HISAM: An Accident Data Base Manager

DAVID L. HARKEY AND RAFAEL RUIZ

The Highway Safety Analysis and Monitoring software was developed under a Federal Highway Administration research contract and is designed to aid local agencies with data base development and accident analysis. The package of programs is designed to enter, retrieve, process, and analyze traffic accident report data, link description data, and node description data. The city of Charlotte, North Carolina, with a population of 350,000, was chosen as the test site for the developed softwware. The average 20,000 accidents per year provided an excellent data base with which the software could be tested. By the end of the test phase the city was able to identify high-accident locations based on accident frequencies, accident rates, or EPDO indexes and rates. This paper describes both the data input and the report output of HISAM, as well as the city's experience with the software.

The extent and sophistication of efforts associated with improving highway safety vary considerably with the resources (i.e., equipment, manpower, budget, computer facilities) available to an agency. Computer facilities are particularly important in the process because most agencies deal with thousands of accidents, extensive roadway networks with varying features, and widely differing traffic conditions. This means a considerable amount of information must be processed as part of highway safety improvement efforts.

The ultimate effectiveness of a safety program depends on the availability of comprehensive and integrated data bases encompassing accident, traffic, and highway data elements. The capability to merge this information is critical to effective safety analysis. For example, the merging of accident data and volume counts permits the ranking of sites on the basis of accident rates. The review of highway features data provides a means to compare similar locations such as right-turn-on-red intersections, high-speed roadways, and narrow bridges. Traffic and highway data are also valuable to the detailed investigation of safety deficiencies at specific locations.

The Highway Safety Analysis and Monitoring (HISAM) software is designed to aid local agencies with data base development and accident analysis. The package of programs is designed to enter, retrieve, process, and analyze traffic accident report data, link data, and node data (*1*, *2*).

The development of the HISAM software is oriented toward the following users:

- City or county engineers who may have little or no formal highway safety background yet require a data management tool to conduct accident analyses, and
- City police officers who are familiar with traffic enforcement but may have little knowledge about highway safety analysis.

Therefore, the package

- Requires minimum field data,
- Involves very little human decision,
- Involves simple input and output,
- Does not require a mainframe computer,
- Does not require the user to have a computer programming background, and
- Is a user-friendly system.

HISAM is designed to run on the IBM PC, PC-XT, PC-AT, and 100 percent IBM-compatible microcomputers. The system must have the following:

- DOS, version 2.0 or higher;
- 640K of main memory;
- A 5.25-in. floppy disk drive;
- A hard-disk drive with a minimum of 10MB;
- A monochrome or color monitor; and
- A printer that is compatible with the computer and operating system just listed.

FUNCTIONAL DESIGN

Some of the features that allow the software to be used by many agencies follow.

- The system is modular in design to allow additional future routines to increase program capabilities,
- Programs are menu driven to facilitate their use,
- Programs allow for the integration of accident and inventory data bases,
- Data entry programs have standardized formatted screens to facilitate the data input process,
- Data entry programs have internal validity checks for alphanumerical characters of all data fields,
- Programs provide error messages and interpretation information to facilitate use of the system, and
- Complete documentation provides the user with instructions to facilitate system use.

It is particularly important to provide linkages between files for the integration of data. For example, the length of a link and the average annual daily traffic volume from the link data base can be combined with the number of accidents on the

D. L. Harkey, Analysis Group, Inc., Charlotte, N.C. 28202. R. Ruiz, City of Charlotte, Department of Transportation, Charlotte, N.C. 28202.

link from the accident data base to determine the accident rate for the link. This feature is currently limited or nonexistent in available microcomputer software.

HISAM is made up of five modules (the main program module, data base module, analysis module, system utilities module, and system information module), as shown in Figure 1. The main module serves as the primary operating system for HISAM and links the other modules together. It is entered each time the system is initiated and provides a means for the user to interact with the various subsystems included in HISAM. This is done through use of the main menu, which provides the user with options, each corresponding to the four system modules. The data base module of HISAM is used to store, view, modify, and remove data from the data bases. Three files are incorporated into the data base module: an accident report file, a link description file, and a node description file.

The analysis module contains several programs to perform a number of analyses and produce reports that are useful in highway safety management. Among the reports generated are high-accident location reports, accident rate reports, and equivalent property damage only (EPDO) reports.

The system utilities module contains programs to merge data base files that may have been entered on separate computers. This module also allows the user to reindex files that may have been damaged as a result of operating errors, such as turning off the computer during data entry.

