REPUBLIC OF AZERBAIJAN

On the rights of the manuscript

ABSTRACT

of the dissertation for the degree of Doctor of Philosophy

Application of queueing models in inventory control systems

Specialty:	3338.01 System Analysis, Management and Information Processing
Field of science:	Technical sciences
Applicant:	**Mammad Ogtay Shahmaliyev**

Sumgait - 2021

The work was performed at National Aviation Academy, "Information Technologies" department.

Scientific supervisor:	ANAS correspondent member, professor **Agasi Zarbali Melikov**
Official opponents:	doctor of technical sciences, professor **Sayyaddin Mashadi Jafarov**
	doctor of technical sciences, professor **Alakbar Ali Agha Aliyev**
	doctor of philosophy in technical sciences, associate professor **Shahla Surkhay Huseynzada**

Dissertation council FD 2.25 of Supreme Attestation Commission under the President of the Republic of Azerbaijan operating at Sumgait State University

Chairman of the Dissertation council:	doctor of technical sciences, professor **Agil Hamid Huseynov**
Scientific secretary of the Dissertation council:	doctor of philosophy in technical sciences, associate professor **Turgay Kilim Huseynov**
Chairman of the scientific seminar:	Doctor of technical sciences, professor. **Valeh Azad Mustafayev**

INTRODUCTION

Relevance of the subject and literature review. The classical QS (Queueing Systems) models are based on the assumption that inventory items are unlimited. On the other hand, in inventory management systems the customers are assumed to be served instantaneously. However, these assumptions are not satisfied in many real-world systems. Therefore, we need to consider QIS (Queueing-Inventory Systems) models with limited inventory items to create better and realistic real-world models. QIS theory is widely applied in modelling and mathematical analysis of the service and logistic related systems.

QIS theory was founded at the end of the past century by Azerbaijani and American scientists independently. This subject has been widely investigated and developed by the researchers.

First QIS models mostly considered non-perishable inventories. Therefore, Perishable QIS (PQIS) models has been less studied. PQIS models could be applied in the analysis of the real systems with perishable inventory like blood reserves, food, pharmacy and chemical industry etc.

Secondly, in wide range of QIS models the customers are assumed to always acquire an inventory item after the service completion. But this assumption does not hold true for many real-world systems. In some cases, inventory remains unchanged after the service completion. For example, the customer may leave the system without acquiring an inventory item due to the bad service quality. This concept has been introduced recently and considered in recent works only.

QIS model are mostly modelled using multidimensional Markov Chains (MC). Then we need to calculate the stationary distribution of MC to analyze it Quality of Service (QoS) measures. The calculation of stationary distributions for infinite or large models may encounter practical computational issues. The known exact methods for the calculation of stationary distribution are of the polynomial complexity. Therefore, there is a need for the development of the new, efficient and less complex algorithms.

The available scientific literature review proves that the PQIS models are less studied. Therefore, only models with perishable inventory are considered and introduced in the dissertation, the exact and approximate algorithms were developed for the analysis and calculation of the corresponding QoS measures.

Research goals and tasks. The main goal of the dissertation is to develop realistic PQIS models that fit the real-world systems, establish ergodicity conditions, calculate stationary distributions, define and improve QoS measures and solve the optimization problems.

Research methodology. Markov processes, QIS theory, linear programming and simulation algorithms, as well as, scientific software tools and libraries were applied to achieve the above goals.

Major claims to be defended: The major claims to be defended are followings:

1. New realistic PQIS models;
2. New methods and algorithms for the calculation of stationary distribution and corresponding QoS measures of infinite or large models.

Scientific results. The following results were obtained in the dissertation:

1. The new PQIS models were developed with the application of different inventory replenishment policies. The models without feedback component were modelled by two-dimensional MC.
2. The new PQIS models with delayed feedback were proposed and modelled by three-dimensional MC.
3. Corresponding transition matrices and balance equations were constructed, ergodicity conditions were specified for the developed models. Approximate formulas for the calculation of stationary distributions and QoS measures were proposed.
4. Space merging algorithm was developed and applied to the proposed models in order to eliminate the practical computational complexities. Its efficiency were demonstrated in comparison with known exact algorithms.

5. Proposed algorithms were applied in numerical experiments to analyze the developed models, to solve the corresponding optimization problems. The high accuracy of Space merging method were comparatively demonstrated and experimentally proven in numerical results.
6. Simulation algorithms were applied to analyze, calculate the stationary distributions and QoS measures of the models with infinite state space.

Theoretical and practical importance. The developed models could be applied in real-world systems with perishable inventory (food, chemical and pharmaceutical industry, blood reserves etc.) to calculate the stationary distributions and QoS measures, to solve optimization problems, to improve QoS measures, to minimize the perishability rate, inventory waste and overall system expenses. Moreover, the proposed approximate algorithm could be applied to calculate the QoS measures of infinite and large models with high accuracy.

Approbation. The obtained results in dissertation were discussed in the following international and local scientific conferences:

1. IV International Scientific Conference of Young Researchers Devoted to the 93rd Anniversary of Azerbaijani National Leader Heydar Aliyev April 29-30, 2016, Baku/Azerbaijan.
2. 15th International Conference named after A. F. Terpugov: Information Technologies And Mathematical Modelling (Tomsk, October 2016).
3. 16th International Conference named after A. F. Terpugov: Information Technologies And Mathematical Modelling (Kazan, September-October 2017).
4. 17th International Conference named after A. F. Terpugov: Information Technologies And Mathematical Modelling (Tomsk, September 2018).
5. International Scientific Conference For The Information Systems And Technologies Achievements And Perspectives

dedicated to 100th anniversary of Azerbaijan Democratic Republic (Sumgait, 15-16 November, 2018).

6. “International Conference on Advances in Applied Probability and Stochastic Processes” international conference (Kottayam, India, 7-10 January 2019).

The dissertation work was performed at National Aviation Academy of Azerbaijan Republic.

The structure and volume of dissertation. The volume of the dissertation is approximately distributed among the chapters as follows:

- Total character count – 180000 symbols
- Introduction – 11000 symbols
- First chapter – 55000 symbols
- Second chapter – 60000 symbols
- Third chapter – 52000 symbols
- Conclusion – 6500 symbols

SUMMARY

The first chapter was dedicated to the review analysis of the existing scientific literature on QIS and PQIS models. According to the analysis, the PQIS models were found mostly in recent articles and were less studied. The history of QIS theory, types of different models and information about known algorithms and methods were given. The different types of inventory replenishment policies, their pros and cons were described as well. The application of Markov chains for the PQIS modelling and the general problem statement was introduced. Different exact and approximate methods and algorithms for the solution of the given problem were classified. The types of numerical experiments, optimization problems were introduced and classified at the end of the chapter.

The main components of the studied QIS models are incoming customers, waiting buffer or queue, server(s), inventory, supplier, feedback (or orbit). Simple single server QIS structure is depicted in Figure 1.

