

УДК 004.9

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ СИСТЕМЫ МАССОВОГО ОБСЛУЖИВАНИЯ С КАНАЛАМИ РАЗНОЙ ПРОИЗВОДИТЕЛЬНОСТИ

Нуриев Н.К., Печеный Е.А., Старыгина С.Д.

ФГБОУ ВО «Казанский национальный исследовательский технологический университет»,
Казань, e-mail: nurievnk@mail.ru

В статье представлена модель системы массового обслуживания (СМО) с каналами разной производительности. В работах по данной тематике, как правило, пренебрегают этим фактом, полагая это обстоятельство несущественным. Однако в ряде случаев такой подход дает искаженную оценку эксплуатационных свойств системы. В работе предположено, что как входной поток, так и все потоки обслуживания пуассоновские, построена модель СМО с произвольным, но конечным числом каналов и получена формула для вычисления общего количества ее возможных состояний. В качестве базового инструмента моделирования была использована система уравнений Колмогорова. Неэквивалентность каналов порождает проблему администрирования: рационального распределения входного потока заявок между рабочими органами так, чтобы добиться желаемых значений эксплуатационных характеристик СМО. Этот вопрос был детально исследован в рамках описанной модели с помощью численных экспериментов для двухканальной СМО при различных соотношениях интенсивностей обслуживания и различных уровнях нагрузки на систему. Установлено, что чем больше различие интенсивностей, тем больше оказывается перегружен «слабый» канал. Диспропорция загрузки каналов может быть частично устранена путем применения предложенных приемов администрирования. Из полученных результатов также следует, что объединение в составе одной СМО каналов с существенно различными возможностями (более чем на порядок) нецелесообразно.

Ключевые слова: система массового обслуживания, каналы разной производительности, поток заявок, пуассоновский поток

MATHEMATICAL SIMULATION OF A MASS SERVICE SYSTEM WITH DIFFERENT CAPACITY CHANNELS

Nuriev N.K., Pechenyu E.A., Starygina S.D.

Kazan National Research Technological University, Kazan, e-mail: nurievnk@mail.ru

The article presents a model of a queuing system (QS) with channels of different performance. In works on this topic, as a rule, this fact is neglected, considering this circumstance insignificant. However, in some cases, this approach gives a distorted assessment of the operational properties of the system. Assuming that both the input flow and all service flows are Poisson, a QS model with an arbitrary but finite number of channels is constructed and a formula for calculating the total number of its possible states is obtained. The Kolmogorov system of equations was used as a basic modeling tool. The inequality of the channels gives rise to the problem of administration: rational distribution of the input flow of applications between the working bodies so as to achieve the desired values of the operational characteristics of the QS. This issue was investigated in detail within the framework of the described model using numerical experiments for a two-channel QS at various ratios of service intensities and different levels of load on the system. It was found that the greater the difference in intensities, the more the «weak» channel is overloaded. The disproportion of the channel load can be partially eliminated by applying the proposed administration techniques. It also follows from the results obtained that combining channels with significantly different capabilities (more than an order of magnitude) in one QS is impractical.

Keywords: queuing system, channels of different performance, flow of applications, poisson flow

Математические модели, для построения которых используется аппарат теории массового обслуживания, имеют широчайшее применение в самых различных исследовательских и прикладных задачах. Анализ работы сложных телекоммуникационных систем [1–3], решение задач управления материальными и информационными потоками [4–7], оптимизация логистических структур и т.п. Кроме классических моделей с пуассоновскими потоками, известны и достаточно обстоятельно исследованы модели с групповым обслуживанием заявок [8], модели с ограниченным временем жизни заявок [9–10], так называемые поллинговые системы [11–13], в которых одно устройство обслуживания

принимает заявки из нескольких независимо друг от друга формирующихся очередей и ряд других. В подавляющем большинстве случаев при построении моделей принимается гипотеза об эквивалентности всех обслуживающих устройств (каналов обслуживания), входящих в состав системы. На самом деле это, конечно, не так. Невозможно набрать бригаду рабочих абсолютно одинаковой квалификации или сформировать парк автопредприятия из машин, обладающих одинаковой грузоподъемностью и скоростью и т.п. Наоборот, разнообразие рабочих органов сообщает системе гибкость в управлении и вариативность. Если различия между интенсивностями работы каналов невелики, то, как правило, ими

можно пренебречь, не нанося при этом ущерба адекватности модели. Однако в тех случаях, когда возможности рабочих органов системы массового обслуживания существенно различны, эти различия должны быть обязательно учтены в ходе математического моделирования, в противном случае полученные результаты, скорее всего, будут недостоверны.

