Knowledge-Based Geographic Information System for Safety Analysis at Rail-Highway Grade Crossings

SRIRAM PANCHANATHAN AND ARDESHIR FAGHRI

The development of a knowledge-based geographic information system for managing and analyzing safety-related information for rail-highway grade crossings is discussed in this paper. The allocation of federal funding for safety improvements at public, at-grade rail-highway crossings is made based on the performance of the states with respect to accident reduction. The motivation behind the work was to establish guidelines and to develop an integrated system that would ultimately result in accident reduction through better access and management of safety information. This was accomplished by using geographic information system (GIS) technology and decision support tools through integration of the GIS application with a statistical model and a knowledge-based expert system (KBES). The work continued an ongoing project that resulted in the integration of rail-highway grade crossing safety data from various sources, such as the FRA and Delaware Department of Transportation (DelDOT), into a data base management system. The selection and integration of the U.S. Department of Transportation (USDOT) accident prediction model into the system was also required. This paper describes the conversion of rail-highway grade crossing safety attribute data into a GIS-acceptable format, the development of the GIS application including the spatial analysis, visual display, and query capabilities, and the development of a KBES to account for site-specific and qualitative factors. The KBES is also capable of suggesting safety upgrade action(s) at the crossing. Interfacing of the KBES and the program for the USDOT model with the GIS and the framework of the complete package developed for safety analysis at rail-highway grade crossings are also discussed in the paper.

Transportation agencies are faced with an ever-increasing need for not only readily available information but also the capability to analyze, manipulate, access, and operate on that information on a broad basis to enable the basic functions of planning, safety, and operations management.

Developments in the field of information technology such as GIS, Expert Systems, and Database Management Systems are especially relevant to the transportation field because of the spatially distributed nature of transportation data, the use of engineering judgment, the often approximate and uncertain nature of transportation information, and the large amounts of data involved.

The application of GIS, in particular, has relevance to transportation due to the essentially spatially distributed nature of transportation-related data, and the need for various types of network-level analysis, statistical analysis, and spatial analysis and manipulation.

Most of the analysis of transportation data is related to quantifiable and statistically significant factors, and most of the models, routines, and algorithms process mathematically quantifiable factors only. The use of KBES in conjunction with mathematical models, algorithms, and procedures will allow for more complete analysis with the incorporation of engineering judgment and heuristic knowledge into the decision-making process and by accounting for qualitative and statistically insignificant, yet important, data.

This paper discusses the development of an integrated package for the management of rail-highway grade crossings safety information. The GIS application was developed using TransCAD (*1*), and the KBES was developed using CLIPS (*2*) shell, which provides a forward-chaining inference engine. The paper focuses on the development of the above-mentioned components, interfacing of the components, and the working mechanism of the integrated system.

RAIL-HIGHWAY GRADE CROSSING SAFETY PROGRAM

The State Transportation Assistance Acts of 1978 and 1982 mandate the provision of federal funding to states for safety improvement projects at rail-highway grade crossings. States are required to establish procedures to rank crossings and to use the rankings in an allocation process to achieve effective use of funds and the greatest possible accident reduction.

In the state of Delaware there are 548 rail-highway grade crossings, of which 265 are public. During the period from 1981 to 1991 there were 71 train-automobile accidents, 10 of which resulted in fatalities (*3*). The responsibility for identification and inventory of rail-highway grade crossings is currently assigned to the Delaware Department of Transportation (DelDOT). An earlier part of the ongoing study resulted in the identification of the most feasible of several existing empirical models and the development of a data base management system that includes all safety-related inventory information for public at-grade rail-highway crossings in Delaware (*3*).

Limitations in the Existing Rail-Highway Grade Crossing Program for Delaware

The existing rail-highway grade crossing safety program for Delaware consisted of a data base management system that maintained safety-related attribute information for all public, rail-highway grade crossings. The USDOT accident prediction model was chosen out of several nationally recognized empirical models for the prediction of accident hazard at a crossing to be implemented for the existing program (*4*). The USDOT model is documented in the work of Hitz and Cross (*5*). The data base management system

Department of Civil Engineering University of Delaware, Newark, Del. 19716

also included a computer program to execute the USDOT model on the data base and append an accident index value to it. The data base was, however, not in a location-referenced format that would enable it to be compatible for a GIS application. Additionally, the site-specific and qualitative information that may be used by the KBES to comprehensively evaluate safety and provide decision support was not available and was maintained in a written form (*3*).

The existing program included the ability to update, modify, and prioritize the data base for rail-highway grade crossings and to calculate the accident hazard index. However, this accident index does not present a true picture of accident hazard at any crossing as it does not account for several site-specific and qualitative factors such as sight distances, truck and hazardous materials-carrier traffic, land use in the crossing vicinity, and other factors. The program also does not have data analysis and manipulation capability and no indication of what action has to be taken at the crossing.

Lack of decision support and inability to consider heuristic and judgmental information were the main deficiencies of the program. The expert, when presented with a text file output of the data base, cannot have a good perception of the existing conditions at individual crossings as he/she has to look up a map every time he/she desires to relate the information to the geographic location of the crossing.

APPLICABILITY OF GIS TECHNOLOGY

Since the crossings safety attribute data is spatially distributed, this problem is suitable for GIS. The main advantage of using GIS is its ability to access and analyze spatially distributed data with respect to its actual spatial location overlaid on a base map of the area of coverage that allows analysis not possible with other data base management systems. The main benefit of using the GIS is not merely the user-friendly visual access and display, but also the spatial analysis capability and the ability to apply standard GIS functionalities such as thematic mapping, charting, network-level analysis, simultaneous access to several layers of data and overlayment of same, as well as the ability to interface with external programs and software for decision support, data management, and user-specific functions.

