ELSEVIER

Accurate calculation of hazardous materials transport risks

B.Y. Kara[a], E. Erkut[b,*], V. Verter[c]

[a] Department of Industrial Engineering, Bilkent University, Bilkent, Ankara 06533, Turkey
[b] School of Business, University of Alberta, Edmonton, Alberta, Canada T6G 2R6
[c] Faculty of Management, McGill University, Montreal, Quebec, Canada H3A 1G5

Abstract

We propose two path-selection algorithms for the transport of hazardous materials. The algorithms can deal with link impedances that are path-dependent. This approach is superior to the use of a standard shortest path algorithm, common in the literature and practice, which results in inaccuracies.
© 2003 Elsevier Science B.V. All rights reserved.

Keywords: Hazardous materials transport; Shortest path

1. Introduction

Almost all papers that deal with the selection of a minimum risk path for transport of a hazardous material (hazmat) reduce the path selection problem to a shortest path problem. The use of a standard shortest path algorithm requires that the impedance of each transport link be known and independent of the impedances of other links. However, this is not the case for hazmat transport problems and simplifications are necessary to render the problem solvable using a standard shortest path algorithm. In this paper we discuss the inaccuracies that result from imposing the standard shortest path model on the hazmat transport problem. We also present two methods, which negate the need to make simplifications. One

of the proposed procedures is a modified version of a well-known shortest path algorithm, and the other is an adaptation of a link-labeling algorithm developed for urban transportation.

Consider the problem of selecting a path for a shipment of dangerous goods between a pre-specified origin-destination pair. Let $N = \{1, \ldots, n\}$ denote the node set of the transportation network. Link (i, j) connects nodes i and j in N. Let P denote a feasible path for this shipment. For the ease of exposition, we assume that the nodes in P are indexed sequentially, i.e. $P = \{1, 2, 3, \ldots, r\}$, where node 1 represents the origin and node r represents the destination.

What differentiates hazmat transport models from other transport models is the explicit modeling of transport risk which usually consists of one or both of the following two factors: incident (i.e. spill, fire) probability and population impacted. Let p_{ij} denote the probability of having an incident on link (i, j), and

* Corresponding author. Tel.: +1-780-492-3068; fax: +1-780-492-3325.
E-mail address: erhan.erkut@ualberta.ca (E. Erkut).

let C_{ij} denote the total number of people who live within a given threshold distance of link (i,j).

Using this notation we list two objectives for selecting a hazmat path that are popular among academics and practitioners.

(1) Minimize total incident probability (for example [8]):

$$p_{12} + (1 - p_{12})p_{23} + (1 - p_{12})(1 - p_{23})p_{34}$$
$$+ \cdots + (1 - p_{12})\cdots(1 - p_{r-2,r-1})p_{r-1r}. \quad (1)$$

(2) Minimize total population exposure (for example [7]):

$$C_{12} + C_{23} + C_{34} + \cdots + C_{r-1r}. \quad (2)$$

We now discuss the simplifications made when using these objectives, the resulting inaccuracies, and how one could avoid these.

2. Modeling and minimizing incident probability

According to (1), the incident probability of a path is computed by adding the incident probabilities along each link of that path. The probability of incident on a given link, however, depends on the incident probabilities of all links leading up to that link. Hence, the incident probabilities on links are path-dependent. While it is possible to formulate a nonlinear integer programming model for selecting the minimum probability path, such a formulation is of little use for practical purposes. Most papers that deal with (1) use a simplification that turns this complicated optimization problem to a shortest path problem: product of incidence probabilities can be approximated by zero. This assumption is justified by the magnitude of incident probabilities (usually on the order of 10^{-6} incidents per mile). Erkut and Verter [4] point out that this approximation is likely to result in a very small error (less than 0.25% in most cases) in measuring the incident probability along a path. While this error is acceptable for most practical purposes, it is rather easy to produce error-free estimates.

