Analysis of one type of communication systems using software and probabilistic methods

A L Reznik [1], A A Soloviev [1], A V Torgov [1]

[1] Probability Research Methods for Information Processing Lab, Institute of Automation and Eletrometry SB RAS, Academician Koptyug ave. 1, Novosibirsk, 630090, Russia.

Abstract. The study is devoted to one of the tasks related to information and computer systems with a different discipline of establishing connections between its nodes. Cases of systems with and without queues are considered. The efficiency of the systems (the relative load, the number of busy nodes, etc.) is calculated. To calculate the probability of the communication system being in one of the possible enlarged states, the corresponding Markov process is introduced and its asymptotic characteristics are determined.

1. Introduction

In the development and design of telecommunication nodes in data transmission systems, as well as in telephone and computing servers, there are many tasks associated with determination the optimal parameters of its work in various conditions (the level of relative system load, average call length, etc.) [1,2]. Since the operation and maintenance of a telecommunication system is expensive, one of the important issues is to determine the minimum set of required equipment (for example, the number of servers) which allows to hold the percentage of customers experiencing communication troubles below a certain acceptable level [3]. Particularly acute are the problems of stable operation of equipment when labor-intensive applied and communication tasks are solved in autonomous computer systems (for example, in satellite airborne navigation and computer systems when processing trajectory and navigation information [4,5]).

Note that even when using the same equipment, the overall performance of the communication and computing system and the waiting time for the service depend on the queue processing method [6] (the so-called "queue discipline"). Let's consider this question in more detail. There are a number of customers. Any of them may demand to establish a communication with any other. Communication is established if the requested customer is free, otherwise the customer who has requested for the communication is queued. It is assumed that there are restrictions on the length of the queues, therefore, in the general case, communication request may be lost. A customer is considered to be busy if he is connected with some other customer or is waiting in line. Busy customers do not do communication requests. Schematically, the possible state of the system is presented in Figure 1.

There are many ways the queue can be processed: FCFS (first-come-first-serve discipline), LCFS (last-come-first-served), or even SIRO (service-in-random-order). In fact, the queuing discipline practically does not affect the relative system load [7] (but can greatly increase the waiting time of some customers in line). On the other hand, a complete queue prohibition significantly vary all the operation parameters of the system. This article discusses a special case of the stated problem, when all customers are "symmetrical" in the sense of the nature of their submission of communication requirements, maintenance of communications between customers, as well as the organization and length of the queues for each of the customers. Research is carried out both for the case when there are no restrictions on the length of the queues, and for the case when the queues are prohibited.
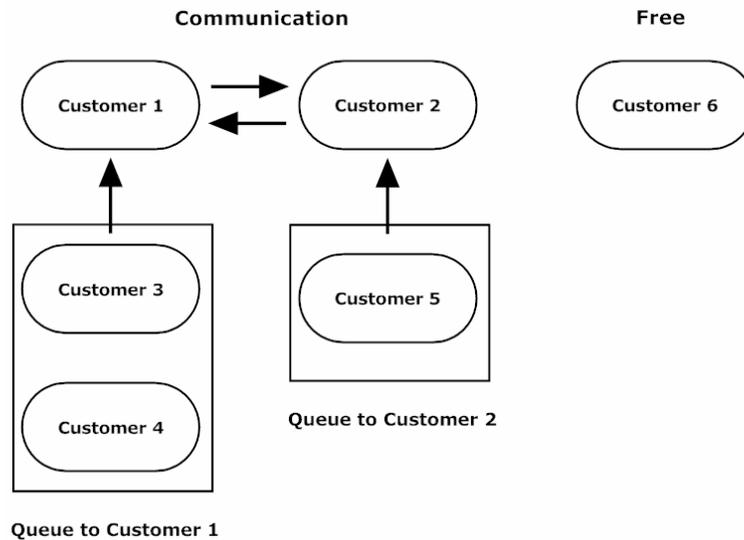
Figure 1. Possible state of the system.

