Development of Criteria to Identify Pedestrian High Crash Locations in Nevada

Quarterly Progress Report

Submitted to

Nevada Department of Transportation (NDOT) Research Division
1263 South Stewart Street
Carson City, NV 89712

Srinivas S. Pulugurtha
Krishna Kumar K. Vanjeeswaran
Shashi S. Nambisan

Transportation Research Center
Howard R. Hughes College of Engineering
University of Nevada, Las Vegas
4505 S. Maryland Parkway Box 454007
Las Vegas, NV 89154-4007
Telephone: (702) 895-1338
Fax: (702) 895-4401

July 15, 2003
Nevada has experienced over 40 pedestrian fatal crashes per year over the last six years. Likewise, Nevada also has experienced over 800 pedestrian injury crashes per year during the same period. More than 70 percent of these pedestrian fatal crashes and pedestrian injury crashes are in Clark County, Nevada. There is a critical pedestrian safety issue on many urban streets in Nevada, in general, and in the Las Vegas metropolitan area in Clark County, Nevada, in particular. The Las Vegas metropolitan area is ranked among the worst urban areas in terms of pedestrian safety. Crashes in such an environment also result in adverse publicity, which can linger long after the incidents themselves. Besides the adverse publicity, these crashes result in significant health and human life consequences, and monetary impacts.

The main objective of this research project is to develop criteria to identify “pedestrian high crash locations” in order to allocate resources including Federal Safety Funds, for safety improvements. The criteria will help in the development of a “Pedestrian Safety Program”, as a part of the Nevada Department of Transportation’s (NDOT) Federal Highway Safety Improvement Program (HSIP). The developed criteria will assist the system managers not only in Las Vegas and Nevada, but also nationally, in better understanding the cause of the crashes and identifying appropriate operating strategies to enhance pedestrian safety.

The proposed research is divided into the following main tasks:

1. Task 1: Literature Research
2. Task 2: Data Collection and Geocoding
3. Task 3: Analysis of Data
4. Task 4: Develop Criteria to Identify “High Crash Locations”
5. Task 5: Recommendations and Scope for Further Research

**Literature Research**

Several analytical tool and techniques are available to analyze crash data.
Combination of Statistical Methods

Most GIS packages have very sophisticated database operations. However, they do not have statistical methods other than means and standard deviations of variables. Therefore, the statistical spatial method is needed when a more sensitive quantitative method is required (Levine 1999). Schneider, Khatk, and Ryznar (2002) adopted statistical method on their cluster analysis in order to integrate and evaluate data from two different sources (crash reports and survey data). CrimeStat, a spatial statistics software package, was used to perform the cluster analysis. In addition, several other spatial statistics such as Chi-squared, Ripley’s K-function, and G-function tests were also computed.

In short, pedestrian safety studies involve data collection and spatial analysis. The basic data needs for this analysis are crash reports, street centerline coverage, and demographic data. The spatial analyses include use of zone guide for pedestrian safety, and integration statistical methods with GIS.

Schneider et al. (2001) explains the importance of methods to identify where the pedestrian crash problem exists so that a greater number of pedestrian crashes can be prevented in the future. Studies in recent years have focused on the issue of safety analysis using GIS techniques. Even though GIS techniques are not extensively used in safety field, they have greater potential to improve crash location evaluation. Several studies have cited the benefits for using GIS to plot automobile crash locations and identify high-risk areas for motorized-vehicle crashes, though fewer have applied the technique to analyze pedestrian or bicycle crashes. Simple crash plotting, or geocoding crash locations, is the most common GIS technique used for safety studies (1). GIS turns statistical data, such as crashes, and geographic data, such as roads and crash locations, into meaningful information for spatial analysis and mapping (2). Using GIS it is relatively simple to combine information received annually on crashes and determines any correlation such as type of street and adjacent land use. GIS also assists in identifying any factors that were contributing to those crashes and/or potential solutions to reduce those crashes (3). GIS-based crash data analysis can influence the four E's of traffic safety: engineering, enforcement, education, and emergency response (4).
Analyses of Pedestrian Crashes

