

УДК 519.872

А.З. Меликов, Л.А. Пономаренко, И.А. Алиев

МАРКОВСКИЕ МОДЕЛИ СИСТЕМ ОБСЛУЖИВАНИЯ–ЗАПАСАНИЯ С РАЗНОТИПНЫМИ ПОВТОРНЫМИ ЗАЯВКАМИ

Ключевые слова: система обслуживания–запасания, повторные заявки, политика пополнения запасов, двумерная цепь Маркова, метод расчета.

Введение

В последние годы интенсивно изучаются модели систем обслуживания–запасания (Queuing-Inventory Systems, QIS) с разнотипными заявками [1–10]. Достаточно подробный обзор работ, посвященных таким QIS, можно найти в [10]. Большой интерес к изучению QIS с разнотипными заявками объясняется тем, что в реальных системах заявки отличаются друг от друга различными показателями, например объемом покупаемого запаса, важностью и т.д. В зависимости от специфической особенности в QIS применяются различные схемы введения приоритетов для входящих заявок. Наиболее часто используется следующая схема: вводится некоторый критический уровень запасов, и когда этот уровень запасов превышен, обслуживаются заявки всех типов, а если он ниже критического, то обслуживаются лишь приоритетные заявки. При этом критический уровень запасов определяется исходя из принятой политики пополнения запасов. Так, например, если в системе принята политика двух уровней (s, S) , то в качестве критического уровня запасов выбирается точка заказа s .

Настоящая работа идейно близка работе [9]. Здесь, как и в [9], изучается модель QIS с двумя типами заявок и мгновенным обслуживанием. Однако, в отличие от [9], в этой работе полагается, что в системе могут быть использованы три различные политики пополнения запасов. Кроме того, используемый здесь подход к исследованию рассматриваемых систем отличается от подхода, применяемого в [9]. Он позволяет разработать простые алгоритмы расчета и оптимизации характеристик подобных моделей.

Описание моделей и постановка задачи

Изучаемая система имеет склад размера $S, S < \infty$, для хранения определенных запасов. Для получения этих запасов поступают пуассоновские потоки заявок двух типов. При этом интенсивность обычных заявок (поток заявок первого типа) равна λ_1 , а интенсивность приоритетных заявок (поток заявок второго типа) — λ_2 . Считается, что время обслуживания заявок обоих типов равно нулю и разнотипные заявки являются идентичными по размеру требуемого запаса, т.е. после обслуживания заявки любого типа уровень запасов системы уменьшается на единицу.

© А.З. МЕЛИКОВ, Л.А. ПОНОМАРЕНКО, И.А. АЛИЕВ, 2019

*Международный научно-технический журнал
«Проблемы управления и информатики», 2019, № 4*

Приоритетное обслуживание заявок осуществляется согласно следующей схеме. Если в момент поступления заявки второго типа уровень запасов больше нуля, то она получает запас и покидает систему. Если в момент поступления заявки первого типа уровень запасов больше критического уровня s , $s < S/2$, то она также получает запас и покидает систему; если уровень запасов системы в момент поступления заявки первого типа меньше или равен s , эта заявка согласно схеме Бернулли либо с вероятностью α уходит в орбиту в целях повторения запроса для получения запаса, либо с дополнительной вероятностью $1 - \alpha$ не получает запас и покидает систему.

Здесь изучаются модели с бесконечным размером орбиты, т.е. любая заявка первого типа может приниматься в орбиту. Заявки с орбиты независимо друг от друга повторяют запросы через случайные моменты времени, которые имеют общее показательное распределение с параметром η , $0 < \eta < \infty$. При этом, если в момент поступления повторной заявки уровень запасов больше критического уровня s , она мгновенно получает требуемый запас и покидает орбиту; иначе (т.е. если уровень запасов системы в этот момент меньше или равен s) она согласно схеме Бернулли либо с вероятностью β уходит с орбиты, либо с дополнительной вероятностью $1 - \beta$ остается в орбите.

Считается, что заказы на пополнения запасов выполняются с определенными случайными задержками, которые имеют общее показательное распределение с параметром $\nu > 0$.

В системе может быть использована одна из следующих политик пополнения запасов:

- (s, S) -политика, т.е. заказ для поставки запасов делается тогда, когда их уровень опускается до величины s , $s < S/2$, при этом объем заказа равен $S - s$;
- политика переменного объема (Variable Order Size, VOS), т.е. заказ для поставки запасов делается тогда, когда их уровень опускается до величины s , $s < S/2$, при этом объем заказа равен $S - m$, где m — текущий уровень запасов;
- $(S - 1, S)$ -политика, т.е. заказ для поставки запасов делается после каждого акта отпуска запасов.

Задача заключается в нахождении совместного распределения уровня запасов системы и числа заявок в орбите. Требуется также найти усредненные характеристики системы: средний уровень запасов (S_{av}); среднюю интенсивность заказов (RR); вероятности отказа в обслуживании заявок каждого типа при поступлении в систему (PB_1, PB_2); среднее число заявок в орбите (L_o); средние интенсивности успешного (RSR) и безуспешного повторения ($RuSR$) запросов с орбиты.

Методы решения задачи

При использовании каждой из указанных выше политик пополнения запасов функционирование изучаемой системы описывается двумерной цепью Маркова с состояниями вида (m, n) , где m — уровень запасов системы, n — число заявок в орбите. Пространство состояний (ПС) этой цепи имеет бесконечную размерность и определяется как декартово произведение двух множеств, т.е. $E = \{0, 1, \dots, S\} \times \{0, 1, \dots\}$.

