

# Robust Hazmat Network Design Problems Considering Risk Uncertainty

Longsheng Sun, Mark H. Karwan, and Changhyun Kwon\*

Department of Industrial and Systems Engineering  
University at Buffalo (SUNY), Buffalo, New York, USA

July 8, 2015

## Abstract

We study robust network design problems for hazardous materials transportation considering risk uncertainty. Risk uncertainty is considered in two ways: (1) uncertainty on each link for each shipment, and (2) uncertainty on each link across all shipments. We extend an existing heuristic framework to solve these two robust network design problems and propose a Lagrangian relaxation heuristic to solve subproblems within the framework. We present our computational experiences and illustrate general insights based on real networks.

**Keywords:** hazardous materials transportation; network design; robust optimization

## 1 Introduction

In this paper, we consider robust versions of the *hazardous materials (hazmat) network design problem* (HNDP) with the design decision being that government authorities can ban certain roads from transporting various hazardous materials. Kara and Verter (2004) define HNDP as follows: (1) given an existing road network, the hazmat network design problem involves selecting the road segments that should be closed so as to minimize total risk, and (2) the carriers will then choose the minimum cost routes on the resulting network. Hence the government should consider the behaviour of the carriers when designing the road network.

---

\*Corresponding Author: [chkwon@buffalo.edu](mailto:chkwon@buffalo.edu), +1-716-645-4705

Kara and Verter (2004) formulate HNBP as a bilevel model with the government as a leader in the upper level and the carriers as followers in the lower level. They transform the bilevel model into a single level mixed integer problem by substituting the lower level problem with its optimality conditions and solve the single model with a standard optimization solver (CPLEX). Erkut and Alp (2007) consider HNBP as a tree selection problem to guarantee only one route desired by the government is available for each origin destination (OD) pair. Erkut and Gzara (2008) generalize the problem considered by Kara and Verter (2004) to the undirected case and propose a heuristic method. They also formulate the problem as a biobjective bilevel model to include trade-offs between risk and cost. Verter and Kara (2008) present a path-based formulation to identify paths that are mutually acceptable by the government and the carriers. Amaldi et al. (2011) provide an exact formulation for the HNBP and use duality theory to transform the problem to a single level mixed integer problem. Gzara (2013) proposes a family of valid cuts and incorporates them within an exact cutting plane algorithm to solve the HNBP.

Besides banning certain road segments, government can also set tolls to regulate hazmat transportation. Marcotte et al. (2009) first propose the use of tolls in mitigating hazardous materials transport risk. Wang et al. (2012) extend the approach to a dual toll pricing method to simultaneously control both regular and hazmat vehicles to reduce risk. Bianco et al. (2015) consider toll policies to regulate hazardous material transportation to not only minimize the total risk but also to spread the risk in an equitable way. In order to avoid confusion, we note here that we consider only the regulation of banning certain road segments in this paper. By default, when we mention hazmat network design problems, we mean the hazmat network design problems with the regulation method of banning certain road segments.

In the current literature, very few works deal with uncertainty in HNBP. Only recently, Xin et al. (2013) consider the hazmat network design problem for risk uncertainty with interval data. However, generally, parameters have uncertainty and ignoring uncertainty can lead to an undesirable solution. This is especially critical in hazardous materials transportation since risk is complex due to different weather conditions, road conditions, potential population exposed etc. Hence, dealing with uncertainty is essential. In fact, how to measure risk is a main research topic in hazmat transportation (readers can refer to Section 3 of Erkut et al. (2007) for a review on risk assessment). While Xin et al. (2013) apply the maximum regret criterion on HNBP that may lead to a too conservative network design, we consider the worst-case risk-measures with uncertainty budget for flexible decision making.

Robust optimization (Ben-Tal and Nemirovski, 1998; El Ghaoui and Lebret, 1997) is a method proposed to deal with uncertainty. It has received much attention due to its tractability, conservativeness,

probability guarantees, and flexibility. This method constructs a solution that is feasible to any realization of the uncertainty in a given set and behaves well over all the likely uncertain outcomes. It does so by solving a problem that is no harder than the deterministic problem in many cases. Generalizing robust optimization to integer linear programming cases, Bertsimas and Sim (2003) considered robust discrete optimization problems and network flows using a cardinality uncertainty set for cost coefficients, and showed that those problems may be solved by enumerating a finite number of dual variable values.

Robust optimization has been utilized in many applications. In transportation, for example, robust optimization has been applied to two-stage network flow and design under demand uncertainty (Atamtürk and Zhang, 2007), capacity expansion of network flows (Ordóñez and Zhao, 2007), network design problem under transportation cost and demand uncertainty (Mudchanatongsuk et al., 2007), and facility location problem for hazardous waste transportation (Berglund and Kwon, 2014). A detailed introduction on the theory and application of robust optimization is found in Bertsimas et al. (2011).

In this paper, we examine robust versions of HNBP. Our contributions are as follows. First, we model the risk uncertainty in the hazmat network design problem using a cardinality uncertainty set in a robust optimization framework. Second, for large networks, we provide a Lagrangian relaxation algorithm with projection and a Golden Section Search method to obtain solutions efficiently. While the Lagrangian relaxation method is developed here to deal with large networks, it can be used for a class of robust optimization problems that cannot be solved by the dual-variable enumeration method of Bertsimas and Sim (2003). Third, we modify an existing algorithm (Erkut and Gzara, 2008) to solve the robust hazmat network design problem.

The remainder of the paper is organized as follows. The next section introduces the deterministic HNBP. In Section 3, we present the mathematical formulation of the robust HNBP using a cardinality uncertainty set. Section 4 provides the existing algorithmic framework and shows how to modify this algorithm to robust cases. In Section 5, we show how a Lagrangian heuristic and golden section search methods can be applied to solve larger networks. Computational results are shown in Section 6. Finally, Section 7 provides a conclusion and suggestions for future research.

## 2 Nominal Problem Formulation

We consider HNBP in which the government determines the available road segments by minimizing total risk and carriers choose routes on the resulting network to minimize cost. Suppose we have a transportation network that is defined by a graph  $\mathcal{G} = (\mathcal{N}, \mathcal{A})$ , where  $\mathcal{N}$  denotes the set of nodes (road intersections) and  $\mathcal{A}$  denotes the set of arcs (road segments). HNBP involves transporting  $\mathcal{S}$  shipments

between different origins and destinations. For each shipment  $s \in \mathcal{S}$ ,  $n_s$  is the corresponding number of shipments,  $r_{ijs}$  and  $c_{ijs}$  are the risk and cost associated with arc  $(i, j) \in \mathcal{A}$ . Let  $x_{ijs} = 1$  if arc  $(i, j)$  is used to transport shipment  $s$  and  $y_{ij} = 1$  if arc  $(i, j)$  is open to hazmat traffic. Then the problem can be formulated using a bilevel integer linear programming model (Kara and Verter, 2004) as

$$\min_{y_{ij} \in \{0,1\}} \sum_{(i,j) \in \mathcal{A}} \sum_{s \in \mathcal{S}} n_s r_{ijs} x_{ijs}, \quad (1)$$

where  $x_{ijs}$  is obtained by

$$\min_{x_{ijs}} \sum_{(i,j) \in \mathcal{A}} \sum_{s \in \mathcal{S}} n_s c_{ijs} x_{ijs}, \quad (2)$$

subject to

$$\sum_{i,k \in \mathcal{A}} x_{iks} - \sum_{k,i \in \mathcal{A}} x_{kis} = \begin{cases} +1 & i = o(s) \\ -1 & i = d(s) \\ 0 & \text{otherwise} \end{cases} \quad \forall i \in \mathcal{N}, s \in \mathcal{S}, \quad (3)$$

$$x_{ijs} \leq y_{ij} \quad \forall (i, j) \in \mathcal{A}, s \in \mathcal{S}, \quad (4)$$

$$x_{ijs} \in \{0, 1\} \quad \forall (i, j) \in \mathcal{A}, s \in \mathcal{S}. \quad (5)$$

The objective in (1) is the total risk on the entire network, which should be minimized by the government by choosing  $y_{ij}$  values to decide open arcs. The lower level problem (2)–(5) decides the routes with corresponding arcs  $x_{ijs}$  based on open segments. Here we assume carriers choose the shortest (least cost) path. The objective for the lower level problem in (2) is the cost for the carriers. Constraints (3) are the flow conservation requirements and constraints (4) restrict carriers from choosing arcs that are not open to hazmat transportation. Note this is a formulation for directed networks. For the undirected case, an additional constraint  $y_{ij} = y_{ji}$  for all  $(i, j) \in \mathcal{A}$  should be added to the upper level problem to ensure both arcs  $(i, j)$  and  $(j, i)$  are open to use if either direction is used for hazmat traffic (Erkut and Gzara, 2008).

### 3 Robust Formulations

In this section, we first briefly introduce robust optimization theory and show how it can be applied to avoid uncertainty. Then we propose several robust formulations for HNDP and describe their meanings.

### 3.1 Robust Optimization

Suppose we model a real-world problem using optimization techniques, and the parameters in our model are not deterministic but are estimated. Therefore we have to consider the uncertainty in our model to be more realistic. Robust optimization is a method to deal with this uncertainty. It allows us to construct an uncertainty set with our knowledge about the parameters. Then it generates a solution that is feasible to any realization of the specified uncertainty set. There are many ways to construct the uncertainty set. Here we give some details on using the cardinality uncertainty set to model uncertainty in the cost coefficients for a combinatorial optimization problem. We consider the following nominal problem:

$$\min_{x \in X} c^\top x \quad (6)$$

where  $c^\top x = \sum_{j=1}^n c_j x_j$  and  $X$  is a feasible set defined on  $\{0, 1\}^n$ . We consider uncertainty on the objective coefficients. For each  $j \in N = \{1, 2, \dots, n\}$ ,  $c_j$  can take any value in the interval  $[c_j, c_j + d_j]$ . Then the robust counterpart for (6) is

$$\min_{x \in X} \left\{ c^\top x + \max_{\{S | S \subseteq N, |S| \leq \Gamma\}} \sum_{j \in S} d_j x_j \right\}. \quad (7)$$

$\Gamma$  in model (7) is to control the level of conservativeness in the objective. It is referred to as the *budget of uncertainty*. We are interested in finding an optimal solution which optimizes against all scenarios under which a number  $\Gamma$  of cost coefficients can vary in such a way as to maximally influence the objective (Bertsimas and Sim, 2003). When  $\Gamma = 0$ , model (7) is equivalent to the nominal problem (6). As  $\Gamma$  increases, we consider more worst cases and become more conservative in generating the robust solution. A more explicit formulation for model (7) is as follows:

$$\min_{x \in X} \left\{ c^\top x + \max_{u \in \mathcal{U}} \sum_{j \in N} u_j d_j x_j \right\}, \quad (8)$$

where

$$\mathcal{U} = \left\{ u : \sum_{j \in N} u_j \leq \Gamma, 0 \leq u_j \leq 1 \forall j \in N \right\}. \quad (9)$$

An optimal solution to the inner maximization problem for (8) consists of  $\lfloor \Gamma \rfloor$  variables equal to 1 and one at  $\Gamma - \lfloor \Gamma \rfloor$ . So set  $\mathcal{U}$  equals set  $S$  and formulation (8) is equivalent to (7).

### 3.2 Robust Hazmat Network Design Problem

Now we use the methodology introduced in Section 3.1 to formulate the robust counterpart of HNDP. In this problem, we are particularly interested in the uncertainty in risk. In the hazmat network design problem, there are several ways to consider uncertainty and construct the *budget of uncertainty*  $\Gamma$ .

One way to model the uncertainty is to consider the number of arcs that are subject to uncertainty across all shipments. The risks on all arcs for different shipments are different and can take any value in  $[r_{ijs}, r_{ijs} + d_{ijs}]$  for all  $(i, j) \in \mathcal{A}, s \in \mathcal{S}$  where  $d_{ijs}$  is the uncertainty upper bound. The uncertainty budget  $\Gamma$  value is the sum of the number of arcs exposed to uncertainty for all shipments. We refer to this problem as the *Robust Arc-Shipment-Risk HNDP* and its formulation is:

$$(R1) \quad \min_{y_{ij} \in \{0,1\}} \left( \sum_{(i,j) \in \mathcal{A}} \sum_{s \in \mathcal{S}} n_s r_{ijs} x_{ijs} + \max_{u \in \mathcal{U}} \sum_{(i,j) \in \mathcal{A}} \sum_{s \in \mathcal{S}} n_s d_{ijs} u_{ijs} x_{ijs} \right), \quad (10)$$

where

$$\mathcal{U} = \left\{ u : \sum_{(i,j) \in \mathcal{A}} \sum_{s \in \mathcal{S}} u_{ijs} \leq \Gamma, 0 \leq u_{ijs} \leq 1 \forall (i, j) \in \mathcal{A}, s \in \mathcal{S} \right\}$$

and  $x_{ijs}$  solves the lower level problem defined by (2)–(5). This formulation is flexible to model the uncertainty and computationally easier, but may not be as easy for decision makers to understand to choose an appropriate  $\Gamma$  value. For each shipment, the route consists of a certain number of arcs and this is the maximum number of arcs that can be considered to be exposed to uncertainty. Because uncertainty on the unused arcs will make no difference, you have to get an idea of how many arcs could be used for each shipment or the average number of arcs used for each shipment to decide the  $\Gamma$  value.

