Operator's dilemma: Converging Passenger Demand and Supply using Electronic Ticketing Machine Data

Kartik Agarwal, Garvit Juneja, Satyam Singh, Rupam Fedujwar, and Amit Agarwal Indian Institute of Technology Roorkee Correspondance: amitfce@iitr.ac.in

Abstract. The effectiveness of the bus system often relies not only on the availability of buses but also on how well those buses are allocated across different routes and time periods based on the passenger demand. Uneven distribution of buses can lead to overcrowding on certain high-demand routes, while other routes may be underutilized, leading to wastage of resources. Operators face such issues in day-to-day scheduling. This study addresses this gap by addressing the operator's dilemma of the number of buses allocated for maximum utilization of the resources. The study developed an optimization-based approach to minimizing the demand-supply gap by leveraging the potential of Electronic Ticketing Machine (ETM) data. The model considers three levels of occupancy (or convenience) in the model to see the effect of optimization. The ETM data from Indore, India, is used to develop the model. The results confirm that the strategic distribution of service levels improves capacity utilization by assigning appropriate occupancy in transit vehicles based on the demand profiles of different routes. Notably, these improvements were achieved without changing the fleet size. The findings demonstrate that data-driven optimization can substantially reduce the supply-demand gap in public transit systems by effectively utilizing existing resources.

Keywords: Public transport; Demand-supply gap; Optimization; ETM

Preferred citation style: Agarwal, Kartik, Garvit Juneja, Satyam Singh, Rupam Fedujwar and Amit Agarwal (2025). "Operator's dilemma: Converging passenger demand and supply using electronic ticketing machine data". In: 8th Conference of Transportation Research Group of India (CTRG). Guwahati, India. See WP-104 at: https://faculty.iitr.ac.in/~amitfce/publications.html.

1 Introduction

Urban public transportation systems play a pivotal role in shaping the mobility, accessibility, and sustainability of modern cities. As urban populations continue

to grow, especially in developing countries, public transport becomes an essential service for ensuring equitable access to jobs, education, health services, and social activities [1]. In most cities, buses constitute the backbone of the public transport system due to their operational flexibility, lower infrastructure requirements, and relative affordability compared to other modes of transport [12]. Cities across the globe are actively investing in improving their bus networks to accommodate increasing passenger volumes, reduce congestion, and lower carbon emissions. However, the effectiveness of bus systems hinges not only on the availability of buses but also on how well those buses are allocated across different routes and time periods based on passenger demand.

Despite the critical role of buses in urban transport, many cities face persistent challenges related to service quality, especially on individual routes [5]. Uneven distribution of buses can lead to overcrowding on certain high-demand routes, while other routes may be underutilized, leading to wastage of resources. From the passenger's perspective, the quality of service is judged by factors such as frequency [8], travel time, and its reliability [14], comfort, and availability of seating [4]. A poorly allocated system, where buses do not match passenger demand, often results in long waiting times, crowded buses, and reduced user satisfaction, discouraging the use of public transport [13]. On the other hand, routes that are overserved waste operational resources such as fuel, labor, and maintenance capacity. Therefore, improving service quality requires a deep understanding of passenger demand and an efficient allocation of supply in response to it.

The challenge lies in the demand-supply gap, the difference between the number of passengers requiring service and the available bus capacity on a particular route during a given time period. When this gap is large, passengers may be unable to board buses or face severe discomfort, leading to loss of ridership and shifting to private modes of transport. Conversely, a negative gap implies underutilization of buses, increasing per-passenger operational costs. Reducing this gap is thus crucial not only for improving the passenger experience but also for enhancing the efficiency and financial sustainability of public transport agencies. A well-balanced demand-supply scenario ensures that each route is served by an optimal number of buses that align with actual usage patterns, thereby contributing to both operational performance and passenger satisfaction.