The existing data base does not allow the user to manipulate, access, and query the data base other than in a very limited way. The user is limited to textual queries only, and selection and viewing of crossings attribute data with respect to spatial and topological relationships is not possible. Other related data, such as land use, population, and the road network characteristics of the area in the crossings vicinity, cannot be accessed in the present data base. This ability of the GIS, along with the final presentation of results on a digital base map, will allow the user a better perception of the problem, enable better decisions, and allow a better understanding of what is to be achieved in a broader sense. The ability to define conditional queries, perform statistical analysis, create thematic maps, and provide charting enhances the crossings safety program by allowing for better understandability of data.

Furthermore, the ability of most GIS software to provide many basic transportation models and algorithms may also be useful in specific situations. The ability to link up to external procedures and software also provides flexibility, as these procedures can access data within the GIS and present the results of analysis to the GIS for viewing and analysis.

DEVELOPMENT OF GIS APPLICATION

Source of Graphical Data

In general, there exist several sources for spatial data from which the digital base map for the application can be created. The base map can be created by scanning or digitizing from hard copy maps. The other approach is to use map data from some primary source such as the Bureau of Census Geographic Base Files (GBF), the United States Geological Survey (USGS), Digital Line Graphs (DLG), and the United States Census TIGER files.

TIGER was found to be a suitable source for building the base map for this project primarily because of the level of detail available about the street centerlines and the rail lines and, also, because of the valuable attribute information tied into it. Since the main requirement of the application was for data retrieval, spatial analysis, and visual display of analysis results from external procedures, and not network-level analysis, the inaccuracies of TIGER would not affect its use in the development of the GIS application.

Nature of Attribute Data

The existing program consisted of data from various sources integrated into one data base. This included inventory data maintained by the FRA, accident data from the DelDOT Bureau of Traffic, and data from other sources, such as railroad companies, in tabulated inventory sheets. Each crossing has a unique DOT-AAR (Association of American Railroads) identification number that serves as a key to access each crossing record. Additionally, site-specific and qualitative information was not in the data base but was maintained in written form (*3*).

The attribute information is the data required as input for the USDOT model (*3*). The location reference is given in the form of city and county codes, railroad ID (identification) number, road and railroad names at the crossing, and the milepoint on the approach road.

Location Referencing of Attribute Data

If the location reference is in the form of latitude and longitude coordinates, the input data for a point data base needs to be contained in a single data file. TransCAD (*1*) has the ability to convert data referenced in any other coordinate system when provided with the local and world coordinates of any three points. However, in the case of the current data base, the only location reference information consisted of city, county, street and railroad names, and street milepoint information. Milepoint information on the approach roads could not be used for location referencing because of the unavailability of accurate milepoint information. Interactive referencing was done to tag records to their associated spatial locations.

Some of the crossings records could be location referenced by a simple program that searched for a match in the strings consisting of the railroad and street names for a crossing record with the respective fields in the node layer data base. If a match was found the coordinate information from the node layer record was attached to the crossings record in question. Many of the records could not be geocoded in this manner because of the inconsistencies in the rail and road name strings between the TIGER files and the crossings records for the Delaware rail-highway grade crossings data base.

These records were location referenced by locating the approximate location on the base map, zooming into the location of the crossing, keying in the attribute data into the identified point in the node layer, and finally, downloading the information into an ASCII file and building the crossings data base.

The advantage of location referencing to latitude-longitude coordinates is the referencing of data to a global and universal coordinate system that can be accessed and accepted by most other GIS software.

Analysis and Interpretation of Data

This system provides several options for analysis and interpretation of data and results. Procedures and conditional query abilities are available, including the ability to classify information by theme through color coding and representative icons. Another important spatial analysis capability is buffering. The user can specify a circular area around a crossing point and count the number of objects of a specific type that fall within that area. It is possible for the user to do selective spatial querying to estimate the number of point objects within a specified area, for instance, the determination of the location of another crossing within a quarter mile of a particular crossing, which is essential to the decision to close a crossing. Other objects that can be identified are accident locations, the presence of an intersection within a specified range of the crossing, and the presence of a certain type of land use in the vicinity of the crossing. Overlaying can be performed to gather land use data in the vicinity of crossings; this is one of the inputs to the expert system developed for the program.

Some of the conditions created for query are shown in Figure 1. Conditions were created to reflect the resource allocation strategies to be considered and the strategies for identification of deficient, hazardous, or significant attributes, as the case may be. The user could create thematic maps to classify crossings based on hazard index, traffic, or accidents.

Linkage to external procedures allows for flexibility and the selection of records from data base based on the specific nature of requirements of the type not possible within the functional capabilities of the GIS. An external class of procedures, including a program for executing the USDOT model and a KBES, for decision support were created using the external procedure interface in TransCAD (*1*).

DEVELOPMENT OF THE KNOWLEDGE-BASED EXPERT SYSTEM

Though the USDOT model gives a reasonably accurate estimate of the hazard potential, it does not consider several site-specific and qualitative factors that are important in determining the overall safety status of a crossing. Furthermore, the process of identification of the cause for deficiency or hazard at a crossing and the associated remedial action, is, by and large, one involving engineering judgment on the part of the field expert as well as compliance to codes, mandates, and procedures established for this purpose. The objective of this study was to develop a KBES that accounts for site-specific and judgmental factors that are not considered by the USDOT model and that also provides a guideline for selecting the most feasible combination(s) of actions that can be considered to

