Timetable Optimization of Mixed Transit Networks with User Heterogeneity

Arantxa De La Hoz*[1] and Yu Jiang, Jesper Ingvardson, Guido Cantelmo[2]

[1]PhD Student, DTU Management, Technical University of Denmark, Denmark
[2]Associate Professor, Technical University of Denmark, Denmark

Short summary

This study proposes a methodological framework for timetable optimization in mixed transit networks combining frequency-based and schedule-based services. A linearized mixed-integer linear programming (MILP) model is developed, accounting for hard capacity constraints, user heterogeneity, and differences in passenger behaviour across service types. The model is applied to a large-scale case study of the Copenhagen transit network, evaluating scenarios with varying user composition and cost weights. Results show that timetable settings are mainly driven by trade-offs between operational and passenger costs, while equity gaps persist, with vulnerable users consistently experiencing higher generalized costs.
Keywords: Equity, Heterogeneity, Mixed Transit Network, Route choice modelling, Timetable Optimization.

1 Introduction

Most European cities rely on mixed transit networks that combine two types of services: schedule-based and frequency-based services. However, in the context of timetable optimization, most studies focus on optimizing only one type of service at a time. Schedule-based studies typically emphasize synchronization at transfer points (Ibarra-Rojas & Rios-Solis (2012)), while purely frequency-based optimization is often integrated with line planning (Szeto & Jiang (2014)).
A limited number of studies consider both schedule-based and frequency-based services simultaneously. Among these, Jiang et al. (2022) proposed an integrated optimization model for mixed transit networks. The objective function aimed to minimize both operational costs and total passenger generalized travel cost. However, several limitations remain, including scalability challenges, and the lack of hard capacity constraints.
The literature shows that travellers exhibit different behaviours when using schedule and frequency-based services. For low-frequency (schedule-based) services, they tend to consult timetables and arrive in a planned, non-random manner. Conversely, for high-frequency services, arrivals are mostly random. Ingvardson et al. (2018) showed this dynamic can be modelled using a mixture of uniform (random) and beta (planned) distributions, showing that the share of random arrivals decreases from 57% at 5-minute headways to 10% at 30 minutes. Consequently, the common assumption that average waiting time equals half the headway does not hold for the first station of the journey.
Another issue is that, while several optimization studies assume homogeneous demand, demand is heterogeneous and diverse by nature. This may lead to overlooking the needs of sub-populations, particularly vulnerable users such as the elderly and people with mobility disabilities (PwMD). Liu et al. (2023) showed that PwMD display different travel behaviours, including distinct stop usage, higher sensitivity to weather, and a preference for off-peak travel. Despite this fact, existing evidence suggests that vulnerable users perceive transfers differently: while regular users are willing to accept transfers with relatively small reductions in travel time or cost (16–33%), PwMD require substantially larger reductions—23% when already transferring and up to 40.6% when switching from a direct route—to accept additional transfers (Chowdhury et al. (2015), Park & Chowdhury (2022)). Incorporating heterogeneity into timetable optimization however requires models that capture these behavioural differences. However, to the best of our knowledge, such models remain underdeveloped.
Therefore, this study aims to fill all the above-mentioned gaps. First, we propose a MILP model for timetable optimization that integrates both schedule-based and frequency-based services, minimizing operational and passenger costs. Second, user heterogeneity is incorporated through two

user groups (regular and vulnerable), each with distinct perception parameters in the route choice model. Third, passenger behaviour is differentiated by service type, with waiting time at the first boarding stop modelled as a function of headway, while transfer waiting time is approximated as half the headway. Finally, the model accounts for hard capacity constraints and combines discrete and continuous representations of headways for schedule-based and frequency-based services, respectively.

