

# Probability model for non-homogeneous multiprogramming computer system

By J. SZTRIK

## 1. Introduction

In this paper we deal with a special queueing problem, which is of considerable practical importance in the field of computer applications. A new mathematical model for FIFO multiprogramming system is given in the following way.

A number of  $n$  jobs are permitted simultaneous access to the resources of the system in such a way that the Central Processor Unit (CPU) is busy processing one job while various input-output (I/O) peripheral units (e.g. a rotating disk memory, a swapping drum, a magnetic tape unit, a data cell, a card reader and so on) are processing some of the others. In such a multiprogrammed environment jobs do indeed circulate among these devices in such a way that they require the attention of the CPU followed by the need for some peripheral units, and such a cycle is repeated several times.

We assume that the CPU services the jobs according to FIFO (first in-first out) discipline. In addition the system is supposed to have enough peripheral device, so no queueing for I/O operation occurs and a waiting line can be formed at the CPU only.

The jobs are assumed to be stochastically different; the  $i$ -th program is characterised by exponentially distributed I/O time with rate  $\lambda_i$  and CPU time with an arbitrary distribution function  $F_i(x)$ . The I/O and compute times for all jobs are mutually independent. Similar models with less complexity were discussed by GAVER [5], TOMKÓ [3].

The interested reader is referred to KLEINROCK [1, 2] for further models and a good bibliography on this subject.

The purpose of the present paper is to give the main characteristics of the system in steady state, namely CPU productivity, expected CPU busy period, mean response time of the jobs. Finally numerical example illustrates the problem in question.

## 2. Description of the model

Let the random variable (abbreviated by r.v.)  $v(t)$  denote the number of jobs at the CPU at time  $t$  and let  $(\alpha_1(t), \dots, \alpha_{v(t)}(t))$  indicate their indices in the order of their arrival.

Introduce the process

$$\mathbf{Y}(t) = (v(t), \alpha_1(t), \dots, \alpha_{v(t)}(t)).$$

The stochastic process  $(\mathbf{Y}(t), t \geq 0)$  is not Markovian unless the distribution functions  $F_i(x)$  are exponential,  $i=1, 2, \dots, n$ .

Let the r.v.  $\xi_t$  denote the attained compute time the job under service has got till time  $t$ .

Putting

$$\mathbf{X}(t) = (v(t), \alpha_1(t), \dots, \alpha_{v(t)}(t); \xi_t) \quad (1)$$

the process  $(\mathbf{X}(t), t \geq 0)$  has already the Markov property.

Let  $V_k^n$  denote the set of all variations of order  $k$  of the integers  $1, 2, \dots, n$  ordered lexicographically. The points  $(i_1, \dots, i_k; x)$  form the state space of process (1), where  $(i_1, \dots, i_k) \in V_k^n$ ,  $x \in \mathbf{R}_+$ ,  $1 \leq k \leq n$ . The process is in state  $(i_1, \dots, i_k; x)$  if  $k$  jobs need the CPU their indices in the order of arrival are  $(i_1, \dots, i_k)$  and the attained service time of job  $i_1$  is  $x$ . The state when the CPU is idle is denoted by  $\{0\}$ .

In order to derive the Chapman—Kolmogorov equations we should consider the transitions that can occur in an arbitrary time interval  $(t, t + \Delta t)$ . The transition probabilities are the following.

$$\begin{aligned} P\{x(t + \Delta t) = (i_1, \dots, i_k; x + \Delta t) / x(t) = (i_1, \dots, i_k; x)\} = \\ = (1 - \sum_{j \neq i_1, \dots, i_k} \lambda_j \Delta t) \frac{1 - F_{i_1}(x + \Delta t)}{1 - F_{i_1}(x)} + o(\Delta t), \end{aligned} \quad (i)$$

$$\begin{aligned} P\{x(t + \Delta t) = (i_1, \dots, i_{k+1}; x + \Delta t) / x(t) = (i_1, \dots, i_k; x)\} = \\ = \lambda_{i_{k+1}} \Delta t \frac{1 - F_{i_1}(x + \Delta t)}{1 - F_{i_1}(x)} + o(\Delta t), \end{aligned} \quad (ii)$$

$$\begin{aligned} P\{x(t + \Delta t) = (i_2, \dots, i_k; 0) / x(t) = (i_1, \dots, i_k; x)\} = \\ = \frac{F_{i_1}(x + \Delta t) - F_{i_1}(x)}{1 - F_{i_1}(x)} + o(\Delta t). \end{aligned} \quad (iii)$$

