Redesigning Bus Networks at Scale

J. Carlos Martínez Mori, Stavros Roditis
CU Denver, {carlos.martinezmori, stavros.roditis}@ucdenver.edu

Saigopal Rangaraj
Greater Denver Transit (GDT), rsaigopal@outlook.com

1. Introduction

The fundamental challenge of any transit agency is to design a service plan that simultaneously serves a broad range of people, with different origins, destinations, and trip purposes. However, there is no single answer for what exactly constitutes an ideal service plan. At its core, the main product of transit (and more generally of transportation as a public good) is accessibility through ease of movement (Martens 2012). Nevertheless, its provision is complicated by the need to balance heterogeneous transportation needs alongside geometric, physical, regulatory, economic, and political constraints. The role of a transit planner is to elicit and navigate this information to design a service plan that reflects the values of the communities it serves.

Our goal in this work is to develop mathematical tools to assist transit planners in this process, specifically in the context of redesigning bus networks. We combine mixed-integer programming-based modeling with established principles of transit planning (Walker 2024) to address a prevalent difficulty in this space: scale. Specifically, prior works have focused on column-generation approaches to produce a "rich" set of *candidate lines*, i.e., geospatial traces of potential bus routes. This is a provably hard problem (Borndörfer et al. 2007) that involves multi-commodity flow formulations—typically on large graphs representing road networks with thousands of nodes. As a result, prior works have seen their computational limit in the order of hundreds or at most a few thousand origin-destination (OD) pairs (Borndörfer et al. 2007, Bertsimas et al. 2021); made possible only after pruning away those pairs that experience (or are expected to experience) little demand. However, this approach makes the at times implicit assumption that *the* goal of a transit system is to maximize ridership. And while maximizing ridership is indeed a principal goal, in practice it competes with the often conflicting goal of maximizing geographical coverage (Walker 2024). Pruning away OD pairs with little yet non-zero demand invalidates the coverage goal, whereas any transit agency that adopts a role as a social service is not in a position to do so. To illustrate, a graph with 1,000 nodes has 499,500 unordered OD pairs, so restricting to the top 1,000 OD pairs only accounts for 0.2% of all possible distinct transportation needs.

Now, given a "rich" enough set of candidate lines, the selection problem is comparatively easy. Here is where our key insight appears: in practice, transit planning involves generating candidate lines through the current service plan, public engagement, and the consultants' experiential knowledge. In other words, riders and seasoned planners know which lines may work well individually, based on factors that may or may not be captured by automated line-generation techniques. Instead, their main bottleneck is to identify a subset selection that works well as a system.

Therefore, in this work we pivot from generating a "rich" set of candidate lines to generating a "rich" selection based on principles of transit planning (Walker 2024). We assume we are given a set $\mathcal{L}$ of candidate lines, a finite set $\mathcal{H}$ of headway options, and the current service plan $\mathcal{C} \subseteq \mathcal{L} \times \mathcal{H}$. The goal is to propose a service plan $\mathcal{P} \subseteq \mathcal{L} \times \mathcal{H}$ that improves on $\mathcal{C}$ subject to various practical considerations. Incidentally, this re-focus on leveraging the given $\mathcal{L}$ rather than generating it allows us to reach a practical scale, with hundreds of thousands of OD pairs.

2. Problem Formulation

Our main variables are $x_\ell^h \in \{0,1\}$: set to 1 if we operate line $\ell \in \mathcal{L}$ at headway $h \in \mathcal{H}$. From these we derive two further sets of variables: $0 \le u_{od} \le 1$ and $y_{od} \in \{0,1\}$, which quantify the level and existence of service (conceptualized as $u_{od} \ge \frac{1}{\max \mathcal{H}}$) for an unordered OD pair $(o,d) \in \mathcal{D}$. We also introduce three auxiliary sets: $\mathcal{L}_{od} \subseteq \mathcal{L}$ is the subset of lines that can serve (o,d) (e.g., based on walking accessibility, on-board trip duration), $\mathcal{T} \subseteq \mathcal{L} \times \mathcal{L}$ is the subset of unordered pairs of lines that form a transfer candidate (e.g., pairs that intersect), and $\mathcal{T}_{od} \subseteq \mathcal{T}$ is the subset of transfer candidates that can serve (o,d) given that neither of their constituent lines can do so individually.

