



Ссылка на статью:

// Ученые записки УлГУ. Серия Математика и информационные технологии. 2024, № 1, с. 51-59.

Поступила: 28.12.2023

Окончательный вариант: 11.01.2024

© УлГУ

УДК 519.872.3, 519.872.8

Модель замкнутой СМО с повторным обслуживанием

Савинов Ю.Г., * Лазарева В.А.

[*uras@aport.ru](mailto:uras@aport.ru)

Ульяновский государственный университет, Россия

В работе развивается траекторный подход к описанию и моделированию замкнутых систем массового обслуживания. Приведена модель с повторным обслуживанием. Показан переход от математической модели к итерационным формулам, по которым проводится имитационное моделирование.

Ключевые слова: система массового обслуживания, имитационное моделирование, повторное обслуживание, точечный процесс, компенсатор

Введение

Наряду с повторными вызовами (публикаций по данной тематике достаточно много, см., например, [1-10]) иногда имеет смысл рассматривать системы массового обслуживания (СМО) с повторным обслуживанием [11-15]. Например, в системе документооборота некоторый документ (например, отчет) поступает на согласование. Он согласовывается, если нет замечаний. В случае наличия замечаний документ отправляется на доработку. При этом, повторное согласование, как правило, происходит намного быстрее, чем первичное, поскольку проверяющие уже знакомы с документом. Математической моделью этой системы можно считать СМО с повторным обслуживанием (при этом в общем случае повторное согласование-обслуживание документа будет иметь другое распределение времени обслуживания, чем первичное). Ещё один пример, пусть есть фиксированный автопарк (например, машины такси). Эти автомобили по мере выхода из строя отправляются на ремонт и возвращаются из ремонта. Причем после ремонта автомобили отправляются на проверку (ОТК), где проходят контрольные испытания. Если автомобиль про-

верку проходит, машина отправляется работать в автопарк, иначе машина отправляется на повторный ремонт, при этом в силу экономических соображений, для машин, отправленных на повторный ремонт, имеет смысл установить приоритет в очереди на ремонт. Также в общем случае повторный ремонт будет иметь другое распределение времени ремонта (например, по сравнению с первичным ремонтом скорее среднее время ремонта будет существенно меньше). В качестве математической модели в этом случае также можно рассматривать СМО с повторным обслуживанием (ремонтом) заявок (машин). Модели с различными распределениями для первичного и повторного времени обслуживания и посвящена данная работа.

Для математического описания СМО в данной работе использован аппарат точечных (считающих) процессов и их компенсаторов [16]. Классическое марковское описание и исследование первых задач с ограниченным источником заявок было дано Пальмом [17]. К сожалению, для многоканальных замкнутых систем с приоритетами (именно такая модель рассматривается в данной работе), моделирование и исследование в марковском описании затруднено. Сложность анализа резко возрастает особенно для СМО, в которых средние длительности обслуживания различны для заявок разных типов/классов (см., например, [18-20]). Более подробное обоснование выбора в пользу траекторного подхода при описании многоканальных замкнутых СМО с приоритетами дано в [21-22]. В частности там показано, что при траекторном (семимартингальном) описании сложность как математической, так и компьютерной моделей практически не растет как с ростом емкости источника заявок, так и числа ремонтников [21].

1. Постановка задачи

Рассмотрим многоканальную СМО с ограниченным источником заявок емкости $n_1 > 0$ (см. рис.1). В СМО циркулируют заявки двух типов: первый тип – машины, требующие ремонта (первичного), второй тип – машины, требующие вторичного ремонта (они не прошли проверку ОТК с вероятностью $q > 0$ и отправлены на повторное обслуживание-ремонт). Вероятность поломки одной машины за малый промежуток времени Δ равна $\lambda_1 \Delta + o(\Delta)$, $\lambda_1 > 0$. Заявки и на первичный и на вторичный ремонт обслуживают r операторов, имеющих одинаковую квалификацию со средней длительностью $1/\mu_1 > 0$ для первичного ремонта и $1/\mu_2 > 0$ для вторичного. Для каждого типа заявок формируется отдельная неограниченная очередь, на первичный ремонт очередь q^1 , для вторичного ремонта q^2 . В начальный момент времени все n_1 машин находятся в исправном состоянии. После первичного и вторичного ремонта заявка-машина отправляется на проверку

