An optimal model for selection of stops for bus routes in busy times: Based on General Transit Feed Specification

Qinglin Zhu*

Department of Applied Mathematics, Xi'an Jiaotong Liverpool University, Suzhou, Jiangsu, China

*Corresponding author: qinglin.zhu19@student.xjtlu.edu.cn

Abstract. The bus routes model optimization dramatically influences the modern urban transport system because a better bus routes model could relieve traffic pressure, shorten the commuting time for citizens and effectively increase the number of people using public transport. This article proposes a new optimal model for bus routes during busy hours based on solving a linear programming problem, curtailing passenger travel time through a linear programming problem. It also helps solve traffic jams during busy hours and can act in different application scenarios due to its simplicity. Furthermore, this model was based on General Transit Feed Specification as much as possible to simplify the model and reduce the cost of its application, while at the same time, some alternative plans were discussed, which also pointed out the model optimizability.

Keywords: Bus routes, optimal model, general transit feed specification, linear programming.

1. Introduction

Regardless of the city size and economic level, city buses, including Bus Rapid Transit system, have become an inherent mode of transportation in the most modern city. However, severe traffic congestion is still a serious problem hindering the operation of the bus, incredibly huge traffic flow in rush hour. As one of the critical issues in modern social infrastructure, bus routes arrangement could solve the commuting problem of citizens and relieve traffic pressure. Furthermore, an optimal bus routes model could also significantly impact the city's development and the citizens' living standards.

To satisfy the dramatically blow-up demand for public transit during busy times, especially during rush hour and holidays, one of the traditional solutions is to increase the frequency of bus departures and opening new bus lines [1]. Although this method can meet the requirement, the imbalance between different stops, such as passenger flow and transfer needs, will bring massive waste of public resources. A better way to solve this problem is to apply the express mode [2] on the city bus. Different from the limitation on metro lines, for instance, the need for extra tracks and the restriction on overtaking [3], city bus has more degrees of freedom. According to DeWeese et al. [4], express routes mode could significantly contribute to society, including reducing commuting time and improving the employment environment. Furthermore, a preferable way to reduce public resource waste and increase utilization is to skip some unimportant stops for each bus line [5]. Such a method would lead to shorter passenger travel time, shorter runtime for each route and more efficient use of bus routes while also satisfying the passengers simultaneously by increasing the average waiting time for passengers slightly. Further development on this method is to dynamically determine the stop for each route, which satisfies the passenger demand better and further reduces the runtime.

This paper will discuss an optimal model for dynamically determining the stops using linear programming. This research will be based entirely on General Transit Feed Specification (GTFS), a kind of data specification that allows public transit agencies to publish their transit data in a format that can be consumed by a wide variety of software applications [6]. The model aims to minimize passenger travel time while satisfying passenger demand on expected waiting time at every stop at the same time. The rest of this paper is organized as follows: Section 2 presents the related work; Section 3 describes the problem in detail, introduces the assumption and constructs the model; Section 4 gives a summary and discusses the future research directions.

2. Related Work

The problem of optimizing transit routes with stop skipping strategy has been widely discussed nowadays. Li et al. [3] introduced integrated optimization to metro line stop planning to minimize passenger travel time. In the bus field, Deng and Zhu [2] discussed bus rapid transit and express mode with Hotelling Nash equilibrium in conventional bus systems in detail, providing several valuable propositions for this model construction. On the other hand, An and Zhang [7] also have proposed the stop-skipping strategy to minimize the cost by using mixed integer programming. Nesheli, Srikukenthiran and Shalaby [5] took the idea further and proposed a Nexus-based planning limited‑stop transit operations model to optimize transit paths.

Although GTFS started in 2006, the studies based on GTFS started to appear until the late 2010s. Based on the data format, several specific models have been set up. Chen et al. [8] built a reliability-based transit trip planning model by minimizing travel time to find reliable bus routes. Elliott and Lumley [9] proposed a travel time model based on GTFS and GPS vehicle locations, which is a wonderful way to make time prediction of bus in actual practice if GPS location is available. Moreover, the paper on the model for planning limited‑stop transit operations proposed by Nesheli, Srikukenthiran, Shalaby [5] and bus express mode evaluation declared by DeWeese et al. [4] are also based on GTFS. These increasingly GTFS-based study has revealed the reliability and value of GTFS.

3. Model Construction

3.1. Assumption and Notation

To simply and facilitate the model construction, the following assumptions are needed.

Assumption 1: All buses travel at an ideal speed with enough space and arrive at the station in the expected arrival time, that is, the model will ignore the influence of traffic conditions and vehicle space restrictions.

