1-1-2008

Development of a GIS-based safety analysis system

Vidhya Kumaresan
University of Nevada, Las Vegas

DEVELOPMENT OF A GIS-BASED SAFETY ANALYSIS SYSTEM

by

Vidhya Kumaresan E.I

Bachelor of Engineering
Anna University, Chennai, India
2006

A thesis submitted in partial fulfillment
of the requirements for the

Master of Science in Engineering
Department of Civil and Environmental Engineering
Howard R. Hughes College of Engineering

Graduate College
University of Nevada, Las Vegas
August 2008

UNLV
UNIVERSITY OF NEVADA LAS VEGAS

Thesis Approval

The Graduate College
University of Nevada, Las Vegas

22 August, 20 08

The Thesis prepared by

Vidhya Kumaresan

Entitled

Development of a GIS-based Safety Analysis System

is approved in partial fulfillment of the requirements for the degree of

Master of Science in Engineering

Examination Committee Chair

Dean of the Graduate College

Examination Committee Member

Examination Committee Member

Graduate College Faculty Representative

1017-53

ABSTRACT

Development of a GIS-based Safety Analysis System

by

Vidhya Kumaresan

Shashi Nambisan, Ph.D., P.E., Examination Committee Chair
Adjunct Research Professor (Research Advisor)
Department of Civil and Environmental Engineering
University of Nevada, Las Vegas

Edward Neumann, Ph.D. P.E., Examination Committee Co-Chair
Professor (Academic Advisor)
Department of Civil and Environmental Engineering
University of Nevada, Las Vegas

The objective of this study is to develop a safety analysis system that integrates crash data and roadway related information. This system is developed in a Geographic Information System (GIS) environment. It includes customized user interfaces to support queries to analyze data and to display results either in graphical or tabular formats. The system also affords the capabilities to export such results. The system permits analyses to be performed either at individual locations such as intersections, or for roadway segments. These analyses are based on data fields included in the crash database. The queries may be based on individual attributes recorded in the crash database or by combining multiple attributes from the database. The system also contains a module to identify high crash locations based on methods identified from the published literature. The methods range from those based on simple crash frequency to more complex

methods which incorporate different weights for crashed based on the crash outcomes. An application of the system is illustrated using data from the Las Vegas metropolitan are in the state of Nevada.

The system can be used to identify safety issues in a region, and to plan and deploy appropriate countermeasures to enhance safety. It also can be used to monitor the effectiveness of traffic safety programs. Such a system could also be used to screen projects and operational strategies to be funded in through appropriate funding processes.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I am grateful to many individuals who contributed throughout the course of my research project and Master's degree. I appreciate each of their participation in the completion of this project and my thesis. I wish to express my sincere thanks to Dr. Shashi Nambisan, my advisor, for having initiated and for having given me the opportunity of working on this interesting project. His guidance and suggestions in facing the challenges during the course of this research are invaluable. I also wish to appreciate him for his advice regarding the coursework required for the completion of this degree. I would also like to extend my thanks to Dr. Edward Neumann, Dr. Mohamed Kaseko and Dr. Ashok Singh for their support and for having served on my Thesis Committee.

I thank Mr. Vinod Vasudevan for his invaluable guidance and support over the duration of the research project that formed the basis of this thesis. I am deeply grateful to the Transportation Research Center (TRC) and the Department of Civil and Environmental Engineering at UNLV, in particular to Dr. Nambisan, and Mr. Vasudevan, for giving me a research assistantship. The accompanying financial support over the past two years made my graduate studies at UNLV possible. The research project at the TRC was funded by the Regional Transportation Commission of Southern Nevada, and this support is gratefully acknowledged. The advice, information, and insights provided by several individuals at the RTC (including Mr. Jerry Duke, Mr. Tom Wolch, Mr. Brian Hoeft, and Mr. Shital Patel), and other members of the project advisory committee helped develop the system presented in this thesis. I am grateful to these individuals for their contributions. Ms. Xin Li at the TRC and ESRI technical support staff are thanked for their help during the various stages of my research Also, Dr. Mukund Dangeti and the other researchers, faculty, and students associated with the TRC have earned my deepest appreciation for their support and cooperation.

I would like to take this opportunity to acknowledge my friends and other individuals who have contributed in some way toward the completion of this thesis and my graduate studies at UNLV. Most importantly, I thank my family and friends for their affection and strong support all this while. I dedicate with all my love and respect this thesis to my parents.

CHAPTER 1

INTRODUCTION

Roadway safety has been a source of significant concern to transportation planners, engineers and public policy makers, particularly over the past couple of decades. The identification of traffic safety related concerns, and establishment of programs to improve safety are the subject of various efforts around the world. Analyses of relevant data are essential to develop strategies to improve safety. However, data required for such analyses often come from multiple sources, in varied formats, with various levels of accuracy and reliability. Computerized tools and systems offer great potential to combine such data and to perform the required analyses. This chapter provides an introduction to the need for the development of an automated safety analysis system.

1.1 Motivation

The World Health Organization says that about 1.2 million people are killed in road accidents every year throughout the world. If left unattended, a WHO report predicts that the number will increase to 2.3 million by the year 2020 (WHO, 2004). Over the past decade, traffic crashes have resulted in more than 42000 fatalities annually in the United States (Figure 1.1). Fatal crashes (Figure 1.2) increased by 1.9 percent from 2004 to 2005, and the fatality rate rose to 1.45 fatalities per 100 million vehicle miles of travel in 2005 (NHTSA). The *National Highway Traffic Safety Administration* states that in 2004, motor vehicle crashes were the leading cause of deaths for the people in the age group 2 to 34 (Subramanian 2007). These facts reinforce the need for an effective traffic safety program to reduce the number, rate, and varied impacts of traffic crashes. Identification of safety issues is necessary to implement appropriate countermeasures to enhance safety. These measures in the long run would be able to provide beneficial impacts in the form of reduced individual and societal costs from crashes, reduced delays on roadways, improved travel time reliability, and reduced undesirable emissions from motor vehicles.

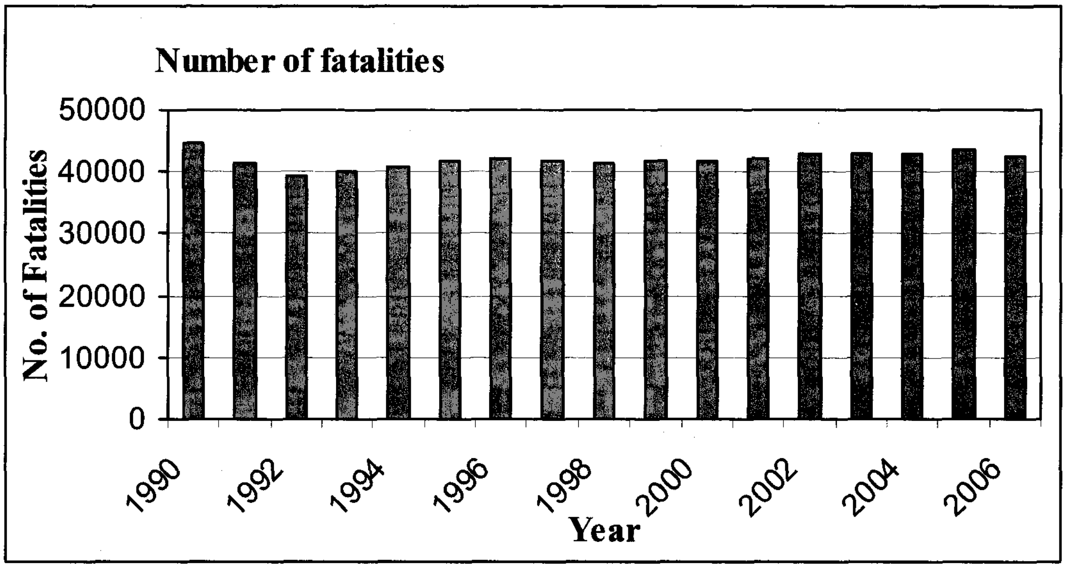


Figure 1.1 Road fatalities in the US, 1990-2006
Source: http://www.bts.gov/publications/national_transportation_statistics/

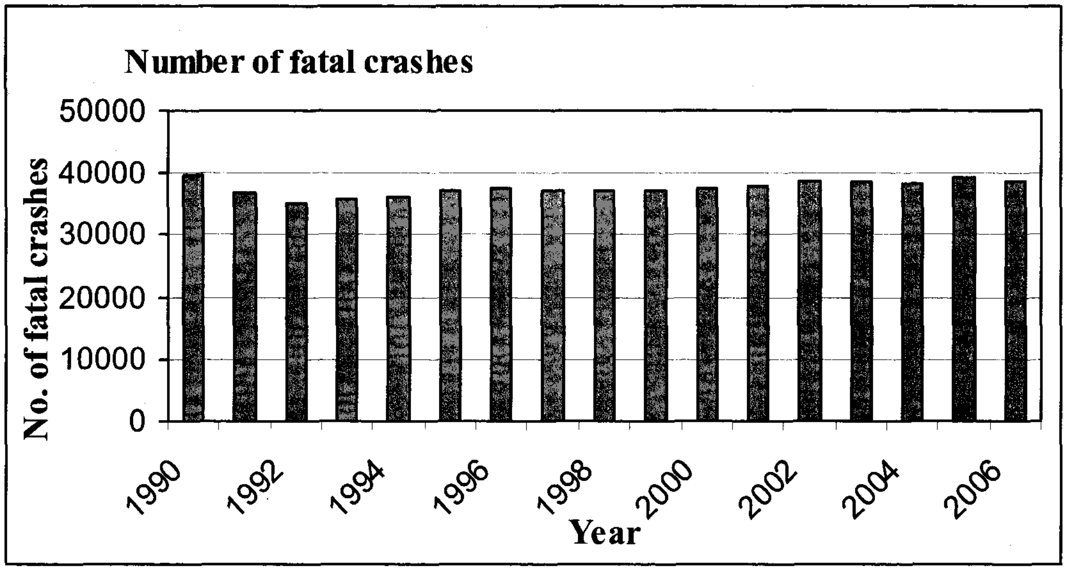


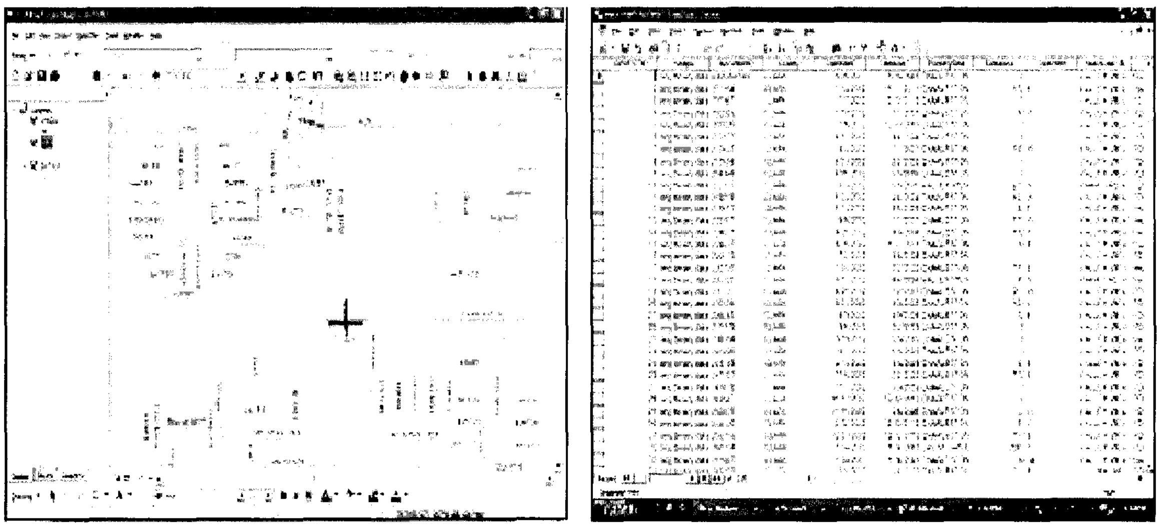
Figure 1.2 Fatal Crashes in the US, 1990-2006
Source: http://www.bts.gov/publications/national_transportation_statistics/

1.1.1 Integration of Databases

A vital element of traffic safety programs is access to accurate data pertaining to crashes and roadway characteristics. There are numerous agencies that record and store crash related information. They include the local and state law enforcement agency systems, Police Accident Reports, National Highway Traffic Safety Administration's Fatality Accident Reporting System (FARS), General Estimates System, State Department of Transportation surveys. Therefore, a reliable database has to be compiled after integrating crash data from such sources. For a good traffic safety analysis system, such historical crash records need to be combined with roadway network information. Similar to crash data, roadway data are also gathered and maintained at various levels of accuracy by local, stat, and federal agencies. Integrating such databases offers additional benefits in terms of better representation of data and decision-making.

1.1.2 Use of GIS platform

The data collected vary widely in the quantity and quality and to the extent to which the different data sets could be integrated. The use of Geographic Information Systems (GIS) is a solution to this issue. Due to the spatial nature of disparate datasets needed for transportation safety analysis, the GIS platform would be a good option to bring them together. Further, the GIS platform offers a natural environment for development of a safety analysis system using the consolidated database due to its data integration and mapping capabilities. The advantage of viewing data in spatial format is perhaps a very important benefit of using GIS. Presenting data in a graphical view rather than as a table (e.g., a spreadsheet) helps with faster interpretation as shown in Figure 1.3.



(a) Map view (b) Table view

Figure 1.3 Viewing data on a map as opposed to a spreadsheet

1.2 Research Objective, Tasks and Scope

The objective of this research is to develop a safety analysis system by integrating in a GIS environment crash data with information related to the roadway network characteristics. The integration of crash information with the roadway network database is to help facilitate evaluations of roadway and operational characteristics with crash data records. In turn, this will enhance the ease of access to such disparate but related datasets and expedite analyses. Such a system is expected to aid decision makers in reviewing historical information, screening projects, and developing design and operational strategies.

The safety analysis system is intended to be used to perform analyses based on crash data to extract and review data for crashes which match user specified criteria so as to better understand traffic safety issues at a particular location or along a corridor. In addition to aiding decision-making, the system is also aimed to improve data representation by enabling graphical displays of the crashes with the street network as the background. The intended users of the system include elected officials, managers, engineers, planners, and engineering technicians. The system has been developed and demonstrated using data from the Las Vegas metropolitan area. The scope of the study also includes development of graphical user interfaces to perform analyses at various levels of detail (ranging from a macro level for policy analyses to a micro level for technical analyses).

1.3 Study Area

The Las Vegas Metropolitan area that includes the city of Las Vegas and unincorporated parts of the county of Clark in the state of Nevada area is chosen as the study area for this system. This area, also known as the Las Vegas valley, covers an area of nearly 600-square-miles and holds the largest concentration of the population in the state of Nevada. The metropolitan area consists of the cities of Boulder City, Henderson Las Vegas, and North Las Vegas, apart from parts of unincorporated Clark County (see Figure 1.4). Las Vegas is considered to be one of the fastest growing cities in the United States with a population growth of nearly 29.2% from 2000 to 2006 (http://www.census.gov/Press-Release/www/releases/archives/population/009865.html). A 2006-estimate of U.S. Census Bureau shows that the population of the Las Vegas metropolitan area was over 2 million.

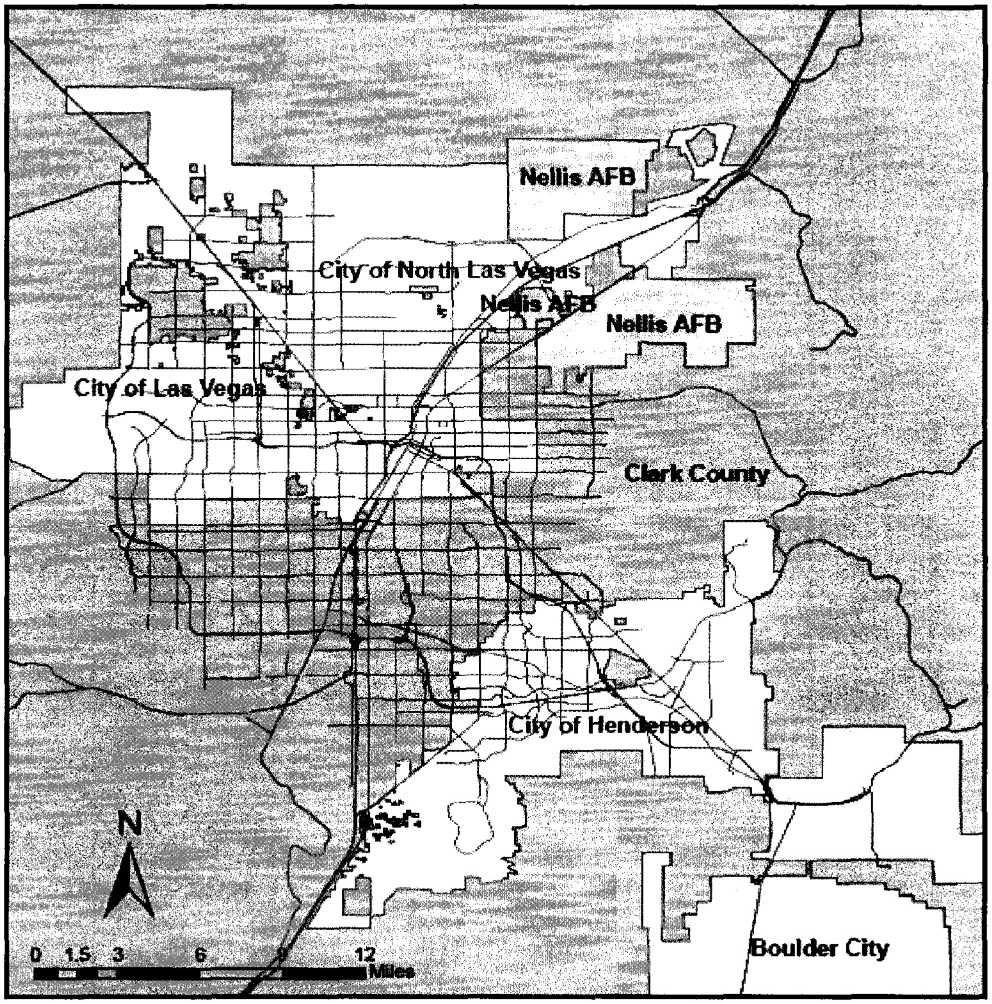


Figure 1.4 A map of the Las Vegas metropolitan area

1.4 Layout of the Thesis

This thesis report consists of five chapters. An introduction to the study and its scope is presented in Chapter 1. Chapter 2 presents a brief review of the literature on the use of GIS in traffic safety, data management and analysis, estimating road risk and hazard level, and existing traffic safety systems that involve integration of disparate datasets. The collection of data, integration of datasets, development of the system architecture, creation of user-interface and description of the tool are discussed in Chapter 3. Chapter 4 illustrates the application of the developed methodology and analysis for the Las Vegas Metropolitan area. The results are summarized along with conclusions in Chapter 5. The final chapter also describes recommendations for future research.

CHAPTER 2

LITERATURE REVIEW

Traffic analysts use historical crash data to perform statistical and spatial analyses for gathering information that is mainly directed towards improvement of traffic safety in the region/location under consideration. Numerous researchers and traffic safety professionals have applied the results from these analyses to concepts like ranking the high crash locations in a particular area, evaluating the statistics to identify the traffic safety issue and recognize the cause for the crashes in terms of roadway geometry design, traffic operation concerns or other reasons that need to be addressed. Also, after identification of high crash location, the analysis results are used to propose and develop appropriate countermeasures to improve safety at the location.

In recent times, GIS is used in analyses of crash data owing to its mapping capabilities. The user has the ability to view the results in a graphical from along with tabulated values by the use of mapping software. Further, GIS facilitates the integration of crash data along with other information such as traffic volume, roadway details and weather conditions. This chapter discusses some of the research and published literature in the above context. Studies that were conducted in the area of analysis of crash data, identification of high crash locations and the use of GIS in crash analyses have been described.

2.1 Analysis of Crash Data

The use of crash data and their analyses to help make decisions related to traffic improvements have been pursued by transportation agencies for many years. These include efforts across the United States that have attempted to create crash reporting systems which allow other local agencies to obtain information such crash data and their attributes specific to their area. The Fatality Analysis Reporting System (FARS) developed by the National Highway Traffic Safety Administration (NHTSA) is one such system that aims to provide a measure of overall highway safety and thereby to evaluate the effectiveness of highway safety programs. FARS has been operational since 1975. It is a unique nation-wide source for highway crash census involving fataities (http://www-nrd.nhtsa.dot.gov/departments/nrd-01/summaries/FARS_98.html). The FARS is accessible online (http://www-fars.nhtsa.dot.gov/Main/index.aspx) and the website hosts a dynamic query builder that allows the user to query data for a specific year at a time. The data fields included for querying are categorized according to crashes, persons, vehicles, and drivers. Results of the user's query are displayed in tabular format and allowed to be exported as a text file or a spreadsheet document.

Parrish et. al. (2003) explain the Critical Analysis Reporting Environment (CARE), developed by the University of Alabama. CARE was developed to analyze automobile crash data to aid in the development of strategies to reduce crashes. Crash data can be queried in any of the data fields for which the data are included in the system. Crash records that match the criteria set by the user are displayed in a tabular format and also a chart format. CARE is adopted by a number of states since its development (Parrish et al. 2003, 20-80). Similarly, the University of North Carolina's

Highway Safety Research Center (HSRC) has developed a crash data query website (http://www.hsrc.unc.edu/crash/index.cfm). Crash data over the period 2001-2006 in North Carolina can be queried on crash, vehicle, and person information records. The data can be made specific to any area in the state of North Carolina. The data filters that can be included in the query are driver age, bicycle, pedestrian, motorcycle, speeding, truck/bus or alcohol involvement. Data display in this case is also in a tabular format.

The Highway Safety Research Group (HSRG) of the Louisiana State University has its own website that provides information related to crash statistics for the state of Louisiana (http://lhsc.lsu.edu/default.asp). The website hosts a dynamic query generator that allows data query for location, vehicle, time, and people. The data currently available on the website are for the period 1999 to 2006. The data are displayed as a table on the website upon running the query. An option to export the results to a spreadsheet is also offered along with a print tool. In addition, a chart displaying the number of crashes in the various entries for the query data field chosen is displayed.

2.2 Suitability of GIS for Analysis

The suitability of a GIS software environment for evaluating traffic safety is discussed by Affum and Taylor (Affum and Taylor 1997, 93-119). They describe SELATM, which is a GIS-based program for analyzing crash patterns over time, and the evaluation of the safety benefits of local area traffic management schemes. The program involved integration of network data with crash data, installed devices data to generate crash statistics summaries for the various network levels. This also allowed before and after comparison with a control area. A GIS program was used due to the need for spatial analysis to check the clustering of crash locations over the different time periods. The program used the data integration capabilities of a GIS to integrate data from crashes, the street network infrastructure, and the installed devices to produce an output of crash statistics assembled into a format allowing for before and after comparisons. The main program runs in the form of pull down menus with six main options. Selecting an option pulls down its submenus, some of which require further selection of various options before the system can perform the required operation. A case study describing the application of the program to the Enfield LATM scheme in metropolitan Adelaide was also described.

Roche discussed the capability of using GIS as a tool to perform macroscopic and microscopic analysis on crash data to influence the four Es of traffic safety: engineering, education, enforcement and emergency response (Roche 2000, 110-122). Apart from assisting in crash mapping, a GIS program can be used to perform queries on isolated datasets and therefore process large amounts of data quickly. The GIS environment offers an indirect impact on traffic safety by assisting decision makers to analyze crash data and make decisions about adopting countermeasures based on the analysis. Analysis such as creation of spatial queries, collision diagrams, and identification of high crash locations can be carried out using GIS.

Pawlovich and Souleyrette discuss the limitations of a previously developed accident analysis system and the solutions obtained by the use of a GIS based analysis system instead of the former. The system was originally developed to assist users in obtaining accident statistics for specific locations. One of the limitations of the system was that the location of data required the entry of the node number which had to be identified from paper maps and entered into the system. The use of GIS eliminated this constraint and enabled graphical location of nodes. Further enhancements using GIS were in displaying the results. Graphical as well as tabular results could be generated using GIS (Pawlovich and Souleyrette 1996). Kim and Levine describe the development of a prototype for spatial crash analysis using GIS for Honolulu, Hawaii (Kim and Levine 1996, 289-302). The issues related to selection of hardware and software components and determination of tools related to the type of traffic analyses are discussed. Spatial analyses tools like point analyses, segment analyses and zonal analyses developed for the Hawaii crash data are explained.

Austin et al. describe uses of GIS in addition to answering crash queries such as methods to better identify errors in the police crash report forms, improve the selection of routes and areas suitable for remedial treatment, provide additional information on the safety of school based journeys, and identify the home location of road crash casualties (Austin et. al. 1997, 249-266).The identification of errors in the crash database was of use in estimating if the system understated or overestimated the crash data. The number of crashes that occurred within each section was found to be affected by factors other than route length and traffic flow. For example, considerable differences existed in the crash rates, with the 30 mph limit having almost four times as many crashes per vehicle kilometer than any other speed limit type. Based on such analysis, the areas which were suitable for remedial treatment were identified.

Typical GIS programs also offer various analytical abilities. Miller (2000) stresses the need to integrate spatial and statistical queries for enabling engineering improvements. The geographical relationship in a GIS system enables a common locating system for diverse datasets, rendering many categories of analyses possible. At the macroscopic level, GIS analysis methods such as grid cell modeling, network applications, pattern identification and risk computation are feasible. Crash data can be combined with other datasets such as schools, population densities and roadway characteristics and the possible results from such combinations are discussed (Miller 2000, 21-28).

The application of a Geographic Information System platform in modeling hazardous roadway segments is addressed by Li and Zhang (2007). The study incorporates both the variability in observed risks and spatial autocorrelation by using a Bayesian approach for identification of hazardous roadway segments. A GIS program was used for analysis and visualizing in 3D format (Li and Zhang 2007). This was to enable the allocation of funds to sites that required countermeasures. The advantage of the Bayesian approach is that it can aid in predictive forecasting of risks even in the presence of sparse data. A study area in Houston, Texas was segmented and investigated. Crash and traffic data for the study area were collected and mapped. In this study, linear referencing was used to geocode the crashes. The crash prediction model used for this study incorporated Annual Average Daily Traffic (AADT) as a variable. It also states the possible use of other factors such as number of lanes, geometry, and vehicle type as variables. Crashes due to speeding and Driving Under Influence (DUI) were analyzed in detail. Risk values were computed and the values if greater than 1 related to a risk that was more than expected. Thus a relative risk map was plotted using a GIS program. These values represented a truer value of risk as compared to other methods of estimation since it took into account changes in volumes while smoothing out variability on segments with low traffic or short lengths.

