GIS APPLICATION FOR BETTER IDENTIFICATION OF HAZARDOUS HIGHWAY LOCATIONS

Supornchai UTAINARUMOL
Ph.D, lecturer
Department of Civil Engineering
King Mongkut's Institute of Technology
North Bangkok
1518 Pibunsongkram Rd.
Bangsue, Bangkok 10800 Thailand
Tel/Fax: (662) 587-4337
E-mail: paul@kmitnb05.kmitnb.ac.th

Robert E. STAMMER, Jr.
Associate Professor
Department of Civil Engineering
Vanderbilt University
Box 90, Station B
Nashville, TN 37235, USA
Tel: (615) 343-8061
Fax: (615) 343-0670
E-mail: docbob@vuse.vanderbilt.edu

Abstract: The identification of hazardous highway locations is an important first step in any highway safety plan. The technique used to identify these locations should be sufficiently accurate to instill a high degree of confidence in the reported results.

At present, Geographic Information System technology is a very popular information management technique used for more effectively managing accident data and performing accident analyses [Filian an Kim in 1995]. GIS provides several capabilities that can identify hazardous highway locations graphically. Therefore, the objective of this paper is to illustrate how GIS and other analytical tools can improve the identification of hazardous highway locations.

The derived system was implemented for the Tennessee interstate highway. This system emphasized flexible use of existing methods such as accident frequency, accident rate, accident severity, rate quality control, and easy user selection of a combination of any of these methods.

The prototype system applications using various analytical methods in this paper was developed to identify hazardous highway locations more comprehensively. The system was developed with flexibility in mind so that the user can identify hazardous highway locations by each method alone or by combining two or more methods together. Additional advantages of the system are the system's capabilities and the ease of use.

1. INTRODUCTION

Accidents cause the loss of life and money. The extremely high costs associated with highway crashes cause highway safety improvements to be an important objective of transportation engineering. General safety measures such as speed limit regulation, increased law enforcement, and education, or more localized measures relating to local traffic control and geometry improvements can affect highway safety. The latter, more localized methods are used at individual road facilities such as intersections and along roadway segments.

Every year, local government provides budgeted funds for safety improvements. Portions of these safety funds are used to improve specific roadway facilities in order to reduce roadway accidents. However, the budgeted funds are constrained and limited. Therefore, the locations truly requiring improvement must be identified correctly both to minimize future accidents and to receive the highest benefits.

Several years ago, the technology of GIS emerged and became extremely popular in many fields. Today, GIS provides many useful capabilities to manage accident data and produce helpful graphic maps. Moreover, GIS allows users to design queries and programs which are important when identifying hazardous highway locations. Thus, GIS is a logical choice when identifying hazardous highway locations. The implemented system can determine hazardous highway locations accurately for an entire state and present those locations clearly on maps. This system provides many benefits and much improved identification since users can spend only a few minutes performing their desired analysis, but still obtain extremely reliable results.

This paper presents how GIS and other analytical techniques were used to identify hazardous highway locations in a more detailed manner. The created system was then used to identify hazardous highway locations throughout the entire Tennessee interstate highway system by various techniques. This system emphasized flexible use of existing methods such as accident frequency, accident rate, accident severity, rate quality control, and easy user selection of a combination of any of these methods.

2. GIS SOFTWARE SELECTION

TransCAD was the selected GIS software used for developing the accident database, and for identifying hazardous highway locations because TransCAD provided all the tools, maps, road networks, coordinates, milepost conversions, queries and other features necessary for identifying potentially hazardous highway locations [TransCAD in 1996]. TransCAD is operated on Windows, thus making its operation easy and not costly. In addition, TransCAD accepts the format of the accident data file and other required data which were downloaded from Tennessee Roadway Information Management System (TRIMS) database.

GIS Developer's Kit (GISDK) was the programming software used for creating the user interface [GISDK in 1995]. The GISDK is a programming software which supports TransCAD. The GISDK was used to develop queries and macro programs, such as developing user interfaces, and to identify hazardous highway locations in this research.

3. SYSTEM DATA

The accident data were downloaded by selecting required data from TRIMS database and were saved as a tab delimited text format file. The accident data file was imported into TransCAD without changing the format because TransCAD accepts the delimited text format file. Figure 1a shows the accident data structure. The accidents were identified by their field *nbr_case* which is a unique number. Other descriptive attributes included *nbr_tenn_cnty*, *nbr_rte*, *spcl_cse*, *cnty_seq*, *log_mle*, *dte_accd*, *tot_injr*, *tot_killed*, *tot_vhcl*, *typ_collision*, and *mslink*. The fields *nbr_rte* and *log_mle* are important because they identify the location of an accident and they were used to attach accident records into interstate routes. The fields *tot_injr* and *tot_killed* were used to describe the severity of accident. These fields were used to calculate accident frequency, accident rate (R), severity index (SI), hazardous index (HI), and other indices in the analysis program.

Additional required data included traffic data, pavement surface data, roadway geometric data, interstate route data, and group of interstate routes data. The structure of these data is described in Figure 1b to 1f.

