Computers & Operations Research 65 (2016) 164–173

Contents lists available at ScienceDirect

Computers & Operations Research

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

Time-dependent performance approximation of truck handling operations at an air cargo terminal

Gregor Selinka [a,*], Axel Franz [b], Raik Stolletz [a]

[a] Business School, University of Mannheim, Germany
[b] Center for Doctoral Studies in Business, Graduate School of Economic & Social Sciences, University of Mannheim, Germany

ARTICLE INFO

Available online 16 June 2014

Keywords:
Airport operations
Heterogeneous queueing model
Non-stationary performance analysis

ABSTRACT

This paper provides an analytical solution for the time-dependent performance evaluation of truck handling operations at an air cargo terminal. The demand for loading and unloading operations is highly time-dependent and stochastic for two classes of trucks. Two heterogeneous handling facilities with multiple servers are available to handle trucks assuming exponentially distributed processing times. Trucks are routed to a handling facility depending on the current state of the system upon arrival. To approximate the time-dependent behavior of such heterogeneous queueing systems, we develop a stationary backlog-carryover (SBC) approach. A numerical study compares this approach with simulations and demonstrates its applicability to real-world input data.

1. Introduction

The demand for air cargo transportation services is cyclical in nature. This demand is characterized by strong interdependencies between the economic situation and long-term airfreight volumes [15]. Moreover, considerable peaks and off-peaks in air cargo transportation activities occur within a day [18]. These dynamics are reflected in the demand for freight handling capacity at air cargo terminals. Air cargo terminals serve as cross-docking facilities for sorting, (re-)consolidation, and short-term storage before and after transportation by air (e.g., [24]). Airfreight shipments are delivered and picked up by trucks (e.g., [22]). Such road transportation takes a fundamental position in the air cargo logistics chain. Freight forwarders provide trucking services for air cargo shipments from the shipper to the origin airport and from the destination airport to the consignee (e.g., [29]). Furthermore, cargo airlines themselves operate scheduled intra-continental road feeder services between airports in their hub-and-spoke networks (e.g., [4]). Especially within Europe, such trucking services have increased significantly at an annual growth rate of 20% between 2002 and 2012, amounting to nearly 20,000 scheduled intra-European frequencies per week [5].

In this paper, we analyze the truck handling operations at the hub of one of Europe's largest combination carriers. An evaluation of such a system's time-dependent performance provides crucial information for various managerial decisions. Operations managers of air cargo terminals have to evaluate the time-dependent operational performance to adjust capacity levels, to change operational handling procedures, and (if possible) to schedule truck arrivals. Thus far, the performance of air cargo operations (e.g., [17,23]) and of truck handling operations in other contexts (e.g., [11,12]) under non-stationary conditions has mainly been analyzed by simulation. The objective of this work is to develop and evaluate an accurate and fast analytical approximation method for the time-dependent performance evaluation of truck handling operations at an air cargo terminal.

The corresponding system features two handling facilities for loading and unloading activities of unit load devices (ULDs), such as pallets and containers used for consolidated transportation. While facility 1 is equipped with a single truck dock, facility 2 features two parallel truck docks. Because of different operational requirements, such as requirements regarding shape, size, and weight, we distinguish two heterogeneous classes of trucks according to the type of airfreight carried: (1) export deliveries, which can be handled only at handling facility 1, and (2) import and transit shipments, which can be handled at both facilities. The number of truck arrivals is highly time-dependent, resulting in significant variations in activity level throughout a day with peaks typically occurring at night. Such fluctuations are somewhat predictable, the actual extent, however, is subject to uncertainty. Processing times are stochastic and facility-dependent, but independent of the truck class, as empirical analyses revealed. We assume that the processing times are time-independent. Since the

* Corresponding author.
E-mail addresses: selinka@bwl.uni-mannheim.de (G. Selinka), axel.franz@gess.uni-mannheim.de (A. Franz), stolletz@bwl.uni-mannheim.de (R. Stolletz).

http://dx.doi.org/10.1016/j.cor.2014.06.005
0305-0548/

two handling facilities lie some distance apart, arriving trucks are assigned to one of the available facilities upon arrival. Trucks with export shipments are exclusively routed to handling facility 1. For trucks with import or transit shipments, the routing decision depends on the current numbers of trucks being handled or waiting at each handling facility. Trucks waiting for cargo handling services are processed on a first-come, first-served basis at both facilities.

Prior to this study, similar queueing systems with heterogeneous servers and heterogeneous jobs that join a queue directly upon arrival have only been analyzed in steady state. Static routing decisions are analyzed by Ross and Yao [25], Ansell et al. [2], Argon et al. [3], and Liu and Righter [19]. In the case of state-dependent routing, threshold policies based on a particular facility may be applied (e.g., [26,28]) or routing decisions may be based on the state of several stations; e.g., an arriving job may be routed to the facility with the shortest queue (e.g., [6,1]). Furthermore, in contrast to our setting, all these references primarily restrict the scope of analysis to parallel single-server queues. The term "N-system" is often used to describe similar queueing systems in call centers. However, while trucks are routed directly at arrival in the considered truck handling system, calls are routed just before being served in call center systems, but wait in job specific queues (e.g., [7,8]).