Here deserve attention such characteristics as the distribution of the number of simultaneously established connections and the number of waiting in the queue, the distribution of the waiting time for the customer to provide him with the required connection, and in case of limitation on the queue length, also the probability of request's loss.

2. No-limit queue length

There are N customers; all of them can request for the communication with any (but only one) of $N\text{-}1$ other customers. If the i-th customer has made a request for communication with the j-th customer, then the communication between them begins immediately if the j-th customer is free, otherwise the i-th customer is put on the waiting list (waiting line) of j-th customer and waiting for his request to be satisfied in a fair (sequential) order. If the customer has made a request for the communication, then the probability that this is a request for a communication with a specific customer is equal to $1/(N\text{-}1)$. The queue length for each customer is unlimited (it is enough to set the number of places in the queue equal to $N\text{-}2$). The duration of the communication session between any two customers is exponentially distributed with the parameter μ.

The event that any free customer will not make a request for the communication with any other customer during time τ has a probability of $\exp(-\lambda\tau)$, $\lambda>0$. It is required to determine the parameters characterizing the number of simultaneous communications and the number of busy customers depending on the total number of system customers, the intensity of generation and the duration of the service of incoming connection requests. Due to the lack of accurate analytical formulas describing the parameters of this kind of communication systems, in practice, researchers often resort to methods of numerical or simulation modeling [8]. In our case, to study the described class of communication systems, it is proposed to use a probabilistic software approach based on the analysis of their enlarged states.

The state of the system at an arbitrary point of time will be characterized by the vector $\alpha=(\alpha_1, \alpha_2,…, \alpha_k, \alpha_m)$, $m = [N/2]$, where $[x] = entire\ (x)$; α_k – the number of busy customers at time t, «belonging» to k-th communication (that includes customers who are in the process of the k-th communication and those who are in line with above customers). It is assumed that $N\geq\alpha_1\geq…\geq\alpha_m\geq 0$ (so, communications are ordered by the number of customers they own). The random process introduced by the above method is Markovian with the transition probability matrix determined from the relations:

$$P\{(\alpha_1, \alpha_2, \ldots, \alpha_k, \alpha_m)_{t+\Delta t} / (\beta_1, \beta_2, \ldots, \beta_m)_t\} =$$

$$= \begin{cases}
1 - \lambda\gamma(\overline{\beta})\Delta t - \mu\omega(\overline{\beta})\Delta t, if\ \alpha_k = \beta_k\ (k = \overline{1,m}); & (1.1) \\[2mm]
\dfrac{\lambda\beta_i\gamma(\overline{\beta})\xi_i(\overline{\beta})\Delta t}{N-1}, if\ \alpha_k = \beta_k\ (k = 1,\ldots,i-1,i+1,\ldots,m), & (1.2) \\[1mm]
\qquad\qquad \alpha_i = \beta_i + 1,\ 2 \le \beta_i \le N-1 & \\[2mm]
\dfrac{\lambda(\gamma(\overline{\beta})-1)\gamma(\overline{\beta})\Delta t}{N-1}, if\ \alpha_k = \beta_k\ (k = 1,\ldots,i-1,i+1,\ldots,m),\ \alpha_i = 2, \beta_i = 0; & (1.3) \\[2mm]
\dfrac{\mu\xi_i(\overline{\beta})(2-\delta(q,\beta_i - q))\Delta t}{\beta_i - 1}, & (1.4) \\[1mm]
\text{if the set of nonzero vector's coordinates } (\alpha_1,\ldots,\alpha_m) \text{ coincides with the set of nonzero} & \\
\text{components of the selection } \{\beta_1,\ldots,\beta_{i-1},\beta_{i+1},\ldots,\beta_m,q,\beta_i\text{-}q\}, \beta_i \le 4, 2 \le q \le [\beta_i/2]; & \\[2mm]
\dfrac{2\mu\xi_i(\overline{\beta})\Delta t}{\beta_i - 1}, if\ \alpha_k = \beta_k\ (k = 1,\ldots,i-1,i+1,\ldots,m), \alpha_i = \beta_i - 1, \beta_i \ge 3; & (1.5) \\[2mm]
\mu\xi_i(\overline{\beta})\Delta t, if\ \alpha_k = \beta_k\ (k = 1,\ldots,i-1,i+1,\ldots,m),\ \alpha_i = 0, \beta_i = 2; & (1.6) \\[2mm]
0, \qquad otherwise & (1.7)
\end{cases}$$