ОТК. Емкость ОТК превышает общее число машин n_1 , поэтому очереди перед ОТК не возникает. Средняя длительность проверки в модуле ОТК равна $1/\mu_0 > 0$ для обоих типов заявок (набор проверок регламентирован и не зависит от типа ремонта). Далее машина с вероятностью $p = 1 - q > 0$ отправляется работать в автопарк, а с вероятностью $q > 0$ подается на повторное обслуживание-ремонт в очередь q^2 . У заявок из очереди q^2 относительный приоритет [17] перед заявками из очереди q^1 .

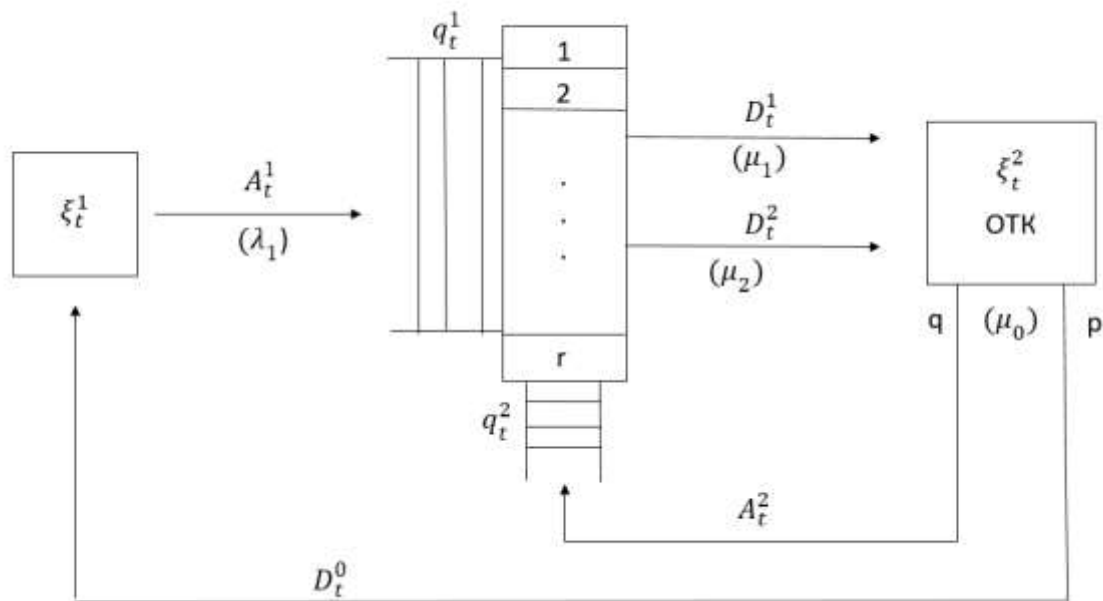


Рис. 1. Схема СМО

2. Математическая модель

2.1. Для описания математической модели будем использовать следующие процессы и параметры:

ξ_t^1 – число исправных автомобилей в автопарке в момент $t \geq 0$, $\xi_0^1 = n_1 > 0$;

A_t^1 – число первичных поломок автомобилей за время $t \geq 0$, $A_0^1 = 0$;

$r > 0$ – число ремонтников;

$\lambda_1 > 0$ – интенсивность первичных поломок автомобиля;

D_t^1 – число первично отремонтированных автомобилей время $t \geq 0$, $D_0^1 = 0$;

D_t^2 – число повторно отремонтированных автомобилей время $t \geq 0$, $D_0^2 = 0$;

r_t^1 – число автомобилей на первичном ремонте в момент $t \geq 0$, $r_0^1 = 0$;

r_t^2 – число автомобилей на повторном ремонте в момент $t \geq 0$, $r_0^2 = 0$;

