Development of a Prototype Traffic Safety Geographic Information System

KARL KIM, NED LEVINE, AND LAWRENCE NITZ

A prototype geographic information system (GIS) for the analysis of motor vehicle collisions in Honolulu, Hawaii, is described. An overview of GIS hardware and software is provided along with criteria utilized in the development of the Hawaii system. Mapping and spatial data sources relevant to traffic safety are described and evaluated. The usefulness of Census Bureau Topologically Integrated Geographic Encoding and Referencing (TIGER) files for spatial traffic safety analyses is highlighted. Spatial analyses and potential applications of this technology in traffic safety are also outlined. Recommendations for enhancing the uses of GIS in traffic safety are offered.

In recent years there have been numerous developments in geographic information system (GIS) technologies. Not only is the technology well established but it is here to stay (*1–3*). The emergence of desktop GIS packages has helped to make the technology more widespread, affordable, and accessible. Yet, surprisingly, there have been relatively few accounts of traffic safety programs that use this technology. One exception is a demonstration project developed in Haifa, Israel (*4*). Although GISs are used by no fewer than 29 state departments of transportation, many use these systems solely for engineering drafting and design, and only a few use them for mapping and geographic analysis (*5*). Several developments suggest that this will change. The passage of the Intermodal Surface Transportation Efficiency Act (ISTEA) in 1991 mandates the development of various information systems that could benefit directly from GIS technologies. The federal government has been studying the applicability of GIS technologies in transportation planning for a number of years (*6*). Many states have invested heavily in GIS hardware and software. The GIS has been routinely utilized in transportation planning. More universities are providing training for a cadre of GIS technicians and spatial analysts. GIS technologies have become quite widespread in numerous federal agencies and private-sector organizations (*5*).

In this paper the development of a GIS and spatial analysis system for motor vehicle crashes is described. The system was developed as a tool for research problem identification, and policy formation. In 1992 Hawaii was one of seven states selected by NHTSA to develop a crash outcome data evaluation system (CODES) project. In addition to building data bases containing information on motor vehicle collisions and various health and economic outcomes, the CODES sites were to conduct various analyses on the effectiveness of seatbelts, motorcycle helmets, and other traffic safety interventions. Hawaii's effort included the development of a prototype GIS. Different types of hardware and software as well as sources of spatial data were considered. Geocoding procedures specific to the analysis of collisions were developed.

University of Hawaii at Manoa, Honolulu, Hawaii 96822.

Here various applications of GIS technology in traffic safety are suggested. Recommendations for improving the use of GIS technology are provided in the concluding section.

BACKGROUND OF PROTOTYPE GIS

The prototype traffic safety GIS was developed for Honolulu, Hawaii. Located on the island of Oahu, the city and county of Honolulu consists of a single metropolitan area. With a 1990 population of more than 840,000, Honolulu is one of the largest metropolitan areas in the United States. Honolulu has the largest concentration of population in Hawaii, accounting in 1990 for 75 percent of the state's population and 74 percent of all motor vehicle collisions. In that year there were 19,598 crashes in Honolulu, of which 74 involved fatalities and another 6,733 involved injuries. Problems associated with police crash reports have been widely documented (*7*). In Hawaii, however, there is reason to believe that the quality of police reports is better than in other places. One crash report form is used statewide. In Honolulu all the data are collected by one police department, which uses the standard crash form. Special accident investigators are utilized for collisions involving a fatality or a serious injury. Moreover, the state of Hawaii as well as the city and county of Honolulu have been actively involved in the development and implementation of GIS technology (*8*).

OVERVIEW OF HARDWARE AND SOFTWARE

The GIS is a collection of hardware and software for entering, managing, retrieving, analyzing, and displaying spatial data. Typical hardware includes various input devices (digitizers, scanners), storage devices (disk drives, CD-ROM, tape drives), various processing configurations depending on the operating environment, and a range of different output devices (CRT displays, printers, plotters, and image projection devices). It is often convenient to think of two types of information that can be captured and manipulated with a GIS: graphic and nongraphic, or "attribute," data. The graphic information consists primarily of various elements, such as points, segments or lines, and zones or polygons, that are used in mapping. Attribute data are spatially referenced to the various graphic features and are typically captured and manipulated in a record format by some type of data-base management program and topological algorithm. Together, graphic information and attribute data are especially powerful in traffic safety research. For example, points could be used to identify the locations of events such as head-on collisions or perhaps collisions with blunt-end guardrails. Segments could be used to reference different roadways and the frequency of collisions during peak or off-peak times, or perhaps before and after the installation of lane markers. Zones or polygons can be used to describe or analyze the patterns of crashes that occur within various neighborhoods or travel zones and to relate these patterns to population or employment characteristics.

