BOUNDS FOR THE EXPECTED NUMBER OF HOSPITAL BED OCCUPANCY

J. G. Shanthikumar, Department of Industrial Engineering and Operations Research, Syracuse University, Syracuse, New York, 13210.

Abstract

In this paper we model the hospital bed occupancy by an infinite server queueing system with time varying Poisson arrival and service processes. Upper and lower bounds for the mean number of beds occupied are derived and simple conditions on the patient admittance and discharge rates are given to guarantee the boundedness of the mean number of beds occupied.

Introduction

In this paper we consider the hospital bed occupancy problem. This problem cannot be modelled as a stationary queueing system, since patient admissions and discharge rates vary over time. To be precise, one may expect these rates to vary cyclically over time with a 24-hour cycle. Very few people may be admitted or discharged during nights. Therefore, we will model this problem as a queueing system with time varying arrival and service rates. It may be more realistic to assume a finite number of beds and then look at the adequacy of that number to handle the arrival and service pattern. However, the analysis of such a system is complex. Therefore, we will model this bed occupancy problem assuring the availability of sufficient number of beds. That is, we will model this problem as an infinite server queue. We will further assume that both arrival and service processes are non-homogeneous Poisson processes. Hence, we will represent this bed occupancy problem by a Poisson arrival, exponential service time infinite server queue ($M/M/\infty$) with time varying arrival and service rates.

The $M/M/\infty$ Queue with Time Dependent Arrival and Service Rates

Consider an $M/M/\infty$ queue with an arrival rate $\lambda(t)$ and service rate $\mu(t)$ that varies with time t. Let N(t) be the number in the system at time t and its mean be L(t). Now define

$$\lambda_t^* = \max_{0 \le x \le t} (\lambda(x)) \quad \text{and} \quad \mu_t^* = \min_{0 \le x \le t} (\mu(x))$$

and assume that $\lambda_t^* < \infty$ and $\mu_t^* > 0$. Let us now construct another $M/M/\infty$ queue (shown in Figure 1 as system 2) with an arrival rate of $\lambda_2(x) = \lambda_t^* - \lambda(x) \ge 0$ and a service rate of $\mu_2(x) = \mu_t^* > 0$, parallel to our $M/M/\infty$ queue (shown as system 1 in Figure I). As shown in Figure 1, it is assumed that a customer leaving

