

Федеральное государственное бюджетное образовательное учреждение
высшего образования «Казанский национальный исследовательский
технологический университет»

На правах рукописи



ХАСАНОВ НИЯЗ АДЕЛЕВИЧ

МОДЕЛИРОВАНИЕ ЗАМКНУТЫХ СИСТЕМ МАССОВОГО
ОБСЛУЖИВАНИЯ С ТРЕХКОМПОНЕНТНЫМ ПОТОКОМ ЗАЯВОК И
ОГРАНИЧЕНИЕМ ПО ВРЕМЕНИ ОЖИДАНИЯ ЗАЯВКИ В ОЧЕРЕДИ

1.2.2. Математическое моделирование,
численные методы и комплексы программ

ДИССЕРТАЦИЯ
на соискание учёной степени
кандидата технических наук

Научный руководитель:
доктор технических наук, доцент
Титовцев Антон Сергеевич

Казань – 2025

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ	5
1 ЛИТЕРАТУРНЫЙ ОБЗОР И ПОСТАНОВКА ЗАДАЧИ ИССЛЕДОВАНИЯ	16
1.1 Литературный обзор.....	16
1.2 Постановка задачи.....	22
Выводы:	28
2 МАТЕМАТИЧЕСКАЯ ФОРМАЛИЗАЦИЯ МОДЕЛИ ЗАМКНУТОЙ СИСТЕМЫ МАССОВОГО ОБСЛУЖИВАНИЯ С ОГРАНИЧЕНИЕМ ПО ВРЕМЕНИ ОЖИДАНИЯ ЗАЯВКИ В ОЧЕРЕДИ.....	29
2.1 Вероятностные характеристики стационарного режима СМО	29
2.2 Числовые характеристики стационарного режима СМО.....	32
2.3 Временные характеристики СМО	35
Выводы:	38
3 НЕСТАЦИОНАРНЫЙ РЕЖИМ РАБОТЫ СМО С ТРЕХКОМПОНЕНТНЫМ ПОТОКОМ ЗАЯВОК И ОГРАНИЧЕНИЕМ ПО ВРЕМЕНИ ОЖИДАНИЯ ЗАЯВКИ В ОЧЕРЕДИ.....	39
3.1 Имитационное моделирование	39
3.2 Метод производящих функций	44
Выводы:	53
4 ЧИСЛЕННОЕ МОДЕЛИРОВАНИЕ РЕЖИМОВ РАБОТЫ СМО С ТРЕХКОМПОНЕНТНЫМ ПОТОКОМ ЗАЯВОК И ОГРАНИЧЕНИЕМ ПО ВРЕМЕНИ ОЖИДАНИЯ ЗАЯВКИ В ОЧЕРЕДИ.....	54
4.1 Постановка задачи.....	54
4.2 Численная оценка и анализ параметров СМО	55
4.3 Показатели загрузки системы	58
4.4 Пропускная способность СМО.....	74
4.5 Временные характеристики СМО	79
Выводы:	85

5 ПОИСК ЭФФЕКТИВНЫХ РЕЖИМОВ РАБОТЫ ЗАМКНУТОЙ СМО С ОГРАНИЧЕННЫМ ВРЕМЕНЕМ ОЖИДАНИЯ И ПОЛИКОМПОНЕНТНЫМ ПОТОКОМ ЗАЯВОК	86
5.1 Постановка задачи.....	86
5.2 Исследование эффективных режимов СМО по параметру относительной пропускной способности q при изменении параметров μ и $\lambda_{0,1}$	87
5.3 Анализ влияния интенсивности ухода заявок (ν) на фиксированный параметр относительной пропускной способности q	101
5.4 Исследование эффективных режимов СМО по параметру относительной пропускной способности при изменении параметров μ и m	105
5.5 Исследование эффективных режимов СМО по параметру относительной пропускной способности при изменении параметров $\lambda_{0,1}$ и m	109
5.6 Исследование эффективных режимов СМО по параметру среднего времени ожидания при изменении параметров m и λ_0, λ_1	115
5.7 Исследование эффективных режимов СМО по параметру среднего времени ожидания при изменении параметров m и μ	118
5.8 Исследование эффективных режимов СМО по параметру среднего времени ожидания при изменении параметров m и ν	121
5.9 Исследование эффективных режимов СМО по параметру среднего времени ожидания при изменении параметров μ и ν	124
Выводы:	127
6 ПРАКТИЧЕСКИЕ РЕКОМЕНДАЦИИ ПО ВНЕДРЕНИЮ РЕЗУЛЬТАТОВ ДИССЕРТАЦИОННОЙ РАБОТЫ	129
6.1 Общие положения	129
6.2 Проверка соответствия входного потока заявок пуассоновскому потоку	130
6.3 Проверка соответствия времени обслуживания экспоненциальному распределению.....	134
6.4 Проверка соответствия ограничения на время ожидания экспоненциальному распределению	136
6.5 Математическое моделирование работы СМО в стационарном режиме	137

Выводы:	137
ЗАКЛЮЧЕНИЕ	139
СПИСОК ЛИТЕРАТУРЫ	141
ПРИЛОЖЕНИЯ	154
Приложение 1	155
Приложение 2	161
Приложение 3	162

ВВЕДЕНИЕ

Актуальность темы исследования. В современную эпоху цифровой трансформации вопросы, связанные с системами массового обслуживания (СМО), приобретают особую актуальность и значимость. В соответствии с указом Президента Российской Федерации № 145 «О стратегии научно-технического развития Российской Федерации», в соответствии с п. 21 а, е, и п. 22 стратегии, развитие современных интеллектуальных телекоммуникационных систем относится к одному из приоритетных направлений научно-технического развития страны.

Цифровизация и автоматизация процессов приводят к существенному увеличению нагрузки на системы обработки запросов, что требует разработки новых подходов и методов эффективного управления. Это особенно критично для таких областей, как логистика, где необходимо эффективно управлять потоками товаров, телекоммуникации, где требуется распределение потоков в высокоскоростных сетях, и промышленность, где важна балансировка ресурсов для обеспечения непрерывной работы с требуемым уровнем нагрузки.

Современные технологические системы требуют эффективного управления ресурсами. Одним из ключевых направлений является минимизация времени простоя и потерь из-за перегрузок. Это обуславливает актуальность разработки математических моделей, учитывающих разнородные требования к обслуживанию. Постановка таких задач часто связана с накладыванием разнородных ограничений на количество обслуживающих устройств и время обслуживания. В качестве примера можно привести ожидание транспорта или ответа оператора, очереди в торговых центрах, движение автомобилей на дорогах и т.д. Управление очередями остается актуальным для всех уровней: от уровня индивидуума до корпораций. Для предприятий различных отраслей минимизация задержек в обслуживании является ключевым фактором повышения рентабельности и

удержания клиентской базы. Решением подобных задач с применением научного подхода занимается теория массового обслуживания.

Наиболее изученными в теории массового обслуживания (ТМО) являются СМО открытого типа с упрощенными предположениями о поведении заявок, в то время как СМО с ограниченным источником остаются значительно менее исследованными, особенно в части разработки математических моделей. В частности, для замкнутых СМО отсутствуют комплексные модели, адекватно описывающие системы, в которых заявки демонстрируют разнородное поведение: часть готова ждать, часть требует немедленного обслуживания, а часть уходит из очереди при превышении заданного временного порога. Зачастую анализ и конфигурация подобных систем вынужденно опираются на упрощенные представления о поведении заявок и системных ограничениях. Однако, такие допущения не всегда соответствуют реальным условиям. Сочетание свойств замкнутости системы, ограничения на время ожидания и трехкомпонентного потока является ключевой особенностью многих существующих систем, но слабо отражено в существующих исследованиях. Отсутствие аналитических решений для замкнутых СМО с таким комплексом ограничений создает значительный разрыв между теорией и практикой, вследствие чего затрудняется проектирование и эффективное управление соответствующими системами. На практике это вынуждает использовать либо упрощенные модели, дающие большую погрешность, либо прибегать к трудозатратному имитационному моделированию для каждой отдельной конфигурации.

В данной работе впервые представлена математическая модель СМО с разнородным характером входного потока, ограничением на ресурс источника и на время ожидания в очереди. Подобного рода системы могут быть использованы для мониторинга производственных процессов (станки, прессы, автоматизированные линии, системы контроля качества), в IT инфраструктуре (облачные вычисления, системы облачного сканирования данных), в логистике и транспорте, в медицинских учреждениях и

телекоммуникациях, call-центрах, а значит, разработка комплекса инструментов для анализа характеристик замкнутой СМО с ограничением на время ожидания и трехкомпонентным потоком заявок предоставит недостающий научно обоснованный инструмент для оценки функционирования и конфигурирования подобных систем.

Степень разработанности темы. Первые работы в сфере массового обслуживания принадлежат датскому математику и инженеру А.К. Эрлангу, который заложил основы ТМО. Работая в международной телефонной компании, Эрланг столкнулся с проблемой обеспечения определения оптимального количества обслуживающих каналов. Будучи специалистом в области математики и статистики, А.К. Эрланг был первым практическим исследователем задач массового обслуживания и посвятил значительную часть своей жизни разработке теории телефонного трафика и исследованиям в области анализа телефонных сетей. Его работы заложили основу дальнейшего развития ТМО. В частности, А.К. Эрланг разработал формулы для определения минимального количества телефонных каналов, необходимых для обеспечения заданного уровня обслуживания. Он также ввел понятие "время ожидания" и разработал методы для оценки вероятности отказа в обслуживании. Эти работы оказали огромное влияние на развитие теории массового обслуживания и стали основой для дальнейших исследований.

Позднее, советский математик и выдающийся ученый А.Я. Хинчин в сотрудничестве со специалистами московского телефонного дела провел ряд исследований в области ТМО. Основное направление его научной деятельности связано с изучением строения потока входящих требований. Именно решению этих задач посвящена его монография и последующие статьи. В своих исследованиях А.Я. Хинчин ввел понятие пуассоновского потока случайных событий, позволяющее моделировать случайные потоки поступающих заявок в СМО. Он также разработал методы для оценки вероятности отказа в обслуживании и среднего времени ожидания.

Значительный вклад в ТМО внес А.Н. Колмогоров, разработавший теорию эргодических свойств марковских процессов, что позволило более точно моделировать поведение СМО в долгосрочной перспективе. Он также ввел понятие марковских цепей, которые стали основой для дальнейших исследований в области ТМО. Совместно с А.Я. Хинчиным, А.Н. Колмогоров разработал теорию случайных процессов, что дало возможность моделировать случайные процессы в СМО.

Строгое математическое изложение основ ТМО с применением численных методов впервые было изложено в работах Б.В. Гнеденко и Б.В. Коваленко. Применение основ ТМО для моделирования простейших марковских процессов в различных областях экономики описаны в трудах Л.Г. Лабскера и Л.О. Бабешко. Л.Г. Лабскер и Л.О. Бабешко также представили разработанные для персонального компьютера программы, позволяющие производить расчеты простейших СМО.

В настоящее время классические модели $M/M/n$ широко используются для описания однородных потоков заявок, не учитывая при этом возможность наличия в потоке заявок разного типа. Последние исследования, в частности, работы А.П. Кирпичникова рассматривают системы с ограничениями по времени ожидания, но они не включают гетерогенные потоки заявок. Таким образом, данная работа «заполняет этот пробел», предлагая математическую модель замкнутой системы с трехкомпонентным потоком заявок и ограничением по времени ожидания.

Цель и задачи исследования.

Целью настоящего исследования является разработка математических моделей, алгоритма и комплекса программных средств для анализа характеристик замкнутых СМО с трехкомпонентным потоком заявок и ограничением по времени ожидания заявки в очереди.

Задачи исследования:

1. Разработать математическую модель замкнутой СМО с учетом трех типов входящих заявок и ограничения по времени ожидания заявки в очереди, получить в общем виде формулы для расчета характеристик модели.

2. Разработать численный алгоритм исследования СМО на предмет определения параметров модели, обеспечивающих требуемый уровень эффективности функционирования системы, а также разработать рекомендации для применения и адаптации полученной модели к реальным объектам.

3. Разработать имитационную модель замкнутой СМО с учетом трех типов входящих заявок и ограничения по времени ожидания заявки в очереди.

4. Разработать комплекс программных средств и провести вычислительные эксперименты, необходимые для оценки характеристик СМО.

Объектом исследования являются СМО, функционирование которых предполагает наличие очередей, к примеру call-центры, логистические комплексы, телекоммуникационные сети, контрольно-пропускные пункты, сборочные производства и пр.

Предметом исследования являются модели, алгоритмы и методы, описывающие замкнутые СМО, в которых поступающие заявки подразделяются на три типа: заявки, требующие немедленного обслуживания и покидающие систему при отсутствии свободных каналов, заявки, дожидаящиеся обслуживания до конца, и заявки, у которых времени ожидания ограничено.

Соответствие исследования паспорту научной специальности. Область диссертационного исследования соответствует следующим пунктам паспорта научной специальности 1.2.2. Математическое моделирование, численные методы и комплексы программ (технические науки): п.6. Разработка систем компьютерного и имитационного моделирования, алгоритмов и методов имитационного моделирования на основе анализа

математических моделей (технические науки); п.8. Комплексные исследования научных и технических проблем с применением современной технологии математического моделирования и вычислительного эксперимента; п.9. Постановка и проведение численных экспериментов, статистический анализ их результатов, в том числе с применением современных компьютерных технологий.

Научная новизна представленных результатов характеризуется развитием методов математического моделирования СМО, методов и алгоритмов исследования СМО, и заключается в следующем:

1. *Разработана новая математическая модель* замкнутой СМО с трехкомпонентным входным потоком заявок, *отличающаяся* сочетанием свойств известных моделей замкнутых СМО с ограниченным временем ожидания заявки в очереди, модели Энгсета, а также классической модели многоканальной замкнутой СМО. Аналитическими методами *получены* выражения для расчета вероятностных, числовых и временных характеристик СМО в стационарном режиме функционирования, *позволяющие* получать численные значения ключевых характеристик СМО (п. 8 паспорта 1.2.2).

2. *Разработан численный алгоритм* поиска эффективных режимов работы СМО, *отличающийся* возможностью определения требуемых сочетаний вводных данных для достижения заданной эффективности функционирования СМО, критерием которой является относительная пропускная способность системы. *Выявлено*, что при добавлении во входной поток «нетерпеливых» заявок характер зависимости интенсивности обслуживания от суммарной интенсивности входного потока при фиксированной величине относительной пропускной способности меняется со степенного на линейный (п. 8 паспорта 1.2.2).

3. *Разработана имитационная модель* СМО с использованием языка программирования Python, основанная на методе Монте-Карло и

позволяющая производить вычисление характеристик системы в нестационарном режиме функционирования (п. 6 паспорта 1.2.2).

4. *Разработан комплекс специализированных программных средств для проведения вычислительных экспериментов, отличающийся возможностью расчета характеристик замкнутых СМО и оценки её эффективности (п. 9 паспорта 1.2.2).*

Теоретическая значимость исследования заключается в расширении математического аппарата ТМО. В частности, в работе представлены:

- неисследованная ранее математическая модель замкнутой СМО с трехкомпонентным потоком заявок и ограничением по времени ожидания заявки в очереди;
- математическая формализация первых и вторых моментов числовых и временных характеристик модели в стационарном режиме работы;
- методы и алгоритмы поиска эффективных режимов работы замкнутых СМО с двухкомпонентным и трехкомпонентным потоком заявок;
- предложены рекомендации по внедрению математической модели для адаптации к реальным техническим системам.

Практическая значимость. Область применения результатов работы охватывает широкий спектр отраслей, включая производственно-техническую сферу, торговлю, управление производством и производственными процессами, телефонию и связь, а также логистику. Полученные результаты применимы для систем, где разнородные типы заявок могут существенно влиять на эффективность обслуживания. В таких системах часто встречаются различные типы требований, направленных на применение различного подхода к организации обслуживания. Например, модель может быть использована для организации работы call-центров, где входящие запросы делятся на срочные (жалобы), стандартные (консультации) и отложенные (обратная связь). Также полученная математическая модель может служить основой для проведения дальнейших исследований в рамках

ТМО, что позволит разрабатывать более эффективные и надежные системы для различных областей применения.

Методология и методы исследования. В работе при построении математической модели применены методы теории вероятностей, а также теории случайных процессов и цепей Маркова. Для решения трансцендентных уравнений применялись соответствующие численные методы. Имитационное моделирование, основанное на методе Монте-Карло, осуществлялось с использованием языка программирования Python.

Положения, выносимые на защиту.

1. Математическая модель замкнутой СМО с трехкомпонентным потоком заявок и ограничением по времени ожидания заявки в очереди.
2. Разработанный алгоритм поиска сочетаний параметров математической модели СМО, обеспечивающих требуемое качество обслуживания и полученные численные результаты расчетов.
3. Имитационная математическая модель СМО, позволяющая получить численные результаты ключевых характеристик СМО в нестационарном режиме.
4. Программное обеспечение, применяемое для расчетов и изучения реакций основных характеристик СМО на изменение исходных данных и оценки ее эффективности.

Степень достоверности результатов диссертационной работы обеспечивается корректной постановкой задач, строгим выполнением математических выкладок, проверкой полученных решений на соответствие известным случаям, описанным другими авторами.

Апробация. Разработанные методы и алгоритмы внедрены в ООО «НПФ Геоник», с. Набережные Моркваши Верхнеуслонского района РТ, в части решений для анализа загруженности сервисных ресурсов и выявления факторов, ограничивающих эффективность обслуживания оборудования.

Результаты исследований используются в учебном процессе ФГБОУ ВО КНИТУ при обучении студентов по направлениям подготовки 01.03.02

«Прикладная математика и информатика» и 02.03.03 «Математическое обеспечение и администрирование информационных систем».

Результаты диссертационной работы докладывались на 1-й международной научной конференции «Школа теории массового обслуживания» (г. Томск, 2025), XXXVIII международной конференции «Математические методы в технике и технологиях – ММТТ-38» (г. Самара, 2025), научной конференции КНИТУ (г. Казань, 2023)

Личный вклад автора. Все представленные результаты получены автором лично. В зарегистрированном электронном ресурсе автору принадлежит ведущая роль в разработке программного обеспечения для расчета необходимых характеристик.

Публикации. По теме диссертации опубликовано самостоятельно и в соавторстве 9 печатных работ: 4 работы в рецензируемых научных изданиях, рекомендованных ВАК РФ, 5 – в прочих изданиях. Зарегистрирован 1 программный продукт.

Структура и объем работы. Диссертационная работа состоит из введения, шести глав, заключения, списка литературы и приложений. Полный объём диссертации составляет 162 страницы, включая 65 рисунков и 22 таблицы. Библиографический список включает 96 литературных источников.

Структура диссертации.

Во **введении** показана актуальность темы диссертационной работы, приведены её цель, задачи, научная новизна, теоретическая и практическая значимость, приводится краткий обзор глав.

В первой главе рассмотрены актуальные исследовательские работы, имеющие отношение к СМО с ограничениями, проведен обзор результатов последних исследований. Выполнена постановка задачи диссертационной работы.

Во второй построена математическая модель замкнутой СМО с трехкомпонентным потоком заявок и ограничением по времени ожидания заявки в очереди. Получены математические формулы для нахождения

вероятностных, числовых и временных характеристик СМО в стационарном режиме работы.

В третьей главе с использованием языка программирования Python и метода Монте-Карло разработана имитационная модель рассматриваемой СМО, позволяющая определять числовые характеристики СМО в нестационарном режиме. Также получено аналитическое решение системы уравнений Колмогорова для частного случая $P_N(t) \approx 0$, $\lambda_0 = 0$, $\nu = \mu$.

В четвертой главе проведено численное моделирование различных режимов работы замкнутой СМО с трехкомпонентным потоком заявок и ограничением по времени ожидания заявки в очереди в стационарном режиме. Определены вероятностные, числовые и временные характеристики СМО. Рассмотрены случаи наличия и отсутствия в потоке «нетерпеливых заявок», на основании чего сделаны выводы о влиянии 0-й компоненты потока на основные характеристики системы.

В пятой главе представлен алгоритм, использующий численный метод Ньютона с адаптивной релаксацией и итерационный метод поиска по сетке для поиска эффективных режимов работы замкнутой СМО с трехкомпонентным потоком заявок и ограничением по времени ожидания заявки в очереди. По результатам вычислительных экспериментов построены графики зависимостей для ключевых параметров системы при обеспечении требуемого уровня обслуживания по параметру относительной пропускной способности и среднему времени ожидания.

В шестой главе даны практические рекомендации, которых следует придерживаться при приложении результатов диссертационной работы к реальным объектам.

В заключении в структурированном виде приводятся результаты проделанной работы и возможные перспективы дальнейшей разработки темы.

В приложении приводятся пример оценки соответствия входного потока распределению Пуассона и времени обслуживания

экспоненциальному распределению, акт о внедрении результатов диссертационной работы, а также свидетельство о регистрации электронного ресурса.

1 ЛИТЕРАТУРНЫЙ ОБЗОР И ПОСТАНОВКА ЗАДАЧИ ИССЛЕДОВАНИЯ

1.1 Литературный обзор

Теория массового обслуживания (далее – ТМО), являясь разделом прикладной теории вероятностей, изучает стохастические процессы в системах, где возникают конкуренция за ресурсы между поступающими требованиями (заявками) и их обслуживанием. Одной из ключевых задач ТМО является синтез оптимальных структур реальных систем на основе анализа характеристик входных потоков, дисциплин обслуживания и ограничений. Например, внедрение инновационных автоматизированных систем управления транспортом в мегаполисах позволяет сократить среднее время ожидания маршрутных автобусов почти вдвое. Такая оптимизация не только повышает эффективность пассажироперевозок, но и приводит к снижению выбросов углекислого газа автобусными парками на 40–50 % [1].

Несмотря на то, что в ТМО используются методы теории вероятностей и математической статистики, данная дисциплина имеет свою специфику и терминологию, способствующую грамотной формулировке обзореваемых задач. При исследовании операций в рамках данной теории нередко встречаются системы, предназначенные для решения широкого класса задач. Такие системы называют системами массового обслуживания (далее – СМО).

Для описания характеристик эффективности функционирования СМО используются различные параметры и функции, в частности, средний процент необслуженных заявок, среднее время простоя, среднее время ожидания обслуживания, вероятность немедленного обслуживания, закон распределения длины очереди и др. Каждая из этих характеристик несет в себе информацию о пропускной способности системы.

Абсолютная пропускная способность – метрическая характеристика, показывающая предельное количество заявок, которое система может обслужить в единицу времени. Выделяют также относительную пропускную

способность – отношение числа обслуженных заявок к числу поступающих заявок.

Для детерминированных систем, где моменты поступления заявок и длительности обслуживания заданы явно, расчет пропускной способности сводится к решению алгебраических уравнений. Однако в реальных системах случайный характер событий, описываемый распределениями случайных величин, требует применения вероятностных методов, таких как анализ марковских цепей. Нагруженность такой системы в разные моменты времени различна. В результате чего, в зависимости от выбранного временного промежутка, в системе могут образоваться очереди или простой оборудования. Таким образом, функционирование СМО носит случайный характер и является случайным процессом. В таком случае для расчета характеристик данной системы необходимо изучить и описать ее математически с использованием элементов теории вероятности, чем в свою очередь и занимается ТМО.

Немаловажным элементом математического представления СМО является ее математическая модель. Принято разделять модели СМО на марковские и немарковские. Наиболее изученными являются марковские модели [2-6], т.е. такие модели процесс функционирования которых относят к классу марковских процессов.

Каждая СМО состоит из обслуживающих устройств (устройств, пунктов, аппаратов и т.д.), называемых каналами обслуживания и входящего потока требований на обслуживание. По числу используемых для обслуживания каналов СМО подразделяются на одноканальные и многоканальные. Периодичность поступления заявок в систему в наиболее распространенном и изученном случае описывается пуассоновским потоком [7-9], который обладает свойствами стационарности и ординарности, а также отсутствием последействия. Системы с более сложными потоками заявок [10-14] практически не изучены и для них получены лишь частные решения.

Классификация СМО на открытые и замкнутые определяется поведением заявок после завершения обслуживания. В **открытых системах** заявки покидают систему окончательно, что характерно, к примеру, для серверов обработки запросов. В **замкнутых системах** (например, производственные циклы с повторяющимися операциями) заявки возвращаются в источник и могут повторно поступать на обслуживание [15-18], образуя циклический процесс.

В связи с тем, что класс существующих СМО весьма обширен, в 1953 г. Д. Г. Кендалл предложил описывать классы СМО в соответствии с разработанной им нотацией [19]. Классификация (нотация) Кендалла состоит в общем случае из пяти элементов: $X/Y/s/q/D$, где X и Y – характеризуют распределение входящего потока заявок и времени обслуживания соответственно, s обозначает количество обслуживающих устройств, q – размер накопителя, D – дисциплина обслуживания, указывается в случаях отличных от FIFO («Первым пришел – первым ушел»).

В рамках классификации Кендалла наиболее изученными являются системы с пуассоновским входящим потоком заявок ($X=M$) и экспоненциальным распределением времени обслуживания ($Y=M$). Например, модель $M/M/1$ описывает систему с одним каналом обслуживания, где заявки поступают по пуассоновскому потоку и время обслуживания распределено по экспоненциальному закону.

В СМО с организацией очереди, заявка, поступающая в систему при занятых каналах обслуживания, переходит в состояние ожидания. Однако, в условиях современных приложений ТМО к реальным объектам (например, в телекоммуникациях или производственных линиях) возникает необходимость учета дополнительных ограничений, которые не были заложены в классические модели, описанные в трудах А.К. Эрланга [20-22]. Одним из актуальных направлений модернизации СМО стало введение временных и емкостных ограничений на очередь. В зависимости от организации ожидания выделяют два типа ограничений: ограничение по времени ожидания и

ограничение по размеру очереди. В случае СМО с ограничениями по времени ожидания заявка покидает систему, если время пребывания в очереди превышает заданное значение. СМО с ограничением по длине очереди предполагает фиксированную емкость накопителя: при его заполнении новые заявки получают отказ. Впервые такие модели были рассмотрены в исследованиях Дж. Козна [23-34] и впоследствии получили название «СМО смешанного типа».

