Available online at www.sciencedirect.com
ScienceDirect
Computers & Operations Research 34 (2007) 1304–1327
computers & operations research
www.elsevier.com/locate/cor

Finding minimum and equitable risk routes for hazmat shipments☆

Pasquale Carotenuto[a], Stefano Giordani[b,*], Salvatore Ricciardelli[b]

[a]*Istituto di Tecnologie Industriali e Automazione - Sezione di Roma, Consiglio Nazionale delle Ricerche, Via del fosso del cavaliere 100, I-00133 Roma, Italy*
[b]*Dipartimento di Ingegneria dell'Impresa, Università di Roma "Tor Vergata", Via del Politecnico 1, I-00133 Roma, Italy*

Available online 25 July 2005

Abstract

This paper deals with the generation of minimal risk paths for the road transportation of hazardous materials between an origin–destination pair of a given regional area. The main considered issue is the selection of paths that minimize the total risk of hazmat shipments while spreading the risk induced on the population in an equitable way. The problem is mathematically formulated, and two heuristic algorithms are proposed for its solution. Substantially, these procedures are modified versions of Yen's algorithm for the *k*-shortest path problem, which take into due consideration the risk propagation resulting from close paths and spread the risk equitably among zones of the geographical region in which the transportation network is embedded. Furthermore, a lower bound based on a Lagrangean relaxation of the given mathematical formulation is also provided. Finally, a series of computational tests, referring to a regional area is reported.
© 2005 Elsevier Ltd. All rights reserved.

Keywords: Hazardous materials; Transportation planning; Lagrangean relaxation; Heuristics

☆ This work has been partially supported by Grant CNR 02.00171.37 from the Italian National Research Council (CNR-GNDRCIE), and commissioned by *Protezione Civile* the Italian Civil Protection Body.
* Corresponding author. Tel.: +39 06 7259 7358; fax: +39 06 7259 7305.
E-mail addresses: p.carotenuto@itia.cnr.it (P. Carotenuto), giordani@disp.uniroma2.it (S. Giordani), ricciardelli@disp.uniroma2.it (S. Ricciardelli).

0305-0548/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.cor.2005.06.003

1. Introduction

The transportation of hazardous materials (hazmat shipment) is an important problem in industrialized societies. Among the different means used to transport hazardous materials, the road system represents an increasingly pressing problem due to the constant increase of the amount of hazmat shipments.

In this regard, one of the main objectives of research in this field is to provide appropriate answers to the safety management of dangerous goods shipments, in collaboration with the principal parties involved in the goods transportation process.

Research in this area focuses on two main issues. The first one is related to assessing the risk induced on the population by hazmat vehicles traveling on various segments of the road network, and the second one involves the selection of the safest routes to take.

A lot of work in risk assessment has already been done by modeling risk probability distribution over given areas, for example, by taking into account the risk related to the transported substance and the transport modality [1] as well as the environmental conditions [2]. Moreover, as useful tools, map algebra techniques from Geographic Information Systems allow us to combine mathematically the concentration of hazardous material releases into the environment with population distribution in order to estimate the risk when airborne contamination happens [3]. Given the incident probabilities on unit segments of a network, an analysis of different risk models associated to a route is given in [4]. In that paper, it is highlighted that one of the most popular risk model used by researchers and practitioners is the *societal risk*, this being the product between the incident probability per unit length and the incident consequence, which is evaluated as the population in the *impact area*. One widely used assumption, based on the λ-neighborhood concept, is that the impact area is a circle centered in the incident location with a substance-dependent radius λ [5].

The main problem related to this issue is that of finding minimum risk routes while limiting and spreading the risk equitably over any zone. As a matter of fact, risk equity has to be taken into account whenever several hazmat shipments take place from a given origin to a given destination. In this situation, the planning effort not only has to be directed towards minimizing the total risk, but also has to be devoted to distributing risk uniformly over all the zones of the geographical crossed region. This concept is well defined by Keeney in [6], where a measure of the collective risk is determined with explicit reference to equity.

In the literature, some models have been proposed for determining paths of minimum total risk while guaranteeing equitable risk spreading; see, for example, the models in [7] and [8]. Another model has been proposed by Gopalan et al. in [9], where the risk is computed by considering an exposure area around each link (impact area) depending on the shipped substance, and an equitable distribution of the risk is assured by partitioning the region into zones and by limiting, for each couple of zones, the difference between the risks induced on the population of the two zones. Moreover, the authors also show that a high degree of equity may be achieved by modestly increasing the total risk.

The concept of dissimilar paths has also been considered in order to guarantee the spreading of risk [10]. In this regard, different methods have been proposed in the past, also in contexts different from that of hazmat shipments, to generate a number of spatially dissimilar paths.

The Iterative Penalty Method is based on an iterative application of an appropriate shortest path algorithm. After each iteration of the procedure, a cumulative penalty is applied to each selected link.

As a result, repeated selection of the same links is discouraged, and dissimilar paths are generated [11,12]. This method is simple to implement, but not very simple to apply, since one has to arbitrarily choose one of several dimensions for the implementation of the penalty mechanism; moreover, the method has no way of measuring the quality of the produced dissimilar paths, in terms of spatial differences and lengths.

The Gateway Shortest Path method is based on the generation of the shortest paths, between an origin and a destination, constrained to go through defined nodes called "gateways". Therefore, by generating paths which are constrained to go through different gateways, a set of dissimilar paths can be obtained. However, the quality of the paths, unfortunately, depends heavily on the selected gateways. This method also provides a criterion to estimate the similarity between two paths, using the concept of "area under a path" [13].

The Minimax method selects a subset of dissimilar paths, starting from a set of k assigned paths, taking into account both their lengths and dissimilarities. Given a set of k-shortest paths, the algorithm iteratively selects a given number of such paths. The choice is done by using a particular dissimilarity index [14].

Finally, the p-dispersion method is based on modeling the problem as a p-dispersion location problem, which consists in selecting p points in some space, so as to maximize the minimum distance between any couple of selected points. In this method, a large set of candidate paths is generated and modeled as a set of points, with the distance between two points representing the dissimilarity between the two relative paths [15,16]. Recently, a bi-criteria approach has been proposed for generating the candidate path set before applying the p-dispersion method [17].

Nevertheless, it seems from the literature that the problem of spreading the risk equitably over the zones of the geographical region in which the transportation network is embedded still needs further attention. This is exemplified by the case in which two selected paths with very few common links (and hence highly dissimilar ones) are geographically very close to each other, such that the intersection of the two exposure zones around the paths is not negligible. This means that the population living in the intersection of the two exposure zones is affected by the risk of both paths, implying a low degree of risk equity. To cope with this problem, Akgün et al. [10] introduced a different definition of similarity based on the concept of buffer zone around a path. This new definition allows us to consider also similar paths that do not have common links, but whose buffer zones are not disjointed.

In our approach, in addition to our main objective of selecting a set of paths of minimum total risk, we also try to address risk equity, by bounding the maximum risk sustained by the population living along (in the proximity of) each populated link of the network. In particular, in our model we refer to the societal risk, and, in order to estimate the incident consequence (and, hence, the risk) for humans, we extend the λ-neighborhood approach proposed by Batta and Chiu in [5], by weighting the incident consequence with a distance-sensitive damage function that generalizes the zero-one function used in [5].

We propose a mathematical formulation for our model which is exploited so as to get a Lagrangean relaxation in order to achieve an effective lower bound on the optimal solution value. Two heuristic algorithms that are modified versions of Yen's algorithm [18] are proposed, and experimentally evaluated on an Italian geographical region. Besides the total risk to be minimized, different performance measures are also considered to evaluate the generated paths.

The paper is organized as follows. In Section 2, the problem under consideration is defined in detail. In Section 3, the mathematical formulation is given together with the Lagrangean relaxation. In Section 4, the heuristic algorithms are described. Finally, in Section 5, we present our experimentation results.

2. Problem definition

Let $G = (N, A)$ be a directed graph representing the transportation network of the regional area under consideration, with N and A being the set of n nodes, and the set of m directed links of the network, respectively. Each link $h \in A$, according to a given level of data aggregation of the network, corresponds to a road segment of the network, and each node $i \in N$ corresponds to a road crossing. Let there be given an origin $o \in N$ and a destination $d \in N$, and let P be the set of simple paths in G from origin o to destination d.

We assume that the population is mostly located in the neighborhood of the links of the network; hence, we consider populated links. This is typical for the kind of population distribution within a regional area in some European countries, such as Italy. Moreover, this is not a severe limitation (e.g., see [5]), since in the network model we can incorporate links that are not part of the transportation network, but whose populations living in the proximity of them are exposed to the risk induced by hazmat vehicles traveling on other (close) links.

The problem we consider is to find a set of k paths (routes) for hazmat shipments from the origin o to the destination d that minimizes the total risk induced by traveling on these paths on the population of the regional area in which the transportation network is embedded while spreading the risk equitably among the populated links of the network.

2.1. Risk estimation

We suppose that an incident on a link may impact not only on the population residing on that link, but also significantly on inhabitants living along other (close) links. In other words, the extent of the (significantly) impacted zone may involve also populated links that are close to the link where the incident occurs. Clearly, the entity of the effect of the incident decreases with the physical distance from the point where the incident occurs. As a consequence, risk propagation on another close link decreases with the distance between this link and the link along with the hazmat vehicle is traveling. Since risk associated to hazmat transportation depends also on the characteristics of the road type along the selected routes and on population distribution, we will introduce the adopted risk model starting from the risk of a hazmat transportation on a unit-length segment of the network.