Pedestrian crashes can be categorized into three major areas (Baltes 1998). They are:

1. Pedestrian characteristics - which explains characteristics of persons involved in these crashes (gender, age and ethnicity)
2. Crash types – which explains elements that lead to crashes (for example, alcohol related, failed to yield the right-of-way and stepped into the path of an oncoming vehicle, disregarded a traffic signal, or made some improper action that contributed to the crash like crossing not at intersection, crossing at mid block crosswalk, crossing at intersection, walking along road with traffic, walking along road against traffic, working on vehicle in road, standing playing in road, standing in pedestrian island, etc), and
3. Crash event – which explains when and where did these crashes occur (date, time of the day, day of the week, location (urban or rural), weather and lighting conditions, roadway number of lanes, road system identifier, and road surface conditions).

This way of categorizing the pedestrian crashes helps develop effective and practical countermeasures to reduce the pedestrian injuries and fatalities. It is important that crash types are analyzed for different pedestrian age groups. For example, studies have shown that alcohol impaired pedestrian problem is high among some racial and ethnic groups which points to another set of characteristics to be analyzed (NHTSA 1998; Leaf and Preusser).

Analysis based on the number of pedestrian crashes in a particular age group, ethnicity, or gender group is useful, but insufficient for determining whether a specific group is more or less prone to be in a crash. This can only be obtained by considering crash rate per capita (from census data) and crash rate per kilometers walked (from Nation Wide Personal Transportation Survey, NTPS information). These crash rates by different age groups will show which age group are most likely to involve in a crash.

Analysis of crashes based on severity is another critical element. Higher vehicle speeds are strongly associated with both a greater likelihood of pedestrian crash occurrence and more serious resulting pedestrian injury (Leaf et al. 1999; IIHS 2000).
The population density is not a good replacement for pedestrian exposure as it does not account for the amount of walking people do (Qin and Ivan 2001). The number of lanes, area type and sidewalk system are some of the factors that affects the pedestrian exposure.

Crash studies are generally based on reported crash records. Schneider, Khattk, and Ryznar (2002) state that reported crash data alone may not be a good predictor of future crash locations, especially for infrequently - occurring pedestrian crashes. To solve this problem, Schneider, Khattk, and Ryznar presented the idea of combining the crash data with perception survey method. The study concluded that perception survey data helps improve site selection and recommendations for pedestrian safety treatment (for example, gather large quality of data about locations that may have pedestrian problems, and study differences in the perceptions of people with specific traits). However, surveying method may not be appropriate for large study areas such as city or metropolitan areas as it is a time consuming and expensive process.

Braddock et al. (1994) identifies two high pedestrian crash locations which account for 30 percent of all pedestrian crashes in Hartford County, Connecticut based on address matched crash data for analysis (15).

**Tools and Techniques**

Several analytical tool and techniques are available to analyze crash data. However, questions such as “where are most of the crashes occurring and why?” is difficult to answer. These questions can be easily achieved in a GIS environment. Using GIS to geocode crash locations and plot the locations is the most common first step (Anadaluz, Robers, and Tina 1997). In order to ensure a reasonable stable measure, experience has shown that a minimum of one year's data or at least 100 crash records should be available for establishing pedestrian safety zones (NHTSA, 1998). For data analysis, various techniques were used to create zones, identify hotspot locations, and rank the study locations.

