Transportation Research Part A 111 (2018) 227–251

Contents lists available at ScienceDirect

Transportation Research Part A

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

Inter-arrival time distribution of passengers at service facilities in underground subway stations: A case study of the metropolitan city of Chengdu in China

Yanru Chen[a], Bing Yi[b], Yangsheng Jiang[c,*], Jidong Sun[d], M.I.M. Wahab[e]

[a] School of Economics and Management, Southwest Jiaotong University, 111, 1st Section, Northern 2nd Ring Road, Chengdu, Sichuan 610031, China
[b] ChongQing Urban Integrated-Transportation Pivot Investment Co., Ltd, 128, Zhongshan 3rd Road, Yu Zhong District, Chongqing 400000, China
[c] School of Transportation and Logistics, Southwest Jiaotong University, 111, 1st Section, Northern 2nd Ring Road, Chengdu, Sichuan 610031, China
[d] Department of Intelligent Transportation, Beijing Aerospace Dacheng Wisdom System Technology Development Center, 5 Minzuyuan Road, Chaoyang District, Beijing 100086, China
[e] Department of Mechanical and Industrial Engineering, Ryerson University, 350 Victoria Street, Toronto M5B2K3, Canada

A R T I C L E I N F O

Keywords:
Inter-arrival time distribution of passengers
Hyper-Erlang distribution
Parameter estimation model
Service facility
Underground subway station

A B S T R A C T

Inter-arrival time distribution of passengers plays an important role in the capacity design of service facilities such as, fare gate, ticket vending machine, and passageways, in an underground subway station. An inaccurate inter-arrival time distribution likely causes traffic congestion or resource wastage at service facilities. In this study, to obtain accurate inter-arrival time distribution, we collected an inter-arrival time of passengers at existing service facilities in three underground subway stations of a metropolitan city, Chengdu, China. We fitted eight types of distributions, including Hyper-Erlang distribution (HErD), which is firstly introduced in the capacity design of service facilities in an underground subway station, to the observed data set based on maximum likelihood estimation. Results showed that the HErD works the best in terms of fitting quality and flexibility. We also fitted eight types of distribution to the observed inter-arrival time data at service facilities of seventy-seven underground subway stations of another metropolitan city—Shenzhen, China—to confirm our findings. Results also showed the HErD still performs the best. Simulation is also conducted to examine the effect of inter-arrival distributions on the performance of service facilities. To estimate future inter-arrival time based on HErD for the capacity design of service facilities to be constructed in the planning period, we developed a basic parameter estimation model according to two given design parameters namely, long-term peak-hour volume and peak-hour factor. However, the proposed model did not work well because the HErD has many free parameters to be estimated. Thus, we derived a method to reduce the number of free parameters, and then we proposed an improved parameter estimation model of HErD to describe future inter-arrival time distribution based on given long-term peak-hour volume and peak-hour factor.

1. Introduction

Subway construction is rapidly increasing in China. According to the survey of National Development and Reform Commission of

* Corresponding author.
E-mail addresses: chenyanru@swjtu.cn (Y. Chen), 147298717@qq.com (B. Yi), jiangyangsheng@swjtu.cn (Y. Jiang), sunjd503@163.com (J. Sun), wahab@ryerson.ca (M.I.M. Wahab).

https://doi.org/10.1016/j.tra.2018.03.009
Received 18 December 2015; Received in revised form 4 March 2018; Accepted 7 March 2018
0965-8564/

China (2015), the total mileage of subway tracks is 3500 km in 2015, which indicates an increase of 1134 km since 2013 and will increase by 2500 km in 2020. The investment for subway construction is 300 billion yuan in 2015, which indicates an increase of 14.3 billion yuan since 2014.

In subway construction, the capacity design of an underground subway station is an important issue (e.g., Hu et al., 2013), including the capacity design of service facilities such as subway passageway, fare gate, and ticket vending machine. The capacities of service facilities should be designed according to the passenger arrival pattern to ensure that the capacities of the facilities, such as the number of fare gate or the width of the passageway, adequately match the number of passengers and to avoid excessively long waiting time or serious wastage of service resources. Therefore, it is necessary and important to select an inter-arrival time distribution that can accurately describe the passenger arrival pattern at service facilities.

