ITC19/ ITU & ITC Workshop for Developing Countries 2081-2090
LIANG X.J. and XIN Z.H.(Editors)
V.B. IVERSEN and KUO G.S.(Editors)
Beijing University of Posts and Telecommunications Press

Performance evaluation of discrete time queueing system with variable buffer capacity

Vladimir Efimushkin[a] and Tatiana Ledovskikh[b]

Central Science Research Telecommunication Institute (ZNIIS)
8, 1st Proezd Perova Polya, Moscow, 111141, Russia
[a]ef@zniis.ru
[b]tledovskikh@ zniis.ru

Abstract: The paper is devoted to the study of a finite buffer capacity queueing system $Geom|Geom|1|r_{\text{var}} < \infty$ functioning in discrete time with geometric-type input stream, geometric-type service and variable buffer capacity in steady state. The solution of system's balance equations in matrix form is obtained. The formulas for performance characteristics are derived and some numerical results are discussed. The considered case studies take into account impact of lost customers, used buffer capacity and switching-over operations on system aggregate costs.
Keywords: queueing system of complex structure, variable buffer capacity, system's balance equations, performance characteristics

1. INTRODUCTION

The modern switching equipment of packet networks has advanced procedures of buffer memory allocation that minimize the aggregate loss of packets. The study of these issues have been conducted from the beginning of 80s [1-3], i.e. the period of packet network development and the problem statement of network nodes memory optimization.

At present actuality of the issue have only risen with the development of the Next Generation Networks (NGN), which are oriented to packet technologies and services. For NGN the intermediate storage of huge volume of information (distributed videoconferencing, high quality audio, video on demand and etc.) is required [4]. As a result the study of communication systems with buffer memory allocated to specific direction packets and varying during the system functioning has become actual [5-7]. Thus in queueing systems modelling the communications systems the situations of temporary decrease or increase of buffer capacity to a certain nearest value from an array of beforehand given values appears.

The discrete time queueing systems mostly correspond to the discreteness of the real processes in packet network equipment [8,9]. Significant number of papers are devoted to the study of the discrete time queueing systems (see, e.g. [9-13] and references there).

The paper is organized as follows. The formal description of the system is given in section 2. In section 3 the system's balance equations are obtained, and in section 4 its matrix-recursive type solution is derived. In section 5 and 6 the performance characteristics of the queueing system and some numerical results are presented respectively. Finally, the main results are accumulated in summary.

2. QUEUEING SYSTEM DESCRIPTION

Let us consider a single-server queueing system $Geom|Geom|1|r_{var}<\infty$ with variable buffer capacity r_{var}, $0<r_{var}<\infty$, functioning in discrete time, Figure.1.

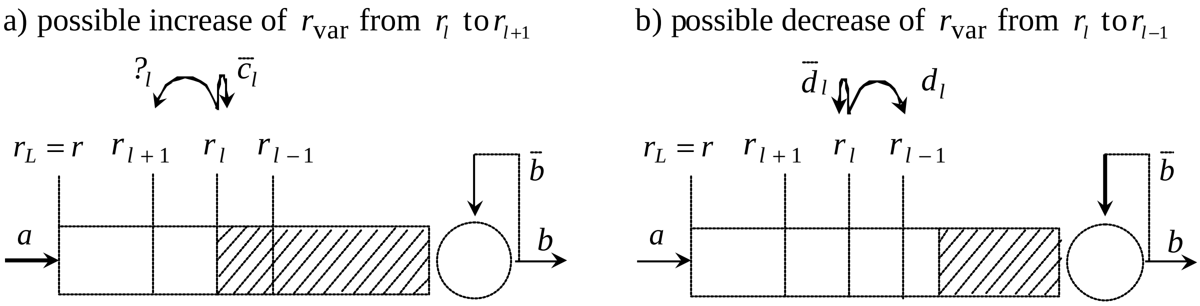

Figure 1 $Geom|Geom|1|r_{var}<\infty$ queueing system.