A more interpretable way to model  $\Gamma$  is to consider it as the number of arcs that can be subject to uncertainty. This is meaningful and can be understood by decision makers easily but is computationally harder. We refer to this problem as the *Robust Arc-Risk HNDP* and it can be formulated as follows:

$$(R2) \quad \min_{y_{ij} \in \{0,1\}} \left( \sum_{(i,j) \in \mathcal{A}} \sum_{s \in \mathcal{S}} n_s r_{ij} x_{ijs} + \max_{u \in \mathcal{U}} \sum_{(i,j) \in \mathcal{A}} \sum_{s \in \mathcal{S}} n_s d_{ij} u_{ij} x_{ijs} \right), \quad (11)$$

where

$$\mathcal{U} = \left\{ u : \sum_{(i,j) \in \mathcal{A}} u_{ij} \leq \Gamma, 0 \leq u_{ij} \leq 1 \forall (i, j) \in \mathcal{A} \right\}$$

and  $x_{ijs}$  solves the lower level problem defined by (2)–(5). We assume for each arc  $(i, j)$ , its associated risk can take any value in  $[r_{ij}, r_{ij} + d_{ij}]$  where  $d_{ij}$  is the uncertainty upper bound for each arc. Note here we have  $r_{ij}$  instead of  $r_{ijs}$  since we consider the uncertainty budget as the number of arcs in the

network. In this problem, we assume all shipments are homogeneous in terms of the accident risk. This scenario can represent transferring only a single kind of hazmat and not differentiating carriers for various OD pairs. In order to apply the model to the case of multiple kinds of hazmat, this method can be used by assigning a  $\Gamma$  value to each kind of hazmat.

We can also consider the uncertainty as the number of arcs with uncertainty on each specified shipment (route). The formulation for this problem is

$$(R3) \quad \min_{y_{ij} \in \{0,1\}} \left( \sum_{(i,j) \in \mathcal{A}} \sum_{s \in \mathcal{S}} n_s r_{ijs} x_{ijs} + \max_{u \in \mathcal{U}} \sum_{(i,j) \in \mathcal{A}} \sum_{s \in \mathcal{S}} n_s d_{ijs} u_{ijs} x_{ijs} \right), \quad (12)$$

where

$$\mathcal{U} = \left\{ u : \sum_{(i,j) \in \mathcal{A}} u_{ijs} \leq \Gamma_s, 0 \leq u_{ijs} \leq 1 \forall (i,j) \in \mathcal{A}, s \in \mathcal{S} \right\}$$

and  $x_{ijs}$  solves the lower level problem defined by (2)–(5). This formulation is only of use when the government wants to consider the uncertainty for several chosen shipments. For example, the government only wants to consider the shipments with very large flow and assign a specified uncertainty level  $\Gamma$  for each shipment. Then for each shipment, it is the same with the robust shortest path problem considered in Bertsimas and Sim (2003). If the number of shipments the government wants to consider is large, it is not practical to use this formulation since deciding  $\Gamma$  values for each shipment is excessive. Thus in this paper we consider only the Robust Arc-Shipment-Risk (R1) and Robust Arc-Risk (R2) HNDPs.

By using model R1, the government can assign different risk and uncertainty values for various shipments. Model R2 assumes shipments are homogeneous. This means that the risk and its associated risk uncertainty are the same for all shipments. The government will not be able to consider the possible difference of risk uncertainty due to different hazardous material types and carriers in one formulation. Moreover, the difference of the two models lies in the interpretation of uncertainty level  $\Gamma$ . For model R1,  $\Gamma$  is considered as the number of arcs that are subject to uncertainty over all shipments. If one arc is considered for having uncertainty on one certain shipment, it does not affect other shipments. However, for model R2,  $\Gamma$  is interpreted as the number of arcs that are subject to uncertainty in the network. If one arc is considered for having uncertainty, it is considered for all shipments using this arc. Thus model R1 provides a flexible way of modeling uncertainty but its uncertainty level  $\Gamma$  is difficult to interpret. On the other hand, model R2 provides a model with homogeneous shipments while providing a more interpretable meaning of uncertainty level.

## 4 Heuristic Framework

In this section, we propose a heuristic solution framework that is suitable for solving the hazmat network design problem with various risk measures, which is based on the method of Erkut and Gzara (2008). While Erkut and Gzara (2008) consider the expected risk measure in their hazmat network design problems, this framework can be generalized to solve the hazmat network design problem with any risk measure.

We first discuss the method of Erkut and Gzara (2008) for a general risk measure  $\gamma(x)$  for traffic pattern  $x$ . By considering risk with uncertainty as a robust risk measure, the heuristic framework is able to solve the robust HNDPs. The method begins with the full arc set  $\mathcal{A}$  and eliminates an arc in each iteration. Let  $\mathcal{A}_k$  denote the set of available arcs in the  $k$ -th iteration. We first solve the following problem:

$$\text{MinRisk}(\mathcal{A}_k) = \min_x \gamma(x),$$

subject to

$$\sum_{(i,k) \in \mathcal{A}_k} x_{iks} - \sum_{(k,i) \in \mathcal{A}_k} x_{kis} = \begin{cases} +1 & i = o(s) \\ -1 & i = d(s) \\ 0 & \text{otherwise} \end{cases} \quad \forall i \in \mathcal{N}, s \in \mathcal{S},$$

$$x_{ijs} \in \{0, 1\} \quad \forall (i, j) \in \mathcal{A}_k, s \in \mathcal{S}.$$

Let  $x^k$  denote a solution of the Min Risk problem. The obtained traffic pattern  $x^k$  is the *desired* traffic pattern in terms of the hazmat risk and the risk measure  $\gamma$ . Our objective is to introduce a network design so that the *selfish* hazmat carriers who minimize their own travel costs use the desired paths described by the best traffic pattern. However, such a pattern may often not be obtainable, so we will try to reduce the risk as much as we can.

We now introduce network design variable  $y^k$ , based on  $x^k$ . We let  $y_{ij}^k = 1$  and  $y_{ji}^k = 1$  if  $x_{ijs}^k = 1$  or  $x_{jis}^k = 1$ . Then we construct a new set of arcs  $\mathcal{A}(y^k) = \{(i, j), (j, i) : y_{ij}^k = 1, (i, j) \in \mathcal{A}, (j, i) \in \mathcal{A}\}$ . Given  $\mathcal{A}(y^k)$ , we obtain the traffic pattern that actually describes the selfish carriers route choices by solving the following problem

$$\text{MinCost}(\mathcal{A}(y^k)) = \min_z \sum_{s \in \mathcal{S}} \sum_{(i,j) \in \mathcal{A}(y^k)} n_s c_{ij} z_{ijs},$$

subject to

$$\sum_{(i,k) \in \mathcal{A}(y^k)} z_{iks} - \sum_{(k,i) \in \mathcal{A}(y^k)} z_{kis} = \begin{cases} +1 & i = o(s) \\ -1 & i = d(s) \\ 0 & \text{otherwise} \end{cases} \quad \forall i \in \mathcal{N}, s \in \mathcal{S},$$

$$z_{ijs} \in \{0, 1\} \quad \forall (i, j) \in \mathcal{A}(y^k), s \in \mathcal{S},$$

which is a collection of shortest-path problems. We identify all solutions of the Min Cost problem, and call the set of all solutions  $\mathcal{Z}_k$ . As Erkut and Gzara (2008) pointed out, it is possible that  $\mathcal{Z}_k$  is not a singleton, and therefore the risk induced by the actual carriers' traffic pattern may be as high as the following:

$$\text{MaxRisk}(\mathcal{Z}_k) = \max_{z \in \mathcal{Z}_k} \gamma(z).$$

We let  $z^k$  denote the solution, which means the traffic pattern with the highest risk among all possible carriers' route choices. If we compare the two solutions  $x^k$  and  $z^k$ , we can observe which arcs are used in the actual carriers' traffic pattern  $z^k$  but not included in the desired pattern  $x^k$ . Let us denote the set of such arcs by  $\overline{\mathcal{A}}_k = \{(i, j) : z_{ijs}^k = 1, x_{ijs}^k = 0, \forall s \in \mathcal{S}, (i, j) \in \mathcal{A}\}$ . We will choose an arc in  $\overline{\mathcal{A}}_k$  and remove it from the network, i.e. close the arc to hazmat traffic.

When there is more than one element in  $\overline{\mathcal{A}}_k$ , we can choose an arc that has the biggest contribution to increase  $\gamma$ . Such elimination is not guaranteed to reduce the total risk in the next iteration, but it is a good heuristic rule. A simple rule is to choose the link with the largest risk value and we will adopt this rule in the implementation of the heuristic algorithm.

We repeat this procedure until the lower bound of  $\gamma(z^k)$  does not improve. The following statement summarizes the algorithm:

**Step 0.** (Initialization) Set  $k \leftarrow 1$ ,  $\mathcal{A}_k \leftarrow \mathcal{A}$ ,  $\gamma_0 \leftarrow \infty$ , and  $y^0 \leftarrow \{y_{ij} = 1 : (i, j) \in \mathcal{A}\}$ . Choose a positive integer  $N$ , and an arc selection rule.

**Step 1.** (Min Risk Problem) Compute  $\text{MinRisk}(\mathcal{A}_k)$  on the network  $\mathcal{A}_k$ , and call the obtained min risk traffic pattern  $x^k$ . Let  $y_{ij}^k = 1$  and  $y_{ji}^k = 1$  if  $x_{ijs}^k = 1$  or  $x_{jis}^k = 1$  for any  $s \in \mathcal{S}$ , and

$$\mathcal{A}(y^k) = \{(i, j), (j, i) : y_{ij}^k = 1, (i, j) \in \mathcal{A}, (j, i) \in \mathcal{A}\}.$$

**Step 2.** (Min Cost Problem) Identify the set of all solutions,  $\mathcal{Z}_k = \arg \text{MinCost}(\mathcal{A}(y^k))$  on the network  $\mathcal{A}(y^k)$ .

**Step 3.** (Max Risk Problem) Find the maximum risk value by solving  $\text{MaxRisk}(\mathcal{Z}_k)$ , which is a simple comparison, and call the maximal traffic pattern  $z^k$ .

**Step 4.** (Stopping Criteria) If  $\gamma_0$  does not improve for  $N$  consecutive iterations, stop and declare  $y^0$  as the heuristic network design solution. Otherwise, do the following:

(a) If  $\gamma(z^k) > \gamma(x^k)$ , then determine the set

$$\overline{\mathcal{A}}_k = \{(i, j) : z_{ijs}^k = 1, x_{ijs}^k = 0, \forall s \in \mathcal{S}, (i, j) \in \mathcal{A}\}$$

and select an arc  $(i, j) \in \overline{\mathcal{A}}_k$  by the arc selection rule. Set  $\mathcal{A}_{k+1} \leftarrow \mathcal{A}_k \setminus \{(i, j), (j, i)\}$ ,  $\gamma_0 \leftarrow \min(\gamma_0, \gamma(z^k))$ ,  $y^0 \leftarrow y^k$  and  $k \leftarrow k + 1$ , and go to Step 1.

(b) If  $\gamma(z^k) = \gamma(x^k)$ , then declare  $y^k$  as the heuristic network design solution.

There are three problems involved: Min Risk, Min Cost, and Max Risk. While the Max Risk problem is a simple comparison, we need to discuss how to solve the Min Risk problem depending on the risk measure of choice,  $\gamma$ , and how to identify all solutions of the Min Cost problems.

The Min Cost problem may have multiple minimum cost routes with different risk values. If the heuristic framework does not consider the maximum risk route among those minimum cost routes, the resulting traffic pattern may exhibit a higher level of risk than the Min Risk problem intended; such a solution of HNDP is called *unstable* (Erkut and Gzara, 2008). The HNDP literature (Erkut and Gzara, 2008; Amaldi et al., 2011; Gzara, 2013) assumes carriers choose the worst risk route among all minimum cost routes to guarantee that the resulting risk measure is single valued and the solution is stable. This approach is referred to as the pessimistic formulation in bilevel optimization literature (Colson et al., 2005; Xu and Wang, 2014). For stable solutions, we solve a Max Risk problem in each iteration to obtain the worst risk route among the alternative optimal lower-level minimum-cost routes.

In order to solve the Max Risk problem, we need to identify all solutions of the Min Cost problem first. We propose to slightly revise the  $K$ -shortest path algorithm of Yen (1971). Instead of stopping after finding  $K$  paths, we stop when the cost of a new path is greater than  $\text{MinCost}(\mathcal{A}(y^k))$ . This is possible because Yen's algorithm finds paths in ascending order of path cost. After identifying all minimum cost paths, we can do a simple comparison of the  $\gamma$  values of those paths to find  $\text{MaxRisk}(\mathcal{Z}_k)$ .

Note that the original algorithm by Erkut and Gzara (2008) solves the Max Risk problems only when the risk value by a solution of  $\text{MinCost}(\mathcal{A}(y^k))$  is no greater than  $\gamma(z^k)$ ; that is, the original algorithm

does not find all solutions when it is unnecessary. On the other hand, in the proposed algorithm in this paper, we always identify all solutions to track  $\gamma_0$  and to terminate the algorithm if it does not improve. The proposed algorithm obtains a singleton in each step, which ensures the solution obtained will be stable as defined in the HNRP literature. In most cases, there would only be one shortest path, i.e.  $\mathcal{Z}_k$  is a singleton, for most realistic road networks. Note also that the original algorithm by Erkut and Gzara (2008) requires a Mixed-Integer-Program solver for the Max Risk problem.