$1/\mu_1 > 0$ – среднее время первичного ремонта;

$1/\mu_2 > 0$ – среднее время вторичного ремонта;

q_t^1 – очередь на первичный ремонт в момент $t \geq 0$, $q_0^1 = 0$;

q_t^2 – очередь на вторичный ремонт в момент $t \geq 0$, $q_0^2 = 0$;

$1/\mu_0 > 0$ – средняя длительность проверки в модуле ОТК;

ξ_t^2 – число автомобилей на проверке в модуле ОТК в момент $t \geq 0$, $\xi_0^2 = 0$;

$q > 0$ – вероятность повторного ремонта;

$p = 1 - q > 0$;

A_t^2 – число автомобилей, отправленных на повторное обслуживание за время $t \geq 0$, $A_0^2 = 0$;

D_t^0 – число автомобилей, успешно прошедших ОТК и отправленных в автопарк за время $t \geq 0$, $D_0^0 = 0$;

2.2. Опишем сначала работу по обработке заявок первого типа (первичный ремонт).

Процессы A_t^1 , D_t^1 – это точечные (считающие) процессы. Опишем их подробнее: $A^1 = (A_t^1)_{t \geq 0}$ – число заявок первого типа, поступивших на ремонт за время $t \geq 0$, $A_0^1 = 0$; $D^1 = (D_t^1)_{t \geq 0}$ – число обслуженных заявок первого типа за время $t \geq 0$, $D_0^1 = 0$.

Точечные процессы A^1 , D^1 определяются своими компенсаторами $\tilde{A}^1 = (\tilde{A}_t^1)_{t \geq 0}$ и $\tilde{D}^1 = (\tilde{D}_t^1)_{t \geq 0}$ в соответствии с разложением Дуба-Мейера для субмартингалов [16]:

$$\tilde{A}_t^1 = \int_0^t \xi_s^1 \lambda_1 ds, \quad \lambda_1 > 0. \quad (1)$$

Здесь ξ_t^1 – число заявок первого типа в СМО в момент времени $t \geq 0$, $\xi_0^1 = n_1$ (находящихся в рабочем состоянии). Компенсатор для процесса $D^1 = (D_t^1)_{t \geq 0}$ определяется следующим соотношением:

$$\tilde{D}_t^1 = \int_0^t r_s^1 \mu_1 ds. \quad (2)$$

Соответственно r_t^1 – число заявок первого типа, находящихся непосредственно на обслуживании (первичном ремонте) в момент времени $t \geq 0$, можно определить следующим образом:

$$r_t^1 = n_1 - \xi_t^1 - q_t^1 - \xi_t^2 - r_t^2 - q_t^2, \quad r_0^1 = 0. \quad (3)$$

Кроме этого, для ξ_t^1 в момент времени $t \geq 0$ можно написать следующее основное балансовое соотношение:

$$\xi_t^1 = n_1 - A_t^1 + D_t^0. \quad (4)$$

2.3. Аналогично опишем работу по обработке заявок второго типа.

Точечные процессы A_t^2 , D_t^2 , D_t^0 также определяются своими компенсаторами $\tilde{A}^2 = (\tilde{A}_t^2)_{t \geq 0}$ и $\tilde{D}^2 = (\tilde{D}_t^2)_{t \geq 0}$, $\tilde{D}^0 = (\tilde{D}_t^0)_{t \geq 0}$. Логику работы «виртуального роутера», который с