2.1.1. Unit-length segment risk

Considering that each link is partitioned into a sequence of unit-length segments, we assume that the risk σ_x^y of traveling on a unit-length segment x of a link for the population living in the proximity of another unit-length segment y is equal to the product between the *probability* per unit-length $prob_x^{\text{inc}}$ of an incident occurring on x, and the *estimated consequence* $cons_{x\to y}$ on the population pop_y living in the proximity of segment y. The estimated consequence $cons_{x\to y}$ is given by population pop_y multiplied by a distance-sensitive *weight* (*damage*) function $u(x, y) = \mathrm{e}^{-\alpha[d(x,y)]^2}$. As a matter of fact, we assume that the consequence decreases exponentially with the square of the Euclidean distance $d(x, y)$ between the centers of segment x and segment y, where α is the *impact factor* depending on the hazardous material under consideration. The smaller the α, the less the steepness of the damage function decay (with the distance), and, hence, the larger the size of the (significantly) impacted zone. Indeed, the incident consequence, in general, decreases as we move away from the source of the risk (as also Batta and Chiu state in the

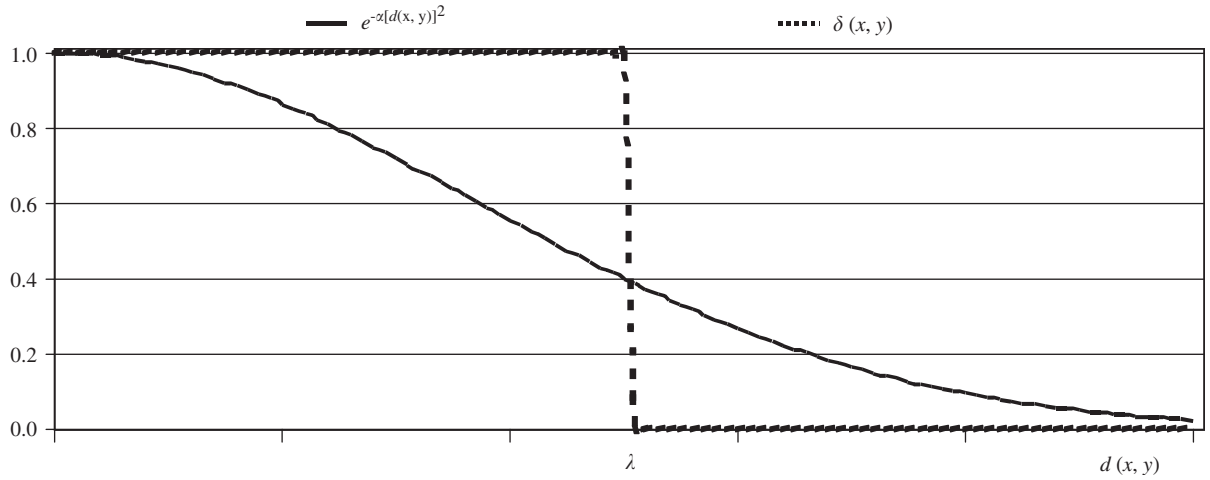

Fig. 1. Distance-sensitive damage functions: $e^{-\alpha[d(x,y)]^2}$ vs. $\delta(x, y)$, with $\alpha = (-\ln 0.4)/\lambda^2$.

closing section of their paper [5]); that is, the probability of injury decreases with distance, as is shown in [19] where examples of so-called risk contours are given.

We adopt the damage function $e^{-\alpha[d(x,y)]^2}$ to model this issue for the following reasons:

1. The function generalizes the $\delta(x, y)$ damage function adopted in the λ-neighborhood approach proposed by Batta and Chiu in [5], where $\delta(x, y)$ is equal to 1 if $d(x, y) \leqslant \lambda$, and 0 otherwise;
2. The function does not have discontinuity points;
3. From a modeling point of view it can be easily calibrated.

Fig. 1 shows the difference between the distance-sensitive damage function $e^{-\alpha[d(x,y)]^2}$ adopted in this paper and the $\delta(x, y)$ function of Batta and Chiu.

In summary, we have

$$\sigma_x^y = prob_x^{\text{inc}}\ pop_y\, e^{-\alpha[d(x,y)]^2}.$$

Therefore, the *unit-length segment risk* σ_x of traveling on a unit-length segment x for the population of the regional area is

$$\sigma_x = prob_x^{\text{inc}} \sum_{y \in S} pop_y\, e^{-\alpha[d(x,y)]^2},$$

where S is the set of all link segments of the network.

2.1.2. Link risk

Now we give the definition of the *link risk* r_h of link $h \in A$, which is the risk for the whole population due to hazmat transportation on link h. Assuming that link h is partitioned into a sequence $(h_1, \ldots, h_{q_h})$ of q_h unit-length segments, to compute the risk r_h we should multiply the probability $prob_{h_s}^{\text{inc}}$ of incident in a segment h_s with the probability of no-incident in the previous segments of that link. Since incident

probabilities are very small (on the order of one-in-a-million-miles), we can reasonably assume that the probability of no-incident is very close to 1, and hence approximate r_h as the sum of the unit-length segment risks of the segments of link h (e.g., see [4]). That is, the *link risk* r_h of link h is

$$r_h = \sum_{s=1}^{q_h} \sigma_{h_s}.$$

2.1.3. Path risk

We now address the calculation of the *path risk* c_j of a path j in G, which is the risk for the population due to hazmat transportation on path j. Since a path is a sequence of links, using a similar assumption to the one for link risk calculation, we may reasonably approximate the path risk c_j calculation as

$$c_j = \sum_{h \in A_j} r_h,$$

where $A_j \subseteq A$ is the set of links of path j.

2.2. Problem formulation

In order to spread the risk equitably as much as possible over the network, we consider an upper limit on the total hazmat transportation risk over populated links by considering a *risk threshold* ρ_h for each (populated) link $h \in A$; therefore, the total risk for the population living along link h induced by hazmat shipments on the network cannot be greater than ρ_h.

According to the definition of (link) risk r_h of link h, let

$$r_h^i = \sum_{s=1}^{q_h} \left(prob_{h_s}^{\text{inc}} \sum_{t=1}^{q_i} pop_{i_t} \, \mathrm{e}^{-\alpha[d(h_s, i_t)]^2} \right)$$

be the risk induced on the population living along populated link i (partitioned into q_i segments $i_1, \ldots, i_{q_i}$), by a hazmat vehicle traveling on link h (partitioned into q_h segments $h_1, \ldots, h_{q_h}$). Clearly, $r_h = \sum_{i \in A} r_h^i$.

Again, assuming the risks to be additive, let $a_{ij} = \sum_{h \in A_j} r_h^i$ be the risk for the population along link i due to a hazmat shipment on path j, where $A_j \subseteq A$ is the set of links of path j.

The optimization problem we consider is to select a set P^* of k distinct simple paths on the network $G = (N, A)$, from the origin $o \in N$ to the destination $d \in N$, so as to minimize the total path risk of the selected paths while satisfying risk threshold constraints on populated links.

Notice that, if a link h belongs to $q \geqslant 2$ (distinct) selected paths in P^*, then the objective function counts the risk r_h of traveling on link h (induced on the population) q times.

Let us introduce a binary variable x_j, such that x_j is equal to 1 if path $j \in P^*$, and 0 otherwise. Then, the problem can be formulated as follows:

$$z^* = \min \sum_{j \in P} c_j x_j \tag{1}$$

s.t.

$$\sum_{j\in P} a_{ij}x_j \leqslant \rho_i \quad i \in A, \tag{2}$$

$$\sum_{j\in P} x_j = k \tag{3}$$

$$x_j \in \{0,1\} \quad j \in P, \tag{4}$$

The objective function (1) to be minimized is the total risk. Constraint set (2) ensure for each link $i \in A$ that the risk threshold ρ_i not be exceeded by the sum of risks induced on the population in the proximity of link i by all the selected paths in the solution. Constraint (3) ensures that the number of selected paths has to be equal to k.

In the next section, a Lagrangean relaxation of the above formulation is proposed in order to provide a lower bound on the optimal solution value.

3. A Lagrangean relaxation

We dualize the risk threshold constraints (2) by using the Lagrangean dual variables (penalties) $\boldsymbol{\lambda} \geqslant 0$. We obtain the following Lagrangean relaxation problem $LR(\boldsymbol{\lambda})$:

$$g(\boldsymbol{\lambda}) = \min \sum_{j\in P} \left[c_j + \sum_{i\in A} \lambda_i a_{ij} \right] x_j - \sum_{i\in A} \lambda_i \rho_i \tag{1'}$$

s.t.

$$\sum_{j\in P} x_j = k \tag{3}$$

$$x_j \in \{0,1\} \quad j \in P. \tag{4}$$

From the definition of c_j and a_{ij}, we have the following expression of the cost (or length) of a path $j \in P$ in the objective function (1′) of the Lagrangean relaxation:

$$\begin{aligned} c_j + \sum_{i\in A} \lambda_i a_{ij} &= \sum_{h\in A_j} r_h + \sum_{i\in A} \lambda_i \sum_{h\in A_j} r_h^i \\ &= \sum_{h\in A_j} \sum_{i\in A} r_h^i + \sum_{i\in A} \lambda_i \sum_{h\in A_j} r_h^i = \sum_{h\in A_j} \sum_{i\in A} (1+\lambda_i) r_h^i. \end{aligned} \tag{5}$$

Therefore, the $LR(\boldsymbol{\lambda})$ problem is that of determining the set $P'(\boldsymbol{\lambda})$ of k-shortest (simple) distinct paths, where the length l_h of each link $h \in A$ is $l_h = \sum_{i\in A} (1+\lambda_i) r_h^i$.

It is easy to observe that $l_h \geqslant 0$, since $\lambda_i \geqslant 0$; hence, for a given assignment of Lagrangean penalties, the problem may be solved, for example, by Yen's algorithm [18]. To obtain a good lower bound, we have to adequately define the Lagrangean penalties. In theory, the best lower bound can be achieved by solving the dual problem

$$g^* = \max\ g(\boldsymbol{\lambda}) \tag{6}$$

s.t.

$$\lambda_i \geqslant 0 \quad i \in A. \tag{7}$$

The dual problem is in general difficult to solve exactly, and may be heuristically solved by means of the well known subgradient technique [20]. That is, for example, starting with all penalties $\{\lambda_i\}$ equal to zero and, at each iteration $r > 1$, updating them as

$$\lambda_i^r = \max\left\{0, \lambda_i^{r-1} + \theta^{r-1}\frac{z^H - g(\boldsymbol{\lambda}^{r-1})}{\sum_{h\in A}\left(\sum_{j\in P'(\boldsymbol{\lambda}^{r-1})} a_{hj} - \rho_h\right)^2}\left(\sum_{j\in P'(\boldsymbol{\lambda}^{r-1})} a_{ij} - \rho_i\right)\right\},$$

where z^H is an upper bound on z^* (e.g., the (best) heuristically obtained feasible solution value) and θ^r is a scalar satisfying $0 < \theta^r \leqslant 2$, starting with $\theta^1 = 2$ and halving its value whenever $g(\boldsymbol{\lambda})$ has failed to increase in some fixed number of iterations.

