Robust Shortest Path Problems with Two Uncertain Multiplicative Cost Coefficients

Changhyun Kwon

Department of Industrial and Systems Engineering
University at Buffalo, SUNY, USA

Taehan Lee

Department of Industrial and Information Systems Engineering
Chonbuk National University, Korea
Corresponding author. E-mail: `myth0789@jbnu.ac.kr`

Paul Berglund

Department of Industrial and Systems Engineering
University at Buffalo, SUNY, USA

Feburary 15, 2012

Abstract

We consider a robust shortest path problem when the cost coefficient is a multiplication of two uncertain parameters. We first show that the robust problem can be solved by a dual variable enumeration with with shortest path problems as subproblems. We also propose a path enumeration solution approach using a K-shortest paths finding algorithm that may be efficient in many real cases. An application in hazardous materials transportation is discussed and the solution methods are illustrated by numerical examples.

Keywords: robust shortest path, budgeted uncertainty, hazardous materials transportation

1 Introduction

For a directed and weighted graph $G(\mathcal{N}, \mathcal{A})$ we are interested in the following shortest path problem:

$$\min_{x \in \Omega} \sum_{e \in \mathcal{A}} p_{ij} c_{ij} x_{ij} \tag{1}$$

where

$$\Omega \equiv \left\{ x : \sum_{(i,j) \in \mathcal{A}} x_{ij} - \sum_{(j,i) \in \mathcal{A}} x_{ji} = b_i \quad \forall i \in \mathcal{N}, \text{ and } x_{ij} \in \{0,1\} \quad \forall (i,j) \in \mathcal{A} \right\}$$

The parameter b_i has the following values:

$$b_i = \begin{cases} 1 & \text{if } i = \text{origin} \\ -1 & \text{if } i = \text{destination} \\ 0 & \text{otherwise} \end{cases}$$

While the nonnegative cost parameters p_{ij} and c_{ij} appear as a multiplicative form in the objective function, this nominal problem (1) can be solved as a regular shortest path problem when p_{ij} and c_{ij} are known. The problem (1) arises in hazardous materials transportation where p_{ij} and c_{ij} represent the accident probability and the accident consequence, respectively. However, the solution methods proposed in this paper can be applied to any general model in the form of (1). We will discuss an application to hazardous materials transportation in Section 6.

In many realistic cases, however, accurate estimates of the parameters p_{ij} and c_{ij} may be unavailable. When only one of the parameters p_{ij} and c_{ij} is unknown and its values are given as a set of uncertainty, a general form of robust shortest path problems reported in the literature (Bertsimas and Sim, 2004, 2003; Chaerani et al., 2005; Kouvelis and Yu, 1996) minimizes the worst case cost as follows:

$$\min_{x \in \Omega} \max_{\tilde{c} \in \mathcal{C}} \tilde{c}^T x \tag{2}$$

where $\mathcal{C}$ is the set of possible realizations of the uncertain cost parameter $\tilde{c}$. The problem (2) minimizes the worst-case path cost under the set of uncertainty $\mathcal{C}$. When $\mathcal{C}$ is a box-constrained set or a budgeted box-constrained set, the problem (2) can be solved in polynomial time (Bertsimas and Sim, 2003). On the other hand, when $\mathcal{C}$ is an ellipsoid (Bertsimas and Sim, 2004; Chaerani et al., 2005) or a set of scenarios (Kouvelis and Yu, 1996), the problem (2) becomes NP-hard in general.

In this paper, we consider a class of robust shortest path problems, when the cost coefficient is a multiplication of two uncertain values. In such problems, the worst-case path cost can become bigger than when the multiplication of two uncertain parameters is considered as a single uncertain coefficient. In particular, we consider the uncertain parameters $\tilde{p}_{ij}$ and $\tilde{c}_{ij}$ represented in the following budgeted box-constrained uncertainty set:

$$\tilde{p}_{ij} = p_{ij} + q_{ij}u_{ij}$$

$$\tilde{c}_{ij} = c_{ij} + d_{ij}v_{ij}$$

where

$$u_{ij} \in U = \left\{ u : 0 \le u_{ij} \le 1 \quad \forall (i,j), \quad \sum_{(i,j)} u_{ij} \le \Gamma_u \right\}$$

$$v_{ij} \in V = \left\{ v : 0 \le v_{ij} \le 1 \quad \forall (i,j), \quad \sum_{(i,j)} v_{ij} \le \Gamma_v \right\}$$

and Γ_u and Γ_v are positive integers. This structure extends Bertsimas and Sim (2003). The parameters Γ_u and Γ_v are called the budget of uncertainty and represent the risk attitude of decision makers: the larger the budget of uncertainty is, the more risk-averse the decision maker is.

We show that the above robust shortest path problem with two uncertain coefficients can be solved by a method based on dual variable enumeration and regular shortest path problems. We provide another algorithm based on enumeration that uses a K-shortest path finding algorithm whose worst-case complexity is exponential, but may be efficient in real cases. We find an application in hazardous materials transportation to find the safest path and illustrate the algorithms by numerical examples.

2 The Robust Problem

In this paper, we consider a robust optimization model of the form:

$$\min_{x\in\Omega} \max_{u\in U, v\in V} \sum_{(i,j)} (p_{ij} + q_{ij}u_{ij})(c_{ij} + d_{ij}v_{ij})x_{ij} \tag{3}$$

$$= \min_{x\in\Omega} \left[\sum_{(i,j)} p_{ij}c_{ij}x_{ij} + \max_{u\in U, v\in V} \sum_{(i,j)} (q_{ij}c_{ij}x_{ij}u_{ij} + p_{ij}d_{ij}x_{ij}v_{ij} + q_{ij}d_{ij}x_{ij}u_{ij}v_{ij}) \right]$$

We note that the inner maximization problem is a disjoint bilinear program (DBP) for any given x. Although the problem is not a convex optimization problem, an optimal solution exists at an extreme point (Floudas and Visweswaran, 1994). While DBP is NP-hard in general (Vicente et al., 1992), this special case can be solved efficiently with transformation to a linear program. The inner problem is equivalent to: for any x

$$\max_{u,v,w} \sum_{(i,j)} (q_{ij}c_{ij}x_{ij}u_{ij} + p_{ij}d_{ij}x_{ij}v_{ij} + q_{ij}d_{ij}x_{ij}w_{ij}) \tag{4}$$

subject to

$$\begin{aligned}
u_{ij} &\le 1 \qquad (\rho_{ij}) \\
v_{ij} &\le 1 \qquad (\mu_{ij}) \\
-u_{ij} + w_{ij} &\le 0 \qquad (\eta_{ij}) \\
-v_{ij} + w_{ij} &\le 0 \qquad (\pi_{ij}) \\
\sum_{(i,j)} u_{ij} &\le \Gamma_u \qquad (\theta_u) \\
\sum_{(i,j)} v_{ij} &\le \Gamma_v \qquad (\theta_v) \\
u_{ij}, v_{ij}, w_{ij} &\ge 0
\end{aligned}$$

When Γ_u and Γ_v are positive integers, we can easily show that the polytope defined by the constraints of problem (4) is integral. Therefore, the optimal u, v and w are binary, given that Γ_u and Γ_v are positive integers.

Let us consider the dual problem of problem (4) with the corresponding dual variables in parentheses:

$$\min_{\theta_u,\theta_v,\rho_{ij},\mu_{ij},\eta_{ij},\pi_{ij}} \Gamma_u\theta_u + \Gamma_v\theta_v + \sum_{(i,j)}(\rho_{ij} + \mu_{ij}) \tag{5}$$

subject to

$$\rho_{ij} - \eta_{ij} + \theta_u \geq q_{ij}c_{ij}x_{ij} \tag{6}$$

$$\mu_{ij} - \pi_{ij} + \theta_v \geq p_{ij}d_{ij}x_{ij} \tag{7}$$

$$\eta_{ij} + \pi_{ij} \geq q_{ij}d_{ij}x_{ij} \tag{8}$$

$$\rho_{ij}, \mu_{ij}, \eta_{ij}, \pi_{ij}, \theta_u, \theta_v \geq 0 \tag{9}$$

Using strong duality, we can write the robust optimization problem (3) as:

$$\min \Gamma_u\theta_u + \Gamma_v\theta_v + \sum_{(i,j)}(p_{ij}c_{ij}x_{ij} + \rho_{ij} + \mu_{ij}) \tag{10}$$

subject to

$$\begin{aligned} x &\in \Omega \\ \rho_{ij} - \eta_{ij} + \theta_u &\geq q_{ij}c_{ij}x_{ij} \\ \mu_{ij} - \pi_{ij} + \theta_v &\geq p_{ij}d_{ij}x_{ij} \\ \eta_{ij} + \pi_{ij} &\geq q_{ij}d_{ij}x_{ij} \\ \rho_{ij}, \mu_{ij}, \eta_{ij}, \pi_{ij}, \theta_u, \theta_v &\geq 0 \end{aligned}$$

which is a mixed integer linear program (MILP). It has been shown that the problem (3) can be solved by solving a finite number of shortest path problems when either Γ_u or Γ_v is zero (Bertsimas and Sim, 2003).