вероятностью $q > 0$ распределяет заявку на повторное обслуживание и с вероятностью $p = 1 - q > 0$ возвращает автомобиль работать в автопарк можно описать следующим образом. Пусть случайная величина $\eta \sim R[0; 1]$ – генерируется в момент окончания проверки автомобиля на ОТК. Тогда

$$\widetilde{A}_t^2 = \int_0^t \xi_s^2 \mu_0 I(\eta \in [0; q]) ds, \quad (5)$$

аналогично

$$\widetilde{D}_t^0 = \int_0^t \xi_s^2 \mu_0 I(\eta \in (q; 1]) ds, \quad (6)$$

где $I(\cdot)$ – индикаторная функция. Компенсатор для процесса $D^2 = (D_t^2)_{t \geq 0}$ определяется следующим соотношением:

$$\widetilde{D}_t^2 = \int_0^t r_s^2 \mu_2 ds. \quad (7)$$

Соответственно r_t^2 – число заявок второго типа, находящихся непосредственно на обслуживании в момент времени $t \geq 0$, можно определить следующим образом:

$$r_t^2 = n_1 - \xi_t^1 - q_t^1 - \xi_t^2 - r_t^1 - q_t^2, \quad r_0^2 = 0. \quad (8)$$

Кроме этого, для ξ_t^2 в момент времени $t \geq 0$ можно написать следующее основное балансовое соотношение:

$$\xi_t^2 = D_t^1 + D_t^2 - A_t^2 - D_t^0. \quad (9)$$

2.4. Опишем управление относительным приоритетом через регулирование размеров очередей.

Для размера очереди из заявок на первичный ремонт q_t^1 в момент времени $t \geq 0$ можно написать следующее балансовое уравнение:

$$dq_t^1 = I(r_t^1 + r_t^2 = r) dA_t^1 - I(q_t^1 > 0, q_t^2 = 0) dD_t^1 - I(q_t^1 > 0, q_t^2 = 0) dD_t^2. \quad (10)$$

Построение уравнения (10) имеет следующую логику. Во-первых, размер очереди заявок q_t^1 может увеличиться на одну заявку, если поступит заявка на первичный ремонт ($dA_t^1 = 1$) в момент, когда все r операторов заняты ($r_t^1 + r_t^2 = r$). Во-вторых, размер очереди q_t^1 может уменьшиться на одну заявку, если есть заявки в очереди q^1 , обслужилась заявка ($dD_t^1 = 1$ или $dD_t^2 = 1$) и нет приоритетных заявок в очереди q_t^2 .

Аналогично, для размера очереди из заявок на повторное обслуживание q_t^2 в момент времени $t \geq 0$ можно написать следующее балансовое уравнение:

$$dq_t^2 = I(r_t^1 + r_t^2 = r)dA_t^2 - I(q_t^2 > 0)dD_t^1 - I(q_t^2 > 0)dD_t^2, \quad (11)$$

Логика построения уравнения (11) аналогична (10). Разница в том, что с учетом относительного приоритета заявок в очереди q^2 нет необходимости учитывать число заявок в очереди q^1 . Таким образом, размер очереди заявок вида q_t^2 может увеличиться на одну заявку, если поступит заявка на повторный ремонт в момент, когда все r операторов заняты ($r_t^1 + r_t^2 = r$). Размер очереди заявок второго вида q_t^2 может уменьшиться на одну заявку, если обслужилась заявка ($dD_t^1 = 1$ или $dD_t^2 = 1$) и есть заявки в очереди ($q_t^2 > 0$).

3. Итерационные формулы для компьютерного моделирования

Выведем формулы для имитационного моделирования. На стохастическом базисе $\mathbb{V} = (\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ из формул (1)-(11) можно получить следующие инфинитезимальные соотношения:

$$P\{A_{t+\Delta}^1 - A_t^1 = 1 | \mathcal{F}_t\} = \lambda_1 \cdot \xi_t^1 \cdot \Delta + o(\Delta), \quad (12)$$

$$P\{A_{t+\Delta}^2 - A_t^2 = 1 | \mathcal{F}_t\} = \mu_0 \cdot \xi_t^2 \cdot I(\eta \in [0; q]) \cdot \Delta + o(\Delta), \quad (13)$$

$$P\{D_{t+\Delta}^0 - D_t^0 = 1 | \mathcal{F}_t\} = \mu_0 \cdot \xi_t^2 \cdot I(\eta \in (q; 1]) \cdot \Delta + o(\Delta), \quad (14)$$

$$P\{D_{t+\Delta}^1 - D_t^1 = 1 | \mathcal{F}_t\} = \mu_1 \cdot r_t^1 \cdot \Delta + o(\Delta), \quad (15)$$

$$P\{D_{t+\Delta}^2 - D_t^2 = 1 | \mathcal{F}_t\} = \mu_2 \cdot r_t^2 \cdot \Delta + o(\Delta). \quad (16)$$