COLUMN	DATA TYPE	DESCRIPTION
nbr case	Number(8.0)	accident case number
nbr tenn cnty	Number(2.0)	county number
nbr rte	Varchar2(8)	roadway route number
spcl cse	Varchar2(1)	special case
cnty seq	Number(2)	county sequence
log mle	Number(5.3)	accident location described in log mile
dte accd	Date	the date of occurred accident
tot injr	Number(2.0)	number of person injured
tot killed	Number(2.0)	number of person killed
tot vhcl	Number(2.0)	number of vehicles involved
typ collision	Varchar2(1)	characteristic of vehicles collision
mslink	Number(10)	missing link

a) Accident Data Structure (trimsacc.txt)

COLUMN	DATA TYPE	DESCRIPTION
nbr tenn cnty	Number(2.0)	county number
nbr rte	Varchar2(8)	roadway route number
spcl cse	Varchar2(1)	special case
cnty seq	Number(2)	county sequence
BGLM	Number(5.3)	beginning log mile
ELM	Number(5.3)	ending log mile
DATE	Date	year of traffic
AADT	Number(6)	average annual daily traffic
DIRDST	Number(2)	distribution of traffic in each direction
mslink	Number(10)	missing link

b) Traffic Data Structure (trimstfc.txt)

COLUMN	DATA TYPE	DESCRIPTION
nbr tenn cnty	Number(2.0)	county number
nbr_rte	Varchar2(8)	roadway route number
spcl_cse	Varchar2(1)	special case
cnty seq	Number(2)	county sequence
BGLM	Number(5.3)	beginning log mile
ELM	Number(5.3)	ending log mile
POINT	Number(3)	point rating
SKDCOEF	Number(2)	skid coefficient
DATE	Date	date of collection
mslink	Number(10)	missing link

c) Pavement Surface Data Structure (trimssuf.txt)

COLUMN	DATA TYPE	DESCRIPTION
nbr tenn cnty	Number(2.0)	county number
nbr_rte	Varchar2(8)	roadway route number
spcl_cse	Varchar2(1)	special case
cnty seq	Number(2)	county sequence
BGLM	Number(5.3)	beginning log mile
ELM	Number(5.3)	ending log mile
TERRAIN	Varchar2(1)	type of terrain
SPDLMT	Number(3.0)	speed limit
NO.LANES	Number(1)	number of lanes
mslink	Number(10)	missing link

d) Roadway Geometric Data Structure (trimsgmt.txt)

COLUMN	DATA TYPE	DESCRIPTION
nbr rte	Varchar2(8)	roadway route number
BGLM	Number(5.3)	beginning log mile
ELM	Number(5.3)	ending log mile

e) Interstate Routes Data Structure (tnintst.txt)

COLUMN	DATA TYPE	DESCRIPTION
nbr rte	Varchar2(8)	roadway route number
BGLM	Number(5.3)	beginning log mile
ELM	Number(5.3)	ending log mile
group	Number(1)	group number

f) Group of Interstate Route Data Structure (trimsgrp.txt)

Figure 1 System Data Structure

4. METHODS OF IDENTIFYING HAZARDOUS HIGHWAY LOCATIONS

After the accident database was completed, the hazardous highway locations were identified based on the following methods [Zegeer in 1982 and Gharaybeh in 1991]:

4.1 Accident Frequency Method

The accident frequency method is used to search the accident file for concentrations of accidents within variable segment length. Usually one or more segment length (0.1-mile, 0.2-mile, 0.3-mile, 0.4-mile, 0.5-mile, etc.) are used to "float" through the accident file in which accidents are ordered by location, and sections that meet or exceed a predefined accident criterion are identified. Such floating segments generally advance in 0.1-mile increments through the file. The roadway segments are ranked based on the accident frequency. The segments which have the same accident frequency are given the same ranked value.

4.2 Accident Rate Method

The accident rate method consists of simply dividing the accident frequency at a location by the vehicle exposure to determine the number of accidents per million vehicle-miles of travel at highway segments. Segments are priority ranked in order of descending accident rate. The accident rate for highway segment that is currently used by TDOT is calculated as follows:

$$R = A*1,000,000/(365*T*V*L) \tag{1}$$

where:

- R = accident rate for highway segment (in accidents per million vehicle miles),
- A = number of accidents for given analysis period,
- T = time of analysis period (in years or fraction of years),
- V = average annual daily traffic (AADT) during study period, and
- L = length of highway segment (in miles).

4.3 Rate Quality Control Method

The rate quality control method not only entails the calculation of the accident rate at each location, but also a statistical test to determine if that rate is significantly higher than accident rates for other locations with similar characteristics. The statistical test is based on the commonly accepted assumption that accidents follow a Poisson distribution. For each location, a critical rate which is currently used by TDOT is computed as follows:

$$R_c = R_a + K(R_a/E)^{0.5} + 1/(2E) \tag{2}$$

where:

- R_c = critical accident rate for highway segment (accidents per million vehicle-miles),
- R_a = average accident rate for all highway segments of similar characteristics or on similar road types,
- E = million vehicle-miles of travel on the highway segment during the study period $= (365*T*V*L)/1,000,000$, and
- K = a probability factor determined by the desired level of significance for the equation. The value of K corresponding to 99% confidence level is 2.327. This value is recommended and used by TDOT.