Рассмотрим систему массового обслуживания, имеющую n независимых каналов, на вход которой поступает пуассоновский поток заявок интенсивности λ . Потоки заявок, покидающих систему, также являются пуассоновскими, а интенсивности их для разных каналов различны и равны μ_i , $i = 1, n$. Будем полагать, что заявки попадают в любой свободный канал с одинаковой вероятностью. Граф такой системы представлен на рис. 1. Числа в левой части прямоугольников имеют смысл количества каналов, задействованных в обслуживании в данный момент, а в правой части – номера этих каналов. Поскольку каналы неэквивалентны, это суть различные состояния. Весовые коэффициенты ребер имеют смысл интенсивностей соответствующих потоков.

Обращаясь к рис. 1, заметим, что число различных состояний системы быстро растет с увеличением количества каналов. Выделим из множества всех возможных состояний системы подмножество, характеризующиеся одинаковым числом занятых каналов и назовем их *наборами состояний*.

Количество элементов в составе набора, где число занятых каналов равно k , находится с помощью известной комбинаторной формулы

$$Q_k = C_n^k, \quad (1)$$

где n – общее число рабочих органов (каналов) системы массового обслуживания; k – номер набора состояний.

Для системы с отказами, не предусматривающей накопление очереди, общее число возможных состояний системы Q определится как

$$Q = \sum_{k=0}^n C_n^k, \quad (2)$$

Именно это соотношение будет определять размерность системы обыкновенных дифференциальных уравнений Колмогорова, которая используется для математического описания поведения систем массового обслуживания с пуассоновскими потоками.

Аналитическое исследование систем такого типа с большим количеством каналов

затруднительно, несмотря на наличие модели, в силу очевидных технических сложностей. Здесь мы ограничиваемся анализом решений, полученных для стационарного состояния двухканальной системы массового обслуживания с рабочими органами различной интенсивности.

Прежде всего заметим, что наличие различий между рабочими органами значительно повышает роль управления входным потоком заявок, которое предполагает отыскание наиболее рационального распределения входного потока заявок между каналами обслуживания. Ниже представлены три возможные схемы администрирования и системы уравнений Колмогорова для отыскания вероятностей стационарных состояний.

1. Поступление заявок на оба канала происходит с одинаковой вероятностью.

$$\begin{cases} -\lambda P_0 + \mu_1 P_1^1 + \mu_2 P_1^2 = 0 \\ \frac{\lambda}{2} P_0 - (\lambda + \mu_1) P_1^1 + \mu_2 P_2 = 0 \\ \frac{\lambda}{2} P_0 - (\lambda + \mu_2) P_1^2 + \mu_1 P_2 = 0, \\ \lambda P_1^1 + \lambda P_1^2 - (\mu_1 + \mu_2) P_2 = 0 \\ P_0 + P_1^1 + P_1^2 + P_2 = 1 \end{cases}, \quad (3)$$

где P_0 – вероятность того, что система свободна; P_1^1 – вероятность того, что в системе одна заявка и ее обслуживанием занят канал с интенсивностью μ_1 ; P_1^2 – вероятность того, что в системе одна заявка и ее обслуживанием занят канал с интенсивностью μ_2 ; P_2 – вероятность того, что на обслуживании находятся две заявки и, так как накопление очереди не предусмотрено, $P_2 = P_{\text{отк}}$. Граф распределения потоков по такой схеме для случая n каналов представлен на рис. 1.

