International Journal of
Mathematical Modelling & Computations
Vol. 04, No. 02, 2014, 167- 176

Modelling and Analysis of a Discrete-time Priority Queuing Computer Network with Priority Jumps using Probability Generating functions

Deepak C. Pandey[a,*] and Arun K. Pal[b]

[a] *Department of Mathematics, Statistics and Computer Science*
G. B. Pant University of Ag. and Technology, Pantnagar
U. S. Nagar, India
[b]*Department of Mathematics, Statistics and Computer Science*
G. B. Pant University of Ag. and Technology, Pantnagar
U. S. Nagar, India

Abstract. Priority queues have a great importance in the study of computer communication networks in which different types of traffic require different quality of service standards. The discrete-time non-preemptive priority queuing model with priority jumps is proposed in this paper. On the basis of probability generating functions mean system contents and mean queuing delay characteristics are obtained. The effect of jumping mechanism is analysed which clearly shows that the queuing system provides better results when the fraction of class-1 arrivals in the overall traffic mix is small.

Received: 28 June 2013; Revised: 28 November 2013; Accepted: 5 December 2013.

Keywords: Discrete-time queuing theory, Priority scheduling, Generating functions, Delay analysis

Index to information contained in this paper

1. Introduction

Different QoS (Quality of Service) standards are required by the different types of traffic in modern computer communication networks. We always face two types of traffic namely delay-sensitive traffic (i.e., voice and video) and delay-insensitive traffic (i.e., data) in these networks. Since response time or delay is a crucial performance measure for

*Corresponding author. Email: deechp2006@gmail.com

delay-sensitive applications, time delays in priority queues have been studied extensively in recent years. Due to the sensitivity of the delay-sensitive traffic, it is a basic requirement for designing andconstructing an efficient communication network with a very small mean delay and delay jitter. While the values of the important performance measures of the delay-insensitive traffic, namely loss ratio and throughput should also be small. Priority scheduling scheme is an important way by which we can achieve these requirements.

Since the research works in the field of two-class discrete queues have begun to proliferate only in recent years. Analysis of two-class discrete queues is not available too much. Two-class discrete-time systems, with service time assumed deterministically 1 slot have been analysed in [1-2] as well as in [4]. The systems with non-preemptive priority and general service times with nonhomogeneous packet arrivals have been studied in [8] and [11]. In a HOL non-preemptive priority scheduling, we assume that delay-sensitive traffic has priority over delay-insensitive traffic without preemption i.e., when a server becomes idle, a packet of delay-sensitive traffic, when available, will always be scheduled next but the service of a delay-insensitive packet which has already in service cannot be interrupted by a freshly arriving delay-sensitive packet. This priority scheduling scheme is not only very easy to implement but also provides relatively low delays for the delay sensitive traffic ([3] and [13]). Rubin and Tsai [9] discussed the HOL non-preemptive priority queues taking a variety of arrival and service time distributions (see also in [10] and [12]).

In this paper, we will consider a head-of-line priority scheme with priority jumps (HOL-PJ) in which class-2 i.e., delay-insensitive packets (at the HOL position) jump to (the end of) the class-1 i.e., delay-sensitive traffic queue. In literature, many jumping schemes with different criteria are discussed (see in [5-7]).

2. Formulation of Mathematical Model

A discrete-time single-server queueing system with two queues having infinite buffer space is proposed for study. Two types of traffic are arriving in the system; viz. packets of class-1 that are stored in the first queue and packets of class-2 which are lined in the second. It is assumed that the time is slotted. We denote the number of arrivals of class-j during slot k by $a_{j,k} (j = 1, 2)$. Both types of packet arrivals are assumed to be i.i.d. from slot-to-slot and are characterized by the joint probability mass function

$$a(m, n) \triangleq Prob\left[a_{1,k} = m, a_{2,k} = n\right],$$

and joint probability generating function (pgf) $A(z_1, z_2)$,

$$A(z_1, z_2) \triangleq E\left[z_1^{a_{1,k}} z_2^{a_{2,k}}\right] = \sum_{m,n=0}^{\infty} a(m,n)\, z_1^m z_2^n$$

It is assumed that the service times of the class j packets are i.i.d. and are characterized by the

probability mass function $s_j(m)$and probability generating function$S_j(z)$,

$$S_j(z) = \sum_{m=1}^{\infty} s_j(m) z^m$$

with $j = 1, 2$. Also the mean service time of a class j packet is denoted by $\mu_j = S_j'(1)$. The total load is given by $\triangleq \rho_1 + \rho_2$, where $\rho_j \triangleq \lambda_j \mu_j (j = 1, 2)$.