A study by Saccomano et al. (1997) reports on the development of a GIS-based crash risk model for Ontario's Highway network. A GIS program was used owing to the controls that could be exercised on risk factors explaining variations in crash involvement and injury severity. Risk estimation was done at 4 levels of segregation: route-specific, site-specific, route-section-specific, and network-wide. GIS was used to overlay the point-based crash information on the arc-based roadway network coverage. The data collected included details of the highway inventory on roadway geometry, existing/future AADT, road crash data and road-environment data. The objective of the project was to identify 'blackspots' based on roadway geometry as well as volumes using GIS for risk assessment. Results were obtained considering the Ontario Highway network in Canada and identifying blackspots in the study area (Saccomano et. al. 1997, 18-26).

Bapna and Gangopadhyay (2005) describe the design of a GIS for the state of Maryland that provides online crash data and statistical information for commercial vehicle crashes (Bapna and Gangopadhyay 2005, 1-12). The tool's mapping function displays the crashes on a map for visualization. Different layers of depth were used to extend the query capabilities. Reports of the form Summary Report, Factor Analysis Report, Crash Information Report and Identified Crash Report could be generated using the system reporting options. The Online Analytical Processing (OLAP) system allowed multiple user operations. Interagency collaboration issues were stated as a prominent problem in the development of the system.

2.3 Safety Analysis System

Work performed at the Institute of Transportation Studies at the University of California, Berkeley used GIS for maintaining and analyzing a safety database (http://www.techtransfer.berkeley.edu/newsletter/00-2/GIS.php). The ability of a GIS system to show the relationship between objects, all objects in a certain radius of a given location and all features that intersect that region was used. A GIS can help to visually locate many distinctive features of the roadway and traffic landscape, which are then displayed on maps together with details described in a linked table. Information about each location on the map is entered into fields much like in traditional databases. For a particular street segment, information related to citations issued, lighting, road conditions, number of fatalities, and types of injuries may be noted. There is an interface between the tables and the maps, which allows for the creation of a classification hierarchy. The article also states that software packages utilizing GIS are currently available for municipalities to develop in-house automated safety databases, in which various kinds of data commonly used in safety analysis can be readily incorporated: crashes, roadway inventory, traffic data, and citations. Such in-house databases would enable municipalities to swiftly access and analyze safety data. Furthermore, data analysis, display, and reporting capabilities can also be automated, which would help to save time and reduce staff workload in conducting safety studies. The biggest advantage is that municipalities can tailor such database and analysis capabilities to meet their own needs at relatively modest costs. With such automated databases in operation, it is also possible for police officers to enter crash reports directly into the database, and at the same time eliminate inaccurate reporting of crash locations. Such a system to aid agencies and other analysts as well as the general public to perform analyses based on crash data in a spatial background was hoped to be achieved through the research described in this thesis.

Thomanna (2006) describes the traffic accident data analysis efforts undertaken by Arlington County of Virginia. The objective was to provide end users with easy access to data and reports that assist in making funding decisions for safety improvements. To address the need for a suitable product that would enable performing operations such as electronic access to collision data, creation of collision diagrams and a user-friendly data-entry interface for obtaining crash rates based on traffic volume, the Department of Transportation (DOT) of Arlington adopted the use of Accident Information Management System (AIMS). The customized features used by the DOT facilitated defining a crash to be intersection-based or otherwise, and the use of a data entry screen that helped store the data into a database to enable generation of reports and analyze them. The advantage of using this system was the cost effectiveness and the reduction in time taken to perform the various tasks without using AIMS. The limitations included staff training for entry of data which resulted in reduced accuracy (16% of the crashes were unable to be mapped due to inadequate data). The use of this GIS-based technology was found to be especially helpful while conducting corridor analysis, pedestrian safety studies and crash analysis. It provided the tools needed to look at area-wide safety and to identify, and prioritize safety improvements that provide maximum impact (Thomanna 2006, 75-78). The research described in this thesis was designed to incorporate the crash data without much modification to any individual records, thereby eliminating the labor for data alteration.

The Nevada Department of Transportation has embarked on developing a Safety Management System (SMS) whose objectives are to record inventory, monitor, and improve overall roadway safety across the state of Nevada (http://www.ndotsms.com/index.htm). This would also provide better analysis capabilities for improved safety countermeasure design and improve decision making based upon better and more relevant information. It includes roadway safety related data like the Nevada Citation and Accident Tracking System (NCATS) database, NDOT Road Inventory statewide. The information is being used to assess and classify the safety of the road network (compute crash rates for roads and intersections, develop safety rating index for all roads and intersections etc.). The results expected are to provide improved incident and location analysis for developing safety countermeasures - location statistical analysis, countermeasure development, safety benefit determination etc. Ranking of high crash locations based on crash data alone and performance of analyses not specific to an intersection but also a street segment were addressed in the research work for this thesis.

The Washington State DOT's report on the Local Agency Safety Management System (WDOT 1998) makes use of 8 elements for easy management of data. The heart of this system is the data collection, analysis and system output capabilities that provide the information decision-makers use to identify and prioritize safety needs, select the most effective countermeasures, and monitor the performance of the decisions. The sources of information include collision records and roadway inventory data. The analysis procedure uses data sorting for identification of a safety concern, analysis followed by field review, and evaluating the roadway network for risk. Statistically, significance testing could be performed to identify the effects of a change.

Crash Outcome Data Evaluation System (CODES) involves the linkage and analysis of police crash data, emergency medical services transport data, hospital data, and insurance claims. A report by Kim et al. summarizes information about implementation of the GIS based CODES tool in the states of Hawaii, New Hampshire, Maryland, and South Carolina. Issues regarding integrating different types of map data and resources needed for developing a traffic safety GIS have also been discussed. Questions were posed in the following forms: point locations (which intersections produce the highest number of fatal crashes); or segment and roadway queries (which roads have the highest incidence of bicycle or pedestrian crashes); or zonal or areal tabulations and analyses (how do cities or towns or census tracts or block groups compare in terms of the frequencies of various types of crashes). With the use of linked data, the questions and topics of inquiry could be expanded. One of the challenges for traffic safety and GIS that was faced involved the integration of different spatial databases. The potential GIS applications included mapping of pedestrian crashes and injury outcomes, installation of traffic calming devices where they will have the most impact on reducing injuries and health care costs and mapping locations with serious injuries and high health care costs in order to prioritize installation of red-light running cameras (Kim et. al. 2001, *DOT HS 809 201*). Sando et. al. explains the tool developed by Florida A&M University – Florida State University College of Engineering that aids Florida law enforcement officers in performing crash analyses using information from the TraCS (Traffic Criminal Software) database. Some of the types of analyses supported by the system were Display/Query Analysis, Intersection/Spot Analysis and Segment/Strip Analysis. Through these analyses, the Florida Department of Transportation and other law enforcement agencies could identify problematic areas and apply countermeasures to improve safety.

The Federal Highway Administration has undertaken an effort to use GIS for traffic safety analysis. Smith et al. (2007) describe the development of a common interface for organizations that desire to implement GIS highway safety analysis and GIS analysts for the development of a Geographic Information System – Transportation (GIS-T) infrastructure to support their needs (Smith et. al. 2007). Some of the analytical capabilities of GIS include display/query analysis, spatial analysis, network analysis, and cell-based modeling. Some of the features of GIS that aid in traffic safety analysis are overlay techniques that combine spatial data, such as features combined to simply add one spatial data set to another, or to update or replace portions of one data set with another data set. Proximity analysis tools like spot/intersection analysis, strip analysis, and cluster analysis are also available. The applications from the already developed 'GIS Safety Analysis Tools' system are discussed for explanation about the various analytical capabilities of the Geographical Information System. The GIS based safety analyses for the current research does not restrict itself to the highway system alone but incorporates the entire roadway network for the location.

Dangeti et. al. describes a methodology to develop a spatial data tool to aid in transportation network management. The authors discuss the issues related to designing the architecture of the system including identifying the most appropriate hardware and software components for the system. Some of the aspects of hardware for field use considered were portability, durability and effectiveness. As for the software, the program had to be easy-to-use with the ability to handle numerous data formats. Some commercial software like Autodesk MapGuide, ESRI ArcIMS, ESRI ArcPad and Terrasync were considered and a pilot study was described to illustrate the use of the system in managing and processing data related to transportation infrastructure. The concepts as described in the paper regarding the requirements of software were used in the current research. An easy-to-use system was developed with the ability to handle transportation-related data and perform analyses based on the data to obtain results.

2.4 High Crash Ranking

Ranking of high crash locations is a method of listing the locations in the order of their ranks assigned based on a chosen procedure. There are numerous procedures that are identified in the published literature. Some of the procedures that are used in high crash ranking are crash frequency, crash rates, crash loss, crash score and mapping software for density computation.

Pulugurtha et. al describes numerous methods to identify high pedestrian crash zones based on crash data and traffic volume data. Two types of crash indices based on the number of pedestrian crashes in the vicinity of the zone by severity and length of the zone were selected results were obtained for the Las Vegas metropolitan area (Pulugurtha et. al. 2004, 115-119). Based on the results, the appropriate countermeasures to improve safety in the highest ranked high crash zones were suggested.

Researchers at the Iowa State University document the methods that are used for high crash ranking projects in the state of Iowa. One of the methods identified is the crash frequency, which is just the computation of frequency of occurrence of crashes in locations (http://www.ctre.iastate.edu/research/hcl/documents/ranking.htm). On the other hand, crash rate method relies on a formula that involves the traffic volume forthe locations. Another method had been described as the crash loss method that is based on the computation of the financial loss that is caused by the crashes based on a dollar value estimate.

Pulugurtha et. al. describes a GIS methodology to identify high crash pedestrian zones and evaluates the proposed methodology of geocoding the crashes, creating crash concentration maps and then identifying high crash zones. A five year database of crash records was analyzed using numerous ranking procedures. Based on the variation in ranking results, the authors recommend the use of composite ranking methods such as sum-of-the-ranks method and crash score method instead of individual methods like crash frequency, crash rate and crash density (Pulugurtha et. al. 2007, 800-811).

Krishnakumar et. al. (2005) describes the use of GIS to rank high crash locations. They address safety concerns and ranking methods for the Las Vegas area with an emphasis on for pedestrian crashes. The kernel density was used to develop a density map for the Las Vegas area using different levels of crash concentration. Circular and linear zones created and crash score method was used to assign ranks based on the normalization of the crashes to the same scale to obtain a score for each location. A case study of the Las Vegas area is presented for illustration (Krishnakumar et. al 2005).

Vasudevan et. al summarizes many methods to identify pedestrian high crash locations and presents a statistical analysis relating the various methods. Crash indices based on frequency, weighted severity, vehicle volume and age group were considered among other methods. Thirty HCLs from the Las Vegas metropolitan area were selected and statistical analyses were performed correlating the various methods (Vasudevan et. al. 2007, 39-54). Results indicated strong correlation between the various methods for the top few ranks and the authors suggest the use of simple frequency or weighted frequency based methods as a first step.

The literature review focused on the applications of GIS in the filed of traffic safety analyses. Research efforts conducted in the analyses of crash data and proposed high crash ranking methods were reviewed. Various tools and techniques are used previously to enhance traffic safety analyses. Identification of high crash locations could be performed using the various analysis applications available in widely used commercial GIS software. The literature review suggested that though some studies are conducted in analyzing crash data using GIS, the features have to be tailored to fit the current analyses requirement. The aim of this thesis is to facilitate both micro and macro- analyses for a specific location after identification of a location using a suitable ranking procedure.

Using GIS to integrate crash data with the street network data would aid in spatial analyses of any particular location. Also, the literature review documents the existence of various methods of ranking high crash locations. Referring to Vasudevan et. al. 2007, 39-54, this thesis adopted a simple crash frequency and weighted ranking procedure for high crash ranking of locations. After identification of a particular location, the crash data for pertaining to that particular location could be queried to obtaining the summary of results in a tabular or graphical format. The development of an automated GIS system to facilitate identification of a location by means of high crash ranking procedures and performance of spatial and query analyses based on crash records on a macro and micro-level is the subject of this thesis.

CHAPTER 3

METHODOLOGY AND SYSTEM DEVELOPMENT

This chapter discusses the methodology followed to develop the safety analysis system. The chapter describes the system needs, the architecture created to develop the system, data used and the ways in which they were used to develop the various features of the tool.

3.1 Data Assimilation

The functionalities of a safety analysis tool depend on the data available. There are various types of data that are available from numerous sources. The ways in which such data can be used depend on understanding the users' needs. The initial step in the development of such a system is to identify the data required to support the desired analyses. The format in which such data are available also will affect the system development. Since this thesis aims at providing the end user with a system that would assist the querying and extraction of safety related information, the use of a detailed database of historical crash records was indispensable. It also was critical that such data be available in digital formats.

Crash data are used widely in most crash analyses to understand the occurrence of crashes and to interpret the reasons behind the crashes at various locations. This project required the distinct integration of disparate databases like crash data, street data and intersection data which enabled exploration of the relationships among the available information about the road system. Examples of analyses based solely on only crash data follow:

- All the crashes that resulted in one or more fatalities
- All crashes that had at least one pedestrian involved in the crash

Examples of analyses relating multiple databases are:

- Crashes that occurred on a speed zone of 50 mph
- Crashes per million vehicles entering an intersection.

Transportation engineers use crash data as one of the key datasets for safety analyses. In addition, crash data can be used for identification of high risk locations and factors that led to those crashes. The results of these analyses can be further used to allocate safety funds in order to reduce the occurrence of crashes and improve transportation safety at such high risk locations. Crash data are also used to perform problem identification, establish goals and performance measures, allocate resources, determine the progress of specific programs, and support the development and evaluation of highway and vehicle safety countermeasures. Furthermore, it is also useful to map the crashes for spatial analyses to improve the understanding of these data. Historical crash data and the street network had to be assimilated for the current research. One of the issues related to data assimilation in this research was the compatibility of both the data after incorporating it in a GIS interface. Data format along with the projection system had to be modified to display the crashes as points and streets as lines for a graphical view and performance of spatial analyses.

3.1.1 Data Quality

The quality of data analysis depends directly on the quality of the data. Therefore it is important to have good quality, reliable crash data. Some of the important characteristics of data quality are:

- Accuracy – The information within the database should be reliable in possessing descriptive entries for each field. Accuracy is normally enhanced by the data collecting agencies by means of conducting periodical consistency checks in the data being entered into the database.
- Accessibility and Currency – The data must be readily available for use by authorized users in a prompt and timely manner. Availability of data in a timely manner would help to identify and evaluate safety issues for that region.
- Completeness – In order to assure unbiased results, the database should have comprehensive entries of all individual reportable records that occurred during the time period of interest. Incomplete details may not be brought out while running queries on a database.
- Consistency – Information accumulated in the database should be consistent with the standards established by the national or international agencies such as the Model Minimum Uniform Crash Criteria (MMUCC) for the crash databases in the U.S. The objective of the MMUCC is to make available a de-facto standard data set for describing crashes of motor vehicles that will generate the information necessary to improve highway safety within each state and nationally (http://www.mmucc.us/). This would ensure that the data collected by various agencies within a state or within the US are consistent for overall analysis.

- Integration – The database should have at the least one common identifier to enable integration with another database from another source. An example of data integration would be the linkage of a crash with the roadway network database for spatial representation and analyses.

A brief description of the data and its sources is provided below.

3.1.2 Crash Data

A typical motor vehicle crash report includes information that describes characteristics of the events, vehicles, and persons (drivers, injured and uninjured occupants, injured pedestrians and bicyclists, etc.) involved in the crash. Law enforcement officers investigate the crash at the scene and document the information on a crash report. By using evidence found at the scene, and by interviewing participants and witnesses, the investigating officer may answer questions such as the directions in which the vehicles or pedestrians were traveling before impact or the factors that may have contributed to the impact. The Department of Transportation (DOTs) maintain such a database that has detailed entries for each crash. It is important to note that not all records have complete details. Any querying using this database will depend solely on the recorded information.

Some of the most important fields in the crash database and their description as specified in the MMUCC data element description are as shown in Table 3.1.

Table 3.1 Attributes of crash data

Attributes	Description	Details of further classification
CollisionD	Date of Collision	
CollisionT	Time of Collision	
PrimaryRoa	Primary Road	
Distance	Distance	
SecondaryR	Secondary Road	
Severity	Severity of Crash	Fatal, Injury, PDO, Other
HighestInj	Highest Injury level	K,A,B,C,O
NumofInjur	Number of Injuries	
NumofFatal	Number of Fatalities	
NumberofVe	Number of Vehicles involved	
CRCrashTyp	Manner of Crash	Angle, Animal, Fixed Object, Head-on, Pedalcycle, Pedestrian, Rear-end, Rear-to-rear, Sideswipe – meeting, overtaking
Veh1Direct	First vehicle's direction	N,S,E,W
Veh1Type	Type of the First Vehicle	Sedan, SUV, Truck, Bus, Emergency vehicle
Veh1Act	Action of First Vehicle	Aggressive, Apparently Normal, Driving too fast, Had been drinking, Inattention, Unsafe lane change,
Veh2Direct	Second vehicle's direction	N,S,E,W
Veh2Type	Type of the Second Vehicle	Sedan, SUV, Truck, Bus, Emergency vehicle
Veh2Act	Action of Second Vehicle	Aggressive, Apparently Normal, Driving too fast, Had been drinking, Inattention, Unsafe lane change,
NonMotActs	Non Motorist Action	Pedestrian walking, running, Waiting to cross
TurnLaneNu	Turn Lane Number	
RoadwayCon	Roadway surface condition	Dry, Wet, Snow
WeatherCon	Condition of Weather	Clear, Cloudy, Rain, Fog
LightCondi	Lighting Condition	Dawn, Daylight, Dusk
HwyEnvFact	Highway Environment Factor	Work zone, Debris, Glare, Icy slush
TravelLane	Lane of travel	

3.1.3 Street Network Data

In order to relate the crashes to the street network, it is necessary to accurately represent the street network information. A map format is a good mechanism to represent the street network. One method is to represent the network on maps using the centerline of the streets, The centerline representation helps accurately show the geographic layout of the streets. It can then be augmented with attributes that describe the physical and administrative characteristics as well as operational aspects of individual links and nodes. This is a common method used to link crash data to street networks.

3.2 System Architecture and Design

Almost always, transportation analyses involve working with numerous records stored in huge databases. Involving multiple databases in a single operation and integration of the data to obtain results require special design. Use of multiple databases is indispensable for querying crashes based a whole range of criteria as well as understanding the crashes and reasons associated with it. Apart from querying, using the system to locate high crash locations would involve computation of the number of crashes and use of a ranking methodology to assign ranks. Mapping this information would assist users to perform good quality analyses and retrieve data from the system.

This research involved the development of a system to perform safety analyses in terms of querying crash data and ranking of high crash locations .These processes would involve integration of multiple datasets like the crash data, street data and intersection location data. Establishment of the functional needs is an important element of system design.

Improvement of transportation safety is of top interest to a wide variety of people like transportation engineers, elected officials, decision makers and the general public. Locating areas that have safety related issues by means of monitoring the number of crashes and analyses of the crashes to identify the cause forms the first part of a transportation safety program. Identification of the locations that require attention would then be followed by development of specific countermeasures to rectify the issues. Introducing and monitoring the effectiveness of the countermeasure or the traffic safety program would follow the safety improvement procedure.

Analyses of the historical crash data recorded and maintained by the local agencies as well as the state Department of Transportation (DOTs) are essential for transportation safety improvement. Performance of analyses can be broadly classified into two levels of examination – micro and macro level. Macro level analyses involve analyses on a larger spatial scale such as a city while micro analyses are used to focus at a smaller spatial extent such as an intersection. The system should apply a common database accessible to all users of the system. The data should also be relatively easily available for future updates. Another requirement of the system would be representation of the results in a form that would be acceptable for most users. Furthermore, the software and hardware requirements should be easily available to the users.

Software and hardware requirements for using the Safety Analysis tool are described in detail in this section.

3.2.1 ArcGIS

Environmental Systems Research Institute, more commonly known as ESRI has developed numerous products in GIS (http://support.esri.com). The software that was used for development for the Safety Analysis System was ESRI's ArcGIS 9.2 version. The coding to set up the tool was scripted using Visual Basic for Application (VBA), which is a built-in feature in the ArcGIS software. The ArcGIS Desktop framework covers ArcView, ArcEditor and ArcInfo. A key application in ArcGIS is ArcMap which is used for all mapping and editing tasks as well as for running queries. Data are displayed in maps by means of layers which have associated attributes. ArcCatalog is an application that allows organization and management of the GIS data required by the user. ArcToolbox is a collection of built-in tools such as buffer, clip, overlay and export that are used in the tool development programmatically. The software used in this effort has the following system requirements:

- CPU Speed - 1.6 GHz or higher.
- Memory - 1 GB minimum.
- Space for data storage – 1 GB for the crash and streets database.
- Display and pointing device.

3.2.2 VBA Coding

ArcObjects are the actual building blocks of the ArcGIS software which can be used to customize the application as desired. ArcGIS can be customized in many ways such as

- Addition of toolbars and commands in a way suitable to the user
- Using VBA scripts and developing custom commands to perform extensive operations on features.

Extensions and custom tools can be developed using standard programming like Component Object Model (COM), .NET and Java interfaces. ArcObjects supports Visual Basic for Applications (VBA) which can be used to create forms and modules that enable by the design of a customized tool. These tools can be accessed using the ArcMap application framework. VBA is embedded in the software and provides an excellent programming environment using which a VB macro can be compiled and tested in ArcMap. VBA codes can be integrated easily with ArcObjects controls. The Graphical User Interface (GUI) for this thesis was developed using VBA by introducing forms and controls and adding codes to perform operations on the data collected. After ArcGIS installation, VBA programming can be accessed through the Tools -> Macros -> Visual Basic Editor menu choice available on any map document. Figure 3.1 shows a Visual Basic Editor window. This window can be used to create forms that have controls such as textbox and radio buttons that enable the development of a user interface frame. For each control, codes can be written to specify the function it should perform. Figure 3.1 also shows the coding window using which the actual tool functionalities can be developed.

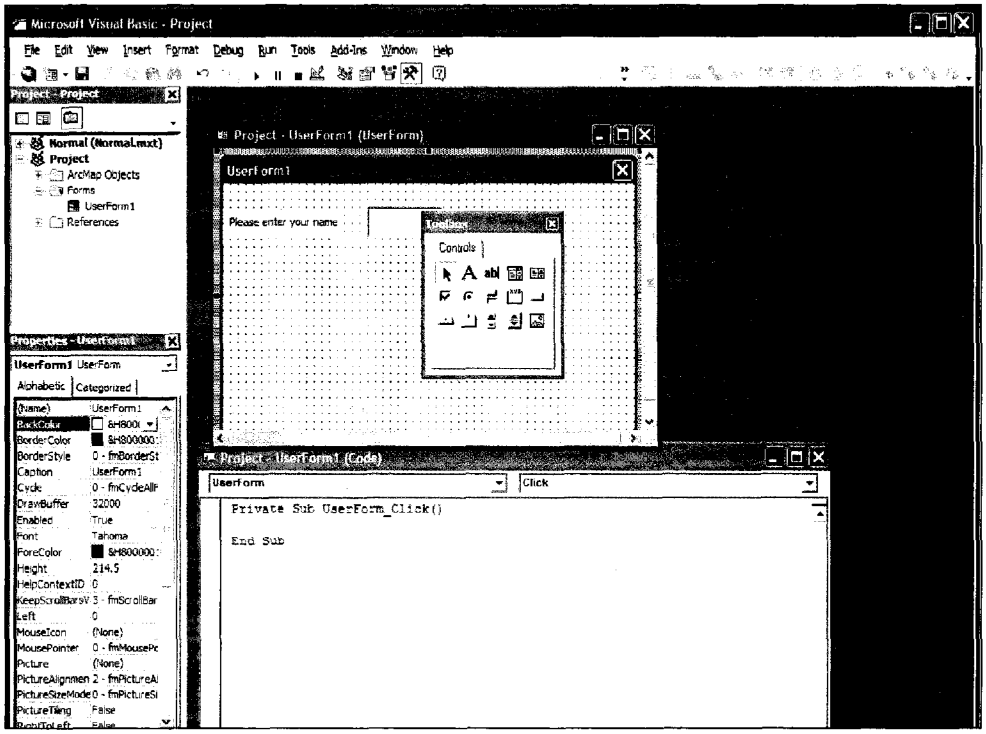


Figure 3.1 Microsoft Visual Basic Editor in ArcGIS

A basic structure that shows the user interaction with ArcGIS using VBA is shown in Figure 3.2.

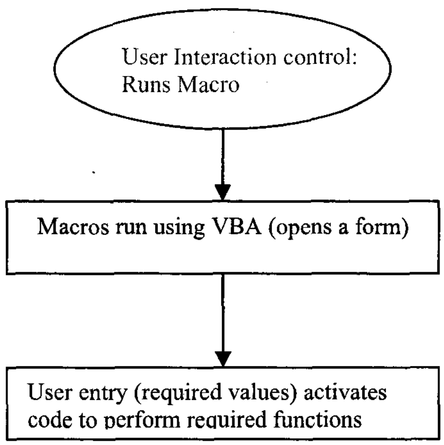


Figure 3.2 Interaction with ArcGIS – general structure

3.2.3 Microsoft Excel

Excel is a spreadsheet application from Microsoft, which helps with calculations and graphing operations. Microsoft Excel was used as one of the options for displaying results obtained from the Safety Analysis System that was developed. Microsoft Excel 2002 version was used for this application. The program has the following system requirements (http://office.microsoft.com/en-us/excel/HA102132381033.aspx):

- Windows 98, Me, 2000 Professional or XP operating system
- 140 MB of available hard disk space
- A CD-ROM drive
- Display monitor and pointing device (Mouse)

3.3 System Outputs

The main objective in developing the safety analysis tool was that to enable the end users to perform crash related queries and analyses. The tool incorporated query options for the entire crash data as well as a section solely dedicated to pedestrian and bike crashes. The desired output components of this tool were:

1. Display records that match user-defined queries in map format.
2. Analysis tool – High Crash Ranking that displays the ranking of crashes for locations based on total crashes as well as user introduced weights for severity level.
3. The tool also provides the user with a choice to export the results as a shapefile or Excel spreadsheet.