Now with the heuristic framework introduced above, in order to solve the robust versions of the hazmat network design problem, we need to modify the method accordingly. First, we need to modify the selection rule. In the heuristic framework, one common selection rule is to select the link with largest risk  $r_{ij}$  value. For the robust versions, with the existence of risk uncertainty, we can choose the link with the largest  $r_{ij} + d_{ij}$  value where  $d_{ij}$  is the maximum risk associated with link  $(i, j)$ . We can consider the same approach in solving the Max Risk problem. Second, in calculating the risk for the routes chosen by the carriers, we should also consider risk uncertainty. For the cardinality uncertainty set, we can choose  $\Gamma$  number of links which contribute most to the risk value. Third, since the heuristic algorithm does not guarantee an optimal solution, we add a check to see if the obtained robust solution has lower worst risk than the nominal solution. If not, we terminate with the nominal solution as best incumbent or restart the algorithm with the nominal solution as the initial. Finally, we need to solve a different Min Risk problem with a risk measure that considers robustness. Solving the Min Risk problem with robustness consideration can be complex. Thus in the next section, we will discuss how to solve the Min Risk problem for different robust cases.

## 5 Solving the Min Risk Problem

In this section, we will formulate the Min Risk problem separately for each of the Robust Arc-Shipment-Risk and Robust Arc-Risk HNDRPs. The Min Risk problem with the risk measure considering robustness has a min-max structure and can be reformulated as a single level problem using dualization. Optimization solvers like CPLEX can solve moderate size problems. However when solving the robust network design problem, we need to solve the Min Risk problem for each iteration so we also provide methods to solve the Min Risk problem more efficiently.

## 5.1 Min Risk Problem for Robust Arc-Shipment-Risk HNDP (R1)

First we consider the Min Risk problem for the Robust Arc-Shipment-Risk HNDP which considers the budget of uncertainty  $\Gamma$  on the number of arcs that are subject to uncertainty for all shipments. The formulation is

$$\min_x \max_u \sum_{(i,j) \in \mathcal{A}_k} \sum_{s \in \mathcal{S}} n_s (r_{ijs} + d_{ijs} u_{ijs}) x_{ijs}, \quad (13)$$

subject to

$$\sum_{(i,j) \in \mathcal{A}_k} \sum_{s \in \mathcal{S}} u_{ijs} \leq \Gamma, \quad : \theta \quad (14)$$

$$0 \leq u_{ijs} \leq 1 \quad \forall (i,j) \in \mathcal{A}_k, s \in \mathcal{S}, \quad : \rho_{ijs} \quad (15)$$

$$\sum_{i,k \in \mathcal{A}_k} x_{iks} - \sum_{k,i \in \mathcal{A}_k} x_{kis} = \begin{cases} +1 & i = o(s) \\ -1 & i = d(s) \\ 0 & \text{otherwise} \end{cases} \quad \forall i \in \mathcal{N}, s \in \mathcal{S}, \quad (16)$$

$$x_{ijs} \in \{0, 1\} \quad \forall (i,j) \in \mathcal{A}_k, s \in \mathcal{S}. \quad (17)$$

This formulation can be transformed into a single level problem by using the dual of the inner problem. However, since the uncertainty number is dependent on both arcs and shipments, we are able to apply the method for solving robust combinatorial problems proposed by Bertsimas and Sim (2003). First we consider only the inner maximization problem and dualize it. It then becomes

$$\min_{\theta, \rho} \Gamma \theta + \sum_{(i,j) \in \mathcal{A}_k} \sum_{s \in \mathcal{S}} \rho_{ijs}, \quad (18)$$

subject to

$$\theta + \rho_{ijs} \geq n_s d_{ijs} x_{ijs} \quad \forall (i,j) \in \mathcal{A}_k, s \in \mathcal{S},$$

$$\rho_{ijs} \geq 0 \quad \forall (i,j) \in \mathcal{A}_k, s \in \mathcal{S},$$

$$\theta \geq 0.$$

Because  $x_{ijs} \in \{0, 1\}$ , we can transform the constraints of model (18) and obtain

$$\rho_{ijs} = \max(n_s d_{ijs} x_{ijs} - \theta, 0) = \max(n_s d_{ijs} - \theta, 0) x_{ijs}.$$

Let  $\mathcal{X}$  denote the set of constraints (16) and (17), then the original problem becomes

$$Z^* = \min_{x \in \mathcal{X}, \theta \geq 0} \Gamma\theta + \sum_{(i,j) \in \mathcal{A}_k} \sum_{s \in \mathcal{S}} n_s r_{ijs} x_{ijs} + \sum_{(i,j) \in \mathcal{A}_k} \sum_{s \in \mathcal{S}} \max(n_s d_{ijs} - \theta, 0) x_{ijs}. \quad (19)$$

For the notational simplicity, we introduce a tuple  $h = (i, j, s)$  and the set  $\mathcal{H}_k = \{(i, j, s) : (i, j) \in \mathcal{A}_k, s \in \mathcal{S}\}$ . We let  $r_h = n_s r_{ijs}$ ,  $d_h = n_s d_{ijs}$ ,  $x_h = x_{ijs}$ . Then problem (19) becomes

$$Z^* = \min_{x \in \mathcal{X}, \theta \geq 0} \Gamma\theta + \sum_{h \in \mathcal{H}_k} r_h x_h + \sum_{h \in \mathcal{H}_k} \max(d_h - \theta) x_h. \quad (20)$$

In order to solve the problem, we arrange  $d_h$  in descending order and let  $d_{|\mathcal{H}_k|+1} = 0$ . Now, according to Bertsimas and Sim (2003), problem (20) can be solved by solving  $|\mathcal{H}_k| + 1$  nominal problems:

$$Z^* = \min_{l=1, \dots, |\mathcal{H}_k|+1} G^l, \quad (21)$$

where for  $l = 1, \dots, |\mathcal{H}_k| + 1$ ,

$$G^l = \Gamma d_l + \min_{x \in \mathcal{X}} \left( \sum_{h \in \mathcal{H}_k} r_h x_h + \sum_{h=1}^l (d_h - d_l) x_h \right)$$

and  $x_{|\mathcal{H}_k|+1} = 0$ . In the Robust Arc-Shipment-Risk HNBP, the number of nominal problems needs to be solved is  $|\mathcal{H}_k| + 1 = |\mathcal{S}| \cdot |\mathcal{A}_k| + 1$  in each iteration  $k$ . However, in the case of large networks or a large number of shipments, even if it is polynomially solvable, the time of solving the problem could be excessive. Although the objective function cannot be shown to be strictly quasiconvex in general, all of our analysis showed it to be quasiconvex with possible exceptions in the tails far from the optimal solution. Thus it is practical to use a line search algorithm and again we always find the optimum. Here we use a modified golden section search algorithm (Bazaraa et al., 2013) by rounding the interval to integers. A summary of the method is shown below.

**Step 0.** (Initialization) Choose an allowable final length of uncertainty  $\Delta > 0$ . Let  $[\alpha_1, \beta_1]$  be the initial interval where  $\alpha_1 = 1$ ,  $\beta_1 = |\mathcal{S}| \cdot |\mathcal{A}_k| + 1$ . Initialize  $\eta_1 = \lfloor \alpha_1 + (1 - \delta)(\beta_1 - \alpha_1) \rfloor$  and  $\xi_1 = \lfloor \alpha_1 + \delta(\beta_1 - \alpha_1) \rfloor$ , where  $\delta = 0.618$ . Evaluate  $G^{\eta_1}$  and  $G^{\xi_1}$ , and let the iteration number  $t = 1$ .

**Step 1.** If  $\beta_t - \alpha_t < \Delta$ , stop; the optimal solution lies in the interval  $[\alpha_t, \beta_t]$ . Otherwise, if  $G^{\eta_t} > G^{\xi_t}$ , go to step 2; and if  $G^{\eta_t} \leq G^{\xi_t}$ , go to step 3.

**Step 2.** Let  $\alpha_{t+1} = \eta_t$ ,  $\beta_{t+1} = \beta_t$ . Furthermore, let  $\eta_{t+1} = \xi_t$ , and let  $\xi_{t+1} = \lfloor \alpha_{t+1} + \delta(\beta_{t+1} - \alpha_{t+1}) \rfloor$ . Evaluate  $G^{\xi_{t+1}}$  and go to step 4.

**Step 3.** Let  $\alpha_{t+1} = \alpha_t$ ,  $\beta_{t+1} = \xi_t$ . Furthermore, let  $\xi_{t+1} = \eta_t$ , and let  $\eta_{t+1} = \lfloor \alpha_{t+1} + (1 - \delta)(\beta_{t+1} - \alpha_{t+1}) \rfloor$ . Evaluate  $G^{\eta_{t+1}}$  and go to step 4.

**Step 4.** Replace  $t$  by  $t + 1$  and go to step 1.

## 5.2 Min Risk Problem for Robust Arc-Risk HNDP (R2)

Now we consider the Min Risk problem for the Robust Arc-Risk HNDP where the budget of uncertainty is considered as how many arcs are subject to uncertainty. It can be formulated as

$$\min_x \max_u \sum_{(i,j) \in \mathcal{A}_k} \sum_{s \in \mathcal{S}} n_s (r_{ij} + d_{ij} u_{ij}) x_{ijs}, \quad (22)$$

subject to

$$\begin{aligned} \sum_{(i,j) \in \mathcal{A}} u_{ij} &\leq \Gamma, \\ 0 &\leq u_{ij} \leq 1 \quad \forall (i,j) \in \mathcal{A}_k, \\ \sum_{i,k \in \mathcal{A}_k} x_{iks} - \sum_{k,i \in \mathcal{A}_k} x_{kis} &= \begin{cases} +1 & i = o(s) \\ -1 & i = d(s) \\ 0 & \text{otherwise} \end{cases} \quad \forall i \in \mathcal{N}, s \in \mathcal{S}, \\ x_{ijs} &\in \{0, 1\} \quad \forall (i,j) \in \mathcal{A}_k, s \in \mathcal{S}. \end{aligned}$$

This formulation can be transformed into one single model by dualizing the inner maximization problem. Similar to the Robust Arc-Shipments-Risk HNDP (R1) case, we first consider only the inner maximization problem:

$$\max_u \sum_{(i,j) \in \mathcal{A}} \sum_{s \in \mathcal{S}} n_s d_{ij} u_{ij} x_{ijs}, \quad (23)$$

subject to

$$\begin{aligned} \sum_{(i,j) \in \mathcal{A}_k} u_{ij} &\leq \Gamma, & : \theta \\ 0 &\leq u_{ij} \leq 1 \quad \forall (i,j) \in \mathcal{A}_k. & : \rho_{ij} \end{aligned}$$

By dualizing the inner maximization problem to an equivalent minimization problem and merging it into problem (22), we have the single level formulation as:

$$\min_{x, \theta, \rho} \Gamma\theta + \sum_{(i,j) \in \mathcal{A}_k} \sum_{s \in \mathcal{S}} n_s r_{ij} x_{ijs} + \sum_{(i,j) \in \mathcal{A}_k} \rho_{ij}, \quad (24)$$

subject to

$$\theta + \rho_{ij} \geq \sum_{s \in \mathcal{S}} n_s d_{ij} x_{ijs} \quad \forall (i, j) \in \mathcal{A}_k, \quad (25)$$

$$\sum_{i, k \in \mathcal{A}_k} x_{iks} - \sum_{k, i \in \mathcal{A}_k} x_{kis} = \begin{cases} +1 & i = o(s) \\ -1 & i = d(s) \\ 0 & \text{otherwise} \end{cases} \quad \forall i \in \mathcal{N}, s \in \mathcal{S}, \quad (26)$$

$$x_{ijs} \in \{0, 1\} \quad \forall (i, j) \in \mathcal{A}_k, s \in \mathcal{S}, \quad (27)$$

$$\rho_{ij} \geq 0 \quad \forall (i, j) \in \mathcal{A}_k, \quad (28)$$

$$\theta \geq 0. \quad (29)$$

However, different from Robust Arc-Shipment-Risk HNDP (R1), if we perform a similar transformation of constraint (25), it becomes  $\rho_{ij} = \max\left(\sum_{s \in \mathcal{S}} n_s d_{ij} x_{ijs} - \theta\right)$ , which cannot be decomposed by  $x_{ijs}$  because of the summation. Thus we cannot apply the method in Bertsimas and Sim (2003). The Min Risk problem for Robust Arc-Risk HNDP (R2) is a mixed integer linear programming problem and has the integer multi-commodity network flow problem form. Although it may be solved by a solver such as CPLEX, it is difficult in general to solve large-scale networks. Also we need to solve this problem for each iteration, so we need to consider a more efficient algorithm. In the remaining, we develop a Lagrangian relaxation algorithm with projection.