A GIS should provide the following: (a) information about the location of objects in space (geographical data), (b) information about what occurs at those locations (attribute data), (c) an analysis system for querying features from both the geographical data and the attribute data, and (d) a means for retrieving the information on a display terminal, a printer or plotter, or an external file. Many different GIS programs are available from commercial vendors as well as government sources.

CHOICE OF GIS PROGRAM

GIS programs can be classified as either vector or raster systems. In vector systems, geographical objects are defined by combinations of points, lines (or arcs), and polygons (or areas). Vector programs store the coordinates of the points, lines, and polygons and are programmed to draw lines between the stored coordinates. Attribute data are stored separately but are connected through a data-base management program. Vector systems are efficient in terms of storage and are useful for spatial phenomena that can be meaningfully linked to the basic geometric primitives. For example, motor vehicle crashes are easily linked to a vector system. The crash locations are defined by points, and the streets are defined by lines. Neighborhoods or areas are represented by polygons.

A raster system divides the geographical area into a uniform grid structure. Each cell in the grid is independently coded with regard to an attribute. The geographical and attribute data are intrinsically related. However, every variable that is mapped has to be stored as a separate map layer, although all are referenced to the same grid system. Spatial operations are carried out between the corresponding cells of each layer. There are advantages and disadvantages to raster systems. For spatial phenomena that are continuous and varying (e.g., topography, land uses, soil types), a raster representation is more useful. Operations between layers are more easily conducted for raster than for vector systems; the operations are applied only to individual cells. Raster systems can become very large because the geographical characteristics have to be stored for every data layer. Unlike in a vector system, in which one geographical file can be linked to multiple variables, each data layer represents a separate geographical and attribute file. Storage and processing speed requirements can be demanding with such a system. The major raster GIS programs utilize workstation or mainframe computer technology, for example, Erdas, GRASS, and Vicar. An exception is Idrisi, which runs on a personal computer. Consequently, for a GIS representing motor vehicle crashes a vector system may be more practical for agencies just starting out with this technology.

Another way of distinguishing among GIS programs is to consider whether they are designed to run either on a personal computer or in a workstation or larger computer environment. This distinction between desktop GIS systems and larger GIS programs is more than just a semantic one. Desktop GIS programs are limited by the memory capacity and processing speeds of the personal computers. The complexity and size of the geographic and attribute files are limited. Although large files can be analyzed with a desktop system, there is a price in terms of processing speed. With workstation-based GIS programs, size, complexity, and speed are less important issues, though there are, of course, limits to these systems as well.

There are some other important distinctions between desktop and larger systems. With desktop programs the underlying data structures are much simpler. Most desktop programs, such as ATLAS*GIS, GIS PLUS, and MAPINFO, utilize a simple data structure called a whole polygon structure. With this data syntax, each geographical object (points, segments, polygons) exists and can be manipulated independently of other objects. Changes to the objects (e.g., the addition of a new street) are implemented independently of the existing objects. For a desktop system this is an advantage in that fairly large files can be handled quickly. Most larger systems, however, such as ArcInfo and Intergraph, use a hierarchical data structure in which polygons refer to previously defined lines, which in turn refer to previously defined points. Any change to an object requires a reindexing of all related files in the system, because internal data integrity must be retained. This reindexing requires more data processing steps, more memory, and more processing time. On the other hand, a hierarchical data structure has the advantage of allowing meaningful links to be established among the different objects. For example, building a transportation network that would allow routing decisions to be made requires that all road links be connected and understood by the program; the program has to know that Street A in one part of the map is the same as Street A in another part of the map and also which street links are interconnected. Most of the desktop systems do not have such data structures; they cannot be used as transportation network programs.

