

Министерство науки и высшего образования Российской Федерации
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ (НИ ТГУ)
Институт прикладной математики и компьютерных наук
Кафедра прикладной информатики

ДОПУСТИТЬ К ЗАЩИТЕ В ГЭК

Зав.каф. прикладной информатики
д-р.физ.-мат.н., профессор
_____ Е.М. Семёнов
« ____ » _____ 2021 г.

БАКАЛАВРСКАЯ РАБОТА

ИССЛЕДОВАНИЕ ДВУМЕРНОГО ВЫХОДЯЩЕГО ПОТОКА МАРКОВСКОЙ МОДЕЛИ УЗЛА ОБРАБОТКИ ЗАПРОСОВ С ПОВТОРНЫМИ ОБРАЩЕНИЯМИ И ВЫЗЫВАЕМЫМИ ЗАЯВКАМИ

по основной образовательной программе подготовки бакалавров
направление подготовки
090303 - Прикладная информатика

Благинин Алексей Леонидович

Руководитель ВКР
доцентр.каф. прикладной информатики
_____ И.Л. Лапатын
« ____ » _____ 2021 г.

Автор работы
студент группы № 931704
_____ А.Л. Благинин

СОДЕРЖАНИЕ

1	Введение	3
2	Математическая модель	10
3	Уравнения Колмогорова	11
4	Метод асимптотического анализа	13

1 Введение

Многие ситуации с очередями имеют особенность, заключающуюся в том, что клиенты, обнаруживающие, что зона обслуживания занята по прибытии, должны временно покинуть ее и присоединиться к группе неудовлетворенных клиентов, но они повторяют свой запрос через некоторое случайное время. Говорят, что между испытаниями заказчик находится на орбите. Эти модели массового обслуживания возникают при стохастическом моделировании многих протоколов связи, локальных сетей и повседневных жизненных ситуаций. Самый простой и очевидный пример - это человек, который звонит по телефону. Если линия занята, значит, он не может стоять в очереди, но через некоторое время снова испытывает удачу. Для систематического изложения основных методов и основных результатов по этой теме мы отсылаем читателя к книгам [1], [2] и библиографической информации, приведенной в [3], [4], [5]. Чтобы проиллюстрировать активную роль очередей на повторное рассмотрение за последние несколько лет, упомянем некоторые недавние статьи, опубликованные в этом журнале [6], [7], [8], [9], [10], [11], [12].

В большинстве публикаций, посвященных очередям повторного запроса, сервер обслуживает только тех, кто поступает от постоянных клиентов. Однако существуют реальные ситуации (например, сценарий call-центра), когда серверы имеют возможность совершать исходящие телефонные звонки, когда они не участвуют в разговоре. Эта функция организации очереди известна как парная коммутация или модели двусторонней связи. В ранней литературе Фалин [13] вывел интегральные формулы для частичных производящих функций и явные выражения для ожидаемого значения некоторых характеристик производительности очереди повторных вызовов с двухсторонней связью в предположении, что длительности входящих и исходящих вызовов соответствуют одинаковому распределению времени обслуживания. Однако на практике это предположение носит ограничительный характер, поскольку разные типы клиентов обычно демонстрируют разное поведение и, следовательно, у них должны быть разные потребности в обслуживании. Арталехо и Ресинг [14] использовали метод анализа среднего значения для получения некоторых ожидаемых значений, связанных со временем ожидания в очереди на повторное обращение с двусторонней связью и различным распределением времени обслуживания входящих и исходящих вызовов. Недавно Artalejo и Phung-Duc [15] провели подробное исследование очереди на повторное рассмотрение с двухсторонней связью. Полученные результаты включают явные выражения для совместного стационарного распределения состояния сервера и количества заявок на орбите, а также для частичных факториальных моментов. Большинство явных формул в [15] выражены в терминах гипергеометрических рядов, что согласуется с той особой ролью, которую играют эти специальные функции при выводе аналитических решений для многих других очередей повторных попыток [16], [17], [18], [19], [20], [21].

Наша главная цель в этой статье - дать представление об исследовании очередей по-

вторного запроса на одном сервере с двусторонней связью. С этой целью мы сначала проводим тщательное исследование очереди повторного запроса - типа с различным распределением времени входящего и исходящего обслуживания. Для этой модели получены как стационарные, так и асимптотические результаты. Стационарный анализ основан на хорошо известных математических инструментах, но они предоставляют новые выражения в замкнутой форме и стабильные вычислительные схемы для производительности системы. Таким образом, мы надеемся, что наши явные формулы будут полезны в приложениях. С другой стороны, асимптотический подход, используемый в этой статье, подразумевает заметное алгебраическое упрощение по сравнению с выводами, необходимыми для получения аналоговых асимптотических результатов для классической очереди повторных вызовов (без исходящих вызовов) [22]. Еще одна цель этой статьи - использовать марковский процесс прибытия (МАР) и распределение типа фазы (РН) для создания более сложных очередей повторных запросов с двусторонней связью, допускающей неэкспоненциальные прибытия и корреляцию между временами прибытия.