2.1 Comparison with the Existing Approach

We can compare the robust problem (3) with the regular approach (Bertsimas and Sim, 2003) considering a single uncertain cost vector. Let us define the single uncertain cost vector $\tilde{m}_{ij} = \tilde{p}_{ij}\tilde{c}_{ij}$

so that

$$\tilde{m}_{ij} \in [p_{ij}c_{ij}, (p_{ij} + q_{ij})(c_{ij} + d_{ij})] \equiv [m_{ij}, m_{ij} + n_{ij}] \tag{11}$$

In addition we define $z^{UV}(\Gamma_u, \Gamma_v)$ to denote the optimal objective function value of the proposed approach in (3) with budgets of uncertainty Γ_u and Γ_v, and $z^W(\Gamma_w)$ to denote the optimal objective function value of the regular approach with the single uncertain cost vector in (11) with the single budget of uncertainty Γ_w. That is,

$$z^{UV}(\Gamma_u, \Gamma_v) = \min_{x \in \Omega} \max_{u \in U, v \in V} \sum_{(i,j)} (p_{ij} + q_{ij}u_{ij})(c_{ij} + d_{ij}v_{ij})x_{ij}$$

$$z^W(\Gamma_w) = \min_{x \in \Omega} \max_{w \in W} \sum_{(i,j)} (m_{ij} + n_{ij}w_{ij})x_{ij}$$

where

$$W = \left\{ w : 0 \leq w_{ij} \leq 1 \quad \forall (i,j), \quad \sum_{(i,j)} w_{ij} \leq \Gamma_w \right\}$$

When the regular approach is used, a challenge is how to determine the budget Γ_w. Determining Γ_u and Γ_v is easier, because we could directly observe the data source for each data type to determine the budgets of uncertainty. On the other hand, the new single parameter m is *manipulated* after data collection. Therefore, it is not obvious how we should determine how uncertain the new manipulated data is. We may determine $\Gamma_w = \Gamma_u = \Gamma_v$; then we always have $z^W(\Gamma_w) \leq z^{UV}(\Gamma_u, \Gamma_v)$. That is, the worst-case may not be captured by the regular approach; therefore, we need the proposed approach to consider the real worst-case even in the special case. We further provide the following results without proof:

Lemma 1. Depending on the budgets of uncertainty, we can compare the proposed approach with the regular approach (Bertsimas and Sim, 2003) as follows:

1. If $\Gamma_w = \Gamma_u = \Gamma_v$, we always have $z^W(\Gamma_w) \leq z^{UV}(\Gamma_u, \Gamma_v)$.
2. If $\Gamma_w \leq \max(\Gamma_u, \Gamma_v)$, we always have $z^W(\Gamma_w) \leq z^{UV}(\Gamma_u, \Gamma_v)$.
3. If $\Gamma_w = \Gamma_u + \Gamma_v$, we always have $z^W(\Gamma_w) \geq z^{UV}(\Gamma_u, \Gamma_v)$.

In the second case of Lemma 1, one would determine Γ_w at the maximum budget of Γ_u and Γ_v, but the worst-case still cannot be captured with the regular approach. Setting $\Gamma_w = \Gamma_u +$

Γ_v as in the third case would capture the worst-case with the regular approach, but in general it leads to unnecessarily conservative results. Therefore, one may choose Γ_w in the interval of $[\max(\Gamma_u, \Gamma_v), \Gamma_u + \Gamma_v]$, but a proper choice is ambiguous to make. Therefore, to capture the real worst-case in two multiplicative uncertain coefficient cases, the regular approach proposed by Bertsimas and Sim (2003) is hardly justified, and the proposed approach considering two separate uncertainties should be used. We further discuss the justification of the robust model proposed in this paper later.

3 A Dual Variable Enumeration Approach

In this section, we present a dual variable enumeration method. First, we will represent the solution of (5)-(9) in terms of θ_v and x. We first observe that there exists an optimal solution such that

$$\eta_{ij} + \pi_{ij} = q_{ij} d_{ij} x_{ij} \qquad \forall (i,j) \in \mathcal{A} \tag{12}$$

for any given x. This is because minimizing η_{ij} and π_{ij} as much as possible may lead to smaller values of θ_u, ρ_{ij} and μ_{ij} which decrease the objective function value, and η_{ij} and π_{ij} are not present in the objective function. Therefore there exists an optimal solution such that

$$0 \leq \eta_{ij} \leq q_{ij} d_{ij} x_{ij}$$
$$0 \leq \pi_{ij} \leq q_{ij} d_{ij} x_{ij}$$

for all $(i,j) \in \mathcal{A}$ and we can determine η_{ij} and π_{ij} by some allocation of $q_{ij} d_{ij} x_{ij}$ between η_{ij} and π_{ij} such that other constraints are satisfied.

Now suppose that θ_u and θ_v are fixed. Then we can write the constraints (6) and (7) as

$$\rho_{ij} = \max(q_{ij} c_{ij} x_{ij} - \theta_u + \eta_{ij}, 0)$$
$$\mu_{ij} = \max(p_{ij} d_{ij} x_{ij} - \theta_v + \pi_{ij}, 0)$$

Increasing η_{ij} and π_{ij} may increase ρ_{ij} and μ_{ij} and consequently increase the objective function value. Therefore, we need to find a way to allocate $q_{ij} d_{ij} x_{ij}$ to η_{ij} and π_{ij} without increasing the objective function value. We observe that, if $q_{ij}c_{ij}x_{ij} - \theta_u < 0$ for some $(i,j) \in \mathcal{A}$, we can increase η_{ij} without increasing ρ_{ij}. Therefore, an optimal solution in this case is to allocate $q_{ij}d_{ij}x_{ij}$ to η_{ij} as much as we can, that is, until $q_{ij}c_{ij}x_{ij} - \theta_u + \eta_{ij} = 0$. We can apply a similar argument to μ_{ij} and π_{ij}. On the other hand, if $q_{ij}c_{ij}x_{ij} - \theta_u \geq 0$ and $p_{ij}d_{ij}x_{ij} - \theta_v \geq 0$, then any allocation to η_{ij} and π_{ij} will have the same impact to the objective function value, since the cost coefficients of ρ_{ij} and μ_{ij} are identical.

From the above observations, we can consider the following optimal allocation rules for η_{ij} and π_{ij} when θ_u and θ_v are fixed:

Case 1. For $(i,j) \in \mathcal{A}$ such that

$$q_{ij}c_{ij}x_{ij} - \theta_u \geq 0$$
$$p_{ij}d_{ij}x_{ij} - \theta_v \geq 0$$

any combination of $\eta_{ij} \geq 0$ and $\pi_{ij} \geq 0$ such that $\eta_{ij} + \pi_{ij} = q_{ij}d_{ij}x_{ij}$ is optimal.

Case 2. For $(i,j) \in \mathcal{A}$ such that

$$q_{ij}c_{ij}x_{ij} - \theta_u < 0$$
$$p_{ij}d_{ij}x_{ij} - \theta_v \geq 0$$

we can first allocate some of $q_{ij}d_{ij}x_{ij}$ to η_{ij} until $q_{ij}c_{ij}x_{ij} - \theta_u + \eta_{ij} = 0$ in advance, and then allocate any remaining amount to η_{ij} and π_{ij} by any combination. For example, we can consider $\eta_{ij} = q_{ij}d_{ij}x_{ij}$ and $\pi_{ij} = 0$.

Case 3. For $(i,j) \in \mathcal{A}$ such that

$$q_{ij}c_{ij}x_{ij} - \theta_u \geq 0$$
$$p_{ij}d_{ij}x_{ij} - \theta_v < 0$$

we can first allocate some of $q_{ij}d_{ij}x_{ij}$ to π_{ij} until $p_{ij}d_{ij}x_{ij} - \theta_v + \pi_{ij} = 0$ in advance, and then allocate any remaining amount to η_{ij} and π_{ij} by any combination. For example, we can consider $\pi_{ij} = q_{ij}d_{ij}x_{ij}$ and $\eta_{ij} = 0$.

Case 4. For $(i, j) \in \mathcal{A}$ such that

$$q_{ij}c_{ij}x_{ij} - \theta_u < 0$$
$$p_{ij}d_{ij}x_{ij} - \theta_v < 0$$

we can first allocate some of $q_{ij}d_{ij}x_{ij}$ to η_{ij} until $q_{ij}c_{ij}x_{ij} - \theta_u + \eta_{ij} = 0$, and then allocate any remaining amount to π_{ij} until $p_{ij}d_{ij}x_{ij} - \theta_v + \pi_{ij} = 0$. If there is any remaining amount, we can allocate additionally to η_{ij} and π_{ij} by any combination.