To effectively reduce the demand-supply gap, a variety of models and tools have been developed over the years. Traditional methods often relied on manual surveys and empirical rules of thumb to allocate buses based on average daily ridership. Further research on transit frequency optimization approaches addressing different aspects of the problem. Early work by [10] analyzed route frequencies and trip distributions across zones, proposing an iterative solution method. Later, [6] introduced a frequency-setting model focused on reducing passenger wait times and overcrowding. Their approach involved a two-step algorithm: first optimizing fleet size, then adjusting allocations to limit peak crowding levels. Further advancements came from [11], who developed a method to determine both vehicle sizes and route frequencies based on existing bus net-

works and origin-destination demand data. Several additional studies are also cited in Table 1. However, these approaches may fail to capture temporal variations, peak-hour congestion, or spatial inequities across routes. With the advent of fare collection system's and smart card ticketing systems, it has become possible to estimate demand at a much finer spatial and temporal resolution.

Among these, optimization models have emerged as powerful tools for systematically allocating limited bus resources in the most efficient manner. An optimization model provides a mathematical framework to minimize or maximize an objective function, such as minimizing the total demand-supply gap across all routes, subject to a set of constraints like total fleet size, maximum buses per route, or operational costs. These models can incorporate real-world constraints and trade-offs, making them highly applicable to practical transit planning scenarios. However, implementing such solutions on a recurrent basis can be challenging for operators, who may prefer adjustments, such as moving a few buses between routes, rather than undertaking comprehensive rescheduling for long-term planning. In this study, an optimization-based approach is used to address the critical issue of bus allocation across transit routes, leveraging observed passenger demand derived from Electronic Ticketing Machine (ETM) data. ETM data, with its high granularity and accuracy, provides a reliable source for estimating route-level demand, enabling data-driven decision-making in transit planning. The objective is to minimize the divergence between demand and supply by determining the optimal number of buses to be assigned to each route, while adhering to constraints such as the total available fleet size and minimum service frequency. The study leverages demand data aggregated at route and time-of-day levels and formulates an optimization model. By focusing on reducing the difference between the capacity and demand, the model aims to ensure better utilization of buses.

2 Case study: Indore

The method examines the fixed-route bus services operated by Atal Indore City Transport Service Limited (AICTSL). AICTSL, established in 2005, is the primary transit provider throughout Indore city, in Madhya Pradesh, one of India's fastest-growing urban centers. With its 34 bus routes, the city bus network serves as a vital mode of transportation in Indore, catering to a population of over 2 million residents. The dataset used in this study originates from the Electronic Ticketing Machines (ETMs) deployed across public buses operating in Indore, India. ETMs automatically capture rich, operational, and service data, including trip-level details, timestamps, route identifiers, passenger counts, and revenue information. This comprehensive ticketing data provides a detailed representation of the transit system's performance and passenger demand patterns, making it invaluable for optimization analysis.

Table 1: Summary of past studies on optimization models

Study	objective function	decision vari- ables	constraints	model	outcome
[3]	total cost	quantities, inventory levels	Capacity, demand fulfillment, inventory balance	Mixed-integer linear pro- gramming (MILP)	Efficient supply chain optimization under uncertainty
[9]	_	Facility locations, transportation routes	ity capacity,	_	Robust optimization for disaster relief logistics
[7]	profit in	Production quantities, recycling rates	straints, en-	programming	Sustainable supply chain with recycling benefits
[15]		Power generation, storage allocation		•	Improved energy efficiency in smart grid systems
[16]	senger waiting	Bus frequencies, dispatching times	,	Mixed-integer programming	Improved bus scheduling
[2]		Route allocations, resource distribution	resource	Two-stage stochastic programming	Enhanced disaster response
[17]		Traffic signal timing, route assignments	ity, safety		

2.1 Data preprocessing

The ETM data went through multiple transformations to convert them into suitable values for optimization problem.