We propose an extension of Dijkstra's [2] node-labeling shortest path algorithm to find a minimum incident probability path. Let $q(i)$ denote the probability of safely arriving at node i of path P. Note that this probability is dependent on the previous links

of the path. Observe that

$$q(i + 1) = q(i)(1 - p_{ii+1}).$$

Our algorithm adjusts the link impedances (incident probabilities) at each iteration by multiplying them with the probability of safely arriving at the starting node of the arc. Although shortest path algorithms for dynamically adjusted link lengths have been proposed in the operational research literature for other problems (for example [3,9]) we know of no reference to them in the hazmat transportation literature.

At a given iteration of the algorithm, let $\theta(i)$ denote the incident probability of the current minimum incident probability path to node i. Since $\theta(i) + q(i) = 1$, it is straight forward to show that the optimality principle holds. This allows us to determine the optimal path via a node-labeling algorithm. Let a_{ij} denote the increase in total incident probability due to the addition of link (i,j) to the current path. In standard shortest path applications a_{ij} is a problem parameter that does not depend on the predecessor node. For our problem a_{ij} depends on the predecessor node: $a_{ij} = q(i)p_{ij}$. The algorithm must compute the $q(i)$ values at each iteration since they depend on the path leading up to node i. Let $pred(i)$ denote the predecessor of node i.

Impedance-Adjusting Node-Labeling Shortest Path Algorithm:

Initialize $S = \{\}, S' = N, \theta(i) = \infty \ \forall i,$
　$\theta(1) = 0, pred(1) = 1, q(1) = 1.$
While $r \notin S$;
　　Let i in S' be such that $\theta(i) = \min\{\theta(j) : j \in S'\}$
　　$S = S \cup \{i\}, S' = S' \setminus \{i\}$
　　For each adjacent link (i,j) to node i
　　　　Calculate $a_{ij} = q(i)p_{ij}$
　　　　If $\theta(j) > \theta(i) + a_{ij}$ then $\theta(j) = \theta(i) + a_{ij},$
　　　　　$pred(j) = i$, and $q(j) = q(i)(1 - p_{ij}).$

The minimum incident probability for the origin-destination path is $\theta(r)$ at termination. Note that, while this algorithm computes the a_{ij} and $q(j)$ values on-the-fly, its computational complexity is the same as that of a standard node-labeling algorithm.

3. Modeling and minimizing population exposure

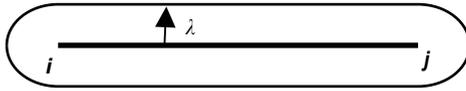Consider a hazmat truck moving on a link. The impact area of an incident is usually assumed to be a

Fig. 1. Semicircular exposure zone around link (i, j).

circle with a substance-dependent radius centered at the incident location. Consider the union of all such circles centered at all points on link (i, j). We refer to this area as the "semicircular exposure zone". Fig. 1 shows the semicircular exposure zone of radius λ for link (i, j).

Note that it is possible to preprocess the population density data and compute C_{ij} for all links in the transport network. This would facilitate the use of a standard shortest path algorithm, such as Dijkstra [2], to find the path that minimizes the population exposure. This is rather tempting since it can be accomplished easily using a geographical information system (GIS) such as ArcView [5]. However, such an approach would overestimate the population exposure of all paths in the network to varying degrees and may result in the selection of a suboptimal path.

To demonstrate the overestimation, consider a simple example: a path with only two links, $P = \{1, 2, 3\}$ where the two links intersect forming an acute angle. The semicircular exposure zone of this path is given in Fig. 2a and the semicircular exposure zones of the two links are shown in Fig. 2b.

Observe that the semicircular exposure zone of the path, given in Fig. 2a, is not the sum of the semicircular exposure zones of the two links (1,2) and (2,3), given in Fig. 2b. In finding the population exposure of this path, a standard shortest path algorithm would add the population exposure figures of the two links. Clearly,

this would double-count the population in the shaded area in Fig. 2b resulting in an overestimation.

It is possible to reduce the error by using exposure zones in the shape of rectangles around the links as shown in Fig. 3. This is the method used by PC*HazRoute [1], a special software developed for hazmat transportation, to compute population exposure.

