ARTICLE IN PRESS

Performance Evaluation ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

Performance Evaluation

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

Analysis of a semi-open queueing network with Markovian arrival process

Jiseung Kim [a], Alexander Dudin [b,c], Sergey Dudin [b,c], Chesoong Kim [d,*]

[a] Kyungil University, Gyeongsan, 38428, Republic of Korea
[b] Belarusian State University, 4, Nezavisimosti Ave., Minsk, 220030, Belarus
[c] Peoples' Friendship University of Russia (RUDN University), 6 Miklukho-Maklaya St, Moscow, 117198, Russia
[d] Sangji University, Wonju, Kangwon, 220-702, Republic of Korea

ARTICLE INFO

Article history:
Received 2 August 2016
Received in revised form 15 November 2017
Accepted 18 December 2017
Available online xxxx

Keywords:
Queueing network
Restricted access
Markovian arrival process
Wireless telecommunication networks

ABSTRACT

A semi-open queueing network having a finite number of nodes is considered. The nodes are modeled by single-server queueing systems with a finite buffer and an exponential service time distribution. Customers arrive to the network according to a Markovian arrival process. The number of customers, which can be processed in the network simultaneously, is restricted by a threshold. If the number of customers in the network is less than this threshold, when a new customer arrives, the customer is processed in the network. Choice of the first and the subsequent nodes for service is performed randomly according to a fixed stochastic vector and a transition probability matrix. If the number of customers in the network at the customer arrival epoch is equal to the threshold, the customer is queued into an input buffer with an infinite capacity. Customers in the input buffer are impatient. The stationary behavior of network states is analyzed. The Laplace–Stieltjes transform of the distribution of the customer's waiting time in the input buffer is obtained. Expressions for computing performance measures of the network are derived. Numerical results are presented. The model is suitable, e.g., for analysis and optimization of wireless telecommunication networks and manufacturing systems with a finite number of machines and workers.

1. Introduction

The theory of queueing networks is an important part of the queueing theory. For history of the theory of queueing networks and a short overview of the existing results see, e.g., [1–3]. This theory has found extensive applications, e.g., for modeling telecommunication networks and manufacturing systems. There is a huge number of publications devoted to the analysis of queueing networks, see, e.g. [4,5], etc.

The queueing networks can be broadly classified (see, e.g., [6]) into three categories: (1) open network (where the customers arrive at a specified rate), (2) closed network (where the number of customers circulating in the network is fixed) and (3) semi-open network (where the customers arrive at a specified rate, but the maximum number of customers that can be processed simultaneously is fixed). The semi-open queueing networks exhibit properties of both open and closed queueing networks, i.e. the customers still arrive from an infinite population set, but the number of 'circulating' or 'movable' resources (that need to be matched with an incoming customer) available to process customers is constant (below we denote this constant as N). In recent years, semi-open queueing networks are adopted for performance analysis of both

* Corresponding author.
E-mail addresses: jiskim@kiu.ac.kr (J. Kim), dudin@bsu.by (A. Dudin), dudins@bsu.by (S. Dudin), dowoo@sangji.ac.kr (C. Kim).

https://doi.org/10.1016/j.peva.2017.12.005
0166-5316/

Please cite this article in press as: J. Kim, et al., Analysis of a semi-open queueing network with Markovian arrival process, Performance Evaluation (2018), https://doi.org/10.1016/j.peva.2017.12.005.

manufacturing and service systems. For the review of the state of the art and the relevant references the reader is addressed to [6–9].

In this paper, we present an analysis of a semi-open queueing network. The topology of the network is arbitrary. A matrix defining routing of customers is also arbitrary, loops are not forbidden. The nodes are represented by the single-server queues with finite buffers having enough capacity to exclude the possibility of blocking the customers arriving to a node from a previous node. Our model has the certain disadvantages in comparison to some models considered in [6–9], e.g., we assume an exponential service time distribution at the nodes of the network, homogeneous customers, etc.

