Service Network Design for Same-Day Deliveries using Freight on Urban Public Transit

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

November 1, 2024

Keywords: service network design; urban freight; integration; freight on transit; dynamic discretization discovery method

Abstract: We present a tactical planning model for designing a service network for same-day delivery of courier, express, and parcels (CEPs) in an urban environment using public transit. Given a set of origin depots, destination deports, public transit lines with scheduled departures, the goal is to use the transit lines to move CEPs alongside passengers. The problem setting is characterized by time-dependent demand, time windows, limited storage capacity at transit terminals, and synchronization of the fleet and the public transit lines at these terminals operating on different routes. The problem is modeled on a time-expanded network and formulated as an mixed-integer program. We present a dynamic discretization discovery-based exact solution method for the problem. Preliminary computational experiments demonstrate the effectiveness of the mathematical formulation and the proposed solution method. A case study for the city of Montreal, Canada is presented and managerial insights are derived.

1 INTRODUCTION

The same-day delivery market has seen the biggest growth in freight transportation and is expected to grow 36% per year (Deloison & et al, 2020). Courier, Express, and Parcel (CEP) deliveries have grown significantly in the recent years, a direct consequence of short deadlines to be met by the delivery companies and consequential smaller vehicles being deployed for customers. Integration and collaboration between transportation modes is an alternative for the CEP market, with a focus on optimizing the emptiness of freight vehicles delivering products in the city. Public transit agencies possess assets such as vehicles and stations that are highly underutilized, particularly in off-peak periods. The freight on transit (FOT) operations, where the spare capacity of transit vehicles is used to move freight, has been reviewed by (Elbert & Rentschler, 2021), who observes that there is a promising opportunity in combining freights and people's movements, utilizing the same infrastructure and city network.

2 PROBLEM DESCRIPTION AND MODEL

We present a tactical planning model for designing a service network for same-day delivery of CEPs in an urban environment using public transit. Transportation goods and passengers can be carried at the same time on the shared scheduled vehicle, which has a reserved capacity to carry a certain amount of goods based on the average spare capacity of the trips. The transit lines considered in this problem can be seen as any fixed scheduled line, like a bus, a subway, or a tram. It is considered that the characteristics of the scheduled line are known, like the average boarding and disembarking time at each station, the travel time between each point, and the associated costs for operating this line. The freight demand originates and is destined for warehouses (hubs), where small vehicles are entitled to route and distribute the goods for the last mile. We are interested in the problem of connecting these hubs and incorporating partial or the totality of the consolidated demand for trips in the public transit lines, in order to reduce the number of trucks dispatched between warehouses. The goal is to design an integrated network of hubs and selected transit routes that can accommodate the freight demand during the planning horizon and therefore minimize the costs of the operations. A novel feature of this problem is the incorporation of existing shared services (transit routes) into a typical service network design (SND).

Model Formulation: Consider a network with $D = (N, A)$, where N is the set of nodes and A is the set of arcs. Let $k \in K$ denote a set of commodities to be served, characterized by an origin (o_k), a destination (d_k), availability at its origin at time (e_k), due at its destination (l_k), revenue for service (q_k) and demand units (w_k). Transportation costs associated with the arc (i, j) are denoted by c_{ij}, the fixed cost for accessing the transit routes are denoted by f_r, and the operational costs of transit lines are denoted by c_p. Vehicles on arc (i, j) have a capacity of Q_{ij} whereas Q'_p denotes the capacity of the transit line p. The transit line has the departures between nodes given by the parameter v_r, which indicates the OD pair and the departure and arrival of the transit vehicles. We use a time-expanded network to model the service network design problem of the FOT operations by creating a set of discrete time stamps, $t \in T$, over the planning horizon. The timed node set $\mathcal{N}$ has a node (i, t) for each node $i \in N$ and $t \in T_i$. Arcs $(i, t), (j, \bar{t})$ models the possibility of sending freight from origin i to destination j with the loading of freight starting at i at time t and the unloading of freight finishing at j at time $\bar{t}$. Consider that travel time in arcs is represented by τ. Let $x_{ij}^{k,t\bar{t}}$ be a binary variable representing whether a commodity k $(k \in K)$ is transported on the arc $(i, t)(j, \bar{t})$, $y_{ij}^{t\bar{t}}$ a non-negative integer variable represent the number of vehicles (trucks) used on arc $(i, t)(j, \bar{t})$, z_r a binary variable that represents if transit line $r \in R$ is used (giving access to predetermined trips v_r), and finally, s_k a binary variable that represents if commodity $k \in K$ is served on the network. The service network design problem is presented below:

$$[\text{M}] : \max \sum_{k \in K} q_k s_k - \sum_{(i,t),(j,\bar{t}) \in \mathcal{A} \setminus \mathcal{A}_B} c_{ij} y_{ij}^{t\bar{t}} - \sum_{r \in R} (f_r z_r + \sum_{p \in P_r} c_p v_p) \tag{1}$$