The larger, more complex systems are more costly. Most desktop programs cost less than $2,000. Larger systems are frequently leased by the producers at a cost of $5,000 or more per year. Desktop GIS manufacturers can usually provide some data at minimal cost, whereas the manufacturers of larger systems usually do not (most will, however, contract with an agency to construct such files). There is also a user cost. The desktop systems are easy to learn. Typically, there are menus that guide the user through the steps. A traffic safety analyst can become familiar with these systems in a few weeks, sufficiently to be able to develop applications. The larger systems, on the other hand, are typically command driven and are more difficult to use. Because they are typically implemented in shared-user environments, there may be additional networking and system administration requirements. The experience in Hawaii has shown that it takes a user several years to become fully competent with a large system such as ArcInfo. Consequently, there may be a functional difference between a GIS specialist capable of using one of the larger systems and a traffic safety analyst who can use one of the desktop GIS packages.

The criteria used to develop a prototype system for Hawaii involved identifying the various functional capabilities that would be needed. Several key decisions were made early in the effort. First, it was decided that, given the modest resources dedicated to this initiative, existing digital files would be used to the greatest extent possible. The approach was to avoid having to digitize paper maps or to create new base maps from scratch. This also eliminated the necessity for expensive digitizing and scanning equipment. Second, the prototype system had to be working in a matter of months and needed to be functional in terms of both mapping capabilities and analysis. For this reason the simpler desktop packages had a certain appeal. Finally, the GIS work had to be integrated into other project activities. The GIS developed in Hawaii struck a balance between desktop and larger systems. The attribute data were stored and managed in a workstation environment; SAS was the principal software engine for data-base management and statistical analysis. However, the spatial data were manipulated and displayed on a desktop vector GIS using both ATLAS*GIS and MAPINFO.

The setup in Hawaii permits cross-platform and cross-software transfers, thereby taking advantage of the strengths of various packages and sidestepping some of the weaknesses. It is somewhat of an eclectic setup, bringing together conventional GIS packages as well as a number of different specialized tools and utilities. In addition to utilizing SAS for analysis, SpaceStat, a program for zonal analysis (*9*), was acquired. A special software program for analyzing point distributions, called Hawaii PointStat, was written by project staff (*10*).

BUILDING GEOGRAPHICAL FILES FOR REFERENCING CRASHES

A reference system for identifying the location of crashes needed to be constructed. The majority of motor vehicle collisions occur on roads. This necessitated a spatially referenced road file. The Bureau of the Census Topologically Integrated Geographic Encoding and Referencing (TIGER) System was selected (*11*). This is a comprehensive data base of line segments for the entire United States. Each record in the data base represents an individual line segment. An individual record is identified by the type of feature that it represents, for example, a street or a freeway or a railroad line or a pipeline. The TIGER dictionary gives the breakdown of the individual street segments in what is called a census feature classification code.

The data structure of TIGER is topological in that directionality and adjacency are built into the data. Each link has a direction, defined as a line occurring between a "from" node and a "to" node. The nodes are defined by their latitudes and longitudes. There is also a record that defines a shape grammar, which is a separate identifier giving up to 10 additional coordinates for the segment. Because a segment is defined as the line between two nodes, it is graphically drawn as a straight line. However, many segments are not straight but curved. The 1990 TIGER represented an integration of the basic street segment data base (previously called the DIME file) with the U.S. Geological Survey's 1:100,000 quadrangle maps. The scale of TIGER is set at 1:100,000.

A reference system for linking crashes to neighborhoods or small areas was also devised. For several reasons, census geography rather than transportation geography was utilized. First, the census geography is sensitive to residential population. Most trips are homebased, rather than workbased, so a geography relevant to most travel behavior seemed appropriate. Second, there is consistency between one census and another; the census geography is changed only when the population changes considerably. A crash analysis system based on census geography will have continuity for many years. Third, a large amount of data is available for the census geography, particularly from the decennial population and housing census. Fourth, many organizations use census geography and data. Therefore, there is consistency across organizations.

Different levels of census geography, for example, census tracts, block groups, and blocks, were considered. Block groups were selected as the primary spatial units of analysis. A block group is made up of 7 to 15 individual blocks, and a census tract is made up of 3 to 6 block groups. In theory, a census tract should have a population of 3,000 to 5,000, and a block group should have a population of approximately 1,000. In 1990, however, the census geography for Hawaii was too crudely differentiated. On Oahu, for example, for 1990 there were 199 census tracts, each with average of 4,500 persons. However, there were 363 block groups, each averaging slightly fewer than 2,500 persons. Using census tracts would produce too crude a zonal structure. Individual crash locations were aggregated to the block-group level.