The inter-arrival time distribution of passengers in existing literature can be divided into four categories:

The first category is fixed-length distribution. Several design standards, such as NO.100 TCRP of the "Public Transport Capacity and Service Quality Manual" of the USA (Transportation Research Board of the National Academies, 2003) and GB 50157-2003 of the "Design Specifications" of China (e.g., Shi et al., 2003), assume that inter-arrival time of passengers and service time of facilities follow a fixed-length distribution. For simplicity, several studies also used a fixed-length distribution to describe passenger arrival pattern. For example, Zhong et al. (2008) generated the number of passengers arriving at a station based on a fixed-length distribution in the simulation of the evacuation process of an underground subway station. Wong et al. (2008) and Xu et al. (2016) also assumed that the passenger arrival rate at each station is constant. A fixed-length distribution is easy to handle because it assumes that the inter-arrival time of passengers or service time is unchanged. However, at an underground subway station in a city, particularly in a metropolitan city, inter-arrival time of passengers usually is not constant but random, and hence a fixed-length distribution is unsuitable for describing the inter-arrival time of passengers precisely (e.g., Hu et al., 2015).

The second category is a uniform distribution. It is also used to describe inter-arrival time of passengers. A uniform distribution is a probability distribution in which all inter-arrival times within a specified rage have the same probability. Sims (1995) generated the number of passengers arriving at a station based on a uniform distribution function when simulating evacuation in an underground subway station in the UK. Xu and Xu (2011) and Si et al. (2016) assumed that inter-arrival time of the passengers follows a uniform distribution and then predicted the number of passengers arriving at a station. A uniform distribution may be unsuitable to represent the uncertainty associated with the inter-arrival time of the passengers in an underground subway station in a metropolitan city because of the significant randomness of inter-arrival time of passengers.

The third category is a normal distribution. Liu and Yu (2008) and Liao et al. (2013) assumed that inter-arrival time of passengers follows a normal distribution. Liu and Zhang (2004) used the normal distribution to fit inter-arrival time of passengers at the fare gate in an underground subway station in Shanghai. The normal distribution is a common and remarkably useful continuous probability distribution because of the central limit theorem. Nevertheless, the normal distribution, which is symmetrical, may be an unsuitable model for the inter-arrival time of passengers that may be strongly skewed.

The last category is exponential distribution. It has been extensively applied to describe inter-arrival time of passengers in previous studies. Jiang et al. (2010) assumed that inter-arrival time of passengers follows an exponential distribution when designing the width of a passageway in an underground subway station. Løvås (1994) developed the M/G/1 model to describe passenger queuing behavior during an emergency in an underground subway station. Cheah and Smith (1994), Cruz et al. (2005), and Chen et al. (2012) used the M/G/C/C model to investigate passengers and to design the capacities of service facilities. Dessouky et al. (2003), Yalçınkaya and Bayhan (2009), Fernández (2010), and Xu et al. (2014) proposed that the passenger arrival process can be assumed to follow a Poisson distribution. The exponential distribution is used to describe the time between events in a Poisson process in which the value of the variation coefficient of the inter-arrival time distribution is equal to 1. However, in the real world, the value of the variation coefficient of passenger flow in an underground subway station may not meet this requirement. Thus, the exponential distribution may not be the best choice to describe the inter-arrival time of passengers in an underground subway station (e.g., Hu et al., 2015).

In summary, existing literature assumed that inter-arrival time of passengers follows a fixed-length, uniform, normal, or exponential distribution. However, these distributions may not adequately describe the inter-arrival time of passengers at service facilities in an underground subway station because inter-arrival time of passengers in underground subway stations is strongly skewed and random (e.g., Jiang et al., 2013). Besides, most of the distributions may not be algorithmically tractable and cannot obtain explicit solutions in a queueing system.

Therefore, in this study, a new distribution—phase-type (PH) distribution—is introduced to describe inter-arrival time of the passengers at service facilities in an underground subway station.

A PH distribution is constructed by using a combination of exponential distributions (e.g., Harchol-Balter, 2013). The distribution results from a system of one or more interrelated Poisson processes that occur in sequence or phase, where each sequence or phase may be a stochastic process itself. PH distribution has two main advantages as statistical models. The first one is that PH distribution can be the computational vehicle of much of applied probability for the problems which need explicit solutions because it is algorithmically tractable as well as exponential distributions (Asmussen et al., 1996). The second one is that the class of pH distributions is dense and hence it can be approximated arbitrarily to any distribution on $[0, \infty)$ (Asmussen et al., 1996). The PH distribution has been used in different queuing systems (e.g., Latouche and Ramaswami, 1997; Alfa and Zhao, 2000; Miyazawa et al., 2007; Krishnamoorthy et al., 2008; Hu et al. 2015).

At present most approaches in fitting and application of pH distributions focus on the acyclic phase-type (APH) distribution because it offers better tractability than the general PH (Reinecke et al., 2012). One of the important sub-classes of APH is the class of Hyper-Erlang distributions (HErD). Jiang et al. (2013) fitted HErD to the observed inter-arrival time of the passengers at the entrance