To analysed the system contents at the beginning of start slots, i.e., slots at the beginning of which a service of a packet (if present) can begin. Let us denote the system contents of class-j at the beginning of the l-th start slot by $n_{j,l} (j = 1, 2)$ and their joint pgf is denoted by $N_l(z_1, z_2)$, i.e.,

$$N_l(z_1, z_2) \triangleq E\left[z_1^{n_{1,l}} z_2^{n_{2,l}}\right].$$

If s_l^* denotes the service time of the packet that enters service at the beginning of start slot *l*. Then the system contents of both types of packets can be obtained according to the following system equations:

1. If $n_{1,l} = n_{2,l} = 0$:

$$n_{1,l+1} = a_{1,k};$$

$$n_{2,l+1} = a_{2,k},$$

2. If $n_{1,l} = 0\ and n_{2,l} > 0$:

$$n_{1,l+1} = \sum_{i=0}^{s_l^*-1} a_{1,k+i};$$

$$n_{2,l+1} = n_{2,l} + \sum_{i=0}^{s_l^*-1} a_{2,k+i} - 1,$$

3. If $n_{1,l} > 0$:

   I. $a_{1,k} = 0$:

$$n_{1,l+1} = n_{1,l};$$

$$n_{2,l+1} = n_{2,l} + \sum_{i=0}^{s_l^*-1} a_{2,k+i} - 1,$$

II. $a_{2,k} > 0$:

$$n_{1,l+1} = n_{1,l} + \sum_{i=0}^{s_l^*-1} a_{1,k+i} - 1;$$

$$n_{2,l} = n_{2,l} + \sum_{i=0}^{s_l^*-1} a_{2,k+i},$$

where the value of s_l^* in the three cases is given respectively as follows:

$$s_l^* = \begin{cases} 0; sincethebufferisempty. \\ s_2(m); sinceaclass\ 2\ packetentersintotheservice\ , \\ s_1(m);\ sinceaclass\ 2\ packetentersintotheservice. \end{cases}$$

With the help of above system equations, we can derive a relation between $N_l(z_1, z_2)$ and $N_{l+1}(z_1, z_2)$. Taking into account that the random variables s_l^* , $(n_{1,l}, n_{2,l})$ and$(a_{1,k+i}, a_{2,k+i})$,$i \geq 0$; are statistically independent, the relation is obtained as follows :

$$N_{l+1}(z_1, z_2) \triangleq E\left[z_1^{n_{1,l+1}} z_2^{n_{2,l+1}}\right]$$

$$= E\left[z_1^{a_{1,k}} z_2^{a_{2,k}}\{n_{1,l} = n_{2,l} = 0\}\right] + E\left[z_1^{\sum_{i=0}^{s_l^*-1} a_{1,k+i}} z_2^{n_{2,l}+\sum_{i=0}^{s_l^*-1} a_{2,k+i}-1}\{n_{1,l} = 0, n_{2,l} > 0\}\right]$$

$$+E\left[z_1^{n_{1,l}} z_2^{n_{2,l}+\sum_{i=0}^{s_l^*-1} a_{2,k+i}-1}\{n_{1,l} > 0, a_{1,k} = 0\}\right]$$

$$+E\left[z_1^{n_{1,l} + \sum_{i=0}^{s_l^*-1} a_{1,k+i}-1} z_2^{n_{2,l}+\sum_{i=0}^{s_l^*-1} a_{2,k+i}}\{n_{1,l} > 0, a_{1,k} > 0\}\right]$$

Solving this equation and using the steady state condition, $N(z_1, z_2) \triangleq \lim_{l\to\infty} N_l(z_1, z_2)$,we obtain the following formula for $\mathrm{N}(z_1, z_2)$:

$$N(z_1, z_2) = \frac{\begin{array}{c}[z_1 z_2 A(z_1, z_2) - z_1 s_2(A(z_1, z_2))]N(0,0) + \\ [z_1\{s_2(A(z_1, z_2)) - s_2(A(0, z_2))\} - z_2 s_1(A(z_1, z_2))]N(0, z_2)\end{array}}{z_1 z_2 - z_1 s_2(A(0, z_2)) - z_2 s_1(A(z_1, z_2))} \quad (1)$$

