Designing Relay-Hub Networks for Consolidation Planning Under Demand Uncertainty

Onkar Kulkarni, Mathieu Dahan, and Benoit Montreuil

H. Milton Stewart School of Industrial & Systems Engineering,
Georgia Institute of Technology, Atlanta, GA, USA
onkar.kulkarni@gatech.edu, mathieu.dahan@isye.gatech.edu, benoit.montreuil@isye.gatech.edu

February 15, 2025

Keywords: Relay network design; stochastic demand; consolidation-based routing; branch-and-Benders-cut; integer L-shaped method

1 INTRODUCTION

The trucking industry forms the backbone of any major logistics system: In 2019 in the U.S., 64.7% of the total domestic tonnage—which corresponded to $940.8 billion in gross freight revenues—were transported across the country through trucks (U.S. Bureau of Transportation Statistics, 2023). However, to maximize truck utilization and reduce costs, trucking companies consolidate shipments between locations, requiring drivers to travel for more than 2,000 miles in a single journey. These unsustainable working conditions have been shown to take a toll on the drivers' quality of life, leading to 80-90% annual turnover rates and a chronic driver shortage expected to reach 160,000 by 2030 in the U.S. (American Trucking Associations, 2019).

One potential solution identified by companies such as Amazon.com and UPS is to build a network of relay hubs to fulfill long-haul demand via single-day driver trips. Although there has been an increase of academic interest in this area, the investigations often come with restrictive assumptions. First, most models on relay logistics network design assume deterministic commodity demands (Cabral *et al.*, 2007, Yıldız *et al.*, 2018, Leitner *et al.*, 2019). Furthermore, a large number of models pre-allocate demand origins and destinations to specific hubs, constraining commodity flows along partial routes, thus leading to an increase in traveled distances and potential hub congestion (Üster & Kewcharoenwong, 2011, Kewcharoenwong & Üster, 2017). Although Hu *et al.* (2019) considered stochastic commodity demands, one of their major limitations, along with other investigations considering fixed-charge relay network design, lies in assuming continuous flow routing decisions that do not account for the important consolidation incentives in logistics planning. Finally, the corresponding developed models have been solved only for small- to medium-size instances and are less applicable to real-world industry issues.

In this work, we formulate a two-stage stochastic optimization program to design a relay-based transportation network. In the first stage, we locate relay logistics hubs and decide their capacities. Then, we design a minimum-cost consolidation plan with back-hauling to transport commodities upon realization of the demand uncertainty in the second stage. To exactly solve this challenging problem with second-stage mixed-integer recourse, we devise and accelerate a decomposition-based branch-and-cut algorithm with nested Benders decomposition and integer L-shaped method. We validate our model and methodology by designing large-scale relay networks for finished vehicle deliveries from our U.S.-based car manufacturing partner.

2 PROBLEM DESCRIPTION

We consider a truck carrier (or a consortium of such carriers) interested in designing a relay-hub network to deliver commodities between a set $\mathcal{P}$ of origin-destination (O-D) pairs. As the carrier faces uncertain demand, the demand variability is modeled through a finite set of demand scenarios $\omega \in \Omega$ associated with commodity demand volume $d^{p,\omega}$ for each O-D pair p. The carrier intends to build relay logistics hubs with capacities belonging in a set of possible configurations $\mathcal{K}$ within a pre-selected set of candidate locations $\mathcal{H}$. Let $\mathcal{A}$ represent the available directed transportation legs that satisfy the driving time regulations to ensure a daily return for all drivers to their respective homes. The goal is to select a subset of hub locations and their respective capacities so as to minimize hub opening and expected transportation costs.