The system information module of HISAM is accessed to determine the amount of space available on the hard disk and the number of records stored in the HISAM program.

DATA BASE CAPABILITIES

The HISAM data bases are structured to allow for the effective management and monitoring of collected data as well as integration between data files. The three HISAM data bases—accident report, link description, and node description—are menu driven for easy access by the user. The menu for the link description data base is shown in Figure 2. The functional capabilities of each data base (add, view, modify, and remove) allow for the efficient storage and retrieval of data while they minimize the chance of operation errors.

The accident report file is used to maintain the accident record system. Each accident that occurs within the specified system of links and nodes is recorded under a separate report number. Figure 3 shows the HISAM accident data entry screen with the variables that can be entered for each accident.

The format for the link and node description data bases is similar to that shown in Figure 3. Contained in these link and node description files is information describing the physical and operational characteristics of the system. Some of these variables include

- Link and node location codes;
- Street name;
- Length of link;
- Highway type (divided, undivided, etc.);
- Administrative class (state, federal, etc.);
- Number of lanes;
- Speed limit;
- Pavement type;
- Parking;
- Roadway width;
- Curb, median, and shoulder characteristics; and
- Traffic volumes (used for accident rate calculations).

ANALYSIS CAPABILITIES

The analysis module contains programs that generate reports used in highway safety analysis. They are as follows:

- Link Accident Location Report,
- Node Accident Location Report,
- Total Accident Frequency Report,
- Accident Frequency by Accident Type Report,

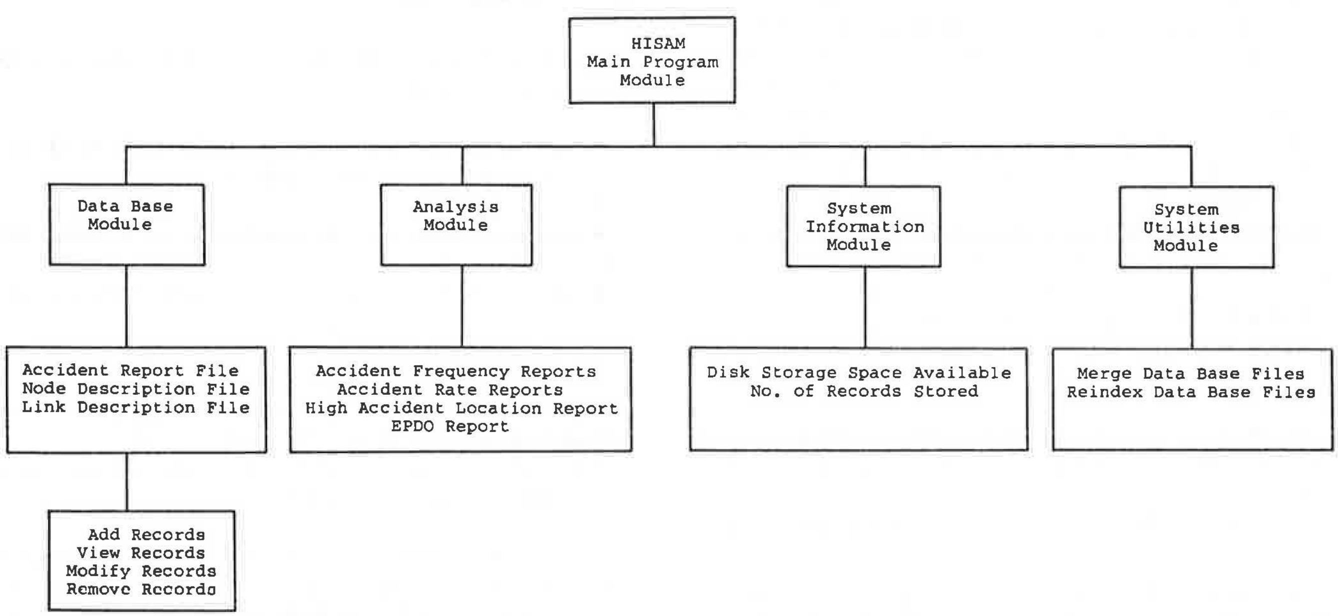

FIGURE 1 HISAM fundamental structure.

```
L I N K   D E S C R I P T I O N   D A T A B A S E
DATA BASE MAINTENANCE MENU

F1: Add Link Description Reports to Data Base
F2: View an Existing Link Description Report
F3: Modify an Existing Link Description Report
F4: Remove an Existing Link Description Report
F5: Link Data Base Information
F9: Return to Master Menu

Press Desired Key:
```

