REPUBLIC OF AZERBAIJAN

On the rights of the manuscript

ABSTRACT

of the dissertation for the degree of Doctor of Philosophy

DEVELOPMENT OF MATHEMATICAL MODELS AND METHODS FOR ANALYSIS OF SERVICE-STORAGE SYSTEMS WITH DIFFERENT TYPES OF REQUESTS

Specialty:	3338.01 – "System Analysis, Management and Information Processing"
Field of science:	Technical sciences
Applicant:	**Ismayil Alakbar Aliyev**

BAKU – 2022

The work was performed at the department of "Information Technologies and programming" of Baku State University.

Scientific adviser:	Corresponding Member of ANAS, Doctor of Engineering, Professor **Aghasi Zarbali Melikov**
	Doctor of Engineering, Professor **Leonid Anatolyevich Ponomarenko**
Official opponents:	Doctor of Engineering, Professor **Ramin Rza Rzayev**
	Doctor of Engineering, Assoc. Professor **Latafat Abbas Gardashova**
	PhD in Engineering, Assoc. Professor **Farhad Firudin Yusifov**

Dissertation council ED 1.20 of Supreme Attestation Commission under the President of the Republic of Azerbaijan operating at İnstitute of Control Systems of ANAS

Chairman of the Dissertation council: ______________	Academician, Doctor of Engineering, Professor **Ali Mahammad Abbasov**
Scientific secretary of the Dissertation council: ______________	Doctor of Engineering, Professor **Naila Fuad Musaeva**
Chairman of the scientific seminar: ______________	Doctor of Engineering, Assoc. Professor **Fahrad Haydar Pashayev**

GENERAL DESCRIPTION OF WORK

The relevance of the problem and the degree of its development. In the last three decades, queuing-inventory systems (QIS) have been intensively studied, in which the service time of requests is a positive random variable. In such systems, the formation of a queue of incoming requests is possible, and in them the level of stocks decreases only after the completion of their service. In addition to them, systems are also intensively studied in which the service time of requests is equal to zero, i.e. classical inventory management systems (IMS). In contrast to QIS, in IMS the stock level of the system is reduced instantly at the time of receipt of the request.

An analysis of the existing literature has shown that in the vast majority of studies, QIS and IMS models are studied, in which applications are identical in all respects - in terms of the volume of reserves they require, in terms of service time, in importance, etc. However, in real life, suppliers of goods distinguish their customers. So, for example, clients can have different sizes; they can be constant and episodic; some customers may pay more for the same product than other customers, etc. In such cases, suppliers of goods to encourage profitable customers use various schemes for their priority service. Despite the fact that QIS and IMS with different types of applications often occur in real life, they are little studied in the available literature. Based on these facts, in this dissertation work, mathematical models of QIS and IMS with different types of applications are developed and studied. These circumstances explain the relevance of the topic of this work.

The main goal and objectives of the study. The purpose of the work is to develop and study mathematical models of QIS and IMS with different types of applications. Based on this goal, the following tasks were solved in this work:

- development of mathematical models of IMS and QIS with different types of applications;

- development of methods and algorithms for calculating and optimizing the characteristics of QIS and IMS with different types of applications.

Object and subject of research. The object of the dissertation work are queuing-inventory systems with different types of requests. The subject of the dissertation work is the creation of methods for finding and improving the characteristics of the objects under study.

Research methods. To achieve this goal, the methods of inventory management theory, the theory of queuing systems, graph theory, the theory of multidimensional Markov chains and mathematical modeling and numerical optimization methods were used.

Defense provisions. The author defends the following positions.

1. Mathematical models of IMS with two types of requests and without their repetition when using deterministic and randomized restocking policies.

2. New priority service schemes in IMS with repetition of low priority requests, which are based on the values of the current level of stocks of the system at the time of receipt of requests from out of orbit and from orbit.

3. Mathematical models of QIS with two types of requests in the presence of a finite and infinite queue, where the priorities of requests are determined either depending on the length of the queue of different types of requests or on the current inventory level of the system.

4. Methods and algorithms for calculating and optimizing the quality-of-service indicators of the studied IMS and QIS models when using various restocking policies.

Scientific novelty. The main results of this work, which have scientific novelty, are as follows:

1. Mathematical models of IMS with two types of requests and without repetition of requests were developed using four deterministic and one randomized inventory restocking policy (IRP). The scheme of priority servicing of different types of requests is determined, which is based on the value of the current stock level of the system at the time of receipt of requests. Exact formulas are proposed for finding the characteristics of the studied IMS when

using each IRP.

2. A new scheme of priority servicing of applications in the IMS with two types of applications and repetition of low priority applications using different IRP is proposed. According to this scheme, low priority claims according to the Bernoulli scheme can either go into orbit or leave the system if, at the time of their arrival, the stock level of the system is below a certain critical level. In this case, if at the moment of arrival of the request from the orbit, the level of reserves is again below the critical level, then the repeated request, also according to the Bernoulli scheme, either finally leaves the orbit, or remains there for repetition. It is shown that the corresponding two-dimensional Markov chains (MC) are the mathematical models of the IMS under study, and the conditions for the ergodicity of these chains are established. Formulas are found for exact and approximate calculation of the characteristics of the systems under study.

3. Two schemes are proposed for priority servicing of applications in the QIS with a finite or infinite queue, in which some of the applications do not receive reserves after the completion of the service. In both schemes, high priority applications are accepted if there is at least one free place in the queue. However, in one scheme, low priority claims are only accepted when the total number of claims in the system is below a certain threshold. In another scheme, low priority orders are only accepted when the stock level of the system is above a certain critical value.

4. Two-dimensional Markov chains are constructed that describe the operation of the studied QIS using both access schemes and various deterministic IRP. Conditions for the ergodicity of these systems in the presence of a queue of different types of requests of infinite length are obtained. A unified approach is proposed for studying QIS with finite and infinite queues. Exact and approximate formulas for calculating their characteristics are found and problems of maximizing their profits are solved.

Scientific and practical value. The proposed mathematical models more adequately describe the operation of a wide class of systems for providing material resources, including complex logistics networks. The obtained formulas and the developed package of applied programs make it possible to calculate and improve the characteristics of these systems.