Сначала предположим, что в системе принята (s, S) -политика пополнения запасов. Рассмотрим задачу построения производящей матрицы данной двумер-

ной цепи Маркова. Переход из состояния $(m, n) \in E$ в состояние $(m', n') \in E$ обозначим $(m, n) \rightarrow (m', n')$. Тогда с учетом принятой политики пополнения запасов и системы приоритетов заключаем, что искомые интенсивности определяются так:

- интенсивность перехода $(m, n) \rightarrow (m-1, n)$ при выполнении условия $m > s$ равна λ , где $\lambda = \lambda_1 + \lambda_2$;
- интенсивность перехода $(m, n) \rightarrow (m-1, n)$ при выполнении условия $m \leq s$ — λ_2 ;
- интенсивность перехода $(m, n) \rightarrow (m, n+1)$ при выполнении условия $m \leq s$ — $\lambda_1 \alpha$;
- интенсивность перехода $(m, n) \rightarrow (m-1, n-1)$ при выполнении условий $m > s, n > 0$ равна $n\eta$;
- интенсивность перехода $(m, n) \rightarrow (m, n-1)$ при выполнении условий $m \leq s, n > 0$ равна $n\eta\beta$;
- интенсивность перехода $(m, n) \rightarrow (m+S-s, n)$ при выполнении условия $m \leq s$ — ν .

Стационарную вероятность состояния $(m, n) \in E$ обозначим $p(m, n)$ (условия существования стационарного режима устанавливаются ниже). Эти величины удовлетворяют следующей системе уравнений равновесия (СУР):

Случай $s < m \leq S$:

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

Случай $0 \leq m \leq s$:

$$(\lambda_2 I(m > 0) + \lambda_1 \alpha + m\eta\beta + \nu)p(m, n) = \lambda p(s+1, n)\delta(m, s) + \lambda_2 p(m+1, n) \times \\ \times I(m < s) + \lambda_1 \alpha p(m, n-1)I(n > 0) + (n+1)\eta\beta p(m, n+1). \quad (2)$$

Здесь и далее $I(A)$ — индикаторная функция события A , $\delta(i, j)$ — символы Кронекера.

К этой СУР (1), (2) добавляется условие нормировки:

$$\sum_{(m, n) \in E} p(m, n) = 1. \quad (3)$$

Искомые характеристики системы вычисляются с помощью стационарных вероятностей состояний. Так, средний уровень запасов системы определяется следующим образом:

$$S_{av} = \sum_{m=1}^S m \sum_{n=0}^{\infty} p(m, n). \quad (4)$$

Средняя интенсивность заказов определяется так:

$$RR = \lambda \sum_{n=0}^{\infty} p(s+1, n). \quad (5)$$

Заявка первого типа покидает систему без получения запасов, если в момент ее поступления уровень запасов системы меньше $s+1$ (тогда она теряется с вероятностью $1-\alpha$). Отсюда заключаем, что вероятность потери заявок первого типа при поступлении в систему вычисляется следующим образом:

$$PB_1 = (1-\alpha) \sum_{m=0}^s \sum_{n=0}^{\infty} p(m, n). \quad (6)$$

Заявки второго типа покидают систему без получения запасов лишь тогда, когда уровень запасов системы равен нулю, т.е. вероятность потери заявок второго типа при поступлении в систему определяется так:

$$PB_2 = \sum_{n=0}^{\infty} p(0, n). \quad (7)$$

Среднее число заявок в орбите находится следующим образом:

$$L_o = \sum_{m=0}^S \sum_{n=1}^{\infty} np(m, n). \quad (8)$$

Средние интенсивности успешного и безуспешного повторений запроса вычисляются так:

$$RSR = \eta \sum_{m=s+1}^S \sum_{n=1}^{\infty} np(m, n). \quad (9)$$

$$RuSR = \eta\beta \sum_{m=0}^s \sum_{n=1}^{\infty} np(m, n). \quad (10)$$

Следовательно, для нахождения характеристик системы (4)–(10) потребуется найти решение СУР (1)–(3), которая представляет собой бесконечную систему линейных алгебраических уравнений.

Из-за сложной структуры производящей матрицы не удастся разработать эффективный метод нахождения точного решения указанной СУР. Исходя из этого ниже предлагается приближенный метод нахождения стационарных вероятностей состояний данной двумерной цепи Маркова. Этот метод может быть корректно применен для моделей QIS, в которых большую долю входящего потока составляют заявки высокого приоритета, при этом их интенсивность существенным образом превышает интенсивности поступления повторных заявок с орбиты.

При выполнении этого допущения интенсивности переходов между состояниями внутри каждого класса $E_n = \{(m, n) : m = 0, 1, \dots, S\}$, $n = 0, 1, 2, \dots$, существенно превышают интенсивности переходов между различными классами E_n и $E_{n'}$, $n \neq n'$. Исходя из этого рассмотрим следующее разбиение исходного ПС:

$$E = \bigcup_{n=0}^{\infty} E_n, \quad E_m \cap E_{m'} = \emptyset, \quad m \neq m', \quad (11)$$

где $E_n = \{(m, n) \in E : m = 0, 1, \dots, S\}$, $n = 0, 1, \dots$.

Все состояния $(m, n) \in E_n$ объединяются в одно укрупненное состояние $\langle n \rangle$, и на исходном ПС определяется функция укрупнения $U : E \rightarrow \Omega$, где $U(m, n) = \langle n \rangle$, если $(m, n) \in E_n$. Обозначим $\Omega = \{\langle n \rangle : n = 0, 1, \dots\}$.

На основе разбиения (11) можно составить бесконечное число расщепленных моделей с ПС E_n , $n = 0, 1, \dots$, где в расщепленной модели с ПС E_n учитываются лишь переходы между состояниями, входящими в E_n .