The lagrangean relaxation has been introduced just to define a lower bound on the optimal solution so that one is able to evaluate the maximum distance from optimality of the solutions produced by the heuristics we describe in the following.

4. Description of the heuristic algorithms

We propose two heuristic algorithms, all based on Yen's algorithm [18] for the k-shortest path problem. They are, in practice, constrained k-shortest path algorithms. The first one, called *GD*, is a greedy heuristic algorithm, and the second one, called *RGD*, is a randomized version of the former.

Yen's procedure is a classic algorithm for ranking the k-shortest simple paths among the set P of paths between a pair (s, t) of nodes in a network G. It is an iterative algorithm that, at each iteration $h \leqslant k$, selects the shortest (i.e., minimum risk in our case) path j^h from a set $C^h \subseteq P$ of candidate simple paths from s to t. The candidate set C^h contains the shortest simple paths from s to t that deviate from the paths selected in the previous $h-1$ iterations. Initializing C^1 with the shortest (minimum risk) simple path j^1 from s to t, for $h > 1$ we have $C^h = C^{h-1} \cup Dev(j^{h-1})$, where $Dev(j^{h-1})$ is the set of all the new shortest simple paths from s to t that deviate from the selected path j^{h-1}. A brief step-by-step description of Yen's algorithm is given below. For the details of determining $Dev(j^{h-1})$ see [18].

Yen's algorithm (input: s, t, k, G; output: S)
Let $S := \emptyset$.
Let j^1 be the shortest (minimum risk) simple path from s to t in G.
Let $C^1 := \{j^1\}$, and h :=1.
While ($C^h \neq \emptyset$ and $h \leqslant k$) do
Step 1: Select the shortest path j^h among the paths in C^h.
Let $C^h := C^h \backslash \{j^h\}$.
Let $S := S \cup \{j^h\}$.
Step 2: Generate the set $Dev(j^h)$ of all the possible new shortest simple paths from s to t that deviate from j^h.
Let $C^{h+1} := C^h \cup Dev(j^h)$.
Let $h := h + 1$.

4.1. The greedy algorithm GD

Algorithm *GD* implements Yen's algorithm, and greedily constructs a feasible solution adding one path j at a time to a *partial* feasible solution S made up of a subset of $h < k$ feasible paths, such that the new partial solution $S \cup \{j\}$ is also feasible, that is, it fulfills the risk threshold constraints (2). The algorithm stops either providing the solution S whenever the partial solution S contains exactly k feasible paths (i.e., S is a *complete* feasible solution), or with no solution whenever the set of candidate paths is empty. Algorithm *GD* differs from Yen's algorithm in the "while do" condition which is replaced with ($C^h \neq \emptyset$ and $|S| < k$) and in Step 1 which is replaced with the following step:

Step 1′: Select the shortest path j^h among the paths in C^h.
Let $C^h := C^h \backslash \{j^h\}$.
If $S \cup \{j^h\}$ is a partial feasible solution, then let $S := S \cup \{j^h\}$.

4.2. The randomized greedy algorithm RGD

Algorithm *RGD* also implements Yen's algorithm and slightly differs from algorithm *GD* in the way it selects a path from the candidate set in order to augment the feasible partial solution S. At each iteration h, algorithm *RGD* randomly selects a path j^h from a restricted candidate set $RC^h \subseteq C^h$ of paths of minimum risk, and adds j^h to the partial solution S only if $S \cup \{j^h\}$ is feasible. The algorithm stops either providing the solution S whenever the partial solution S contains exactly k feasible paths (i.e., S is a *complete* feasible solution), or with no solution whenever the set of candidate paths is empty. Therefore, algorithm *GD* differs from Yen's algorithm in the "while do" condition which is replaced with ($C^h \neq \emptyset$ and $|S| < k$) and in Step 1 which is replaced with the following step:

Step 1″: Select at random a path j^h among the paths in RC^h.
Let $C^h := C^h \backslash \{j^h\}$.
If $S \cup \{j^h\}$ is a partial feasible solution, then let $S := S \cup \{j^h\}$.

RC^h is the subset of candidate paths in C^h, whose risk is not greater than $risk^{\min} + \gamma(risk^{\max} - risk^{\min})$, where $risk^{\min}$ and $risk^{\max}$ are the minimum and maximum risk values, respectively, among paths in C^h, and γ is a real value between 0 and 1 that controls the size of the restricted candidate set RC^h. Obviously, if $\gamma = 0$ we have that $RC^h = \{j^h\}$, with j^h being the minimum risk path in C^h, and *RGD* is a pure greedy algorithm (i.e., it is the same as algorithm *GD*); on the contrary, if $\gamma = 1$ *RGD* operates completely at random. Finally, *RGD* is run several times and the best solution found is taken.

The value of parameter γ is not fixed a priori but is self-tuned, according to the quality of the solutions found during the previous runs. In particular, this is accomplished by using the method proposed in [21], where parameter γ is chosen randomly according to a probability distribution which is periodically updated, by assigning higher probabilities to those values of γ which allow us to achieve solutions whose values are, on average, closer to the best solution found. In detail, the method works as follows. Let $\Gamma = \{\gamma_1, \gamma_2, \ldots, \gamma_m\}$ be a set of m potential values of γ (e.g., we consider $m = 10$, and $\Gamma = \{0.1, 0.2, \ldots, 1.0\}$). The probabilities associated with the choice of each value γ_i are all initially made equal to $p_i = 1/m$, with $i = 1, 2, \ldots, m$. Furthermore, let z^+ be the current best solution value, and $\bar{z}_i$ the average solution

value found using parameter value γ_i, with $i = 1, 2, \ldots, m$. The selection probabilities are updated every 10 runs, such that $p_i = q_i / \sum_{j=1}^{m} q_j$, with $q_i = z^+/\bar{z}_i$, for $i = 1, 2, \ldots, m$.

5. Experimental analysis

In this section, we give some experimental results obtained by testing the proposed model and algorithms on a real network. First, let us define the performance indicators we consider for a better evaluation of the solutions.

5.1. Performance indicators

Let $\underline{P}$ be the set of k-selected paths in a given solution. Besides the total (societal) risk $TotR = \sum_{j \in \underline{P}} c_j$, we also evaluate solutions with respect to some other performance measures. In particular, we are interested in evaluating how equitably the risk is distributed over the population.

In this regard, we consider four indices: *risk equity index* (*EI*), *restricted risk equity index* (*REI*), *average path dissimilarity index* (*AvgDI*), and *minimum path dissimilarity index* (*MinDI*). The former two indices aim to quantify the equity of the risk distribution over the population. The last two ones quantify the dissimilarity of the selected paths. Note that one alternative way to spread the risk equitably is by ensuring that the chosen paths are quite dissimilar [10].

We define the risk equity index *EI* as the variation coefficient of the (average per inhabitant) risk for the population along populated links of the network; that is,

$$EI = \frac{1}{\mu} \sqrt{\sum_{i \in A} (\sigma_i - \mu)^2 / |A|},$$

where $\mu = \left(\sum_{j \in \underline{P}} c_j \right) / \sum_{h \in A} pop_h$ is the average risk per inhabitant of the geographical region where the transportation network $G = (N, A)$ is embedded, and $\sigma_i = \left(\sum_{j \in \underline{P}} a_{ij} \right) / pop_i$ is the average risk per inhabitant of the populated link $i \in A$. The smaller the *EI*, the more equitable the risk distribution.

As *EI* is computed with respect to all the links of G, we expect its value to be very high, since in any solution there are many populated links that are far from the selected paths and, hence, the risk for the population of these links should be negligible compared to the risk for the population located in the neighborhood of the links of the paths. Therefore, we also compute the value of the restricted risk equity index *REI* in order to evaluate the equity distribution of the risk only over populated links belonging to at least one of the selected paths in $\underline{P}$. Index *REI* is accordingly defined as

$$REI = \frac{1}{\mu_R} \sqrt{\sum_{i \in A_R} (\sigma_i - \mu_R)^2 / |A_R|},$$

where $A_R = \bigcup_{j \in \underline{P}} A_j$ is the subset of links of G belonging to the selected paths, and $\mu_R = \left(\sum_{i \in A_R} \sum_{j \in \underline{P}} a_{ij} \right) / \sum_{h \in A_R} pop_h$ is the average risk per inhabitant among the population located in the neighborhood of populated links in A_R. Again, the smaller the *REI*, the more equitable the risk distribution over populated links belonging to the selected paths.

The average dissimilarity index *AvgDI*, and the minimum dissimilarity index *MinDI* are, respectively, the average and the minimum values among the dissimilarities $D(p_q, p_r)$ between pairs (p_q, p_r) of the selected paths.

The (spatial) *dissimilarity* $D(p_q, p_r)$ between the pair of paths p_q, p_r is defined as $D(p_q, p_r) = 1 - S(p_q, p_r)$, where $S(p_q, p_r)$ is the (spatial) *similarity* between these two paths, which we define as follows:

$$S(p_q, p_r) = \left[\left(\sum_{h \in A_{pq}} len_h \, e^{-\alpha[\min_{l \in A_{pr}} \{d(h,l)\}]^2} \right) \Big/ \sum_{h \in A_{pq}} len_h + \left(\sum_{l \in A_{pr}} len_l \, e^{-\alpha[\min_{h \in A_{pq}} \{d(l,h)\}]^2} \right) \Big/ \sum_{l \in A_{pr}} len_l \right] /2,$$

where len_h (len_l) is the length of link $h \in A_{p_q}$ ($l \in A_{p_q}$), $d(h, l)$ is the Euclidean distance between the centers of link h and link l, and α is the impact factor depending on the hazardous material under consideration.