Approbation of work and implementation of work. The main scientific and practical results of the work were reported and discussed:

- at the XXII Republican scientific conference of doctoral students and young researchers, Baku, Azerbaijan, November, 2018;
- at the XVII and XVIII International Conferences "Information Technologies and Mathematical Modeling", Tomsk, Russia, September, 2018 and June, 2019, respectively;
- at the XXV-th International Open Science Conference «Modern Informatization Problems in Economics and Safety», Yelm, WA, USA, January 2020;
- at the XII International Scientific and Practical Conference "Internet, Education, Science" (ION-2020), Vinnitsa, Ukraine, May, 2020;
- at the seminars of the Institute of Control Systems of the National Academy of Sciences of Azerbaijan and at the seminars of the department "Information Technologies and Programming" of the Faculty of Applied Mathematics and Cybernetics of Baku State University.

The results of the dissertation work were used and implemented in the existing system of the company OFFICE LINES.

The name of the institution where the work was performed. The dissertation work was carried out at the Department of Information Technologies and Programming of the Faculty of Applied Mathematics and Cybernetics of the Baku State University.

Number of published scientific articles. According to the results of the research, a total of 12 papers were published: 7 scientific articles, including 6 abroad, of which 3 articles with international scientific citation indices from the Web-Science and Scopus archives; 5 abstracts of reports at conferences.

Structure and scope of work. The dissertation work consists of an introduction, four chapters, a conclusion, a list of references and an appendix. The work contains 165 pages, including 26 figures, 9 tables and 107 references. The volume of the general and structural sections of the dissertation is distributed approximately as follows: Total - 170,000 characters, Table of contents – 2,000 characters, Introduction – 14,000 characters, Chapter one – 34,500 characters, Chapter Two - 31,500 characters, Chapter Three – 26,000 characters Chapter Four - 58,000 characters, Conclusion – 4,000 characters

MAIN CONTENT OF THE WORK

The introduction shows the relevance of the ongoing research, formulates the goals and objectives of the dissertation, lists the main provisions submitted for defense, lists scientific novelties, indicates the practical significance of the work, and provides a list of scientific forums where the results of this work were reported. It also provides a summary of each chapter of the dissertation.

The first chapter indicates that in the classical theory of a queuing system it is assumed that the system has a certain number of stationary channels (servers, devices, lines, etc.) that are designed to perform the process of servicing requests arriving at random times.

In works on the theory of the queuing system, the fundamental assumption is that queuing systems, in addition to stationary channels, also have unlimited reserves, so the loss of customers is associated only with the lack of a free channel and / or free space in the buffer to wait for claims, as well as with impatience applications in the queue. At the same time, in a number of real queuing systems, to service an incoming request, it is not enough just to have a free channel, but it also requires the presence of certain resources (reserves) of the system, since in them servicing an application is associated with the release (sale) of certain things to the incoming request. It is important to note that there are systems in which inventory levels are also reduced as a result of spoilage. Such systems are called systems with perishable stocks. However, here we will consider only systems with non-perishable reserves, i.e. we will assume that the lifetime of stocks is infinite.

The classical theory of the queuing system was developed in the fundamental works of A.N. Kolmogorova, A.Ya. Khinchin, B.V.

Gnedenko, Palm, Cox and a number of other scientists. Subsequently, this theory was rapidly developed in the works of G.P. Basharina, V.S. Korolyuk, I.N. Kovalenko, B.V. Sevastyanov, L. Kleinrock, K. Ross and several other famous scientists. In recent years, a significant contribution to the theory of the queuing system has been made by V.M. Vishnevsky, V.V. Rykov, A.M. Gortsev, A.A. Nazarov, A.N. Dudin, K.E. Samuylov and their numerous students.

Queuing systems, in which, in addition to a free server, certain reserves are also required to service an application are called queuing-stock systems. These systems can also be considered from the point of view of classical inventory management systems, which do not consider the possibility of forming a queue of applications even if the required amount of system inventory is available, i.e. they assume that the system has an infinite number of channels for the release of the required inventory. In other words, it is considered that the service time of applications is equal to zero. Therefore, the IMS can be represented as a QIS with instant service of applications, i.e. in IMS, the stock level of the system decreases at the time of receipt of requests.

This chapter outlines the main specific features of IMS and QIS. It is noted that the IMS theory has a rather deep history, while the QIS theory was created almost thirty years ago. It is indicated that the founders of the theory of QIS, along with the American scientists Sigman and Simch-Levi, are the Azerbaijani scientist A.Z. Melikov. Here is a detailed analysis of the review of works devoted to the study of models of IMS and QIS. It is noted that in well-known works, models of systems with identical applications are mainly studied. The importance of studying IMS and QIS models with different types of applications is noted. At the same time, an increase in the complexity of studying the models of IMS and QIS with different types of applications is indicated, since considering the fact of heterogeneous applications leads to the need to study models of multidimensional Markov chains. Certain classifications of these models are carried out and the main characteristics of the systems under study are indicated. Three classes of models with two types of requests are distinguished, while two classes of systems belong to the IMS type, and the third class belongs to the QIS. The first class is the IMS without repeated applications; the second class is the IMS with repeated applications and the third class is the IMS, where a queue of incoming applications is possible. Block diagrams of these systems are given, and their main characteristics are determined. It also provides a detailed description of the stages of constructing mathematical models of the systems under study.

As in the theory of the queuing system, here it is also possible to introduce certain classifications of QIS and IMS. These classifications can be based on various factors. Among them, the most important are the following factors: lifetime of system inventory; types of distribution functions for incoming flows and the time of their service; application types; the behaviour of applications in the absence of stocks of the system; the behaviour of applications after the completion of service; restocking policy; types of distribution functions of the delay time in the fulfillment of ordered stocks; types of cost function and income.

The second chapter examines IMS models with two types of claims. They can use different IRP.

Here we study IMS models in which Poisson flows of requests of two types are serviced: the intensity of ordinary requests (the flow of requests of the first type) is equal to λ_1, and the intensity of priority claims (the flow of claims of the second type) is equal to λ_2. In all models, the following scheme for servicing heterogeneous requests is adopted.

- If the inventory level is greater than zero at the moment the priority order arrives, it receives the inventory and leaves the system.
- If at the time of receipt of a regular order, the level of stocks is greater than the critical level $s, s < S/2$, where S stands for the system's maximum stock, then it also gets stock and leaves the system.
- If the stock level of the system at the time of receipt of a regular order is less than or equal to s, this order, according to the Bernoulli scheme, either receives a stock with a probability α and leaves the system, or with an additional probability $1-\alpha$ leaves the

system without receiving a stock. One of the following four types of deterministic IRP can be used in the system:

- Two-level policy, i.e. (s,S)-policy. This means that an order for the supply of stocks is made when their level drops to the value $s, s < S/2$, while the order volume is a constant value and equals $S-s$;
- Variable Order Size (VOS) policy an order for the supply of stocks is made when their level drops to the value $s, s < S/2$, while the volume of the order is a variable value and equals $S-m$, where m the level of stocks at the time of dispatch of stocks;
- $(S-1,S)$-policy, i.e. an order (of a single size) for the supply of stocks is made after each stock issue act;
- "Up to *S*" policy, i.e. if the inventory level in the system is equal to zero, then an order is made to fill the warehouse to the maximum level.