Let us introduce some further notations.

$$A_{i_1, \dots, i_k} = \sum_{j \neq i_1, \dots, i_k} \lambda_j, \quad A = \sum_{j=1}^n \lambda_j,$$

$$S_{i_1, \dots, i_k} = \sum_{j=1}^k \lambda_{i_j}, \quad \beta_i = \int_0^\infty x dF_i(x), \quad \Phi(s; i) = \int_0^\infty e^{-sx} dF_i(x),$$

$$\Pi_l^{(i_1, \dots, i_k)} = \prod_{r=l+1}^k \lambda_{i_r} / \prod_{q=l+1}^k S_{i_{l+1}, \dots, i_q}, \quad 1 \leq l \leq k, \quad 1 \leq k \leq n.$$

For the distribution of  $\mathbf{x}(t)$  consider the functions given below.

$$P_0(t) = P(v(t) = 0),$$

$$P_{i_1, \dots, i_k}(x, t) = P(v(t) = k, \alpha_1(t) = i_1, \dots, \alpha_k(t) = i_k; \xi_t \cong x),$$

$$(i_1, \dots, i_k) \in V_k^n, \quad k = 1, 2, \dots, n.$$

**Theorem 1.** If  $\beta_i < \infty$  ( $i=1, 2, \dots, n$ ) then the limits

$$P_0 = \lim_{t \rightarrow \infty} P_0(t)$$

$$P_{i_1, \dots, i_k}(x) = \lim_{t \rightarrow \infty} P_{i_1, \dots, i_k}(x; t).$$

exist and satisfy the equation

$$P_0 + \sum_{k=1}^n \sum_{V_k^n} \lim_{x \rightarrow \infty} P_{i_1, \dots, i_k}(x) = 1$$

*Proof.* Note that  $(\mathbf{x}(t), t \geq 0)$  is a linear Markov process treated in GNEDENKO—KOVALENKO [7] in details. Our statement follows from a theorem on page 211 of this monograph.

Our next task is to give a procedure to determine the ergodic probabilities

$$(P_0, P_{i_1, \dots, i_k}), \quad (i_1, \dots, i_k) \in V_k^n, \quad k = 1, 2, \dots, n.$$

To do so, first of all we show that the ergodic functions

$$P_{i_1, \dots, i_k}(x), \quad (i_1, \dots, i_k) \in V_k^n, \quad x \in \mathbf{R}_+, \quad k = 1, 2, \dots, n$$

are differentiable at common continuity points of  $F_i(x)$ .

Then we introduce the so-called normed density functions:

$$p_{i_1, \dots, i_k}^*(x) = \frac{\frac{d}{dx} P_{i_1, \dots, i_k}(x)}{1 - F_{i_1}(x)}. \quad (2)$$

We derive a system of integro-differential equations for these functions, and by the help of its solution we can give an algorithm for calculating the stationary distribution.

Let  $V_{i_1, \dots, i_r}^{i_1, \dots, i_k}(\tau)$  denote the probability of the event that at an arbitrary instant jobs with indices  $(i_1, \dots, i_r)$  are in compute period and from this epoch during a period of time  $\tau$  additional jobs  $(i_{r+1}, \dots, i_k)$  finish their I/O operation in this order. One can readily verify that

$$V_{i_1, \dots, i_r}^{i_1, \dots, i_k}(\tau) = e^{-A_{i_1, \dots, i_k} \tau}.$$

$$\int_{0 < z_1 < z_2 < \dots < z_{k-r}} \lambda_{i_{r+1}} e^{-\lambda_{i_{r+1}} z_1} \dots \lambda_{i_k} e^{-\lambda_{i_k} z_{k-r}} dz_1 \dots dz_{k-r},$$

which can be expressed by the help of exponential functions. In the homogeneous case ( $\lambda_1 \equiv \lambda$ )

$$V_{i_1, \dots, i_r}^{i_1, \dots, i_k}(\tau) = \frac{1}{(k-r)!} (1 - e^{-\lambda\tau})^{k-r} e^{-(n-k)\lambda\tau}.$$

Now we prove the following theorem.

