



Ссылка на статью:

// Ученые записки УлГУ. Сер. Математика и информационные технологии. УлГУ. Электрон. журн. 2020, № 1, с. 95-102.

Поступила: 24.05.2020

Окончательный вариант: 08.06.2020

© УлГУ

УДК 519.218.5

Стратегия адаптивного разделения ресурсов в мультисервисных сетях

Савинов Ю.Г.^{*}, Рослов М.Э., Куманина Я.А.

^{*}uras@aport.ru

УлГУ, Ульяновск, Россия

В работе представлена математическая и компьютерная модель, реализующая стратегию адаптивного разделения ресурсов в мультисервисной сети. Математическая модель построена в терминах точечных процессов и их компенсаторов. Показан переход от математической модели к итерационным формулам, по которым проводится имитационное моделирование.

Ключевые слова: система массового обслуживания, стратегия подвижной границы, мультисервисные сети, точечный процесс, компенсатор, имитационное моделирование.

Введение

Мультисервисные сети используются для одновременной передачи обычного трафика (пакеты данных) и трафика другой информации (речи, видео и др.). За последние годы мультисервисные сети обрели большую популярность [1]. В таких сетях особое значение имеет анализ параметров качества обслуживания: надежность доставки, задержки и скорость передачи [2]. При этом, например, речь и видеоконференции критичны к задержкам, а пакеты данных критичны к надежности доставки. Впервые необходимость интеграции таких видов нагрузки встретилась при разработке сетей ISDN [3]. Эта же проблема стояла и для сетей GSM/GPRS, а также актуальна для современных сетей связи. Возникает задача управления ресурсом (полосы пропускания) для предоставления каждому из видов трафика. В книге [2] проанализированы три стратегии: стратегия на основе обслуживания в порядке поступления, стратегия на основе абсолютного приоритета и стратегия подвижной границы. Аналитическими методами (методом производящих функций) при определенных ограничениях получены приближенные формулы для определения задержек пакетов в сети, которые говорят про существенный выигрыш стратегии подвижной границы в

характеристиках качества обслуживания. Другие модели с различными адаптивными стратегиями разделения ресурса можно найти в работах [4-6] и ссылках в них. Цель данной работы – построить модель сети с адаптивным разделением ресурса и методами имитационного моделирования оценить характеристики качества обслуживания для стратегии подвижной границы, для случаев, когда не выполняются ограничения, позволяющие использовать аналитическое решение из [2].

1. Постановка задачи и описание стратегии

Рассмотрим два класса нагрузок – обслуживание с коммутацией каналов (звонки, видеозвонки) и нагрузка, требующая обслуживания с коммутацией пакетов – данные. Назовем первую из этих видов нагрузка 1-го класса. Будем считать, что интенсивность поступлений ВЫЗОВОВ (требований на соединение) $\lambda_1 \geq 0$, а среднее время занятия $1/\mu_1 > 0$. Этот тип нагрузки требует одного временного канала и при отсутствии свободных каналов – блокируется [2]. Второй из видов нагрузки назовем нагрузкой 2-го класса. Интенсивность поступлений ПАКЕТОВ положим равной $\lambda_2 \geq 0$, а средняя длина пакета определяет время обслуживания $1/\mu_2 > 0$. Эта группа заявок ставится в очередь при отсутствии свободных каналов.

Предположения и допущения:

1. поступающие в систему заявки 1 и 2 класса образуют простейшие потоки [7-9] с интенсивностями $\lambda_1 \geq 0, \lambda_2 \geq 0$;
2. длительность обслуживания заявок 1 и 2 класса распределена по экспоненциальному закону с интенсивностями $\mu_1 \geq 0, \mu_2 \geq 0$;
3. дисциплина обслуживания - в порядке поступления: заявка, поступившая в систему, принимается на обслуживание, если есть хотя бы один свободный прибор. Если заявка застала свободными несколько приборов, то она направляется в один из них случайным образом [8].

Опишем стратегию подвижной границы. При этом методе интеграции общий ресурс из n каналов делится на две части. Одна часть, содержащая n_1 каналов, предназначается для обслуживания нагрузки первого класса (запросов на соединение). Другая часть, содержащая $n_2 = n - n_1$ каналов, резервируется для пакетов – обслуживания нагрузки второго класса. Пакеты могут занимать также любой из n_1 каналов, отведенных для обслуживания нагрузки первого класса, если он не используется в данный момент времени. Однако при поступлении заявки первого класса она имеет абсолютный приоритет перед нагрузкой второго класса и сбрасывает при необходимости пакет, занимающий один из n_1 каналов. Ниже приведена иллюстрация метода (см. рис. 1).