Формулы (12)-(16) позволяют, основываясь на понятии геометрической вероятности, провести имитационное моделирование. А именно, введя дискретизацию (шаг по времени) Δ из условия $\lambda_1 \cdot \xi_t^1 \cdot \Delta \ll 1$, $\mu_0 \cdot \xi_t^2 \cdot \Delta \ll 1$, $\mu_i \cdot r_t^i \cdot \Delta \ll 1$, $i = 1, 2$ получим следующие итерационные формулы (для вычисления значений процессов в момент времени $t + \Delta$ через значения процессов в момент t):

$$A_{t+\Delta}^1 = A_t^1 + \delta(\lambda_1), \quad (17)$$

$$A_{t+\Delta}^2 = A_t^2 + \delta(\mu_0 \cdot \xi_t^2 \cdot I(\eta \in [0; q])), \quad (18)$$

$$D_{t+\Delta}^0 = D_t^0 + \delta(\mu_0 \cdot \xi_t^2 \cdot I(\eta \in (q; 1])), \quad (19)$$

$$D_{t+\Delta}^1 = D_t^1 + \delta(\mu_1 \cdot r_t^1), \quad (20)$$

$$D_{t+\Delta}^2 = D_t^2 + \delta(\mu_2 \cdot r_t^2), \quad (21)$$

где $\delta(\gamma) = \begin{cases} 1, & \text{с вероятностью } \gamma \cdot \Delta, \\ 0, & \text{с вероятностью } 1 - \gamma \cdot \Delta. \end{cases}$

Здесь $\Delta A_t^1 = A_{t+\Delta}^1 - A_t^1$, $\Delta A_t^2 = A_{t+\Delta}^2 - A_t^2$, $\Delta D_t^0 = D_{t+\Delta}^0 - D_t^0$, $\Delta D_t^1 = D_{t+\Delta}^1 - D_t^1$, $\Delta D_t^2 = D_{t+\Delta}^2 - D_t^2$.

Итерационные формулы для очередей имеют следующий вид:

$$q_{t+\Delta}^1 = q_t^1 + \Delta A_t^1 \cdot I(r_t^1 + r_t^2 = r) - \Delta D_t^1 \cdot I(q_t^2 = 0, q_t^1 > 0) - \Delta D_t^2 \cdot I(q_t^2 = 0, q_t^1 > 0), \quad (22)$$

$$q_{t+\Delta}^2 = q_t^2 + \Delta A_t^2 \cdot I(r_t^1 + r_t^2 = r) - \Delta D_t^1 \cdot I(q_t^2 > 0) - \Delta D_t^2 \cdot I(q_t^2 > 0). \quad (23)$$

Заметим, что ξ_t^1 , r_t^1 , r_t^2 – определяются алгебраическими уравнениями и не требуют итерационных формул для моделирования.

4. Результаты компьютерного моделирования

Практическая реализация СМО осуществлена с помощью языка программирования высокого уровня C# в среде разработки Visual Studio 2019.