FIGURE 2 Link description data base menu.

```
Move Forward: Tab    | Move Backward: Ctrl-E     | Save: Ctrl-Z Aftr Rsp Rq Fs
Clear Field: Ctrl-Y  | Clear to left: BackSpace  | Clear to right: Ctrl-G
A C C I D E N T   D A T A   E N T R Y    Report No.: ______
Accident Location:      Location Code:        Reference Code:
Accident Date:          Day of Week:          Time of Accident:
Distance:               Accident Type:        Total Prsns Invl:
No. Injured:            No. Killed:           Accident Severity:

             Vehicle No. 1         Vehicle No. 2         Vehicle No. 3
          Driver   Passenger    Driver   Passenger    Driver   Passenger

Inj. Class |
Belt Use   |

           | Dr   Dr   Veh   Veh/Ped  Drink  Travel  Veh  Driv   Vio
           | Age  Sex  Type  Manvr    Cond   Speed   Dir  Fault  Ind

Vehicle #1 |
Vehicle #2 |
Vehicle #3 |

Rd.Char. :    Surface Cond:    Light Cond:    Weather Cond:

Enter a number with leading zeros if any.
```

FIGURE 3 Accident data entry screen.

- Missing Location Codes Report,
- Link Accident Rate Report,
- Node Accident Rate Report,
- Equivalent Property Damage Only (EPDO) Report, and
- Accident Report List.

The Link Accident Location Report is used to determine the distances (in feet) at which accidents occurred along a given highway segment during a specified time period. The Node Accident Location Report is used to determine the distances at which accidents occurred within varying radii from a particular intersection during a specified time period. The Total Accident Frequency Report is used to rank the links and/or nodes in descending order of number of accidents occurring at each location during the specified time period. The Accident Frequency by Accident Type Report ranks accidents in the same manner as the Accident Frequency Report but also lists the accidents by type for up to four user codes. A typical report of this type is shown in Figure 4.

The Missing Location Codes Report is used to determine which links and nodes have not been entered in their respective data bases but have been entered on at least one accident report. This aids the user in maintaining a complete data base at all times.

The Link and Node Accident Rate Reports rank locations in descending order of accident rates. The accident rate for links is calculated as follows:

$$R = \frac{N \times 1{,}000{,}000}{L \times \text{AADT} \times 365 \times n}$$

where

R = accident rate per million-vehicle miles (mvm),
N = number of accidents on the link during the year,
L = length of the link (miles),
AADT = annual average daily traffic on the link, and
n = number of years of accident data being considered.

The calculation for the node accident rate is as follows:

$$R = \frac{N \times 1{,}000{,}000}{\text{AAEDT} \times 365 \times n}$$

where

R = accident rate per million entering vehicles,
N = number of accidents at the node during the year,
AAEDT = total annual average daily entering volume at the node, and
n = number of years or accident data being considered.

A typical Link Accident Rate Report is shown in Figure 5. It is important to note that the accident rates for nodes and links cannot be directly compared because the calculations do not use the same variables; thus the rates do not have the same units.

In addition to the accident rate listing, the reports also list all links or nodes for which no volume is found in the data base. This helps the user to identify which links and nodes need volume counts.

The Equivalent Property Damage Only (EPDO) Report calculates the EPDO index and the EPDO rate for accidents and ranks accident locations (links and nodes) on the basis of the index. A sample report is shown in Figure 6. The EPDO index for a given location is calculated as follows:

$$\text{EPDO index} = F(\text{C1}) + A(\text{C2}) + B(\text{C3}) + C(\text{C4})$$
$$= \text{PDO}$$

where

F = number of fatality accidents,
A = number of Class A accidents
B = number of Class B accidents,
C = number of Class C accidents,
PDO = number of property-damage-only accidents, and
C1, C2, C3, C4 = constants by which the accident totals are multiplied (input by the user).

The EPDO rate is calculated as follows:

$$R = \frac{\text{EPDO index} \times 1{,}000{,}000}{\text{ADT} \times 365 \times n}$$

where

R = EPDO accident rate per million vehicles (links) or per million entering vehicles (nodes),

Accident Frequencies by Acc. Type From 010182 To 010187

Location Code	Frequency	22 Type	08 Type	11 Type	06 Type	Other Type
36D09	10	2	4	2	0	2
47A04	23	2	8	4	3	6
36D19	5	0	0	0	5	0
36D16	26	1	0	15	10	0
36D15	24	10	4	6	0	4

FIGURE 4 Accident frequency by accident type report.