2. Вероятности поступления заявок в каналы обслуживания пропорциональны интенсивностям потоков обслуживания соответствующих каналов.

$$\begin{cases} -\lambda P_0 + \mu_1 P_1^1 + \mu_2 P_1^2 = 0 \\ \frac{\lambda \mu_1}{\mu_1 + \mu_2} P_0 - (\lambda + \mu_1) P_1^1 + \mu_2 P_2 = 0 \\ \frac{\lambda \mu_2}{\mu_1 + \mu_2} P_0 - (\lambda + \mu_2) P_1^2 + \mu_1 P_2 = 0, \\ \lambda P_1^1 + \lambda P_1^2 - (\mu_1 + \mu_2) P_2 = 0 \\ P_0 + P_1^1 + P_1^2 + P_2 = 1 \end{cases}, \quad (4)$$

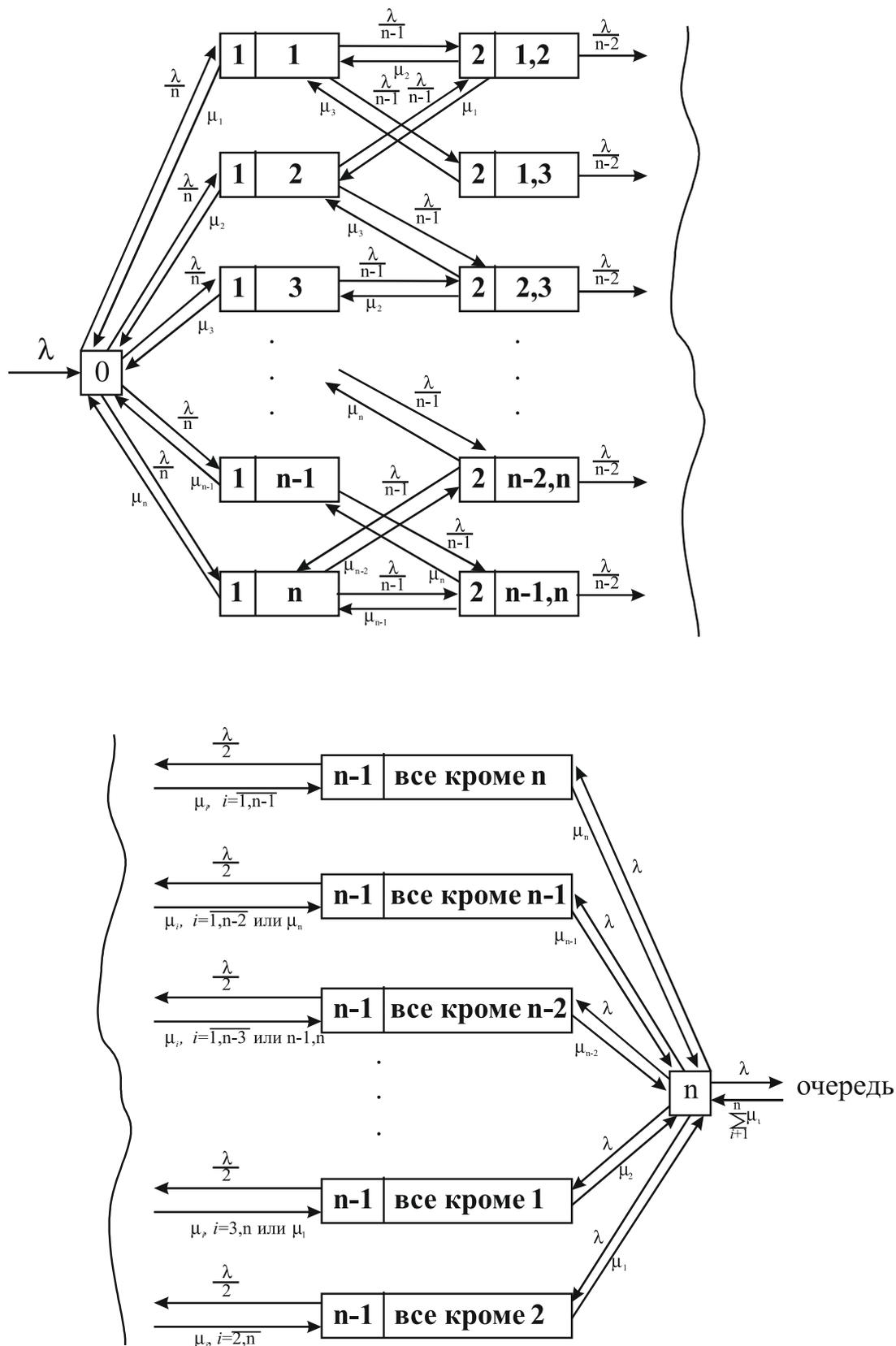


Рис. 1. Система массового обслуживания с n каналами разной производительности и равновероятным распределением входного потока

Граф распределения потоков по такой схеме представлен на рис. 2.

3. Если оба канала свободны, поступившая заявка отправляется на «сильный» канал. (Приоритет «сильного» канала).

$$\begin{cases} -\lambda P_0 + \mu_1 P_1^1 + \mu_2 P_1^2 = 0 \\ -(\lambda + \mu_1) P_1^1 + \mu_2 P_2 = 0 \\ \lambda P_0 - (\lambda + \mu_2) P_1^2 + \mu_1 P_2 = 0, \\ \lambda P_1^1 + \lambda P_1^2 - (\mu_1 + \mu_2) P_2 = 0 \\ P_0 + P_1^1 + P_1^2 + P_2 = 1 \end{cases} \quad (5)$$

Граф распределения потоков по такой схеме представлен на рис. 3.

Системы уравнений (3)–(5) были использованы для сравнительной оценки возможностей всех вышеописанных схем администрирования при различных соотношениях интенсивностей рабочих органов. Для обеспечения корректности сравнения

интенсивности входных потоков заявок λ выбирались так, чтобы величина нагрузки на систему $\gamma = \frac{\lambda}{\sum_{i=1}^n \mu_i}$ оставалась неизмен-

ной и равной 0,8. Результаты вычислений сведены в табл. 1.

В процессе анализа материалов табл. 1 обращает на себя внимание значительная диспропорция загруженности каналов обслуживания, заметно возрастающая с увеличением различий интенсивностей потоков обслуживания. Так, при распределении входного потока заявок по схеме 1 и величины отношения $\frac{\mu_2}{\mu_1} = 4$ время

работы «слабого» канала оказывается примерно в 1,6 раза больше «сильного». По мере увеличения отношения $\frac{\mu_2}{\mu_1}$ различие нарастает и при $\frac{\mu_2}{\mu_1} = 12$ приближается к 2.