It is worth noting that if one uses rectangular exposure zones one would be underestimating the exposure at the origin and the destination points due to the lack of the half circles at these points. However, this omission (of a constant) does not impact the optimization problem for path selection, and the error can be corrected easily by adding the omitted exposure after finding the optimal path.

More importantly this representation also overestimates the population exposure at intersection points, though not as badly as the semicircular representation. To demonstrate the overestimation, consider the example in Fig. 4 where two links (1,2) and (2,3) are perpendicular to each other.

With the rectangular exposure zones, the people living in the area marked with A are not counted as being impacted by the transport activity, whereas those living in the area marked with B are counted twice (one for each rectangle). If we assume the population density around Node 2 is uniform, the double-counting in B negates the exclusion in A. However, there is still double-counting in C; the area that is double counted is: $\lambda^2 - \pi\lambda^2/4$.

We can express the errors resulting from using a semicircular or rectangular exposure zone at the intersection of two links (k, i) and (i, j) as a function of the population density at the intersection $(\rho(i))$, the
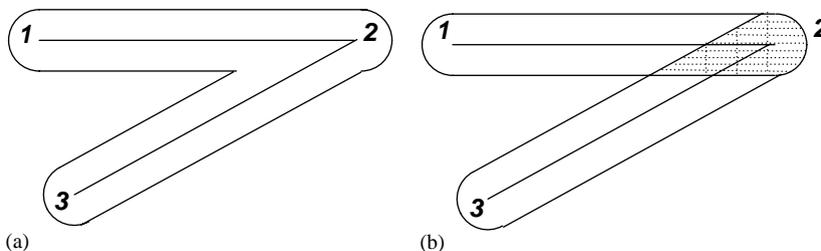

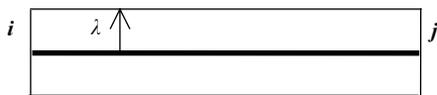Fig. 2. (a) Semicircular exposure zone around the path and (b) semicircular exposure zones of the two links.
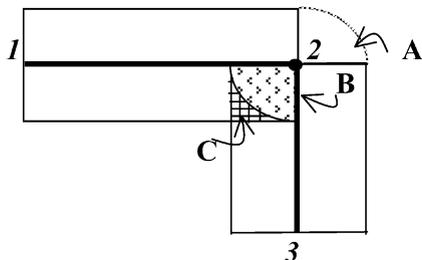
Fig. 3. Rectangular exposure zone around link (i, j).

Fig. 4. Rectangular exposure zones for a path.

threshold distance for exposure (λ) and the angle between the two links (α). For the rectangular exposure zone, the error is

$$\varepsilon_{ki,ij} = [\lambda^2/\tan(\alpha/2) - (180 - \alpha)\pi\lambda^2/360]\rho(i) \qquad (3)$$

and for the semicircular exposure zone the error is equal to the above expression plus the number of people living in the two semicircles at the intersecting node, namely $\rho(i)\pi\lambda^2$.

As per (3) the increase of the error is linear in ρ and quadratic in λ. As one would expect, the error for the rectangular representation goes to zero as the angle approaches 180° (no double-counting if the links are lined up perfectly), and it increases drastically as the

angle becomes smaller. Fig. 5 shows the errors as a function of α for $\lambda = 1$ and $\rho = 100$ (1 km threshold and a population density of $100/\text{km}^2$).

The semicircular representation can lead to fairly significant errors due to the double-counting at the nodes. It is clear that the rectangular representation is associated with lower error figures and in many cases the resulting errors may be negligible. However, the errors associated with rectangular representation can reach nontrivial values depending on the population density around the link intersections and the angle between the link pairs. For example, within Montreal there is a major highway junction where six links intersect. Using population figures of 1996, the overestimation for a link pair at this junction can be as high as 797 people for $\lambda = 800$ m resulting in a relative error of 1.3%. The intersection of four highway segments at Sainte-Foy (Quebec) offers an example where the overestimation reaches a relative error of 3% for $\lambda = 800$ m. We provide a complete numerical example in the appendix that demonstrates the errors associated with the two population exposure zone representations on the highway network of Southwestern Ontario.