Исходя из описанного выше алгоритма построения производящей матрицы исходной двумерной цепи Маркова заключаем, что во всех расщепленных моделях с ПС E_n интенсивности переходов между состояниями определяются идентичным образом и не зависят от индекса n , $n = 0, 1, \dots$. Обозначим интенсивность

перехода из состояния $(m, n) \in E_n$ в состояние $(m', n) \in E_n$ как $q_n(m, m')$. Тогда имеем

$$q_n(m, m') = \begin{cases} \lambda, & \text{àñëè } m > s, m' = m - 1, \\ \lambda_2, & \text{àñëè } 0 < m \leq s, m' = m - 1, \\ \nu, & \text{àñëè } 0 \leq m \leq s, m' = m + S - s, \\ 0 & \text{â ññòàëèèíüò ñëó-àÿò.} \end{cases} \quad (12)$$

Вероятность состояния $(m, n) \in E_n$ внутри расщепленной модели с ПС E_n обозначим $\rho_n(m)$. Поскольку эти вероятности не зависят от индекса n , $n = 0, 1, \dots$, ниже индекс n у этих величин $\rho_n(m)$ опускается. Тогда из (12) получаем, что указанные величины определяются так (см. [11]):

$$\rho(m) = \begin{cases} a_m \rho(s+1), & \text{àñëè } 0 \leq m \leq s, \\ \rho(s+1), & \text{àñëè } s+1 \leq m \leq S-s, \\ b_m \rho(s+1), & \text{àñëè } S-s+1 \leq m \leq S. \end{cases} \quad (13)$$

Здесь и далее приняты следующие обозначения:

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

Вероятность $\rho(s+1)$ находится из условия нормировки, т.е.

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

Исходя из описания исходной системы заключаем, что в пространстве укрупненных состояний Ω возможны лишь переходы $\langle n \rangle \rightarrow \langle n+1 \rangle$ и $\langle n \rangle \rightarrow \langle n-1 \rangle$. Интенсивность перехода $\langle x \rangle \rightarrow \langle y \rangle$ в пространстве Ω обозначим $q(\langle x \rangle, \langle y \rangle)$.

Переходы $\langle n \rangle \rightarrow \langle n+1 \rangle$ осуществляются при поступлении заявок первого типа с вероятностью α , если в этот момент система находится в одном из состояний типа $(m, n) \in E_n$, где $m \leq s$. Следовательно, интенсивность перехода $\langle n \rangle \rightarrow \langle n+1 \rangle$ вычисляется следующим образом:

$$q(\langle n \rangle, \langle n+1 \rangle) = \lambda_1 \alpha \sum_{m=0}^s \rho(m). \quad (14)$$

Переходы $\langle n \rangle \rightarrow \langle n-1 \rangle$ выполняются при поступлении повторных заявок с орбиты. При этом такой переход осуществляется с вероятностью единица, если в этот момент система находится в одном из состояний типа $(m, n) \in E_n$, где $m > s$; иначе, т.е. если в этот момент $m \geq s$, переход выполняется с вероятностью β . Таким образом, имеем

$$q(\langle n \rangle, \langle n-1 \rangle) = n\eta \left(\sum_{m=s+1}^S \rho(m) + \beta \sum_{m=0}^s \rho(m) \right). \quad (15)$$

Для упрощения дальнейшего изложения введем следующие обозначения:

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

Тогда из соотношений (14), (15) получаем, что вероятности состояний укрупненной модели с пространством состояний Ω в случае линейной интенсивности поступления повторных заявок с орбиты (т.е. если число повторных заявок в орбите равно n , то интенсивность их поступления равна $n\eta$) совпадают со стационарным распределением модели системы обслуживания $M/M/\infty$ с нагрузкой $\Psi = \Lambda/M$. Это означает, что при любых (положительных) значениях параметров системы существует стационарный режим и вероятности укрупненных состояний определяются так:

$$\pi(\langle n \rangle) = \frac{\Psi^n}{n!} e^{-\Psi}, \quad n = 0, 1, 2, \dots \quad (16)$$

Замечание 1. Если предположить, что интенсивность поступления повторных заявок является постоянной величиной (т.е. независимо от числа $n > 0$ повторных заявок в орбите интенсивность их поступления равна η), то для существования стационарного режима потребуется выполнение условия $\Psi < 1$ (это условие при $\beta = 1$ может быть заменено более простым и гарантированным условием $\lambda_1 < \eta$). В этом случае вероятности состояний укрупненной модели совпадают со стационарным распределением модели системы обслуживания $M/M/1/\infty$ с нагрузкой Ψ , т.е. при выполнении условия эргодичности $\Psi < 1$ стационарные вероятности укрупненных состояний определяются так:

$$\pi(\langle n \rangle) = \Psi^n (1 - \Psi), \quad n = 0, 1, 2, \dots \quad (17)$$

Используя (13) и (16) или (17), получаем, что вероятности состояний исходной двумерной цепи Маркова вычисляются (приближенно) следующим образом:

$$p(m, n) \approx \rho(m) \pi(\langle n \rangle), \quad (m, n) \in E. \quad (18)$$

