PERGAMON

Transportation Research Part C 8 (2000) 337–359

TRANSPORTATION
RESEARCH
PART C

www.elsevier.com/locate/trc

Spatial decision support system for hazardous material truck routing

William C. Frank [a], Jean-Claude Thill [b,*], Rajan Batta [a]

[a] *Department of Industrial Engineering and National Center for Geographic Information and Analysis, State University of New York at Buffalo, Buffalo, NY, USA*

[b] *Department of Geography and National Center for Geographic Information and Analysis, State University of New York at Buffalo, Buffalo, NY 14261, USA*

Abstract

Shipping hazardous material (hazmat) places the public at risk. People who live or work near roads commonly traveled by hazmat trucks endure the greatest risk. Careful selection of roads used for a hazmat shipment can reduce the population at risk. On the other hand, a least time route will often consist of urban interstate, thus placing many people in harms way. Route selection is therefore the process of resolving the conflict between population at risk and efficiency considerations. To assist in resolving this conflict, a working spatial decision support system (SDSS) called Hazmat Path is developed. The proposed hazmat routing SDSS overcomes three significant challenges, namely handling a realistic network, offering sophisticated route generating heuristics and functioning on a desktop personal computer. The paper discusses creative approaches to data manipulation, data and solution visualization, user interfaces, and optimization heuristics implemented in Hazmat Path to meet these challenges.

Keywords: Hazardous material; Routing; Spatial decision support system; GIS; Highway network

1. Introduction

Transportation of hazardous materials (hazmat) is necessary for an industrial economy to exist. For example, gasoline is consumed by millions of cars, trucks and buses in the US, but needs to be transported from refineries/shipment points to gas stations. Gasoline is clearly

* Tel.: +1-716-645-2722 ext. 24; fax: +1-716-645-2329.
E-mail address: jcthill@acsu.buffalo.edu (J.-C. Thill).

PII: S0968-090X(00)00007-3

hazardous in the sense that a truck carrying gasoline, if involved in a collision, can result in personal casualty and serious damage to other vehicles and property. However, it begs us to the question: what is hazardous material? The Federal Hazardous Materials Transportation Law (October '94) states that "The secretary of transportation shall designate material (including an explosive, radioactive material, etiologic agent, flammable or combustible liquid or solid, poison, oxidizing or corrosive material, and compressed gas) or a group or class of materials as hazardous when the secretary decides that transporting the material in commerce in a particular amount and form may pose an unreasonable risk to health and safety or property".

To understand the magnitude of the problem, one simply has to note that there are approximately 500,000 daily hazmat shipments (Nozick et al., 1997). Hazmat is transported by air, highway, railway, and water. In a perfect world, incidents would not occur when transporting hazmat. However, due to the risks associated with transportation, hazmat incident costs to health and property are quite substantial. Highway incidents alone were responsible for 15 deaths, 511 injuries and $19,800,000 of property damage in 1993 (US DoT, 1994).

Often, a hazmat accident results in the spillage of a dangerous chemical that necessitates an evacuation of the population in the surrounding area. Evacuations are costly and are often used as a surrogate measure of unreasonable risk to health and safety. US DoT (1994) statistics indicate that the average number of people evacuated per highway incident is above one hundred, which is substantial since people evacuated during an incident are placed at an unreasonable risk.

There are several strategies that come to mind to mitigate the risk associated with such a situation. First, by judiciously selecting the route we can reduce the probability of an accident. Also, choosing a route through less populated areas could also be used to reduce the number of people placed at unreasonable risk. Next, vehicle and container design could be altered to reduce the severity of a release once an accident has occurred. Finally, accident probability could also be reduced by improved driver training.

In this paper, we focus on risk mitigation via route selection. We present a spatial decision support system (SDSS) that recommends one or more routes compatible with stated objectives and constraints. While previous hazmat route solutions have often been advanced for small illustration networks, our goal is to operationalize our SDSS on a large-scale transportation network. For this purpose, the implementation makes use of a realistic road network of about 57,000 intersections maintained by the Bureau of Transportation Statistics. The sheer size of the network calls for innovative design solutions discussed in the rest of the paper.

In Section 2, the problem of hazmat routing is discussed in greater detail and our general approach to its solution is presented. A brief literature review follows in Section 3. The remaining sections discuss in detail the components of the proposed DSS called Hazmat Path. The first part of Section 4 describes the user interface, which includes the routing parameters, network intersections and display options. The second part of Section 4 discusses the data structures that include the map objects and link attribute costs. The methods used to generate link attributes that are SDSS inputs (time-dependent population at risk, accident rates and link travel speeds) are presented in Section 5. Section 6 is devoted to the route generation algorithm of the SDSS: the mathematical formulation of the problem is presented and solution methods are outlined. The paper concludes in Section 7 with an overview of the main properties of the proposed routing system and of future extensions that it can accommodate.

2. Route selection strategies: problem statement

The goal of this paper is to develop a working SDSS capable of handling a realistic transportation network covering a multi-state region or an entire country and their numerous loading and delivery points. DSSs are computer-based systems that share several key characteristics, including (Sprague and Carlson, 1982; Geoffrion, 1983; Turban, 1995):

- Assisting users in their decision making in a flexible and interactive manner.
- Solving all classes of problems, including ill-structured ones.
- Having a powerful and user-friendly interface.
- Having a data analysis and modeling engine.

Conceptually, SDSSs are special cases of DSSs. Densham (1991) effectively argues that they differ markedly from general DSSs in some key respects. They need spatial capabilities for data input, display of complex relations and structures, analysis, and cartographic output. The architecture of our Hazmat Path SDSS is depicted in Fig. 1. Hazmat Path is a full-featured SDSS allowing for interactive problem editing, comparison of solutions, and evaluation of decision criteria.

There is a difference between conceptualizing and actually developing a working SDSS. When developing a working SDSS, numerous trade-offs exist between level of effort and SDSS quality. A conceptual SDSS does not have this real world trade-off. Commercial geographic information systems (GIS) software applications, which often provide the spatial information processing engine of SDSSs, are designed with tremendous flexibility but at a cost to time for producing results. Displaying results in a timely manner is imperative for a SDSS. Therefore, the choice has been

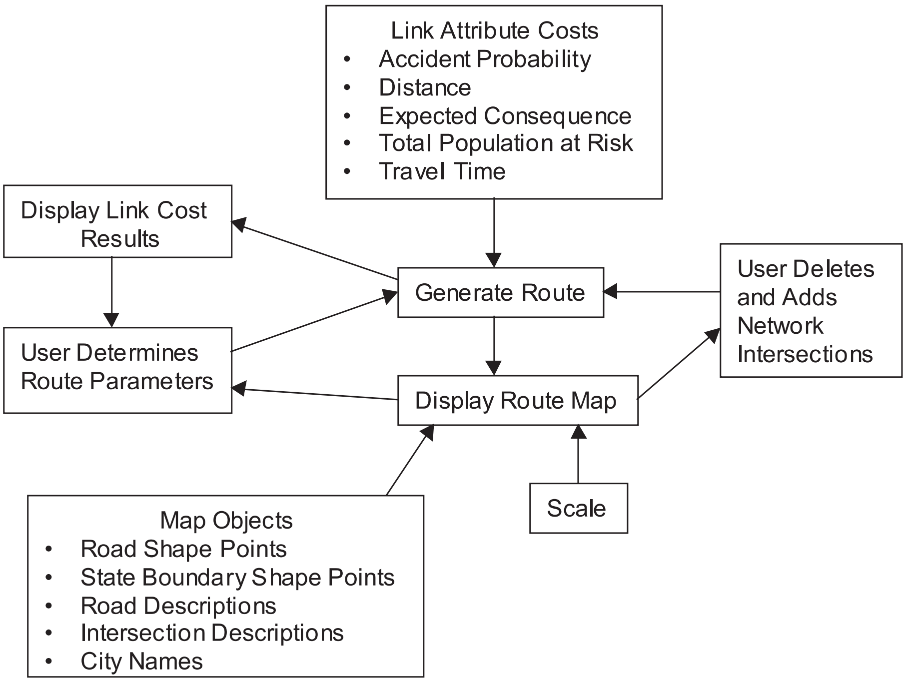


Fig. 1. Flow of information in Hazmat Path.

made to design a Windows-based software application tailored to the SDSS and running on mid-range desktop computers instead of using an existing commercial application.

Since most long-distance travel occurs on highways, they compose the network on which travel time is minimized. They primarily connect and transverse large population centers. Therefore, minimizing travel time puts a large population at risk. Population centers can be avoided by using slower and less direct non-interstate roads. Thus, the strategy of selecting a route to minimize travel time and the strategy of minimizing population at risk are conflicting in nature.