The following notation is introduced here: $\delta(i,j) = \dfrac{1, i = j}{0, i \ne j}$ – Kronecker symbol;

$\gamma(\overline{\beta}) = (N - \sum\limits_{k=1}^{m}\beta_k)$ – the total number of free customers; $\omega(\overline{\beta}) = (m - \sum\limits_{k=1}^{m}\delta(\beta_k,0))$ – the total

number of current communications; $\xi_i(\overline{\beta}) = \sum\limits_{k=1}^{m}\delta(\beta_i,\beta_k)$ – the multiplicity of occurring of the

component (numerically equal to β_i) in the vector $\overline{\beta} = (\beta_1,\ldots,\beta_m)$.

The rationale of the elements (1.1)-(1.7) for the transition probability matrix is given below. The transition probability (1.1), i.e., the probability that the system remains in the same state for an infinitely small time interval Δt, is

$$\exp(-\lambda\gamma(\overline{\beta})\Delta t)\exp(-\mu\omega\Delta t) = 1 - \lambda\gamma(\overline{\beta})\Delta t - \mu\omega(\overline{\beta})\Delta t + o(\Delta t).$$

Transition probability (1.2), that is, the probability that none of the communications $\omega(\beta)$ will be over during time Δt, and one of the free customers will make a request for communication with one of the customers belonging to the i-th communication, is equal to

$$\exp(-\mu\omega(\overline{\beta})\Delta t)\exp(-\lambda(\gamma(\overline{\beta})-1)\Delta t)(1-\exp(-\lambda\Delta t)\times$$

$$\times\left(\frac{\gamma(\overline{\beta})}{1}\right)\frac{\beta_i}{N-1}+o(\Delta t)=\frac{\lambda\beta_i\gamma(\overline{\beta})\Delta t}{N-1}+o(\Delta t).$$

The transition probability (1.3), that is, the probability that no communications will be completed during time Δt and one of the free customers will make a request for communication with another free customer is

$$\exp(-\mu\omega(\overline{\beta})\Delta t)\exp(-\lambda(\gamma(\overline{\beta})-1)\Delta t)(1-\exp(-\lambda\Delta t)\times$$

$$\times\left(\frac{\gamma(\overline{\beta})}{1}\right)\frac{\gamma(\overline{\beta})-1}{N-1}=\frac{\lambda(\gamma(\overline{\beta})-1)\gamma(\overline{\beta})\Delta t}{N-1}+o(\Delta t).$$

The transition probability (1.4), that is, the probability that i-th communication will end within a period of time Δt, and the end of this communication will immediately lead to the organization of two new ones (special cases when the end of a communication leads to the creation of only one new or a new communication does not form at all, are described by transitional probabilities (1.5) and (1.6), respectively).

To calculate the transition probability (1.4), we use the following technique. We assign each of customers β_i to one of two subgroups – left or right (it depends on the line in which it stands – to the first or second customer doing i-th communication). Then the group of β_i customers can be formed in (β_i-1) different ways: (j,β_i-j), $j=1,\ldots,\beta_i-1$. We prove by induction that all methods are equally probable, that is, the probability of each of them is equal to $1/(\beta_i-1)$. So, let $P(j,n-j)$ be the probability that a group of n customers belonging to one communication consists of j left and $n-j$ right ($j=1,\ldots,n-1$) customers. Let us note (this will be needed later) that $P(n,0)=P(0,n)=0$ for any $n\geq 2$. Now suppose that for some n and arbitrary $j=1,\ldots,n-1$ probability $P(j,n-j)$ is equal $1/(n-1)$. For $n=2$ the validity of the assumption is obvious. Then the group of $(n+1)$ customers consisting of j left and $(n+1-j)$ right could be constructed in two ways: either the last customer was added to $(j-1)$ left, or to $(n-j)$ right. Considering that all customers can be demanded with equal probability,