The error term is a function of three parameters, and we could calculate it for all pairs of adjacent links during preprocessing. This allows us to use initial population exposures as algorithm input and then correct for the double-counting "on the fly." We note that we are only correcting for the overlaps between adjacent links, and not for those between nonadjacent links,
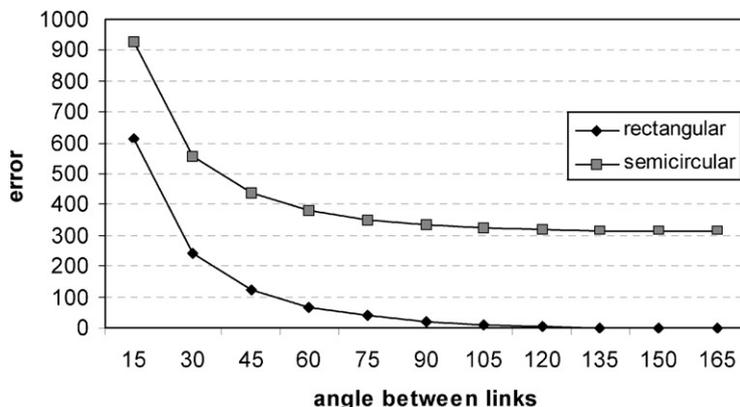

Fig. 5. The double-counting error (in persons) for the semicircular and the rectangular exposure zones as a function of the angle between two adjacent links for a fixed distance threshold ($\lambda = 1$ km) and fixed population density ($\rho = 100/\text{km}^2$).
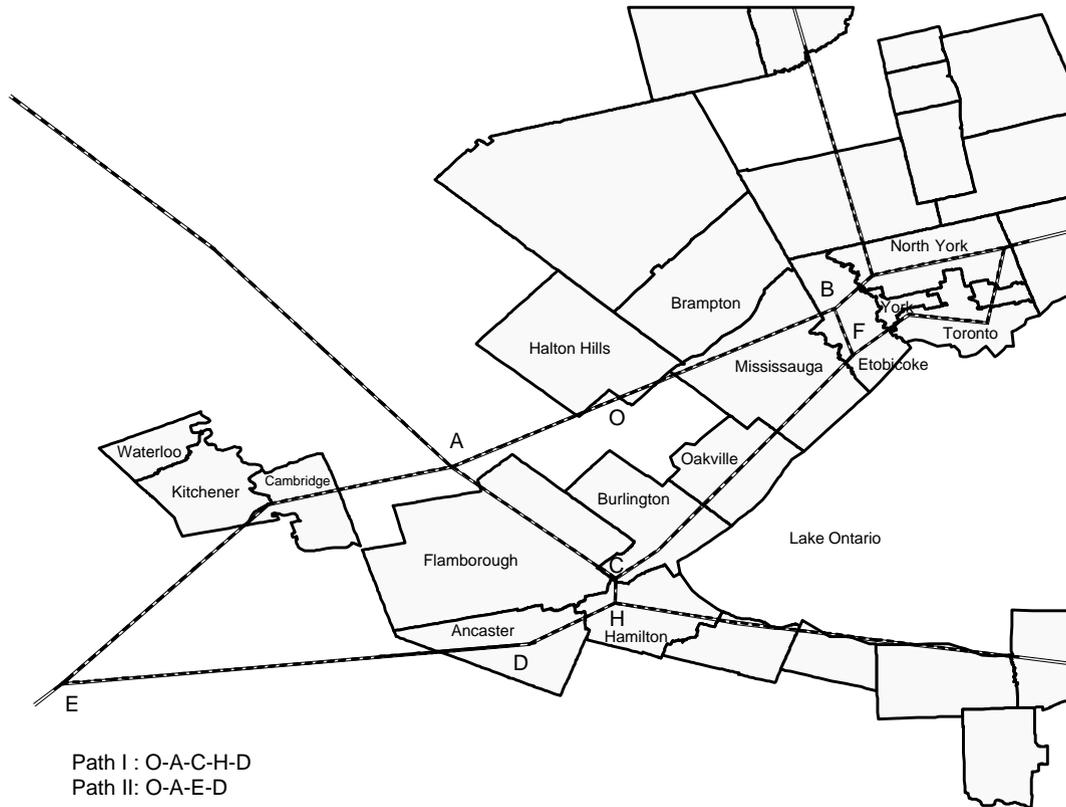
Fig. 6. The population centers and highway network of Southwestern Ontario.