In particular, our definition of the similarity $S(p_q, p_r)$ between two paths p_q and p_r generalizes the one given by Erkut and Verter [4] that measures similarity only as a function of the length of common links between two paths. Moreover, our definition extends the concept of spatial similarity in order to consider, in some ways, as quasi-similar also paths that are very close, even if they do not have any links in common. In fact, in the hazmat transportation context, it may not be sufficient to select paths that are (link) disjointed in order to assure risk equity; it is necessary that any two paths also be sufficiently far apart in order to prevent imposing risk on residents living in areas located in the proximities of both these two paths. The impact factor α allows us to weight the similarity with respect to the distance between the links of paths p_q and p_r, according to the hazardous material under consideration. Note that for α that tends to infinity, both $\sum_{h \in A_{pq}} len_h \, e^{-\alpha[\min_{l \in A_{pr}} \{d(h,l)\}]^2}$ and $\sum_{l \in A_{pr}} len_l \, e^{-\alpha[\min_{h \in A_{pq}} \{d(l,h)\}]^2}$ tend to the total length of the common links of paths p_q and p_r, and, hence, for the similarity $S(p_q, p_r)$ we have the same definition given by Erkut and Verter [4].

Here is the complete set of performance indicators we compute in our experimentation for each solution (i.e., the set $\underline{P}$ of k selected paths):

- *total risk*: $TotR = \sum_{j \in \underline{P}} c_j$,
- *average path risk*: $AvgR = \left(\sum_{j \in \underline{P}} c_j \right) / k$,
- *maximum path risk*: $MaxR = \max_{j \in \underline{P}} \{c_j\}$,
- *risk index*: $RI = kc_{\min} / \sum_{j \in \underline{P}} c_j$,
- *length index*: $LI = kLen_{\min} / \sum_{j \in \underline{P}} Len_j$,
- *risk equity index*: $EI = 1/\mu \sqrt{\sum_{i \in A} (\sigma_i - \mu)^2 / |A|}$,
- *restricted risk equity index*: $REI = 1/\mu_R \sqrt{\sum_{i \in A_R} (\sigma_i - \mu_R)^2 / |A_R|}$,
- *average path dissimilarity index*: $AvgDI = \left(\sum_{p_q, p_r \in \underline{P}} D(p_q, p_r) \right) / (k(k-1)/2)$,
- *minimum path dissimilarity index*: $MinDI = \min_{\{(p_q, p_r): p_q \neq p_r \in \underline{P}\}} \{D(p_q, p_r)\}$,

where Len_j is the length of path j, $Len_{\min}$ is the length of the minimum length path and $c_{\min}$ is the risk of the minimum risk path among paths from node o to node d in G.

In particular, the risk index RI, equal to the ratio between the minimum risk $c_{\min}$ and the average risk $AvgR$, gives evidence of the quality of the solution with respect to the risk attribute; the closer this index is to 1, the closer the average risk is to the minimum value. The length index LI, defined as the ratio between the minimum length $Len_{\min}$ and the average length of the selected paths, gives evidence of the quality of the solution with respect to the path length attribute; again, the closer this index is to 1, the better the solution with respect to total path length.

5.2. Experimental results

All algorithms have been implemented in Java language, and run on a Pentium III PC with 800 MHz and 256 MB of RAM.

We have experimented the proposed algorithms on an Italian regional road network, using realistic data for the population and incident probabilities. In particular, we consider the region of Lazio containing the city of Rome, which covers an area of 17,203 km^2, and has a population of about 5 million inhabitants. The resulting network $G = (N, A)$ has 311 nodes and 441 links, and contains both highways and local regional roads (see Fig. 5). Obviously, the roads and streets of Rome (i.e., the ones inside the ring road encompassing Rome) are not considered in the network. We consider each link $h \in A$ as being segmented into a sequence $(h_1, \ldots, h_{q_h})$ of q_h link segments about 1 km in length, and, for each segment h_s, the available data are the incident probability $prob^{\text{inc}}_{h_s}$ on that segment, and the population pop_{h_s} located in its proximity area. As for the risk threshold ρ_h, for each link $h \in A$, we consider its value as being M times the risk induced on the population living along populated link h by a hazmat vehicle traveling on that link; that is, $\rho_h = Mr^h_h$.

Note that according to the definition of σ_x given in Section 2, the risk induced by a hazmat vehicle traveling on link segment x over the population pop_y, residing in the proximity area a_y of link segment y, is counted as many times as the number of link segments close to y and whose proximity areas almost completely overlap with a_y. This could imply a non-negligible risk overestimation that obviously could increase with the network density. However, the density of the network we use for the experimentation is not so high, and the error due to the risk multiple counting is negligible. Nevertheless, also for high density network the risk multiple counting error may be reduced by preprocessing the population data before running the algorithms. As a matter of fact, the value of σ_x, which is independent on the selected paths, may be estimated with a negligible error by reducing the population values of the link segments whose proximity areas intersect so as to make the total population around these segments equal to the total population inside the union of their proximity areas.

For our test problems, we consider the Orvieto–Latina origin–destination pair case study, where the origin node (Orvieto) is located in the north, and the destination one (Latina) is in the south of Lazio, with a shortest path between them of 166 km. The proposed algorithms are experimented on 16 different scenarios, with $k = 2, 4, 6, 8$, and $\alpha = 0.03, 0.06, 0.15, 0.92$, respectively. Recall that α, in practice, models the extent of the significantly impacted zone depending on the shipped hazardous material: the smaller the α, the larger the size of this zone. In particular, the α values are chosen according to the following rationale. We consider the significantly impacted zone as being a circle of radius λ centered in the incident point x, such that the distance-sensitive damage function $u(d(x, y)) = e^{-\alpha[d(x,y)]^2}$ computed on each point

Table 1
Algorithm comparison

k	α	M	*GD*			*RGD*			*LB*
			TotR	*gap%*	*CPU*(sec)	*TotR*	*gap%*	*CPU*(sec)	*TotR*
2	0.92	2.01	140588.0	23.22	32.46	**126497.0**	10.87	10.31	114091.4
		2.02	122484.0	7.03	5.18	**122209.0**	6.79	23.93	114436.5
		2.12	117293.0	2.76	1.05	117293.0	2.76	16.65	114137.7
		2.17	114103.0	0.07	0.33	114103.0	0.07	17.90	114023.2
2	0.15	2.07	252849.0	13.57	21.11	252849.0	13.57	17.02	222619.3
		2.10	249295.0	12.26	14.68	249295.0	12.26	21.53	222067.8
		2.13	235902.0	6.38	1.51	235902.0	6.38	33.83	221742.8
		2.16	227055.0	2.43	0.64	227055.0	2.43	28.25	221665.3
2	0.06	2.23	455031.0	15.08	16.85	455031.0	15.08	31.42	395398.8
		2.30	446378.0	12.51	11.68	446378.0	12.51	34.14	396710.5
		2.35	413798.0	4.32	1.82	413798.0	4.32	24.85	396647.9
		2.45	409397.0	3.64	1.45	409397.0	3.64	35.76	394985.0
2	0.03	2.80	900829.0	17.71	56.70	900829.0	17.71	38.85	765261.6
		2.85	715868.0	6.08	4.75	**713822.0**	5.77	39.66	674837.6
		2.90	689641.0	2.35	1.67	689641.0	2.35	42.16	673773.3
		2.95	689641.0	2.45	1.67	**680708.0**	1.13	32.96	673094.7
4	0.92	4.03	240600.0	4.43	11.92	**240161.0**	4.24	52.28	230373.8
		4.04	238142.0	3.39	8.00	**235873.0**	2.40	54.73	230334.9
		4.05	237927.0	2.89	6.74	**234755.0**	1.52	64.88	231230.2
		4.20	232736.0	0.86	1.18	232736.0	0.86	54.78	230734.8
4	0.15	4.07	482711.0	6.34	75.79	482711.0	6.34	67.93	453901.1
		4.09	479157.0	5.44	14.95	479157.0	5.44	51.13	454433.3
		4.10	478901.0	5.56	5.39	478901.0	5.56	65.14	453675.9
		4.20	474500.0	4.84	1.23	474500.0	4.84	65.13	452587.9
4	0.06	4.45	855068.0	5.92	51.82	**849744.0**	5.26	88.46	807254.2
		4.50	850667.0	5.45	22.72	**848082.0**	5.13	99.95	806632.4
		4.70	842014.0	4.58	10.78	**821904.0**	2.08	82.06	805123.2
		4.80	809434.0	0.73	1.30	809434.0	0.73	82.97	803525.6
4	0.03	5.48	1495580.0	9.67	67.86	1495580.0	9.67	123.21	1363697.0
		5.50	1463895.0	7.34	41.52	1463895.0	7.34	106.91	1363458.1
		5.90	1394669.0	2.27	3.80	1394669.0	2.27	113.72	1363697.0
		6.30	1394669.0	2.69	3.80	**1368442.0**	0.76	90.63	1358082.6
6	0.92	6.04	359031.0	2.90	10.83	**356043.0**	2.04	89.13	348899.7
		6.05	355912.0	1.67	7.27	355912.0	1.67	102.74	350056.4
		6.10	351167.0	0.33	1.48	**350506.0**	0.14	92.84	350004.3
		6.15	350506.0	0.17	0.81	350506.0	0.17	97.54	349910.7