Если в системе все n_1 каналов, предназначенных для обслуживания нагрузки первого класса, заняты обслуживанием заявок 1 класса, то новая заявка 1 класса получает отказ и покидает систему не обслуженной. Застав все приборы занятыми, заявка 2 класса становится в очередь и ждет до тех пор, пока одно из устройств не освободиться.

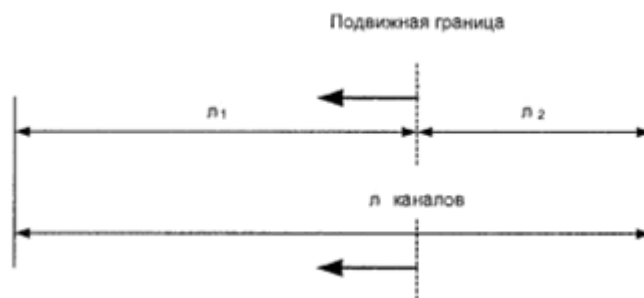


Рис. 1. Схема метода подвижной границы, адаптировано из [2].

2. Математическая модель

Аналитическое исследование данной модели возможно только при существенных ограничениях и предположениях. Поэтому основным способом исследования выбрано имитационное моделирование. Для имитационного моделирования наиболее подходит не марковское описание из [2], а семимартингальное (траекторное) описание систем массового обслуживания (СМО) (подробнее см. работы [10-15] и ссылки в них). Модель в терминах считающих процессов [16] позволяет легко переходить от математической модели к итерационным формулам, по которым проводится имитационное моделирование. Выбор в пользу семимартингального описания подкрепляется еще и тем фактом, что любой случайный процесс в дискретном времени является семимартингалом [16]. Поэтому в данной работе выбрано семимартингальное описание СМО в терминах точечных (считающих) процессов и их компенсаторов [16].

Введем считающие (точечные) процессы [16]:

$A^i = (A_t^i)_{t \geq 0}$ - число заявок i -го класса ($i=1,2$), поступивших в систему за время $t \geq 0$, $A_0^i = 0$,

$D^i = (D_t^i)_{t \geq 0}$ - число обслуженных заявок i -го класса ($i=1,2$) за время $t \geq 0$, $D_0^i = 0$,

$F^1 = (F_t^1)_{t \geq 0}$ - число заявок 1-го класса, получивших отказ в обслуживании за время $t \geq 0$, $F_0^1 = 0$.

Рассмотрим многоканальную СМО, реализующую, в первом приближении, стратегию подвижной границы (см. рис. 2).

Для ξ_t^i - числа заявок i -го класса ($i=1,2$) в СМО в момент времени $t \geq 0$ можно написать следующие основные балансовые соотношения

$$\xi_t^1 = \xi_0^1 + A_t^1 - D_t^1 - F_t^1, \quad (1)$$

$$\xi_t^2 = \xi_0^2 + A_t^2 - D_t^2, \quad (2)$$

где $\xi_0^i \geq 0$ - число заявок i -го класса ($i=1,2$) в СМО при $t = 0$ и $A_0^i = D_0^i = F_0^i = 0$.

Точечные процессы A^1, A^2, D^1, D^2, F^1 определяются своими компенсаторами в соответствии с разложением Дуба-Мейера для субмартингалов [16]. Компенсатор для процесса $A^i = (A_t^i)_{t \geq 0}$, ($i = 1,2$) определяется соотношениями (3)-(4):

$$A_t^i = \widetilde{A}_t^i + m_t^{A^i}, \quad (3)$$

$$\widetilde{A}_t^i = \lambda_i \cdot t, \quad \lambda_i > 0, \quad (4)$$

где \widetilde{A}_t^i ($i=1,2$) - неубывающие предсказуемые процессы, m^{A^i} и m^{D^i} – мартингалы [16].