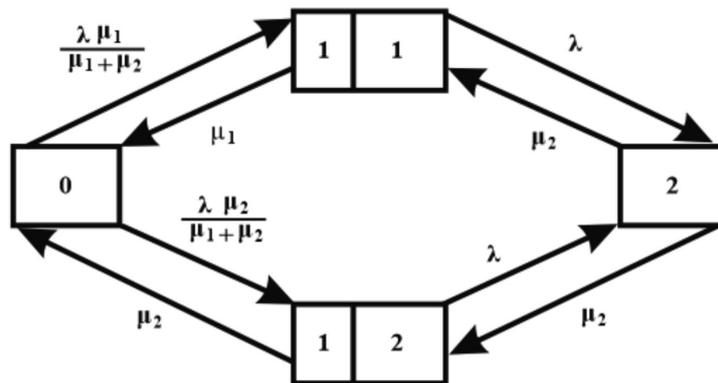


Рис. 2. Распределение потоков пропорционально их интенсивности

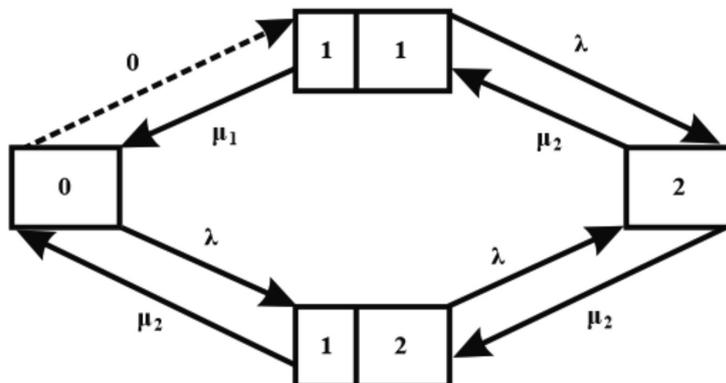


Рис. 3. Распределение потоков с приоритетом на «сильный» канал

Таблица 1

Вероятности стационарных состояний двухканальной системы с отказами

	$\lambda = 4 \mu_1 = 1 \mu_2 = 4$			$\lambda = 7,2 \mu_1 = 1 \mu_2 = 8$			$\lambda = 10,4 \mu_1 = 1 \mu_2 = 12$			Случай одинаковых каналов
	Схема 1	Схема 2	Схема 3	Схема 1	Схема 2	Схема 3	Схема 1	Схема 2	Схема 3	
P_0	0,182	0,205	0,224	0,121	0,152	0,163	0,090	0,120	0,127	0,2577
P_1^1	0,364	0,315	0,276	0,434	0,383	0,363	0,467	0,420	0,408	0,2062
P_1^2	0,091	0,126	0,155	0,054	0,089	0,102	0,039	0,070	0,077	0,2062
P_2	0,363	0,354	0,345	0,391	0,376	0,372	0,404	0,390	0,388	0,3299

Таблица 2

Вероятности стационарных состояний двухканальной системы при различной нагрузке

	$\mu_1 = 1 \mu_2 = 4 \lambda = 2 \gamma = 0,4$				$\mu_1 = 1 \mu_2 = 4 \lambda = 8 \gamma = 1,6$			
	Схема 1	Схема 2	Схема 3	одинаковые каналы	Схема 1	Схема 2	Схема 3	одинаковые каналы
P_0	0,364	0,417	0,461	0,471	0,071	0,078	0,082	0,107
P_1^1	0,363	0,278	0,205	0,189	0,286	0,266	0,251	0,172
P_1^2	0,091	0,139	0,180	0,189	0,071	0,089	0,102	0,172
P_2	0,182	0,166	0,154	0,151	0,571	0,567	0,565	0,549

Использование других схем администрирования позволяет уменьшить диспропорцию при $\frac{\mu_2}{\mu_1} = 4$ до 1,25 раза, а при $\frac{\mu_2}{\mu_1} = 12$ –

до 1,7 раза, но не устранить полностью. Это может оказаться существенным и повлиять на работоспособность системы, если рабочие органы имеют элементы, характеристики которых подвержены временному дрейфу, или требуют длительной настройки и регулярного профилактического ремонта.

Неэквивалентность каналов заметно влияет и на технико-экономические показатели. Так, если для системы с каналами одинаковой интенсивности доля потерянных заявок равна примерно 33 %, то при $\frac{\mu_2}{\mu_1} = 12$

она достигнет 40 % и практически не корректируется с помощью рассмотренных приемов администрирования. Это ставит под сомнение целесообразность объединения в систему каналов, интенсивность работы которых существенно различна.

В рамках данной работы авторы ограничились только рассмотрением систем с отказами, не имеющими опции накопления очереди, поскольку в системах с ожиданием доля обслуженных заявок растет по мере

увеличения допустимого числа мест в очереди независимо от свойств каналов обслуживания. При отсутствии ограничений на длину очереди и выполнении условия $\gamma < 1$ все заявки рано или поздно будут обслужены, т.е. $P_{отк.} = 0$ для любого соотношения интенсивностей потоков обслуживания.

Представляет интерес вопрос о том, как изменяются эксплуатационные характеристики системы с неэквивалентными каналами при изменении величины нагрузки γ . С помощью модели (3)–(5) были вычислены вероятности стационарных состояний системы, находящейся в «недогруженном» положении: $\gamma = 0,4$ и «перегруженной» системы: $\gamma = 1,6$. Результаты расчетов представлены в табл. 2.