$$\text{maximize} \begin{bmatrix} \sum_{(o,d)\in\mathcal{D}} \rho_{od} u_{od} \\ \sum_{(o,d)\in\mathcal{D}} y_{od} \end{bmatrix} \tag{1a}$$

$$\text{subject to} \sum_{h\in\mathcal{H}} x_\ell^h \le 1, \qquad \forall \ell \in \mathcal{L} \tag{1b}$$

$$u_{od} \le y_{od} \le 1 - \frac{1}{\max \mathcal{H}} + u_{od}, \qquad \forall (o,d) \in \mathcal{D} \tag{1c}$$

$$z_t \le \min\left\{ \sum_{h\in\mathcal{H}} \frac{1}{h} x_{\ell_1}^h, \sum_{h\in\mathcal{H}} \frac{1}{h} x_{\ell_2}^h, 2 - \sum_{h\in\mathcal{H}: h>h^*} x_{\ell_1}^h + x_{\ell_2}^h \right\}, \quad \forall (\ell_1,\ell_2) \in \mathcal{T} \tag{1d}$$

$$\alpha \cdot \left(\sum_{(\ell,h)\in\mathcal{C}:\ell\in\mathcal{L}_{od}} \frac{1}{h} + \sum_{t\in\mathcal{T}_{od}} \chi_t(\mathcal{C}) \right) \le u_{od} \le \min\left\{ \frac{1}{\min \mathcal{H}}, \sum_{\ell\in\mathcal{L}_{od}} \sum_{h\in\mathcal{H}} \frac{1}{h} x_\ell^h + \sum_{t\in\mathcal{T}_{od}} z_t \right\}, \qquad \forall (o,d) \in \mathcal{D} \tag{1e}$$

$$\sum_{\ell\in\mathcal{L}} \sum_{h\in\mathcal{H}} c_\ell^h x_\ell^h \le \sum_{(\ell,h)\in\mathcal{C}} c_\ell^h, \quad \sum_{\ell\in\mathcal{L}} \sum_{h\in\mathcal{H}} x_\ell^h \le |\mathcal{C}| \tag{1f}$$

$$x \in \{0,1\}^{\mathcal{L}\times\mathcal{H}}, \quad u \in \mathbb{R}_{\ge 0}^{\mathcal{D}}, \quad y \in \{0,1\}_{\ge 0}^{\mathcal{D}}, \quad z \in \mathbb{R}_{\ge 0}^{\mathcal{T}}. \tag{1g}$$

The objectives in (1a) are to maximize ridership as the weighted sum $\sum_{(o,d)\in\mathcal{D}} \rho_{od} u_{od}$ and coverage as the unweighted sum $\sum_{(o,d)\in\mathcal{D}} y_{od}$. (1b) ensures a physically coherent selection while (1c) implements our conceptualization of coverage. Based on the planning literature (Walker 2024), we also design constraints that enforce various practical considerations. (1d) quantifies the frequency at which a transfer is offered as long as at least one of its constituent lines is operated at high frequency, i.e., with headway under some choice of $h^* > 0$. The right-hand side of (1e) models the diminishing marginal value of service frequency. Meanwhile, its left-hand side ensures the selection preserves the current level of service across OD pairs within some choice of incentive-compatibility (IC) factor $0 \le \alpha \le 1$, where $\chi_t(\mathcal{C})$ is a constant computed as in (1d) for the current plan. Finally, (1f) ensures the selection respects the current budget and operational complexity.