### 5.2.1 Lagrangian Relaxation for the Min Risk Problem of Robust Arc-Risk HNDP

In order to solve large-scale problems, we propose a Lagrangian relaxation method. We relax constraint (25) which is  $\theta + \rho_{ij} \geq \sum_{s \in \mathcal{S}} n_s d_{ij} x_{ijs}$ ,  $\forall (i, j) \in \mathcal{A}_k$ . By introducing Lagrangian multiplier vector

$\mu = \{\mu_{ij} : (i, j) \in \mathcal{A}_k\}$  for constraint (25), the objective (24) becomes

$$\begin{aligned} & \Gamma\theta + \sum_{(i,j) \in \mathcal{A}_k} \sum_{s \in \mathcal{S}} n_s r_{ij} x_{ijs} + \sum_{(i,j) \in \mathcal{A}_k} \rho_{ij} + \sum_{(i,j) \in \mathcal{A}_k} \mu_{ij} \left( \sum_{s \in \mathcal{S}} n_s d_{ij} x_{ijs} - \theta - \rho_{ij} \right) \\ & = \left( \Gamma - \sum_{(i,j) \in \mathcal{A}_k} \mu_{ij} \right) \theta + \sum_{(i,j) \in \mathcal{A}_k} \sum_{s \in \mathcal{S}} n_s (r_{ij} + \mu_{ij} d_{ij}) x_{ijs} + \sum_{(i,j) \in \mathcal{A}_k} (1 - \mu_{ij}) \rho_{ij}. \end{aligned} \quad (30)$$

Since our problem is to minimize the risk, if  $\Gamma - \sum_{(i,j) \in \mathcal{A}_k} \mu_{ij} < 0$ ,  $\theta$  will go to  $+\infty$ ; if  $1 - \mu_{ij} < 0$ ,  $\rho_{ij}$  will go to  $+\infty$ . This makes the objective goes to  $-\infty$ . In order to get bounded results, we must have  $\Gamma - \sum_{(i,j) \in \mathcal{A}_k} \mu_{ij} \geq 0$  and  $1 - \mu_{ij} \geq 0$  for all  $(i, j) \in \mathcal{A}_k$ . These restrictions will be realized in the projection step of our subgradient search algorithm which we will discuss in Section 5.2.3. If  $\Gamma - \sum_{(i,j) \in \mathcal{A}_k} \mu_{ij} \geq 0$  and  $1 - \mu_{ij} \geq 0$  for all  $(i, j) \in \mathcal{A}_k$ , we will have  $\theta = 0$  and  $\rho_{ij} = 0$  for all  $(i, j) \in \mathcal{A}_k$  since it is a minimization problem. Then the objective in equation (30) will become  $\sum_{(i,j) \in \mathcal{A}_k} \sum_{s \in \mathcal{S}} n_s (r_{ij} + \mu_{ij} d_{ij}) x_{ijs}$  as  $\theta = 0, \rho_{ij} = 0$  for all  $(i, j) \in \mathcal{A}_k$ . Thus we have the equivalent Lagrangian relaxation problem  $P_\mu$  for any fixed  $\mu$  formulated as:

$$(P_\mu) \quad \min_x \sum_{(i,j) \in \mathcal{A}_k} \sum_{s \in \mathcal{S}} n_s (r_{ij} + \mu_{ij} d_{ij}) x_{ijs}, \quad (31)$$

subject to

$$\sum_{i,k \in \mathcal{A}_k} x_{iks} - \sum_{k,i \in \mathcal{A}_k} x_{kis} = \begin{cases} +1 & i = o(s) \\ -1 & i = d(s) \\ 0 & \text{otherwise} \end{cases} \quad \forall i \in \mathcal{N}, s \in \mathcal{S},$$

$$x_{ijs} \in \{0, 1\} \quad \forall (i, j) \in \mathcal{A}_k, s \in \mathcal{S},$$

where the Lagrange multipliers  $\mu_{ij}$  must satisfy  $0 \leq \mu_{ij} \leq 1$  and  $\Gamma - \sum_{(i,j) \in \mathcal{A}_k} \mu_{ij} \geq 0$ .

The Lagrangian Relaxation problem  $P_\mu$  is easy to solve since it can be decomposed as a series of shortest path problems. Let  $v(P_\mu)$  denote the value of an optimal solution to problem  $P_\mu$ , then the Lagrangian dual problem  $(D_L)$  is

$$(D_L) \quad \max_\mu v(P_\mu), \quad (32)$$

subject to

$$\begin{aligned}\Gamma - \sum_{(i,j) \in \mathcal{A}_k} \mu_{ij} &\geq 0, \\ \mu_{ij} &\geq 0 \quad \forall (i,j) \in \mathcal{A}_k, \\ \mu_{ij} &\leq 1 \quad \forall (i,j) \in \mathcal{A}_k.\end{aligned}$$

The dual problem can be solved by a customized subgradient method by modifying the set of multipliers. The Lagrangian dual problem provides a lower bound to the original problem (24). Below, we will show how to obtain an upper bound.

### 5.2.2 Obtaining an Upper Bound

In the Lagrangian relaxation, we can also obtain an upper bound during the subgradient process for fixed values of  $x_{ijs}$  by solving the following linear optimization problem,

$$\min_{\theta \geq 0, \rho_{ij} \geq 0} \Gamma\theta + \sum_{(i,j) \in \mathcal{A}_k} \sum_{s \in \mathcal{S}} n_s r_{ijs} x_{ijs} + \sum_{(i,j) \in \mathcal{A}_k} \rho_{ij},$$

subject to

$$\theta + \rho_{ij} \geq \sum_{s \in \mathcal{S}} n_s d_{ijs} x_{ijs} \quad \forall (i,j) \in \mathcal{A}_k.$$

Since  $x_{ijs}$  are fixed, the objective for the linear optimization problem becomes  $\min_{\theta, \rho} \Gamma\theta + \sum_{(i,j) \in \mathcal{A}_k} \rho_{ij}$ . Then the linear problem  $U$  that obtains the upper bound becomes

$$(U) \quad \min_{\theta \geq 0, \rho_{ij} \geq 0} \Gamma\theta + \sum_{(i,j) \in \mathcal{A}_k} \rho_{ij}, \quad (33)$$

subject to

$$\theta + \rho_{ij} \geq \sum_{s \in \mathcal{S}} n_s d_{ijs} x_{ijs} \quad \forall (i,j) \in \mathcal{A}_k.$$

The linear problem  $U$  can be reformulated as

$$\min_{\theta} \varphi = \Gamma\theta + \sum_{(i,j) \in \mathcal{A}_k} \max \left\{ 0, \sum_{s \in \mathcal{S}} n_s d_{ijs} x_{ijs} - \theta \right\} \text{ for } \theta \geq 0.$$

Now suppose we rank  $\sum_{s \in \mathcal{S}} n_s d_{ij} x_{ijs}$  in ascending order as  $\{s_1, s_2, \dots, s_{|\mathcal{A}_k|}\}$  and let  $s_0 = 0$ . Then for  $s_l < \theta < s_{l+1}$ ,  $0 \leq l \leq |\mathcal{A}_k|$ , the objective  $\varphi$  becomes

$$\varphi = \Gamma\theta + s_{l+1} - \theta + s_{l+2} - \theta + \dots + s_{|\mathcal{A}_k|} - \theta. \quad (34)$$

By taking the derivative of  $\varphi$  in (34), we have

$$\frac{\partial \varphi}{\partial \theta} = \Gamma - (|\mathcal{A}_k| - l)$$

By letting  $\frac{\partial \varphi}{\partial \theta} = 0$  we have  $l^* = |\mathcal{A}_k| - \Gamma$ . If  $l < l^*$ , then  $\frac{\partial \varphi}{\partial \theta} < 0$ , which means the objective  $\varphi$  is decreasing in interval  $(s_l, s_{l+1})$ . Similarly, if  $l > l^*$ , then  $\frac{\partial \varphi}{\partial \theta} > 0$ , which means the objective  $\varphi$  is increasing in interval  $(s_l, s_{l+1})$ . Besides, the objective is continuous. So  $l^*$  is optimal. Then  $\theta^*$  can be any value in  $[s_{l^*}, s_{l^*+1}]$ . We let  $\theta^* = s_{l^*}$  here. Then  $\rho_{ij}^* = \max \left\{ 0, \sum_{s \in \mathcal{S}} n_s d_{ij} x_{ijs} - \theta^* \right\}$ .

### 5.2.3 Lagrangian Heuristic Algorithm Framework

Now we introduce the proposed algorithm. The idea is that we try to find an upper bound and a lower bound so that the gap between them falls into a certain range. The detailed process for the algorithm is described in the following:

**Step 0:** Arbitrarily choose  $\mu_{ij}^t$  for all  $(i, j) \in \mathcal{A}_k$  such that  $0 \leq \mu_{ij}^t \leq 1$  and  $\Gamma - \sum_{(i,j) \in \mathcal{A}_k} \mu_{ij}^t \geq 0$ . Set  $t = 1$ , let an initial incumbent for problem  $D_L$  be  $v^0 = -\infty$  (lower bound) and let an initial value for problem  $U$  be  $w^0 = +\infty$  (upper bound).

**Step 1:** Solve the Lagrangian relaxation problem  $P_\mu^t$  and obtain an optimal solution  $x_{ijs}^t$ . Put the obtained optimal solution  $x_{ijs}^t$  into the linear optimization problem  $U^t$  and solve it to obtain optimal solution  $\theta^t$  and  $\rho_{ij}^t$ .

**Step 2:** Let  $v(\cdot)$  denote the value of an optimal solution to problem  $(\cdot)$ . If  $v(P_\mu^t) > v^{t-1}$  then set  $v^t = v(P_\mu^t)$  to obtain a new incumbent for  $v(D_L)$  (improved lower bound). Otherwise set  $v^t = v^{t-1}$ . If  $v(U^t) < w^{t-1}$ , then set  $w^t = v(U^t)$  to obtain a better upper bound. Otherwise set  $w^t = w^{t-1}$ . If  $(w^t - v^t)/v^t \leq \tau$  (the upper bound is close enough to the lower bound), or a certain time limit is reached, stop.

**Step 3:** Find a new vector of Lagrange multipliers  $\mu_{ij}^{t+1}$  by setting

$$\mu_{ij}^{t+1} = \mu_{ij}^t + \lambda_t \frac{\sum_{s \in \mathcal{S}} n_s d_{ij} x_{ijs}^t - \theta - \rho_{ij}}{\left\| \sum_{s \in \mathcal{S}} n_s d_{ij} x_{ijs}^t - \theta - \rho_{ij} \right\|} = \mu_{ij}^t + \lambda_t \frac{\sum_{s \in \mathcal{S}} n_s d_{ij} x_{ijs}^t}{\left\| \sum_{s \in \mathcal{S}} n_s d_{ij} x_{ijs}^t \right\|}, \quad (35)$$

since  $\theta = 0, \rho_{ij} = 0$  for all  $(i, j) \in \mathcal{A}_k$  must hold in the Lagrangian relaxation problem. Note that  $\lambda_t$  is a scalar that satisfies  $\lambda_t \geq 0$  for all  $t$ ,  $\lim_{t \rightarrow \infty} \lambda_t = 0$ ,  $\sum_{t=1}^{\infty} \lambda_t = \infty$ .

**Step 4:** Project  $\mu^{t+1}$  onto  $\left\{ \mu^{t+1} : 0 \leq \mu_{ij}^{t+1} \leq 1 \forall (i, j) \in \mathcal{A}_k, \Gamma - \sum_{(i,j) \in \mathcal{A}_k} \mu_{ij}^{t+1} \geq 0 \right\}$ . Go to **Step 1**.

The above algorithm falls into a standard subgradient search method except for the last step of projection. The standard projection is to project  $\mu$  onto  $\{\mu \geq 0\}$ ; however, here we will project onto  $\{\mu : 0 \leq \mu_{ij} \leq 1 \forall (i, j) \in \mathcal{A}_k, \Gamma - \sum_{(i,j) \in \mathcal{A}_k} \mu_{ij} \geq 0\}$ . We denote the current values as  $\mu_{ij}$  and the projected values as  $\bar{\mu}_{ij}$ . The projection procedure is as follows:

**Step 0:** Project  $\mu$  onto  $\{\mu : 0 \leq \mu_{ij} \leq 1 \forall (i, j) \in \mathcal{A}_k\}$ , *i.e.*  $\bar{\mu}_{ij} = \max\{0, \min\{1, \mu_{ij}\}\}$ . Check whether  $\sum_{(i,j) \in \mathcal{A}_k} \bar{\mu}_{ij} \leq \Gamma$  or not. If satisfied, we obtain  $\bar{\mu}_{ij}$ ; otherwise go to **Step 1**.

**Step 1:** Rank  $\mu_{ij}$  and  $\mu_{ij} - 1$  in non-increasing order as  $b_t$  where  $1 \leq t \leq 2|\mathcal{A}_k|$ . Starting from  $t = 1$ , obtain  $\lambda$  that satisfies  $\lambda \geq 0$ ,  $\lambda \in [b_{t+1}, b_t]$  and

$$\sum_{\{(i,j) \in \mathcal{A}_k : \mu_{ij} - 1 \leq b_t < \mu_{ij}\}} (\mu_{ij} - \lambda) + \sum_{\{(i,j) \in \mathcal{A}_k : b_t < \mu_{ij} - 1\}} (1) = \Gamma.$$

Then assign  $\bar{\mu}_{ij} = \max\{0, \min\{1, \mu_{ij} - \lambda\}\}$ .