Случай, когда в рамках одной СМО действуют оба эти ограничения рассмотрен в диссертационной работе Нгуена Т.Б. [35]. Для математической формализации вероятностных характеристик СМО автором введена неполная функция Г. Миттаг–Леффлера первого порядка, применяемая для суммирования конечных рядов, не сводящихся к геометрическим прогрессиям. В результате проведения численных экспериментов установлено, что квазистационарный режим функционирования в СМО рассматриваемого типа устанавливается приблизительно в два раза быстрее, чем у аналогичных СМО без ограничений на длину накопителя. Полученные в диссертационной работе результаты, могут быть использованы для решения задач управления СМО с ограничениями, наложенными как на предельный объём накопителя, так и на среднее время пребывания заявок в очереди на обслуживание.

На некоторых объектах, работающих по принципу СМО, входящий поток заявок может содержать различные типы заявок, отличающихся по дисциплине обслуживания и интенсивности. Модель такой СМО будет отличаться от классической. Методы исследования подобных открытых моделей СМО и их характеристик рассмотрены в работе И. Н. Валеева [36]. Ввиду неоднородности компонентов потока входящих заявок, автором предложена новая структура обслуживающей системы в виде двух различных групп обслуживающих устройств. Выявлено, что квазистационарный режим для моделей селективного обслуживания с двухкомпонентным и трехкомпонентным потоками заявок устанавливается

приблизительно за 100000 и 1000 единиц модельного времени соответственно, равных среднему времени обслуживания заявки одним обслуживающим устройством. В рамках работы проведено исследование вероятностных, числовых и временных характеристик стационарного и нестационарного режимов работы открытых СМО с поликомпонентным входным потоком заявок, разработаны численные алгоритмы и методы организации обслуживания.

В силу ограниченности ёмкости накопителя заявки разных типов имеют различные ограничения на длину очереди, зависящие от интенсивностей входных потоков заявок и определяющиеся общим количеством заявок всех типов, которое может вместить накопитель в каждом конкретном случае. Например, в компьютерных сетях с коммутацией пакетов заявками являются передаваемые пакеты данных, длина которых может варьироваться в широких пределах. Похожую картину можно наблюдать, рассматривая объекты транспортной инфраструктуры.

Работа А.С. Титовцева [37] направлена на подробное рассмотрение подобных систем. Разработана многоканальная модель массового обслуживания смешанного типа с трехкомпонентным потоком требований, которая сочетает в себе свойства известных многоканальных моделей Эрланга, классической, а также СМО с ограниченной очередью. Введено фундаментальное понятие поликомпонентного потока. Предложена концепция очередей высших порядков, в рамках которой разработан общий алгоритм вычисления моментов очередей N -го порядка на базе математической модели. Обобщены формулы Литтла и выделено 2 класса очередей высших порядков по скорости движения заявок в них. Установлен линейный характер поведения граничных значений числа обслуживающих устройств, соответствующих стабильному режиму по длине очереди и времени ожидания, с изменением приведенной интенсивности входного потока требований.

Немаловажным фактором при проектировании СМО является дисциплина обслуживания. Наиболее распространенными дисциплинами обслуживания в рамках изученных на данный момент моделей СМО являются дисциплины обслуживания FIFO и обслуживание в обратном порядке LIFO («Последним пришел – первым ушел»). Также возможны системы со случайным выбором требования для обслуживания. На реальных объектах задача выбора заявки на обслуживание не всегда так проста.

В работе Н.А. Смирновой [38] представлена модель СМО, в которой учитывается приоритет входящих заявок применительно к реальному производству. Предложен способ описания приоритетов заявок в очереди многомерным случайным вектором. Разработана имитационная модель технологических процессов с учетом приоритета обслуживания в зависимости от сочетания нескольких случайных факторов.

К настоящему времени при анализе многих реальных систем замечено, что использование моделей СМО с марковским входящим потоком не всегда приемлемо. К тому же среди новейших инфотелекоммуникационных систем встречаются системы, требующие применение моделей СМО с полумарковским потоком, обобщающим рекуррентный и марковский потоки.

В работе В.В. Чаплыгина [39] рассмотрены некоторые математические методы расчета стационарных характеристик немарковских моделей СМО, в частности $SM/MSP/n/r$, $SM/MSP/n/\infty$, $G/MSP/n/r$, $G/MSP/n/\infty$, $SM/MSP/1/r$, $G/BMSP/1/r$. Разработаны алгоритмы для вычисления основных стационарных характеристик систем по полученным математическим соотношениям.

Проблемами изучения систем массового обслуживания в настоящее время в нашей стране занимается ряд ученых: Т.И. Алиев, С.П. Моисеева, А.Н. Моисеев, А.А. Назаров, С.В. Пауль, В.А. Романенко, В.В. Рыков, К.Е. Самуйлов, С.П. Сущенко, В.Н. Тарасов, А.С. Титовцев, М.П. Фархадов, М.А. Шнепс-Шнеппе и др. [40-61].

1.2 Постановка задачи

В качестве теоретического базиса разрабатываемой модели был выбран марковский случайный процесс $x(t)$, описывающий последовательности событий, где вероятность наступления каждого последующего события в момент времени $t+\Delta t$ зависит только от текущего состояния системы в момент времени t для любой последовательности моментов времени $t_0, t_1, \dots, t+\Delta t$, такой, что, $t_0 < t_1 < \dots < t+\Delta t$.

Цепи Маркова находят применение в самых различных областях, включая статистику [62, 63], экономику [64-66], инженерные науки [67-70] и биологию [71-73]. В рамках моделирования очередей марковские цепи используются для анализа СМО телефонных станций, серверов, производственных линий и др.

Основной компонент цепи Маркова – это состояние системы в данный момент времени $x \in \{0, 1, \dots, N\}$, которое может изменяться с определенной вероятностью. Такой подход позволяет моделировать и анализировать сложные системы, где взаимодействие между элементами может быть случайным, но при этом подчиняется определенным закономерностям.

Для описания дискретного входящего потока требований выбран пуассоновский поток случайных событий, т.к. он эффективно моделирует случайные и независимые события и позволяет адекватно описать процесс поступления требований в систему в рамках марковского процесса. Вероятность поступления в СМО k заявок за время t для простейшего (пуассоновского) потока подчиняется закону:

$$p_k(t) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}.$$

Вероятность того, что прибытие заявки произойдет за время t :

$$P(t) = 1 - e^{-\lambda t}.$$

Тогда плотность распределения вероятности случайной величины T_λ , описывающей время между поступлениями двух соседних заявок в систему, задается формулой:

$$f_{T_\lambda}(\tilde{t}) = \lambda e^{-\lambda \tilde{t}},$$

где \tilde{t} – фактическое значение времени между поступлением заявок.

Следует отметить, что в рамках данной работы рассматриваемое время обслуживания и предельное время ожидания предполагаются распределенными по экспоненциальному закону. Тогда, с учетом количества станций обслуживания m , вероятность обработки системой j заявок за время t подчиняется закону Пуассона:

$$G_{ij}(t) = \frac{((m\mu + i\nu)t)^j}{j!} e^{-(m\mu + i\nu)t}, \quad (1.1)$$

где i – текущая длина очереди;

μ – интенсивность обслуживания заявки одним обслуживающим устройством;

ν – интенсивность ухода заявок по истечении предельного времени ожидания.

Результатом выступает марковская цепь с дискретным числом состояний и непрерывным временем, в которой навстречу поступающим заявкам движется поток обслуженных и покинувших очередь заявок.

Пусть на вход замкнутой многоканальной СМО с числом доступных станций обслуживания m , поступает пуассоновский входной поток заявок трех типов: 0-го, 1-го и 2-го. Поведение заявок в системе зависит от их типа и регламентируется следующими правилами:

– заявки 0-го типа требуют немедленного обслуживания. В случае отсутствия свободных каналов такая заявка покидает систему без постановки в очередь.

– заявки 1-го типа обслуживаются немедленно при наличии свободных каналов. Если все каналы заняты, заявка помещается в очередь с

ограниченным временем ожидания. При превышении допустимого срока ожидания заявка покидает систему необслуженной;

– заявки 2-го типа всегда принимаются в систему, независимо от занятости каналов. При отсутствии свободных каналов они становятся в очередь и ожидают начала обслуживания неограниченное время.

Каждая из обозначенных заявок характеризуется собственной интенсивностью появления в системе: λ_0 ; λ_1 и λ для заявок 0-го, 1-го и 2-го типа соответственно. Следует отметить, что в случае отсутствия свободных обслуживающих устройств, заявки, имеющие интенсивность λ_0 – будут получать отказ в обслуживании, а интенсивность входящего потока будет снижена на величину данной компоненты. Процесс выхода заявки с интенсивностью λ_1 из системы возможен либо после проведения обслуживания, либо по истечении предельного времени ожидания. Заявки с интенсивностью λ будут находиться в системе до момента завершения их обслуживания. Таким образом, в случае отсутствия очереди входящий поток заявок содержит в себе все три компоненты и имеет интенсивность: $\lambda_0 = \lambda_0 + \lambda_1 + \lambda$. Интенсивность входящего потока заявок при отсутствии свободных обслуживающих устройств будет иметь вид: $\lambda_1 = \lambda_1 + \lambda$.

В связи с тем, что СМО является замкнутой, а максимальное количество заявок в источнике ограничено значением N , увеличение числа заявок в системе приводит к сокращению количества заявок в источнике. В результате наблюдается соответствующее уменьшение потока входящих заявок, вплоть до полного отсутствия входящих требований на обслуживание для состояния системы N . К примеру, все обслуживающие устройства заняты, очередь отсутствует, тогда в источнике остается $(N-m)$ поступающих в систему заявок. Если все заявки находятся в системе, то интенсивность входящего потока будет состоять лишь из заявок, получающих отказ в обслуживании. Приход такой заявки в систему не способствует ее переходу из одного состояния в другое. Структурная схема СМО представлена на рис.

1.1

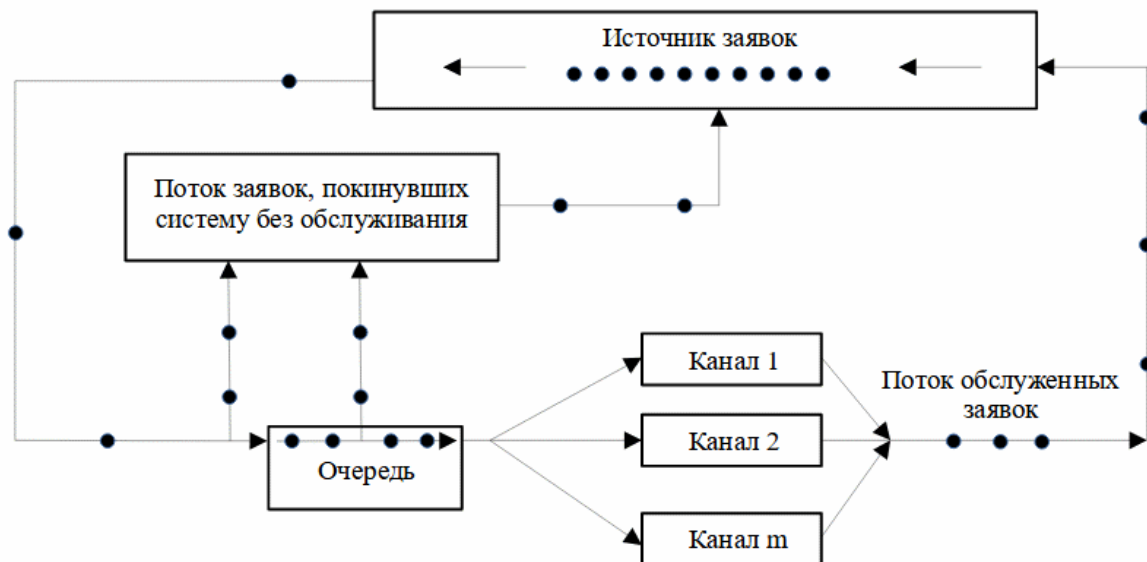


Рис. 1.1 – Структурная схема СМО

Поток обслуживания рассматриваемой СМО, очевидно, зависит лишь от числа занятых обслуживающих устройств и скорости обслуживания « μ » и определяется однозначно при наличии очереди как $m \cdot \mu$. Поток заявок, покидающих систему необслуженными « ν » ввиду ограничения времени ожидания, в свою очередь, будет возрастать при увеличении количества заявок в накопителе от величины ν до величины $(N-m) \cdot \nu$, когда все заявки источника находятся в системе. В данном выражении компонента $(N-m)$ определяет максимальное количество заявок в очереди. Для отображения процессов, происходящих в описанной системе, используется граф состояний СМО.

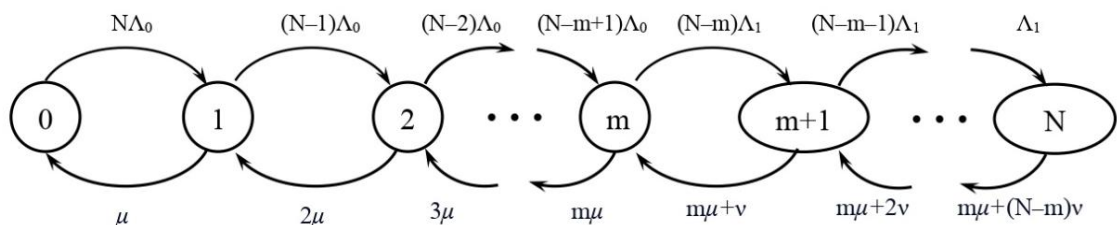


Рис. 1.2 – Граф состояний СМО

Рассмотрим возможные состояния системы более подробно:

«0» – входящих требований нет;

«1» – в системе одна заявка, которая обслуживается одним каналом обслуживания, очереди нет;

«2» – в системе две заявки, которые обслуживаются двумя каналами обслуживания, очереди нет;

«m» – количество заявок в системе равно количеству каналов обслуживания, очереди нет. Вновь прибывшая заявка с интенсивностью λ_0 получит отказ в обслуживании;

«m+1» – количество заявок в системе превышает количество каналов обслуживания на единицу, одна заявка ожидает в очереди;

«N» – все заявки источника находятся в системе, количество заявок в очереди (N–m).

Рассмотрим, как определяются вероятности состояний по графу, считая данный случайный процесс однородным марковским. Пусть в момент времени t система находится в одном из возможных состояний x ($x = 0, 1, \dots, N$). Тогда вероятность того, что на интервале $(t, t + \Delta t)$ произойдет переход $x \rightarrow x + 1$, равна $\lambda_x \Delta t + o(\Delta t)$, где:

$$\lambda_x(x) = \begin{cases} (N-x)\lambda_0, & \text{при } 0 \leq x < m, \\ (N-x)\lambda_1, & \text{при } m \leq x < N. \end{cases} \quad (1.2)$$

Вероятность того, что на интервале $(t, t + \Delta t)$ произойдет переход $x \rightarrow x - 1$, равна $\mu_x \Delta t + o(\Delta t)$, где:

$$\mu_x(x) = \begin{cases} x\mu, & \text{при } 0 < x < m, \\ (m\mu + (x-m)v), & \text{при } m \leq x \leq N. \end{cases} \quad (1.3)$$

Вероятность отсутствия изменения состояния равна $1 - (\lambda_x + \mu_x) \Delta t + o(\Delta t)$. Следовательно, уравнение полной вероятности для состояния x с учетом приведенных постулатов запишется в виде:

$$P_x(t + \Delta t) = \lambda_{x-1} P_{x-1}(t) \Delta t + [1 - (\lambda_x + \mu_x) \Delta t] P_x(t) + \mu_{x+1} P_{x+1}(t) \Delta t + o(\Delta t),$$

откуда разделив обе части на Δt и устремив Δt к нулю, получим систему дифференциальных уравнений в виде:

$$\frac{dP_0(t)}{dt} = -N\lambda_0 P_0(t) + \mu P_1(t), \quad x=0$$

$$\frac{dP_x(t)}{dt} = \lambda_{x-1}P_{x-1}(t) - (\lambda_x + \mu_x)P_x(t) + \mu_{x+1}P_{x+1}(t), x=1, 2, \dots, N-1 \quad (1.4)$$

$$\frac{dP_N(t)}{dt} = -[m\mu + (N-m)v]P_N(t) + A_1P_{N-1}(t), x=N$$

В предельном случае при $t \rightarrow \infty$ в системе устанавливается стационарный режим, в ходе которого система бесконечно блуждает по всем возможным состояниям, но их вероятности не зависят от времени. Поэтому существуют предельные вероятности стационарных состояний системы $p_x = \lim_{t \rightarrow \infty} p_x(t)$, $x = 0, 1, 2, \dots, N$. При этом система дифференциальных уравнений Колмогорова превращается в систему алгебраических уравнений, которая решается относительно вероятностей возможных состояний системы $p_x = p_i$.

Наиболее значимые параметры функционирования исследуемой СМО, рассматриваемые в данной работе:

- вероятность нахождения СМО в свободном состоянии (P_0);
- вероятности состояний СМО (P_i);
- вероятность немедленного обслуживания ($P_{но}$);
- вероятность ожидания заявки в очереди ($P_{ож}$);
- вероятность отказа ($P_{отк}$);
- полное число заявок в системе (\bar{k});
- средняя длина очереди (\bar{l});
- средняя число заявок на обслуживании (\bar{m});
- абсолютная пропускная способность системы (A);
- относительная пропускная способность системы (q);
- среднее время ожидания ($\bar{t}_{ож}$);
- среднее время нахождения заявки в системе ($\bar{t}_{сис}$);
- дисперсии параметров системы (δ_x^2);
- коэффициенты вариации параметров системы (k_x).

Выводы:

1. Анализ современных исследований в области ТМО свидетельствует о том, что замкнутые СМО остаются малоизученными, особенно в сравнении с открытыми системами. Несмотря на их широкое применение в реальных объектах, количество научных работ, направленных на разработку комплексного математического аппарата для моделирования таких систем, ограничено.

2. Гетерогенность входного потока характерна для множества реальных систем и требует учета разнородности заявок при моделировании СМО.

3. Проблемой в теории массового обслуживания для замкнутых СМО с трехкомпонентным потоком заявок является отсутствие строгой математической формализации вероятностных характеристик, а также первых и вторых моментов числовых и временных параметров.

2 МАТЕМАТИЧЕСКАЯ ФОРМАЛИЗАЦИЯ МОДЕЛИ ЗАМКНУТОЙ СИСТЕМЫ МАССОВОГО ОБСЛУЖИВАНИЯ С ОГРАНИЧЕНИЕМ ПО ВРЕМЕНИ ОЖИДАНИЯ ЗАЯВКИ В ОЧЕРЕДИ

2.1 Вероятностные характеристики стационарного режима СМО

В соответствии с выделенными в разделе 1.2 состояниями СМО получен процесс с дискретными состояниями и непрерывным временем. Стационарный режим замкнутой СМО описывается системой уравнений Колмогорова, выведенной для непрерывной марковской цепи с дискретными состояниями $x \in \{0, 1, \dots, N\}$. Для определения вероятностей состояний системы в стационарном режиме функционирования необходимо записать систему уравнений Колмогорова для стационарного режима, и решить ее относительно вероятностей возможных состояний системы P_i .

Обозначим **приведенные интенсивности входящего потока требований** (среднее количество поступающих требований за среднее время обслуживания одной заявки) как:

$$R_i = \lambda_i / \mu$$

В свою очередь **приведенная интенсивность ухода заявок из системы** (среднее количество необслуженных заявок, покидающих систему ввиду ограничения на время ожидания, за среднее время обслуживания одной заявки) будет определяться как:

$$\beta = \nu / \mu$$

Система дифференциальных уравнений Колмогорова с учетом условия непрерывности Марковской цепи имеет вид:

$$\left\{ \begin{aligned}
\frac{dP_0(t)}{dt} &= \mu P_1(t) - N\Lambda_0 \cdot P_0(t) \\
\frac{dP_1(t)}{dt} &= N\Lambda_0 \cdot P_0(t) + 2\mu \cdot P_2(t) - ((N-1)\Lambda_0 + \mu) \cdot P_1(t) \\
\frac{dP_2(t)}{dt} &= (N-1)\Lambda_0 \cdot P_1(t) + 3\mu \cdot P_3(t) - ((N-2)\Lambda_0 + 2\mu) \cdot P_2(t) \\
&\dots \\
\frac{dP_{m-1}(t)}{dt} &= (N-m+2)\Lambda_0 \cdot P_{m-2}(t) + m\mu \cdot P_m(t) - ((N-m+1)\Lambda_0 + (m-1)\mu) \cdot P_{m-1}(t) \\
\frac{dP_m(t)}{dt} &= (N-m+1)\Lambda_0 \cdot P_{m-1}(t) + (m\mu + \nu) \cdot P_{m+1}(t) - ((N-m)\Lambda_1 + m\mu) \cdot P_m(t) \\
\frac{dP_{m+1}(t)}{dt} &= (N-m)\Lambda_1 \cdot P_m(t) + (m\mu + 2\nu) \cdot P_{m+2}(t) - ((N-m-1)\Lambda_1 + m\mu + \nu) \cdot P_{m+1}(t) \\
&\dots \\
\frac{dP_{N-1}(t)}{dt} &= 2\Lambda_1 \cdot P_{N-2}(t) + (m\mu + (N-m)\nu) \cdot P_N(t) - (\Lambda_1 + m\mu + (N-m-1)\nu) \cdot P_{N-1}(t) \\
\frac{dP_N(t)}{dt} &= \Lambda_1 \cdot P_{N-1}(t) - (m\mu + (N-m)\nu) \cdot P_N(t)
\end{aligned} \right.$$

Для стационарного режима при $t \rightarrow \infty$ получим:

$$\left\{ \begin{aligned}
P_0 N\Lambda_0 &= P_1 \mu \\
P_1 (N-1)\Lambda_0 + P_1 \mu &= P_0 N\Lambda_0 + 2P_2 \mu \\
P_2 (N-2)\Lambda_0 + 2P_2 \mu &= P_1 (N-1)\Lambda_0 + 3P_3 \mu \\
&\vdots \\
P_{m-1} (N-m+1)\Lambda_0 + P_{m-1} (m-1)\mu &= P_{m-2} (N-m+2)\Lambda_0 + P_m m\mu \\
P_m (N-m)\Lambda_1 + P_m m\mu &= P_{m-1} (N-m+1)\Lambda_0 + P_{m+1} (m\mu + \nu) \\
P_{m+1} (N-m-1)\Lambda_1 + P_{m+1} (m\mu + \nu) &= P_m (N-m)\Lambda_1 + P_{m+2} (m\mu + 2\nu) \\
&\vdots \\
P_{N-1} \Lambda_1 + P_{N-1} (m\mu + (N-m-1)\nu) &= 2P_{N-2} \Lambda_1 + P_N (m\mu + (N-m)\nu) \\
P_N (m\mu + (N-m)\nu) &= P_{N-1} \Lambda_1
\end{aligned} \right.$$

Выразим каждое из возможных состояний системы через P_0 :

$$P_1 = NR_0 P_0 = \frac{N! R_0}{(N-1)!} P_0$$

$$P_2 = \frac{(N-1)R_0}{2} P_1 = \frac{N(N-1)R_0^2}{2} P_0 = \frac{N! R_0^2}{(N-2)! 2!} P_0$$

$$P_3 = \frac{(N-2)R_0}{3} P_2 = \frac{N(N-1)(N-2)R_0^3}{3!} = \frac{N!R_0^3}{(N-3)!3!} P_0$$

⋮

$$P_m = \frac{(N-m+1)R_0}{m} P_{m-1} = \frac{N(N-1)\dots(N-(m-1))R_0^m}{m!} P_0 = \frac{N!R_0^m}{(N-m)!m!} P_0$$

$$P_{m+1} = \frac{(N-m)\Lambda_1}{\mu(m+\beta)} P_m = \frac{(N-m)R_1}{\beta\left(\frac{m}{\beta}+1\right)} \frac{N!R_0^m}{(N-m)!m!} P_0 = \frac{R_1}{\beta\left(\frac{m}{\beta}+1\right)} \frac{N!R_0^m}{(N-m-1)!m!} P_0$$

$$P_{m+2} = \frac{(N-m-1)\Lambda_1}{\mu(m+2\beta)} P_{m+1} = \frac{(N-m-1)(N-m)R_1^2}{\beta^2\left(\frac{m}{\beta}+1\right)\left(\frac{m}{\beta}+2\right)} \frac{N!R_0^m}{(N-m)!m!} P_0 =$$

$$= \frac{R_1^2}{\beta^2\left(\frac{m}{\beta}+1\right)\left(\frac{m}{\beta}+2\right)} \frac{N!R_0^m}{(N-m-2)!m!} P_0$$

⋮

$$P_N = \frac{\Lambda_1 P_{N-1}}{m\mu + (N-m)\nu} = \frac{R_1}{\beta} \frac{P_{N-1}}{\left(\frac{m}{\beta} + N - m\right)} =$$

$$= \frac{R_1}{\beta} \frac{R_1^{N-m-1}}{\beta^{N-m-1} \left(\frac{m}{\beta}+1\right)\left(\frac{m}{\beta}+2\right)\dots\left(\frac{m}{\beta}+N-m\right)} \frac{N!R_0^m}{m!} P_0 =$$

$$= \left(\frac{R_1}{\beta}\right)^{N-m} \frac{\Gamma\left(\frac{m}{\beta}+1\right)}{\Gamma\left(\frac{m}{\beta}+N-m+1\right)} \frac{N!R_0^m}{m!} P_0 = \left(\frac{R_1}{\beta}\right)^{N-m} \frac{N!R_0^m}{\left(\frac{m}{\beta}+1\right)_{N-m} m!} P_0.$$

Используя условие нормировки:

$$\sum_{i=0}^{\infty} P_i = 1,$$

решим данное уравнение относительно P_0 :

$$P_0 = \left[\sum_{i=0}^m \frac{N!R_0^i}{(N-i)!i!} + \sum_{i=m+1}^N \left(\frac{R_1}{\beta}\right)^{i-m} \left[\left(\frac{m}{\beta}+1\right)_{i-m} \right]^{-1} \frac{N!R_0^m}{(N-i)!m!} \right]^{-1}.$$

Распределение состояний в рассматриваемой СМО будет иметь вид:

$$P_i = \begin{cases} \frac{N!R_0^i}{(N-i)!i!} P_0, & i \leq m \\ \left(\frac{R_1}{\beta}\right)^{i-m} \left[\left(\frac{m}{\beta} + 1\right)_{i-m}\right]^{-1} \frac{N!R_0^m}{(N-i)!m!} P_0, & m < i \leq N \end{cases}. \quad (2.1)$$

Вероятность немедленного обслуживания – вероятность того, что вновь поступившая заявка будет немедленно обслужена.

$$P_{но} = \frac{1}{N-\bar{k}} \sum_{k=0}^{m-1} (N-k) P_k.$$

Вероятность отказа в обслуживании – вероятность отказа для заявок 0-го типа, которая рассчитывается как сумма вероятностей, где все каналы обслуживания заняты.

$$P_{отк} = \frac{\lambda_0}{\Lambda_0} \frac{1}{N-\bar{k}} \sum_{k=m}^{N-1} (N-k) P_k = \frac{\lambda_0}{\Lambda_0} \frac{1}{N-\bar{k}} \sum_{k=m}^N (N-k) P_k.$$

Вероятность ожидания вновь прибывшей заявки – вероятность того, что вновь поступившая заявка найдет обслуживающее устройство занятым и будет ожидать в очереди (для заявок 1-го типа и 2-го типа).

$$P_{ож} = \frac{\Lambda_1}{\Lambda_0} \frac{1}{N-\bar{k}} \sum_{k=m}^{N-1} (N-k) P_k = \frac{\Lambda_1}{\Lambda_0} \frac{1}{N-\bar{k}} \sum_{k=m}^N (N-k) P_k.$$

2.2 Числовые характеристики стационарного режима СМО

Числовые характеристики стационарного режима работы СМО включают в себя показатели, которые помогают оценить эффективность и производительность системы.