The export of output (Figure 3.3) in a shapefile format and in spreadsheet format is shown in Figure 3.4 and 3.5, respectively. The map format is displayed in the ArcMap window and table uses Microsoft Excel.

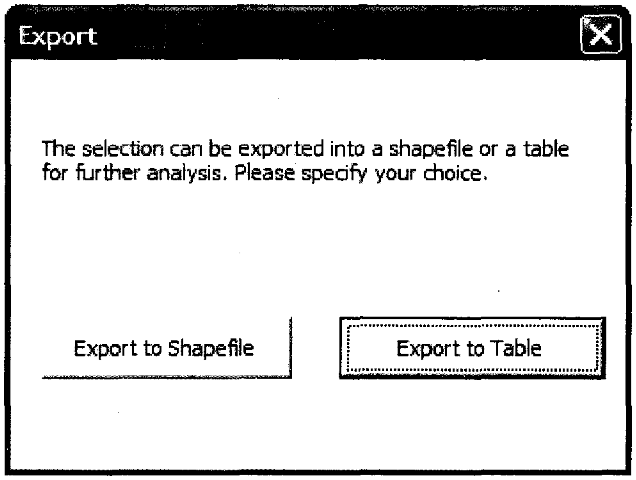


Figure 3.3 Export Options

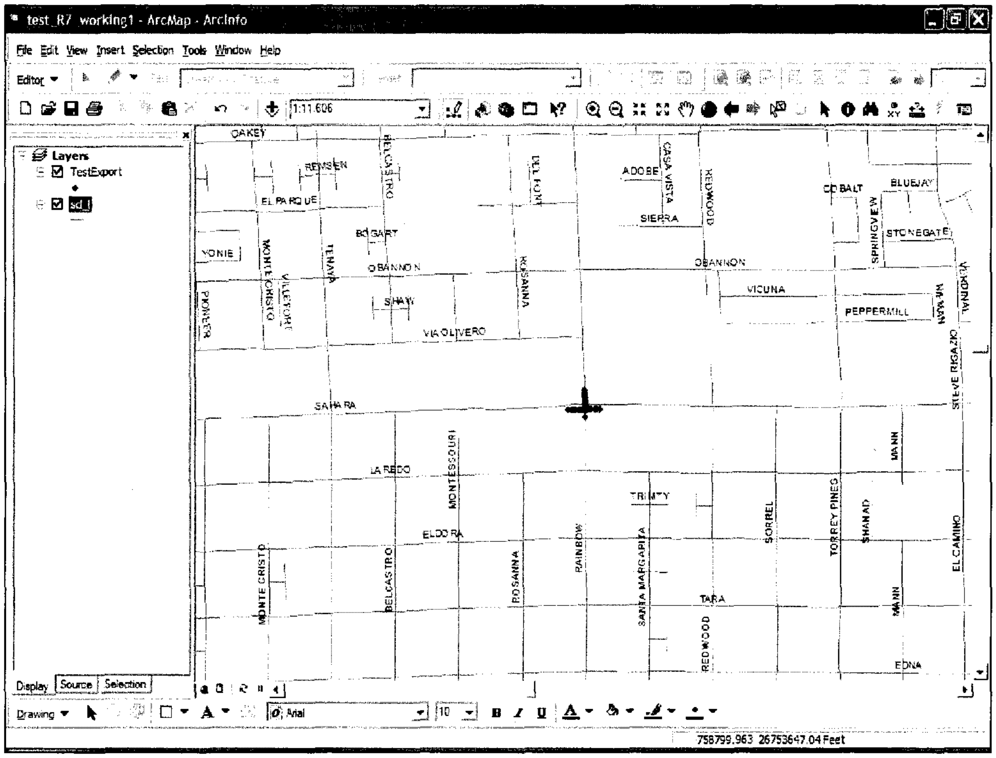
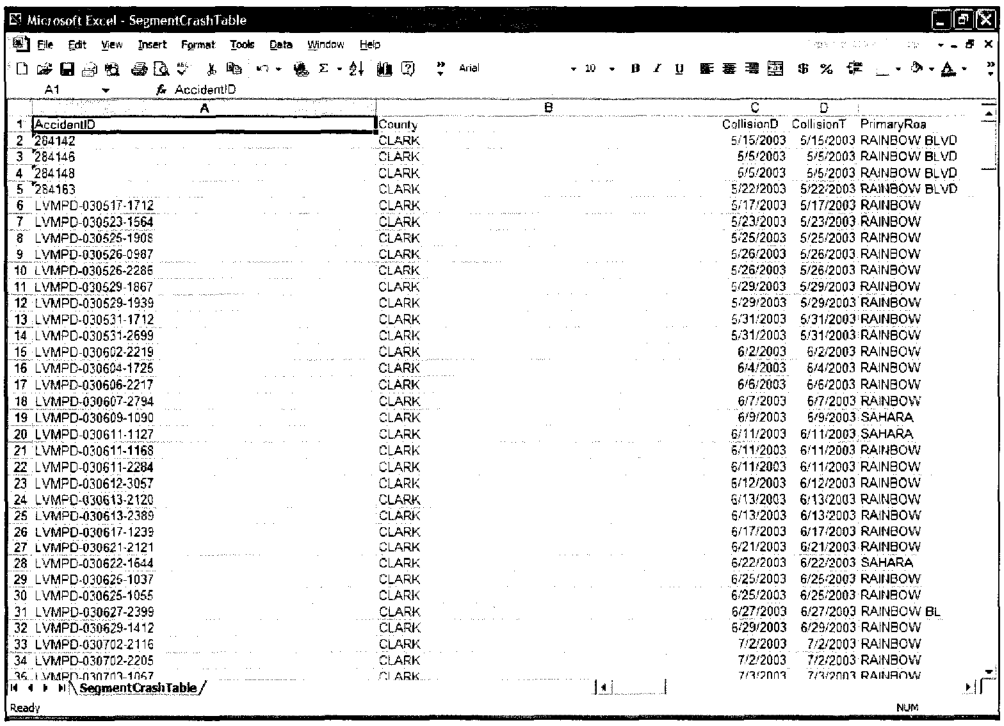


Figure 3.4 Results viewed in map format for a sample location choice



	A	B	C	D	
1	AccidentID	County	CollisionD	CollisionT	PrimaryRoa
2	284142	CLARK	5/15/2003	5/15/2003	RAINBOW BLVD
3	284146	CLARK	5/5/2003	5/5/2003	RAINBOW BLVD
4	284148	CLARK	5/5/2003	5/5/2003	RAINBOW BLVD
5	284163	CLARK	5/22/2003	5/22/2003	RAINBOW BLVD
6	LVMPD-030517-1712	CLARK	5/17/2003	5/17/2003	RAINBOW
7	LVMPD-030523-1564	CLARK	5/23/2003	5/23/2003	RAINBOW
8	LVMPD-030525-1908	CLARK	5/25/2003	5/25/2003	RAINBOW
9	LVMPD-030526-0987	CLARK	5/26/2003	5/26/2003	RAINBOW
10	LVMPD-030526-2286	CLARK	5/26/2003	5/26/2003	RAINBOW
11	LVMPD-030529-1867	CLARK	5/29/2003	5/29/2003	RAINBOW
12	LVMPD-030529-1939	CLARK	5/29/2003	5/29/2003	RAINBOW
13	LVMPD-030531-1712	CLARK	5/31/2003	5/31/2003	RAINBOW
14	LVMPD-030531-2699	CLARK	5/31/2003	5/31/2003	RAINBOW
15	LVMPD-030602-2219	CLARK	6/2/2003	6/2/2003	RAINBOW
16	LVMPD-030604-1725	CLARK	6/4/2003	6/4/2003	RAINBOW
17	LVMPD-030606-2217	CLARK	6/6/2003	6/6/2003	RAINBOW
18	LVMPD-030607-2794	CLARK	6/7/2003	6/7/2003	RAINBOW
19	LVMPD-030609-1090	CLARK	6/9/2003	6/9/2003	SAHARA
20	LVMPD-030611-1127	CLARK	6/11/2003	6/11/2003	SAHARA
21	LVMPD-030611-1168	CLARK	6/11/2003	6/11/2003	RAINBOW
22	LVMPD-030611-2284	CLARK	6/11/2003	6/11/2003	RAINBOW
23	LVMPD-030612-3057	CLARK	6/12/2003	6/12/2003	RAINBOW
24	LVMPD-030613-2120	CLARK	6/13/2003	6/13/2003	RAINBOW
25	LVMPD-030613-2389	CLARK	6/13/2003	6/13/2003	RAINBOW
26	LVMPD-030617-1239	CLARK	6/17/2003	6/17/2003	RAINBOW
27	LVMPD-030621-2121	CLARK	6/21/2003	6/21/2003	RAINBOW
28	LVMPD-030622-1644	CLARK	6/22/2003	6/22/2003	SAHARA
29	LVMPD-030625-1037	CLARK	6/25/2003	6/25/2003	RAINBOW
30	LVMPD-030625-1055	CLARK	6/25/2003	6/25/2003	RAINBOW
31	LVMPD-030627-2399	CLARK	6/27/2003	6/27/2003	RAINBOW BL
32	LVMPD-030629-1412	CLARK	6/29/2003	6/29/2003	RAINBOW
33	LVMPD-030702-2116	CLARK	7/2/2003	7/2/2003	RAINBOW
34	LVMPD-030702-2205	CLARK	7/2/2003	7/2/2003	RAINBOW
35	LVMPD-030703-1067	CLARK	7/3/2003	7/3/2003	RAINBOW

Figure 3.5 Results viewed using Microsoft Excel for the same location choice

3.4 Tool Development

The Safety Analysis was developed for the crash data as a whole and another section was dedicated solely to bicycle and pedestrian crashes for those users who wish to look at improvements to those particular categories of road users. Therefore the initial window of the tool allowed the user the choice of working with the complete crash data or restricting the analysis specifically to bicycle and pedestrian crashes. The "welcome" or opening window of the Safety Analysis System is shown in Figure 3.6.

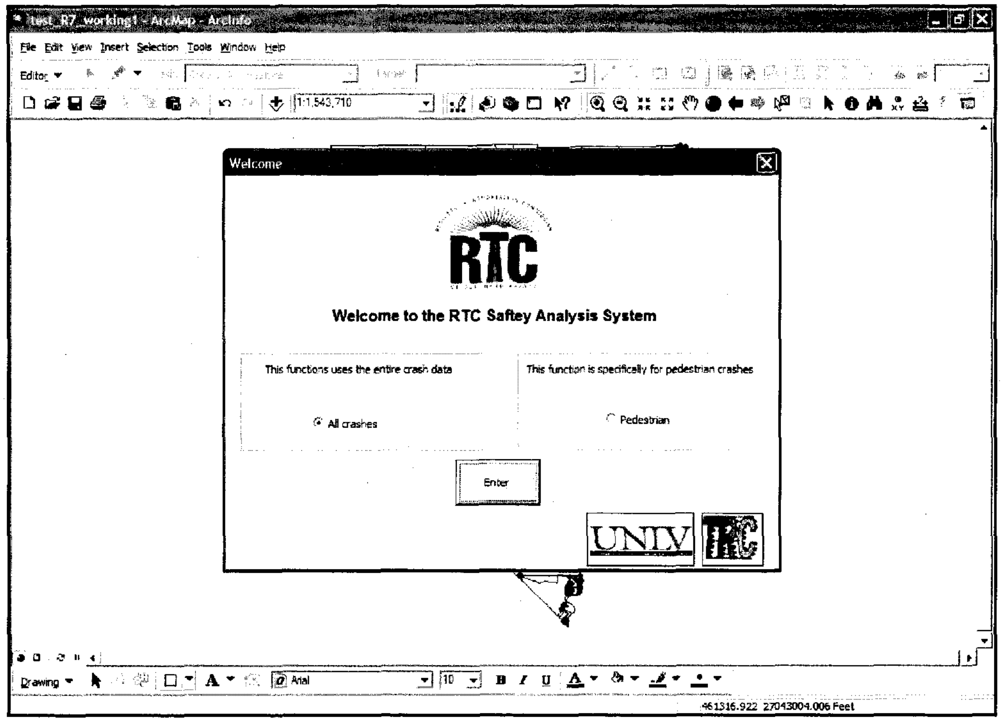


Figure 3.6 Opening window of the SAS tool

The development of the entire Safety Analysis system can be described in two parts. First, the Crash Data analysis system and then the High Crash Ranking analysis system for various locations, each of which is discussed in detail in the following sections. Both of these tools were common to the two options shown in Figure 3.6 that allowed the database selected to be the entire crash database or the pedestrian data alone.

3.4.1 Crash Data Analysis

This section describes the features of the tool that are dedicated primarily to analysis of crash data. The structure of the Crash analysis tool is illustrated in Figure 3.7

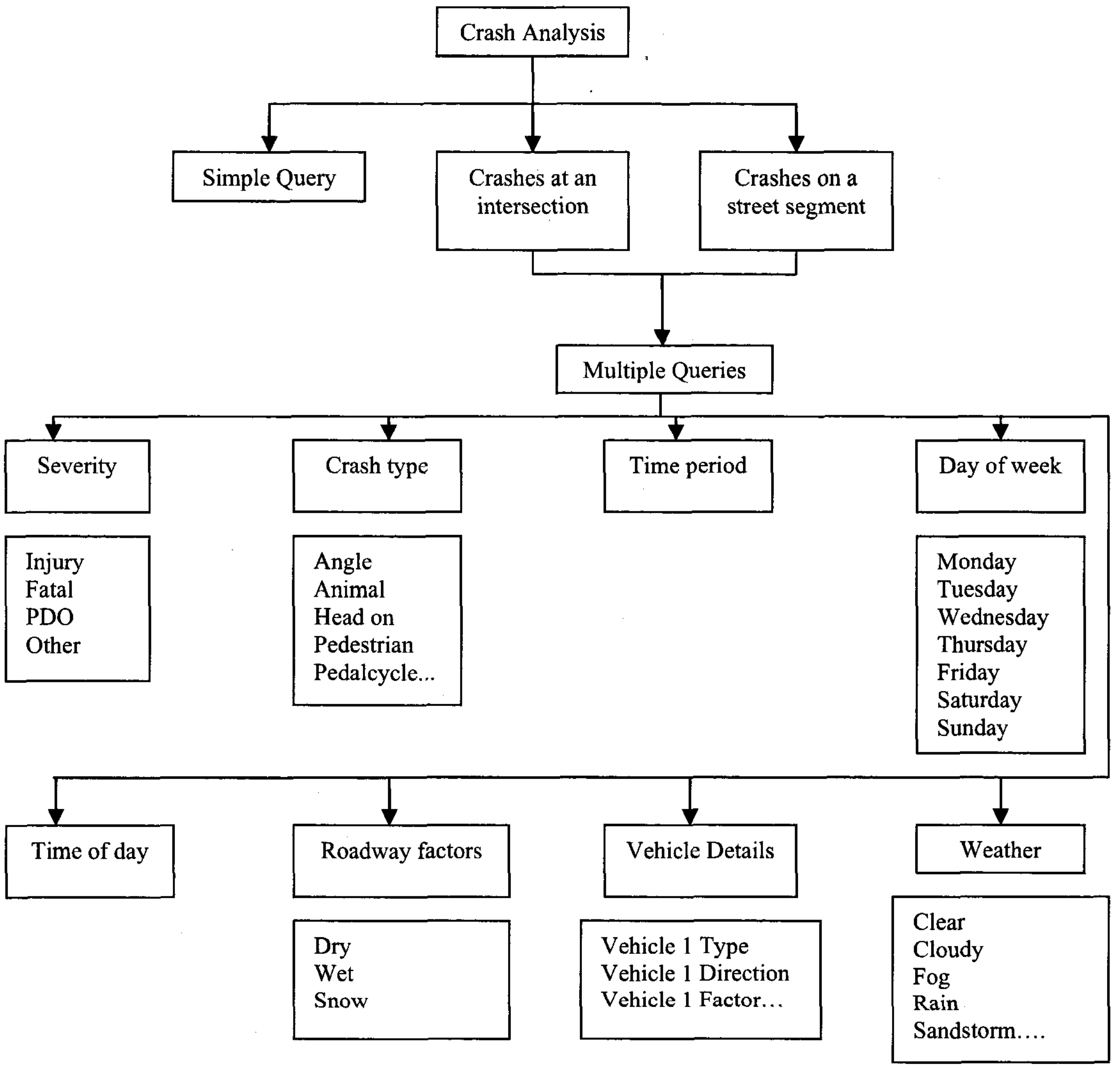


Figure 3.7 Crash Analysis Tool structure

The main window of the crash analysis system that was developed using VBA coding has three search options to enable the user to either perform area wide analyses or to establish a location of interest and confine analysis to crashes specific to that location. The three options provided to the user are:

1. Simple search
2. Crashes at an intersection
3. Crashes on a street segment

3.4.2 Simple Search

This is the basic search option that is provided to the user to perform a quick search of the crash data for any criterion that the database supported. This search window resembles the one portrayed in Figure 3.8.

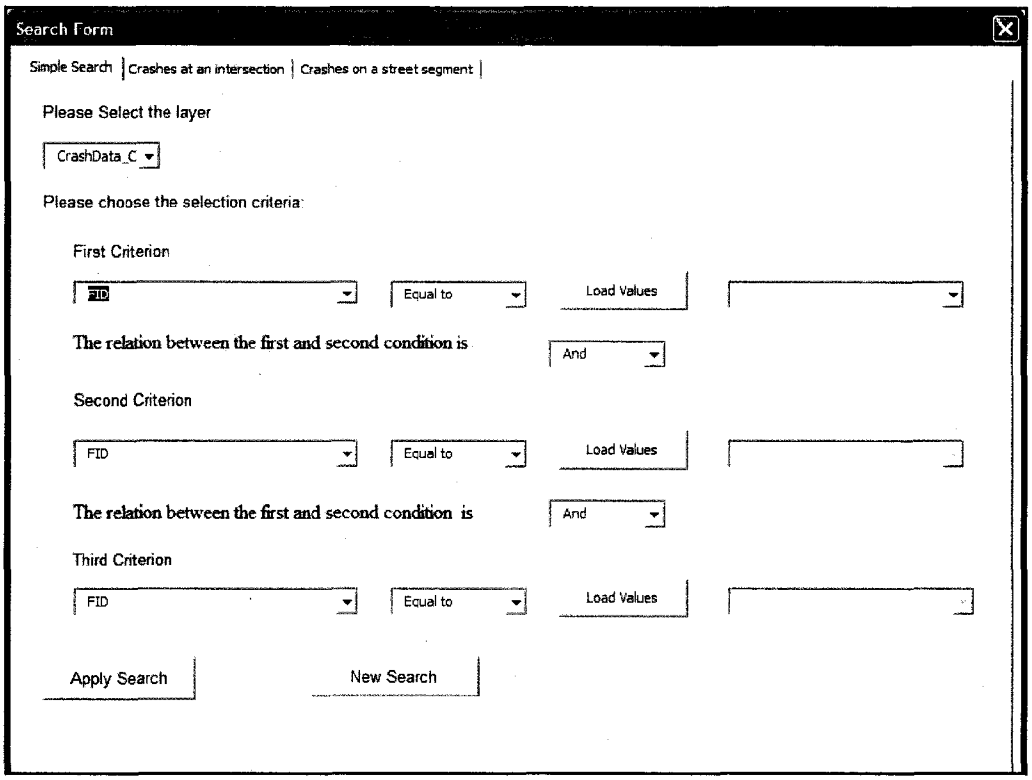


Figure 3.8 Simple Search Window

The user entry in this section is restricted to three entries. Each of these allows the user to query the crash database and display the results in the map window. The tool also allows the use of operators between the data thus enabling the user to build a combination of criteria to narrow results as desired. For example, if the user wished to find all fatal crashes that were classified under the "Head-On" collision type, then the user would choose "Severity" from the first criterion combo-box. Upon clicking the 'Load Values' button, the different values for Severity that are available in the database would be loaded to the combo-box beside it. The user then would choose the 'Fatal' option from this combo box and repeat the same procedure with the second criterion set for Crash type of 'Head-On' with an 'And' operator between them. These options as loaded into the search form are illustrated in Figure 3.9.

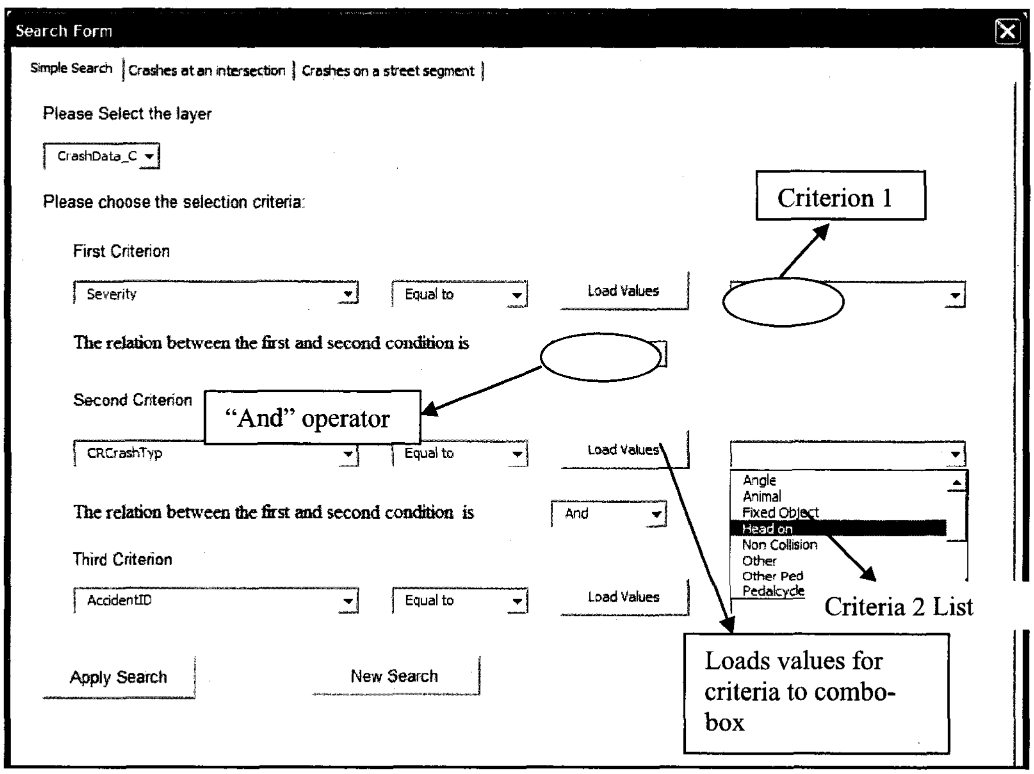


Figure 3.9 Search Window with user-specified choices

The main objective of this section of the system is to accommodate those users who wish to quickly examine data that met specific criteria. The following are examples of the types of questions that can be addressed:

- Crashes that occurred between 1:00 PM and 2:00 PM
- Crashes that involved at least one pedestrian
- Fatal crashes involving DUIs
- Crashes that occurred while pavement conditions were wet

Two additional search options are provided for users who wish to perform other analyses regarding specific locations. These are described next.

3.4.3 Crashes at an Intersection

A tool to locate all crashes that occurred at an intersection is one among the two features developed to provide the ability to permit the user to focus on a study location or area. The main objective of the intersection level analysis option is to analyze safety aspects at an intersection location. The user inputs required for this analysis include the following data fields, and a screenshot of this interface is shown in Figure 3.10.

- Name of the first cross street (selected from a drop-down menu)
- Name of the second cross street (selected from a drop-down menu)
- A user-defined buffer distance (numeric field)
- Units for the buffer distance (feet or meters)

The first required data field entry for operating this tool is the name of one of the streets that intersect to form the intersection. The user interface form is programmed to upload all the available street names so as to assist the user in data entry. Upon clicking the dropdown button (see Figure 3.11) next to the 'Name of the First Street' box on the form, the system would load all the possible names that are available from the street network (e.g., Street Centerline) database. There are two ways in which data choice could be made. The first option for the user is to choose one entry from the loaded list directly, and the second option is to spell out the street name. In the latter case the system would prompt the user the correct way to spell the street from the list that was already loaded. This was done to ensure that the users do not mistype the street names. Only street names included in the street network database are acceptable inputs for street names.

