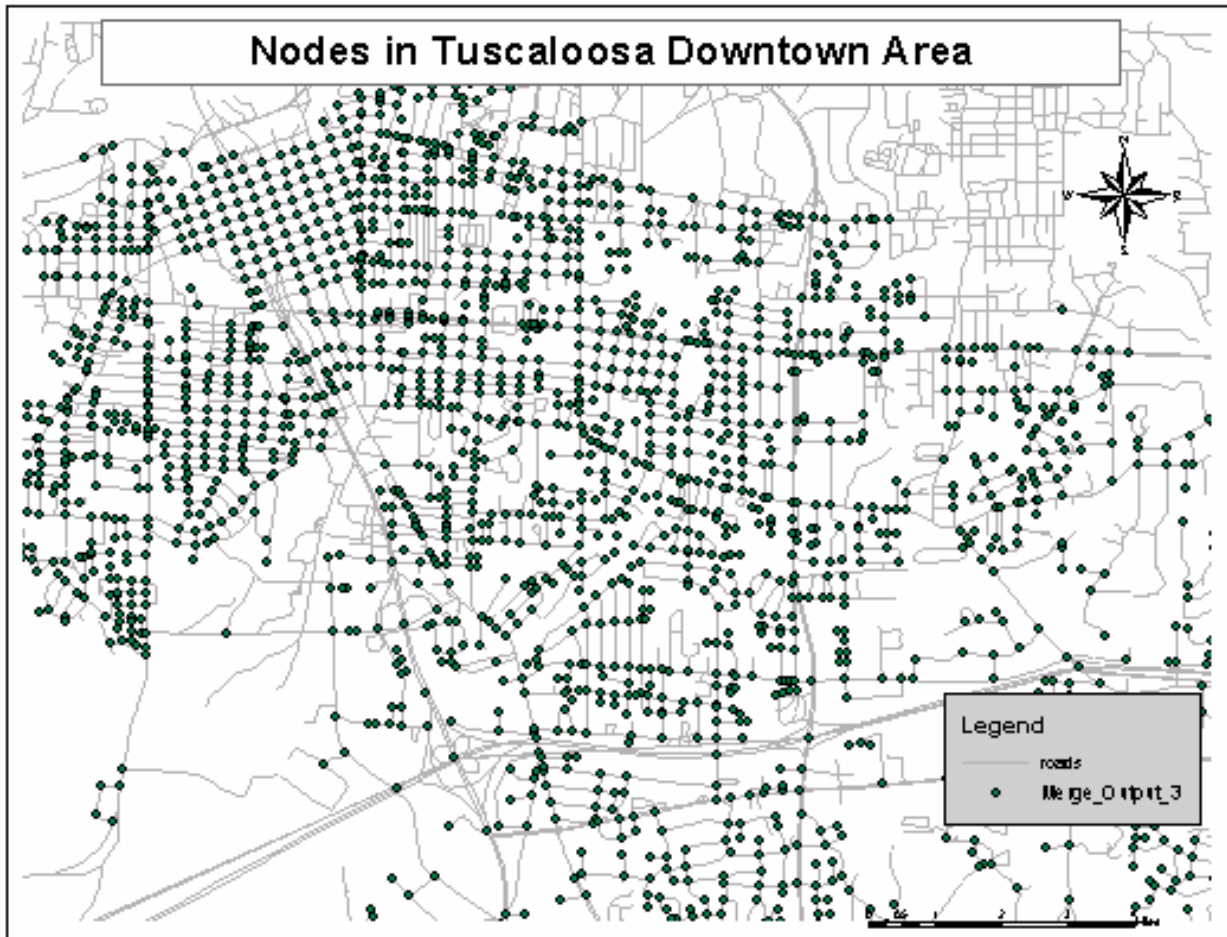


Figure 3-2 Map showing nodes digitized in the Tuscaloosa downtown area

3.3 Crashes Occurring on Road Segments and State Highways in Urban Areas

Crashes on state highways are recorded in the DPS database in two ways depending on whether the crash occurred in a rural area or an urban area. The crash location in a rural area is completely described by the state route number and the milepost at which the crash occurred. Crashes on state highways in urban areas are recorded in the same manner as crashes on city roads, i.e., according to Link_ID, Node_1, and Node_2. Urban area crashes on state highways were mapped in the same manner as the crashes on road segments. The crashes in rural areas were mapped using the route-milepost system, which is explained in Section 3.4 “Crashes Occurring on Interstate and State Highways in Rural Areas.”

3.3.1 Road Segment Crash Recorded in DPS

The location of a crash that occurred along a road segment is described by three fields: Link_ID (road segment on which the crash occurred), Node_1, and Node_2 (two intersections at the ends of the road segment). All three fields are required to describe the location completely.

Evaluation of the DPS location data for crashes along road segments revealed these facts:

- Node_1 value is always lower than Node_2.
- State highways always use the Link_ID as the state highway number. For example, State highway 6 has Link_ID 6.
- County roads do not have the same Link_ID as the county road number.

3.3.2 Road Segments in GIS

The base map for this project contained a shapefile of roads in Tuscaloosa County. As described in Appendix A, it was necessary to modify the roads layer and break each road at each intersection so an individual road segment between two intersections could be uniquely identified. To uniquely identify a road segment, columns named Link_ID, Node_1 and Node_2 were added to the attribute table of the roads layer in GIS. While entering the node numbers for each link, care was taken to enter the lower number in the Node_1 field and the higher number in the Node_2 field. This allowed connection between the crash attribute table and the attribute table of the roads layer. Editing this layer to geocode the map was the most time consuming aspect of this project, using approximately 330 person-hours of effort.

Geocoding was done in GIS using values from the node-link paper maps provided by ALDOT. Four different problems arose during geocoding the links of the road layer in GIS. First, in some cases, the line features representing the road between two intersections were in multiple segments. Before the link and node numbers were inserted for these roads, the line features were merged into a single segment so that the crashes associated with that segment were mapped to one line feature. The second problem involved missing nodes between two different road segments. In this case, the line representing two different road segments existed as a single entity. To solve this problem, the segment was manually broken into two separate segments, and a node was inserted at the intersection location. The third case involved missing road segments on the roads layer. In this case, the road segments were manually digitized at the location where they exist on the paper maps. The fourth case involved road segments existing on the road layer but not on the paper maps. In this case, the columns for link and node numbers in the attribute table were left empty, and the segments were kept on the layer.

3.3.3 Connecting Road Shapefile to DPS Database

As mentioned previously, the location descriptions for road segments in the GIS and the crash location in the DPS database were defined by three fields: Link_ID, Node_1 and Node_2. Connecting the crash attribute table to the roads shapefile attribute table required matching three different fields in the database with three different fields in the GIS attribute table. If the three fields match, then the data will be connected. Unfortunately, the “Join” tool in GIS is not designed to connect two attribute tables based on three common fields.

To overcome this problem, the attribute fields Link_ID, Node_1, and Node_2 were concatenated into a single entity, and the concatenated values were placed in a new column called ConnectID for both the attribute table of roads layer and the database table. To distinguish between the three values in a concatenated field, the three fields were separated by underscores. For example, a segment of road along link 500 between nodes 1001 and 1002 was coded in the database as 500_1001_1002. In this way, the two tables have a common “ConnectID” field. The far right column of Figure 3-1 shows the “ConnectID” column and the concatenated values. The procedure to concatenate fields of data is described in Appendix B.

A small program was coded to determine if a crash occurred at an intersection or along a road segment. Crashes that have no value for Node_2 are crashes that occurred at intersections. The records representing intersection crashes are not required to be concatenated with a road segment to uniquely identify the crash location: the node number alone is sufficient to represent an intersection crash. Thus, the location data in the ConnectID column of the crash database table was selected so that if the value of Node_2 for a particular record was “0”, the program only entered the value for Node_1. If the value of Node_2 was not “0”, the concatenation of Link_ID, Node_1, and Node_2 was entered into the ConnectID column. Therefore, the ConnectID column can be used to connect crashes on both nodes and links.

3.4 Crashes occurring on Interstates and State Highways in Rural Areas

3.4.1 Interstate and State Highway Crashes (Rural Area) Recorded in DPS

A crash location on an Interstate or on a rural state highway is completely described by a route number and the milepost at which the crash occurred. The milepost distance can have up to two digits of accuracy after the decimal. The data for Interstate and state highway crashes are mapped separately for intersection and segmental crashes on rural highways. Figure 3-3 shows a sample of Interstate crash data with location data highlighted.

OID	ACCIDENT	DATE	COUNTY	CITY	LINK	MILEPOST	NODE1
205	3009129	1/31/1993	TUSCALOOSA	TUSCALOOSA	I359	1	96
206	3015052	3/2/1993	TUSCALOOSA	TUSCAL RUR	I359	1	95
207	3016726	2/26/1993	TUSCALOOSA	TUSCALOOSA	I359	71	96
208	3019789	3/17/1993	TUSCALOOSA	TUSCALOOSA	I359	2	3
209	3026749	4/8/1993	TUSCALOOSA	TUSCALOOSA	I359	2	97
210	3027134	4/8/1993	TUSCALOOSA	TUSCALOOSA	I359	2	97
211	3041336	5/17/1993	TUSCALOOSA	TUSCAL RUR	I359	2	91
212	3044682	5/14/1993	TUSCALOOSA	TUSCALOOSA	I359	1	91

Figure 3-3. Table showing crashes on Interstate I-359 with milepost

3.4.2 Interstate and State Highways in GIS

The GIS road layer contains a line feature for every Interstate and rural state highway. Each line feature is one single continuous entity. At the ends of the line feature, a start and end milepost value were added manually. Crashes that occurred at any milepost within the range of the start and end milepost numbers were interpolated and plotted at the distance along the road.

Mapping was accomplished using a feature in GIS called *dynamic segmentation*. This tool requires each line feature to have the “M” measure turned to the “active” mode so that the software identifies the milepost values. The default for the “M” measure is “inactive.” The setting for the “M” measure cannot be changed to “active” after a shapefile is created. Therefore, to use the *dynamic segmentation* feature, a new shapefile with the “M” measure activated was required. The procedure for creating a shapefile with “M” active is described in Appendix A.

Once a shapefile is created with the “M” measure activated, the roads are ready to map crashes as events represented by points. The attribute table of the shapefile must have a column containing the route number, which will be matched with the route number in the crash attribute table.

3.4.3 Connecting Route File to DPS Database for Interstates and Rural State Highways

Crashes on Interstates and rural highways were mapped to line features by adding the crash attribute table to the GIS map as described in Appendix B. Figure 3-4 shows a map of all the events, rounded to the nearest full milepost, on the Interstates and rural state highways in Tuscaloosa County.