После определенных математических преобразований с учетом (18) из (4)–(10) получаем следующие приближенные формулы для вычисления характеристик изучаемой системы при использовании (s, S) -политики пополнения запасов:

$$S_{av} = \sum_{m=1}^S m \rho(m); \quad (19)$$

$$RR = \lambda \rho(s+1); \quad (20)$$

$$PB_1 = (1 - \alpha) \sum_{m=0}^s \rho(m); \quad (21)$$

$$PB_2 = \rho(0); \quad (22)$$

$$L_o = \Psi; \quad (23)$$

$$RSR = \eta \Psi \sum_{m=s+1}^S \rho(m); \quad (24)$$

$$RuSR = \eta \beta \Psi \sum_{m=0}^s \rho(m). \quad (25)$$

Замечание 2. Формулы (23)–(25) определяют значения соответствующих характеристик при линейной интенсивности поступления повторных заявок; для случая постоянной интенсивности поступления повторных заявок в этих формулах величина Ψ заменяется величиной $\Psi/(1-\Psi)$. Из формул (19)–(22) видно, что средний уровень запасов системы, средняя интенсивность заказов и вероятности потери заявок не зависят от интенсивности поступления повторных заявок. Это объясняется тем, что при выводе этих приближенных формул принято допущение о том, что интенсивность поступления повторных заявок намного меньше, чем интенсивности поступления первичных заявок. Вместе с тем другие характеристики системы, а именно среднее число заявок в орбите, а также средние интенсивности успешного и безуспешного повторений, существенным образом зависят от всех параметров системы (см. (23)–(25)).

Предложенный выше подход может быть применен и для изучения данной системы при использовании других политик пополнения запасов.

Предположим, что в системе используется VOS-политика. Тогда элементы производящей матрицы соответствующей двумерной цепи Маркова также определяются согласно алгоритму, описанному выше, но единственное отличие состоит в том, что здесь в случаях $m \leq s$ в момент поступления заказа происходит переход $(m, n) \rightarrow (S, n)$ с интенсивностью ν .

В отличие от (s, S) -политики, объем поставляемого заказа является переменной величиной, и потому необходимо ввести новую характеристику — средний объем заказа, V_{av} . Эта величина определяется так:

$$V_{av} = \sum_{m=S-s}^S m \sum_{n=0}^{\infty} p(S-m, n). \quad (26)$$

Замечание 3. Здесь и далее для упрощения изложения для всех параметров и характеристик системы при использовании различных политик пополнения запасов сохраняются одинаковые обозначения.

Другие характеристики системы при использовании VOS-политики вычисляются из (4)–(10). Точные значения этих характеристик также определяются через стационарные вероятности состояний системы, которые удовлетворяют соответствующей СУР. Для вычисления приближенных значений вероятностей состояний и характеристик системы может быть использован предложенный выше метод.

Опуская подробное изложение этапов указанного метода, отметим лишь отличающиеся моменты. Так, в данной модели во всех расщепленных моделях с ПС E_n интенсивности переходов между состояниями определяются аналогично (12), но с единственным отличием: в третьей строке правой части формулы (12) условие $0 \leq m \leq s$, $m' = m + S - s$ заменяется $0 \leq m \leq s$, $m' = S$. С учетом этого находим, что в данной модели вероятности состояний внутри расщепленной модели с ПС E_n определяются так (см. [12]):

$$\rho(m) = \begin{cases} c_m \rho(0), & \text{если } 1 \leq m \leq s+1, \\ d \rho(0), & \text{если } s+1 < m \leq S, \end{cases} \quad (27)$$

где

$$c_m = \begin{cases} \frac{\nu}{\lambda_2} \left(1 + \frac{\nu}{\lambda_2}\right)^{m-1}, & 1 \leq m \leq s, \\ \frac{\nu}{\lambda_2} \left(1 + \frac{\nu}{\lambda_2}\right)^{s-1} \frac{\nu + \lambda_2}{\lambda}, & m = s+1; \end{cases}$$

$$d = \frac{v}{\lambda_2} \left(1 + \frac{v}{\lambda_2} \right)^{s-1} \frac{v + \lambda_2}{\lambda}.$$

Вероятность $\rho(0)$ находится из условия нормировки, т.е.

$$\rho(0) = \left(1 + (S - s - 1)d + \sum_{m=1}^{s+1} c_m \right)^{-1}.$$

Далее с помощью формул (14)–(18) и (27) определяются вероятности состояний модели при использовании VOS-политики, а характеристики системы вычисляются из (19)–(25). Новая характеристика (26) приближенно вычисляется так:

$$V_{av} \approx \sum_{m=S-s}^S m \rho(S-m). \quad (28)$$

Теперь предположим, что в системе используется $(S-1, S)$ -политика пополнения запасов. Элементы производящей матрицы соответствующей цепи Маркова также определяются согласно алгоритму, описанному выше, но единственное отличие состоит в том, что в случаях $0 \leq m < S$ в моменты поступления заказа происходит переход $(m, n) \rightarrow (m+1, n)$ с интенсивностью $(S-m)v$.

Характеристики системы при использовании $(S-1, S)$ -политики вычисляются из (4), (6)–(10). Однако при использовании данной политики средняя интенсивность заказов вычисляется по следующей формуле:

$$RR = \lambda \sum_{m=s+1}^S \sum_{n=0}^{\infty} p(m, n) + \lambda_2 \sum_{m=1}^s \sum_{n=0}^{\infty} p(m, n). \quad (29)$$

В данной модели во всех расщепленных моделях с ПС E_n интенсивности переходов между состояниями определяются аналогично (12), но с единственным отличием: в третьей строке правой части формулы (12) условие $0 \leq m \leq s$, $m' = m + S - s$ заменяется $0 \leq m < S$, $m' = m + 1$, при этом интенсивность перехода $(m, n) \rightarrow (m+1, n)$ равна $(S-m)v$.