There are different approaches for the non-stationary analysis of homogeneous queueing systems. The numerical solution of the respective set of ordinary differential equations (e.g., [16,21]) and the randomization approach [10] are applicable to Markovian systems. Although these methods provide (nearly) exact results, the numerical solution is rather time-consuming [13]. Deterministic fluid approaches approximate discrete events through continuous processes. These approaches are fast and suitable for the time-dependent analysis of overloaded systems (e.g., [20,14]). However, any queue in an underloaded system is not considered. Another class of approximations is based upon the application of steady-state models. Comparing various approximation methods, Ingolfsson et al. [13] show that the stationary independent period-by-period (SIPP) approximation achieves good results within a reasonable time. This method divides the observed time horizon into multiple smaller periods and then analyzes each period independently using a stationary model [9]. In contrast, the stationary backlog-carryover (SBC) approach considers the dependencies between successive periods [27]. This method builds backlogs of non-served arrivals and carries them over to the succeeding period. Numerical studies indicate better approximation results than the SIPP approach for $M(t)/M/c(t)$ systems.

The contribution of this paper is the analysis of a queueing system with two heterogeneous classes of trucks, two separate handling facilities with multiple servers, and state-dependent routing upon arrival. Based on a stationary Markov model, we develop an SBC approach for the time-dependent performance evaluation. The approximation method is applied to arbitrary state-dependent routing policies.

The remainder of this paper is organized as follows. Section 2 introduces the queueing model of the analyzed truck handling system. The corresponding Markov chain and the calculation of the steady-state performance are presented in Section 3. The first part of Section 4 provides a brief introduction to the SIPP approach to analyze non-stationary systems. The SBC approach for the heterogeneous queueing system is developed in the second part of Section 4. In Section 5, a numerical study is conducted for the purpose of comparing the SIPP and SBC approximations with simulation results. Furthermore, a sensitivity analysis with respect to handling capacities, demand, and routing policies is presented to gain insights into the real-world behavior of the system. A

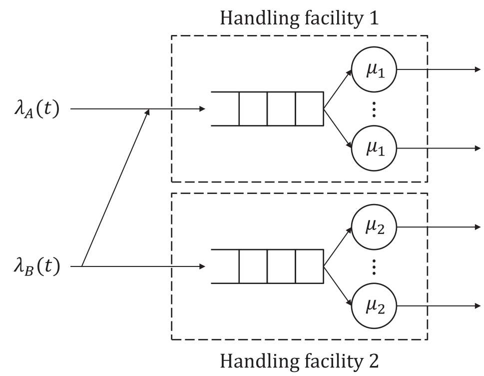

Fig. 1. The model of the truck handling system.

conclusion and suggestions for further research are provided in Section 6.

2. The queueing model

The truck handling system is represented by a queueing model with heterogeneous jobs (i.e., truck classes), with heterogeneous servers (i.e., truck docks) at two parallel stations (i.e., handling facilities), and with routing decisions before entering a queue (see Fig. 1).

We distinguish between two independent inhomogeneous Poisson arrival processes with instantaneous arrival rates $\lambda_A(t)$ and $\lambda_B(t)$, respectively. Trucks of class A carry export shipments, whereas trucks of class B are dedicated to import and transit shipments. Depending on the truck handling facility, the servers represent flexible or specialized truck docks for loading and unloading activities. Handling facility 1 features c_1 flexible servers, which are able to handle trucks of classes A and B. Handling facility 2 is equipped with c_2 parallel specialized truck docks, which are only able to handle trucks of class B. The truck docks are assumed to operate with exponentially distributed service times at constant rates μ_1 and μ_2 independent of truck class. In front of each handling facility, there is a single queue that is served on a first-come, first-served basis. We assume an infinitely large waiting room.

The state of the system is described by a tuple (n_1, n_2), where n_1 denotes the overall number of trucks at facility 1, i.e., the trucks being processed at a server or waiting, and where n_2 denotes the overall number of trucks at facility 2. All possible states are included in the infinite state space:

$$\mathcal{S} = \{(n_1, n_2) | n_1 \in \{0, 1, 2, \ldots\}; n_2 \in \{0, 1, 2, \ldots\}\} \tag{1}$$

Immediately upon arrival, trucks are assigned to one of the two handling facilities. An arriving truck of class A is always served at handling facility 1, whereas a truck of class B can be handled at either facility. Let $R(n_1, n_2)$ define the state-dependent routing decision for an arriving class B truck, i.e.,

$$R(n_1, n_2) = \begin{cases} 1, & \text{if an arriving truck of class } B \text{ is routed} \\ & \text{to the flexible facility 1,} \\ 0, & \text{if an arriving truck of class } B \text{ is routed} \\ & \text{to the specialized facility 2.} \end{cases} \tag{2}$$

For example, in the truck handling system at the considered air cargo hub, an arriving truck of class B is routed to handling facility 1 if the following two conditions are met:

- There is no server available at specialized handling facility 2.
- The ratio of the numbers of trucks at handling facilities 1 and 2 is smaller than a predefined parameter ω.