2 Methodology

Model Formulation

Let W be the set of origin–destination pairs, where each $w \in W$ is defined by origin o^w and destination d^w. Let $G = \{g1, g2\}$ be the set of user groups, where $g1$ and $g2$ denote regular and vulnerable users, respectively. The model is formulated as a minimization problem composed of three main components: operational cost C^{op}, passenger cost C^{pass}, and crowding cost C^{crowd} Equation (1). The weights ρ_{op}, ρ_{pass}, and ρ_{crowd} are used to measure their relative impact on the objective function. The decision variables of the model are the headway H_l of each transit line $l \in L$ and the start time $X_{l,d}$ of the service for line l in direction d.

$$\min Z = \rho_{op} \cdot C^{op} + \rho_{pass} \cdot C^{pass} + \rho_{crowd} \cdot C^{crowd} \tag{1}$$

subject to:

$$t_{l,d,k,s} = X_{l,d} + kH_l, \quad \forall l \in L,\ d \in D_l, s \in S_l,\ k = 0, 1, 2, \dots \tag{2}$$

$$X_{l,d} \leq H_l \tag{3}$$

$$X_{l,d} \leq UB_l \tag{4}$$

$$N_l H_l = 60, \quad \forall l \in L \tag{5}$$

$$\sum_{l \in L_{s,d}} N_l \leq C^{corr}_{s,d}, \quad \forall s \in S \tag{6}$$

$$|t_{A,d,k,s} - t_{B,d,m,s}| \geq \Delta, \quad \forall A, B \in L,\ A \neq B \tag{7}$$

$$P_r^{w,g} = \frac{\exp(-\theta\, GC_r^{w,g})}{\sum_{r' \in R_w} \exp(-\theta\, GC_{r'}^{w,g})}, \quad \forall w \in W,\ r \in R_w,\ g \in G \tag{8}$$

$$P^g_{l,i,j} = \sum_{w \in W} \sum_{r \in R_w} D^{w,g}\, P_r^{w,g}\, I_{r,l,i,j}, \quad \forall l \in L,\ (i,j) \in S_l,\ g \in G \tag{9}$$

$$P_{l,i,j} = \sum_{g \in G} P^g_{l,i,j}, \quad \forall l \in L,\ (i,j) \in S_l \tag{10}$$

$$P_{l,i,j} \leq C^{tot}_l\, N_l, \quad \forall l \in L,\ (i,j) \in S_l \tag{11}$$

$$E_{l,i,j} = \max\left\{0,\ P_{l,i,j} - N_l C^{seat}_l\right\}, \quad \forall l \in L,\ (i,j) \in S_l \tag{12}$$

$$GC_r^{w,g} = \alpha^g\, WT_r + \sum_{l \in L_r} \beta^g_l\, TT_{r,l} + \gamma^g\, NT_r, \quad \forall w \in W,\ r \in R_w,\ g \in G \tag{13}$$

$$C^{op} = \sum_{l \in L} C^{op}_l \cdot N_l \tag{14}$$

$$C^{pass} = \sum_{g \in G} \sum_{w \in W} \sum_{r \in R_w} D^{w,g}\, GC_r^{w,g}\, P_r^{w,g} \tag{15}$$