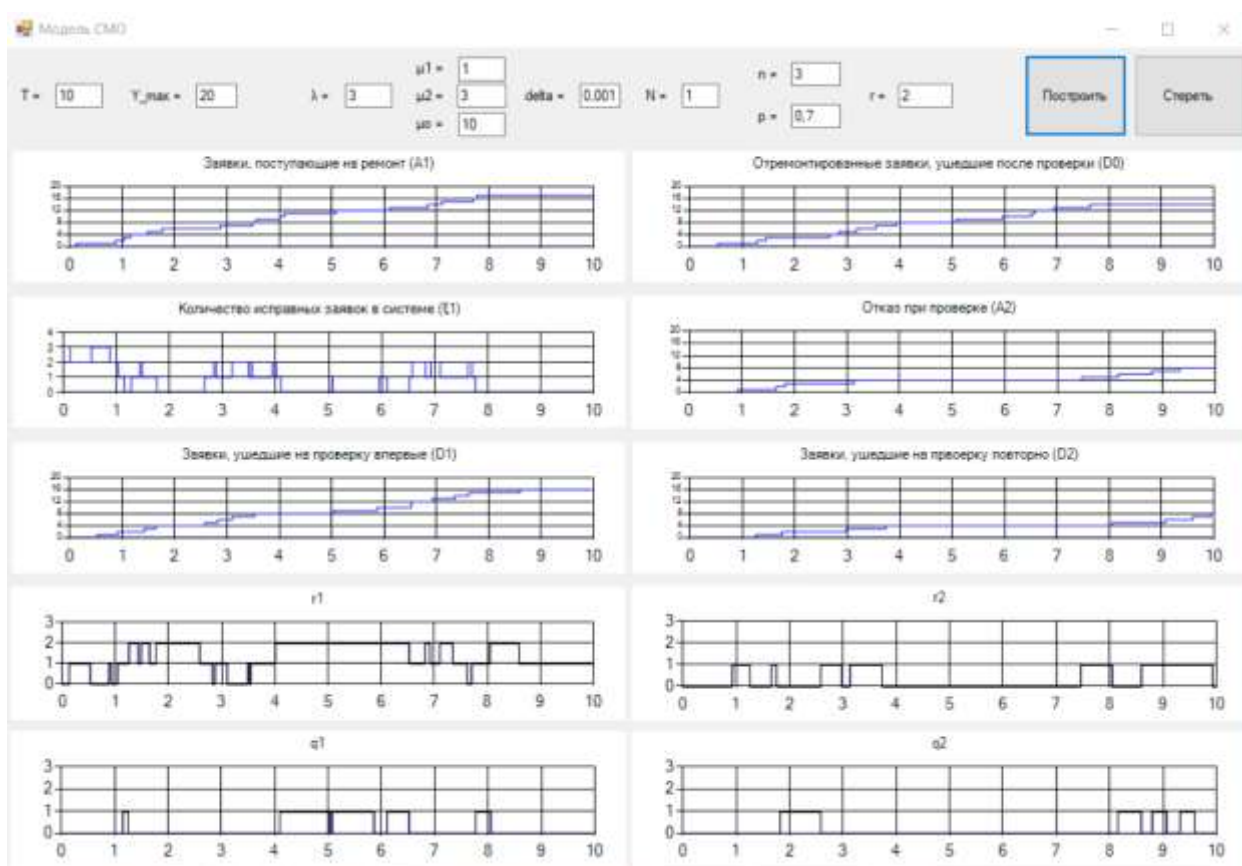


Рис. 2. Диалоговое окно программы

Список литературы

1. Yang T., Templeton J. G. C. A survey on retrial queues // *Queueing Systems*. 1987, no 2, p. 201-233.
2. Kim C.S., Klimenok V.I., Dudin A.N. Optimization of Guard Channel Policy in Cellular Mobile Networks with Account of Retrials // *Computers and Operation Research*. 2014, v. 43, p. 181–190.
3. Yamamuro K. The queue length in an M/G/1 batch arrival retrial queue // *Queueing Systems*. 2012, v. 70, p.187-205.