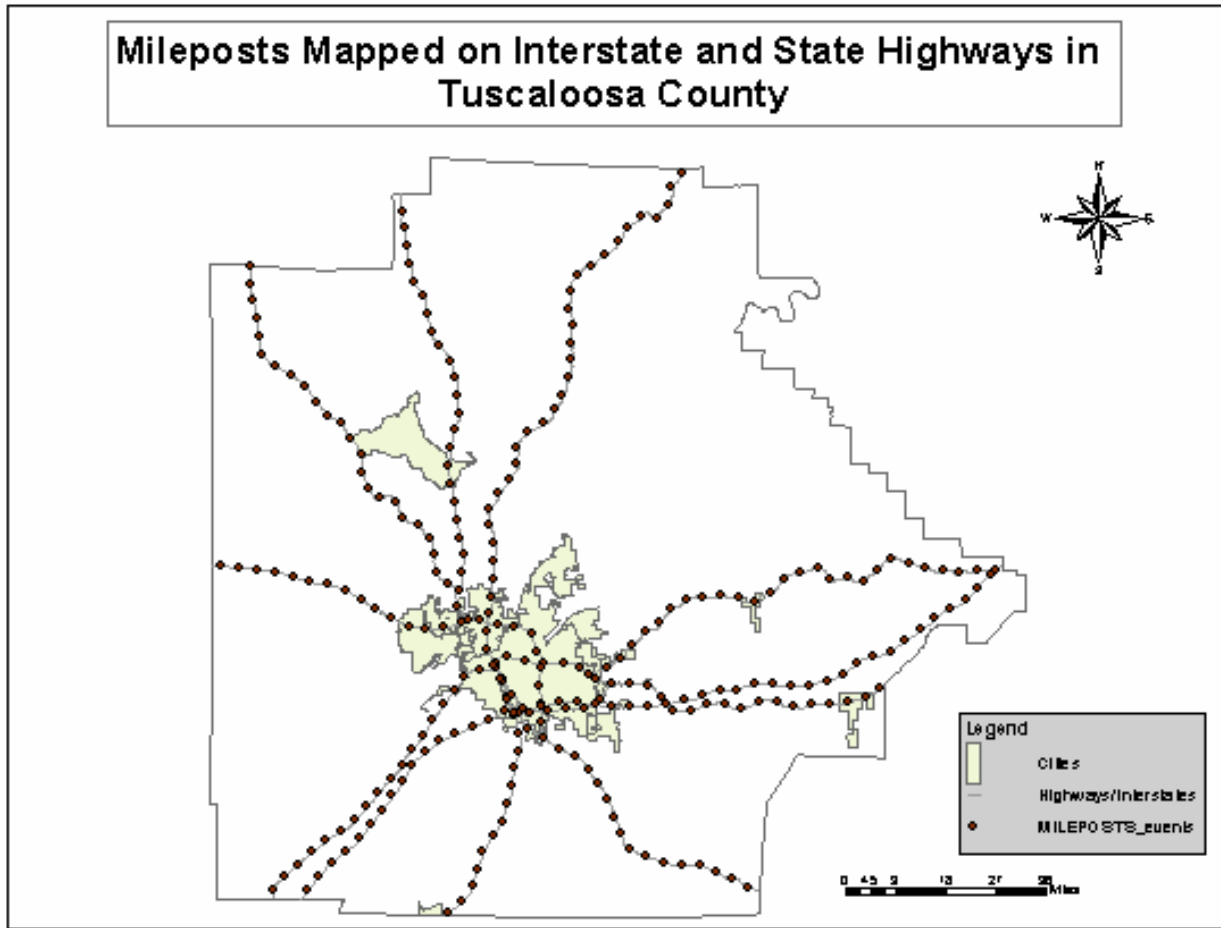


Figure 3-4. Milepost mapped along highways using linear referencing

3.5 Summary

Three different types of crash locations are stored in the DPS database: nodes, links, and route-milepost. All three types of data were connected to GIS and mapped in this project. Intersection crashes were mapped by connecting the crash attribute table to the “nodes” shapefile created by screen digitizing. The location information for road segments (Link_ID, Node_1, and Node_2) were inserted manually in the “roads” shapefile and combined together in a new column called “Connect_ID”, which was then matched to a similar field in the crash attribute table. After matching the fields, crashes were mapped based on the DPS crash information connected to the GIS table. The two joined tables enabled spatial analyses to be performed that will be discussed in the following chapter. Crashes on Interstate and rural state highways were mapped in GIS according to the route and milepost at which they occurred.

Section 4.0

Spatial Analysis and Display of Crash Data

4.1 Introduction

This section explains spatial analyses performed in GIS on crash data, and the results are displayed as maps. Multi-year crash data for Tuscaloosa County were mapped and spatially compared to existing roadway features such as bridges, crossroads, railroad grade crossings, etc.

4.2 Analyses and Display of Intersection Crashes

Forty-four percent of crashes are reported to have occurred at an intersection (Frank et al. 2001). This supports a growing concern in the transportation and safety community towards intersection safety. The first step in improving intersection safety is to identify “hotspots” or intersections that are possible candidates for safety treatment.

The GIS application developed in this project can play an important role in identification of hotspots. Hotspots can be identified with respect to factors such as number of crashes, crash severity, areas with vision obstruction problems, etc. The following subsections present examples of hotspot identification using those factors.

4.2.1 Total number of intersection crashes

The number of crashes occurring at an intersection is the simplest way to identify hotspot intersections. However, every intersection has different average daily traffic (ADT), and streets with high ADT have a larger chance for crashes compared to streets with low ADT. Even with this caution, the total number of crashes provides a good representation of problem areas across a city. Figure 4-1 shows a thematic map of crashes at intersections in Tuscaloosa for the year 2001. Larger circles represent increasing numbers of crashes at an intersection. Intersections with no crashes are not assigned a symbol. This map provides a good visual comparison of intersections along a corridor or major street with similar ADT values.

The number of crashes can also be mapped depending on factors such as crashes during rush hour, crashes occurring on Fridays, crashes due to vision obstruction, crashes on rainy days, crashes due to deteriorated pavement condition, crashes due to DUI, etc. Figure 4-2 shows the percentage of crashes that occurred at intersections in the same area during evening rush hour and helps in identifying hotspots during times of high traffic flows and maximum congestion.

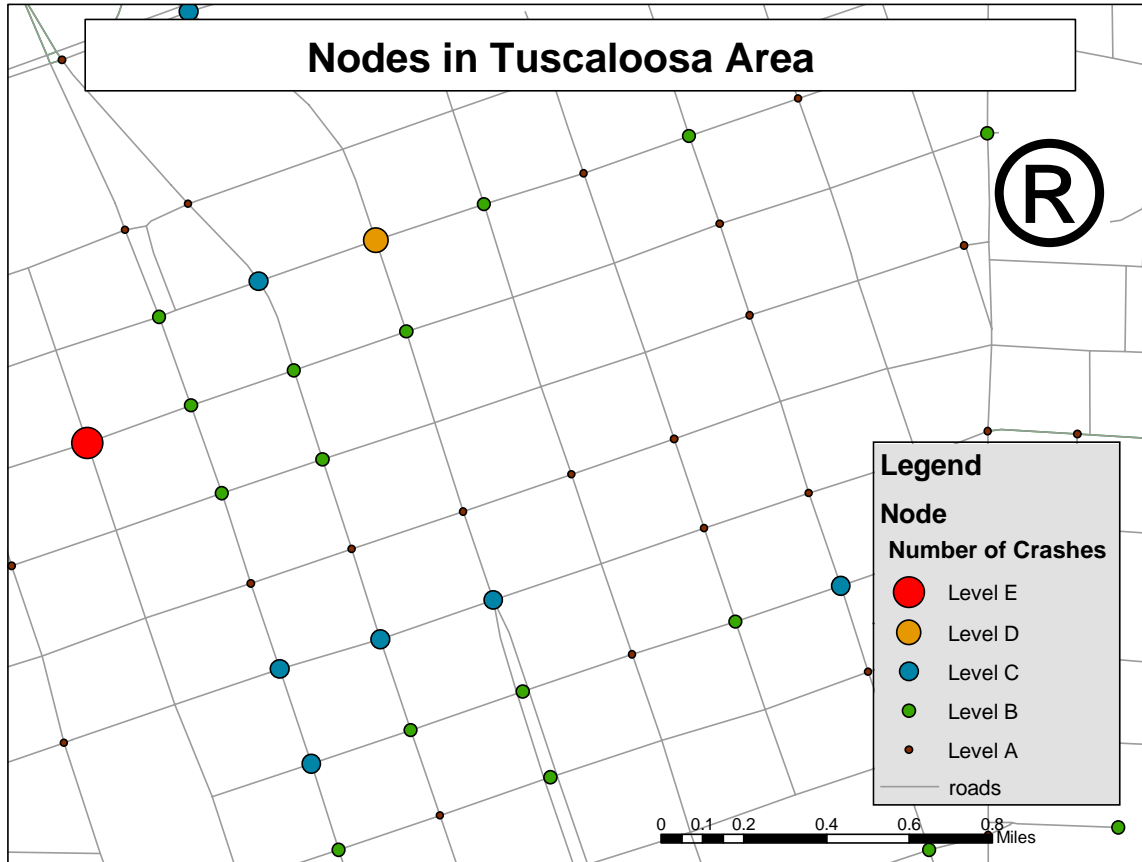


Figure 4-1. Thematic map of intersection crashes in Tuscaloosa area

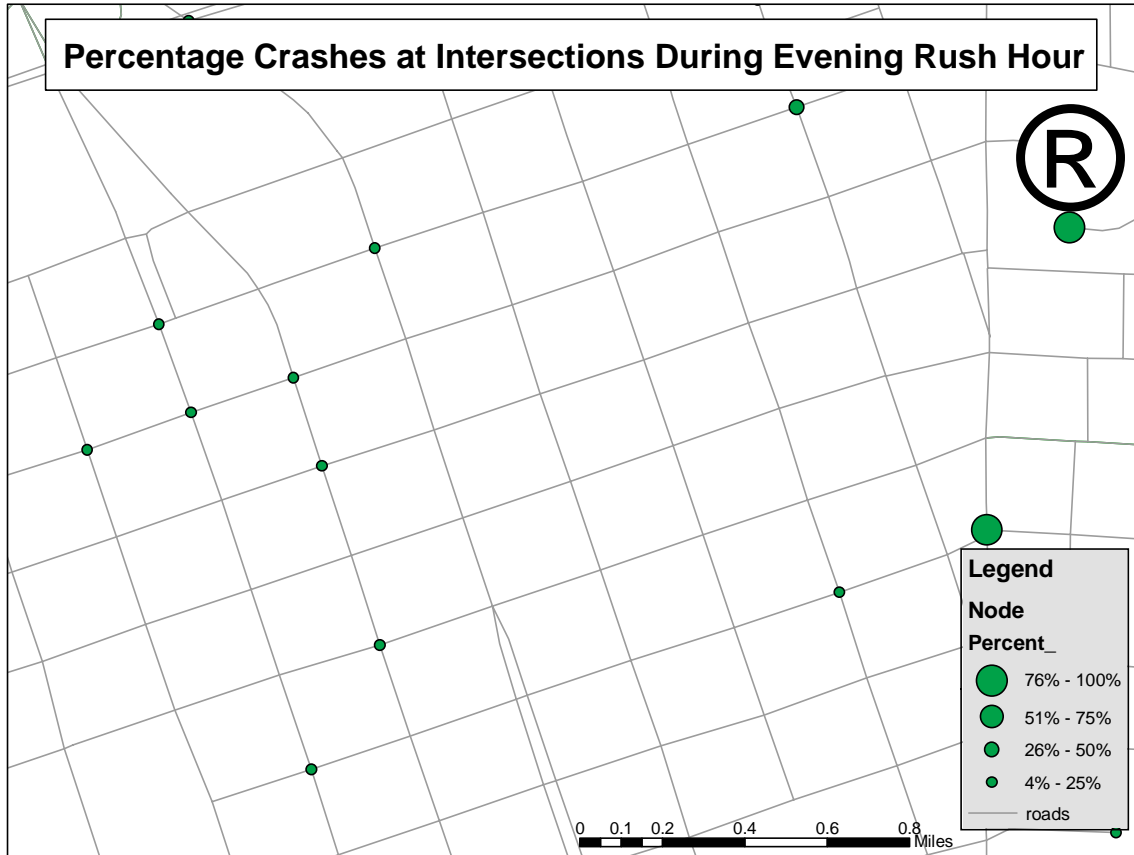


Figure 4-2. Percentage of total crashes occurring at intersections during evening rush hour

4.2.2 Crash Severity

A vehicle crash may cause fatality, injury, or property damage only (PDO). Hence, a technique was developed to map crash severity as 3 points for fatalities, 2 points for injuries, and 1 point for PDO. (Note: this scale was used for example purposes only and does not represent a recommended severity scale.) These severity numbers can be added together and summarized in a single field. Figure 4-3 summarizes the severity of crashes at intersections in Tuscaloosa. This thematic map identifies the locations of highest severity by giving more emphasis to fatality and injury crashes than to PDO crashes.

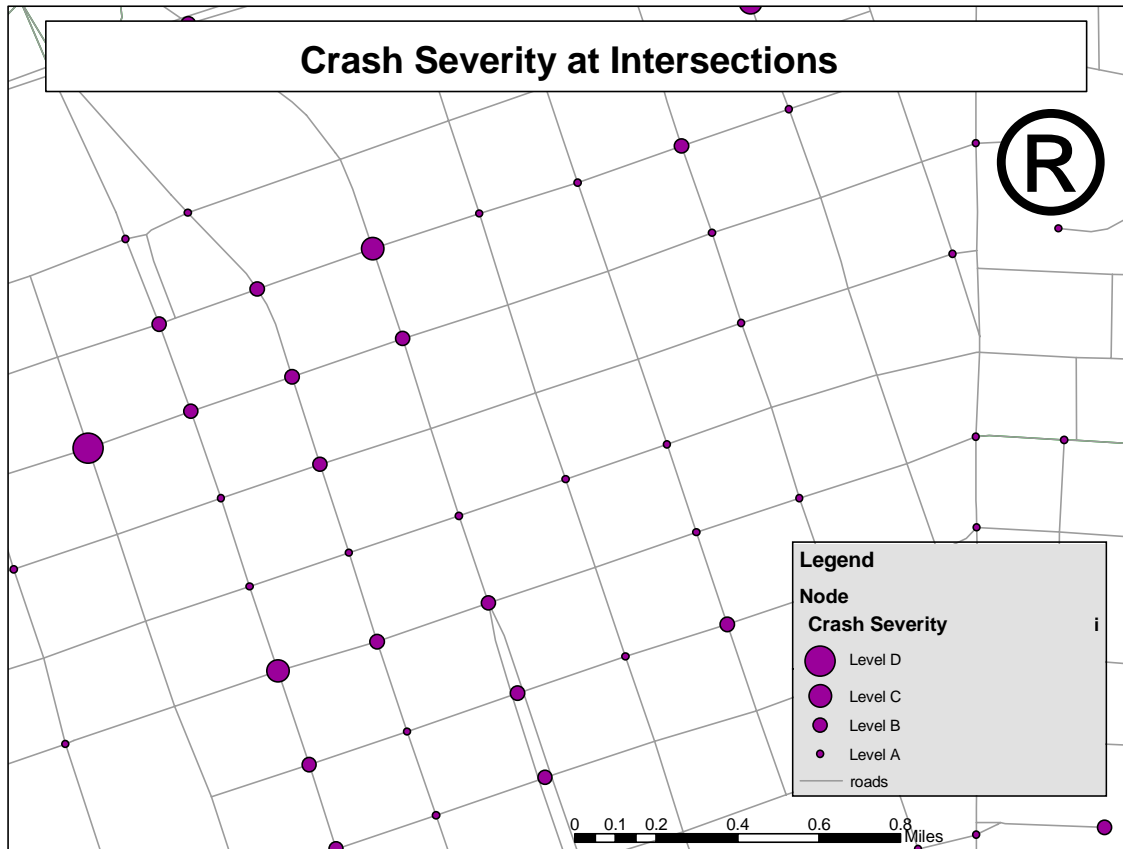


Figure 4-3. Thematic map for crash severity at intersections

For intersections having a similar ADT range and a similar number of crashes, maps such as Figure 4-3 should be of great help, as they identify candidates for safety treatment based on fatality, injury, and PDO.