Using the above optimal allocation rules, we can obtain an optimal solution for given θ_u and θ_v. From Cases 3 and 4, when $p_{ij}d_{ij}x_{ij} - \theta_v < 0$, we can consider allocating $\min(q_{ij}d_{ij}x_{ij}, \theta_v - p_{ij}d_{ij}x_{ij})$ to π_{ij} without worrying about η_{ij}. From Cases 1 and 2, when $p_{ij}d_{ij}x_{ij} - \theta_v \geq 0$, we can simply put $\pi_{ij} = 0$. Therefore, an optimal allocation to π_{ij} is

$$\pi_{ij} = \min(q_{ij}d_{ij}x_{ij}, \max(\theta_v - p_{ij}d_{ij}x_{ij}, 0)) \tag{13}$$

for all $(i, j) \in \mathcal{A}$ and any given $\theta_v \geq 0$. Using the condition (12), we have the corresponding allocation to η_{ij} as

$$\eta_{ij} = q_{ij}d_{ij}x_{ij} - \min(q_{ij}d_{ij}x_{ij}, \max(\theta_v - p_{ij}d_{ij}x_{ij}, 0)) \tag{14}$$

for all $(i, j) \in \mathcal{A}$ and any given $\theta_v \geq 0$. Note that the expression (14) is not dependent on θ_u.

From (13) and (14), we can now determine ρ_{ij} and μ_{ij} as follows:

$$\begin{aligned}
\rho_{ij} &= \max(q_{ij}c_{ij}x_{ij} - \theta_u + \eta_{ij}, 0) \\
&= \max(q_{ij}c_{ij}x_{ij} + q_{ij}d_{ij}x_{ij} - \min(q_{ij}d_{ij}x_{ij}, \max(\theta_v - p_{ij}d_{ij}x_{ij}, 0)) - \theta_u, 0) \\
\mu_{ij} &= \max(p_{ij}d_{ij}x_{ij} - \theta_v + \pi_{ij}, 0) \\
&= \max(p_{ij}d_{ij}x_{ij} + \min(q_{ij}d_{ij}x_{ij}, \max(\theta_v - p_{ij}d_{ij}x_{ij}, 0)) - \theta_v, 0)
\end{aligned}$$

We can express $\rho_{ij} + \mu_{ij}$ as a function of x_{ij} with cost coefficients dependent on θ_u and θ_v.

Lemma 2. The sum $\rho_{ij} + \mu_{ij}$ can be expressed as follows:

$$\rho_{ij} + \mu_{ij} = \begin{cases} 0 \cdot x_{ij} \\ \qquad\qquad \text{if } \theta_u \geq q_{ij}c_{ij},\ \theta_v \geq p_{ij}d_{ij} + q_{ij}d_{ij} \\ \qquad\qquad \text{or } p_{ij}d_{ij} \leq \theta_v \leq p_{ij}d_{ij} + q_{ij}d_{ij},\ \theta_u + \theta_v \geq p_{ij}d_{ij} + q_{ij}c_{ij} + q_{ij}d_{ij} \\ \\ (q_{ij}c_{ij} - \theta_u)x_{ij} \qquad\qquad\qquad\qquad \text{if } \theta_u \leq q_{ij}c_{ij}, \theta_v \geq p_{ij}d_{ij} + q_{ij}d_{ij} \\ \\ (p_{ij}d_{ij} + q_{ij}c_{ij} + q_{ij}d_{ij} - \theta_u - \theta_v)x_{ij} \\ \qquad\qquad \text{if } p_{ij}d_{ij} \leq \theta_v \leq p_{ij}d_{ij} + q_{ij}d_{ij},\ \theta_u + \theta_v \leq p_{ij}d_{ij} + q_{ij}c_{ij} + q_{ij}d_{ij} \\ \qquad\qquad \text{or } \theta_u \leq q_{ij}c_{ij} + q_{ij}d_{ij}, \theta_v \leq p_{ij}d_{ij} \\ \\ (p_{ij}d_{ij} - \theta_v)x_{ij} \qquad\qquad\qquad\qquad \text{if } \theta_u \geq q_{ij}c_{ij} + q_{ij}d_{ij}, \theta_v \leq p_{ij}d_{ij} \end{cases}$$

for each $(i, j) \in \mathcal{A}$ and all $\theta_u \geq 0$ and $\theta_v \geq 0$.

Proof. For the simplicity of notation, we drop the subscript ij and let $c_0 = pc$, $c_1 = pd$, $c_2 = qc$, $c_3 = qd$. We obtain

$$\begin{aligned} \mu &= \max[c_1x - \theta_v + \min\{c_3x, \max(\theta_v - c_1x, 0)\}, 0] \\ &= \max[\min\{(c_1 + c_3)x - \theta_v, \max(c_1x - \theta_v, 0)\}, 0] \\ &= \begin{cases} \max[\max(c_1x - \theta_v, 0)\}, 0] & \text{if } (c_1 + c_3)x - \theta_v \geq 0 \\ \max[(c_1 + c_3)x - \theta_v, 0] & \text{if } (c_1 + c_3)x - \theta_v \leq 0 \end{cases} \\ &= \begin{cases} \max(c_1x - \theta_v, 0) & \text{if } (c_1 + c_3)x - \theta_v \geq 0 \\ 0 & \text{if } (c_1 + c_3)x - \theta_v \leq 0 \end{cases} \end{aligned}$$

$$
\begin{aligned}
&= \begin{cases} c_1x - \theta_v & \text{if } c_1x - \theta_v \geq 0 \\ 0 & \text{if } -c_3 \leq c_1x - \theta_v \leq 0 \\ 0 & \text{if } c_1x - \theta_v \leq -c_3 \end{cases} \\
&= \max(c_1x - \theta_v, 0) \\
&= \max(c_1 - \theta_v, 0)x
\end{aligned}
$$

where the last equality is due to the binarity of x. Again using the binarity of x, we can express ρ as follows:

$$
\begin{aligned}
\rho &= \max[(c_2 + c_3)x - \theta_u - \min\{c_3x, \max(\theta_v - c_1x, 0)\}, 0] \\
&= \max[(c_2 + c_3) - \theta_u - \min\{c_3, \max(\theta_v - c_1, 0)\}, 0]x
\end{aligned}
$$

For three intervals of θ_v, we obtain the following expressions of ρ:

$$
\rho = \begin{cases} \max(c_2 - \theta_u, 0) & \text{if } \theta_v \geq c_1 + c_3 \\ \max(c_1 + c_2 + c_3 - \theta_u - \theta_v, 0) & \text{if } c_1 \leq \theta_v \leq c_1 + c_3 \\ \max(c_2 + c_3 - \theta_u, 0) & \text{if } \theta_v \leq c_1 \end{cases}
$$

Therefore, we obtain the lemma. □

Lemma 2 means that we can obtain an optimal solution to (10) by solving a nominal shortest path problem when θ_u and θ_v are given.

Figure 1 illustrates how the value of $\rho_{ij} + \mu_{ij}$ varies according to Lemma 2 in each region of (θ_u, θ_v) for each $(i, j) \in \mathcal{A}$. We observe that in each shaded region of (θ_u, θ_v), the cost coefficient of x_{ij} in $\rho_{ij} + \mu_{ij}$ becomes a linear function of θ_u and θ_v. We extend this observation to all links in $\mathcal{A}$.

Let $\{a_k\}$ be the ordered sequence of $q_{ij}c_{ij} + q_{ij}d_{ij}$ and $q_{ij}c_{ij}$ for all $(i, j) \in \mathcal{A}$ and 0, which is the set of θ_u-values where nonlinearlity occurs in $\rho_{ij} + \mu_{ij}$. Similarly, let $\{b_l\}$ be the ordered sequence of $p_{ij}d_{ij}$ and $p_{ij}d_{ij} + q_{ij}d_{ij}$ for all $(i, j) \in \mathcal{A}$ and 0 , for such θ_v-values. Let $\{f_m\}$ be the ordered sequence of $p_{ij}d_{ij} + q_{ij}c_{ij} + q_{ij}d_{ij}$ for all $(i, j) \in \mathcal{A}$ and 0 , for such $(\theta_u + \theta_v)$-values.