Additional geographical files for use in analysis, modeling, or interpretation were generated. All levels of census geography—census tracts, block groups, blocks, and streets—for both the 1980 and the 1990 censuses were constructed. Each of these represents a separate data layer that had to be extracted.

The TIGER files follow the Census Bureau's policy of acknowledging political jurisdictions. In Hawaii this creates special problems because each census unit that borders the coastline actually extends 5 km (3 mi) into the ocean (because this is where the jurisdictional boundary ends). Such maps, although they are correct in terms of boundaries, are not appropriate for analysis of motor vehicle collisions. This problem was corrected by use of a Census Bureau convention that defines the ocean as an individual block and any beach as a separate block. Additional blocks are allocated to harbors, marinas, river outlets, and other features. Deleting the ocean, beach, harbor, and other features produced a reasonably familiar representation of Oahu. Figure 1 shows the block-group structure for Honolulu.

Compared with other mapping sources, such as the U.S. Geological Survey's digital line graph files, the TIGER files are more up to date and comprehensive, particularly for urban Honolulu. Other mapping sources such as U.S. Geological Survey quad maps, the highway inventory, and other spatial files were considered, but problems with referencing census data as well as concerns about scale, accuracy, and accessibility also arose. Because the TIGER files were used to create the base files, data from other sources, such as zone-to-zone trip tables and various trip attractors and generators used in the regional travel demand model, were easily integrated into the system. Data on the location of schools, bars, hotels, tourist sites, and rainfall levels were obtained and used in various analyses. Because all these files are referenced to the same coordinate system (latitude and longitudes in a certain projection), data from additional geographies can be referenced to the baseline block-group geography.

GEOCODING CRASH LOCATIONS

The original motor vehicle accident file did not contain geographic coordinates (latitude and longitude). Crashes were referenced by descriptive information such as the street name and the nearest cross-street name, mile marker, or street address number. Converting the descriptive information into geographic coordinates was done by a semiautomated process known as geocoding.

A standardized dictionary of street names was developed. Crash reports typically list accidents by the intersection nearest to where they occurred (e.g., "at King St. and University Ave."). The sites of freeway crashes are often identified by the nearest freeway exit or entrance (e.g., "H1 at the Vineyard Blvd. off ramp"). However, collisions with only mile marker references could not be automatically matched because milepost locations are not in the TIGER files and are otherwise not available in digital form. Further, other identifiers are often used, such as a street or a bridge crossing (but not actually intersecting) a freeway or the transition road between two freeways, tunnels, or access roads to the freeway. In addition, police reports often use local slang to refer to locations that are not technically identified on a map (e.g., referring to a parallel street as if it were