4.3 Analyses and Display of Segmental Crashes

Road segment crashes occur at a lower frequency than crashes at intersections because the number of conflicting movements of vehicles is lower on road segments. Hotspot identification for road segments can be mapped just as it was mapped for intersections. In these maps, links representing road segments are highlighted with different colors or line weights to indicate the number of crashes. The following subsections describe road segment thematic maps and how they are used for hotspot identification. As with intersections, hotspots for road segments can be identified based on factors such as number of crashes, time of day, and injury severity.

4.3.1 Total Number of Segmental Crashes

Mapping the total number of crashes that occur on road segments provides an overview of road safety in an area. Thematic mapping of crash frequency is effective for corridors because road segments within a corridor frequently have similar ADT values. Crashes on road segments can

be mapped based on different factors such as weekend crashes, rush hour crashes, rainy-day crashes, DUI crashes, crash severity, crashes at a distance from highways, etc.

Figure 4-4 displays a thematic map of crashes on road segments that occurred on Fridays in Tuscaloosa. The total number of crashes that occurred on a segment is represented by the width of the segment. High-crash segments are readily apparent. Road segments with no crashes are represented as thin gray lines.

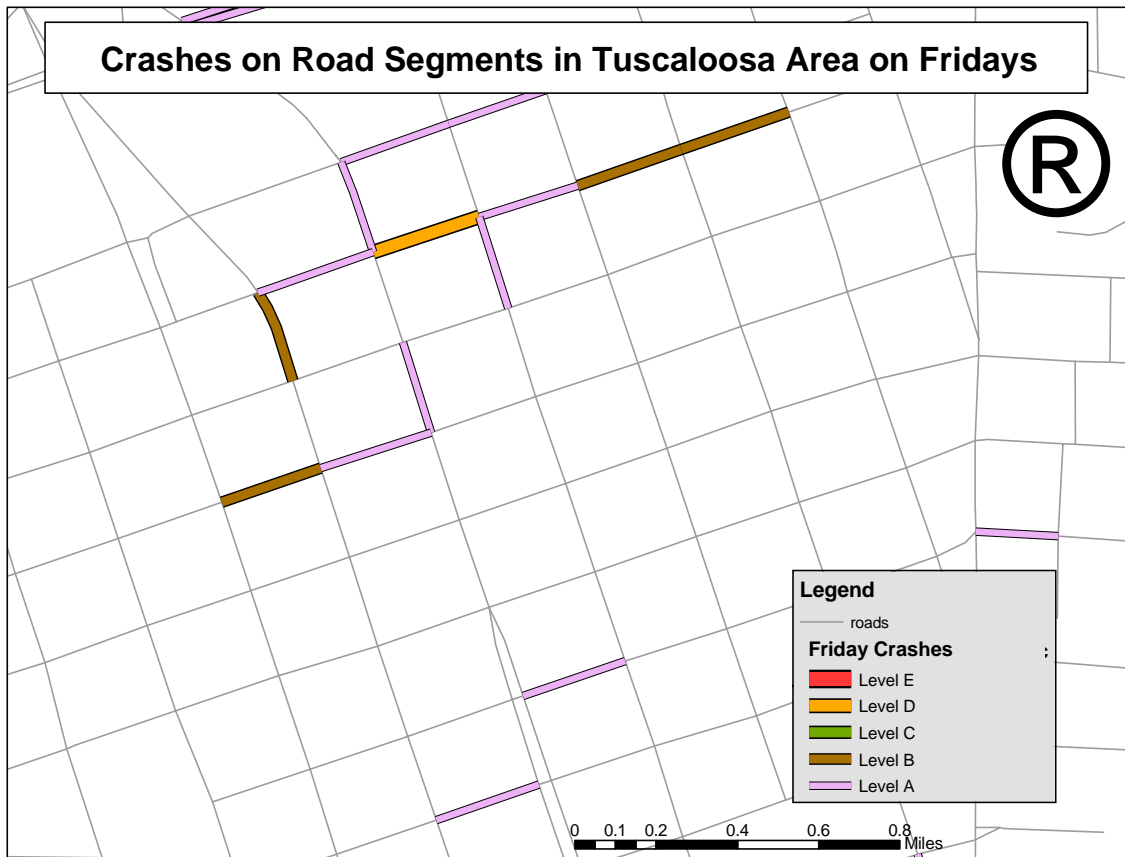


Figure 4-4. Crashes on road segments in Tuscaloosa area on Fridays

A GIS analysis feature called buffering can be employed to select crashes that occur within a specific distance of existing roadway feature such as railroad grade crossings, bridges, etc. This analysis has the potential to identify roadway features that may contribute excessively to crash occurrences. As an example, Figure 4-5 is a map displaying a 0.25-mile buffer around major roadways. All of the road segments that fall within the buffer distance are selected, and the crashes related to only these road segments are thematically represented. A similar buffering technique can be used to select nodes instead of links.

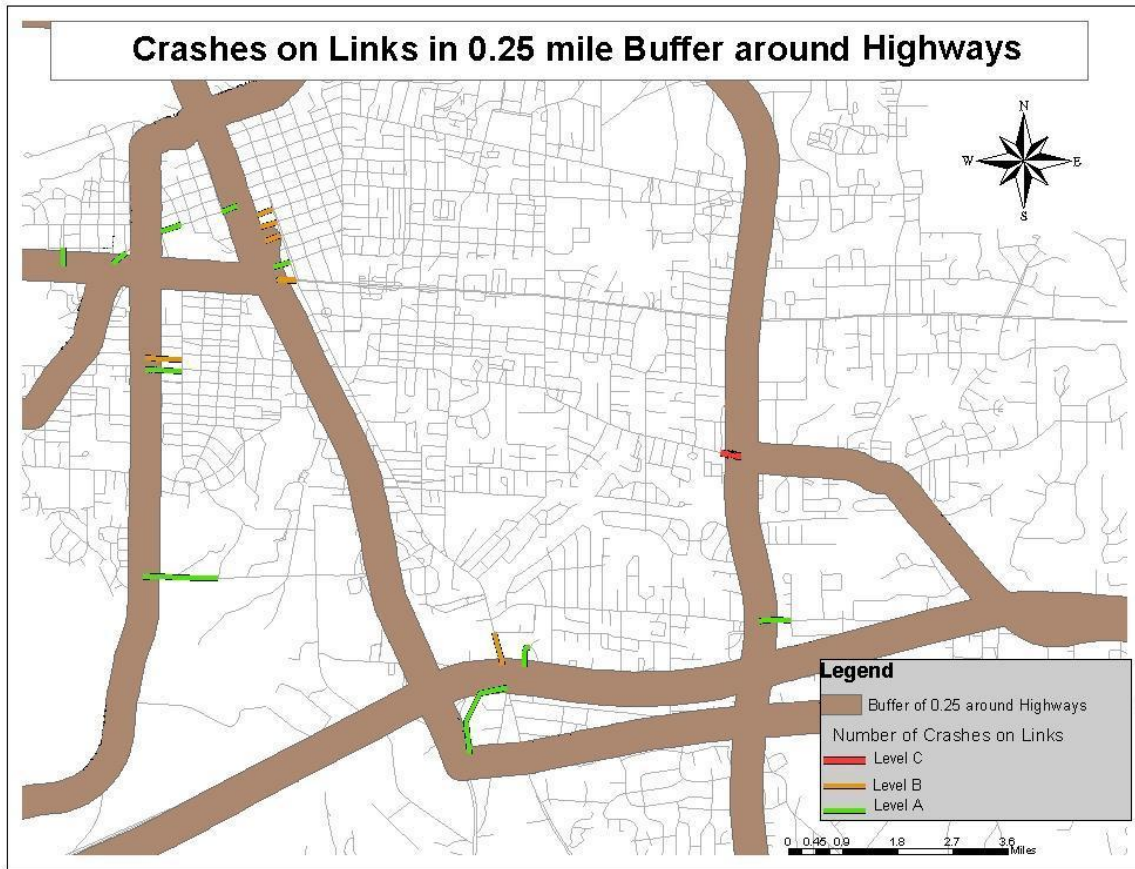


Figure 4-5. Crashes on road segments within 0.25-mile buffer around highways

4.4 Analyses and Display of Highway Crashes

Crashes that occur on Interstate or state highways are generally more severe than intersection or segmental crashes due to the high speed of vehicles on highways. Crashes on highways are stored in the DPS database according to the route and milepost at which they occur. The Interstate highways that run through Tuscaloosa County are I-20/59, starting at milepost 45 and ending at milepost 101, and I-359, starting at milepost 0 and ending at milepost 2.5. Other United States highways and state highways in Tuscaloosa County are US-43, US-11, US-82, S-6, S-7, S-13, S-215, S-216, S-69, and S-171.

The following subsections describe maps and analyses that were performed for highway crashes. As mentioned in Section 3, the database table for highway crashes was added to the GIS, and then the locations were mapped as events.

4.4.1 Total Number of Crashes

The CARE system supplied mileposted crash locations that were truncated after the decimal point. Thus, all highway crashes in this report are mapped at a milepost. Figure 4-6 shows the

total number of crashes at each milepost for all Interstate and state highways in Tuscaloosa. Larger circles represent a larger number of crashes. Areas of higher total crashes are easily identified.

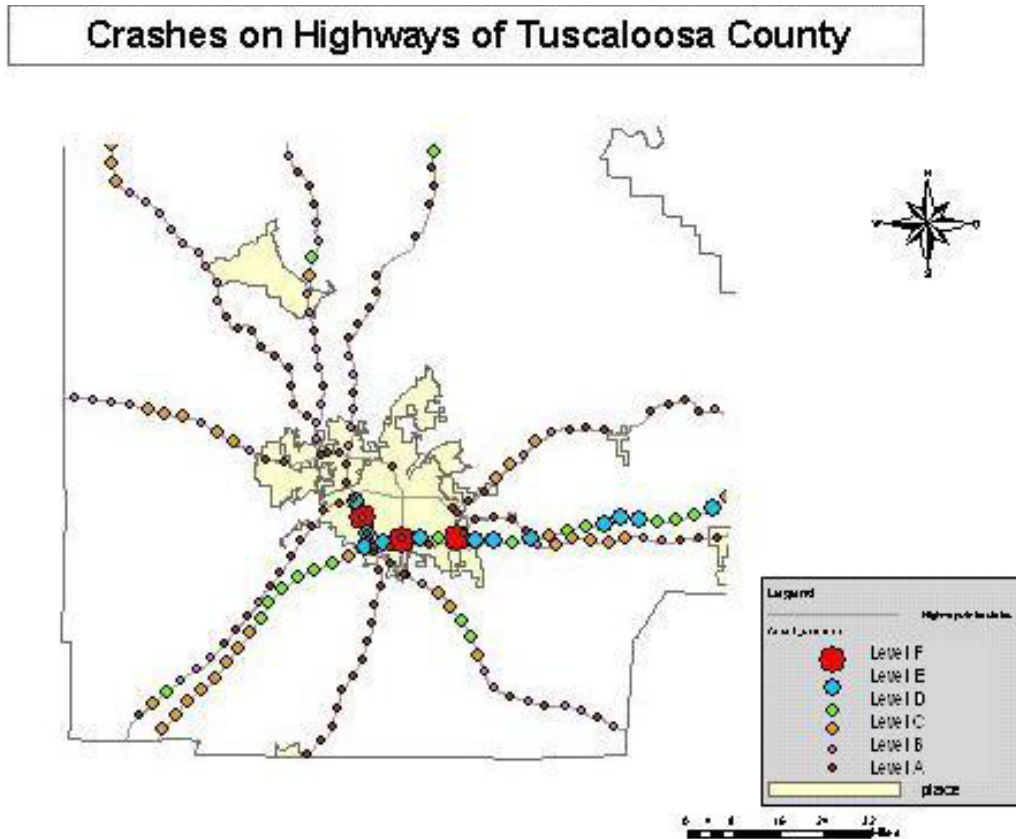


Figure 4-6. Crashes on Interstate and state highways according to mileposts

The number of crashes that occurred due to driving under the influence during 1996-2000 is shown in Figure 4-7. This map may indicate the location of roadway conditions, such as sharp curves, that are more likely to cause a crash involving intoxicated drivers.