3. Experimental Results

We implement (1) for Ithaca, NY due to its size and our familiarity with its bus service. Despite Ithaca being a small city, its road network retrieved through `osmnx` (Boeing 2017) contains 37,022 nodes. Instead of considering all nodes, we find a hitting set of 647 nodes such that each node is within 500 meters (i.e., walking distance) of the hitting set. After discarding any OD pairs that would involve a walking trip, we obtain 198,879 pairs. We retrieve the current service plan from the local operator's publicly available `GTFS` to obtain 25 lines. We complement these with 225 randomly generated lines to obtain a candidate set with $|\mathcal{L}| = 250$. Finally, we let $\mathcal{H} = \{5, 10, 15, 20, 30, 45, 60\}$ in minutes, $h^* = 15$, c_ℓ^h be the length of line ℓ divided by h, and $\rho_{od} = \rho_o \rho_d$, where ρ_u is one plus the number of amenities, shops, and offices near u (retrieved through `osmnx`). This models service priority based on the creation of access to goods and services. We ran our experiments on a `MacBook Air` with an `Apple M2` chip and 8 GB of memory, with a 1-hour solver timeout.

The current plan is shown in Figure 1 (left); the lines are color-coded, their width is proportional to their frequency, and the size of each node u is proportional to $\log \rho_u$. If we set the IC factor to $\alpha = 0$, the ridership plan in Figure 1 (center) drops service for certain peripheral regions and augments it for the urban core. Conversely, the coverage plan in Figure 1 (right) maintains service for peripheral regions while installing an `X`-shaped high-frequency corridor to enable transfers. We

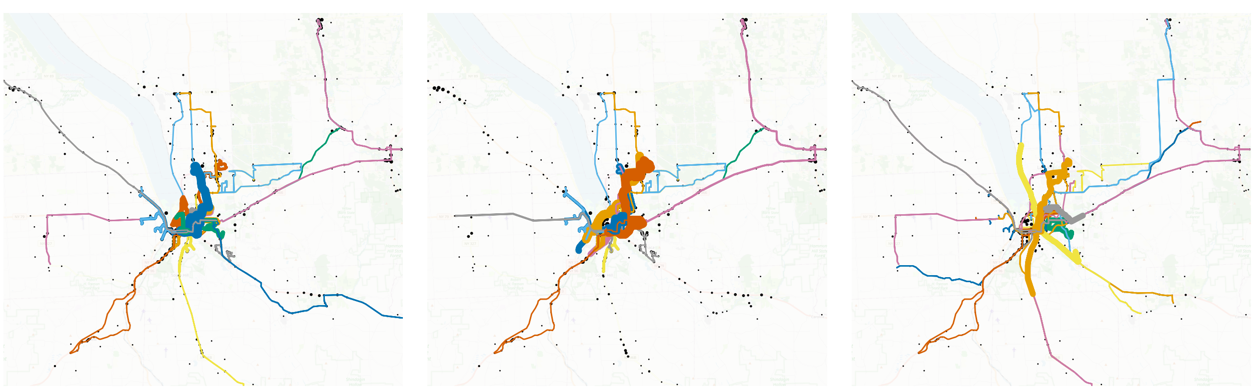

Figure 1 Current service plan (left), ridership service plan (center), and coverage service plan (right).

further illustrate this in Figure 2. The ridership plan dominates the current plan regardless of whether one considers transfers. Without transfers, the coverage plan decreases the level of service for most well-served OD pairs. However, with transfers, it alleviates this effect while more than doubling the number of pairs with basic service. Setting $\alpha = 0.5$ repeats the trends with decreased impact, whereas $\alpha = 1$ leaves no room for optimization: the optimal solution is the current plan.