Среднее число заявок в системе – полное число требований, находящихся в системе (в очереди и на обслуживании):

$$\bar{k} = \sum_{k=0}^N k P_k = \sum_{k=0}^m k P_k + \sum_{k=m+1}^N k P_k = \sum_{k=0}^m k P_k + m \sum_{k=m+1}^N P_k + \sum_{k=m+1}^N (k-m) P_k = \bar{m} + \bar{l}.$$

Второй момент числа заявок в системе – математическое ожидание квадрата числа заявок в системе.

$$\overline{k^2} = \sum_{k=0}^N k^2 P_k .$$

Дисперсия числа требований в системе – среднее значение квадратов отклонений каждого возможного значения от среднего значения (математического ожидания) числа требований в системе. Она показывает, насколько сильно варьируется число требований, поступающих в систему.

$$\sigma_k^2 = \sum_{k=0}^N k^2 P_k - \left(\sum_{k=0}^N k P_k \right)^2 = \overline{k^2} - \overline{k}^2 .$$

СКО и коэффициент вариации числа требований в системе. Численное значение стандартного отклонения – это квадратный корень из дисперсии. Для текущего параметра оно показывает, насколько в среднем значения числа требований отклоняются от своего математического ожидания. Коэффициент вариации – относительная мера изменчивости, показывающая насколько СКО соотносится с математическим ожиданием.

Среднее число заявок в очереди – число заявок, ожидающих обслуживания.

$$\bar{l} = \sum_{k=m+1}^N (k - m) P_k = \overline{k} - \overline{m} .$$

Второй момент числа заявок в очереди – математическое ожидание квадрата числа заявок в очереди.

$$\overline{l^2} = \sum_{k=m+1}^N (k - m)^2 P_k .$$

Среднее число заявок на обслуживании – число заявок, находящихся на обслуживании.

$$\overline{m} = \sum_{k=0}^m k P_k + m \sum_{k=m+1}^N P_k = m - \sum_{k=0}^{m-1} (m - k) P_k .$$

Второй момент числа заявок на обслуживании – математическое ожидание квадрата числа заявок на обслуживании.

$$\overline{m^2} = \sum_{k=0}^m k^2 P_k + m^2 \sum_{k=m+1}^N P_k .$$

Абсолютная пропускная способность – теоретическая величина, показывающая максимальное количество требований, которое может обработать система за единицу времени при идеальных условиях.

$$\begin{aligned}
 A &= \Lambda_0 \sum_{k=0}^{m-1} (N-k) P_k + \Lambda_1 \sum_{k=m}^{N-1} (N-k) P_k - \nu \bar{l} = \\
 &= \lambda_0 \sum_{k=0}^{m-1} (N-k) P_k + \Lambda_1 \sum_{k=0}^N (N-\bar{k}) P_k - \nu \bar{l} = \\
 &= \lambda_0 \sum_{k=0}^{m-1} (N-\bar{k}) P_k + \Lambda_1 \sum_{k=0}^N (N-\bar{k}) P_k = \\
 &= \lambda_0 P_{ho} (N-\bar{k}) + \Lambda_1 (N-\bar{k}) - \nu \bar{l} = \\
 &= (\lambda_0 P_{ho} + \Lambda_1) (N-\bar{k}) - \nu \bar{l} .
 \end{aligned}$$

Относительная пропускная способность – отношение фактической пропускной способности системы к ее максимально допустимому значению.

$$\begin{aligned}
 q &= 1 - P_{отк} - \frac{\nu \bar{l}}{\Lambda_0 (N-\bar{k})} = P_{ho} + P_{ож} - \frac{\nu \bar{l}}{\Lambda_0 (N-\bar{k})} = \\
 &= \frac{\Lambda_0}{\Lambda_1} \frac{1}{N-\bar{k}} \sum_{k=0}^{m-1} (N-k) P_k + \frac{\Lambda_0}{\Lambda_1} \frac{1}{N-\bar{k}} \sum_{k=m}^N (N-k) P_k - \frac{\nu \bar{l}}{\Lambda_0 (N-\bar{k})} = \\
 &= \frac{\lambda_0}{\Lambda_0} P_{ho} + \frac{\Lambda_1}{\Lambda_0} \frac{1}{N-\bar{k}} (N-\bar{k}) - \frac{\nu \bar{l}}{\Lambda_0 (N-\bar{k})} = \\
 &= \frac{(\lambda_0 P_{ho} + \Lambda_1) (N-\bar{k}) - \nu \bar{l}}{\Lambda_0 (N-\bar{k})} = \frac{A}{\Lambda_0 (N-\bar{k})} .
 \end{aligned}$$

Таким образом:

$$q = \frac{A}{\Lambda_0 (N-\bar{k})} . \quad (2.2)$$

2.3 Временные характеристики СМО

Временные характеристики СМО описывают временные аспекты функционирования системы. Они позволяют оценить эффективность работы СМО и спрогнозировать поведение системы в различных условиях нагрузки.

Среднее время ожидания – это среднее время, которое требование проводит в очереди, ожидая обслуживания. Функция распределения времени ожидания одной заявки с учетом (1.1) определяется формулой:

$$\begin{aligned} \frac{N-\bar{k}}{N} [1-F_{ож}(t)] &= \frac{N-m}{N} P_m G_{10}(t) + \frac{N-m-1}{N} P_{m+1} [G_{20}(t)] + \\ &+ G_{21}(t) + \dots + \frac{1}{N} P_{N-1} [G_{N-m,0}(t) + G_{N-m,N-m-1}(t)] = \\ &= \sum_{k=m}^{N-1} \frac{N-k}{N} P_k \sum_{j=0}^{k-m} G_{k-m+1,j}(t), \end{aligned}$$

тогда

$$\begin{aligned} F_{ож}(t) &= 1 - \frac{1}{(N-\bar{k})} \sum_{k=m}^{N-1} (N-k) P_k \sum_{j=0}^{k-m} G_{k-m+1,j}(t) = \\ &= 1 - \frac{1}{(N-\bar{k})} \sum_{k=m}^{N-1} (N-k) P_k e^{-[m\mu + (k-m+1)\nu]t} \sum_{j=0}^{k-m} \frac{([m\mu + (k-m+1)\nu]t)^j}{j!} = \\ &= 1 - \frac{1}{(N-\bar{k})} \sum_{k=m}^{N-1} (N-k) P_k e^{-[m\mu + (k-m+1)\nu]t} e_{k-m}([m\mu + (k-m+1)\nu]t), \end{aligned}$$

дифференцируя данное выражение, получим плотность распределения времени ожидания:

$$\begin{aligned} f_{ож}(t) &= 1 - \frac{1}{(N-\bar{k})} \sum_{k=m}^{N-1} (N-k) P_k [m\mu + (k-m+1)\nu] e^{-[m\mu + (k-m+1)\nu]t} \cdot \\ &\cdot e_{k-m}([m\mu + (k-m+1)\nu]t) - \frac{1}{(N-\bar{k})} \sum_{k=m}^{N-1} (N-k) P_k [m\mu + (k-m+1)\nu] \cdot \\ &\cdot e^{-[m\mu + (k-m+1)\nu]t} \cdot e_{k-m-1}([m\mu + (k-m+1)\nu]t) = \frac{1}{(N-\bar{k})} \sum_{k=m}^{N-1} (N-k) P_k \cdot \end{aligned}$$

$$\cdot \frac{[m\mu + (k-m+1)\nu]^{k-m+1} t^{k-m}}{(k-m)!} \cdot e^{-[m\mu + (k-m+1)\nu]t}.$$

В итоге:

$$\begin{aligned} \bar{t}_{ож} &= \int_0^\infty t f_{ож}(t) dt = \frac{1}{(N-\bar{k})} \sum_{k=m}^{N-1} (N-k) P_k \frac{[m\mu + (k-m+1)\nu]^{k-m+1}}{(k-m)!} \cdot \\ &\cdot \int_0^\infty t^{k-m+1} e^{-[m\mu + (k-m+1)\nu]t} dt = \frac{1}{(N-\bar{k})} \sum_{k=m}^{N-1} (N-k) P_k \cdot \\ &\cdot \frac{[m\mu + (k-m+1)\nu]^{k-m+1}}{(k-m)!} \cdot \frac{(k-m+1)!}{[m\mu + (k-m+1)\nu]^{k-m+2}} = \\ &= \frac{1}{(N-\bar{k})} \sum_{k=m}^{N-1} (N-k) P_k \frac{k-m+1}{m\mu + (k-m+1)\nu} = \\ &= \frac{P_0}{\mu\beta(N-\bar{k})} \sum_{k=m}^{N-1} \frac{(N-k)(k-m+1)}{\frac{m}{\beta} + k - m + 1} \cdot \frac{\left(\frac{R_1}{\beta}\right)^{k-m}}{\left(\frac{m}{\beta} + 1\right)_{k-m}} \cdot \frac{N! R_0^m}{(N-k)! m!} = \\ &= \frac{P_0}{\Lambda_1(N-\bar{k})} \sum_{k=m}^{N-1} \frac{(N-k)(k-m+1) \left(\frac{R_1}{\beta}\right)^{k-m+1}}{\left(\frac{m}{\beta} + 1\right)_{k-m+1}} \cdot \frac{N! R_0^m}{(N-k)! m!} = \\ &= \frac{P_0}{\Lambda_1(N-\bar{k})} \sum_{k=m}^{N-1} \frac{(k-m+1) \left(\frac{R_1}{\beta}\right)^{k-m+1} N! R_0^m}{\left(\frac{m}{\beta} + 1\right)_{k-m+1} (N-k-1)! m!} = \\ &= \frac{1}{\Lambda_1(N-\bar{k})} \sum_{k=m}^{N-1} (k-m+1) P_{k+1} = \frac{1}{\Lambda_1(N-\bar{k})} \sum_{k=m}^N (k-m) P_k = \\ &= \frac{\bar{t}}{\Lambda_1(N-\bar{k})}. \end{aligned}$$

Таким образом,

$$\bar{t}_{ожс} = \frac{\bar{l}}{\Lambda_1(N-\bar{k})}.$$

Второй момент времени ожидания

Среднее значение квадрата времени ожидания определим как:

$$\begin{aligned} \overline{t_{ожс}^2} &= \int_0^\infty t^2 f_{ожс}(t) dt = \frac{1}{(N-\bar{k})} \sum_{k=m}^{N-1} (N-k) P_k \cdot \\ &\cdot \frac{[m\mu + (k-m+1)\nu]^{k-m+1}}{(k-m)!} \cdot \frac{(k-m+2)!}{[m\mu + (k-m+1)\nu]^{k-m+3}} = \\ &= \frac{1}{(N-\bar{k})} \sum_{k=m}^{N-1} (N-k) P_k \frac{(k-m+2)(k-m+1)}{[m\mu + (k-m+1)\nu]^2}. \end{aligned}$$

Моменты времени ожидания в общем виде:

$$t_{ожс}^{(j)} = \frac{1}{(N-\bar{k})} \sum_{k=m}^{N-1} (N-k) P_k \frac{(k-m+1)_j}{[m\mu + (k-m+1)\nu]^j}, \quad j = \overline{1, 2}.$$

Среднее время обслуживания – это среднее время, которое требуется системе для обслуживания одного входящего требования.

$$\bar{t}_{обсл} = \frac{1}{\mu}.$$

Второй момент среднего времени обслуживания:

$$\sigma_{обсл}^2 = \frac{1}{\mu^2}.$$

Среднее время нахождения заявки в системе – полное время нахождения заявки в системе с учетом ожидания и обслуживания.

$$\bar{t}_{сист} = \bar{t}_{ожид} + \bar{t}_{обсл}.$$

Дисперсия времени нахождения заявки в системе:

$$\sigma_{сист}^2 = \sigma_{обсл}^2 + \sigma_{ожид}^2.$$

Выводы:

1. Получено аналитическое решение системы уравнений Колмогорова в стационарном режиме относительно состояния P_0 . Получены аналитические выражения для расчета основных вероятностных, числовых и временных характеристик работы СМО. Получены общие математические формулы для расчета вторых моментов числовых и временных характеристик.

2. Полученные выражения расширяют математический аппарат ТМО, дополняя существующие модели учетом гетерогенности входного потока и временных ограничений. Параметры, рассчитываемые по формулам главы 2, находят прямое применение для описания и корректировки функционирования реальных систем [74-76].

3 НЕСТАЦИОНАРНЫЙ РЕЖИМ РАБОТЫ СМО С ТРЕХКОМПОНЕНТНЫМ ПОТОКОМ ЗАЯВОК И ОГРАНИЧЕНИЕМ ПО ВРЕМЕНИ ОЖИДАНИЯ ЗАЯВКИ В ОЧЕРЕДИ

3.1 Имитационное моделирование

В рамках настоящего исследования с использованием метода Монте-Карло [77-80] и языка программирования Python разработана имитационная модель многоканальной СМО с тремя типами заявок, таймаутами и ограниченным числом заявок в источнике, позволяющая определять характеристики СМО в нестационарном режиме.

Рассматриваемая СМО характеризуется состояниями $x \in \{0, 1, \dots, N\}$, где x – текущее число заявок в системе. На время начала моделирования $t = 0$ система находится в состоянии простоя $x = 0$. Для каждого состояния системы x в соответствии с (1.2) и (1.3) определена интенсивность переходов:

$$I_{x \rightarrow x+1} = \begin{cases} (N-x)\lambda_0, & 0 \leq x < m, \\ (N-x)\lambda_1, & m \leq x < N. \end{cases}$$
$$I_{x \rightarrow x-1} = \begin{cases} x\mu, & 0 < x \leq m, \\ m\mu + (x-m)\nu, & m < x \leq N. \end{cases}$$

Для формирования временных параметров системы использован метод обратного преобразования, позволяющий генерировать значения, распределенные по экспоненциальному закону. Функция распределения вероятности прибытия заявки за промежуток Δt имеет вид:

$$F(t) = 1 - e^{-\lambda \Delta t},$$

где Δt – время до следующего события;

λ – интенсивность события (прибытия, обслуживания или таймаута);

$F(t)$ – вероятность того, что событие произойдет за время Δt или раньше.

Если случайная величина U , равномерно распределена на интервале $(0,1)$, то случайная величина $X = F^{-1}(U)$ имеет распределение с функцией распределения $F(t)$. Таким образом, найденное методом обратного преобразования значение Δt также будет являться экспоненциально

распределенным. Для экспоненциального закона решим уравнение $U = F(t)$ и найдем значение Δt :

$$\begin{aligned} U &= 1 - e^{-\lambda \Delta t}, \\ \ln(1-U) &= -\lambda \Delta t, \\ \Delta t &= -\frac{\ln(1-U)}{\lambda}, \end{aligned}$$

что эквивалентно:

$$\Delta t = -\frac{\ln(U)}{\lambda}.$$

Определение следующего события системы на временном интервале $(t, t + \Delta t)$ производится путем генерации равномерно распределенного случайного числа $r \in (0, I_{\text{сумм}})$, где $I_{\text{сумм}}$ – суммарная интенсивность переходов для текущего состояния системы, $I_{\text{сумм}} = I_{x \rightarrow x+1} + I_{x \rightarrow x-1}$. Если полученное значение r находится на интервале $(0, I_{x \rightarrow x+1})$, система совершает переход $x \rightarrow x + 1$, в обратном случае $x \rightarrow x - 1$.

В процессе моделирования, при нахождении системы в состоянии x в момент времени t , и при генерации времени до следующего события Δt , рассматриваются все моменты измерения τ_j , удовлетворяющие условию $t \leq \tau_j < t + \Delta t$. Для каждого τ_j фиксируется состояние системы x . После обработки всех таких моментов измерения система переходит в следующее состояние в момент времени $t + \Delta t$. Моменты измерения τ и их количество выбираются адаптивно на основе предварительного анализа времени стабилизации системы. Блок-схема алгоритма имитационного моделирования СМО приведена на рис. 3.1.

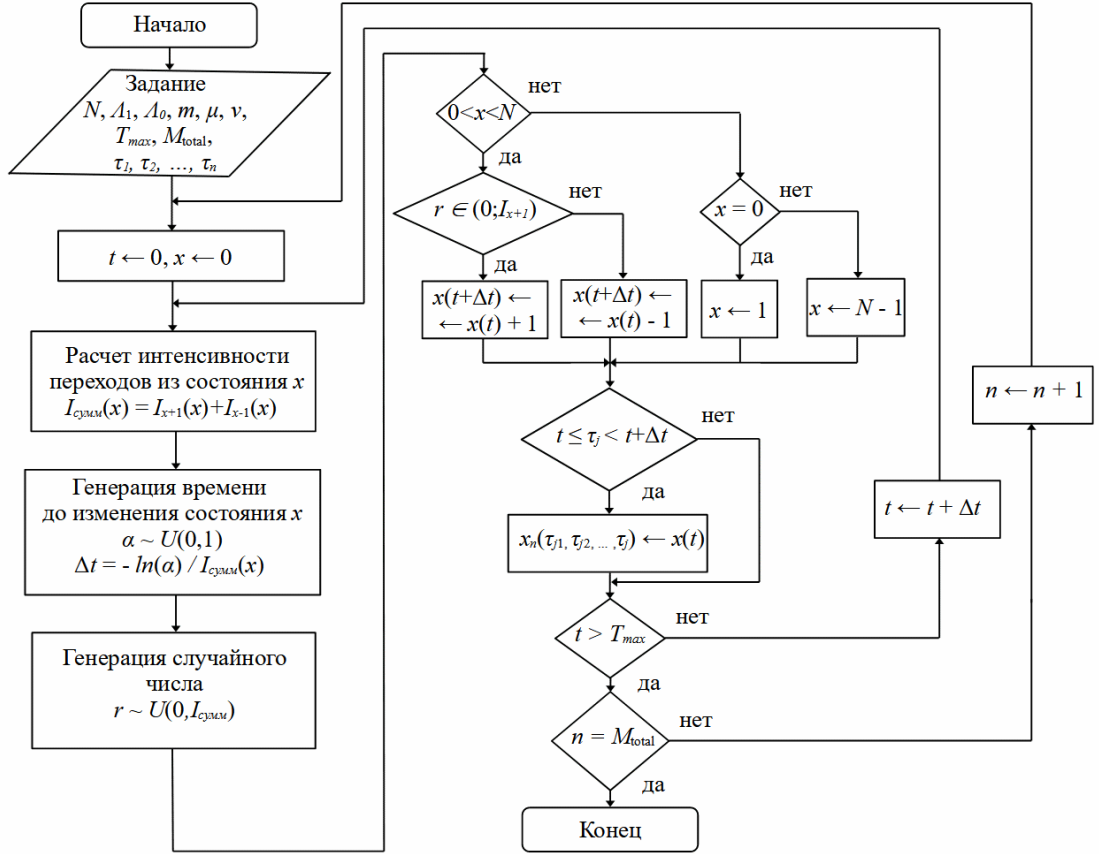


Рис. 3.1 – Блок-схема алгоритма имитационного моделирования

После числа симуляций M_{total} производится статистическая оценка математического ожидания вероятностей системы в состояниях $x = (0, 1, \dots, N)$ по формуле:

$$\hat{P}_x(\tau) = \frac{n_x(\tau)}{M_{total}},$$

где $n_x(\tau)$ – число прогонов, в которых система находилась в состоянии x в момент времени τ ;

M_{total} – общее число прогонов.

Квазистационарное значение для каждого состояния $P_{st}(x)$ определяется как среднее арифметическое оценок вероятностей $\hat{P}_x(\tau)$ по всем моментам измерения τ , принадлежащим последним 20 % времени моделирования:

$$P_{st}(x) = \frac{1}{n_{last}} \sum_{i=1}^{n_{last}} \hat{P}_x(\tau_i^{last}),$$

где n_{last} – число моментов времени τ_i^{last} на интервале от $0,8 \cdot T_{max}$ до T_{max} ;

T_{max} – максимальное время моделирования;

τ_i^{last} – моменты измерения, принадлежащие интервалу $[0,8 \cdot T_{max}, T_{max}]$.

Для каждого заданного момента τ вычисляется относительное отклонение для состояний со значимой вероятностью:

$$\delta(\tau) = \max_{x \in S} \frac{|\hat{P}_x(\tau) - P_{st}(x)|}{P_{st}(x)},$$

где $S = \{x \in \{0, 1, \dots, N\} | P_{st}(x) > 0,01\}$.

Система считается вышедшей на квазистационарный режим, в момент τ_s , если значение $\delta(\tau_s) < 0,05$ и данное условие выполняется не менее чем для трех последующих временных меток τ .

Применение разработанного алгоритма к системе с параметрами $N=30$, $m=5$, $\lambda_0=0,0416$ (1 заявка в 24 часа), $\lambda_1=0,048$ (1 заявка в 25 часов), $\lambda=0,0208$ (1 заявка в 48 часов), $\mu=0,28$ (1 заявка в 3,5 часа), $\nu=0,166$ (1 заявка в 6 часов) показало, что время выхода системы на квазистационарный режим составляет $\tau_s = 11,1$ часов.

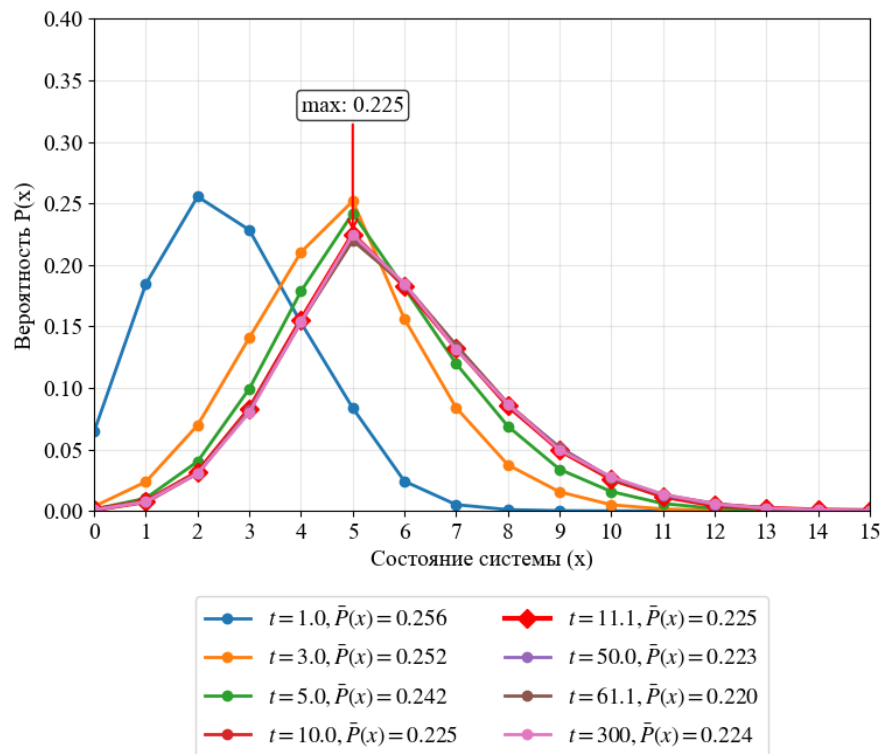


Рис. 3.2 – Результаты имитационного моделирования методом Монте-Карло для числа симуляций $M_{total} = 2 \cdot 10^6$

Сопоставление результатов с аналитической моделью показало высокую степень соответствия, что подтвердило корректность реализации. Относительное отклонение между максимальными значениями вероятностей состояний, определенными методом Монте-Карло и аналитическим методом, не превышает $\pm 1,5\%$.