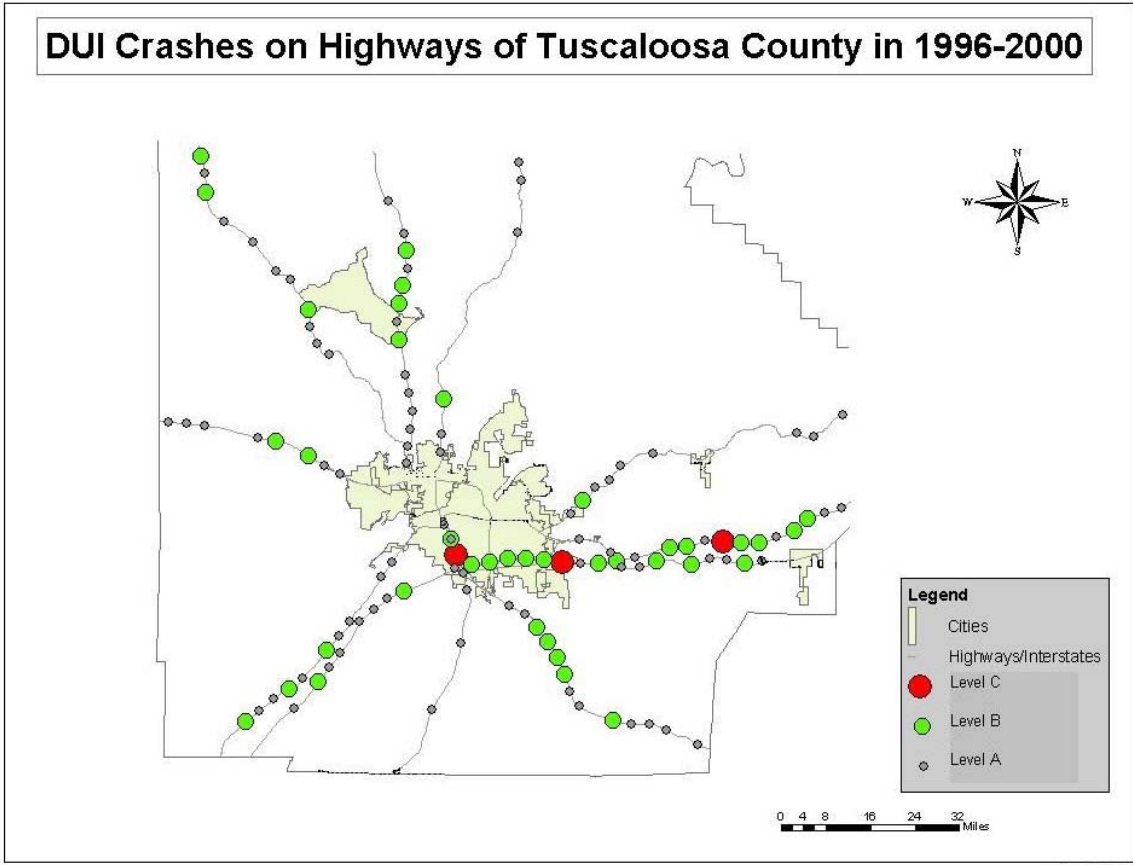


Figure 4-7. Crashes on highways due to driving under influence during 1996-2000

4.4.2 Crash Severity

Figure 4-8 shows a map displaying crash severity on Tuscaloosa County highways from 1996 to 2000. This map expresses the severity of crash in terms of fatality, injury, or PDO. Larger circles represent fatal crashes, medium circles represent injury crashes, and the smallest circle represents PDO crashes. Very few highway road segments did not sustain some form of crash during this time frame.

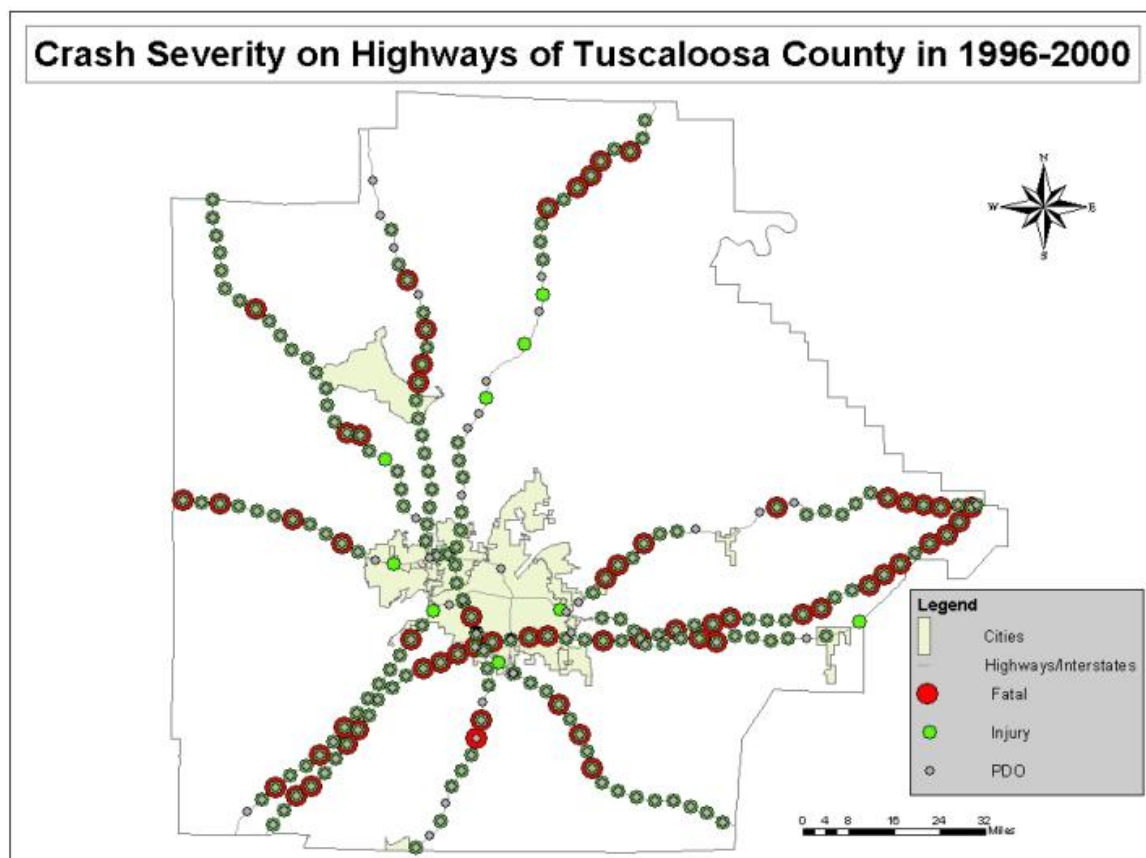


Figure 4-8. Crash severity on highways of Tuscaloosa County during 1996-2000

4.4.3 Crashes on I-20/59 normalized by ADT Data

The shapefile for highways originally did not contain ADT data, but displaying it on maps along with the number of crashes facilitates a better understanding of crash frequency. ADT for a highway changes approximately every two to three miles. Adding ADT data required breaking every highway into segments and entering the ADT value for each segment. For this work, ADT data was added only to Interstate I-20/59.

Figure 4-9 shows the total number crashes in 2001 at each milepost on I-20/59 (points) along with ADT data for each segment of the interstate (adjusted line width). As seen on the map, segments of I-20/59 passing through the City of Tuscaloosa have high ADT, in the range of 41,500-46,500 vehicles per day. These locations also have a high number of crashes. Figure 4-10 normalizes crashes by ADT, highlighting segments that have a lower ADT but a higher number of crashes. Such a map would be helpful to traffic safety personnel when they are identifying hotspots using the “crash rate” method

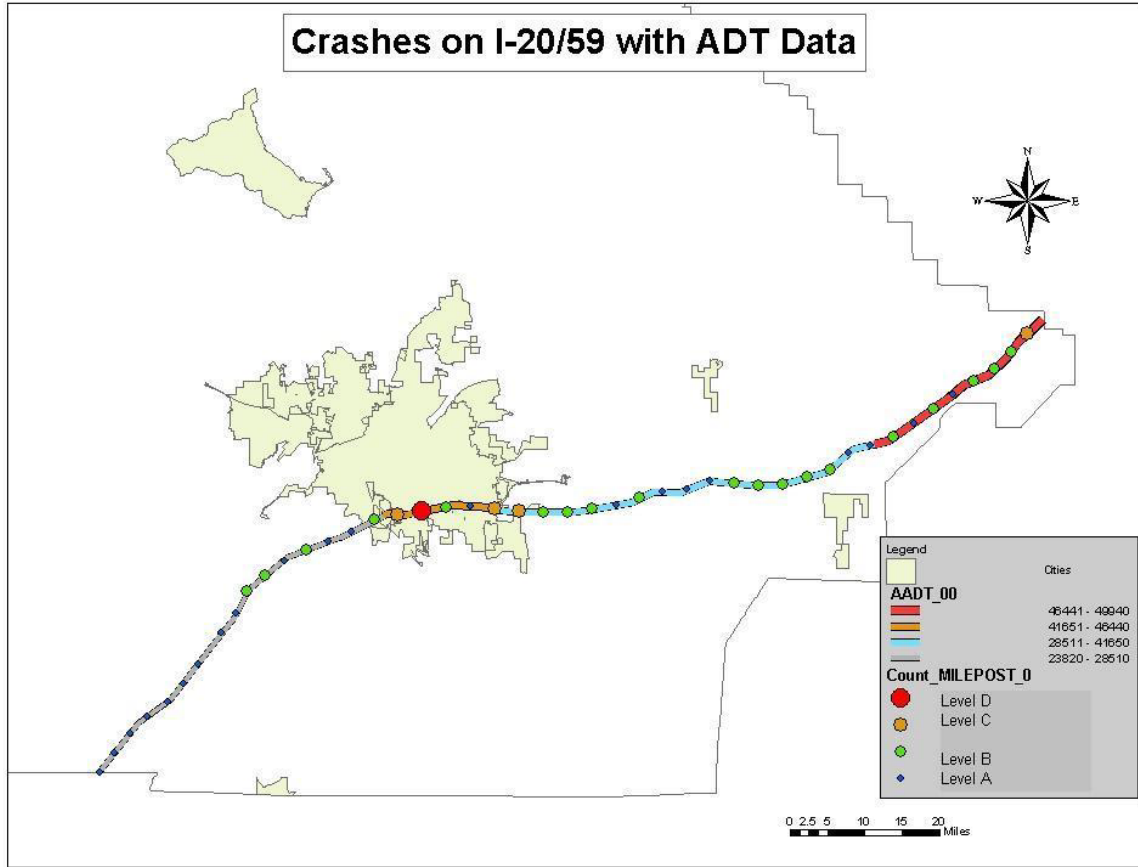


Figure 4-9. Crashes on I-20/59 with ADT data for 2001

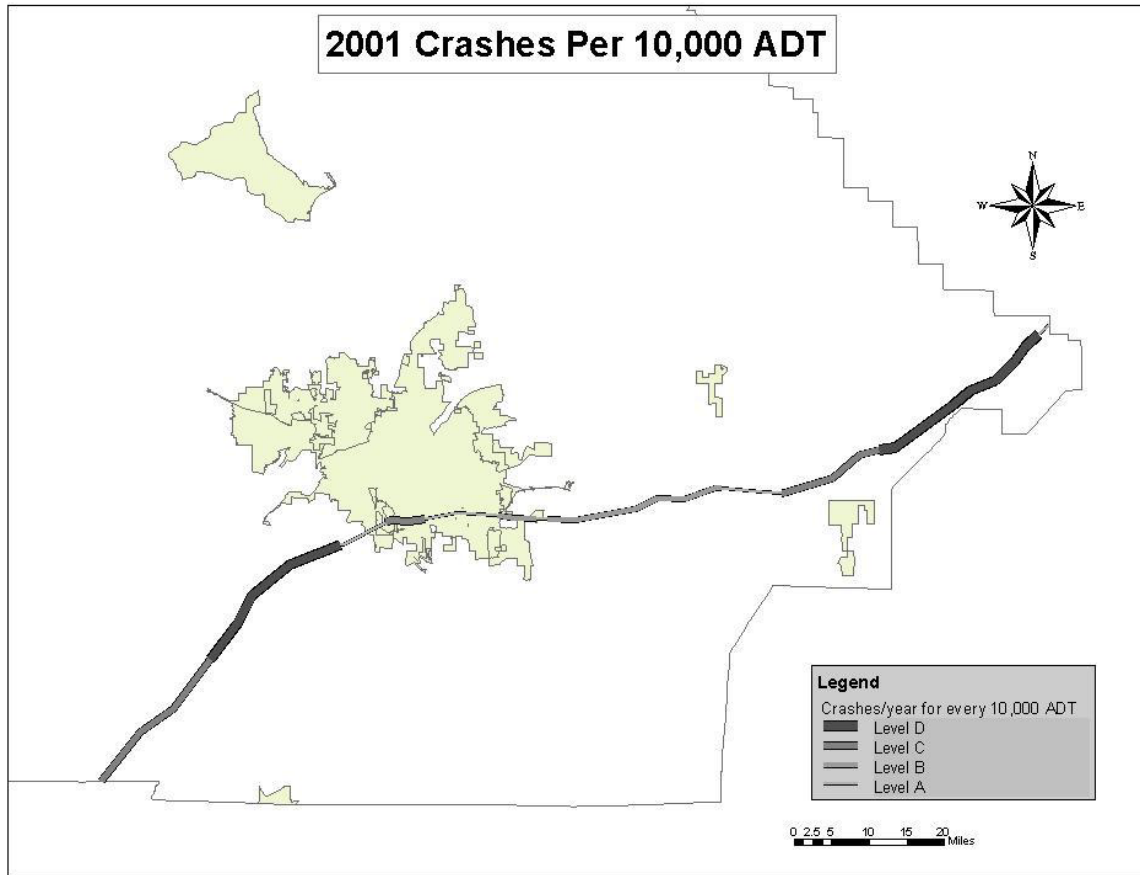


Figure 4-10 Crashes on I-20/59 in 2001 normalized by ADT

4.5 Summary

This section of the report presented methods for representing crash data using the GIS applications developed during this project. Techniques such as thematic mapping and buffering were used to display and analyze crashes spatially. These maps add value to the DPS data by providing a spatial framework in which to view and analyze the data.