The critical rate (R_c) is computed for each location and compared to the actual accident rate (R). If the actual accident rate exceeds the critical rate, then the location may be considered for improvement. Therefore, the highway locations are ranked based on their R/R_c ratio.

According to this method, the study area has to be divided into several groups. In this paper, the interstate highways were divided into four groups based on the characteristics of roadway and regional offices which are responsible for maintaining the state highway system. These four regions are shown in Figure 2.

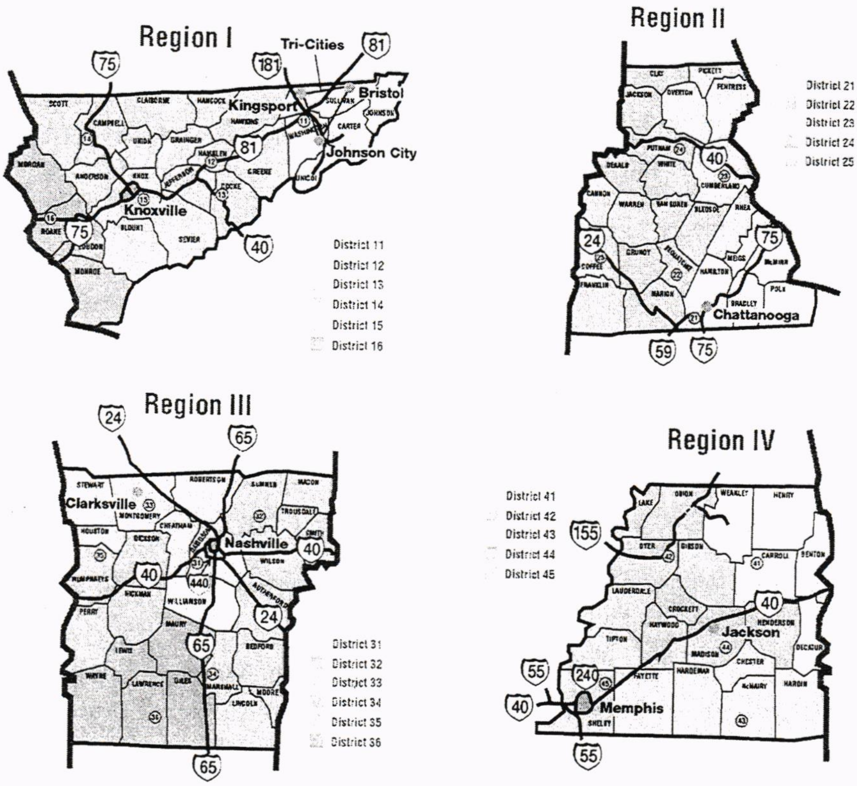

Figure 2 Four Regions of Tennessee

4.4 Accident Severity Methods

The accident severity methods are used to identify and/or rank locations based on the number of severe accidents at each location. Accident severity is defined as a fatal accident, personal injury accident, or property damage only (PDO) accident.

The severity index (SI) is calculated for each roadway segment. The roadway segments are ranked based on SI values. The severity index (SI) that is used to rate roadway segments in the Tennessee Department of Transportation (TDOT) is determined using the following equation:

$$SI = (F + PI) / Total\ Accidents \tag{3}$$

where:

SI = severity index,
F = number of fatal accidents during study period,
PI = number of personal injury accidents during study period, and
Total = total number of all types of accidents for this segment.

4.5 Combination Method

The major contribution of recent research was the development of an improved decision management tool. A combination method combines all four methods described previously together (or other various selected combinations), and their inherent individual strengths, in order to create a better decision-making tool for selecting hazardous highway locations. The hazardous highway locations are ranked on the basis of a newly developed Hazardous Index (HI) determined by the following equation:

$$HI = (F\_Rank + R\_Rank + S\_Rank + Q\_Rank) / 4 \quad (4)$$

where:

HI	=	hazardous index,
F_Rank	=	rank of location by accident frequency method,
R_Rank	=	rank of location by accident rate method,
S_Rank	=	rank of location by accident severity method, and
Q_Rank	=	rank of location by rate quality control method,

(Note: The denominator value of "4" must correspond to the number of methods whose ranks are totaled in the numerator.)

The location that has the lowest Hazardous Index will be ranked first. The location which has the next highest Hazardous Index will be second and so on.

The C program used to identify hazardous highway locations was based on the methods described above. The program was written following the flowchart logic presented in Figure 3. In order to explain clearly how the program works, the following hypothetical example is presented. For example, there are two roadways (I-440 and I-65). I-440 starts at milepost 0.0 and ends at milepost 15.0 and I-65 starts at milepost 1.5 to milepost 7.0. First, the user must input the roadway segment length (L). In this case, 0.5-mile segments were specified. The search will begin along I-440 at milepost 0.0 and the first length reviewed on the program will be milepost 0.0 to 0.5. The number of accidents located on this segment will be counted and accident frequency, accident rate, critical rate, and other accident indices calculated based on the above equations and then stored in memory. After the first segment is finished, the second segment (0.1 to 0.6) will be searched by using a 0.1-mile increment to create data for the next segment. The calculated values for this second segment will also be saved in memory. The next segment to be reviewed will be milepost 0.2 to 0.7. The calculations are performed on this segment and also saved in the memory until the final segment is reviewed (milepost 14.5 to 15.0). Then the next roadway will be selected (I-65). The first segment of I-65 from milepost 1.5 to 2.0 will be reviewed and calculations performed. The next segment tested after 1.5 to 2.0 will be 1.6 to 2.1. The process will continue until all segments of all roadways have been reviewed. Then the final step is to compare and rank the calculated values which were saved in memory. The segments meeting certain accident criterion will be printed and presented in the thematic map.