Проведен анализ влияния основных параметров системы на среднюю длительность переходного процесса τ_s . Результаты вычислительных экспериментов сведены в таблицу 3.1.

Таблица 3.1 – Значения времени достижения квазистационарного режима при различных комбинациях управляющих параметров

Параметр системы	Базовое значение	Обновленное значение	Время выхода на квазистационарный режим τ_s , ч.	Отклонение параметра τ_s , %
L_0	0,110	0,220	17,1	+ 54,05
		0,055	9,6	– 13,51
μ	0,400	0,800	5,1	– 54,05
		0,200	20,1	+ 81,08
ν	0,167	0,334	9,6	– 13,51
		0,084	15,1	+ 36,04

Анализ результатов моделирования выявил, что ключевым фактором, определяющим продолжительность переходного процесса в системе, является интенсивность обслуживания заявок μ , что обусловлено ролью данного параметра в динамике замкнутой СМО. В отличие от открытых систем, где заявки окончательно покидают систему, в рассматриваемой модели происходит постоянная рециркуляция заявок. Так, низкое значение скорости обслуживания приводит не только к росту очереди, но и к сокращению числа заявок в источнике, что значительно увеличивает время нахождения заявки в системе и, как следствие, время выхода на квазистационарный режим.

3.2 Метод производящих функций

Для построения аналитического решения системы уравнений Колмогорова (1.4) использован метод производящих функций [81]. Для свертки системы использована производящая функция вида:

$$F(z, t) = \sum_{x=0}^{\infty} z^x P_x(t),$$

и следующие начальные условия:

$$P_x(0) = \begin{cases} 1, & \text{при } x = 0, \\ 0, & \text{при } x > 0, \end{cases}$$

где z – формальная переменная, $|z| < 1$.

Таким образом:

$$F(z, 0) = 1.$$

Умножим каждое уравнение системы (1.4) на соответствующую степень z^x . Для $x = 0$:

$$\frac{dP_0(t)}{dt} \cdot z^0 = \mu P_1(t) \cdot z^0 - N A_0 P_0(t) \cdot z^0,$$

для $0 < x < m$:

$$\begin{aligned} \frac{dP_x(t)}{dt} \cdot z^x &= (N - x + 1) A_0 \cdot P_{x-1}(t) \cdot z^x + (x + 1) \mu P_{x+1}(t) \cdot z^x - \\ &- ((N - x) A_0 + x \mu) P_x(t) \cdot z^x, \end{aligned}$$

для $x = m$:

$$\begin{aligned} \frac{dP_m(t)}{dt} \cdot z^m &= (N - m + 1) A_0 \cdot P_{m-1}(t) \cdot z^m + (m \mu + \nu) P_{m+1}(t) \cdot z^m - \\ &- ((N - m) A_1 + m \mu) P_m(t) \cdot z^m, \end{aligned}$$

для $m < x < N$:

$$\begin{aligned} \frac{dP_x(t)}{dt} \cdot z^x &= (N - x + 1) A_1 \cdot P_{x-1}(t) \cdot z^x + (m \mu + (x - m + 1) \nu) P_{x+1}(t) \cdot z^x - \\ &- ((N - x) A_1 + m \mu + (x - m) \nu) P_x(t) \cdot z^x, \end{aligned}$$

для $x = N$:

$$\frac{dP_N(t)}{dt} \cdot z^N = A_1 \cdot P_{N-1}(t) \cdot z^N - (m\mu + (N-m)\nu)P_N(t) \cdot z^N.$$

Перед преобразованием системы уравнений (1.4) получен ряд вспомогательных тождеств, которые будут использованы в дальнейшем анализе:

$$\begin{aligned} \sum_{x=1}^{N+1} z^x P_{x-1}(t) &= zP_0(t) + z^2P_1(t) + \dots + z^NP_{N-1}(t) + z^{N+1}P_N(t) = \\ &= z \sum_{x=0}^N z^x P_x(t) = zF(z, t), \end{aligned} \quad (3.1)$$

$$\begin{aligned} \sum_{x=1}^N xz^x P_x(t) &= zP_1(t) + 2z^2P_2(t) + \dots + Nz^NP_N(t) = \\ &= z[P_1(t) + 2zP_2(t) + \dots + Nz^{N-1}P_N(t)] = z \frac{\partial F(z, t)}{\partial z}, \end{aligned} \quad (3.2)$$

$$\begin{aligned} \sum_{x=0}^{N-1} (x+1)z^x P_{x+1}(t) &= P_1(t) + 2zP_2(t) + 3z^2P_3(t) + \dots + \\ &+ Nz^{N-1}P_N(t) = \frac{\partial F(z, t)}{\partial z}. \end{aligned} \quad (3.3)$$

Сумма левых частей уравнений представляет собой производную производящей функции по времени:

$$\sum_{x=0}^N \frac{dP_x(t)}{dt} = \frac{\partial F(z, t)}{\partial t}.$$

Представив $A_0 = A_1 + \lambda_0$, отдельно просуммируем правые части уравнений в зависимости от вида переходов. Для переходов $x-1 \rightarrow x$:

$$\begin{aligned} &zN(A_1 + \lambda_0)P_0(t) + z^2(N-1)(A_1 + \lambda_0)P_1(t) + \\ &+ z^{m-1}(N-m+2)(A_1 + \lambda_0)P_{m-2}(t) + z^m(N-m+1)(A_1 + \lambda_0)P_{m-1}(t) + \\ &+ z^{m+1}(N-m)A_1P_m(t) + \dots + \\ &+ z^{N-1}2A_1P_{N-2}(t) + z^NA_1P_{N-1}(t) = \\ &= zNA_1 \sum_{x=0}^{N-1} z^x P_x(t) + zN\lambda_0 \sum_{x=0}^{m-1} z^x P_x(t) - zA_1 \sum_{x=0}^{N-1} xz^x P_x(t) - z\lambda_0 \sum_{x=0}^{m-1} xz^x P_x(t). \end{aligned}$$

Для $x+ \rightarrow x$:

$$\begin{aligned}
& \mu P_1(t) + 2z\mu P_2(t) + 3z^2\mu P_3(t) + \dots + mz^{m-1}\mu P_m(t) + z^m(m\mu + \nu)P_{m+1}(t) + \\
& + z^{m+1}(m\mu + 2\nu)P_{m+2}(t) + \dots + z^{N-1}(m\mu + (N-m)\nu)P_N(t) = \\
& = \mu \sum_{x=0}^N xz^{x-1}P_x(t) + \sum_{x=m+1}^N z^{x-1}(\nu - \mu)(x-m)P_x(t).
\end{aligned}$$

Для $x \rightarrow x+1$:

$$\begin{aligned}
& NA_0P_0(t) + z(N-1)A_0P_1(t) + z^2(N-2)A_0P_2(t) + \dots + \\
& + z^{m-1}(N-m+1)A_0P_{m-1}(t) + z^m(N-m)A_1P_m(t) + \\
& + z^{m+1}(N-m-1)A_1P_{m+1}(t) + \dots + \\
& + z^{N-1}A_1P_{N-1}(t) = NA_1 \sum_{x=0}^{N-1} z^x P_x(t) + N\lambda_0 \sum_{x=0}^{m-1} z^x P_x(t) - \\
& - A_1 \sum_{x=0}^{N-1} xz^x P_x(t) - \lambda_0 \sum_{x=0}^{m-1} xz^x P_x(t).
\end{aligned}$$

Для $x+1 \rightarrow x$:

$$\begin{aligned}
& zmP_1(t) + 2z^2\mu P_2(t) + \dots + (m-1)z^{m-1}\mu P_{m-1}(t) + m\mu z^m P_m(t) + \\
& + (m\mu + \nu)z^{m+1}P_{m+1}(t) + \dots + (m\mu + (N-m-1)\nu)P_{N-1}(t)z^{N-1} + \\
& + (m\mu + (N-m)\nu)z^N P_N(t) = z\mu \sum_{x=0}^N xz^{x-1}P_x(t) + z \sum_{x=m+1}^N z^{x-1}(\nu - \mu)(x-m)P_x(t).
\end{aligned}$$

Просуммируем полученные выражения:

$$\begin{aligned}
\frac{\partial F(z,t)}{\partial t} = & (z-1)NA_1 \sum_{x=0}^{N-1} z^x P_x(t) + (z-1)N\lambda_0 \sum_{x=0}^{m-1} z^x P_x(t) - \\
& - (z-1)A_1 \sum_{x=0}^{N-1} xz^x P_x(t) - (z-1)\lambda_0 \sum_{x=0}^{m-1} xz^x P_x(t) - \\
& - (z-1)\mu \sum_{x=0}^N xz^{x-1}P_x(t) - (z-1)(\nu - \mu) \sum_{x=m+1}^N (x-m)z^{x-1}P_x(t).
\end{aligned}$$

Система уравнений (1.4) имеет кусочно-линейную структуру интенсивностей из-за наличия порогового значения m , при превышении которого изменяется структура входного потока, ввиду чего нахождение

аналитического решения для общего случая затруднено. Рассмотрим частный случай, приняв следующие допущения:

- 1) вероятность нахождения в системе N заявок стремится к нулю $P_N(t) \approx 0$;
- 2) 0-й тип заявок исключен из входного потока $\lambda_0 = 0$, в результате чего интенсивность входного потока одинакова при $x < m$ и $x \geq m$ и равна λ_1 ;
- 3) Интенсивности обслуживания и ухода нетерпеливых заявок равны между собой $\nu = \mu$.

$$\sum_{x=0}^{N-1} z^x P_x(t) \approx F(z, t).$$

В соответствии с (3.1) – (3.3):

$$(z-1)N\lambda_1 \sum_{x=0}^{N-1} z^x P_x(t) = (z-1)N\lambda_1 F(z, t),$$

$$(z-1)\lambda_1 \sum_{x=0}^{N-1} xz^x P_x(t) = (z-1)\lambda_1 \left[z \frac{\partial F}{\partial z} - Nz^N P_N(t) \right] \approx z(z-1)\lambda_1 \frac{\partial F(z, t)}{\partial z},$$

$$(z-1)\mu \sum_{x=0}^N xz^{x-1} P_x(t) = (z-1)\mu \sum_{x=1}^N xz^x P_x(t) \cdot \frac{1}{z} = (z-1)\mu \frac{\partial F(z, t)}{\partial z}.$$

Таким образом:

$$\frac{\partial F(z, t)}{\partial t} = (z-1)N\lambda_1 F(z, t) - (z-1)\lambda_1 \left(z \frac{\partial F(z, t)}{\partial z} \right) - (z-1)\mu \left(\frac{\partial F(z, t)}{\partial z} \right),$$

или

$$\begin{aligned} \frac{\partial F(z, t)}{\partial t} &= (z-1)N\lambda_1 F(z, t) - (z-1)(\lambda_1 z + \mu) \frac{\partial F(z, t)}{\partial z}, \\ \frac{\partial F(z, t)}{\partial t} + (z-1)(\lambda_1 z + \mu) \frac{\partial F(z, t)}{\partial z} &= (z-1)N\lambda_1 F(z, t). \end{aligned}$$

Решим полученное уравнение методом характеристик:

$$\frac{dt}{1} = \frac{dz}{(z-1)(\lambda_1 z + \mu)} = \frac{dF(z, t)}{(z-1)N\lambda_1 F(z, t)}.$$

Первая характеристика:

$$\frac{dt}{1} = \frac{dz}{(z-1)(A_1 z + \mu)}.$$

Разложим на простые дроби:

$$\frac{1}{(z-1)(A_1 z + \mu)} = \frac{A}{z-1} + \frac{B}{A_1 z + \mu}$$

Определим коэффициенты:

$$A = \frac{1}{A_1 + \mu}, \quad B = -\frac{A_1}{A_1 + \mu}$$

Таким образом:

$$\frac{dz}{(z-1)(A_1 z + \mu)} = \frac{1}{A_1 + \mu} \left(\frac{dz}{z-1} - \frac{A_1 dz}{A_1 z + \mu} \right),$$

или

$$dt = \frac{1}{A_1 + \mu} \left(\frac{dz}{z-1} - \frac{A_1 dz}{A_1 z + \mu} \right).$$

Интегрируем:

$$(A_1 + \mu)t = \ln|z-1| - \ln|A_1 z + \mu| + C_1$$

$$(A_1 + \mu)t = \ln \left| \frac{z-1}{A_1 z + \mu} \right| + C_1$$

Обозначим z_0 как значение z при $t=0$, тогда:

$$C_1 = -\ln \left| \frac{z_0-1}{A_1 z_0 + \mu} \right|$$

$$(A_1 + \mu)t = \ln \left| \frac{z-1}{A_1 z + \mu} \right| - \ln \left| \frac{z_0-1}{A_1 z_0 + \mu} \right|$$

$$(A_1 + \mu)t = \ln \left| \frac{(z-1)(A_1 z_0 + \mu)}{(A_1 z + \mu)(z_0-1)} \right|$$

Решим относительно z_0 :

$$e^{(A_1 + \mu)t} = \frac{(z-1)(A_1 z_0 + \mu)}{(A_1 z + \mu)(z_0-1)}$$

$$e^{(A_1 + \mu)t}(A_1 z + \mu)(z_0-1) = (z-1)(A_1 z_0 + \mu)$$

$$e^{(A_1 + \mu)t}(A_1 z + \mu)z_0 - e^{(A_1 + \mu)t}(A_1 z + \mu) = (z-1)A_1 z_0 + (z-1)\mu$$

$$z_0 [e^{(A_1 + \mu)t}(A_1 z + \mu) - (z-1)A_1] = e^{(A_1 + \mu)t}(A_1 z + \mu) + (z-1)\mu$$

$$z_0 = \frac{e^{(A_1+\mu)t}(A_1z+\mu)+(z-1)\mu}{e^{(A_1+\mu)t}(A_1z+\mu)-(z-1)A_1}$$

$$z_0 = \frac{(A_1z+\mu)+(z-1)\mu e^{-(A_1+\mu)t}}{(A_1z+\mu)-(z-1)A_1 e^{-(A_1+\mu)t}}$$

Вторая характеристика:

$$\frac{dt}{1} = \frac{dF}{(z-1)NA_1F}$$

Тогда:

$$\frac{dF}{F} = NA_1(z-1)dt$$

Заметим из первой характеристики, что:

$$(A_1+\mu)t = \ln \left| \frac{z-1}{A_1z+\mu} \right| - \ln \left| \frac{z_0-1}{A_1z_0+\mu} \right|$$

ЭКВИВАЛЕНТНО

$$\frac{z-1}{A_1z+\mu} = \frac{z_0-1}{A_1z_0+\mu} e^{(A_1+\mu)t}$$

таким образом z меняется со временем вдоль характеристики и зависит от t , в связи с чем интегрирование затруднено. Из первого уравнения характеристики:

$$\frac{dz}{dt} = (z-1)(A_1z+\mu)$$

$$(z-1) = \frac{1}{(A_1z+\mu)} \frac{dz}{dt}$$

Из уравнения второй характеристики:

$$\frac{dF(z,t)}{dt} = (z-1)NA_1F(z,t)$$

Подставим полученное выражение

$$\frac{dF(z,t)}{dt} = \left(\frac{1}{(A_1z+\mu)} \frac{dz}{dt} \right) NA_1F(z,t)$$

или

$$\frac{dF(z,t)}{F(z,t)} = \frac{NA_1}{(A_1z+\mu)} \frac{dz}{dt} dt$$

$$\frac{dF(z,t)}{F(z,t)} = \frac{N\lambda_1}{(\lambda_1 z + \mu)} dz$$

Полученное уравнение зависит только от z , без явной зависимости от t .

Интегрируем:

$$\ln F(z,t) = N\lambda_1 \int \frac{dz}{(\lambda_1 z + \mu)} + C_2$$

получим

$$F(z,t) = e^{C_2} (\lambda_1 z + \mu)^N$$

Используя начальные условия $t = 0$, $z = z_0$ и $F(z,0)=1$:

$$e^{C_2} = \frac{1}{(\lambda_1 z_0 + \mu)^N}$$

Решение имеет вид:

$$F(z,t) = \left[\frac{\lambda_1 z + \mu}{\lambda_1 z_0 + \mu} \right]^N$$

где $z_0(z,t)$ определяется из первой характеристики:

После преобразований получим:

$$F(z,t) = \left[\frac{\mu + \lambda_1 z + \lambda_1 (1-z)e^{-(\lambda_1 + \mu)t}}{\lambda_1 + \mu} \right]^N$$

или в другом виде

$$F(z,t) = \left[\frac{\mu + \lambda_1 e^{-(\lambda_1 + \mu)t}}{\lambda_1 + \mu} + \frac{\lambda_1 (1 - e^{-(\lambda_1 + \mu)t})}{\lambda_1 + \mu} z \right]^N$$

Заметим, что это биномиальная производящая функция:

$$F(z,t) = [q(t) + p(t)z]^N$$

где:

$$p(t) = \frac{\lambda_1 (1 - e^{-(\lambda_1 + \mu)t})}{\lambda_1 + \mu}, \quad q(t) = \frac{\mu + \lambda_1 e^{-(\lambda_1 + \mu)t}}{\lambda_1 + \mu}$$

Причем $p(t) + q(t) = 1$ для всех t .

Разлагая по биному Ньютона:

$$F(z,t) = \sum_{x=0}^N \frac{N!}{x!(N-x)!} q(t)^{N-x} p(t)^x z^x$$

Следовательно коэффициенты при z^x дают вероятности состояний:

$$P_x(t) = \frac{N!}{x!(N-x)!} q(t)^{N-x} p(t)^x$$

Или:

$$P_x(t) = \frac{N!}{x!(N-x)!} \left(\frac{\mu + \lambda_1 e^{-(\lambda_1 + \mu)t}}{\lambda_1 + \mu} \right)^{N-x} \left(\frac{\lambda_1 (1 - e^{-(\lambda_1 + \mu)t})}{\lambda_1 + \mu} \right)^x \quad (3.4)$$

Для верификации корректности аналитического решения, полученного для частного случая $P_N(t) \approx 0$, $\lambda_0 = 0$, $\nu = \mu$, проведено численное моделирование методом Монте-Карло. На рис. 3.3 представлены графики временной динамики вероятностей $P_x(t)$, полученные аналитическим методом для частного случая (3.4) и методом Монте-Карло.

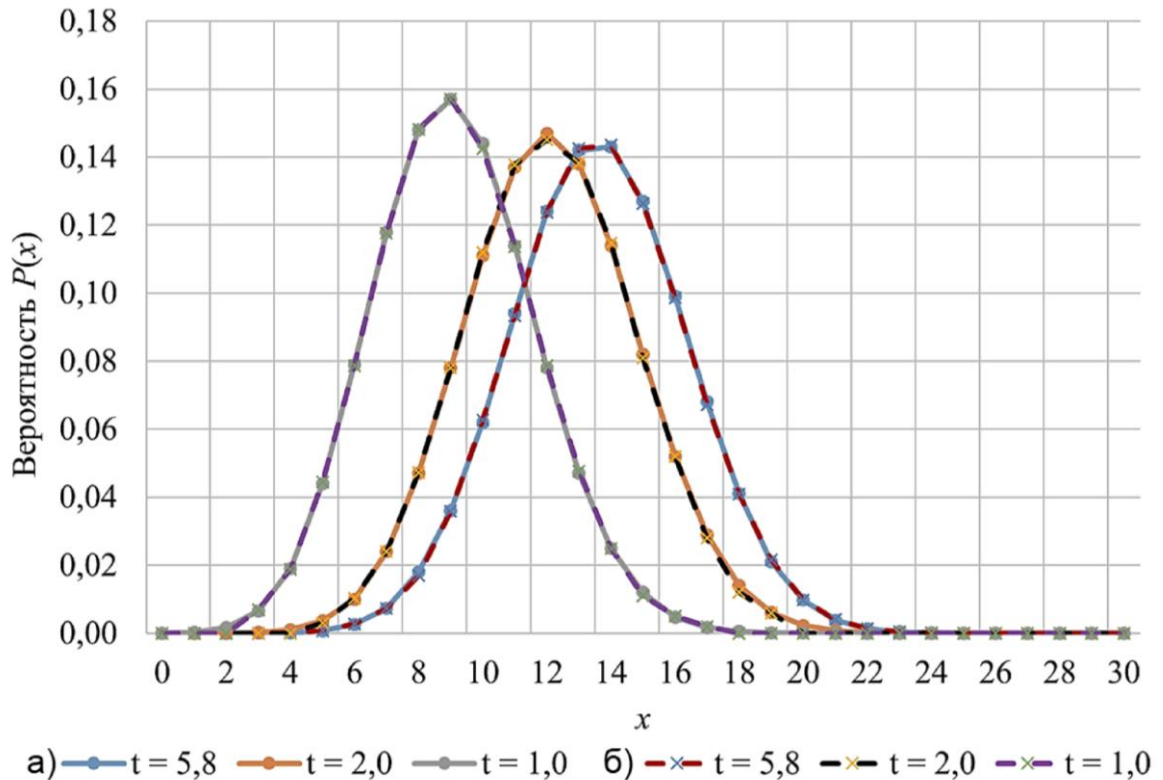


Рис. 3.3 – Сравнение временных зависимостей $P_x(t)$, полученных:

а) в аналитическом виде для частного случая и

б) методом Монте-Карло

Для оценки полученных результатов аналитического решения вычислены следующие метрики:

– максимальное значение вероятности $P_{max}(t)$,

- среднее число заявок в системе $\bar{x}(t)$.
- время выхода на квазистационарный режим t_{cm} , ч;
- относительное отклонение между максимальными значениями вероятности, %:

$$\delta_P(t) = \frac{|P_{max}(t)_a - P_{max}(t)_m|}{P_{max}(t)_m} \cdot 100,$$

где $P_{max}(t)_a$ – максимальное значение вероятности в момент времени t , полученное в аналитическом решении для частного случая;

$P_{max}(t)_m$ – максимальное значение вероятности в момент времени t , полученное методом Монте-Карло.

- относительное отклонение между средним числом заявок в системе, %:

$$\delta_{\bar{x}}(t) = \frac{|\bar{x}(t)_a - \bar{x}(t)_m|}{\bar{x}(t)_m} \cdot 100,$$

где $\bar{x}(t)_a$ – среднее число заявок в системе в момент времени t , полученное в аналитическом решении для частного случая;

$\bar{x}(t)_m$ – максимальное значение вероятности в момент времени t , полученное методом Монте-Карло.

Таблица 3.2 – Полученные характеристики системы по времени

Характеристика	Аналитическое решение (частный случай)	Метод Монте-Карло
$P_{max}(t), t = 1,0$ ч	0,157	0,157
$P_{max}(t), t = 2,0$ ч	0,147	0,145
$P_{max}(t), t = 5,8$ ч	0,143	0,144
$\bar{x}(t), t = 1,0$ ч	9,097	9,083
$\bar{x}(t), t = 2,0$ ч	12,125	11,930
$\bar{x}(t), t = 5,8$ ч	13,613	13,615
t_{cm} , ч	5,8	5,1

Таблица 3.3 – Результаты количественной оценки точности по времени

Характеристика	Время		
	$t = 1,0$ ч	$t = 2,0$ ч	$t = 5,8$ ч
$\delta_P(t)$, %	0,1	1,3	0,4
$\delta_{\bar{x}}(t)$, %	0,2	1,6	0,1

Полученные результаты демонстрируют высокую степень соответствия между полученным для частного случая аналитическим решением и результатами численного моделирования. Наблюдаемые предельные отклонения максимального значения вероятности в $\pm 1,3$ % и среднего числа заявок в системе в $\pm 1,6$ % объясняются статистической погрешностью метода Монте-Карло, конечным числом реализаций вычислительного эксперимента и принятым допущением о нулевой вероятности состояния системы N .

Выводы:

1. В данной главе проведено имитационное моделирование с использованием метода Монте-Карло и языка программирования Python. Получены результаты расчета распределения вероятностей состояний системы в нестационарном режиме и времени выхода системы на квазистационарный режим. Анализ полученных результатов показал высокую степень соответствия аналитической модели, что подтвердило корректность реализации.

2. В аналитическом виде получено решение системы уравнений Колмогорова методом производящих функций для частного случая $P_N(t) \approx 0$, $\lambda_0 = 0$, $\nu = \mu$. Верификация методом Монте-Карло подтвердила корректность полученного решения. Полученные результаты подтверждают применимость аналитической модели для описания функционирования систем с аналогичной структурой в нестационарном режиме.

4 ЧИСЛЕННОЕ МОДЕЛИРОВАНИЕ РЕЖИМОВ РАБОТЫ СМО С ТРЕХКОМПОНЕНТНЫМ ПОТОКОМ ЗАЯВОК И ОГРАНИЧЕНИЕМ ПО ВРЕМЕНИ ОЖИДАНИЯ ЗАЯВКИ В ОЧЕРЕДИ

4.1 Постановка задачи

В данной главе рассматриваются числовые и временные характеристики замкнутых СМО с трехкомпонентным потоком заявок и ограничением по времени ожидания заявки в очереди в стационарном режиме. Исследуется влияние изменения входных параметров системы на динамику работы СМО.

В процессе синтеза исходных данных для вычислительного эксперимента в рамках данной работы использованы следующие принципы:

- число обслуживающих устройств m не должно превышать число заявок в источнике N ;
- число заявок в источнике должно быть ограниченным, поскольку при значительном их объеме в системе не наблюдается падения интенсивности входящего потока из-за уменьшения числа заявок (подобную систему целесообразно рассматривать как открытую);
- суммарная интенсивность входящего потока, интенсивность ухода заявок из системы и интенсивность обслуживания должны соотноситься по численным значениям.

Параметры системы, используемые для численного моделирования, приведены в таблице 4.1. Приведенная таблица в полной мере задает исходные данные для моделирования СМО с тремя типами заявок. Изменение данных параметров позволяет анализировать динамику изменения ключевых параметров эффективности работы системы и принимать решения о необходимости внесения корректировок для организации более эффективного обслуживания.