We prove that the above procedure indeed achieves the projection onto  $\{\mu : 0 \leq \mu_{ij} \leq 1 \forall (i, j) \in \mathcal{A}_k, \Gamma - \sum_{(i,j) \in \mathcal{A}_k} \mu_{ij} \geq 0\}$  in Appendix A using a method similar to Held et al. (1974) and Peinhardt (2003). The basic idea is to consider its KKT conditions and derive an optimal solution. The Lagrangian relaxation method is developed in the robust HNDP context. However, we believe that this method can be applied to many other general robust optimization problems when the problem cannot be solved by enumerating a finite number of uncertain parameters.

## 6 Experiments

In this section, we provide some computational results for the methods we introduced. We perform the numerical analysis on two datasets. The first set of data is from the city of Ravenna, Italy (Bonvicini

and Spadoni, 2008; Erkut and Alp, 2007). The data consists of 105 nodes and 134 arcs. Risks are carefully measured as functions of both the accident frequency and its damage effects. 12 nodes of the entire network can be origin or destination nodes. 35 origin-destination (OD) pairs are formed to transport four kinds of hazmat, namely, chlorine, LPG, gasoline, and methanol. Demand is the number of shipments between each OD pair. Also various number of OD pairs can be generated on this network data. We also test our methods on a larger network. The dataset is constructed based on the Barcelona Network (Bar-Gera, 2013). The transportation cost is generated using the free flow time. For this data set, we only use the network structure. There are 1020 nodes and 2522 links. We simulate the risks that are uniformly distributed on  $[0, 50]$  and demands that are uniformly distributed on  $[100, 500]$ . Experiments are performed using C++ and CPLEX 12.2 on a generic computer with 6GB memory and Intel Core i5 processor. Due to memory issues, a higher performance computer with a Xeon processor and 32GB memory is used for the Barcelona network to solve the Robust Arc-Risk HNDP.

We will first briefly introduce some results on the heuristic framework procedures. Then we will provide results on the two versions of robust HNDPs: Robust Arc-Shipment-Risk HNDP (R1) and Robust Arc-Risk HNDP (R2).

## 6.1 Heuristic Framework

Here we briefly show some results on the heuristic framework. We illustrate the results on a representative case of Robust Arc-Shipment-Risk HNDP with  $\Gamma = 50$  using the Ravenna data.

First we show the necessity of using network design by considering three cases: (a) the network obtained by minimizing risk designed by government with full regulation, (b) the network obtained by minimizing cost by carriers without any regulation, and (c) the network chosen by carriers if given the network of case (a). The networks are shown in Figure 1. We can see that without any regulation, the carriers can choose the minimum cost routes and this results in higher risk, approximately 20.2% higher in this case. If government can force the carriers to choose any route, it will result in the minimum risk. However this is not applicable since government can only give carriers a network over which to choose routes. But the carriers' choices can deviate from the routes that the government wants carriers to choose. For example, in Figure 1(c), the carriers' chosen network based on the minimum risk network is different and results in higher risk. So a network design model is needed in order to reduce the risk.

To illustrate how the heuristic works, we have shown four steps of the heuristic procedure for illustration in Figure 2. In each iteration,  $R_g$  and  $C_g$  denote the risk and cost for the routes chosen by the government and  $R_c$  and  $C_c$  denote the risk and cost for routes chosen by the carriers. By closing

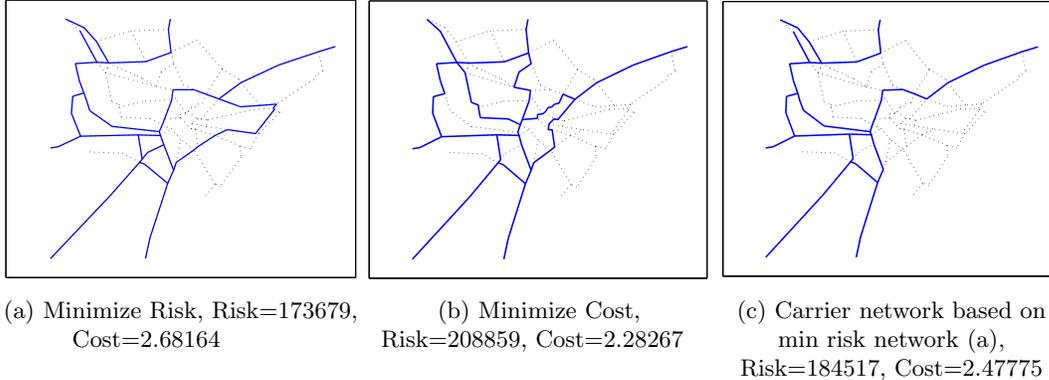


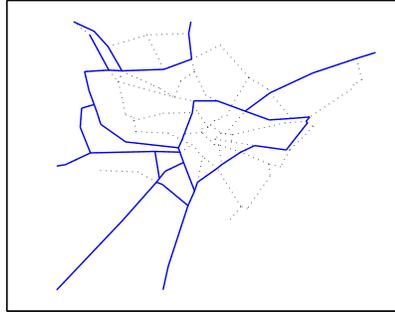
Figure 1: Network for different cases.

a certain link at each step, the risk for the carriers can decrease or increase but we can regard the minimum risk case as the solution. For this case, the carriers' risk goes as follows:  $184517 \rightarrow 184544 \rightarrow 177226 \rightarrow 174476$ . The best solution is obtained at iteration 4. The procedure stops when  $R_g$  and  $R_c$  are close enough or the risk for carriers cannot be reduced for a certain number of consecutive steps.

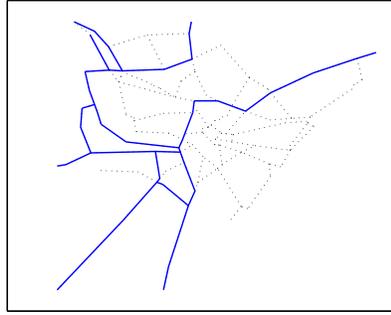
## 6.2 Robust Arc-Shipment-Risk HNDP (R1)

For the Robust Arc-Shipment-Risk HNDP version, we first show the efficiency of the Min Risk problem using the method described in Bertsimas and Sim (2003) (referred to as the Bertsimas-Sim method) and the Golden Section Search method. The Bertsimas-Sim method is polynomially solvable; however, the number of nominal problems needed to be solved can be  $|\mathcal{S}| \cdot |\mathcal{A}_k| + 1$ , which makes solving larger datasets inefficient. Although the objective function cannot be shown convex in general, it is practical to use a search algorithm like the Golden Section Search method in the context of HNDP. A comparison of time and gap for the Bertsimas-Sim method and the Golden Section Search method for the Min Risk problem can be found in Table 1. We observe that both methods find the same optimal solution, while the Golden Section Search method saves considerable computational time. Additionally, we test different  $\Gamma$  values on the Ravenna dataset and the objectives are still the same. Thus, the Golden Section Search method can be a very good practical approach.

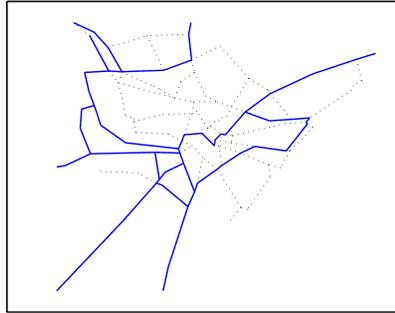
After examining the efficiency of the algorithms, we now show the effectiveness of the robust network design model solutions. The solutions are obtained by the heuristic framework using our Golden Section Search method to solve the Min Risk problem. We assume that 40% of all links—which are randomly chosen in experiments—are subject to uncertainty and the possible risk uncertainty upper bound  $d_{ijs}$  is randomly generated using a uniform distribution between 0 and  $2r_{ijs}$  for all  $(i, j) \in \mathcal{A}$  and  $s \in \mathcal{S}$ . We test different  $\Gamma$  values and a visualization of the solutions with  $\Gamma = 0$  (nominal), 25, 50, 75 for the



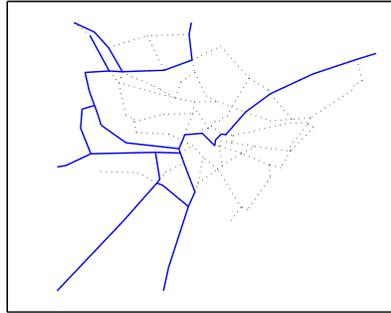
(a) Government network in iteration 1,  $R_g = 173679$ ,  $C_g = 2.68164$



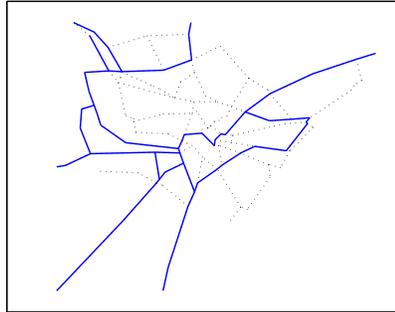
(b) Carrier network in iteration 1,  $R_c = 184517$ ,  $C_c = 2.47775$



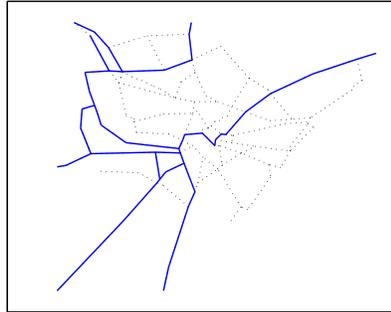
(c) Government network in iteration 2,  $R_g = 173768$ ,  $C_g = 2.68155$



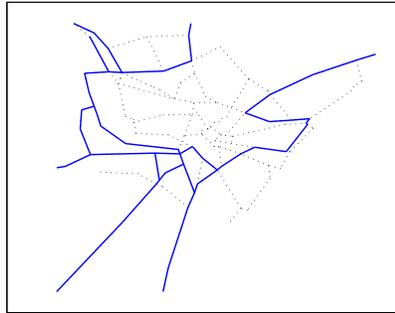
(d) Carrier network in iteration 2,  $R_c = 184544$ ,  $C_c = 2.47665$



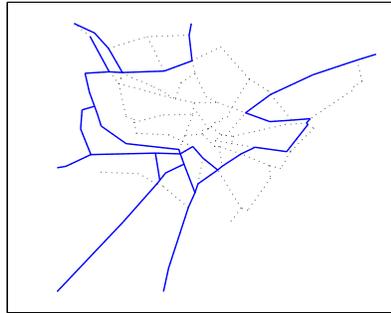
(e) Government network in iteration 3,  $R_g = 174423$ ,  $C_g = 2.69179$



(f) Carrier network in iteration 3,  $R_c = 177226$ ,  $C_c = 2.66276$



(g) Government network in iteration 4,  $R_g = 174476$ ,  $C_g = 2.68697$



(h) Carrier network in iteration 4,  $R_c = 174476$ ,  $C_c = 2.68697$

Figure 2: Government and carrier networks at each iteration of the heuristic procedure.

Table 1: Time (sec) and gap analysis for the Min Risk problem of Robust Arc-Shipment-Risk HNDP (R1)

Data	OD Number	Bersimas-Sim		Golden Section Search	
		Time	Objective	Time	Objective
Ravenna	20	246	104050	2.3	104050
	40	1078	173335	4.5	173335
	60	2135	284209	6.5	284209

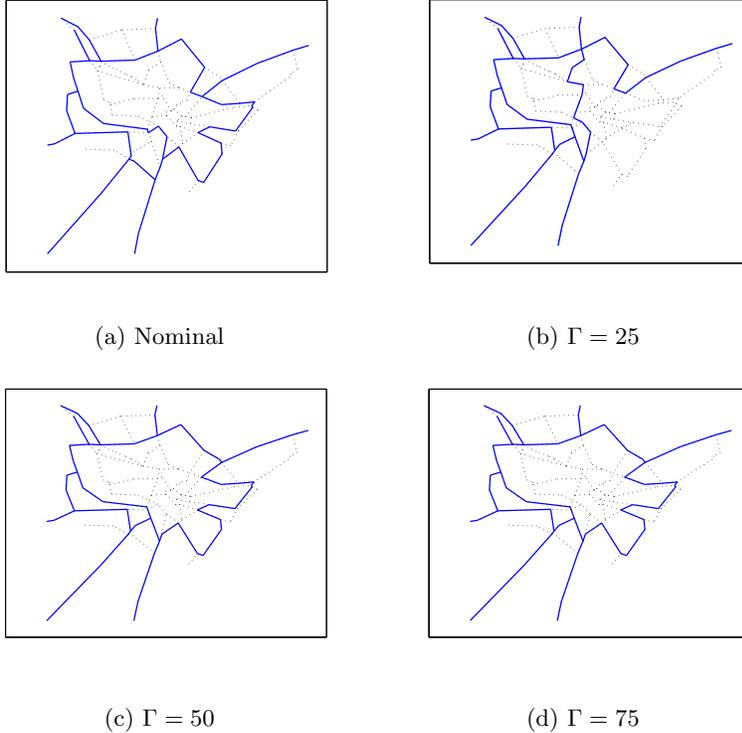


Figure 3: Network solutions for different  $\Gamma$  values for Robust Arc-Shipment-Risk HNDP (R1)

Ravenna network can be found in Figure 3. A larger  $\Gamma$  value denotes that we are considering more uncertain links and are more conservative. Note in Figure 3, solutions with  $\Gamma = 50$  and  $\Gamma = 75$  are the same.