- 1. **Peak hour filtration**: While the suggested method can be applied to any time interval, the current research is centered on the peak demand hours (cf. Figure 1), during which optimal bus allocation is most important. To this end, the Electronic Ticketing Machine (ETM) data, i.e., passenger-level ticketing data, was preprocessed to keep only those transactions that fall within the peak hours.
- 2. **Boardings aggregation**: For every bus stop and route pair, the total number of passenger boardings $(B_{i,s})$ during the period is calculated, where i represents the route and s represents the stop. This helped in the analysis

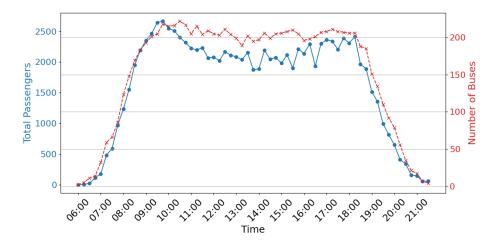


Fig. 1: Hourly demand (passengers) and supply (number of buses)

of the distribution of demand along each route. The demand on each route can be represented mathematically as:

$$D_i = \sum_{s \in S_i} B_{i,s} \quad \forall i \in I \tag{1}$$

where S_i represents the set of all stops along route i, and D_i represents the total passenger demand for route i during the given period.

3. **Initial bus assignment**: To determine the initial number of buses assigned to each route, the number of unique buses operating on that route during the given period is calculated. This value is denoted by N_i , representing the total number of buses assigned to route i. The total number of buses operating across the entire network during the study period is constant and given by:

$$N_f = \sum_{i \in I} N_i \tag{2}$$

where N_f represents the total number of distinct buses in operation and I is the set of all routes.

4. **Supply calculation**: The current supply for each route is calculated as the sum of seating capacity (S_b) and the standing capacity (S_b) over all routes. Standing capacity depends on the person density in the bus; higher is the density, higher is the capacity and lower is Level of Service (LoS).

$$C_i = \sum_{b=1}^{N_i} (S_b + ST_b) \quad \forall i \in I$$
 (3)

where N_i is the number of buses operating on route i and C_i represents the total passenger capacity on route i.

The preprocessing resulted in the following variables for our optimization problem:

- Route number (i)
- Total passenger demand for each route (D_i)
- Total number of buses for each route (N_i)
- Total current capacity of each route (C_i)

3 Methodology

This paper presents an optimization approach for allocating buses across multiple routes during the peak period. For the current study, the 9-11 AM peak time interval is taken as discussed in the previous section, in order to enhance capacity and demand matching of passengers. The approach involves a mathematical optimization model, comprehensive implementation procedures, and performance indicators to measure the performance of the suggested optimization model.

3.1 Problem formulation

The bus allocation problem is addressing how to allocate a fleet of buses to routes in such a manner that the capacity offered minus demand is minimized. It is an optimization problem whose main parameters are as follows:

- Routes: A set of routes in which each route is a separate bus route. There are 34 working routes during the 9-11 AM time period.
- **Bus fleet**: There are a total of 273 (N_f) buses in the ETM data that operate during the selected peak hours.
- Service types:
 - High convenience service (C_H) having $ST_b = 0$.
 - Moderate convenience service (C_M) having $0 < ST_b \le 15$.
 - Low convenience service (C_L) having $15 < ST_b \le 30$.
- Initial deployment: The original number of buses allocated to a route according to the initial bus assignment.

3.2 Mathematical Model

Objective function: The objective is to minimize the total of absolute differences between the capacity and passenger demand assigned to each route. The individual capacity for every route is estimated by adding up the total of all the assigned buses multiplied by their respective capacities. The aim is to minimize the overall gap between capacity and demand so that equitable distribution and optimal utilization of resources can be provided.

$$\min_{x_{b,i}^H, x_{b,i}^M, x_{b,i}^L} \sum_{i \in I} z_i = \min_{x_{b,i}^H, x_{b,i}^M, x_{b,i}^L} \sum_{i \in I} \left| C_i^{\text{total}} - D_i \right|$$
(4)

Where:

$$C_i^{\text{total}} = \sum_{b=1}^{N_f} \left(C_H \cdot x_{b,i}^H + C_M \cdot x_{b,i}^M + C_L \cdot x_{b,i}^L \right) \tag{5}$$