$$\text{s.t.} \quad \sum_{(i,t),(j,\bar{t}) \in \mathcal{A}^k} x_{ij}^{k,t\bar{t}} - \sum_{(j,\bar{t}),(i,t) \in \mathcal{A}^k} x_{ji}^{k,t\bar{t}} = \begin{cases} s_k & \text{for}(i,t) = (o_k, e_k) \\ -s_k & \text{for}(i,t) = (d_k, l_k) \ \forall \ (i,t) \in \mathcal{N}, k \in K \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

$$\sum_{k \in K_{ijt\bar{t}}} w_k x_{ij}^{k,t\bar{t}} \leq Q_{ij} y_{ij}^{t\bar{t}} \qquad \forall \quad ((i,j),(j,\bar{t}) \in \mathcal{A} \setminus \mathcal{A}_B, k \in K \tag{3}$$

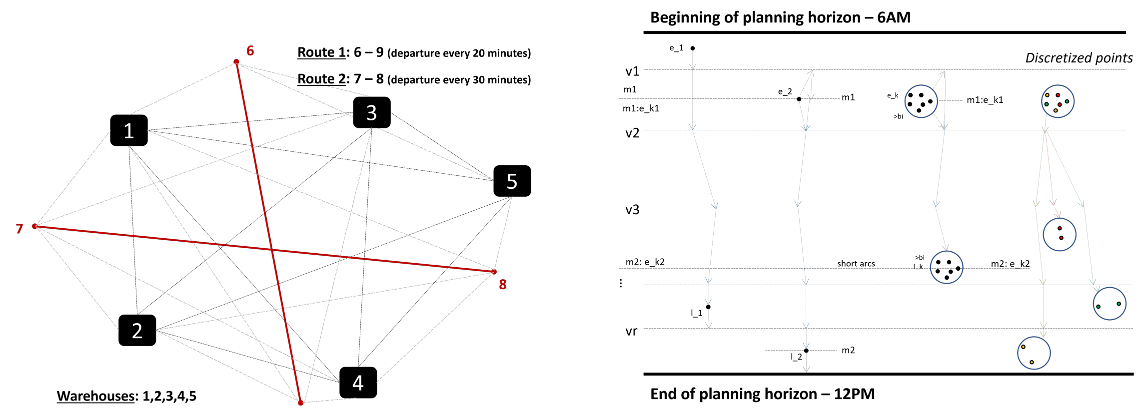
$$x_{o_r d_r}^{k,(t_s - \overline{\tau_l})(t_r + \overline{\tau_u})} \leq v_p \qquad \forall \quad r \in R \tag{4}$$

$$z_r \geq v_p \qquad \forall r \in R, p \in P_r \tag{5}$$

$$\sum_{k \in K_{ijt\bar{t}}} w_k x_{ji}^{k,t\bar{t}} \leq Q_p \qquad \forall \quad ((i,t),(j,\bar{t}) \in \mathcal{A}_B, k \in K \tag{6}$$

$$x_{ij}^{k,t\bar{t}} \in \{0,1\}, \quad z_r \in \{0,1\}, \quad s_k \in \{0,1\}, \quad y_{ij}^{t\bar{t}} \in Z^+ \tag{7}$$

The objective function (1) maximizes the profit for supplying each commodity minus the travel costs using trucks and the fixed and variable costs for utilizing the transit routes. The constraints represent flow conservation (2), selection of the routes r that will be inserted in the network (4-5), the capacity of the moving arcs, representing the vehicular capacity between nodes (Q_{ij}) for the time-expanded network excluding the transit lines $(\mathcal{A} \setminus \mathcal{A}_B)$ (3), and the capacity of the transit lines itself (Q_p) over the time-expanded network $\mathcal{A}_B$ (5).



(a) *Network example* (b) *Discretization DDD*

Figure 1 – *Physical network (a) and Time-expanded discretization (b)*

3 METHODOLOGY

The problem is $\mathcal{NP}$-hard, and formulated over a time-expanded network, making it very difficult to be solved for large periods of time. Detailed solutions are required for such a problem, discretizing the network in a granular manner may lead the model to intractability and infeasibility, given the exponential number of variables. We use a dynamic discretization discovery algorithm to solve the problem. The DDD algorithm relies on relaxing the time-expanded network by iteratively building and updating a partially time-expanded network (PTEN) with timed nodes and arcs to yield a lower bound on the value of the problem and refine the PTEN to obtain an upper bound until converge to an exact solution. The algorithm was introduced by Boland *et al.* (2017) for a Travelling Salesman Problem with Time Windows (TSPTW) and the Continuous Service Network Design Problem, both formulated over a time-expanded network (TEN).