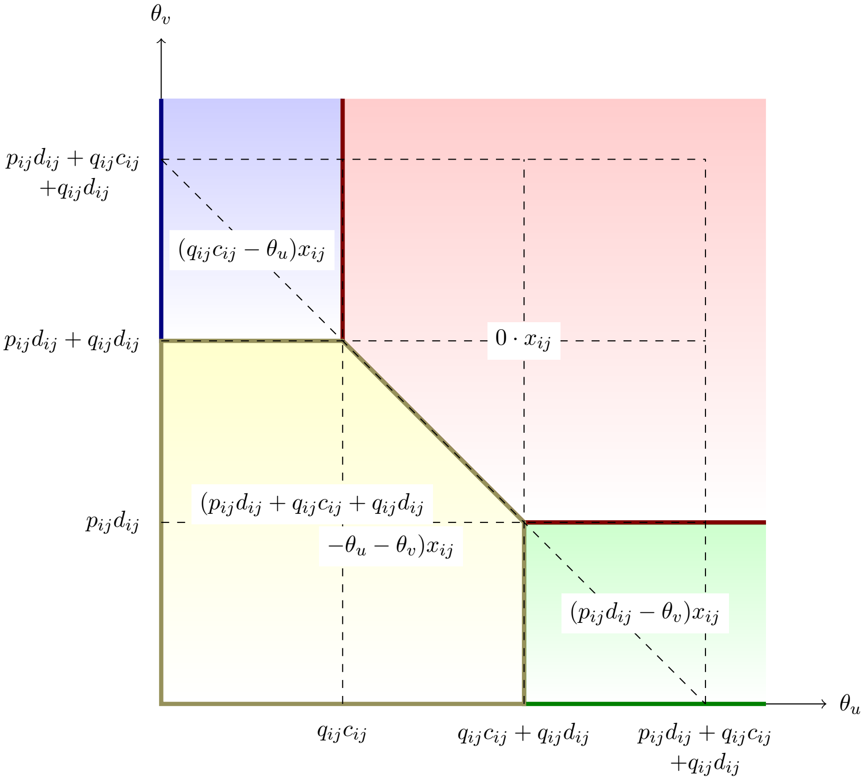


Figure 1: The value of $\rho_{ij} + \mu_{ij}$ for each interval of (θ_u, θ_v) for each link $(i, j) \in \mathcal{A}$

We consider the following problem:

$$Z_{klm} = \min_{x,\theta_u,\theta_v} \Gamma_u\theta_u + \Gamma_v\theta_v + \sum_{(i,j)}(p_{ij}c_{ij}x_{ij} + \rho_{ij} + \mu_{ij}) \tag{15}$$

subject to

$$x \in \Omega \tag{16}$$

$$a_k \leq \theta_u \leq a_{k+1} \tag{17}$$

$$b_l \leq \theta_v \leq b_{l+1} \tag{18}$$

$$f_m \leq \theta_u + \theta_v \qquad \text{if } f_m \in [a_k + b_l, a_{k+1} + b_{l+1}] \tag{19}$$

$$f_{m+1} \geq \theta_u + \theta_v \qquad \text{if } f_{m+1} \in [a_k + b_l, a_{k+1} + b_{l+1}] \tag{20}$$

We note that within the (θ_u, θ_v)-space defined by (17)-(20) for each tuple (k, l, m), the cost coefficient of x becomes linear in θ_u and θ_v for all links $(i, j) \in \mathcal{A}$ and for all x. Therefore, for the

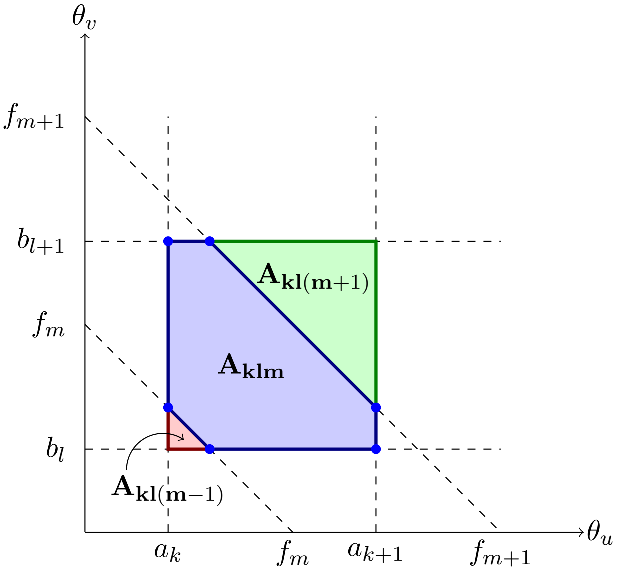


Figure 2: An illustration of decomposed feasible (θ_u, θ_v)-space on the mesh

problem Z_{klm}, we always obtain a solution at an extreme point of the feasible space defined by (17)-(20). This idea is represented in Figure 2. For the feasible region $\mathbf{A_{klm}}$ defined by (17)-(20) in Figure 2, we know one of the six extreme points is a solution of the problem (15). Therefore, we can solve the problem (15) by examining those six points; for each point, the problem is a regular shortest-path problem.

If we extend this idea to the entire (θ_u, θ_v)-space, we know that we can solve the robust shortest path problem by examining the following points:

- intersections of $\theta_u = \{a_k\}$ and $\theta_v = \{b_l\}$
- intersections of $\theta_u = \{a_k\}$ and $\theta_u + \theta_v = \{f_m\}$
- intersections of $\theta_v = \{b_l\}$ and $\theta_u + \theta_v = \{f_m\}$

Accordingly, we define the following three sets of (θ_u, θ_v):

$$\Theta_1 = \Big\{ (\theta_u, \theta_v) : \theta_u \in \{0\} \cup \{q_{ij}c_{ij} + q_{ij}d_{ij}, q_{ij}c_{ij} : (i,j) \in \mathcal{A}\}, \\ \theta_v \in \{0\} \cup \{p_{ij}d_{ij}, p_{ij}d_{ij} + q_{ij}d_{ij} : (i,j) \in \mathcal{A}\} \Big\}$$

$$\Theta_2 = \Big\{ (\theta_u, \theta_v) : \theta_u \in \{0\} \cup \{q_{ij}c_{ij} + q_{ij}d_{ij}, q_{ij}c_{ij} : (i,j) \in \mathcal{A}\}, \\ \theta_u + \theta_v \in \{0\} \cup \{p_{ij}d_{ij} + q_{ij}c_{ij} + q_{ij}d_{ij} : (i,j) \in \mathcal{A}\}, \quad \theta_v \geq 0 \Big\}$$

$$\Theta_3 = \Big\{ (\theta_u, \theta_v) : \theta_v \in \{0\} \cup \{p_{ij}d_{ij}, p_{ij}d_{ij} + q_{ij}d_{ij} : (i,j) \in \mathcal{A}\}, \\ \theta_u + \theta_v \in \{0\} \cup \{p_{ij}d_{ij} + q_{ij}c_{ij} + q_{ij}d_{ij} : (i,j) \in \mathcal{A}\}, \quad \theta_u \geq 0 \Big\}$$

We obtain the following theorem:

Theorem 1. Let us define the following problem with an arbitrary constraint set Θ:

$$Z(\Theta) = \min_{x \in \Omega, (\theta_u, \theta_v) \in \Theta} \Gamma_u \theta_u + \Gamma_v \theta_v + \sum_{(i,j)} (p_{ij}c_{ij}x_{ij} + \rho_{ij} + \mu_{ij}) \tag{21}$$

Then the robust shortest path problem (3) is equivalent to the following problem:

$$Z^* = \min\{Z(\Theta_1), Z(\Theta_2), Z(\Theta_3)\} \tag{22}$$

Now we obtain the computational complexity of (22).

Theorem 2. The computational complexity of (22) is $O(|\mathcal{N}|^6)$ and the number of shortest-path problems to be solved is $O(|\mathcal{N}|^4)$.

Proof. First we note that the sizes of feasible sets are

$$|\Theta_1| = O(|\mathcal{A}|^2), \qquad |\Theta_2| = O(|\mathcal{A}|^2), \qquad |\Theta_3| = O(|\mathcal{A}|^2)$$

When we consider $|\mathcal{A}| = O(\mathcal{N}^2)$, the number of shortest path problems we need solve to obtain Z^* is $O(|\mathcal{N}|^4)$. Since the complexity of Dijkstra's algorithm is $O(|\mathcal{N}|^2)$, the computational complexity of (22) is $O(|\mathcal{N}|^6)$. Since there are other algorithms with better worst-case complexity (Ahuja et al., 1990), this provides the upper bound. □

Obviously, the number of shortest-path problems to be solved in this method is significantly less than in the full primal variable (u and v) enumeration for which it is

$$\binom{|\mathcal{A}|}{\Gamma_u} \times \binom{|\mathcal{A}|}{\Gamma_v}$$

which is 1.0651×10^{20} when $|\mathcal{A}| = 150$, $\Gamma_u = 4$, and $\Gamma_v = 8$. In the same network, the dual variable enumeration method solves 151,148 shortest path problems (see Section 6.1).