Таблица 4.1 – Параметры системы, используемые для численного моделирования

№ п/п	Наименование величины	Обозначение	Значение
1	Количество заявок в источнике	N	30
2	Количество обслуживающих устройств	m	5
3	Интенсивность обслуживания	μ	0,6
4	Интенсивность ухода заявок ввиду ограничения на время ожидания	ν	0,6
5	Интенсивность входящего потока заявок 0-го типа	λ_0	0,25
6	Интенсивность входящего потока заявок 1-го типа	λ_1	0,25
7	Интенсивность входящего потока заявок 2-го типа	λ	0,25

4.2 Численная оценка и анализ параметров СМО

Результаты расчетов начальных параметров системы приведены в таблице 4.2.

Таблица 4.2 – Результаты расчетов начальных параметров системы.

№ п/п	Наименование величины	Обозначение	Значение
1	Общий поток заявок в систему	A_0	0,75
2	Поток заявок в систему без учета заявок 0-го типа	A_1	0,5
3	Отношение интенсивности общего входящего потока к интенсивности обслуживания	R_0	1,25

№ п/п	Наименование величины	Обозначение	Значение
4	Отношение интенсивности входящего потока без учета заявок 0-го типа к интенсивности обслуживания	R_1	0,833
5	Отношение интенсивности ухода заявок из системы ввиду ограничения на время ожидания к интенсивности обслуживания	β	1
6	Вероятность простоя системы	P_0	$1,669 \cdot 10^{-9}$

Более детально значение каждого из параметров для системы рассмотрено в главе 2. При анализе полученных методом математического моделирования численных значений следует учесть, что в данном случае параметры R_0 и R_1 меньше единицы, что свидетельствует о том, что интенсивность обслуживания заявок одним обслуживающим устройством больше интенсивности поступающих в систему заявок. Значение β , равное единице, в свою очередь, говорит о равенстве потока заявок, уходящих по времени ожидания и потока обслуживания.

Вероятности состояний СМО, рассчитанные по формуле (2.1), приведены на рис. 4.1

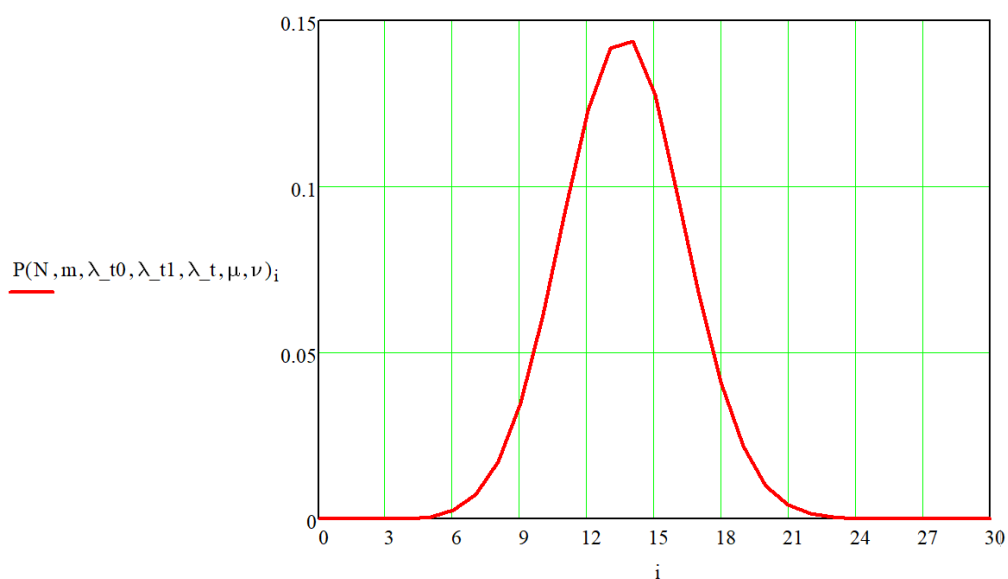


Рис. 4.1 – Вероятности состояний СМО

График демонстрирует все возможные вероятности состояний функционирования СМО в стационарном режиме. Ось абсцисс – количество клиентов в системе на данный момент, ось ординат – вероятность этого состояния. Состояние «0» соответствует простоя системы, в то время как состояние «1», «2» и далее отражают увеличивающееся количество клиентов в системе вплоть до состояния «30», когда все заявки находятся в системе.

Ключевые вероятностные характеристики СМО приведены в таблице 4.3.

Таблица 4.3 – Ключевые вероятностные характеристики СМО

№ п/п	Наименование величины	Обозначение	Значение
1	Вероятность немедленного обслуживания	$P_{обсл}$	$2,013 \cdot 10^{-4}$
2	Вероятность отказа	$P_{отк}$	0,333
3	Вероятность ожидания	$P_{ож}$	0,667

Очевидно, сумма значений вероятностей, указанных в таблице 4.3, составляет 1, что говорит об адекватности полученной модели в части вероятностных характеристик. Рассчитанные вероятностные характеристики позволяют проанализировать эффективность работы реальной СМО и выявить потенциально возможные проблемы. Высокие вероятности отказа и ожидания могут служить сигналом для поиска более эффективных режимов работы системы. В данном примере 33,3 % заявок получают отказ в обслуживании. Следовательно, заявки 0-го типа с вероятностью 33,3 % будут получать отказ в обслуживании ввиду наличия очереди.

4.3 Показатели загрузки системы

Таблица 4.4 – Ключевые числовые характеристики, касающиеся среднего числа требований в системе

№ п/п	Наименование величины	Обозначение	Значение
1	Среднее число требований в системе	\bar{k}	13,637
2	Второй момент числа требований в системе	\bar{k}^2	193,402
3	Дисперсия числа требований в системе	δ_k^2	7,431
4	СКО числа требований в системе	δ_k	2,726
5	Коэффициент вариации числа требований в системе	$K_{\bar{k}}$	0,2

Таблица 4.5 – Ключевые числовые характеристики, касающиеся среднего числа требований на обслуживании

№ п/п	Наименование величины	Обозначение	Значение
1	Среднее число требований на обслуживании	\bar{m}	5
2	Второй момент числа требований на обслуживании	\bar{m}^2	24,999
3	Дисперсия числа требований на обслуживании	δ_m^2	$1,758 \cdot 10^{-4}$
4	СКО числа требований на обслуживании	δ_m	0,013
5	Коэффициент вариации числа требований на обслуживании	$K_{\bar{m}}$	$2,652 \cdot 10^{-3}$

Таблица 4.6 – Ключевые числовые характеристики, касающиеся среднего числа требований в очереди

№ п/п	Наименование величины	Обозначение	Значение
1	Среднее число требований в очереди	\bar{l}	8,637
2	Второй момент числа требований в очереди	\bar{l}^2	82,031
3	Дисперсия числа требований очереди	δ_l^2	7,428
4	СКО числа требований в очереди	δ_l	2,726
5	Коэффициент вариации числа требований в очереди	K_l	0,316

В соответствии с полученными в таблицах 4.4-4.6 данными $\bar{l} + \bar{m} = \bar{k}$, что подтверждает адекватность полученной модели. Второй момент величины определяется как математическое ожидание квадрата этой величины. Дисперсия характеризует меру разброса значений случайной величины от ее математического ожидания. СКО, в свою очередь, также является мерой разброса, но в тех же единицах измерения, что и исходная величина. Полученные значения дисперсии и СКО позволяют получить информацию о надежности и достоверности полученных значений при создании модели реальных СМО. Коэффициент вариации в данном случае позволяет оценить степень изменчивости величины. Он показывает меру отклонения СКО от математического ожидания и может быть выражен в процентах. В случаях, при которых данный коэффициент превышает 100 %, анализируемый параметр можно считать неоднородным, а характер функционирования системы с точки зрения эксплуатации – нестабильным [82], ввиду большого разброса величины относительно среднего значения.

В ряде современных СМО эффективность работы напрямую зависит от подбора необходимого числа обслуживающих устройств m . Изменение этого параметра существенно влияет на производительность системы, и в ряде случаев, является рациональным решением для поиска более эффективных режимов работы СМО. Более того, в большом количестве реальных систем увеличение числа обслуживающих устройств является единственным возможным способом повышения производительности. Увеличение количества станций обслуживания дает существенные результаты, в то время как кратное изменение других параметров системы дает меньший прирост эффективности. Ввиду своей дискретности изменение данного параметра является наиболее наглядным способом отображения влияния этой величины на параметры эффективности работы СМО.

В рамках исследования реакций СМО важное значение имеют параметры, характеризующие распределение заявок в системе. На представленном на рис. 4.2 графике показана зависимость \bar{k} , \bar{m} и \bar{l} от числа доступных станций обслуживания (m).

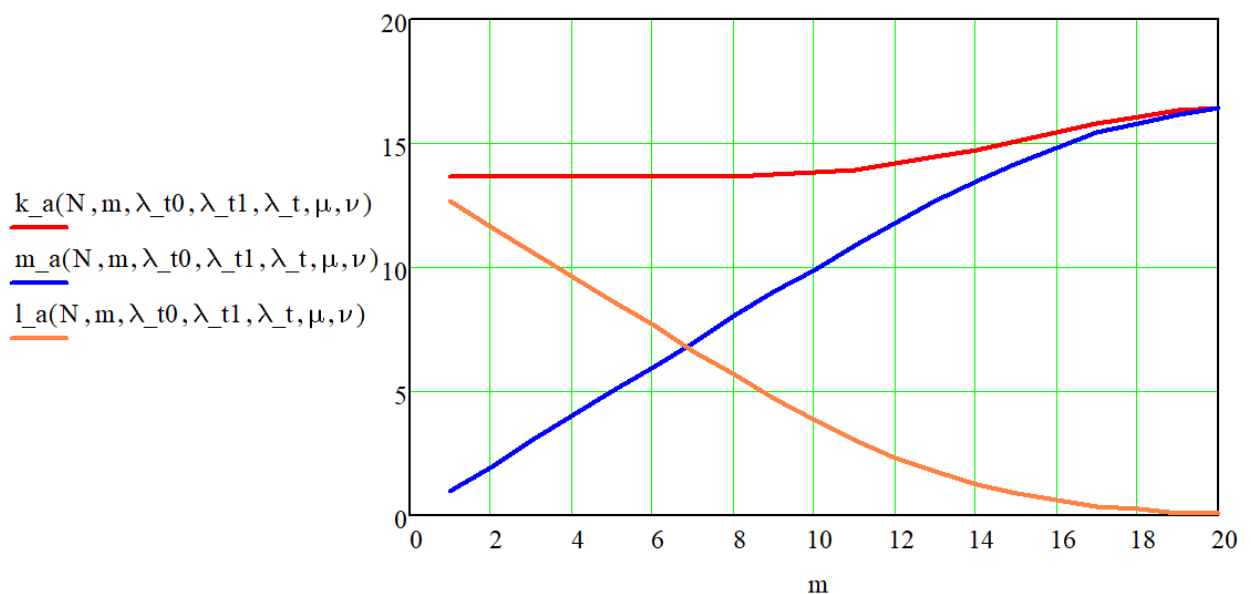


Рис. 4.2 – Зависимость \bar{k} , \bar{m} и \bar{l} от числа доступных станций обслуживания (m)

Среднее число заявок в системе \bar{k} в численном виде представляет собой количество клиентов, находящихся как на обслуживании, так и ожидающих в

очереди. Этот параметр позволяет определить, насколько эффективно система справляется с потоком клиентов. Среднее число заявок на обслуживании \bar{m} , в свою очередь, помогает оценить, насколько эффективно станции обслуживания справляются с текущим спросом. По графику видно, что этот показатель линейно увеличивается с ростом числа обслуживающих устройств. Однако в диапазоне значений $m \in [14 \dots 20]$ рост замедляется, свидетельствуя о простое оборудования, т.к. при таком количестве обслуживающих устройств новые заявки не успевают поступать в систему. На этих же значениях m параметр \bar{l} стремится к нулю, что, в свою очередь, свидетельствует о несоответствии между возможностями обслуживания и интенсивностью потока входящих заявок, что продемонстрировано на рис. 4.3.

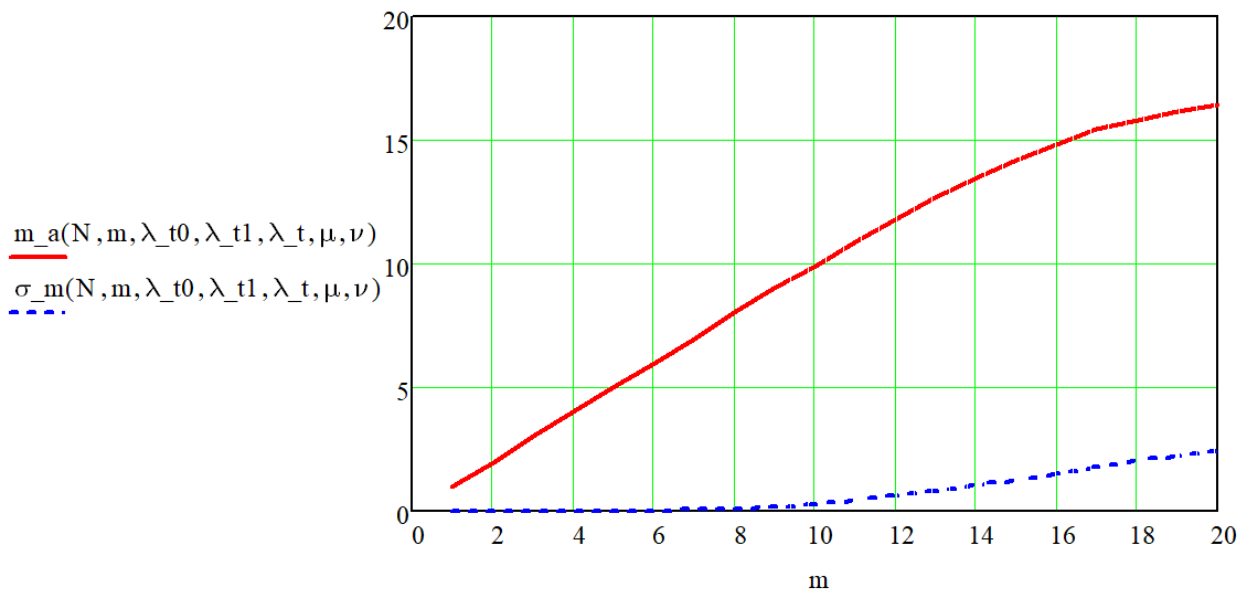


Рис. 4.3 – Зависимость \bar{m} и $\delta_{\bar{m}}$ от числа доступных обслуживающих устройств m

Так как параметр $\delta_{\bar{m}}$ в данном случае значительно меньше параметра \bar{m} , можно сделать вывод о достоверности этого параметра. Рассмотрим отдельно полученную зависимость для \bar{l} от числа доступных станций обслуживания m .

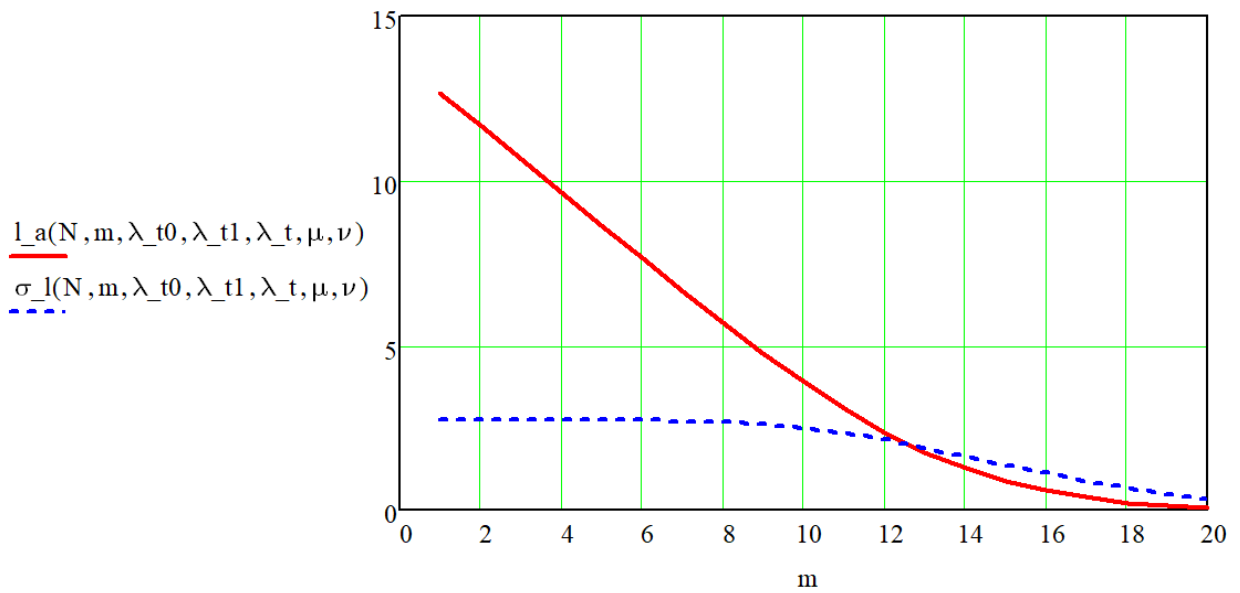


Рис. 4.4 – Зависимость \bar{l} и δ_l от числа обслуживаемых устройств m

По графику видно, что при m равном 12 и более значения δ_l превышают значения \bar{l} , что свидетельствует о переходе значения $K_{\bar{l}}$ порога в 100 % и должно быть принято во внимание для каждого отдельно взятого реального объекта и числа обслуживаемых устройств. Так, для данного примера получим, что действительное значение средней длины очереди при $m=12$ находится в пределах от 0 до 4,8. Однако, с ростом количества обслуживаемых устройств, \bar{l} и δ_l стремятся к нулю, в связи с чем, значение $K_{\bar{l}}$, превышающее 100 %, не будет являться мерой существенного разброса величины \bar{l} .

Теперь рассмотрим изменение параметров интенсивности входного потока и результирующее воздействие на параметры эффективности работы СМО. Для упрощения задачи и более наглядного анализа влияния различных компонент потока на параметры системы, примем начальное допущение о равенстве интенсивностей $(\lambda_0, \lambda_1, \lambda)$, которые в свою очередь составляют входящий поток заявок λ_0 .

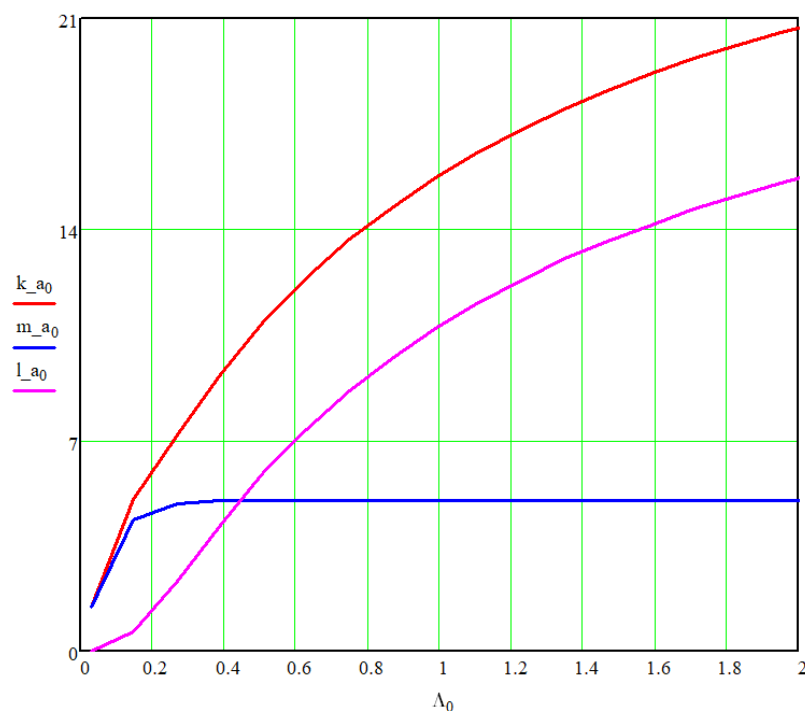


Рис. 4.5 – Зависимость \bar{k} , \bar{m} и \bar{l} от Λ_0

В данном случае число обслуживающих устройств фиксировано ($m=5$).

При увеличении суммарной интенсивности входного потока Λ_0 значения \bar{k} , \bar{m} и \bar{l} также возрастают, что свидетельствует о положительной корреляции между этими параметрами и Λ_0 .

Рассмотрим также случай отсутствия во входящем потоке заявок 0-го типа. При такой конфигурации входного потока рассматриваемая модель СМО легко сводится к модели, описанной в трудах А.П. Кирпичникова [3] путем подстановки в соответствующие формулы значения $\lambda_0=0$. Примем дополнительно, что суммарная величина интенсивности входящего потока находится на том же уровне за счет увеличения интенсивности потоков заявок 1-го и 2-го типа, что позволит построить графики зависимостей характеристик системы от Λ_0 в той же размерности. Для наглядности здесь и в дальнейшем значения величины, соответствующие отсутствию 0-й компоненты потока обозначены с индексом 1, а при наличии этой компоненты с индексом 0. Полученные зависимости приведены на рис. 4.6 – 4.9.

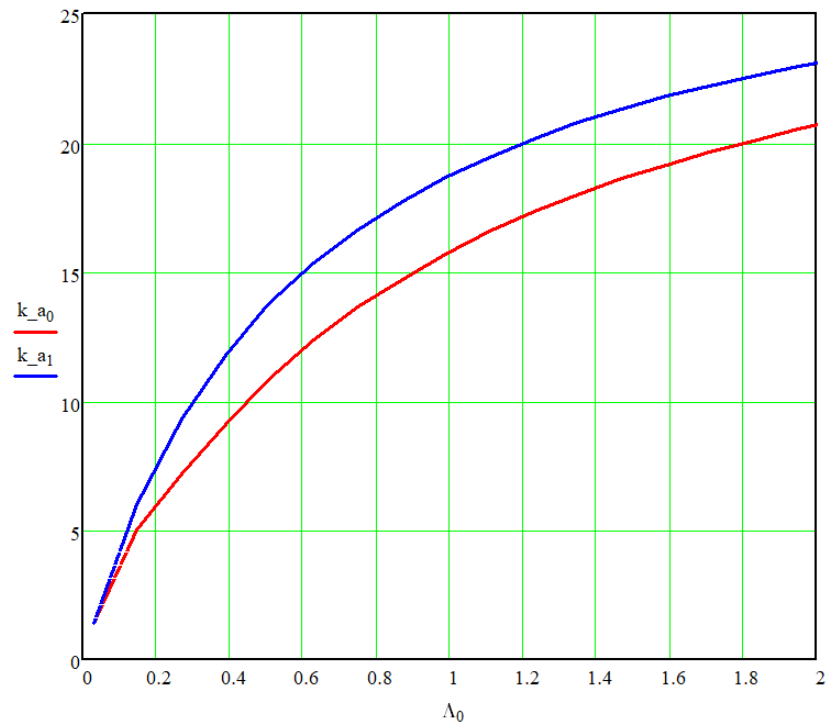


Рис. 4.6 – Зависимость \bar{k} от Λ_0 ;

- красная кривая (k_{a0}) соответствует трехкомпонентному потоку;
- синяя кривая (k_{a1}) соответствует двухкомпонентному потоку.

По полученной зависимости можно оценить влияние 0-й компоненты потока на характер поведения системы. Так как при наличии очереди 0-я компонента потока получает отказ в обслуживании, график \bar{k}_0 находится ниже графика \bar{k}_1 . В рассматриваемом примере для $\Lambda_0=0,75$, среднее число заявок в системе \bar{k} при отсутствии отказов в обслуживании составит 16,7, а при наличии в потоке заявок 0-го типа составит лишь 13,6. Таким образом, можно сделать вывод, что наличие в суммарном входящем потоке заявок 0-го типа ведет к снижению параметра \bar{k} ввиду отказов в обслуживании.

Зависимости $\delta_{\bar{k}}$ от Λ_0 представлены на рис. 4.7. Очевидно, что на всем промежутке значений Λ_0 , значение $K_{\bar{k}}$ менее 100 %.

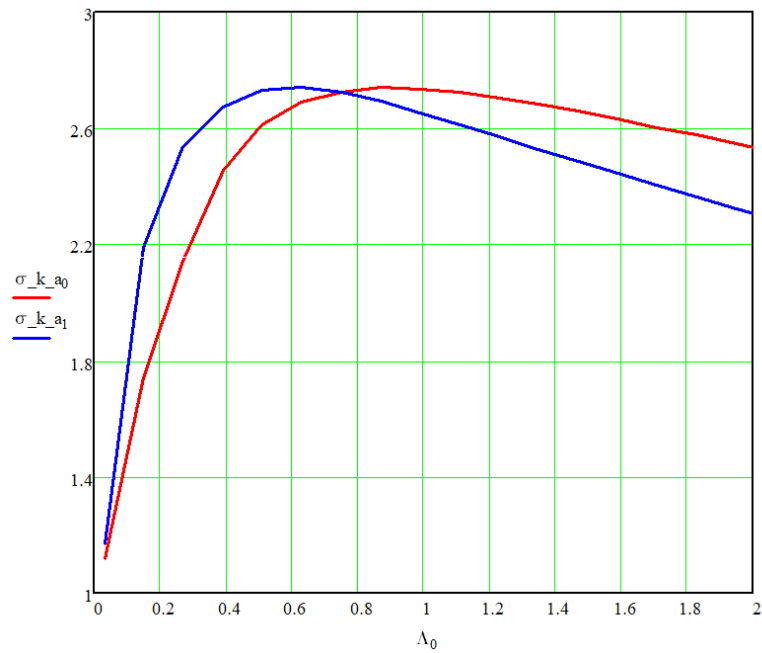


Рис. 4.7 – Зависимость $\delta_{\bar{k}}$ от Λ_0 :

- красная кривая ($\sigma_k_a_0$) соответствует трехкомпонентному потоку;
- синяя кривая ($\sigma_k_a_1$) соответствует двухкомпонентному потоку.