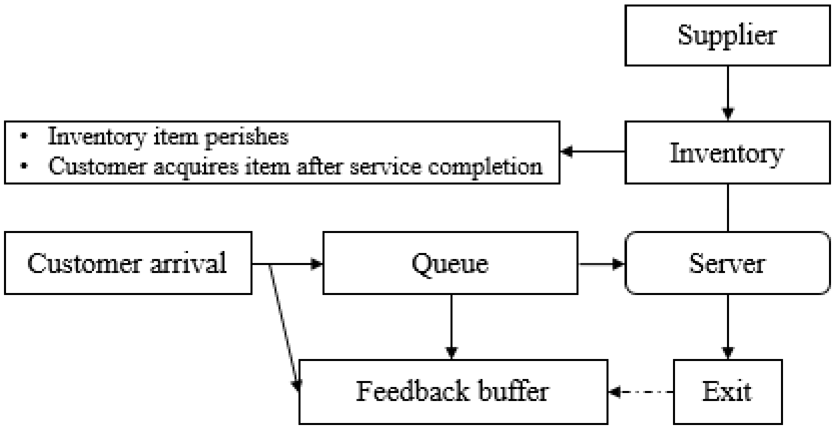


Figure 1. Structure of the single server QIS model

The inventory component manages inventory in the system and provides item from inventory to the served customers. The inventory level decreases after the served customer acquires item or the inventory item perishes. Therefore, it is important to maintain the necessary inventory level and periodically replenish the inventory. The replenishment of the inventory is performed by the corresponding supplier services (or supplier). The inventory is replenished according to the selected replenishment policy. The selection of the replenishment policy is very important for the overall system optimization and improvement of QoS measures in the long run. There are majorly four types of replenishment policies based on the replenishment volume: fixed, dynamic, random and hybrid.

One of the types of QIS models is the perishable QIS (PQIS) which is the major subject of the dissertation. The main assumption in PQIS models is that inventory level decreases not only after the served customer acquires the item but also after the item perishes. We concluded from the literature review analysis that the PQIS models are less studied and this is a relatively new subject.

The main tool applied in the mathematical modeling of PQIS is Markov chains. When modeling the PQIS with Markov chains the state of the system is described by the inventory level and the number of the customers in the system in the two-dimensional case, additionally with the number of the customers in the feedback buffer for the three-dimensional case. The set of all the possible states of the system is called the *state space*. The transitions between the states of the system is described by the transition matrix and is denoted by $Q(i,j)$.

The main problem in the analysis of the Markov models is calculation of the stationary distribution of the states of the system. If the system satisfies the ergodicity condition(s) there exists a single limiting stationary distribution vector which does not depend on the initial state of the system.

To find the stationary distribution vector we need to solve the system of linear equations which is based on the global balance equations.

The known methods used for the calculation of stationary distribution mostly apply matrix decompositions, eigenvalues and eigenvectors and other elements of matrix algebra. These algorithms have polynomial complexity and could be effective only for small or medium-sized models.

One of the algorithms that imposes restrictions on the structure of transition matrix is Spectral expansion (SE) method. To apply the SE method the initial transition matrix is divided into three sub-matrices based on the transitions of the second infinite component of the state vector (horizontal, one-step upward and one-step downward). Then it is required that these sub-matrices do not change after some positive threshold M (spectral threshold).

With the growth of the length of the state vector, for the large or infinite sized models the application of exact algorithms becomes practically inefficient and encounters computational complexities. In that case we should apply approximate algorithms. So the space merging algorithm was applied in the dissertation that has been extensively applied in recent works. The space merging algorithm uses “divide and conquer” principle to divide the original two-dimensional model into one-dimensional models. Then the stationary distribution of the initial model is approximately calculated via stationary distributions of the obtained one-dimensional models. The stationary distribution of the three or larger multi-dimensional models is calculated by successively applying space merging algorithm (hierarchical space merging).

When the exact and approximate algorithms are not suitable for the given model we could apply simulation algorithms. Gillespie's Direct method is used in the dissertation to perform numerical experiments due to its simplicity of implementation and practical efficiency.

After calculating the stationary distribution it is possible to analyze the system's behavior and its QoS measures in the long run. QoS analysis plays an important role in system optimization and service improvement.

The next step in the PQIS analysis after the calculation of stationary distribution and QoS measures is solving optimization problems. The optimization problems consider the minimization of the overall expenses of the system (Total Cost, TC). The main task is to determine and minimize the TC function. One and two-parameter TC functions were considered and optimized in the dissertation.

The **second chapter** of the dissertation covers the analysis of two-dimensional PQIS models. PQIS models with different types of customers were introduced and analyzed under the application of different inventory replenishment policies. Both finite and infinite models were considered and corresponding optimization problems were solved.

The main assumption in classical QIS models is that the inventory level decreases after the service completion. But there are recent works that considered new models where this assumption does not hold[1]. In these models, it is assumed that the inventory level does not change after the service completion in some cases. For example, a customer may refuse to acquire the item due to the low service quality. We applied this concept to the models with perishable inventory and introduced PQIS models with different types of

---

[1] Krishnamoorthy, A., Manikandan, R., Shajin, D. Analysis of a Multiserver Queueing-Inventory System // – London: Advances in Operations Research, – 2015. – 16 p.

customers. The main assumptions of the introduced model are followings:

1. The model with perishable inventory was considered.
2. Customers are assumed to enter the system even when the inventory level is zero.
3. The service time intensities for the customers that acquire the inventory item or leave the system empty-handed are different. This assumption holds true for many real systems, because the service process of the customer acquiring the item may require more time (for example, extra time may be needed for item packaging).
4. The space merging method was applied to the analysis of the model. High accuracy of the approximate formulas was demonstrated in numerical experiments.

The problem is to find the stationary distribution of the inventory level and the number of customers in the system. QoS analysis is also included in the problem statement. The QoS measures of the studied model are followings:

- S_{av}– average inventory level;
- Γ_{av} – average perishability rate;
- RR – inventory replenishment rate;
- PL – customer loss probability;
- L_{av} – average number of customers in the system.

The system is modeled with a two-dimensional Markov chain with state vector (m, n), where m and n denote inventory level and the number of customers in the system accordingly. The state space of the system is defined as follows:

$$E = \{(m, n): m = 0,1, \dots, S; n = 0,1, \dots, N\} \quad (1)$$

The state transition graph of the model with (s, S) replenishment policy is depicted in Figure 2. In (s, S) policy, the system initiates replenishment request of volume $(S - s)$ when the inventory level reaches the predefined threshold s.