In order to test the effectiveness of the robust solutions, we use a simulation to obtain the risk distributions of the nominal and robust solutions. We generate 10,000 samples with risk values  $r_{ijs} + u_{ijs}d_{ijs}$  where  $u_{ijs}$  is uniformly distributed in the interval  $[0, 1]$ . For the Ravenna network, the budget of uncertainty for generating the samples is 1,000 while for the Barcelona network, all links of the shipments are associated with risk uncertainty values. Due to the limited variation of solutions, we only record robust solutions with  $\Gamma$  up to 100 for the Ravenna network. Based on the samples, we then obtain the risk profile for each network design solution with different  $\Gamma$  values. The statistics of the

mean for overall risk (MOR), standard deviation (SD), mean worst case risk (MWCR), worst case risk (WCR) are recorded in Table 2 to measure solution performance. MWCR is the mean for the highest 3% risk values and WCR is the risk corresponding to the worst (highest) risk scenario. Additionally we calculate the corresponding quantitative gap of the robust solutions compared to the nominal solution. The value is negative if there is risk reduction and vice versa.

From the MOR Gap column in Table 2, the robust solutions do not necessarily have lower MOR values since the objective of the robust models is minimizing the worst case risk. For example, the cases when  $\Gamma = 25$  for the Ravenna network and  $\Gamma = 200, 250, 300$  for the Barcelona network have higher MOR values. However, for the WCR and MWCR, the robust solutions mostly reduce risk compared to the nominal solution. The highest risk reduction is 3.26% for the WCR and 3.42% for MWCR. Note here for the case of  $\Gamma = 250$  for the Barcelona network, the WCR has higher risk value than that of the nominal solution. The reason could be that the generated samples failed to capture the worst case scenario. Thus we make comparisons of robust solutions to the nominal one regarding the theoretical worst case risk (TWCR) and the results are shown in Table 3. The robust TWCR is the objective value of the robust Arc-Shipment-Risk HNRP model with budget of uncertainty  $\Gamma$ . For each robust solution, we calculate the worst risk possible for the nominal solution accordingly to obtain the nominal TWCR. This is achieved by adding the possible highest  $\Gamma$  number of uncertainty values ( $n_s d_{ijs} x_{ijs}^0$ ) to the nominal HNRP objective risk value ( $\sum_{s \in \mathcal{S}} \sum_{(i,j) \in \mathcal{A}} n_s r_{ijs} x_{ijs}^0$ ) where  $x_{ijs}^0$  is the nominal solution. We also calculate the gap and record the time of obtaining each robust solution. We can see for all the theoretical worst cases, the robust solutions have smaller risk values. The highest risk reduction can be 4.71% for the Ravenna network and 6.33% for the Barcelona network. Additionally, from Table 2, we observe that compared to the nominal solution, the robust solutions have significant SD reductions, which can be as high as 11.79% and 31.87% among the test cases for the Ravenna and Barcelona network respectively. The SD represents the variability for a certain solution and denotes the risk preference of the government. For robust solutions considering uncertainty, they might sacrifice the chances of having even lower risk values. However, they gain a lower probability of having higher risk values due to the smaller SD. Based on the analysis of the generated statistics, the government can decide an appropriate  $\Gamma$  to mitigate risk uncertainty.

For the Ravenna network, we also record the overall and tail distributions (5% highest risks) for  $\Gamma = 0$  (nominal),  $\Gamma = 25$ ,  $\Gamma = 50$  and  $\Gamma = 75$  solutions in Figure 4. The results are consistent with Table 2 and provide a visual distribution of the risk under uncertainty. The tail distribution can clearly show the difference of the worst case risk. Note here the cases of  $\Gamma = 50$  and  $\Gamma = 75$  have the same solution.

Table 2: Statistics of nominal and robust solutions for Robust Arc-Shipments-Risk HNDR (R1)

Dataset	$\Gamma$	MOR	SD	MWCR	WCR	MOR Gap	SD Gap	MWCR Gap	WCR Gap
Ravenna	0	$8.51380 \times 10^4$	1881.5	$9.06982 \times 10^4$	$9.30860 \times 10^4$	NA	NA	NA	NA
	25	$8.58621 \times 10^4$	1683.1	$9.07777 \times 10^4$	$9.21784 \times 10^4$	0.85%	-11.79%	0.09%	-0.98%
	50	$8.37950 \times 10^4$	1710.1	$8.87698 \times 10^4$	$9.12604 \times 10^4$	-1.58%	-10.03%	-2.13%	-1.96%
	75	$8.37950 \times 10^4$	1710.1	$8.87698 \times 10^4$	$9.12604 \times 10^4$	-1.58%	-10.03%	-2.13%	-1.96%
	100	$8.37950 \times 10^4$	1710.1	$8.87698 \times 10^4$	$9.12604 \times 10^4$	-1.58%	-10.03%	-2.13%	-1.96%
Barcelona	0	$1.07567 \times 10^7$	60582.3	$1.08934 \times 10^7$	$1.09595 \times 10^7$	NA	NA	NA	NA
	50	$1.05331 \times 10^7$	52476.9	$1.06520 \times 10^7$	$1.07207 \times 10^7$	-2.08%	-13.38%	-2.22%	-2.18%
	100	$1.05892 \times 10^7$	54657.3	$1.07140 \times 10^7$	$1.08068 \times 10^7$	-1.56%	-9.78%	-1.65%	-1.39%
	150	$1.07194 \times 10^7$	53424.2	$1.08414 \times 10^7$	$1.09211 \times 10^7$	-0.35%	-11.82%	-0.48%	-0.35%
	200	$1.07684 \times 10^7$	50209.9	$1.08828 \times 10^7$	$1.09592 \times 10^7$	0.11%	-17.12%	-0.10%	0.00%
	250	$1.08746 \times 10^7$	41275.4	$1.09682 \times 10^7$	$1.10335 \times 10^7$	1.10%	-31.87%	0.69%	0.68%
	300	$1.07627 \times 10^7$	51422.6	$1.08789 \times 10^7$	$1.09463 \times 10^7$	0.06%	-15.12%	-0.13%	-0.12%
	350	$1.04369 \times 10^7$	52503.4	$1.05567 \times 10^7$	$1.06490 \times 10^7$	-2.97%	-13.34%	-3.09%	-2.83%
	400	$1.04126 \times 10^7$	48557.8	$1.05209 \times 10^7$	$1.06027 \times 10^7$	-3.20%	-19.85%	-3.42%	-3.26%
	450	$1.04126 \times 10^7$	48557.8	$1.05209 \times 10^7$	$1.06027 \times 10^7$	-3.20%	-19.85%	-3.42%	-3.26%
	500	$1.04126 \times 10^7$	48557.8	$1.05209 \times 10^7$	$1.06027 \times 10^7$	-3.20%	-19.85%	-3.42%	-3.26%

Table 3: Theoretical worst case comparison and solution time for Robust Arc-Shipments-Risk HNDR (R1)

Dataset	$\Gamma$	Nominal TWCR	Robust TWCR	Gap	Time(s)
Ravenna	25	$9.5592 \times 10^4$	$9.2952 \times 10^4$	-2.76%	61.7
	50	$1.0000 \times 10^5$	$9.6389 \times 10^4$	-3.61%	53.4
	75	$1.0191 \times 10^5$	$9.7789 \times 10^4$	-4.04%	41.2
	100	$1.0301 \times 10^5$	$9.8592 \times 10^4$	-4.29%	40.8
	125	$1.0363 \times 10^5$	$9.9041 \times 10^4$	-4.43%	45.7
	150	$1.0401 \times 10^5$	$9.9265 \times 10^4$	-4.56%	32.6
	175	$1.0420 \times 10^5$	$9.9359 \times 10^4$	-4.65%	31.8
	200	$1.0429 \times 10^5$	$9.9376 \times 10^4$	-4.71%	27.2
Barcelona	50	$1.00391 \times 10^7$	$9.94587 \times 10^6$	-0.93%	4988
	100	$1.06920 \times 10^7$	$1.05943 \times 10^7$	-0.91%	4915
	150	$1.11663 \times 10^7$	$1.11194 \times 10^7$	-0.42%	6817
	200	$1.15111 \times 10^7$	$1.13875 \times 10^7$	-1.07%	11033
	250	$1.17718 \times 10^7$	$1.15975 \times 10^7$	-1.48%	11382
	300	$1.19921 \times 10^7$	$1.18250 \times 10^7$	-1.39%	6013
	350	$1.21711 \times 10^7$	$1.16695 \times 10^7$	-4.12%	6226
	400	$1.23084 \times 10^7$	$1.16341 \times 10^7$	-5.48%	5125
	450	$1.24110 \times 10^7$	$1.16777 \times 10^7$	-5.91%	5053
500	$1.24856 \times 10^7$	$1.16953 \times 10^7$	-6.33%	5026	

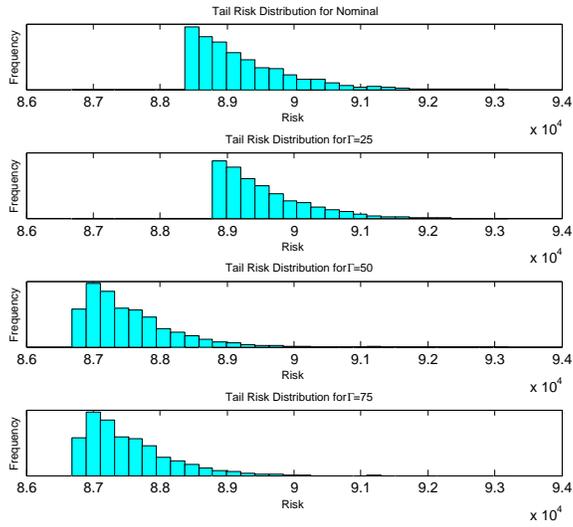
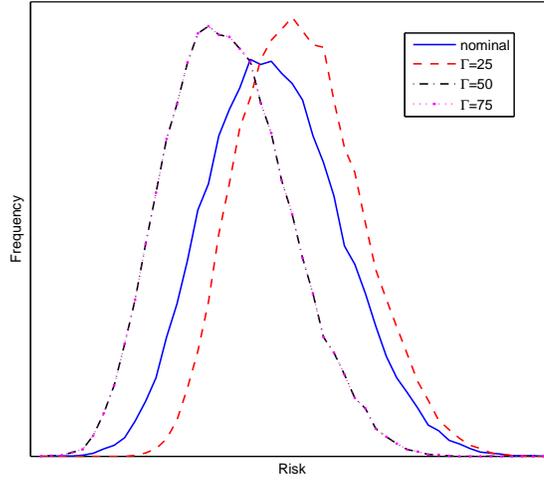


Figure 4: Performance comparison of solutions generated with different  $\Gamma$  values for Robust Arc-Shipment-Risk HNDP (R1) on the Ravenna network

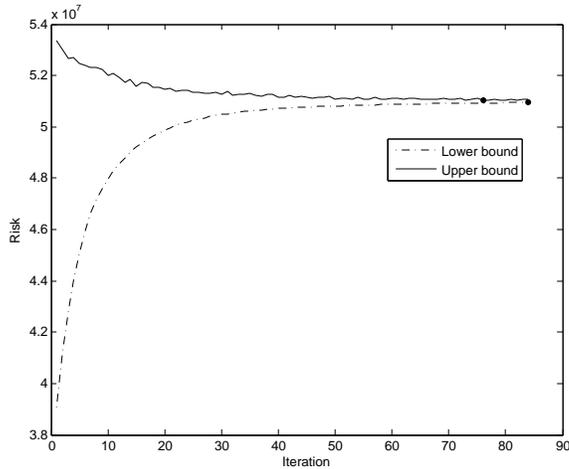


Figure 5: Lagrangian heuristic search process illustration.

Table 4: Comparison between CPLEX solver and Lagrangian heuristic for Robust Arc-Risk HNBP

OD Number	CPLEX		Lagrangian heuristic			GapA	GapB
	Time(s)	Objective	Time(s)	Upper Bound	Lower Bound		
500	86.8	$5.10344 \times 10^7$	26.1	$5.10477 \times 10^7$	$5.09256 \times 10^7$	0.24%	0.026%
1500	323.7	$1.55458 \times 10^8$	35.4	$1.55559 \times 10^8$	$1.55066 \times 10^8$	0.32%	0.065%
2500	613.2	$2.58295 \times 10^8$	43.2	$2.58520 \times 10^8$	$2.57608 \times 10^8$	0.35%	0.087%
3500	1566.0	$2.38922 \times 10^8$	48.1	$2.39175 \times 10^8$	$2.38255 \times 10^8$	0.38%	0.106%
4500	2678.9	$4.66045 \times 10^8$	56.4	$4.66482 \times 10^8$	$4.64807 \times 10^8$	0.36%	0.094%
5500	3034.1	$5.68190 \times 10^8$	61.9	$5.68788 \times 10^8$	$5.66545 \times 10^8$	0.39%	0.105%

### 6.3 Robust Arc-Risk HNBP (R2)

For the Robust Arc-Risk HNBP, it is satisfactory to use CPLEX to solve small and medium size networks. For large networks, we use the proposed Lagrangian heuristic method. An example of the search process for the Barcelona network with 500 OD pairs is shown in Figure 5. The dashed line shows how the lower bound improves during the sub-gradient search process. We can also obtain an upper bound for each lower bound solution, as shown by the solid line. The best upper and lower bounds are shown by the large dots. Figure 5 shows the process of how we can obtain improved upper and lower bounds using the Lagrangian heuristic method and achieve a good solution.