$$C^{crowd} = \delta \sum_{l \in L} \sum_{(i,j) \in S_l} E_{l,i,j} \tag{16}$$

$$H_l \in \mathbb{R}_+, \qquad \forall l \in L_{fr} \tag{17}$$

$$H_l \in \mathbb{Z}_+, \qquad \forall l \in L_{sch} \tag{18}$$

$$X_{l,d} \in \mathbb{R}_+, \qquad \forall l \in L,\ d \in D_l \tag{19}$$

The model considers two types of services: schedule-based lines $l \in L_{sch}$ and frequency-based lines $l \in L_{fr}$, where $L_{fr} \cup L_{sch} = L$. Schedule-based lines are defined using headways H_l that are replicated within an hour, taking only certain discrete values (e.g., 10, 15, or 20 minutes). In contrast, headway H_l variables for frequency-based lines are modelled as a continuous variable with lower LB_l and upper bounds UB_l. The transit network consists of a set of lines $l \in L$, where each line is represented as an ordered sequence of stops $s \in S_l$. The structure of the timetable is defined in Equation (2), where the term $t_{l,d,k,s}$ represents the departure time of the k-th trip of line l in direction d from its first station $s = 1 \in S_l$ during the simulation period.
Constraints (3) and (4) ensure that the first vehicle of each line departs before the minimum value between the upper bound UB_l and the line headway H_l.
Equation (5) represents the inverse relationship between the number of vehicles per hour and the headway H_l. Similarly, constraint (6) ensures that the capacity of a corridor at stop s, does not exceed the maximum number of vehicles per direction d per hour. Constraint (7) guarantees a minimum safety separation time Δ between the arrivals of vehicles from different lines at a given station s, preventing from scheduling simultaneous arrivals.
Passenger route choice in the transit network is modelled using a simple Multinomial Logit Model (MNL) formulation. The probability of choosing route $r \in R_w$ for user group g, denoted by $P_r^{w,g}$, is defined in Equation (8) as a function of the generalized cost $GC_r^{w,g}$. Equations (9) (10) define the passenger load on each segment (i,j) of line l, where $(i,j) \in S_l$. This expression aggregates passenger demand by route and user group, considering demand $D^{w,g}$, route choice probability $P_r^{w,g}$, and the binary indicator $I_{r,l,i,j}$, equal to 1 if route r uses segment (i,j) of line l, and 0 otherwise. Similarly, constraint (11) represents the hard capacity constraint, ensuring that the total hourly passenger flow on each segment (i,j) does not exceed the total hourly capacity $C_l^{tot} N_l$ for line l.

Crowding Cost

Equations (12) and (16) define the number of standing passengers and the corresponding crowding cost. In particular, $E_{l,i,j}$ denotes the number of standing passengers on line l between stops (i,j). The total crowding cost is obtained by aggregating this excess across all lines and segments and multiplying it by the crowding parameter δ.

Operational Cost

Equation (14) defines the total operational cost C^{op} of the system. For each line l, the number of trips N_l is multiplied by the corresponding operating cost per trip C_l^{op}. The operational line costs C_l^{op} are given inputs.

Route Choice Model

Equations (13) and (15) present the generalized cost of a route and the passenger cost respectively. The generalized cost of each route is defined for each user group based on in-vehicle travel time, waiting time, and transfers (Nielsen et al. (2019)). Crowding is not directly incorporated into the route cost to avoid fixed-point formulation.
While in-vehicle travel time and transfers are model inputs, waiting time depends on headway. To capture differences between schedule-based and frequency-based services, waiting time is modelled separately for the first boarding and subsequent lines using piecewise linear formulations with different breakpoints.

Linearization

The non-linear constraints (5) (8) (7) must be linearized in order to solve the MILP model by using a conventional linear optimization solver such as CPLEX.

The bilinear constraint in equation (5) defining an hyperbolic relationship between headway H_l and hourly vehicles N_l, is linearized using a piecewise formulation over a relevant range of headway values (Vielma (2015)).
Constraint (7) is an absolute value constraint that is linearized using a Big-M formulation (Vielma (2015)).
The MNL constraint (8) of route choice probability $P_r^{w,g}$ is linearized through logarithmic transformations and introduction of auxiliary variables (Jiang et al. (2022)).

3 Results and discussion

Case Study: Copenhagen

The network includes seven suburban rail and four metro bidirectional lines. An hourly demand of 24,420 trips was derived from smart card data for 2023, covering 970 OD pairs, of which only 17% have two alternative routes. Travel times were obtained from DSB and Google Maps websites for rail lines and metro lines respectively. Indicative operational costs are assumed proportional to line length. Perception factors for route choice are reported in Table 1 (source: Jiang et al. (2022)).

Table 1: Perception factors for route choice

Perception Factor	**Regular Users**	**Vulnerable Users**
α^g (Wait time)	0.575	0.575
β^g_{STrain} (In-vehicle time)	0.430	0.430
β^g_{Metro} (In-vehicle time)	0.420	0.420
γ^g (Transfers)	2.150	4.300 (*)

Note: (*) Parameter derived from a sensitivity analysis.

Scenarios

A total of nine scenarios were analysed (Table 2). Passenger heterogeneity was varied across three cases: 100% regular users, a 95–5% split between regular and vulnerable users, and a 90–10% split. These scenarios are referred to as homogeneous demand, low heterogeneity, and moderate heterogeneity, respectively. Additionally, three weights of operational cost were considered: parameter ρ_{op} was set to 1.50, 0.80, and 0.50, corresponding to operational-oriented, balanced, and passenger-oriented scenarios, respectively.
Model complexity increases with user groups and route alternatives, from 12,626 variables and 15,758 constraints for homogeneous demand, to 15,950 variables and 20,534 constraints for heterogeneous demand. Optimality gaps across scenarios remain below 0.60% in all but two cases (up to 2.20%). CPU times were below 2,000 seconds, except for one case reaching 14,300 seconds.