The two unknown quantities, namely $N(0, z_2)$ and$N(0,0)$, are determined by using Rouche's theorem and putting $z_1 = z_2 = 1$ with the use of de l'Hospital's rule respectively (see also *Walraevens et al.* 2003) and are given by the following expressions:

$$N(0,z_2)=\frac{Y(z_2)\left[s_2\big(A(Y(z_2),z_2)\big)-z_2A(Y(z_2),z_2)\right]}{Y(z_2)\left\{s_2\big(A(Y(z_2),z_2)\big)-s_2\big(A(0,z_2)\big)\right\}-z_2s_1\big(A(Y(z_2),z_2)\big)}N(0,0) \qquad (2)$$

and

$$N(0,0)=\frac{1-\rho}{1-\rho+\lambda_1+\lambda_2} \qquad (3)$$

Using these two equations in eqn. (1), we finally obtain the joint pgf of the system contents at the beginning of a start slot in the steady state with

$$Y(z)=\frac{Y(z)s_2\big(A(0,z)\big)+\ zs_1\big(A(Y(z),z)\big)}{z}$$

The marginal pgfs$N_T(z)$, $N_1(z)$ and $N_2(z)$ of the total system contents and the system contents of class-j at the start slot are obtained by putting appropriate values of z_1 and z_2, which are given as follows:

$$N_T(z) \qquad (5)$$

$$\triangleq \lim_{k\to\infty} E[z^{n_{T,l}}]=N(z,z)$$

$$=\frac{1-\rho}{1-\rho+\lambda_1+\lambda_2}\left[\frac{\begin{array}{c}\{zA_T(z)-s_2\big(A_T(z)\big)\}\Big\{Y(z)\Big(s_2\big(A(Y(z),z)\big)-s_2\big(A(0,z)\\ +\{s_2\big(A_T(z)\big)-s_2\big(A(0,z)\big)-s_1\big(A_T(z)\big)\}\Big\{Y(z)\Big(s_2\big(A(Y(z\end{array}}{\{z-s_2\big(A(0,z)\big)-s_1\big(A_T(z)\big)\}\Big\{Y(z)\Big(s_2\big(A(Y(z),z)\big)-s_2\big(A(}\right.$$

$$N_1(z)\triangleq \lim_{k\to\infty} E[z^{n_{1,l}}]=N(z,1) \qquad (6)$$

$$=\frac{1-\rho_1}{1-\rho_1+\lambda_1}\left[\frac{A_1(z)-s_2\big(A_1(z)\big)}{z-s_1\big(A_1(z)\big)}\right.$$

$$\left.+\lambda_2\left\{\frac{zs_2\big(A_1(z)\big)-s_1\big(A_1(z)\big)}{s_1\big(A_1(z)\big)\Big(z-s_1\big(A_1(z)\big)\Big)}-\mu_2\right\}\right]$$

$$N_2(z) \qquad (7)$$

$$\triangleq \lim_{k\to\infty} E[z^{n_{2,l}}]=N(1,z)$$

$$=\frac{1-\rho}{1-\rho+\lambda_1+\lambda_2}\left[\frac{\begin{array}{c}\{zA_2(z)-s_2\big(A_2(z)\big)\}\Big\{Y(z)\Big(s_2\big(A(Y(z),z)\big)-s_2\big(A(0,z\\ +\{s_2\big(A_2(z)\big)-s_2\big(A(0,z)\big)-s_1\big(A_2(z)\big)\}\Big\{Y(z)\Big(s_2\big(A(Y(\end{array}}{\{z-s_2\big(A(0,z)\big)-zs_1\big(A_2(z)\big)\}\Big\{Y(z)\Big(s_2\big(A(Y(z),z)\big)-s_2\big(A}\right.$$

3. Calculation of Performance Measures

In this section the performance measures such as mean value of the system contents and mean packet delay of class $j,\ j = 1, 2$ packets at the start slot are analysed.

Let us define $\lambda_{i,j}$ and μ_{jj} as

$$\lambda_{i,j} = \left.\frac{\partial^2 A(z_1, z_2)}{\partial z_i \partial z_j}\right|_{z_1 = z_2 = 1}$$

And

$$\mu_{jj} = \left.\frac{d^2 S_j(z)}{dz^2}\right|_{z=1}$$

respectively, with $i, j = 1, 2$.