We formulate this network design problem using a two-stage stochastic optimization model with mixed-integer recourse. The first-stage decisions consist of binary variables y_i^k equal to 1 if a hub at location $i \in \mathcal{H}$ with capacity configuration $k \in \mathcal{K}$ is opened. Then, for each demand scenario ω, we define continuous variables $f_{i,j}^{p,\omega}$ to represent the volume of commodity between O-D pair $p \in \mathcal{P}$ transported on leg $(i,j) \in \mathcal{A}$, and discrete variables $x_{i,j}^{\omega}$ to determine the number of trucks traveling on leg $(i,j) \in \mathcal{A}$. The resulting optimization problem is given as follows:

$$\min \quad \sum_{i \in \mathcal{H}} \sum_{k \in \mathcal{K}} C_i^k \cdot y_i^k + \mathbb{E}_\omega \Big[\sum_{\{(i,j) \in \mathcal{A} \,|\, i<j\}} A_{i,j} \cdot x_{i,j}^{\omega} + \sum_{p \in \mathcal{P}} \sum_{(i,j) \in \mathcal{A}} B_{i,j} \cdot f_{i,j}^{p,\omega} \Big] \tag{1a}$$

$$\text{s.t.} \quad \sum_{k \in \mathcal{K}} y_i^k \leq 1, \qquad \forall i \in \mathcal{H}, \tag{1b}$$

$$\sum_{\substack{\{j \in \mathcal{H} \cup \{t\} \\ |\, (i,j) \in \mathcal{A}\}}} f_{i,j}^{p,\omega} - \sum_{\substack{\{j \in \mathcal{H} \cup \{s\} \\ |\, (j,i) \in \mathcal{A}\}}} f_{j,i}^{p,\omega} = \begin{cases} d^{p,\omega} & \text{if } i = s \\ 0 & \text{if } i \in \mathcal{H} \\ -d^{p,\omega} & \text{if } i = t \end{cases}, \quad \forall p = (s,t) \in \mathcal{P},\ \forall i \in \mathcal{H} \cup \{s,t\}, \tag{1c}$$

$$\sum_{p \in \mathcal{P}} f_{i,j}^{p,\omega} \leq Q \cdot x_{i,j}^{\omega}, \qquad \forall (i,j) \in \mathcal{A}, \tag{1d}$$

$$\sum_{\{j \in \mathcal{S} \cup \mathcal{H} \cup \mathcal{T} \,|\, (i,j) \in \mathcal{A}\}} x_{i,j}^{\omega} \leq \sum_{k \in \mathcal{K}} S_k \cdot y_i^k, \qquad \forall i \in \mathcal{H}, \tag{1e}$$

$$x_{i,j}^{\omega} = x_{j,i}^{\omega} \in \mathbb{Z}_{\geq 0}, \qquad \forall (i,j) \in \mathcal{A} \mid i < j, \tag{1f}$$

$$f_{i,j}^{p,\omega} \geq 0, \qquad \forall (i,j) \in \mathcal{A},\ \forall p \in \mathcal{P}, \tag{1g}$$

$$y_i^k \in \{0,1\}, \qquad \forall i \in \mathcal{H},\ \forall k \in \mathcal{K}. \tag{1h}$$

The objective function (1a) minimizes the total cost incurred by the carrier, while accounting for the hub opening and sizing costs C_i^k, as well as the fixed ($A_{i,j}$) plus linear ($B_{i,j}$) transportation costs in each demand scenario. Constraints (1c) route commodities across the network, while Constraints (1d) ensure that enough trucks, each of volume capacity Q, are scheduled to transport commodities on each leg. Capacity constraints (1e) limit the number of trucks visiting hubs based on their capacities S_k. Finally, Constraints (1f) facilitate back-and-forth trips by balancing the number of trucks in both directions of each leg.

3 SOLUTION METHODOLOGY

The two-stage stochastic program with mixed-integer recourse (1) is challenging to solve, as it embeds $|\Omega|$ minimum-cost consolidation planning problems within a facility location problem. Thus, we develop a decomposition-based branch-and-cut algorithm that uses nested Benders decomposition and integer L-shaped method. We employ Benders decomposition at two stages: Similarly to classical stochastic programs, we solve the dual of the linear programming (LP) relaxation of our second-stage subproblem for each stochastic scenario to add Benders feedback cuts to the first-stage master problem. However, as the LP relaxation of the second-stage problem entails solving a capacitated multi-commodity minimum-cost network flow problem, solving its dual by directly feeding it to an off-the-shelf optimization solver is not computationally efficient. Instead, we use Benders decomposition (again) to solve the relaxed second-stage subproblem by decomposing it for each O-D pair. Here, we leverage the underlying network-flow structure to generate Benders cuts in polynomial time using shortest-path routines in the newly defined third stage. In order to further enhance the computational efficiency of our solution approach, we add valid inequalities guaranteeing enough hub capacity adjacent to commodity origins and destinations. Finally, to guarantee the exactness of our solution approach, we add integer L-shaped cuts by solving the second-stage subproblem exactly through Benders decomposition as well. Figure 1 summarizes our solution approach.