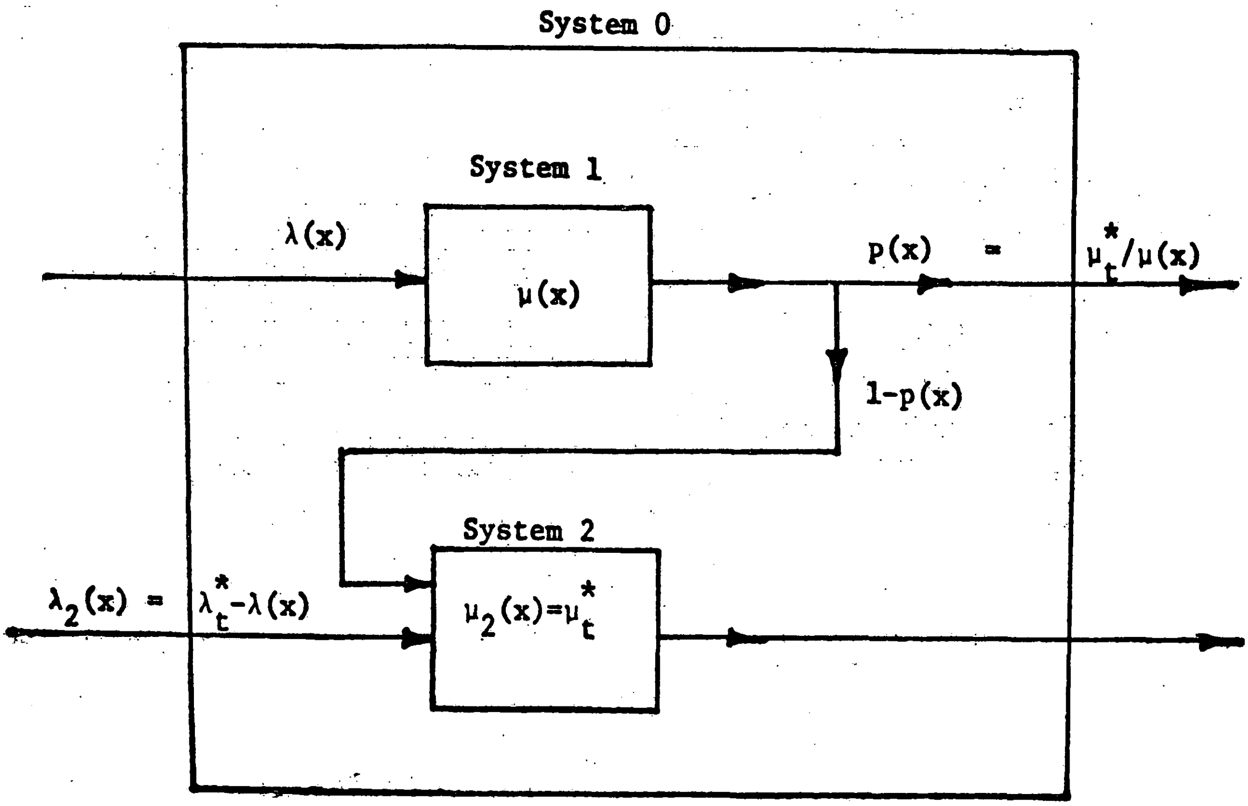


FIGURE 1

system 1 will either leave the overall system 0 with probability $p(x)=\mu_t^*/\mu(x)$ or enter system 2 with probability 1 — p(x) at time x ($0 \le x \le t$). Let $N_2(x)$ and $N_0(x)$ be the number of customers in the systems 2 and 0, respectively, at time x. Then N)x) = $N_0(x) - N_2(x)$.

It can be easily varified that the probability of an arrival to the overall system 0 in (x, x + Δx) is $(\lambda(x) + \lambda_2(x))\Delta x + O(\Delta x) = \lambda_t^* \Delta x + 0(\Delta x)$ and departure probability in (x, x + Δx) is $N(x)\,\mu(x)\,p(x) + N_2(x)\mu_t^*\Delta x + 0(\Delta x) = N_0(x)\mu_t^*\Delta x + 0(\Delta x)$. That is the overall system 0 behaves like an M/M/∞ with an arrival rate λ_t^* and service rate μ_t^* during (0, t). Since λ_t^* and μ_t^* are constants during (0,t). from the results available in Collings and Stoneman (1) we have for the mean $L_0(x)$ of $N_0(x)$, assuming N(0) = 0

$$L_0(x) = L(0)\exp(-\gamma_t^* x) + \frac{\lambda_t^*}{\mu_t^*}(1-\exp(-u_t^* x)), \quad 0 \le x \le t \qquad (1)$$

Since $\lambda_t^* > \infty$ and $\mu_t^* > 0$, we have $L_0(x) > \infty$,

$0 \le x \le t$ as long as $L(0) \le \infty$. We have

$N(x) = N_0(x) - N_2(x)$ and $N_2(x) \ge 0$

Then $N(x) \le N_0(x)$ and therefore

$$L(x) \le L_0(x) \qquad (2)$$

Then from (1) and (2) for x = t

$$L(t) \le L(0)\exp(-\mu_t^* t) + \frac{\lambda_t^*}{\mu_t^*}(1-\exp(-\mu_t^* t)) \qquad (3)$$

Then in the limit as $t \to \infty$

$$\lim L(t) = L \le \lambda^*/\mu^*,$$

where $\lambda_t^* \ge \lambda^*$ and $\mu_t^* \ge \mu^*$ for all $t > 0$,

Equation (3) can be used to develop an iterative bound for L(t). Suppose we choose intervals ε_n, n = 1,2,...., $\varepsilon_0 = 0$, from (3) we can show by induction that

$$L_u(t) = L_u(t_{n-})\exp(-\mu_n^*\varepsilon) + \frac{\lambda_n^*}{\mu_n^*}(1-\exp(-\mu_n^*\varepsilon)),$$

if $\mu_n^* > 0$ (4)