Table 1 (*continued*)

k	α	M	*GD*			*RGD*			*LB*
			TotR	*gap%*	*CPU*(sec)	*TotR*	*gap%*	*CPU*(sec)	*TotR*
6	0.15	6.14	712141.0	3.35	107.27	712141.0	3.35	130.96	689065.7
		6.20	711885.0	3.55	42.10	**711079.0**	3.43	108.74	687453.6
		6.40	707484.0	2.85	19.05	707484.0	2.85	109.85	687857.8
		6.60	690872.0	0.74	2.25	690872.0	0.74	112.56	685775.8
6	0.06	6.60	1272374.0	4.17	132.56	**1267050.0**	3.74	149.53	1221396.1
		6.70	1267973.0	3.92	27.66	**1264465.0**	3.64	111.83	1220049.9
		7.30	1250704.0	2.93	12.91	**1219940.0**	0.40	123.96	1215007.5
		7.60	1218124.0	0.42	2.09	1218124.0	0.42	100.38	1212931.0
6	0.03	8.55	2107370.0	2.61	85.96	**2064169.0**	0.50	163.75	2053767.3
		8.60	2059990.0	0.23	5.85	2059990.0	0.23	159.24	2055115.4
		9.00	2053329.0	0.33	1.56	2053329.0	0.33	175.77	2053042.5
		10.60	2051057.0	0.62	1.26	2051057.0	0.62	180.61	2047123.7
8	0.92	8.04	481091.0	2.43	11.37	**476324.0**	1.42	117.57	469637.1
		8.05	476324.0	1.27	9.18	476324.0	1.27	106.47	470341.0
		8.10	476109.0	1.26	8.86	**472511.0**	0.49	117.57	470173.0
		8.15	470164.0	0.02	1.64	470164.0	0.02	124.18	470068.5
8	0.15	8.17	950528.0	2.54	82.15	**947566.0**	2.22	162.65	926959.6
		8.18	946974.0	2.15	42.18	946974.0	2.15	155.84	927021.0
		8.20	946718.0	2.07	6.32	**945912.0**	1.98	163.25	927480.3
		8.80	942317.0	2.23	3.63	**932666.0**	1.19	135.17	921704.8
8	0.06	8.85	1672413.0	2.15	135.42	1672413.0	2.15	183.14	1637247.1
		9.00	1669826.0	2.04	14.14	1669826.0	2.04	180.24	1636301.8
		9.50	1669826.0	2.25	14.14	**1663922.0**	1.89	178.04	1633001.0
		10.00	1630631.0	0.10	2.18	1630631.0	0.10	173.91	1628902.1
8	0.03	11.20	2827100.0	2.55	96.40	**2762061.0**	0.19	258.26	2756786.8
		11.50	2798601.0	1.61	19.62	**2755400.0**	0.04	256.69	2754251.8
		12.00	2755400.0	0.00	2.33	2755400.0	0.00	265.39	2755295.5
		13.00	2751221.0	0.02	1.85	2751221.0	0.02	257.35	2750526.8

y inside this zone is not less than 0.4 (e.g., see Fig. 1). For example, by considering $\lambda = 2.4$ km (1.5 miles), as for the benzyl chloride hazmat type (e.g., see [9]), we get $\alpha = 0.16$. Hence, the $\alpha = 0.03, 0.06, 0.15, 0.92$ values correspond to the $\lambda = 5.5, 4.0, 2.5, 1.0$ km values, respectively.

For each of the 16 scenarios (identified by the pair (k, α)), we have carried out a long series of tests varying the risk threshold parameter M, spanning a set of instances ranging from high (for small values of M) to low (for large values of M) constrained test problems. For each scenario we present the results obtained with different M values such that for the smaller values the problem was unfeasible, while for the larger values the solution was the same to that for the unconstrained problem

Table 2
Results with $k = 2$

α	M	*TotR*	*AvgR*	*MaxR*	*RI*	*LI*	*EI*	*REI*	*Min DI*	*Avg DI*
0.92	2.01	126497.0	63248.5	68995.0	0.88	0.66	3.13	0.21	0.63	0.63
	2.02	122209.0	61104.5	63589.0	0.92	0.76	3.37	0.29	0.37	0.37
	2.12	117293.0	58646.5	61131.0	0.95	0.81	3.45	0.29	0.32	0.30
	2.17	114103.0	57051.5	57941.0	0.98	0.88	4.00	0.51	0.17	0.17
	∞	113664.0	56832.0	57502.0	0.99	0.90	4.10	0.51	0.15	0.15
			AveMaxR = 62914.0		*MaxMaxR* = 68995.0					
0.15	2.07	252849.0	126424.5	142689.0	0.87	0.51	1.83	0.37	0.56	0.56
	2.10	249295.0	124647.5	139135.0	0.88	0.52	1.96	0.39	0.45	0.45
	2.13	235902.0	117951.0	125742.0	0.93	0.63	2.26	0.53	0.16	0.16
	2.16	227055.0	113527.5	116895.0	0.97	0.72	2.35	0.52	0.12	0.12
	∞	221651.0	110825.5	111491.0	0.99	0.73	2.37	0.52	0.03	0.03
			AveMaxR = 131115.2		*MaxMaxR* = 142689.0					
0.06	2.23	455031.0	227515.5	258967.0	0.86	0.51	1.17	0.23	0.59	0.59
	2.30	446378.0	223189.0	250314.0	0.87	0.52	1.22	0.28	0.52	0.52
	2.35	413798.0	206899.0	217734.0	0.94	0.59	1.30	0.30	0.45	0.45
	2.45	409397.0	204698.5	213333.0	0.95	0.59	1.35	0.37	0.37	0.37
	∞	394713.0	197356.5	198649.0	0.99	0.73	1.43	0.53	0.09	0.09
			AveMaxR = 236187.2		*MaxMaxR* = 258967.0					
0.03	2.80	900829.0	450414.5	566690.0	0.74	0.44	0.73	0.37	0.77	0.77
	2.85	713822.0	356911.0	358320.0	0.93	0.59	1.01	0.34	0.52	0.52
	2.90	689641.0	344820.5	355502.0	0.97	0.59	0.99	0.27	0.43	0.43
	2.95	680708.0	340354.0	342132.0	0.98	0.69	1.10	0.44	0.21	0.21
	∞	672715.0	336357.5	338576.0	0.99	0.73	1.11	0.52	0.06	0.06
			AveMaxR = 405661.0		*MaxMaxR* = 566690.0					

(i.e., with an unlimited value for M). In particular, we only report the results considering four values of M that were chosen in a such a way the performance indicators exhibit significant changes. As a matter of fact, the changes observed do not appear to have a linear dependence on the M values. One reason for this can surely be the risk definition that includes an exponentially decreasing damage function.

In Table 1, we compare the solution values (i.e., the values of the total path risk *TotR*) provided by algorithms *GD* and *RGD* with the Lagrangean lower bound *LB*. The *RGD* solution is the best one obtained after 200 runs of this algorithm. For both algorithms *GD* and *RGD*, the table lists also the values of the relative *gap* (in percentage) of the solution value *TotR* with respect to *LB*, (i.e., $(TotR - LB)/LB)\%$). Comparing algorithms *GD* and *RGD* in terms of solution values (*TotR*), it can be noted that *RGD* is more effective than *GD*, providing better solutions (exhibited in bold in the table) in 26 cases out of 64, and the same solution as *GD* in the other cases. The quality of the solutions is also acceptable if compared to *LB*, with an average relative gap of 3.49% and a maximum value of 23.22%.

Table 3
Results with $k = 4$

α	M	*TotR*	*AvgR*	*MaxR*	*RI*	*LI*	*EI*	*REI*	*Min DI*	*Avg DI*
0.92	4.03	240161.0	60040.2	66537.0	0.93	0.76	3.18	0.49	0.21	0.41
	4.04	235873.0	58968.2	63589.0	0.95	0.82	3.22	0.47	0.21	0.38
	4.05	234755.0	58688.7	61131.0	0.95	0.84	3.28	0.48	0.21	0.35
	4.20	232736.0	58184.0	61131.0	0.96	0.85	3.33	0.47	0.17	0.31
	∞	230225.0	57556.2	58620.0	0.97	0.87	3.54	0.64	0.08	0.26
			AveMaxR = 63097.0		*MaxMaxR* = 66537.0					
0.15	4.07	482711.0	120677.7	147346.0	0.91	0.55	1.76	0.58	0.03	0.45
	4.09	479157.0	119789.2	143792.0	0.92	0.56	1.82	0.58	0.03	0.37
	4.10	478901.0	119725.2	143536.0	0.92	0.60	1.82	0.57	0.03	0.35
	4.20	474500.0	118625.0	139135.0	0.92	0.60	1.91	0.63	0.03	0.31
	∞	451454.0	112863.5	116089.0	0.97	0.70	2.11	0.70	0.03	0.14
			AveMaxR = 143452.2		*MaxMaxR* = 147346.0					
0.06	4.45	849744.0	212436.0	250314.0	0.91	0.60	1.24	0.57	0.02	0.32
	4.50	848082.0	212020.5	245913.0	0.92	0.59	1.28	0.61	0.08	0.29
	4.70	821904.0	205476.0	213333.0	0.95	0.63	1.34	0.54	0.02	0.26
	4.80	809434.0	202358.5	213333.0	0.96	0.64	1.36	0.61	0.08	0.25
	∞	800074.0	200018.5	203973.0	0.98	0.69	1.39	0.71	0.08	0.11
			AveMaxR = 230723.2		*MaxMaxR* = 250314.0					
0.03	5.48	1495580.0	373895.0	416158.0	0.89	0.57	0.97	0.51	0.13	0.35
	5.50	1463895.0	365973.7	415890.0	0.91	0.58	1.01	0.53	0.04	0.30
	5.90	1394669.0	348667.2	381729.0	0.96	0.64	1.02	0.54	0.06	0.27
	6.30	1368442.0	342110.5	355502.0	0.97	0.64	1.02	0.55	0.06	0.26
	∞	1355072.0	338768.0	342132.0	0.99	0.69	1.08	0.63	0.06	0.12
			AveMaxR = 392319.7		*MaxMaxR* = 416158.0					

Regarding algorithm efficiency, as one might expect computational times, on average, increase with k for both algorithms. For a given k value, they increase, on average, if α decreases; this is due to the larger size of the significantly impacted zone for smaller α, which implies the search for quite dissimilar paths. Algorithm *RGD*, on average, requires much greater CPU times than algorithm *GD*: obviously, this is due to the fact that *RGD* is run several times. Nevertheless, for high constrained instances (i.e., with small values of M), *GD* sometimes takes much more time than *RGD*; indeed, for these instances, *GD* may require the generation of a very large number q (with $q \gg k$) of simple paths in order to find k feasible paths since many of the generated paths may be unfeasible, and, hence, are discarded. On the contrary, for quasi-unconstrained instances (i.e., with large values of M), almost all the paths generated by algorithm *GD* happen to be feasible, and, therefore, a feasible solution may be found in a very short time by generating $q \cong k$ paths.