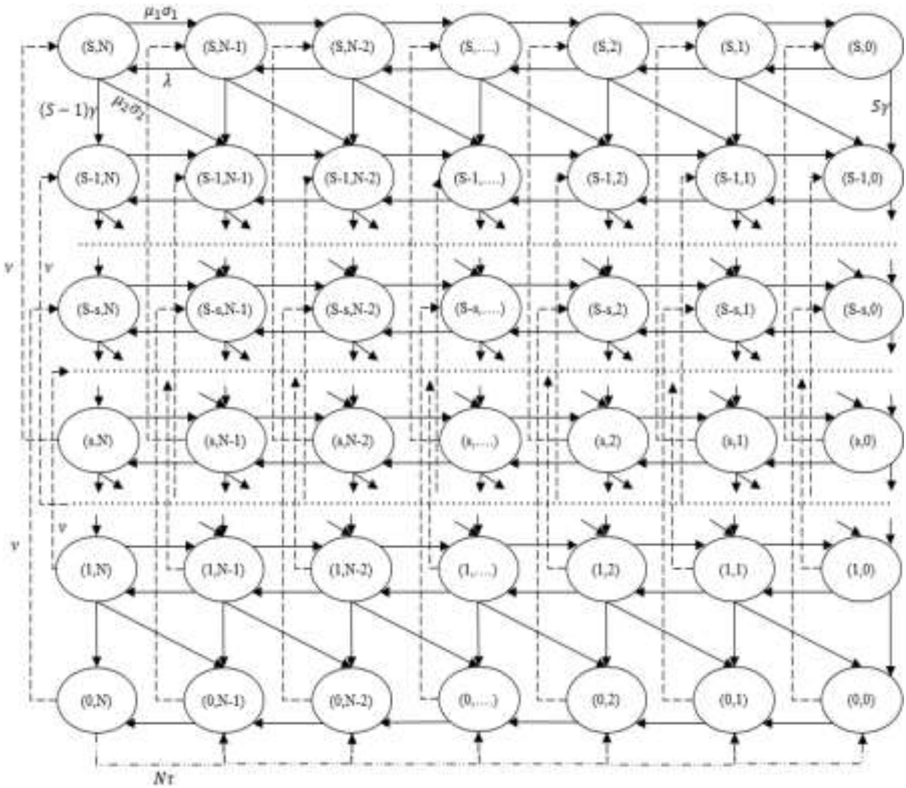


Figure 2. The state transition graph of (s, S) model

The space merging algorithm was applied to calculate the stationary distribution of the model with finite queue length ($N < \infty$). To apply the space merging algorithm it assumed that the customer arrival intensity is much greater than perishing and replenishment intensities, i.e. $\lambda \gg \max\{\gamma, \nu\}$. Additionally, as mentioned above the service intensities for every type of customer are different, i.e. $\mu_2 \ll \mu_1$.

In the next step, taking into consideration the above conditions the initial state space of the model (1) is divided into sub-spaces as follows:

$$E = \bigcup_{m=0}^{S} E_m\,,\ E_{m_1} \cap E_{m_2} = \emptyset,\ \ when\ \ m_1 \neq m_2 \tag{2}$$

Here $E_m = \{(m,n) \in E: n = 0,1,\dots,N\}$.

Next, based on the partitioning (2) we define the merge function as follows:

$$U\big((m,n)\big) = \langle m \rangle \tag{3}$$

Here $\langle m \rangle \in E_m, m = 0,1,\dots,S$ is the merged state and the merged state space is denoted as follows: $\Omega = \{\langle m \rangle: m = 0,1,\dots,S\}$.

The stationary distribution of the studied model could be approximately calculated with the following formula[2]:

$$p(m,n) \approx \rho_m(n)\pi(\langle m \rangle) \tag{4}$$

The $\rho_m(n)$ in formula (4) is stationary distribution within the merged state space $E_m, m = 0,1,\dots,S$ and $\pi(\langle m \rangle)$ is stationary distribution of the merged state $\langle m \rangle \in \Omega$.

We conclude from (2) that every state within E_m partition is dependent only on the 2nd component. Therefore, the states within merged partitions further will be denoted by $n, n = 0,\dots,N$ for the simplicity.

According to Figure 2, the stationary distribution within the merged classes $E_m, m = 1,\dots,S$ is the same as in $M/M/1/N$ model with load parameter $a = \lambda/(\mu_1\sigma_1)$.

$$\rho_m(n) = a^n \frac{1-a}{1-a^{N+1}};\ \ m = 1,\dots,S \tag{5}$$

[2] Melikov, A.Z., Ponomarenko, L., Shahmaliyev, M.O. Analysis of perishable queuing-inventory systems with different types of requests // – New York: Journal of Automation and Information Sciences, – 2017. Vol. 49, Issue 9, – p. 42-60

Finally, when the quantities $\rho_m(n)$ and $\pi(\langle m\rangle)$ are known the initial stationary distribution is calculated with the formula (4). After the formula for the stationary distribution is developed the QoS measures of the system could be calculated as follows:

$$\left.\begin{array}{c} S_{av} \approx \sum_{m=1}^{S} m\,\pi(\langle m\rangle) \\ \Gamma_{av} \approx \gamma\left[\sum_{m=1}^{S} \pi(\langle m\rangle)(m\rho(0)+(m-1)(1-\rho(0)))\right] \\ RR \approx \pi(\langle s+1\rangle)((s+1)\gamma\rho(0)+(\mu_2\sigma_2+s\gamma)(1-\rho(0))) \\ PL \approx \rho(N)\big(1-\pi(\langle 0\rangle)\big)+\pi(\langle 0\rangle)(\rho_0(N)+ \\ +\sum_{n=1}^{N-1}\rho_0(n)\frac{n\tau}{\lambda+n\tau}) \\ L_{av} \approx \pi(\langle 0\rangle)\sum_{n=1}^{N} n\rho_0(n)+(1-\pi(\langle 0\rangle))\sum_{n=1}^{N} n\rho(n) \end{array}\right\} \quad (6)$$

Likewise, the space merging method was successfully applied to the model with infinite queue and $(S-1,S)$ replenishment policy, the corresponding formulas were developed both for stationary distributions and QoS measures. Additionally, the spectral expansion algorithm was applied in the calculation of stationary distribution for the similar model with $(s,S-m)$ replenishment policy.

The results of the space merging method were compared with the results of exact algorithm to evaluate the former's accuracy in MS/(s, S) model. $\|N\|_1$ – maximum absolute difference, $\|N\|_2$ - Jaccard similarity and $\|N\|_3$ – cosine norms were used for the comparison.

The following system parameter values were chosen in numerical experiments:

$$\mu_1=15,\mu_2=3,\gamma=2,\nu=1,\tau=0.5,\sigma_1=0.3,\phi_1=0.6 \quad (7)$$

Numerical experiments were performed in Python programming language using SciPy library and the results are displayed in Table 1.