Dr. J. G. Shanthikumar is an Assistant Professor of Industrial Engineering and Operations Research at Syracuse University. He holds a B.Sc degree in Mechanical Engineering from the University of Sri Lanka, and an M.A. Sc. and Ph.D. in Industrial Engineering from the University of Toronto, Toronto. He received the E.O.E. Pereira Gold Medal for the outstanding student graduating from the College of Engineering, Canadian Commonwealth Scholarship from 1975-1979 and the Outstanding Young Man of America Award for 1980. His research interests are in the areas of production systems modelling and analysis, queueing theory, reliability, scheduling and simulation.

and $L_u)t) = L(t_{n-1}) + \lambda_u^*\varepsilon$ if $\mu_n^* = 0$ for $t = t_{u-u} + \varepsilon$ and $0 \le \varepsilon \le \varepsilon_n$,

where $t_u = \sum_{i=0}^{n}(\varepsilon_n)$, $\mu_n^* = \min_{t_{n-1} \le t \le t_n}(\mu(t))$,

$\lambda_n^* = \max_{t_{n-1} \le t \le t_n} > (\lambda(t))$ and $L_u(t)$ is an upper bound for L(t). From the above result we can further reduce the restriction on λ(t) and μ(t) to the following: as long as λ(t) < ∞ and the total length of time during which $\mu_n^* = 0$ is finite in (0,∞) the mean number L(t) < ∞ for all t ≤ 0 assuming L(0) < ∞.

Now define

$$\tilde{\lambda}_t = \min_{0 \le x \le t}(\lambda(x)) \quad \text{and} \quad \tilde{\mu}_t = \max_{0 \le x \le t}(\mu(x))$$

To obtain a lower bound let us construct another system. Here the arrival rate to the overall system 0 is λ(x) and the departure rate is N(x)μ(x) (shown in Figure 2) at time x. That is, the overall system 0 behaves like our M/M/∞ queue with arrival rate Λ(x) and service rate u(x). However, within this system 0, there are two sub-systems 1 and 2. The arrival stream to the system 0 is divided into two such that the arrival rate to system 1 is $\tilde{\lambda}_t$ and to system 2 is $\lambda_2(x) = \lambda(x) - \tilde{\lambda}_t$ during (0,t). The service rate of system 1 is $\tilde{u}_t$ and of system 2 is u(x). A job leaving system 1 will either leave the system 0 with probability $p(x) = u(x)/\tilde{u}_t$ or join system 2 with probability l-p(x) at time x e (0,t). Let Nl(x)