С учетом этого находим, что в данной модели вероятности состояний внутри расщепленной модели с ПС E_n определяются так (см. [13]):

$$\rho(m) = \begin{cases} \frac{S!}{(S-m)!} \left(\frac{v}{\lambda_2} \right)^m \rho(0), & \text{ãñèè } 0 \leq m \leq s, \\ \left(\frac{\lambda}{\lambda_2} \right)^s \frac{S! v^m}{(S-m)!} \left(\frac{v}{\lambda} \right)^m \rho(0), & \text{ãñèè } s+1 \leq m \leq S, \end{cases} \quad (30)$$

где вероятность $\rho(0)$ находится из условия нормировки, т.е.

$$\rho(0) = \left(\sum_{m=0}^s \frac{S!}{(S-m)!} \left(\frac{v}{\lambda_2} \right)^m + \left(\frac{\lambda}{\lambda_2} \right)^s \sum_{m=s+1}^S \frac{S!}{(S-m)!} \left(\frac{v}{\lambda} \right)^m \right)^{-1}.$$

Окончательно из (14)–(18) и (30) определяются вероятности состояний модели при использовании $(S-1, S)$ -политики и характеристики системы вычисляются из (19), (21)–(25). С учетом (30) характеристика (5) в данной модели приближенно вычисляется так:

$$RR \approx \lambda \sum_{m=s+1}^S \rho(m) + \lambda_2 \sum_{m=1}^s \rho(m). \quad (31)$$

Численные результаты

Рассмотрим результаты численных экспериментов, выполненных с помощью разработанных выше алгоритмов. При этом основными целями экспериментов являются, во-первых, сравнительный анализ характеристик изучаемой системы при использовании различных политик пополнения запасов и, во-вторых, нахождение оптимального значения критического уровня запасов при использовании каждой политики.

Сразу же отметим, что, как и в работах [10–13], здесь нами также изучены вопросы точности разработанных приближенных формул вычисления стационарных распределений соответствующих двумерных цепей Маркова и характеристик системы. При этом использованы различные меры близости (подробнее об этих мерах см. [10–13]). Однако из-за ограниченности объема работы соответствующие результаты здесь не приводятся. Укажем только, что проведенные эксперименты показали высокую точность приближенных алгоритмов, при этом время их выполнения практически равно нулю.

Предположим, что нагрузочные параметры системы, а также размер склада системы являются постоянными величинами, и единственный регулируемый параметр — критический уровень запасов, т.е. s . Следовательно, приведенный ниже сравнительный анализ характеристик изучаемой системы при использовании различных политик пополнения запасов осуществляется при изменении указанного параметра в отрезке $[0, \lceil S/2 \rceil]$, где $\lceil x \rceil$ — целая часть x .

Часть результатов проводимых численных экспериментов показана на рис. 1, где символы \circ , \diamond и \times указывают политики (s, S) , VOS и $(S-1, S)$ соответственно. Исходные данные выбирались так: $S = 120$, $\lambda_1 = 50$, $\lambda_2 = 200$, $\nu = 20$, $\eta = 3$, $\alpha = 0,3$, $\beta = 0,6$.

Из рис. 1, а видно, что характер и значения функции S_{av} , при использовании различных политик пополнения запасов существенным образом отличаются друг от друга. Так, при использовании (s, S) -политики функция выпуклая, при VOS-политике — вогнутая, а при использовании $(S-1, S)$ -политики — постоянная. Функция RR принимает близкие значения при использовании политик (s, S) и VOS, при этом при использовании политики VOS она является постоянной и равна 2,45, а при использовании политики (s, S) она растет с малой скоростью от 2 до 5 (рис. 1, б); при использовании политики $(S-1, S)$ эта функция является постоянной и равна 250 (поскольку значения данной функции при использовании политики $(S-1, S)$ существенно больше, чем при использовании остальных политик, на графике эта линия не показана). Такая существенная разница в значениях этой функции при использовании политики $(S-1, S)$ объясняется тем, что в этой политике заказ на пополнение осуществляется после каждого акта обслуживания.

При использовании политики (s, S) функция PB_1 является возрастающей, в то время как для политики VOS она становится постоянной величиной (рис. 1, в); функция PB_2 является убывающей при использовании обеих политик пополнения запасов, при этом ее значения существенно меньше при использовании политики (s, S) (рис. 1, г). Отметим, что при использовании политики $(S-1, S)$ значения этой функции являются бесконечно малыми величинами (практически равны нулю), потому ее график при данной политике не показан на рис. 1, в, г) (далее графики соответствующих характеристик при использовании этой политики не показаны).

Это объясняется тем, что для выбранных исходных данных в большом диапазоне изменения параметра s многие вероятности состояний $\rho(m)$ при использовании этой политики практически равны нулю, и лишь при нескольких значениях этого параметра, близких к максимально возможному значению $\lceil S/2 \rceil$, некоторые вероятности состояний принимают достаточно большие значения (см. (30)).

Функция L_0 является возрастающей при использовании политик (s, S) и VOS, при этом в обеих политиках ее значения схожи; здесь также эта функция практически равна нулю при использовании политики $(S-1, S)$ (рис. 1, *д*).

При малых значениях параметра s для функции RSR предпочтительна политика VOS, однако при $s > 35$ благоприятна политика (s, S) (рис. 1, *е*). В отличие от функции RSR , функция $RuSR$ является возрастающей при использовании обеих политик — (s, S) и VOS, при этом здесь также при $s > 40$ предпочтительна политика (s, S) (рис. 1, *ж*).

Теперь рассмотрим задачу нахождения оптимального значения критического уровня запасов при использовании каждой политики. Отметим, что возможны различные постановки этой задачи. Для определенности изложения здесь рассматривается задача максимизации прибыли системы за счет выбора надлежащих значений параметра s .