$$P(j,n+1-j)=\frac{P(j-1,n+1-j)(j-1)}{n}+\frac{P(j,n-j)(n-j)}{n}.$$

Since by assumption of induction $P(j+1,n+1-j)=P(j,n-j)=1/(n-1)$, then, $P(j,n+1-j)=1/n$ for any $j=1,\ldots,n$. Thus, the validity of the induction hypothesis is proved for the arbitrary n.

In the transition probability formula (1.4), a factor $(2-\delta(q,\beta_i-q))$ requires some explanation: since in the group of βi customers the subgroups (q left, β_i-q right) and (β_i-q left, q right) do not differ, the corresponding probability should be doubled, except for the case $q=\beta_i-q$.

Transitional probability (1.5) is, as already noted, a special case of transitional probability (1.4) with the only difference that the i-th communication that ends during interval Δt has either a single left or a single right customer. In this case, instead of one ending, one new communication is formed.

Transitional probability (1.6) is the probability of termination of a communication between two customers of the system, where both customers had no queue. In this case, of course, a new communication does not form. Since the transition probabilities are independent of t, the introduced Markov process with a finite number of states is homogeneous and there are limit probabilities for it

$$\pi(\alpha_1,...,\alpha_m) = \lim_{t \to \infty} P(\alpha_1,...,\alpha_m)_t,$$

being a solution to a system of equations [9]

$$
\begin{cases}
\pi(\alpha_1,...,\alpha_m) = \sum_{(\beta_1,...,\beta_m)} \pi(\beta_1,...,\beta_m) P\{(\alpha_1,...,\alpha_m)/(\beta_1,...,\beta_m)\}, \\
\sum_{(\alpha_1,...,\alpha_m)} \pi(\alpha_1,...,\alpha_m) = 1.
\end{cases}
\tag{2}
$$

The transition probabilities in (2) are determined from (1.1)-(1.7). The problem of estimating the dimension of system (2) can be formulated in terms of distributing balls into boxes [10]. Two possible placements are considered indistinguishable if one of them can be obtained from the other by rearranging the boxes. It is required to determine

$$\Lambda(N) = \sum_{n=0}^{N} s\left(n, \left[\frac{N}{2}\right]\right),$$

where $s(n, [N/2])$ – the number of distinguishable locations of n balls in $[N/2]$ boxes, and any box may contain $0,2,3,...,t$ balls, but it is forbidden to place one ball in the box. To estimate the complexity of the system of algebraic equations (2) (i.e. various states of the Markov process) we used the relation

$$\Lambda(N) = 1 + \sum_{n=2}^{N}\left\{\left[\frac{n}{2}\right] + \sum_{r=3}^{\left[\frac{n}{2}\right]}\left(\sum_{v_1=2}^{\left[\frac{n}{r}\right]}\sum_{v_2=v_1}^{\left[\frac{n-v_1}{r-1}\right]}... \sum_{v_i=v_{i-1}}^{\left[\frac{n-v_1-...-v_{i-1}}{r-(i-1)}\right]} ... \sum_{v_{r-1}=v_{r-2}}^{\left[\frac{n-v_1-...-v_{r-2}}{2}\right]}1\right)\right\}
\tag{3}$$

The number of states of the system according to formula (3) is growing very quickly. So, it is necessary to solve a system of 42 equations for $N=10$, a system of 5604 equations for $N=30$, and a system of 204226 equations for $N=50$.

It should be noted that estimation of the number of system states in this task is equivalent to the well-known problem of partition of the integer number N. Indeed, it can be easily verified if, in order to obtain all possible enlarged states of the communication system described by the transition probability matrix (1), in each of the partitions of the integer N replace "1" by "0". Example for $N = 4$ is shown in table 1.