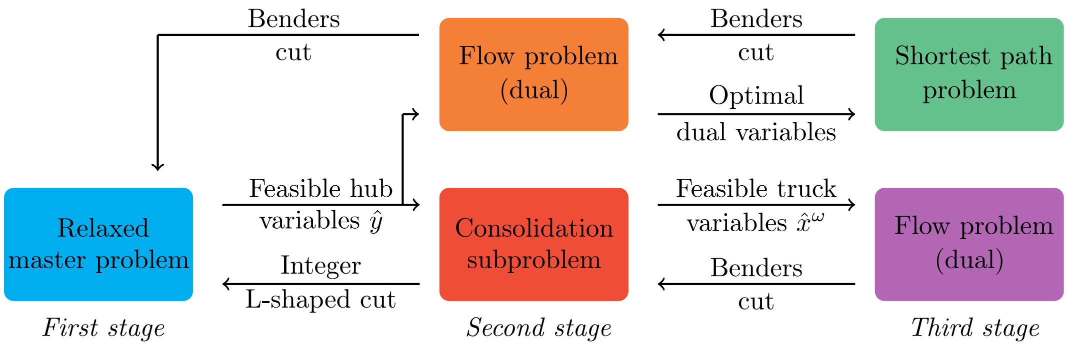

Figure 1 – *Overview of the proposed tri-stage decomposition-based branch-and-cut algorithm*

4 CASE STUDY

We utilize the data from a major car manufacturer in the U.S. that partnered with our research team to create 5 representative problem instances of increasing size and complexity in the South-East region of the U.S., ranging from 20 O-D pairs, 25 candidate hub locations, and 486 legs for Instance 1, to 460 O-D pairs, 88 candidate hub locations, and 3,626 legs for Instance 5. We consider major cities and highway intersections as potential locations for relay-hub construction. Furthermore, transportation legs are selected with a travel time no greater than 5.5 hours to permit truck drivers to return back home without exceeding the 11-hour daily driving limit imposed by the Federal Motor Carrier Safety Administration. Each instance is associated with 30 stochastic demand scenarios received from the car manufacturer, where demand originates at production plants, ports, or rail-heads, and is destined for car dealerships.

Algorithms are implemented in Python v.3.8 and optimization problems are solved using Gurobi v.10.0.3 on an AMD EPYC processor (with IBPB) 2.50 GHz (1 core), with 255 GB assigned RAM and Windows Server 2012 R2 standards, 64-bit operating system, and x64-based processor. To benchmark our developed solution approach, we compare it with a classical decomposition-based branch-and-cut (B&C) algorithm where at every integer feasible first-stage solution, the second-stage subproblem (red in Figure 1) and the dual of the LP relaxation of the second-stage subproblem (orange in Figure 1) are solved exactly using Gurobi (as opposed to further decomposing them in our approach). Table 1 compares the resulting optimality gaps for the instances after a 96-hour time limit.

Table 1 – *Optimality gaps after a 96-hour time limit*

Data instance	**Classical decomposition-based B&C (%)**	**Decomposition-based B&C with nested Benders (%)**
1	32.74	0.00
2	47.91	0.00
3	60.02	3.11
4	69.77	7.94
5	81.32	10.29

From that table, we observe that our proposed method significantly outperforms a classical decomposition-based branch-and-cut that generates Benders and L-shaped cuts. Our approach optimally solves the first two instances, and achieves small optimality gaps for the largest instances that cannot be solved by previous methods. This highlights the value of further decomposing the second-stage subproblems and the duals of their LP relaxations, as well as leveraging their network structures, resulting in our faster tri-stage solution approach.