In Tables 2–5 we give the whole performance results evaluating the performance indicators only for the best algorithm. The tables list the results for the test cases with $k = 2$, 4, 6 and 8, and their columns, respectively, show the values of the impact factor α, the risk threshold parameter M, the total path risk

Table 4
Results with $k = 6$

α	M	*TotR*	*AvgR*	*MaxR*	*RI*	*LI*	*EI*	*REI*	*Min DI*	*Avg DI*
0.92	6.04	356043.0	59340.5	66537.0	0.94	0.79	3.25	0.62	0.08	0.35
	6.05	355912.0	59318.7	66537.0	0.94	0.80	3.27	0.61	0.08	0.34
	6.10	350506.0	58417.7	61131.0	0.96	0.85	3.35	0.61	0.08	0.30
	6.15	350506.0	58417.7	61131.0	0.96	0.85	3.35	0.61	0.08	0.30
	∞	348656.0	58109.3	59281.0	0.96	0.88	3.41	0.73	0.08	0.26
			AveMaxR = 63834.0		*MaxMaxR* = 66537.0					
0.15	6.14	712141.0	118690.2	143792.0	0.92	0.60	1.91	0.68	0.04	0.29
	6.20	711079.0	118513.2	139135.0	0.92	0.62	1.91	0.68	0.03	0.29
	6.40	707484.0	117914.0	139135.0	0.93	0.63	1.98	0.73	0.03	0.25
	6.60	690872.0	115145.3	122523.0	0.95	0.69	2.14	0.78	0.03	0.17
	∞	685762.0	114293.7	117413.0	0.96	0.69	2.21	0.79	0.03	0.15
			AveMaxR = 136146.2		*MaxMaxR* = 143792.0					
0.06	6.60	1267050.0	211175.0	245913.0	0.93	0.59	1.23	0.68	0.02	0.32
	6.70	1264465.0	210744.2	245913.0	0.93	0.59	1.23	0.68	0.02	0.32
	7.30	1219940.0	203323.3	213333.0	0.96	0.65	1.35	0.68	0.02	0.24
	7.60	1218124.0	203020.7	213333.0	0.96	0.65	1.38	0.73	0.02	0.20
	∞	1210580.0	201763.3	205789.0	0.97	0.69	1.44	0.75	0.02	0.14
			AveMaxR = 229623.0		*MaxMaxR* = 245913.0					
0.03	8.55	2064169.0	344028.1	355502.0	0.97	0.61	1.00	0.69	0.06	0.26
	8.60	2059990.0	343331.7	353595.0	0.97	0.65	1.04	0.69	0.01	0.20
	9.00	2053329.0	342221.5	353595.0	0.97	0.64	1.05	0.70	0.06	0.20
	10.60	2051057.0	341842.8	351323.0	0.97	0.70	1.12	0.76	0.06	0.11
	∞	2046303.0	341050.5	346569.0	0.98	0.68	1.17	0.76	0.05	0.10
			AveMaxR = 353503.7		*MaxMaxR* = 355502.0					

TotR, the average path risk *AvgR*, the maximum path risk *MaxR*, the risk index *RI*, the length index *LI*, the risk equity index *EI*, the restricted risk equity index *REI*, the minimum path dissimilarity index *MinDI*, and of the average path dissimilarity index *AvgDI*.

For each couple (k, α), besides the results obtained with the four chosen M values, we also give the performance results of the solution for the problem instance with a very large (let us say "unlimited") value for M (e.g., $M \to \infty$). In practice, for this latter case we may consider the problem without risk threshold constraints (2) which, therefore, can solved by simply applying Yen's algorithm [18]. Finally, the average (*AveMaxR*) and the maximum (*MaxMaxR*) values among the maximum path risk (*MaxR*) values, obtained for the four cases with limited M values, are given.

The tables show that both the average *AvgR* and the maximum *MaxR* path risks increase with decreasing values for the impact factor α (i.e., by increasing the size of the significantly impacted zone), while there are no significant differences among their values for the cases with the same α and different k values. For a given couple of (k, α) values, as one may expect, the risk index *RI* and the length index *LI* values, in general, increase with increasing M values, since this implies higher risk threshold values and, thus,

Table 5
Results with $k = 8$

α	M	*TotR*	*AvgR*	*MaxR*	*RI*	*LI*	*EI*	*REI*	*Min DI*	*Avg DI*
0.92	8.04	476324.0	59540.5	66537.0	0.94	0.80	3.14	0.69	0.08	0.34
	8.05	476324.0	59540.5	66537.0	0.94	0.80	3.14	0.69	0.08	0.34
	8.10	472511.0	59063.9	61353.0	0.95	0.84	3.31	0.70	0.08	0.31
	8.15	470164.0	58770.5	61131.0	0.95	0.84	3.36	0.70	0.08	0.30
	∞	468993.0	58624.1	60377.0	0.95	0.86	3.39	0.79	0.08	0.27
			AveMaxR = 63889.5		*MaxMaxR* = 66537.0					
0.15	8.17	947566.0	118445.7	139391.0	0.93	0.61	1.94	0.75	0.03	0.28
	8.18	946974.0	118371.7	143792.0	0.93	0.61	1.94	0.75	0.03	0.27
	8.20	945912.0	118239.0	139135.0	0.93	0.63	1.95	0.75	0.03	0.27
	8.80	932666.0	116583.2	128892.0	0.94	0.67	2.14	0.83	0.03	0.18
	∞	921194.0	115149.2	118012.0	0.95	0.68	2.18	0.84	0.03	0.16
			AveMaxR = 137802.5		*MaxMaxR* = 143792.0					
0.06	8.85	1672413.0	209051.6	248500.0	0.94	0.59	1.27	0.76	0.02	0.28
	9.00	1669826.0	208728.2	245913.0	0.94	0.60	1.28	0.76	0.02	0.27
	9.50	1663922.0	207990.2	233064.0	0.94	0.64	1.37	0.78	0.02	0.22
	10.00	1630631.0	203828.9	213333.0	0.96	0.66	1.37	0.76	0.02	0.20
	∞	1625672.0	203209.0	208374.0	0.96	0.68	1.42	0.81	0.02	0.14
			AveMaxR = 235202.5		*MaxMaxR* = 248500.0					
0.03	11.20	2762061.0	345257.6	355502.0	0.94	0.63	1.01	0.76	0.01	0.24
	11.50	2755400.0	344425.0	355502.0	0.97	0.62	1.02	0.77	0.05	0.23
	12.00	2755400.0	344425.0	355502.0	0.97	0.62	1.02	0.77	0.05	0.23
	13.00	2751221.0	343902.6	353595.0	0.97	0.65	1.06	0.76	0.01	0.18
	∞	2749378.0	343672.2	351752.0	0.97	0.68	1.07	0.81	0.01	0.11
			AveMaxR = 355025.2		*MaxMaxR* = 355502.0					

less constrained instances. For the same reason, we may observe that the risk equity index *EI* values increase (i.e., risk equity decreases) and the average path dissimilarity index *AvgDI* values decrease with increasing values of *M*.

For the case with $k = 2$, Fig. 2 shows in more detail the impact of α and *M* on risk equity, quantitatively evaluated by the risk equity *EI* and the average path dissimilarity *AvgDI* indices. Distinct curves are depicted for different *M* values ranging from the "*smallest*" to the "*largest*" among the set of four limited *M* values chosen for each specific α, and for the case of "unlimited" (e.g., "∞") *M* value (see Table 2). As far as index *EI* is concerned, Fig. 2 shows that *EI* decreases (i.e., risk equity increases) with decreasing values of α (i.e., with increasing impacted zone sizes). The values of *EI* are only slightly greater for larger *M* values. As for the *AvgDI* index, we may note a larger variability of its values with respect to *M*, showing higher values for smaller *M* values, whereas *AvgDI* does not vary significantly with α. For the cases with $k = 4$, 6, and 8, we have experimented analog behaviors to the case with $k = 2$, although the variability of both the *EI* and *AvgDI* indices is smaller than that with $k = 2$.

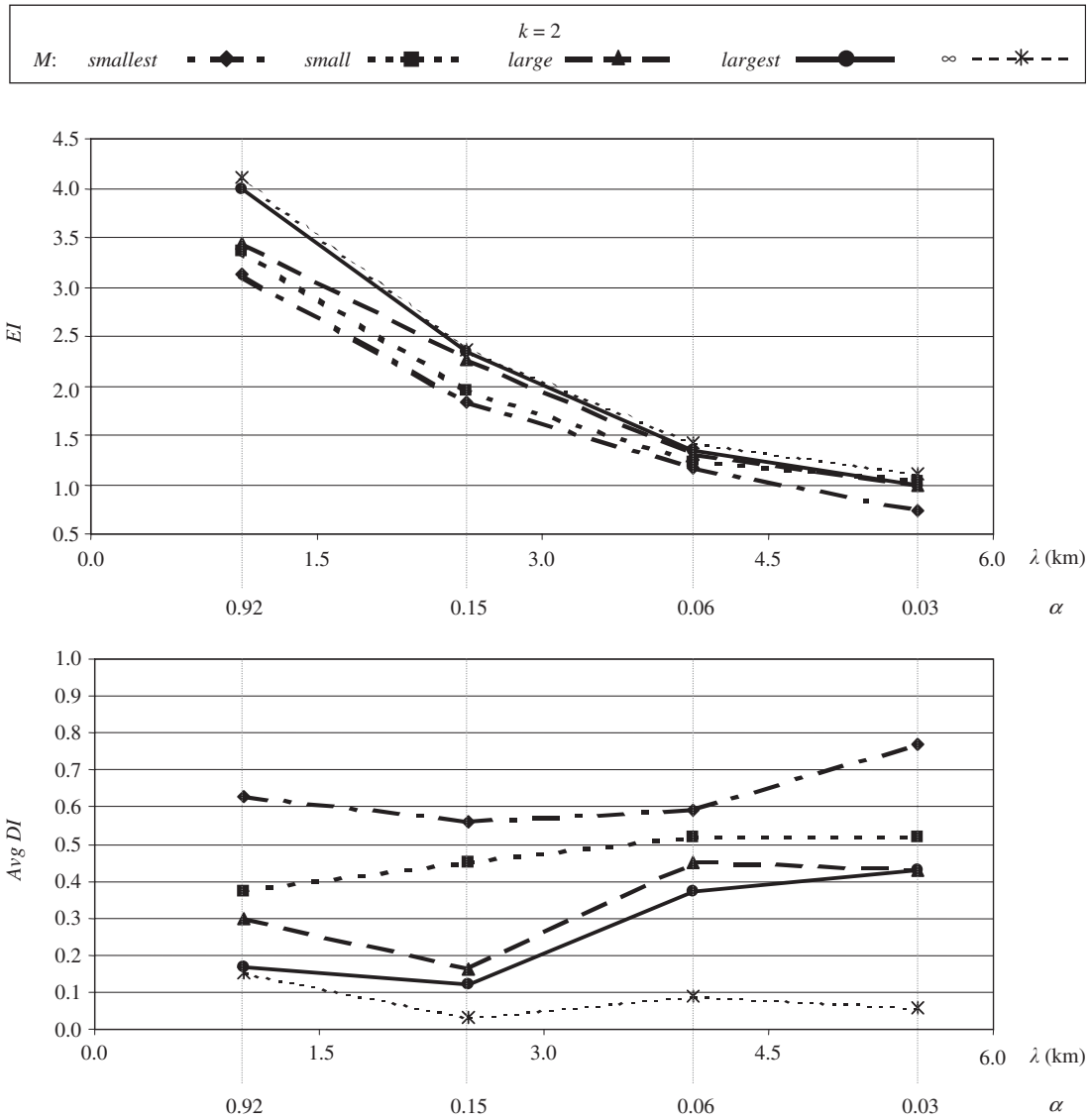

Fig. 2. The impact of α and M on the *EI* and *AvgDI* performance indices: the case with $k = 2$.