Section 5.0 Network Analysis

5.1 Introduction

This section describes a network analysis developed in this project to select safe travel routes. Safe routes are links and nodes (roads and intersections) with the least number of crashes between two locations. Safest routes were determined by spatially analyzing two shapefiles: 1) roads represented as lines and 2) intersections represented as points. The method used for mapping crashes according to streets and intersections is explained in detail in Section 3.

A typical purpose for conducting a network analysis is to determine a path between two locations that exhibits minimum travel time or minimum distance. These analyses can be achieved by setting the travel time or the distance as the impedance in a network. The analysis selects a path with the minimum impedance and gives the fastest or shortest route from origin to destination.

This project utilized a network analysis to determine the safest path between two locations (origin and destination) selected by the user. Recent crashes that occurred on roads and at intersections are considered impedance in this analysis. The final output identifies a path with the minimum number of crashes. This procedure allows a traveler to select a historically safe route between an origin and a destination.

5.2 Methodology

For any network analysis, the first step is to build a geometric network with feature classes such as roads and intersections. Such networks are called geometric networks for the simple reason that feature classes have geometry and can be mapped. A geometric network contains connectivity information between links (road segments) and nodes (intersections) and defines rules of behavior for connectivity, e.g., what class of roads may be connected to a particular class of intersections. The *Build Geometric Network Wizard* and the *Utility Network Analyst* found in the ArcGIS software are the tools used to create and analyze the network.

Creating a geometric network requires three steps: editing the link attribute table, converting the data into a feature dataset within a geometric network (if necessary), and using the *Build Geometric Network Wizard*. Each step is described in detail in Appendix C.

The *Utility Network Analyst* toolbar chooses a path according to the weights assigned to links and nodes. Weights are designated attributes of the links and nodes that are used to compare different paths. The GIS looks at each value in the designated attribute column and chooses the

lowest value. If “length of links” is chosen as the weight, the *Utility Network Analyst* will select the shortest path between two points. If “travel time” is chosen as the weight, the *Utility Network Analyst* will select the fastest path if speed information is included in the attribute table. For the current application, “number of crashes” was chosen as the weight, which enabled the *Utility Network Analyst* to select the path of links and nodes with the fewest recent crashes, which is the safest path.

5.3 Network Analysis

Analysis of the network can be performed once the geometric network has been created, and is accomplished by using the *Utility Network Analysis* toolbar and choosing the number of crashes as the weight. Two paths may have the same number of crashes for the same period of time. To select between them, the length of the path is considered as a secondary factor for selection of route, and the route with shorter length will be selected. This results in the selection of the shortest - safest path.

Most drivers use length as the most important parameter when selecting routes; therefore, runs were made using “length” as the weight. A comparison was made between selected routes with length as the weight and selected routes with the number of crashes as the weight. The comparison revealed that the route with absolutely the lowest number of crashes might be far longer than a route with just a few crashes more. For example, if there are two routes with 10 and 11 crashes, and the route with 10 crashes is 10 miles longer than the route with 11 crashes, it may not be logical to travel 10 miles extra to avoid one additional crash. A proper balance is required between the number of crashes and the length of the selected route.

More realistic routes can be selected with an optimum combination of route length and number of crashes. This can be achieved through trial and error by using different combinations for the balance of weights between the number of crashes and the length of a route. It is important to note that weights cannot be added to an existing network, and only one network can be created per feature dataset. In order to “add” weights, the entire process must be repeated, and a new network must be created.

Figures 5-1, 5-2, and 5-3 are maps showing routes produced with the geometric network developed in this study. Figure 5-1 exhibits the shortest route from the MIB building on the University of Alabama campus to a popular restaurant at the corner of University Boulevard and 21st Street in Tuscaloosa, AL. Figure 5-1 used length of route as the weight. The selected path has 164 crashes and is 5.01 miles long.

Figure 5-2 shows one of the safest routes between the two locations. This route was selected using number of crashes as the weight. The selected path has 62 crashes and is 9.27 miles long. Note that this path has over one hundred crashes fewer than the shortest path but is almost twice as long.

Figure 5-3 shows the shortest of the safest routes between the locations. This route was produced using a combination of the number of crashes and the length of the path as the weight. The selected route has 62 crashes and is 8.51 miles long. Note that this path has the same number of crashes as the path in Figure 5-2, but is 3/4 of a mile shorter. Table 5-1 shows a comparison of number of crashes and path lengths for different weights used in Figures 5-1 through 5-3.

In real life, most drivers would still choose the path in Figure 5-1. However, the exercise demonstrates that safer route selection can be performed using DPS data.



Figure 5-1. Map showing path selected based on length of route as weight

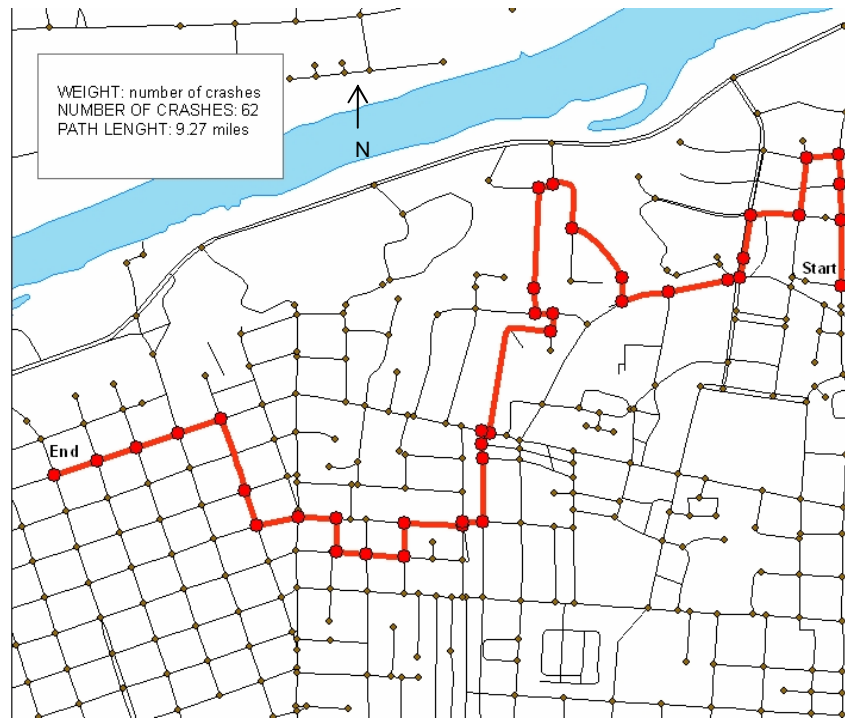


Figure 5-2. Map showing path selected based on number of crashes as weight

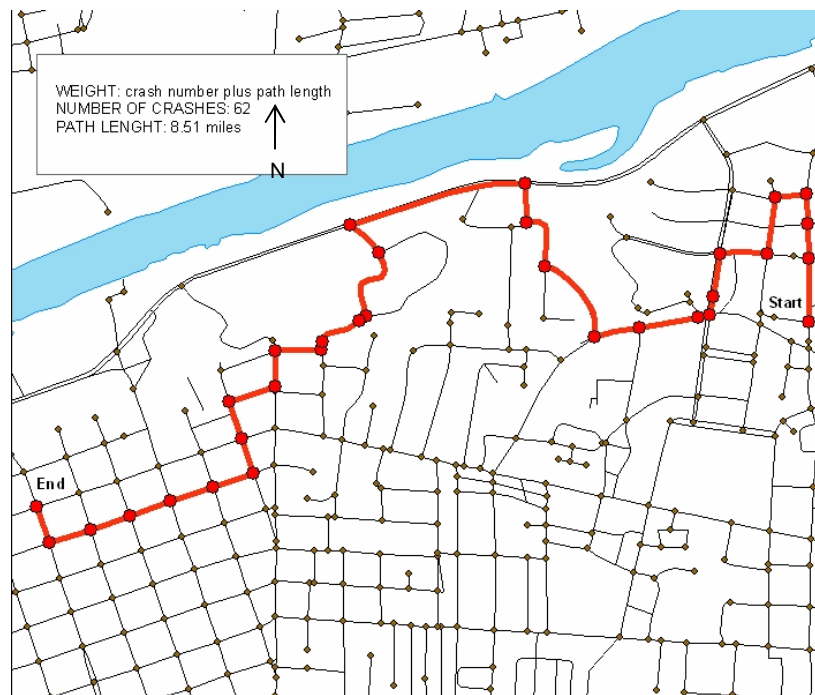


Figure 5-3. Map showing path selected based on a combination of length of route and number of crashes as weight

Table 5-1. Summary of safe route maps

Map Figure Number	Weight	Number of Crashes	Route Length (miles)
Figure 5-1	Path length	164	5.01
Figure 5-2	Crashes	62	9.27
Figure 5-3	crash_length	62	8.57

5.4 Summary

This project employed a network analysis to determine the safest, shortest, and shortest of the safe paths between two user-defined locations. A combination of the number of crashes and the length of the route was considered as impedance during the analysis. The result of the analysis consisted of maps illustrating selected routes with minimum number of crashes. This procedure allows the traveler to select a route that is relatively short and less susceptible to crashes.

Section 6.0

Duplicate Node Numbers

6.1 Introduction and Problem Statement

Much of the work for this project involved geocoding a GIS basemap with node and link information. Upon completion, it was noticed that 215 intersections improperly had been assigned previously-existing node numbers by the City of Tuscaloosa. These duplicate node numbers were compared to both digital and paper maps provided by ALDOT. Figure 6-1 shows a map of intersections with duplicate and even triplicate node numbers in downtown Tuscaloosa. Two intersections with the same node number are shown as a small circle, while three intersections with the same node numbers are shown with large circles. Each of these intersections was investigated to ascertain the cause of the duplication error.

6.2 Matching Paper Maps, Digital Maps, and DPS Data

To determine the cause of duplicate intersection node numbers, a combination of paper maps, ALDOT digital maps, and DPS data were examined. It was found that over time new node numbers are added to the existing node-link system when a crash occurred at an intersection that previously did not have a crash, and sometimes node numbers were assigned twice. Additionally, some nodes kept the same number for many years but were then assigned new numbers, so that older crashes at an intersection may have one node number while newer crashes at the same intersection have a different node number. In the DPS database, every crash at an intersection is assigned a node number and a link number indicating the direction of travel. For every intersection with duplicate node numbers, crash records were pulled from the DPS database, and the true location of the crash was determined.