Table 1. Accuracy of space merging method in MS/(s, S) model

Parameter values				Norms		
s	*S*	*N*	*λ*	$\Vert N \Vert_1$	$\Vert N \Vert_2$	$\Vert N \Vert_3$
6	20	10	40	0.00796	0.999029	0.915987
		10	60	0.005631	0.999577	0.942885
		30	40	0.004257	0.998506	0.911695
		30	60	0.002826	0.999538	0.944291
		50	60	0.002826	0.998905	0.928744
		70	60	0.00403	0.992309	0.859102
11	30	10	40	0.00639	0.998987	0.91286
		10	60	0.004521	0.99956	0.940802
		30	40	0.003417	0.998503	0.909424
		30	60	0.002268	0.999518	0.941928
		50	40	0.004682	0.987063	0.825411
		50	60	0.002268	0.998962	0.929457
		70	60	0.003235	0.993323	0.87311
16	40	10	20	0.00875	0.995336	0.831666
		10	40	0.005497	0.998962	0.911085
		10	60	0.003889	0.99955	0.939618
		30	40	0.00294	0.998504	0.908135
		30	60	0.001951	0.999507	0.940586
		50	60	0.001951	0.998995	0.929864
		70	60	0.002783	0.993855	0.881173
21	50	10	20	0.007799	0.995252	0.829281
		10	40	0.004899	0.998946	0.9099
		10	60	0.003466	0.999543	0.938827
		30	40	0.00262	0.998505	0.907273
		30	60	0.001739	0.9995	0.939689
		50	60	0.001739	0.999016	0.930135
		70	60	0.00248	0.994197	0.886607

We conclude from Table 1. that the accuracy of the space merging algorithm is very high. So that for the higher values of the arrival intensity, the $\|N\|_1$ norm approaches 0, while the other norms approach 1 which implies the increase in accuracy. This is explained by the fact that with the increase of arrival intensity the inter-transition intensities between merged sub-spaces decrease. This state, in turn, is the necessary condition for the correct application of the space merging method.

We used the results of exact algorithms in numerical experiments to evaluate the accuracy of the developed approximate formulas. The complexity of the exact algorithms used to solve the balance equations generally equals to $O(n^3)$. This means that for the larger values of S and N the requirement for the resources and time increases in a polynomial manner. For example, in the computer with CPU Core i7 2.40 Ghz and 8GB RAM it took 3-4 hours to calculate stationary distribution and QoS measures of the system. On the other hand, it took 3-4 seconds for the space merging method to do the same job in the same environment. These results demonstrates the practical efficiency of the space merging method for the large sized or infinite models.

We observed very high accuracy in the QoS measures related to the inventory processes, while there were practically negligible errors for service related QoS measures.

We concluded from the comparison of MS/(s, S) and MS/(s, S-m) models that the former was better in terms of optimization and inventory management.

Different optimization problems were considered with respect to the replenishment level and supplier selection, corresponding one and two-parameter TC functions were introduced. The developed TC functions were minimized using the brute force method and it was concluded that maintaining minimal replenishment level is better for optimizing replenishment expenses. Optimal supplier selection problem was considered for MS/(S,S-1) model and it was concluded that for larger inventory volumes the supplier with lower and higher rates should be chosen in finite and infinite cases correspondingly.

Three-dimensional PQIS models were studied in the **third chapter**. The feedback component was added to the state vector of the model presented in the previous chapter. So the state of the new model is described by the inventory level, the number of customers in the queue and feedback buffer (or orbit).

Two different models were studied based on the use case and source of the feedback buffer. In the first model the served customers are assumed to join the feedback buffer to re-join the system in the future. In the second model, the feedback buffer is filled by the impatient or by the customers that arrived when the inventory was empty.

In the models with different types of requests, it is assumed that the served customer according to Bernoulli scheme either acquires the item or leaves the system without purchasing the item and inventory level does not change. This assumptions holds true for many real systems and was not considered in classical QIS models.

The other assumption that misses in the classical models is retrial requests or the possibility of the customer to repeat its request in the future (r-customers). This event is referred to as feedback in QIS literature. This assumption is also common in many real systems. For example, the customer may repeatedly apply to the system because of good service. There are two kinds of feedback. If the served customer re-joins the system instantly this is called the Instantaneous Feedback (IFB), on the other hand, if the customer returns after some positive time interval then the model with delayed feedback (DFB) should be considered.

The **models with delayed feedback** were studied in the dissertation. The studied model is denoted by PQIS/DFB. The following assumptions were considered in the developed model:

- Models with perishable inventory were considered.
- Served customer does one the followings with different probabilities:
    - Leaves the system without acquiring an item.
    - Acquires the item and leaves the system.
    - Joins the orbit for “decision making” without acquiring an item.

- r-customers may also acquire the item.
- Both finite and infinite models were considered.
- The customers in the queue become impatient and leave the system independently at exponentially distributed random times when the inventory level is zero.

The structure and working scheme of the presented model is depicted in Figure 3.

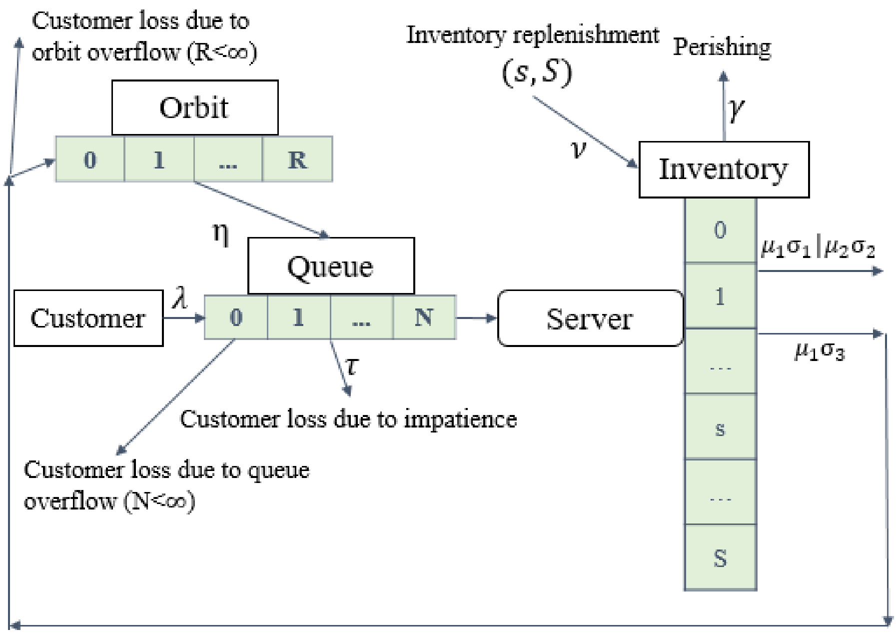


Figure 3. The structure of PQIS model with delayed feedback

The system is modeled using the three-dimensional Markov chains and the state vector is denoted by (m, n, k). Here m, n and k denote inventory level, the number of customers in the queue and orbit correspondingly. The state space of the system is defined as follows:

$$E = \{(m, n, k): m = 0,1, \dots, S; n = 0,1, \dots, N; k = 0,1, \dots, R\} \quad (8)$$