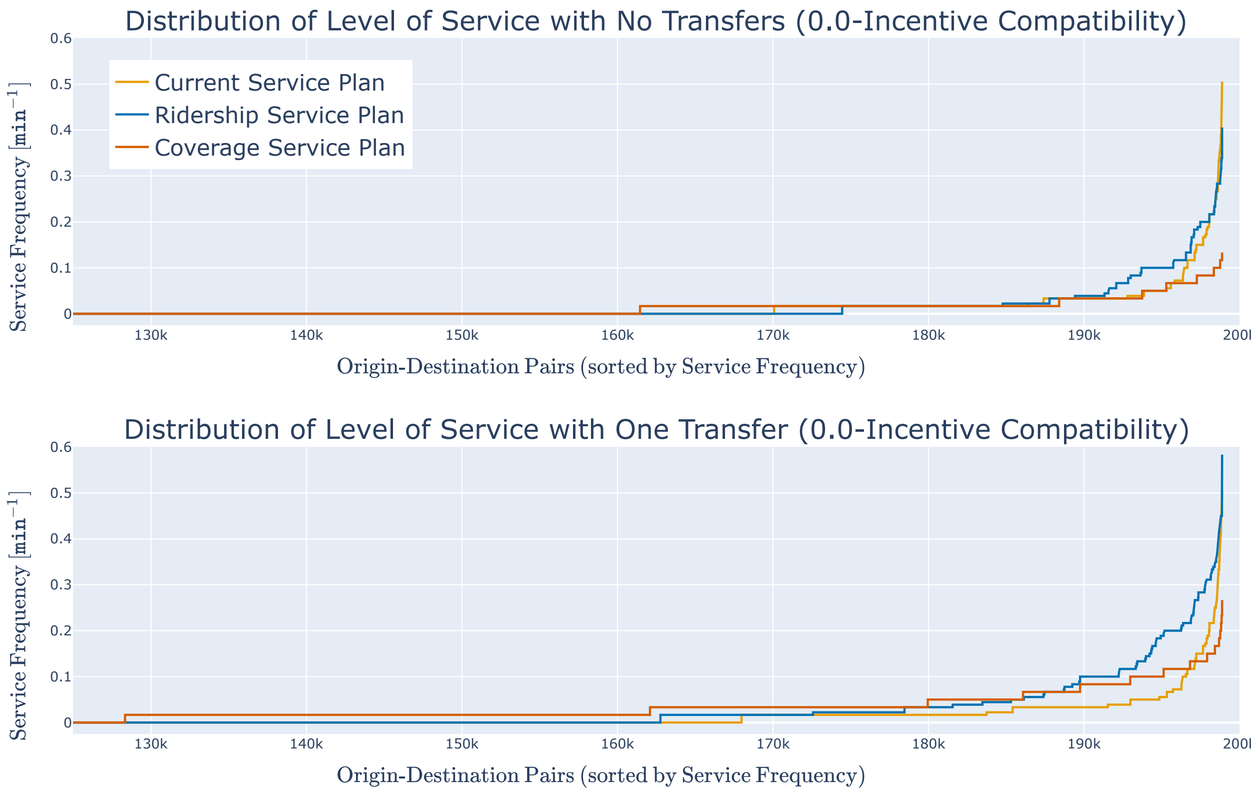

Figure 2 **Lorenz curves of level of service across OD pairs, without (top) and with (bottom) transfers.**

4. Conclusion

Our next goal is to replicate this work on the Denver, CO area to leverage our local knowledge (e.g., to better inform $\mathcal{L}$) and to push the limit on scale. This would be timely given the local operator's ongoing comprehensive operational analysis (RTD 2025) and GDT's advocacy mission.

References

D. Bertsimas, Y. S. Ng, and J. Yan. Data-driven transit network design at scale. *Operations Research*, 69 (4):1118–1133, 2021.

G. Boeing. Osmnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks. *Computers, environment and urban systems*, 65:126–139, 2017.

R. Borndörfer, M. Grötschel, and M. E. Pfetsch. A column-generation approach to line planning in public transport. *Transportation Science*, 41(1):123–132, 2007.

K. Martens. Justice in transport as justice in accessibility: Applying Walzer's "spheres of justice" to the transport sector. *Transportation*, 39:1035–1053, 2012.

RTD. 2026 Comprehensive Operational Analysis 10% SBE Goal, 2025. URL `https://procurement.opengov.com/portal/rtd-denver/projects/209164`.

J. Walker. *Human Transit: How Clearer Thinking about Public Transit Can Enrich Our Communities and Our Lives*. Island Press, Washington, DC, 2 edition, 2024.