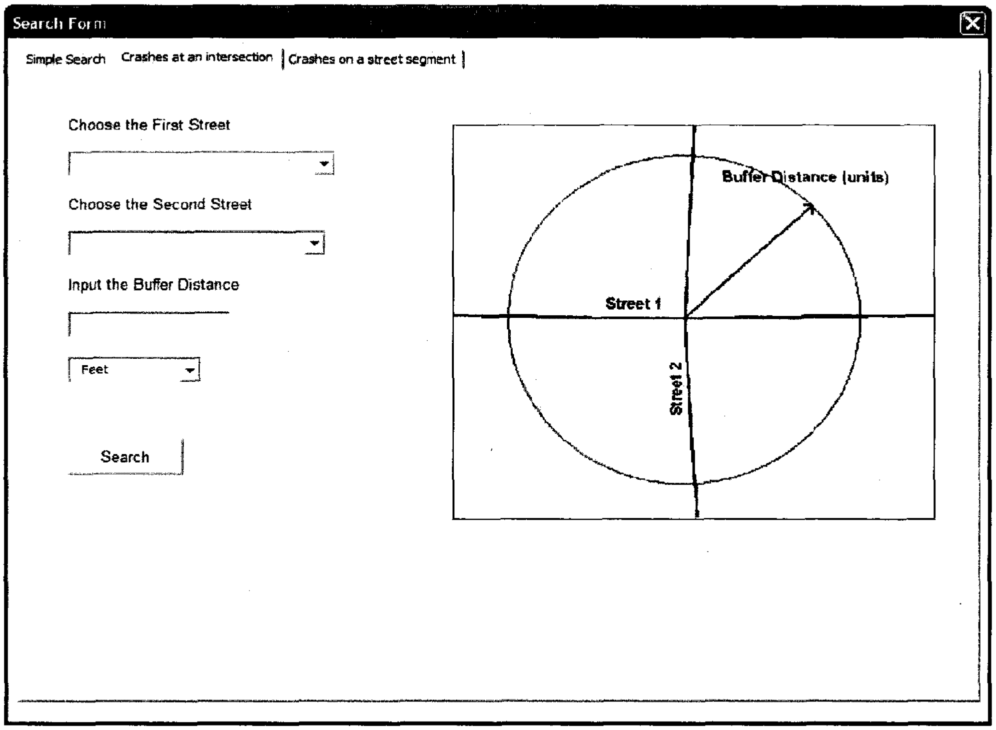


Figure 3.10 User interface to identify an intersection for crash data analysis

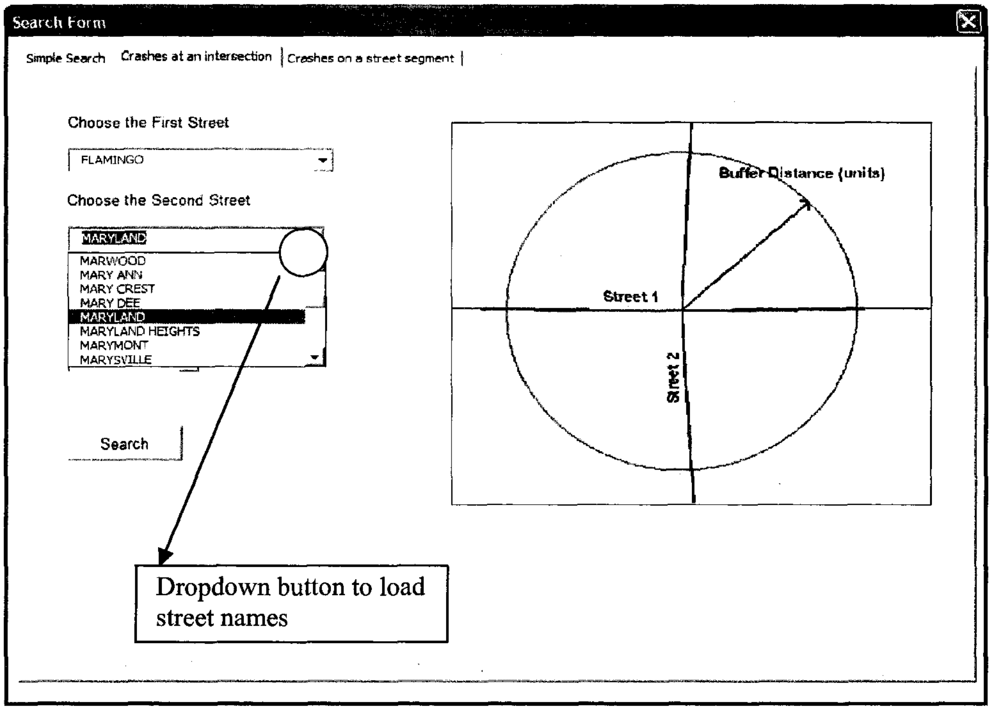


Figure 3.11 Identifying street names to select an intersection location

The aforementioned process also applies for the data entry of second street name. As the data for these two fields were necessary to carry out this operation, the system is programmed to display a message to prompt the entry of acceptable entries (names) for both the street names if the user failed to enter acceptable names before clicking the 'Search' command button. The system was also coded to check if the identified streets actually do intersect on the street network map. If they did not intersect, a pop-up box is displayed showing the following error message 'Please enter intersecting street names.'

Another data entry required from the user is the distance around the intersection for which the crashes are to be analyzed. This distance is called the 'Buffer distance.' The buffer is defined as the area within a specified distance (radius) around the identified location. The tool allows the user to enter a desired numeric value along with its representative units (either feet or meters). The system draws a buffer zone with the specified radius around the selected intersection. Upon clicking the 'Search' button, the program first checks if all the required fields have indeed been provided by the user. Else, it would not proceed to the actual function of drawing the buffer and selecting the crashes within the given area. Upon the user providing the required inputs and clicking the 'Search' button, the system completes the analysis and displays as points all the crashes within the selected buffer area. Figure 3.12 summarizes sequence of steps followed by the program to perform this analysis.

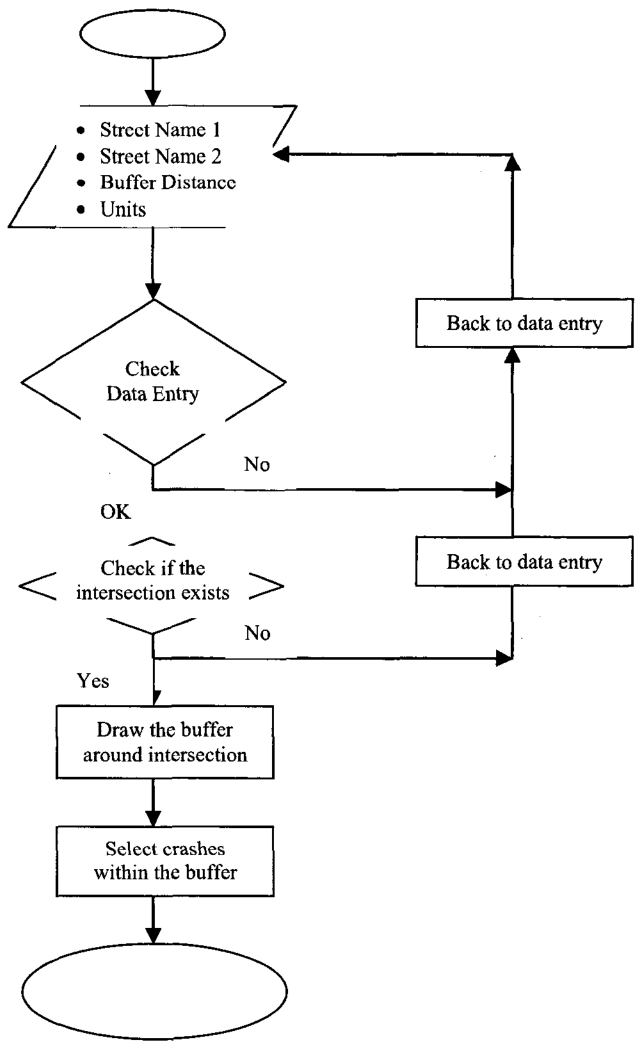


Figure 3.12 Flow chart for selecting crashes for intersection level analysis

3.4.4 Crashes on a Street Segment

The third option for crash analysis is based on identifying street segments. An user interface help to identify a segment of a given street between two cross streets that intersected the former street. This option permits evaluation of long segment of a street comprised of several smaller sections or just one section. A customized program was developed to facilitate this process. The inputs required from the user are somewhat similar to that for the intersection based analysis. The 'crashes on a street segment' screen on the whole looked like the one shown in Figure 3.13. The user-entered data fields in this case were as follows:

- Name of the main street
- Name of the first cross street
- Name of the second cross street

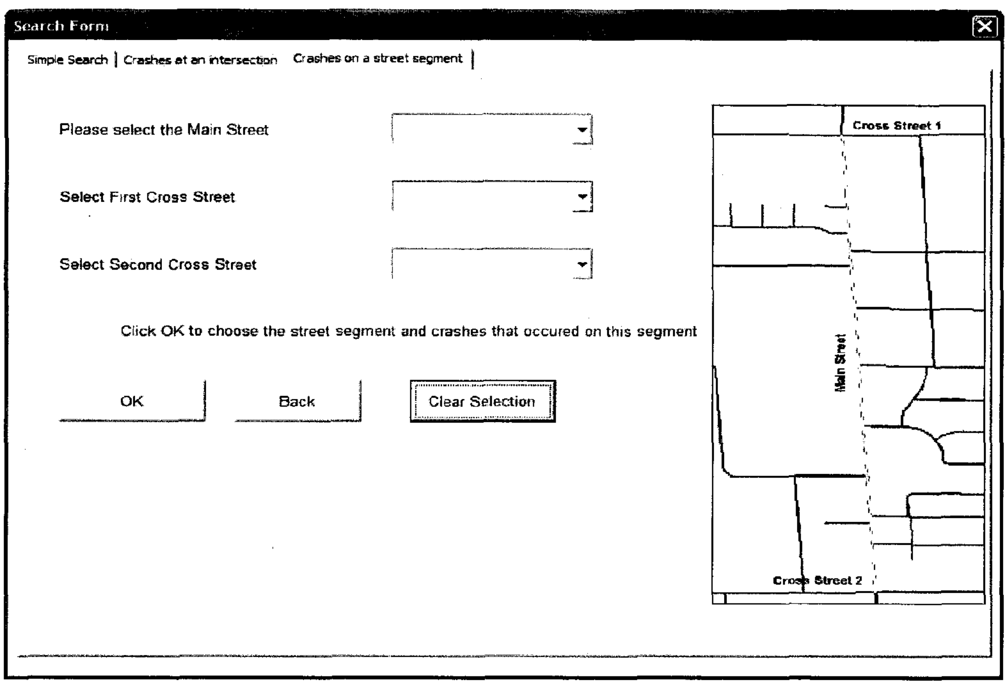


Figure 3.13 User interface to identify a street segment for crash data analysis

The first required data field is to the name of the main street. This choice denotes the street on which a segment is to be defined to identify the crashes for further analysis. Similar to the 'crashes at an intersection' tool, the list of street names from the street centerline database are loaded onto the combo-box next to the label 'Please enter the name of Main Street.' The user is allowed to choose one name from the list or type it in as the system provided prompts of street names with matching spelling.

For the next two entries, the user has to choose the names of the two cross streets which bound the desired street segment. It is important to select the bounding streets so as to ensure that the segment defined exists in reality. If a street segment is not identified by the system, it would just create an empty file and a blank display to reflect that the input specification was incorrect. As in the case of the main street, the lists next to the cross streets are programmed to automatically load the names of the streets stored in the database to facilitate data entry. This is shown in Figure 3.14.

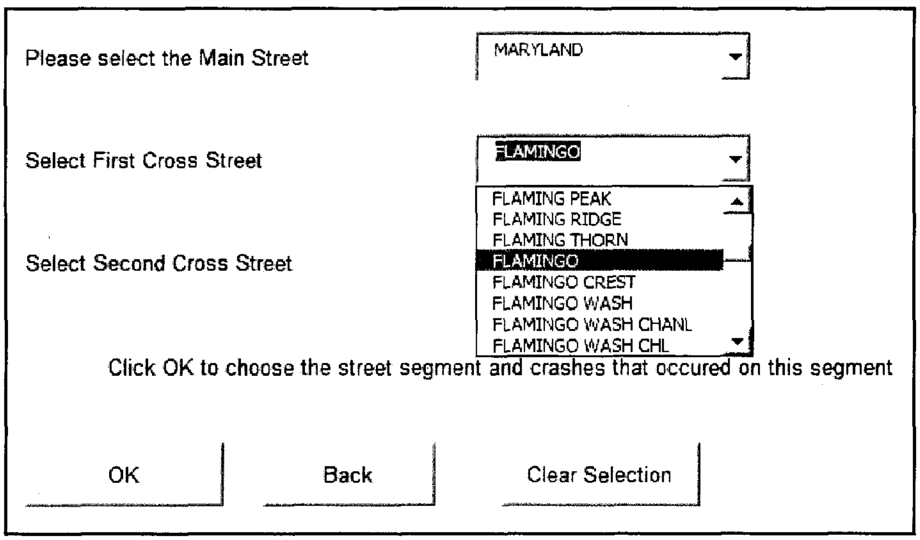


Figure 3.14 Identifying street names to select a street segment for analysis

The procedure adopted to extract the crashes that occurred on the selected street segment is described next. Upon clicking the 'OK' button to define the desired street segment, (shown in Figure 3.14) the system was programmed to automatically perform the following actions. First, the selected main street is identified in the street network the map and database. Then, to account for potential spatial errors in locating the street sections, the 'Buffer' tool is used to draw a buffer along the street of a pre-specified size of 100 feet. This results in an area of 100 feet on either side of the street centerline to be considered a part of the street for the ensuing analysis. Next, all the streets that were located within this buffered area are extracted. Thus, a buffer with all the streets inside it is created on the street network map. A sample of this is shown in Figure 3.15.

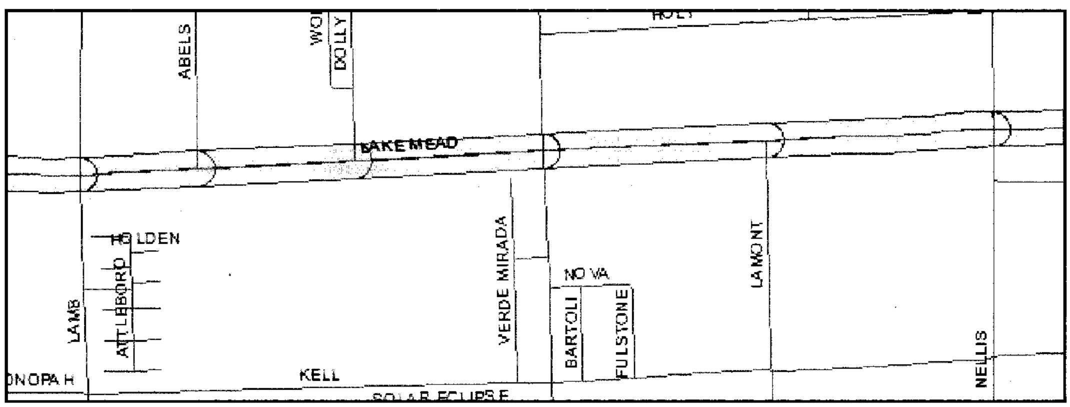


Figure 3.15 Clipped streets within 100-foot main street buffer

From the file containing the buffered area around the main street and all the clipped cross streets, the street segment defined by the user is selected. Thus, a file containing the segment of the main street between the two specified cross streets is created. A buffer of 50 meters is drawn around the main street to identify the crashes that fell within this buffer. This buffering process is used to help address potential errors in coding locations of crashes on to the street network. It results in any crash located within 50 feet of the centerline of a street to be considered to have been on the street. A flow chart summarizing the steps followed by the program to perform this analysis is shown in Figure 3.16.

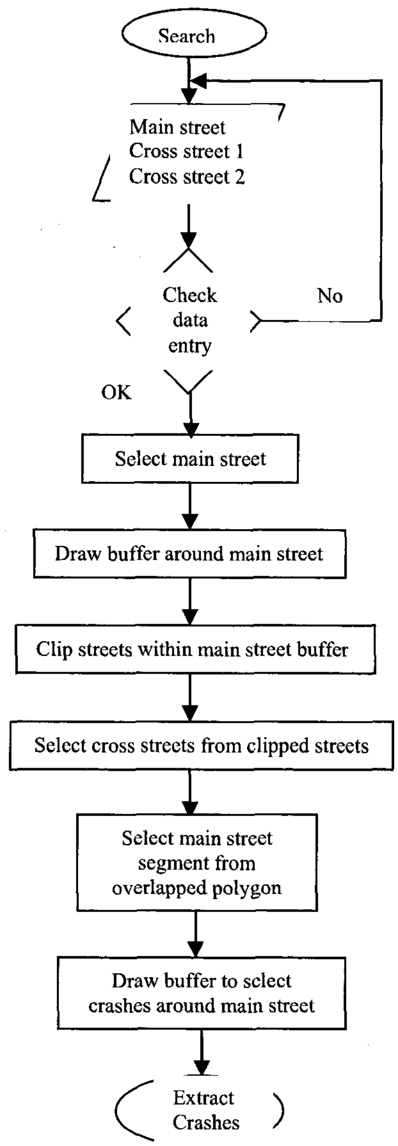


Figure 3.16 Flow chart for selecting crashes for street segment based analysis

Upon completing the simple search or after defining an intersection or street segment for detailed analysis, the system prompts the user for inputs required for the detailed analysis. This is described next.

3.4.5 Multiple Query Criteria

Based on the selected intersection or street segment, the system provides the user an interface to build a query based on various user defined criteria using data fields from the crash database. The user inputs to select the criteria for the multiple query analysis are placed on four tabs in four broad categories as follows: date & time, crash characteristics, vehicle characteristics, and road characteristics. Figures 3.17 to 3.21 show the user interfaces for each of these input groups. The window was programmed to have four tabs to include all the important datasets into the query builder. The four tabs corresponding to the datasets contained in each group are as follows:

- Date/Time information (as shown in Figure 3.17)
- Crash information (as shown in Figure 3.19)
- Vehicle information (as shown in Figure 3.20)
- Road information (as shown in Figure 3.21)

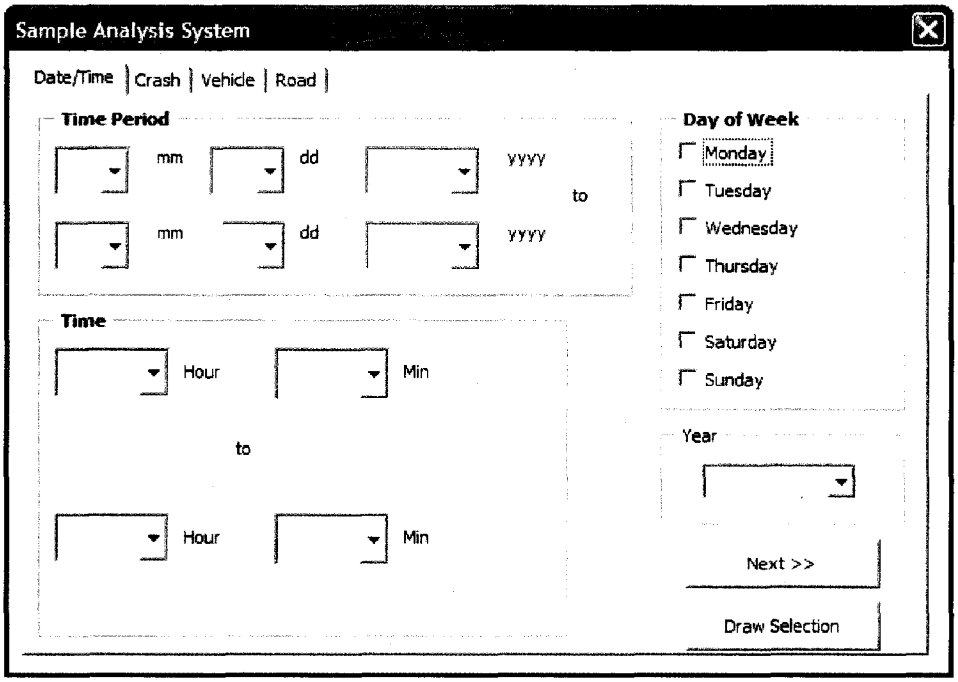


Figure 3.17 Multiple query window – Date/Time choices

The date/time tab allows the user to choose either a specific year or a time period (of multiple years) using the combo-boxes as shown in the frame titled 'Time Period.' If a choice for a specific year was made, then the time period option was disabled and similarly if a time period was chosen, the year choice was deactivated to prevent data complication. Day of week choices were also provided as shown in the data frame titled 'Day of Week' to the right of the window. The data were modified from the 'Collision Date' information from the database to be compared with a calendar to find the weekday on which the crash occurred. Also, the time of crash information could be included in the analysis provided that the original crash database actually contains the time of crash as entered by the officer on duty. The details of of the data frames are shown in Tables 3.1 to 3.4.

Table 3.2 Date/Time options – Year

Data Field	Entries					
Year	2002	2003	2004	2005	2006	2002-2006

Table 3.3 Date/Time options – Time Period

<table>
<tr><th>Data Field</th><th colspan="3">Entries</th></tr>
<tr><td rowspan="3">Time Period</td><td>mm (1)</td><td>dd (1)</td><td>yyyy (1)</td></tr>
<tr><td colspan="3">to</td></tr>
<tr><td>mm (2)</td><td>dd (2)</td><td>yyyy (3)</td></tr>
</table>

Table 3.4 Date/Time options – Day of Week

Data Field	Entries						
Day of Week	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday

Table 3.5 Date/Time options – Time of Day

<table>
<tr><th>Data Field</th><th colspan="2">Entries</th></tr>
<tr><td rowspan="3">Time of Day</td><td>Hour (1)</td><td>Minute (1)</td></tr>
<tr><td colspan="2">to</td></tr>
<tr><td>Hour (2)</td><td>Minute(2)</td></tr>
</table>

For the aforementioned choices, the 'Day of Week' alone has checkboxes to permit multiple entries by the user. For the rest, combo-boxes are used to allow the user to include only a single combination to prevent complications with data manipulations needed for the analyses. For example, the time period option allows the user to develop only one statement like the one shown in Figure 3.18.

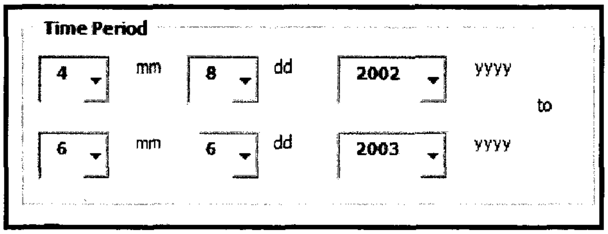


Figure 3.18 Development of a query statement

The second tab is the data group with information related to the crash occurrence. The data frames included in this tab are the type of crash and its severity. A screenshot of the user interface for this tab is shown in Figure 3.19. Tables 3.6 and 3.7 show the data fields that stored in the database for the 'Crash Type' and 'Severity' fields, respectively. These data fields are in the form of checkboxes to allow multiple selections.

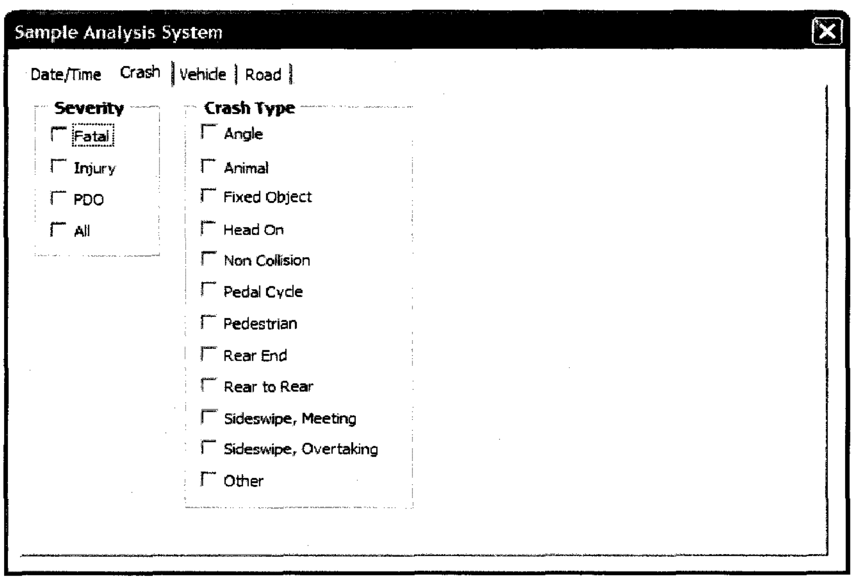


Figure 3.19 Muliple query window – Crash Information User Interface

Table 3.6 Crash Details – Severity

Data Field	Entries			
Severity	Fatal	Injury	PDO	Other

Table 3.7 Crash Details – Crash Type

Data Field	Entries
Crash Type	Angle
	Animal
	Fixed Object
	Head On
	Non Collision
	Pedal Cycle
	Pedestrian
	Rear End
	Rear to Rear
	Sideswipe, Meeting
	Sideswipe, Overtaking
	Other

The next tab is for information related to the vehicles that were involved in the crash. The vehicle type, the direction in which it had been traveling, and the factor contributing to the crash (as noted in the crash report by the law enforcement officer) are provided in this tab for the vehicles involved. Figure 3.20 shows the Vehicle information and the data frames contained within it. It is to be noted that from the crash database that each of these fields had numerous unique entries under them. For example, a summary of crash database for the Vehicle 1 Type showed nearly 200 entries for the vehicle types. Since it was impractical to give these many options to the user, a list that covered at least 90% of the data was created and the tool was designed to include these options. Tables 3.8, 3.9 and 3.10 show the various data entries for the first vehicle's details. The same combinations are used for the second vehicle that was involved in the crash.