The pseudo-code for the transition matrix is demonstrated in Figure 4.

function QELEM$(m_1, n_1, k_1, m_2, n_2, k_2)$ $\triangle\, q((m_1, n_1, k_1), (m_2, n_2, k_2))$

  define $q := 0$

  if $k_2 = k_1$ **and** $m_1 > 0$ **then**

    if $m_2 = m_1$ **and** $n_2 = n_1 + 1$ **then** $q := \lambda$

    else if $m_2 = m_1$ **and** $n_2 = n_1 - 1$ **then** $q := \mu_1 \sigma_1$

    else if $m_2 = m_1 - 1$ **and** $n_2 = n_1 - 1$ **then** $q := \mu_2 \sigma_2$

    else if $m_2 = m_1 - 1$ **and** $n_2 = n_1 = 0$ **then** $q := m_1 \gamma$

    else if $m_2 = m_1 - 1$ **and** $n_2 = n_1 > 0$ **then** $q := (m_1 - 1)\gamma$

    else if $m_1 \le s$ **and** $m_2 = m_1 + S - s$ **and** $n_2 = n_1$ **then** $q := \nu$

  else if $k_2 = k_1$ **and** $m_1 = 0$ **then**

    if $m_2 = 0$ **and** $n_2 = n_1 + 1$ **then** $q := \lambda \phi_1$

    else if $m_2 = 0$ **and** $n_2 = n_1 - 1$ **then** $q := n_1 \tau$

    else if $m_2 = S - s$ **and** $n_2 = n_1$ **then** $q := \nu$

  else if $k_2 \ne k_1$ **and** $m_2 = m_1 > 0$ **then**

    if $n_2 = n_1 - 1$ **and** $k_2 = k_1 + 1$ **then** $q := \mu_1 \sigma_3$

    else if $n_2 = n_1 + 1$ **and** $k_2 = k_1 - 1$ **then** $q := k_1 \eta$

  return q

Figure 4. Pseudo-code of the transition matrix

In the next step, we divide the initial state space into sub-spaces as follows:

$$E = \bigcup_{k=0}^{R} E_k\,, \quad E_{k_1} \cap E_{k_2} = \emptyset, \quad k_1 \neq k_2 \; olduqda \tag{9}$$

Here $E_k = \{(m, n, k) \in E : m = 0,1, \dots, S;\ n = 0,1, \dots, N\}$, $k = 0,1, \dots, R$.

To calculate the stationary distribution and QoS measures of the given model we need to solve the $(S+1)(N+1)(R+1)$ dimensional system of linear equations. As this is practically inefficient, we applied the hierarchical space merging algorithm to solve the problem. To correctly apply the method we assume that the inter-transition intensities between the merged classes E_k are very small compared to the transition rates between the classes.

Conclusively, the approximate formulas for the stationary distributions and QoS measures were developed both for finite and infinite PQIS/DFB models.

In M/M/1/∞ model with infinite queue the impatient customers and arrived customers when the inventory is empty are assumed either to join the orbit or leave the system according to the Bernoulli scheme[3]. The r-customers in the orbit repeat their requests after exponentially distributed random times. Unlike PQIS/DFB model, in M/M/1/∞ model the source of orbit consists of the impatient customers and the customers that arrive when there are no items left in inventory. The stationary distributions and QoS measures of the models with (s, S) and $(s, S-m)$ replenishment policies were calculated using the simulation algorithm. The following formulas for QoS measures were developed in (s, S) model:

$$\left.\begin{array}{c} S_{av} \approx \sum_{m=1}^{S} m\,\pi_2(\langle m\rangle) \\ L_s \approx \mathrm{b}\pi_2(\langle 0\rangle) + \big(1-\pi_2(\langle 0\rangle)\big)\dfrac{a^2}{1-a} \\ L_o \approx \mathrm{c} \\ \Gamma_{av} \approx \gamma\left[\sum_{m=1}^{S} \pi_2(\langle m\rangle)(m-a)\right] \\ RR \approx \pi_2(\langle s+1\rangle)\big((s+1)\gamma(1-a) + (\mu_2\sigma_2 + s\gamma)a\big) \\ RL_s \approx \lambda\phi_1\pi_2(\langle 0\rangle) \end{array}\right\} \qquad (10)$$

The implementation of the Gillespie's direct simulation algorithm in python programming language for the model that applies $(s, S-m)$ replenishment policy is presented in Figure 5.

[3] Melikov, A.Z., Shahmaliyev, M.O. Queueing System M/M/1/∞ with Perishable Inventory and Repeated Customers // – Berlin: Automation and Remote Control, – 2019. Vol. 80, Issue 1, – p. 53-65.

```python
import random
def nextStateSpace(self,m1,n1,k1):
    nextStateSpace = {}
    for m2 in (m1, m1 + 1, m1 - 1, S):
        for n2 in (n1, n1 + 1, n1 - 1):
            for k2 in (k1, k1 + 1, k1 - 1):
                tr = self.Q_elem(m1, n1, k1, m2, n2, k2)
                if tr > 0: nextStateSpace[state] = tr
    return nextStateSpace
def simulate(initialState = 0, simulationTime = 5000):
    currState = initialState
    timePassed = 0
    reachedStates = {} # dict(state: sojournTime)
    nextStatesAll = {} # dict(state: nextStates)
    while timePassed < simulationTime:
        if currState not in reachedStates:
            reachedStates[currState] = 0
        if currState not in nextStatesAll:
            nextStatesAll[currState] = nextStateSpace(currState)
        nextStates = nextStatesAll[currState]
        rateSum = sum(nextStates.values()) # transition rate sum
        nextInterval = random.expovariate(rateSum)
        r = random.uniform(0,1)
        prevRateSum = 0
        for st in nextStates:
            p_sum = prevRateSum / rateSum
            n_sum = (prevRateSum + nextStates[st]) / rateSum
            if r > p_sum and r < n_sum:
                reachedStates[currState]  += nextInterval
                currState = st
                break
            prevRateSum += nextStates[st]
        timePassed += nextInterval
    totalTime = sum(reachedStates.values())
    return {s: reachedStates[s] / totalTime for s in reachedStates}
```

Figure 5. The implementation of the Gillespie's direct method for $(s, S - m)$ model

The results of the numerical experiments for the developed models were processed and visualized. The accuracy of the approximate formulas, the behavior of QoS measures and their dependence on the system parameters, different optimization problems were considered and analyzed.