782 ИИСУС Р. АРТАЛЕХО И ТУАН ПХУНГ-ДУК В последнее время большое внимание уделяется очередям на повторное рассмотрение, потому что они имеют приложения для анализа производительности различных систем, таких как центры обработки вызовов, компьютерные сети и телекоммуникационные системы [3, 12, 17, 28]. Очереди на повторное рассмотрение характеризуются тем, что клиенты (т.е. звонки), которые не могут получить услугу по прибытии выходят на виртуальную орбиту и повторяют попытку обслуживания через некоторое случайное время. Поток прибытия с орбиты делает ниже лежащую марковскую цепочку очередей на повторное рассмотрение быть неоднородным. В результате анализ очередей на повторное рассмотрение намного сложнее, чем у соответствующих моделей массового обслуживания без повторных попыток и явных результатов получают лишь в некоторых частных случаях [3, 12, 23, 24]. Гипергеометрические функции и их специальные версии играют важную роль в вывод аналитических решений для очередей на повторное рассмотрение. Фактически стационарные характеристики состояния системы обычной очереди повторных запросов $M / M / 1/1$ выражается через специальные гипергеометрические функции [3, 12, 24]. Обзор 2000 Математическая классификация предметов. Первичный: 68M20, 90B22; Вторичный: 60K25. Ключевые слова и фразы. Очереди повторных вызовов, двусторонняя связь, смешанные центры обработки вызовов, стационарное распределение, факторные моменты, рекурсивные формулы, асимптотический анализ. Рецензированием статьи занимались Уи Юэ и Ютака Такахаши в качестве гостя. Редакторы. существующая литература показывает, что гипергеометрические функции также являются ключевым инструментом для анализировать стационарные характеристики (т.е. предельные вероятности состояния системы и их частичные производящие функции) широкого спектра очередей повторных попыток, включая одиночные серверные очереди с отказом Бернулли [12, 24], очереди на повторное рассмотрение $M / M / 1/1$ тип с отказом Бернулли и обратной связью [9], очереди повторных запросов на одном сервере с

орбитальный поиск и непостоянные заявки [19], очередь повторных запросов $M / M / 1/1$ с линейная политика повторного рассмотрения [2] и очередь повторного рассмотрения $M / M / 2/2$ [14]. Может быть, последний пример принадлежит Киму [16], который изучает очередь повторных запросов на одном сервере с конфликтом и нетерпение при использовании гипергеометрических функций. В большинстве публикаций по очередям повторного вызова сервер обслуживает только входящие вызовы. После обслуживания вызова сервер ждет либо следующего поступления основного вызова, либо для повторного вызова. Однако бывают ситуации из реальной жизни, когда у серверов есть возможность совершать исходящие телефонные звонки. Наиболее очевидное применение возникает в повседневной жизни, потому что каждый использует телефонную линию или мобильный телефон для приема звонков, но также для звонков на улицу. Более того, в различных системах обслуживания, таких как звонок центр, оператор не только обслуживает входящие звонки, но и совершает исходящие звонки. звонит, если он свободен. Пока сервер занят, входящие звонки не могут получить услугу. Мы предполагаем, что эти вызовы присоединяются к орбите, и снова пытаемся занять сервер после некоторого экспоненциально распределенного времени независимо от других вызовов. В настоящее время бизнес колл-центра очень важен, потому что он обеспечивает канал для двусторонней связи между компаниями и их клиентами [1, 18, 26]. Как правило, существует два типа центров обработки вызовов: центры обработки вызовов для входящих и исходящих вызовов. Первый используется для поддержки клиентов, когда клиенты звонят извне для некоторых такие запросы, как бронирование билетов и подтверждение данных кредитной карты или жалоба на продукцию и т. д. [27]. С другой стороны, последний используется для телефонный маркетинг, при котором система набора номера случайным образом направляет призывы к потенциальным клиентам для рекламы или продажи новых продуктов [26]. Недавно, современные центры обработки вызовов объединяют как входящие, так и исходящие функции, чтобы увеличить продуктивность [7, 10]. Они называются смешанными центрами обработки вызовов, в которых оператор не только принимает входящие звонки, но и звонит клиентам, когда он или она простаивает. Бхулаи и Кул [7] предлагают модель организации массового обслуживания с несколькими серверами с бесконечным буфером для смешанные центры обработки вызовов, для которых разработаны оптимальные и почти оптимальные политики для случай, когда входящие и исходящие вызовы имеют одинаковое экспоненциальное распределение и в противном случае соответственно. Deslauriers et al. [10] разработать пять марковских очередей модели смешанных центров обработки вызовов, в которых различаются входящие и исходящие вызовы и незаметно. Как указано в [10], модели входящего и исходящие вызовы с разными распределениями сложнее, чем с одинаковое распределение времени обслуживания для обоих типов вызовов. В этих работах [7, 10] повторные испытания не учитываются. Фалин [11] выводит интегральные формулы для частичных производящих функций и явные выражения для некоторых ожидаемых показателей эффективности повторного испытания $M / G / 1/1$. очередь с двусторонней