По аналогии построим график зависимости \bar{l} от Λ_0 .

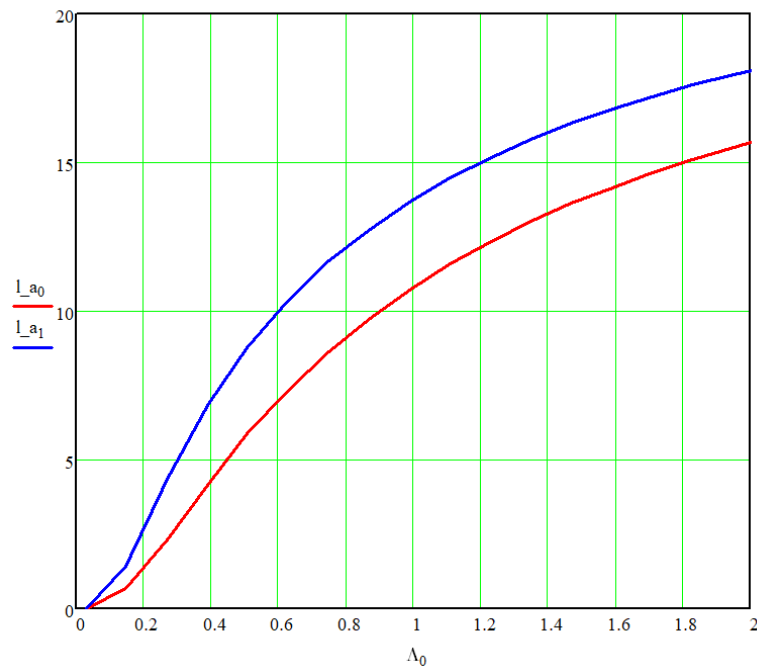


Рис. 4.8 – Зависимость \bar{l} от Λ_0 :

- красная кривая (l_a_0) соответствует трехкомпонентному потоку;
- синяя кривая (l_a_1) соответствует двухкомпонентному потоку.

Получена обратная ситуация, когда при отсутствии в потоке заявок 0-го типа увеличивается потенциальная длина очереди. Это вызвано тем, что источник содержит уже не три, а два типа заявок, не получающих отказ в обслуживании. При этом интенсивность входящего потока остается неизменной.

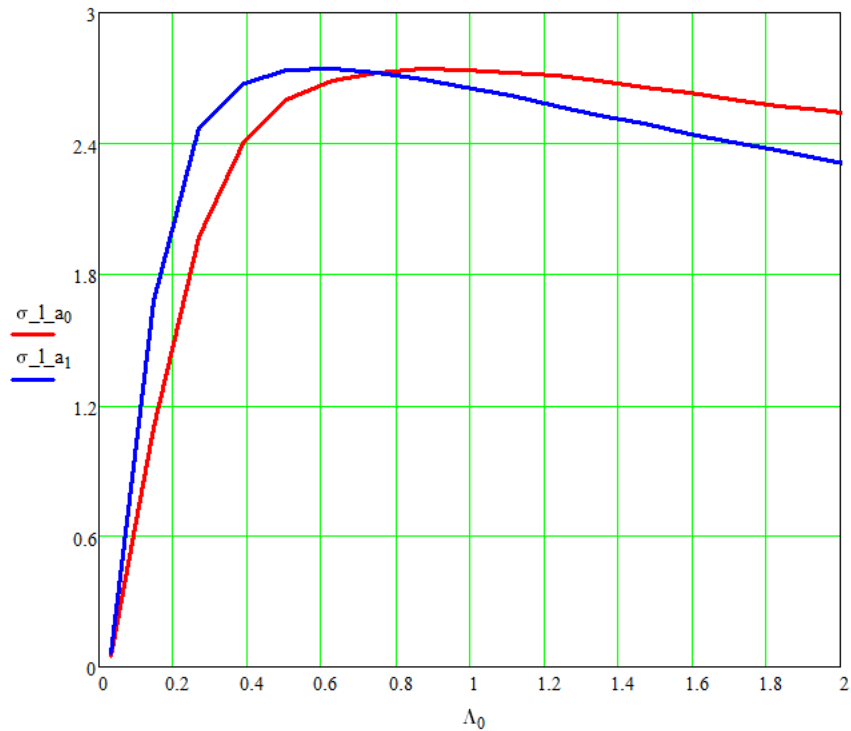


Рис. 4.9 – Зависимость δ_l от Λ_0 :

- красная кривая (σ_{l_a0}) соответствует трехкомпонентному потоку;
- синяя кривая (σ_{l_a1}) соответствует двухкомпонентному потоку.

Теперь рассмотрим ситуацию со средним количеством заявок, находящихся на обслуживании \bar{m} .

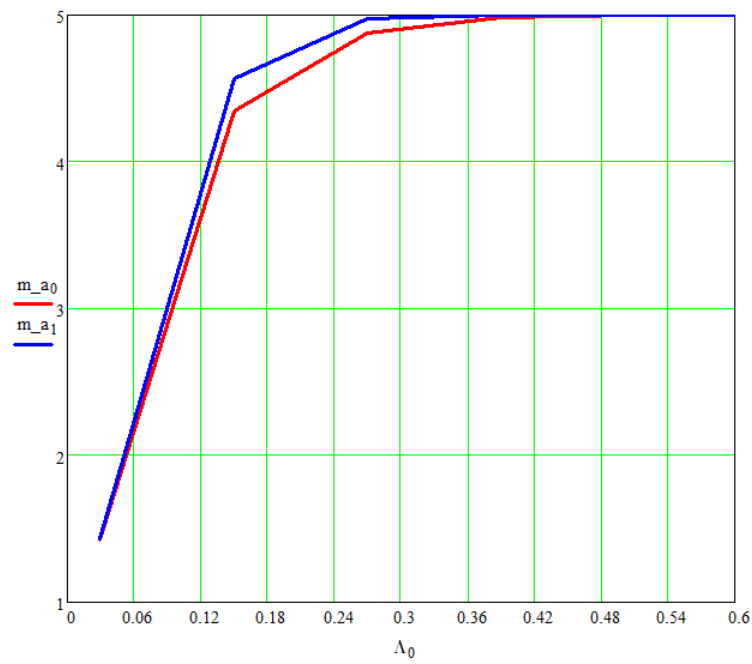


Рис. 4.10 – Зависимость \bar{m} от Λ_0 :

- красная кривая (m_{a0}) соответствует трехкомпонентному потоку;
- синяя кривая (m_{a1}) соответствует двухкомпонентному потоку.

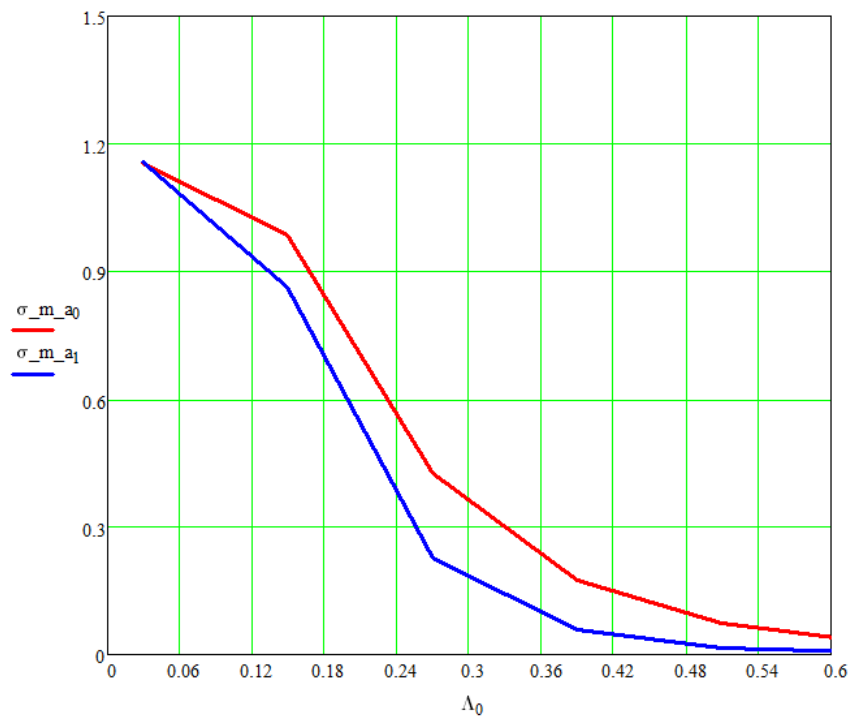


Рис. 4.11 – Зависимость $\delta_{\bar{m}}$ от Λ_0 :

- красная кривая ($\sigma_{m_{a0}}$) соответствует трехкомпонентному потоку;
- синяя кривая ($\sigma_{m_{a1}}$) соответствует двухкомпонентному потоку.

Как видно по графику, изображенному на рис. 4.10, в случае отсутствия 0-й компоненты в потоке заявки незначительно быстрее попадают на обслуживание при малых значениях параметра λ_0 , что, однако, является несущественным, т.к. при увеличении величины λ_0 все обслуживающие устройства оказываются заняты, и заявки вынуждены ожидать в очереди.

Следует отметить, что здесь и далее рассмотренная ситуация актуальна лишь при малом количестве обслуживающих устройств m по сравнению с общим количеством заявок в источнике N . В случае, если количество обслуживающих устройств будет соизмеримо с количеством заявок в системе, данный параметр будет существенно видоизменяться ввиду наличия в системе простоя оборудования. Приведенная ситуация в рамках текущей работы не рассматривается.

В рамках анализа СМО важным аспектом является оценка интенсивности обслуживания, которая непосредственно влияет на эффективность функционирования системы. Интенсивность обслуживания (μ) отражает количество заявок, которое обрабатывает одно обслуживающее устройство за единицу времени. На практике изменение интенсивности обслуживания не всегда возможно ввиду специфики реальных систем. Данный процесс так или иначе связан с изменением режима функционирования обслуживающей станции, что, в свою очередь, может вызвать существенные и необоснованные экономические затраты. Рассмотрим изменение данного параметра на предмет исследования реакций СМО. На рис. 4.12 представлена зависимость уже рассмотренных выше параметров эффективности функционирования системы от интенсивности обслуживания μ .

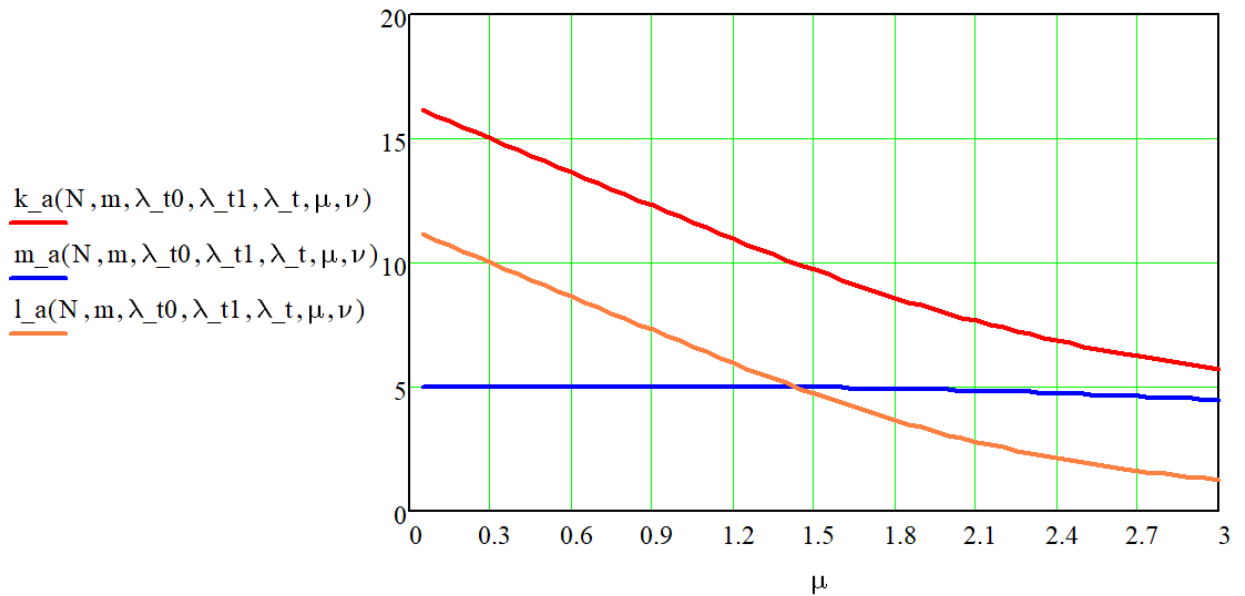


Рис. 4.12 – Зависимость \bar{k} , \bar{m} и \bar{l} от μ

По графику на рис. 4.12 видно, что изменение μ в рассматриваемом примере практически не влияет на число занятых обслуживающих устройств \bar{m} . Анализ зависимостей $\bar{k}(\mu)$ и $\bar{l}(\mu)$ показал, что оба параметра демонстрируют обратную корреляцию с интенсивностью обслуживания μ .

Так, если рассмотреть отрезок $\mu [0,9 \dots 2,7]$ получим, что μ увеличилась в 3 раза, при этом получим снижение \bar{l} с 7,3 до 1,6 заявки (в 4,6 раз).

Рассмотрим отдельно параметр \bar{l} , однако, для оценки реакций системы, дополним график кривой со значительным изменением параметра ν . Ввиду специфики рассматриваемой математической модели, заявка, находящаяся в очереди, может покинуть систему двумя способами – после обслуживания и по истечении граничного времени ожидания в очереди.

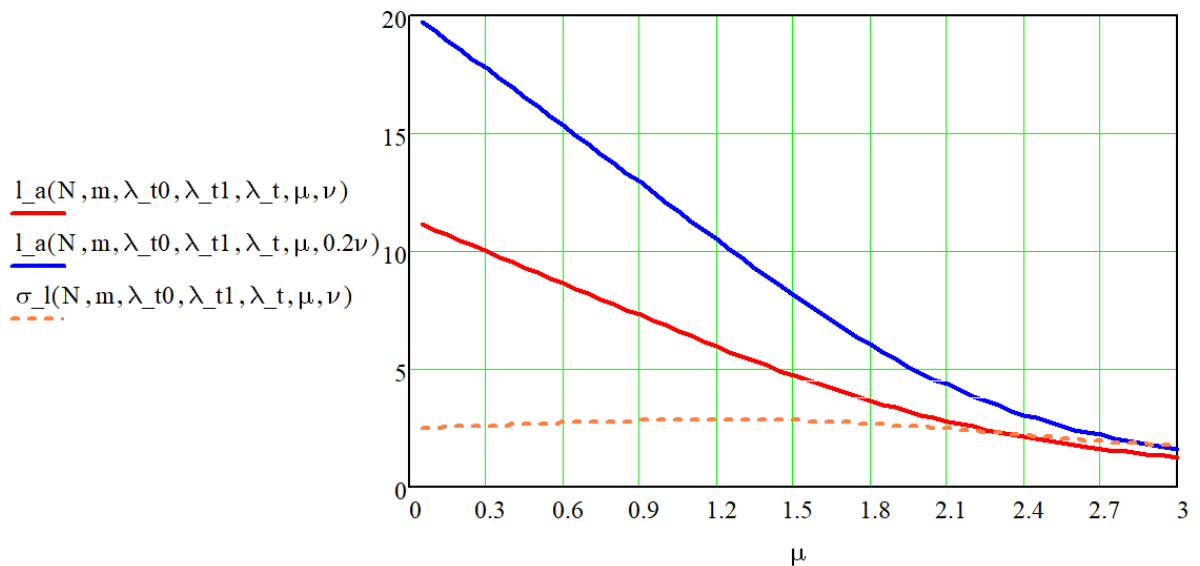


Рис. 4.13 – Зависимость \bar{l} и δ_l от μ

Интенсивность вылета заявок из системы ν в данном случае является мерой фиксированного предела времени ожидания, и ее значительное уменьшение в соответствии с графиком $(0, 2 \cdot \nu)$ способствует значительному росту \bar{l} ввиду факторов, описанных выше.

На следующем графике приведена зависимость \bar{m} от параметра μ . Результатом увеличения данного параметра выше значения 1,2 является режим функционирования системы, при котором μ превышает количество поступающих заявок, и возможен простой оборудования, что нехарактерно для СМО с ожиданием.

Также на графике рассмотрена ситуация с параллельным уменьшением интенсивности вылета заявок $(0, 2 \cdot \nu)$, результатом чего является незначительная реакция системы. Более наглядно эта ситуация представлена на рис. 4.14.

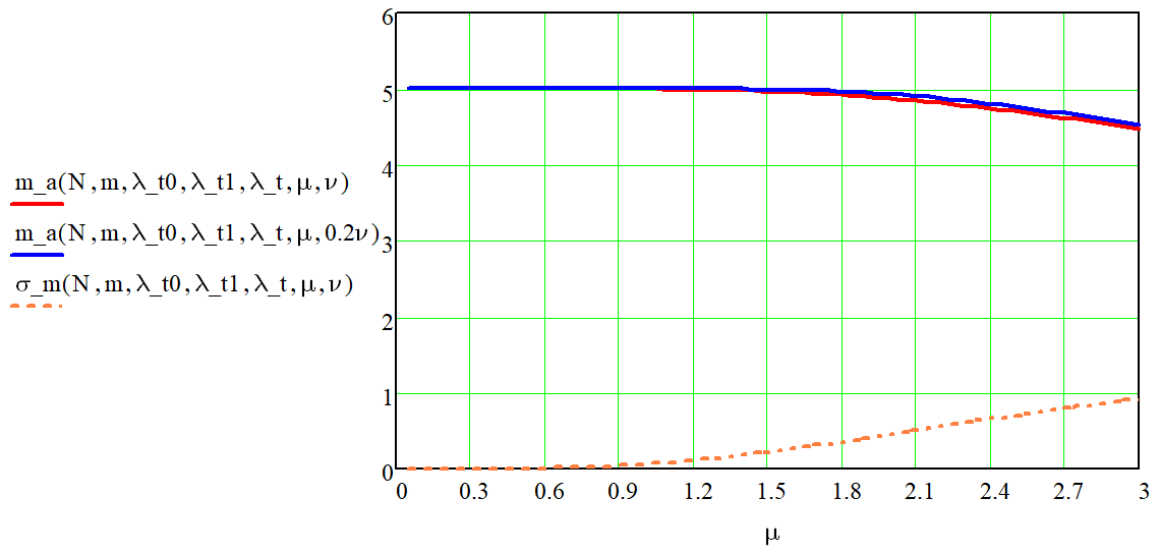


Рис. 4.14 – Зависимость \bar{m} и $\delta_{\bar{m}}$ от μ

Рассмотрим более подробно изменение параметра ν . На рис. 4.15 представлена зависимость показателей загруженности системы от ν , которая демонстрирует, как колебания граничного времени ожидания в очереди заявок 1-го типа влияют на метрики системы. Следует подчеркнуть, что интенсивность ухода заявок из очереди в рамках данной работы не является инструментом прямого управления системой. Ее значение формируется эмпирически на основе поведения заявок типа 1, поскольку именно их статистические характеристики определяют динамику очереди. Таким образом, система реагирует на внешние факторы, а не задаёт их искусственно.

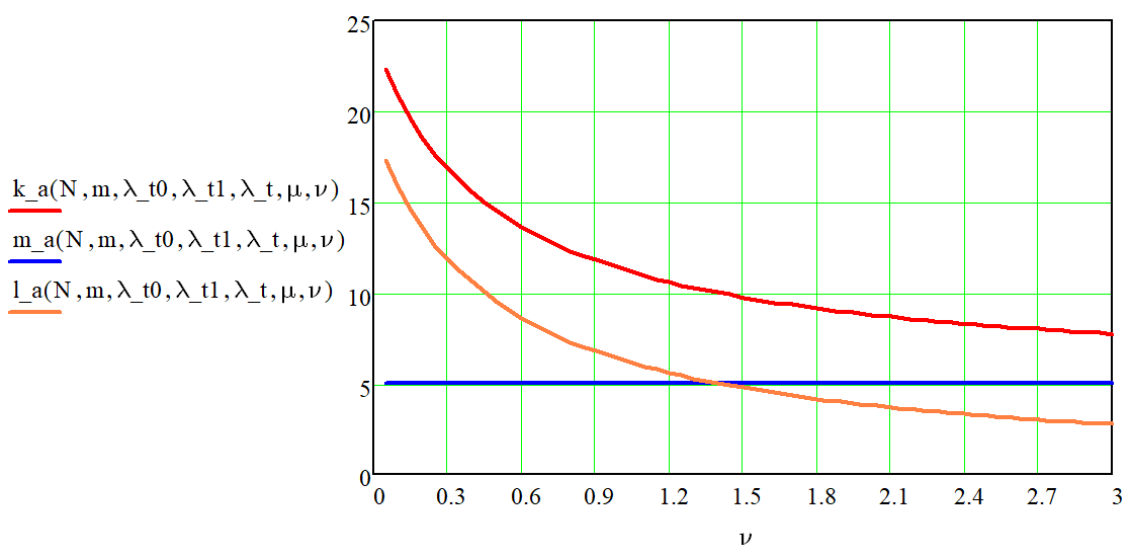


Рис. 4.15 – Зависимость \bar{k} , \bar{m} и \bar{l} от ν

Из сравнения данного графика с графиком, представленным на рис. 4.13 следует, что уменьшение значения параметра граничного времени ожидания, как и в случае увеличения интенсивности обслуживания, сопровождается снижением средней длины очереди. При этом в случае с граничным временем ожидания будет иметь место рост доли необслуженных заявок 1-го типа, что объясняется их ускоренным уходом из очереди.

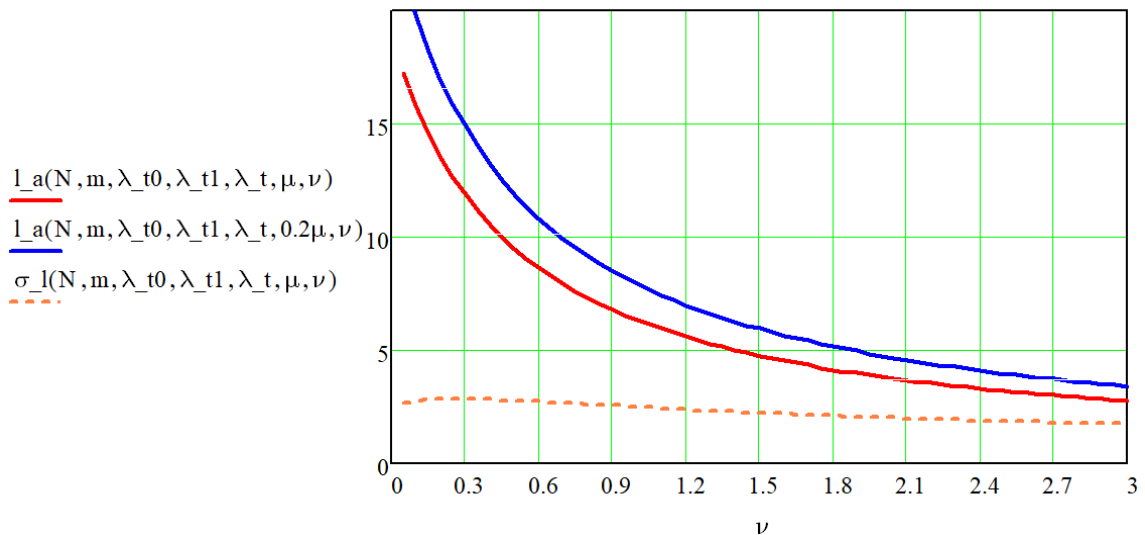


Рис. 4.16 – Зависимость \bar{l} и δ_l от ν

По приведенному на рис. 4.16 графику можно сделать вывод о том, что при сокращении интенсивности обслуживания ($0,2 \cdot \mu$) кривая среднего числа заявок в очереди будет находиться незначительно выше кривой среднего числа заявок в очереди при неизменной интенсивности вылета заявок ν .

Таким образом, одновременное снижения интенсивности обслуживания и интенсивности уход заявок из очереди приводят к увеличению средней длины очереди \bar{l} . При этом важно отметить, что изменение количества обслуживающих устройств носит дискретный характер, тогда как изменение данных параметров допускает непрерывное воздействие на среднюю длину очереди. Несмотря на аналогию в их воздействии на систему, указанные факторы демонстрируют принципиальные функциональные различия и асимметрию последствий. Снижение граничного времени ожидания приводит к сокращению очереди за счет роста числа необслуженных заявок

типа 1, обусловленного их преждевременным уходом из системы. Увеличение интенсивности обслуживания μ также снижает \bar{l} , однако этот эффект достигается за счет повышения количества обслуженных заявок, что положительно влияет на пропускную способность системы и увеличивает уровень ее загрузки. При проектировании СМО критически важно учитывать эти различия и обеспечивать оптимальное соотношение между параметрами μ и ν .

Рассмотрим отдельно влияние параметра μ и ν на среднюю длину очереди на примерах графиков рис. 4.16 и 4.13. Очевидно, что степень влияния скорости обслуживания и интенсивности уходов заявок на параметр средней длины очереди различна на разных участках графика на рис. 4.13. По графику, представленному на рис. 4.13, можно определить, что параметр \bar{l} при $\mu \in [0; 1,2)$ более чувствителен к изменению параметра ν , нежели к изменению параметра μ . На участке $\mu \in (1,2; 3)$ ситуация противоположна, и равнозначное изменение скорости обслуживания μ вызывает большие колебания величины средней длины очереди \bar{l} . Полученный вывод визуально подтверждается графиком, построенным для значения $\mu = 0,6$ и представленным на рис. 4.16, где на всем промежутке изменения параметра ν параметр \bar{l} более чувствителен к равнозначному изменению параметра ν , нежели к изменению параметра μ .