4 A Path Enumeration Approach

In this section, we provide another algorithm whose worst-case complexity is exponential, but may be efficient in many real cases. Let us denote the nominal shortest path by l_1 and its corresponding objective function value and x-value by z_1 and x^1, respectively. That is,

$$z_1 = \min_{x \in \Omega} \sum p_{ij} c_{ij} x_{ij} = \sum p_{ij} c_{ij} x^1_{ij} \tag{23}$$

The corresponding maximum possible path cost is obtained by

$$z^R_1 = \max_{u \in U, v \in V} \sum (p_{ij} + q_{ij} u_{ij})(c_{ij} + d_{ij} v_{ij}) x^1_{ij}$$

which can be solved as a linear program; see (4). Similarly, we define z_k and z^R_k for any path l_k. Then, we note that the following relationship holds:

$$z_k \leq z^R_k \qquad \forall l_k$$

The objective of the robust problem is to find a path l_* that attains $z^R_* = \min_{l_k \in \mathcal{P}} z^R_k$. After solving the nominal shortest path problem (23), we know

$$z^R_* \leq z^R_1$$

Therefore, any path l_k whose nominal path cost is greater than z^R_k is not a solution to the robust problem, because for such path l_k, we have $z^R_* \leq z^R_1 < z_k \leq z^R_k$.

Let us now define the set of all paths whose nominal objective function value (z_k) is less than or equals to z_1^R

$$\mathcal{P}_c = \{l_1, l_2, \ldots, l_{|\mathcal{P}_c|}\}$$

We know that the set $\mathcal{P}_c$ contains the robust path. Therefore, once we have the set $\mathcal{P}_c$, computing z_k^R for all paths in the set requires solving $|\mathcal{P}_c|$ linear programs, whose dimension is as small as three times the number of arcs contained in each path l_k.

To determine the set $\mathcal{P}_c$, we can use any K-shortest paths finding algorithm that provides paths in the ascending order of path length and allows termination at any point when the path length exceeds a certain value. Therefore the complexity of finding the set $\mathcal{P}_c$ depends on the size of the set and the complexity of finding K-shortest paths. For example, we can use Yen's algorithm (Yen, 1971) whose complexity is $O(K|\mathcal{N}|(|\mathcal{A}| + |\mathcal{N}| \log |\mathcal{N}|))$. However, the number K for our algorithm is unknown *a priori*, and in the worst case K is equal to the number of all available paths; therefore the worst case complexity becomes exponential.

We can further reduce the computational effort by stopping the algorithm as soon as the nominal path cost exceeds the minimum value of z_k^R among the paths found so far. That is, when we construct the set $\mathcal{P}_c$, we update the reference cost value by the current value of $\min_{l_j \in \mathcal{P}_c} z_j^R$. The algorithm is summarized as follows:

Step 0. Find the nominal shortest path l_1 by solving a shortest path problem, and obtain the worst-case path cost z_1^R by solving the $3|l_1|$-dimensional linear program of the form (4). Set $k = 1$, $z^R = z_1^R$ and $l^* = l_1$.

Step 1. Find the next best nominal shortest path l_{k+1} with $z_{k+1} \leq z^R$. If no such path exists, stop. l^* is the optimal robust path.

Step 2. Obtain the worst-path cost z_{k+1}^R by solving the $3|l_{k+1}|$-dimensional linear program of the form (4). If $z_{k+1}^R \leq z^R$ then set $z^R = z_{k+1}^R$ and $l^* = l_{k+1}$. Set $k = k + 1$. Go to Step 1 and repeat.

5 An Abstract Example

In this section, we provide an example application of the proposed robust method on an abstract network presented in Figure 3. The test network consists of 15 nodes and 33 links with randomly

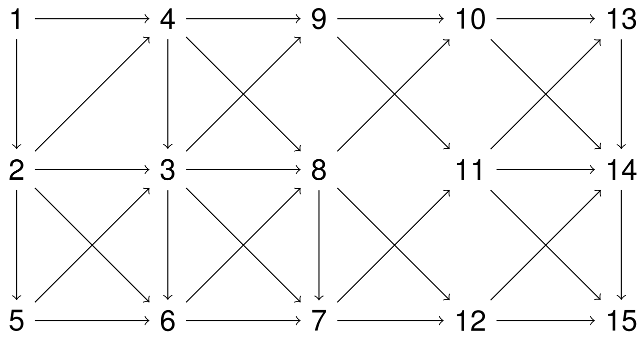


Figure 3: An abstract test network with 15 nodes and 33 links.

generated cost coefficients in Table 1. In addition, we used Node 1 and Node 15 as origin and destination, respectively.

We intend to compare the performance of the nominal shortest path, the robust shortest path by the Bertsimas-Sim (B-S) model (Bertsimas and Sim, 2003) with a single cost vector, and the robust shortest path proposed in this paper. We used $\Gamma_u = 2$ and $\Gamma_v = 3$ for the proposed model, and we tested both $\Gamma = 2$ and $\Gamma = 3$ for the B-S robust model.

Table 1: Cost coefficients used for the test network

(i,j)	p_{ij}	c_{ij}	q_{ij}	d_{ij}	(i,j)	p_{ij}	c_{ij}	q_{ij}	d_{ij}
$(1,4)$	1	67	11	32	$(7,11)$	85	95	40	37
$(1,2)$	97	67	34	44	$(7,12)$	85	24	34	27
$(2,4)$	28	44	48	4	$(8,7)$	23	59	30	45
$(2,3)$	70	81	3	16	$(8,10)$	42	39	26	47
$(2,5)$	74	57	10	30	$(8,12)$	42	72	46	44
$(2,6)$	23	4	34	41	$(9,10)$	14	65	32	33
$(3,9)$	37	71	50	20	$(9,11)$	72	85	29	42
$(3,8)$	96	0	8	24	$(10,13)$	62	68	33	49
$(3,6)$	24	13	46	38	$(10,14)$	100	20	30	34
$(3,7)$	46	77	32	6	$(11,14)$	45	58	35	14
$(4,3)$	59	82	33	21	$(11,15)$	46	49	35	31
$(4,9)$	77	73	19	42	$(11,13)$	89	39	44	15
$(4,8)$	56	44	14	30	$(12,15)$	22	27	37	44
$(5,3)$	18	23	9	35	$(12,14)$	21	53	0	23
$(5,6)$	31	39	39	4	$(13,14)$	32	21	1	35
$(6,7)$	100	9	47	28	$(14,15)$	16	56	26	8
$(6,8)$	4	53	19	10					

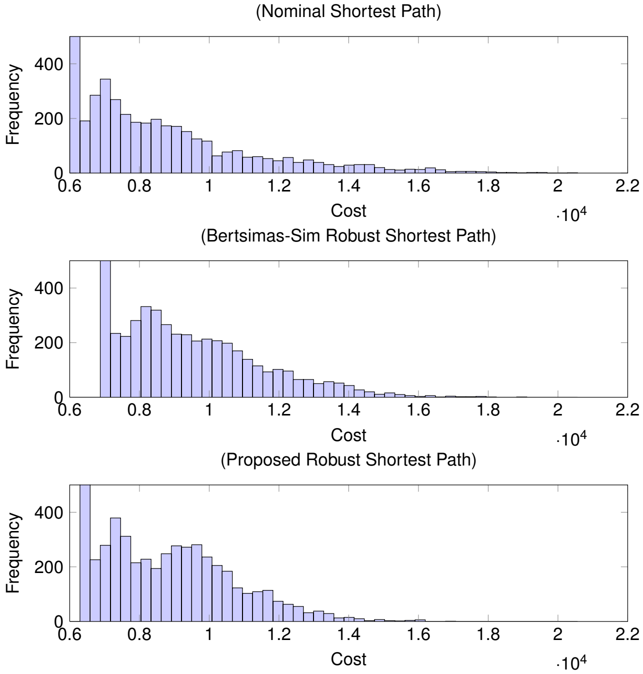


Figure 4: Histogram of the Realized Path Costs (10,000 samples)

The obtained nominal shortest path is

$$\{1, 4, 8, 12, 15\}$$

the B-S path with $\Gamma = 2$ is

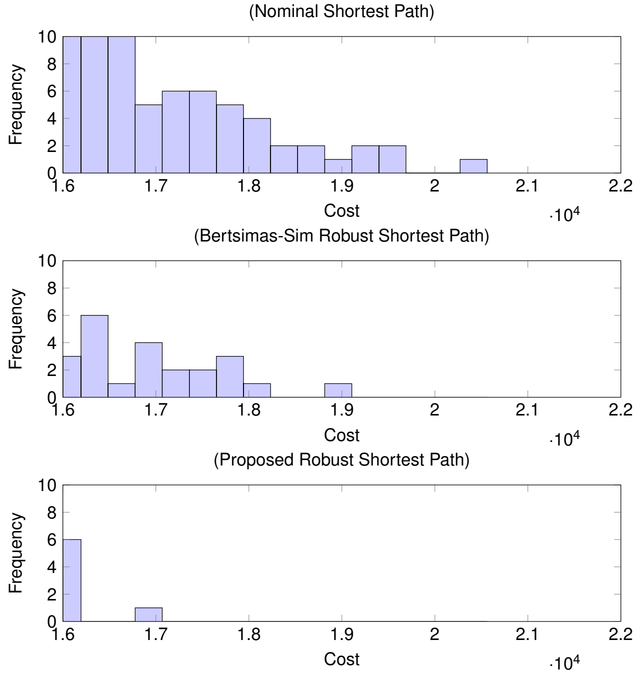
$$\{1, 4, 8, 10, 14, 15\}$$

and the proposed robust path in this paper with $\Gamma_u = 2$, $\Gamma_v = 3$ is

$$\{1, 4, 8, 7, 12, 15\}$$

The B-S path with $\Gamma = 3$ is same as the one with $\Gamma = 2$.