Decision variables: The model uses binary decision variables to represent the assignment of each bus to specific routes and vehicle service types:

$$x_{b,i}^H \in \{0,1\}$$
 $\forall b \in \{1,2,\dots,N_f\}, \forall i \in I$ (6)

$$x_{b,i}^{M} \in \{0,1\}$$
 $\forall b \in \{1,2,\dots,N_f\}, \forall i \in I$ (7)

$$x_{b,i}^{L} \in \{0,1\}$$
 $\forall b \in \{1,2,\dots,N_f\}, \forall i \in I$ (8)

Where:

- $-x_{b,i}^{H}=1$ if bus b is assigned to route i under the high-convenience service level (i.e., no standing passengers); 0 otherwise.
- $-x_{b,i}^{M}=1$ if bus b is assigned to route i under the moderate-convenience service level (i.e., limited standing capacity); 0 otherwise.
- $-x_{b,i}^{L}=1$ if bus b is assigned to route i under the low-convenience service level (i.e., higher standing capacity allowed); 0 otherwise.
- $-N_f$ denotes the total number of buses available in the network on a given operational day.

Additionally, auxiliary continuous variables is defined $z_i \in \mathbb{R}^+ \cup \{0\}, \forall i \in I$, to represent the absolute deviation between the total passenger-carrying capacity allocated to route i and the empirically observed passenger demand for route i.

Constraints: The model takes into account the following constraints:

Assignment limitation: Each bus can be assigned to at most one route and one service category.

$$\sum_{i \in I} \left(x_{b,i}^H + x_{b,i}^M + x_{b,i}^L \right) \le 1 \quad \forall b \in \{1, 2, \dots, f\}$$
 (9)

Minimum bus constraint: All routes must be assigned at least a single bus.

$$\sum_{b=1}^{N_f} \left(x_{b,i}^H + x_{b,i}^M + x_{b,i}^L \right) \ge 1 \quad \forall i \in I$$
 (10)

Deviation linearization: For optimization convenience, linear constraints are employed to model the absolute deviation in a computationally tractable way.

$$z_{i} \ge \sum_{b=1}^{N_{f}} \left(C_{H} \cdot x_{b,i}^{H} + C_{M} \cdot x_{b,i}^{M} + C_{L} \cdot x_{b,i}^{L} \right) - D_{i} \quad \forall i \in I$$
 (11)

$$z_{i} \ge D_{i} - \sum_{b=1}^{N_{f}} \left(C_{H} \cdot x_{b,i}^{H} + C_{M} \cdot x_{b,i}^{M} + C_{L} \cdot x_{b,i}^{L} \right) \quad \forall i \in I$$
 (12)

Total fleet: The total number of buses assigned across all routes should not exceed the available fleet, denoted as N_f .

$$\sum_{b=1}^{N_f} \sum_{i \in I} \left(x_{b,i}^H + x_{b,i}^M + x_{b,i}^L \right) \le N_f \tag{13}$$

3.3 Implementation

The optimization model was solved using the Gurobi optimization solver through its Python interface. The procedure is as follows:

Model construction and solution: The implementation creates a Gurobi optimization model with:

- Binary decision variables for bus-route assignments.
- Support variables for demand-capacity imbalance.
- Constraints imposed to ensure valid and effective allocation.
- Objective to minimize total deviation.

The optimization problem was addressed using Gurobi's branch-and-bound algorithm to determine the optimal solution.

3.4 Comparative analysis

To evaluate the advantages of the optimization model, the original and optimized allocations were compared: (a) changes in route-level bus allocation, (b) reductions in the supply-demand gap, and (c) bus service levels are reevaluated across routes based on passenger demand to show capacity and comfort. This evaluation identifies areas for improvement and demonstrates the value of the optimization strategy in practical transit planning scenarios.