NHTSA (1998) recommended the guide to identify study zones for pedestrian safety. The zone process provides a systematic method for targeting pedestrian safety improvements in a cost effective manner. Zoning identifies a subset of a jurisdiction containing as much of the
pedestrian problem of interest in as little land area as possible. The first step is to select an initial shape for the zones and to define the target rate i.e., the number of events that must fall in an area for it to be defined as a zone. The approach suggested is to search for circular zones, then to search for linear zones, then to determine their final shape. The initial circular zones could be created by using one mile radius, as generally pedestrian crashes occur within one mile of the victim’s home or work place. Risk zones could be identified using a target rate of 10 crashes per zone for total 200 crashes data. For linear zones, it could be determined for the segments where six or more crashes occur in a two miles for total 200 crash data. In addition, if total crash data that are used in analysis is higher, the target rate should be adjusted upward as necessary. The final step is to identify the final zone shape, as it may be useful to combine zones, add more radiuses, change zones’ shape, or reduce zones’ size. Finally, to define zones, areas with some clustering and some dispersion throughout a land area should be identified. However, such a methodology may not work if no clustering is apparent in the study area.

A few GIS based studies are briefly discussed next. Braddock, Lapidus, Cromly, Burke, and Banco (1994) identified two high pedestrian crash locations which account for 30 percent of all pedestrian crashes in Hartford County, Connecticut based on address matched crash data for analysis. In a different context, three-mile buffer zones were created around 3 clustered areas using GIS to study moped safety in Hawaii. The temporal variations, environmental characteristics, and crash characteristics of these spatially distributed moped crashes were then studied (Kim, Takeyama, and Nitz 1995).

A GIS based crash analysis tool developed by FHWA (1999) uses five different types of analysis to evaluate crashes. The Spot/Intersection Analysis program is used to evaluate crashes at a user-designated spot or an intersection within a given search radius. The Strip Analysis program is used to study crashes along a designated length of roadway as opposed to a spot or an intersection. The Cluster Analysis program is used to study crashes clustered around a given roadway feature such as a bridge, railroad crossing, or traffic signal. The Sliding-Scale Analysis program is used to identify roadway segments with a high crash occurrence. The Corridor Analysis program is used to locate high crash concentrations within a corridor. Using traditional
methods, segments along a specific route could be examined, but multiple routes within a corridor could not be easily linked and analyzed as a group, which is possible using this tool.

A simple method, called nearest neighborhood analysis, was used to identify hot spot locations in a mid-block pedestrian safety study (Cui 2000) The analysis used grid cells with a dimension of 100 feet per site and a circular radius of 500 feet. The resultant scores were grouped and ranked based on the distribution of number of pedestrian crashes.

Steiner et al. (2002) discusses about three steps for developing GIS crash mapping: (1) identification and collection of data (both crash data and the map data layer); (2) selection of a program for processing of crashes; and (3) analysis of data collected by the system.

**Identify High Crash Locations**

One of the most common macroscopic applications of GIS is the determination of high crash locations, HCLs (Roche 2000). HCLs identify the areas that would potentially receive the largest benefit if safety funds were allocated. These locations can be analyzed in many different ways. One method of HCL identification includes crashes within a specified distance of a major roadway. Another method determines the crash frequency within a specified distance. One of the drawbacks of identifying locations with high crash frequencies is that traffic volume or exposure is not taken into account. This can be accomplished by the crash rate method. The crash rate method for roadway segments divides the total number of crashes by the annual average daily traffic (AADT) and the length of the segment to obtain crashes per vehicle miles traveled.

Cui (2000) used nearest neighborhood analysis, to identify hot spot locations in a mid-block pedestrian safety study. The study used grid cells with a dimension of 100 feet per site and a circular radius of 500 feet. The resultant scores were grouped and ranked based on the distribution of number of pedestrian crashes.

South East Michigan Council of Government (SEMCOG) Crash Analysis Manual (SEMCOG 2001) explains five ways of locating high crash locations. Spot map method, the simplest method of identifying high-crash locations, is to examine a map showing clusters of symbols at those
spots and on those segments in the road network having the greatest numbers of total crashes. Crash Frequency Method is a method to rank locations by the number of reported crashes (or crashes per mile), with frequencies listed in descending order. Locations having crash frequencies greater than or equal to a critical crash frequency are considered to be high-crash locations. Crash Rate Method ranks locations by descending crash rate. Locations with above-average rates are tested for significance. The Crash Rate Method compares the number of crashes to the volume of traffic, with the later measured either as the number of vehicles crossing a spot in a given time period, or as the number of vehicle-miles of travel along a segment in that period. In Crash Severity Method, crash frequencies or rates are weighted by severity whereas in the Crash Probability Index (CPI) Method frequency, rate, and severity results are combined.

