NON-PREEMPTIVE SERVICE OF A FINITE NUMBER OF JOBS

L. LAKATOS
(Budapest)
(Received March 2, 1978)

In [1] systems are considered in which the jobs are treated preferentially, according to the priorities assigned to them. Conventionally, the priority of a job is a nonnegative integer index fixed at arrival time, and lower index means higher priority. Given any priority index m, then the jobs with this index m form a priority class (of index m), and each priority class is ordered into a queue, according to arrival. For the service discipline it is assumed, that each job, once begun on a processor, is allowed to run to completion, and after completing a job the processor next takes the one at the head of the highest-priority (lowest-level) nonempty queue. Independent Poisson arrivals are assumed for the different priority classes with arrival rate $\lambda_i (i = 1, 2, \dots, n)$. The service time distributions for the different priority classes may be arbitrary and different. In [1] the mean waiting time in queue for the jobs of priority i during statistical equilibrium are computed.

Here we shall consider a system, different from the aforementioned one. We assume that the number n of priority classes and the number k of jobs present in the system, are restricted. We want to get the ergodic probability of servicing a job of given type or that of being in free state, and to derive the distribution function $K(x)$ of the lengths of busy periods.

Under such restrictions the functioning of our system may be described by of a certain piecewise linear process introduced in [4]. By Kovalenko's theorem, if the number of states of the piecewise linear process as well as the mean values of the continuous components, are finite, the the sought ergodic distribution exists, and our purpose is to determine it. Hereafter by job and demand we shall mean the same.

To get the ergodic probabilities we shall use the procedure described in §4.11 of [4]. By general ergodic considerations the sought ergodic probabilities may be determined on the basis of a period between two neighbouring transitions of our system from busy state to free state, using the formulas

$$p_0 = \frac{t_0}{t_0 + t_1 + \dots + t_n}, \quad p_1 = \frac{t_1}{t_0 + t_1 + \dots + t_n}, \dots,$$

$$p_n = \frac{t_n}{t_0+t_1+\ldots+t_n},$$

where t_0 is the mean value of the durations of free state and $t_i (i = 1, 2, \ldots, n)$ is the mean value of service time spent on a job of type (priority) i in a busy period.

As the arrivals at the system are independent Poisson processes with given parameters, the distribution of the durations of free states is exponential with a known parameter, the mean value is

$$t_0 = \frac{1}{\lambda_1+\lambda_2+\ldots+\lambda_n}.$$

Now we denote by t the mean value of the lengths of busy periods. Because of the additivity of mean values it is clear that

$$t = t_1+t_2+\ldots+t_n.$$

To get $t_i (i = 1, 2, \ldots, n)$ we give a system of integral equations for determining $K(x)$. This system may be solved using the Laplace-Stieltjes transform. On the basis of Laplace-Stieltjes transform t can be computed. Let us denote by $B_i(x)$ $(i = 1, 2, \ldots, n)$ the distribution function of the service time of type i jobs. According to Wald's equality

$$t = k_1\tau_1+k_2\tau_1+\ldots+k_n\tau_n,$$

where $\tau_i = -b_i'(O)$ $\big(i = 1, 2, \ldots, n$ and $b_i(s)$ is the Laplace-Stieltjes transform of the distribution function $B_i(x)\big)$ is the mean value of service time of a type i job; k_i $(i = 1, 2, \ldots, n)$ the mean value of the number of type i jobs, serviced for a busy period. Determining t (*it* is necessary to differentiate the Laplace-Stieltjes transform of $K(x)$ and to take it minus sign at the point O (if we do not substitute the numerical value of τ_i we obtain for t an expression consisting of n summands, each of them being in the form $k_i\tau_i$, where k_i is a concrete numerical value. Substituting now the numerical value of τ_i we get $t_i = k_i\tau_i$.

In the following we give a system of integral equations for determining $K(x)$. As we mentioned earlier, we put no restrictions on $B_i(x)$. At first we assume that the service runs simply according to the FIFO rule, after that we shall consider the case of priorities, so at each waiting position the waiting job may be of arbitrary type (priority).

Let us denote by $K_i(x)$ the specialization of the distribution function $K_i(x)$ to the case when the service starts with a job of type $i (i = 1, 2, \ldots, n)$. Then we have

$$K(x) = \sum_{i=1}^{n} \frac{\lambda_i}{\sum_{j=1}^{n} \lambda_j} K_i(x).$$

For $i_1, \ldots, i_j = 1, 2, \ldots, n$; $j = 1, \ldots, k$ let $K_{i_1 \ldots i_j}(x)$ be a further specialization of $K(x)$, belonging to the case when up to the starting of service j jobs have entered the system, their types and order of occurence are given by the sequence $i_1, i_2, \ldots, i_j$. We introduce the notation