Assumption 2: Due to the lack of data on passenger flow on GTFS, the model will measure these data by bus flow, that is, assigned a weight (usually a percentage), which describe how many vehicles pass through this site, to each stop and use the weight to determine the passenger flow, and the origin-destination (OD) matrix can be evaluated correctly. A natural thought of the algorithm is to normalize the weight to a percentage and multiply it by total passenger flow to get the elements of the OD matrix.

Assumption 3: No passenger restriction on the vehicle, which means that vehicles always have ample space for passengers. Although this assumption is unlikely to occur in practice due to cost effects, the operator can meet this assumption by increasing the bus frequency in an ideal model.

Also, in order to promote the formulation, the notations with its descriptions are as below:

Table 1. Notation and Descriptions

Notation	Description
S	set of stops, by denoting each stop with $1,2,3,\dots,\lvert S\rvert$
S'_t	set of stops of bus shift t will stop
T	set of bus shift, by denoting each shift with $1,2,3,\dots,\lvert T\rvert$
$P_{t,o,d}$	passenger numbers of shift t, from stop o to stop d
WT_S	waiting time of the bus at stop s
TTS_S	travel time saved by bus if skipped stop s
TI_t	the time interval between shift t to shift $t+1$

3.2. Model Construction

In this model, two different kinds of passengers should be considered separately to measure the travel time. The first kind of passengers are whose origin and destination station fit the optimal route of the coming bus, that is, they are unnecessary to wait for extra time because of the skipped stop.

Another kind of passengers is just the opposite, with their origin and destination station not suitable for the coming bus, and therefore, they should wait for extra time until a consort route coming.

By assumption 2, the OD matrix can be obtained. Through this upper triangular matrix, $P_{t,o,d}$ can be measured by the elements of the OD matrix. To build a simple model, only a single unidirectional route from the starting station to the terminal station will be considered. Then, instead of calculating the value of waiting time directly, maximizing the time difference between two plans will be easier.

The first kind of passengers has two influences on passenger waiting time. One is the waiting time in stop s, denoted by WT_S, which would be naturally saved if the bus skips the stop. The other is the travel time saved by bus, denoted by TTS_S, including the delay caused by traffic conditions, acceleration and slowing down. The data of this part is usually hard to determine and is expected to obtain through the collection of operators. The second kind of passengers is more straightforward, with their only influence being extra waiting time, which can be considered as the time interval between this shift and next if we set a constraint that two consecutive shifts cannot skip the same station. Furthermore, this restriction can be relaxed in actual practice as one of three or four adjacent shifts must go through each site.

The figures below reveal the difference between optimized route and original route. In Fig 2, the stop s will be skipped in the bus route. Therefore, passengers whose destination is stop s, from stop 1 to stop $s-1$ will not be available to take this shift. These passengers should wait until next shift coming simultaneously.

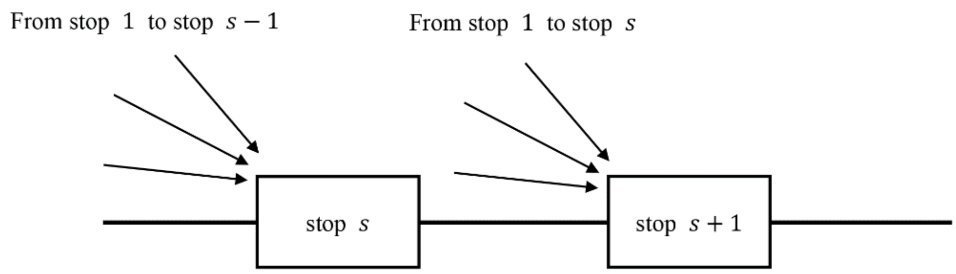

Fig. 1 Original route sketch at stop s

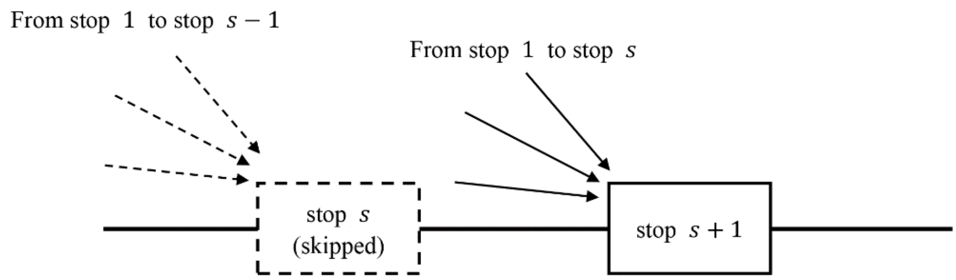

Fig. 2 optimized route sketch at stop s

By assumption and description, the objective function for shift t is obtained by maximizing

$$TS_t = \sum_{d=2}^{|S|}\sum_{o=1}^{d-1} P_{t,o,d}\cdot x_{d,t}\cdot \sum_{i=o}^{i=d}(WT_i + STT_i) - \sum_{d=2}^{|S|}\sum_{o=1}^{d-1} P_{t,o,d}\cdot TI_t\cdot \left(1 - x_{d,t}\right). \quad (1)$$