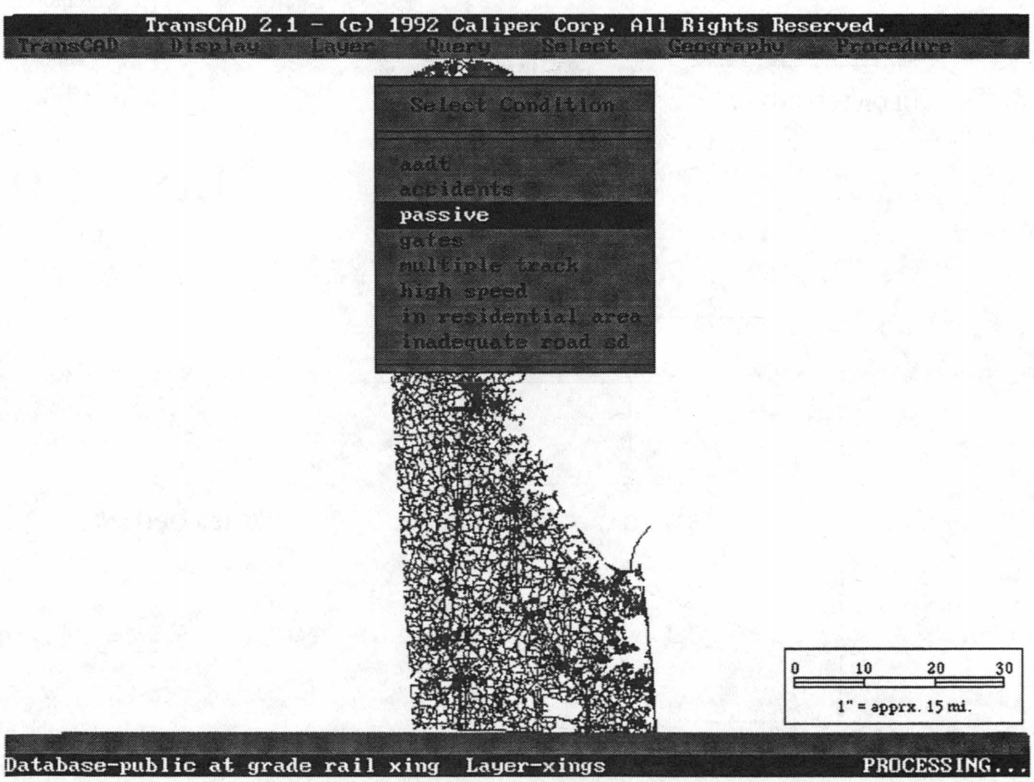

FIGURE 1 Some conditions created for query on attribute data fields.

TABLE 1 Site-Specific and Qualitative Factors Considered by KBES in Safety Decision Making

Track sight distance ratio (actual/available)
Road sight distance ratio (actual/available)
Night visibility
Illumination
Type of land use at the crossing
School bus traffic
Hazardous materials carrier traffic
Pedestrian traffic
Signing near the crossing and sign visibility
Speed limit compliance for the approach road
Proximity to intersection on approach road
Proximity to another crossing (for closure warrant)
Occurence of fatalities
Accident history specifics (train or nontrain related, severity, etc.)
Percentage of trucks at crossing

bring about accident reduction at the crossing. Table 1 consists of the significant site-specific and qualitative factors that are considered by the KBES.

The KBES uses these site-specific factors in conjunction with the USDOT accident index and the basic safety and inventory data (*3*) to assign an indicator of the danger level at a crossing. It further suggests a remedial action that could be considered for safety improvement at the crossing. Fifteen possible basic safety improvement alternatives were identified and established with the help of interviews with field engineers and from the codes and relevant literature establishing guidelines and procedures for safety improvement at crossings (*6,7,8,*). Some of these are standard safety device installation mandates that cover the three basic types of devices, namely, passive protection, flashing lights, and gates. Several other options, which are nonstandard and do not classify as safety devices but are nevertheless important for safety improvement and accident reduction, are installation of Stop signs, reduction of speed limits, improvement of pavement condition, illumination at the crossing, warrant for a pedestrian overpass, and finally grade separation and closure leading to elimination of crossings. Each of these options has a cost and effectiveness factor associated with it that the KBES considers for the particular case being evaluated.

Figure 2 shows the inference network for the KBES. The system is a forward-reasoning system developed in CLIPS (*2*) that goes through a phase-by-phase evaluation. The system modifies the accident hazard index for the crossing and finally suggests a set of possible action(s) for safety improvement. Some of these actions are due to mandates such as the closure of a crossing that has time-table speeds greater than 125 mph, categorized as a high-speed line, and the mandatory installation of gates at every multiple track crossing.

Extensive documentation and on-line help is also available to the user. Documentation includes providing the user with the inference