**Theorem 2.** The ergodic distribution function  $P_{i_1, \dots, i_k}(x)$  possesses density function  $p_{i_1, \dots, i_k}(x)$ ,  $(i_1, \dots, i_k) \in V_k^n$ ,  $1 \leq k \leq n$ , and at almost every  $x \in \mathbf{R}_+$ . In addition, the normed d.f.

$$p_{i_1, \dots, i_k}^*(x) = \frac{p_{i_1, \dots, i_k}(x)}{1 - F_{i_1}(x)}$$

is differentiable at every  $x \in \mathbf{R}_+$ .

*Proof.* We first prove the existence of densities  $p_{i_1, \dots, i_k}(x)$ . Let the  $(\mathbf{x}(t), t \geq 0)$  be in state  $(i_1, \dots, i_k; x)$  at an arbitrary  $t$ . The process is in this state iff some epoch  $u$ ,  $t - x < u < t$ , the CPU completes a computation and immediately starts servicing job  $i_1$ . If the indices of tasks in compute period are  $(i_1, \dots, i_r)$  at time  $u$  then the unexpired service time of  $i$  must exceed  $t - u$  and during the time interval  $(u, t)$  jobs  $(i_{r+1}, \dots, i_k)$  should arrive in this order. For the sake of easier understanding we notice that the process  $\mathbf{x}(t)$  is of regenerative type.

The regenerative periods can be defined in several ways. Let us consider for example the epochs when the CPU completes a computation and starts to serve the program  $i_1$  while the others with indices  $(i_2, \dots, i_r)$  are already waiting for their turn. If the initial state  $(j_1, \dots, j_s; z)$  differs from  $(i_1, \dots, i_r; 0)$  then the renewal process in question is a so-called delayed one.

Let us denote by  $H_{j_1, \dots, j_s, z}^{i_1, \dots, i_r}(t)$  the renewal function of the process considered above.

Denote by

$$(R_0, R_{j_1, \dots, j_s}(z); z \geq 0, (j_1, \dots, j_s) \in V_s^n, s = 1, 2, \dots, n)$$

the initial distribution of  $(\mathbf{x}(t), t \geq 0)$ . Keeping in mind the behaviour of the recurrent process, by using the theorem of total probability, we get

$$P_{i_1, \dots, i_k}(x, t) = \left( \sum_{s=1}^n \sum_{V_s^n} \int_0^t dR_{j_1, \dots, j_s}(z) + R_0 \right) \cdot \sum_{r=1}^k \int_{t-x}^t V_{i_1, \dots, i_r}^{i_1, \dots, i_k}(t-u) [1 - F_{i_1}(t-u)] dH_{j_1, \dots, j_r, z}^{i_1, \dots, i_k}(u).$$

Applying the key renewal theorem of Smith we have

$$\lim_{t \rightarrow \infty} P_{i_1, \dots, i_k}(x, t) = P_{i_1, \dots, i_k}(x) = \sum_{r=1}^n \frac{1}{m_{i_1, \dots, i_r, 0}} \int_0^x [1 - F_{i_1}(u)] V_{i_1, \dots, i_r}^{i_1, \dots, i_k}(u) du,$$

where  $m_{i_1, \dots, i_r, 0}$  denotes the expected recurrence time into state  $(i_1, \dots, i_r, 0)$  which is finite since the process is ergodic. The functions  $V_{i_1, \dots, i_r}^{i_1, \dots, i_k}(\tau)$  can be expressed with

the aid of exponential ones hence the d.f.  $P_{i_1, \dots, i_k}(x)$  is differentiable at every continuity point of  $F_{i_1}(x)$ .

This implies that the density  $f. p_{i_1, \dots, i_k}(x)$  is defined at almost everywhere and

$$p_{i_1, \dots, i_k}(x) = \sum_{r=1}^k \frac{1}{m_{i_1, \dots, i_r}} V_{i_1, \dots, i_r}^{i_1, \dots, i_r}(x) [1 - F_{i_1}(x)].$$

Therefore the normed functions are differentiable at  $x \in \mathbf{R}_+$ .