Accident Frequencies and Rates From 010182 To 010187

Link Code	Frequency	AADT	Rate
29A0809	14	2500	4.824
36D0910	10	9000	3.754
28B0315	22	12000	2.890
286CA3370	16	100000	2.870
28D1112	11	20000	2.24
28C0237	9	32110	1.68

FIGURE 5 Link accident rate report.

Accident Equivalent Property Damage Only (EPDO) From 010182 To 010187

Fatality	Type A Injury	Type B Injury	Type C Injury	PDO Co.
9.50	3.50	3.50	3.50	1.00

Location Code	Fatal Accident	A-type Accident	B-type Accident	C-type Accident	PDO Accident	EPDO Index	EPDO Rate
36D09	0	0	3	14	29	88.50	4.82
36D09	0	0	3	14	29	88.50	3.68
47A04	0	0	2	9	20	58.50	6.42
47A04	0	0	2	9	20	58.50	7.89
46B09	0	0	5	8	10	55.50	2.60
46B09	0	0	5	8	10	55.50	3.20

FIGURE 6 Equivalent property damage only report.

ADT = annual average daily traffic (links) or annual average total entering volume (nodes), and
n = number of years of accident data being considered.

The final report, the Accident Report List, lists all accidents at a given link or node along with selected data on each accident (e.g., time and date of accident, severity, distance of the accident from the point of reference).

CHARLOTTE TESTING AND ENHANCEMENTS

Once the software was developed, a thorough test of its capabilities was conducted by the Charlotte Department of Transportation (CDOT). The city of Charlotte currently averages 20,000 accidents per year and, until this test of the HISAM software, manually handled all of the accident data and safety analyses. The system consisted of keeping three years' worth of accident data (approximately 60,000 accident reports) on file by intersection. This limited any analyses to a specific location. To monitor accident data or conduct safety analyses for an artery, thoroughfare section, planning area, or the whole city was impossible. The HISAM software not only allowed the city to create an interactive microcomputer accident data base but also increased city capabilities in accident analyses.

The 1985 accident data base, consisting of 23,522 accidents reported by the Charlotte Police Department, was used for testing the software. A total of 15,306 accidents were entered in the HISAM accident file, requiring a total storage of 1.98 megabytes. In addition, a street network was developed requiring data to be input for 4,800 nodes and 7,500 links.

CDOT was very satisfied with the data entry process and the amount of information that could be stored in the HISAM data base. However, the reports produced by HISAM were limited and actually used only 5 of the 28 variables entered in the accident report. Although these 5 variables (accident frequency, accident type, injury class, total entering volume, and average daily traffic volume) are used to produce the reports most commonly used in accident analysis, it was felt that output could be greatly enhanced by

- Expanding the data analysis and report capabilities using a proprietary data base manager,
- Creating a computerized accident location file interactive with the HISAM link and node files, and
- Generating computerized collision diagrams with a CAD software package using information from the HISAM accident file.

DATA BASE MANAGEMENT

Concerned citizens and communities are often interested in accidents and roadway safety and continuously contact the CDOT with questions about safe speeds, drunk driving, seat belt effectiveness, and more. Answering these questions requires a complete data base and the ability to analyze many variables. The first part, a complete data base, already existed within the HISAM accident file; however, the ability to analyze the data was limited.

The reports produced by HISAM are adequate for most accident analyses, but there are some limitations as a result of the Federal Highway Administration policy that does not allow the use of proprietary software in the development of any computer programs. Therefore, an "off-the-shelf" data base manager could not be incorporated into the HISAM software package by the contractor, Analysis Group, Inc. Thus the ability to search, sort, and analyze any variable or combination of variables was not achievable.

The CDOT, as part of its expansion of the HISAM software, incorporated a data base manager, dBASE-III Plus, with the HISAM accident files. This particular data base management package was chosen because the city already used this package; note, however, that any data base manager would work. To incorporate the two packages, a program was written by the CDOT in Turbo-Pascal (the language in which HISAM was developed) to convert the HISAM accident file to an ASCII text file. dBASE III-Plus was then used to read the ASCII file and convert it to a dBASE file. Because of the size of this file, it was split into two smaller files to speed the data analysis process. One file contained the driver and passenger data (seat belt usage, driver at fault, etc.) on each accident report, and the other contained the rest of the information on the report. As a result of the dBASE programming, detailed analyses using the accident data were obtained; these analyses included the generation of tables and charts describing the 1985 accident figures and trends.