To observe the realized path costs, we randomly allocate the budget of uncertainty Γ_u and Γ_v to u_{ij} and v_{ij}, not necessarily 0 or 1; instead, we allow any value between 0 and 1. The distribution of 10,000 realized path costs for each path is presented in Figure 4. Since the most budget of uncertainty is allocated in the links that are not part of the shortest paths, we truncated the frequency of the smallest realized costs to observe the distirubtion better. We observe that in the


(Nominal Shortest Path)

(Bertsimas-Sim Robust Shortest Path)

(Proposed Robust Shortest Path)

Figure 5: Closer look to histogram of the realized path costs

nominal shortest path the distribution of the realized path costs is most widely spreaded, while it is least widely spreaded in the proposed robust shortest path.

Since we are most interested in the worst-case scenarios in this paper, we provide a closer-look in the long tail in Figure 5. We can clearly observe that we are not protected to the worst-case scenarios with the single budget of uncertainty without considering the multiplicative nature of uncertain cost coefficients. The biggest realized cost is 20,417 in the nominal shortest path, 18,950 in the B-S robust shortest path, and 16,800 in the proposed robust shortest path. A summary for the performance measures is provided in Table 2.

However, we do not intend to overly exaggerate the performance of the proposed robust model over the B-S robust model. First, the performance is largely dependent on the cost coefficient values, the network structure, and the choice of the budget of uncertainty. For example, as discussed in earlier in this paper, if we choose $\Gamma = 5$ for the B-S model, it will produce a path that is more robust than the path by the proposed robust model, although it may be too conservative. However, in some applications, it may be unclear how to choose the combined budget of uncertainty (Γ) for the multiplication of two cost coefficients, as opposed to two separated budgets of uncertainty (Γ_u and Γ_v). We will discuss this point in depth later in this paper. Second, the performance also depends on what measure we use. The performance test may lead to a different result with a different measure, for example, value-at-risk and conditional value-at-risk. Third, it is obvious that the B-S model has better computational complexity. For this test network, the computation time for the B-S model was 0.006 seconds, while for the proposed model it was 0.8110 seconds by the dual variable enumeration and 1.8272 seconds by the path enumeration. Algorithms were implemented with MATLAB and executed in a generic PC.

Table 2: Summary of path performances in the simulation

	Nominal Path	B-S Path	Proposed Path
worst-case (maximum) cost	20,417	18,950	16,800
mean cost	7,199	8,189	7,669
variance of cost	4,172,954	3,319,902	3,053,113

6 An Application to Hazardous Materials Transportation

Accidents involving hazardous materials (hazmat) are low-probability, high-consequence incidents. While the probability of hazmat accidents is very low, the consequences can be catastrophic. The U.S. had about 15,000 hazmat accidents in the year 1998 only 429 of which were classified as serious accidents (Kara and Verter, 2004). There are two important facts that make hazmat problems a proper application of the proposed robust optimization approach. First, historical data sufficient to construct a stochastic distribution are rarely available in realistic problems. Hazmat accident probabilities are hard to obtain because hazmat accidents, especially serious ones, are rare events. In real routing decision-making, the accident probabilities of general traffic are used to estimate the accident probabilities of hazmat trucks, but the two kinds of probabilities might be very different. We cannot know if they are different or the same, due to the lack of available data. In addition, the consequences of hazmat accidents depend on weather conditions like wind speed and direction, the number of people present at the time of an accident, the effectiveness of evacuation, the seriousness of the accident, etc. Therefore, the impacted number of people at a hazmat accident is very difficult to estimate and subject to uncertainty. Again, historical data is not usually available for consequences.

Second, the sources of the two types of data, probabilities and consequences, are different.

Accident probabilities are obtained from organizations like the U.S. Department of Transportation and its sub-divisions including the Pipeline and Hazardous Materials Safety Administration and the Federal Emergency Management Agency, while consequence data may be computed based on population information and travel pattern information available through the U.S. Census Bureau and the U.S. Commodity Flow Survey by the U.S. Bureau of Transportation Statistics. Therefore it is hard to determine a single budget of uncertainty for the regular approach with a single uncertain cost vector (Bertsimas and Sim, 2003) as described in Section 2.1.

In addition, it is unclear if the two data are correlated. One might think that the congestion level would be high in the links with high consequence level, and the accident probability is an increasing function of the congestion level; hence there exists positive correlation between the accident probability and the accident. However Martin (2002) observed that accident rates are highest when traffic is lightest and lowest when traffic level is modest. It is found that the accident rate decreases and then increases as the traffic volume increases. In light traffic, the accident rate is higher on weekends, while it is higher on weekdays in heavy traffic. Lord et al. (2005) observed that the relationship between accident rates and traffic flow cannot be described fully without considering vehicle density, level of service, vehicle occupancy, volume/capacity ratio, and speed distribution. In addition, these studies are for all vehicles, not exclusively for hazmat vehicles. The relationship between the accident probability of hazmat vehicles and the hazmat accident consequence is hard to study, because of lack of historical data. This indicates that it is unclear how the uncertainty of the accident probability and the accident consequence can be modeled as a single data type.

Although we used population data as the measure of the accident consequence for illustration purpose in the subsequent section, depending on how the consequence is measured, the relationship varies. In an uncongested road with low population, the accident consequence by the population measure is small. However, if a nuclear power plant is located nearby, a hazmat accident would bring unwanted catastrophe. Therefore, the relationship between the accident probability and the accident consequence also depends on the measure of consequence and it is unclear what the nature of correlation between two data is, if it exists.

Suppose that we have some estimates of hazmat accident probability and accident consequence, denoted by p_{ij} and c_{ij}, respectively, in each road segment (i, j). The expected consequence of a hazmat truck traveling along path l is as follows (Alp, 1995):

$$R^l = \sum_{(i_k,j_k)\in\mathcal{A}_l} \prod_{(i_h,j_h)\in\mathcal{A}_l, h<k} (1 - p_{i_h j_h}) p_{i_k j_k} c_{i_k j_k} \tag{24}$$

where $\mathcal{A}_l$ is the set of all arcs in path l, and (i_k, j_k) is the k-th arc in path l. The expression (24) assumes that the shipment terminates once an accident happens in any road segment. It is noted that accident probabilities p_{ij} are extremely small, usually in the range of 10^{-8} to 10^{-6} per mile traveled (Abkowitz and Cheng, 1988). Therefore we can approximate as

$$\prod_{(i_h,j_h)\in\mathcal{A}_l, h<k} (1 - p_{i_h j_h}) \approx 1$$

Consequently, we obtain the following approximation (Jin and Batta, 1997):

$$R^l \approx \sum_{(i,j)\in\mathcal{A}_l} p_{ij} c_{ij} \tag{25}$$

The resulting nominal problem to minimize the expected consequence is a shortest path of the form (1).

Due to lack of data, the accident probability p_{ij} is subject to uncertainty. In addition, the accident consequence c_{ij} is also hard to estimate. In many hazmat accidents, especially those involving explosives or poisonous gas, the weather conditions are important factors (Akgun et al., 2007). One needs to consider uncertain factors involving weather conditions to determine the safest path in hazmat transportation. However, accurate weather conditions can be difficult to obtain and the resulting accident consequences may be computed as interval data at best. Therefore, the robust optimization model (3) is a natural approach to minimize the expected consequence in the worst-case scenario in hazmat transportation.

6.1 Numerical Results

We provide numerical results of the proposed algorithms based on Albany, New York, USA and its nearby highway network. The transportation network considered consists of 90 nodes and 150 links as presented in Figure 6.