which may arise if the links are short and the exposure zone is wide. We now propose a link-labeling shortest path algorithm to find the minimum population exposure path. Our algorithm is an adaptation of Namkoong et al. [6], which was developed for computing shortest paths in urban networks with turn penalties. We replace the turn penalty with our error term and modify the link-label update step accordingly. It is important to point out that a node-labeling algorithm would not work here, because the optimality principle is violated due to the error terms.

At a given iteration of the algorithm, let $\theta(i, j)$ denote the length of the current minimum exposure path from the origin to link (i, j). Let $pred(i, j)$ denote the predecessor of link (i, j). Recall that C_{ij} is the population exposure for link (i, j) and $\varepsilon_{ki, ij}$ is the error term for links (k, i) and (i, j). The algorithm works for both exposure zones: rectangular or semicircular. We compute $C_{ij} = 2\lambda l_{ij}\rho_{ij}$ for rectangular exposure zones and $C_{ij} = 2\lambda l_{ij}\rho_{ij} + 1/2\pi\lambda^2(\rho(i) + \rho(j))$ for semicir-

cular zones, where l_{ij} is the link length and ρ_{ij} is the population density around the link. Likewise, the error terms depend on the type of exposure zone used (see (3)).

Impedance-adjusting link-labeling shortest path algorithm:

Create an artificial source node s, an artificial destination node d, and two additional links $(s, 1)$ and (r, d) with $C_{s1} = C_{rd} = 0$.

Initialize $L = \{(s, 1)\}, \theta(i, j) = \infty \ \forall (i, j)$,
  $\theta(s, 1) = 0, \varepsilon_{s1, 1j} = 0 \ \forall j$ adjacent to 1,
     Next-Link $= (s, 1)$. Label $(s, 1)$ as "permanent" and every other link as "temporary".

Step 1: For every temporary link (j, k) adjacent to
  Next-Link(i, j),
     If $\theta(j, k) > \theta(i, j) + C_{ij} - \varepsilon_{ij, jk}$ then
     update $\theta(j, k) = \theta(i, j) + C_{ij} - \varepsilon_{ij, jk}$ and
        $pred(j, k) = (i, j) \ L = L \cup (j, k)\}$

Table 1
The population exposure calculations

Links	Path I	Path II	Exposed population within 800 m	
			Semicircular	Rectangular
153	*	*	226	93
152	*	*	216	153
151	*	*	100	22
18	*	*	0	0
21	*		52	19
176	*		2092	1752
293	*		2469	2013
67	*		905	347
178	*		689	268
295	*		575	446
23	*		1655	1512
301	*		888	698
195		*	40	2
194		*	1535	946
302		*	1716	1217
190		*	1957	1131
189		*	1580	858
188		*	1063	509
269		*	0	0
28		*	0	0
184		*	0	0
185		*	0	0
186		*	0	0
187		*	0	0
299		*	129	59
300		*	1609	1519

Link pairs	Path I	Path II	Corrections		
			Semicircular	Rectangular	Angle
153–152	*	*	109		
152–151	*	*	99		
151–18	*	*	0		
21–176	*		48		
176–293	*		401		
293–267	*		859	16	123
267–178	*		436		
178–295	*		103	2	119
295–23	*		150		
23–301	*		193	0	162
195–194		*	40		
194–302		*	877	3	151
302–190		*	448		
190–189		*	961		
189–188		*	874		
299–300		*	96		