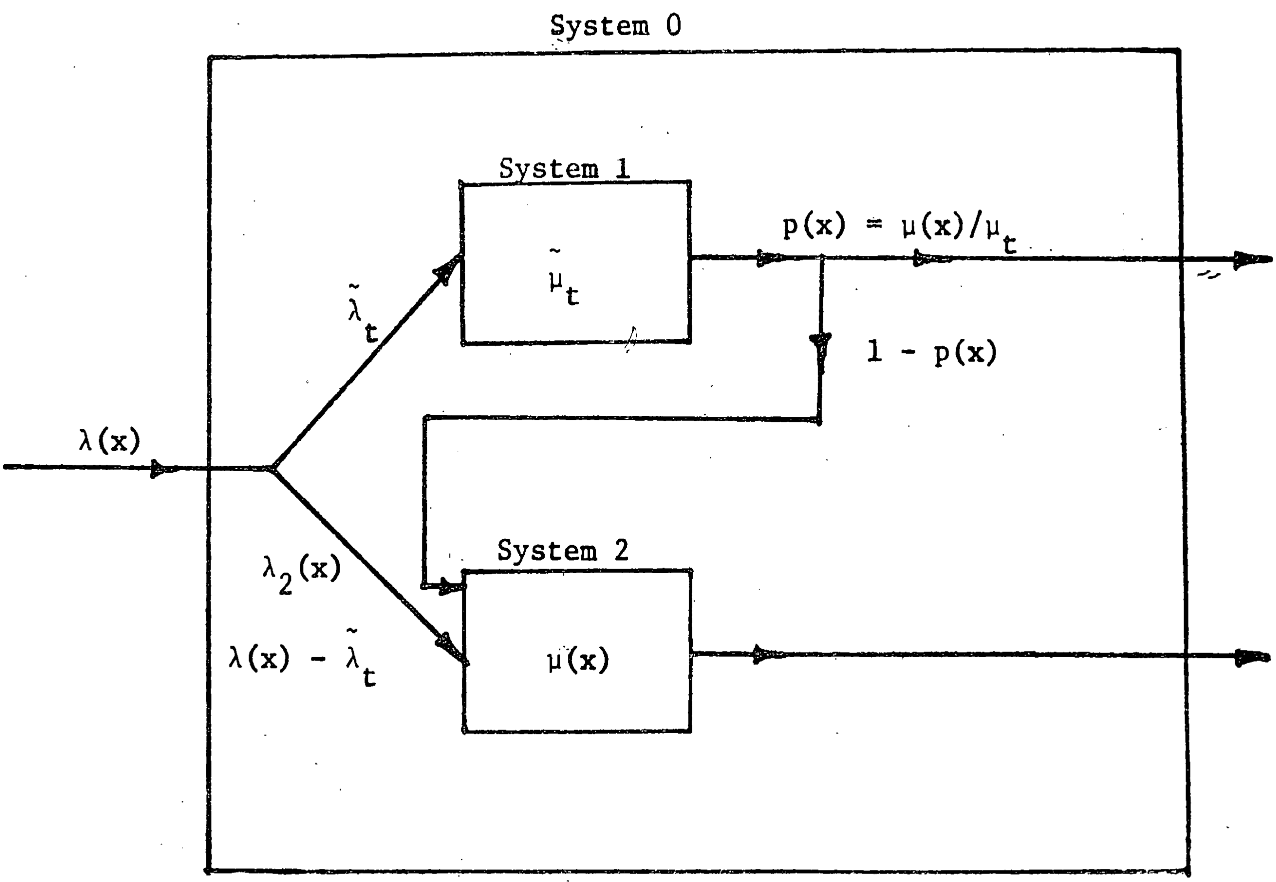


FIGURE 2

and $N_2(x)$ be the number in the systems 1 and 2 at time x and let $L_1(x)$ and $L_2(x)$ be their mean values. Then $N(x)$ — $N_1(x) + Nx(x)$ and $L(x)$ — $L_1(x) + L_2(x)$. It can be easily varified that the overall system behaves like our M/M/∞ queue. Inspection will also reveal that system 1 behaves like an M/M/∞ with constant arrival rate$\lambda +$ yt and service rate ut. Then similar to the derivation of the upper bound we can show that for t / $t = t_{n-1} + \varepsilon$ and $0 \leq \varepsilon \leq \varepsilon_n$,

$$L_e(t) = L_e(t_{n-1}) \exp(-\mu_n \varepsilon) + \frac{\tilde{\lambda}_n}{\tilde{\mu}_n} (1 - \exp(\tilde{\mu}_n 8))$$

if $\tilde{\mu}_n > 0$ (5)

and $L_e(t) = L_e(t_{n-1}) + \tilde{\lambda}_n \varepsilon$ if $\tilde{\mu}_n = 0$ is a lower bound for L(t). That is