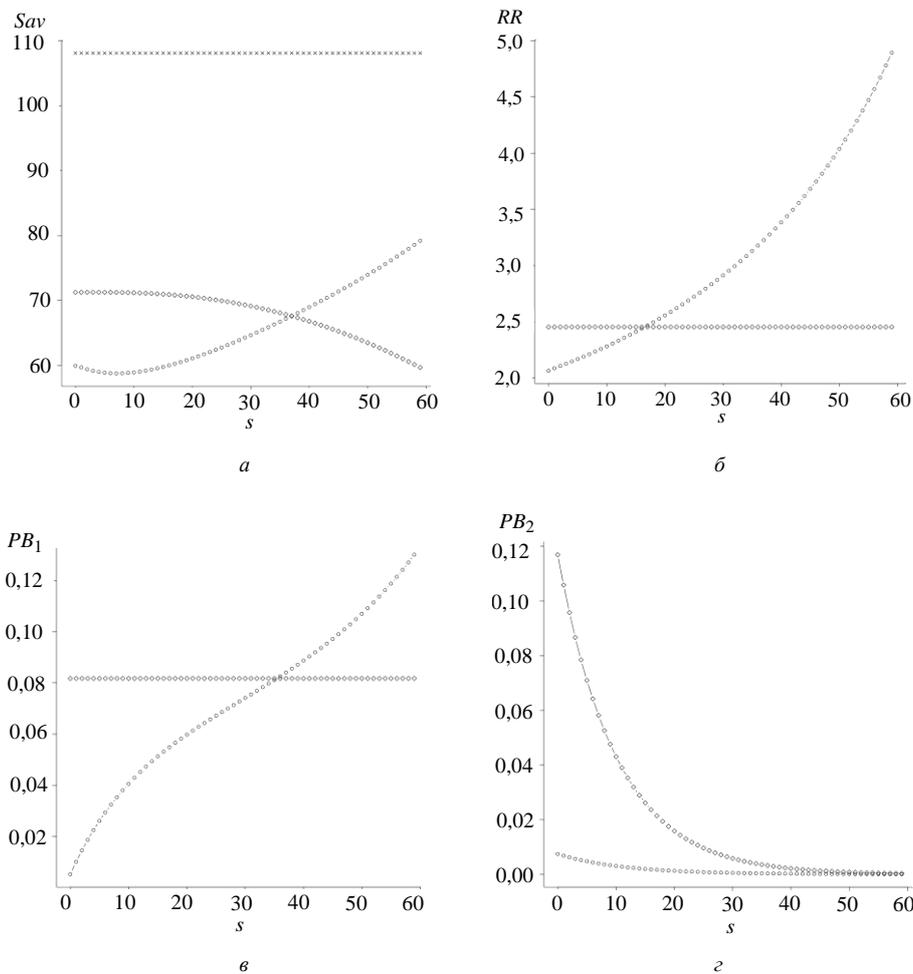
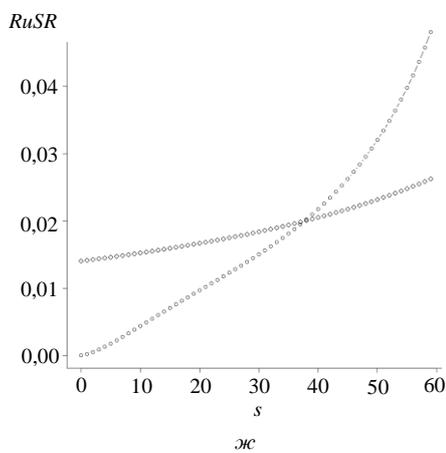
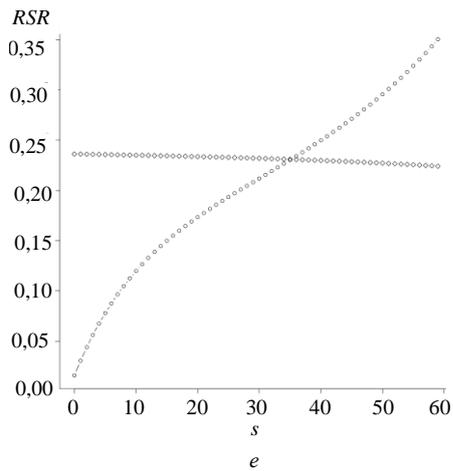
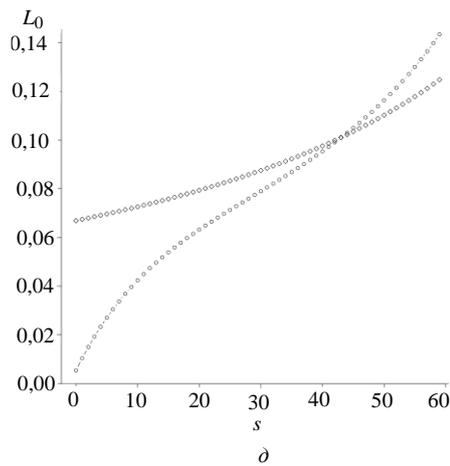


Рис. 1



Доходы системы (RV) формируются из продажи запасов и определяются так:

$$RV(s) = (\lambda_1(1 - PB_1(s)) + RSR(s)) \times \\ \times C_{rev}^1 + \lambda_2(1 - PB_2(s))C_{rev}^2, \quad (32)$$

где C_{rev}^k — доходы системы от продажи единицы запаса по заявкам k -го типа, $k = 1, 2$; при этом следует ожидать, что доходы от обслуживания приоритетных заявок намного больше, чем от обслуживания обычных, т.е. $C_{rev}^2 \gg C_{rev}^1$.

Замечание 4. Здесь и далее в обозначениях характеристик системы и функций в скобках указан оптимизируемый параметр s .