4 Results and Discussion

4.1 Value of objective function

The value of the objective function over iterations is shown in Figure 2. As required, the value of objective function drops significantly in the beginning and then stabilizes over the iterations. At iteration 40, the objective value is 182.0, which indicates a nearly optimal bus allocation with negligible mismatch.

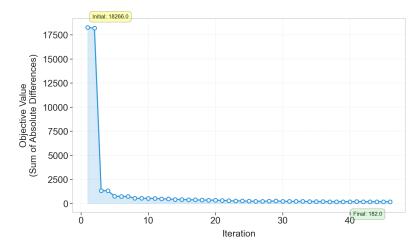


Fig. 2: Convergence of optimization function with iterations

4.2 Bus reallocation

Figure 3a illustrates the initial and optimized allocations for each route, high-lighting the routes that experienced an increase or decrease in the assigned buses after optimization. Further, Figure 3b provides the net changes in bus allocation resulting from the optimization process to perfectly illustrate how buses were taken from routes with excessive supply and assigned to routes with excessive demand. Overall, the most significant changes occurred on routes that had extreme differences in demand and supply values in the initial assignment. For instance, Route R-5, having the highest demand, received an additional 10 buses in the optimized allocation, while routes like M-4 and E-3 had their allocation reduced by 4 and 3 buses, respectively, allowing reallocation of excess buses.

4.3 Capacity utilization

Figure 4a shows the average demand at each route initially and the number of buses assigned on each route after optimization for all routes. On a few routes, where the demand is low, minimally a bus is assigned to respect the accessibility (e.g., see last four routes in Figure 4a). Figure 4b demonstrates a comparison of the allocation of buses to various transit routes according to convenience or occupancy. The first set of stacked chart for each route shows the occupancy or convenience before optimization and second set of stacked chart exhibits the results after optimization. Most of the buses have low convenience, i.e., higher occupancy to meet the demand, particularly for crowded routes R-5, M-6, and M-27. High Convenience buses and moderate convenience buses are allocated to some routes, on the basis of certain requirements and capacity for such routes, which makes them more occupied and less convenient. A combination of various

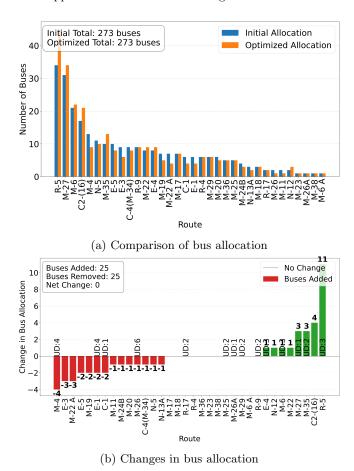
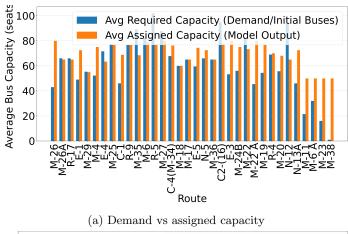


Fig. 3: Optimization results showing a number of buses allocated on each route before and after

types of buses on some routes and a single type on others reflects a thoughtful use of resources in the transit system.

Figure 5 shows the bus composition of the network for the original and optimized fleets. The original fleet had 151 low-convenient buses, 44 moderate-convenient buses, and 78 high-convenient buses. After optimization, the same numbers are 199, 37 and 37, respectively, which clearly highlights the effect of optimization. This shift implies that there will be more buses with higher occupancy that are less convenient. Although the overall number of buses remains constant at 273, the overall number of spaces (for passengers) will increase from 18,840 to 20,175. The optimization maximizes the use of available capacity by allocating the more passengers in the transit vehicles on different routes to have a better fit and lesser gap between demand and supply. Clearly, from the operators perspective, more crowded buses are good but from the passengers' perspective,



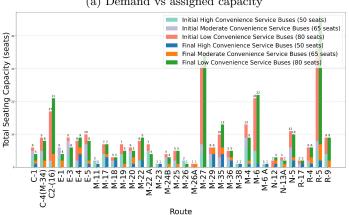


Fig. 4: Optimization results (a) showing a comparison of demand and supply (b) number of buses on each route differentiated by the convenience levels

(b) Number of buses by convenience

it is inconvenient. Thus, in the future, an increase in the occupancy may be considered as a penalty and model may be re-optimized.