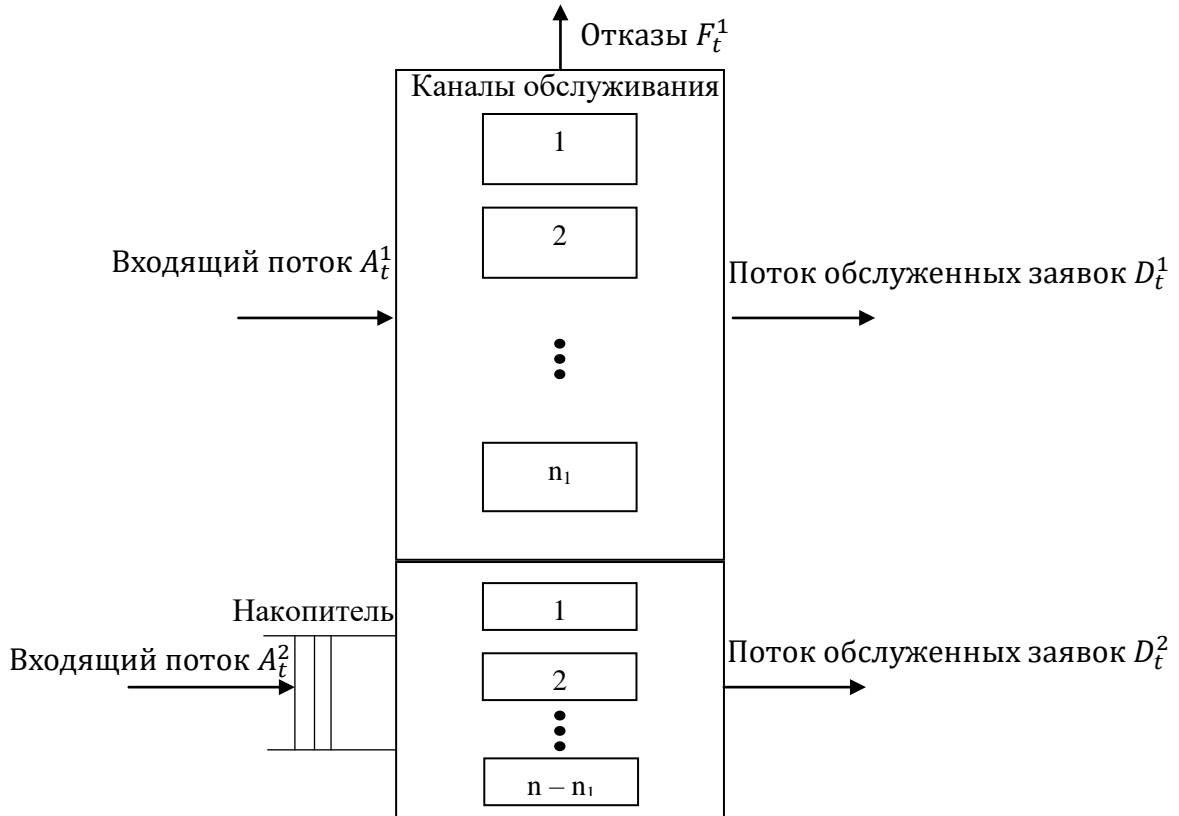


Рис.2. Схема СМО, реализующей стратегию подвижной границы.

Мгновенная интенсивность обслуживания заявок первого типа определяется числом заявок первого типа, находящихся в СМО, поскольку в СМО отсутствует очередь для заявок первого типа, и, следовательно, все заявки первого типа, находящиеся в СМО, обслуживаются. Поэтому, компенсатор для процесса $D^1 = (D_t^1)_{t \geq 0}$ определяется следующим соотношением:

$$\widetilde{D}_t^1 = \int_0^t \mu_1 \cdot \xi_s^1 \cdot I(n_1 \geq \xi_s^1 > 0) ds. \quad (5)$$

Компенсатор для процесса $D^2 = (D_t^2)_{t \geq 0}$ определяется следующим соотношением:

$$\widetilde{D}_t^2 = \int_0^t \mu_2 \cdot \beta_s ds, \quad (6)$$

$$\text{где } \beta_s = \begin{cases} (n - \xi_s^1), & \text{если } \xi_s^1 + \xi_s^2 > n, \\ \xi_s^2, & \text{если } \xi_s^1 + \xi_s^2 \leq n, \end{cases}$$

то есть интенсивность обслуживания заявок второго типа, определяется числом заявок первого и второго типа. Заявки 2 типа обслуживаются, если в момент ее поступления есть свободные приборы, если свободных приборов нет, но есть заявки 1 типа на обслужива-

нии, то заявка 2 типа ожидает окончания обслуживания этой заявки и после этого встает на обслуживание.