Arbitration between these conflicting costs makes use of the capability of the SDSS to generate several alternative truck routing solutions based on single optimization criteria. A widely used approach is to combine several attribute costs into a single cost. The new cost is often taken to be a linear function of population at risk, distance, time and accident probability. With a single link cost, a simple solution method (e.g., Dijkstra's shortest path algorithm) can be used to determine a vehicle route. By varying the weights of the attribute costs, different routes can be generated. The process of varying the weights indicates the sensitivity that the attribute costs have on route selection.

Another approach to route selection is having multiple objectives. Minimizing travel time and total population at risk is an example of multiple objectives. Many routes can be in the solution set because it contains all of the non-dominated routes. The number of non-dominated paths can become very large in networks typical of real-world applications, thus rendering the approach unpractical.

Still another approach is to minimize one cost attribute while limiting the sums of other cost attributes. This type of problem is called a constrained shortest path (CSP). It is used in this research as follows. Travel time is minimized while four other criteria – total population at risk, distance, accident probability, and consequence – are constrained. Travel time is chosen for the objective function because, it best represents financial cost. As noted earlier, minimizing time for a long route often produces a path, which places large numbers of people at risk. Consequently, the CSP problem is a method that resolves the minimum time and minimum population conflict.

An added complication to the CSP problem is that some link attributes are assumed in this research to be a function of the time of day. Link travel time is one of the temporal link attributes. The level of traffic congestion influences link travel time and, congestion is affected by time of day. This problem is called the time-dependent constrained shortest path (TCSP).

3. Literature review

The literature relevant to our hazmat routing problem can be organized into studies that implement a solution within a GIS and those that use a SDSS. Both groups use GIS techniques for data storage, data manipulation (for instance, to generate link attribute costs) and to display solutions on a map. SDSS-based implementations have the added capability of allowing the user to easily specify shipment origins and destinations, along with removing intersections and roads from route consideration. GIS provides an ideal environment for design and management of hazmat routes because of its ability to integrate multi-theme and multi-source data into an operational information system. Souleyrette and Sathisan (1994) advocate the use of GIS for the comparative study of pre-defined, alternative routes on selected characteristics. This type of analysis is illustrated by a case study of Nevada highway and rail routes for shipment of high-level radioactive materials.

Abkowitz et al. (1990) envision GIS to fulfill functions in hazmat transportation that reach beyond those of input data storage, data manipulation and output map display. They propose a GIS application of hazmat routing on a large-scale network of size similar to the one used in this research. The routing algorithm handles a single routing criterion, but compromise or negotiated solutions can be achieved by ex-post comparison of solutions generated on different routing criteria. In their implementation, the authors use criteria of distance (a measure of efficiency) and populations at risk (a measure of safety). The latter is measured by the tally of people within a given bandwidth (0.25, 0.5, 1, 3, 10, or 25 miles) along highway segments. A gradient method is used in conjunction with Thiessen polygons to allocate enumeration district population to pre-defined buffers.

In their study of transportation of aqueous hazardous waste in the London, UK, area, Brainard et al. (1996) apply weighting schemes to identify routing solutions that compromise between criteria. A single link attribute is calculated from a weighted combination of link attributes (population at risk, groundwater vulnerability, and accident likelihood computed from historical records). A labeling algorithm is then used to minimize this new attribute combination. Solutions associated with alternative weights can be compared visually (map display) and statistically (for instance, travel time, expected number of accidents, highway mileage versus mileage on local roads) for risk assessment purposes. A similar approach is suggested by Lepofsky et al. (1993). These authors stress that the network data model used by GIS to represent individual highway segments allows for their detailed attribute characterization and for the efficient modeling of segment-specific risk of hazmat shipping. Following established practices in the matter, they define risk by a combination of accident likelihood, probability of a release, consequence of an incident measured in terms of population exposed, and risk preference of affected interest groups. Generation of the accident likelihood and consequence factors is considerably enhanced in a GIS environment. Lepofsky et al. (1993) present the case study of a shipment from the California/Arizona border to Vandenberg, CA, with a rather small-scale network representing the highway system of Southern California. A route through Los Angeles is produced when travel time was minimized. On the down side, this route has the highest population exposure. A compromise solution is retained with weights of 25% travel time and 75% accident likelihood, allowing for an acceptable trade-off between efficiency and safety.

Useful components of a hazmat SDSS are outlined by Baaj et al. (1990) and discussed in greater detail by Erkut (1996). They include using different routing solution methods. Also, interactive post-editing of generated and displayed routes gives control of the process to the analyst: a user may wish to create a detour around a sensitive location, or it may be deemed desirable to remove some links from the network and have a new route generated. Usually there is no "best" route. Minimizing one criterion typically conflicts with minimizing another. By an iterative process of displaying routes, using different solution methods and creating detours, a compromise route can be developed.

A working SDSS is developed by Lassarre et al. (1993). Their network covers 600 km^2 in the region of Haute-Normandie, France. The SDSS has the capability of loading geographical overlays that include hydrology, railways and population densities. Dijkstra's algorithm is used to compute a route with the lowest risk. Risk is defined to be the product of accident rate and people affected in an accident. The latter is composed of the day population in adjacent polygons (buffers of given width around highway segments), on links (traffic of road segments), and nearby points (children at school). Their SDSS also has the ability of removing links and nodes or areas with particular characteristics (i.e., single-lane roadways) from the network.

Coutinho-Rodrigues et al. (1997) propose a personal computer-based SDSS for multi-objective hazmat location and routing problems. The routing solution generator offers several alternative multi-objective optimization techniques, including the weighted method, the constraint method, and goal programming. The user interacts with the SDSS through multiple graphical and numerical solution displays. Time-dependence of attributes and solutions are not accommodated by the SDSS in its current form. Geographic visualization of attributes and solution paths is limited to graph windows that hardly allow for geographic reasoning on solutions and other geo-referenced information.

4. User interface and data structures

4.1. Map objects display

The user may elect to display several relevant map objects. Fig. 2 shows the options available to the user. In this dialogue box, truck route selection is gray because no route has been generated so far and therefore, cannot be displayed. The user can also specify the size and color of many of these objects.

Displaying individual road attributes is useful when evaluating truck routes. Questions such as "Are there any alternative routes available with a lower total population at risk without greatly increasing travel time"? can be explored and answered by visual examination of suitably attributed map objects. The geo-visualization method used here to display link costs is by road width.

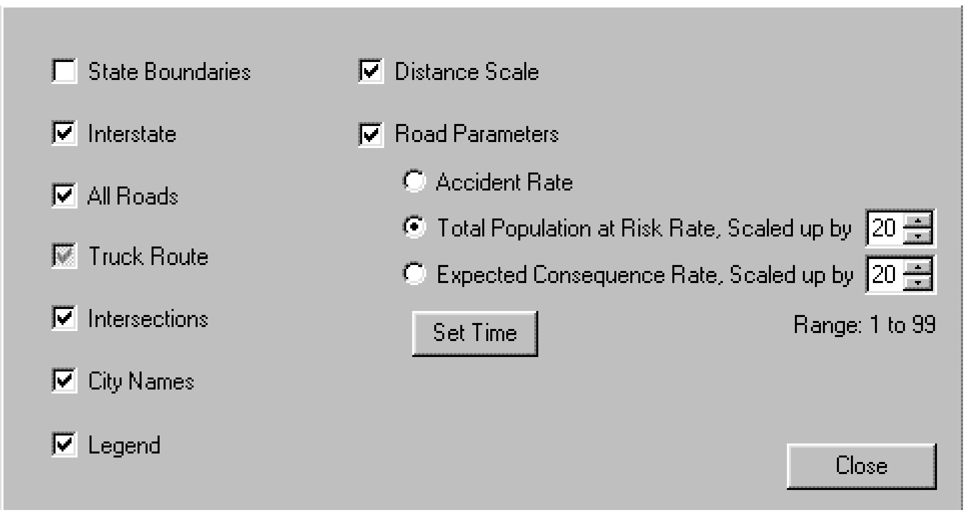


Fig. 2. Map objects selected dialogue box.

In Fig. 2, the attribute "total population at risk" is selected for display and the resulting map is shown in Fig. 3. The truck route is shown with intersection arrival times.

A linear relationship between attribute cost and width of the representation of the road feature is used. The maximum road width is seven pixels. Each incremental pixel width corresponds to a range of the attribute cost, with the requirement that the range of the single pixel start at zero. Therefore, the area of a link on the monitor is a measure of its cost. Furthermore, the origin to destination path displayed with the least area (smallest number of colored pixels) is also the least cost path for that attribute.