Table 1. Correspondence between tasks

Partitions of integer number 4	Possible system's state for 4 customers)
4	4
3+1	3+0
2+2	2+2
2+1+1	2+0+0
1+1+1+1	0+0+0+0

Generally speaking, the problem of decomposing the natural number N into natural terms was formulated by Gottfried Leibniz as early as 1654, and the recurrence formula was obtained and proved by Leonard Euler in 1740. Nevertheless, a closed analytical solution to this problem is unknown today. Accordingly, the exact analytical formula for expression described by relation (3) is unknown too. Therefore, to evaluate the complexity of the system of linear equations arising in the calculation of transition probabilities matrix (1.1)-(1.7), we used the asymptotic formula (Hardy, Ramanujan) [11]:

$$s(n) \approx \frac{1}{4n\sqrt{3}} e^{\pi\sqrt{2n/3}} .$$

Calculations were made on modern computing systems based on Matlab using parallel computing up to $N=20$, and for systems with $N>20$, the results were determined using extrapolation of the obtained data. Since a proportional increase of the values λ and μ does not affect the final solution of system (2), the initial data for the program calculating the probabilities $\pi(\alpha_1,\ldots,\alpha_m)$ were N and μ/λ.

Let us briefly consider the main ideas of the algorithm. After entering the data, the matrix is calculated IA(i,j), each line of which characterizes one of the possible states of the system. The structure of the string is shown in Figure 2.

Figure 2. Structure of the string

Here NRC_i – total number of active customers in this particular state (in communication and in line), KRC_i – number of communications in this state, n_j^i – the number of groups of customers related to one communication, each of which contains exactly j customers. Some redundancy of information (cells NRC_i and KRC_i) is used to increase the speed of the calculation. For the same purpose, the rows of the matrix IA are ordered in ascending NRC_i order, and inside the rows with the same NRC_i value – in ascending KRC_i order.

Next, a pointer matrix IND (NRC, KRC) is created, it's elements are the minimum row numbers i, for which IA(i,0)=NRC, IA(i,1)=KRC. To be precise, there are several such lines, but they are all arranged sequentially, starting from i-th string. The construction of the matrix of infinitesimal coefficients (7) is more convenient (in the terms of the efficiency of the resulting program), not by rows (and not by direct enumeration, as it is formally written in (1)-(6)), but by columns. To archive this, a software simulation of the transition of each rows of matrix IA(i,j) to all adjacent states is done. In this case, the current values NRC_i and KRC_i can be changed, as well as the values of some of n_j^i (i=2,3,...,N). The infinitesimal coefficient itself is calculated according to one of the rules (1.1) - (1.6). A significant reduction in the counting time during the computational procedure organized in this way is ensured by the fact that the line number corresponding to some adjacent state (into which the current line can transform) can be found trivially using a pointer matrix.

The calculation of stationary probabilities $\pi(\alpha_1,\ldots,\alpha_m)$ is completed by solving the system of linear equations (2). In addition to calculating the probabilities $\pi(\alpha_1,\ldots,\alpha_m)$, a software calculation was made of the relative load of the system, that is, the ratio of the average number of communications $\overline{\omega}$ to the maximum possible [N/2], as well as the calculation of the relative number of busy customers, that is,