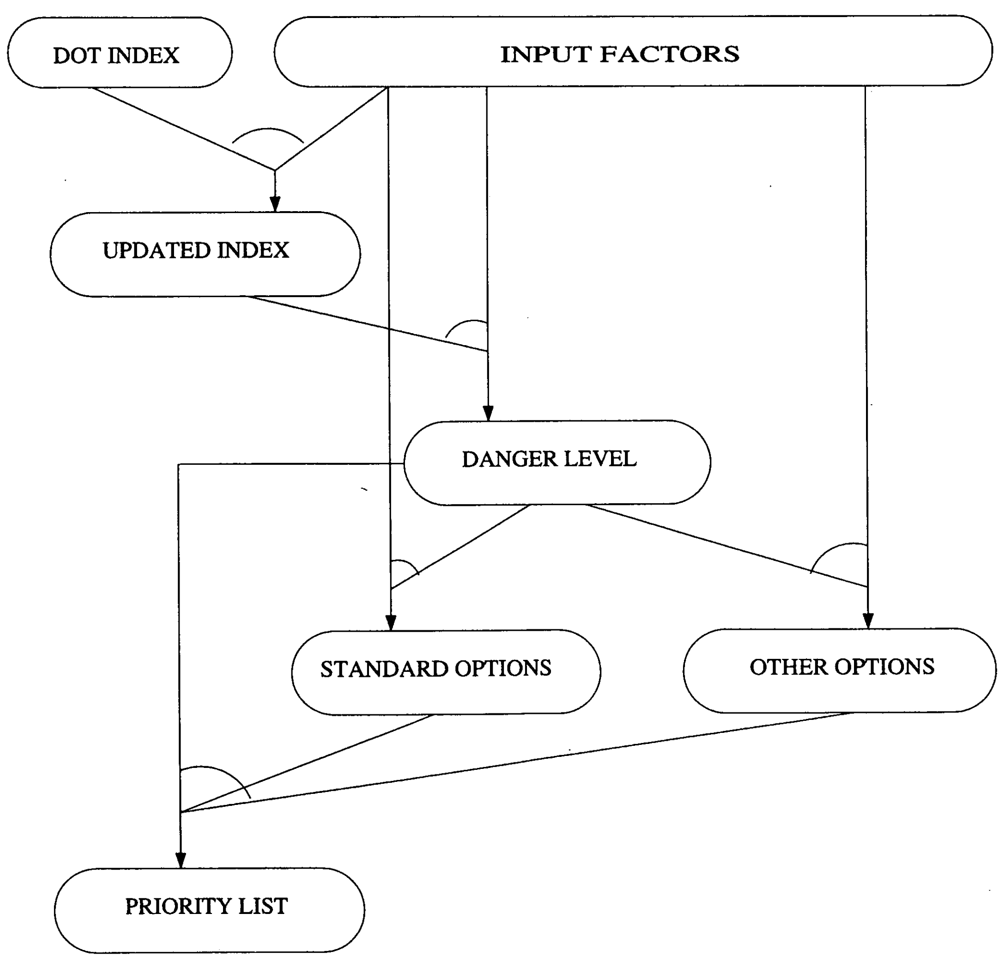

FIGURE 2 Inference mechanism of the rail-highway grade crossing KBES.

chain that the KBES uses in selecting a particular decision alternative. On-line help includes screens containing information on how to input data, run the system, view results, and the options available to the user.

The KBES can run in two modes. It can be user-interactive, with the user responding to queries by the system and supplying the input data, or the system can directly extract the input data from the GIS data base. Results of the system are also provided to the user in detailed descriptive form or can be converted into a text file that can be accessed and viewed within the GIS environment.

The user-interactive mode can be used to evaluate the effect of change in conditions for any single crossing while the data-acquisitive mode can be used to analyze several crossings at a time. This would be especially useful when processing a large crossings data base. Figure 3 shows a sample result of the KBES for a crossing giving the danger, current protection, and improvement, recommendations at the crossing.

Verification and Validation

"Verification is the process of building the system right, validation is the process of building the right system" (*9*).

Verification of the system was performed on similar lines to that done for conventional software programs. Verification in this case is restricted to that of the knowledge base, as the inference engine used is that of the CLIPS shell (*2*), verification of which is thus assumed. The two major steps involved in the verification process were checking the system for compliance with requirements, and checking for syntactical, logical, and other errors that affect the consistency and completeness of the knowledge base.

The interface between the system components was checked for proper functioning through passing of data with the GIS data base. The specification was for a user interface that permitted user interaction, automated extraction of data from the GIS data base, and finally, the presentation of results within the GIS environment. Compliance of the developed user interface to this specification was found to be satisfactory. Other aspects that were considered in system verification included the adequacy of explanation and help facilities and the modularity and maintainability of the system. The system was also checked for syntactical errors in the rule base including conflicting, redundant, subsumed, circular, dead-end, missing, and unreachable rules.

The validation methodology established was the presentation of a set of test cases and analysis of the results provided by the system in response to these test cases. Some of these test cases were from the crossings data base, and the results for these were compared with those of a human expert's evaluation. Other test cases were random and hypothetically formulated to test for theoretically possible and extreme combinations of input data. For the problem under consideration there is no clear-cut solution. The KBES suggests a set of options, as seen in Figure 3. Agreement of the experts with these options provided by the KBES for a series of test cases was the criterion for validation. In the majority of the cases tested on a random basis, the expert was satisfied with the results provided by the KBES. Some minor modifications were made as a result of the validation process.

The system was also tested for adequacy and completeness of the knowledge base by presenting random test cases with different combinations of input data to catch any errors of omission. Omission is a situation in which the system fails to reach a conclusion for a given input case, indicating incompleteness of the domain knowledge.