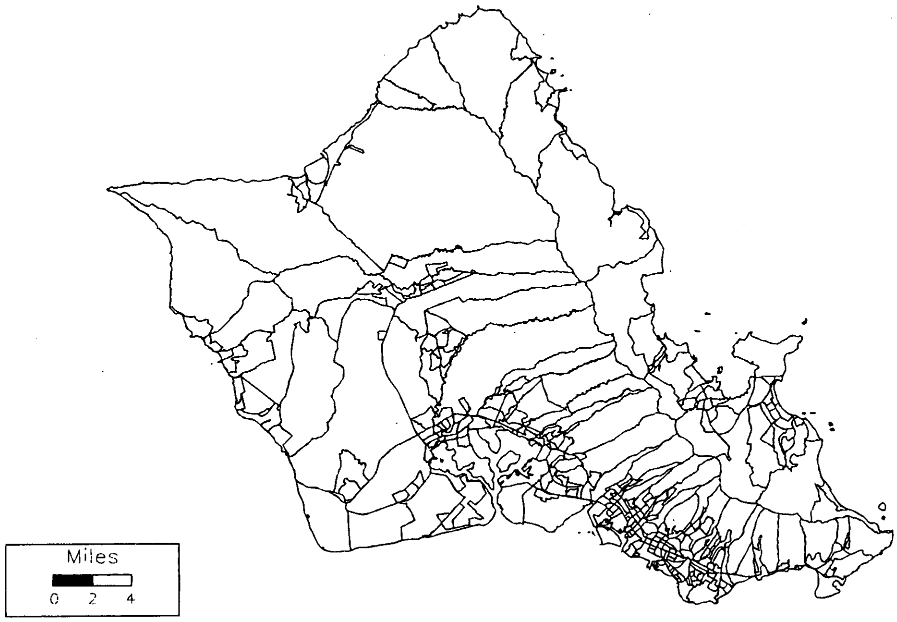

FIGURE 1 Oahu: 1990 block groups (from TIGER files).

intersecting a freeway segment when in fact it does not or referring to a tunnel by the tunnel name rather than by the road name). There are also the usual spelling and typographical errors that affect almost all data bases.

To georeference the crash locations, the 1990 and 1992 TIGER files were used as a basis. The locations were matched to the crash file with a probabilistic linkage software package known as Automatch (*12*). Another program by the same company that developed Automatch, called AutoStan, was used to standardize address references in the crash file for matching against TIGER references. Alternative spellings were included in the dictionary. Artificial intersections had to be created for those crashes for which the reported intersections were not actually intersections (e.g., when a bridge crossed a freeway or when an exit ramp on the freeway accessed a parallel road). Mile marker locations had to be manually processed from paper maps and entered into the dictionary. The process was iterative. It involved, first, a run through all crash locations that were in the dictionary. For those locations that were not identified, specific identifiers were inserted into the dictionary, and the search procedure was rerun. As the dictionary grew, successive codings were able to identify a higher proportion of all names.

Specialized software for matching street names was developed. For every intersection the program would take the standardized TIGER name and search through the street segments to select every link with the name. These would be placed into two matrices, A and B (for each of the two street names). For example, for a collision at King St. and University Ave., all TIGER segments with the name King St. and all TIGER segments with the name University Ave. were selected. Each TIGER segment has two nodes, with a specific longitude and latitude attached to each (a "from" node and a "to" node). The program would compare the longitude of each link in the A list with the longitude of each link in the B list for both nodes (i.e., four separate comparisons for each pair of links from the A list and the B list). For those pairs for which the longitudes in both the A and B lists were the same the program would then compare the latitudes of each pair. Only one node from a single pair would produce a match between the A list and the B list. The longitude and the latitude of the matched node would then be assigned as the crash location. Approximately 2 percent of the crashes had to be identified manually because not all streets in the TIGER system have street identifiers.

The entire process took several months. Approximately 19,213 of the 19,598 locations (or 98.0 percent) on Oahu were identified. For each geocoded crash location, latitude and longitude coordinates were assigned, as were the census tract, block group, and block number for accidents occurring on Oahu. The data were written back into the SAS crash database. Accuracy with this method is reasonable for Oahu. TIGER longitudes and latitudes up to four decimal places were used. The indeterminacy of the fourth decimal place is approximately 6 m (18 ft) [i.e., for latitude, 0.0001 represents approximately 11 m (37 ft); the error of measurement is half of this]. Approximately 43 percent of all collisions occurred at intersections; the error for these is at most a few meters. Of the remaining 57 percent that did not occur at intersections, most occurred on street segments and were assigned to the nearest intersection. For these the error is at most half a block. For freeway collisions, usually the nearest on or off ramp was selected; therefore, 0.02 to 0.4 m (1/8 to 1/4 mi) might be a typical error. Finally, for a few road segments, usually in mountainous or rural areas, assigning the collision to the nearest intersection might produce an error of up to 0.8 m (0.5 mi) or so; there were only a handful of these locations.

ATLAS*GIS and MAPINFO were used for mapping the data. Figure 2 shows the location of the 19,213 crashes on Oahu, each represented by a dot. Since many crashes occur at the same location, the dots are printed over one another. It is clear that crashes are highly concentrated in the built-up urban areas, as the crash pattern overlaps that of urban concentration. Figure 3 shows a detail of the central part of Honolulu, indicating how the crash locations are assigned to the intersections of streets.