**Theorem 3.** The stationary density  $f$  of the process  $(\mathbf{X}(t), t \geq 0)$  satisfies the following system of differential equations

$$\begin{aligned} \frac{dp_{i_1}^*(x)}{dx} + A_{i_1} p_{i_1}^*(x) &= 0, \\ \vdots \\ \frac{dp_{i_1, \dots, i_k}^*(x)}{dx} + A_{i_1, \dots, i_k} p_{i_1, \dots, i_k}^*(x) &= \lambda_{i_k} p_{i_1, \dots, i_{k-1}}^*(x), \\ \vdots \\ \frac{dp_{i_1, \dots, i_n}^*(x)}{dx} &= \lambda_{i_n} p_{i_1, \dots, i_{n-1}}^*(x). \end{aligned} \tag{3}$$

The boundary conditions are

$$\begin{aligned} AP_0 &= \sum_{j=1}^n \int_0^\infty p_j^*(x) dF_j(x), \\ \vdots \\ p_{i_1, \dots, i_k}(0) &= \sum_{j \neq i_1, \dots, i_k} \int_0^\infty p_{j, i_1, \dots, i_k}^*(x) dF_j(x), \quad p_{i_1, \dots, i_n}(0) = 0. \end{aligned} \tag{4}$$

*Proof.* Considering the transition probabilities, using the equations of Chapman—Kolmogorov the first part of the theorem can easily be verified. The second part can be proved by the theorem of total probability. The details of the proof is to be found in SZTRIK (6).

It is quite easy to see that the solution of (3), (4) is

$$\begin{aligned} p_{i_1, \dots, i_k}^*(x) &= \sum_{l=1}^k (-1)^{k-l} c_{i_1, \dots, i_l} e^{-A_{i_1, \dots, i_l} x} \prod_{i=1}^l (i_1, \dots, i_k), \\ (i_1, \dots, i_k) &\in V_k^n, \quad k = 1, 2, \dots, n, \end{aligned}$$

where the constants  $C_{i_1, \dots, i_l}$  are to be determined from the boundary condition (4).

In the following we describe an iterative method to calculate these coefficients. Let  $\mathbf{c}_k$  denote the vector

$$\begin{pmatrix} c_{1, 2, \dots, k} \\ \vdots \\ c_{i_1, \dots, i_k} \\ \vdots \\ c_{n, \dots, n-k+1} \end{pmatrix}$$

of dimension  $\binom{n}{k} k!$ . The components of  $\mathbf{c}_k$  are listed in the lexicographic order of their indices,  $k=1, \dots, n$ . Notice, that boundary condition  $p_{i_1, \dots, i_n}(0)=0$  is equivalent to equation

$$\mathbf{c}_n = A_{n-1}^{(n)} \mathbf{c}_{n-1} + \dots + A_1^{(n)} \mathbf{c}_1,$$

with a suitably chosen matrix  $A_k^{(n)}$  of order  $n! x \binom{n}{k} k!$ . The  $k$ -th boundary condition, where  $2 \leq k \leq n-1$ , gives the relation

$$\sum_{i=1}^k (-1)^{k-l} c_{i_1, \dots, i_l} \Pi_i^{(i_1, \dots, i_k)} = \sum_{j \neq i_1, \dots, i_k} C_{j, i_1, \dots, i_l} \Pi_i^{(i_1, \dots, i_k)} \int_0^\infty e^{-A_{j, i_1, \dots, i_l} x} dF_j(x).$$

In term of the Laplace—Stieltjes transform this becomes

$$\sum_{i=1}^k (-1)^{k-l} c_{i_1, \dots, i_l} \Pi_i^{(i_1, \dots, i_k)} = \sum_{j \neq i_1, \dots, i_k} C_{j, i_1, \dots, i_l} \Pi_i^{(i_1, \dots, i_k)} \Phi(A_{j, i_1, \dots, i_l}; j).$$

More succinctly

$$\mathbf{c}_k = A_{k+1}^{(k)} \mathbf{c}_{k+1} + \dots + A_1^{(k)} \mathbf{c}_1.$$