связью, в которой входящие и исходящие вызовы предполагается, что они следуют тому же распределению услуг. Choi et al. [8] расширяют модель Фалина в очереди повторного вызова $M / G / 1 / K$, где входящие и исходящие вызовы также предполагается следуют тому же распределению времени обслуживания. Однако с точки зрения приложения точки зрения, это предположение носит ограничительный характер, поскольку входящие и исходящие вызовы могут иметь разное распределение времени обслуживания. Арталехо и Ресинг [5] получают первые ПОВТОРНЫЕ ОЧЕРЕДИ С ДВУСТОРОННЕЙ СВЯЗЬЮ 783 частичные моменты для очереди повторных вызовов $M / G / 1/1$ с различным распределением времени обслуживания входящих и исходящих вызовов с использованием подхода анализа среднего значения. Это Следует отметить, что анализ среднего значения не может использоваться для получения стационарного распределения, а также высшие факторные моменты. В этом документе термин двухсторонняя связь относится к тому факту, что сервер может совершать исходящие звонки, пока не участвует в разговоре. Есть ряд моделей повторного судебного разбирательства, которые связаны с этим определением двухстороннего коммуникационная функция. Фактически, с аналитической точки зрения, два пути модель коммуникации можно даже рассматривать как частный случай других существующих модели, которые по своему происхождению были разработаны для моделирования других различных очередей Особенности. Это случай ссылок [21, 6, 12]. Мартин и Арталехо [21] рассматривают очередь $M / G / 1/1$ с двумя типами нетерпеливых единицы, которые можно рассматривать как очередь на повторное рассмотрение с двухсторонней связью. В [21] а заблокированный заказчик хранится в орбитальной очереди, из которой только заказчик Глава очереди может повторить попытку после экспоненциально распределенного времени. Авраченко и другие. [6] использовать методы матричного анализа для изучения очереди повторных запросов на одном сервере с два класса клиентов, чьи повторные попытки и распределения времени обслуживания разные. Прибытие происходит в соответствии с ярко выраженным марковским процессом прибытия. Несомненно, что рассмотрение обобщенных марковских приходов, допускающих корреляция - интересная цель. Однако следует отметить, что методика использованный в [6] не дает явных решений. Фалин и Темплтон [12] представляют предварительный анализ многоклассовых очередей на повторное рассмотрение $M / G / 1/1$, для которых система представлены уравнения для среднего числа заявок на орбите. В авторы в [12] также указывают на некоторые открытые проблемы для модели, которые требуют дальнейшего изучения. Существующая библиография по очередям на повторное рассмотрение обширна и обширна. В результате в В дополнение к вышеупомянутым ссылкам можно было бы найти другие повторные судебные разбирательства модели, относящиеся к двухсторонней очереди связи, исследуемой здесь. В целом, среди ближайших вариантов повторного исследования отметим мультикласс, приоритетный и нетерпеливый повторные модели. Для общего обзора читатель отсылается к разделу 2.3 в Artalejo и Gomez-Corral [3], а также к обновленной библиографии [4]. Первая и основная цель этой статьи - предоставить бо-

более обширный анализ очереди на повторное рассмотрение $M / M / 1/1$ с двухсторонней связью и другим сервисом временное распределение входящих и исходящих звонков. В частности, мы предоставляем явные решения для совместного стационарного распределения состояния сервера и количество заявок на орбите, частные факториальные моменты и их порождающие функции. Мы также приводим рекурсивные формулы для стационарного распределения и частные факториальные моменты, на основе которых как символьные, так и численные алгоритмы могут быть реализованы. Кроме того, мы выводим несколько простых асимптотических формул для стационарного распределения и частные факториальные моменты. Вторая цель этой статьи - обсудить расширение мультисерверного повторного тестирования. очереди с двусторонней связью и различным распределением входящих и исходящих вызовов, для которых мы получаем некоторые явные результаты. В частности, мы устанавливаем необходимое и достаточное условие устойчивости системы и вывести явную формулу для среднего количества входящих вызовов на серверах. Кроме того, мы сформулируйте мультисерверную модель с помощью зависимой от уровня квазिवорождения и смерти (QBD) процесс, который может быть использован для численного исследования. Надеемся, что наша модель полезен для анализа производительности смешанных центров обработки вызовов. 784 ИИСУС Р. АРТАЛЕХО И ТУАН ПХУНГ-ДУК

Оставшаяся часть теста организована следующим образом. Раздел 2 описывает модель в деталь. Раздел 3 посвящен основным результатам данной статьи, в которой обширный представлено исследование очереди повторных попыток $M / M / 1/1$ с двусторонней связью. В разделе 4 мы обсуждаем расширение многосерверной очереди повторных запросов с двухсторонней общение и получить некоторые явные результаты.