Table 2
Summary of the results

	Without correction		With correction	
	Semicircular	Rectangular	Semicircular	Rectangular
Path I	9867	7485	7469	7467
Path II	10 171	6671	6667	6668

Step 2: If $L = \emptyset$ or *Next-Link* $= (r, d)$ then STOP.
  Else
    Remove *Next-Link* from L,
    Select new *Next-Link* $= (j, k)$ where $\theta(j, k) = \min\{\theta(i, j) : (i, j) \in L\}$.
    Label *Next-Link* permanent. Go to Step 1.

The minimum number of people exposed on the origin-destination path is $\theta(r, d)$ for the semicircular zones and $\theta(r, d) + \pi\lambda^2(\rho_1 + \rho_r)/2$ for rectangular zones.

4. Concluding remarks

To conclude, we point out that impedance-adjusting shortest path algorithms can effectively remove errors in quantifying path impedances for the popular hazmat route selection methods considered in this paper. In the case of incident probabilities the errors result from a simplification in the original model, whereas in the case of population exposure, the errors result from the mistreatment of input data (compounded by the network topology). While both types of errors are rather small (unless the semicircular exposure representation is used to quantify population exposure), there is no reason for tolerating such errors since it is very easy to eliminate them by using appropriate algorithms.

Acknowledgements

Appendix

Fig. 6 depicts the population centers and the highway network of Southwestern Ontario, Canada. As an illustrative example, we focus on the shipments between Halton Hills and Ancaster. In 1998, there were 577 fuel oil trucks, 200 gasoline trucks and 1382 petroleum trucks shipped between these two population centers. We consider the gasoline and fuel oil trucks, which require an 800-m evacuation zone in case of fire.

Although there are a number of alternative paths between Halton Hills and Ancaster, two of them dominate the others when a shortest path algorithm is implemented by using population exposure as the arc impedance. These two paths are denoted as Path I (O-A-C-H-D) and Path II (O-A-E-D) in Fig. 6. Table 1 shows the details of population exposure calculations. The links with zero population exposure correspond to rural segments of Highway 401. Table 2 summarizes the results of our comparative analysis. When population exposure estimates are not corrected, Path I is selected with semicircular exposure zones, and Path II is selected with rectangular exposure zones. Not only does the use of the semicircular zone result in a significant overestimation, it also results in the selection of the wrong path. In contrast, using the appropriate corrections produces error-free results regardless of the type of exposure zone used—the differences between the figures for the two types of exposure zones are due to round-off errors.

References

[1] ALK Associates, ALK's PC*HazRoute (Version 2.0), ALK Associates, Inc., 1000 Herrontown Road, Princeton, NJ, USA, 1994.

[2] E. Dijkstra, A note on two problems in connection with graphs, Numer. Math. 1 (1959) 269–271.

[3] S.E. Dreyfus, An appraisal of some shortest-path algorithms, Oper. Res. 17 (1969) 395–412.

[4] E. Erkut, V. Verter, Modeling of transport risk for hazardous materials, Oper. Res. 46 (5) (1998) 625–642.

[5] ESRI, ArcView 3.1, ESRI Inc., 380 New York Street, Redlands, CA, USA, 1996.

[6] S. Namkoong, J.H. Rho, J.U. Choi, Development of the tree-based link labeling algorithm for optimal path-finding in urban transportation networks, Math. Comput. Modelling 27 (9–11) (1998) 51–65.

[7] C. ReVelle, J. Cohon, D. Shobrys, Simultaneous siting and routing in the disposal of hazardous wastes, Transp. Sci. 25 (2) (1991) 138–145.

[8] F.F. Saccomanno, A. Chan, Economic evaluation of routing strategies for hazardous road shipments, Transp. Res. Record 1020 (1985) 12–18.

[9] K. Sung, M.G.H. Bell, M. Seong, S. Park, Shortest paths in a network with time dependent flow speeds, Eur. J. Oper. Res. 121 (2000) 32–39.