The comparative analysis of the QoS measures for the PQIS/DFB model was presented for the following cases:

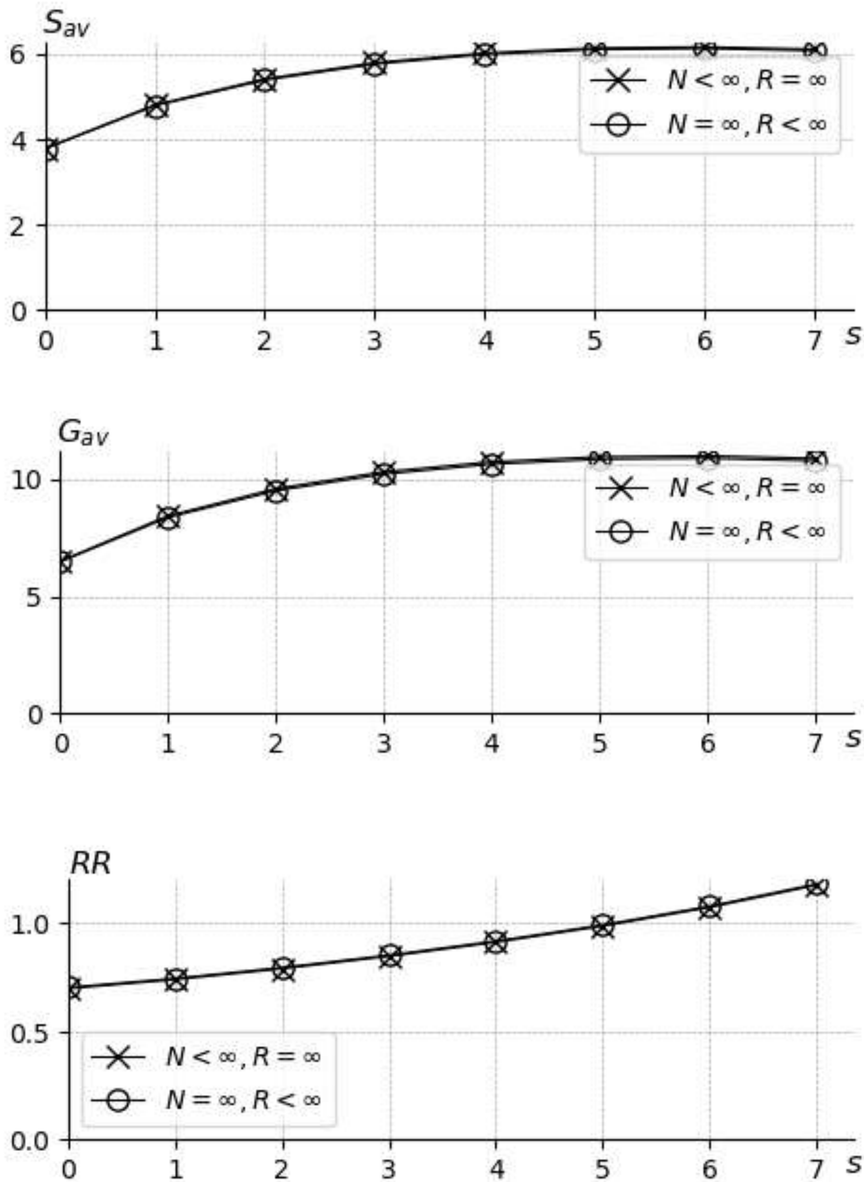
1. Finite queue and infinite orbit: $N < \infty, R = \infty$.
2. Infinite queue and finite orbit: $N = \infty, R < \infty$.

The system parameter values were chosen as follows in numerical experiments:

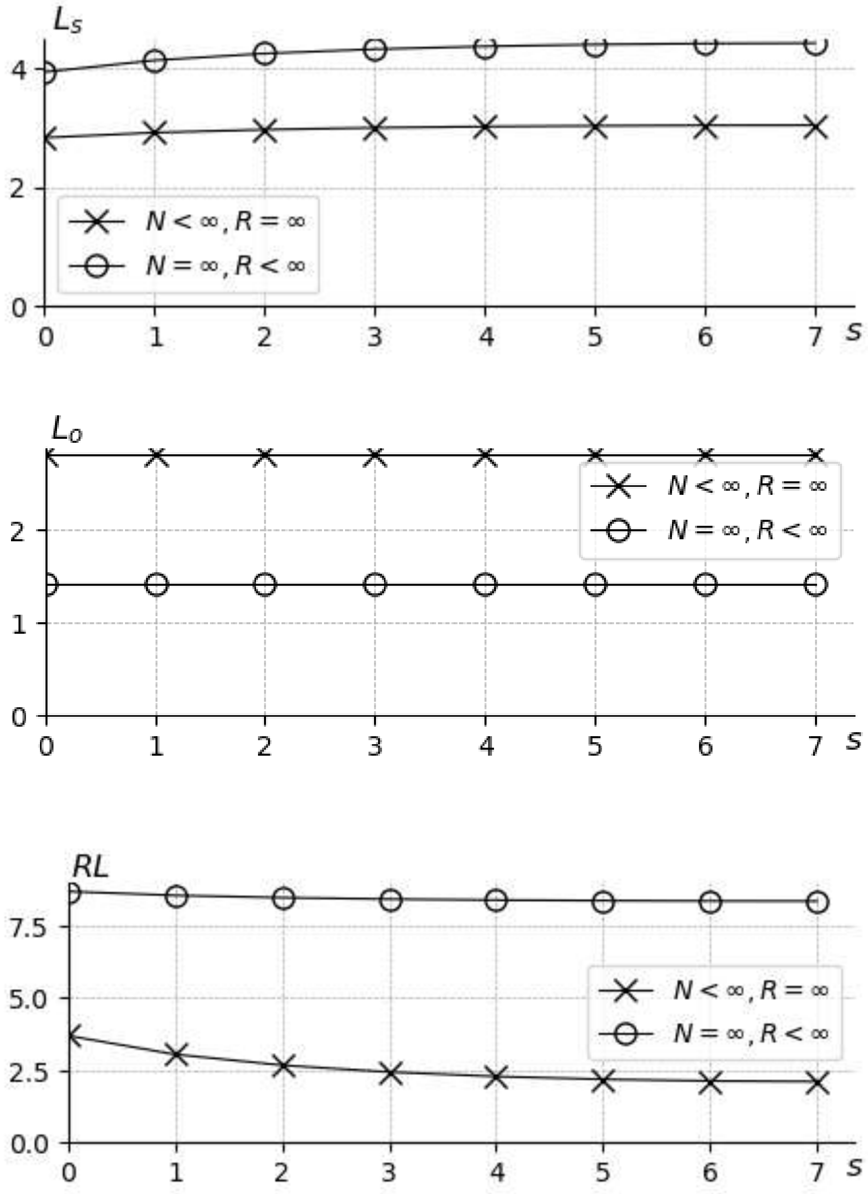
$$= 10, \eta = 5, \mu_1 = 60, \mu_2 = 15, \gamma = 2, \nu = 2, \tau = 1.5, \\ \sigma_1 = 0.2, \sigma_2 = 0.5, \phi_1 = 0.3, S = 15 \qquad (11)$$

We assume that $N = 10$ for finite queue and $R = 2$ for finite orbit cases correspondingly. We used approximate formulas in numerical experiments to calculate the QoS measures. The results of the experiments are depicted in Graph 1 and Graph 2. The x and o symbols were used in presented graphs to denote the first and second cases accordingly.

