Journal of Computational and Applied Mathematics 236 (2012) 3948–3955

Contents lists available at SciVerse ScienceDirect

Journal of Computational and Applied Mathematics

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

Boundary value methods for transient solutions of queueing networks with variant vacation policy

Raymond H. Chan [a], Spike T. Lee [a], Hai-Wei Sun [b,*]

[a] Department of Mathematics, The Chinese University of Hong Kong, Shatin, Hong Kong, China

[b] Department of Mathematics, University of Macau, Macao, China

ARTICLE INFO

Article history:
Received 5 October 2011
Received in revised form 17 January 2012

Keywords:
Queueing systems
Heterogeneous servers
Variant vacation policy
Boundary value method
Crank–Nicolson preconditioner

ABSTRACT

Boundary value methods are applied to find transient solutions of $M/M/2$ queueing systems with two heterogeneous servers under a variant vacation policy. An iterative method is employed to solve the resulting large linear system and a Crank–Nicolson preconditioner is used to accelerate the convergence. Numerical results are presented to demonstrate the efficiency of the proposed method.

1. Introduction

Queueing networks with vacations were proposed in the 1970s to overcome the deficiency of classical queueing networks in modeling complex hi-tech systems [1]. Server vacations may literally mean a lack of work, or figuratively stand for server failure, server maintenance, or a server taking another assigned job, and hence the introduction of server vacations makes waiting-line systems more lifelike. The applications of vacation queueing systems lie in various areas such as flexible manufacturing systems, lane control at border-crossing stations, and data transfer in telecommunication systems. The thorough development of queueing networks with vacations can be found in survey papers in [2–4], and the monographs in [1,5].

Homogeneity of service rates is a general assumption in the study of multiserver queueing system, and it ensures that all servers in the system provide services at an identical rate. However, the hypothesis of homogeneous systems is feasible only when the service process is mechanically or electronically controlled. In reality, human servers are more likely to perform the same assignment at different service rates. Therefore *heterogeneous* servers are introduced and their service time distributions may be different for different servers. The combination of server vacations and heterogeneous servers is more practical in real-life situations [6–8].

In [9], Yue et al. proposed an $M/M/2$ queueing system with one queue and two heterogeneous servers under a variant vacation policy, in which the two servers will *simultaneously* take *at most* J vacations when the system becomes empty. They carried out a steady-state analysis and obtained the stationary distributions of system size and mean system size. Moreover, they studied the distribution of the amount of vacations taken, and the conditional stochastic decomposition properties of the queue length and the waiting time. The analytical results in [9], however, are based on the assumption of infinite queueing spaces, which may not be practical in general.

* Corresponding author.
E-mail addresses: rchan@math.cuhk.edu.hk (R.H. Chan), thlee@math.cuhk.edu.hk (S.T. Lee), HSun@umac.mo (H.-W. Sun).

0377-0427/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.cam.2012.01.027

For the sake of practicality, in this paper, we aim to consider the problem with finite queueing spaces and find the transient solution of the queueing system in [9]. It is well known that the transient solution of a queueing system can be numerically approximated by discretizing the Kolmogorov backward equation and solving the resulting ordinary differential equation (ODE) system [10,11]. Classical initial value methods (IVMs) such as Runge–Kutta methods are natural candidates, but they are computationally more expensive than multistep methods of comparable accuracy. In this work, we follow the idea in [12] and apply boundary value methods (BVMs) to solve the ODE systems. BVMs are the generalization of implicit linear multistep formulas (LMFs), and by using those with unconditional stability [13] one can disregard the restrictions on the step sizes for stability reasons. However, temporal discretization with BVMs requires solutions of larger linear systems than with Runge–Kutta methods or LMFs used as IVMs. Fortunately, owing to the block tridiagonal structure of the transition rate matrix, the resulting linear system is sparse, and therefore we can resort to iterative methods.

Preconditioning techniques have long been used to speed up the convergence process of iterative methods when solving large sparse linear systems produced by BVMs [14]. Over the years, different preconditioners have been proposed, including the T. Chan's circulant preconditioner [15,16], the P-circulant preconditioner [15], the Strang's circulant preconditioner [15,17], the skew-circulant preconditioner [18], and recently the Crank–Nicolson (CN) preconditioner [19,20]. In [20], the CN preconditioner is paired with BVMs for pricing options in the jump-diffusion model, and the numerical results therein show that the CN preconditioner contributes to smaller computational cost and fewer iterations than the Strang-type preconditioner. In this paper, we mainly discuss the use of the CN preconditioner because we will see from the numerical results that it involves cheaper computational cost than other methods.

The rest of the paper is organized as follows. In Section 2, we outline the two-server queueing system and its variant vacation policy. In Section 3, we briefly introduce the BVMs and apply them to discretize the Kolmogorov backward equation to obtain an ODE system. In Section 4, we form the CN preconditioner and study some of its properties when used in the iterative method. In Section 5, we present the numerical results. In Section 6, we give some concluding remarks and ideas for possible future work.

2. Transient solution for a queueing network with variant vacation policy

In this paper, we consider an $M/M/2$ queueing system with two heterogeneous servers under the variant vacation policy proposed in [9]. Customers arrive and join a single queue according to the first-come first-served (FCFS) principle. The arrival of customers is modeled by a Poisson process with rate λ. When the system becomes empty, the two servers will simultaneously take a vacation of length V, where V is an exponentially distributed variable with parameter θ. When the servers are back from their vacation, they either resume working immediately if they find at least one customer waiting in the queue, or leave for another vacation of the same length V. The two servers will only take at most J vacations, and after that, they will stay active in the system to provide services even when the system becomes empty.

The service rates of the two heterogeneous servers are modeled by exponential distributions with rates μ_1 and μ_2 for Server 1 and Server 2, respectively. Note that $\mu_1 \neq \mu_2$, since the two servers are heterogeneous. When both servers are free at any moment, Server 1 will step up to serve the newly arriving customer. Finally, all the stochastic processes involved in the system are assumed to be independent.

It is noted that $M/M/2$ vacation queueing systems are modeled by quasi-birth-and-death (QBD) processes, the generalization of the birth-and-death process from a one-dimensional state space to a multidimensional state space [1]. Let $X(t)$ be the number of customers in the system at time t, and let $L(t) = j,\ j = 0, 1, \ldots, J+1$ be the status of the servers at time t. The state (i, j) means that $i \geq 0$ customers are in the system and both servers are taking the $(j+1)$th vacation for $j = 0, 1, \ldots, J-1$. Moreover, the state $(0, J)$ means that the system is empty while both servers are free. The state $(1, J)$ means that one customer is in the system while Server 1 is busy and Server 2 is free. The state $(1, J+1)$ means that one customer is in the system while Server 2 is busy and Server 1 is free. The state (i, J) means that $i \geq 2$ customers are in the system while both servers are busy.

For the two-dimensional Markov process $\{(X(t), L(t)),\ t \geq 0\}$ with the state space

$$\Omega_0 = \{(1, J+1)\} \cup \{(i, j),\ i \geq 0,\ j = 0, 1, \ldots, J\},$$

the infinitesimal generator of the process is given by [9]

$$Q_0 = \begin{bmatrix} B_{00} & B_{01} & & & & & \\ B_{10} & B_{11} & B_{12} & & & & \\ & B_{21} & A_1 & A_0 & & & \\ & & A_2 & A_1 & A_0 & & \\ & & & A_2 & A_1 & \ddots & \\ & & & & \ddots & \ddots & \ddots \\ & & & & & \ddots & \ddots \end{bmatrix}.$$