the ratio of the average number of busy customers $(N - \overline{\gamma})$ to N. The calculation results showed that these parameters as a function of μ/λ almost do not depend on N. Relevant graphs are presented in Figures 3 and 4.
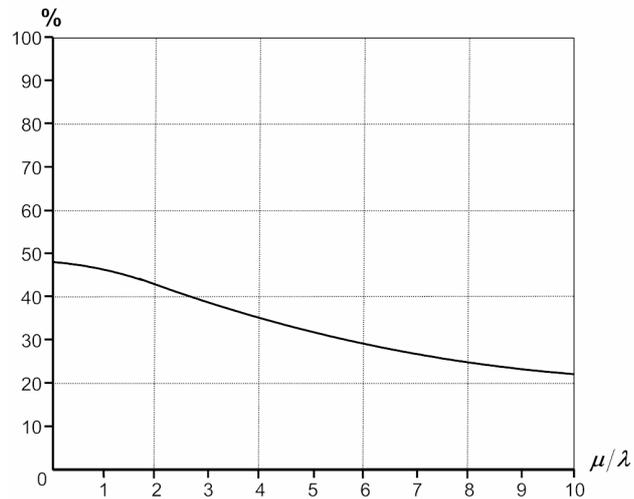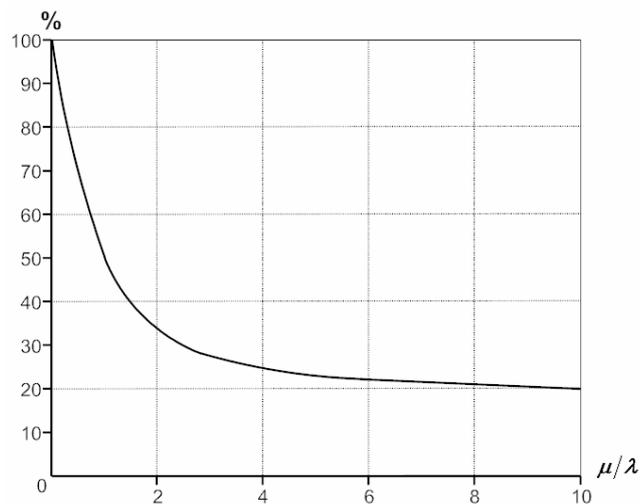
Figure 3. Relative system's load.

Figure 4. Relative percentage of busy customers.

3. Prohibition of Queues

The state of the system in this case, in contrast to the situation described in paragraph 2, is completely determined by the number of ongoing communications ω. We denote by $P_t(\omega)$ the probability that the system is in state ω at time t. Then the following relations are valid:

$$P_{t+\Delta t}(\omega) = \frac{N - 2(\omega-1))(N - 2(\omega-1)-1)\lambda}{N-1} P_t(\omega-1)\Delta t +$$

$$+ P_t(\omega)(1 - \lambda(N-2\omega)\frac{N-2\omega-1}{N-1}\Delta t - \mu\omega\Delta t) + P_t(\omega-1)\Delta t + o(\Delta t),$$

$$\omega = 0,1,...,\left[\frac{N}{2}\right], P_t(-1) = P_t\left(\left[\frac{N}{2}\right]+1\right) = 0.$$

To calculate the stationary probabilities $\pi(\omega) = \lim\limits_{t\to\infty} P_t(\omega)$, we have a system of linear algebraic equations:

$$\begin{cases} \pi(\omega-1)\dfrac{N - 2(\omega-1))(N - 2(\omega-1)-1)\lambda}{N-1} - \pi(\omega)(\lambda(N-2\omega)\dfrac{N-2\omega-1}{N-1} + \mu\omega) + \\ + \pi(\omega+1)\mu(\omega+1) = 0 \qquad (\omega = 0,1,...,[N/2]) \\ \sum\limits_{k=0}^{[N/2]} \pi(k) = 1 \end{cases} \qquad (4)$$

In contrast to the system of equations (2), the matrix of the system of equations (4) is tridiagonal, therefore, calculations on a computer of stationary probabilities can be carried out for sufficiently large *N*.