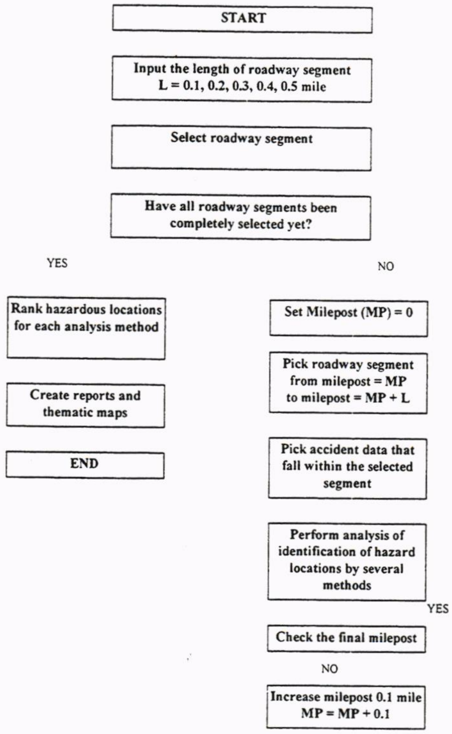

Figure 3 Flowchart of a C Program for Identifying Hazardous Highway Locations

5. SYSTEM IMPLEMENTATION

In developing the system for identifying hazardous highway locations, all accident records need to be associated with a roadway location on a map. An accident location in the accident data was recorded in milepost format. Therefore, the roadway routes and mile reference system had to be developed in order to attach accident data into the roadway route.

The Tennessee interstate routes were created based on a network which contains nodes and links. The network was created first, then routes were developed. Therefore, the following steps were performed during the process of network and route development.

5.1 TN Interstate Network Development

A network must be built if routing is to be performed. Networks are a series of links in which all elements are connected. In this paper, a network consisting of the interstate system of Tennessee was created. Before building the network, a base map of the system was needed. A base map was obtained by importing the Interstate Highways file from Caliper's CD ROM into TransCAD. This geographic file contained the interstate system for the entire United States. Therefore, only the interstates within Tennessee were copied

and then exported to a new geographic file. This new geographic file was used as a base map.
The next step was to develop a TN interstate layer. The TN interstate layer contained nodes and links based on the base map. To develop the TN interstate layer, a line geographic file must be created. Then this layer was added to the based map. The line editing tools were used to create, add, and modify nodes and links on TN interstate layer. The nodes must be placed on all exit points and interchanges between interstates and secondary highways in order that the milepost of these nodes can be calibrated in a later step (see milepost calibration in section 5.3).

After the TN interstate layer was done, the TN interstate network file is built. This network file is used to create interstate routes in the following section.

5.2 TN Interstate Routes Development

In developing the interstate routes, the TN interstate network has to be activated. After completion of setting the active network, route system file is created. The route system file containing the designated line layer was added to a map in the TN interstate routes layer.

All TN interstate routes resulted in 66 routes as shown in Figure 4. These TN interstate routes were created in each county through which an interstate highway passed. The starting milepost 0.0 was located at the west or south boundary of each county and ended at the east or north boundary of the same county. The routes were coded in 8 characters, where the first 5 characters describe the name of interstate route and the last three characters represent the county number. For example, interstate route coded "I0440C19" is the interstate I-440 located in Davidson County.

5.3 Locating Accident Data

Before assigning or locating accident data to interstate routes, the mileposts of each interstate route had to be calibrated in order to make the accident locations more accurate. In calibrating mileposts, all interstate routes located in each county were checked against the length of interstate routes data obtained from the TRIMS database. This was performed because mileposts must be adjusted occasionally to match the TRIMS database.

After calibrating mileposts for all interstate routes in all 95 TN counties, all accident records were ready to be located into the map using the *Attach* tool in TransCAD. The accident data were located on TN interstate routes in a different layer named "Accident Locations" layer.

5.4 Thematic Layers

The eventual layers in this research totaled six. This includes the (1) TN Interstate Route Layer, (2) Accident Locations Layer, (3) Hazardous Location(s) Layer, (4) Traffic Volume Layer, (5) Roadway Geometry Layer, and (6) Road Surface Characteristic Layer. The details of these layers are described in the following section.