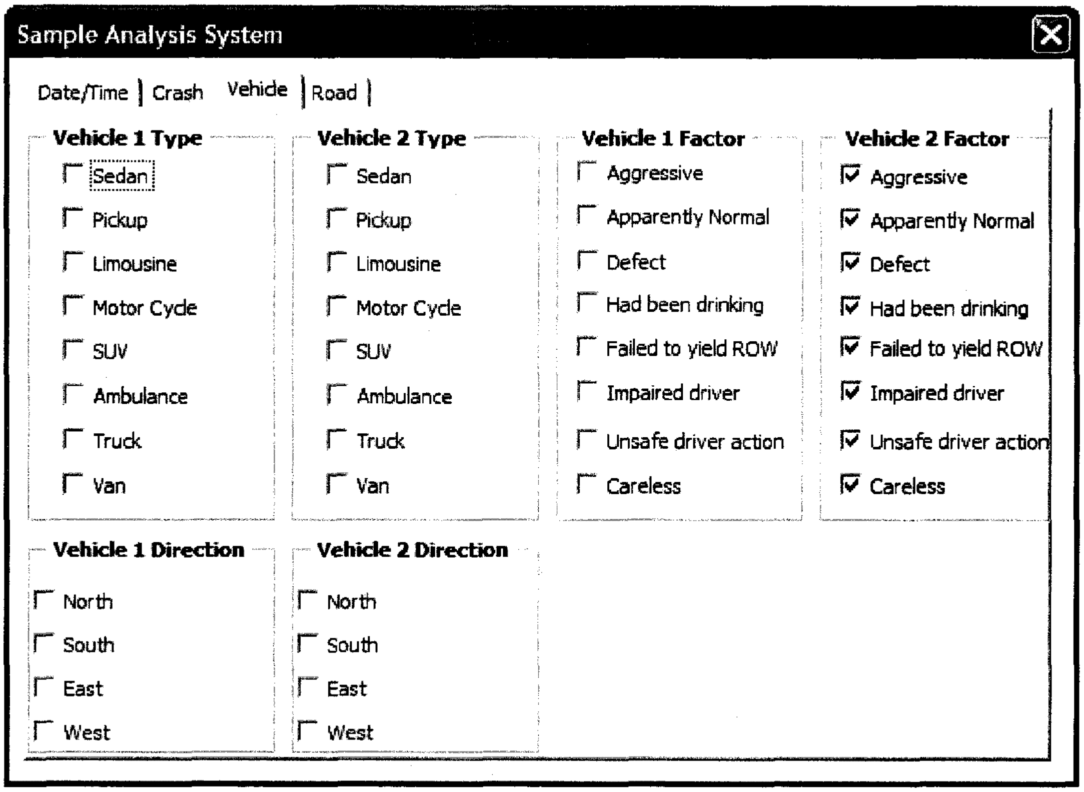


Figure 3.20 Multiple query window – Vehicle Information choices

Table 3.8 Vehicle Details – Vehicle Type

Data Field	**Entries**
Vehicle 1 Type	Sedan
	Pickup
	Limousine
	Motorcycle
	SUV
	Ambulance
	Truck
	Van

Table 3.9 Vehicle Details – Vehicle Direction

Data Field	**Entries**			
Vehicle 1 Direction	North	South	East	West

Table 3.10 Vehicle Details – Vehicle Factor

Data Field	**Entries**
Vehicle 1 Factor	Aggressive
	Apparently Normal
	Defect
	Had been drinking
	Failed to yield
	Impaired driver
	Unsafe driver action
	Careless

The final category of criteria that could be used in a user-defined query is under the tab titled 'Road' information. This window has four data frames: Highway Environment, Lighting Condition, Road Factors, and Weather Conditions. As in the case of Vehicle details, some of the Road characteristics such as like Highway Environment also had numerous possible entries. The list was shortened to include those values that covered at least 90% of the total possible entries. All of the four data frames contain checkboxes as the mode of chosen display in order to allow multiple choices to be included in the query statement. The data entries for each of these data frames are shown in Tables 3.11, 3.12, 3.13 and 3.14.

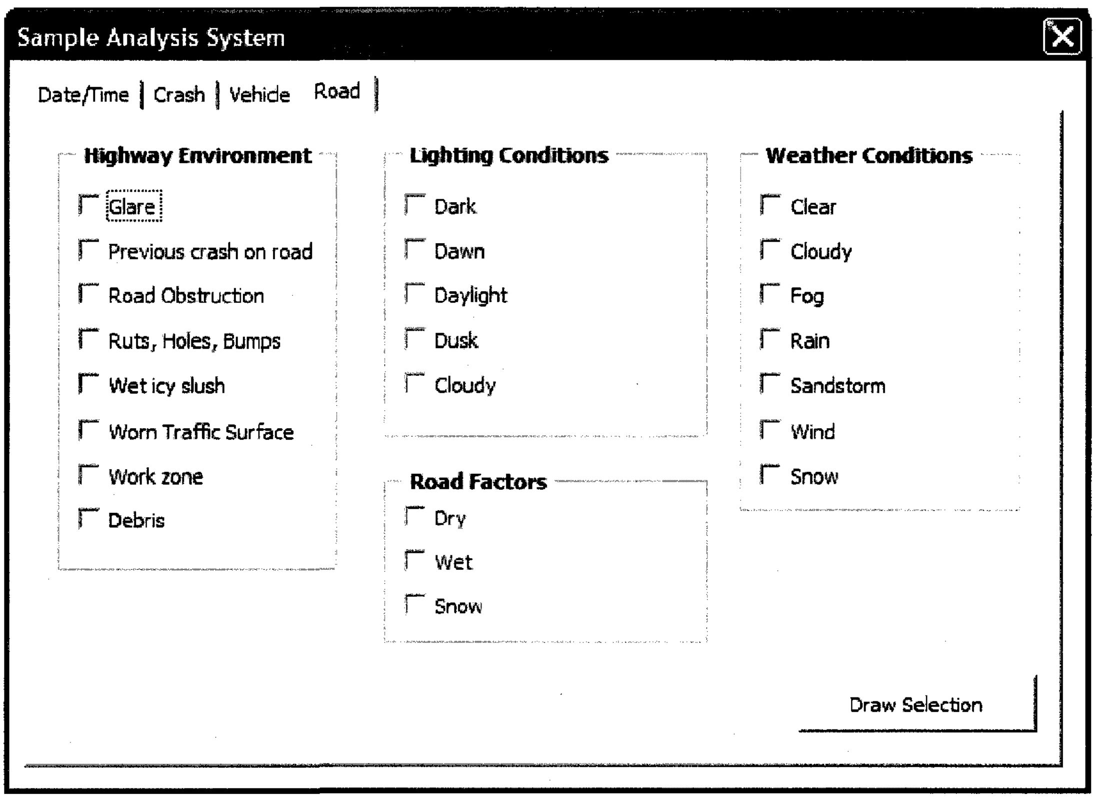


Figure 3.21 Multiple query window – Road details

Table 3.11 Road details – Highway Environment

Data Field	Entries
Highway Environment	Glare
	Pedestrian in X-walk
	Previous crash on road
	Roadway Obstruction
	Ruts, holes and bumps
	Wet icy slush
	Worn traffic surface
	Work zone
	Debris

Table 3.12 Road details – Lighting Condition

Data Field	Entries
Lighting Condition	Dark
	Dawn
	Daylight
	Dusk
	Cloudy

Table 3.13 Road details – Road Factors

Data Field	Entries		
Road Factors	Dry	Wet	Snow

Table 3.14 Road details – Weather condition

Data Field	Entries
Weather Condition	Clear
	Cloudy
	Fog
	Rain
	Sandstorm
	Wind
	Snow

Customized software codes help convert the user inputs registered through the user interfaces are to analytical queries which would be recognized by the GIS software selected (ArcGIS) for the system. Inconsistencies in the data entry by the law enforecment officers and errors in spelling meant that the system should be capable of distinguishing many combinations of spellings and data entries. For example 'Sedan,' 'SEDAN', 'Sedan 2D' and 'Sedan 4D' essentially should be accounted for if the users chooses a vehicle type of 'Sedan' as one criterion. Efforts were made to include typical and common combinations for each data field. The system is programmed to ensure that each query would cover at least 90% of the likely possibilities. The fields that had very few entries were ignored and more commonly represented entries were accounted for.

After the query statement is formed, the target file to be queried is set as the output file that was created during the operations described in the previous sections – crashes that occurred at an intersection or a street segment. All the crash features that had the same combination as the user-specified criteria are selected and displayed on the map for data export or more analysis. Figure 3.22 shows these options along with an illustrative selection done during the sample run of the tool.

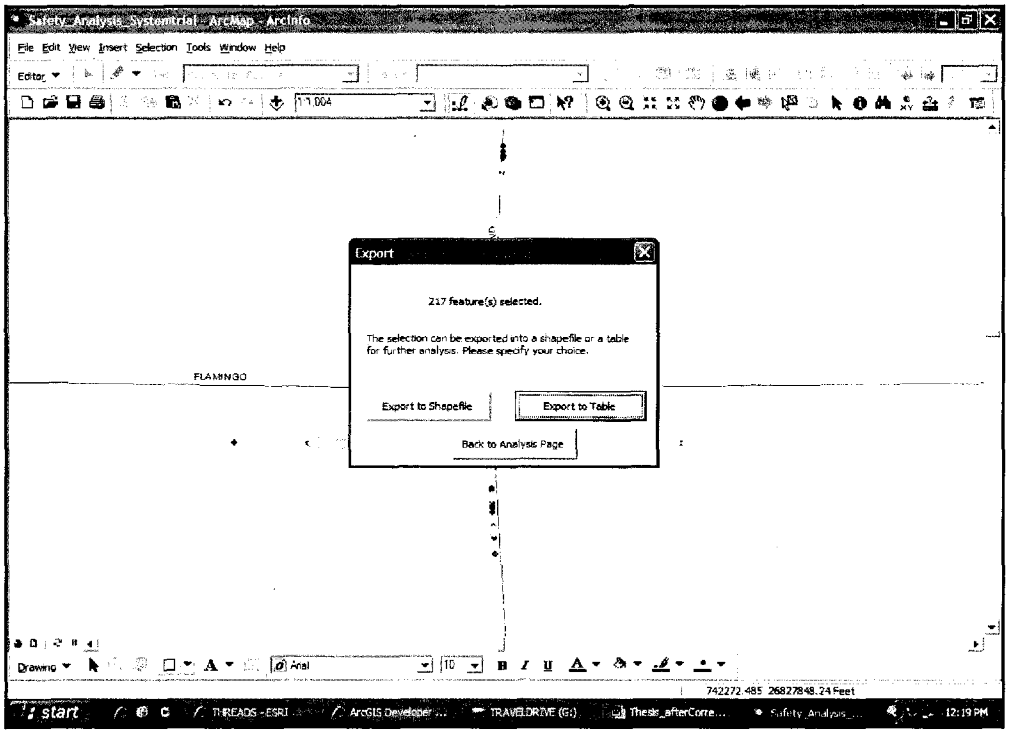


Figure 3.22 Selection after querying and export options

3.4.6 High Crash Ranking

The Safety Analysis system also includes a tool to rank high crash locations. This is based on a spreadsheet document containing a list of the locations (intersections) created from the existing crash data. The crashes within the specified buffer distances of 100 feet, 200 feet, and 500 feet around the locations are extracted for each of the intersection. Using customized software coding (Excel macros), the totals for each of the interactions were calculated. A VBA macro was specifically developed using Microsoft Excel to compute a total of the crashes occurred in each location with a column listing for fatal, injury, and PDO type of crashes. Combinations of years (2002 to 2006) and buffer sizes (100, 200, and 500 feet) are used to create a simple set of files and minimize the program execution time. These files when added onto ArcMap can be viewed because the original file has the coordinate location of each value. Thus the files can be displayed as point features denoting the total number of crashes and its attribute table along with the number of fatal, injury, and PDO crashes at that location based on the year of occurrence and buffer distance.

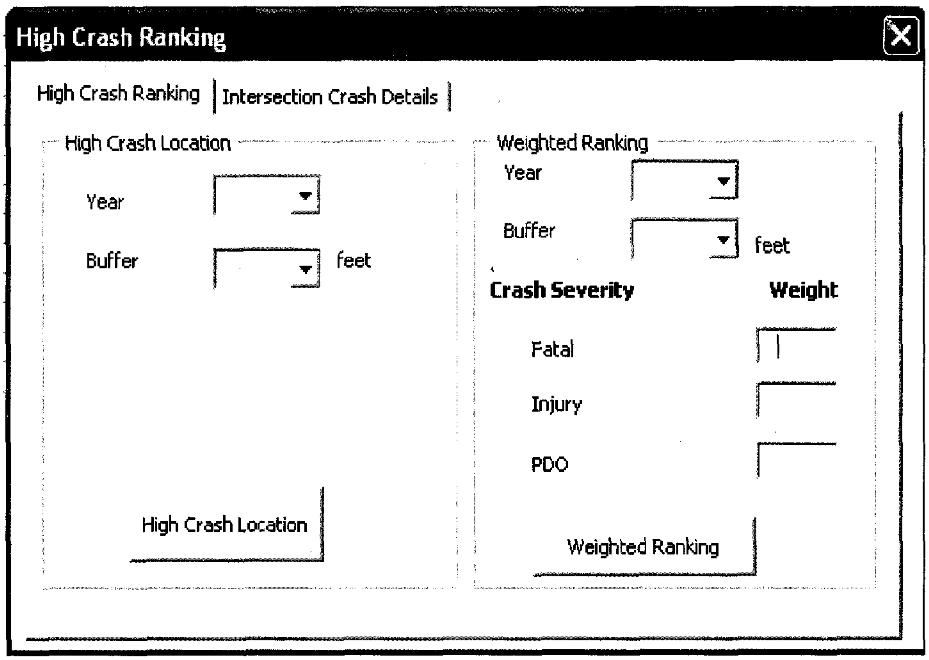


Figure 3.23 High Crash Ranking tool

The user interface window for the High Crash Ranking tool is shown Figure 3.23. The window provides the user two options of ranking the locations: the first option is based on the simple crash frequency method where ranks were assigned to the locations based solely on the total number of crashes that occurred at that location irrespective of the severity levels of the crashes. The user input entry required for this option is the year for which ranking is to be performed and the buffer size to be considered around the intersection. The second option is based on using different weights based on crash severity The frame to the right side of the user interface provides the user options to include weights for the severity levels in the calculation of ranks for the various locations. The user can provide an integer value to be considered for the calculation of the total for which ranks needed to be assigned for each of the severity levels: fatal, injury and PDO. The user could enter any value based upon the user's requirement of the scale in which severity needs to be included. These weights are included in the following formula.

$$\text{Weighted Sum of Crashes} = N_f \text{ x } X_f + N_i \text{ x } X_i + N_p \text{ x } X_p$$

Where

N_f = number of fatalities

X_f= Fatality factor (user-entered)

N_i= number of injuries

X_i= Injury factor (user-entered)

N_p= number of PDO

X_p= PDO factor (user-entered)

A default value of one (1) is provided if the user fails to enter any one of the weights. Based on the sum computed using the above mentioned formula, new ranks are assigned to each of the locations and the table with the location details and its rank is displayed for the user.

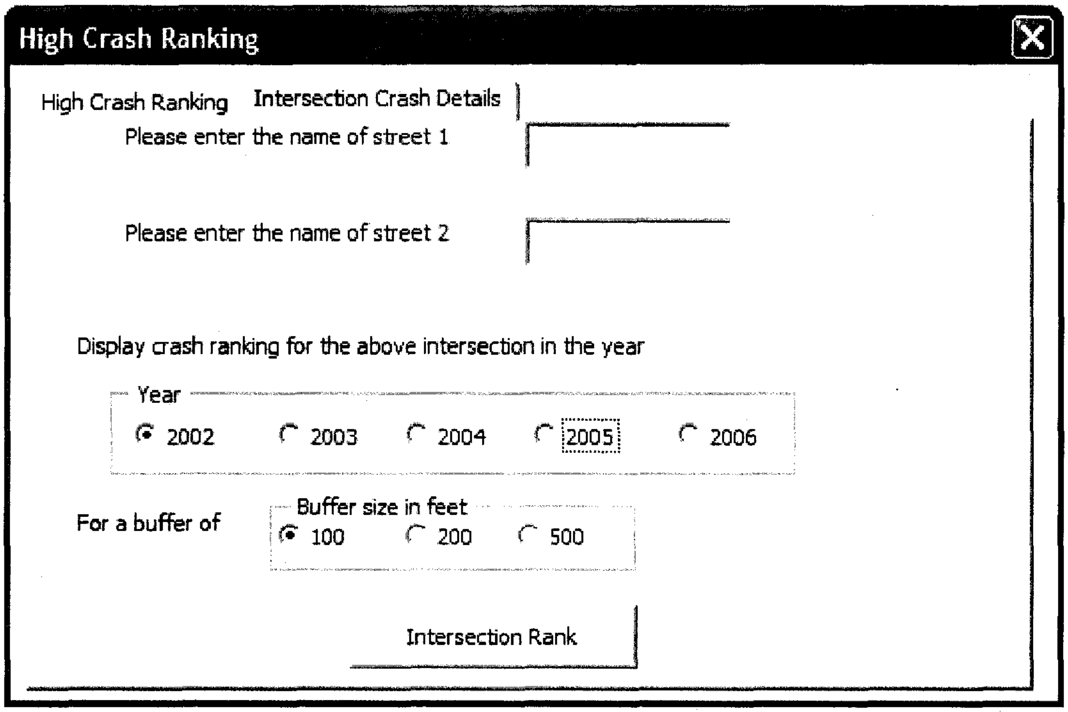


Figure 3.24 Rank for a particular intersection

Also included in this window was an option to view just the rank of a particular location in comparison with the rest of the locations. Figure 3.24 shows the user interface for this feature. This option was designed for the user who wished to see the rank of a particular location and not all the other locations.

CHAPTER 4

CASE STUDY

This chapter provides an illustration of the application of the Safety Analysis System using data from the Las Vegas Metropolitan area. The system's key features and key functions are illustrated with an example selected from the Las Vegas area. The following sub-sections describe in detail the operation of the Safety Analysis System.

The historical crash data used in this thesis are acquired from the Nevada Department of Transportation (NDOT). The data contained the information about crashes that had been recorded by law enforcement agencies in the Las Vegas area from 2002 to 2006. NDOT data provided crash data in the shapefile (.shp) spatial data format to facilitate ease in processing using the software selected for developing the system (ArcGIS software). The advantages of using a shapefile are that they can handle features that overlap and a shape can be defined to them. They also have faster drawing speed and editing ability as they do not have the processing overhead of a topological data structure. They typically require less disk space for storage. The attributes of a shapefile are stored in dBASE format file. The crash database that was obtained from NDOT had over 216,000 records and was of the point shapefile format which was displayed as points in a map window.

Street Centerline database is a mapped collection of all streets in the Las Vegas valley. This can be downloaded from Clark County's GIS Management Office (GISMO).

The GISMO has maintained a record of the streets database since 1993. The database is public and can be downloaded from the GISMO website (http://gisgate.co.clark.nv.us/gismo/gismo.htm). The street centerline database is in the polyline shapefile format, which is displayed as a series of lines on the map window of the ArcGIS software.

The map window is customized and a User Interface (UI) control added on to the toolbar to facilitate access to the Safety Analysis System from an ArcMap window. Upon clicking the button shown in Figure 4.1 which appears at the upper right corner of the top toolbar, the Safety Analysis System would be initialized. The first window was designed to help the user to specify his choice of data inclusion – the entire crash database or pedestrian data only. The following screens depended on the user choice of data. Though most of the operations were similar for both options, the Simple Search tool was made available only for the entire crash data.

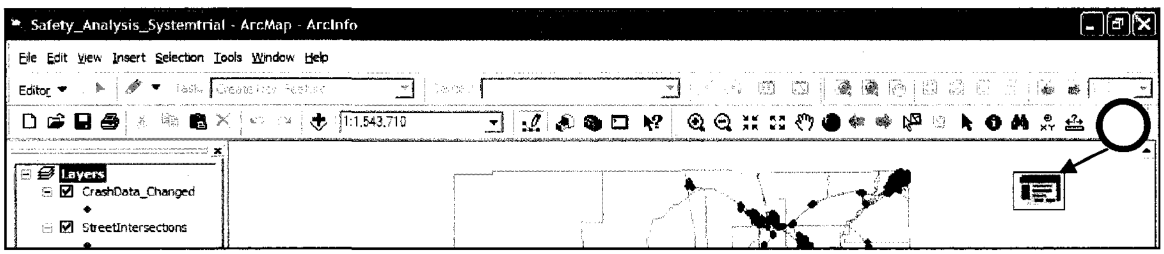


Figure 4.1 The Opening User Interface for the Safety Analysis System

4.1 Simple Search

This option, as explained in the methodology, was intended to provide the user with a basic search option so as to obtain a set of crash records with specified criteria. The user was given the data options to be included in the search functions. The user interface was developed to facilitate choice of criteria by means of combo-boxes programmed to designate all the available data choices within the database.

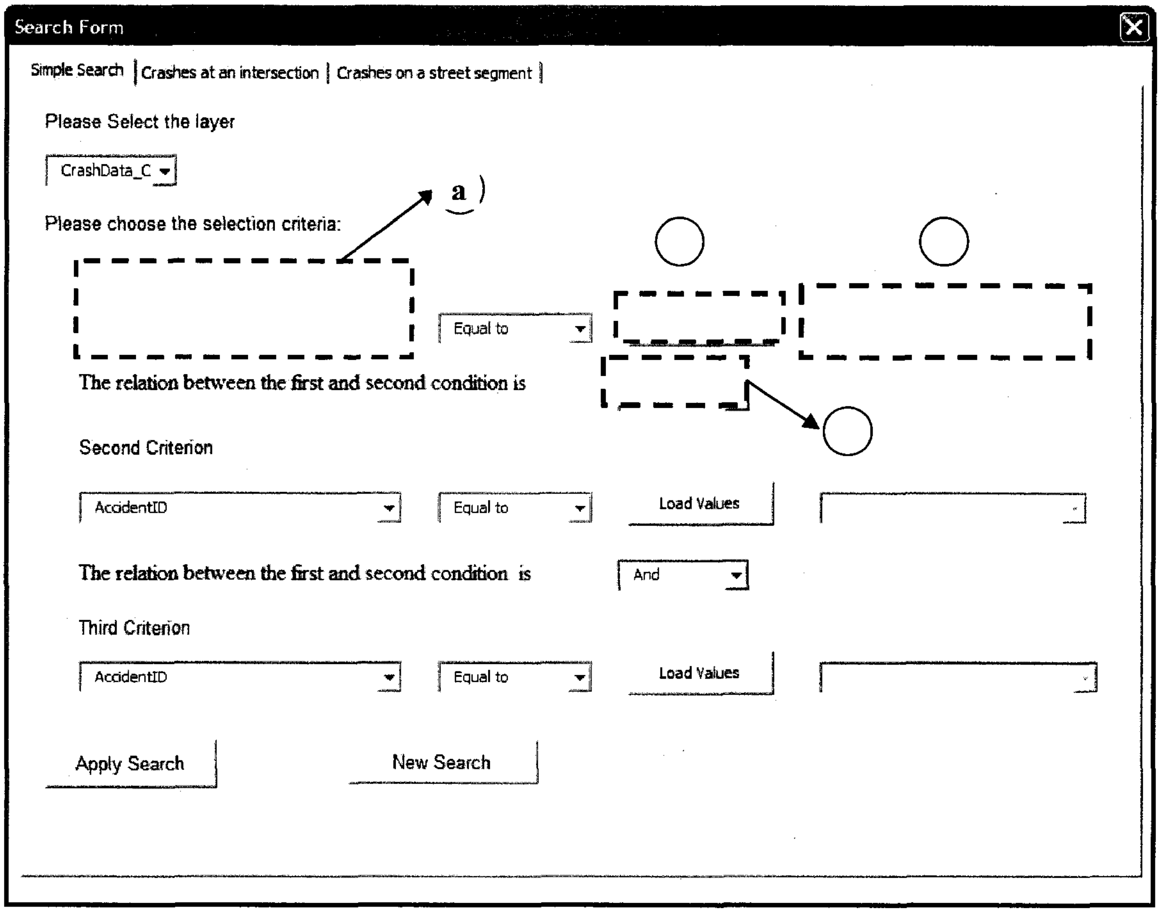


Figure 4.2 Operation of Simple Search tool

The following section describes a sample demonstration of the Simple Search tool with the objective of explaining the purpose and functions of the various items on the user interface and the performance of a search option after data entry. The main components of the Simple Search window were the combo-boxes to display the various data fields using which the user could perform a query. The user had the ability to utilize a maximum of three data fields in the same search query. Figure 4.2 (a) shows the 'First

Criterion' combo-box. Upon clicking the drop down button beside the combo-box, a list of available data fields for the crash data querying were loaded and displayed. The user could thus provide his first criterion. In the same manner, the user could specify up to three criteria to be included in the search query. In order to illustrate this Simple Search option, an example to query the crash database for angle type collisions that resulted in injuries during the year 2005 is explained. The query statement can be broken down into three parts:

- First Criterion: Crash type – Angle, and
- Second Criterion: Severity – Injury, and
- Third Criterion: Year – 2005

The first step for a user to perform this sample query was to click on the 'First Criterion' combo-box as shown in Figure 4.2 (a.) Upon doing so, the list with all possible data fields would be displayed. Next, the user would choose 'Crash Type' from the list displayed. This was to set the first criterion to crash type being angle. After choosing 'Crash Type' as the data field to be queried on, the user had to click on the 'Load unique values' button as shown in Figure 4.2 (b.) This would activate the program to load all the available data entries for the data field 'Crash Type' that could be found by the system in the crash database. The values would be loaded on to the combo-box beside the 'Load unique values' button shown in Figure 4.2 (c.) The next step was for the user to choose 'Angle' from the list that could be viewed on clicking the combo-box's drop down button. Thus, the first criterion is to set to crash type to "angle." Next, the connection between the first and second statement was to be mentioned. The relationship between the first and second criteria could be specified in the box shown in Figure 4.2 (d.) The operators available for use are: ‘And’ and ‘Or’, and they would be displayed on clicking the down arrow button next to the box. For this example, the use of the ‘And’ operator was required. After choosing ‘And’ as the operator, the user could move on to select the second criterion.