$$\Lambda = \lambda_1 + \lambda_2 + \ldots + \lambda_n .$$

Now we prove the following

Theorem. Consider an $M/G/1$ nonpreemptive service system which has n priority levels and in which at any moment can be at most k demands. Then the distribution functions $K_i(x)$ $(i = 1, 2, \ldots, n)$ defined above can be obtained from the following system of equations:

$$(1) \qquad K_i(x) = \int_0^x e^{-\Lambda y}\, dB_i(y) + \sum{}' \sum_{l=2}^{k-1} \int_0^x K_{i_2 \ldots i_l}(x-y)\, P_{i_2 \ldots i_l}(y)\, dB_i(y) + \sum{}'' \int_0^x K_{i_2 \ldots i_k}(x-y)\, P_{i_2 \ldots i_k}(y)\, dB_i(y),$$

$$i = 1, 2, \ldots, n, \quad l = 2, \ldots, k-1,$$

where

$$P_{i_2 \ldots i_l}(y) = e^{-\Lambda y} \frac{y^{l-1}}{(l-1)!} \prod_{j=2}^{l} \lambda_{i_j} \quad \text{and}$$

$$P_{i_2 \ldots i_k}(y) = \frac{1}{\lambda} \prod_{j=2}^{k} \lambda_{i_j} \left\{ \int_0^y e^{-\Lambda u_{k-1}} \frac{u_{k-1}^{k-3}}{(k-3)!}\, du_{k-1} - e^{-\Lambda y} \frac{y^{k-2}}{(k-2)!} \right\};$$

$$(2) \qquad K_{i_1 \ldots i_m}(x) = \int_0^x K_{i_2 \ldots i_m}(x-y)\, e^{-\Lambda y}\, dB_{i_1}(y) + \sum{}' \sum_{l=m-1}^{k-1} \int_0^x K_{i_2 \ldots i_m i_{m+1} \ldots i_l}(x-y)\, P_{i_{m+1} \ldots i_l}(y)\, dB_{i_1}(y) + \sum{}'' \int_0^x K_{i_2 \ldots i_m i_{m+1} \ldots i_k}(x-y)\, P_{i_{m+1} \ldots i_k}(y)\, dB_{i_1}(y),$$

where

$$P_{i_{m+1}\dots i_l}(y) = e^{-\Lambda y}\frac{y^{l-m}}{(l-m)!}\prod_{j=m+1}^{l}\lambda_{i_j},$$

$$P_{i_{m+1}\dots i_k}(y) = \frac{1}{\Lambda}\prod_{j=m+1}^{k}\lambda_{i_j}\left\{\int_0^y e^{-\Lambda u_{k-1}}\frac{u_{k-1}^{k-m-2}}{(k-m-2)!}\,du_{k-1} - e^{-\Lambda y}\frac{y^{k-m-2}}{(k-m-2)!}\right\},$$

$$1 \le i_1, i_2, \dots, i_m, i_{m+1}, \dots, i_k \le n, \; 2 \le m \le k-2;$$

$$K_{i_1 i_2 \dots i_{k-1}}(x) = \int_0^x K_{i_2 \dots i_{k-1}}(x-y)\,e^{-\Lambda y}\,dB_{i_1}(y) + \sum_{j=1}^{n}\frac{\lambda_j}{\Lambda}\int_0^x K_{i_2\dots i_k}(x-y)[1-e^{-\Lambda y}]\,dB_{i_1}(y),$$

$$i_k = j, \quad 1 \le j \le n.$$

Proof. We show how we can get (2), on the basis of which the other equations can be derived easily. Suppose in the system there are m demands with the first of them just being on service. According to the theorem of total probability we have some mutually exclusive possibilities:

1. During the service of the first demand other demands do not enter. Then the integral

$$\int_0^x K_{i_2\dots i_m}(x-y)\,e^{-\Lambda y}\,dB_{i_1}(y)$$

gives the probability of the following event: "The service of the first demand takes y time units ($0 \le y \le x$), during this period other demands do not enter, and the whole service of the next $m-1$ demands, being present already at the beginning of the service of the first one, takes not nore than $x-y$ time units."

2. At the beginning there are m demands in the system and again the first demand requires for its completion y time units during which $l-m$ ($1 \le \le l-m \le k-2$) new demands enter, so the distribution function of the remaining part of the busy period is $K_{i_2\dots i_m\, i_{m+1}\dots i_l}(y)$.
Denoting by $P_{i_{m+1}\dots i_l}(y)$ the probability of the event that the sequence

of types (according to arrival) of the entering $l-m$ new demands in i_{m+1}, i_{m+2}, ..., $i.$, we have

$$P_{i_{m+1}\dots i_l}(y) = \int_0^y \frac{\lambda_{i_l}}{\Lambda} \int_0^{u_l} \frac{\lambda_{i_{l-1}}}{\Lambda} \int_0^{u_{m+2}} \Lambda e^{-\Lambda u_{m+1}} \cdot$$

$$\cdot \Lambda e^{-\Lambda(u_{m+2}-u_{m+1})} du_{m+1} \dots \Lambda e^{-\Lambda(u_l-u_{l-1})} du_{l-1} \cdot e^{-\Lambda(y-u_l)} du_l,$$

where u_r ($r = m+1, m+2, \dots, l$) is the moment when the r-th demand enters, so the last demand enters at u_l, and between u_l and y no further demands arrive. After simplifications and integration we get

$$P_{i_{m+1}\dots i_l}(y) = e^{-\Lambda y} \frac{y^{l-m}}{(l-m)!} \prod_{j=m-1}^{l} \lambda_{i_j}.$$