In all types of IRP, lead times for replenishment orders are random variables that have a common exponential distribution function with the parameter $\nu > 0$.

It is shown that the mathematical models of the systems under study using various IRP are one-dimensional MC. The states of the indicated MC at an arbitrary point in time are described by a scalar value *m*, which indicates the current level of the system's reserves, *m*=0, 1, …, *S*.

Explicit formulas for calculating the characteristics of the system were found in all models: average inventory level (S_{av}); average intensity of orders (RR); probabilities of denial of service for requests of each type when entering the system (PB_1, PB_2).

So, to calculate the characteristics of the system when using (s,S)-inventory restocking policy the following formulas were obtained:

$$S_{av} = \sum_{m=1}^{S} mp(m); \tag{1}$$

$$RR = \lambda p(s+1); \tag{2}$$

$$PB_1 = (1-\alpha)\sum_{m=0}^{s} p(m);\qquad (3)$$

$$PB_2 = p(0).\qquad (4)$$

In (1) - (4) the values $p(m)$ indicate the probabilities that the stock level is equal $m, m = 0,1,\ldots,S$, and are defined as follows:

$$p(m) = \begin{cases} a_m p(s+1), & 0 \le m \le s, \\ p(s+1), & s+1 \le m \le S-s, \\ b_m p(s+1), & S-s+1 \le m \le S, \end{cases}$$

where $p(s+1) = \left(\sum_{m=0}^{s} a_m + S - s + \sum_{m=S-s+1}^{S} b_m \right)^{-1}$;

$a_m = \prod_{i=m+1}^{s+1} \frac{x_i}{\nu + x_{i-1}}$, $x_0 = 0$; $x_m = \begin{cases} \lambda_1\alpha + \lambda_2, & 1 \le m \le s, \\ \lambda, & s+1 \le m \le S; \end{cases}$

$b_m = \frac{\nu}{\lambda} \sum_{i=m-S+s}^{s} a_i$.

Here, the IMS model was also studied using a randomized IRP, which is defined as follows. The service of requests continues until the system warehouse is empty, and when the system warehouse is empty, with probability $\sigma(m)$ volume stocks are ordered $m, m = 1,2,\ldots,S$, where $\sum_{m=1}^{S} \sigma(m) = 1$, $\sigma(S) > 0$.

For this model, the state probabilities are calculated as follows:

$$p(m) = \begin{cases} \frac{\nu}{\tilde{\lambda}} \left(1 - \sum_{i=1}^{m-1} \sigma(i) \right) p(0), & 1 \le m \le s, \\ \frac{\nu}{\lambda} \left(1 - \sum_{i=1}^{m-1} \sigma(i) \right) p(0), & s+1 \le m \le S, \end{cases}$$

where $p(0) = \left(1 + \nu \left(\frac{1}{\tilde{\lambda}} \sum_{m=1}^{s} \left(1 - \sum_{i=1}^{m-1} \sigma(i) \right) + \frac{1}{\lambda} \sum_{m=s+1}^{S} \left(1 - \sum_{i=1}^{m-1} \sigma(i) \right) \right) \right)^{-1}$.

The desired characteristics of this model are calculated as follows:

$$S_{av} = \sum_{m=1}^{S} mp(m) = \nu p(0)\left(\frac{1}{\tilde{\lambda}} \sum_{m=1}^{s} \left(1 - \sum_{i=1}^{m-1} \sigma(i)\right) + \frac{1}{\lambda} \sum_{m=s+1}^{S} \left(1 - \sum_{i=1}^{m-1} \sigma(i)\right)\right);$$

$$RR = \tilde{\lambda} p(1) = \nu p(0);$$

$$PB_1 = (1-\alpha)\sum_{m=0}^{s} p(m) = (1-\alpha)\left(1 + \frac{\nu}{\tilde{\lambda}} \sum_{m=1}^{s} \left(1 - \sum_{i=1}^{m-1} \sigma(i)\right)\right) p(0);$$

$$PB_2 = p(0).$$

The paper presents the results of numerical experiments on the calculation of models. Some results of numerical experiments for the system using (s,S)-inventory restocking policy are shown in the Fig.1; it is assumed here $S = 120, \lambda_1 = 50, \lambda_2 = 200, \nu = 20, \alpha = 0.3$.

This chapter also solves the problems of maximizing the system profit when using various restocking policies: for each restocking policy, it is required to maximize the profit of the system $(PT(s))$ by choosing the appropriate values of the critical stock level s, i.e. you need to solve the following problem:

$$s^* = \arg\max_{s} PT(s), \tag{5}$$

where $PT(s) = RV(s) - TC(s)$. System revenue (*RV*) is generated from the sale of inventory, i.e.

$$RV(s) = (\lambda_1 (1 - PB_1(s)))C_{rev}^1 + \lambda_2 (1 - PB_2(s))C_{rev}^2 ,$$

where C_{rev}^k – system income due to the sale of a unit of stock for orders of the *k* type, $k = 1, 2$.

Total fines (*TC*) in the system are defined as follows:

$$TC(s) = (K + c_r V_{av}(s))RR(s) + c_h S_{av}(s) + c_l^1 \lambda_1 PB_1(s) + c_l^2 \lambda_2 PB_2(s),$$

where K – is the fixed price of one order; c_r – price per unit of order volume; c_h – the price of storing a unit of inventory volume per unit of time; c_l^k – penalty for the loss of one application of the k type.

Note that problem (5) always has a solution, since the set of feasible solutions $\{0 \le s \le \lceil S/2 \rceil\}$ is finite and discrete. This chapter presents the results of solving problem (5) for each IRP.

