European Journal of Operational Research 000 (2015) 1–9

Contents lists available at ScienceDirect

European Journal of Operational Research

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

Optimization and strategic behavior in a passenger–taxi service system

Ying Shi [a], Zhaotong Lian [b,*]

[a] School of Information, Zhejiang University of Finance and Economics, China
[b] Faculty of Business Administration, University of Macau, Macau SAR, China

ARTICLE INFO

Article history:
Received 17 July 2014
Accepted 11 July 2015
Available online xxx

Keywords:
Double-ended queueing system
Equilibrium
Optimization
Strategic behavior
Threshold

ABSTRACT

We study a passenger–taxi problem in this paper. The objective to maximize the social welfare and optimize the allocation of taxi market resources. We analyze the strategic behavior of passengers who decide whether to join the system or balk in both observable and unobservable cases. In observable case, we obtain the optimal selfish threshold that maximizes their individual revenues and give the conditions of the existence of the optimal selfless threshold that maximize the social welfare. In unobservable case, we discuss the equilibrium strategies for the selfish passengers and derive the optimal arrival rate for the socially concerned passengers. Further, we analyze how the government controls the number of taxis by subsidizing taxis or levying a tax on taxis.

1. Introduction

Many people have the experience that sometimes it takes long time to wait for a taxi during peak hour in an airport, a ferry terminal, or a railway station. For example, in Changi Airport of Singapore, “The long queue for taxis observed on Sunday was due to the combined effects of a peak in weekend flight arrivals at T1, as well as the crunch in taxi supply generally observed island wide between 3 p.m. and 5 p.m., when taxi drivers change shift”. Passengers complained that “the taxis just vanished before the peak hour”.[1] In some cities, the problem becomes worse with growing passenger and visitor numbers. However, during non-peak hour, there are many taxis waiting in a long queue but no passengers. As Anwar, Volkov, and Rus (2013) mentioned, when too many taxis wait at the airport, it reduces the number of taxis available to service the rest of the city and reduces the income of taxis waiting in queue because they could be more productively finding fares elsewhere. When too few taxis are available, this results in travelers having to wait in line for long periods of time. The above issue motivates us to study the matching problem between passengers and taxis. What is the optimal taxi buffer size? In what case the government should subsidize the taxi drivers? And in what case the government should levy a tax on the taxi drivers? Our purpose is to answer those questions and provide some efficient strategies to coordinate passengers and taxis and improve the system performance in the point of view of social welfare. This paper is also motivated by a transportation project being studied in Macau.

We study this passenger–taxi issue by modeling it as a double-ended queueing problem. Queueing theory is a mathematical method that studies the random phenomena and the working processes of random service systems, which have been widely used in many areas. Among a variety of queueing systems, double-ended queueing system are very common and have a wide applications in computer science, perishable inventory system, organ allocation system.

The double-ended or matching queue problem was first proposed by Kendall (1951), who introduced an example about passengers and taxis: in an orderly taxi rank, on one side a queue is formed by the arrival of a stream of passengers who wait for taxis, while on the other side, a queue of taxis wait for passengers. Kendall (1951) pointed that the size of the queue was the difference between the passenger arrival rate and the taxi arrivals. The author also pointed that the mean of queue length was equal to zero when the two arrival rates were equal and the variance of queue length increases infinitely with the time. Moreover, based on results of Skellam (1946), Kendall (1951) gave the probability distribution of the queue length and proved that the double-ended queue did not have steady-state probability under the constant arrival rates. Dobbie (1961) considered the nonhomogeneous Poisson arrival of passengers and taxis and obtained transient probability by using Laurent generating function. Jain (1962) examined a double-ended queueing system in which only finite capacity for taxis were available and the inter-arrival distribution of taxis was general. The author obtained the time-dependent probability generating function. Similar to the model by Jain (1962), Bhat (1970) considered the control problem in which the system is controlled

* Corresponding author.
E-mail address: lianzt@umac.mo (Z. Lian).

[1] See http://www.sgcarmart.com/news/article.php?AID=11116.

http://dx.doi.org/10.1016/j.ejor.2015.07.031
0377-2217/

by calling extra taxis whenever the total number of passengers lost reaches a certain predetermined number. The author gave the transient and steady-state probability of the process by using renewal theoretic arguments and obtained the optimal value of the control variable to minimize the total cost. Giveen (1963) further discussed the double-ended queueing system with nonhomogeneous Poisson arrival process and showed that the steady-state probabilities exist only if the mean and the variance of the queue length remain finite as time become infinite. Kashyap (1966) 1967) considered the general arrival of passengers and the Poisson arrival of taxis and assumed that each taxi can take a fixed number of passengers. The probability of an empty state was obtained by the supplementary variable technique. Wong, Wong, Bell, and Yang (2005) adopted an absorbing Markov chain to model the searching process of taxi movements and proposed a useful formulation for describing the urban taxi services in a network. Crescenzo, Giorno, Kumar, and Nobile (2012) considered a double-ended queue with catastrophes and repairs and obtained steady-state and failure-state probabilities.

The above literature mainly studied the performance measures of double-ended queue models. Mendoza, Sedaghat, and Yoon (2009) imposed thresholds on both sides of the queues to assure stability in the absence of abandonment and investigated minimization of the expected total cost. Kim, Yoon, Mendoza, and Sedaghat (2010) utilized a simulation approach to investigate the system. The authors also conducted sensitivity analysis and found that batch size considerably affected the performance of the model. Moreover, if expectations of the batch size distributions were the same, the model was insensitive to the types of such distributions.