The population at risk road parameters (total population at risk and "expected consequence" in Fig. 2) exhibit heavily skewed frequency distributions, where very large values appear in a few large urban areas. A strict application of the rule of proportional symbolic representation leads to the depiction of most roads outside these areas with a width of one pixel. While this poses no problem when the area of interest includes these few large urban areas, for proper visual differentiation, the pixel width should be scaled up as shown in Fig. 2 when they are not included. With default scaling of 20, roads with 1/20 to 20/20 of the true maximum population at risk are displayed with the same maximum width. Roads with zero to 1/20 of the true maximum population at risk have pixel width proportional to population at risk.

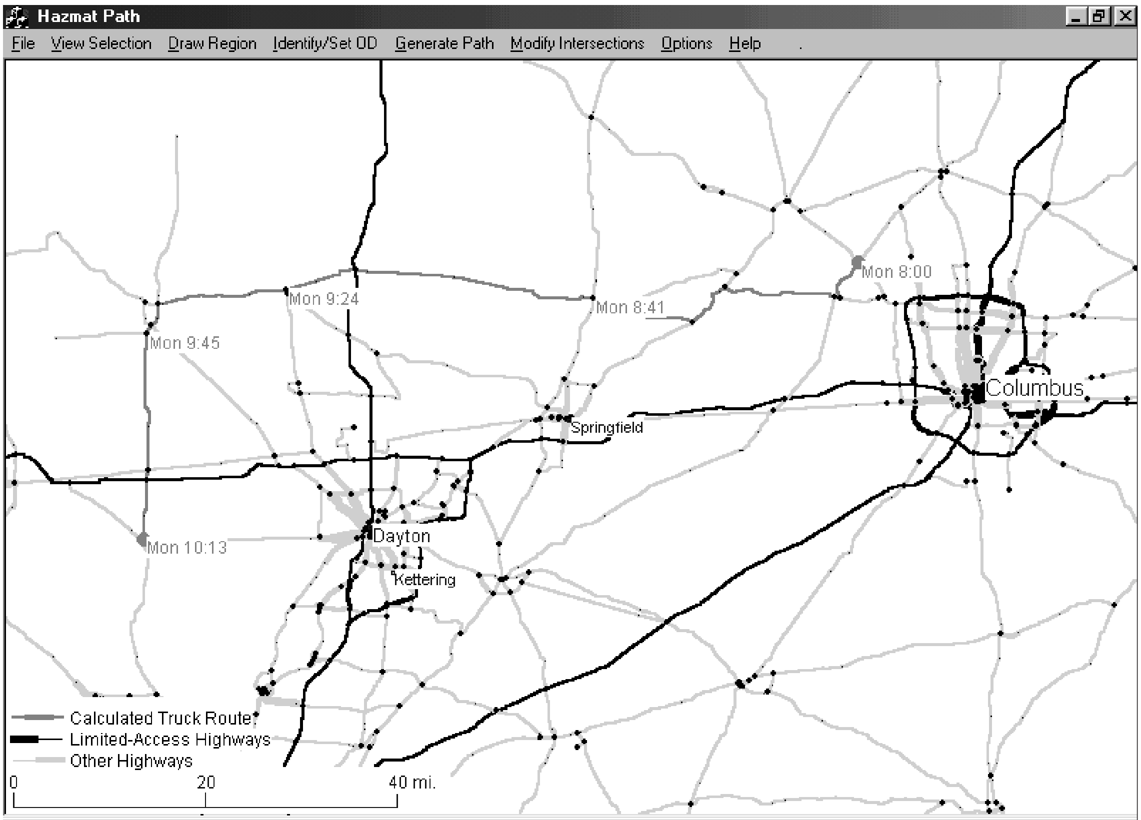


Fig. 3. Display of total population at risk.

Since the area of a link is linearly proportional to one of its attribute costs, the link length times the attribute cost needs to represent the cost to travel that link. The probability of an accident pixel width is a linear function of probability of an accident per mile. Expected consequence width pixel width is also a rate, expected consequence per mile. The same is true for total population. Average population at risk has no equivalent rate and therefore, cannot be displayed.

4.2. Deleting and adding network intersections

Hazmat Path uses the national highway planning network (NHPN) maintained by the Federal Highway Administration (Bureau of Transportation Statistics, 1999). The NHPN has 95,000 nodes. Pre-processing of the original network brings it down to 57,000 nodes. This operation involved removing nodes outside the continental US as well as most of the nodes with only two incident links.

The user has the option of temporarily removing intersections from the network. Intersections temporarily removed from the network cannot be part of the truck route being planned. There are two reasons for removing intersections, namely to reduce solution run-time and to prevent a solution from traversing a particular region for equity reasons or to comply with local or state hazmat traffic regulations.

In our SDSS, three different methods are available to exclude or deactivate intersections. The first method is by selecting individual intersections by point and click with the mouse or cursor. Intersections can be added to the network by the same operation. Also, the cursor can be dragged over a rectangular region to add or remove all intersections within the circumscribed area. Fig. 4 displays a rectangle of intersections being deleted: the intersections represented by clear circles have been removed from the network. The last method is associated with a constraint imposed by the user. Only intersections that are reachable while traveling from the specified origin to the specified destination without violating the said constraint are included in the network. In the current implementation, the constraint is based either on travel time or distance.

Often many hazmat shipments take place between the same origin–destination pairs. Following the same route for multiple shipments could concentrate the risk among a localized group of people above and beyond what is deemed acceptable. Therefore, with such equity considerations in mind, one may consider spreading the risk more evenly over a large population. This can be accomplished by finding D differentiated routes. The metric for total population at risk would be a set of D numbers (total population placed at risk by only one of the routes, total population placed at risk by any two of the routes… total population placed at risk by all D routes). This requires the calculation of the combined total population at risk of any two links along with developing a solution method, both of which are beyond the scope of this paper.

A less computationally intensive approach can be applied by utilizing the network editing features of the SDSS. The user can develop differentiated routes by reducing the overlap of route buffers in populated areas. For this purpose, the total population at risk along each road segment is displayed along with the mile scale so as to reveal the population density in regions of interest. The mile scale is needed for the user to estimate the distance between highway links or more importantly the degree of overlap between link buffers. The approach consists in manually subtracting and adding intersections and routes generated until an appropriate level of differentiation is obtained.

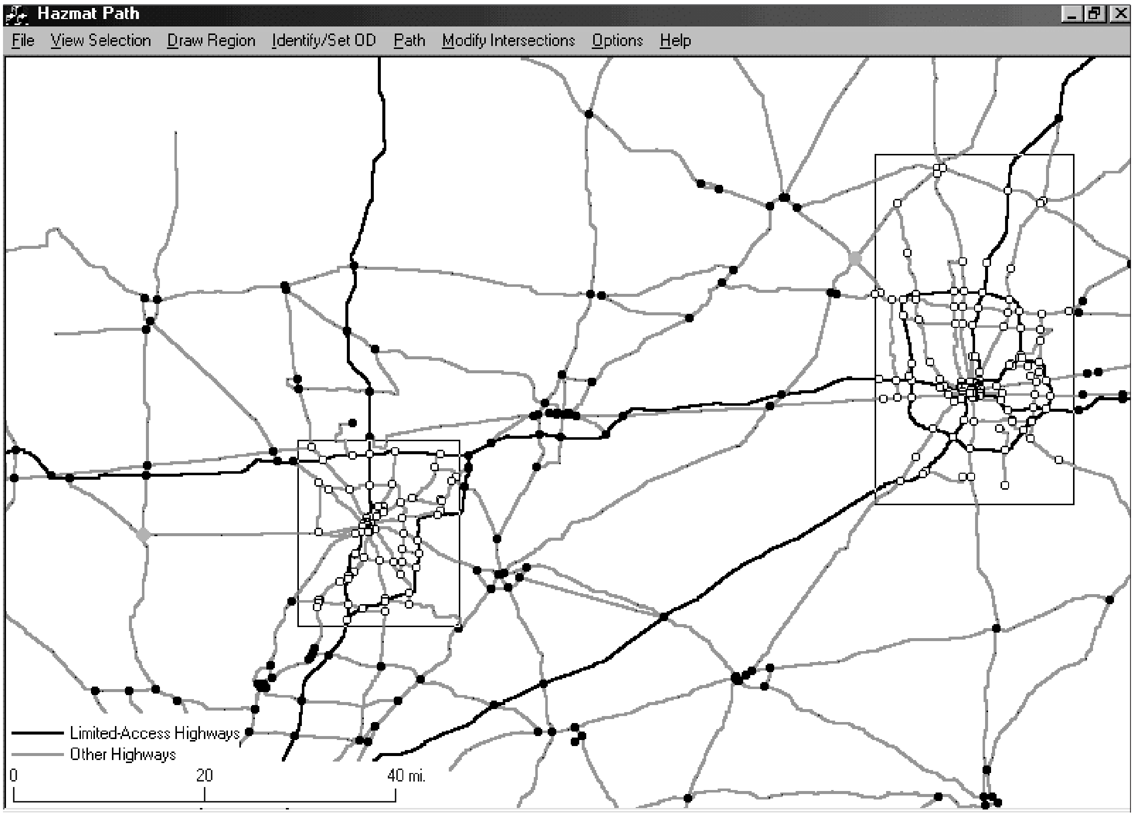


Fig. 4. Display of removal of intersections.