Different categories of node numbering errors were identified. Each intersection having a duplication error was assigned an error category number. Figure 6-2 shows a thematic map of the 18 different categories of errors. Each intersection numbering error has two or three intersections involved. Therefore, either none, one, two, or three intersections match either the most recent ALDOT map, the DPS data, neither, or both. Details of each error category are listed below:

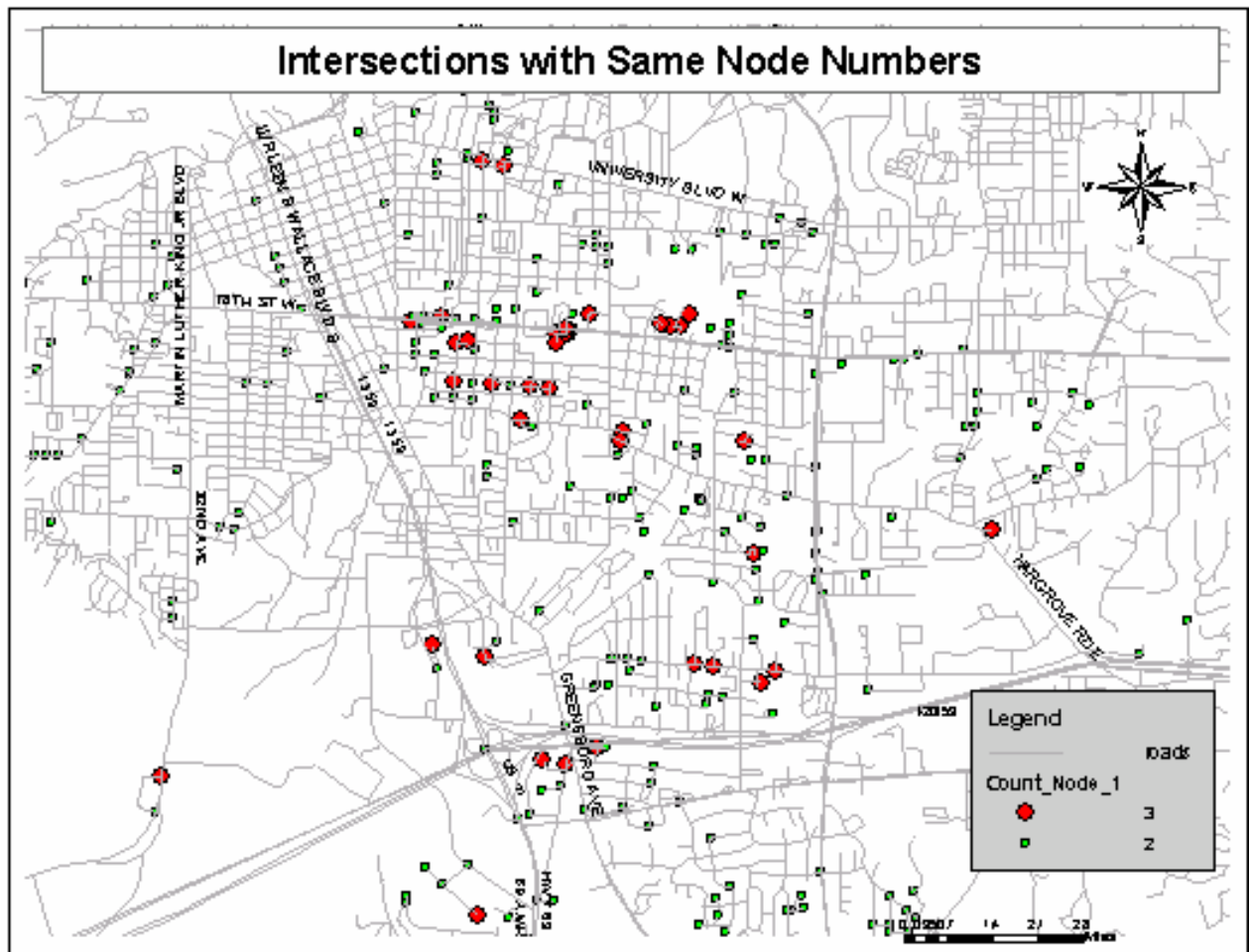


Figure 6-1. Map showing intersections with same node numbers

- Category 1: None of the node numbers match the ALDOT data or the DPS system. (2 nodes)
- Category 2: Digitizing error. (2 or 3 nodes)
- Category 3: One of the node numbers matches the ALDOT data, but none of the nodes match the DPS database. (2 nodes)
- Category 4: One of the node numbers matches the ALDOT data, and the node that does not match matches the DPS database. (2 nodes)
- Category 5: Both node locations match the ALDOT data, and only one node location matches the DPS database. (2 nodes)
- Category 6: One of the node numbers does not match the ALDOT data, and the other node number does not match the DPS database. (3 nodes)
- Category 7: Two node numbers match the ALDOT data, and two of the node numbers do not match the DPS database. (3 nodes)
- Category 8: One of the node numbers matches the ALDOT data, and only one node number matches the DPS database. (3 nodes)
- Category 9: Two of the node locations match the ALDOT data, and only one location matches the DPS database. (3 nodes)

- Category 10: Both node numbers match the ALDOT data, and none of the numbers match the DPS database. (2 nodes)
- Category 11: Both node numbers match the ALDOT data, and both node numbers also match the DPS database. (2 nodes)
- Category 12: One of the node numbers matches the ALDOT data, while the other number does not exist on the ALDOT map, and the one that matches the ALDOT digital map number also matches the DPS database. (2 nodes)

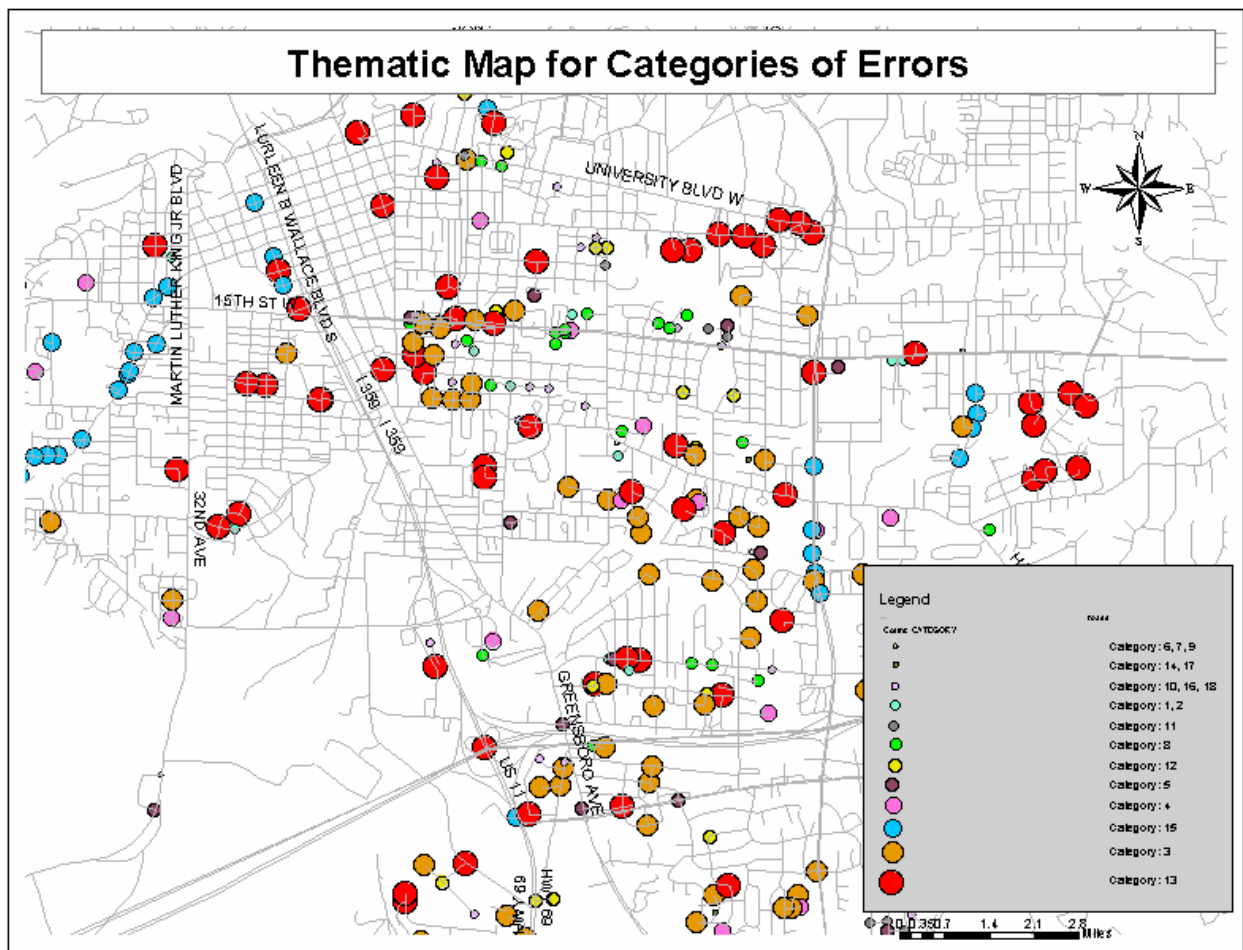


Figure 6-2. Thematic map for categories of errors

- Category 13: One of the nodes numbers matches the ALDOT data, and the same number matches the DPS database. (2 nodes)
- Category 14: Neither of the node numbers matches the ALDOT data, and one of the numbers matches the DPS database. (2 nodes)
- Category 15: One of the node numbers matches the ALDOT data, and both of the node numbers match the DPS database. (2 nodes)
- Category 16: A node number for one of the node locations is changed, but the new node number matches another node number at a different location, and the node

location which matches with the ALDOT data also matches the DPS database. (2 nodes)

Category 17: One of the node numbers matches the ALDOT data, and the second node was not found on the map. (2 nodes)

Category 18: One of the node numbers matches the ALDOT data, and none of the other numbers matches the DPS database. (3 nodes)

Categories 1 through 5 and categories 10 through 17 are for cases where two nodes have the same node number. Categories 6 through 9 and category 18 are for cases where three nodes have the same node number. Figure 6-2 shows a thematic map of the number of cases for each category of error. The maximum numbers of errors were for category 13 (59 intersections) followed by category 3 (51 intersections). There were just five cases where the error was due to improper digitizing, and these errors were corrected.

Category 1 errors were left unchanged. These errors should be corrected when ALDOT assigns a new number to the nodes. Category 2 errors were rectified when detected so that appropriate crashes could be mapped. Category 3 node numbers that matched the ALDOT data were kept unchanged, and the node that did not match was deleted. The node that was deleted will be recreated once a new node number is assigned by ALDOT. Nodes with category 4 through 12 errors and categories 14 through 18 were left unchanged, which will misrepresent some crashes. The node numbers will be corrected after ALDOT has assigned a new number to one of the nodes. Until ALDOT assigns new node numbers, records for crashes associated with these node numbers should be studied to match the links associated to the nodes. The node numbers associated with category 13 errors which matched the ALDOT data and the DPS data were retained, and the others which did not match with any data were deleted. The node number for the deleted node will be updated after ALDOT assigns a new node number.

6.3 Summary

This section of the report demonstrates the error checking advantage of the GIS application developed during this project. The investigators found that there are duplicate node numbers that need to be corrected to avoid confusion caused by misrepresentation of the data. This GIS application proved to be an effective tool for the detection of such errors, and displaying the errors on a map provides a clear understanding of the error locations. Moreover, the type and number of errors can be easily presented through thematic mapping.

Section 7.0

Conclusions and Recommendations

7.1 Introduction

This final report described the development and testing of a new GIS application for the display and analysis of crash data stored in the DPS database. The primary objective was to map crashes and correlate them with existing roadway features such as crossroads, curves, and railroad grade crossings. This application was developed to enable DPS data to be presented graphically on maps, which is not possible in CARE alone. The application takes advantage of both the CARE and GIS software capabilities. Reports can be created in CARE for detailed study, while representative maps can be generated in GIS. Crashes can be selected according to parameters such as time of day, DUI, pavement conditions, number of lanes, etc. This selection can occur in either CARE or GIS.

This project accomplished the following tasks:

- Developed techniques to better represent and analyze multi-year crash data on maps that also correlate with roadway features,
- Performed spatial analysis on crash data that are mapped using the GIS application to identify “hot spots” at intersections, road segments, and mileposts along highways,
- Identified the route with least impedance (crashes) between an origin and destination using the *Network Analysis Feature* within ArcGIS, and
- Identified node numbering errors in the system.

7.2 Conclusions

From this study, the following conclusions are drawn:

- The GIS crash application provided an alternative for the display and analysis of crash data from DPS as compared to the linear representation in CARE. The new GIS application allowed the user to view crash locations on a map, which gives a spatial representation of the data. The application provides an overview of crash locations at intersections, on road segments, or at mileposts along highways.
- With the buffering feature, the GIS application allowed DPS data to be correlated with existing roadway features like cross roads and railroad grade crossings, which cannot be done with CARE alone. This ability enables identification of locations where roadway features may cause an increase in the number of crashes. Crashes could also be correlated with schools, malls, theaters, bars, and other public places that experience increased traffic at specific times of the day.

- Thematic mapping of crashes for intersections, road segments, and mileposts along highways can play an important role in identifying possible candidates for safety treatment, which are often called “hot spots.” These hotspots are in the form of nodes, links, or milepost along a state or Interstate highway. Crash data may be mapped by total crashes, by rate of crashes, by severity of crashes, by weather conditions, and by many other parameters in the DPS database.
- Network analyses can be performed on a network of roads to find the safest travel path. Here, crashes can be considered as impedance, and road segments having the least number of crashes will be selected.
- There are 215 intersections in Tuscaloosa County that were assigned non-unique node numbers. This GIS application worked as an excellent tool to detect these errors by thematically mapping intersections with the same node number. Numbering errors were categorized into 18 different groups that were thematically mapped.
- The use of multi-year data enables “before and after” analysis to be performed and the results represented graphically. This type of data representation could be more effective than tables and charts to identify crash patterns through time.