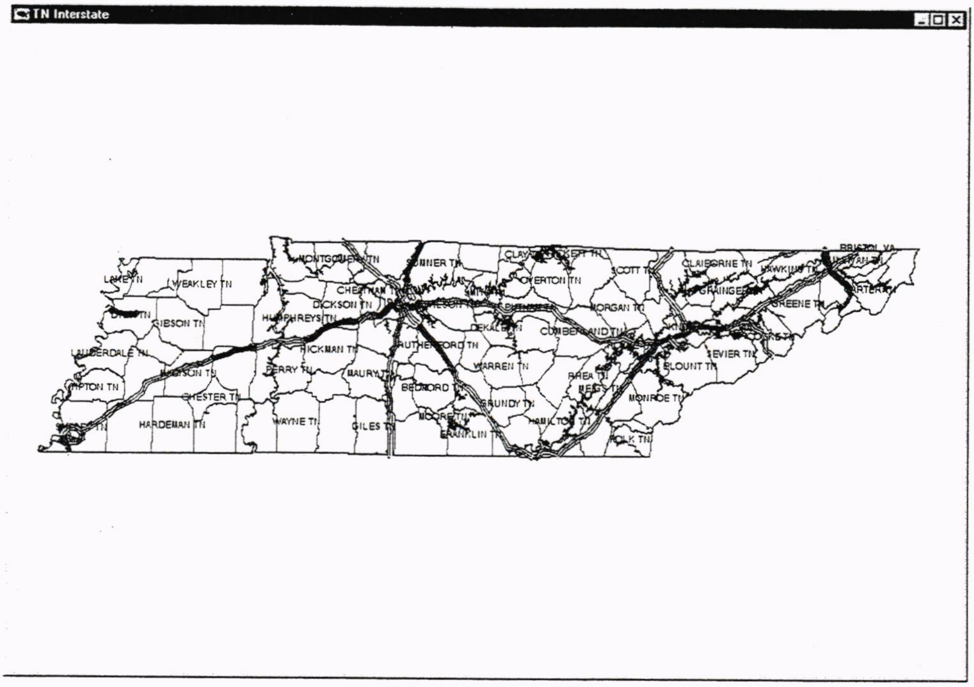

Figure 4 Tennessee Interstate Routes

5.4.1 TN Interstate Route Layer

The TN interstate route layer is comprised of route name and milepost reference of the interstate routes in Tennessee. This layer was developed based on the interstate network and was used to locate all accident records. The interstate routes data were grouped into the following two categories: (1) Identification route including route ID and route name, and (2) Length of route including beginning log mile and ending log mile.

5.4.2 Accident Locations Layer

The accident locations layer includes all accidents that occurred on TN interstate highways for three years from 1993 to 1995. The accidents were located to the interstate routes using the milepost reference system.

The non-spatial accident data were divided into three major groups: (1) Identification data including accident case number and date of accident, (2) Location information including route number and milepost, and (3) Type of accident including severity type, collision type, and vehicles involved.

5.4.3 Hazardous Location(s) Layer

The hazardous location(s) layer contains interstate route identification, milepost reference along the interstate routes, accident data, and traffic volume (AADT) data for the TN interstate highways. This layer was developed after accident analyses were performed and the output files were created. The output files were attached to the interstate route layer and displayed on a different layer named the "Hazardous Location(s)" layer. The hazardous location(s) layer presents the high accident locations that need to be improved. The data contained in this layer were grouped into four categories:

(1) Identification route including route ID and route name,
(2) Length of route including beginning log mile and ending log mile,
(3) Traffic volume including AADT, and
(4) Accident analysis results including accident frequency (F), accident rate (R), R/Rc, severity index (SI), hazardous index (HI), and rank of location calculations.

5.4.4 Traffic Volume Layer

The traffic volume layer is comprised of traffic volume (AADT) data for all TN interstate highways. Non-spatial traffic volume data on the TN interstate highways are grouped into three major categories:

(1) Identification route including route ID and route name,
(2) Length of route including beginning log mile and ending log mile, and
(3) Traffic volume and date including AADT, directional distribution of traffic, and observation date.

5.4.5 Roadway Geometric Layer

The roadway geometry layer includes the geometric characteristics of the TN interstate highways. This layer provides the information of speed limit, number of lanes, and type of terrain for the entire TN interstate highway system.

5.4.6 Road Surface Characteristic Layer

The road surface characteristic layer contains the information of roadway pavement conditions such as skid coefficient and rating of pavement, date of observation, and more.

5.5 User Interface Design

The User Interface was designed primarily to aid the user by simplifying the analysis process. Someone not familiar with the TransCAD software and its icons can use the system easily by simply clicking on the user interface icons to perform analyses as desired.

In designing a user interface for identification of hazardous highway locations, the intention is to make the system easier to use by limiting the selection options and to make the new set of options understandable. That is, the user interface should allow the user to perform the most common procedures with no outside instruction. The set of common procedures that were considered in designing the user interface are (1) accident analysis calculations, (2) searching accident locations, and (3) searching hazardous highway locations.

5.5.1 Accident Analysis Calculations

When performing accident analysis calculations, the C program described in section 4 is run in order to determine accident frequency (F), accident rate (R), severity index (SI), critical accident rate (Rc), average accident rate (Ra), R/Rc ratio, hazardous index (HI), and other factors. Then the locations are ranked based on these values. The user must input the accident data file (*trimsacc.txt*), AADT data file (*trimstfc.txt*), TN interstate information data (*tnintst.txt*), TN interstate group data (*trimsgrp.txt*), roadway geometry data (*trimsgmt.txt*), road surface characteristics data (*trimssuf.txt*), analysis time periods (years), and the length of roadway segment (miles). After a few minutes of program time for compiling, the output files are created in the comma delimited text format that can be opened as data viewed by TransCAD.