Today, the CDOT intensively monitors on a regular basis its dBASE-generated accident files. Concepts for improvement of defined network sections or planning areas and benefit-cost analyses are periodically developed on the basis of their accident experience. Besides being interactive with community interest, the file has also been used to provide data analyses reports to other local departments, including the police department.

LOCATION DIAGRAM FILE

In the HISAM accident data entry, the location of an accident is defined by a location (link or node) and a reference code (nearest node). The city did not have a network coding system that would meet the HISAM requirements. Thus a new coding system was developed. The new system codes each pair of nodes with an alphanumeric five-digit code and codes each link with an alphanumeric nine-digit code using part of the code from each end-node. To ensure that the node codes were unique, a Lotus-123 file that included the node code and the crossing street names was developed.

A location diagram form was designed, and a manually kept accident location file was generated to aid in the coding process. The location diagram (Figure 7) describes an intersection by its code, the incident link codes, the adjacent intersection codes, and all corresponding street names. The node code is obtained by the concatenation of map area number, the alpha subarea reference, and the intersection number. The link code is obtained by the combination of the two end-node codes.

Although location listings were obtained from the Lotus-123 file sorted either by location codes or by street names, the entry of the location and reference codes in the HISAM accident data screen became slower as the location diagram file grew.

The CDOT designed and wrote a Turbo-Pascal program named NALDD to generate the location diagrams in a microcomputerized file. The screen generated by the program is the same as the location diagram in the designed form.

Initially, this new program required the entry of data already available in three different existing files: the Lotus-123 file described earlier, the HISAM node file that includes the node code and the codes of the incident links, and the HISAM link file that includes the link code, the street name, and the codes of the end-nodes.

To avoid the reentry of available data, the design of the program included the merging of these three files. Once these files were merged, the resulting location diagram file was integrated with the HISAM software. An option was added to the HISAM main menu to update the location diagram