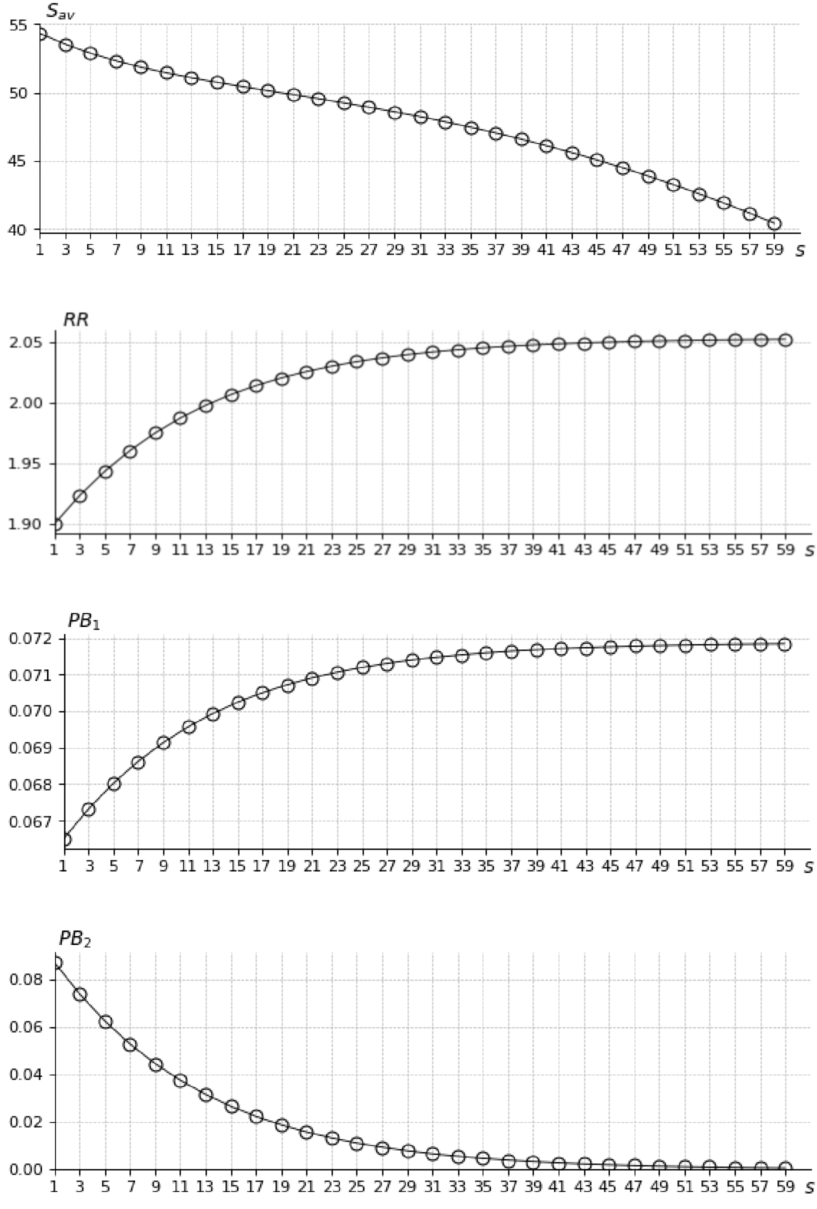

Fig. 1. Dependence of the characteristics of the system on the parameter s when using (s, S)- inventory restocking policy.

The third chapter proposes priority IMS models with two types of applications and repeated low-priority applications. Priority servicing of applications is carried out according to the following scheme: if at the moment of receipt of a high-priority application the stock level is greater than zero, then it receives stock and leaves the system; if at the moment of receipt of a low-priority application the stock level is higher than the critical level $s, s < S/2$, then it also receives a stock and leaves the system; otherwise, i.e. if the stock level of the system at the time of receipt of a low-priority request is less than or equal to s, then this request, according to the Bernoulli scheme, either goes into orbit with a probability α in order to repeat the request to receive a stock, or with an additional probability does not receive a stock and leaves the system. Here models with finite and infinite orbit sizes are studied. For a model with an infinite orbit size, any claim of the first type can be accepted into the orbit. In the case of a finite size of the orbit, the first type notice that is sent to the orbit can be accepted into the orbit if there is at least one free seat there; otherwise, it is lost with probability one.

Requests from the orbit repeat requests independently of each other at random times that have a common exponential distribution with the parameter $\eta, 0 < \eta < \infty$. At the same time, if at the time of receipt of a repeated request the level of reserves is greater than the critical level s, then it instantly receives the required reserve and leaves the orbit; otherwise (i.e., if the stock level of the system at this moment is less than or equal to s) according to the Bernoulli scheme, it either leaves the orbit with probability β or remains in orbit with an additional probability.

It is believed that orders for replenishment of stocks are fulfilled with certain random delays, which have a common exponential distribution with the parameter $\nu > 0$.

One of the following restocking policies can be used in the system: (1) (s,S)-policy, (2) VOS- policy и (3) $(S-1,S)$- policy.

Here the task is to find the joint distribution of the stock level of the system and the number of requests in orbit.

It is shown that the operation of a system with a finite size (R) of the orbit is described by a two-dimensional Markov chain with states of the form (m,n), where m – is the system inventory level, n – is the number of customers in the orbit. The state space is defined as the Cartesian product of two sets, i.e. $E=\{0,1,\dots,S\}\times\{0,1,\dots,R\}$. A system of equilibrium equations (SEE) has been developed for the probabilities of states $p(m,n)$ using the (s, S)-policy:

Case $s<m\le S$:

$$(\lambda+n\eta)p(m,n)=\lambda p(m+1,n)+(n+1)\eta p(m+1,n+1)+\nu p(m-S+s,n).$$

Case $0\le m\le s$:

$$(\lambda_2 I(m>0)+\lambda_1\alpha+n\eta\beta+\nu)p(m,n)=\lambda p(s+1,n)\delta(m,s)+$$
$$+\lambda_2 p(m+1,n)I(m<s)+\lambda_1\alpha p(m,n-1)I(n>0)+(n+1)\eta\beta p(m,n+1).$$

$$\sum_{(m,n)\in E}p(m,n)=1.$$

To solve the obtained SEE, which is a system of linear algebraic equations (SLAE) of dimension $(S+1)(R+1)$, existing application packages (AP) can be used. Next, the characteristics of the system are calculated:

average inventory level

$$S_{av}=\sum_{m=1}^{S}m\sum_{n=0}^{R}p(m,n);$$

average order intensity

$$RR=\lambda\sum_{n=0}^{R}p(s+1,n);$$

the probability of denial of service for requests of each type

$$PB_1=(1-\alpha)\sum_{m=0}^{s}\sum_{n=0}^{R-1}p(m,n)+\sum_{m=0}^{s}p(m,R);$$

$$PB_2=\sum_{n=0}^{R}p(0,n);$$

average number of applications in orbit

$$L_o=\sum_{m=0}^{S}\sum_{n=1}^{R}np(m,n);$$

average rate of successful repetition of requests from orbit

$$RSR = \eta \sum_{m=s+1}^{S} \sum_{n=1}^{R} np(m,n);$$

average intensity of unsuccessful repetition $(RuSR)$ of requests from orbit

$$RuSR = \eta\beta \sum_{m=0}^{s} \sum_{n=1}^{R} np(m,n).$$

A model with an infinite orbit size is also studied here. It is shown that the mathematical model of this system is a two-dimensional MC with an infinite state space. To calculate the approximate values of the probabilities of the states of the resulting circuit, the method of phase enlargement of the states of two-dimensional MCs developed by A.Z. Melikov. Next, the approximate values of the system characteristics are calculated in a standard way:

$$S_{av} = \sum_{m=1}^{S} m\rho(m);$$

$$RR = \lambda\rho(s+1);$$

$$PB_1 = (1-\alpha)\sum_{m=0}^{s} \rho(m);$$

$$PB_2 = \rho(0);$$

$$L_o = \Psi;$$

$$RSR = \eta \Psi \sum_{m=s+1}^{S} \rho(m);$$

$$RuSR = \eta\beta \Psi \sum_{m=0}^{s} \rho(m).$$

In the last formulas, the parameters are defined as follows:

$$\rho(m) = \begin{cases} a_m \rho(s+1), & \text{если} \quad 0 \le m \le s, \\ \rho(s+1), & \text{если} \quad s+1 \le m \le S-s, \\ b_m \rho(s+1), & \text{если} \quad S-s+1 \le m \le S, \end{cases}$$

where $\rho(s+1) = \left(S - 2s + \sum_{m=0}^{s} a_m + \sum_{m=S-s+1}^{S} b_m \right)^{-1};$

$$a_m = \left(\frac{\lambda_2}{\nu + \lambda_2} \right)^{s+1-m}; \; b_m = \frac{\nu}{\lambda} \sum_{i=m-S+s}^{s} \left(\frac{\lambda_2}{\nu + \lambda_2} \right)^{s+1-i};$$

$\Psi = \Lambda / \mathrm{M}$, where $\Lambda = \lambda_1 \alpha \sum_{m=0}^{s} \rho(m)$; $\mathrm{M} = \eta \left(\sum_{m=s+1}^{S} \rho(m) + \beta \sum_{m=0}^{s} \rho(m) \right)$.

Models are studied in a similar way for other IRP. Numerical experiments were carried out to calculate the models and the problems of maximizing the profit of the systems under study were solved using various restocking policies.

The fourth chapter proposes two-dimensional Markov models of QIS with two types of requests, which use different IRP. Two schemes for organizing priority servicing of different types of requests are considered. In both schemes, there are no restrictions on the access of high priority claims. However, in the first scheme, low priority claims are accepted only when the total number of claims in the system is less than a given threshold value. In the second scheme, low priority orders are accepted only when the current stock level is above a certain critical value.

Unlike the IMS models that have been studied in the previous chapters, this considers the fact that after the completion of service, some customers may not buy inventory due to various reasons, for example, due to the poor quality of inventory, because of its high cost, etc. To consider such situations, it is considered here that after the completion of the service of an application of any type, it with a probability $\sigma_1 > 0$ does not receive a stock and with an additional probability $\sigma_2 = 1 - \sigma_1$ receives a stock. The service time for both types of requests depends on whether she received the stock or not; in both cases, this value has an exponential distribution function (d.f.), and if the application did not receive the stock, then the average value of the service time is equal to μ_1^{-1}, otherwise it is equal to μ_2^{-1} .

Here, models with a finite and infinite common queue of heterogeneous requests are studied. In the case of a final queue, requests of both types are waiting in the queue of maximum length $N, N < \infty$. The first priority service scheme is defined as follows: if at the time of receipt of a high-priority application the inventory level is positive, then it is accepted into the system, while this application is accepted for service if the queue is empty; otherwise, it is accepted into the queue if there is at least one free seat. However, a low-priority request is accepted only when, at the time of its arrival, the total queue length is less than the specified threshold value $r, 1 \le r \le N-1$. In the model with an infinite queue, customers of the second type are not lost, and in the model with a finite queue, these customers are lost only when the queue is full.

Here it is considered that if at the moment of receipt of a request of any type of the stock level is equal to zero, then it joins the queue with probability φ_1, and with an additional probability $\varphi_2 = 1 - \varphi_1$ it leaves the system unserved. In addition, impatience of applications is taken into account here. This means that if, while they are waiting in the queue, the inventory level of the system drops to zero, i.e. if, in the presence of a queue, the inventory level of the system drops to zero, then the order of each type leaves the system after a random time, which has an exponential d.f. with parameter $\tau > 0$.

One of two IRP can be used in the system (*s*, *S*)-policy or VOS-policy.

First, we consider a model with a finite common queue of dimension *N*. When both IRP are used, the operation of a system with a common finite queue is described by a two-dimensional MC with states of the form (m, n), where m – is the system stock level, n – is the total number of requests in the system. The state space is defined as $E = \{0,1,...,S\} \times \{0,1,...,N\}$. SEE was developed for state probabilities $p(m,n)$ when using the (*s*, *S*)-policy:

Case $m = 0$:

$$(\lambda_1 \varphi_1 I(n < N) + n\tau) p(0,n) = \lambda_1 \varphi_1 p(0,n-1) I(n > 0) +$$

$$(n+1)\tau p(0,n+1) + \mu\sigma_2 p(1,n+1) I(n < N).$$

Case $0 < m < s$:

$$\left(\lambda I(n<r)+\lambda_2 I(n\ge r)+\mu+\nu\right)p(m,n)=\lambda p(m,n-1)I(0<n<r)+$$

$$+\lambda_2 p(m,n-1)I(n\ge r)+\mu\sigma_1 p(m,n+1)I(n<N)+$$

$$+\mu\sigma_2 p(m+1,n+1)I(n<N).$$

Case $m \ge s$:

$$\left(\lambda I(n<r)+\lambda_2 I(n\ge r)+\mu\right)p(m,n)=\lambda p(m,n-1)I(0<n<r)+$$

$$\lambda_2 p(m,n-1)I(n\ge r)+\mu\sigma_1 p(m,n+1)I(n<N)+$$

$$+\mu\sigma_2 p(m+1,n+1)I(n<N)+\nu p(m-S+s,n).$$

The normalization condition is added to them:

$$\sum_{(m,n)\in E} p(m,n)=1.$$

The resulting SEE is a SLAE of dimension $(S+1)(N+1)$, and known AP can be used to solve this system. It is shown that when using both IRP, the following relations take place:

$$S_{av}=\sum_{m=1}^{S} m\sum_{n=0}^{N} p(m,n); \tag{6}$$

$$RR=\mu_2\sigma_2\sum_{n=1}^{N} p(s+1,n); \tag{7}$$

$$PB_1=\sum_{m=1}^{S}\sum_{n=r}^{N} p(m,n)+\theta_1\sum_{n=1}^{N} p(0,n)\frac{n\tau}{\lambda\varphi_1+n\tau}; \tag{8}$$

$$PB_2=\sum_{m=0}^{S} p(m,N)+\theta_2\sum_{n=1}^{N} p(0,n)\frac{n\tau}{\lambda\varphi_1+n\tau}. \tag{9}$$