The time scale is divided into slots of the length h, $h>0$. Let the intervals between the moments of customer arrivals and duration of service be the whole random variables multiple to h, and all events in the system occur at moments nh, $n=1,2,\ldots$ Let the n-th slot be a semi-interval $[(n\text{-}1)h,\ nh)$, and events in the slot n, i.e. at the moment nh, occur in a following sequence: changing of buffer capacity, completion of service, customer arrival and buffering, choice for service, fixing of a system state. The system state in the slot $n+1$ concurs with the state at the moment nh on the completion of all active events happened in slot n.

The input stream and service process are described by geometric distribution with parameters a and b, $0<a<1$, $0<b\le 1$, respectively. The customer arrived when for its buffering there is no free space, is lost, is not renewed and does not affect the queueing system and the input stream.

Let be $r_{var}\in\{r_l, l=\overline{1,L}\}$, $r_l\in N$, where $1\le r_1<r_2<\ldots<r_L=r$, $r<\infty$, $1<L\le r$. The value r_{var} is a buffer capacity at a given moment. Let us consider that the customer under service occupies one buffer place and leaves it at the moment of service completion.

Let hereafter q be a number of customers, and l be a buffer capacity value number. The buffer capacity r_{var} may change its value in a slot n with probability

- c_l, $0<c_l\le 1$, by r_{l+1}, if at the moment $(n-1)h$ $q=r_l$, $l=\overline{1,L-1}$, Figure 1, (a);
- d_l, $0<d_l\le 1$, by r_{l-1}, if at the moment $(n-1)h$ $q<r_{l-1}$, $l=\overline{2,L}$, Figure1, (b).

3. THE SYSTEM'S BALANCE EQUATIONS

Let x_n be a number of customers in the queueing system, and h_n be a number of current buffer capacity value at the moment nh, $n\ge 1$. The operation of the queueing system $Geom|Geom|1|r_{var}<\infty$ can be described by the homogeneous Markov chain $z_n=(x_n,h_n)$, $n\ge 1$, with a set of states $X=\bigcup_{l=1}^{L}X_l$, where $X_l=\{(q,l),\ q=\overline{0,r_l}\}$.

Accordingly to assumptions it follows that the Markov chain z_n is indecomposable and aperiodic. Therefore the final probability distribution $\mathbf{p}^T = (\mathbf{p}_1^T, \mathbf{p}_2^T, \ldots, \mathbf{p}_L^T)$, $\mathbf{p}_l^T = (p_{ql})$, $q = \overline{0, r_l}$, $l = \overline{1, L}$, exists, is unique, $\mathbf{p} > 0$, $\mathbf{p}$ does not depend on initial distribution and equals to the stationary distribution [14].

Let $\mathbf{A}$ be a matrix of transition probabilities of the considered Markov chain, $\mathbf{0}^T = (0, \ldots, 0)$, $\mathbf{I}$ is an unitary matrix. Then it is possible to find $\mathbf{p}$ from the system's balance equations of $|X| = \sum_{l=1}^{L} (r_l + 1) = r_\bullet + L$ order and |*X*|-1 rank:

$$\mathbf{p}^T (\mathbf{A} - \mathbf{I}) = \mathbf{0}^T,$$

with the normalization condition $\mathbf{p}^T \mathbf{1} = 1$. Here the point instead of a subscript is used for marking the full sum of this index.