4.3. Route generation parameters

The SDSS can generate two different categories of routes, namely those along the least cost path and those along the least time path with limits on the other attributes. In both instances, the user is prompted to input the distance from a road within which an accident is expected to have harmful impacts on human populations. The area affected by the release of toxic materials and population exposure depends heavily on the properties of the material being shipped and the spill characteristics (Lepofsky et al., 1993). The user-supplied distance is therefore case-specific. It defines the radius of the moving circular buffer used to identify average population at risk when calculating the expected consequence of an accident as well as the bandwidth of buffers around highway segments used when determining total population at risk. More details on the calculation of the population at risk attributes is found in Section 5.

4.4. Least cost path

With the least cost path option, each of the five link attributes is weighted to produce a single composite link cost. As in Lepofsky et al. (1993) and Brainard et al. (1996), Dijkstra's algorithm is then applied to this cost to generate a single path. The user inputs the weights using a dialogue box. Due to the wide variation of measurement scales of each link attribute, weights are normalized. The normalization consists in dividing non-zero weighted attributes of each path by the minimum origin–destination path cost for that attribute. As weights are applied linearly, the solution path is invariant with respect to a scaling of the weight vector.

The least cost path solution method can conceivably use non-temporal accident and population at risk attributes only. If temporal attributes were used without a constraint on time, then a likely solution would involve excessive parking to avoid links with high cost during specific time windows. For instance, a vehicle assigned to a 500 mile ride could optimally be scheduled to travel for a few hours each day so as to take advantage of lower cost temporal attributes, thus unrealistically and impractically delaying delivery for several days. The approach proposed here avoids such unrealistic scenario.

4.5. Constrained least time path

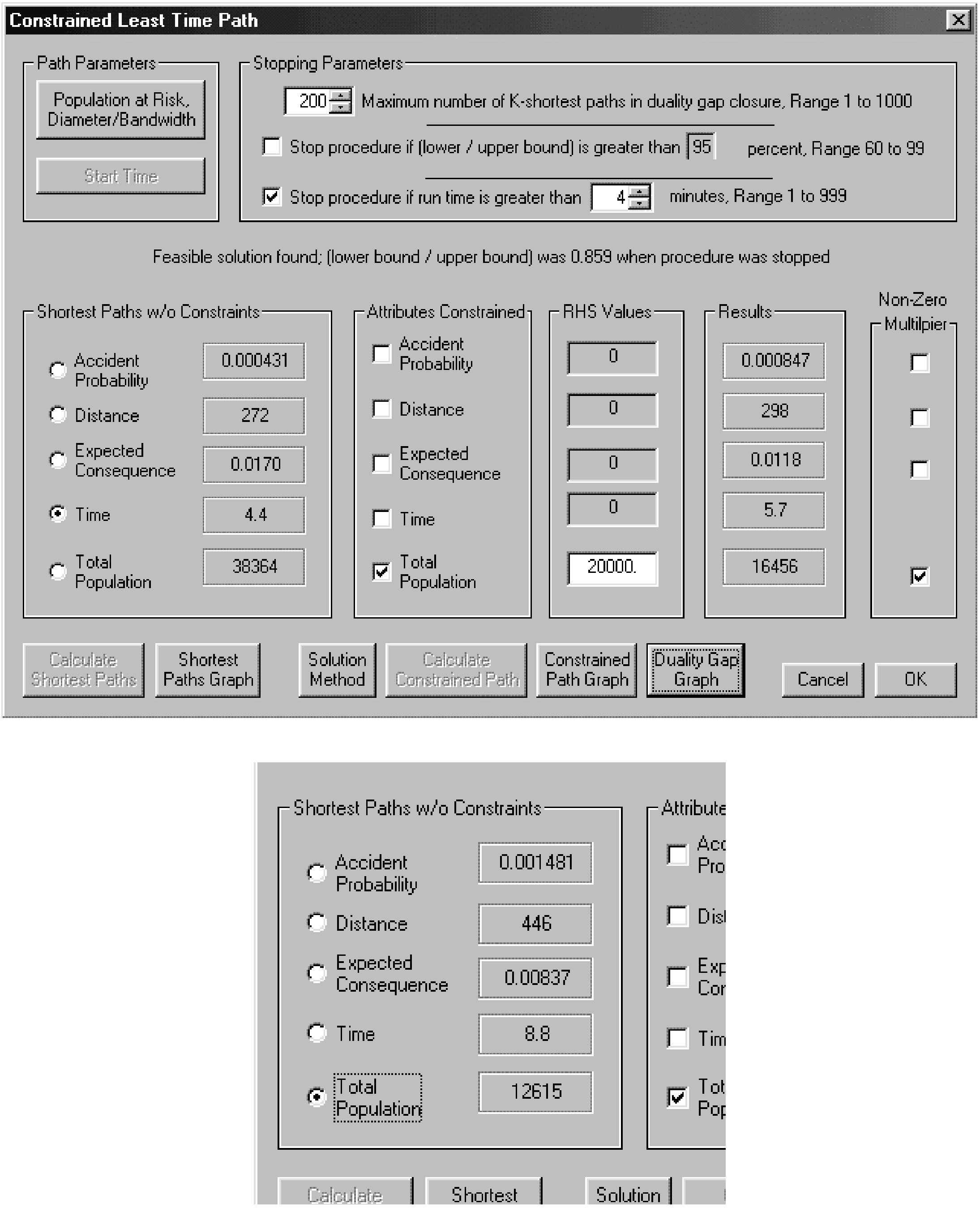
The constrained least time path problem has three different solution methods, all of which are extensions of Handler and Zang (1980). The first approach assumes non-temporal attributes. The remaining two solution methods allow for temporal attributes. They are the intermediate node method and the weight-guided solution method. Discussion of the latter methods is conducted in Section 6.

If a solution method with temporal link attributes is selected, then start time influences the route selected. Start time can be input manually in a dialogue box. The month of the year can also be supplied to account for seasonal fluctuations in the level of congestion. The dialogue box through which the constrained least time path problem is specified is shown in Fig. 5(a).

The lower left panel is labeled shortest paths without constraints. With the exception of the minimum time, non-temporal attributes are used to calculate the minimum link attributes. Fig. 5(a) displays the results describing the solution path in terms of each of the five attributes (accident probability, distance, expected consequence, travel time, and total population) when travel time is minimized, while Fig. 5(b) displays the corresponding results when total population is minimized. By clicking on the appropriate radio button, the numerical results can be displayed for different attribute minima. This feature assists the user in determining appropriate constraint bounds.

Solution paths can be depicted in the decision space by displaying the appropriate map object in the map window. As an aid to decision making, the performance of solutions vis-a-vis the objectives and constraints of the routing problem can be visualized by displaying them in the objective space. Spider webs are used in this research (see also Coutinho-Rodrigues et al., 1997). Routing performance results with travel time minimization are depicted in Fig. 6. In this graph, the corners of the dark gray polygon represent the minima of the five attributes. The corners of the light gray polygon represent the attribute sums when travel time is minimized. Fig. 7 shows the solution diagram for a scenario where total population is minimized. It should be pointed out that the dark gray polygon is the same as in Fig. 6 but the light gray polygon is very different. These graphs illustrate the conflict that typically exists between minimizing travel time and minimizing population at risk.

(a)



(b)

Fig. 5. (a) Constrained least time path dialogue box. (b) Total population minimized.