Число заявок 1 типа, получивших отказ в обслуживании за время t , определяется по формуле:

$$F_t^1 = \int_0^t I(\xi_s^1 = n_1) dA_s^1, \quad (7)$$

то есть заявка 1 типа теряется, если в момент ее поступления (когда $dA_s^1 = 1$), все приборы заняты обслуживанием заявок 1 типа (соответственно $I(\xi_s^1 = n_1) = 1$).

3. Компьютерная модель стратегии подвижной границы

Выведем формулы, позволяющие произвести имитационное моделирование. На стохастическом базисе $B = (\Omega, \mathcal{F}, F = (\mathcal{F}_t)_{t \geq 0}, P)$ из формул (1)-(6) нетрудно получить следующие инфинитезимальные соотношения (8)-(11):

$$P\{A_{t+\Delta}^1 - A_t^1 = 1 | \mathcal{F}_t\} = \lambda_1 \cdot \Delta + o(\Delta), \quad (8)$$

$$P\{D_{t+\Delta}^1 - D_t^1 = 1 | \mathcal{F}_t\} = \mu_1 \cdot \xi_s^1 \cdot I(n_1 \geq \xi_s^1 > 0) \cdot \Delta + o(\Delta), \quad (9)$$

$$P\{A_{t+\Delta}^2 - A_t^2 = 1 | \mathcal{F}_t\} = \lambda_2 \cdot \Delta + o(\Delta), \quad (10)$$

$$P\{D_{t+\Delta}^2 - D_t^2 = 1 | \mathcal{F}_t\} = \mu_2 \cdot \beta_s \cdot \Delta \cdot o(\Delta). \quad (11)$$

Формулы (8) - (11) позволяют, основываясь на понятии геометрической вероятности, провести имитационное моделирование. А именно, введя дискретизацию (шаг по времени) Δ из условия $\lambda_i \cdot \Delta \ll 1, n\mu_i \cdot \Delta \ll 1, i=1,2$ получим следующие итерационные формулы (для вычисления последующих значений через предыдущие):

$$A_{t+\Delta}^1 = A_t^1 + \delta(\lambda_1), \quad (12)$$

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

$$D_{t+\Delta}^1 = D_t^1 + \alpha_s \cdot \delta(\mu_1), \quad (13)$$

$$F_{t+\Delta}^1 = F_t^1 + I(\xi_t^1 = n) \Delta A_t^1, \quad (14)$$

где $\Delta A_t^1 = A_{t+\Delta}^1 - A_t^1$,

$$A_{t+\Delta}^2 = A_t^2 + \delta(\lambda_2), \quad (15)$$

$$D_{t+\Delta}^2 = D_t^2 + \beta_s \cdot \delta(\mu_2). \quad (16)$$

Наиболее важной характеристикой для рассматриваемой СМО являются вероятность блокировки вызова для заявок 1 класса. В качестве оценок естественно принять усреднение (по N траекториям) отношения числа заявок, получивших отказ F_t^1 ($i = 1,2$) к числу поступивших заявок A_t^1 ($i = 1,2$):

$$\bar{\pi}_1(t) = \frac{1}{N} \sum_{i=1}^N \left(\frac{F_t^1(\omega_i)}{A_t^1(\omega_i)} \right). \quad (17)$$

В стационарном режиме вероятность блокировки вызова заявок 1 класса, при $\rho_1 < 1$ рассчитывается по формуле Эрланга [6]:

$$P_B = \frac{\frac{1}{n_1!} \rho_1^{n_1}}{\sum_{k=0}^{n_1} \frac{\rho_1^k}{k!}}, \quad (18)$$

где $\rho_i = \frac{\lambda_i}{\mu_i}$ – нагрузка, ($i = 1, 2$), n_1 – каналы, обслуживающие заявки 1 класса.

Вероятность блокировки вызова для заявок 1 класса рассчитывается по формуле Эрланга (18) и представлена на рисунке 3 черным цветом, модельно по формуле (17) красным цветом (сходимость говорит об адекватности компьютерной модели).