5 Limitations

Although there are great improvements evident through the optimization model, there are necessary restrictions to point out:

1. Static demand assumption: The model assumes demand for passengers to be static, while real demand varies depending on the quality of the service, frequency, hour of the day, day of the week, and season. A dynamic method with demand elasticity may yield more realistic results.

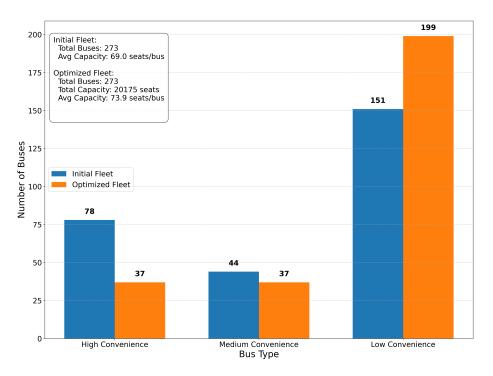


Fig. 5: Distribution of bus types across the network after optimization

- 2. **Operational restrictions**: Even though the model has significant operational restrictions like the total capacity of the fleet and availability of different bus types, it does not include driver schedules, maintenance needs, and other operational complexities that could change the feasibility of the model
- 3. **Fixed route network**:The optimization is for a fixed route network and consists of bus allocation alone. Combined bus allocation and route optimization could lead to even higher increases in system effectiveness. Optimizing both route configuration and bus allocation simultaneously would lead to dynamic adjustments in route frequency and sequence of stops in the route depending on the demand.
- 4. **Representation of aggregate demand**: The model is aggregate passenger volume and does not represent the geographical distribution of demand across routes or the passenger comfort choice, which can affect the perceived quality of service.

In the future, more research must break these limitations to build more sophisticated optimization models that can provide more meaningful suggestions for planning and operating public transport systems.

6 Conclusion

The study presented an optimization-focused approach to improve bus distribution in city bus routes, effectively balancing demand and supply. The applied model on Indore buses significantly enhanced efficiency by reallocating the existing fleet without adding new buses. This data-driven approach highlights how transit systems can maximize efficiency within existing constraints. The study showed the possibility of better bus utilization across different routes, i.e., buses from unde-utilized routes are re-assigned to over-utilized (or under-supplied) routes. Further, it also demonstrated that after optimization highly occupied buses increased from 55% to 73%, which is desirable from operator's perspective. Future work could incorporate time-varying demand patterns for adaptive scheduling. Integrate operational constraints like driver shifts and maintenance, and expand the framework for multi-model integration and inter-route transfers. Additionally, passenger comfort and environmental metrics could further refine optimization, promoting sustainable urban mobility. The findings of this study have important real-world implications for transit operators and planners, demonstrating that meaningful service improvement is achievable through efficient utilization even within existing budget and fleet constraints. Through careful supply and demand trend matching management, transit agencies are well-positioned to improve service quality, improve operating efficiency, and even increase ridership, leading to more efficient urban mobility systems.