Для текущих исходных данных $\delta_{\bar{l}}$ в соответствии с графиком на рис. 4.16 не превышает величину \bar{l} и практически не изменяется с ростом параметра ν , что говорит о стабильности значений параметра средней длины очереди в рассматриваемом примере.

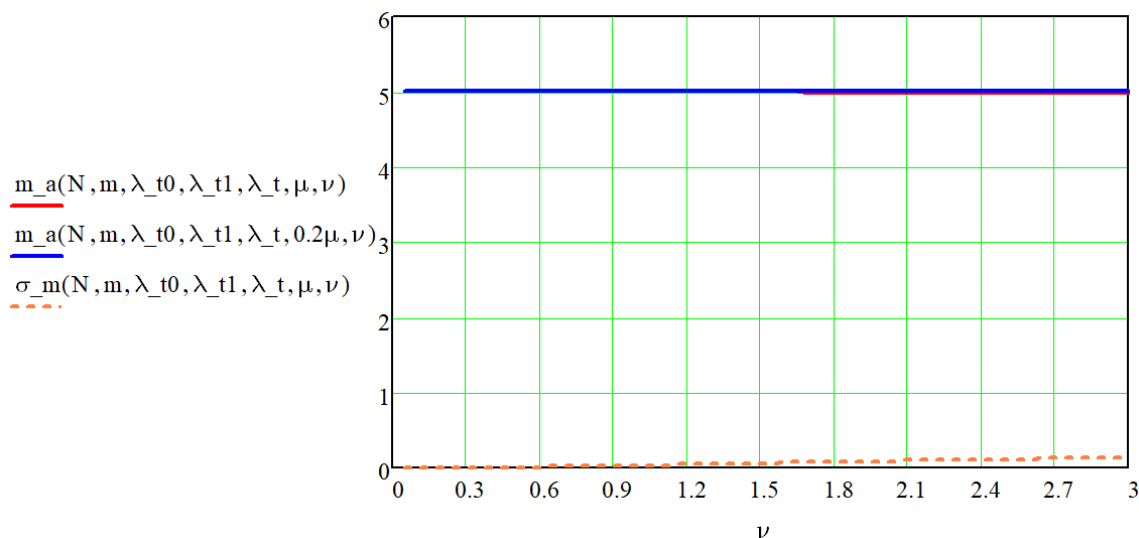


Рис. 4.17 – Зависимость \bar{m} и $\delta_{\bar{m}}$ от ν

Изменение параметра ν , в отличие от изменения интенсивности обслуживания μ , практически не влияет на изменение \bar{m} ввиду того, что параметр μ управляет заявками всех трех типов, а ограничение на длину очереди действует лишь на заявки 1-го типа.

4.4 Пропускная способность СМО

В таблице 4.7 приведены результаты расчета абсолютной и относительной пропускной способности СМО.

Таблица 4.7 – Результаты расчета абсолютной и относительной пропускной способности СМО

№ п/п	Наименование величины	Обозначение	Значение
1	Абсолютная пропускная способность	A	3,0
2	Относительная пропускная способность	q	0,244

В соответствии с полученными результатами, можно сделать вывод, что при указанных в таблице 4.1 исходных данных рассматриваемая СМО обрабатывает около 3 заявок в единицу времени, что соотносится с количеством обслуживающих устройств и потоком обслуживания. В относительных же величинах получается, что система обслуживает порядка 24 % входящих заявок, а остальные либо получают отказ, либо покидают очередь ввиду ограничения на время ожидания. Рассмотрим возможные способы повышения данного показателя.

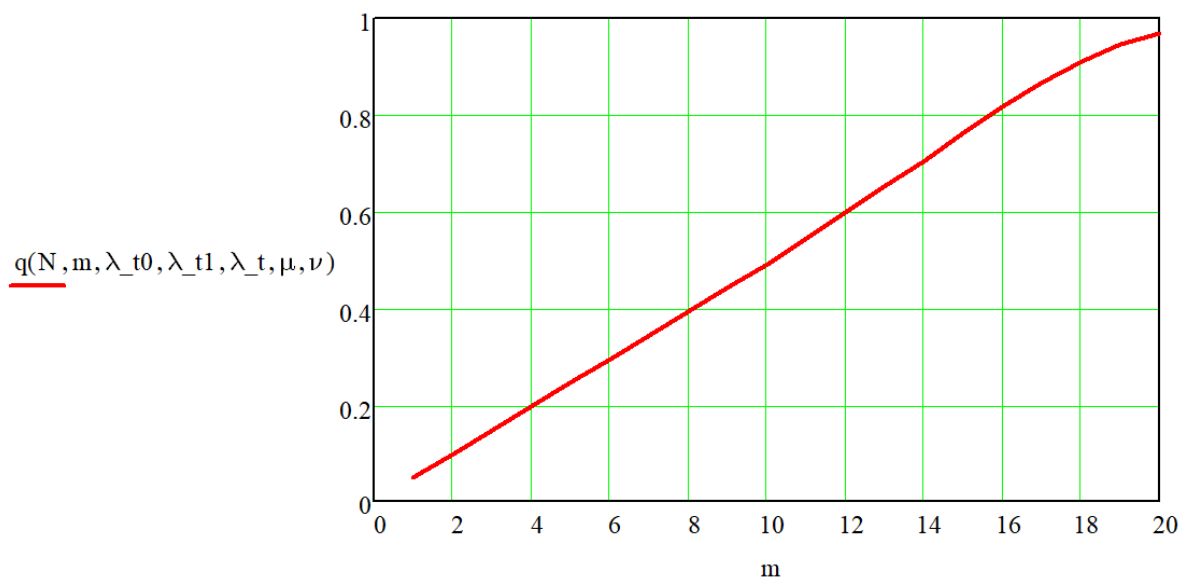


Рис. 4.18 – Зависимость q от числа доступных станций обслуживания m

В соответствии с рис. 4.18 увеличение количества станций обслуживания способствует линейному росту относительной пропускной способности до значения в 97 %.

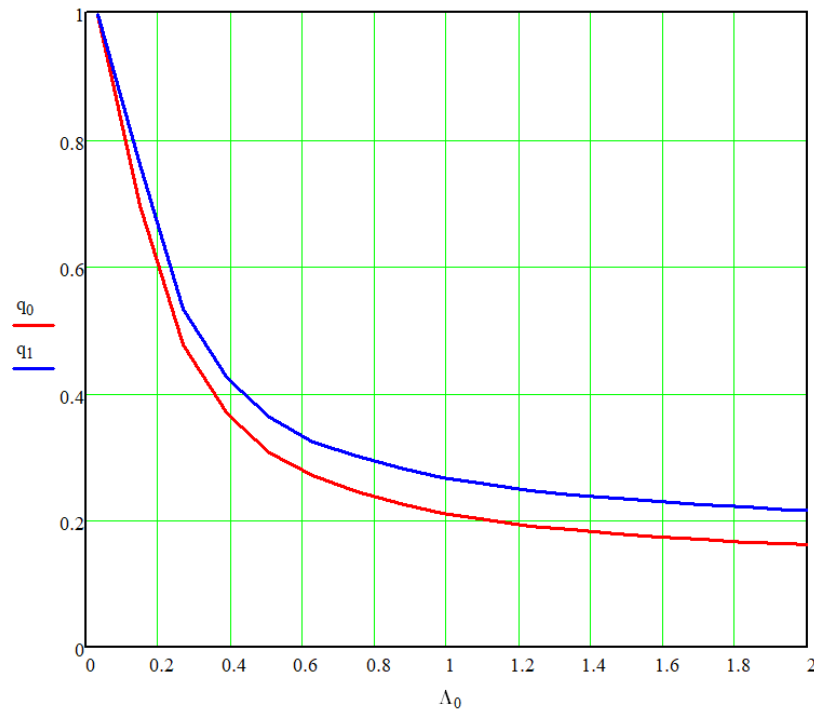


Рис. 4.19 – Зависимость q от интенсивности входящего потока заявок Λ_0 :

- синяя кривая (q_0) соответствует трехкомпонентному потоку;
- красная кривая (q_1) соответствует двухкомпонентному потоку.

В соответствии с рис. 4.19 очевидно, что с ростом количества заявок, поступающих в систему в единицу времени, относительная пропускная способность системы падает. Остановимся более подробно на интенсивности входящего потока Λ_0 . На графике представлены две зависимости параметра q . Для относительной пропускной способности, обозначенной индексом 0, суммарный входящий поток заявок Λ_0 состоит из трех типов заявок $(\lambda_0, \lambda_1, \lambda)$, а для относительной пропускной способности, обозначенной индексом 1 (λ_1, λ) . Исключены заявки 0-го типа с сохранением величины суммарной интенсивности. При этом $\lambda_0 = \lambda_1 = \lambda$. Из полученных зависимостей можно сделать вывод, что наличие заявок 0-го типа в потоке снижает относительную пропускную способность q , т.к. при наличии очереди данный тип заявок получает отказ в обслуживании. Для полученных на рис. 4.19 зависимостей можно сказать, что результатом отсутствия во входном потоке заявок 0-го типа будет повышение потенциально возможной пропускной способности

приблизительно на 10-25 %. Это пример поведения системы при переходе 0-го типа заявок в заявки 1-го и 2-го типа с сохранением суммарной интенсивности λ_0 . Очевидно, если просто исключить заявки 0-го типа из суммарного потока, рост относительной пропускной способности будет более значительным.

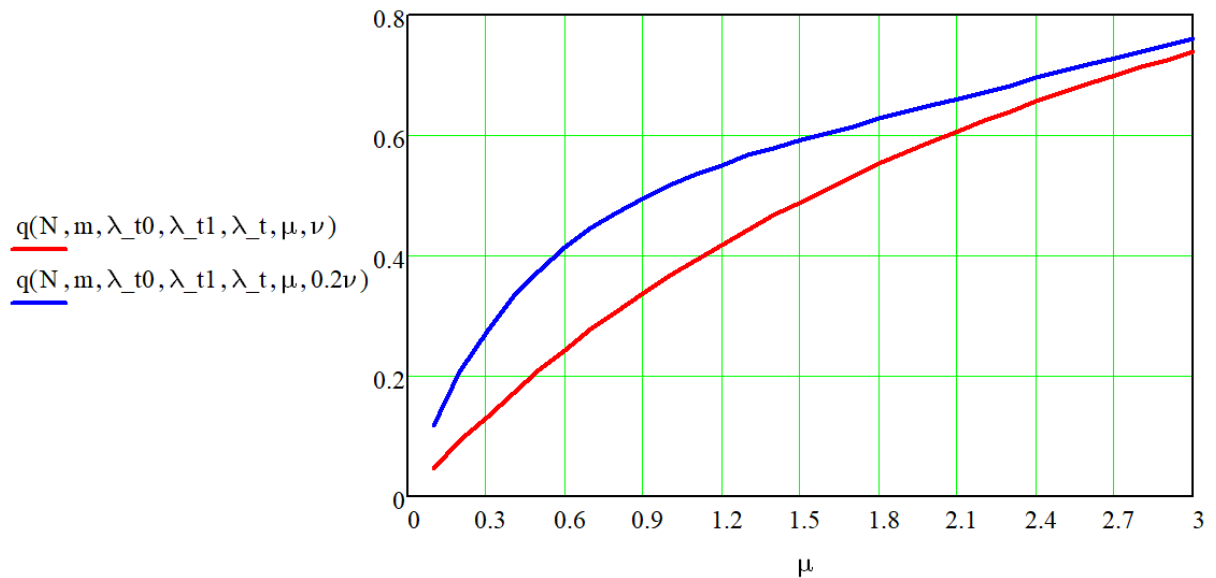


Рис. 4.20 – Зависимость q от μ

В соответствии с графиком на рис. 4.20 увеличение параметра μ для рассматриваемой системы в 2 раза даст прирост относительной пропускной способности приблизительно на 65 %. Приведенную зависимость для рассматриваемой конфигурации СМО можно считать линейной. С уменьшением параметра ν величина входного потока снижается, ввиду роста очереди, что напрямую влияет на величину относительной пропускной способности q , т.к. в таком случае нагрузка на систему снижается.

Рассмотрим более подробно изменение параметра ограничения на время ожидания заявки в очереди.

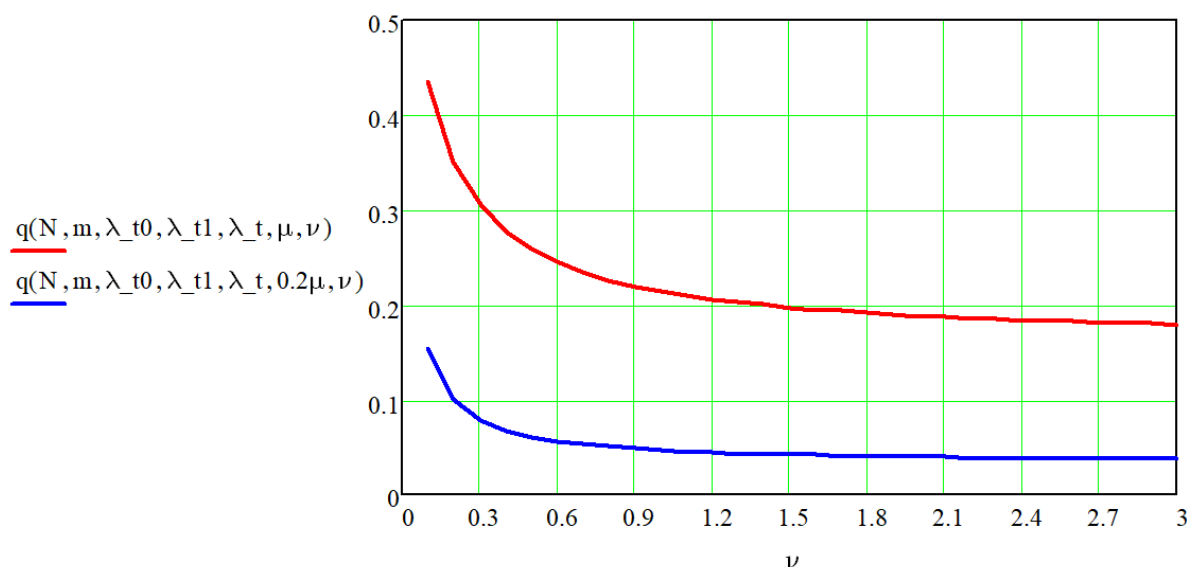


Рис. 4.21 – Зависимость q от ν для потока A_0

Полученная зависимость позволяет судить о том, что уменьшение параметра ν провоцирует рост нагрузки на систему, что, однако, наблюдается лишь в узком диапазоне значений ν . Данный параметр незначительно влияет на относительную пропускную способность и для рассматриваемой системы оказывает влияние на параметр q лишь в небольшом диапазоне значений.

Рассмотрим ситуацию, в которой во входящем потоке отсутствует 0-я компонента. Нагрузка входящего потока на систему в таком случае будет распределена между 1-й и 2-й компонентой.

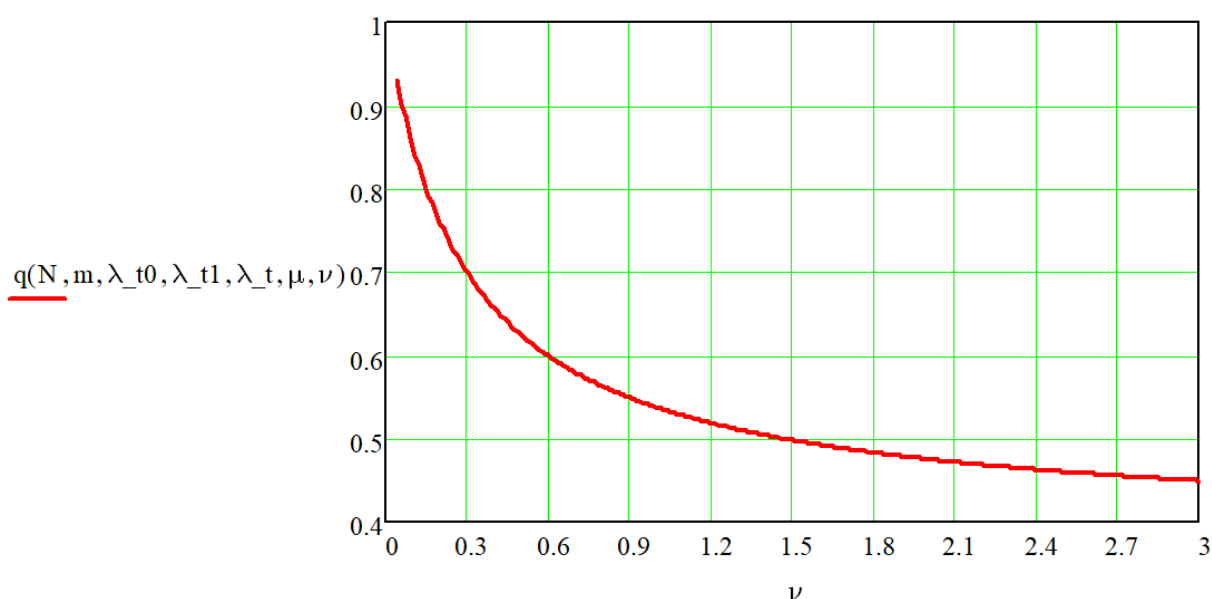


Рис. 4.22 – Зависимость q от ν для потока A_1

В такой ситуации относительная пропускная способность q более чувствительна к изменениям параметра ν . Это вызвано тем, что при отсутствии заявок, которые получают отказ в обслуживании, число заявок в очереди значительно увеличивается, что снижает нагрузку входного канала. Ограничение по времени ожидания в таком случае оказывает значительное влияние на долю необслуженных заявок. Т.к. необслуженные заявки снова поступают на вход системы, входящий поток заявок в таком случае кратно увеличивается, увеличивая нагрузку на обслуживающие устройства и значительно снижая относительную пропускную способность.

4.5 Временные характеристики СМО

В таблицах 4.8, 4.9, 4.10 представлены результаты расчета временных характеристик СМО и параметров, характеризующих их разброс.

Таблица 4.8 – Среднее время ожидания и его ключевые характеристики

№ п/п	Наименование величины	Обозначение	Значение
1	Среднее время ожидания в очереди	$\overline{t_{ож}}$	1,056
2	Дисперсия времени ожидания	$\delta_{\overline{t_{ож}}}^2$	0,142
3	СКО времени ожидания	$\delta_{\overline{t_{ож}}}$	0,377
4	Коэффициент вариации времени ожидания	$K_{\overline{t_{ож}}}$	0,357

Таблица 4.9 – Среднее время обслуживания и его ключевые характеристики

№ п/п	Наименование величины	Обозначение	Значение
1	Среднее время обслуживания	$\overline{t_{обсл}}$	1,667
2	Дисперсия времени обслуживания	$\delta_{\overline{t_{обсл}}}^2$	2,778

№ п/п	Наименование величины	Обозначение	Значение
3	СКО времени обслуживания	$\delta_{\overline{t_{обсл}}}$	1,667
4	Коэффициент вариации времени обслуживания	$K_{\overline{t_{обсл}}}$	1

Таблица 4.10 – Среднее время ожидания и его ключевые характеристики

№ п/п	Наименование величины	Обозначение	Значение
1	Среднее время нахождения заявки в системе	$\overline{t_{общ}}$	2,722
2	Дисперсия времени нахождения заявки в системе	$\delta_{\overline{t_{общ}}}^2$	2,92
3	СКО времени нахождения заявки в системе	$\delta_{\overline{t_{общ}}}$	1,709
4	Коэффициент вариации времени нахождения заявки в системе	$K_{\overline{t_{общ}}}$	0,628

В рамках определения времени нахождения заявки в системе $\overline{t_{общ}}$ можно выделить время обслуживания $\overline{t_{обсл}}$ и время ожидания $\overline{t_{ож}}$, сумма которых составляет полное время нахождения заявки в системе. Время обслуживания в рассматриваемом случае является случайной величиной, которое распределено по экспоненциальному закону, и его математическое ожидание известно. Таким образом, единственной действительно зависимой от изменения начальных переменных величиной служит время нахождения заявки в очереди. Рассмотрим некоторые из возможных вариантов изменения данного параметра.

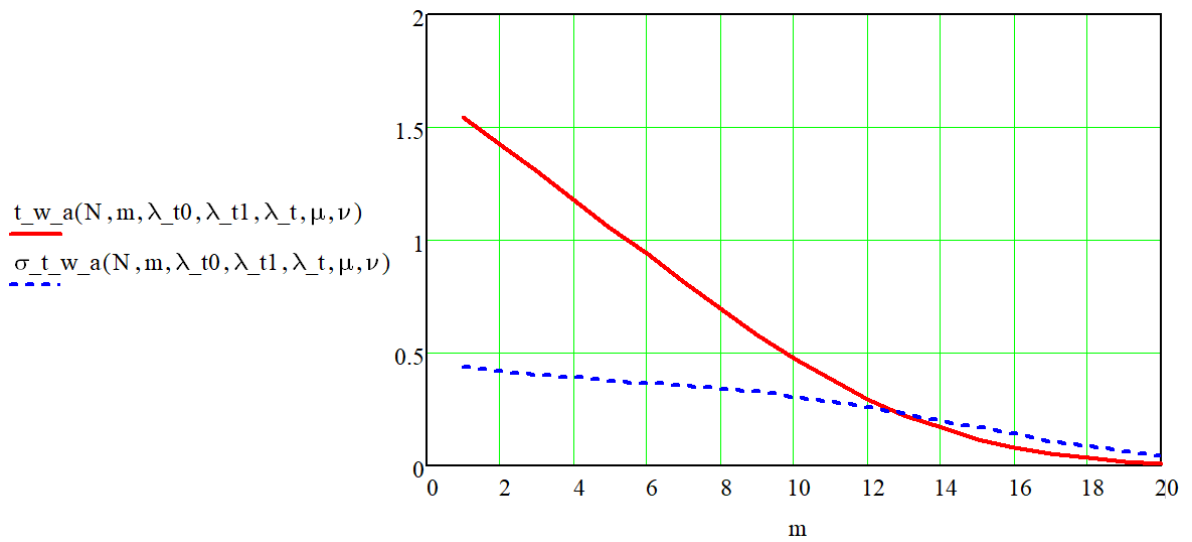


Рис. 4.23 – Зависимость $\overline{t_{ожс}}$ и $\delta_{\overline{t_{ожс}}}$ от числа доступных станций обслуживания m

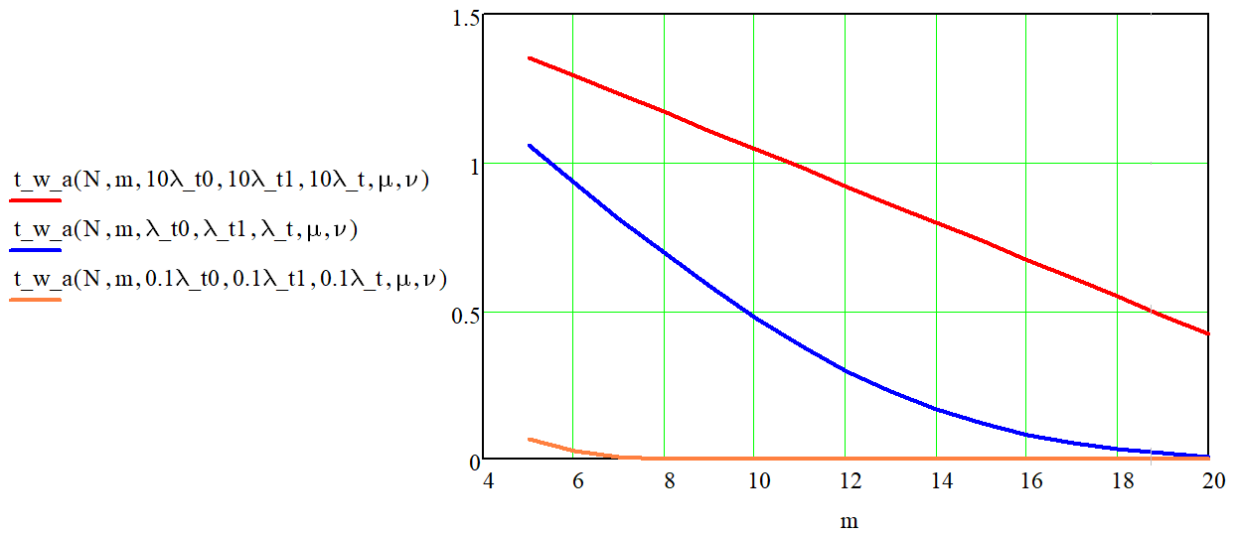


Рис. 4.24 – Зависимость $\overline{t_{ожс}}$ от числа доступных станций обслуживания m при различных параметрах λ_i

На рис 4.23, рис. 4.24 рассмотрен вариант изменения числа обслуживающих устройств m и его влияние на $\overline{t_{ожс}}$. В соответствии с таблицей 4.8 среднее время ожидания составляет 1,056. Ограничение на время пребывания заявки в очереди можно выразить как $1/\nu$. В рассматриваемом случае это значение составит 1,666. Таким образом, среднее время ожидания $\overline{t_{ожс}}$ не превышает ограничение на время нахождения заявки в очереди, в связи с чем в рассматриваемом примере редко имеют место

уходы «нетерпеливых» заявок. Тем не менее, необходимо учитывать, что мы имеем дело со случайным потоком событий и его математическим ожиданием. В результате наблюдается определенный разброс значений определяемых параметров. В данном случае $\delta_{\overline{t_{ож}}}$ составляет 0,377, в связи с чем можно сказать, что $\overline{t_{ож}}=1,056\pm 0,377$. Ограничение на время пребывания заявки в очереди также является случайной величиной с известным математическим ожиданием. Т.к. распределение этой величины экспоненциальное, ее СКО будет равно математическому ожиданию. В результате для ограничения на время ожидания получим $v=1,666\pm 1,666$.