Now we are ready to define our algorithm. We have

$$\mathbf{c}_n = \sum_{j=1}^{n-1} A_j^{(n)} \mathbf{c}_j, \quad (*)$$

$$\mathbf{c}_{n-1} = \sum_{j=1}^{n-2} B_j^{(n-1)} \mathbf{c}_j,$$

where the matrix  $B_j^{(n-1)}$  is defined by

$$B_j^{(n-1)} = (1 - A_n^{(n-1)} A_{n-1}^{(n)} - A_{n-1}^{(n-1)})^{-1} (A_n^{(n-1)} A_j^{(n)} + A_j^{(n-1)}), \quad 1 \leq j \leq n-2.$$

Similarly

$$\mathbf{c}_k = \sum_{j=1}^{k-1} B_j^{(k)} \mathbf{c}_j, \quad (**)$$

where the matrix  $B_j^{(k)}$  is given by

$$B_j^{(k)} = (1 - A_{k+1}^{(k)} B_k^{(k+1)} - A_k^{(k)})^{-1} (A_{k+1}^{(k)} B_j^{(k+1)} + A_j^{(k)})$$

$$2 \leq k \leq n-1, \quad 1 \leq j \leq k-1.$$

For  $\mathbf{c}_1$  we have the equation

$$\mathbf{c}_1 = A_2^{(1)} \mathbf{c}_2 + A_1^{(1)} \mathbf{c}_1 + P_0 \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}.$$

Using the formula for  $C_2$  we obtain

$$(1 - A_2^{(1)} B_1^{(2)}) \mathbf{c}_1 = P_0 \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}.$$

Hence

$$\mathbf{c}_1 = (1 - A_2^{(1)} B_1^{(2)})^{-1} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} P_0. \quad (***)$$

Starting with an arbitrary  $P_0$  and using the relations (\*), (\*\*), (\*\*\*) we can determine the vectors  $\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_n$  (in this order). Following this procedure we obtain all constants and also the density functions  $p_{i_1, \dots, i_k}(x)$ .

Let us denote by  $P_{i_1, \dots, i_k}$  the stationary probability that jobs with indices  $(i_1, \dots, i_k)$  are in compute period. Apparently,

$$P_{i_1, \dots, i_k} = \int_0^\infty p_{i_1, \dots, i_k}(x) dx = \int_0^\infty p_{i_1, \dots, i_k}^*(x) [1 - F_{i_1}(x)] dx.$$

Denote by  $P_k$  the steady state probability that  $k$  program is at the CPU. We have

$$P_k = \sum_{V_k^n} P_{i_1, \dots, i_k}.$$

The value of  $P_0$  can be calculated from the normalising condition

$$P_0 + \sum_k P_k = 1.$$

### 3. Utility investigations

#### (i) CPU utilization.

It is easy to see that the CPU's activity can be divided into two periods, viz. idle and busy ones. Together they form a cycle. The durations of these cycles are independent and identically distributed random variables.

By the virtue of a renewal consideration it follows that

$$P_0 = \frac{1/A}{1/A + M\delta}$$

where  $M\delta$  denotes the mean CPU busy period and  $1/A$  is the average idle period length.

If  $U_{CPU}$  denotes the CPU productivity, which is the long-run fraction of time the CPU is busy, then

$$U_{CPU} = \frac{M\delta}{1/A + M\delta}$$

Consequently

$$M\delta = (1 - P_0)/(AP_0).$$

#### (ii) Mean waiting times.

During the execution a job waits for the CPU, occupies it and takes I/O operations. If one considers these periods as a cycle, then in equilibrium for a fixed job these cycles have identical distribution, but they are not independent.

Let  $P^{(i)}$  denote the steady state probability that job  $i$  is in compute period and let the average period lengths designated by  $W_i, \beta_i, 1/\lambda_i$ , respectively.

Consider now the semi-Markov process  $(Y(t), t \geq 0)$ , with state space  $\bigcup_{k=1}^n V_k^n + \{0\}$ .