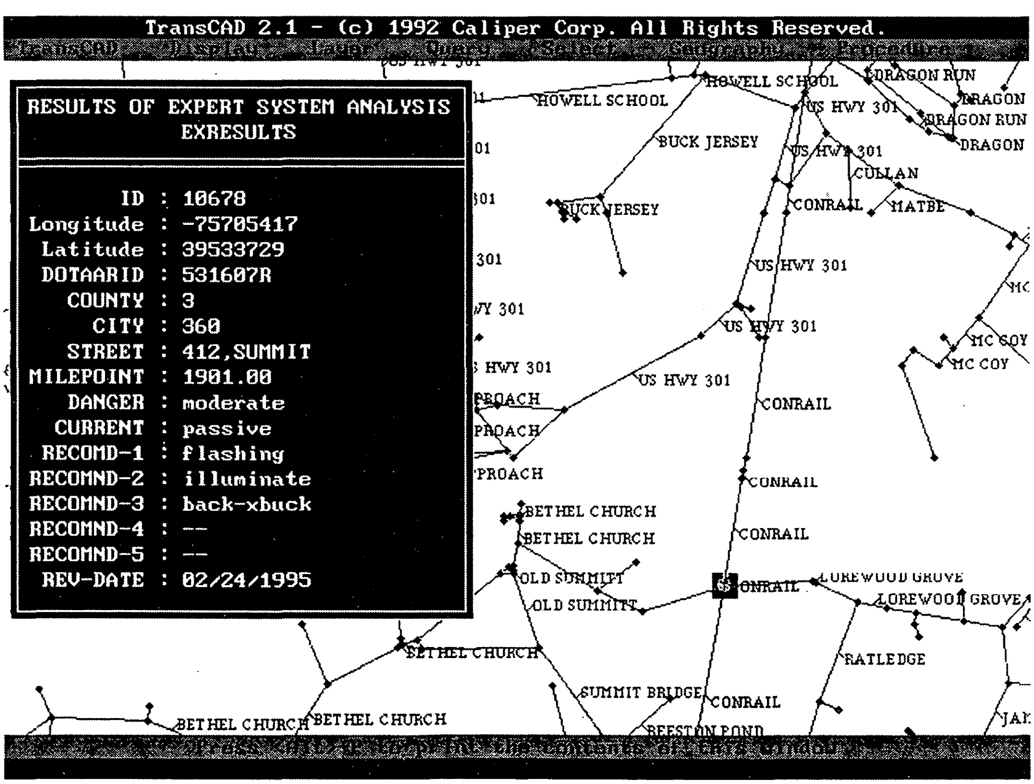

FIGURE 3 Results of KBES analysis for the crossing in response to the changes, viewed in the GIS.

INTEGRATING THE COMPONENTS OF THE CROSSINGS SAFETY PROGRAM

The previous sections described the development of the different components of the rail-highway crossings safety program for Delaware. The different components developed for the program that have been described are:

- A GIS application for rail-highway grade crossing safety in TransCAD (*1*),
- A location-referenced, expanded crossings safety data base that can be accessed by the various components of the program,
- the USDOT model for calculating accident index, and
- A KBES for updating the crossings hazard index, identifying hazards, and suggesting improvement action(s) at each crossing.

The interfacing of these components is the basis of the integrated functioning of the rail-highway safety management program. The format of input and output for each of the components of the system must be compatible with the others in order to interface them. Table 2 shows the input and output format and content specifications for each of the components. The framework of the integrated system is shown in Figure 4. The interface can either be from within the GIS or from outside the GIS depending on the requirements and the level of access desired. For the integration of these components, the required data is accessed by each of the procedures from the GIS data base and the results are passed back to the GIS. The integrated working of all these components is required for complete analysis of the crossings problem. The interface was achieved in two ways:

1. The user can access the KBES and the USDOT model program from within the GIS application environment by selecting the appropriate option from the procedures menu as shown in Figure 5. The GIS environment is used to display crossings, view the results of analysis, and provide input data to the interfaced components.
2. The control can be passed to the user from outside the GIS, as shown in the custom menu in Figure 6, that is, at the operating system level so that the user can access the different components including the data base builder. It also presents the user with documentation and help before using any of the components, assuming the user has minimal knowledge of the software/shells in question.

The crucial part of the interface is to maintain the location reference on any record being processed at every stage of the process. Figure 3 shows a sample result of a typical analysis for a crossing by the KBES. The external program takes in data from the record for that crossing and appends its location reference data to the data file containing the results of the analysis so that it can be passed back into the GIS for viewing and analysis.

When any menu item is selected from the menu shown in Figure 6, the appropriate chain of command substitutions and path changes are performed. Data are converted from the GIS data base to the format required. On-screen messages and documentation are provided to the user about the system component to enter, the specific function to perform, and options to execute within that environment.

TABLE 2 Input/Output Format and Requirements for Interfacing GIS, KBES, and External Procedures

Comp.	Input Specification	Input Contents	Output Specification	Output Contents
KBES	ASCII Format, User Screen Inputs, Worksheet Form	Hazard Index, Safety Factors, Site-Specific, Qualitative Information	ASCII Format, Screen Output, Worksheet Form	Danger Index, List of Decision Alternatives
External Procedure (USDOT Model)	ASCII Format	Crossings Safety Factors, Accident Data	ASCII Format	Updated Hazard Index, Priority List of Crossings
GIS	Spatial Data: TIGER, DLG, Scanned, Digitized Hard Copy Maps, Orthophotos etc.	Point, Line and Area Layers, Roads, Land Use, Census Blocks etc.	PCX, Other Graphic Format	Screen Image at any Display
	Attribute Data: ASCII, Text Files, Worksheet files etc.	Safety, Accident Data etc.	Worksheet, Text, ASCII, Other File Formats	Specific Attribute Data Dumped on Requirement