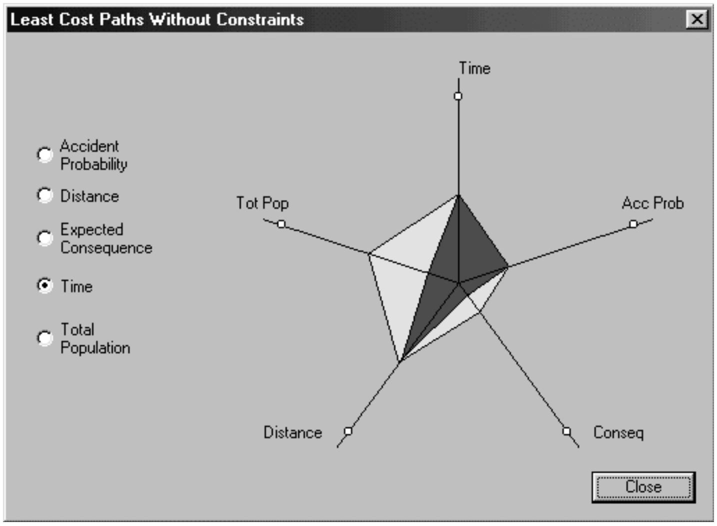


Fig. 6. Solution diagram with time minimized.

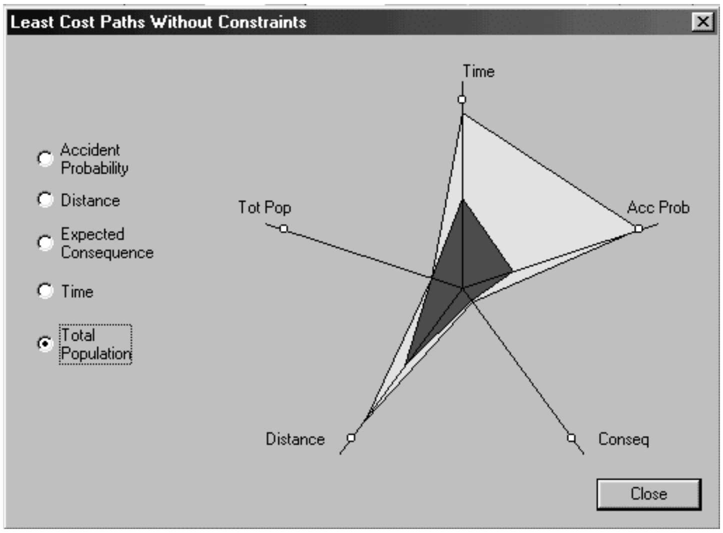


Fig. 7. Solution diagram with total population minimized.

The lower center panel of Fig. 5(a) is used to supply upper bounds on attribute sums. The numerical results are displayed in the lower right side of Fig. 5(a). Fig. 8 displays the results graphically. This figure is similar to Fig. 6 except that the constraint is added in the form of a dot on the total population axis. The attribute sums are shown in the form of a polygon without a colored interior. The differences between the light gray polygon and the uncolored polygon are caused by constraints.

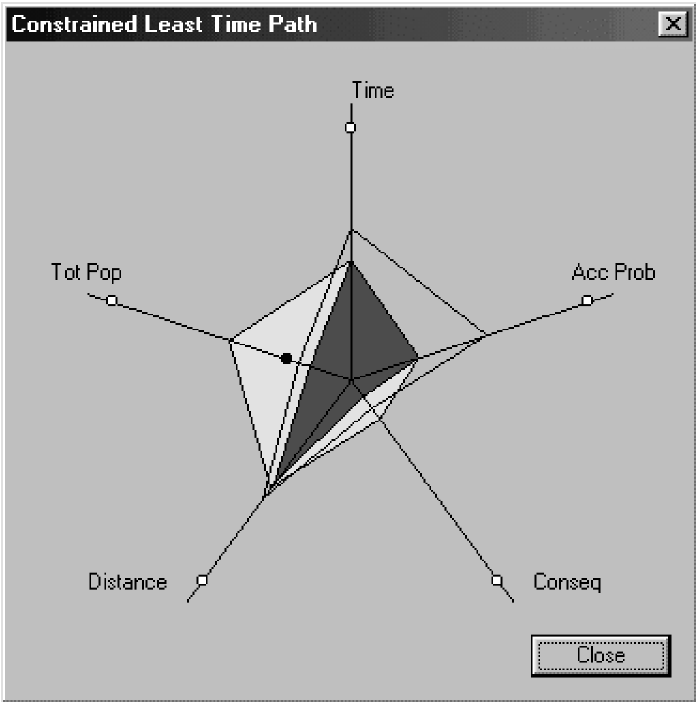


Fig. 8. Solution diagram of constrained least time path.

Parameters controlling the computer runtime can be specified by any of three different approaches in the upper center right panel of Fig. 5(a). The first method limits the number of paths generated when using Yen's K least cost algorithm (Yen, 1971). The second method stops the heuristic if the present solution is close enough to the best solution. The third method limits the total time in minutes.

4.6. Data structures

The Hazmat Path SDSS utilizes two data structure sets, namely link costs and map objects. The link costs set is used to determine the truck route. The development of these costs is presented in Section 5. To produce one TCSP route, Dijkstra's algorithm might be performed thousands of times. Therefore, to reduce computer runtime, this data is stored in RAM in a binary tree data structure. The tree is constructed only once when the program is initially started. The user has the option of choosing different solution methods. This option dictates the assumptions applied to the link attribute costs.

The second data structure set consists of information needed to draw a map. Most of this data are in the form of shape points used to draw roads and state boundaries. This data is stored on disk because of its large size. There are three complete sets of state boundaries. Each set has the same number of chains but the number of shape points varies dramatically. As the user zooms in or out, the resolution is adjusted by displaying a set with more or less shape points. The sets range in total file size from 300 KB to 8 MB. This decreases map drawing time with little noticeable decrease in map quality. Without different sets, multiple shape points may represent one pixel on a monitor. The state boundary chains are divided among 10 files, each of which covers a different region of the country. A file is read when at least one chain from that file is displayed. This prevents reading an entire state boundary data set every time state boundaries are redrawn. Road data sets are handled in a similar manner. There are three road data sets each divided into 45 regions. The sets range in total file size from 23 to 67 Mb. To reduce file read time, interstate chains are placed at the beginning of each of the files. With this structure, the entire file need not be read if the user wants to display interstate roads only.

The resolution-specific data sets are created from the same master data set by selectively removing shape points. Generalization of features in a data set proceeds as follows. If the change in direction at a shape point is below a given threshold, then the shape point is removed. Another approach for the pre-processing of chain features is based on the distance between adjacent shape points. If this distance is below a pre-defined threshold, then one of the shape points is removed.

5. Link attribute costs

In this section, we discuss how link costs are determined for the SDSS. While distance is a primary link cost, total and average population at risk, probability of an accident, and link travel time are secondary costs calculated from the distance attribute combined with appropriate domain-specific knowledge. The expected consequence link attribute is derived from average population at risk and probability of an accident. All secondary costs are time dependent.

5.1. Population at risk

There are two different approaches to determining population at risk:

1. Calculate total population at risk for a given bandwidth and link (Abkowitz et al., 1990).
2. Calculate average population at risk for a given radius of affected area and link (Saccomanno and Chan, 1985; Erkut and Verter, 1995).

With the first method, total population at risk, the population count within a buffer of given bandwidth is calculated for each link. The individual link values of total population at risk are summed to determine the total population at risk along a route from end to end. This method of calculating total population at risk has a flaw: people are double counted if they are located within half the bandwidth of the link end points. To avoid this double counting, a more elaborate method is developed as follows. A standard buffer is divided into three regions as shown in Fig. 9: regions A and C are the buffer drawn around the end points of the link, while region B is the remaining part of the standard buffer.

The total population at risk for a link is

$$\frac{1}{2} \text{ population in region A} + \text{ population in region B} + \frac{1}{2} \text{ population in region C.}$$

This method greatly reduces double counting but does not eliminate it. Indeed, if the B region of two links overlap, double counting occurs.

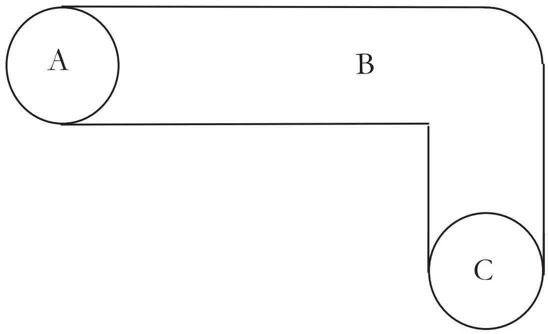


Fig. 9. Partitioned buffer.

