Transportation Research Part C 88 (2018) 1–16

Contents lists available at ScienceDirect

Transportation Research Part C

journal homepage: www.elsevier.com/locate/trc

# Reinforcement learning approach for coordinated passenger inflow control of urban rail transit in peak hours

Zhibin Jiang[a], Wei Fan[a,b,*], Wei Liu[c], Bingqin Zhu[d], Jinjing Gu[a]

[a] *College of Transportation Engineering, Key Laboratory of Road and Traffic Engineering of the State Ministry of Education, Tongji University, 4800 Cao'an Rd., Shanghai 201804, PR China*

[b] *USDOT Center for Advanced Multimodal Mobility Solutions and Education, Department of Civil and Environmental Engineering, University of North Carolina at Charlotte, 9201 University City Boulevard, Charlotte, NC 28223, USA*

[c] *Technical Center of Shanghai Shentong Metro Group Co., Ltd., 909 Guilin Road, Xuhui, Shanghai 201103, PR China*

[d] *Operation Management Center of Shanghai Shentong Metro Group Co., Ltd., 222 Hengtong Road, Jing'an, Shanghai 200070, PR China*

ARTICLE INFO

*Keywords:*
Urban rail transit
Train capacity
Operation safety
Coordinated passenger inflow control
Reinforcement learning

ABSTRACT

In peak hours, when the limited transportation capacity of urban rail transit is not adequate enough to meet the travel demands, the density of the passengers waiting at the platform can exceed the critical density of the platform. Coordinated passenger inflow control strategy is required to adjust/meter the inflow volume and relieve some of the demand pressure at crowded metro stations so as to ensure both operational efficiency and safety at such stations for all passengers. However, such strategy is usually developed by the operation staff at each station based on their practical working experience. As such, the best strategy/decision cannot always be made and sometimes can even be highly undesirable due to their inability to account for the dynamic performance of all metro stations in the entire rail transit network. In this paper, a new reinforcement learning-based method is developed to optimize the inflow volume during a certain period of time at each station with the aim of minimizing the safety risks imposed on passengers at the metro stations. Basic principles and fundamental components of the reinforcement learning, as well as the reinforcement learning-based problem-specific algorithm are presented. The simulation experiment carried out on a real-world metro line in Shanghai is constructed to test the performance of the approach. Simulation results show that the reinforcement learning-based inflow volume control strategy is highly effective in minimizing the safety risks by reducing the frequency of passengers being stranded. Additionally, the strategy also helps to relieve the passenger congestion at certain stations.

## 1. Introduction

With the rapid development of urban rail transit (URT) in China, the limited transportation capacity is not adequate enough to meet the booming travel demands, especially during peak hours. It is common to see a number of people gather at the platform and cannot get aboard when the train arrives during the rush hours. In particular, if the density of the passengers waiting at the platform exceeds the critical density of the platform, it will be extremely challenging to ensure the safety of passengers and efficient daily operations.

Increasing the capacity is a straightforward solution. However, it is not always feasible in practice due to potential long

* Corresponding author at: Department of Civil and Environmental Engineering, University of North Carolina at Charlotte, 9201 University City Boulevard, Charlotte, NC 28223, USA.

*E-mail address:* wfan7@uncc.edu (W. Fan).

https://doi.org/10.1016/j.trc.2018.01.008
Received 6 April 2017; Received in revised form 27 December 2017; Accepted 6 January 2018
Available online 30 January 2018
0968-090X/ 

**Table 1**
Passenger inflow control for URT in major cities in China.

| City | The number of operating lines | The number of inflow control lines | The number of inflow control stations | Time Periods for passenger inflow control |
|---|---|---|---|---|
| Beijing | 18 | 14 | 75 | Morning peak 7:00–9:00<br>Evening peak 17:00–19:00 |
| Shanghai | 14 | 13 | 50 | Morning peak 7:30–9:00<br>Evening peak 17:30–19:00 |
| Guangzhou | 9 | 6 | 38 | Morning peak 8:00–9:00<br>Evening peak 17:30–19:00 |