For the scenarios with ($k = 2$, $\alpha = 0.15$) and ($k = 2$, $\alpha = 0.03$), respectively, whose results are listed in Table 2, Figs. 3 and 4 show in detail the impact of M on the solutions with respect to the total path risk and risk equity. The figures also show the solutions' performance for the case with $M \to \infty$ (i.e., without risk threshold constraint). In particular, from both figures, it can be noted that *TotR* decreases and *EI* increases (i.e., risk equity decreases) by increasing M (i.e., moving from high constrained problem instances to low constrained ones).

Finally in Figs. 5 and 6, we show the structure of the computed solutions of the most constrained test problems (i.e., with the smallest M values) among the scenarios analyzed in Figs. 3 and 4, respectively.

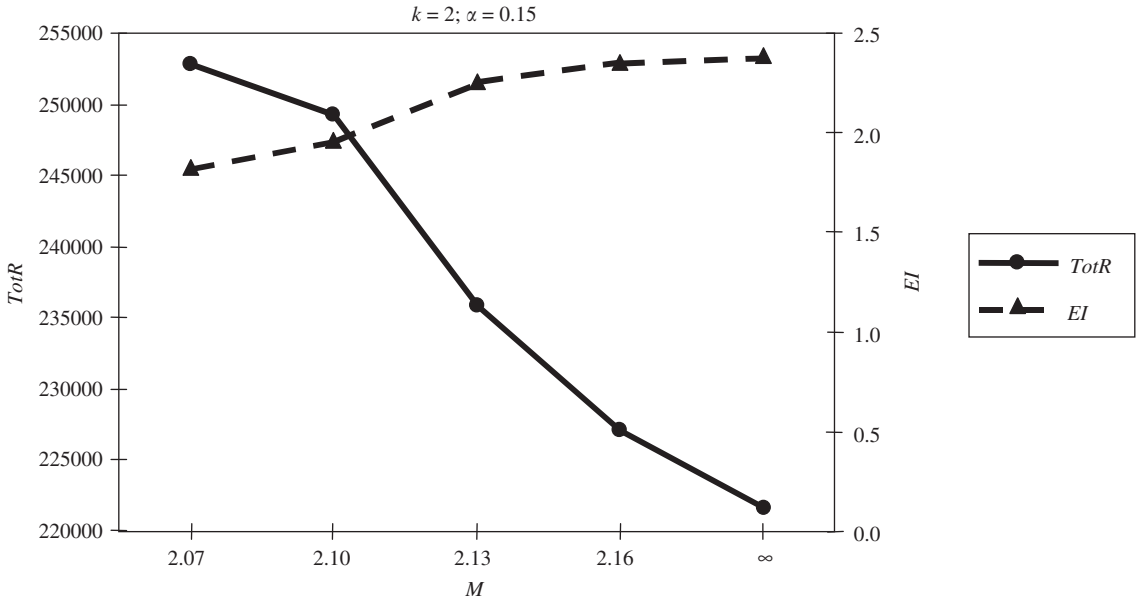

Fig. 3. Impact of M on *TotR* and *EI* solution measures: the scenarios with $k = 2$ and $\alpha = 0.15$.

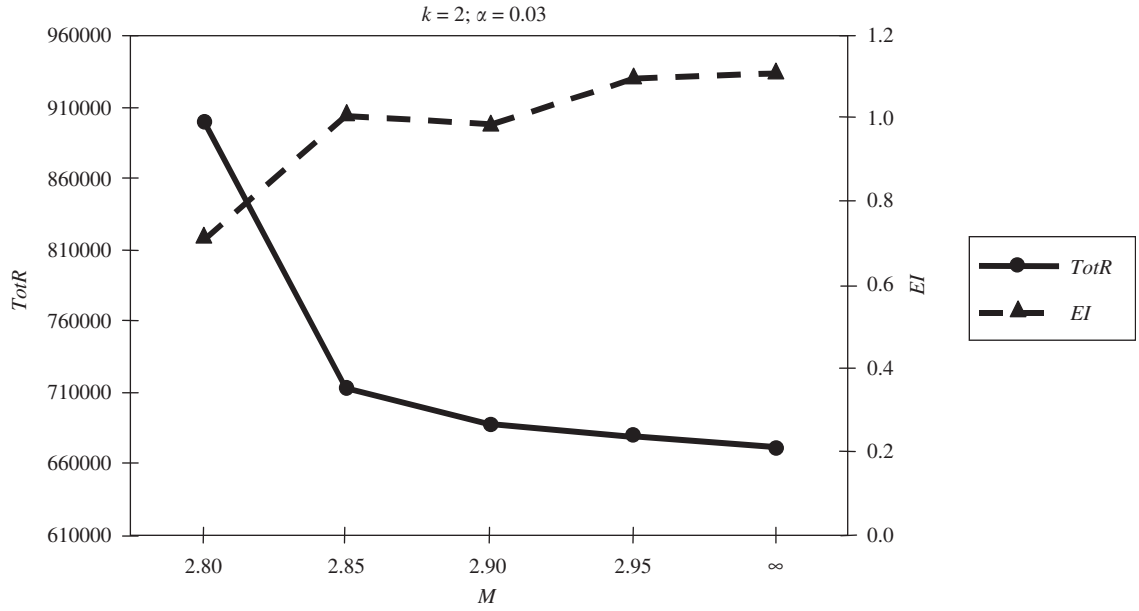

Fig. 4. Impact of M on *TotR* and *EI* solution measures: the scenarios with $k = 2$ and $\alpha = 0.03$.

Let us call "Scenario A" (with $k = 2$, $\alpha = 0.15$, $M = 2.07$) the test case whose solution is depicted in Fig. 5, and "Scenario B" (with $k = 2$, $\alpha = 0.03$, $M = 2.80$) the latter one whose solution is depicted in Fig. 6. In both figures, the $k = 2$ selected solution paths from Orvieto to Latina are represented in dashed and bold type, while the shortest path of 166 km is depicted in plain and bold type. Note that in both

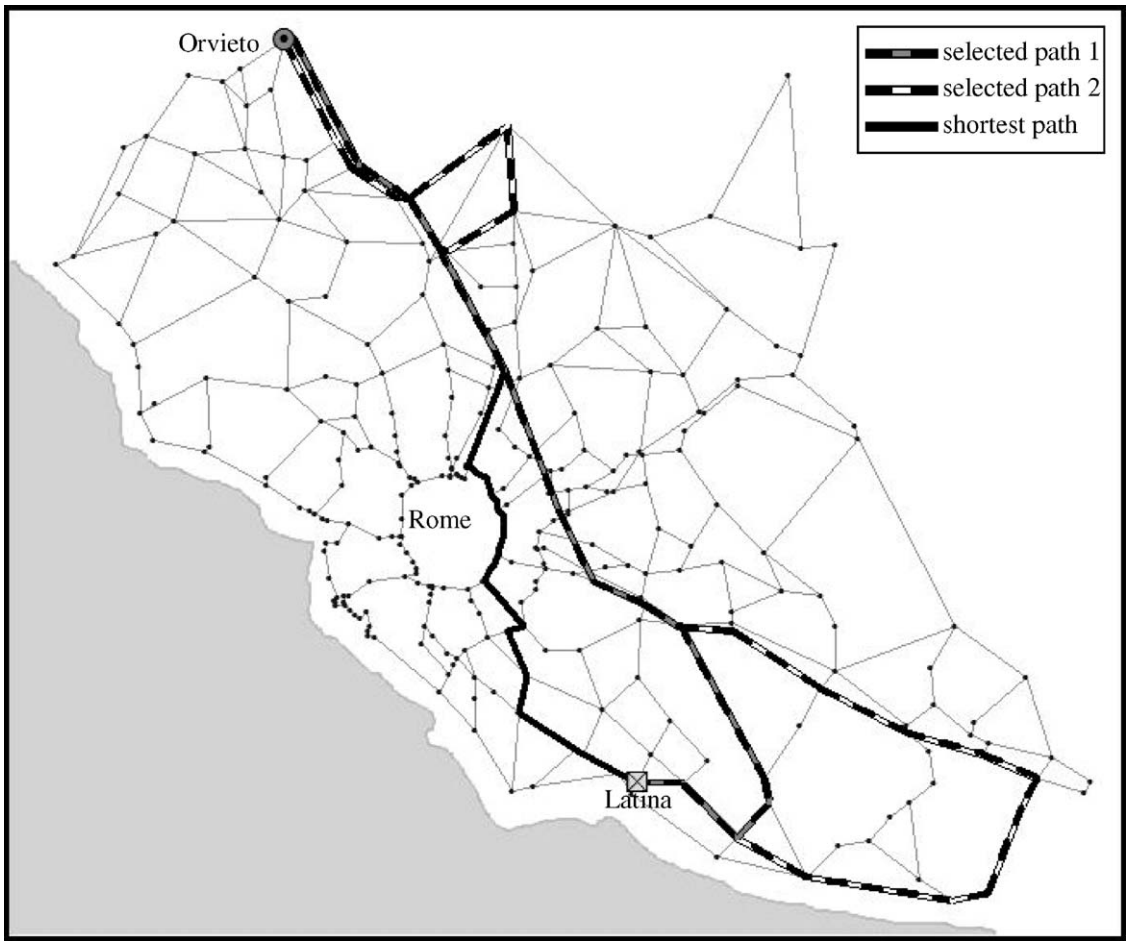

Fig. 5. Shortest path and minimum risk 2-paths for "Scenario A" ($k = 2$, $\alpha = 0.15$, $M = 2.07$).

scenarios, while the two solution paths are outside the city of Rome, the shortest path passes through the ring road encompassing Rome. The figures show that, for both scenarios, the two selected paths may be considered sufficiently dissimilar, observing that the main segment where they overlap corresponds to a highway. Nevertheless, it is easy to observe that the two selected paths for Scenario A (see Fig. 5) are much closer to each other than the selected paths for Scenario B (see Fig. 6), due to a smaller impacted zone (higher α) for Scenario A. This is confirmed by the *AvgDI* values (see Table 2) which are equal to 0.56 and 0.77, respectively, for Scenarios A and B. At the same time, for Scenario B we also have a more equitable risk distribution than for Scenario A ($EI = 1.83$ for Scenario A, and $EI = 0.73$ for Scenario B). Moreover, comparing the values of indices *RI* and *LI* ($RI = 0.87$ and 0.74, and $LI = 0.51$ and 0.44, for Scenarios A and B, respectively), we have that in Scenario B both the average path risk and the average path length are higher than their minimum values compared to Scenario A, whereas in Scenario B the selected paths imply a higher degree of risk equity.