In the last formulas, the parameters $\theta_1=\eta_1/(\eta_1+\eta_2)$ and $\theta_2=1-\theta_1$ estimate the probability that, in the type states $(m,n), n>0$, a randomly selected request is a request of the first and second types, respectively, where η_1 and η_2 denote the intensity of requests of the first and second types accepted into the system, respectively. It is shown that $\eta_2 = \lambda_2$ and

$$\eta_1 = \sum_{k=1}^{r} k \cdot \frac{\lambda_1^k}{k!} e^{-\lambda_1} + r \sum_{k=r+1}^{\infty} \frac{\lambda_1^k}{k!} e^{-\lambda_1} = e^{-\lambda_1} \sum_{k=1}^{r} \frac{\lambda_1^k}{(k-1)!} + r\left(1 - e^{-\lambda_1} \sum_{k=0}^{r} \frac{\lambda_1^k}{k!}\right).$$

To study the model with an infinite queue, the method of spectral expansion, as well as the method of phase enlargement, are used.

It is shown that the condition for the ergodicity of a model with an infinite queue is the fulfillment of the following condition: $\lambda_2 < \mu_1 \sigma_1$. When this condition is met, the characteristics of the system are calculated as follows:

$$S_{av} \approx \sum_{m=1}^{S} m \pi(< m >);$$

$$RR \approx \mu_2 \sigma_2 \pi(< s+1 >)(1 - \rho(0));$$

$$PB_1 \approx (1 - \pi(< 0 >))\left(1 - \sum_{n=0}^{r-1} \rho(n)\right) + \theta_1 \pi(< 0 >) \sum_{n=1}^{\infty} \rho_0(n) \frac{n\tau}{\lambda \varphi_1 + n\tau};$$

$$PB_2 \approx \theta_2 \pi(< 0 >) \sum_{n=1}^{\infty} \rho_0(n) \frac{n\tau}{\lambda \varphi_1 + n\tau}.$$

The following notation is adopted in the last formulas:

$$\rho(0) = \left(\sum_{n=0}^{r} \left(\frac{\lambda}{\mu_1 \sigma_1} \right)^n + \left(\frac{\lambda}{\mu_1 \sigma_1} \right)^r \frac{\lambda_2}{\mu_1 \sigma_1 - \lambda_2} \right)^{-1};$$

$$\rho_0(n) = \frac{a^n}{n!} e^{-a}, \ n = 0,1,....;$$

$$\pi(< m >) = \begin{cases} \alpha_m \pi(< s+1 >), & \text{если} \quad 0 \le m \le s, \\ \pi(< s+1 >), & \text{если} \quad s+1 \le m \le S - s, \\ \beta_m \pi(< s+1 >), & \text{если} \quad S - s + 1 \le m \le S, \end{cases}$$

where $\pi(< s+1 >) = \left(S - 2s + \sum_{m=0}^{s} \alpha_m + \sum_{m=S-s+1}^{S} \beta_m \right)^{-1};$

$$\alpha_m = \left(\frac{\mu_2 \sigma_2 (1-\rho(0))}{\nu + \mu_2 \sigma_2 (1-\rho(0))} \right)^{s-m+1} ; \ \beta_m = \frac{\nu}{\mu_2 \sigma_2 (1-\rho(0))} \sum_{i=m-S+s}^{s} \alpha_i .$$

Huge computational experiments have been carried out to calculate and optimize this model. These experiments have shown that the accuracy of the developed approximate algorithms for calculating the probabilities of states and characteristics of the system turns out to be quite high, while the approximate algorithms significantly exceed their exact counterparts in terms of execution time. In this case, the accuracy of approximate algorithms is estimated using two proximity norms: the maximum of differences, i.e.

$$\|\mathrm{N}\|_1 = \max_{\boldsymbol{n} \in E} \left| p(\boldsymbol{n}) - \tilde{p}(\boldsymbol{n}) \right| ;$$

and similar cosine:

$$\|\mathrm{N}\|_2 = \frac{\sum_{\boldsymbol{n} \in E} p(\boldsymbol{n}) \tilde{p}(\boldsymbol{n})}{\left(\sum_{\boldsymbol{n} \in E} (p(\boldsymbol{n}))^2 \right)^{\frac{1}{2}} \left(\sum_{\boldsymbol{n} \in E} (\tilde{p}(\boldsymbol{n}))^2 \right)^{\frac{1}{2}}} .$$

Here, the following optimization problem is solved: it is required to maximize the profit of the system by choosing the appropriate values of the critical inventory level and the threshold parameter that regulates the access of low priority requests, i.e. you need to solve the following problem:

$$\left(s^*, r^*\right) = \arg \max_{(s,r) \in X} PT(s,r), \tag{10}$$

where

$$PT(s,r) = RV(s,r) - TC(s,r).$$

It is shown, that

$$RV(s,r) = \left(\lambda_1 (1 - PB_1(s,r)) C_{rev}^1 + \lambda_2 (1 - PB_2(s,r)) C_{rev}^2 \right) PS(s,r),$$

where

$$PS(s,r) = \frac{\mu_2 \sigma_2}{\lambda + \mu_1 \sigma_1 + \mu_2 \sigma_2} (1 - \rho(0))(1 - \pi(<0>));$$

$TC(s,r) = (K + c_r V_{av}(s,r)) RR(s.r) + c_h S_{av}(s,r) + c_l^1 \lambda_1 PB_1(s,r)$
$+ c_l^2 \lambda_2 PB_2(s,r)$.