Data source: Beijing Metro Operation Co., Ltd. & Guangzhou Metro Operation Co., Ltd.: Collected in April 2016; Shanghai Metro Operation Co., Ltd.: Collected in December 2015.

construction time period, budget limitations, right-of-way constraints, operational and safety restrictions including line capacity, and the maximum possible amount of rolling stock, etc. Under such circumstances, passenger inflow control can be a potentially effective short-term alternative to ensure operational efficiency and safety while also lowering the pressure on stations. By setting railings outside metro stations, shutting down part of the ticket vending machines (TVM) or entrance gates, and/or closing off partial entrances, the speed and flow rate of passengers entering the metro stations during a certain period of time can be effectively limited so that the number of passengers waiting at the platform per unit of time is under control. In fact, these measures for passenger inflow control have already been taken in the daily operation of metro lines in major cities of China like Beijing, Shanghai and Guangzhou (see Table 1 below). Unfortunately, such present control strategies are usually developed based on the engineering judgement and subjective work experience of the operation staff at each station without any help from mathematical programming and scientific methods, leading to the potential deficiency in overall and dynamic performance.

The coordinated passenger inflow control is highly needed to maintain the safety of all passengers in which a quick response to the demand flow dynamics is typically required. In this paper, a novel approach is developed which is based on reinforcement learning, more specifically Q-learning. Based on the simulation of the interaction between passengers and trains, the reinforcement learning algorithm automatically learns when and at which stations to enforce inflow control and the optimal control rates at each control station (measured in per unit of time) with the aim of minimizing the safety risks at metro stations of a single line. As a powerful tool for solving complex sequential decision-making problems in control theory (Gosavi, 2009), reinforcement learning is able to make quick responses to the dynamic changes of the network environment and has already been successfully used to solve the similar problem such as traffic flow control optimization problem in highways.

The main contributions of this paper can be summarized as follows:

1. Real-time data about some key indicators, including the number of people waiting at the platform and the frequency of passengers being stranded, are used to evaluate the safety risks in this paper. On the one hand, more people waiting at the platform can certainly pose great challenges while increasing the possibility of pushing some passengers off of the platform and/or stepping on those who may fall onto the ground under overcrowded situations, which may cause significant safety risks of the platform. On the other hand, higher frequency of passengers being stranded can result in more anxious and impatient passengers especially in morning peak hours, many of whom may force/rush into the train (therefore making the doors unclosed and causing significant delays in train departure), and sometimes could even lead to severe accidents. As such, both the number of people waiting at the platform and the frequency of passengers being stranded are used in this paper to evaluate the safety risks in this paper.
2. Model of coordinated passenger inflow control is formulated to minimize the penalty value of passengers being stranded on a whole metro line considering real-time passenger trip chains and states. The interaction between passenger and train allows for a quantitative evaluation of their available capacities respectively. It can be used and applied to identify the entry rate of passengers at each station with dynamical passenger demand.
3. The reinforcement learning developed in this paper can be applied to formulate strategies dictating when and at which stations to enforce inflow control and the corresponding control rates to ensure the safety of the entire metro line.

The rest of the paper is organized as follows. Section 2 reviews the previous studies on passenger inflow control and the application of reinforcement learning. Section 3 describes coordinated passenger inflow control problem for a single metro line during peak hours. Section 4 builds an optimization model for the coordinated passenger inflow control problem on a whole line. Section 5 presents a unique approach to the coordinated passenger inflow control with reinforcement learning by systematically explaining the fundamental components and rationale behind the specific design choices. Section 6 provides a real-world example and discusses how it works to support the coordinated passenger inflow control. Finally, conclusions are made and future research directions are also given in section 7.

## 2. Literature review

The coordinated passenger inflow control of URT is a complex and very challenging optimization problem which involves various essential components including passengers, trains, infrastructure and strategies. Previous studies regarding the passenger inflow