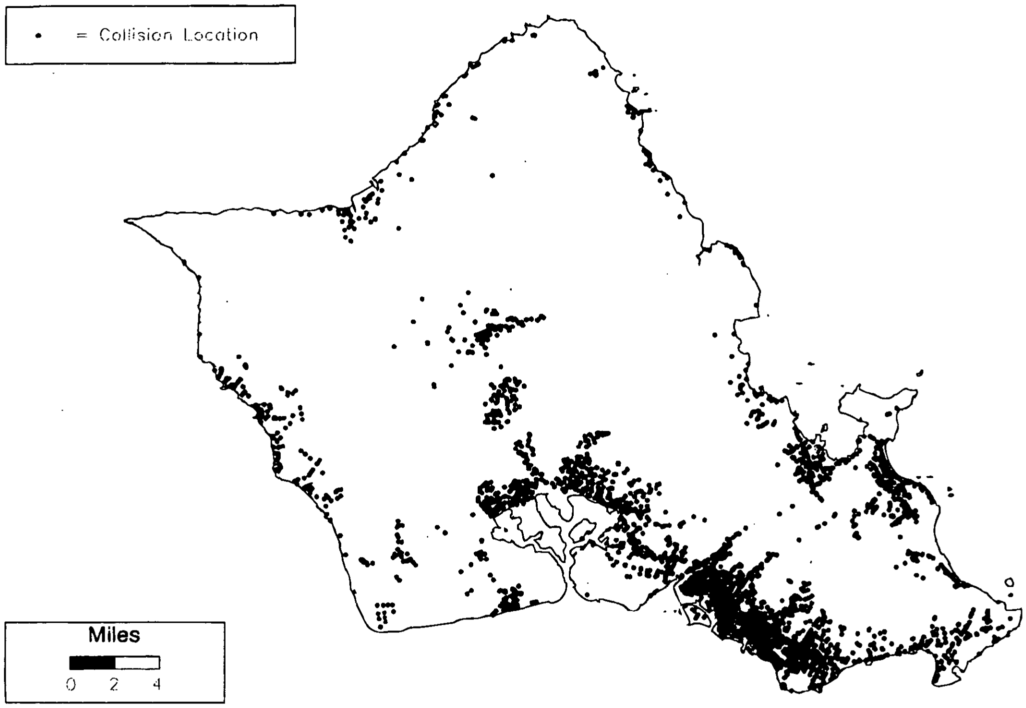

FIGURE 2 Oahu motor vehicle collisions, 1990: location of all collisions.

TRAFFIC SAFETY APPLICATIONS

Many different traffic safety applications were developed. More continue to evolve. The GIS can be used as a tool for description, analysis, and modeling as well as for problem exploration and identification. The information can be used in the formulation of appropriate safety programs.

Based on research facilitated by the GIS, it was determined that the spatial patterns of crashes vary according to time of day and day of week. Presumably these patterns reflect changes in traffic levels. Fatalities and serious injuries were found to have a different spatial pattern from that of noninjury accidents, most likely as a result of differences in speed, alcohol use, time of day (most fatalities occur at nighttime), and home-based activities. Different types of vehicles were found to have different spatial patterns in terms of collisions. The GIS also permitted the examination of relationships between collisions and population size, density, and manufacturing, retail trade, or other employment categories. Moreover, block groups with elementary and junior high schools were found to have a higher share of crashes than would be expected.

Work is also under way to analyze various roadways. Although the study is not yet complete, it appears that freeway ramps and access roads are particularly dangerous locations; block groups with ramps and access roads have much higher likelihoods of crashes when traffic volume and land use are held constant. Freeways by themselves are relatively safe, having much lower accident likelihoods than major arterials, mile for mile. Over time, as more information regarding the roadway inventory is added to the data base, other analyses can be conducted.

The traffic safety GIS has been used to study moped collisions in Honolulu. It helped determine that there are two distinct spatial patterns, one involving out-of-state residents and another involving Hawaii residents. GIS can be a useful tool in identifying special population groups and in locating locations of high incidence of other types of collisions as well. The power of GIS technology is that it lends itself to decision making in terms of spatial features. Certain points, segments, and areas stand out as places for intervention. Intersections (such as those near the University of Hawaii), roadways (Kalanianaole Highway), and districts (Waikiki and downtown) should be examined in more detail to determine the causes and types of traffic conflicts that involve mopeds. A better understanding of the root causes of these collisions could lead to improved policies for enhancing moped safety.

