Hub Location and Service Network Design under Uncertainty

Hao Li[a], Gita Taherkhani[b], Mike Hewitt[b], Sibel A. Alumur[a]

[a] Department of Management Science and Engineering, University of Waterloo, Waterloo, ON, Canada
[b] Quinlan School of Business, Loyola University, Chicago, IL, USA

Extended abstract submitted for presentation at the 12th Triennial Symposium on Transportation Analysis conference (TRISTAN XII)
June 22-27, 2025, Okinawa, Japan

February 28, 2025

Keywords: Freight transportation, hub location, service network design, uncertainty, strategic/tactical planning, Benders decomposition

1 INTRODUCTION

For less-than-truckload (LTL) freight carriers, the cost of a vehicle executing a transportation move is mostly fixed regardless of the total size of the shipments it transports. Consequently, the greater the total shipment size in that vehicle (i.e., higher vehicle utilization), the lower the cost per unit.

When multiple shipments can share the same truck movements at the same time, these shipments can potentially be *consolidated* into the same vehicle via *hubs*, leading to high vehicle utilization. The transportation costs between hubs on a per-vehicle basis are often referred to as *modular links* in the hub location literature (Yaman & Carello, 2005). The carriers typically establish regular service routes and policies to meet the requirements and expectations of contractual customers, with optional shipments to fulfill. The carriers need to make important strategic decisions, such as hub locations, for consolidation purposes (Alumur *et al.*, 2021). The transportation carriers also need to propose *services* and *schedules* and establish a transportation plan to ensure timely pickup and delivery of shipments in a profitable way by solving service network design problems (Crainic, 2000).

This paper proposes a profit-maximizing model for a combined hub location and service network design problem for LTL carriers. The model considers partial contracting of shipments, where a subset of shipments are guaranteed to be fulfilled and other shipments will be fulfilled to a certain degree to ensure profit maximization. The problem formulation involves making strategic/time-invariant decisions about hub locations and tactical decisions regarding the timings of dispatch for trucks, utilizing a time-space network to model events across time and location dimensions. A two-stage stochastic model is introduced to account for the variability of shipment volumes.

This paper offers the following contributions to the study of hub location and service network design problems, especially within the realm of freight transportation. Firstly, we believe this is the first paper to address strategic and tactical planning for freight transportation by proposing an integrated hub location and service network design optimization model using a space-time network. Additionally, this is the first time uncertainty in shipment sizes has been explicitly acknowledged and incorporated in such a setting, with the development of a two-stage stochastic formulation. We also develop an exact Benders-based algorithmic framework, specifically designed to manage the large-scale stochastic aspects of this problem. Our computational experiments indicate that the algorithm, enhanced by valid inequalities, substantially decreases computation time and enables the solving of medium to large-scale instances of the problem, which commercial solvers often fail to find feasible solutions for. Additionally, by applying this model to a realistic case study from a freight carrier in the US, the paper provides important insights for both practitioners and researchers in the field of freight logistics.

2 PROBLEM FORMULATION

To model the time aspect of the problem, we extend the static network, $\mathcal{G}$, to a time-expanded network $\mathcal{G}_\mathcal{T} = (\mathcal{N}_\mathcal{T}, \mathcal{H}_\mathcal{T} \cup \mathcal{A}_\mathcal{T})$, where $\mathcal{T}$, $\mathcal{N}_\mathcal{T}$, $\mathcal{H}_\mathcal{T}$, and $\mathcal{A}_\mathcal{T}$ represent a set of time steps, time-expanded nodes, holding arcs, and transportation arcs, respectively. Let K be the set for all shipments from $o(k)$ to $d(k)$ with time windows $[e(k),\ l(k)]$, $k \in K$, both contracted and non-contracted. To model a problem in which the shipment volume $w_k, k \in K$ is a random variable, we presume a joint distribution Ψ of commodity sizes is known and that associated with realization $\psi \in \Psi$ is a random vector of commodity sizes $w_k(\psi), k \in K$. Let H be a set of potential hubs. The first-stage decision variables y_h and z_a represent hub location decisions for $h \in H$ and truck moves on the time-expanded arc $a \in \mathcal{A}_\mathcal{T}$, respectively. However, as commodity transportation decisions can be made after their sizes are revealed, the commodity path variables are second-stage variables and parameterized by the random variable. Thus, we have the second stage decision variables $x_{ak}(\psi)$ denoting the amount of flow of commodity $k \in K$ on arc $a \in \mathcal{H}_\mathcal{T} \cup \mathcal{A}_\mathcal{T}$ under realization $\psi \in \Psi$. Given this notation, the problem formulation is as follows:

$$\max_{(z,y)\in\mho} \quad E_\psi\Big[\varphi(z,y,\psi)\Big] - \Big[\sum_{a\in\mathcal{A}_\mathcal{T}} \rho_a z_a + \sum_{h\in H} f_h y_h\Big] \tag{1}$$

where $\mho = \{(z, y : z_a \in Z^+; \forall a \in \mathcal{A}_\mathcal{T}, y_h \in \{0,1\}; \forall h \in H\}$ defines the domain of z and y variables, $\varphi(z,y,\psi)$ represents the profits realized in realization $\psi \in \Psi$ for a given set of variable values (z,y), and E_ψ represents the expectation of those profits with respect to the random variable ψ. Computing the profits $\varphi(z,y,\psi)$ requires solving the second stage problem, subject to hubs and trucks assigned in the first stage.

The time-expanded network, which considers hub location decisions, must account for the time dimension, the variability of consolidation points, and hub arcs to generate possible shipment routes. This results in a problem with a large number of flow variables, x_{ak}, especially in large networks, making it computationally challenging. To address the issue of excessive flow variables, a preprocessing technique is developed, involving two steps of reduction based on both the static and time-expanded networks.

We analyze a directed incomplete graph $G = (N, A)$ in the subsequent case study based on a realistic freight network operation. Every route from an origin $o(k)$ to a destination $d(k)$ will pass through either one or two hubs, depicted as $(o(k), h_1, d(k))$ or $(o(k), h_1, h_2, d(k))$, where (h_1, h_2) belongs to the ordered pairs in $H \times H$. The goal of the first stage is to check and enlist all feasible static paths and eliminate potential cycles within the routes. The second stage iterates through the static paths from the first stage and expands them to include the time dimensions. The second stage considers the possible holding of the shipment at the origins and hubs while respecting the delivery deadline for each commodity. Thus, the preprocessing technique eliminates the need to generate all possible flows across the entire planning horizon for each node and arcs in the network. Only the feasible arcs are generated to reduce the problem size. As an example, Figure 1(a) shows a complete time-expanded network, with all nodes and arcs illustrated, whereas Figure 1(b) shows a subset of those arcs but also direct arcs between two non-adjacent time steps.

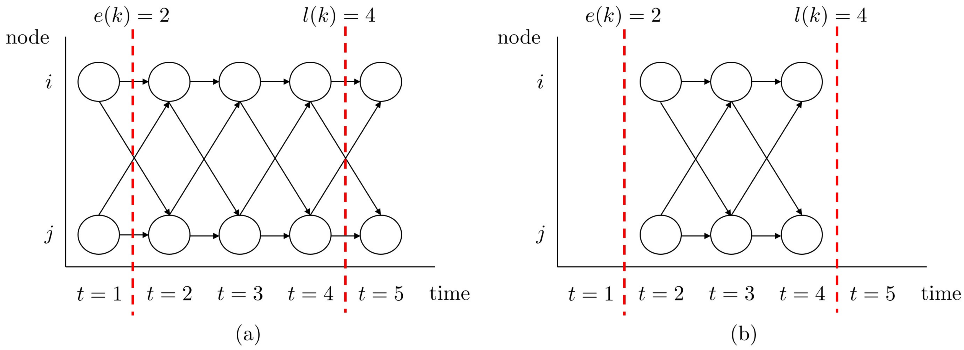

Figure 1 – *The comparison between a complete time expanded network and a reduced network.*

3 SOLUTION APPROACH

Evaluating $E_\psi\Big[\varphi(z, y, \psi)\Big]$ requires assessing a large number of demand scenarios. To approximate this expectation, we employ a Monte Carlo simulation-based algorithm known as the Sample Average Approximation (SAA) scheme (Kleywegt *et al.* 2002, Adulyasak *et al.* 2015, Taherkhani *et al.* 2020, 2021). Specifically, we generate a large set of random demand scenarios $\{\psi_1, \psi_2, \ldots, \psi_S\}$ and approximate the second-stage expectation using the sample average function $E_\psi\Big[\varphi(z, y, \psi)\Big]$. This procedure is repeated for a predetermined number of replications. In each replication, a MILP must be solved; however, due to the problem's large scale, off-the-shelf solvers may struggle to solve them in a reasonable amount of time. To efficiently tackle these MILPs and exploit their combinatorial substructures, we propose using a Benders decomposition (BD) algorithm, where the first-stage location and trucking decisions are handled in the Benders master problem, and the second-stage flow variables are solved in the subproblems. Additionally, a set of valid inequalities is added to the master problem to ensure that enough trucks are allocated at the origins and destinations of contracted shipments.