$$L_e(t) \leq L(t) \leq L_u(t).$$

A plot of L (.), L(.) and L_u(.) for the example

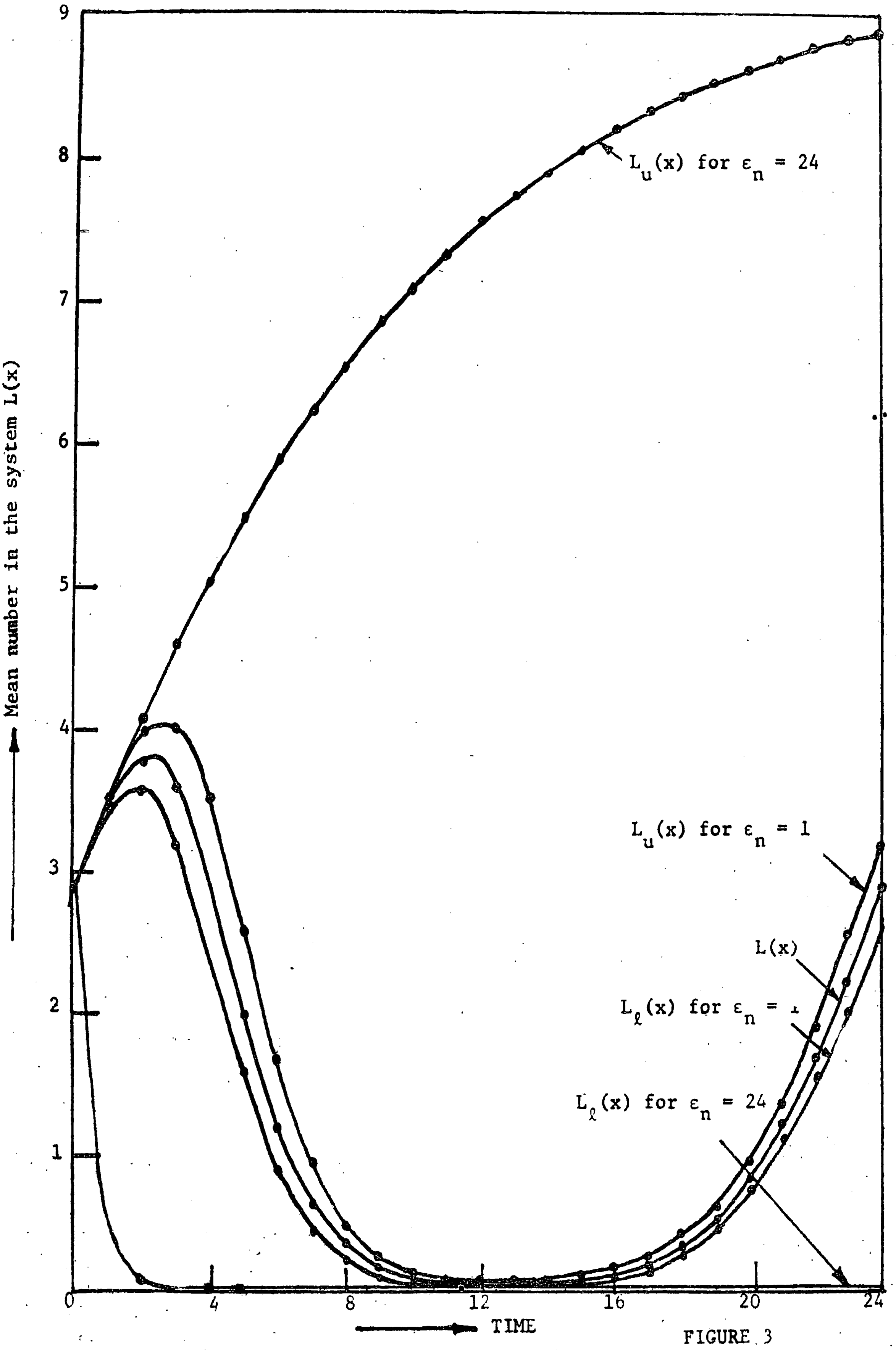
$\lambda(t) = 0.5(1 + 0.9\cos(\pi t/12))$, $\mu(t) = (1 + 0.9\cos(\pi(t-12)/12)$

and $L(0) = 2.865$ considered by Collings and Stoneman is shown in Figure 3. Note that better and better bounds can be obtained by reducing ε_n further and further. Obviously as

$$\varepsilon_n \rightarrow 0, \quad n = 1, 2, \ldots, \quad \begin{array}{l} L_e(t) \rightarrow L(t) \quad \text{and} \\ L_u(t) \rightarrow L(t). \end{array}$$

Reference

1. T. Collings and C. Stoneman, "The M/M/∞ Queue with Varying Arrival and Departure Rates", *Operations Research,* 24 (1976), 760-773.



Mean number in the system L(x)
TIME
0
4
8
12
16
20
24
1
2
3
4
5
6
7
8
9
$L_u(x)$ for ε_n = 24
$L_u(x)$ for ε_n = 1
L(x)
$L_\ell(x)$ for ε_n = 1
$L_\ell(x)$ for ε_n = 24

FIGURE 3