Other applications can be suggested. Areas that have high levels of crashes need to be investigated. There may be unique land uses or activities that generate disorderly traffic flows that, combined with poor signals, traffic routing, or signs, interact to produce dangerous traffic zones.

A traffic safety GIS could be especially useful to community groups in formulating meaningful safety plans. Having a GIS with a visual output can allow groups to develop a common understanding for managing traffic and safety on the neighborhood scale. The availability of maps and spatial data can help to focus discussions among traffic engineers, safety advocates, and concerned citizens.

There are many uses of such a system beyond that which has been described. Such a framework can provide a systematic spatial approach to crash prevention, allowing transportation planners, law enforcement agencies, and other agencies to focus on those areas that have excessively high levels of crashes (beyond that which would be expected). A spatial analysis framework can provide a basis for monitoring future interventions. It can provide a basis for integrating a wide range of information on roadways, vehicles, and drivers.

Another application allows particular crash characteristics to be examined in relation to more general patterns. For example, fatali-

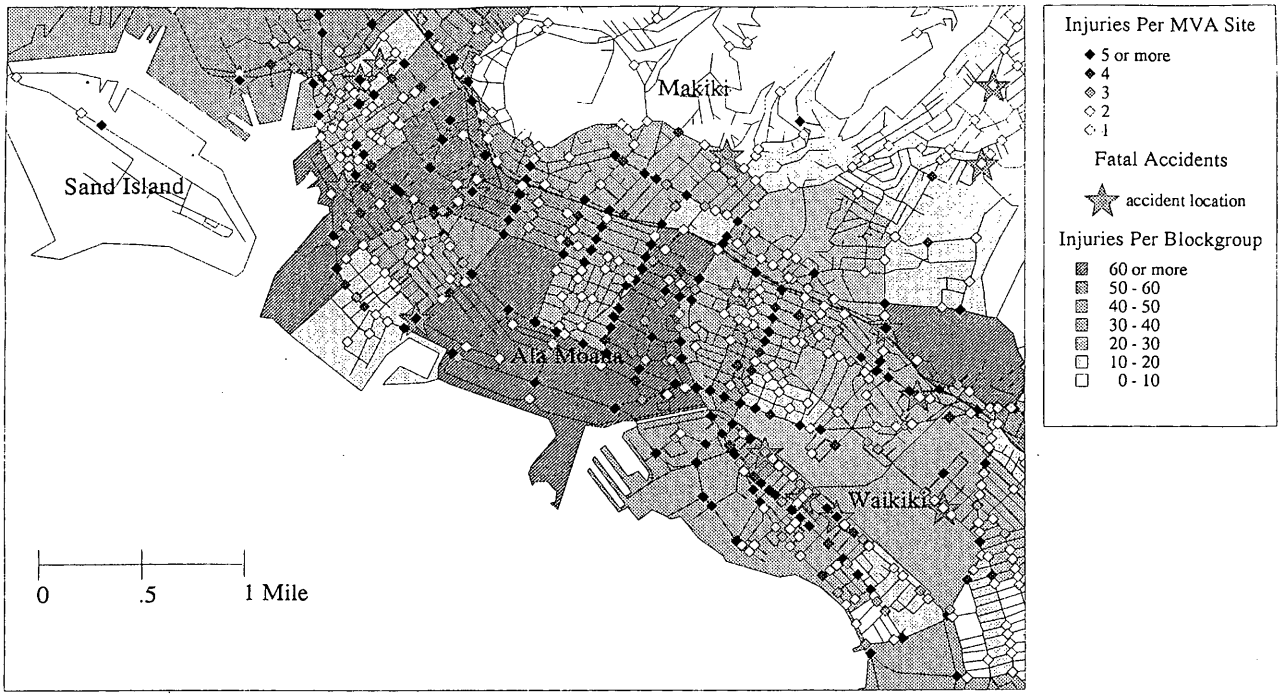

FIGURE 3 Injuries in central Honolulu, 1990 (1 mi = 1.6 km).

ties, alcohol-related crashes, and moped crashes in Hawaii have an essentially different spatial pattern from the majority of crashes. In the case of fatalities there is an interaction among travel speed, alcohol use, and home-based activities. Future research will examine the spatial correlations of crashes with census variables such as unemployment, divorce rates, and income levels. Using the spatial tools developed and utilized here, one can create distinct intervention strategies that target particular populations of vehicle users.