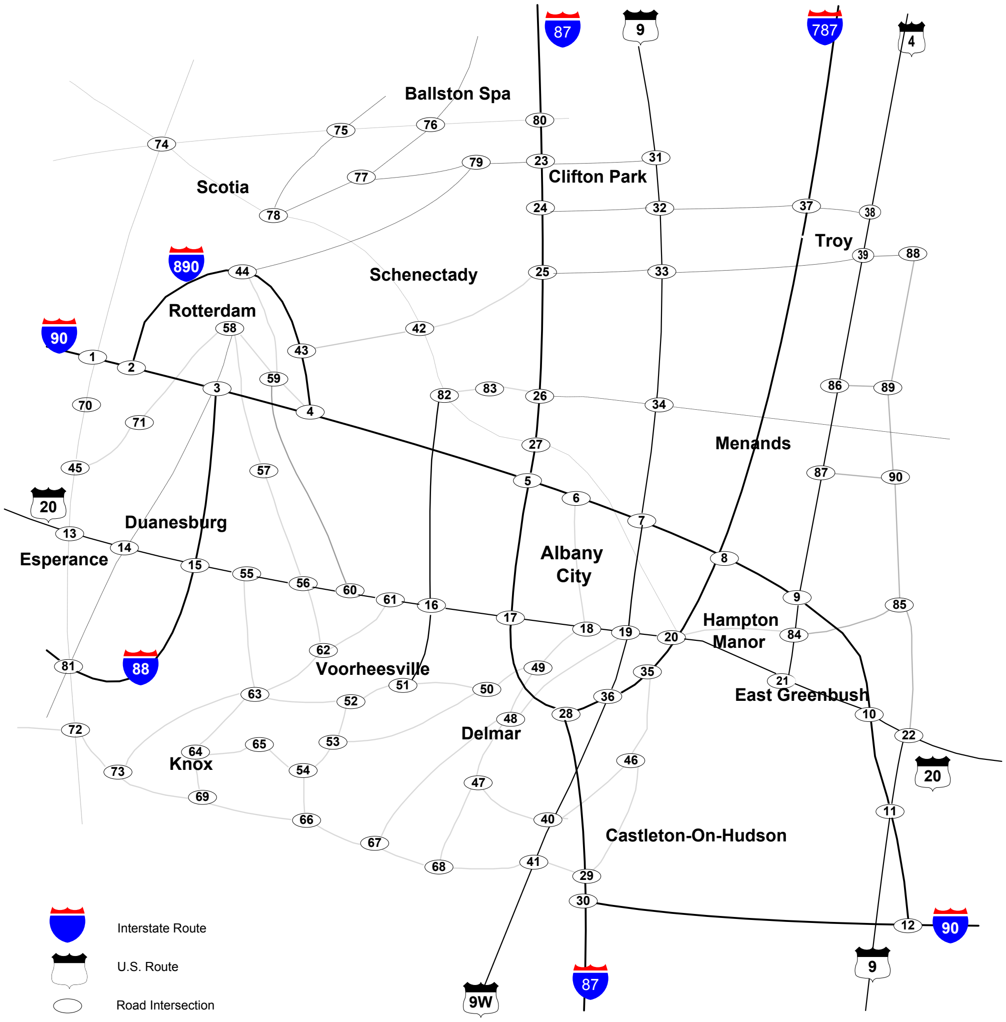


Figure 6: Albany Area Highway Network (90 nodes, 150 links)

The nominal accident probabilities are computed by $p_{ij} = 10^{-6}\times$ (length of link (i, j)) as in Abkowitz and Cheng (1988). The nominal accident consequences c_{ij} are computed using the λ-neighborhood concept developed by Batta and Chiu (1988). The road length and population statistics are obtained from Department of Transportation and Department of Commerce websites. We generated q_{ij} and d_{ij} randomly, but in the same order as the nominal coefficients, p_{ij} and c_{ij}. We used $\Gamma_u = 4$ and $\Gamma_v = 8$. While the randomly generated uncertain intervals do not represent any real scenario, our intention is to show that considering two multiplicative uncertain cost coefficients as a single coefficient may not capture the worst-case properly, and the robustness may not be proportional to the single budget of uncertainty with such consideration.

We find paths from origin node 1 to destination node 12. The nominal shortest path is

$$\text{NSP} = \{1, 70, 45, 13, 81, 72, 73, 69, 66, 67, 68, 41, 29, 30, 12\}$$

the B-S robust shortest path with $\Gamma = 4$ is

$$\text{BS4} = \{1, 70, 45, 13, 14, 15, 81, 72, 73, 63, 52, 51, 50, 49, 48, 47, 40, 41, 29, 30, 12\}$$

the B-S robust shortest path with $\Gamma = 8$ is

$$\text{BS8} = \{1, 70, 45, 13, 14, 15, 81, 72, 73, 69, 66, 67, 68, 41, 29, 30, 12\}$$

and the proposed robust shortest path with $\Gamma_u = 4$ and $\Gamma_v = 8$ is

$$\text{KLB48} = \{1, 70, 45, 13, 81, 72, 73, 63, 52, 51, 50, 49, 48, 47, 40, 41, 29, 30, 12\}$$

(KLB is from the initials of the authors' last names.) The Dijkstra's algorithm for the nominal shortest path took 0.023 seconds, and the B-S robust shortest paths with the corresponing algorithm (Bertsimas and Sim, 2003) were obtained after 0.106 seconds and 0.105 seconds of computing time, respectively. For the proposed robust shortest path computation, the dual variable enumeration method took 99.765 seconds after solving 151,148 shortest path problems, and the path enumeration method took 106.891 seconds after finding 1,344 paths. We used MATLAB at a generic PC running

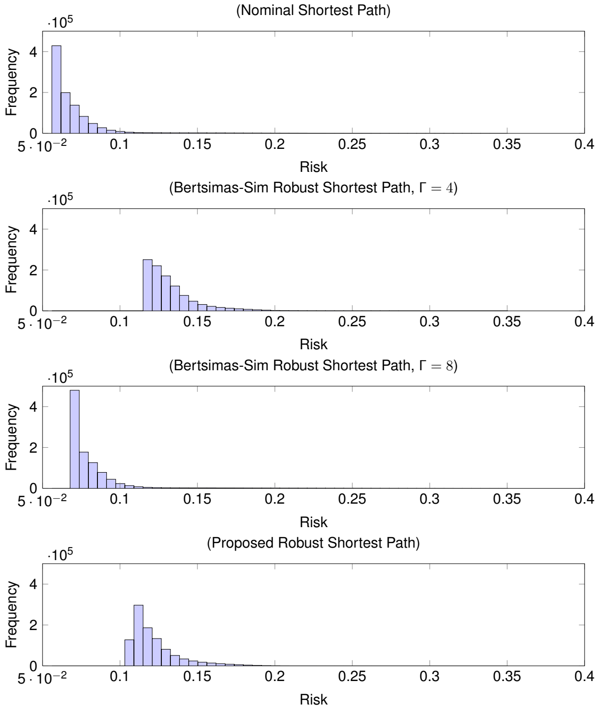
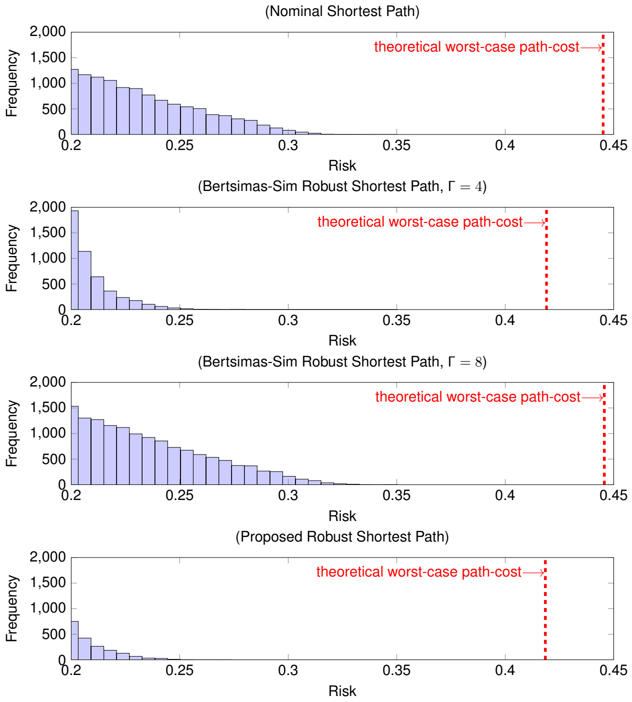


Figure 7: Histogram of the Realized Path Costs (1,000,000 samples)

Windows 7.

To test the path performances of the above four paths, we randomly allocate the budget of uncertainty ($\Gamma_u = 4, \Gamma_v = 8$), to $\{u_{ij}\}$ and $\{v_{ij}\}$, independently from each other. Although the worst-case happens when $\{u_{ij}\}$ and $\{v_{ij}\}$ are binary, we allowed any value between 0 and 1 in the simulation. We generated 1,000,000 samples and the histograms of path-costs are presented for all four paths in Figure 7. In the simulation, we assumed the link costs are distributed uniformly within the uncertain intervals, $\tilde{p}_{ij} \in [p_{ij}, p_{ij} + q_{ij}]$ and $\tilde{c}_{ij} \in [c_{ij}, c_{ij} + d_{ij}]$. Since the budget of uncertainty was allocated to links that are not included in the above four paths in many realizations, the far left path-costs in the histogram represent the nominal path-costs in the four paths. The NSP and BS8 paths show better nominal performances than the BS4 and KLB48 paths. However, higher path-costs, which are of interest in the robust optimization framework, show different patterns.