СТАТЬЯ В этой статье мы рассматриваем двумерный процесс вывода системы массового обслуживания. [6,8] с повторными вызовами [1] и вызванными приложениями [7]. Такую систему можно интерпретировать как узел обработки с множественным произвольным доступом, который в свободное время обработки запросов может запросить самодиагностику или любую другую процедуру, которая будет продолжаться в случайное время. Также может применяться рассматриваемая система для моделирования узлов обработки с разными типами приложений. Приложения одного вида не теряются и обслуживаются в любом случае, а заявки другой тип будет обслуживаться только с бесплатным ресурсом. Отдельные узлы образуют модель сети связи, в которой исходящий поток из одного узла входит в другой. В случае приложений разных типов, после обслуживания на определенном узле, они уходят по своим маршрутам. Таким образом, результаты исследования выходных процессов систем массового обслуживания широко применяются для проектирования реальных систем передачи данных и анализ сложных процессов, состоящих из нескольких этапов. В связи с этим для 2 А.Благинин, И.Лапатин моделируя сети, важно иметь информацию о наличии корреляция между процессами в нем. Слабая корреляция позволяет рассматривать процессы как независимые при моделировании, что может существенно упростить модель и ее исследование. В данной работе мы рассматриваем

влияние параметров системы на значения асимптотического коэффициента корреляции компонент двумерной процесс вывода разных типов приложений. Чтобы изучить систему, метод асимптотического анализа используется для нахождения вида предельного двумерное распределение количества обслуженных заявок ввода процесс и количество обслуженных вызванных заявок за некоторое время t , при условии что на орбите наблюдается большая задержка заявок [10].

КУРСАЧ Человек, в своей повседневной жизни, постоянно сталкивается с необходимостью получить доступ к различным ресурсам, к которым пытаются получить доступ и другие люди. При этом ресурсы могут быть совершенно разной природы – касса в магазине, консультация преподавателя, место в общественном транспорте и многие другие. Во всех этих случаях человек чаще всего сталкивается с очередью, которая может быть как упорядоченной, так и неупорядоченной. Математическим моделированием и изучением закономерностей, возникающих при различных способах организации доступа к ресурсу, занимается теория массового обслуживания [1]. Заметим, что англоязычное название (queuing theory) на русский язык переводится как теория очередей. [2] Теория массового обслуживания возникла еще в начале XX века. А.К. Эрланг [3], решая практические задачи совершенствования работы систем связи, получил ряд формул, которые являются базовыми для теории массового обслуживания. Терминология, используемая в теории массового обслуживания, соответствует исходным задачам – заявка, прибор, требование, буфер и другие. Классическими моделями массового обслуживания являются модели с очередью и модели с отказами. В моделях с очередью заявка (требование), которая обратилась к прибору (ресурсу) и застала его занятым, встает в очередь для ожидания обслуживания. В моделях с отказами заявка при невозможности получить обслуживание теряется. Существует достаточно много модификаций и комбинаций таких моделей. Во второй половине XX века происходит бурное развитие телекоммуникационных сетей, что приводит к необходимости моделировать и проектировать сети передачи данных различной структуры и назначения. Для этого стали применяться модели систем массового обслуживания с повторными вызовами. В системах такого типа заявка, которая не может получить доступ к ресурсу повторяет попытку захвата ресурса через случайную задержку, а не встает в очередь. В англоязычной литературе такие модели называются retrial queue (RQ), большой обзор работ по исследованию которых приведен в монографии [4]. Заметим, что большинство работ посвящены исследованию характеристик состояния самой системы (число заявок в системе, среднее время пребывания в системе) [5, 6], а исследованию выходящих потоков уделяется недостаточно внимания. При этом, в разного рода системах обслуживания, таких как системы телекоммуникационной связи, автоматизированные call-центры, компьютерные сети и др., одной из важнейших и представляющей большой практический интерес характеристик является именно число обслуженных заявок, то есть выходящий поток заявок. Исследование выходящих потоков усложняется тем, что его характеристики зависят от функционирования самой системы. Это увеличивает

размерность задачи и далеко не всегда делает возможным получить решение аналитически. Результаты для выходящих потоков классических систем были получены еще в середине XX века [7, 8]. Обзор работ по исследованию выходящих потоком содержится в [9]. В работе [9] рассматриваются выходящие потоки систем с повторными вызовами. Качественное и структурное изменение реальных систем требует разработки новых математических моделей, которые позволят точно описывать их функционирование. В данной работе предлагается исследовать выходящий поток системы с повторными вызовами и вызываемыми заявками. Такую систему можно интерпретировать как узел сети связи со случайным множественным доступом, который в свободное от обработки запросов время может запросить самодиагностику или другую процедуру, которая будет продолжаться случайное время. Так же модель call-центра, в которой вызываемые заявки могут интерпретироваться как «холодные звонки» клиентам, в то время, как линия не занята. Отдельные узлы образуют модель сети связи, в которой выходящий поток одного узла является входящим для другого, поэтому результаты исследования выходящих потоков сетей массового обслуживания широко применимы для проектирования реальных систем передачи данных и анализа сложных процессов, состоящих из нескольких этапов.