In this process, the mean values of the system contents i.e., $E[n_1]$ and $E[n_2]$ are obtained by taking the first derivative of the respective pgf's (6), (7) for $z = 1$. Taking the first derivative of (6) for $z = 1$ we get the mean class-1 system content as

$$E[n_1] = \frac{\mu_1\lambda_1\lambda_{11} + 4\mu_1^2\lambda_1^2 - 4\mu_1\mu_2\lambda_1^2 + \mu_{11}\lambda_1^3(1 + \mu_1 - \mu_2) + 2(\lambda_1 + \mu_1^2\lambda_1^3)(\mu_2 - \mu_1)}{2(1 - \mu_1\lambda_1)(1 - \mu_1\lambda_1 + \lambda_1)} \qquad (8)$$

Also $E[n_T]$ is the mean total system content which can be obtained by taking first derivative of $N_T(z)$ for $z = 1$ and is given by

$$E[n_T] = \lambda_T + \frac{A_T''(1)}{2(1 - \lambda_T)} \qquad (9)$$

and also the mean value of the system contents of class 2, $\mathrm{E}[\mathrm{n}_2]$, can be calculated by taking the first derivative of the respective pgf for $z = 1$ and also by using the following relation:

$$E[n_T] = E[n_1] + E[n_2] \qquad (10)$$

Then the mean packet delays i.e.,$E[d_1]$ and $E[d_2]$ are obtained by using Little's law given by

$$E[d_j] = E[n_j]/j \; (j = 1, 2) \tag{11}$$

4. Numerical Example

We assume the traffic of the two classes is arriving according to a two-dimensional binomial process. The joint pgf of the arriving traffic, $A(z_1, z_2)$, is given by:

$$A(z_1, z_2) = \left(1 - \frac{\lambda_1}{N}(1 - z_1) - \frac{\lambda_2}{N}(1 - z_2)\right)^N \tag{12}$$

It is also noticed that if $N \to \infty$, the arrival process is supposed to be a superposition of two Poisson variates. The fraction of class-1 arrivals in the overall traffic mix is defined and denoted by α (i.e.,$\alpha = \lambda_1/\lambda_T$). The remaining part of this section is analysed by taking $N = 16$ and assuming that the service times for both classes are deterministic.

In Fig. 1 the mean value of the system contents of class-1 and class-2 packets is shown as a function of the total loadρ, when$\mu_1 = \mu_2 = 2$, while the fraction of the arrival rate of class 1 trafficα =0.25 and 0.5 respectively of the total arrival rate. It can be easily observed the influence of priority scheduling: the mean of the number of class-1 packets in the system is severely reduced by the HOL priority scheduling; the opposite holds for class-2 cells. In addition, it also becomes apparent that increasing the fraction of high priority cells in the overall mix increases the amount of high priority traffic while decreasing the amount of low priority traffic in the buffer. Finally, it is also clear that the impact of priority scheduling is more important if the load is high.

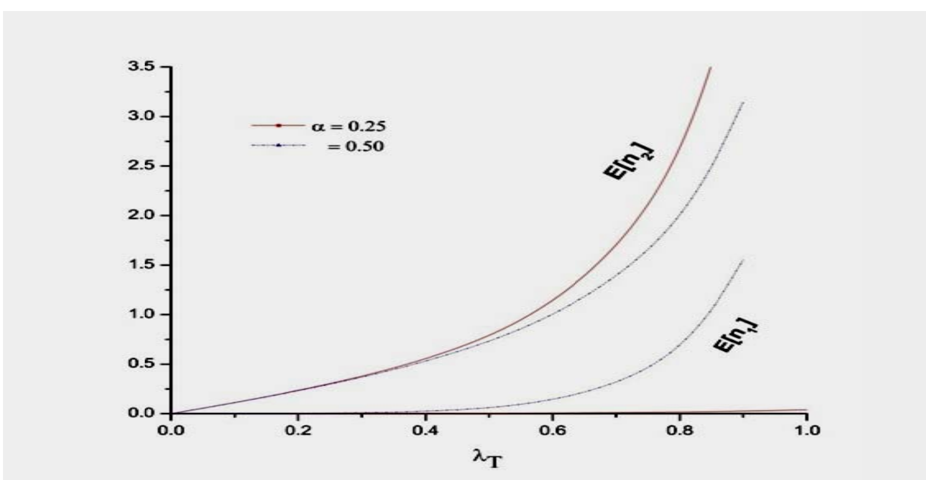

Fig. 1: Mean value of system contents versus the total arrival rate.