Conolly, Parthasarathy, and Selvaraju (2002) studied a double-ended queueing system with impatience. The concept of “impatience” is also known as “reneging”. The taxi queue is shortened by the arrival of passengers or by “impatience” of taxis and vice versa for the passenger queue. The authors assumed that both taxis and passengers arrive based on Poisson streams, and that the taxis impatiently depart with exponential parameter $\mu_{1,n}$ when n taxis are in the queue. Similarly, passengers impatiently depart with exponential parameter $\mu_{2,n}$ when n passengers are in the queue. By constructing a two-dimensional Markov process, the authors derived the moment generating function and the expected first passenger times. Afèche, Diamant, and Milner (2014) first gave the closed form results for a double-ended queueing model with batch arrivals and abandonment. The passengers and taxi drivers in our paper can be also considered “impatient”, but different from the literature, they determine whether to join the system or balk based on their utilities, rather than only based on the number of passengers/taxis in the system.

Double-ended matching systems are very common in the other areas: computer science, perishable inventory system, organ allocation system. For example, different types of components are piled for matching in an assembly line; organs and patients wait for matching in health care. Zenios (1999) considered a double-ended matching problem between several classes of organs and patients who would renege due to death. Gurvich and Ward (2014) considered the optimal control of matching queues with dynamically arriving jobs. Other literature related to double-ended queue areYang, Leung, Wong, and Bell (2010), Yang and Yang (2011) and Wong, Szeto, and Wong (2014).

We can see that most of above literature focus on studying the performance measures of the passenger–taxi queueing system, for example, the passenger/taxi waiting time. Some of them did study the optimization, but they did not consider the passenger/driver behavior. In this paper, we contribute to the literature by introducing the strategic behaviors of passengers which really affect the service provider’s decision on the taxi buffer size. To the best of our knowledge, no studies explicitly considered the optimization problems of the double-ended queueing systems with the strategic behavior of passengers. The main contributions of this paper are three folds:

(1) analyze the behavior of socially concerned and selfish passengers;
(2) derive the optimal taxi buffer size by maximizing the social welfare; and
(3) provide the government a policy to balance the benefit between passengers and taxi drivers by granting subsidy to drivers or levying a tax on drivers.

Several excellent studies on equilibrium strategic behavior, pricing, and the interaction between service quality and congestion. Inspired by the problem raised in connection with the control of vehicular traffic congestion on a road network, Naor (1969) discussed a queue where a uniform toll was imposed on customers who want to join it. The author pointed out that self interest did not ordinarily lead to overall optimality. Shone, Knight, and Williams (2013) considered two different types of optimal customer behaviors (selfishly optimal and socially optimal) in observable and unobservable cases. Under both types of customer behavior, the authors established necessary and sufficient conditions for equality of optimal queue-joining rates between the observable and unobservable cases. And the authors drew a comparison between two cases on performance measure under the equality of optimal joining rates. Similar to Shone et al. (2013), we also consider these two different types of customer behaviors to study their impact on the social welfare.

Yang, Guo, and Wang (2014) investigated customers’ equilibrium queueing strategies in both monopoly market and duopoly competition market. Then, the authors analyzed the servers’ pricing decision. They found that profit- and welfare-maximizing prices were not the same in a monopoly market. Several other studies related to equilibrium strategies include Luski (1976), Allon and Federgruen (2007), Debo, Toktay, and Wassenhove (2008), Veeraraghavan and Debo (2009), Zhou, Chao, and Gong (2014a), and Zhou, Lian, and Wu (2014b).

The rest of the paper is organized as follows. In Sections 2 and 3, we introduce the model in detail and analyze the stationary probability and other system performances. In Section 4, we analyze the impact of taxi buffer size on social welfare and obtain the optimal taxi buffer size; further, we discuss the strategic behavior of selfish and socially concerned passengers in observable and unobservable cases. In Section 5, we present some numerical examples showing the effectiveness and correctness of the theoretical results. Finally, we conclude the paper in Section 6.

2. Model description

In this paper, we consider a passenger–taxi queueing system. Passengers arrive in batches according to a Poisson process with rate λ_1. Taxis arrive according to a Poisson process with arrival rate λ_2. Each batch of passengers consists of one to five persons, who can be exactly taken by a taxi. For convenience, we can consider each group of passengers as a passenger. The passengers and taxis are arranged based on a first-come-first-served discipline. The waiting space for taxis is finite. Denote by N the maximum number of taxis allowed to be in the taxi station. Since the space occupied by a passenger is much smaller than a taxi, it is reasonable to assume that the waiting space for passengers is infinite.

We consider both observable case and unobservable case. The observable case means that passengers or taxi drivers can observe the passenger/taxi queue length upon arrival. The unobservable case means that passengers and taxi drivers do not know the exact queue length upon arrival, but they know the potential passenger/taxi arrival rate, and determine whether join the system or balk based on the their own utilities.

We assume that each passenger needs to pay a taxi fare of p_1 on average and receives a reward of R by taking a taxi. Denote by C_1 the passenger waiting cost per unit time and C_2 the taxi waiting cost per