The partition of X into subsets X_l, $l = \overline{1, L}$, induces the partition of the matrix $\mathbf{A}$ into blocks. Due to ordinary input stream the matrix $\mathbf{A}$ is block three-diagonal with blocks $\mathbf{A}_{lm} = \left\| a_{lm}(i, j) \right\|_{\substack{i=\overline{0,r_l} \\ j=\overline{0,r_m}}}$. The elements of blocks $\mathbf{A}_{lm}$ are as follows:

$$a_{ll}(i,j) = \begin{cases} \overline{d}_l^{\,u(r_{l-1})}\overline{a}, & i = j = 0, \\ \overline{c}_l^{\,d(i,r_l)u(r_L - i)}\,\overline{d}_l^{\,u(r_{l-1}-i)}(\overline{b}\,\overline{a}^{\,u(r_l - i)} + ba), & i = j = \overline{1, r_l}, \\ \overline{d}_l^{\,u(r_{l-1}-i)}\overline{b}^{\,u(i)}a, & j = i+1,\ i = \overline{0, r_l - 1}, \\ \overline{c}_l^{\,d(i,r_l)u(r_L - i)}\,\overline{d}_l^{\,u(r_{l-1}-i)}b\overline{a}, & j = i-1,\ i = \overline{1, r_l}, \\ 0, & \text{in other cases}, \end{cases} \quad l = \overline{1, L},$$

$$a_{l,l+1}(i,j) = \begin{cases} c_l b\overline{a}, & i = r_l,\ j = r_l - 1, \\ c_l(\overline{b}\,\overline{a} + ba), & i = j = r_l, \\ c_l\overline{b}a, & i = r_l,\ j = r_l + 1, \\ 0, & \text{in other cases}, \end{cases} \quad l = \overline{1, L-1},$$

$$a_{l,l-1}(i,j) = \begin{cases} d_l\overline{a}, & i = j = 0, \\ d_l(\overline{b}\,\overline{a} + ba), & i = j = \overline{1, r_{l-1} - 1}, \\ d_l\overline{b}^{\,u(i)}a, & j = i+1,\ i = \overline{0, r_{l-1} - 1}, \\ d_l b\overline{a}, & j = i-1,\ i = \overline{1, r_{l-1} - 1}, \\ 0, & \text{in other cases}, \end{cases} \quad l = \overline{2, L},$$

where

$$u(x) = \begin{cases} 1, & x > 0 \\ 0, & x \le 0, \end{cases} \quad d(i,j) = \begin{cases} 1, & i = j \\ 0, & i \ne j, \end{cases}$$ and, by definition, $r_0 = 0$.

Figure 2 illustrates obtaining of matrix $\mathbf{A}$ elements by the example of $a_{ll}(i, j)$, $l = \overline{1, L}$; we can apply similar considerations to other probabilities.

The state $(0,l)$ will not be changed, if the buffer capacity isn't reduced, and a new customer doesn't arrive; under this condition $\overline{d}_l^{\,u(r_{l-1})} = 1$ for $r_{\mathrm{var}} = r_1$, since $r_0 = 0$.