2 Математическая модель

Рассмотрим RQ-систему, на вход которой поступает простейший поток заявок с интенсивностью λ . Заявка входящего потока, поступая в систему и обнаруживая прибор свободным, занимает его. Прибор, в свою очередь, начинает обслуживание в течение случайного времени, распределенного экспоненциально с параметром μ_1 . Если же при поступлении в систему заявка обнаруживает прибор занятым, она мгновенно уходит на орбиту и осуществляет там случайную задержку в течение экспоненциально распределенного времени с параметром σ . В свободное от обслуживания заявок с входящего потока время прибор сам вызывает заявки с интенсивностью α и обслуживает их в течение экспоненциально-распределенного времени с параметром μ_2 .

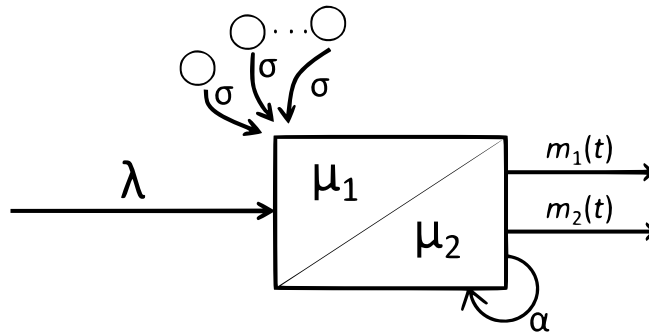


Рисунок 1 — Модель системы

Введем следующие обозначения: $i(t)$ — число заявок на орбите в момент времени t , $k(t)$ — состояние прибора: 0 — прибор свободен, 1 — прибор занят обслуживанием заявки входящего потока, 2 — прибор занят обслуживанием вызванной заявки; $m_1(t)$ — число обслуженных заявок входящего потока в момент времени t , $m_2(t)$ — число обслуженных вызванных заявок t .

3 Уравнения Колмогорова

Итак, мы имеем три характеристики, определяющие результат функционирования системы за некоторое время $i(t)$: состояние прибора – $k(t)$, количество заявок на орбите – $i(t)$, количество обслуженных заявок входящего потока – $m_1(t)$, количество обслуженных вызванных заявок – $m_2(t)$, что можно представить в виде четырех-мерного Марковского процесса

$$\{k(t), i(t), m_1(t), m_2(t)\}$$

Заметим, что именно такая комбинация характеристик будет являться Марковским процессом, так как даёт достаточно информации о том, какое состояние система примет в следующий момент времени. Для этого необходимо знать, в каком состоянии прибор был в предшествующий момент времени, и какое количество заявок находилось в источнике повторных вызовов. Следующее состояние, которое прибор может принять, зависит от состояния, в котором он находился прежде, то есть, каждое из трех состояний $k(t)$ принимается прибором с некоторыми вероятностями. Введем их в рассмотрение

$$P\{k(t) = 0, i(t) = i, m_1(t) = m_1, m_2(t) = m_2\} = P_0(i, m_1, m_2, t)$$

$$P\{k(t) = 1, i(t) = i, m_1(t) = m_1, m_2(t) = m_2\} = P_1(i, m_1, m_2, t)$$

$$P\{k(t) = 2, i(t) = i, m_1(t) = m_1, m_2(t) = m_2\} = P_2(i, m_1, m_2, t)$$

Запишем получившуюся систему уравнений

$$\begin{aligned} \frac{\partial P_0(i, m_1, m_2, t)}{\partial t} &= -(\lambda + i\sigma + \alpha)P_0(i, m_1, m_2, t) + P_1(i, m_1 - 1, m_2, t)\mu_1 + \\ &\quad + P_2(i, m_1, m_2 - 1, t)\mu_2, \\ \frac{\partial P_1(i, m_1, m_2, t)}{\partial t} &= -(\lambda + \mu_1)P_1(i, m_1, m_2, t) + (i + 1)\sigma P_0(i + 1, m_1, m_2, t) + \\ &\quad + \lambda P_0(i, m_1, m_2, t), \\ \frac{\partial P_2(i, m_1, m_2, t)}{\partial t} &= -(\lambda + \mu_2)P_2(i, m_1, m_2, t) + \lambda P_2(i - 1, m_1, m_2, t) + \\ &\quad + \alpha P_0(i, m_1, m_2, t). \end{aligned} \tag{1}$$

Полученная система уравнений – система дифференциальных уравнений Колмогорова, где в левой части каждого уравнения находится производная вероятности состояния рассматриваемого процесса, а в правой – сумма произведений вероятностей состояний, из которых прибор может принять это состояние, на интенсивности соответствующих потоков заявок. Решением данной системы будут являться вероятности всех состояний прибора в виде функций времени. Таким образом, задача сводится к решению данной системы дифференциальных уравнений. Решить данную систему аналитически не получится, так как это система бесконечного числа дифференциальных конечно-разностных уравнений с переменными коэффициентами. Для того, чтобы перейти к конечному числу уравнений,