The other population at risk metric is average population. It measures the average population at risk in case of accident on a link, under the assumption that the probability of an accident is constant along the length of the link. If an accident occurs while traveling a link, then the expected number of people exposed is the population within a given radius λ of the accident location. The value of the radius λ depends on the type of hazmat and the characteristics of the accident. Given that the location of future accidents is not known and that they are expected to occur with a probability that is uniformly distributed over the link, the expected number of people at risk can be approximated as follows. A series of circles are centered and equally spaced along the length of the link as shown in Fig. 10. The circle radius λ is taken to be constant in this implementation, but the methodology proposed here can accommodate a distribution of radii associated with a distribution of hazmat accidents drawn from historical records. The number of people within each circle is calculated by overlay with the layer of block group polygons produced by the Bureau of the Census and incorporated in the TIGER/Lines files. Their statistical mean produces the average population at risk. By design, the average population at risk calculation assigns more weight to people the closer they live to the link. A person nearer the link is more likely to be counted multiple times than a person further from the link. A similar discussion is held in Erkut and Verter (1995).

Average population at risk is of little use by itself. Once multiplied by the probability that an accident occurs on a link, it produces the expected consequence of a single truck accident.

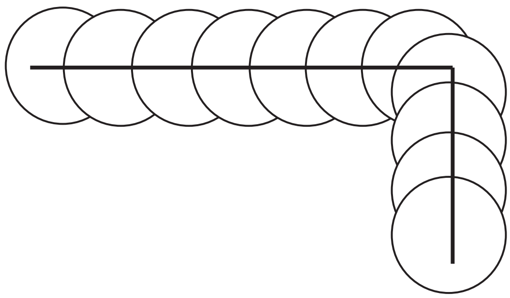


Fig. 10. Average population at risk.

Expected consequence is a measure of the expected number of people exposed on an origin–destination trip. Calculation of accident probabilities is discussed next.

5.2. Accident rates

Considerable research has been devoted to truck accidents, their forecast and measurement. Research usually focuses on a single geographic region, ranging in size from a metropolitan area to a state. Seldom is there any emphasis placed on the variation of the truck accident phenomenon by the hour, the day, or the month.

Mohamedshah et al. (1993) are interested in the non-temporal relationship between highway geometry and accident rates. Using data from Utah, they have identified horizontal curvature and vertical gradient as significant factors of truck accident rates. A highway safety information system summary report (US DoT, 1994a) determines truck accident rates using 1985–1987 Utah and Illinois accident data. The study also concludes that rates are greatly dependent on road type and state.

Hazmat routing in the Toronto area is analyzed by Saccomanno and Chan (1985) on the basis of 1981 Metropolitan Toronto police records. In this study, it is estimated that the probability of an accident on a 100 km/h dry urban expressway to be 2.379×10^{-6} per mile for unrestricted visibility and 4.054×10^{-6} for restricted visibility. Another study by Jovanis and Delleur (1983) also provides statistical estimates of accident rates. From accident records on the Indiana Tollway, they calculate the probability of an accident for a large truck with good weather to be 1.44×10^{-6} per mile for day travel and 1.47×10^{-6} night travel. Other temporal accident rates are derived by Lyles et al. (1991) for the state of Michigan. In addition, a few state agencies have collected, organized and analyzed accident data from their own state. Results of these limited studies exhibit tremendous inconsistencies and cannot reliably be transformed into national accident rates. Deficiencies in truck accident and truck usage data have previously been pointed out by other authors, including Lyles et al. (1991), Lepofsky et al. (1993), and Mohamedshah et al. (1993). Therefore temporal accident rates used in this study are derived from national databases, which offer great consistency and comprehensive geographic coverage at the expense of attribute specificity and detail.

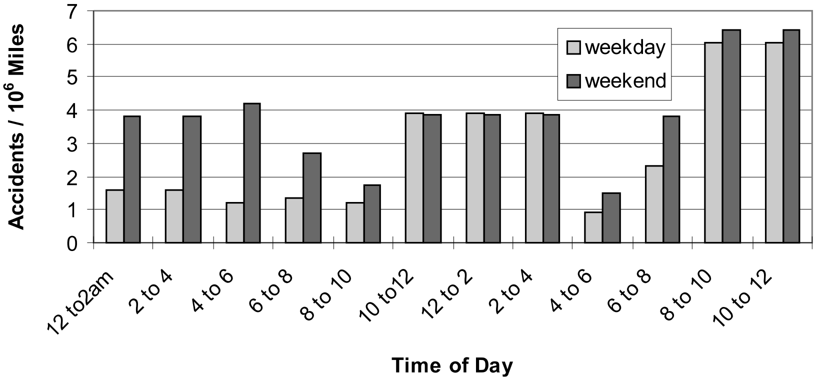


Fig. 11. Truck accident rates.

Three nationwide data sources are integrated to establish temporal accident rates:
1. Annual weighted accident police reports with type of roadway, time of day and day of week (National Highway Traffic Safety Administration, 1993).
2. Truck mileage by type of roadway for a one year period (FHWA, 1996).
3. Annual truck mileage by type of roadway, time of day and day of week (US DoT, 1992).

The second and third source are used to develop a nationally weighted truck mileage by type of roadway, time of day and day of week for a one year period. This statistic is then combined with the first source to determine accident rates by type of roadway, time of day and weekday/weekend. All data are relative to the 1993 calendar year. The results are shown in Fig. 11 for all types of roadways combined.

5.3. Link travel speeds

On most urban and all rural roads, the travel speed is not significantly affected by congestion. Therefore, on these roads, the travel speed can be assumed to be constant. Link travel time is calculated as the link length divided by the speed limit inferred from the road type. On the remaining urban roads where congestion is a problem, travel speed or time delay is time dependent. Ideally, a four-step travel demand methodology should be used to forecast link travel speeds in each metropolitan area afflicted with significant congestion. Given the enormity of such task in the context of the present research, an alternative method incorporating simulations is followed.

The approach to estimating travel time starts with a bivariate classification of links into groups with similar characteristics, namely average annual daily traffic (AADT) and lane width. A simulation is performed for each group to determine temporal travel speeds under expected recurrent and non-recurrent congestion conditions. The simulation follows the procedure outlined in US DoT (1986). Inputs to the model include freeway capacity reduction under incident conditions, incident frequency and average incident duration. There are five different lane widths and five AADT values for a total of 25 simulations. Linear interpolation is used for values between the five average AADT values. The average AADT on any link is estimated by a series of simple proportionality rules given the following link attributes: city name, month of the year, weekday or weekend, time of day, and link orientation with respect to the center of the metropolitan. AADT per lane for the 50 most congested metropolitan areas comes from the US DoT (1997). National averages are used to distribute the latter by month, weekday or weekend and, time of day.

Traffic directionality varies dramatically with the time of the day. More vehicles are heading towards a city in the morning, while more vehicles exit a city during the evening peak hours. These traffic patterns are modeled as a function of link orientation with respect to the city center by means of a directional flow. Equal flow in both directions is associated to an even 50% directional flow. Different directional flows are applied to links for morning and evening peak periods to capture the directionality of commuting patterns.

The directional flow is determined in two steps. The first step is determining the spatial relation between the center of an urban area and the direction of flow. Traffic flow angle is used as the metric in the first step. The traffic flow angle is converted into a directional flow in the second step.

The following algorithm calculates the traffic flow:
1. Draw a line through the link enter node to the link exit node. These are points A and C from the example in Fig. 12.

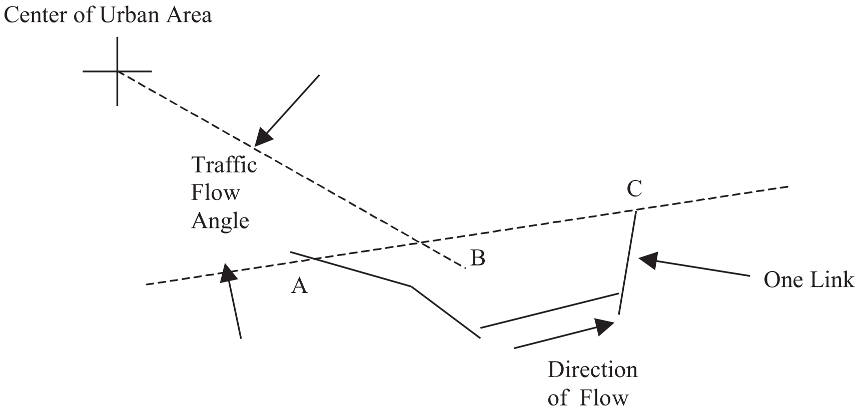


Fig. 12. Traffic flow angle diagram.