Let  $H_i$  be the event that the program  $i$  does not take I/O operations. Introduce the function

$$Z_{H_i}(t) = \begin{cases} 1 & \text{if } Y(t) \in H_i, \\ 0 & \text{otherwise.} \end{cases}$$

**Theorem 4.**

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T Z_{H_i}(t) dt = \frac{W_i + \beta_i}{1/\lambda_i + W_i + \beta_i} = P^{(i)}.$$

The statement is a special case of a theorem concerning mean sojourn time for semi-Markov, processes, see TOMKÓ [4] on page 297. Since the probability  $P^{(i)}$  can be calculated from the distribution  $P_{i_1, \dots, i_k}$ , by

$$P^{(i)} = \sum_{k=1}^n \sum_{V_k^n, i \in (i_1, \dots, i_k) \in V_k^n} P_{i_1, \dots, i_k},$$

the expected waiting time of job  $i$  is

$$W_i = P^{(i)} / [\lambda_i(1 - P^{(i)})] - \beta_i.$$

#### 4. Numerical results

We shall deal with only the case  $n=4$  because the size of matrices involved in the iteration grows rapidly by increasing  $n$ . We assume that the I/O times are identically distributed with parameter  $\lambda=1.2$  and the compute times are mixtures of Erlangian distributions, namely

$$\begin{aligned} F(1, x) &= E(1, 1.2, x), \\ F(2, x) &= 0.8E(2, 3.3, x) + 0.2E(2, 0.9, x), \\ F(3, x) &= 0.2E(2, 3.5, x) + 0.6E(3, 2.6, x) + 0.2E(1, 0.2, x), \\ F(4, x) &= 0.8E(2, 3.3, x) + 0.2E(2, 0.9, x). \end{aligned}$$

$E(k, \lambda, x)$  indicates the  $k$ -stage Erlangian distribution with parameter  $\lambda$ . So the stationary probability that the  $i$ -th job is at the CPU,  $i=1, 2, 3, 4$ , and that the CPU is idle are the following:

$$\begin{aligned} P^{(1)} &= 0.8110664463, \quad P^{(2)} = 0.8130187511, \quad P^{(3)} = 0.8209813440, \\ P^{(4)} &= 0.8130187511, \quad P_0 = 0.005945. \end{aligned}$$

Finally, the main characteristics are:

CPU utilization = 0.994055,  
mean waiting times:

$$W(1) = 3.5773884090, \quad W(2) = 3.623441546, \quad W(3) = 3.821674988, \quad W(4) = 3.623441546$$

*Acknowledgment.* I am very grateful to Dr. J. Tomkó for his helpful discussion.



### Abstract

The aim of the present paper is to give a more adequate model for FIFO multiprogramming computer systems.

The jobs are stochastically different, program  $i$  is characterized by exponentially distributed I/O time with rate  $\lambda_i$  and CPU time with an arbitrary distribution function  $F_i(x)$ .

In stationary case we deal with CPU utilization, mean actual waiting times of the jobs. Finally numerical examples illustrate the problem in question.

**KEYWORDS.** I/O time, CPU time, utilization, mean response time.

DEPARTMENT OF MATHEMATICS  
UNIVERSITY OF DEBRECEN  
PF. 12., DEBRECEN, HUNGARY  
H-4010

### References

- [1] KLEINROCK, L., *Sorbanállás-kiszolgálás*, Műszaki Könyvkiadó, Budapest, 1979.
- [2] KLEINROCK, L., *Queueing systems*, vol. 2, Wiley-Interscience, New York, 1976.
- [3] TOMKÓ, J., Processor utilization study, *Alkalmaz. Mat. Lapok*, v. 1, 1975, pp. 319—331; v. 3, 1977, pp. 83—96.
- [4] TOMKÓ, J., On sojourn times for semi-Markov processes, 14-th European Meeting of Statisticians, Wrocław, Poland, 1981.
- [5] GAVER, D. P., Probability models for multiprogramming computer systems, *J. Assoc. Comput. Mach.*, v. 3, 1967, pp. 423—438.
- [6] SZTRIK, J., On machine interference, Ph. D. Thesis, University of Debrecen, 1980.
- [7] GNEDENKO, B. V., I. N. KOVALENKO, Lectures on Queueing Systems, Kiev, 1963, (in Russian).

(Received Dec. 3, 1981)