The comparison of QoS measures related to the inventory processes is presented in Graph 1. We conclude from the graphs that the average inventory level S_{av}, average replenishment rate RR, average perishability rate Γ_{av} measures are the same in both cases. This is explained by the fact that inventory related process do not depend on queue length or orbit length, instead they are dependent only on inventory volume and replenishment order level. Also with the increase of the parameter s the QoS measure RR naturally increases. The increase of RR, in turn, results in higher values of the average inventory level. Moreover, it is obvious from the graphs that the relation $\Gamma_{av} \approx \gamma S_{av}$ holds true.



Graph 1. Comparison of inventory related QoS measures in PQIS/DFB model



Graph 2. Comparison of service related QoS measures in PQIS/DFB model

The results for service related QoS measures are presented in Graph 2. Intuitively, as concluded from the graphs the average queue length L_s and the average orbit length L_o get higher values when $N = \infty$ and $R = \infty$ correspondingly. Also the average customer loss rate RL is higher when the orbit is finite, $R < \infty$.

Similarly, the optimization problems were considered for the studied models and were solved using the approximate formulas.

The TC function for the PQIS/DFB model is defined as follows:

$$TC(d,s) = (K + c_r(\mathrm{S} - \mathrm{s}))RR + c_s S_{av} + c_p \Gamma_{av} + c_l RL \\ + c_{ws} L_s + c_{wo} L_o \tag{12}$$

Here K– replenishment execution expense,
c_r – replenishment logistics expense per item,
c_s – inventory management expense per item,
c_p – per item perishing expense,
c_l – customer loss expense,
c_{ws} – queue management expense per customer,
c_{wo} – orbit management expense per customer.

We assume that there are multiple suppliers with different rates and service quality. Let's say the number of the suppliers is D and the supplier with index $d: d = 0, \ldots, D$ has the replenishment time intensity and price policy of $(\nu, K, c_r)_d$. To solve the problem (12) we need to find the optimal pair (d, s) that will minimize the TC function.

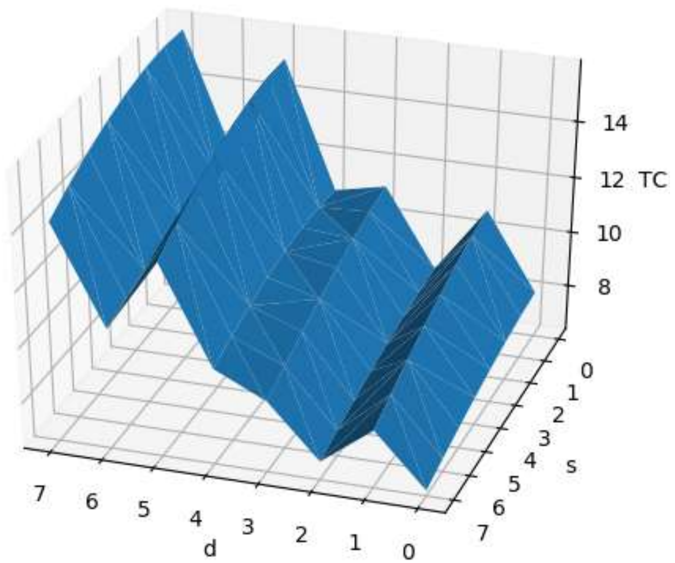
The system parameters and the function coefficients were chosen as follows:

$$c_s = 0.5, c_p = 0.5, c_l = 0.5, c_{ws} = 0.2, c_{wo} = 0.1, \\ \lambda = 5, \eta = 10, \gamma = 2, S = 15, \tau = 0.5, \sigma_1 = 0.2, \\ \sigma_2 = 0.7, \phi_1 = 0.4, \mu_1 = 50, \mu_2 = 30 \tag{13}$$

The existing suppliers are followings: ($D = 7$):

$$\begin{Bmatrix} (0.5,1,0.5), (0.5,1,1), (0.5,2,0.5), (0.5,2,1), \\ (1,1,0.5), (1,1,1), (1,2,0.5), (1,2,1) \end{Bmatrix} \qquad (14)$$

The optimal pair was found using the brute-force method as the set of all the possible pairs is discrete and finite. The results of the experiment is depicted in Graph 3. We concluded that to optimize the system in most cases the minimal replenishment level and the supplier with higher replenishment intensity should be chosen.



Graph 3. Solution of the optimization problem for PQIS/DFB model

The single-parameter TC function dependent on the replenishment level *s* was defined for M/M/1/∞ model as well. The optimization problem was given and solved with respect to the different suppliers using approximate formulas. We concluded that the optimal value of the parameter s does not depend on the supplier selected for the higher replenishment intensities and the minimal value of the parameter is ensured by the selection of “speedy” suppliers.

CONCLUSION

1. The existing scientific literature review and recent trends in QIS theory were presented in detail. The exact, approximate and simulation algorithms for the calculation of stationary distribution and QoS measures were developed and compared. The effectiveness of Spectral expansion method from exact algorithms, space merging method from approximate algorithms and Gillespie's direct method from simulation algorithms was explained. The types of numerical experiments and optimization problems in QIS analysis were summarized. It was noted that the main purpose in the calculation of stationary distribution and QoS measures is to analyze and optimize the overall system performance, increase the quality of customer service and customer satisfaction, decrease the system total run cost and prevent the unnecessary perishability in the long run.

2. It was concluded from the literature analysis that the models with different types of requests were less studied and were not applied in perishable QIS at all. Hence, the new PQIS models with different types of requests were developed. In that models the served customers are assumed either to acquire the item or leave the system empty-handed according to the Bernoulli scheme after the service completion. The models with finite and infinite state spaces using different inventory replenishment policies were studied and analyzed. The transition matrices and corresponding QoS measures were specified for each model. The exact, approximate and simulation algorithms were applied to calculate stationary distribution and QoS measures, the corresponding formulas were developed. The exact methods were noted to be efficient only for small or medium-sized models. The practical effectiveness and high accuracy of the space merging algorithm were proven and demonstrated in numerical experiments. The behavior and dependence of the QoS measures on the system parameters were analyzed in numerical experiments and the results were visualized and demonstrated in graphical and table forms. It was concluded that (s, S) replenishment policy is better in terms of optimization and total cost minimization. The optimal selection of replenishment level

and replenishment policy problems were considered and the corresponding TC functions were introduced. The suggessted optimization problems were solved using the developed approximate formulas and simulation algorithms. It was concluded that the replenishment level should be kept low to minimize overall replenishment expenses. Also for the larger values of inventory volume of MS/(S,S-1) model the selection of suppliers with lower and higher rates was proven to be optimal in finite and infinite cases correspondingly.