7.3 Recommendations

The following recommendations are put forth to enhance the ability of the GIS application developed in this project:

- Only Tuscaloosa County was considered for this project, and this application should be considered for use in the entire state of Alabama. Implementation will not be labor-free: approximately 550 person-hours were consumed to geocode intersections, road sections, and highways in Tuscaloosa County, so innovative geocoding techniques that consume less time should be sought.
- Each of Alabama’s 67 counties may have multiple municipal divisions, and the statewide base map should be geocoded with the municipal codes, which will require considerable coordination effort. The DPS database already stores the municipal data for each crash.
- The State should consider two enhancements to the existing system:
 - Hyper link a digital photograph or video log of each intersection to each intersection node. Such images can be helpful during preliminary crash investigation without traveling to the site.
 - Scan the crash reports for each crash, and add a uniform resource locator (URL) for each report to the attribute table. The information and officer sketches will be helpful in analyzing crash patterns at specific locations.

References

- Abdel Aty and As Saidi, "Using GIS to Locate the High Risk Driver Population," VTI Konferens 13A, Part2 (2000): 111-126.
- Mendoza, A., Uribe, A., Gil, C.Z., and Mayoral, E., "Development of a Relational Accident Database Management Federal Roads," *Transportation Research Record* 1717 (2000): 84-93.
- Mendoza, A., Mayoral, E.F., Vicente, J.L., and Quintero, F.L., "Geographic Information System-Based Accident Data Management for Mexican Federal Roads," *Transportation Research Record* 1746 (2001): 74-83.
- Arthur, R. and Waters N., "Formal Scientific Research of Traffic Collision Data Utilizing a GIS," *Transportation Planning and Technology* (1997): Vol. 21, 121-137.
- Austin, K., Tight, M., and Kirby, H., "The Use of Geographical Information Systems to Enhance Road Safety Analysis," *Transportation Planning and Technology* (1997): Vol. 20, 249-266.
- Bapna, S., Grebele, F., and McCall, M., "Spatially Referenced Crash Data System for Application to Commercial Motor Vehicle Crashes," Maryland Department of Transportation research (2003).
- Brewer, D.J., Grindley, R.D., and Ellison, J.W., "GIS, Collision analysis, and Gaurdrail: A Systematic Approach to Enhancing Roadside Safety," ITE 2000 Annual Meeting and Exhibit, Nashville (2000).
- Breyer, .JP., "Tools to Identify Safety Issues for a Corridor Safety-Improvement Program," *Transportation Research Record* 1717 (2000): 19-27.
- Brown, D.B., Critical Analysis Reporting Environment. Department of Computer Sciences, The University of Alabama, Tuscaloosa, AL, 1997. <http://care.cs.ua.edu>.
- Chakraborty, J. and Armstrong, M. P. (1996) Using Geographic Plume Analysis to Assess Community Vulnerability to Hazardous Accidents. *Computer, Environment, and Urban Systems*, 19: 341-356.
http://www.uiowa.edu/~geog/faculty/armstrong/CEUS_Plumes.pdf
- Harkey, D., "Using GIS in the Analysis of Truck Crashes," Report No. FHWA-RD-99-119 (1999) <http://www.tfhr.gov/safety/hsis/99-119.pdf>.
- Saccomanno, F.F., Fu, L., and Roy, R.K., "Geographic Information System-Based Integrated Model for Analysis and prediction of Road Accidents," *Transportation Research Record* 1768 (2001): 193-202.
- Harkey, D.L., "GIS-Based Crash Referencing and Analysis System," *ITE journal*, Volume 69 Issue 4 (1999): 9-12.
- Harkey, D.L., "Evaluation of Truck Crashes Using a GIS-based Crash Referencing and Analysis System," *Transportation Research Record* 1686 (1999): 13-21.

- Miller, J.S., "What Value may Geographic Information Systems Add to the Art of Identifying Crash Countermeasures," Virginia Transportation Research Council (1999) http://www.virginiadot.org/vtrc/main/online_reports/pdf/99-r13.pdf.
- Kim, K., Levine, N., and Nitz, L., "Development of a prototype Traffic Safety Geographic Information System," *Transportation Research Record 1477* (1995): 41-47.
- Kim, K., Kerns, T., Hettinger, T., Pease, M., Johnson, S., and Utter, D., "Geographic Information Systems Using CODES Linked Data (Crash Outcome Data Evaluation System)," Repot No. DOT HS 809 201 (2001). <http://ntl.bts.gov/data/809-201.pdf>
- The University of North Carolina Highway Safety Research Center, "Mapping the Road to Improved Commercial Truck Safety," *Directions*, Volume 6 Issue 3 (2001): 6-9.
- Miller, J.S., "Geographic Information Systems: Unique Analytic Capabilities for the Traffic Safety Community," *Transportation Research Record 1734* (2001): 21-28.
- Dodd, N.L., Gagnon, J.W., and Schweinsburg, R.E., "Evaluation of Measures to Minimize Wildlife-Vehicle Collisions and Maintain Wildlife Permeability Across Highways in Arizona, USA," *International Conference on Ecology and Transportation, Making Connections* (2003): 353-354 <http://www.icoet.net/downloads/03AnimalVehicleCollision.pdf>
- Smith, R.C., Harkey, D.L., and Harris, B. "Implementation of GIS-Based Highway Safety Analyses: Bridging the Gap," Report No. FHWA-RD-01-039 (2000) <http://www.tfrc.gov/safety/pubs/1039.pdf>.
- Schneider, R.J., Khattak, A.J., Ryznar, R.M., "Identifying Factors Associated with Pedestrian Crash Risk Through Spatial Analysis and Perception-Based Modeling," Submitted to Transportation Research Board (2001). <http://www.planning.unc.edu/program/trbwalk.pdf>
- Roche, J., "Geographic Information Systems-Based Crash Data Analysis and the Benefits to Traffic Safety," 2000 Transportation Scholars Conference (2000): 85-94 <http://www.ctre.iastate.edu/mtc/papers/roche.pdf>.
- Schutle, S.L., and Aultman-Hall, L., "Using Geographic Information Systems to Evaluate the Influence of County Level Alcohol Prohibition on the Distance from Home in Alcohol-Related Vehicle Crashes," *VTI Konferens 13A, Part2* (2000): 217-226.
- Miaou, S., Mallick, B., and Joon Jin Sung, "Tracking Crashes on our Highways," *Texas Transportation Research*, Volume 37, Number 3 (2001): 13 http://tti.tamu.edu/researcher/v37n3/37_3.pdf.
- Souleyrette, R., Kamyab, A., Hans, Z., Knapp, K.K., Khattak, A., Basavaraju, R., and Storm, B., "Systematic identification of High Crash Locations," Iowa DOT TR-442; CTRE 00-59 (2001) <http://www.ctre.iastate.edu/reports/hcl.pdf>.
- Pulugurtha, S.S., and Nambisan, S.S., "Using GIS to Identify Problem Areas and Enhance Pedestrian Safety in Clark County," Conference Title: GIS-T 2002: Melting Down the Stove Pipes (2002) <http://www.gis-t.org/yr2002/gist2002sessions/6.3.2.pdf>.
- Thnay, C., "Comprehensive and Effective High Crash Location Analysis. Digital Database and GIS to a Quicker and More Accurate Solution," ITE 2000 Annual Meeting and Exhibit, Nashville (2000).

Abbreviations

ADT	Average daily traffic
ALDOT	Alabama Department of Transportation
CARE	Critical Analysis Reporting Environment
CODES	Crash Outcome Data Evaluation System
DUI	Driving Under Influence
FHWA	Federal Highway Administration
GIMS	Geographic Information Management System
GIS	Geographic Information System
GIS-ALAS	Geographic Information System Accident Analysis and Location System
GPA	Geographic Plume Analysis
GPS	Global Position System
HSIS	Highway Safety Information System
ITE	Institute of Transportation Engineers
LHRS	Linear Highway Reference System
PDO	Property Damage Only
UTCA	University Transportation Center for Alabama

Appendix A
Procedures for Creating Shapefiles for Intersections, Road Segments,
and Highways

Procedure for digitizing nodes for intersections

The following procedure was followed to create the shapefile for intersections:

1. Create a new feature class in the geodatabase for nodes. Define this feature class as a point feature. Add a new field to the attribute table for Node_ID.
2. Start an editing session and select *create new features* from the drop down menu as shown in Figure A-1. Select the sketch tool icon that is used to create new features.
3. Start inserting points on the GIS base map at the intersection locations corresponding to the link-node paper map. At that time, type in the node number in the Node_ID column of the attribute table as the new node is created. The attribute table for the new node will have the following columns: ObjectID, Shape, and Node_ID.
4. Finally, convert this geodatabase to a shapefile and merge it with other node shapefiles produced by multiple digitizers.

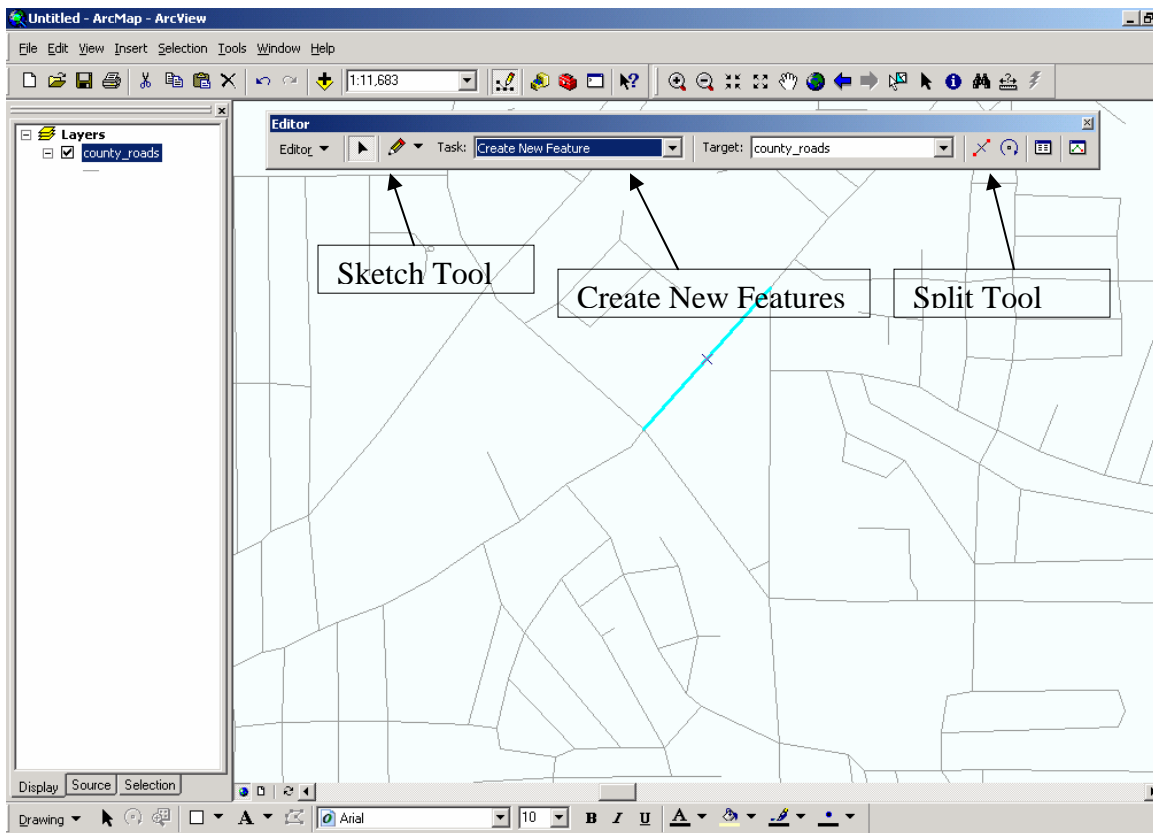


Figure A-1. Edit session open for “county_roads” shapefile

Procedure for digitizing links for road segments

The following procedure was adopted to create a shapefile for road segments:

1. Insert three new fields in the attribute table for roads: Link_ID, Node_1, and Node_2.
2. Select a road segment with a link number/road name from the link-node paper map.
3. Look for the corresponding road on the GIS base map. The entire road will probably be in the GIS as a single entity. Every segment of a road may have the same link number, but a segment can be uniquely identified by the starting and ending node numbers.
4. Select the target segment from the base map. Start the editor to type in the link number and the node numbers in the attribute table for that road segment.
5. If the line between two nodes has more than one segment, these lines must be merged into one segment. Select all the segments that should be merged together. Start the editor session and select *merge* from the drop down menu of editor. All the segments will be merged to form a single segment. Now add the link and node numbers to the attribute table associated with that segment.
6. If a line representing a road in the GIS passes through an intersection, the line must be split at the intersection to divide the road into unique segments. Start the editor, click on the edit icon, and select the line to split. Click on the split icon (as shown in Figure A-1) on the editor toolbar and click the location of the intersection to split the line. Add the link and node numbers to the attribute table associated with each new segment.
7. To add a missing road segment, start the editor and *create a new feature* in the existing shapefile for the roads. Add the segment of road that is missing, and type in the link number and the associated node numbers in the attribute table.
8. Work done by multiple digitizers can be combined in one map. Merging layers combines the work done on two different areas of a map. Before merging layers, all of the roads not worked on should be deleted. After merging two areas of roads on two layers, the attribute fields should be FEANAME, Link_ID, Node_1, Node_2, and Shape_Length.