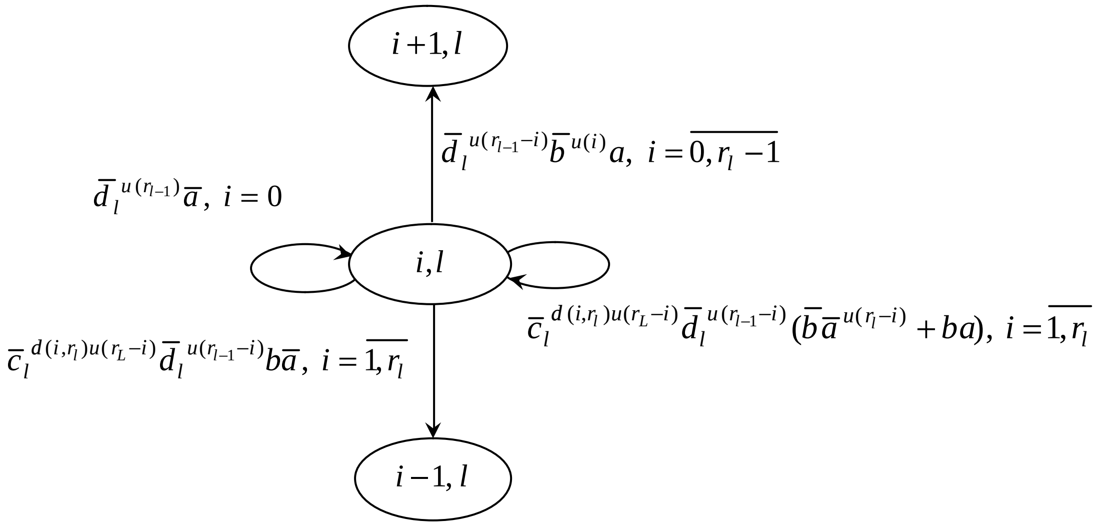

Figure 2 The transition diagram from state (i,l), $i=\overline{0,r_l}$, $l=\overline{1,L}$.

With other values of i the state (i,l) will not be changed

- for $i=r_l$, $l \neq L$, if the buffer capacity doesn't increase (with probability $\overline{c}_l$), and also if a customer on the server leaves the system and a new one arrives (ba), or if a customer doesn't leave the system, and a new one doesn't arrive because of the full buffer ($\overline{b}\,\overline{a}^{\,u(r_l-i)}$). For considered values of i buffer capacity can't be reduced ($\overline{d}_l^{\,u(r_{l-1}-i)}$);

- for $i=r_L$, if a customer on a server leaves the system and a new one arrives, or if a customer doesn't leave the server and a new one can not arrive because of the full buffer. The buffer capacity can neither increase being maximum, nor decrease because of non-fulfilment of reduction condition: number of customers in the system must be less than r_{L-1}.

In case of $i=\overline{0,r_l-1}$ the number of customers can increase if the value of buffer capacity retains ($\overline{d}_l^{\,u(r_{l-1}-i)}$), there is no completion of service ($\overline{b}^{\,u(i)}$) and a customer arriving.

For $i=\overline{1,r_l}$ the number of customers can be reduced due to the service completion with retaining of the buffer capacity value ($\overline{c}_l^{\,d(i,r_l)u(r_L-i)}\overline{d}_l^{\,u(r_{l-1}-i)}$) and not arriving of a customer.

Assume $\mathbf{A}'_{ll}=\mathbf{A}_{ll}-\mathbf{I}$, $l=\overline{1,L}$.

The solution of the system's balance equations has a recursive presentation

$\mathbf{p}_m^T=\mathbf{p}_{m-1}^T\mathbf{W}_{m-1}$, $m=\overline{2,L}$,

$\mathbf{W}_{m-1}=-\mathbf{A}_{m-1,m}(\mathbf{A}'_{mm}+\mathbf{W}_m\mathbf{A}_{m+1,m})^{-1}$, $m=\overline{2,L-1}$,

$\mathbf{W}_{L-1}=-\mathbf{A}_{L-1,L}(\mathbf{A}'_{LL})^{-1}$;

where the vector $\mathbf{p}_1$ is obtained from the system of equations $\mathbf{p}_1^T\mathbf{W}'=\mathbf{e}^T$. Here $\mathbf{W}'$ is the matrix $\mathbf{A}'_{11}+\mathbf{W}_1\mathbf{A}_{21}$ with its right column replaced by vector $\sum_{l=0}^{L-1}(\prod_{m=1}^{l}\mathbf{W}_m)\mathbf{1}$; $\mathbf{e}^T=(0,\ldots,0,1)$.

The solution can also be obtained by the block UL-decomposition method, or the method of exclusions, stated in [15, p.26].

Obtained expressions for calculation of the Markov chain stationary distribution allow carrying out required numerical analysis of introduced performance characteristics which give a possibility to study different buffer capacity control algorithms.

5. CASE STUDIES

Figure 3 shows the plots of the loss probability p as a function of the customer arrival probability a in a slot with the following parameter values: $r = 30$, $L = 5$, $b = 0{,}8$, $c_l = 0{,}5$, $l = \overline{1,4}$, and $d_l = 0{,}5$, $l = \overline{2,5}$. We study three cases of r_{var} buffer capacity values: (a) the case of uniform values allocation; (b) the case of small values; (c) the case of great values.