Из материалов таблицы видно, что с увеличением загруженности системы наблюдается осязаемое снижение диспропорции загруженности «слабого» и «сильного» каналов. В условиях малой нагрузки ($\gamma = 0,4$), отношение вероятностей загрузки слабого канала к сильному при работе по схеме 1 равно 2. С увеличением γ до 0,8 это отношение снижается до 1,6, а при $\gamma = 1,6$ оно составляет 1,33. Вместе с тем при малых значениях нагрузки существенно возрастает эффект администрирования в рамках

предложенных вариантов перераспределения потока поступающих в систему заявок. Особенно отчетливо это проявляется на показателе P_0 , характеризующем долю времени, в течение которого система свободна.

Заключение

1. На базе системы уравнений Колмогорова дано математическое описание поведения систем массового обслуживания с неэквивалентными каналами, функционирующими в режиме пуассоновских потоков.

2. Определены эксплуатационные характеристики двухканальной системы для различных соотношений интенсивностей обслуживания и величины общей нагрузки на систему

3. Предложены схемы администрирования, позволяющие снизить диспропорцию загрузки каналов, которая возникает вследствие их неэквивалентности.

4. Показана нецелесообразность объединения в составе системы каналов, значительно различающихся по показателю интенсивности обслуживания.

Список литературы

1. Трухин М.П. Моделирование сигналов и систем. Система массового обслуживания. М.: ЭБС Лань, 2019. 232 с.
2. Шмелева А.Г., Ладынин А.И., Бахметьев А.В. Построение взвешенных решений управления сложными производственными системами с применением теории массового обслуживания // Информационные технологии. 2018. Т. 24. № 6. С. 421–426.
3. Калинин Р.Н., Щуров К.С. Математическая модель потока заявок на обслуживание потоков уплотняемых ис-

точников в мультиплексах с динамическим уплотнением каналов // Теория и техника радиосвязи. 2017. № 2. С. 93–96.

4. Печинкин А.В., Разумчик Р.В. Системы массового обслуживания в дискретном времени. М.: Физматлит, 2018. 432 с.

5. Кирпичников А.П., Нгуен Тхань Банг, Чан Куанг Кун. Вероятность отказа и вероятность ожидания начала обслуживания в системе с очередью конечной длины и ограниченным средним временем пребывания заявки в очереди // Вестник Казанского технологического университета. Казань: КНИТУ, 2016. Т. 19. № 21. С. 151–154.

6. Ку Дык Тоан, Нуриев Н.К., Печеный Е.А., До Шон Ха, Нгуен Нгок Куинь. Математическое и имитационное моделирование эффективного администрирования широкополосных беспроводных сетей // Современные наукоемкие технологии. 2019. № 6. С. 77–82.

7. Ахметшин Д.А., Нуриев Н.К., Печеный Е.А. Математическое моделирование эффективного администрирования системы доступа в Интернет // Фундаментальные исследования. Технические науки. 2014. № 9. С. 2650–2654.

8. Гайдамака Ю.В., Зарипова Э.Р., Орлов Ю.Н. Анализ зависимости параметров модели сервера протокола установления сессий с групповым поступлением сообщений от распределения длины группы сообщений // Препринты ИПМ им. М.В. Келдыша. 2015. № 27. 16 с.

9. Рыжиков Ю.И. Численные методы теории очередей. М.: ЭБС Лань, 2019. 512 с.

10. Нуриев Н.К., Али А.А., Печеный Е.А. Моделирование одноименного смешанного потока с ограниченным временем обслуживания // Вестник Казанского технологического университета. 2016. Т. 19. № 24. С. 120–122.

11. Вишневецкий В.М., Семенова О.В. Системы поллинга: теория и применение в широкополосных беспроводных сетях. М.: Техносфера, 2007. 312 с.

12. Boxma O.J., Schlegel S., Yechiali U. Two-queue polling models with a patient server. Annals of Operations Research. 2002. V. 112. P. 101–121.

13. Муршед Ф.А., Нуриев Н.К., Печеный Е.А. Моделирование эффективного администрирования поллинговых систем с ограниченным временем жизни заявок // Современные наукоемкие технологии. 2018. № 7. С. 77–83.