In order to see the efficiency of the Lagrangian algorithm, we test the algorithm on various sizes of OD pairs with uncertainty level  $\Gamma = 800$  on the Barcelona network to solve the Min Risk problem using the higher performance computer with 32GB memory. We assume that 40% of all links—which are randomly chosen in experiments—are subject to uncertainty and the possible risk level can deviate 100% from the nominal risk level, i.e.  $d_{ij} = r_{ij}$ . The Lagrangian heuristic method will stop if the lower and upper bounds are within a 0.1% gap or cannot improve for 10 consecutive steps. We record the

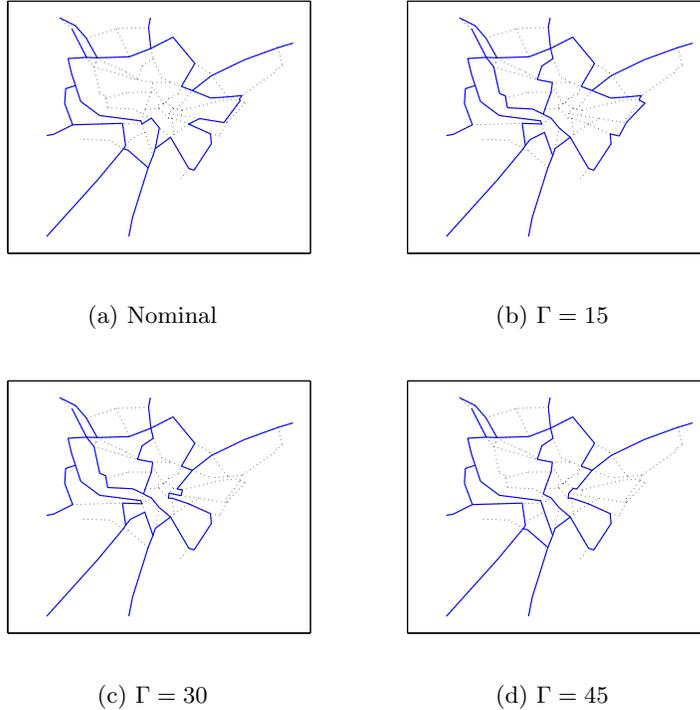


Figure 6: Network solutions for different  $\Gamma$  values for Robust Arc-Risk HNDP (R2)

results in Table 4. GapA is defined as the gap between upper and lower bounds, while GapB is defined as the gap between the Lagrangian heuristic method solution (Upper Bound) and the CPLEX solution. From the table, we can see that the Lagrangian heuristic method can obtain good solutions. Moreover, the Lagrangian heuristic method can run on any ordinary computer while CPLEX requires a computer with large memory. Thus the Lagrangian heuristic method can obtain efficient solutions and can be a good approach especially if time and computer memory are limited.

Next we show results of the nominal and different robust solutions. The solutions are obtained by the heuristic framework using CPLEX to solve the Min Risk problem. We assume 40% of all links are subject to uncertainty and the uncertainty upper bound  $d_{ij}$  is randomly generated using a uniform distribution between 0 to  $2r_{ij}$  for all  $(i, j) \in \mathcal{A}$ . We test different  $\Gamma$  values on the Ravenna data considering transporting gasoline and the Barcelona network with 100 OD number pairs. A visualization of the Ravenna solutions with  $\Gamma = 0$  (nominal), 15, 30, 45 is provided in Figure 6. We can see by considering different numbers of links that are subject to uncertainty,  $\Gamma$ , we obtain different network solutions.

In order to test the effectiveness of robust solutions, we run a simulation to measure the risk for solutions with different  $\Gamma$  values for both networks. We generate 10,000 samples with risk values as  $r_{ij} + u_{ij}d_{ij}$  where  $u_{ij}$  is uniformly distributed in the interval  $[0, 1]$ . The budget of uncertainty for generating the samples is 45 and 1,000 for Ravenna and Barcelona networks respectively. The same

statistics with the Robust Arc-Shipment-Risk HNDP (R1) are recorded in Table 5. The theoretical performance of robust solutions compared to nominal solution is calculated in Table 6. The time of solving each robust solution is also recorded.

By looking at the MOR Gap column in Table 5, the MOR values for the robust solutions have both higher and lower risk than that of the nominal one. A possible reason is that the robust solutions are minimizing the worst case risk and the nominal solution can perform relatively well under some simulated scenarios. However, for the WCR, the robust solutions perform at least as well as the nominal solutions and mostly result in risk reduction. For the Ravenna network the highest WCR reduction is 3.73% and for the Barcelona network, the WCR reduction can be as high as 6.34%. The MWCR has similar results for the Barcelona network by observing the MWCR Gap column. For the Ravenna network, the MWCR risk reduction is limited and has only high risk reduction with  $\Gamma = 40, 45, 50$ . For some cases, there is even a risk increase. This could be due to the limited number of high risk uncertainty cases generated from the samples.

Similar to the Robust Arc-Shipment-Risk HNDP (R1) case, we find a significant reduction of the standard deviation (highest risk reduction of 19.53% for the Ravenna network and 34.01% for the Barcelona network). The standard deviation (SD) measures the risk distribution variability under uncertainty. With smaller SD, the robust solutions can have a lower possibility of having higher risk values even with relatively higher mean risk values. We plot the overall risk distributions with  $\Gamma = 0$  (nominal), 15, 30, 45 for the Ravenna network in Figure 7 as well as the tail distribution (highest 5% risk values) for clarity of the worst case scenario. We can see for the nominal solution, it has a higher probability of obtaining lower risk values. However it also has a higher chance of having higher risk values than the robust solution. Additionally, by looking at the tail distributions of all solutions, the robust solutions have lower worst case risk and the results are consistent with that of Table 5.

After the analysis of results in Table 5, we now look at the theoretical worst case risk from Table 6. The robust TWCR is the objective value of the robust Arc-Risk HNDP model with budget of uncertainty  $\Gamma$ . The nominal TWCR is obtained by summing up the highest  $\Gamma$  number of arc uncertainty values ( $\sum_{s \in \mathcal{S}} n_s d_{ijs} x_{ijs}^0$ ) and the nominal HNDP objective risk value ( $\sum_{s \in \mathcal{S}} \sum_{(i,j) \in \mathcal{A}} n_s r_{ijs} x_{ijs}^0$ ) where  $x_{ijs}^0$  is the nominal solution. We observe that the risk reduction of the robust solutions is consistent. The highest risk reduction in the test cases is 4.07% for the Ravenna network and 6.32% for the Barcelona network.

Table 5: Statistics of nominal and robust solutions for Robust Arc-Risk HNBP (R2)

Dataset		MOR	SD	MWCR	WCR	MOR Gap	SD Gap	MWCR Gap	WCR Gap
Ravenna	0	$8.5739 \times 10^4$	3096.6	$9.2956 \times 10^4$	$9.9441 \times 10^4$	NA	NA	NA	NA
	5	$8.5739 \times 10^4$	3096.6	$9.2956 \times 10^4$	$9.9441 \times 10^4$	0.00%	0.00%	0.00%	0.00%
	10	$8.7080 \times 10^4$	2491.9	$9.2937 \times 10^4$	$9.6437 \times 10^4$	1.56%	-19.53%	-0.02%	-3.02%
	15	$8.7080 \times 10^4$	2491.9	$9.2937 \times 10^4$	$9.6437 \times 10^4$	1.56%	-19.53%	-0.02%	-3.02%
	20	$8.7080 \times 10^4$	2491.9	$9.2937 \times 10^4$	$9.6437 \times 10^4$	1.56%	-19.53%	-0.02%	-3.02%
	25	$8.7080 \times 10^4$	2491.9	$9.2937 \times 10^4$	$9.6437 \times 10^4$	1.56%	-19.53%	-0.02%	-3.02%
	30	$8.7382 \times 10^4$	2493.1	$9.3244 \times 10^4$	$9.6945 \times 10^4$	1.92%	-19.49%	0.31%	-2.51%
	35	$8.7170 \times 10^4$	2699.0	$9.3518 \times 10^4$	$9.7531 \times 10^4$	1.67%	-12.84%	0.60%	-1.92%
	40	$8.4989 \times 10^4$	2620.4	$9.1111 \times 10^4$	$9.5730 \times 10^4$	-0.88%	-15.38%	-1.99%	-3.73%
	45	$8.5040 \times 10^4$	2599.9	$9.1111 \times 10^4$	$9.6665 \times 10^4$	-0.82%	-16.04%	-1.98%	-2.79%
	50	$8.5040 \times 10^4$	2599.9	$9.1111 \times 10^4$	$9.6665 \times 10^4$	-0.82%	-16.04%	-1.98%	-2.79%
Barcelona	0	$1.22884 \times 10^7$	49685.1	$1.23627 \times 10^7$	$1.23822 \times 10^7$	NA	NA	NA	NA
	25	$1.22884 \times 10^7$	49685.1	$1.23627 \times 10^7$	$1.23822 \times 10^7$	0.00%	0.00%	0.00%	0.00%
	50	$1.22884 \times 10^7$	49685.1	$1.23627 \times 10^7$	$1.23822 \times 10^7$	0.00%	0.00%	0.00%	0.00%
	75	$1.19424 \times 10^7$	32788.5	$1.19913 \times 10^7$	$1.20033 \times 10^7$	-2.82%	-34.01%	-3.00%	-3.06%
	100	$1.21379 \times 10^7$	41031.2	$1.21977 \times 10^7$	$1.22091 \times 10^7$	-1.22%	-17.42%	-1.33%	-1.40%
	125	$1.17506 \times 10^7$	33416.7	$1.17994 \times 10^7$	$1.18096 \times 10^7$	-4.38%	-32.74%	-4.56%	-4.62%
	150	$1.20604 \times 10^7$	43366.8	$1.21225 \times 10^7$	$1.21344 \times 10^7$	-1.85%	-12.72%	-1.94%	-2.00%
	175	$1.16019 \times 10^7$	35728.1	$1.16515 \times 10^7$	$1.16612 \times 10^7$	-5.59%	-28.09%	-5.75%	-5.82%
	200	$1.15276 \times 10^7$	39078.8	$1.15851 \times 10^7$	$1.15976 \times 10^7$	-6.19%	-21.35%	-6.29%	-6.34%
	225	$1.16143 \times 10^7$	36511.4	$1.16645 \times 10^7$	$1.16730 \times 10^7$	-5.49%	-26.51%	-5.65%	-5.73%
	250	$1.16143 \times 10^7$	36511.4	$1.16645 \times 10^7$	$1.16730 \times 10^7$	-5.49%	-26.51%	-5.65%	-5.73%

Table 6: Theoretical worst case comparison and solution time for Robust Arc-Risk HNBP (R2)

Dataset	$\Gamma$	Nominal TWCR	Robust TWCR	Gap	Time(s)
Ravenna	5	$9.24334 \times 10^4$	$9.24334 \times 10^4$	0.00%	33.5
	10	$1.00017 \times 10^5$	$9.66095 \times 10^4$	-3.41%	35.7
	15	$1.02334 \times 10^5$	$9.84450 \times 10^4$	-3.80%	33.0
	20	$1.03339 \times 10^5$	$9.94038 \times 10^4$	-3.81%	34.1
	25	$1.03754 \times 10^5$	$9.95311 \times 10^4$	-4.07%	31.0
	30	$1.04006 \times 10^5$	$1.00907 \times 10^5$	-2.98%	29.5
	35	$1.04166 \times 10^5$	$1.01656 \times 10^5$	-2.41%	30.5
	40	$1.04248 \times 10^5$	$1.00425 \times 10^5$	-3.67%	30.1
	45	$1.04290 \times 10^5$	$1.00202 \times 10^5$	-3.92%	31.4
	50	$1.04298 \times 10^5$	$1.00224 \times 10^5$	-3.91%	27.0
Barcelona	25	$1.05992 \times 10^7$	$1.05992 \times 10^7$	0.00%	10262
	50	$1.13054 \times 10^7$	$1.13054 \times 10^7$	0.00%	6142
	75	$1.17215 \times 10^7$	$1.16296 \times 10^7$	-0.78%	6740
	100	$1.19793 \times 10^7$	$1.19612 \times 10^7$	-0.15%	2862
	125	$1.21676 \times 10^7$	$1.17514 \times 10^7$	-3.42%	7020
	150	$1.22856 \times 10^7$	$1.21035 \times 10^7$	-1.48%	4161
	175	$1.23537 \times 10^7$	$1.16620 \times 10^7$	-5.60%	2284
	200	$1.23813 \times 10^7$	$1.15989 \times 10^7$	-6.32%	719
	225	$1.23850 \times 10^7$	$1.16738 \times 10^7$	-5.74%	1195
	250	$1.23850 \times 10^7$	$1.16738 \times 10^7$	-5.74%	1127