Calculations on modern computer systems based on the classical parallel Thomas algorithm [12], allowed us to get results up to *N* = 10000. As well as for a system without restrictions on the length of the queues, the stationary probabilities $\pi(\omega)$ do not change with a proportional increase in the parameters μ and λ. Relative system load (this is, as already noted, the ratio of the average number of stationary communications to the [*N*/2] – maximum possible number of communications) almost doesn't depend on *N*. This graph is presented in Figure 5.
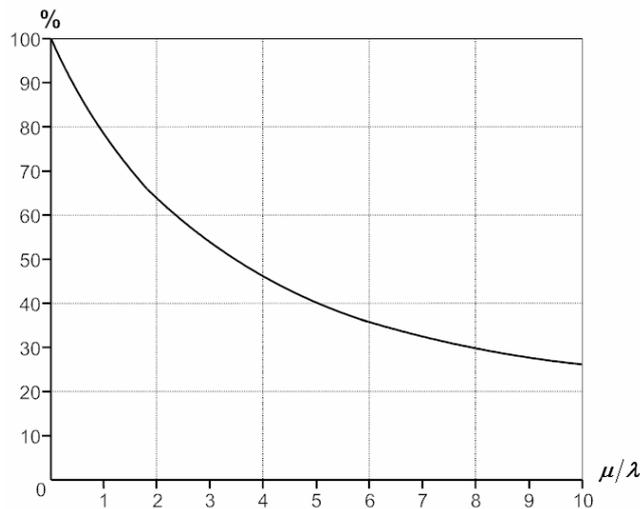

Figure 5. Relative system's load.

A comparison of the two systems described in this paper shows that the operational efficiency (the number of communications per time unit) of the system where queues are prohibited (for the same

values of N and μ/λ) is higher than the efficiency of the system with queues for $\mu/\lambda \leq 1$, that is, when the communication sessions are long enough (the intensity λ is considered to be fixed). At large ratio μ/λ ($\mu/\lambda \rightarrow \infty$), the efficiency of both systems naturally drops to zero, and at moderate ratio μ/λ, the systems with queues are somewhat more efficient.

4. Conclusion
As shown in this paper, with a certain correlation of the parameters of the system, queues are more efficient than systems in which queues are prohibited. Important factor here is average communication session length. If it is known in advance, then this information can be used to configure the parameters of the telecommunication system for its more efficient functioning. It should be noted that the systems discussed here can be proposed as a model of the functioning of a usual telephone network.

An interesting direction for further research is to take into account information (for example, statistics of past periods) about the average time of a communication session of each individual customer. It is also interesting to evaluate the effectiveness of using the "combined scheme": queues are allowed for some customers and forbidden for others.

6. References
[1]   Shannon C, Weaver W 1971 *The Mathematical Theory of Communication* (Illinois: The University of Illinois Press) p 144
[2]   Lakatos L, Szeidl L, Telek M 2013 *Introduction to Queueing Systems with Telecommunication Applications* (Springer) p 388
[3]   Cooper R 1972 *Introduction To Queueing Theory* (London: Macmillan) p 347
[4]   Kirichuk V, Mokin K, Reznik A 2001 Algorithms for processing of a series of digital aerospace images based on automatic search for the conjugate points *Pattern Recognition and Image Analysis (Advances in Mathematical Theory and Applications)* **11** 192
[5]   Efimov V, Kirichuk V, Kolesnikov A, Reznik A 2000 Algorithms for quickly reconstructing the earth's surface using parallel processing of several aerospace images. *Pattern Recognition and Image Analysis (Advances in Mathematical Theory and Applications)* **10** 259
[6]   Zukerman M 2019 Introduction to Queueing Theory and Stochastic Teletraffic Models *Preprint* ArXiv: 1307.2968
[7]   Berry R 2006 *Queuing Theory* (Whitman College press)
[8]   Tawde P 2019 Simulation Tools for Designing Electronics & Telecommunication Devices *Journal of Electronics and Telecommunication Engineering* **4** 1
[9]   Kirkwood J 2015 *Markov Processes* (CRC Press) p 340
[10]   Feller W 1968 *An Introduction to Probability Theory and Its Applications 3rd Edition* (New-York: Wiley) p 509
[11]   Hardy G and Ramanujan S 1918 Asymptotic Formulae in Combinatory Analysis *Proceedings of the London Mathematical Society* **8** 75
[12]   Gallopoulos E, Philippe B, Sameh A 2016 *Parallelism in Matrix Computations* (Springer Netherlands) p 473