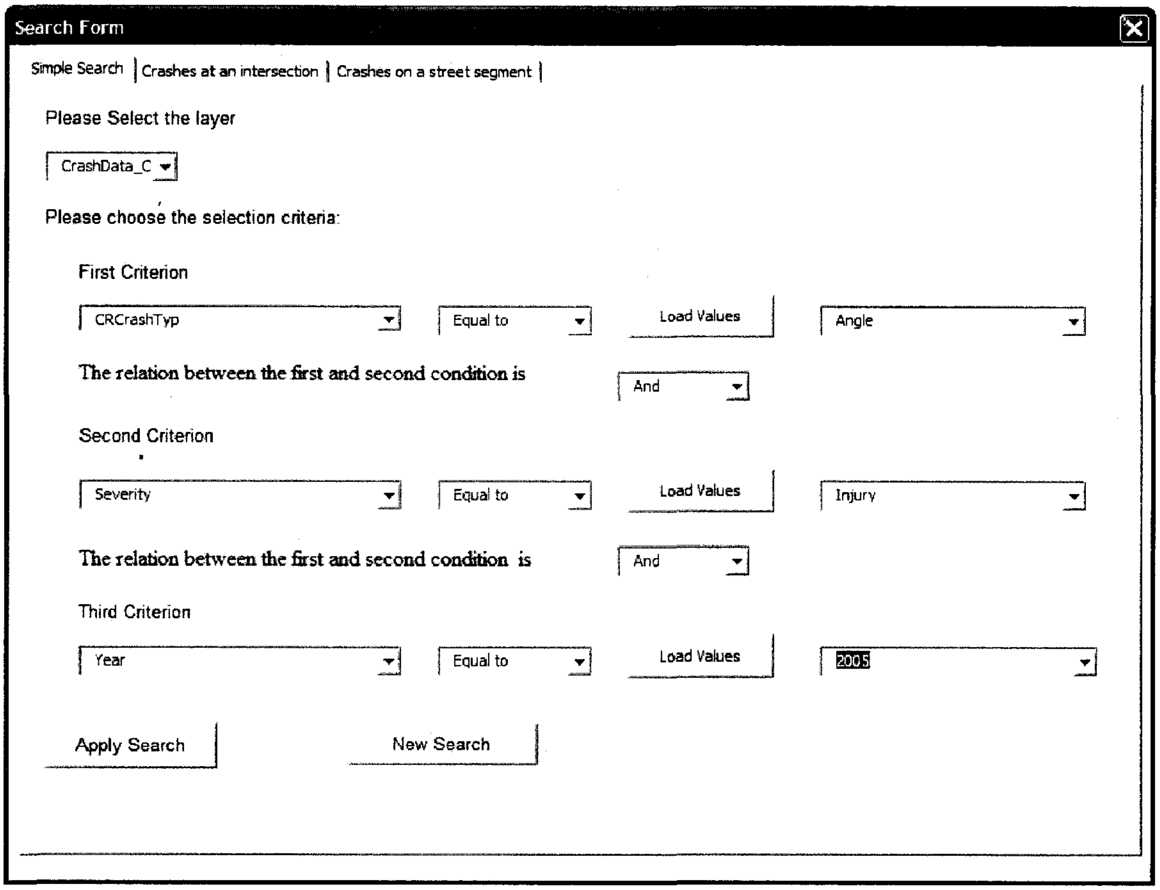


Figure 4.3 Final appearance of sample query

The same procedure of selection of criteria had to be followed for criteria two and three also. For the second criterion, ‘Severity’ was to be chosen from the ‘Second Limitation’ combo-box. After loading the unique values, from the list ‘Injury’ was to be chosen. As mentioned for the first criterion, the connector between the second and third conditions was to be selected next. In this case, the 'And' operator was the choice to program the system to query for all the three criteria. For the final option, the data field was 'Year' and from the loaded values, 2005 was to be selected. This step-by-step procedure provided the user an easy way to mention his choices clearly without any ambiguity. The screen, after data entry would be as shown in Figure 4.3.

After entering the criteria along with the operators, the user needs to click the 'Apply Search' command button in order to initiate the sequence of operations to find the records that matched the query entered by the user. Figure 4.4 shows the results of this sample query on the map along with the option to export the set of records that corresponded to this sample query.

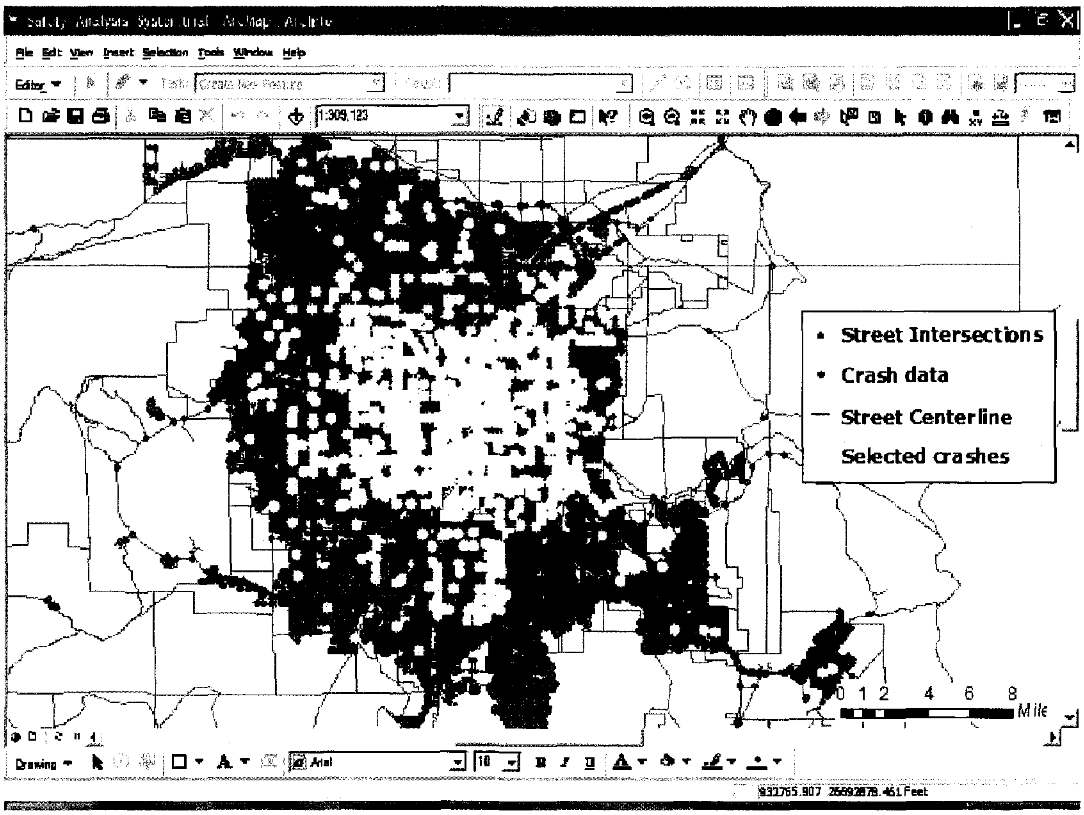


Figure 4.4 Results of the query: Angle crashes that caused injuries in 2005

4.2 Crashes at an Intersection

This section describes the features to identify crashes at an intersection by means of an example from the Las Vegas metropolitan area. In order to explain the operation of the crash analysis tool for an intersection, the following illustration makes use of a sample location, the intersection of Maryland Parkway and Flamingo Road close to University of Nevada Las Vegas. This example aims at finding the crashes that occurred around the Maryland – Flamingo intersection within a model distance of 250 feet. Figure 4.5 shows a map of this location. The step-by-step procedure for carrying out this example is described next.

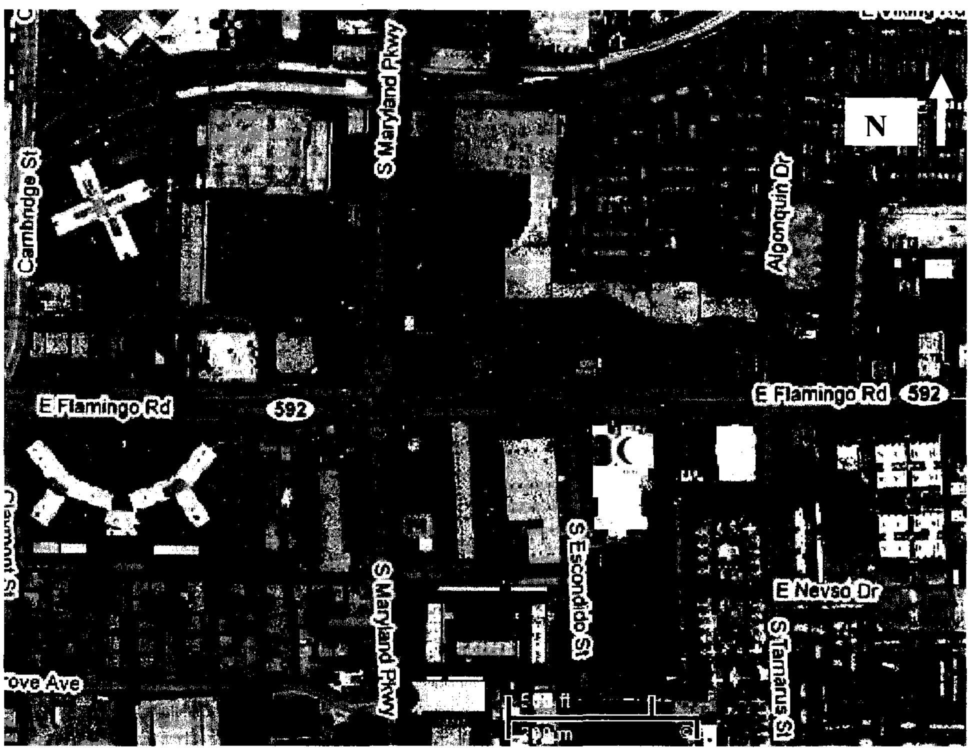


Figure 4.5 Maryland Parkway – Flamingo Road intersection

The first step in this analysis is for the user to choose the names of the intersection streets, i.e., the first and second streets as Maryland Parkway and Flamingo Road or vice versa. Clicking the button shown in Figure 4.6 (a) resulted in the streets names of Las Vegas area from the Street Centerline database being loaded on to the combo-box labeled 'Choose the First Street.' The same was true for the second street combo-box also.

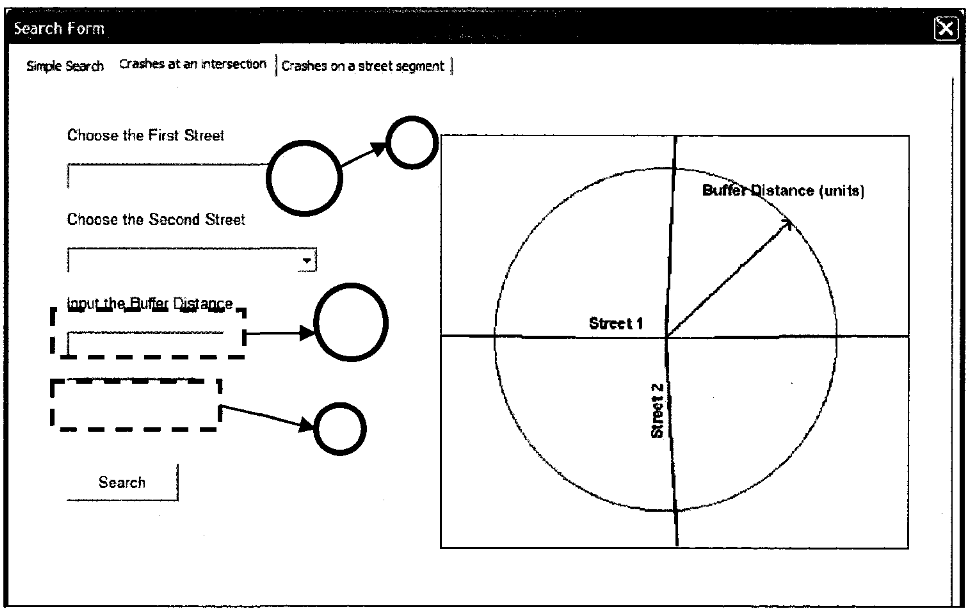


Figure 4.6 Identification of crashes at an intersection

From the list that loaded on to the combo-box, the user had to select 'Maryland' for the name of the first street and 'Flamingo' for the name of the second street. The next step was to enter the required buffer distance into the textbox titled 'Input the Buffer Distance' (Figure 4.6 b) and specify the units for this distance (Figure 4.6 c.) For this example the data entry for the buffer distance was 250 and the unit for this distance was chosen as 'feet.' The screen after entering all the data to perform the sample operation would be as illustrated in Figure 4.7.

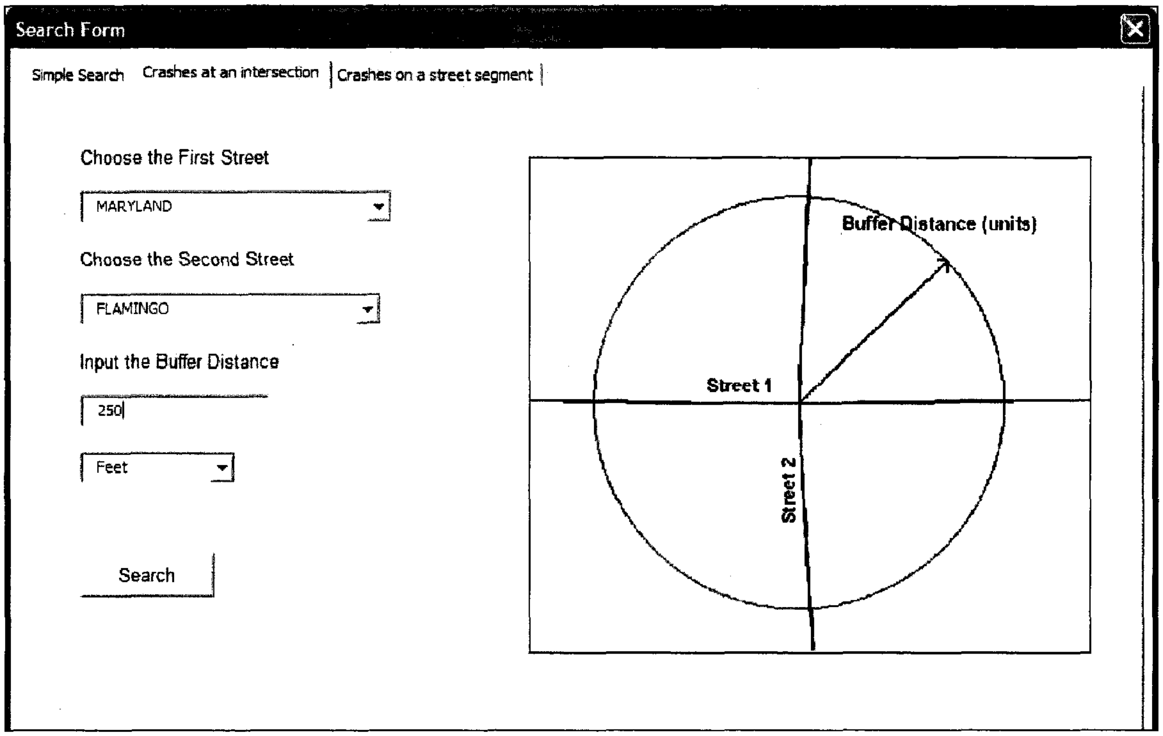


Figure 4.7 Final appearance of intersection crash query

Clicking the 'Search' button would activate the intersection analysis code. The system would check for the existence of the specified intersection location and then performance of the sequence of operations to extract the crashes that occurred around the intersection. The output for the sample query that was run with the data entry as given in the previous figure is shown in Figure 4.8.

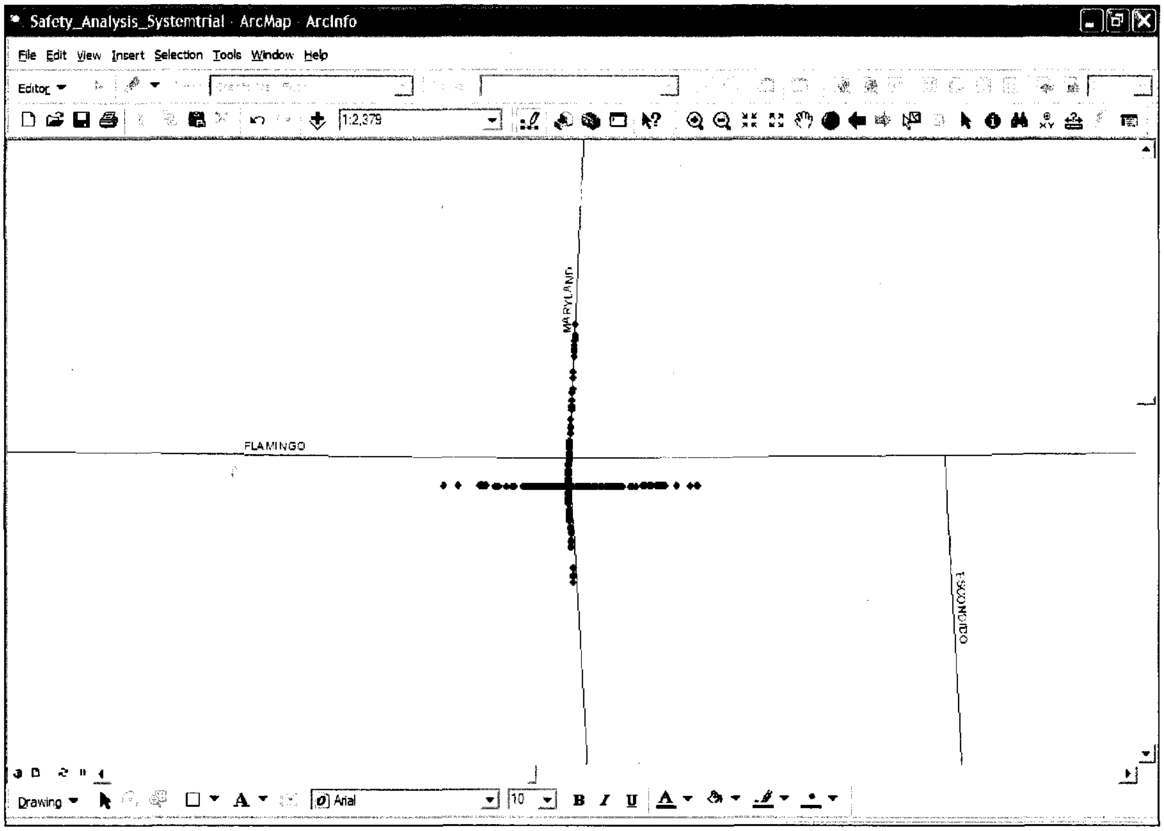


Figure 4.8 Results: Crashes at Maryland Pkwy – Flamingo Rd for a 250 foot radius

Note – There exists a minimal offset of the crash data from the street centerline database owing to the fact that the crashes were geocoded to the street system while creating the database whereas the line represents the centerline of the street and the offset corresponds to the original geocoding process.

4.3 Crashes on a Street Segment

This section describes the process of identification of crashes that occurred on a street segment with the help of an example selected from the Las Vegas area. In order to illustrate the operation extraction of crashes that occurred on a street segment, the segment of Maryland Parkway between Flamingo Road and Tropicana Avenue is chosen and the sequence of operations to carry out this sample query is explained. A map of the sample query location is shown in Figure 4.9. The first step in this operation would be for the user to select the names of the main and cross streets. In Figure 4.10, (a) shows the combo-box item from which the name of the main street needs to be chosen. Upon clicking the drop down button beside the combo-box, the list of all street names would drop down and the name of the main street name, in this case, 'Maryland' would be the name of the main street.

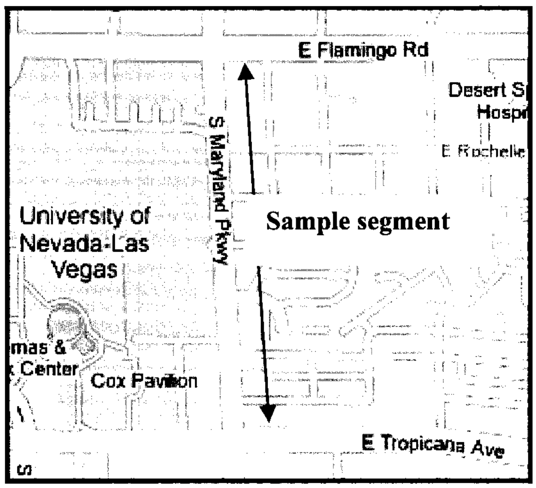


Figure 4.9 Street Segment: Maryland Pkwy between Flamingo Rd and Tropicana Ave

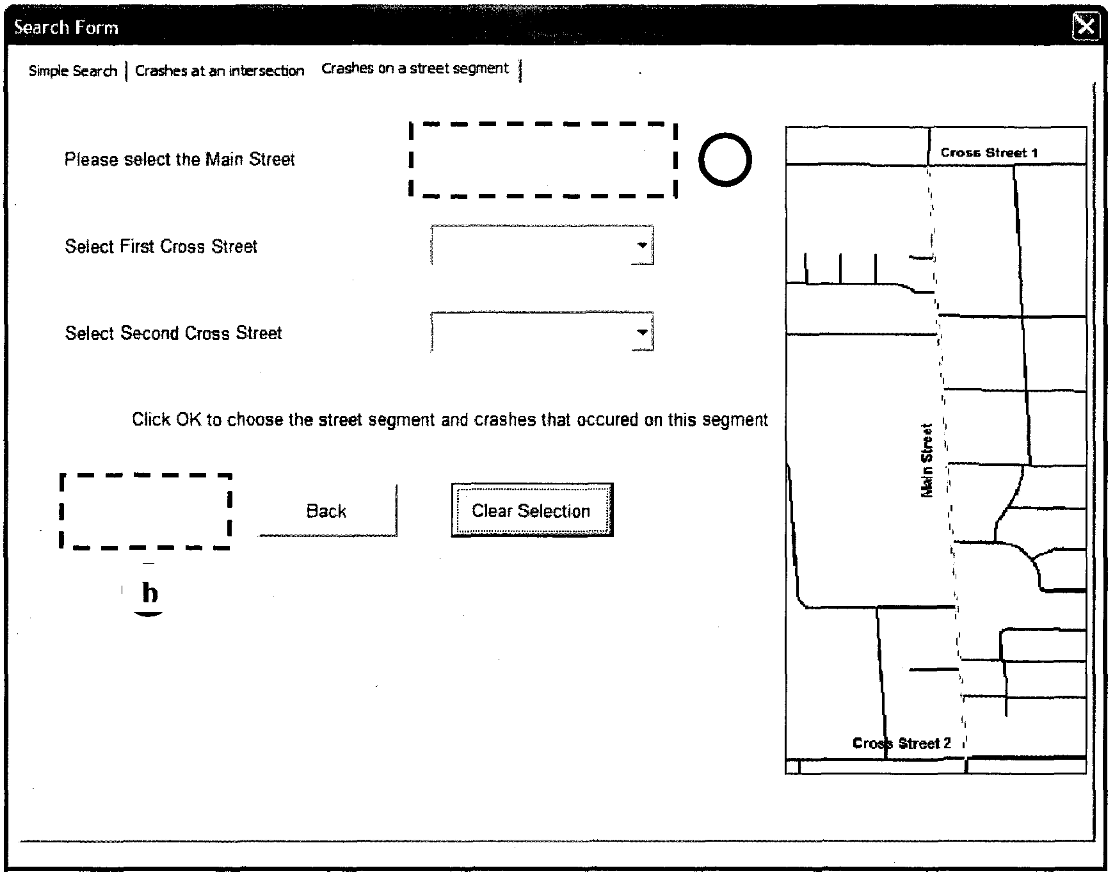


Figure 4.10 Crashes on a street segment tool screen

The next step would be to choose the names of the cross streets from the next two combo-boxes entitled 'Select First Cross Street' and 'Select Second Cross Street' as 'Tropicana' and 'Flamingo' or vice versa. Upon clicking the 'OK' button shown in Figure 4.10 (b) the operation to extract crashes that had occurred on the street segment of Maryland Parkway between Flamingo Road and Tropicana Avenue would be initiated. The final appearance of the screen after the entry of street names is shown in Figure 4.11. The results obtained on the map for the specified street segment are shown in Figure 4.12.

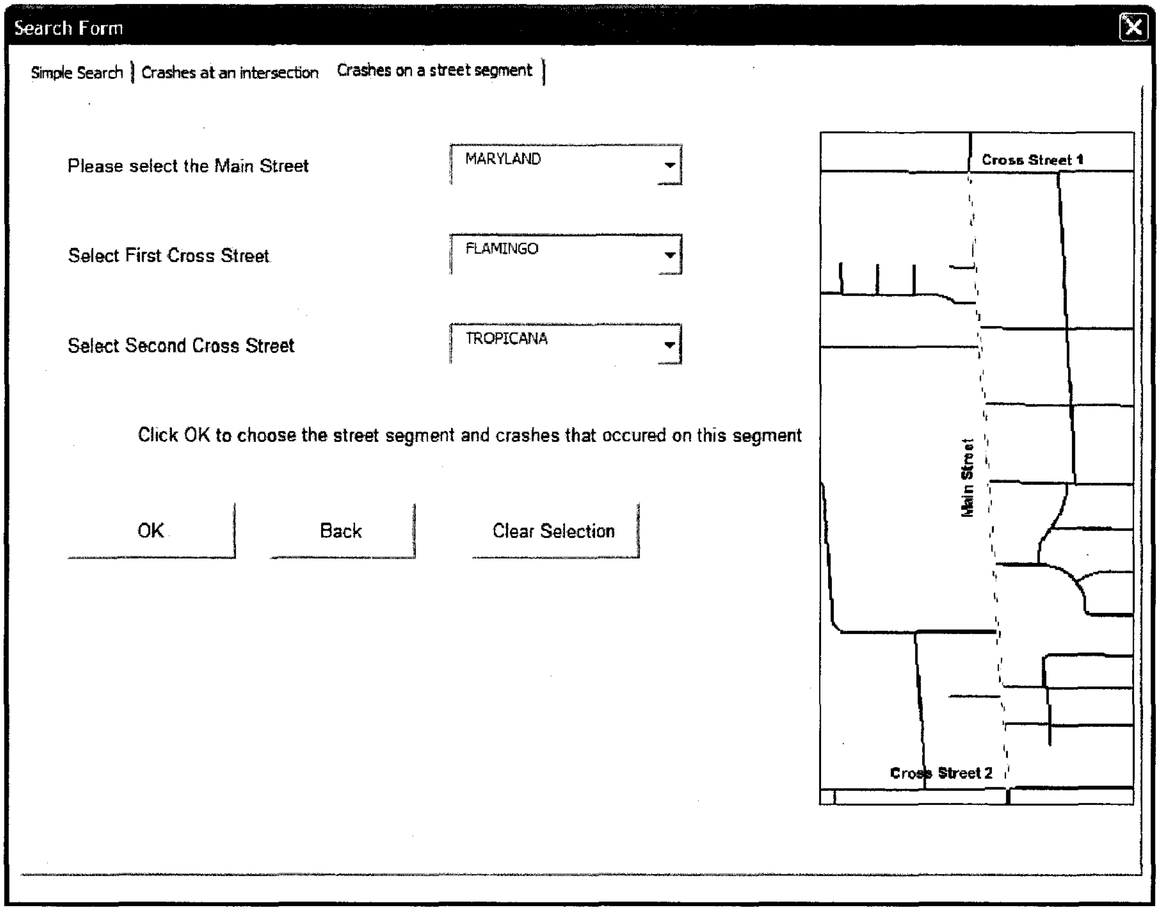


Figure 4.11 Final appearance of screen before running sample street segment query

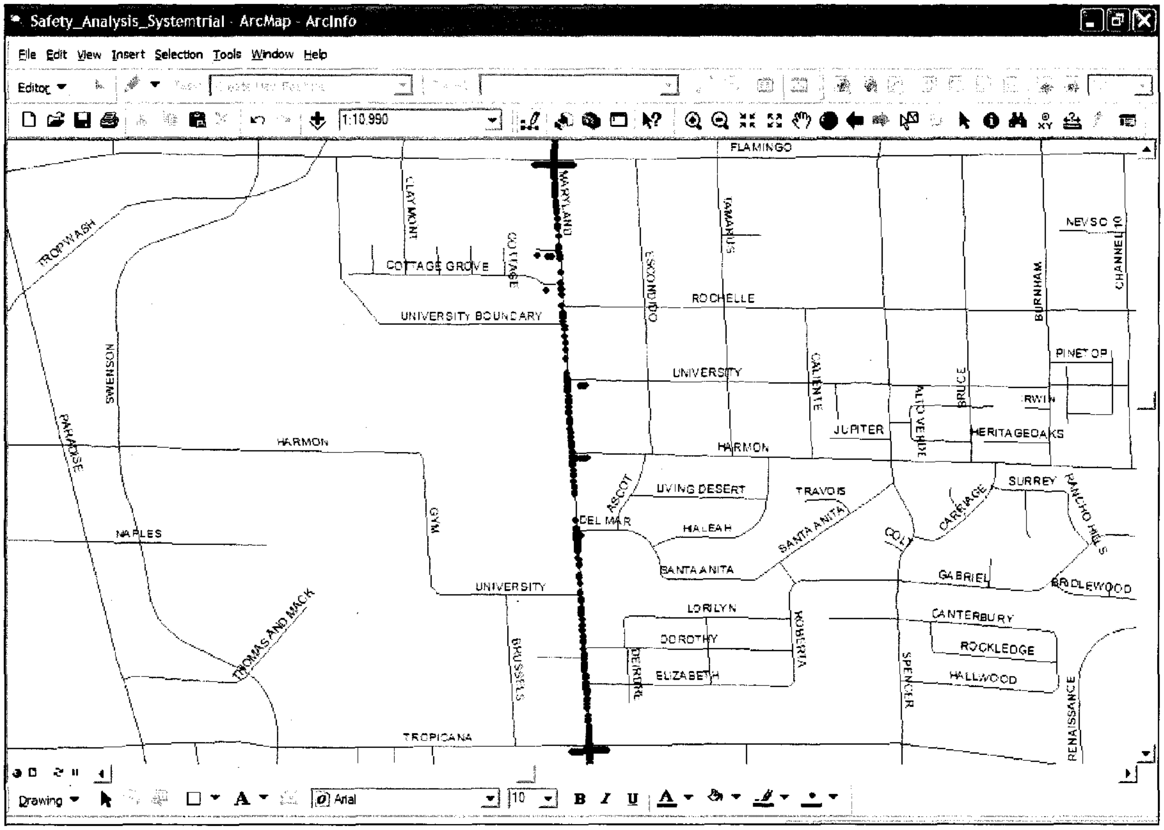


Figure 4.12 Crash Results: Maryland Pkwy between Flamingo Rd and Tropicana Ave

4.4 Multiple Query Criteria

As explained in the methodology section, after either of the two above-mentioned tools – location of crashes at an intersection or on a street segment is performed; the multiple query criteria window pops open to enable the user to include a variety of criteria in his/her query statement. In this section the previously run street segment query, i.e. the segment of Maryland Pkwy between Flamingo Road and Tropicana Avenue and the crashes that occurred on that segment is used as the base to describe the procedure to run a query using multiple criteria. Figure 4.13 shows the appearance of the multiple query criteria screen.