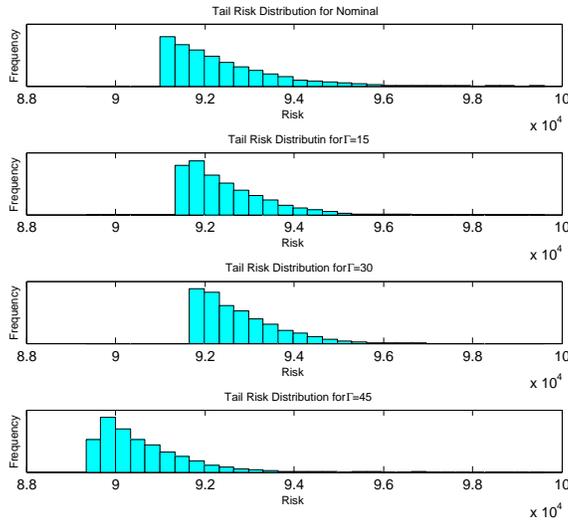
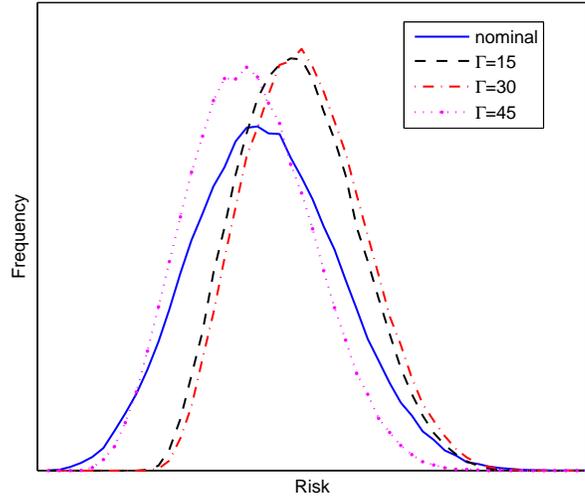


Figure 7: Performance comparison of solutions generated with different  $\Gamma$  values for Robust Arc-Risk HNRP (R2) on the Ravenna network

## 7 Concluding Remarks

We consider the hazardous materials network design problem (HNDP). In this paper, we present two formulations for the HNDP considering edge risk uncertainty using the robust optimization method with a cardinality uncertainty set. The two robust versions are Robust Arc-Shipment-Risk HNDP and Robust Arc-Risk HNDP. We solve these two problems by modifying a heuristic method. Utilizing the  $K$ -shortest path algorithm and adding a feasibility check, we are able to obtain stable solutions, as defined in the HNDP literature. For large networks, we also propose a golden section search approach and a Lagrangian heuristic to solve the Min Risk problem for Robust Arc-Shipment-Risk and Robust Arc-Risk HNDP separately. The efficiency of the proposed algorithms are tested on large networks. Moreover, in order to test the effectiveness of the robust formulations, we test the solutions on a real network (Ravenna). In comparison with the nominal solution, the robust solutions results are superior particularly in high risk uncertainty scenarios and the worst case risk scenario. Robust solutions also provide smaller variability.

There are a few directions for extending this research. One direction is to control the number of links to be closed for hazmat traffic. Closing links costs the government for administration and monitoring. Therefore it would be desirable to minimize the number of links to close. To block a path, it could be possible for the government to close a few key links instead of closing all links in the path. One could obtain the same solution as in the current heuristic method with fewer closed links while requiring greater computational effort.

Another possible direction is to consider other uncertainty risk measures. A popular risk measure is the expected risk which can be computed as the multiplication of hazmat accident probability and consequence. If we consider uncertainty in both of them, we would need to consider a more complicated robust optimization formulation. Kwon et al. (2013) consider the robust shortest path problem with two uncertain multiplicative cost coefficients. We can incorporate this method into the hazmat network design problem.

## References

- Amaldi, E., M. Bruglieri, B. Fortz. 2011. On the hazmat transport network design problem. *Network Optimization* 327–338.
- Atamtürk, A., M. Zhang. 2007. Two-stage robust network flow and design under demand uncertainty. *Operations Research* 55(4) 662–673.

- Bar-Gera, H. 2013. Transportation network test problems. URL <http://www.bgu.ac.il/~bargera/tntp/>.
- Bazaraa, M. S., H. D. Sherali, C. M. Shetty. 2013. *Nonlinear programming: theory and algorithms*. John Wiley & Sons.
- Ben-Tal, A., A. Nemirovski. 1998. Robust convex optimization. *Mathematics of Operations Research* **23**(4) 769–805.
- Berglund, P. G., C. Kwon. 2014. Robust facility location problem for hazardous waste transportation. *Networks and Spatial Economics* **14**(1) 91–116.
- Bertsimas, D., D. B. Brown, C. Caramanis. 2011. Theory and applications of robust optimization. *SIAM Review* **53**(3) 464–501.
- Bertsimas, D., M. Sim. 2003. Robust discrete optimization and network flows. *Mathematical Programming* **98**(1-3) 49–71.
- Bianco, L., M. Caramia, S. Giordani, V. Piccialli. 2015. A game-theoretic approach for regulating hazmat transportation. *Transportation Science* **Articles in Advance**. URL <http://dx.doi.org/10.1287/trsc.2015.0592>.
- Bonvicini, S., G. Spadoni. 2008. A hazmat multi-commodity routing model satisfying risk criteria: A case study. *Journal of Loss Prevention in the Process Industries* **21**(4) 345–358.
- Colson, B., P. Marcotte, G. Savard. 2005. Bilevel programming: A survey. *4OR* **3**(2) 87–107.
- El Ghaoui, L., H. Lebret. 1997. Robust solutions to least-squares problems with uncertain data. *SIAM Journal on Matrix Analysis and Applications* **18**(4) 1035–1064.
- Erkut, E., O. Alp. 2007. Designing a road network for hazardous materials shipments. *Computers & Operations Research* **34**(5) 1389–1405.
- Erkut, E., F. Gzara. 2008. Solving the hazmat transport network design problem. *Computers & Operations Research* **35**(7) 2234–2247.
- Erkut, E., S. A. Tjandra, V. Verter. 2007. Hazardous materials transportation. *Handbooks in Operations Research and Management Science* **14** 539–621.
- Gzara, F. 2013. A cutting plane approach for bilevel hazardous material transport network design. *Operations Research Letters* **41**(1) 40–46.

- Held, M., P. Wolfe, H. P. Crowder. 1974. Validation of subgradient optimization. *Mathematical Programming* **6**(1) 62–88.
- Kara, B. Y., V. Verter. 2004. Designing a road network for hazardous materials transportation. *Transportation Science* **38**(2) 188–196.
- Kwon, C., T. Lee, P. Berglund. 2013. Robust shortest path problems with two uncertain multiplicative cost coefficients. *Naval Research Logistics* **60**(5) 375–394.
- Marcotte, P., A. Mercier, G. Savard, V. Verter. 2009. Toll policies for mitigating hazardous materials transport risk. *Transportation Science* **43**(2) 228–243.
- Mudchanatongsuk, S., F. Ordóñez, J. Liu. 2007. Robust solutions for network design under transportation cost and demand uncertainty. *Journal of the Operational Research Society* **59**(5) 652–662.
- Ordóñez, F., J. Zhao. 2007. Robust capacity expansion of network flows. *Networks* **50**(2) 136–145.
- Peinhardt, M. A. 2003. Integer multicommodity flows in optical networks. Diplomarbeit, Technische Universität Berlin.
- Verter, V., B. Y. Kara. 2008. A path-based approach for hazmat transport network design. *Management Science* **54**(1) 29–40.
- Wang, J., Y. Kang, C. Kwon, R. Batta. 2012. Dual toll pricing for hazardous materials transport with linear delay. *Networks and Spatial Economics* **12**(1) 147–165.
- Xin, C., Q. Letu, Y. Bai. 2013. Robust optimization for the hazardous materials transportation network design problem. *Combinatorial Optimization and Applications*. Springer, 373–386.
- Xu, P., L. Wang. 2014. An exact algorithm for the bilevel mixed integer linear programming problem under three simplifying assumptions. *Computers & Operations Research* **41** 309–318.
- Yen, J. Y. 1971. Finding the  $K$  shortest loopless paths in a network. *Management Science* **17**(11) 712–716.

# Appendices

## A Proof of Projection Method

The procedure introduced in Section 5.2.3 obtains  $\bar{\mu}$  as the projection of  $\mu$  onto  $\{\mu : 0 \leq \mu_{ij} \leq 1 \forall (i, j) \in \mathcal{A}_k, \Gamma - \sum_{(i,j) \in \mathcal{A}_k} \mu_{ij} \geq 0\}$ .

*Proof.* The projection of  $\mu$  onto  $\{\mu : 0 \leq \mu_{ij} \leq 1 \forall (i, j) \in \mathcal{A}_k, \Gamma - \sum_{(i,j) \in \mathcal{A}_k} \mu_{ij} \geq 0\}$  can be formulated as:

$$\min \sum_{(i,j) \in \mathcal{A}_k} \frac{1}{2} (\mu_{ij} - \bar{\mu}_{ij})^2 \quad (36)$$

subject to

$$\begin{aligned} \sum_{(i,j) \in \mathcal{A}_k} \bar{\mu}_{ij} &\leq \Gamma && : \lambda \\ \bar{\mu}_{ij} &\leq 1 \quad \forall (i, j) \in \mathcal{A}_k && : v_{ij} \\ \bar{\mu}_{ij} &\geq 0 \quad \forall (i, j) \in \mathcal{A}_k && : \rho_{ij} \end{aligned}$$

The Karush-Kuhn-Tucker (KKT) conditions for problem (36) are

$$\bar{\mu}_{ij} - \mu_{ij} + \lambda + v_{ij} - \rho_{ij} = 0 \quad (37)$$

$$\lambda \left( \sum_{(i,j) \in \mathcal{A}_k} \bar{\mu}_{ij} - \Gamma \right) = 0 \quad (38)$$

$$v_{ij} (\bar{\mu}_{ij} - 1) = 0 \quad (39)$$

$$\rho_{ij} \bar{\mu}_{ij} = 0 \quad (40)$$

$$\lambda, v_{ij}, \rho_{ij} \geq 0 \quad (41)$$

$$\sum_{(i,j) \in \mathcal{A}_k} \bar{\mu}_{ij} \leq \Gamma \quad (42)$$

$$\bar{\mu}_{ij} \leq 1 \quad (43)$$

$$\bar{\mu}_{ij} \geq 0 \quad (44)$$

Problem (36) is strictly convex over a nonempty bounded region, so there must exist multipliers  $\lambda$ ,  $v_{ij}$  and  $\rho_{ij}$  that satisfy the KKT conditions. We show that we can find a solution to the above KKT conditions by following the projection procedure in Section 5.2.3.

1. For any  $\lambda$ , we set

$$\begin{aligned}\bar{\mu}_{ij} &= \max\{0, \min\{1, \mu_{ij} - \lambda\}\} \\ v_{ij} &= \max\{0, \mu_{ij} - \lambda - 1\} \\ \rho_{ij} &= \max\{0, -(\mu_{ij} - \lambda)\}\end{aligned}\tag{45}$$

for each  $(i, j) \in \mathcal{A}_k$ . We consider the following three cases:

- When  $\mu_{ij} - \lambda \geq 1$ , we have  $\bar{\mu}_{ij} = 1$ ,  $\rho_{ij} = 0$ , and  $v_{ij} = \mu_{ij} - \lambda - 1$ .
- When  $0 \leq \mu_{ij} - \lambda < 1$ , we have  $\bar{\mu}_{ij} = \mu_{ij} - \lambda$ ,  $\rho_{ij} = 0$ , and  $v_{ij} = 0$ .
- When  $\mu_{ij} - \lambda < 0$ , we have  $\bar{\mu}_{ij} = 0$ ,  $\rho_{ij} = -(\mu_{ij} - \lambda)$ , and  $v_{ij} = 0$ .

The above  $\bar{\mu}_{ij}$ ,  $\rho_{ij}$ , and  $v_{ij}$  satisfy conditions (37), (39), (40), (41), (43), and (44) for any given  $\lambda$ . Remaining conditions are (38), (42), and the nonnegativity of  $\lambda$ .

2. When  $\lambda = 0$ , if the resulting  $\bar{\mu}_{ij} = \max\{0, \min\{1, \mu_{ij}\}\}$  satisfies (42), then we are done as described in Step 0 of the projection procedure. If not, we can show that we can obtain  $\lambda > 0$  such that (38) and (42) hold, by following Step 1 of the projection procedure. That is, we can find  $\lambda > 0$  such that

$$\sum_{(i,j) \in \mathcal{A}_k} \bar{\mu}_{ij} = \sum_{(i,j) \in \mathcal{A}_k} \max\{0, \min\{1, \mu_{ij} - \lambda\}\} = \Gamma.\tag{46}$$

Equation (46) can be rewritten as

$$\sum_{(i,j) \in \mathcal{A}_k} \bar{\mu}_{ij} = \sum_{\{(i,j) \in \mathcal{A}_k : \mu_{ij} - 1 \leq \lambda < \mu_{ij}\}} (\mu_{ij} - \lambda) + \sum_{\{(i,j) \in \mathcal{A}_k : \lambda < \mu_{ij} - 1\}} (1) + \sum_{\{(i,j) \in \mathcal{A}_k : \lambda \geq \mu_{ij}\}} (0) = \Gamma.\tag{47}$$

In equation (47),  $\mu_{ij}$  and  $\mu_{ij} - 1$  are critical values. Thus we can find the value of  $\lambda$  by searching the interval of these values as used in Step 1.

This completes the proof. □