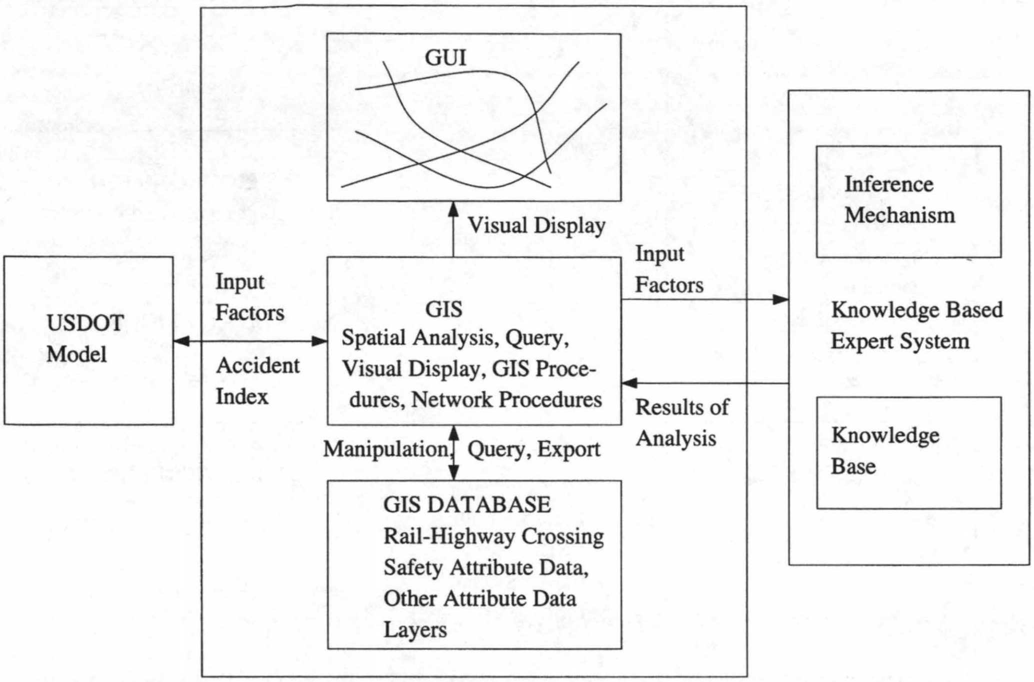

FIGURE 4 Framework of the knowledge-based GIS for safety analysis at rail-highway grade crossings.

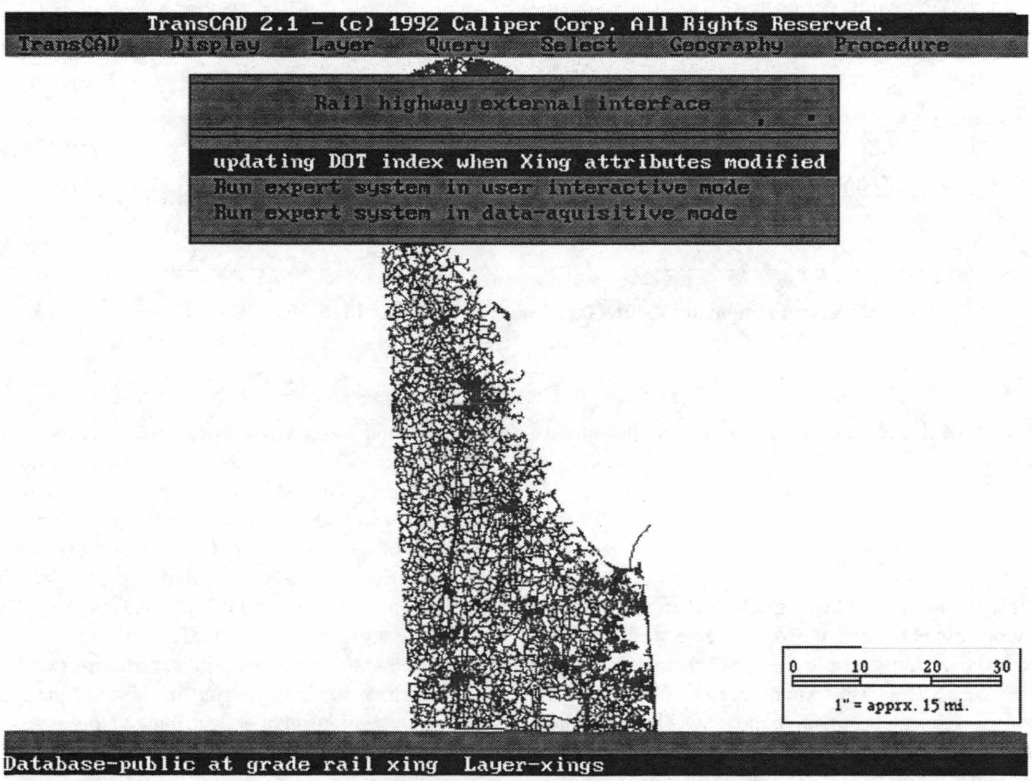

FIGURE 5 Interface menu for KBES in CLIPS and USDOT model from within TransCAD rail-highway application.