5.5.2 Searching Accident Locations

With this option, the user can find accident locations by selecting road name, and/or by county name, and/or by severity of accident. The system then displays, according to user selections, all accident locations and thematic charts on an appropriate map. The associated table information is also displayed concurrently with the geographic element(s).

5.5.3 Searching Hazardous Highway Locations

This option allows the user to find hazardous highway locations based on a number of different and variable criteria. Having earlier performed different methods of analyses, the user can specify their desired method by selecting the method(s) to be used and by inputting the number of top rankings to display. The user uses one of four methods (accident frequency, accident rate, accident severity, and rate quality control) or combines the methods together in any combination (e.g. two, three, or all four) in order to get the desired "best" results. These queries and macro programs that were created during the research were designed for flexible use. This flexibility in choices and uses aids the user in the task of more fully defining and understanding exactly what constitutes a hazardous highway location. This flexibility and the creativity afforded the researcher to define "a hazardous highway location" in numerous ways is truly one of the most important new capabilities resulting from this research.

Macro programs were written for user interface. They were coded in the text editor and saved as resource file. This program includes user interface dialog boxes and macros for running the system applications such as accident analysis and identification of hazardous highway locations mentioned in the previous section. After the resource file was completed, the resource file must be compiled to check for errors within the program. If there are any errors in this step, they have to be removed and the program compiled again until no errors occur.

6. SYSTEM APPLICATIONS

The preceding sections discussed issues involved with the system design and identification of hazardous highway locations. This section is devoted to illustrating real world applications of the user interface capabilities. Step-by-step instructions are provided that will enable another user to duplicate the application results.

This section also serves as somewhat of a user's guide describing the full capabilities of the system. Detailed instructions lead the user through the primary system applications.

The following application procedures begin once the user has: (1) opened TransCAD (by a double click on TransCAD icon), (2) clicked the *Tools* button, then the *Add-ins* option in the main menu to open the Add-ins dialog box, (3) selected *Identification of Hazardous Highway Locations* and clicked OK. After the user has done all the above, the *Identification of Hazardous Location(s)* dialog box will appear on the screen. When this occurs, the user can then proceed with the following application examples.

6.1 Accident Analysis

This type of application allows the user to run a C program described in Section 4 in order to perform model calculations and determine the rank of roadway segments. The advantage of this application is that the user can run a C program in TransCAD without loading other applications.

The user has to input data files including AADT, accidents, interstate information, interstate group, roadway geometry, road surface condition, analysis time periods (*years*), and segment length (*miles*). The next step is to run the executed file of the C program. After a few minutes of program running, the output files of the analysis methods described in Section 4 (Accident Frequency Method, Accident Rate Method, Rate Quality Control Method, Accident Severity Method, and Combination Method) will appear on the screen as shown in Figure 5. These files are in the comma delimited text format containing route name, beginning log mile, ending log mile, calculated values from the various models, rank of segment, and other information.

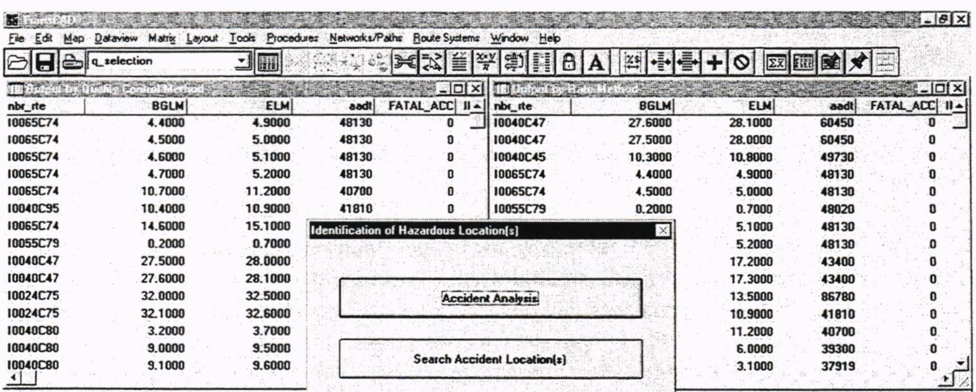

Figure 5 Outputs from Accident Analysis

6.2 Searching Accident Locations

This type of application allows the user to determine the accident location(s) based on road name, or county name, or accident severity type.

This section presents only determination of accident locations by road name. By selecting the Search By Road Name push button from *Search Accident Location(s)* dialog box, the *Input Road Name* dialog box is displayed. This dialog box contains a drop-down list of road names that allow the user to select the road name. The accident(s) that occurred on the selected location will be displayed automatically on the screen including map and dataview when the user clicks the "Show" push button. The accidents are displayed on the selected roads and are shown on a map (in accident locations layer) and dataview table, as shown in Figure 6.