Problem (10) always has a solution, since the set of feasible solutions $X = \{0 \le s < S/2\} \times \{1 \le r \le N-1\}$ is finite and discrete. Table 1 shows the results of solving problem (10) for a model with (s,S) and $(m,S-m)$ policies, where PT^* denotes the maximum value of the functional $PT(s,r)$ and

$C_{rev}^1 = 5, C_{rev}^2 = 10, K = 0.2, c_r = 0.01, c_h = 0.2, c_l^1 = 0.2, c_l^2 = 0.6.$

Table. 1 Problem solution results (10); $\mu_1 = 55$, $\mu_2 = 15$, $\sigma_1 = 0.2$, $\varphi_1 = 0.3$, $\nu = 3$, $\tau = 2$.

(*S*, *N*)	(λ_1, λ_2)	(*s*, *S*) policy		(*m*, *S*-*m*) policy	
		(s^*, r^*)	PT^*	(s^*, r^*)	PT^*
(30,45)	(45,8)	(14,39)	10.45	(1,41)	9.59
	(55,12)	(14,35)	6.52	(1,39)	5.97
	(60,15)	(14,7)	4.87	(1,4)	4.49
(30,50)	(45,8)	(14,44)	10.50	(1,46)	10.00
	(55,12)	(14,41)	6.55	(1,44)	6.24
	(60,15)	(14,36)	4.88	(1,41)	4.63
(40,45)	(45,8)	(19,38)	8.34	(1,41)	8.52
	(55,12)	(19,33)	4.40	(1,38)	4.87
	(60,15)	(19,8)	2.74	(1,4)	3.36
(40,50)	(45,8)	(19,43)	8.35	(1,46)	8.84
	(55,12)	(19,38)	4.41	(1,43)	5.08
	(60,15)	(19,32)	2.74	(1,40)	3.46

The remainder of this chapter explores the dual claim model of the QIS, which uses a second heterogeneous claim access scheme. In contrast to the first access scheme, here the fate of an incoming request is determined based on the current level of stocks, while the QIS models with a common finite and infinite queue of different types of requests are studied.

Assume that different types of requests are waiting in a common buffer, which has a maximum size of $N, N < \infty$. Access of different types of requests to the system is carried out as follows. If at the moment of arrival of the second type request the buffer is not completely filled (i.e. there is at least one free space in the buffer), then this request is accepted into the buffer; otherwise (i.e. when the buffer is full) it is lost. A low-priority order (i.e., orders of the first type) is only accepted into the buffer when, at the time of its entry into the system, the inventory level is not less than s. This means that if at the moment of arrival of the order of the first type the inventory level is less than s, then regardless of the availability of free places in the buffer, it is lost with probability one. Note that priority claims can join the queue even when the inventory level of the system is zero. This process is carried out as follows: if at the time of receipt of a second type request, there are no stocks in the system (i.e., the stock level is zero), then this request is accepted into the buffer with probability φ_1, and with an additional probability $\varphi_2 = 1 - \varphi_1$ it leaves the system unserved.

An analysis of the operation of the last model with a finite queue shows that the average inventory level and average order intensity are determined similarly to (6) and (7), but here the probabilities of losing different types of orders are calculated as follows:

$$PB_1 = \sum_{m=s}^{S} p(m,N) + \sum_{m=0}^{s-1}\sum_{n=0}^{N} p(m,n);$$

$$PB_2 = \sum_{m=0}^{S} p(m,N) + \sum_{n=1}^{N} p(0,n)\cdot\frac{n\tau}{\lambda_1 + n\tau}.$$

A model of a system with an infinite queue is considered using three restocking policies: (s,S)-policy, $(S-1,S)$- policy and $(m,S-m)$- policy. It is shown that the condition for the ergodicity of this model is the fulfillment of the inequality $\psi_2 < 1$, where $\psi_2 = \lambda/\mu_1$. So, when using the $(S-1,S)$-policy the probabilities of aggregated states when using $(S-1,S)$-policy are found from the following relations:

$$\pi(<m>)=\begin{cases}\dfrac{S!}{(S-m)!}\left(\dfrac{\nu}{\theta_1}\right)^m \pi(<0>), & \text{если } 0\le m\le s,\\ \dfrac{S!}{(S-m)!}\left(\dfrac{\theta_2}{\theta_1}\right)^s\left(\dfrac{\nu}{\theta_2}\right)^m \pi(<0>), & \text{если } s+1\le m\le S,\end{cases}$$

where $\pi(<0>)=\left(\sum_{m=0}^{s}\dfrac{S!}{(S-m)!}\left(\dfrac{\nu}{\theta_1}\right)^m+\left(\dfrac{\theta_2}{\theta_1}\right)^s\sum_{m=s+1}^{S}\dfrac{S!}{(S-m)!}\left(\dfrac{\nu}{\theta_2}\right)^m\right)^{-1}$.

Further, for this policy, the approximate values of the characteristics are calculated as follows:

$$S_{av}\approx\sum_{m=1}^{S}m\pi(<m>);$$

$$PB_1\approx\sum_{m=0}^{s-1}\pi(<m>);$$

$$PB_2\approx\pi(<0>)e^{-\omega}\sum_{n=1}^{\infty}\frac{\omega^n}{n!}\cdot\frac{n\tau}{\lambda_1+n\tau}.$$

$$RR\approx\mu_2\left(1-\sum_{m=0}^{S}\rho_m(0)\pi(<m>)\right).$$

The following notation is adopted in the last formulas:

$$\omega=\lambda_2\varphi_1/\tau;\ \rho_0(n)=\frac{\omega^n}{n!}e^{-\varpi},\ n\ge 0;$$

$$\psi_1=\lambda_2/\mu_1 \text{ и } \psi_2=\lambda/\mu_1,$$

$$\rho_m(n)=\begin{cases}\psi_1^n(1-\psi_1), & 1\le m\le s, n\ge 0,\\ \psi_2^n(1-\psi_2), & s+1\le m\le S, n\ge 0.\end{cases}$$

Similarly to the model described above, this scheme of organization of priority access for different types of applications is also studied and the results of the corresponding computational experiments are presented.

CONCLUSION

The research carried out in this dissertation allows us to draw the following conclusions.

1. In the vast majority of studies of IMS and QIS, the basic assumption is that the applications are identical. This assumption significantly reduces the adequacy of the known models. It is shown that in order to increase the adequacy of the developed models of IMS and QIS, it is necessary to take into account the heterogeneity of incoming applications. The main features of the classification of applications in the IMS and QIS are given.