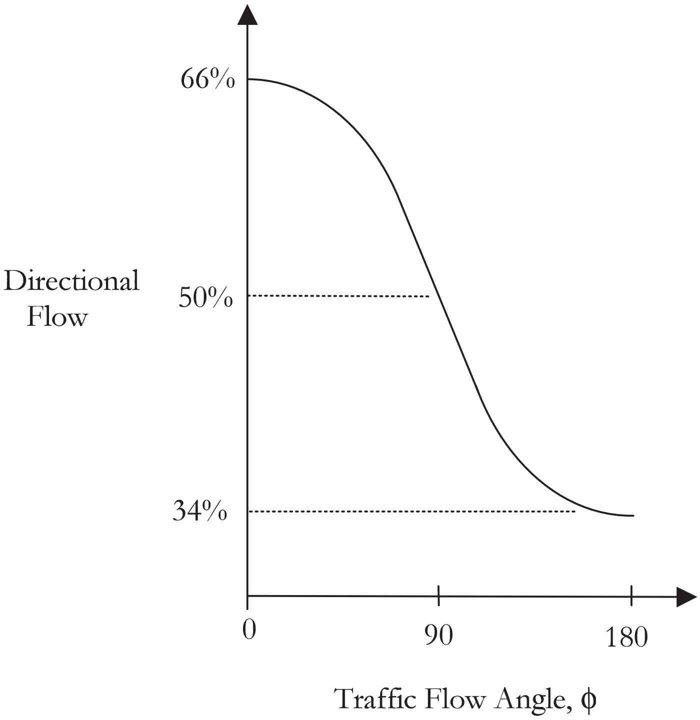


Fig. 13. Evening directional flow vs traffic flow angle.

2. Draw a line from the center of the urban area to a point midway between the entrance and exit nodes, point B.
3. The angle made by these two lines is the traffic flow angle.

For morning traffic, traffic varies directly with the traffic flow angle. The opposite is true for the evening. A sine function is used to convert traffic flow angle to directional factor. Proper scaling of the transform generates directional flows in the 34–66% range (Robinson et al., 1992). The traffic flow angle-directional factor relationship is given by

$$50 + 16\sin(\phi - 90^\circ)\% \quad \text{for mornings,}$$
$$50 + 16\sin(\phi + 90^\circ)\% \quad \text{for evenings,}$$

where ϕ is traffic flow angle, and depicted by Fig. 13. Notice the morning and evening functions are shifted 180°.

6. The temporal constrained shortest path

Costs (approximated by travel distance, travel time and risk exposure metrics) occur when traveling from the origin to the destination of the shipment. This section describes the methods implemented to evaluate these costs once the solution path has been identified. The mathematical formulation of the TCSP problem is also presented along with solution methods to solve it.

6.1. Problem formulation

Let path Π be an origin to destination path. Also, let us denote by δ_{ijt} the travel time on link *ij* when departing node *i* at time *t*. The departure time from node *i* is denoted by γ_i.

If parking is not allowed, then departure time is: $\gamma_j = \gamma_i + \delta_{ij\gamma_i}$. On the other hand, if parking is utilized, then departure time is $\gamma_j = \gamma_i + \delta_{ij\gamma_i} + D_j$, where D_j is the time parked on node *j*.

Let the probability of an accident on link (*i*,*j*) at time *t* be denoted by a_{ijt}. Summing up the costs for all the links on a path, the probability of an accident on Π is $\sum_{\forall(i,j)\in\Pi} a_{ij\gamma_i}$.

The average population at risk along link (*i*,*j*) at time *t* is denoted by s_{ijt}. The expected consequence on link (*i*,*j*) is $a_{ijt}s_{ijt}$. This assumes that an accident anywhere along link (*i*,*j*) is equally likely. The expected consequence on Π becomes $\sum_{\forall(i,j)\in\Pi} a_{ij\gamma_i}s_{ij\gamma_i}$.

The average population at risk along link (*i*,*j*) at time *t* is denoted by b_{ijt}, and the total population attribute associated with path Π is $\sum_{\forall(i,j)\in\Pi} b_{ij\gamma_i}$.

Distance is the only non-temporal link attribute. The total distance on Π is $\sum_{\forall(i,j)\in\Pi} d_{ij} \cdot d_{ij}$ is the distance when traveling link (*i*,*j*).

Total travel time is equal to the destination departure time (γ_{D}) with no parking on the destination.

Therefore, the TCSP takes the following formulation:

$$\begin{aligned}
&\text{Minimize } \gamma_D \\
&\text{Subject to: } \sum_{\forall(i,j)\in\Pi} d_{ij} \leqslant \Lambda_{\text{distance}} \\
&\qquad \sum_{\forall(i,j)\in\Pi} a_{ij\gamma_i} \leqslant \Lambda_{\text{acc prob}} \\
&\qquad \sum_{\forall(i,j)\in\Pi} b_{ij\gamma_i} \leqslant \Lambda_{\text{tot pop}} \\
&\qquad \sum_{\forall(i,j)\in\Pi} a_{ij\gamma_i}s_{ij\gamma_i} \leqslant \Lambda_{\text{conseq}} \\
&\qquad \gamma_j = \gamma_i + \delta_{ij\gamma_i} + D_j \qquad \text{for link } i,j \in \Pi,
\end{aligned}$$

where $\Lambda_{\text{distance}}$, $\Lambda_{\text{acc prob}}$, $\Lambda_{\text{tot pop}}$ and Λ_{conseq} are attribute bounds.

6.2. Solution methods

The TCSP incorporates a number of link attributes whose valuation varies with time of the day, and day of the week. In addition, parking or stopping on a node and continuance of the trip at a later time is allowed in the temporal network. Waiting to enter a link may be advantageous when the cost of traversing this link is expected to decrease in the future. Parking can be thought of as a cycle where the only non-zero link attribute is time. Cycles may exist in the optimum solution if parking is not allowed. If neither cycles nor parking is allowed, then a less obvious method of time consumption may exist. An example is given to illustrate this problem. Consider the network in Fig. 14.

Assume the travel time of subpath 1–3–2 is greater than the travel time from node 1 to node 2. Also assume all other link attribute costs for subpath 1–3–2 are greater than subpath 1–2 attribute costs. Subpath 1–2 will be chosen over subpath 1–3–2 if parking is allowed that subpath 1–2 dominates subpath 1–3–2. The temporal costs of the link from node 2 to the destination have no effect on which subpath is chosen.

Now, consider the case where parking is not allowed and the costs to travel link 2D decrease with time. Traveling subpath 1–3–2 will delay entering link 2D. This delay will decrease the costs of traveling link 2D. However, this delay comes at the increased costs of traveling subpath 1–3–2. Subpath 1–3–2 will be chosen if the total cost is less.

By its sheer size, the 57,000 node network derived from FHWA's NHPN network entails a major challenge to solving the TCSP. Even under the optimistic scenario of a 90% reduction in network size obtained by applying the constraints discussed in Section 4, the network is still too large to apply a temporal dynamic programming solution method. A solution method with exponential complexity would not be practical with such a large network, so a major concern of this research is the development of efficient heuristic procedures.

A workable strategy to get around the network size problem is to break up the problem into manageable parts. The primary difficulty of this approach is relating information between the manageable parts. Two solution methods are implemented in Hazmat Path, namely the so-called weight-guided solution method, and the intermediate nodes solution method. Both approaches are an extension of Handler and Zang (1980) and are now outlined.

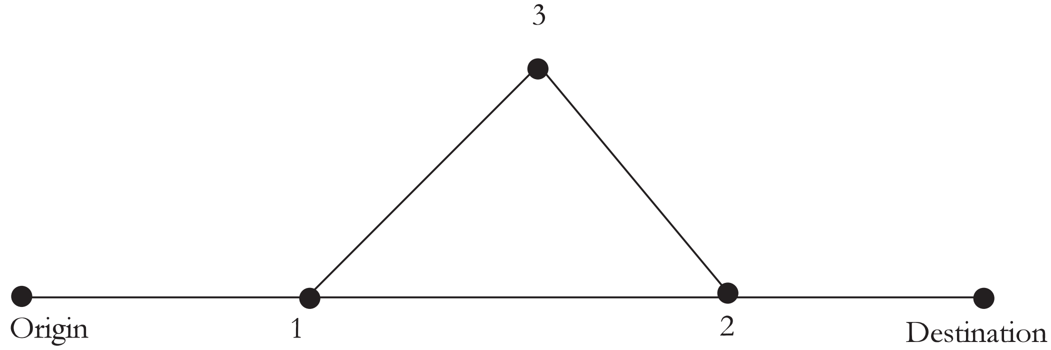


Fig. 14. Simple network.