At last, Σ' in the theorem means that we have to consider all the n^{l-m} possible types of sequences (each demand can have any one of the n types).

3. In the previous cases it was possible to let some waiting places empty. Now we consider the case when each of them is busy. By a reasoning, similar to the previous one, we simply show how $P_{i_{m+1}\dots i_k}(y)$ can be determined. After the arrival of the $(k-1)-th$ demand it is important only whether yet a further demand comes or not. So

$$P_{i_{m+1}\dots i_k}(y) = \frac{\lambda_{i_k}}{\Lambda} \int_0^y \frac{\lambda_{i_{k-1}}}{\Lambda} \int_0^{u_{k+1}} \dots \frac{\lambda_{i_{m+1}}}{\Lambda} \int_0^{u_{m+1}} \Lambda e^{-u\Lambda_{m+1}} \cdot$$

$$\cdot \Lambda e^{-\Lambda(u_{m+2}-u_{m+1})} du_{m+1} \dots \Lambda e^{-\Lambda(u_{k-1}-u_{k-2})} du_{k-2} \cdot [1 - e^{-\Lambda(y-u_{k-1})}] du_{k-1} =$$

$$= \frac{1}{\Lambda} \prod_{j=m+1}^{k} \lambda_{i_j} \left\{ \int_0^y e^{-\Lambda u_{k-1}} \frac{u_{k-1}^{k-m-2}}{(k-m-2)!} du_{k-1} - e^{-\Lambda y} \frac{y^{k-m-1}}{(k-m-1)!} \right\}.$$

Here Σ'' means that the summation is extended over all possible values of i_j.

It can be seen easily that now we have a system of Volterra integral equations which under certain restrictions (see e. g. in [6] and [7]), can be solved using the Laplace-Stieltjes transform. Calculating then then t (the mean value of the lengths of busy periods), the sought ergodic probabilities — taking into account our note concerning t_i — can be computed. The number of equations in our system is

$$\sum_{i-1}^{k-1} n^i,$$

where k is the number of waiting places in the service system (and n is the number of priority levels).

If we assume that in our system the service runs according to the priorities of jobs, then the service is always governed by priority and the order of arrival is taken into account only between jobs of the same priority. So at any moment, if $i_1, i_2 \ldots i_m$ is the sequence of types (priorities) of the jobs waiting at the first, second, ..., *m-th* place respectively in the queue then this sequence is nondecreasing, i. e. lower priority never precedes a higher one. The probabilities $P_{i_1 \ldots i_m}(y)$ can differ from each other only in the order of arrival of jobs and therefore are considered identical and have to be summed up. Of course the number of the equations will be less too, namely

$$\sum_{j-1}^{k-1} S_n^{(j)},$$

which can be obtained (and verified) easily by the following recursion:

$$S_i^{(m)} = S_1^{(m-1)} + S_2^{(m-1)} + \ldots + S_i^{(m-1)},$$

$$S_i^{(1)} = i.$$

LITERATURE

[1] *Coffman E. G., Denning P. J.:* Operating Systems Theory, Prentice Hall, Englewood Cliffs, New Jersey, 1973.

[2] *Lakatos L.:* Vizsgálatok az M/G/1 típusú tömegkiszolgálási rendszerek körében, Egyetemi doktori értekezés, ELTE TTK, 1977.

[3] *Lakatos L.:* On a Queueing Problem in the Theory of Operating Systems, 4th Winterschool of Visegrad on the Theory of Operating Systems, MTA SZTAKI Közlemények, 87/1978, pp. 93–97.

[4] *Гнеденко Б. В., Коваленко И. Н.:* Введение в теорию массового обслуживания, изд. Наука, Москва, 1966.

[5] *Королев Л. Н.:* Структуры ЭВМ и их математическое обеспечение, Изд. Наука, Москва, 1974.

[6] *Краснов М. Л.:* Интегральные уравнения (введение в теорию), Изд. Наука, Москва, 1975.

[7] *Краснов М. Л., Киселев А. И., Макаренко Г. И.:* Интегральные уравнения (задачи и упражнения), Изд. Наука, Москва, 1976.

Mathematical Institute of the UkrSSR Academy of Sciences
252601 KIEV – 4, GSP, ul. Repina 3.
Soviet Union