введем частные характеристические функции, обозначив $j = \sqrt{-1}$,

$$H_k(u, u_1, u_2, t) = \sum_{i=0}^{\infty} \sum_{m_1=0}^{\infty} \sum_{m_2=0}^{\infty} e^{jui} e^{ju_1 m_1} e^{ju_2 m_2} P_k(i, m_1, m_2, t).$$

Тогда перепишем систему (1) в виде

$$\begin{aligned} \frac{\partial H_0(u, u_1, u_2, t)}{\partial t} &= -(\lambda + \alpha) H_0(u, u_1, u_2, t) + j\sigma \frac{\partial H_0(u, u_1, u_2, t)}{\partial u} + \\ &\quad + \mu_1 e^{ju_1} H_1(u, u_1, u_2, t) + \mu_2 e^{ju_2} H_2(u, u_1, u_2, t), \\ \frac{\partial H_1(u, u_1, u_2, t)}{\partial t} &= -(\lambda + \mu_1) H_1(u, u_1, u_2, t) - j\sigma e^{-ju} \frac{\partial H_0(u, u_1, u_2, t)}{\partial u} + \\ &\quad + \lambda H_0(u, u_1, u_2, t) + \lambda e^{ju} H_1(u, u_1, u_2, t), \\ \frac{\partial H_2(u, u_1, u_2, t)}{\partial t} &= -(\lambda + \mu_2) H_2(u, u_1, u_2, t) + \lambda e^{ju} H_2(u, u_1, u_2, t) + \\ &\quad + \alpha H_0(u, u_1, u_2, t). \end{aligned} \tag{2}$$

Таким образом, мы получили ровно три дифференциальных уравнения в частных производных с переменными коэффициентами.

4 Метод асимптотического анализа

Полученную систему дифференциальных уравнений в частных производных (2) будем решать методом асимптотического анализа в предельном условии большой задержки заявок на орбите ($\sigma \rightarrow 0$).

Обозначим $\epsilon = \sigma$, $u = \epsilon w$, $F_k(w, u_1, u_2, t, \epsilon) = H_k(u, u_1, u_2, t)$, тогда система запишется в виде

$$\begin{aligned} \frac{\partial F_0(w, u_1, u_2, t, \epsilon)}{\partial t} &= -(\lambda + \alpha)F_0(w, u_1, u_2, t, \epsilon) + j \frac{\partial F_0(w, u_1, u_2, t, \epsilon)}{\partial w} + \\ &\quad + \mu_1 e^{ju_1} F_1(w, u_1, u_2, t, \epsilon) + \mu_2 e^{ju_2} F_2(w, u_1, u_2, t, \epsilon), \\ \frac{\partial F_1(w, u_1, u_2, t, \epsilon)}{\partial t} &= -(\lambda + \mu_1)F_1(w, u_1, u_2, t, \epsilon) - j e^{-j\epsilon w} \frac{\partial F_0(w, u_1, u_2, t, \epsilon)}{\partial w} + \\ &\quad + \lambda F_0(w, u_1, u_2, t, \epsilon) + \lambda e^{j\epsilon w} F_1(w, u_1, u_2, t, \epsilon), \\ \frac{\partial F_2(w, u_1, u_2, t, \epsilon)}{\partial t} &= -(\lambda + \mu_2)F_2(w, u_1, u_2, t, \epsilon) + \lambda e^{j\epsilon w} F_2(w, u_1, u_2, t, \epsilon) + \\ &\quad + \alpha F_0(w, u_1, u_2, t, \epsilon). \end{aligned} \quad (3)$$

Затем, что используя условие согласованности многомерных распределений, характеристическая функция процессов $m_1(t)$ и $m_2(t)$ будет записана в следующем виде с введенными функциями

$$M\{\exp(ju_1 m_1(t)) \exp(ju_2 m_2(t))\} = \sum_{k=0}^2 H_k(0, u_1, u_2, t) = \sum_{k=0}^2 F_k(0, u_1, u_2, t, \epsilon).$$

Теорема 1. *Асимптотическое приближение двумерной характеристической функции числа обслуженных заявок входящего потока и числа обслуженных вызванных заявок за некоторое время t имеет вид*

$$\begin{aligned} \mathbf{F}(u_1, u_2, t) &= \lim_{\sigma \rightarrow 0} M\{\exp(ju_1 m_1(t)) \exp(ju_2 m_2(t))\} = \\ &= \lim_{\epsilon \rightarrow 0} \sum_{k=0}^2 F_k(0, u_1, u_2, t, \epsilon) = \mathbf{R} \cdot \exp\{G(u_1, u_2)t\} \cdot \mathbf{E} \end{aligned}$$

где

$$\mathbf{G}(u_1, u_2) = \begin{bmatrix} -(\lambda + \alpha + \kappa) & \mu_1 e^{ju_1} & \mu_2 e^{ju_2} \\ \kappa + \lambda & -\mu_1 & 0 \\ \alpha & 0 & -\mu_2 \end{bmatrix}^T,$$