Fig. **2** shows the mean value of the packet delay as a function of the total loadλ_T for$\mu_1 = \mu_2 = 2$ while takingthe fraction of the arrival rate of class 1 traffic α =0.25 and 0.5 respectively of the total arrival rate. It can be observed that the influence of HOL non-preemptive priority scheduling with priority jumps is quite large. It can also be observed from these figures that the delay of high and low priority packets increases with increasing the fraction of high priority packets in the overall traffic mix.

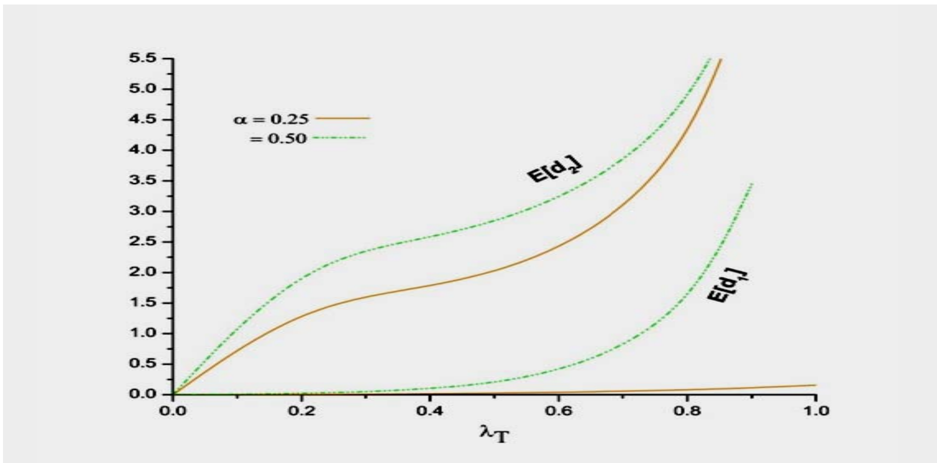

Fig. 2: Mean value of packet delays versus the total arrival rate.

The mean value of the system contents and packet delays of both classes are plotted versus α for λ_T = 0.3 which are shown in **Figs. 3** and **4**. The mean packet delays of the class-1 and class-2 system contents increase with the fraction of class-1 packets in the traffic mix. It is clear from the figure that the difference between class-1 and class-2 packet delays for value of α is large when the load is high.