Table 2: Experimental scenarios combining demand split and objective-function weights

Scenario	**G1 share**	**G2 share**	ρ_{pass}	ρ_{op}	ρ_{crowd}	**Interpretation**
S1	100%	0%	1.0	0.8	1.0	Balanced, homogeneous demand
S2	95%	5%	1.0	0.8	1.0	Balanced, low heterogeneity
S3	90%	10%	1.0	0.8	1.0	Balanced, moderate heterogeneity
S4	100%	0%	1.0	0.5	1.0	Passenger-oriented, homogeneous demand
S5	95%	5%	1.0	0.5	1.0	Passenger-oriented, low heterogeneity
S6	90%	10%	1.0	0.5	1.0	Passenger-oriented, moderate heterogeneity
S7	100%	0%	1.0	1.5	1.0	Operator-oriented, homogeneous demand
S8	95%	5%	1.0	1.5	1.0	Operator-oriented, low heterogeneity
S9	90%	10%	1.0	1.5	1.0	Operator-oriented, moderate heterogeneity

Note: G1: Regular users; G2: Vulnerable users.

Results

Table 3 and Figure 1 present the main transit network performance indicators. Metro headways reach their maximum admissible values across all scenarios, while rail lines exhibit greater variability. Increasing the operational cost weight (ρ_{op}) leads to higher headways in scenarios with the same level of demand heterogeneity. As a result, average waiting times also increase. Passenger-oriented scenarios (S4–S6) achieve average waiting times below or up to 5 minutes, whereas scenarios with higher operational cost weights exceed 5 minutes waiting time. As the importance of operational cost (ρ_{op}) increases, passenger cost also increases. In contrast, when higher emphasis is placed on passenger cost (ρ_{pass}), the indicator Pass-km decreases.
To assess the effect of demand heterogeneity, scenarios are compared under the same operational cost weight. No consistent increasing or decreasing trend is observed as heterogeneity increases, suggesting that demand heterogeneity does not significantly influence these indicators. This may suggest that timetable decisions are primarily driven by trade-offs between operational and passenger cost rather than user composition, partly due to the relatively small share of vulnerable users.
Regarding equity, vulnerable users consistently experience between 4.29% and 4.43% higher generalized costs compared to regular users across all scenarios. In more realistic settings, where demand is spatially heterogeneous, larger differences between user groups could emerge.



(a) Passenger cost

(b) Operational cost

(c) Vehicle-km

(d) Passenger-km

Figure 1: Impact of demand heterogeneity on key performance indicators.

Table 3: Network Performance.

Indicator	Balanced Objective			Passenger Oriented			Operational Oriented		
	S1	S2	S3	S4	S5	S6	S7	S8	S9
Headways									
H_A	20	15	15	10	15	15	20	20	20
H_{Bx}	15	15	15	15	15	15	20	15	15
H_B	20	20	20	20	20	15	20	15	20
H_C	20	20	20	10	20	15	20	20	20
H_E	20	20	20	20	20	20	20	20	20
H_F	10	10	10	10	10	10	20	15	20
H_H	20	15	15	10	10	15	20	20	20
H_{Metro}	3.50	3.50	3.50	3.50	3.50	3.50	3.50	3.50	3.50
Network Performance Indicators									
Avg. Waiting (min)	5.49	5.14	5.14	4.42	5.00	4.87	5.88	5.47	5.70
Load Factor	0.1944	0.1868	0.1868	0.1570	0.1851	0.1760	0.1992	0.1910	0.1966
Objective Value									
Passenger Cost	230,473	225,108	225,613	215,699	223,296	221,388	237,868	230,703	234,861
Operational Cost	271,710	277,790	277,790	303,210	281,130	285,990	266,250	273,630	268,920
Crowding Cost	106	112	101	112	100	0	108	0	101
Objective Value	447,947	447,452	447,945	367,416	363,961	364,383	637,360	641,148	638,341