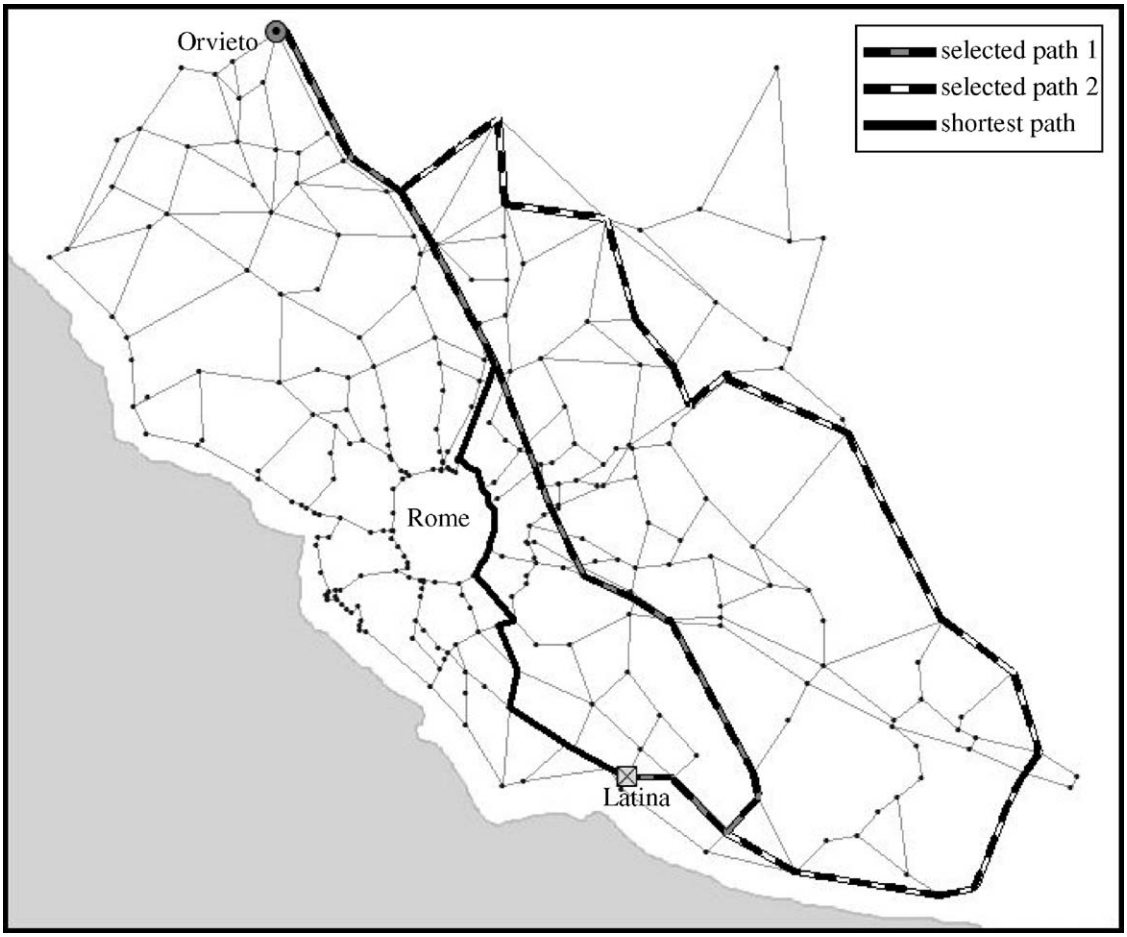

Fig. 6. Shortest path and minimum risk 2-paths for "Scenario B" ($k = 2$, $\alpha = 0.03$, $M = 2.80$).

6. Conclusions

In this paper, we have proposed a new model and new algorithms for generating minimum risk routes for hazmat shipments that exhibit also a sort of risk equity over the exposed population. In particular, risk equity is taken into account by considering the propagation of risks resulting from close paths and by limiting the total risk on each populated link. Besides the total risk to be minimized, different performance indices have also been considered for evaluating the generated paths. The model and algorithms have been experimentally evaluated on an Italian regional road network.

Several refinements may be introduced to the basic model presented in this paper. One of them may be to consider other kinds of distance-sensitive damage functions, that might include, for example, pollutant dispersion issues in the absence or with wind effects.

As a final remark, it should be considered that the issues in the hazmat transportation field are not only about risk assessment or the selection of safest routes, but also deal with the safe scheduling of hazmat

shipments when there are many hazmat shipments executed at the same time in the network. Integrating routing and scheduling decisions may be an important issue that can be addressed in this case also in the planning phase. One way could be that of devising a two-stage approach for routing and scheduling hazmat shipments, where in the first stage, a set of minimum and equitable risk alternative routes is generated (e.g., by the model and methods presented in this paper) for each hazmat shipment request, and in the second stage, a route, among the generated ones in the first stage, and a departure time are assigned to each hazmat shipment request, so as to assure a safe scheduling. For example, a hazmat shipment safe schedule should avoid a hazmat vehicle arriving on a link when the risk is high, due to another hazmat vehicle traveling at the same time in the link neighborhood area. In particular, this two-stage approach and the hazmat shipment scheduling problem to be solved in the second stage are addressed in [22].

Acknowledgements

The authors are grateful to the Guest Editor and the Referees for their valuable comments and suggestions that contributed to improve the quality of the paper. The authors are also indebted with Graziano Galiano for his aid in the algorithm implementation.

References

[1] Abkovitz M, Eiger M, Srinivasan S. Estimating the release rates and costs of transporting hazardous waste. Transportation Research Records 1984;977:22–30.
[2] Patel MH, Horowitz AJ. Optimal routing of hazardous materials considering risk of spill. Transportation Research 1994;28A:119–32.
[3] Zhang J, Hodgson J, Erkut E. Using GIS to assess the risks of hazardous materials transport in networks. European Journal of Operational Research 2000;121:316–29.
[4] Erkut E, Verter V. Modeling of transport risk for hazardous materials. Operations Research 1988;46:625–42.
[5] Batta R, Chiu SS. Optimal obnoxious paths on a network: transportation of hazardous materials. Operations Research 1988;36:84–92.
[6] Keeney RL. Equity and public risk. Operations Research 1980;28:527–34.
[7] Gopalan R, Batta R, Karwan MH. The equity constrained shortest path problem. Computers and Operations Research 1990;17:297–307.
[8] Current J, Ratick S. A model to assess risk, equity and efficiency in facility location and transportation of hazardous materials. Location Science 1995;3:187–201.
[9] Gopalan R, Kolluri KS, Batta R, Karwan MH. Modeling of equity in the transportation of hazardous materials. Operations Research 1990;38:961–73.
[10] Akgün V, Erkut E, Batta R. On finding dissimilar paths. European Journal of Operational Research 2000;121:232–46.
[11] Johnson PE, Joy DS, Clarke DB, Jacobi JM. HIGHWAY 3.01, an enhanced highway routing model: program, description, methodology, and revised user's manual. Oak Ridge National Laboratory, ORNL/TM-12124. Oak Ridge, TN, 1992.
[12] Ruphail NM, Ranjithan SR, ElDessouki W, Smith T, Brill ED. A decision support system for dynamic pre-trip route planning. applications of advanced technologies, in: Transportation engineering: proceedings of the fourth international conference. 1995. p. 325–9.
[13] Lombard K, Church RL. The gateway shortest path problem: generating alternative routes for a corridor location problem. Geographical Systems 1993;1:25–45.
[14] Kuby M, Zhongyi X, Xiaodong X. A minimax method for finding the *k* best differentiated paths. Geographical Analysis 1997;29:315–29.
[15] Erkut E. The discrete *p*-dispersion problem. European Journal of Operational Research, 1990;46:48–60.

[16] Erkut E, Ülküsal Y, Yeniçerioglu O. A comparison of *p*-dispersion heuristics. Computers and Operations Research 1994;21:1103–13.
[17] Dell'Olmo P, Gentili M, Scozzari A. On finding dissimilar Pareto-optimal paths. European Journal of Operational Research, 2005;162:70–82.
[18] Yen JY. Finding the *k* shortest loopless paths in a network. Management Science 1971;17:712–6.
[19] Jonkman SN, van Gelder PHAJM, Vrijling JK. An overview of quantitative risk measures for loss of life and economic damage. Journal of Hazardous Materials 2003;99:1–30.
[20] Held M, Wolfe P, Crowder HP. Validation of subgradient optimization. Mathematical Program 1974;6:62–88.
[21] Prais M, Ribeiro CC. Reactive GRASP: an application to a matrix decomposition problem in TDMA traffic assignment. INFORMS Journal on Computing 2000;12:164–76.
[22] Carotenuto P, Giordani S, Ricciardelli S, Rismondo S. A tabu search approach for scheduling hazmat shipments. Computers and Operations Research, 2004, submitted for publication, doi:10.1016/j.cor.2005.06.004.