In the generic SAA, random sampling from a limited set of scenarios may fail to produce representative results. To address this, we employ a machine learning algorithm, specifically the K-means++ algorithm, which clusters the scenarios into a predetermined number of groups (Crainic *et al.* 2014, Marsland 2014). This approach ensures that the selected scenarios are as widely spaced as possible (Arthur & Vassilvitskii 2007). The enhanced SAA then solves the stochastic problem using different sets of selected scenarios.

4 COMPUTATIONAL RESULTS

We test the model and the solution method on instances configured based on a realistic dataset from a freight carrier that serves a region across multiple states in the US. Specifically, we focus on a region shown in Figure 2. The carrier operates within Illinois, Indiana, Michigan, Ohio, and Wisconsin, encompassing 24 demand points.

The total solving time of the problem formulation using the commercial solver CPLEX, the Benders algorithm provided by CPLEX, and the proposed approach are compared in several instances under different levels of uncertainty where it is shown that the proposed method outperforms CPLEX. The algorithm also successfully solved stochastic instances involving up to 100 commodities optimally, using a sample size of 100 and 20 replications. This indicates that the algorithm can deliver optimal solutions for practical-sized instances with considerable sample sizes and multiple replications, thereby proving its effectiveness.

We then evaluate the impact of different settings and the level of uncertainty incorporated into the problem for planning the transportation system at both strategic and tactical levels. These

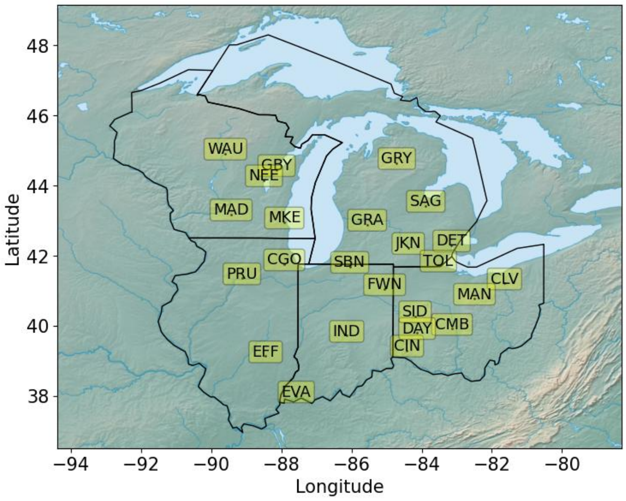

Figure 2 – *Demand points and potential hub locations.*

evaluations are carried out by examining the value of the stochastic solution, performance and structural characteristics of the solutions generated by the proposed model, which include the hub and service network design, truck selection and scheduling, demand fulfillment, consolidation, and net profit. Our experiments underscore the model's capacity to enhance operational efficiency and responsiveness in uncertain environments, providing valuable insights for both practitioners and researchers in freight logistics.

References

Adulyasak, Yossiri, Cordeau, Jean-François, & Jans, Raf. 2015. Benders Decomposition for Production Routing under Demand Uncertainty. *Operations Research*, **63**(4), 851–867.

Alumur, S. A., Campbell, J. F, Contreras, I., Kara, Y. K., Marianov, V., & O'Kelly, M. E. 2021. Perspectives on Modeling Hub Location Problems. *European Journal of Operational Research*, **291**, 1–17.

Arthur, David, & Vassilvitskii, Sergei. 2007. k-means++: the advantages of careful seeding. *Page 1027–1035 of: Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms.* SODA '07. USA: Society for Industrial and Applied Mathematics.

Crainic, T. G. 2000. Service network design in freight transportation. *European Journal of Operational Research*, **122**(2), 272–288.

Crainic, Teodor Gabriel, Hewitt, Mike, & Rei, Walter. 2014. Scenario grouping in a progressive hedging-based meta-heuristic for stochastic network design. *Computers & Operations Research*, **43**, 90–99.

Kleywegt, Anton J., Shapiro, Alexander, & Homem-de-Mello, Tito. 2002. The Sample Average Approximation Method for Stochastic Discrete Optimization. *SIAM Journal on Optimization*, **12**(2), 479–502.

Marsland, Stephen. 2014. *Machine Learning: An Algorithmic Perspective, Second Edition.* 2 edn. New York: Chapman and Hall/CRC.

Taherkhani, Gita, Alumur, Sibel A., & Hosseini, Mojtaba. 2020. Benders Decomposition for the Profit Maximizing Capacitated Hub Location Problem with Multiple Demand Classes. *Transportation Science*, **54**(6), 1446–1470.

Taherkhani, Gita, Alumur, Sibel A., & Hosseini, Mojtaba. 2021. Robust Stochastic Models for Profit-Maximizing Hub Location Problems. *Transportation Science*, **55**(6), 1322–1350.

Yaman, H., & Carello, G. 2005. Solving the Hub Location Problem with Modular Link Capacities. *Computers & Operations Research*, **32**(12), 3227–3245.