Bibliography

- [1] Avishai Ceder. Public transit planning and operation: Modeling, practice and behavior. CRC press, 2016.
- [2] Simin Chai, Jiateng Yin, Andrea D'Ariano, Marcella Samà, and Tao Tang. Train schedule optimization for commuter-metro networks. *Transportation Research Part C: Emerging Technologies*, 155:104278, October 2023. ISSN 0968-090X. https://doi.org/10.1016/j.trc.2023.104278.
- [3] G. Diepen, B.F.I. Pieters, J.M. van den Akker, and J.A. Hoogeveen. Robust planning of airport platform buses. *Computers & Operations Research*, 40(3):747–757, March 2013. ISSN 0305-0548. https://doi.org/10.1016/j.cor.2011.08.002.
- [4] Rupam Fedujwar and Amit Agarwal. A systematic review on crowding valuation in public transport. *Public Transport*, 16(3):743–773, 2024. ISSN 1613-7159. https://doi.org/10.1007/s12469-024-00363-w.
- [5] Rupam Fedujwar and Amit Agarwal. Performance assessment of public transport routes: A framework using revealed data. Research in Transportation Business & Management, 59:101283, 2025. https://doi.org/10.1016/j.rtbm.2024.101283.
- [6] Anthony F. Han and Nigel H.M. Wilson. The allocation of buses in heavily utilized networks with overlapping routes. *Transportation Research Part B: Methodological*, 16(3):221–232, June 1982. ISSN 0191-2615. https://doi.org/10.1016/0191-2615(82)90025-x.
- [7] P. Kuppusamy, S. Venkatraman, C.A. Rishikeshan, and Y.C.A. Padmanabha Reddy. Deep learning based energy efficient optimal timetable rescheduling model for intelligent metro transportation systems. *Physical Communication*, 42:101131, October 2020. ISSN 1874-4907. https://doi.org/10.1016/j.phycom.2020.101131.
- [8] Arnoud Mouwen. Drivers of customer satisfaction with public transport services. *Transportation Research Part A: Policy and Practice*, 78:1–20, August 2015. ISSN 0965-8564. https://doi.org/10.1016/j.tra.2015.05.005.
- [9] Dimitrios Rizopoulos and Georgios K. D. Saharidis. Generic approaches for the rescheduling of public transport services. *Energy Systems*, 15(4): 1341–1370, June 2020. ISSN 1868-3975. https://doi.org/10.1007/s12667-020-00393-w.
- [10] Siv Schéele. A supply model for public transit services. *Transportation Research Part B: Methodological*, 14(1–2):133–146, March 1980. ISSN 0191-2615. https://doi.org/10.1016/0191-2615(80)90039-9.
- [11] Mao-Chang Shih and Hani Mahmassani. Vehicle sizing model for bus transit networks. TRANSPORTATION RESEARCH RECORD, 1995.
- [12] Hemant K. Suman, Nomesh B. Bolia, and Geetam Tiwari. Comparing public bus transport service attributes in delhi and mumbai: Policy implications for improving bus services in delhi. *Transport Policy*, 56:63–74, May 2017. ISSN 0967-070X. https://doi.org/10.1016/j.tranpol.2017.03.002.

- [13] Alejandro Tirachini, David A. Hensher, and John M. Rose. Crowding in public transport systems: Effects on users, operation and implications for the estimation of demand. *Transportation Research Part A: Policy and Practice*, 53:36–52, July 2013. ISSN 0965-8564. https://doi.org/10.1016/j.tra.2013.06.005.
- [14] Santhosh Kumar B V, Ruapm Fedujwar, and Amit Agarwal. Travel time variability of bus routes in delhi using real-time gtfs data. In 2024 16th International Conference on COMmunication Systems & NETworkS (COMSNETS), page 210–215. IEEE, January 2024. https://doi.org/10.1109/comsnets59351.2024.10427234.
- [15] Yanyan Wang, Changjiang Zheng, Jinxing Shen, Junze Ma, and Zhichao Chen. Last train delay management in a metro system: A multi-objective rescheduling optimization model. *IEEE Access*, 10:110821–110834, 2022. ISSN 2169-3536. https://doi.org/10.1109/access.2022.3215961.
- [16] Yonggang Wang, Junxian Chen, Yang Qin, and Xiaofang Yang. Timetable rescheduling of metro network during the last train period. *Tunnelling and Underground Space Technology*, 139:105226, September 2023. ISSN 0886-7798. https://doi.org/10.1016/j.tust.2023.105226.
- [17] Qu Zhen and Shi Jing. Train rescheduling model with train delay and passenger impatience time in urban subway network. *Journal of Advanced Transportation*, 50(8):1990–2014, December 2016. ISSN 2042-3195. https://doi.org/10.1002/atr.1441.