Procedure for creating routes for highways

Crashes on highways were mapped according to the milepost at which the crash occurred. The *dynamic segmentation* feature of ArcGIS was used for this purpose. To create routes, a shapefile for highways must have the “M” measure turned to active mode to identify locations according to mileposts based on the starting and ending milepost assigned to line feature representing the highways.

ArcTools contains a routine called *create routes* that can create a new shapefile with the “M” measure activated from an existing shapefile with “M” measure inactive. The tool

can use the beginning and ending milepost data from the inactive shapefile attribute table to establish a new shapefile with “M” measure active.

Appendix B
Procedure for Connecting DPS Data with GIS Maps

Procedure for preparing DPS data for GIS

For Nodes

Location information for crashes occurring at intersections is stored in the DPS database by way of a single node number in the Node_1 column. DPS also has a Node_2 column that is used for locating crashes that occurred on a road segment; this column stores a “0” if the crash occurred at an intersection. Thus, any crash record showing the Node_2 value as “0” is an intersection crash. In order to map this data, a common field was required in the DPS and GIS data. The following procedure was adopted to create the common field:

1. Open the attribute table of DPS data in the GIS and create a new field called “Connect_ID”.
2. Select Connect_ID column and calculate values to be inserted in each field.
3. A Visual Basic (VB) program was written to calculate Connect_ID value for each record. The program is as shown below:

```
Dim abc As String
If ([Node_2]=0) Then
    abc=[Node_1]
EndIf
```

- The value assigned to “abc” is then stored in the Connect_ID field.
4. The crash attribute table typically contains more than one crash for a particular intersection. It is not possible to join a one-to-many relationship between nodes and crashes. To solve this problem, a summary table was formed which contained one record for each intersection, with the number of crashes that occurred at that intersection as a separate column. This table was then joined to the GIS attribute table for nodes.
 5. After the two tables are joined, thematic maps were created to display the number of crashes occurring at each intersection.

For Links

Crashes occurring on link segments are described in the DPS database by three numbers: link number and the two intersection numbers at each end of the link. It is not possible in GIS to connect two attribute tables based on three fields. To solve this problem, data from all three columns were merged into a single column, which was then used to connect the DPS and GIS data. The following procedure was adopted:

1. The attribute table of DPS data was opened in the GIS, and a new field called “Connect_ID” was created.
2. The Connect_ID column was selected, and calculated values were inserted in each record.
3. A VB program was written to calculate the Connect_ID value for each record. The program is shown below:

```
Dim abc As String
```

```

If ([Node_2]>0) Then
    abc=((([Link_ID]) & "_" & ([Node_1]) & "_" & ([Node_2]))
EndIf

```

4. The value assigned to “abc” is then stored in the “Connect_ID” field.
5. The VB code for links was combined with the VB code for nodes so that the same field can be used to connect intersection crashes as well as link crashes. The combined code is as shown below:

```

Dim abc As String
If ([Node_2]=0) Then
    abc=[Node_1]
Else
    abc=((([Link_ID]) & "_" & ([Node_1]) & "_" & ([Node_2]))
EndIf

```

6. The combined code produces a value equal to Node_1 if the value of Node_2 is zero and a concatenated value of Link number, Node_1 and Node_2 separated by an underscore, “_”, if the value of Node_2 is greater than zero.
7. Similarly, the link number, Node_1, and Node_2 data in the GIS table is also concatenated and stored in a single column which is then used as a common field to join the DPS and GIS information.
8. Thematic maps can then be created for link crashes as was done for intersection crashes by summarizing the number of crashes on each link and displaying the links with different colors and widths.

For Highways

Two fields identify highway crashes: 1) highway number and 2) the milepost at which the crash occurred. DPS data already stores this information in two separate fields in the attribute table. After routes are created for each highway, the DPS file is saved as a .dbf file and added to the map as an event table. This was accomplished by using the *Add Route Event* tool in ArcGIS. This tool asks for the route layer, the route identifier, the event table, and the route id storing crash route information. In addition, the column in the crash data that stores the milepost is entered in the measure field on this box. This tool will identify the highway and then map all the crashes in the event table after interpolating between the start and end milepost for each route.

Appendix C
Procedure for Creating and Using Networks

Procedure for creating network

A Network uses “weights” to choose which path is optimal. Weights are attributes associated to links and nodes that are used to compare different paths. ArcMap compares each value in the “weight” column and chooses the path with the lowest value. For example, if the length of a link is designated as the weight, ArcMap will choose the shortest path. This means that if the safest path is desired, crashes (number of crashes in recent history) for links and nodes should be set as the weight. It is possible to have multiple paths from one location to another that have the same number of crashes. In this case, it is logical to have ArcMap choose the shortest path among the safest paths. For this purpose, a new link attribute can be created that combines the number of crashes in recent history and the link length. The procedure for editing the link attribute table follows:

1. Within ArcGIS, add the links shapefile, start an editing session, and open the attribute table. This may be done without starting an editing session, but some mistakes can only be undone within an editing session.
2. Click the *Options* button and select *Add Field*.
3. Name the field (for example: “crash_length”); set the *field type* as “Double” and the *precision* to “0.000000” and click *OK*.
4. Right click the new field column heading and select *Calculate Values...*
5. In the script box, type “[crashes] + ([shape length] / 1,000,000)” and click *OK* ([shape length] must be divided by a number much larger than the longest shape length in the network to ensure that the shape length only affects the choice when multiple paths have the same number of crashes. If the shape length is divided by a much smaller number, a path that is shorter with more crashes could be chosen over a path that is longer with fewer crashes.)
6. End the editing session and save edits.

If this new attribute is set as the weight, the solver will always choose the shortest between multiple paths with the same number of crashes. For example, if path (a) is 1000 meters long and has 3 crashes, it will have a weight of 3.001000; if path (b) is 100 meters long and has 3 crashes, it will have a weight of 3.000100. Therefore, the weight for path (b) is smaller, and ArcMap will choose it.

Converting Data (if necessary)

The links and nodes must be located in the same feature dataset within the same geodatabase to use the Build Geometric Network Wizard. The procedure for converting the data from a shapefile to a personal geodatabase follows:

1. Within ArcCatalog, use the catalog tree to find the desired location for the new personal geodatabase.
2. Right click in the contents window and select *New* and then *Personal Geodatabase*.

3. Right click the new geodatabase and select *Import* and then *Shapefile to Geodatabase...*
4. Click the browse button next to the *Input shapefile* field and locate and select the link shapefile.
5. In the *Select an existing feature dataset or enter a new one* field, type the name of the feature dataset and click *OK*.
6. Repeat steps 3 through 5 to import the node shapefile.

Using the Build Geometric Network Wizard

The *Build Geometric Network Wizard* is used to create a network. This wizard joins the roads and intersections by automatically snapping them together. In addition, the weights are set for the analysis of the network. The procedure for using the *Build Geometric Network Wizard* follows:

1. Open ArcToolbox. Double-click *Data Management*. Double-click *Geodatabase*. Double-click *Build Geometric Network Wizard* and, when it opens, click *Next*.
2. Click the browse button next to the *Select your input feature dataset* field and locate and select the feature dataset that contains the link and node feature classes.
3. Click the boxes next to the link and node feature class in the *Select the feature classes you want to build your network from* field, name the network in the *Enter a name for your network* field, and click *Next*.
4. Select the *No* button in the *Do you want complex edges in your network?* window and click *Next*.
5. In the *Do your features need to be snapped?* window, select the *Yes* button and set the *Snap tolerance* to 3 (meters). Click the boxes next to the link and node feature classes in the *Select features that can be moved* field and click *Next*.
(Note: The snap tolerance must be set with the scale of the roads in mind. If the snap tolerance is left at the default of 0.000384 (meters), very few links and nodes would move. If the snap tolerance is set too high, links and nodes that should not be connected may be snapped together. A tolerance of 3 meters is very conservative for street center-lines.)
6. Select the *No* button in the *Does your network have sources or sinks?* window and click *Next*.
7. Select the *Yes* button in the *Do you want to assign weights to your network?* window. In the *Enter the names of your weights and their types* table, type “path length” under *Weight Name*, and select “Double” in the drop down menu under *Type*. Click the *Add Row* button on the right, type “crashes” under *Weight Name*, and select “Integer” in the drop down menu under *Type*. Click the *Add Row* button once again, type “crash_length” under *Weight Name*, select “Double” in the drop down menu under *Type*, and click *Next*. (Note: Three weights were created for use in the analysis stage. The weight names “path length”, “crashes”, and “crash_length” are only suggested names.)
8. In the *Assign your weights to fields in your feature classes* window, select “path length: Double” from the drop down menu in the *Available network weights* field. In the *Assign the above weight to the fields you want in the following feature*

- classes* table, select “Shape Length” from the drop down menu in the *Associated Field* column for the link feature class and leave “<none>” for the node feature class.
9. Select “crashes : Integer” from the drop down menu in the *Available network weights* field. In the *Assign the above weight to the fields you want in the following feature classes* table, select “crashes” in the *Associated Field* column for the link and node feature classes.
 10. Select “crash_length : Double” from the drop down menu in the *Available network weights* field. In the *Assign the above weight to the fields you want in the following feature classes* table, select the previously added “crash_length” field from the drop down menu in the *Associated Field* column for the link feature class, select “crashes” from the drop down menu in the *Associated Field* column for the node feature class, and click *Next*.
 11. In the *Summary of your input* window, click *Finish*.

Procedure for using network

After the geometric network has been created, analysis of the network may be performed by using the *Utility Network Analysis* toolbar. To open the toolbar, click *View* in the main menu, place the cursor over *Toolbars*, and then click *Utility Network Analysis*. The procedure for finding the shortest of the safest paths from one location to another follows:

1. Open the new geometric network in ArcGIS and add the Utility Network Analysis toolbar.
2. In the *Analysis* drop down menu on the Utility Network Analysis toolbar, click *Options*.
3. Click the *Weights* tab. In the *Weight for junction features* field select “crashes” from the drop down menu. In the *Weight along digitized direction of edges (from-to weight)* field select “crash_length” from the drop down menu. In the *Weight against digitized direction of edges (to-from weight)* field once again select “crash_length” from the drop down menu and click *OK*.
4. In the *Trace Task* drop down menu on the *Utility Network Analysis* toolbar, select *Find Path*.
5. Click the *Add Junction Flag Tool* button and place a flag at the starting and ending points of the path by clicking the desired nodes. (Note: Starting and ending points of path must be nodes.)
6. Click the *Solve* button.

The path that is highlighted (default color is red) is the shortest of the safest paths. If the shortest path is desired, set the weights to “path length” for the links (from-to and to-from edge weight) and “< None >” for the nodes (junction weight). To check for any other paths with the same number of crashes that may be longer, set the weights to “crashes” for the links (from-to and to-from edge weight) and the nodes (junction weight).

The procedure for finding the total length and the number of crashes along the chosen path follows:

1. Go to the *Analysis* drop down menu on the *Utility Network Analysis* toolbar and click *Options*.
2. Click the *Results* tab and then click the button next to *Selection* under *Return results as*. Solve for the path.
3. Once the path is selected, open the link attribute table and at the bottom of the window after *Show*: click *Selected*.
4. Right click the “Shape Length” column heading and click *Statistics...*
5. The *Sum*: value is the length of the selected path. Close the window.
6. Right click the “crashes” column heading and click *Statistics...*
7. The *Sum*: value is the number of crashes along the links of the path. Close the window.
8. Close the link attribute table, open the node attribute table, and click *Selected* after *Show*: at the bottom of the window.
9. Right click the “crashes” column heading and click *Statistics...*
10. The *Sum*: value is the number of crashes at the nodes along the path. The sum of the crashes along the links and the crashes at the nodes is the total number of crashes along the path. Click OK and close the attribute table.

A variety of paths and the number of crashes along the paths can be found by making several runs using different weights. Sometimes the safest path is much longer than a path that has only a few more accidents. If a geometric network is desired that chooses more realistic paths, paths that are relatively short and are close to being the safest, more paths should be created that give greater weight to path length. This may be accomplished by dividing the “shape length” by a smaller number when adding the “crash_length” field to the link attribute table. Many weights may be created where shape length is given increasing weight, and the network can be solved to find the optimum or desired balance between shape length and number of crashes. It is important to note that weights cannot be added to an existing network, and only one network can be created per feature dataset. In order to “add” weights, the entire process must be repeated, and a new network must be created.