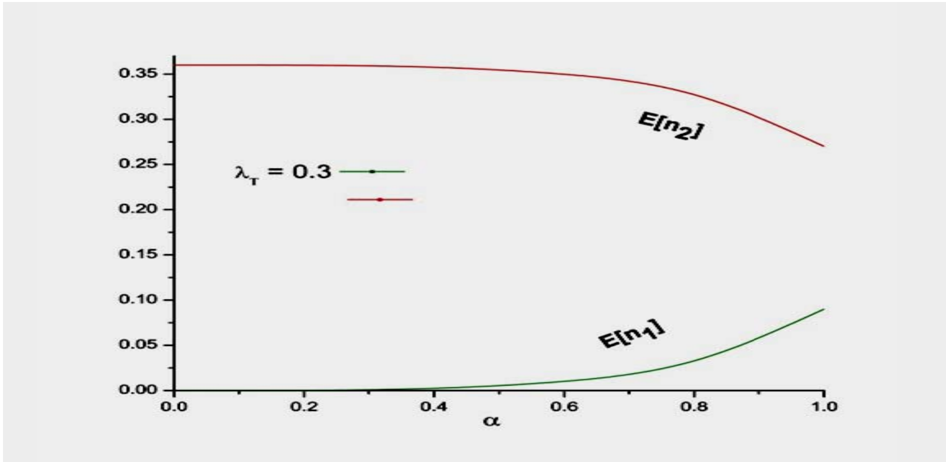

Fig. 3: Mean value of system contents versus the fraction of class-1 arrivals

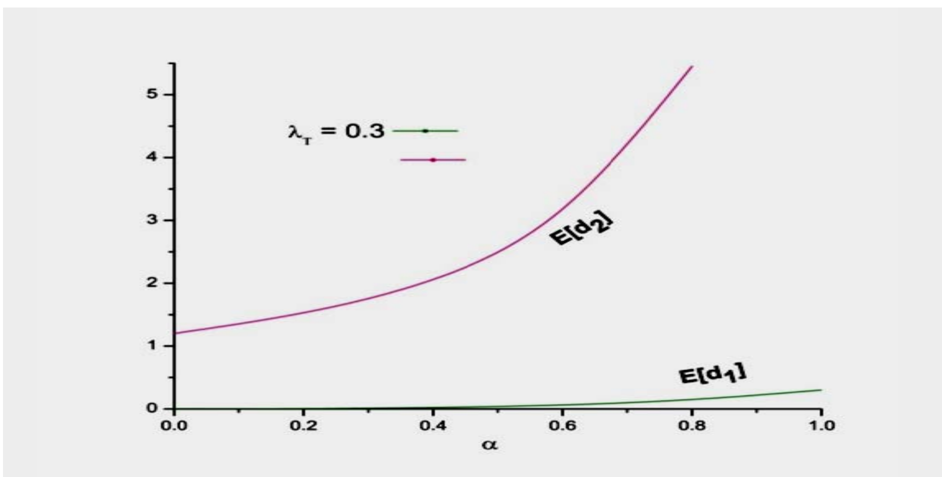

Fig. 4: Mean value of packet delays versus the fraction of class-1 arrivals

5. Conclusion

The system contents in a queueing system with non-preemptive HOL priority jumps scheduling are analysed in this paper. A generating-functions-method was adopted in the study of the performance measures, such as mean of system contents and packet delay of both classes, which are easy to evaluate. The model included possible correlation between the number of arrivals of the two classes during a slot and general service times for packets of both classes. Therefore, the results could be used to analyse performance of an output-queueing switch having Bernoulli arrivals and dynamically prioritized. Finally, the effect of jumping mechanism was analysed which clearly shows that the queueing system provides better results when the fraction of class-1 arrivals in the overall traffic mix is small.

References

[1] Bruneel, H., Steyaert, B., Desmet, E. and Petit, GH. An analytical technique for the derivation of the delay performance of ATM switches with multiserver output queues. *International Journal of Digital and Analog Communication Systems*.**5**(1992) 193-201.

[2]Ishizaki, F. &Takine, T. Analysis of a Discrete-Time Queue with Gated Priority. *Performance Evaluation*.**23**(1995) 121–143.

[3]Kleinrock, L. 1975 *Queueing Systems*. Wiley, New York.

[4] Lee, Y. and Choi, B. D. Queueing System with Multiple Delay and Loss Priorities for ATM Networks. *Information Sciences*, **138:1-4**(2001) **7–29**.

[5]. Lim, Y. and Kobza, J. E. Analysis of a delay-dependent priority discipline in an integrated multiclass traffic fast packet switch. *IEEE Transactions on Communications*.**38(5)**(1990) 659-685.

[6] Maertens,T., Walraevens, J. and Bruneel, H.a On priority queues with priority jumps. *Performance Evaluation*.**63(12)**(2006) 1235-1252.

[7] Maertens, T., Walraevens, J. and Bruneel, H. Performance comparison of several priority schemes with priority jumps. *Ann Oper Res*.**162**(2008) 109-125.

[8] Nassar, H. and Al Mahdy, H. Queueing Analysis of an ATM Multimedia Multiplexer with Nonpreemptive Priority. *IEEE Proceedings on Communications*.**150:3**(2003)**189–196**.

[9] Rubin, I. and Tsai, ZH. Message delay analysis of multiclass priority TDMA, FDMA, and discretetimequeueing systems.*IEEE Transactions on Information Theory*.**35(3)**(1989) 637-647.

[10] Sugahara, A., Takine, T., Takahashi, Y. and Hasegawa, T. Analysis of a nonpreemptive priority queue with SPP arrivals of high class. *Performance Evaluation*.**1(3)**(1995)**215-238**.

[11] Takahashi, M., Osawa, H. and Fujisawa, T. $Geo^X/G/1$ Retrial Queue with Nonpreemptive Priority. *Asia Pacific Journal of Operational Research*.**16:2**(1999)**215–234**.

[12] Takine, T. Sengupta, B. and Hasegawa, T. An analysis of a discrete-time queue for broadband ISDN with priorities among traffic classes.*IEEE Transactions on Communications*, **42(2-4)**(1994) 1837- 1845.

[13] Walraevens, J., Steyaert, B. and Bruneel, H. Performance analysis of a single-server ATM queue with a priority scheduling. *Computers and Operations Research*.**30(12)**(2003) 1807-1829.