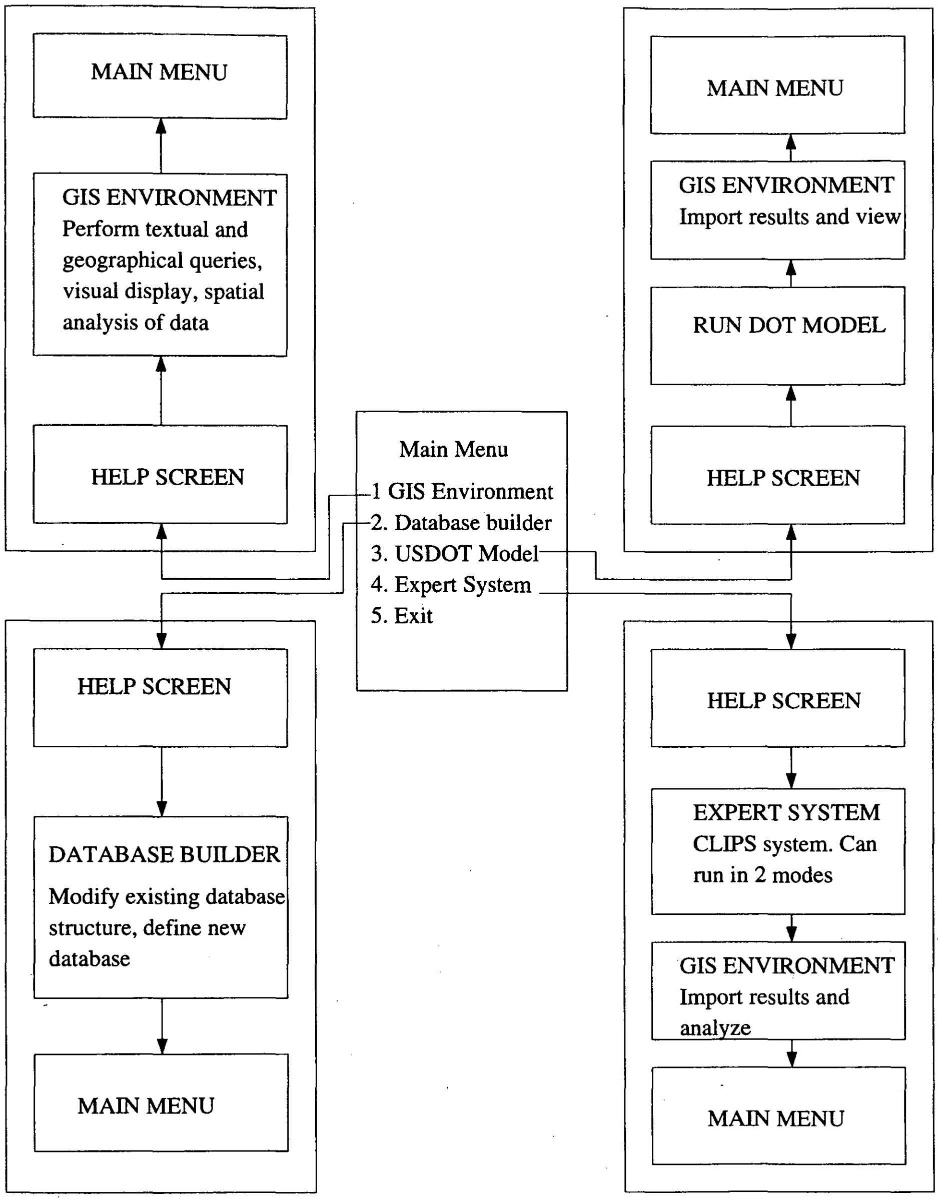

FIGURE 6 Custom menu of the interface, chain of processes in the execution of each menu item.

Figure 6 shows a view of the screens and environments presented to the user after execution of each of the menu items.

CASE STUDY

The case study consists of a rail-highway grade crossing for which some of the safety attributes have changed. The case study is presented to show how the system components could be used to analyze a particular crossing. The case shown here is an example of how the crossings data base can be updated, the USDOT model and the KBES run, and the results passed back to the GIS.

Some of the current safety attribute values on the crossing are seen by querying on the crossing location as presented in Figure 7. Consider a hypothetical change due to a new residential development and an increase in local traffic, as well as in hazardous material-carrier traffic caused by some rerouting in the vicinity of the crossing. The land use changed from predominantly insignificant use to residential use because of a new development in the vicinity of the crossing. This information is directly obtained from the land use layer, if it is regularly updated, or must be updated by keying in the information. The changes due to this new development are an increase in AADT, a change in land use type, and an increase in hazardous materials-carrier traffic.

The changes were updated in the attribute data base. Figure 8 shows some of the changed attribute values upon query at the crossing. The KBES was then run, and the results were passed back into a separate data layer created for that purpose, as can be seen in Figure 3. The danger level at the crossing after the KBES evaluation changed from low to moderate. The current protection is passive.

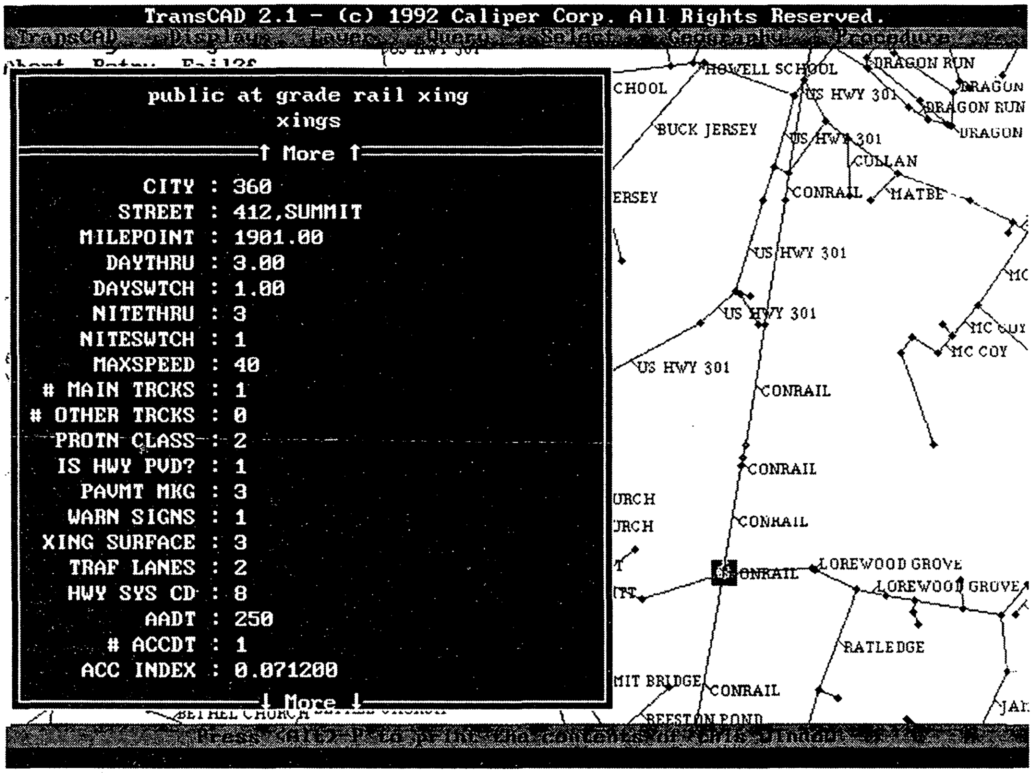

FIGURE 7 Query on crossing showing safety attribute data before changes.