The Crash Rate Method used to find HCLs is a simple statistical test to determine whether the crash rate for a particular location is significantly higher than the average crash rate for other locations in the jurisdiction having similar characteristics. If the crash rate is higher than the average crash rate, the location is classified as a high-crash location. The steps involved in this method are as follows.

1. Determine the location’s crash rate. The spot crash rate is found as annual average number of crashes during the study period divided by the average daily traffic volume during the study period in crashes per million vehicles. The segment crash rate is found as spot crash rate divided by length of segments in crashes per million vehicles per miles.
2. Define the location type. Categorize the location by as many of the following features as possible: area type, roadway functional class (arterial, collector, or local), number of lanes, etc.
3. If a list of previously evaluated locations is being maintained, insert the location into the list of locations ranked in descending order by crash rate
4. Determine the critical crash rate.
5. Compare the location’s crash rate to the critical crash rate. If the crash rate equals or exceeds the critical crash rate, classify the location as a high-crash location.
Iowa Department of Transportation, Office of Traffic and Safety (OTS IDOT 2001) describe the advantage of using crash rate in comparing the crash experience between different time periods or between locations. This provides a basis for more accurate and meaningful conclusions since it accounts for the numbers of vehicles “exposed” to the hazards of driving within a given time period. It also prevents the potentially misleading classification of a relatively safe high-volume location as “high-crash” simply because it has experienced a relatively large number of crashes. However, it tends to unfairly identify low-volume locations having relatively few crashes as high-crash locations.

**Data Collection and Geocoding**

Digitizing crashes on a digital map with street network is not only inaccurate but a time consuming process. On the other hand, the process of automatically creating map features based on address, or similar information exploring the capabilities afforded by GIS software is called geocoding. Crashes can be geocoded using one of the three reference systems (street name / reference street name, mile-post and address). The street name / reference street name reference system and address are most commonly used in urban areas. The advantage of geocoding is that it lets one map locations from crash data that is readily available.

However, a street network in a GIS format with street name and address information is extremely important to geocode crash data. Street centerline (SCL) network in a GIS format are generally developed by public and private agencies. A few of these are commercially available. SCL network attributes include street name, street type (Avenue, Boulevard, and so on), and directional prefixes and suffixes necessary to avoid ambiguity in address location. Each street feature is divided into segments that have beginning and ending addresses, as you see on neighborhood street signs. This makes it possible to estimate the position of an address along the length of a street segment. There may be separate address ranges for each side of the street, so that an address can be geocoded on the correct side of the street.

The Transportation Research Center, UNLV has the SCL coverage for the Clark County developed and maintained by the Clark County Department of Public Works GIS Managers Office (GISMO). The SCL coverage for the Clark County has 61,573 street segments. Street
name and address information is available for all these streets. However, data is not available for other counties in Nevada. A search was conducted to obtain data from other sources. The other common sources for the street network data are: 1) Tiger/Line data from the United States Census Bureau, 2) Geographic Data Technology (GDT) Dynamap U.S Street Data, and 3) Tele Atlas MultiNet.

Census 2000 TIGER/Line data can be downloaded in a shapefile format from United States Census Bureau website free of cost. For the state of Nevada TIGER/Line data contains 345,124 street segments out of which 157,355 are named street segments (45.6%).