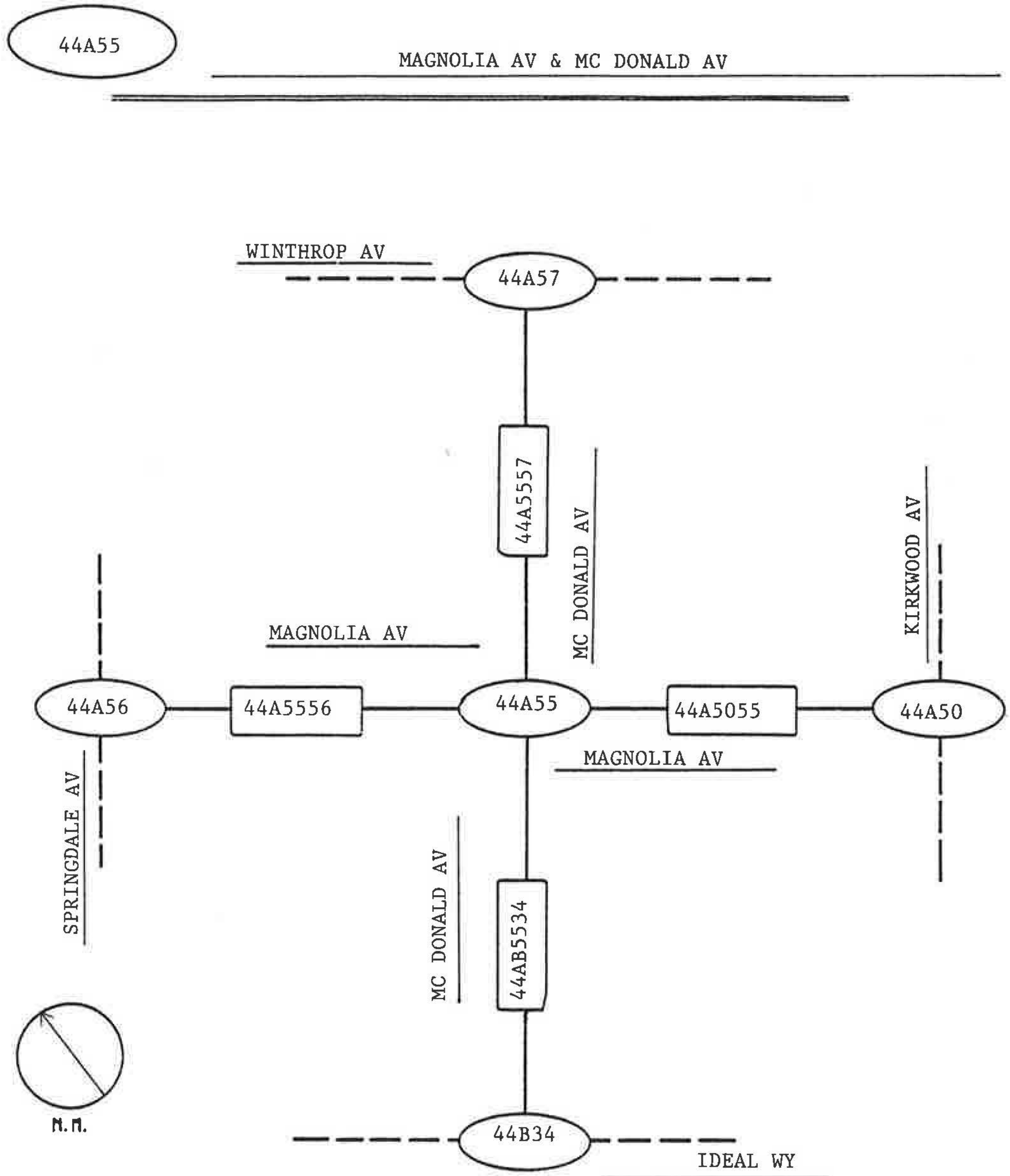

FIGURE 7 Link and node location diagram.

file. From this option, a submenu, including the options to add, modify, view, and delete a location diagram, can be displayed on the monitor system.

The design also incorporated the feature of viewing the location diagram when the "add an accident report" option is selected from the HISAM accident data menu. If the user wants to view the diagram on the monitor screen, this location can be retrieved by the node code or by the names of the crossing streets at the closest intersection. The accident number, location code, and reference codes are displayed on the screen with the location diagram. These data, once entered, are transferred to the HISAM accident data screen where the user can continue and complete the accident data entry.

The creation of the location diagram file and its integration in the HISAM software package have obviously increased the speed of accident data entry and technically improved the system. The location diagram file is also used for planning by other divisions of the CDOT.

COMPUTERIZED COLLISION DIAGRAM

Manually drawn collision diagrams have added a lengthy step to the data analysis processes. The CDOT is currently working to computerize the collision diagram process.

Because the HISAM accident file includes the direction of the vehicles prior to the accident and the vehicle maneuvers, and the location diagram file shows the geographic orientation, an accident collision diagram could be drawn from the available data using a basic CAD software package.

After considering available software packages, the CDOT purchased Prodesign II. A base map form was designed for all collision diagrams and includes a description of the graphic representation for the accident types, travel speed, weather conditions, driver at fault, and the date. All these data can be retrieved from the HISAM accident file and the location diagram files.

With the development of the base map, the CDOT is now working to create a memory bank or template defining the different accident types in all possible directions and maneuvers.

CONCLUSION

Overall, the HISAM software package enables a local municipality to develop and maintain an accident data base and to conduct accident analyses. Detailed information about accidents, links, and nodes can be entered, stored, viewed, and modified. From this information, high-accident locations may be determined based on accident rates, accident frequencies, or EPDO indexes and rates.

The software is complete as packaged and does not require any other software for its operation. It is now available through McTrans Center at the University of Florida. For those cities, counties, and regions interested in conducting more detailed analyses than can be provided by HISAM, the software provides an excellent starting point for an advanced data base and analysis system, as was demonstrated by the city of Charlotte. For more information on these packages, see the references that follow.

REFERENCES

1. D. L. Harkey, A. Hooshmandia, and K. Colpitts. *Computer Programs for Safety Analysis: HISAM Users Manual and Operators Guide*. FHWA/RD-87/073. FHWA, U.S. Department of Transportation, Washington, D.C., Jan. 1987.
2. D. L. Harkey, K. Colpitts, and H. D. Robertson. *Computer Programs for Safety Analysis: Techshare Report*. FHWA/TS-87-211. FHWA, U.S. Department of Transportation, Washington, D.C., Jan. 1987.

Publication of this paper sponsored by Committee on Traffic Records and Accident Analysis.