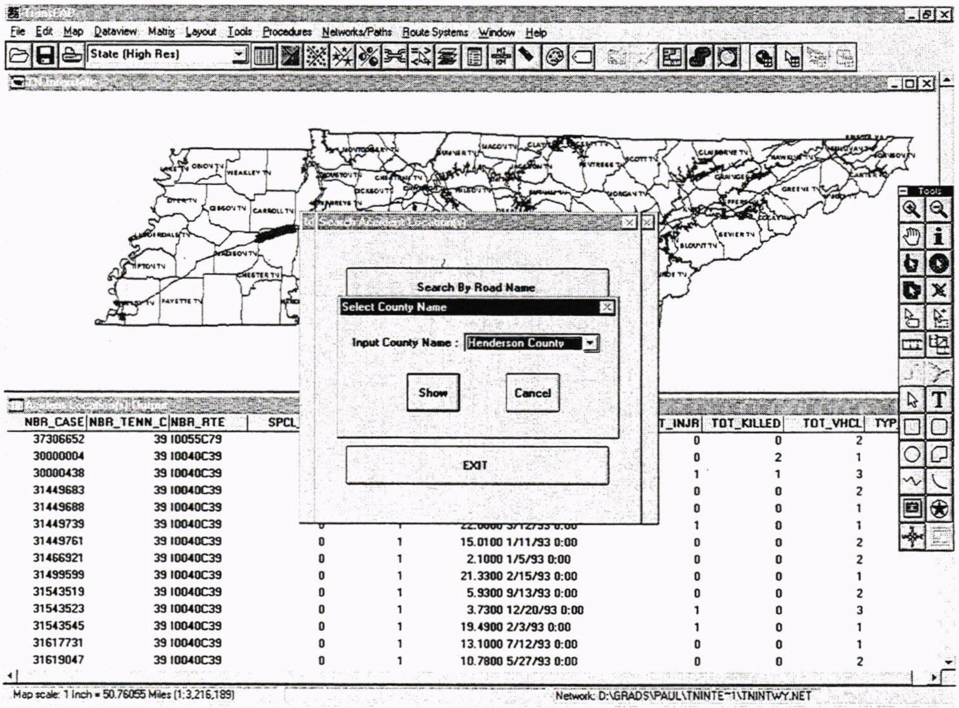

Figure 6 Outputs of Searching Accident Locations by County Name

6.3 Searching Hazardous Highway Locations

This type of visually enhanced application allows the user to determine roadway segment (s) that need to be improved by using an individual method or combination of any methods to identify hazardous highway locations. This section presents only the case of searching hazardous locations based on a combination of all four methods.

From the *Identification of Hazardous Location(s)* dialog box, the *Search Hazardous Location(s)* push button is selected. The *Search Hazardous Location(s)* dialog box is then opened. The *Search Hazardous Location(s)* dialog box contains push buttons labeled

Based On Each Method, Based On Combined Selecting Methods, Based On Combined All Methods, and Exit. By selecting the Based On Combined All Methods push button from *Search Hazardous Location(s)* dialog box, the *Search Hazardous Location(s) Based On Combined All Methods* dialog box is displayed. This dialog box contains an edit item which allows the user to input the top rank of segments. The user enters the top rank of the roadway segments, and then clicks the OK push button. Then the hazardous highway locations are displayed automatically on screen as shown in Figure 7.