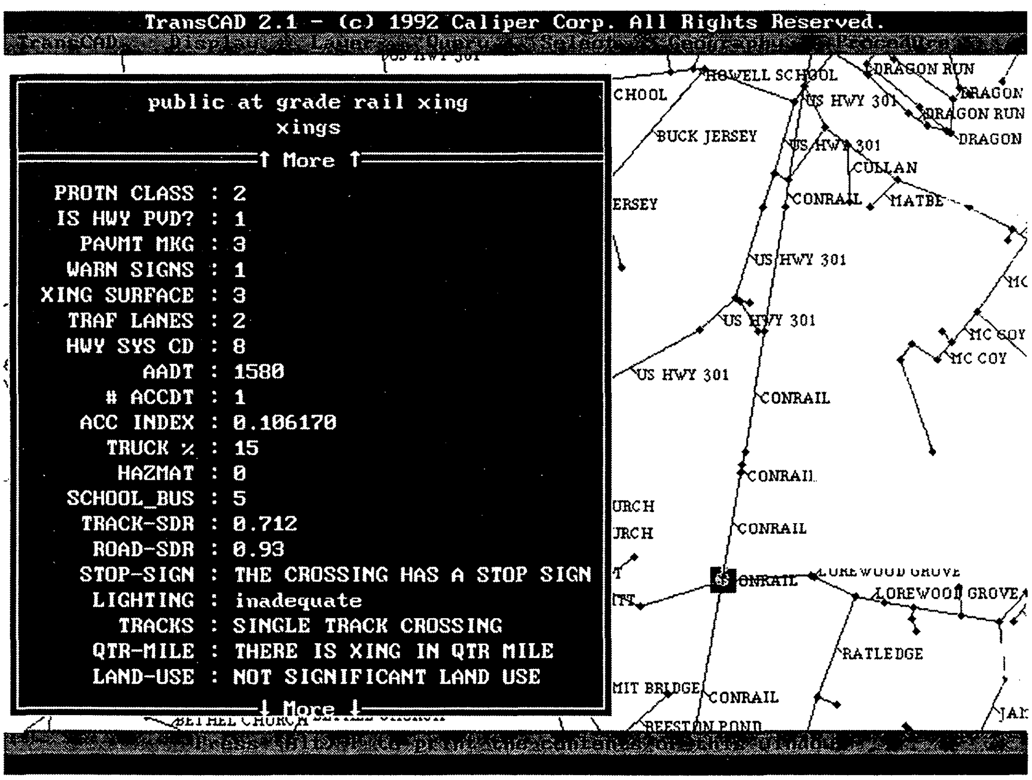

FIGURE 8 Query on crossing after changes in attribute data and USDOT accident index.

The system suggests the installation of a flashing light at the crossing and the improvement of illumination at the crossing. The installation of a backup crossbuck from sight distance consideration is also recommended. Note that all these recommendations only serve as a guideline, and that the field engineer needs to use his or her judgment to make the final decision. For example, in this case, the installation of flashing lights may be examined if the new traffic pattern, because of rerouting, is likely to become a permanent feature. Other options, such as installing a back-up crossbuck, may be considered more strongly because of the lower cost associated with it.

SUMMARY AND CONCLUSION

The work effort described consists of development of an integrated software package for management of safety-related data for rail-highway grade crossings. This involved developing a GIS application for visual display and spatial analysis of safety data and remedial actions; incorporating a program for calculation of an indicator of the accident potential using the USDOT model; and finally, the development of a KBES for modifying and prioritizing the indicator, and suggesting action(s) for safety improvement at the crossing. These components are integrated in such a way that the results and data required by each are compatible with the others, in order to enable visual access and presentation of results and data within the GIS environment.

The benefits of this work are the automation and efficient handling of large amounts of data, typical of a crossings management program for a large state, which can have a few thousand crossings. The use of GIS allows not only better display, spatial analysis, and an overall better visual perception of the problem, but also better data access and management, including access of related data from other layers and sources, easier editing and updating of data, and enabling all analysis to be performed by a user with minimal knowledge of the system.

The complete system results in an aid to resource allocation for safety improvement at rail-highway grade crossings. A more complete and meaningful analysis of the factors involved is achieved through the incorporation of the considerable amount of heuristic reasoning and engineering judgment that goes into the resource allocation process through the KBES. The overall benefits are not only the result of the better analysis and presentation capability of the GIS, but also of the incorporation of a decision support mechanism into the system.

ACKNOWLEDGMENT

Funding for this project was provided by the Delaware Department of Transportation through the Delaware Transportation Institute, Department of Civil Engineering, University of Delaware.

REFERENCES

1. *TransCAD—Transportation GIS Software, Version 2.0, Reference Manual.* Caliper Corporation, Newton, Mass., 1990.
2. *CLIPS Reference Manual, Version 5.1.* Lyndon B. Johnson Space Center, Houston, Tex., Sept. 1991.
3. Faghri, A., and N. Vukadinovic. *Evaluation of Rail-Highway Grade Crossings Program in Delaware.* Report 91-DTC-1. Delaware Transportation Center, University of Delaware, Sept. 1991.
4. Faghri, A., and M. J. Demetsky. A Comparison of Formulae for Predicting Rail-Highway Crossing Hazard. In *Transportation Research Record 1114,* TRB, National Research Council, Washington, D.C., 1980, pp. 152–155.
5. Hitz, J., and M. Cross. *Rail Highway Crossings Resource Allocation Procedure User's Guide.* Report FHWA-IP-82. Transportation Systems Center, U.S. Department of Transportation, Cambridge, Mass., 1982.
6. U.S. Government. *Code of Federal Regulations, 23.* Washington, D.C., 1991.
7. *Railroad-Highway Grade Crossings Handbook.* FHWA, U.S. Department of Transportation, 1986.
8. *Manual on Uniform Traffic Control Devices.* FHWA, U.S. Department of Transportation, 1978.
9. Gonzalez, A., and D. Dankel. *The Engineering of Knowledge Based Systems—Theory and Practice.* Prentice Hall, Engelwood Cliffs, N. J., 1993.

Publication of this paper sponsored by Committee on Artificial Intelligence.