Table 4: Equity: Generalized cost and transfers by user group

Scenario	GC_{g1}	GC_{g2}	**Gap GC (%)**	NT_{g1}	NT_{g2}	**Gap Transfers (%)**
S1	9.57	–	–	0.1970	–	–
S2	9.37	9.79	4.453%	0.1972	0.1945	-1.369%
S3	9.37	9.79	4.453%	0.1972	0.1945	-1.369%
S4	8.96	–	–	0.1970	–	–
S5	9.29	9.71	4.497%	0.1974	0.1946	-1.418%
S6	9.21	9.63	4.531%	0.1971	0.1945	-1.319%
S7	9.80	–	–	0.1954	–	–
S8	9.57	9.98	4.349%	0.1960	0.1941	-0.969%
S9	9.70	10.12	4.287%	0.1954	0.1940	-0.716%

Note: GC_{g1}, GC_{g2} and NT_{g1}, NT_{g2} denote average generalized cost and transfers for regular and vulnerable users. Gaps (%) are relative to regular users.

4 Conclusions

This paper presented a MILP formulation for timetable optimization in mixed transit networks, combining schedule-based and frequency-based services. The model accounts for heterogeneous demand, hard capacity constraints, and differences in passenger behaviour across service types. The methodology was tested on a case study of the Copenhagen transit network, evaluating nine scenarios with varying levels of demand heterogeneity and operational cost weights.

The results indicate that demand heterogeneity does not show a clear effect on overall network performance. In contrast, the weighting of operational and passenger costs has a stronger impact, with higher operational costs leading to longer headways and increased waiting times.

Across all scenarios, vulnerable users experience higher generalized costs than regular users, despite having heterogeneous demand. Future research could integrate equity directly by including a constraint in the formulation or by placing an equity cost in the objective minimization.

Acknowledgements

This research was supported by the Horizon Research and Innovation Actions (Grant No. 101147397, CulturalRoad) and partially funded by the Independent Research Fund Denmark (2035-00111B).

References

Chowdhury, S., Ceder, A., & Schwalger, B. (2015). The effects of travel time and cost savings on commuters' decision to travel on public transport routes involving transfers. *Journal of Transport Geography*, *43*, 151–159.

Ibarra-Rojas, O. J., & Rios-Solis, Y. A. (2012). Synchronization of bus timetabling. *Transportation Research Part B: Methodological*, *46*(5), 599–614.

Ingvardson, J. B., Nielsen, O. A., Raveau, S., & Nielsen, B. F. (2018). Passenger arrival and waiting time distributions dependent on train service frequency and station characteristics: A smart card data analysis. *Transportation Research Part C: Emerging Technologies*, *90*, 292–306. doi: 10.1016/j.trc.2018.03.006

Jiang, Y., Rasmussen, T. K., & Nielsen, O. A. (2022). Integrated optimization of transit networks with schedule- and frequency-based services subject to the bounded stochastic user equilibrium. *Transportation Science*, *56*(6), 1452–1468.

Liu, L., Kar, A., Tokey, A. I., Le, H. T. K., & Miller, H. J. (2023). Disparities in public transit accessibility and usage by people with mobility disabilities: An evaluation using high-resolution transit data. *Journal of Transport Geography*, *109*, 103589.

Nielsen, O. A., Rasmussen, T. K., & Ingvardson, J. B. (2019). Multimodal route choice models of public transport in the greater copenhagen area using smart card data. *Transportation Research Procedia*, *37*, 301–308.

Park, J., & Chowdhury, S. (2022). Investigating the needs of people with disabilities to ride public transport routes involving transfers. *Journal of Public Transportation*, *24*, 100010.

Szeto, W. Y., & Jiang, Y. (2014). Transit route and frequency design: Bi-level modeling and hybrid artificial bee colony algorithm approach. *Transportation Research Part B: Methodological*, *67*, 235–263.

Vielma, J. P. (2015). Mixed integer linear programming formulation techniques. *SIAM Review*, *57*(1), 3–57.