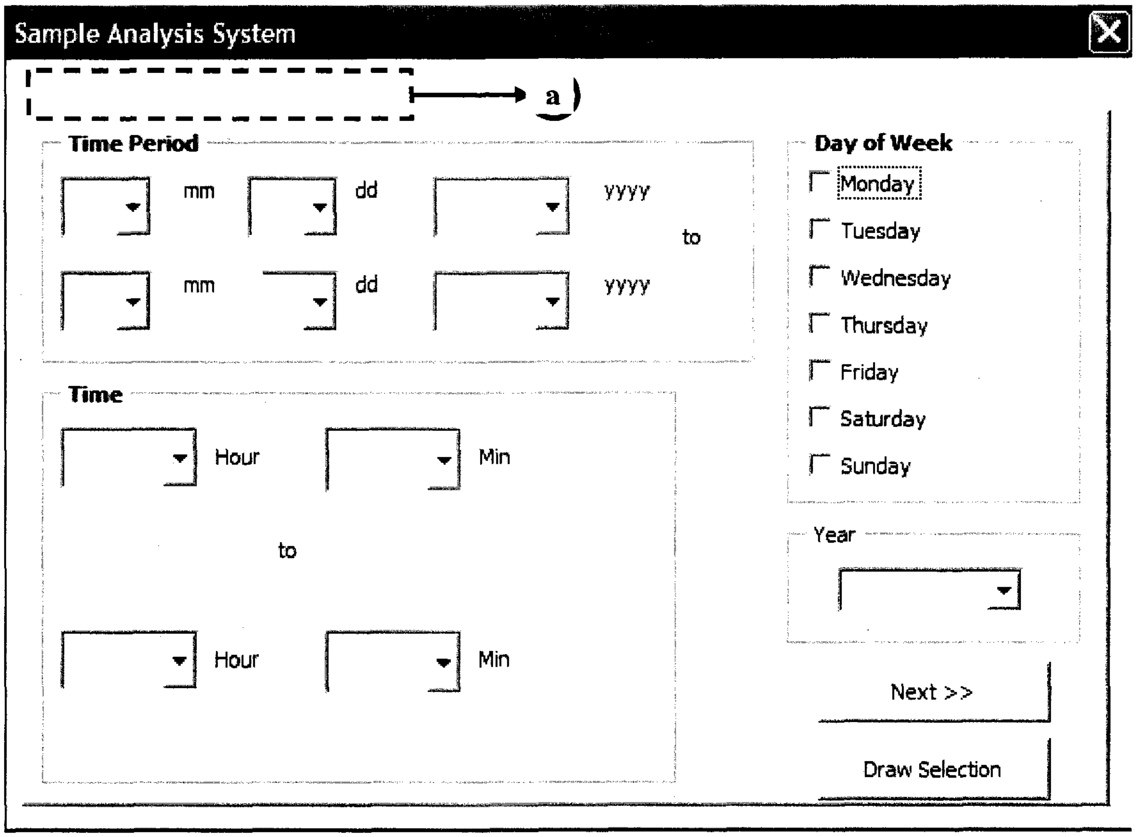


Figure 4.13 Multiple Query Criteria screen

Figure 4.13 (a) shows the various tabs that contain the data frames that are explained in detail in the Methodology chapter. In order to demonstrate the actual process of running a query using this screen, a sample query statement was selected to be run on the crashes that were located on the Maryland Pkwy street segment. For this example, assuming the user wanted to run a query to have the following specific criteria, the procedure to carry out this process is described in the following paragraph.

- Time period: 4, July 2003 to 13, March 2005
- Severity: Fatal or Injury
- Type of Vehicle 1: Sedan
- Weather Condition: Wind or Sandstorm

All these conditions were required to be chosen from the appropriate data frames. Figure 4.14 depicts the screen when the time period criterion was set to the above-mentioned dates.

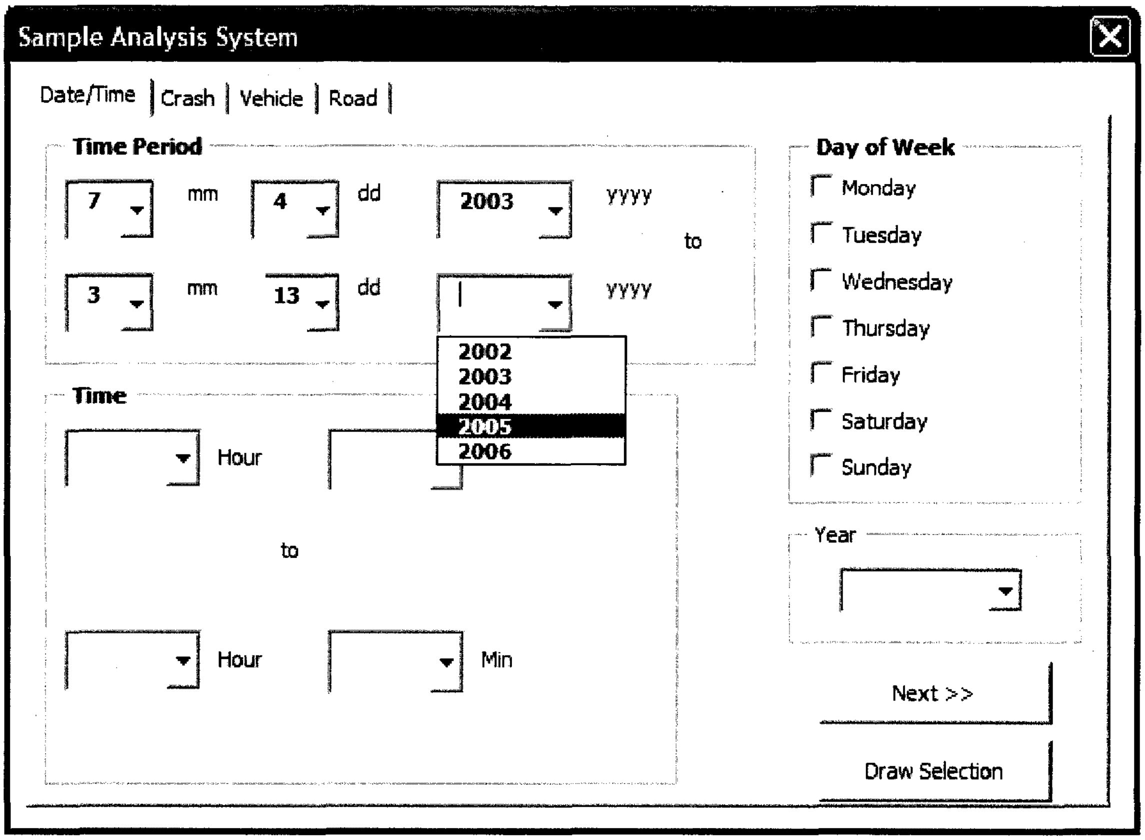


Figure 4.14 Setting desired dates in the Time Period data frame

In order to set the next criterion, the second tab in Figure 4.13 (a) had to be clicked. In this page, under the severity data frame, the severity levels could be set to the desired options. Figure 4.15 shows the severity levels for the sample query been set to fatal or injury.

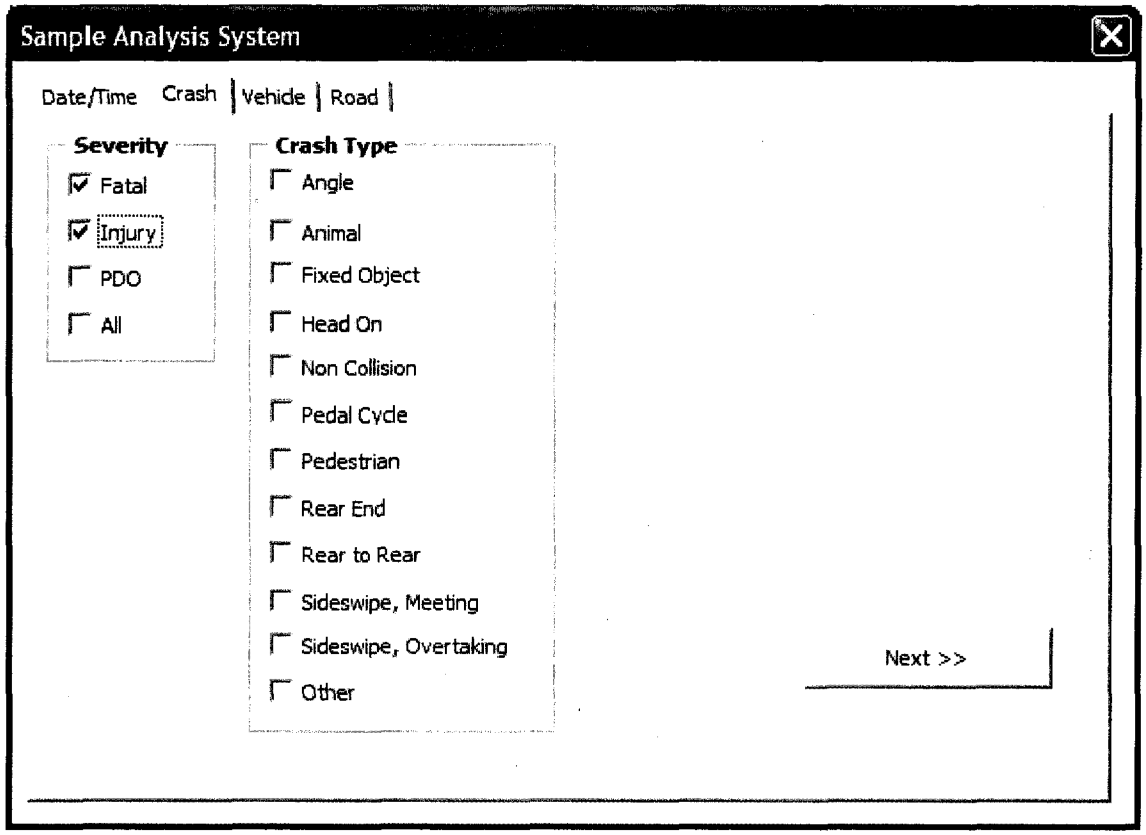


Figure 4.15 Setting severity data field to 'Injury' or 'Fatal'

The third criterion, the type of vehicle involved in the crash required the 'Vehicle 1 Type' data frame option to be set by the user to 'Sedan'. The third tab of the Multiple Query window included the vehicle details and from this, the type of the first vehicle needed to be set to sedan. The screen after selecting the type of first vehicle would look like Figure 4.16.

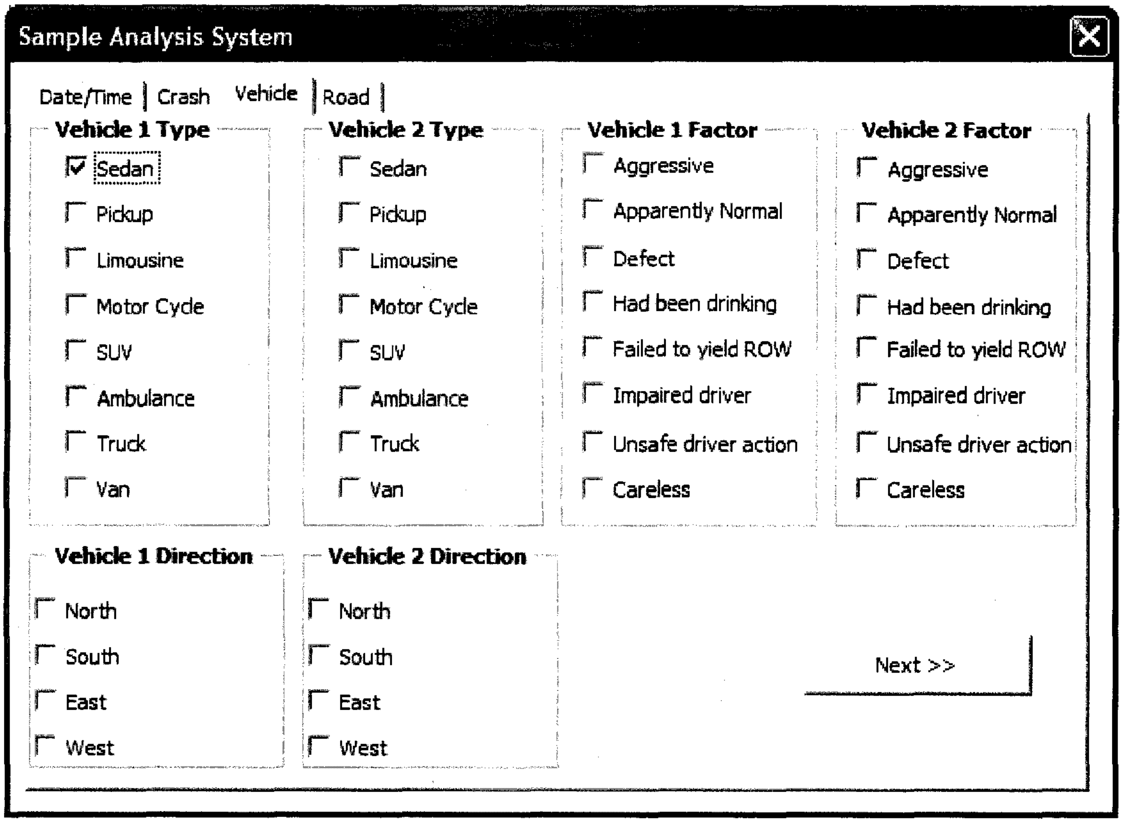


Figure 4.16 Setting Vehicle 1 Type to 'Sedan'

The final option required the Weather Conditions data frame of the 'Road' tab in the Multiple Query window to be set to wind or sandstorm. Figure 4.17 shows the screen after making this selection. After clicking on the 'Draw Selection' command button, the program would search the crashes on Maryland Parkway segment to find matches for the criteria specified by the user. Crashes that matched all of the mentioned factors, if any, would then be selected on the map document and displayed. For the above mentioned criteria, the query was run and the results are displayed in Figure 4.18. It can be seen from the figure that 5 crashes matched the user-specified criteria. The system at this point allows export of features to a shapefile or an Excel file and also allows more analysis.

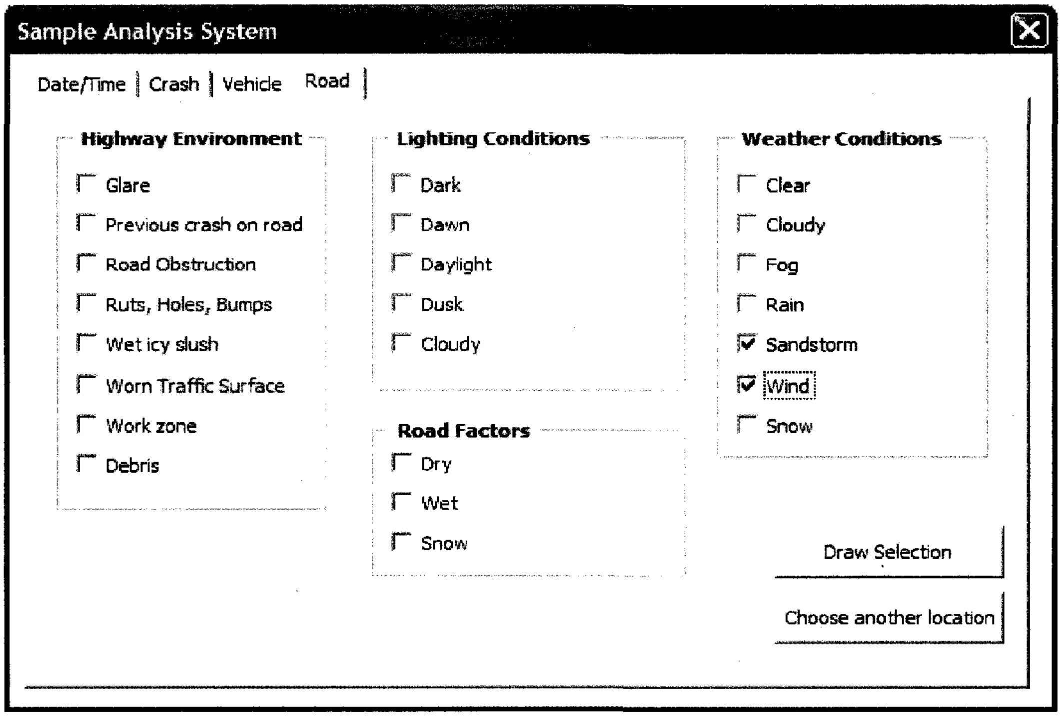


Figure 4.17 Setting Weather Condition to Wind or Sandstorm

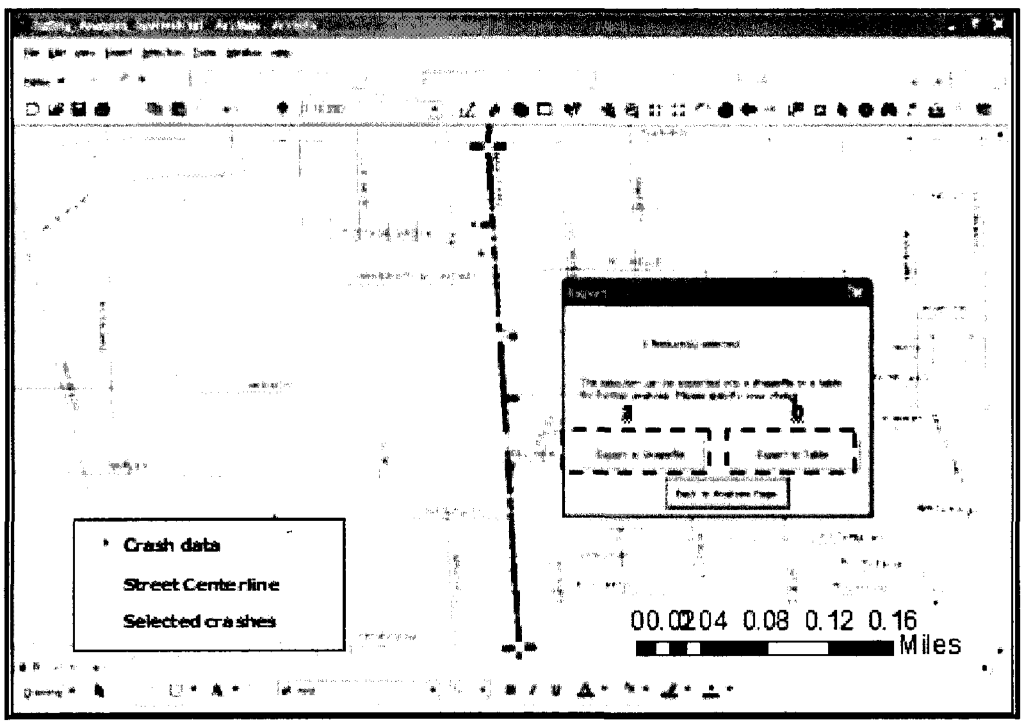


Figure 4.18 Results of Multiple query criteria

4.5 Export Options

This section explains the export feature included in the Safety Analysis system. As can be seen from Figure 4.18, the screen to export the selected features to a shapefile or a spreadsheet document was programmed to show up after a selection had been made using the multiple query window options. From the previous example, five features were found by the system to be matching the user-given criteria. In order to export these results on to a separate empty shapefile for further use, the user was required to click the 'Export to Shapefile' button shown in Figure 4.18 (a). The resulting shapefile would then be created and saved to a pre-specified folder to be accessed later. Figure 4.19 shows the message that would pop-up after copying the features to an empty shapefile. The user can add the shapefile on to the map document by clicking OK.

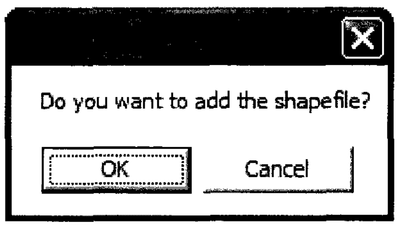


Figure 4.19 Add Shapefile message box

The resulting shapefile after being added on to the map resembled Figure 4.20. Figure Fig 4.18 shows all the crashes that occurred on the segment and highlights the crashes that meet the specified criteria, whereas this Figure 4.20 only shows the latter after the exported shapefile is added onto the map.

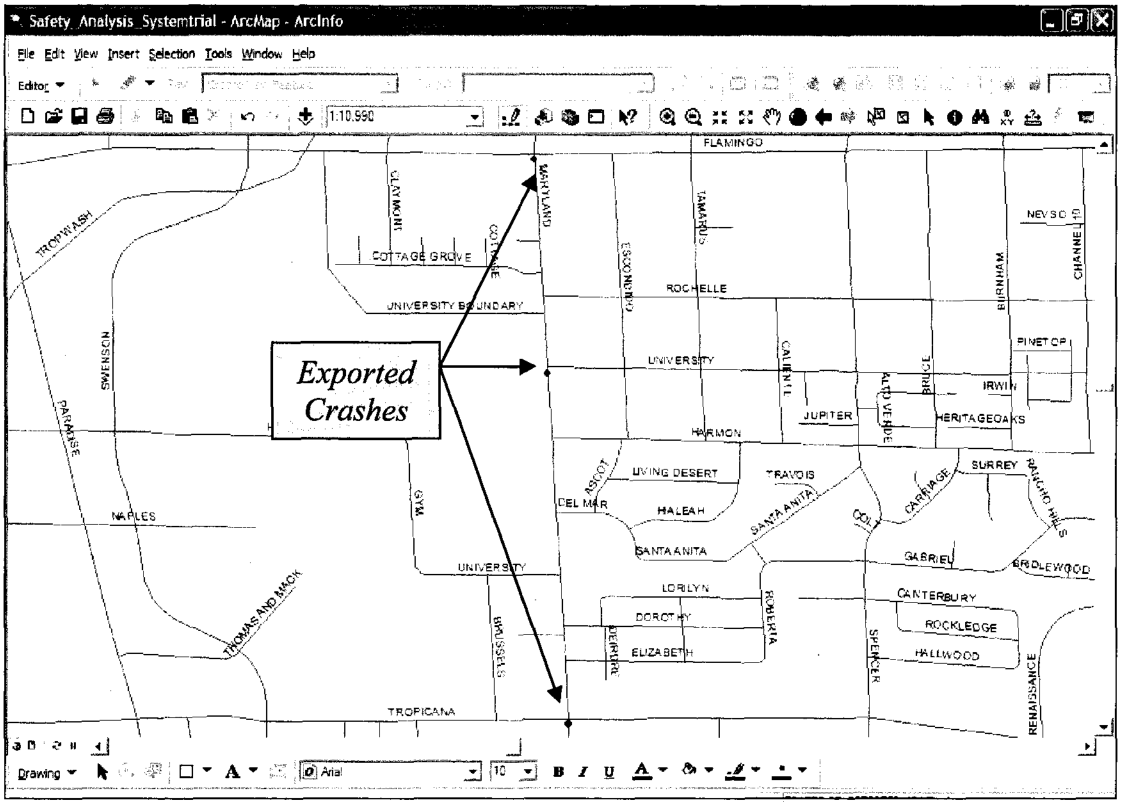


Figure 4.20 Results of exported shapefile

If the user wanted to create a spreadsheet with results obtained from the multiple criteria feature, the button shown in Figure 4.18 (b) would have to be made use of. Upon clicking this button, the system exported the selected results to an Excel file saved to a location that was preset during the programming.