To observe path performances in the long tail, we provide a closer look to the histograms in Figure 8. The KLB48 path shows the best performance in the long-tail, which should not be a surprise since the KLB48 path is supposed so. One interesting observation in Figure 8 is that the BS8 path is weak to uncertainty and its performance is even worse than the NSP and BS4 paths.


(Nominal Shortest Path)

(Bertsimas-Sim Robust Shortest Path, Γ = 4)

(Bertsimas-Sim Robust Shortest Path, Γ = 8)

(Proposed Robust Shortest Path)

Figure 8: Closer look to histogram of the realized path costs

Using dashed lines, we also provide the theoretical worst-case path-costs, which the KLB48 path minimizes. The BS4 path is close to the KLB48 path in terms of the theoretical worst-case cost (the path selections are similar), and is better than the BS8 path.

A detailed summary is provided in Table 3. To compare the worst-case performances in other scenarios, we provide the theoretical worst-case path-costs in three cases: (1) treating two coefficients as a single coefficient with $\Gamma = 4$, denoted by one($\Gamma = 4$), (2) same treatment but with $\Gamma = 8$, denoted by one($\Gamma = 8$), (3) two coefficients separately as proposed in this paper, denoted by two($\Gamma_u = 4, \Gamma_v = 8$). While the theoretical worst-case path-costs do not differ very much among the four paths in the first two cases, there are more significant differences in the third case. This indicates that when more accurate robust path selection is necessary, considering two coefficients as a single coefficient may lead to non-robust path selection. In the later part in Table 3, we provide the worst-case, mean, and variance from the simulation results, which are consistent with the observations already explained.

We also compare the computing performance of the two proposed solution methods. We note that the number of shortest path problems to be solved in the dual variable enumerations method does not vary much with the parameter values, but the number of paths to be found and the number

Table 3: Summary of theoretical worst-case path-costs and the simulation results

	Nominal Path NSP	B-S Path ($\Gamma = 4$) BS4	B-S Path ($\Gamma = 8$) BS8	Proposed Path ($\Gamma_u = 4, \Gamma_v = 8$) KLB48
Theoretical worst-case path-cost				
one($\Gamma = 4$)[a]	0.2443	0.2380	0.2428	0.2386
one($\Gamma = 8$)	0.2843	0.2828	0.2821	0.2849
two($\Gamma_u = 4, \Gamma_v = 8$)[b]	0.4451	0.4190	0.4457	0.4185
Summary of realized path-costs (1,000,000 samples); two($\Gamma_u = 4, \Gamma_v = 8$)				
worst-case	0.3379	0.2793	0.3475	0.2697
mean	0.0714	0.1323	0.0817	0.1220
variance ($\times 10^{-4}$)	6.2520	2.4310	6.2840	2.4000

[a] one(Γ) means when two cost coefficients are treated as a single cost coefficient as in the B-S model.

[b] two(Γ_u,Γ_v)' means when two cost coefficients are treated as proposed in this paper.

of linear programming problems to be solve vary significantly in the path enumeration method. For the test purpose, we set

$$\left.\begin{aligned} \Gamma_u &= \Gamma_v = \tau \\ q_{ij} &= 0.1 \times \tau \times p_{ij} \\ d_{ij} &= 0.1 \times \tau \times c_{ij} \end{aligned}\right\} \quad \text{for } \tau = 1, 2, \ldots, 25$$

in the Albany network. As τ increases, the intervals of uncertainty becomes bigger, and therefore the number of paths to be found in the path enumeration method increases. Until τ is 14, the path enumeration method outperforms the dual variable enumeration method, but for τ bigger than 14, the dual variable enumeration method finds the optimal solution faster. The computation time and the number of paths found are reported in Figure 9.

7 Conclusions

The robust shortest path problem considered in this paper has the cost coefficients as multiplications of two uncertain parameters. We have shown that the problem can be solved using a dual variable enumeration method and a path enumeration method. While we found an application in hazardous materials transportation, we may apply our results to any risk mitigation problems

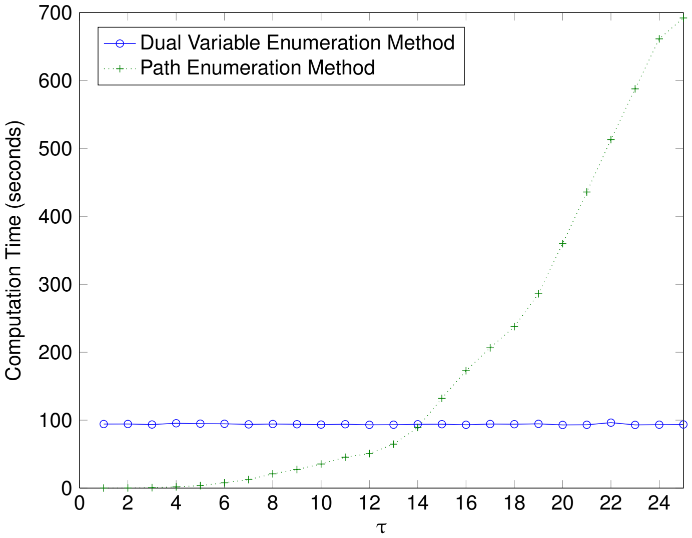


Figure 9: Computation Time for the Two Solution Methods

involving network-structured decision processes.

One drawback of the dual variable enumeration method is that it requires significantly more computation time than the model of Bertsimas and Sim (2003) in large-scale networks. We conjecture that the objective function is quasi-convex (not strictly) in the dual variables θ_u and θ_v. If it is true, we may be able to devise a more efficient search algorithm with a smaller number of shortest-path problems to be solved by reducing the search space on (θ_u, θ_v) and starting from a good choice of the initial (θ_u, θ_v). We leave this as a future research topic. Other potential extensions to this paper include robust risk minimization problem on time-dependent networks, for which this study can provide fundamentals.

References

Abkowitz, M. and Cheng, P. (1988). Developing a risk/cost framework for routing truck movements of hazardous materials. *Accident Analysis and Prevention*, 20(1):39.

Ahuja, R. K., Mehlhorn, K., Orlin, J., and Tarjan, R. E. (1990). Faster algorithms for the shortest path problem. *J. ACM*, 37:213–223.

Akgun, V., Parekh, A., Batta, R., and Rump, C. (2007). Routing of a hazmat truck in the presence of weather systems. *Computers & Operations Research*, 34(5):1351–1373.

Alp, E. (1995). Risk-based transportation planning practice: Overall methodology and a case example. *INFOR*, 33(1):4–19.

Bar-Gera, H. (2011). Transportation network test problems. http://www.bgu.ac.il/~bargera/tntp/. Accessed on April 4, 2011.

Batta, R. and Chiu, S. (1988). Optimal obnoxious paths on a network: Transportation of hazardous materials. *Operations Research*, 36(1):84–92.

Bertsimas, D. and Sim, M. (2003). Robust discrete optimization and network flows. *Mathematical Programming*, 98(1):49–71.

Bertsimas, D. and Sim, M. (2004). Robust Discrete Optimization Under Ellipsoidal Uncertainty Sets.

Chaerani, D., Roos, C., and Aman, A. (2005). The Robust Shortest Path Problem by Means of Robust Linear Optimization. In Fleuren, H., Hertog, D., and Kort, P., editors, *Operations Research Proceedings 2004*, volume 2004 of *Operations Research Proceedings*, chapter 42, pages 335–342–342. Springer Berlin Heidelberg, Berlin/Heidelberg.

Floudas, C. A. and Visweswaran, V. (1994). *Quadratic Optimization*, pages 217–270. Kluwer Academic Publishers.

Jin, H. and Batta, R. (1997). Objectives derived from viewing hazmat shipments as a sequence of independent Bernoulli trials. *Transportation Science*, 31(3):252–261.

Kara, B. and Verter, V. (2004). Designing a road network for hazardous materials transportation. *Transportation Science*, 38(2):188.

Kouvelis, P. and Yu, G. (1996). *Robust Discrete Optimization and Its Applications (Nonconvex Optimization and Its Applications)*. Springer, 1st edition.

Lord, D., Manar, A., and Vizioli, A. (2005). Modeling crash-flow-density and crash-flow-v/c ratio relationships for rural and urban freeway segments. *Accident Analysis & Prevention*, 37(1):185–199.

Martin, J. (2002). Relationship between crash rate and hourly traffic flow on interurban motorways. *Accident Analysis & Prevention*, 34(5):619–629.

Vicente, L. N., Calamai, P. H., and Júdice, J. J. (1992). Generation of disjointly constrained bilinear programming test problems. *Computational Optimization and Applications*, 1(3):299–306.

Yen, J. (1971). Finding the K shortest loopless paths in a network. *Management Science*, 17(11):712–716.