The weight-guided solution method disentangles the original formulation into a master problem and a sub-problem. The master problem produces a path using Dijkstra's algorithm with a single link cost. This link cost is a linear combination of the five link attributes. The sub-problem determines the parking times at nodes. The attribute sums from this path along with the upper bounds on these attribute sums are inputs in determining the new single link cost. The master problem then produces another origin–destination path. This is repeated until there is no change in results between iterations.

The intermediate nodes solution method transforms the TCSP into many time-independent CSP problems. All link attributes are defined over discrete time interval, with the exception of travel time in congested cities which is a continuous function of time. The discrete time intervals range from 1 to 6 h. To implement this method, travel time in large urban areas is converted to a function of discrete time. The algorithm proceeds as follows. A vehicle starts at the origin and reaches a pre-determined intermediate node before any attributes change values. It stays at this parking position until attributes change value. Since no attributes change value while traveling from the origin to the intermediate node, a time-independent solution method can be used in this subpath route selection. The vehicle resumes its trip to another intermediate node or to the destination. This process is repeated as needed until the destination is reached.

7. Conclusions and future extensions

In this paper, we presented a working and easy to use hazmat routing SDSS that overcomes three significant challenges, namely handling a realistic network, offering sophisticated route generating heuristics and functioning on a desktop personal computer. Although many parts of this work can individually be found in previous work, never before have they been combined into one single working system.

A successful SDSS necessitates the development of custom software. Decision making is rendered considerable less cumbersome for several reasons. First, the user follows a logical procedure when developing a route. On the contrary, off-the-shelf software adapted to hazmat routing requires learning the general syntax of the software prior to delving into hazmat routing. Custom software also produces a route in a more timely manner before it incorporates efficient data structures and solution algorithms. The navigational simplicity and efficiency advantages help the decision-maker focus on creating solutions, negotiating trade-offs, and evaluating scenarios.

This paper outlines two solution methods to the TCSP problem implemented in Hazmat Path. Another possible solution method could be developed by combining the work of Handler and Zang (1980) and Lombard and Church (1993) approach to solving the gateway shortest path problem. This method can be outlined as follows. The best route generated from the gateway procedure is used as input to the Handler–Zang procedure, which calculates a set of link weights. These link weights are then applied to the network. The gateway procedure is run again and the best route is determined. This process is repeated until there is no improvement between iterations.

In Section 4, we discussed producing differentiated routes to spread risk over a larger population. In our approach, routes were constructed interactively by the decision-maker on the map window. A possible enhancement of the SDSS could involve adding a route generator to produce differentiated routes. Akgun et al. (1999) evaluated several methods for creating differentiated paths. The computational effort required to generate these paths is very low for some of these methods.[1] However, evaluating the quality of the set of differentiated paths would require considerable computational effort during route selection. This calculation consists of creating polygon overlays, which are used for determining the overlap of link buffers.

The display of temporal link attributes is another area where the current system could benefit from future enhancements. The SDSS presently displays attributes for one user-defined time period whereas a shipment may take a considerable amount of time, possibly several days. Displaying multiple maps of temporal attributes in a time loop would prove to be a useful decision support tool. Such capability would require creating, storing and retrieving multiple bitmaps.

Acknowledgements

The financial support of the National Science Foundation to the National Center for Geographic Information and Analysis (award number SBR-9600465) is gratefully acknowledged. Peer review of this manuscript was handled by Professor Hugh Calkins, Department of Geography, SUNY at Buffalo.

References

Abkowitz, M., Cheng, P.D.-M., Lepofsky, M., 1990. Use of geographic information systems in managing hazardous materials shipments. Transportation Research Record 1261, 35–43.

Akgun, V., Erkut, E., Batta, R., 2000. On finding dissimilar paths. European Journal of Operational Research 121, 232–246.

Baaj, M.H., Ashur, S.A., Chaparrofarina, M., Pijawka, K.D., 1990. Design of routing networks using geographic information systems: applications to solid and hazardous waste transportation planning. Transportation Research Record 1497, 140–144.

Brainard, J., Lovett, A., Parfitt, J., 1996. Assessing hazardous waste transport risks using a GIS. International Journal of Geographic Information Systems 10, 831–849.

Bureau of Transportation Statistics, 1999. National Transportation Atlas Databases, US Department of Transportation, Washington, DC.

Coutinho-Rodriques, J., Current, J., Climaco, J., Ratick, S., 1997. Interactive spatial decision-support system for multiobjective hazardous materials location-routing problems. Transportation Research Record 1602, 101–109.

Densham, P., 1991. Spatial decision Support Systems. In: Maguire, D.J., Goodchild, M.F., Rhind, D.W. (Eds.), Geographic Information Systems: Principles and Applications, vol. 1. Longman, London, pp. 403–412.

Erkut, E., 1996. The road not taken. OR/MS Today 23 (6), 22–28.

Erkut, E., Verter, V., 1995. A framework for hazardous materials transport risk assessment. Risk Analysis 15, 589–601.

FHWA, 1996. Annual Vehicle Miles of Travel and Related Data, US Department of Transportation, Publication No. FHWA-PL-96-024.

Geoffrion, A.M., 1983. Can OR/MS evolve fast enough. Interfaces 13, 10–25.

Handler, G.Y., Zang, I., 1980. A dual algorithm for the constrained shortest path problem. Networks 10, 293–310.

Jovanis, P.P., Delleur, J., 1983. Exposure-based analysis of motor vehicle accidents. Transportation Research Record 910, 1–7.

[1] Lombard and Church (1993) report similar experience.

Lassarre, S., Fedra, K., Weigkricht, E., 1993. Computer-assisted routing of dangerous goods for haute-normandie. Journal of Transportation Engineering 119, 200–210.
Lepofsky, M., Abkowitz, M., Cheng, P.D.-M., 1993. Transportation hazard analysis in integrated GIS environment. Journal of Transportation Engineering 119, 239–254.
Lombard, K., Church, R.L., 1993. The gateway shortest path problem: generating alternative routes for a corridor location problem. Geographical Systems 1 (1), 25–45.
Lyles, R.W., Campbell, K.L., Blower, D.F., Stamatiadis, P., 1991. Differential truck accident rates for Michigan. Transportation Research Record 1322, 62–69.
Mohamedshah, Y.M., Paniati, J.F., Hoheika, A.G., 1993. Truck accident models for interstate and two-lane rural roads. Transportation Research Record 1407, 35–41.
National Highway Traffic Safety Administration, National Accident Sampling System General Estimates System, National Center for Statistics and Analysis, US Department of Transportation.
Nozick, L.K., List, G.F., Turnquist, M.A., 1997. Integrated routing and scheduling in hazardous materials transportation. Transportation Science 31 (3), 200–215.
Robinson, C.C., Levinson, H.S., Goodman, L., 1992. Capacity in transportation planning, In: Edwards, J.D. (Ed.), Transportation Planning Handbook. Prentice-Hall, Englewood Cliffs NJ, pp. 410–446.
Saccomanno, F.F., Chan, A.Y.-W., 1985. Economic evaluation of routing strategies for hazardous road shipments. Transportation Research Record 1020, 12–18.
Souleyrette, R.R., Sathisan, S.K., 1994. GIS for radioactive materials transportation. Microcomputers in Civil Engineering 9, 295–303.
Sprague, R.H., Carlson, E.D., 1982. Building Effective Decision Support Systems. Prentice-Hall, Englewood Cliffs, NJ.
Turban, E., 1995. Decision Support and Expert Systems: Management Support Systems. Prentice-Hall, Englewood Cliffs, NJ.
US Department of Transportation, 1986. Quantification of Urban Freeway Congestion and Analysis of Remedial Measures, Report No. FHWA/RD-87/052.
US Department of Transportation, 1992. 1990 Nationwide Truck Activity and Commodity Survey, ORNL, TM-12188.
US Department of Transportation, 1994a. Highway Safety Information System Summary Report, Truck Accident Models, FHWA-RD-94-022.
US Department of Transportation, 1994b. Research and Special Programs Administration, Biennial Report on Hazardous Materials Transportation Calendar Years 1992–1993.
US Department of Transportation, 1997. Estimates of Urban Roadway Congestion – 1990, DOT-T-94-01, March 1993 (update to 1993 data).
Yen, J., 1971. Finding the K shortest loopless paths in a network. Management Science 17, 712–716.