Рис.3. График вероятности блокировки вызова с параметрами:
 $T = 10, N = 50, \Delta = 0,001, \lambda_1 = 5, \lambda_2 = 20, \mu_1 = 10, \mu_2 = 10, n = 3, n_1 = 1.$

Среднее число малых нагрузок для пакетов 2 класса при $\alpha \gg 1, \rho_2 < 1$ определяется формулой [2]:

$$Z = \mu_2 \cdot T = \frac{4 \cdot \rho_2}{a \cdot (1 + \rho_1) \cdot (2 + \rho_2)} + \frac{\rho_1 \cdot \rho_2}{a \cdot (1 + \rho_1)}, \quad (19)$$

где $\rho_i = \frac{\lambda_i}{\mu_i}$ – нагрузка ($i = 1, 2$), $\alpha = \frac{\mu_2}{\mu_1}$, $a = (1 - \rho_2) + \frac{1}{1 + \rho_1}$. Для среднего числа больших нагрузок для пакетов 2 класса при $\alpha \gg 1, \rho_2 > 1$ в [2] получена формула:

$$Z = \mu_2 \cdot T = \frac{\rho_2}{a} \left(1 + \frac{a}{2 + \rho_2} + \frac{a \cdot (\rho_2 - 1) \cdot \rho_1}{\rho_2 \cdot (1 + \rho_1)^2} \right), \quad (20)$$

где $\rho_i = \frac{\lambda_i}{\mu_i}$ – нагрузка ($i = 1, 2$), $\alpha = \frac{\mu_2}{\mu_1}$, $a = 2 - \rho_2 - \frac{\rho_1}{1 + \rho_1}$. На рисунке 4 представлен результат работы программы, реализующей модель (1)-(7). Для случаев, когда не выполняется ограничение $\alpha \gg 1$, позволяющее использовать аналитическое решение [2] для оценки формул (19)-(20) получены графики для среднего числа пакетов в СМО. Анализ имитационной модели показал, что формулы (19)-(20), полученные в [2], дают нижние оценки для среднего числа пакетов в СМО, когда не выполняется ограничение $\alpha \gg 1$.

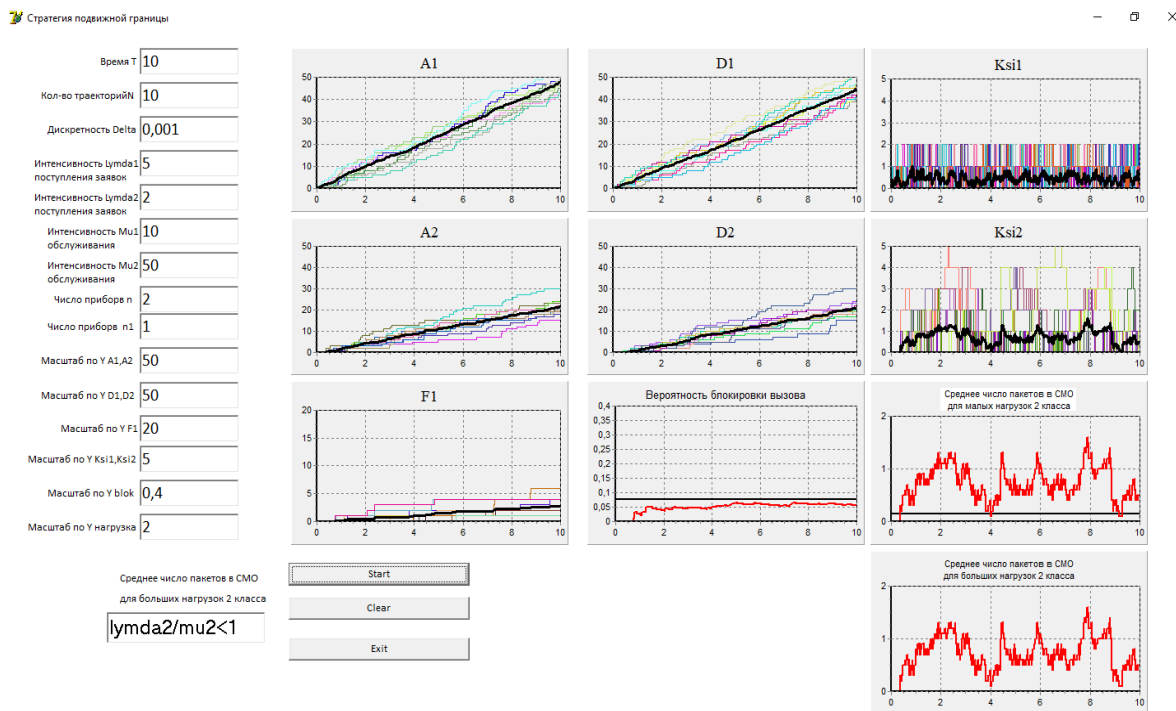


Рис.4. Результаты моделирования стратегии подвижной границы.