Суммарные штрафы (TC) в системе определяются следующим образом:

$$TC(s) = (K + c_r V_{av}(s))RR(s) + c_h S_{av}(s) + c_l^1(\lambda_1 PB_1(s) + RuSR(s)) + c_l^2 \lambda_2 PB_2(s), \quad (33)$$

где K — фиксированная цена одного заказа; c_r — цена единицы объема заказа; c_h — цена хранения единицы объема запасов за единицу времени; c_l^k — штраф за потерю одной заявки k -го типа, $k = 1, 2$. Считается, что штрафы за потерю приоритетных заявок намного больше, чем за потерю обычных, т.е. $c_l^2 \gg c_l^1$.

Замечание 5. При использовании (s, S) -политики в (33) полагается $V_{av} = S - s$, а при использовании $(S - 1, S)$ -политики имеем $V_{av} = 1$. Аналогичным образом при использовании $(S - 1, S)$ -политики функция $RR(s)$ вычисляется с помощью формулы (29) или (31), если используется приближенный подход.

Задача оптимизации ставится следующим образом: для каждой политики пополнения запасов требуется максимизировать прибыль системы за счет выбора надлежащих значений критического уровня запасов, т.е. из (32) и (33) находим, что требуется решить следующую задачу:

$$s^* = \arg \max_s PT(s), \quad (34)$$

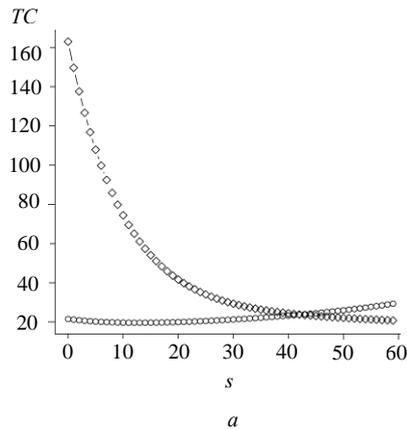
где $PT(s) = RV(s) - TC(s)$.

Задача (34) всегда имеет решение, так как множество допустимых решений $\{0 \leq s \leq \lceil S/2 \rceil\}$ является конечным и дискретным.

Коэффициенты в функционалах (32) и (33) определяются следующим образом:

$$C_{rev}^1 = 5, C_{rev}^2 = 10, K = 0,2, c_r = 0,01, c_h = 0,2, c_l^1 = 2, c_l^2 = 6.$$

Для наглядности на рис. 2 приводятся графики всех трех функций TC , RV и PT . Из этих графиков видно, что для выбранных исходных данных



функция TC возрастает при использовании политики (s, S) , в то время как функция RV убывает при этой политике; обратная картина наблюдается для этих функций при использовании политики VOS (рис. 2, а, б). Поскольку функция PT является возрастающей при использовании политики VOS, а при использовании политики (s, S) она убывает, оптимальное решение задачи (34) при политике VOS равно $s^* = 0$, а при политике (s, S) имеем $s^* = 59$ (рис. 2, в).

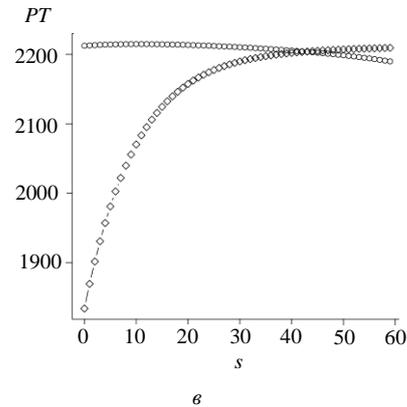
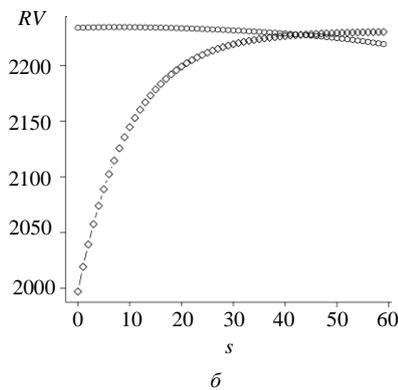


Рис. 2

Заключение

В работе предложены марковские модели системы управления запасами с мгновенным обслуживанием и двумя типами заявок. Заявки высокого приоритета получают запас и покидают систему, если в момент их поступления уровень запасов выше нуля; иначе они покидают систему без получения запаса. Заявки низкого приоритета согласно схеме Бернулли могут либо уходить в орбиту, либо покидать систему, если в моменты их поступления уровень запасов системы ниже определенного критического уровня. При этом, если в моменты поступления заявки с орбиты уровень запасов опять ниже критического уровня, то повторная заявка также согласно схеме Бернулли либо окончательно уходит с орбиты, либо остается там для повторения.

Изучаются системы, в которых возможно использование трех политик пополнения запасов: (s, S) -, $(S-1, S)$ - и VOS-политики. Во всех системах время выполнения заказов имеет показательную функцию распределения с конечным средним значением. Построены соответствующие двумерные цепи Маркова, которые описывают работу изучаемых систем, и разработаны точный и приближенный методы расчета вероятностей состояний этих цепей. Разработаны явные формулы для приближенного вычисления характеристик изучаемых систем, на их основе выполнены численные эксперименты и решена задача максимизации прибыли изучаемых систем при использовании различных политик пополнения запасов.