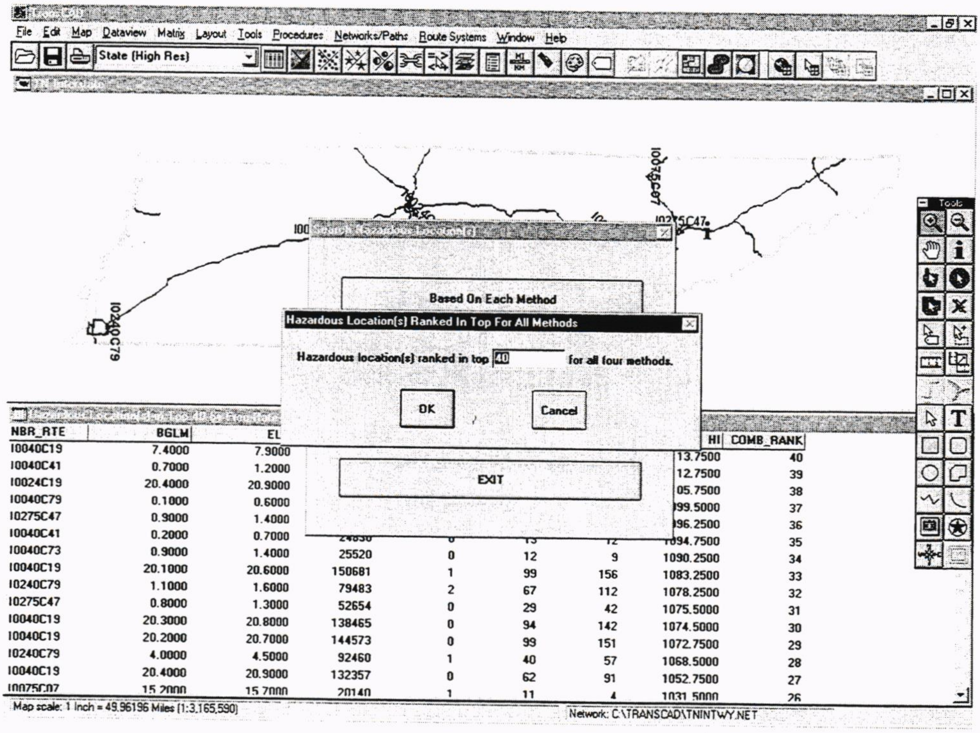

Figure 7 Hazardous Highway Locations

7. CONCLUSIONS AND RECOMMENDATIONS

The prototype system applications using various analytical methods in this paper was developed to identify hazardous highway locations more comprehensively. The system was developed with flexibility in mind so that the user can identify hazardous highway locations by each method alone or by combining two or more methods together. Additional advantages of the system are the system's capabilities and the ease of use.

For the system capabilities, the implemented system can analyze many accident data (about 50,000 accident records) for the entire state or larger area. This ability makes the system reliable and greatly enhances the quality of the results. Furthermore, the floating segment technique was an integral aspect of this research. Using floating segment length can avoid some problems associated with using a fixed segment. Problems with fixed segment analyses arise when a hazard exists near the boundary of two spots or segments. In this case, some accidents would be reported on one side of the segment boundary and others on the adjacent spot. Thus neither of the two spots would be identified as hazardous, and the high accident location would remain undetected.

However, there are some limitations of the current system and future improvements would be beneficial. Suggested enhancements in certain areas are as follows:

- Computer memory, the number of accident records, and running time of the program are important parameters. The degree of accuracy in identifying hazardous locations and the required running time of the analysis program are based on the number of accident records and the computer memory storing those accident data records. A total of approximately 31,000 accident data records were analyzed. The system spent approximately 8 minutes to run the program. The reason for this is because the program must read the input data, calculate the total segments (about 16,100 segments), and compute accident frequency, rate, severity index, R/Rc ratio, and hazardous index for each segment. The program requires about eight minutes to make necessary calculations and then create the output files. Although not instantaneous, nor would this be expected for such a substantial database, eight minutes may be viewed as too long of an analysis period by some. The computer in which the program was installed has a 166 MHz Pentium processor with 32 Mbytes memory. The running time of the program could be reduced if using a computer that has a higher speed and more memory.
- The format of input data is also critical. The system was developed to accept only the input data in tab delimited text format with the specific format of data attributes and then produce the output files with a comma delimited text format after running the analysis program. The correct format of accident input data must be used, or otherwise the system will not run the program and a message showing the errors of data input will display on the screen.
- Street network and routes are also important. The street network and routes must be initially developed by the user. The system cannot create them automatically. The user must import the street geographic files into the system and create the network file based on the geographic files. The street routes have to be created based on the network and use the routing application in TransCAD to develop the routes. For more details, please refer back to the sections detailing the creation of networks and roadway routes in Sections 5.1 and 5.2.

While this paper has demonstrated a technique for identifying hazardous highway locations and should contribute to improving highway safety through the use of more comprehensive hazardous highway location identification techniques, there is still much research left to perform. This section briefly describes some of the future research that can be envisioned.

The accuracy of accident information, particularly with respect to pinpointing the exact accident location, must be increased in order to improve the value and accuracy of accident analysis systems. Analysis techniques cannot overcome poor data quality. The use of more appropriate reference methods, better field referencing techniques by investigating police personnel, and closer interaction and more communication between safety engineers and police personnel can result in more accurate accident information. Recent advancements in global positioning satellite (GPS) techniques may well cause this increased accuracy to occur.

After hazardous highway locations have been identified, these hazardous locations should be reviewed in order to help in the selection of accident countermeasures. A hazardous highway features inventory can provide information on potential accident problems. The final decision to proceed with selection of a selected improvement may be based on the expected cost of the project, the available budget, and the desirability of the project. Economic analysis or cost-effectiveness techniques should be used in selecting the countermeasures with the greatest expected accident savings per unit cost. Priority programming should be conducted to select the projects that will provide the greatest safety benefits within the available fiscal constraints. Although the "best" technique for objectively assigning priorities is very much still debatable, improvement needs exceeding available funding is not very debatable because this is typically universally true.

REFERENCES

C. Zegeer (1982) Highway Accident Analysis Systems, **National Cooperative Highway Research Program**, Synthesis of Highway Practice 91, July 1982, 29-40.

F. Gharaybeh (1991) Identification of Accident-Prone Locations in Greater Amman, **Transportation Research Record No. 1318**, 70-74.

GISDK: Geographic Information System Developer's Kit (1995) Programmer's Guide Version 3.0, Caliper Corporation, USA.

K. Kim, N. Levine, and L. Nitz (1995) Development of a Prototype Traffic Safety Geographic Information System, **Transportation Research Record No. 1477**.

R. Filian, and J. Higelin (1995) Traffic Engineering in a GIS Environment: Highlighting Progress of The County of Riverside Geographic Information System Based Accident Records System (GIS-BARS), **15th Annual ESRI User Conference Proceedings**, California.

TransCAD: Transportation GIS Software (1996) User's Guide Version 3.0 for Windows, Caliper Corporation, USA.