In Eq. (1), $x_{d,t}$ is used to measure the status of station s in optimized bus routes of shift t, that is, if the bus routes skipped the stop s, $x_{d,t} = 0$. Otherwise, $x_{d,t} = 1$. Also, the first term is to measure the time saved by optimal route with skipping some stops and the second one measure the extra waiting time for passengers whose destinations are the skipping stops.

Theoretically speaking, the terms in the formula can be obtained as follows, which are considered executable schemes in practice: waiting time (WT_S) is available through using GTFS stop_time.txt items, which is a mandatory field in the dataset; travel time saving (TTS_S) is expected to be obtained by the data collection of operators or more advanced models; time interval (TI_t) between shifts can be obtained by the pre-prepared schedule, which can be viewed as a constant commonly.

Furthermore, the system constraints of optimal model, which are used to ensure the model feasible is as follow. One is the zero-one constraints: $x_{s,t} \in \{0,1\}$, while the other is the adjacent shifts constraints (negotiable in practice): $x_{s,t} + x_{s,t+1} \geq 1$, for all shift t. In addition, there are also other extra constraints that may work. In actual practice, several factors should be considered, including the traveler's needs and restrictions on bus routes. That is, determined according to the actual requirement of passengers and routes.

4. Conclusion

In this paper, an optimal model for the selection of stops for bus routes based on GTFS has been proposed and discussed, which provides a feasible algorithm to calculate the skipped stop for bus routes and minimize the travel time for passengers. Moreover, this paper also explores the possibility of simplifying the bus route optimization model, points out the necessary conditions and requirements for the simplified model and provides reference suggestions for reducing the complexity of the simplified model in the future. The primary target users of this research are those operators who need to optimize the bus route arrangement, aiming to reduce operating costs and improve passengers' travel experience.

In future studies, the model can be optimized by modifying several assumptions, especially in the assumption part. According to Elliott and Lumley [9], introducing the GPS could evaluate the travel time more precise, and the first assumption could be removed. Passenger flow data is also available to be obtained directly in real action, so the model's accuracy can be promoted. Also, stop spacing and headway, which is a factor of user wait time [10], can also be added as a factor to the optimal model. Furthermore, the optimization degree should also be evaluated, which would afford detailed model performance for the user.

References

[1] Zhang, J. et al. Flexible Express Bus Line Planning and Operating Based on Passenger Flow Analysis. 2018 21st International Conference on Intelligent Transportation Systems (ITSC), 2018, 2511-2518.

[2] Deng, F., Zhu, J. An express mode of urban bus systems. 2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC), Intelligent Transportation Systems (ITSC), 2020, 1–5.

[3] Li, Z., Mao, B., Bai Y., Chen Y. Integrated Optimization of Train Stop Planning and Scheduling on Metro Lines With Express/Local Mode. IEEE Access, 2019, 7: 88534-88546.

[4] DeWeese, J., Palacios M.S., Belikow, A., El-Geneidy, A. Whose express access? Assessing the equity implications of bus express routes in Montreal, Canada. Journal of Transport and Land Use, 2022, 15(1): 25-51.

[5] Nesheli, M.M., Srikukenthiran, S. and Shalaby, A. An optimization model for planning limited-stop transit operations. Public Transport, 2022, 14: 63-83.

[6] https://gtfs.org/

[7] An, S. and Zhang, X. Real-time hybrid in-station bus dispatching strategy based on mixed integer programming. Information (Switzerland), 2016, 7(3): 43.

[8] Chen, Y., Yang, S., Hu, M. and Wu, Y. A reliability-based transit trip planning model under transit network uncertainty. Public Transport: Planning and Operations, 2016, 8(3): 477-496.

[9] Elliott, T., Lumley, T. Modelling the travel time of transit vehicles in real-time through a GTFS-based road network using GPS vehicle locations. Australian & New Zealand Journal of Statistics, 2020, 62(2): 153–167.

[10] Zhao, L., Chien, S. I. Investigating the impact of stochastic vehicle arrivals to optimal stop spacing and headway for a feeder bus route. Journal of Advanced Transportation, 2015, 49(3): 341–357.