4. Shin Y. W., Moon D. H. M/M/c Retrial queue with multiclass of customers // *Methodology and Computing in Applied Probability*. 2014, v. 16, p. 931-949.
5. Cordeiro J.D., Kharoufeh J.P. The unreliable MM/ /1 retrial queue in a random environment // *Stochastic Models*. 2012, v. 28, p. 29–48.
6. Dudin A.N., Dudina O.S. Analysis of multiserver retrial queueing system with varying capacity and parameters // *Mathematical Problems in Engineering*. 2015, v. 2015, p. 1–12.
7. Klimenok V.I., Dudin A.N. Multi-dimensional asymptotically quasi-Toeplitz Markov chains and their application in queueing theory // *Queueing Systems*. 2006, v. 54, p. 245–259.
8. Choi B. D., Chang Y.. Single server retrial queues with priority calls // *Mathematical and Computer Modelling*. 1999, v.30, p.7-32.
9. Dimitriou I. A queueing model with two types of retrial customers and paired services // *Annals of Operations Research*. 2016, v. 238, p.123-143.
10. Dragieva V. I. A finite source retrial queue: number of retrials // *Communications in Statistics-Theory and Methods*. 2013, v. 42, p. 812-829.
11. Dudin A., Lee M., Dudina O., Lee S. Analysis of priority retrial queue with many types of customers and servers reservation as a model of cognitive radio system // *IEEE Transactions on Communications*. 2017, v. 65, p.186-199.
12. Задиранова Л.А. Исследование математических моделей потоков в бесконечнолинейных СМО с повторным обслуживанием требований : *Автореф. дис. ... канд. физ.-мат. наук / Томский государственный университет*. М., 2016.
13. Цициашвили Г.Ш. Пуассоновские потоки в системе с повторным обслуживанием // *Вестник Томского Государственного Университета*, № 4 (37), стр. 83-88.
14. Thomopoulos, N.T. One Server, Repeat Service (M/M/1/θ) // *In: Fundamentals of Queueing Systems*. Springer, Boston, MA, 2012.
15. Baruah M., Madan K.C., Eldabi T. Balking and re-service in a vacation queue with arrival and two types of heterogeneous service // *J. Math. Res*. 2012, v.4, p.1114–1124.
16. Kalita C.R., Choudhury G. A note on reliability analysis of an N-policy unreliable $M^X / \left(\begin{matrix} G_1 \\ G_2 \end{matrix} \right) / 1$ queue with optional repeated service // *RAIRO-Oper. Res*. 2018, v. 52, p. 713–724.
17. Бутов А. А., Раводин К. О. *Теория случайных процессов: учебно-методическое пособие*. Ульяновск: УлГУ, 2009.
18. Гнеденко Б.В., Коваленко И.Н. *Введение в теорию массового обслуживания*. М.:URSS, 2013. 400 с.
19. Рыжиков Ю. И., Уланов А. В. Применение гиперэкспоненциальной аппроксимации в задачах расчета немарковских систем массового обслуживания // *Вестник Томского государственного университета. Управление, вычислительная техника и информатика*. 2016, № 3(36), с. 60–65.

20. Gail H. R., Hantler S. L., Taylor B. A. Analysis of a non-preemptive priority multiserver queue // *Advances in applied prob.* 1988, v. 20, p. 852.
21. Miller D. R. Steady state algorithmic analysis of M/M/c two priority queues with heterogeneous rates // *Applied probability — computer science: the interface*. Boston: Birkhauser, 1982. V.2. P. 207–222.
22. Савинов Ю. Г., Сафиуллов И. Д., Дуняшина М. С. Модель обслуживания парка оборудования с возможной задержкой в начале ремонта // *Ученые записки УлГУ. Сер. Математика и информационные технологии*. 2020, № 1, с. 103-108.
23. Подгорнов М.Д., Савинов Ю.Г. Стратегии управления приоритетами в замкнутых системах массового обслуживания // *Ученые записки УлГУ. Сер. Математика и информационные технологии*. 2022, № 2, с. 64-74.

Model of a closed queuing system with repeated service

Savinov, Y. G.,* Lazareva, V.A.

* uras@aport.ru

Ulyanovsk State University, Russia

The paper develops a trajectory approach to the mathematical description and modeling of closed queuing systems. A model with repeated service is shown. The transition from a mathematical model to iterative formulas, which are used for simulation, is shown.

Keywords: *queuing system, simulation modeling, repeated service, point process, compensator*