A.Z. Melikov, L.A. Ponomarenko, I.A. Aliev

МАРКОВСЬКІ МОДЕЛІ СИСТЕМ ОБСЛУГОВУВАННЯ–ЗАПАСАННЯ ІЗ РІЗНОТИПНИМИ ПОВТОРНИМИ ВИМОГАМИ

Вивчається модель системи обслуговування–запасання із двома типами вимог і миттєвим обслуговуванням. Вважається, якщо в момент надходження вимоги високого пріоритету рівень запасів більший від нуля, то вона одержує запас і покидає систему. Вимога низького пріоритету отримує запас, якщо в момент її надходження рівень запасів є більшим певного критичного рівня; в іншому випадку ця вимога згідно зі схемою Бернуллі або надходить в орбіту, або не отримує запас і залишає систему. Час перебування вимог у необмеженій орбіті є випадковою величиною із показниковою функцією розподілу. Якщо в момент надходження повторної вимоги рівень запасів більший критичного, то вона миттєво отримує необхідний запас і покидає орбіту; інакше вона згідно зі схемою Бернуллі або покидає орбіту, або залишається на орбіті. Розглядається три політики поповнення запасів: політика двох рівнів, політика змінного розміру поповнення і політика, згідно з якою замовлення на поставку запасів роблять після кожного факту видачі запасів. Основними характеристиками системи є середній рівень запасів, середня інтенсивність замовлень, ймовірності відмови в обслуговуванні вимог кожного типу при надходженні в систему, середня кількість вимог в орбіті, середні інтенсивності успішного та невдалого повторення запитів з орбіти. Для математичного аналізу системи, яка вивчається, побудовано відповідний двовимірний ланцюг Маркова та подано алгоритм пошуку його твірної матриці. Знайдено сумісний розподіл рівня запасів системи та кількість вимог в орбіті, а також запропоновано формули для обчислення усереднених характеристик моделей, які вивчались.

Ключові слова: системи обслуговування–запасання, марковська модель, повторна заявка, політика поповнення запасів.

A.Z. Melikov, L.A. Ponomarenko, I.A. Aliev

MARKOV MODELS OF QUEUING–INVENTORY SYSTEMS WITH DIFFERENT TYPES OF RETRIAL CUSTOMERS

In this paper, models of queuing-inventory systems with two kinds of retrial customers and instantaneous service time are considered. It is assumed that if at the time of arrival of a high-priority customer the inventory level is greater than zero then it receives an inventory and leaves the system. Customers of low priority receive inventory if at the time of its arrival the inventory level is above a certain critical level; otherwise, this customer, according to the Bernoulli trials, either goes into orbit or does not receive an inventory and leaves the system. The sojourn time of customers in an infinite orbit is a random variable with an exponential distribution function. If at the time of receipt of the repeated customer the inventory level is more than a critical

level, then it instantly receives the required inventory and leaves the orbit; otherwise, according to the Bernoulli scheme, it either leaves the orbit or remains in orbit. Three replenishment policies are considered — a two-level policy, a variable order size policy, and a policy in which an order is made to supply inventory after each inventory release act. The main characteristics of the system are the average inventory level, the average intensity of orders, the probability of failure in servicing customers of each type when entering the system, the average number of customers in orbit, the average intensities of successful and unsuccessful repetition of customers from orbit. For the mathematical analysis of the system under study, a corresponding two-dimensional Markov chain was constructed and an algorithm was given for finding its generating matrix. Joint distribution of the system's inventory level and the number of customers in orbit as well as formulas for calculating the averaged characteristics of the studied models are developed.

Keywords: queuing-inventory systems, Markov model, retrial customers, replenishment policies.

1. Dekker R., Hill R.M., Kleijn M.J. On the (S–1, S) lost sales inventory model with priority demand classes. *Naval Research Logistics*. 2002. **49**, N 6. P. 593–610.
2. Kranenburg A.A., van Houtum G.J. Cost optimization in the (S–1, S) lost sales inventory model with multiple demand classes. *Operations Research Letters*. 2007. **35**, N 4. P. 493–502.
3. Zhao N., Lian Z. A queuing-inventory system with two classes of customers. *Int. J. Production Economics*. 2011. **129**. P. 225–231.
4. Isotupa K.P.S. Cost analysis of an (S–1, S) inventory system with two demand classes and rationing. *Annals of Operations Research*. 2015. **233**. P. 411–421.
5. Isotupa K.P.S. An (S, Q) inventory system with two demand classes of customers. *International Journal of Operational Research*. 2011. **12**, N 1. P. 12–19.
6. Isotupa K.P.S. An (S, Q) Markovian inventory system with lost sales and two demand classes. *Mathematical and Computer Modeling*. 2006. **43**. P. 687–694.
7. Isotupa K.P.S., Samanta S.K. A continuous review (s, Q) inventory system with priority customers and arbitrary distributed lead times. *Mathematical and Computer Modeling*. 2013. **57**. P. 1259–1269.
8. Sivakumar B., Arivarignan G. A modified lost sales inventory system with two types of customers. *Quality Technology and Quantitative Management*. 2008. **5**, N 4. P. 339–349.
9. Karthick T., Sivakumar B., Arivarignan G. An inventory system with two types of customers and retrial demands. *International Journal of Systems Science: Operations & Logistics*. 2015. **2**, N 2. P. 90–112.
10. Melikov A.Z., Ponomarenko L.A., Aliyev I.A. Markov models of systems with two types of customers and different replenishment policies. *Cybernetics and Systems Analysis*. 2018. **54**, N 6. P. 900–917.
11. Melikov A.Z., Ponomarenko L.A., Bagirova S.A. Analysis of queueing-inventory systems with impatience consume customers. *Journal of Automation and Information Sciences*. 2016. **48**, N 1. P. 53–68.
12. Melikov A.Z., Ponomarenko L.A., Bagirova S.A. Markov models of queueing-inventory systems with variable order size. *Cybernetics and Systems Analysis*. 2017. **53**, N 3. P. 373–386.
13. Melikov A.Z., Shahmaliyev M.O. A perishable queueing-inventory system with positive service time and (S–1, S) replenishment policy. *Communications in Computer and Information Sciences*. Springer. 2017. **800**. P. 83–96.

Получено 08.04.2019