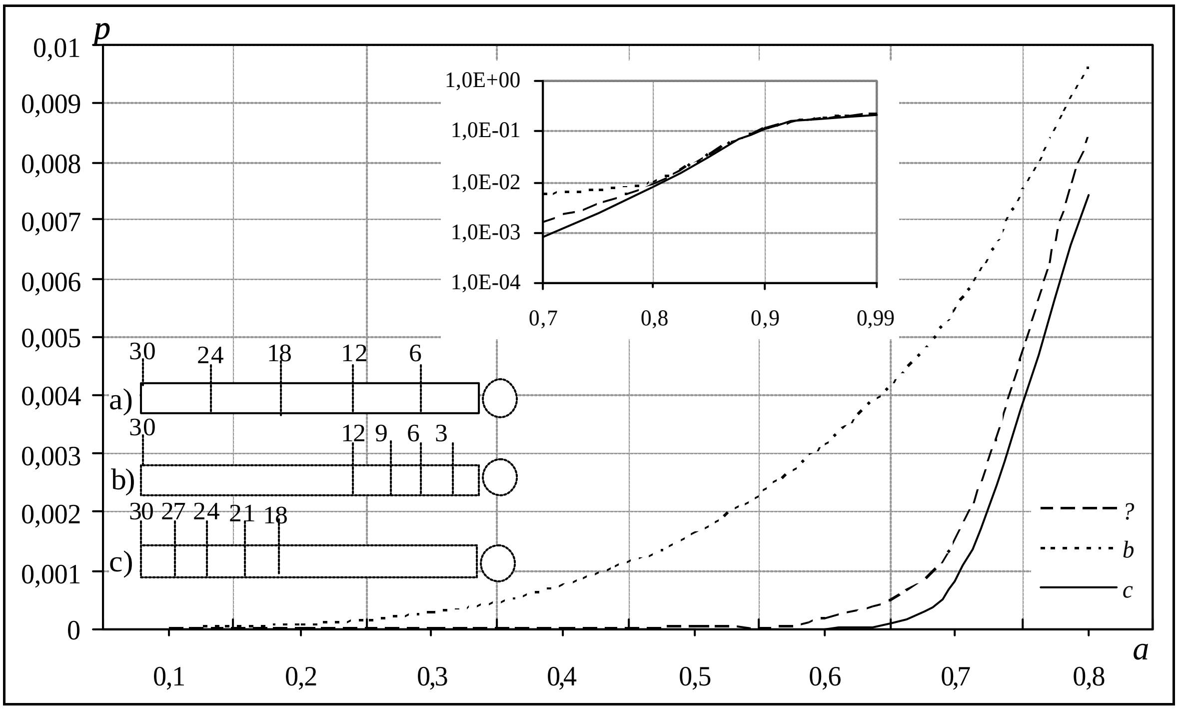

Figure 3 Impact of irregularity of r_{var} values allocation on p .

We can make the following conclusions about the plots behaviour for different variants of r_{var} values allocation. For small values of r_{var} (b) the loss probability has the greatest values, and for great values (c) it is less as compared with the cases (a) and (b) for overall interval of the customer arrival probability. Under the uniform allocation of r_{var} values (a) the plot of loss probability is situated between the considered curves. It can be explained by the following. For small values of r_{var} (b) the availability of free buffer space is essentially less then for (a) and extra less then for (c). While a increases the current value r_{var} verges towards 30 in all cases which leads to fast raise and approaching of curves.

Similar plots are shown in Figure 4 for $L \in \{2,4,5,10,20\}$ and uniform allocation of buffer capacity values; for example, for $L = 2$ – $r_{\text{var}} \in \{10,20\}$, for $L = 4$ – $r_{\text{var}} \in \{5,10,15,20\}$. We consider the following parameter values: $r = 20$, $b = 0{,}5$, $c_l = 0{,}5$ and $d_l = 0{,}5$ for all l. The monotone increasing dependence of p on L is obvious.