Заключение

Рассмотренная стратегия подвижной границы позволяет обеспечить требуемые характеристики обслуживания неоднородного трафика в мультисервисной сети за счет адаптивного распределения нагрузки. Методами имитационного моделирования для случаев, когда не выполняются ограничения, позволяющие использовать аналитическое решение из [2] получены значения для среднего числа пакетов в СМО.

Список литературы

1. Буранова М.А., Киреева Н.В. *Технологии обеспечения качества обслуживания в мультисервисных сетях: учебное пособие*. Самара: ПГУТИ, 2016. 194 с.
2. Крылов В.В., Самохвалова С.С. *Теория телетрафика и ее приложения*. СПб.: БХВ-Петербург, 2005. 288 с.:ил.
3. Шварц М. *Сети связи: протоколы, моделирование и анализ: В 2-х ч.* Пер. с англ. М.: Наука, 1992.
4. Canales, M., Hernández-Solana, Á., Gállego, J.R. et al. Adaptive resource sharing strategies for UMTS multiservice mobiles // *Telecommun. Syst.* 2005, v.28, p. 151–167.
5. Gerasimenko M., Moltchanov D., Andreev S., Koucheryavy Y., Himayat N., Yeh S.-P., Talwar S. Adaptive resource management strategy in practical multi-radio heterogeneous networks // *IEEE Access. Institute of Electrical and Electronics Engineers Inc.* 2017, v. 5, с. 219-235.

6. Antonioli, R.P., Rodrigues, E.B., Maciel, T.F. et al. Adaptive resource allocation framework for user satisfaction maximization in multi-service wireless networks // *Telecommun Syst.* 2018, v. 68, p.259–275.
7. Бутов А.А., Савинов Ю.Г. *Теория массового обслуживания: учебно-методическое пособие*. Ульяновск: УлГУ, 2007.
8. Алиев Т.И. *Основы моделирования дискретных систем*. СПб: СПбГУ ИТМО, 2009. 363 с.
9. Гнеденко Б.В., Коваленко И.Н. *Введение в теорию массового обслуживания*. М.:URSS, 2013. 400 с.
10. Бутов А.А., Галимов Л.А. Оптимальное управление интенсивностью входящего потока многоканальной СМО с роутером при эпизодически наблюдаемой длине очереди на приборах // *Современные проблемы науки и образования*. 2015, № 2, с. 758.
11. Бутов А.А., Галимов Л.А. Стохастическая имитационная модель оценки резерва произошедших, но не заявленных страховых убытков в терминах СМО // *Фундаментальные исследования*. 2016, № 8-2, с. 234-238.
12. Бутов А.А., Галимов Л.А. Оптимальное управление распределением заявок в многоканальной системе массового обслуживания с входящим пуассоновским потоком поступления заявок и экспоненциальным временем обслуживания // *Естественные и технические науки*. 2014, № 9-10 (77), с. 244-247.
13. Столяров И.А., Табакова Е.Д., Савинов Ю.Г. Семимартингальная модель многоканальной СМО с ограниченным временем ожидания // *Прикладная математика и информатика: современные исследования в области естественных и технических наук: материалы IV научно-практической всероссийской конференции (школы-семинара) молодых ученых: 23-25 апреля 2018 г. В двух частях*. – Тольятти: Издатель Качалин Александр Васильевич, 2018, Часть 1, с. 502-506.
14. Столяров И.А., Савинов Ю.Г. Семимартингальная модель СМО с динамическим приоритетом // *Прикладная математика и информатика: современные исследования в области естественных и технических наук: материалы III научно-практической всероссийской конференции (школы-семинара) молодых ученых: 24-25 апреля 2017 г.* – Тольятти: Издатель Качалин Александр Васильевич, 2017, с. 553-557.
15. Савинов Ю.Г., Исмаилова М.В., Рослов М.Э. Траекторные методы моделирования многофазных СМО // *Ученые записки УлГУ. Серия: Математика и информационные технологии*. 2019, № 1, с. 85-91.
16. Бутов А.А., Раводин К.О. *Теория случайных процессов: учебное пособие*. Ульяновск: УлГУ, 2009.