The time-expanded network is formulated by exploiting the properties of the problem, in an attempt to reduce the number of expanded nodes, arcs, and variables associated with them without compromising the optimum solution. For the expanded network of this problem, the departure and arrival times of the scheduled lines are selected and evaluated for each route, defining new time points for the time-expanded network. The commodities are clustered accordingly to their similarities in availability (time windows) and allocated in the network by solving a scheduling problem. This step also helps reduce the number of variables associated with commodities k. If this solution is shown as feasible, the time points are then inserted in the problem, along with the selected routes, and the problem is solved for an upper bound. The infeasibility of the consolidated commodities is then checked by defining the capacities of the nodes. If the commodities are inserted at a time earlier than the discretization, they will be carried until the next time point. Each discretization is a tighter version of the previous one and converges to the optimal solution. The details are shown in Figure 1.b.

For computing upper and lower bounds for the PTEN, we define the earliest arrival (ϵ^r_k) and the latest departure properties (ω^r_k). The earliest arrival correspond to the state that can be reached when the commodity is immediately dispatched when it is inserted into the PTEN, with no holding arcs being attributed to it, since no storage is necessary. This state correspond to higher costs of shipment, since the consolidation is not defined and the direct shipments are assumed to be realized on demand. The latest departure allows consolidation of commodities and considered the dispatch only on the last moment to reach its established deadline, considering the travel time for the path. The holding arcs in the origin node are affected, since the commodities have to wait for being dispatched until the last moment. When including the transit routes into the above properties, we define as well the target earliest arrival ($^{r,P_r}_k$) and the target latest departure ($^{r,P_r}_k$). These values are used to define the trips of the transit time that can be

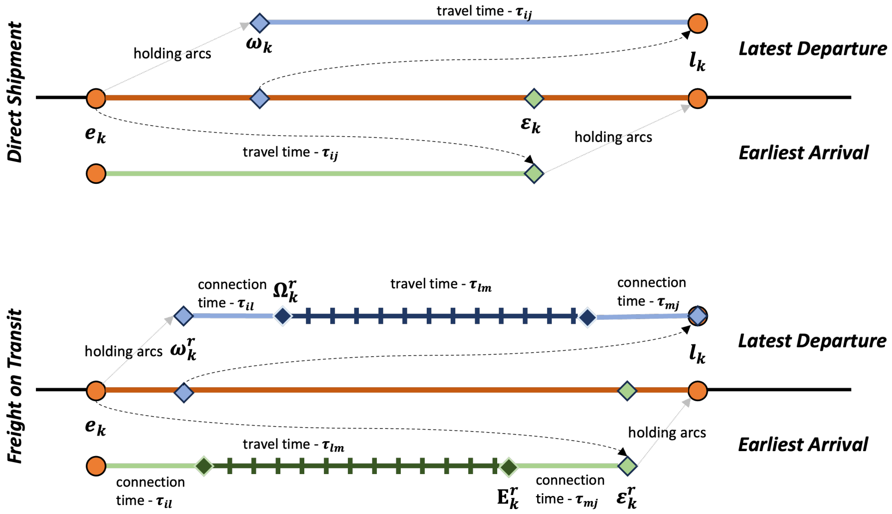


Figure 2 – *Representation of the earliest arrival and latest departure in a timed arc*

associated in a path that utilizes these arcs. For the target earliest arrival, it correspond to the time that the transit arrives at the terminal node m, while for the target latest departure it corresponds to the time the transit departures from the terminal l. Both of them consider the connection time between terminals and depots. Figure 2 illustrates these scenarios.

Preliminary Results: We benchmark the effectiveness of our model by running small instances on CPLEX (DOcplex.CP - 11th Gen Intel(R) Core(TM) i7-11390H @ 3.40GHz, 16 Gb RAM). Only small instances can be dealt with under this formulation, given the size of the time-expanded networks and the number of constraints for the problem. For a time-expanded network with 5 nodes, 2 routes, and 36 time periods of 10-minute intervals (considering a planning horizon from 06h - 12h), we have a total of 8,322,449 variables to evaluate. By running an instance with 100 commodities, the commercial solver took 86 minutes to run, while the DDD provided the same result within 3.05 minutes, both providing the same optimal result. Figure 1.a. displays the tested network.

4 CONCLUSION

The freight on public transit has gained attention over the past years, but tactical and strategic planning models are still in the early stage. The size of the time-expanded networks makes them very difficult to be solved using traditional methods. We proposed a novel approach by developing a service network design model for this problem. A solution method based on dynamic discretization discovery (DDD) is presented to solve this problem. Our algorithm iteratively refines the time points to be inserted in the time-expanded network by evaluating the scheduled lines of each route. Our preliminary results show that this method is competitive for small networks against traditional solvers, by significantly reducing the solving time. Further testing on larger instances is the next step for this problem.

References

Boland, Natashia, Hewitt, Mike, Marshall, Luke, & Savelsbergh, Martin. 2017. The Continuous-Time Service Network Design Problem. *Operations Research*, **65**(5), 1303–1321.

Deloison, T, & et al. 2020. *The Future of the Last-Mile Ecosystem.* Tech. rept. World Economic Forum, Geneva.

Elbert, Ralf, & Rentschler, Johannes. 2021. Freight on urban public transportation: A systematic literature review. *Research in Transportation Business & Management*, June, 100679.