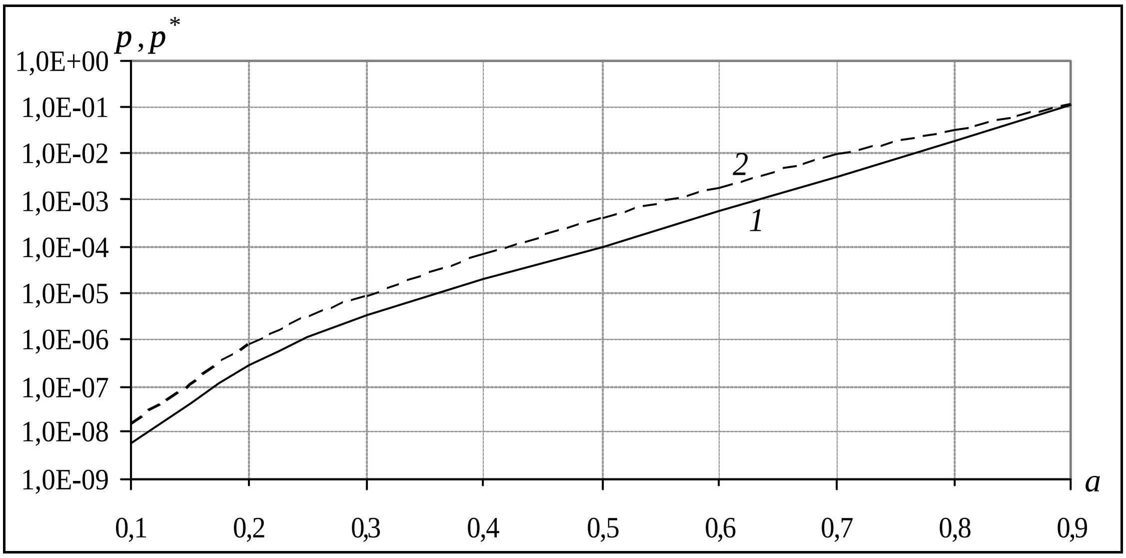

Figure 5 Loss probabilities p and p^* as a function of a customer arrival probability.

In special case of $g_l^+ = g^+$, $g_l^- = g^-$, $f_l = f$ for all values of l, the aggregate cost is equal $S(\vec{r}) = fp + g^+ p^+ + g^- p^-$, and in a case of $g^+ = g^- = g$: $S(\vec{r}) = fp + gp$.

In the second considering system the changes of buffer capacity and corresponding costs are absent so $S(r^*) = f_{r^*} \bar{b} p_{r^*}$. The loss cost f_{r^*} for the buffer capacity r^* can be assumed equal to f_l, where r_l is the nearest integer to r^*.

Figure 7 shows the plots of the aggregate costs $S(\vec{r})$ and $S(r^*)$ as a function of customer arrival probability a per slot under the condition of uniform allocation of buffer capacity values and under the following parameter values: $r = 20$, $L = 4$, $r_{\mathrm{var}} \in \{5,10,15,20\}$, $b = 0{,}8$, $c_l = 0{,}5$, $l = \overline{1{,}3}$; $d_l = 0{,}5$, $l = \overline{2{,}4}$, $g = 10$ and $f = f_{r^*} \in \{10,100,1000\}$. Here (1) corresponds to $S(\vec{r})$, and (2) corresponds to $S(r^*)$. It should be noted that when the cost values of $f = f_{r^*} = 10$ and g are near the same the overall costs for the case (1) are greater then for the case (2).