вектор-строка $\mathbf{R} = \{R_0, R_1, R_2\}$ - стационарное распределение вероятности состояния прибора

$$\mathbf{R} = \left\{ \frac{\mu_2(\mu_1 - \lambda)}{\mu_1(\mu_2 - \alpha)}, \frac{\lambda}{\mu_1}, \frac{\alpha(\mu_1 - \lambda)}{\mu_1(\mu_2 + \alpha)} \right\},$$

κ - нормированное среднее число заявок на орбите

$$\kappa = \frac{\lambda(\lambda\mu_2 + \alpha\mu_1)}{\mu_2(\mu_1 - \lambda)},$$

а \mathbf{E} - единичный вектор-столбец соответствующей размерности.

Доказательство. Делая предельный переход $\lim_{\epsilon \rightarrow 0} F_k(w, u_1, u_2, t, \epsilon) = F_k(w, u_1, u_2, t)$ в полученной системе (3), система уравнений будет записана в виде

$$\begin{aligned}
\frac{\partial F_0(w, u_1, u_2, t)}{\partial t} &= -(\lambda + \alpha)F_0(w, u_1, u_2, t) + j \frac{\partial F_0(w, u_1, u_2, t)}{\partial w} + \\
&\quad + \mu_1 e^{ju_1} F_1(w, u_1, u_2, t) + \mu_2 e^{ju_2} F_2(w, u_1, u_2, t), \\
\frac{\partial F_1(w, u_1, u_2, t)}{\partial t} &= -(\lambda + \mu_1)F_1(w, u_1, u_2, t) - j \frac{\partial F_0(w, u_1, u_2, t)}{\partial w} + \\
&\quad + \lambda F_0(w, u_1, u_2, t) + \lambda F_1(w, u_1, u_2, t), \\
\frac{\partial F_2(w, u_1, u_2, t)}{\partial t} &= -(\lambda + \mu_2)F_2(w, u_1, u_2, t) + \lambda F_2(w, u_1, u_2, t) + \\
&\quad + \alpha F_0(w, u_1, u_2, t).
\end{aligned} \tag{4}$$

Решение системы (4) будет получено в следующей форме

$$F_k(w, u_1, u_2, t) = \Phi(w) F_k(u_1, u_2, t). \tag{5}$$

$\Phi(w)$ - асимптотическое приближение характеристической функции числа заявок на орбите при условии большой задержки на орбите.

Подставив (5) в систему (4) и разделив обе части уравнений на $\Phi(w)$, получим

$$\begin{aligned}
\frac{\partial F_0(u_1, u_2, t)}{\partial t} &= -(\lambda + \alpha)F_0(u_1, u_2, t) + j \frac{\Phi'(w)}{\Phi(w)} F_0(u_1, u_2, t) + \\
&\quad + \mu_1 e^{ju_1} F_1(u_1, u_2, t) + \mu_2 e^{ju_2} F_2(u_1, u_2, t), \\
\frac{\partial F_1(u_1, u_2, t)}{\partial t} &= -(\lambda + \mu_1)F_1(u_1, u_2, t) - j \frac{\Phi'(w)}{\Phi(w)} F_0(u_1, u_2, t) + \\
&\quad + \lambda F_0(u_1, u_2, t) + \lambda F_1(u_1, u_2, t), \\
\frac{\partial F_2(u_1, u_2, t)}{\partial t} &= -(\lambda + \mu_2)F_2(u_1, u_2, t) + \lambda F_2(u_1, u_2, t) + \\
&\quad + \alpha F_0(u_1, u_2, t).
\end{aligned} \tag{6}$$

Заметим, что w содержится только в отношении $\frac{\Phi'(w)}{\Phi(w)}$, а остальные слагаемые и левые части уравнений не зависят от w . Это означает, что $\Phi(w)$ имеет вид экспоненты. Учитывая, что $\Phi(w)$ имеет смысл асимптотического приближения характеристической функции числа заявок на орбите, мы можем конкретизировать вид данной функции

$$\frac{\Phi'(w)}{\Phi(w)} = \frac{e^{j\kappa w} j\kappa}{e^{j\kappa w}},$$

где κ - нормированное среднее число заявок на орбите, которое было получено в [10] и имеет вид

$$\kappa = \frac{\lambda(\lambda\mu_2 + \alpha\mu_1)}{\mu_2(\mu_1 - \lambda)}.$$

Исходя из этого, система (6) примет следующий вид

$$\begin{aligned}
\frac{\partial F_0(u_1, u_2, t)}{\partial t} &= -(\lambda + \alpha + \kappa)F_0(u_1, u_2, t) + \\
&\quad + \mu_1 e^{ju_1} F_1(u_1, u_2, t) + \mu_2 e^{ju_2} F_2(u_1, u_2, t), \\
\frac{\partial F_1(u_1, u_2, t)}{\partial t} &= (\lambda + \kappa)F_0(u_1, u_2, t) - \mu_1 F_1(u_1, u_2, t) + \\
&\quad + 0F_2(u_1, u_2, t), \\
\frac{\partial F_2(u_1, u_2, t)}{\partial t} &= \alpha F_0(u_1, u_2, t) + 0F_1(u_1, u_2, t) - \\
&\quad - \mu_2 F_2(u_1, u_2, t).
\end{aligned} \tag{7}$$