3. The two types of three-dimensional PQIS models with delayed feedback were developed. In the first model, the served customer joins the feedback buffer without acquiring the item for further decision making. In the second model, impatient customers and the customers that arrived when the inventory level was zero join the feedback buffer with some probability and repeat their requests after the exponentially distributed random times. The developed models were modeled using three-dimensional Markov chains, transition matrices were presented using transition diagrams and pseudo-codes, the ergodicity conditions were specified. It was noted that the exact algorithms may become practically inefficient for the calculation of stationary distribution and QoS measures, so the hierarchical space merging algorithm was applied and approximate formulas were developed. The high accuracy of the developed formulas was comparatively proven in numerical experiments. The behavior of QoS measures was analyzed in numerical experiments and the results were given in graphical and table forms. It was concluded that the QoS measures related to the inventory and service processes are directly and inversely proportional to replenishment intensity correspondingly. The independence of the inventory-related QoS measures on the orbit and queue components was demonstrated. Optimization problems and corresponding TC functions were developed and solved for the infinite models. It was concluded that maintaining the low replenishment level improves the overall system performance and the selection of the suppliers with higher intensities keeps the replenishment level low.

4. The corresponding software packages were developed using Python SciPy and Apache Commons Math libraries. The numerical experiment for the demonstration of the high speed of space merging algorithm was implemented using Java Apache Commons Math library. It took several hours for the exact method and several seconds for the space merging algorithm to perform the same job in the same computer. The corresponding software application was developed in Python programming language for the Gillespie's direct algorithm that is used for the simulation of infinite models. The spectral expansion method and Gauss algorithm for the solution of balance equations were implemented and applied as well.

The developed models could be applied in food, chemical and pharmaceutical industries and in other systems with perishable inventory to analyze QoS measures and improve the overall performance and customer service in the long run.

The list of the scientific works published on the dissertation subject

1. Shahmaliyev M. Алгоритмы расчета характеристик системы с конечным временем жизни запасов и повторными заявками // Ümummilli lider Heydər Əliyevin anadan olmasının 93-cü ildönümünə həsr olunmuş Gənc Tədqiqatçıların IV Beynəlxalq Elmi Konfransı. Bakı. 2016. Cild 1. S. 393-394.
2. Shahmaliyev M. Расчет характеристик одной системы обслуживания-запасания с конечным временем жизни запасов // Milli Aviasiya Akademiyası - Elmi Məcmuələr. 2016. Cild 18. № 2. S. 121-130.
3. Melikov A., Ponomarenko L., Shahmaliyev M. Models of Perishable Queueing-Inventory System with Repeated Customers // Journal of Automation and Information Sciences. Begell House. 2016. Vol. 48. Issue 6. P. 22-38.
4. Shahmaliyev M. Расчет характеристик системы с портящимися запасами и бесконечой орбитой для повторных заявок // Journal of Qafqaz University, Mathematic and Computer Science. 2016. Vol. 4. Issue 1. P. 30-39.
5. Shahmaliyev M. Об одной модели системы обслуживания-запасания с конечным временем жизни запасов // XV Международная конференция им. А.Ф. Терпугова «Информационные технологии и математическое моделирование ИТММ – 2016». 2016. С. 143-147.
6. Melikov A., Shahmaliyev M. Analysis of Perishable Queueing-Inventory System with Positive Service Time and (S−1,S) Replenishment Policy // Information Technologies and Mathematical Modelling. Queueing Theory and Applications. International Conference on Information Technologies and Mathematical Modelling (ITMM 2017). Springer, Cham. 2017. P. 83-96.
7. Melikov A., Ponomarenko L., Shahmaliyev M. Analysis of perishable queuing-inventory systems with different types of requests // Journal of Automation and Information Sciences. Begell House. 2017. Vol. 49. Issue 9. P. 42-60.

8. Shahmaliyev M. Müsbət xidmət vaxtlı müxtəlif sorğulu və ehtiyatları xarab olan sonsuz növbəli anbar idarəetmə sisteminin analizi // Milli Aviasiya Akademiyası - Elmi Məcmuələr. 2018. Cild 20. № 2. S. 161-168.
9. Shahmaliyev M. Simulation model of infinite perishable queueing inventory system with feedback // Control Systems and Computers. 2018. Issue 3 (275). P. 60-68.
10. Melikov A., Shahmaliyev M. Markov Models of Inventory Management Systems with a Positive Service Time // Journal of Computer and Systems Sciences International. Pleiades Publishing. 2018. Vol. 57. Issue 5. P. 766-783.
11. Melikov A., Krishnamoorthy A., Shahmaliyev M. Perishable Queuing Inventory Systems with Delayed Feedback // Information Technologies and Mathematical Modelling. Queueing Theory and Applications. International Conference on Information Technologies and Mathematical Modelling (ITMM 2018). Springer, Cham. 2018. P. 55-70.
12. Melikov A., Shahmaliyev M. Analysis of perishable queueing-inventory system with different types of demands and various replenishment policies // Azərbaycan Xalq Cümhuriyyətinin 100 illik yubileyinə həsr olunmuş İnformasiya Sistemləri və Texnologiyalar. Nailiyyətlər və Perspektivlər Beynəlxalq Elmi konfransın materiallari. Sumqayıt. 2018. S. 6-7.
13. Melikov A., Shahmaliyev M. Queueing System $M/M/1/\infty$ with Perishable Inventory and Repeated Customers // Automation and Remote Control. Pleiades Publishing. 2019. Vol. 80. Issue 1. P. 53-65.
14. Melikov A., Krishnamoorthy A., Shahmaliyev M. Numerical Analysis and Long Run Total Cost Optimization of Perishable Queuing Inventory Systems with Delayed Feedback // Queueing Models and Service Management. 2019. Vol. 2. Issue 1. P. 83-111

The contribution of the candidate in the co-authored articles was to develop formulas for the calculation of the steady state distributions and QoS measures of the models, to develop and solve optimization problems, to perform numerical experiments and parse its results.

The defense will be held on 11 May 2021 at 13:00 at the meeting of the Dissertation council FD 2.25 of Supreme Attestation Commission under the President of the Republic of Azerbaijan operating at Sumgait State University

Address: Azerbaijan , Sumgait, 43rd block, AZ5008

Dissertation is accessible at the Sumgait State University Library

Electronic versions of dissertation and its abstract are available on the official website of the Sumgait State University

Abstract was sent to the required addresses on 9 April 2021

Signed for print: 06.04.2021

Paper format: 60*84/1/16

Volume: 34161

Number of hard copies: 40