RECOMMENDATIONS FOR FUTURE DEVELOPMENTS

The system developed in Hawaii is only a prototype. GIS technologies have improved greatly in recent years, but there is still need for improvement in terms of making GISs easier to use and developing useful applications of the technology. The problems with a new technology such as the GIS go beyond hardware and software. There are also significant personnel and organizational challenges involving new responsibilities and new approaches to analysis, training, and information management.

The experience in Hawaii suggests that desktop GIS packages could become for traffic safety analysts what spreadsheet programs have become for accountants and others. Although there will always be a need for larger systems, desktop packages provide a powerful tool for research, problem identification, and policy analysis.

At the same time, the availability of desktop GIS packages will not solve all problems. A better system for geocoding crashes is needed, both at the scene of collisions and in terms of processing information from the crash forms. There have been developments integrating Global Positioning Systems and GIS (*6*). Improvements in police crash reporting, especially in terms of geographic referencing, is needed. Perhaps portable field computers can be adapted to contain standard street dictionaries and utilities for improving locational data.

The use of probabilistic matching software such as that utilized in the CODES project can also greatly enhance geocoding and matching of crash locations to existing geographic files (*13*). GIS technologies not only provide a means of satisfying ISTEA mandates regarding the development of various information systems but can also support management decisions regarding new construction, operations, maintenance, and expansion of transportation facilities on information organized, analyzed, and supported with GIS technologies.

There is a need for more geographic standards specific to traffic safety considerations. Although the TIGER files provide a good starting point, there is a need to consider other spatial files and how they can be integrated. Different state and federal agencies need to share and exchange information and procedures relating to the construction of base maps and geocoding of crashes and spatial analyses.

New partnerships need to be forged between federal and state agencies, between universities and departments of transportation, and between private industry and the public sector around the GIS technologies and applications. Above all, there is a need to devote more resources (human and financial) toward meaningful spatial analysis of traffic collisions. By using GIS technologies, perhaps a new generation of analyses can be developed, a situation that can only help to improve the understanding of highway safety.

REFERENCES

1. Maguire, D., et al. *Geographic Information Systems: Principles and Applications.* John Wiley and Sons, New York, 1991.
2. Star, J. *Geographic Information Systems: An Introduction.* Prentice-Hall, Englewood Cliffs, N.J., 1990.

3. Huxhold, W. *An Introduction to Urban Geographic Information Systems.* Oxford University Press, New York, 1991.
4. Peled, A., A. Hakker, and J. Affum. A PC Oriented GIS Application for Road Safety Improvements. *Traffic Engineering and Control*, Vol. 34, 1993, p.3.
5. Attenucci, J., et al. *Geographic Information Systems. A Guide to the Technology.* Van Nostrand Reinhold, New York, 1991, p. 43.
6. *New and Emerging Technologies for Improving Accident Data Collection.* Publication No. FHWA-RD-92-097. FHWA, U.S. Department of Transportation, March 1993.
7. O'Day, J. *NCHRP Synthesis of Highway Practice 192: Accident Data Quality.* TRB, National Research Council, Washington, D.C., 1993.
8. Kim, K. Geographic Systems in Hawaii. In *Geographic Information Systems: Developments and Applications* L. Worrall, ed., Belhaven Press, London, pp. 87–108.
9. Anselin, L., *SpaceStat: A Program for the Statistical Analysis of Spatial Data.* University of California, Santa Barbara, 1993.
10. Levine, N., et al. *Pointstat: A Statistical Program for Analyzing the Spatial Distribution of Point Locations*, Version 1.0. University of Hawaii, Honolulu, 1994.
11. *TIGER/Line Census Files, 1990, 1991.* Bureau of the Census, U.S. Department of Commerce, 1991.
12. Jaro, M. A. *AutoMatch: Generalized Record Linkage System*, Version 2.0. Data Star, Inc, Silver Spring, Md., 1992.
13. Kim, K., and L. Nitz. Application of Automated Records Linkage Software in Traffic Records Analysis. In *Transportation Research Record 1467*, TRB, National Research Council, Washington, D.C., 1994 pp. 50–55.

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