The main contributions of our paper are as follows:

- We present an *exact algorithmic* analysis of a semi-open queueing network with an arbitrary topology while the *approximate* analysis is provided, sometimes under more general assumptions about the parameters of the network, in the mentioned papers. Therefore, our results can be used for exact performance evaluation and capacity planning of the semi-open queueing networks as well as for validation of the approximate results or results of simulation.
- We assume that an arrival flow to the network is described by the Markovian arrival process (*MAP*), see, e.g., [10,11], which is now popular for the description of real-world arrival processes because they exhibit burstiness, while the overwhelming majority of the existing results are obtained for the stationary Poisson arrival process of customers. As the early paper where the queueing network with the *MAP* arrival process is considered, the paper [12] may be referred. Results for a specific kind of queueing networks, the so-called tandem queues, with the *MAP*, can be found, e.g., in [13,14]. The importance of careful account of correlation in the considered model is illustrated in Section 6.
- We assume that the arriving customers are impatient and can leave the input buffer after an exponentially distributed time. Customer impatience is an important feature of many real-world systems. Account of impatience leads to the fact that the multi-dimensional Markov chain describing the behavior of the network does not belong to the well-studied class of level-independent Quasi-Birth-and Death Processes and the far less known results should be applied for its analysis.

There is a lot of examples of potential applications of the results for semi-open queueing networks to analysis of real-world systems, some of them are relevant to our model. Additionally, we can mention the following practical applications of the model under consideration.

One of the potential applications in the wireless telecommunication area is the following. It is necessary to transmit information between the source node and destination node, e.g. a sensor node and a gateway node. There is no direct way for transmission, but there are L intermediate repeaters (relaying stations), using some chain on which the transmission of information is possible. It is necessary to find an optimal rate of the information transmission regulated via the choice of the number N of messages that can be transmitted via the repeaters simultaneously.

The second possible application area is as follows. We model an enterprise where N workers having identical skills provide service to some items using L, generally speaking, different machines. E.g., we may consider a car repair shop consisting of L machines (workstations, devices, equipments, etc.) implementing different technological operations (diagnostics, repair of engine ignition system, valve control system, fuel supply system, decelerators, speed-change transmission, undercarriage, etc.). An arriving car has to obtain service at some of these workstations. Service of each car is provided by a repairman who is responsible for implementation of all necessary operations. The problem is to choose the optimal number of repairmen who will provide service to cars. If the number of repairmen N is too small, the workstations may be used non-effectively. The workstations starve, stay idle while there is a queue of cars waiting for repair at the gate to the repair facility. In our model, we suggest that customers (cars or their owners) may be impatient and leave the queue without entering the repair facility due to too long waiting time. The owner of the car repair facility loses potential profit from service of these impatient customers. If the number of repairmen is too large, some workstations are often busy and repairmen do not work but wait in the queue to the corresponding workstations. The owner of the car repair facility loses money by payment salaries to the idle repairmen. Thus, the proper choice of the number of repairmen N is important for the owner of the car repair facility. The problem of the proper choice of the number of repairmen can be solved based on the results of an exact algorithmic analysis presented in this paper.

As another example, the model of Flexible Manufacturing Systems (*FMS*) from [3] can be mentioned. The *FMS* consist of automated machines, an automated material handling system to move the jobs between the machines, and the controllers that control the machines and the material handling system. The material handling system typically is comprised of pallets on which the jobs are mounted. In the terminology of the previous example, the automated machines correspond to workstations, jobs to cars and pallets to workmen respectively. In [3], it is suggested that the number of pallets that move the parts in the *FMS* is generally fixed and it is proposed to use close queueing network for modeling such *FMS*. I.e., the *FMS* are assumed to be operating in saturation mode. In real-world systems, the source of jobs is not infinite, thus, sometimes there is nothing to load to a pallet. Therefore, our model that describes operation of *FMS* in non-saturated mode, is more adequate and challenging.

The fourth application is possible in health care systems. There is a center for periodic screening of people. This center has L different specialists (therapist, surgeon, neurologist, ophthalmologist, otolaryngologist, etc.). Each patient should visit sequentially some set of these specialists. To prevent large queues inside this center, it makes sense to regulate the number N of patients that are allowed to be present in the center simultaneously.

Please cite this article in press as: J. Kim, et al., Analysis of a semi-open queueing network with Markovian arrival process, Performance Evaluation (2018), https://doi.org/10.1016/j.peva.2017.12.005.