In Figure 2, we compare the costs of relay networks designed by considering consolidation-based routing (from model (1)), with those of networks designed by continuously approximating routing decisions (i.e., by relaxing second-stage truck integrality constraints (1f)).

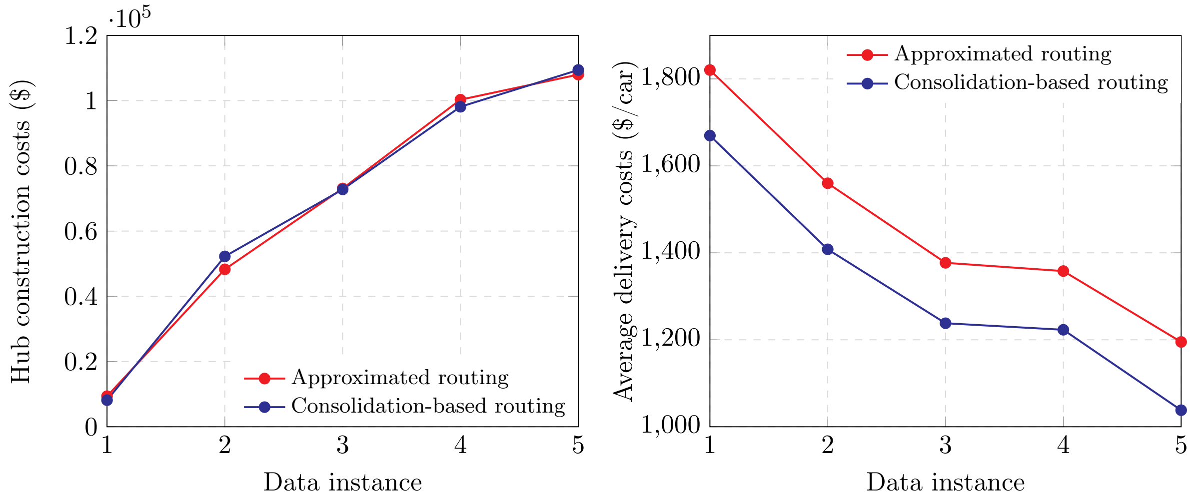

Figure 2 – *Logistics costs of relay networks designed with approximated routing (red) vs. consolidation-based routing (blue)*

Interestingly, we observe that relay networks designed by considering consolidation-based or continuous routing have similar construction costs. However, accounting for consolidation-based routing impacts the networks' topologies, resulting in a 10% decrease in delivery costs on average across instances. This analysis shows the value of integrating tactical consolidation decisions into the strategic facility location and sizing decisions and solving the more computationally challenging problem (1).

REFERENCES

American Trucking Associations. 2019. *Truck driver shortage analysis.* Tech. rept. https://tinyurl.com/jj8ysy2k.

Cabral, Edgar Alberto, Erkut, Erhan, Laporte, Gilbert, & Patterson, Raymond A. 2007. The network design problem with relays. *European Journal of Operational Research*, **180**(2), 834–844.

Hu, Zhengyang, Askin, Ronald G., & Hu, Guiping. 2019. Hub relay network design for daily driver routes. *International Journal of Production Research*, **57**(19), 6130–6145.

Kewcharoenwong, Panitan, & Üster, Halit. 2017. Relay network design with capacity and link-imbalance considerations: A Lagrangean decomposition algorithm and analysis. *Transportation Science*, **51**(4), 1177–1195.

Leitner, Markus, Ljubić, Ivana, Riedler, Martin, & Ruthmair, Mario. 2019. Exact Approaches for Network Design Problems with Relays. *INFORMS Journal on Computing*, **31**(1), 171–192.

U.S. Bureau of Transportation Statistics. 2023. *Transportation Statistics Annual Report.* Tech. rept. U.S. Bureau of Labor Statistics. https://tinyurl.com/zzruzmxb.

Üster, Halit, & Kewcharoenwong, Panitan. 2011. Strategic Design and Analysis of a Relay Network in Truckload Transportation. *Transportation Science*, **45**(4), 505–523.

Yıldız, Barış, Karaşan, Oya Ekin, & Yaman, Hande. 2018. Branch-and-price approaches for the network design problem with relays. *Computers & Operations Research*, **92**, 155–169.