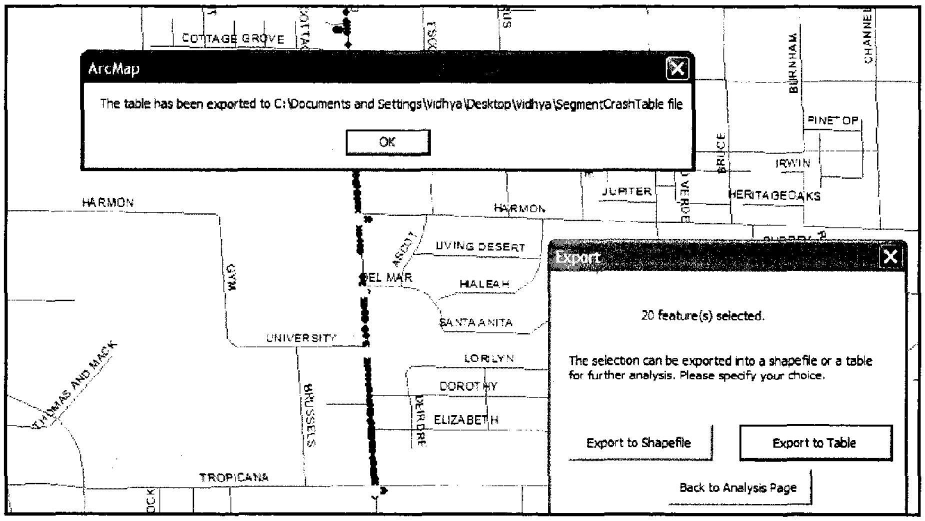


Figure 4.21 Message - Export file path

The resulting spreadsheet document that is exported is shown in Figure 4.22

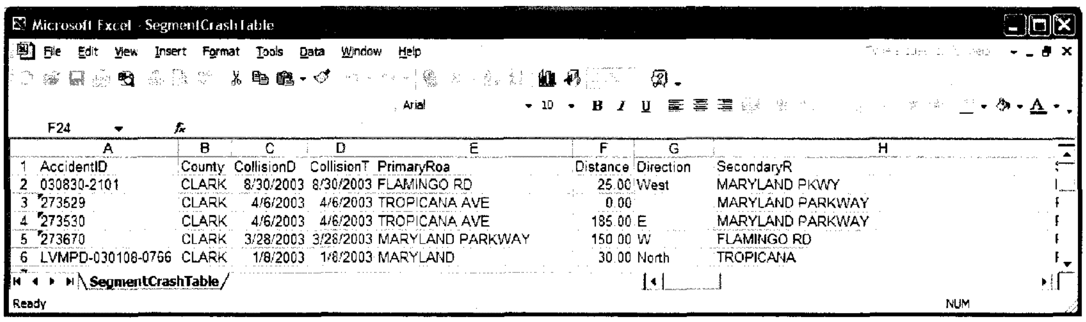


Figure 4.22 Query results exported to Microsoft Excel document

4.6 Ranking of High Crash Locations

In this section, the High Crash Ranking tool is described for a sample run in order to explain the step by step procedure to be followed by the user to get required results.

4.6.1 Ranking based on Total Crashes

Assuming that the user wished to rank the locations in Las Vegas area based on the total number if crashes that occurred at each location in the year 2004 for a buffer distance of 200 feet around the intersection, the High Crash Ranking feature shown in Figure 4.23 (a) had to be used.

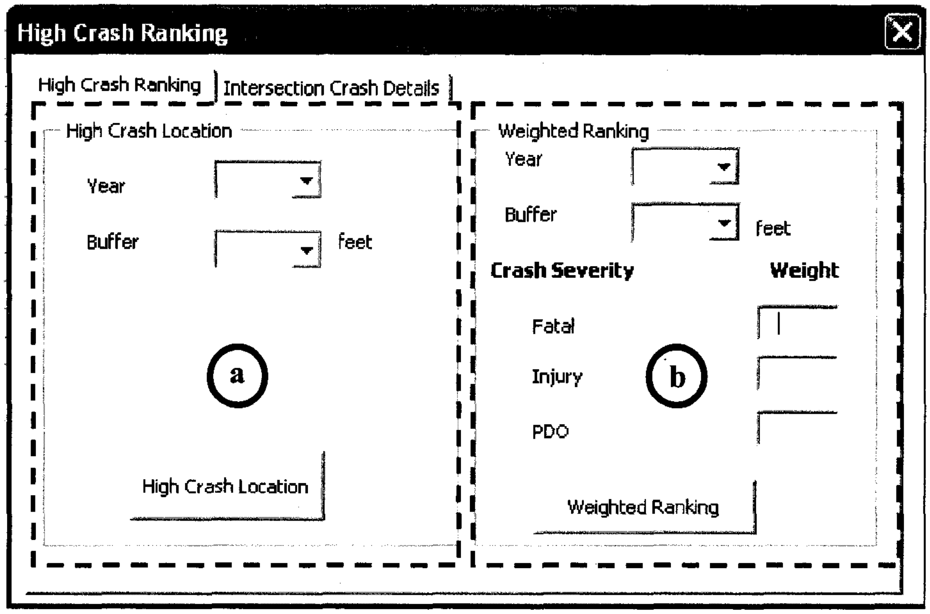


Figure 4.23 High Crash Ranking screen

In order to perform the desired function, the option '2004' had to be chosen from the 'Year' combo-box shown in the figure. Also, the buffer distance of 200 feet had to be selected from the pre-loaded values (100, 200 and 500 feet) in the combo-box labeled 'Buffer'. The data entry is shown in Figure 4.24.

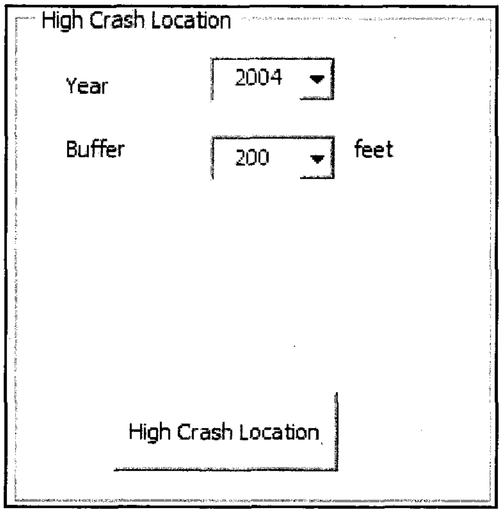
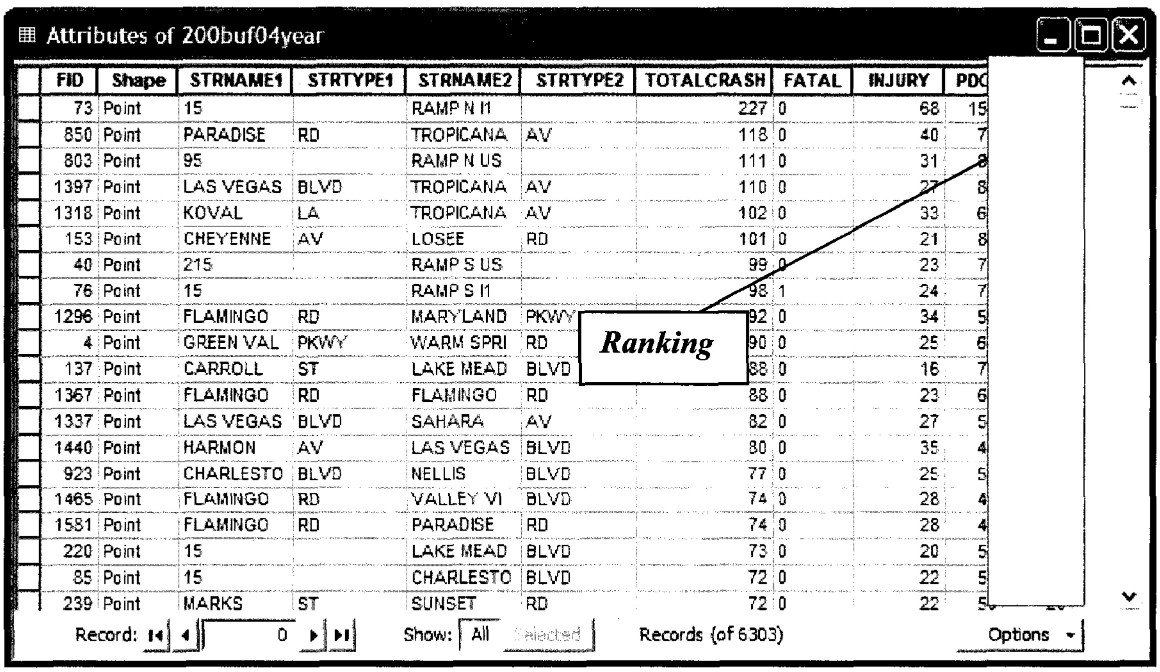


Figure 4.24 Data Entry for High Crash Ranking

After clicking the High Crash Location button, the ranking would be performed by the system. The resulting ranking obtained for the year 2004 and buffer size 200 feet is shown in Figure 4.25.



Attributes of 200buf04year

FID	Shape	STRNAME1	STRTYPE1	STRNAME2	STRTYPE2	TOTALCRASH	FATAL	INJURY	PDC
73	Point	15		RAMP N I1		227	0	68	15
850	Point	PARADISE	RD	TROPICANA	AV	118	0	40	7
803	Point	95		RAMP N US		111	0	31	8
1397	Point	LAS VEGAS	BLVD	TROPICANA	AV	110	0	27	8
1318	Point	KOVAL	LA	TROPICANA	AV	102	0	33	6
153	Point	CHEYENNE	AV	LOSEE	RD	101	0	21	8
40	Point	215		RAMP S US		99	0	23	7
76	Point	15		RAMP S I1		98	1	24	7
1296	Point	FLAMINGO	RD	MARYLAND	PKWY	92	0	34	5
4	Point	GREEN VAL	PKWY	WARM SPRI	RD	90	0	25	6
137	Point	CARROLL	ST	LAKE MEAD	BLVD	88	0	16	7
1367	Point	FLAMINGO	RD	FLAMINGO	RD	88	0	23	6
1337	Point	LAS VEGAS	BLVD	SAHARA	AV	82	0	27	5
1440	Point	HARMON	AV	LAS VEGAS	BLVD	80	0	35	4
923	Point	CHARLESTO	BLVD	NELLIS	BLVD	77	0	25	5
1465	Point	FLAMINGO	RD	VALLEY VI	BLVD	74	0	28	4
1581	Point	FLAMINGO	RD	PARADISE	RD	74	0	28	4
220	Point	15		LAKE MEAD	BLVD	73	0	20	5
85	Point	15		CHARLESTO	BLVD	72	0	22	5
239	Point	MARKS	ST	SUNSET	RD	72	0	22	5

Figure 4.25 High Crash Ranking for the year 2004 and buffer size 200 feet

4.6.2 Weighted Ranking

If the user wishes to make use of the option to include weights while ranking the locations, the feature shown in Figure 4.23 (b) had to be used. In order illustrate this feature, assuming the user wished to append weights for severity to the previously mentioned query year and buffer distance, fatal being 10, injury being 5 and PDO being 2, the procedure for running the option would be thus. The user had to first choose the year and buffer distance as 2004 and 200 feet respectively as described in the previous sub-section. Next, the user had to set the severity levels to the required values as shown in Figure 4.26.

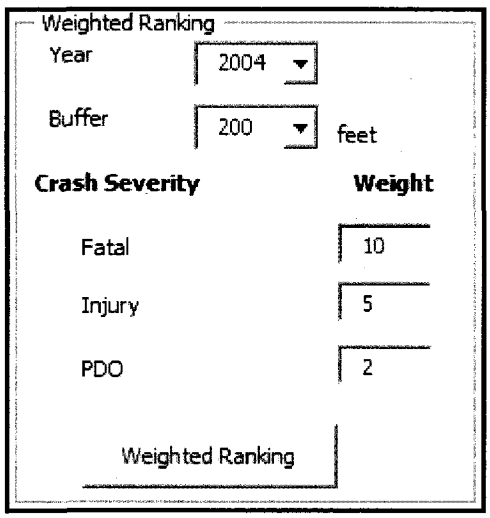


Figure 4.26 Data Entry for Weighted High Crash Ranking

Clicking the 'Weighted Ranking' button triggered the system to calculate the sum of weights according to the procedure described in section 3.4.6. The results obtained for the specified weights of severity levels having been incorporated in the ranking procedure, are shown in Figure 4.27

Attributes of 200buf04year

FID	Shape	STRNAME1	STRTYPE1	STRNAME2	STRTYPE2	FATAL	INJURY	PDO	WeightSum
73	Point	15		RAMP N I1		0	68	159	658
850	Point	PARADISE	RD	TROPICANA	AV	0	40	77	354
803	Point	95		RAMP N US		0	31	80	315
1318	Point	KOVAL	LA	TROPICANA	AV	0	33	68	301
1397	Point	LAS VEGAS	BLVD	TROPICANA	AV	0	27	82	299
1296	Point	FLAMINGO	RD	MARYLAND	PKWY	0	34	58	286
76	Point	15		RAMP S I1		1	24	72	274
40	Point	215		RAMP S US		0	[illegible]	76	267
153	Point	CHEYENNE	AV	LOSEE	RD	0	[illegible]	80	265
1440	Point	HARMON	AV	LAS VEGAS	BLVD	0	[illegible]	84	263
4	Point	GREEN VAL	PKWY	WARM SPRI	RD	0	25	65	255
1337	Point	LAS VEGAS	BLVD	SAHARA	AV	0	27	55	245
1367	Point	FLAMINGO	RD	FLAMINGO	RD	0	23	65	245
1320	Point	RAINBOW	BLVD	SAHARA	AV	0	36	30	240
1329	Point	DECATUR	BLVD	FLAMINGO	RD	1	28	41	232
1465	Point	FLAMINGO	RD	VALLEY VI	BLVD	0	28	46	232
1581	Point	FLAMINGO	RD	PARADISE	RD	0	28	46	232
923	Point	CHARLESTO	BLVD	NELLIS	BLVD	0	25	52	229
1292	Point	CHARLESTO	BLVD	LAMB	BLVD	1	27	41	227
137	Point	CARROLL	ST	LAKE MEAD	BLVD	0	16	72	224

Ranking

Record: 0 Show: All Selected Records (of 6303) Options

Figure 4.27 Ranking Results: Weights for Severity – 10, 5 and 2

CHAPTER 5

CONCLUSIONS AND RECOMMENDATIONS

The main objective of this research was to develop a tool to perform safety analyses on a GIS platform by means of user-friendly interfaces to analyze data and display the results in graphical/tabular format. Such a system could be used to identify safety issues in an area. It could also be used to monitor traffic safety programs and their effectiveness. The ArcGIS software was used as the basic software to develop the tool. The in-built VBA application in ArcGIS was used to develop the user interface. Then codes were composed to design the various functionalities of the tool within the framework of the interface. The Safety Analysis System thus developed was able to perform crash analysis at a specific location in a detailed manner as well as rank locations based on number of crashes. The uses of the Safety Analysis System are demonstrated using the crash data for the Las Vegas Metropolitan area obtained from the Nevada Department of Transportation. The data used for demonstration has crash records from the year 2002 to 2006. Also, another tool with similar functionalities was developed solely analyze the pedestrian crash data from the year 2003 to 2006. Each application is demonstrated for the Las Vegas area.

5.1 Conclusions

The main elements of this research are summarized next.

- The Safety Analysis System helps provide decision support for the various transportation agencies and officials for allocation of resources such as roadway characteristics and traffic control to those locations that require changes to existing controls in order to reduce the number of crashes and thereby reduce the impacts of those crashes.
- Using the ArcGIS software and its integral programming interface, the VBA application was used to develop a user-friendly transportation safety analysis system. The Safety Analysis System incorporates numerous tools to perform specific functions keeping the ease of use as a key requirement.
- The Safety Analysis System included a search tool that allows the user to query the crash data based on any of the fields contained in the original database. The user can query crashes that occurred at a specific location identified as an intersection or a street segment. Identification of crashes on a street segment is a unique feature that can be used to analyze mid-block locations. After identifying a specific location, the user can consider crashes that occurred at that location for additional queries. The multiple query criteria window allows the user to incorporate numerous factors into the query statement thus helping to narrow the results to those required by the user.
- The High Crash Ranking tool can be used to assign ranks to locations based on the total number of crashes that occurred around it. In addition, weights could also

be assigned for the various levels of severity to be included in the ranking of the locations.

- The results can be viewed as maps or tables. The system also allows two forms of export. Results can be exported to a shapefile to be incorporated into ArcGIS or as a Microsoft Excel spreadsheet for further analysis.
- The advantage of the Safety Analysis System is its ability to integrate various formats of data and obtain results based on queries. The display of the results as maps helps enhance understanding of the data.
- The Safety Analysis System can be used to perform before and after studies to determine the effectiveness of a transportation safety program. Also, the user interface helps users who are not familiar with GIS functionalities to make use of the developed analysis tools.

5.2 Recommendations for Future Work

Some of the applications of the Safety Analysis System are that it can be used to obtain a map of any location (intersection or street segment) with results for any query criteria that the user might require. Also, it can be used to evaluate the effectiveness of a traffic safety program by means of before-after studies to monitor the reduction of crashes brought about by the program. This would enable transportation agencies like the Regional Transportation Commission of Southern Nevada to make decisions regarding allocation of resources. Based on the study conducted fort his thesis, the following recommendations are made to improve the Safety Analysis System in the future.

- The tool currently uses crash data obtained from the Nevada Department of Transportation for the time period from years 2002 to 2006. More recent data if used in place of the current data would help improve the results in terms of traffic safety programs introduced after 2006.
- From the crash database, it was found that the records contained numerous misspelled entries. Since the querying of data depended entirely on the individual entries in the database, addressing all possible spellings of a particular factor could not be achieved entirely. Creation of a reliable crash database would help to improve the accuracy of the results obtained by using the various analysis tools.
- Use of Information Technology by means of handheld/PDA type devices while recording crash reports would help in minimizing data inconsistencies and misspells in addition to the time saved in manual inclusion of each report into the crash database.
- Incorporation of other traffic-related information such as traffic volumes/mix, traffic control and operational information, roadway-related information in evaluating traffic safety would help in improving the analyses possible. For example, computation of crash rates such as the crashes per million entering vehicles at an intersection would provide a measure for ranking crashes.

REFERENCES

Affum, Joseph K. and Taylor, Michael A. P. August 1997. SELATM -- A GIS Based Program for Evaluating the Safety Benefits of Local Area Traffic Management Schemes, *Transportation Planning and Technology* 21(1 & 2), 93-119.

Austin, K., Tight, M. and Kirby H. March 1997. The Use of Geographical Information Systems to Enhance Road Safety Analysis. *Transportation Planning and Technology* 20(3), 249-266.

Bapna, Sanjay and Gangopadhyay, Arrya. July 2005. A web-based GIS for analyzing commercial motor vehicle crashes. *Information Resources Management Journal,* 18(3), 1-12.

Clark County GIS Management Office (GISMO), URL - http://gisgate.co.clark.nv.us/gismo/gismo.htm

Dangeti, Mukund R., Pulugurtha, Srinivas S. and Nambisan, Shashi S. 2006. Development of spatial data tools to manage transportation networks. Applications of Advanced Technology in Transportation - Proceedings of the Ninth International Conference on Applications of Advanced Technology in Transportation, August 2006, Chicago, Illinois, 828-833.

ESRI Support Center, ArcInfo Desktop 9.2 System Requirements, URL - http://support.esri.com/index.cfm?fa=knowledgebase.systemRequirements.list&PN=ArcInfo+Desktop&pName=ArcInfo+Desktop&PVName=&PID=43&count=1&pvid=352&VID=945

Kim, K., Kerns, T., Hettinger, T., and Pease, M. March 2001. National Highway Traffic Safety Administration Report Number DOT HS 809 201, National Center for Statistics and Analysis.

Kim, Karl and Levine, Ned. July 1996. Using GIS to improve Highway Safety. *Computer, Environment and Urban Systems,* 20(4), 289-302.

Krishnakumar, Vanjeeswaran K., Pulugurtha, Srinivas S. and Nambisan, Shashi S. 2005. Identification and Ranking of High Pedestrian Crash Zones Using GIS. 2005 Annual ESRI International User Conference, San Diego, California.

Fatality Analysis Reporting System (FARS) Research and Development, URL - http://www-nrd.nhtsa.dot.gov/departments/nrd-01/summaries/FARS_98.html

Iowa DOT High Crash Locations Ranking Procedure, Systematic Identification of High Crash Locations (HCL), Iowa State University, URL - http://www.ctre.iastate.edu/research/hcl/documents/ranking.htm

Li, Linhua and Zhang, Yunlong. 2007. A GIS-Based Bayesian Approach for Identifying Hazardous Roadway Segments for Traffic Crashes. *TRB 2007 Annual Meeting Compendium of papers (CD-ROM), 07-2441, Transportation Research Board, Washington DC.*

Local Agency Safety Management System, *NDOT Safety Management System Reference Documents, Washington State Department of Transportation, 1998*

Microsoft Office Online, Excel 2002 System Requirements, URL - http://office.microsoft.com/en-us/excel/HA102132381033.aspx

Miller, John S. 2000. Geographical Information Systems: Unique Analytic Capabilities for the Traffic Safety Community. *Transportation Research Record* 1734, 21-28.

Model Minimum Uniform Crash Criteria (MMUCC), URL - http://www.mmucc.us/

Nevada DOT - Safety Management System, *Safety Management Homepage, Nevada Department of Transportation* http://www.ndotsms.com/index.htm

Parrish, Allen S., Dixon, B., Cordes, D., Vrbsky, S. and Brown, D. June 2003. CARE: An Automobile Crash Data Analysis Tool. *IEEE Journal,* 36(6), 22-30.

Pawlovich, Michael D. and Souleyrette, Reginald R. May 1996. A GIS-based Accident Location and Analysis System (GIS-ALAS). *Semisequicentennial Transportation Conference Proceedings*, Iowa State University, Ames, Iowa.

Peden, M., Scurfield, R., Sleet, D., Mohan, D., Hyder, A. A., Jarawan, E., Mathers, C. 2004. World Report on Road Traffic Injury Prevention. World Health Organization, Geneva, Switzerland.

Pulugurtha, Srinivas S., Vanjeeswaran, Krishnakumar K. and Nambisan, Shashi S. 2007. New methods to identify and rank high pedestrian crash zones: An illustration. *Accident Analysis and Prevention* 39 (2007) 800–811.

Pulugurtha, Srinivas S., Vasudevan, Vinod, Nambisan, Shashi S. and Wongchavalidkul, Natachai. 2004. Identification of high pedestrian crash sites and selection of countermeasures for the Las Vegas metropolitan area. Proceedings of the Eighth International Conference on Applications of Advanced Technologies in Transportation Engineering, 2004, Beijing, China. 115-119.

Roche, Jerry. 2000. Geographic Information Systems-Based Crash Data Analysis and the Benefits to Traffic Safety. *2000 MTC Transportation Scholars Conference, Ames, Iowa,* 110-122.

Saccomanno, F.F., Chong, K.C. and Nassar, S.A. 1997. Geographic Information System Platform for Road Accident Risk Modeling, *Transportation Research Record* 1581, 18-26.

Sando, Thobias, Mussa, Renatus, Wu, Haitao, Sobanjo, John and Spainhour, Lisa. 2004. A Cost-Effective GIS Safety Analysis Tool for Improving Highway Safety. Proceedings of the Twenty-Fourth Annual ESRI User Conference, 2004, San Diego, California.

Smith, Randy K., Graettinger, Andrew J., Keith, Kerri and Parrish, Allen. January 2007. Identifying High Frequency Crash Locations: Empowering End-Users with GIS Capabilities. *Institute of Transportation Engineers Journal,* 77(1), 22-27.

Subramanian, Rajesh. March 2007. Motor Vehicle Traffic Crashes as a Leading Cause of Death in the United States, 2004. National Highway Traffic Safety Administration.

Thomanna, Jose V. September 2006. Traffic Accident Data Analysis Using Geographic Information Systems (GIS) Based Software. *Institute of Transportation Engineers Journal,* 76(9), 75-78.

Traffic Safety Facts 2005, *USDOT, National Highway Traffic Safety Administration (NHTSA)*

U.S. Census Bureau News, April 2007. URL - http://www.census.gov/Press-Release/www/releases/archives/population/009865.html

Using GIS for Safety Database and Analysis, *Tech Transfer Newsletter, Institute of Transportation Studies at University of California Berkeley.* (2000) http://www.techtransfer.berkeley.edu/newsletter/00-2/GIS.php

Vasudevan, Vinod, Pulugurtha, Srinivas S. and Nambisan, Shashi S. 2007. Methods to Prioritize Pedestrian High Crash Locations and A Statistical Analysis of Relationships Between Them. Transportation Research Record 2002/2007, 39-54.

VITA

Graduate College
University of Nevada, Las Vegas

Vidhya Kumaresan

Home Address:
4217 Grove Circle Apt #2
Las Vegas, NV 89119

Degree:
Bachelor of Engineering, Civil Engineering, 2002
Anna University, Chennai, India

Special Honors and Awards:
Member of honor societies Tau Beta Pi and Phi Kappa Phi
Winner of 2008 ITE Nevada Chapter Student Paper Competition
Serving as Vice President of ITE-UNLV Student Chapter

Thesis Title: Development of a GIS-based Safety Analysis System

Thesis Examination Committee:
Research Chairperson, Dr. Shashi Nambisan, Ph.D., P.E.
Academic Chairperson, Dr. Edward Neumann, Ph.D., P.E.
Committee Member, Dr. Mohamed Kaseko, Ph.D.
Graduate Faculty Representative, Dr. Ashok K. Singh, Ph.D.