GDT Dynamap/2000 United States Street Data can be purchased online in variety of formats including the shapefile format. For the state of Nevada Dynamap/2000 data contains 446,844 street segments out of which 238,716 are named street segments (53.4%). The cost for a 1-5 user internal license, for the state of Nevada, perpetual use of the data, for Dynamap/Transportation is $10,500.00 and a 1-5 user internal license, annual use of the data, is $7,875.00.

Tele Atlas, a private provider of digital maps, offers a product called Tele Atlas MultiNet which is has 40,000 street network segments for the state of Nevada out of which 38,000 are named segments. The cost for up to 5 users of the Tele Atlas MultiNet product in a shapefile format for use on PCs is $7,030.00 without driving directions (routing attributes). With routing attribute information the cost is $14,440.00.

The number of street segments and percent of named street segments in TigerLine data and GDT databases for each County in the State of Nevada are summarized in Table 1. As can be seen from the table, percentage of named street segments is less than 70 percent for most of the counties in the State of Nevada. This might limit the number of crashes that could automatically be geocoded using GIS software. Though, GDT has more percent of named street segments, it is expensive compared to Tiger/Line data which is available free of cost.
TABLE 1 Number of Street Segments and Segments with Street Name by County

<table>
<thead>
<tr>
<th>County</th>
<th>No. of Street Segments</th>
<th>% of Named Street Segments</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Tiger/Line</td>
<td>GDT</td>
</tr>
<tr>
<td>Clark</td>
<td>75,072</td>
<td>108,735</td>
</tr>
<tr>
<td>Carson City</td>
<td>4,785</td>
<td>6,560</td>
</tr>
<tr>
<td>Washoe</td>
<td>34,122</td>
<td>54,110</td>
</tr>
<tr>
<td>Douglas</td>
<td>7,732</td>
<td>8,211</td>
</tr>
<tr>
<td>Lyon</td>
<td>10,991</td>
<td>14,332</td>
</tr>
<tr>
<td>Storey</td>
<td>1,947</td>
<td>2,053</td>
</tr>
<tr>
<td>Churchill</td>
<td>12,736</td>
<td>15,860</td>
</tr>
<tr>
<td>Nye</td>
<td>41,824</td>
<td>53,144</td>
</tr>
<tr>
<td>Elko</td>
<td>41,870</td>
<td>47,038</td>
</tr>
<tr>
<td>Eureka</td>
<td>7,096</td>
<td>9,746</td>
</tr>
<tr>
<td>White Pine</td>
<td>26,095</td>
<td>31,624</td>
</tr>
<tr>
<td>Humboldt</td>
<td>19,809</td>
<td>20,431</td>
</tr>
<tr>
<td>Mineral</td>
<td>10,573</td>
<td>13,178</td>
</tr>
<tr>
<td>Lander</td>
<td>11,150</td>
<td>11,505</td>
</tr>
<tr>
<td>Pershing</td>
<td>13,787</td>
<td>18,176</td>
</tr>
<tr>
<td>Lincoln</td>
<td>17,617</td>
<td>22,157</td>
</tr>
<tr>
<td>Esmeralda</td>
<td>7,918</td>
<td>9,984</td>
</tr>
<tr>
<td>Total</td>
<td>345,124</td>
<td>446,844</td>
</tr>
</tbody>
</table>

PROGRESS

As of June 30 2003 Task 1 has been completed. In the next quarter (July 1 - September 30, 2002), activities will include working on Task 1 and Task 3. This will include 1) collection of street network data for Clark County, Carson City, Washoe County, Douglas County and Lyon County, 2) geocode crash data subject availability of street network, and 3) analyze crash data.

REFERENCES

17. USDOT (1999) GIS-Based Crash Referencing and Analysis System. HSIS Summary Report,
FHWA-RD-99-081, FHWA, U.S. Department of Transportation (U. S. DOT).
Department of Civil & Environmental Engineering, University of Nevada, Las Vegas.
Crash Mapping: A Study of Eight Florida Counties. 2002 Annual Transportation Research
Board Meeting, Pre-print CD-ROM.
Department of Transportation (IDOT), 2001.