Введем следующие обозначения

$$\begin{aligned}
\mathbf{F}(u_1, u_2, t) &= \{F_0(u_1, u_2, t), F_1(u_1, u_2, t), F_2(u_1, u_2, t)\} \\
\mathbf{G}(u_1, u_2) &= \begin{bmatrix} -(\lambda + \alpha + \kappa) & \mu_1 e^{ju_1} & \mu_2 e^{ju_2} \\ \kappa + \lambda & -\mu_1 & 0 \\ \alpha & 0 & -\mu_2 \end{bmatrix}^T,
\end{aligned}$$

$\mathbf{G}(u_1, u_2)$ - транспонированная матрица коэффициентов системы (7). Тогда получим следующее матричное уравнение

$$\frac{\partial \mathbf{F}(u_1, u_2, t)}{\partial t} = \mathbf{F}(u_1, u_2, t) \mathbf{G}(u_1, u_2),$$

общее решение которого имеет вид

$$\mathbf{F}(u_1, u_2, t) = \mathbf{C} e^{\mathbf{G}(u_1, u_2)t}. \tag{8}$$

Для того, чтобы получить единственное решение, которое соответствует поведению рассматриваемой системы, примем в рассмотрение начальное условие

$$\mathbf{F}(u_1, u_2, 0) = \mathbf{R}, \tag{9}$$

где вектор-строка \mathbf{R} - стационарное распределение вероятности состояния прибора, то есть процесса $k(t)$, которое имеет форму [10]

$$\mathbf{R} = \left\{ \frac{\mu_2(\mu_1 - \lambda)}{\mu_1(\mu_2 - \alpha)}, \frac{\lambda}{\mu_1}, \frac{\alpha(\mu_1 - \lambda)}{\mu_1(\mu_2 + \alpha)} \right\}.$$

Описав начальное условие, мы можем перейти к решению задачи Коши (8, 9).

Поскольку нас интересует распределение вероятностей количества заявок в выходных процессах, необходимо найти маргинальное распределение. Для этого суммируем компоненты вектор-строки $\mathbf{F}(u_1, u_2, t)$ по k и умножаем результат на единичный вектор-столбец \mathbf{E} . Получим

$$\mathbf{F}(u_1, u_2, t) \mathbf{E} = \mathbf{R} e^{\mathbf{G}(u_1, u_2)t} \mathbf{E}. \tag{10}$$

Эта формула позволяет найти асимптотическое приближение характеристической функции количества вызванных и входящих заявок, обслуженных системой к некоторому моменту времени t . Другими словами, формула (10) является решением рассматриваемой системы. \square

Список литературы

- [1] Artalejo, J.R., Gómez-Corral, A.: Retrial Queueing Systems: A Computational Approach. Springer-Verlag Berlin Heidelberg (2008)
- [2] Bronson, R.: Matrix methods: An introduction. Gulf Professional Publishing (1991)
- [3] Burke, P.: The output process of a stationary m/m/s queueing system. The Annals of Mathematical Statistics **39**(4), 1144–1152 (1968)
- [4] Daley, D.: Queueing output processes. Advances in Applied Probability **8**(2), 395–415 (1976)
- [5] Gharbi, N., Dutheillet, C.: An algorithmic approach for analysis of finite-source retrial systems with unreliable servers. Computers & Mathematics with Applications **62**(6), 2535–2546 (2011)
- [6] Kendall, D.G.: Stochastic processes occurring in the theory of queues and their analysis by the method of the imbedded markov chain. The Annals of Mathematical Statistics pp. 338–354 (1953)
- [7] Kulkarni, V.G.: On queueing systems with retrials. Journal of Applied Probability pp. 380–389 (1983)
- [8] Lapatin, I., Nazarov, A.: Asymptotic analysis of the output process in retrial queue with markov-modulated poisson input under low rate of retrials condition. In: International Conference on Distributed Computer and Communication Networks. pp. 315–324. Springer (2019)
- [9] Mirasol, N.M.: The output of an m/g/ ∞ queueing system is poisson. Operations Research **11**(2), 282–284 (1963)
- [10] Nazarov, A.A., Paul, S., Gudkova, I., et al.: Asymptotic analysis of markovian retrial queue with two-way communication under low rate of retrials condition (2017)
- [11] Paul, S., Phung-Duc, T.: Retrial queueing model with two-way communication, unreliable server and resume of interrupted call for cognitive radio networks. In: Information Technologies and Mathematical Modelling. Queueing Theory and Applications, pp. 213–224. Springer (2018)