And at last if we take into consideration in costs parameters a cost value v_q of buffer capacity q, then the costs $\hat{S}(\vec{r})$ and $\hat{S}(r^*)$ for corresponding systems are

$$\hat{S}(\vec{r}) = S(\vec{r}) + \sum_{l=1}^{L} v_{r_l} p_{\bullet,l},$$

$$\hat{S}(r^*) = S(r^*) + v_{r^*}.$$

Figure 8 shows the plots of $\hat{S}(\vec{r})$ and $\hat{S}(r^*)$ as a function of probability a with uniform allocation of r_{var} values and the following parameters: $r = 20$, $L = 4$, $r_{\mathrm{var}} \in \{5,10,15,20\}$, $b = 0{,}8$, $c_l = 0{,}5$, $l = \overline{1{,}3}$; $d_l = 0{,}5$, $l = \overline{2{,}4}$, $g = 10$, several values of f and v_{r_l}. Here (1) corresponds to $\hat{S}(\vec{r})$, and (2) corresponds to $\hat{S}(r^*)$. For calculating of $\hat{S}(r^*)$ we use v_{r^*} values according to calculated buffer capacity r^* for the queueing system $Geom \mid Geom \mid 1 \mid r^* < \infty$.

The study has shown the advantage of proposed system as compared with $Geom|Geom|1|r^* < \infty$ queueing system in terms of loss probability. Among $Geom|Geom|1|r_{\text{var}} < \infty$ and $Geom|Geom|1|r^* < \infty$ systems, the former has lower aggregate cost (for lost customers, used buffer capacity and switching-over operations) for considered parameter sets with cost of used buffer capacity prevailing cost of switching-over operations. Further study of the proposed system will consider different probability distributions for buffer capacity switching-over operations.

REFERENCES

1. M. Irland, Buffer Management in a Packet Switch, IEEE Trans. Commun., 1978, No.3, pp.328–337.
2. F. Kamoun and L. Kleinrock, Analysis of shared finite storage in a computer networks under general traffic conditions, IEEE Trans. Comp., 1980, No.7, pp.992-1003.
3. G. Latouch, Exponential Servers Sharing a Finite Storage: Comparison of Space Allocation Policies, IEEE Trans. Commun., 1980, No.6, pp.910–915.
4. N. Wilkinson, Next Generation Network Services: Technologies & Strategies, John Wiley & Sons, Ltd.,Chichester, 2002.
5. M.J. Karam and F.A. Tobagi, Rate and Queue Controlled Random Drop: A Buffer Management Scheme for Internet Routers // Proc. Globecom '2000, San Francisco, CA.
6. G. Nong and M. Hamdi, On the Provision of Quality-of-Service Guarantees for Input Queued Switches, IEEE Commun. Magazine, 2000, No.12, pp.62-69.
7. U. Briem, E. Wallmeier, C. Beck and F. Matthiese, Traffic Management for an ATM Switch with Per-VC Queueing: Concept and Implementation, IEEE Communications Magazine, 1998, No.1, pp.47-52.
8. H. Kobayashi and A. Konheim, Queueing Models for Computer Communications System Analysis // IEEE Trans. Commun., 1977, No.1, pp.2-29.
9. G.P. Basharin and V.A. Efimushkin, The study of the single-server system in discrete time with customers of several types, Problems of Information Transmission, 1984, No.1, pp.95-104 (in Russian).
10. H. Bruneel, Performance of Discrete Time Queueing Systems, Computers and Operations Research, 1993, pp.303-320.
11 M.E. Woodward, Communication and Computer Networks: Modelling with Discrete-Time Queues, Washington, Tokyo: IEE Pres, 1994.
12. H. Daduna, Discrete Time Queueing Networks. Recent Developments, Tutorial Lecture Notes. Performance' 96, Lausanne, 1996.
13. V.A. Efimushkin and T.V. Ledovskikh, Analysis of the discrete time queueing system with burst phase-type input model, Proc. of St. Petersburg Regional Int. Teletraffic Seminar «Telecommunication Network and Teletraffic Theory», 2002, pp.200-214.
14. W. Feller, An introduction to probability theory and its applications. Vol.1, John Wiley and Sons, 1968.
15. P.P. Bocharov, C. D'Apice, A.V. Pechinkin and S. Salerno, Queueing theory. Ultrecht Boston: VSP, 2004.