2. Mathematical models of the IMS with two types of requests and without repeated requests are proposed when using a deterministic and randomized restocking policy. It is shown that one-dimensional Markov chains are used as mathematical models of these systems with the use of the indicated IRP The generators of each CM are constructed, and formulas are obtained for calculating their probabilities of the states of this circuit. Exact formulas are proposed for calculating the characteristics of the studied IMS - the probability of losing requests of each type and the average level of stocks of the system.

3. Mathematical models of IMS are proposed with two types of applications and repetition of only low-priority applications, if at the time of their receipt the level of stocks of the system is less than a certain (threshold) value. It is believed that one of three types of deterministic IRP can be used in the system. Models with finite and infinite orbit dimensions for such applications are considered. Exact and approximate methods have been developed for calculating the characteristics of the systems under study – the average inventory level of the system, the average number of repeated claims in orbit, and the probabilities of losing heterogeneous incoming claims, as well as repeated claims. Problems of their optimization are solved.

4. Mathematical models of QIS with two types of applications and finite and infinite queues have been developed using two restocking policies: in one policy, the volume of supplied stocks is a constant value, and in another policy, it is a variable. It is believed

that after the completion of servicing, some of the orders of both types may refuse to buy stocks, while the service time for orders that receive stocks differs from the time for servicing orders that do not receive stocks.

5. Two schemes for organizing priority service in QIS with applications of two types are proposed. In both schemes, high-priority orders are accepted at any stock level and if there is at least one free space in the waiting buffer (in the case of a model with an infinite queue, these orders are always accepted into the system). In one scheme, low-priority orders are accepted only when the inventory level is above a certain (threshold) value, and in the other, such orders are accepted only when the total queue length is less than a certain (threshold) value.

6. It is shown that QIS models with two types of requests are two-dimensional Markov chains. Algorithms for constructing generating matrices of the corresponding Markov chains have been developed and SEE have been obtained to find their stationary distributions. Formulas are found for calculating the characteristics of the studied QIS - the probability of losing requests of each type, the average level of stocks of the system and the average number of requests in the system.

7. A unified approach is proposed for studying two-dimensional IMS and QIS models with large dimensions of their state space. It is based on the method of phase aggregation of states and allows one to find simple approximate formulas for finding the desired characteristics of these systems using the restocking policies noted above.

8. The results of numerical experiments for all models of IMS and QIS are demonstrated. The results of a comparative analysis of the characteristics of the systems under study with the use of various IRP are given.

The reliability and validity of the scientific and practical conclusions of the dissertation work is confirmed by the correct application of the results of the theories of inventory management and queuing.

The list of references contains the name of the main works, in which the models of inventory management systems and stock-keeping systems are investigated. Attached is the Act on the use of the results.

The main results of the dissertation work are published in the following scientific articles:

1. Алиев, И.А. Об одной модели системы обслуживания-запасания с двумя типами заявок // - Баку: Вестник Бакинского Университета, - 2017. №4, - С. 105-110.
2. Melikov, A.Z., Ponomarenko, L.A., Aliyev, I.A. Analysis and Optimization of Models of Queuing-Inventory Systems with Two Types of Requests // - USA: Journal of Automation and Information Sciences. Begell House, - 2018. Vol.50, №12, - P. 34-50. **(Scopus)**
3. Melikov, A.Z., Ponomarenko, L.A., Aliyev, I.A. Markov Models of Systems with Demands of Two Types and Different Restocking Policies // - USA: Cybernetics and Systems Analysis. Springer, - 2018. Vol.54, №6, - P. 900-917. **(Web of Science)**
4. Алиев, И.А. Модель системы обслуживания-запасания с разнотипными заявками // Материалы XVII Международной конференции «Информационные технологии и математическое моделирование», - Томск: - 10-15 сентября, - 2018, - С. 274-280.
5. Алиев, И.А. Модель системы запасания с разнотипными заявками и мгновенным обслуживанием // Doktorantların və gənc tədqiqatçıların XXII Respublika elmi konfransının materialları, - Bakı: - 22-23 noyabr, - 2018, - S. 218-219.
6. Melikov, A.Z., Ponomarenko, L.A., Aliyev, I.A. Markov Models of Queuing-Inventory Systems with Different Types of Retrial Customers // - USA: Journal of Automation and Information Sciences. Begell House, - 2019. Vol.51, №8, - P. 1-15. **(Scopus)**
7. Алиев, И.А. Расчет характеристик системы обслуживания-запасания с разнотипными заявками и конечной орбитой

для повторных заявок // Материалы XVIII международной конференции «Информационные технологии и математическое моделирование», - Томск: - 26-30 июня, - 2019, - С. 115-120.

8. Aliyev, I.A. Mathematical models of resource management systems with instant service and two types of applications // Proceedings of the XXV-th International Open Science Conference "Modern Informatization Problems in Economics and Safety", - Yelm, WA, USA: - January, - 2020, - P. 4-9.
9. Алиев, И.А. Расчет модели системы обслуживания с различными политиками пополнения запасов // Материалы XII международной научно-практической конференции «Интернет–Образование-Наука» (ИОН-2020), - Винница: - 26-29 мая, - 2020, - С. 120-122.
10. Алиев, И.А. Численное исследование и оптимизация системы управления запасами с мгновенным обслуживанием и двумя типами заявок // - Воронеж: Системы управления и информационные технологии, - 2020. №2 (80), - С. 28-34. **(ВАК Российской Федерации)**
11. Алиев, И.А. Управления запасами одной системы с разнотипными заявками // - Воронеж: Системы управления и информационные технологии, - 2021. №1 (83), - С. 4-8. **(ВАК Российской Федерации)**
12. Алиев, И.А. Модель системы с мгновенным обслуживанием и рандомизированной политикой пополнения запасов // - Киев: Проблемы информатизации и управления, - 2021. Том 2, №66, - С. 4-10. **(ВАК Украины)**

In the jointly published works [2, 3, 6], the author's personal contribution consists in the development of algorithms and formulas for calculating the stationary distributions of the constructed multidimensional Markov chains, performing numerical experiments, and analysing their results.

The defense will be held on 30 september 2022 at 14^{00} at the meeting of the Dissertation council ED 1.20 of Supreme Attestation Commission under the President of the Republic of Azerbaijan operating at Institute of Control Systems of ANAS.

Address: 68, B.Vahabzadeh, Baku, AZ1141.

Dissertation is accessible at the Institute of Control Systems of ANAS Library.

Electronic version of dissertation and its abstract are available on the official website (http://www.isi.az) of the Institute of Control Systems of ANAS.

Abstract was sent to the required addresses on 26 august 2022.

Signed for print: 13.07.2022

Paper format: A5

Volume: 38461

Number of hard copies: 20