Характерно для полученной на рис. 4.23 зависимости то, что на участке $m\in[9;20]$ зависимость среднего времени ожидания от числа обслуживающих устройств m становится нелинейной. Такой результат обусловлен следующими причинами:

1. Наличие заявок 0-го типа не позволяет перенаправить все заявки в систему, полностью истощив источник. Так как данные заявки не встают в очередь, даже при большом количестве обслуживающих устройств они могут получить отказ, подпитывая тем самым входной поток и увеличивая как среднюю длину очереди, так и среднее время ожидания.

2. Ограничение по времени ожидания также препятствует полному исключению времени ожидания, т.к. некоторые заявки все равно будут ожидать в очереди, пока не достигнут предела времени ожидания. Это создает дополнительное давление на систему, т.к. заявки, которые не были обслужены в течение заданного времени, покидают систему, что приводит к увеличению входного потока.

В случае необходимости сокращения доли заявок, покидающих систему необслуженными, $\overline{t_{ож}}$ должно быть как можно меньше ограничения на время нахождения заявки в очереди. Также для обеспечения достоверности полученных результатов $\overline{t_{ож}}$ должно быть больше $\delta_{\overline{t_{ож}}}$.

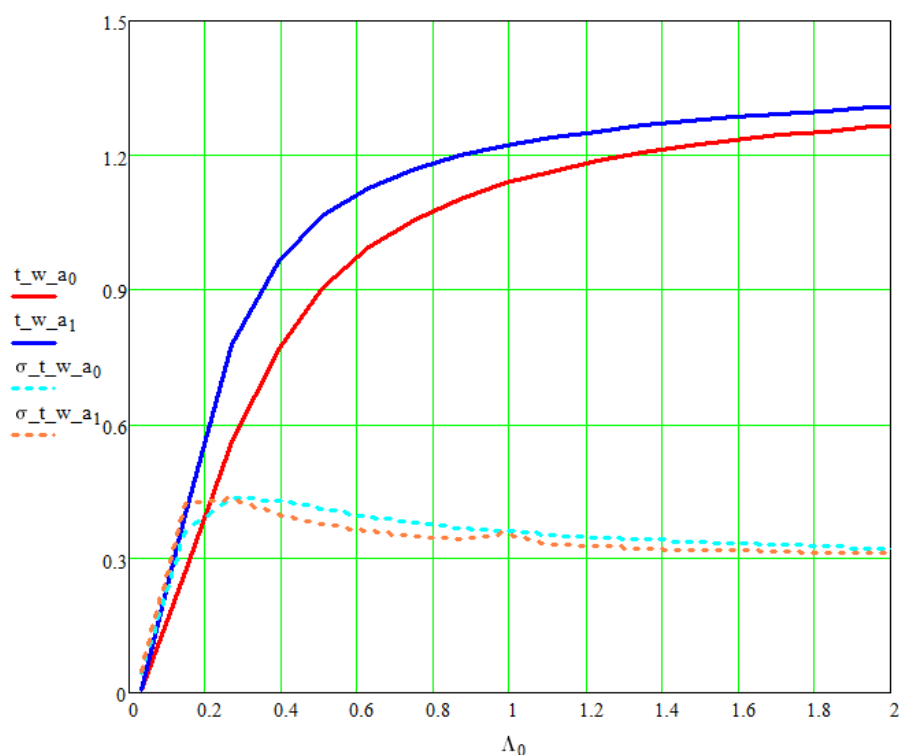


Рис. 4.25 – Зависимость $\overline{t_{ож}}$ и $\delta_{\overline{t_{ож}}}$ от интенсивности

входящего потока заявок Λ_0 :

- красная кривая ($t_w_a_0$) соответствует трехкомпонентному потоку;
- синяя кривая ($t_w_a_1$) соответствует двухкомпонентному потоку.

Рассмотрим полученные графики на рис. 4.25. Для данного примера, как и для случаев, рассмотренных в подпункте 3.2, время ожидания с индексом 1 получено в результате наличия во входящем потоке заявок всех трех компонент входящего потока, при чем $\lambda_0=\lambda_1=\lambda$, а время ожидания с индексом 0 получено для двух компонент $\lambda_1=\lambda$, при этом $\lambda_0=0$. Суммарная интенсивность обоих потоков равнозначна. Во втором случае интенсивность заявок 0-го типа распределена в равных долях на заявки 1-го и 2-го типа.

По приведённым на рис. 4.25 зависимостям можно сделать вывод, что кратное увеличение интенсивности входного потока относительно исходных данных при неизменной скорости обслуживания не ведет к кратному увеличению времени ожидания, т.к. заявки начинают покидать очередь необслуженными ввиду ограничения на время ожидания.

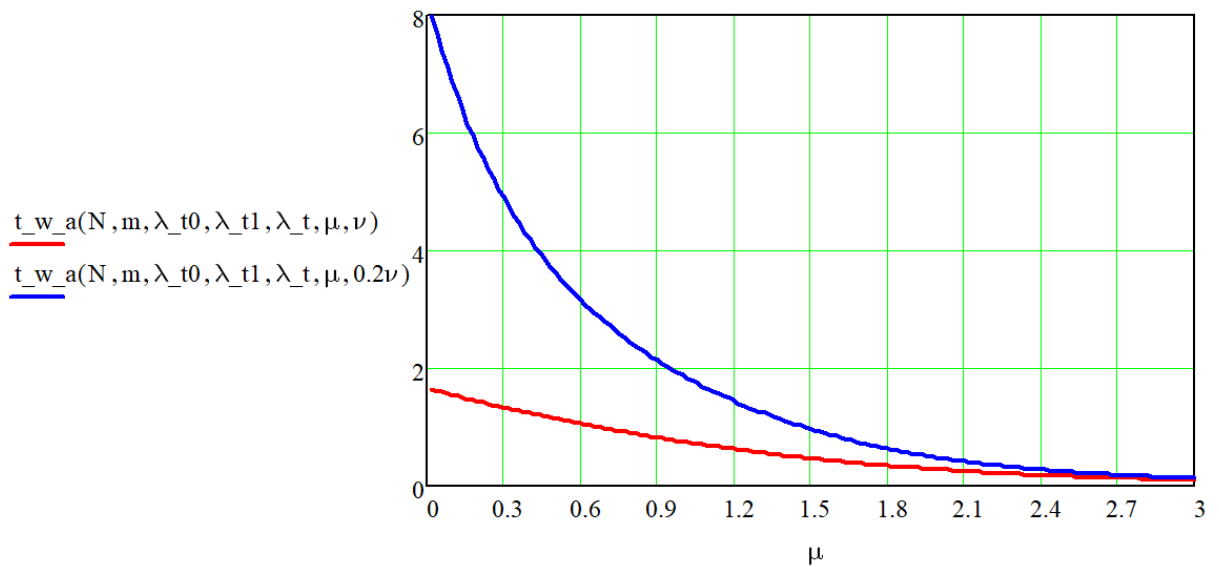


Рис. 4.26 – Зависимость $\overline{t_{ож}}$ от μ

Для кривых, приведенных на графике рис. 4.26, очевидно, что рост интенсивности обслуживания влечет за собой экспоненциальный спад кривой $\overline{t_{ож}}$. Величина значения среднего времени ожидания $\overline{t_{ож}}$ становится меньше на всем промежутке значений μ ввиду повышения производительности системы.

При снижении интенсивности уходов с ν до $0,2 \cdot \nu$ (синий график) у ожидающих заявок становится меньше возможностей покинуть систему необслуженными, в связи с чем среднее время ожидания $\overline{t_{ож}}$ увеличивается.

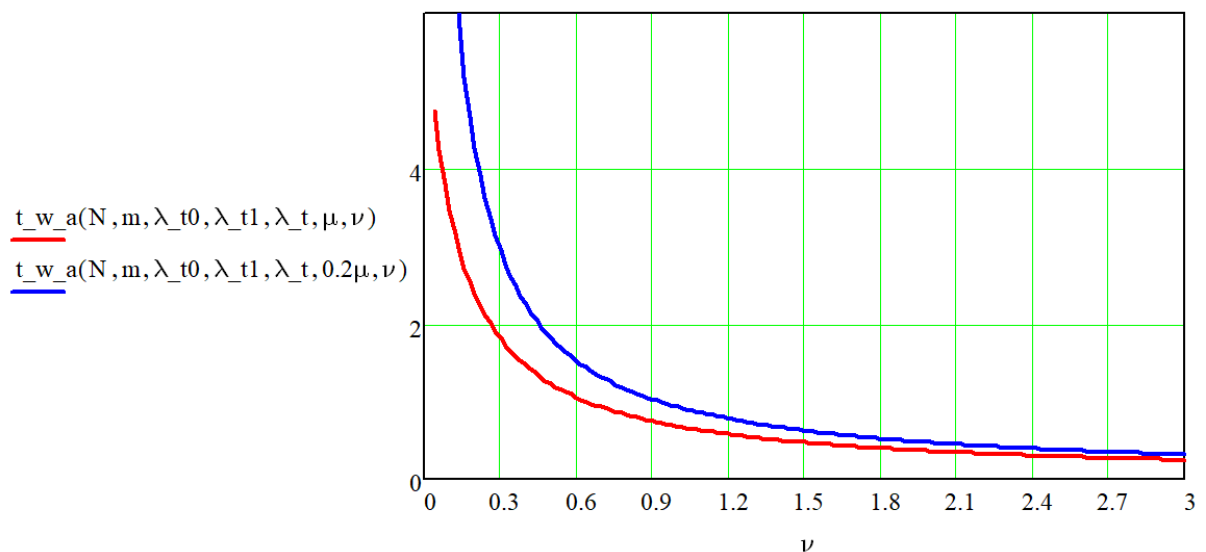


Рис. 4.27 – Зависимость $\overline{t_{ож}}$ от ν

Как можно заметить на графике рис. 4.27, при равнозначном изменении параметров μ и ν , изменение параметра ν , вносит большее влияние на величину среднего времени ожидания заявки в очереди. С падением интенсивности обслуживания в 5 раз, $\overline{t_{ож}}$ увеличивается в среднем в 1,5 раза, в то время как уменьшение ν в 5 раз, ведет к увеличению $\overline{t_{ож}}$ в среднем в 4 раза. Таким образом, полученные реакции параметра $\overline{t_{ож}}$ на изменение параметров μ и ν схожи с результатами, полученными в разделе 3.2 для параметра \bar{l} .

Выводы:

1. В данной главе рассчитаны основные характеристики замкнутой СМО с трехкомпонентным потоком и ограничением по времени ожидания заявки в очереди в стационарном и нестационарном режимах работы.
2. Для верификации модели использован параметр коэффициента вариации, показывающий степень отклонения величины от ее математического ожидания.
3. В ходе ряда вычислительных экспериментов выявлено, что наличие в потоке «нетерпеливых» заявок оказывает комплексное влияние на характеристики СМО. В частности: снижает общее количество заявок в системе, уменьшая среднюю длину очереди и среднее время ожидания, а также понижает пропускную способность СМО.
4. Ввиду замкнутости системы, заявки в системе накапливаются нелинейно.
5. Ограничение на среднее время ожидания для заявок 1-го типа вносит влияние на среднюю длину очереди и среднее время ожидания, сопоставимое со скоростью обслуживания.
6. Вид полученных зависимостей для средней длины очереди и среднего времени ожидания от изменения параметров системы демонстрирует схожие тенденции.

5 ПОИСК ЭФФЕКТИВНЫХ РЕЖИМОВ РАБОТЫ ЗАМКНУТОЙ СМО С ОГРАНИЧЕННЫМ ВРЕМЕНЕМ ОЖИДАНИЯ И ПОЛИКОМПОНЕНТНЫМ ПОТОКОМ ЗАЯВОК

5.1 Постановка задачи

В современных условиях функционирования СМО особое внимание уделяется поиску и анализу эффективных режимов работы, что обусловлено необходимостью обеспечения высокой эффективности и надежности данных систем. Замкнутая СМО представляет собой сложную структуру, в которой ограниченное количество заявок циркулирует между обслуживающими устройствами, создавая уникальные и малоизученные условия для анализа потоков и взаимодействий [83].

Основной задачей данной главы является формулирование условий и критериев, при которых система будет функционировать в эффективном режиме. В качестве основных критериев, характеризующих эффективность работы системы, будут использоваться: относительная пропускная способность (q) и среднее время ожидания заявки в очереди ($\overline{t_{ож}}$). Основные исходные данные, используемые для анализа и моделирования работы СМО, включают количество обслуживающих устройств (m), интенсивность обслуживания (μ), интенсивность ухода заявок ввиду ограничения на время ожидания (ν), интенсивность входящего потока при наличии заявок 0-го типа (A_0), интенсивность входящего потока при отсутствии заявок 0-го типа (A_1).

Примерные виды зависимостей данных величин рассмотрены в главе 3. По полученным графикам можно сделать вывод о том, что существует некое единственное численное значение для группы из двух исходных параметров СМО (m, μ, ν, A_0, A_1), соответствующее требуемому значению выбранного критерия эффективности. Для идентификации эффективных режимов предложен подход, основанный на варьировании пар ключевых параметров и оценке их влияния на критерии эффективности q и $\overline{t_{ож}}$.

Варьирование одновременно несколькими значениями исходных данных, позволяет получить соответствующую матрицу значений критериев эффективности. На основе полученных данных разработаны алгоритмы, позволяющие находить требуемые сочетания параметров системы для достижения эффективных режимов функционирования. Для определения вида зависимостей и оценки реакций системы на изменение исходных данных использованы результаты численных экспериментов, проведенных в главе 2,

Параметры среднего времени ожидания заявки в очереди $\overline{t_{ож}}$ и среднего числа заявок в очереди \bar{l} тесно связаны между собой и имеют схожую реакцию на изменение параметров системы, что подробно рассмотрено в главе 3. В связи с чем, дополнительное проведение численного исследования для поиска эффективных режимов работы по параметру \bar{l} в рамках данной работы не проводится.

5.2 Исследование эффективных режимов СМО по параметру относительной пропускной способности q при изменении параметров

$$\mu \text{ и } \lambda_{0,1}$$

Для решения задачи поиска эффективных режимов работы при изменении параметров λ_0 , λ_1 и μ проведено систематическое исследование влияния этих параметров на относительную пропускную способность системы q . Целью анализа является определение значений λ_0 , λ_1 и μ , при которых критерий эффективности $q=70\%$, и выявление критических точек, в которых система переходит в неэффективный режим.

Для этого, используя программную среду Mathcad, первоначально создан алгоритм поиска матрицы значений q при изменении параметров λ_0 и μ использующий метод вложенного перебора с равномерной дискретизацией. Блок-схема алгоритма представлена на рис. 5.1.

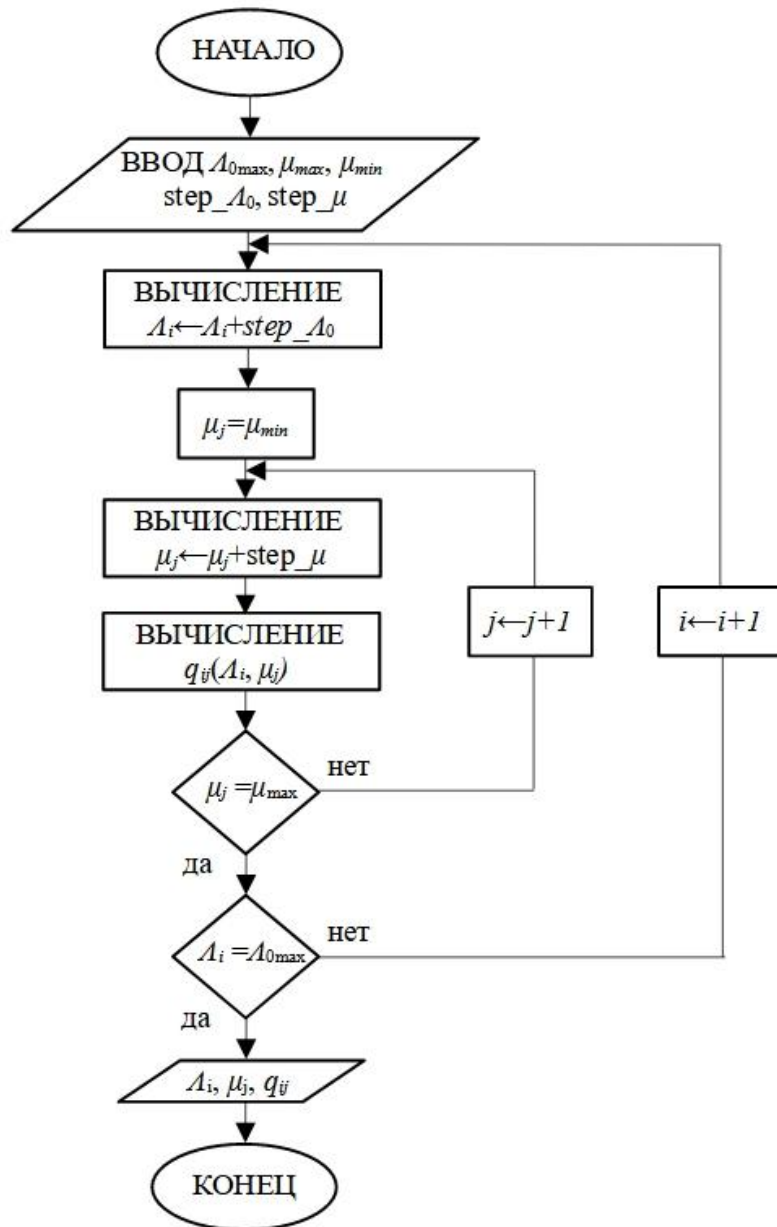


Рис. 5.1 – Алгоритм поиска матрицы значений q при изменении λ_0, μ

Параметризация модели:

- диапазон варьирования μ : от 0,05 до 3 с шагом 0,05;
- диапазон варьирования λ_0 : от 0,15 до 3 с шагом 0,15;
- параметры N и ν фиксированы на основе исходных данных, указанных в главе 3;
- количество доступных станций обслуживания $m=10$ для обеспечения наибольшей наглядности реакций системы;

– относительная пропускная способность q рассчитывается по формуле (2.2).

В соответствии с программной реализацией алгоритма на рис. 5.2, результаты расчетов сохраняются в матрице значений относительной пропускной способности q , строки которой соответствуют значениям Λ_0 , а столбцы – значениям μ .

```

 $\Lambda_0 :=$  |  $\lambda_{t1} \leftarrow 0$ 
          |  $\lambda_t \leftarrow 0$ 
          |  $\lambda_{t0} \leftarrow 0$ 
          |  $\mu \leftarrow 0$ 
          | for  $i_ \in 1.. \left( \frac{\lambda_{t0\_max} \cdot 3}{step} \right) + 1$ 
          |   if  $i_ \neq 1$ 
          |   |  $\lambda_{t1} \leftarrow \lambda_{t1} + step$ 
          |   |  $\lambda_t \leftarrow \lambda_t + step$ 
          |   |  $\lambda_{t0} \leftarrow \lambda_{t0} + step$ 
          |   |  $\mu \leftarrow 0$ 
          |   for  $j_ \in 1.. \left( \frac{\mu_{max}}{step} \right) + 1$ 
          |   |    $q_{i_,j_} \leftarrow \lambda_{t0} \cdot 3$  if  $j_ = 1$ 
          |   |   if  $j_ \neq 1$ 
          |   |   |  $\mu \leftarrow \mu + step$ 
          |   |   |  $q_{i_,j_} \leftarrow \mu$  if  $i_ = 1$ 
          |   |   | if  $i_ \neq 1 \wedge j_ \neq 1$ 
          |   |   |   |  $x \leftarrow q(N, m, \lambda_{t0}, \lambda_{t1}, \lambda_t, \mu, \nu)$ 
          |   |   |   |  $q_{i_,j_} \leftarrow x$ 
          |   |    $q_{1,1} \leftarrow \Lambda_0 / \mu$ 
          |    $q$ 

```

Рис. 5.2 – Программная реализация алгоритма поиска матрицы значений $q(\Lambda_0, \mu)$

Для углубленного анализа реакций системы рассмотрим два сценария. Переменная Λ_0 в данном примере состоит из потока заявок 0-го,

1-го и 2-го типа равной интенсивности. Численные значения максимальной интенсивности 0-й компоненты потока (λ_{t0_max}) и шага (step) могут быть изменены в зависимости от требуемого пространства поиска решений.

Некоторые из полученных значений приведены в таблице 5.1

Таблица 5.1 – Матрица значений $q(\lambda_0, \mu)$

$\lambda_0 \backslash \mu$	0.05	0.1	0.15	0.20	0.25	0.3
0.15	0.186	0.356	0.506	0.634	0.737	0.816
0.3	0.107	0.205	0.296	0.381	0.459	0.53
0.45	0.08	0.154	0.222	0.286	0.345	0.401
0.6	0.067	0.128	0.185	0.238	0.287	0.334
0.75	0.059	0.113	0.163	0.21	0.253	0.293
0.9	0.053	0.103	0.148	0.19	0.23	0.267

Исключим из суммарного потока 0-й тип заявок, перераспределив их между заявками 1-го и 2-го типа. В таком случае поток λ_1 будет состоять из потока заявок 1-го и 2-го типа равной интенсивности. В результате получим новую матрицу значений q . Модификация рассмотренного ранее алгоритма включает:

1. Обнуление λ_0 в инициализации переменных.
2. Корректировку диапазонов λ_1 и μ для сохранения сопоставимости с предыдущими результатами.

Результаты расчётов для двухкомпонентного потока представлены в Таблице 5.2.

Легко заметить, что в случае отсутствия «нетерпеливых» заявок значения относительной пропускной способности выше при той же величине входного потока и скорости обслуживания.

Таблица 5.2 – Матрица значений q (Λ_1, μ)

$\mu \backslash \Lambda_1$	0.05	0.1	0.15	0.2	0.25	0.3
0.15	0.2	0.382	0.544	0.68	0.787	0.865
0.3	0.12	0.231	0.333	0.428	0.515	0.594
0.45	0.093	0.179	0.259	0.333	0.402	0.466
0.6	0.08	0.154	0.222	0.286	0.345	0.4
0.75	0.072	0.138	0.2	0.257	0.31	0.36
0.9	0.067	0.128	0.185	0.238	0.287	0.333

Для программной реализации в данном случае значению λ_{t0} будет присвоено значение 0, а переменная интенсивности 1-й компоненты потока (λ_{t1}) станет мерой суммарного потока. Шаг (step) для равнозначного сопоставления результатов таблиц 5.1 и 5.2 будет равен 0,075.

```

 $\Lambda_1 :=$ 
|  $q \leftarrow 0$ 
|  $\lambda_{t1} \leftarrow 0$ 
|  $\lambda_t \leftarrow 0$ 
|  $\lambda_{t0} \leftarrow 0$ 
|  $\mu \leftarrow 0$ 
| for  $i \in 1.. \left( \frac{\lambda_{t1\_max}}{step} \right) + 1$ 
|   | if  $i \neq 1$ 
|   |   |  $\lambda_{t1} \leftarrow \lambda_{t1} + step$ 
|   |   |  $\lambda_t \leftarrow \lambda_t + step$ 
|   |   |  $\lambda_{t0} \leftarrow 0$ 
|   |   |  $\mu \leftarrow 0$ 
|   | ...
|   |  $q_{1,1} \leftarrow \Lambda_1 / \mu$ 
|  $q$ 

```

Рис. 5.3 – Программная реализация алгоритма поиска матрицы значений q при изменении параметров Λ_1, μ (внешний цикл)

```

for j_ ∈ 1.. ⌈  $\frac{\mu_{\max}}{\text{step}}$  ⌉ + 1
    qi_,j_ ← λt1·2 if j_ = 1
    if j_ ≠ 1
        μ ← μ + step
        qi_,j_ ← μ if i_ = 1
        if i_ ≠ 1 ∧ j_ ≠ 1
            x ← q(N, m, λt0, λt1, λt, μ, ν)
            qi_,j_ ← x

```

Рис. 5.4 – Программная реализация алгоритма поиска матрицы значений q при изменении параметров A_1, μ (внутренний цикл)

Зададимся требуемым минимальным значением относительной пропускной способности системы $q_n=70\%$. На основании вышеизложенного задача состоит в поиске сочетаний числовых значений исходных данных (A_0, A_1 и μ), которые обеспечивают выполнение поставленного условия:

$$q(\mu_i, A_i) = 0,7.$$

Для решения поставленной задачи в работе применен метод Ньютона с адаптивной релаксацией. Выбор данного подхода обусловлен высокой скоростью сходимости и возможностью поиска решения даже в условиях слабой зависимости $q(\mu)$ (например, при $q'(\mu) \rightarrow 0$). Для реализации метода Ньютона использована формула:

$$\mu_{n+1} = \mu_n - \alpha \cdot \frac{f(\mu_n)}{f'(\mu_n)}, \quad (5.1)$$

где $f(\mu_n)$ – вспомогательная функция, $f(\mu_n) = q(\mu_n, A_i) - 0,7$, α – коэффициент релаксации, регулирующий длину шага, $(0, 1]$.

Алгоритм реализации:

1. Задается начальное значение μ_0 близкое к ожидаемому решению (на основе предыдущих вычислений или аналитических оценок).
2. На каждой итерации (внутренний цикл):

2.1 Вычисляется начальное расхождение $f(\mu_0) = q(\mu_0, A_i) - 0,7$ и численное значение производной $f'(\mu_0)$.

2.2 Текущее значение μ_0 обновляется на μ_1 , вычисленное по формуле (21).

2.2.1 В случае, если $\mu_1 < 0$, коэффициенту релаксации α присваивается значение $\alpha/2$ для повторного вычисления значения μ_1 .

2.2.2 Если после очередного вычисления α в соответствии с 2.2.1, значение $\alpha < 10^{-12}$ и $\mu_1 < 0$, регистрируется отсутствие сходимости метода Ньютона для данного начального приближения μ_0 и алгоритм завершает работу.

2.3 Вычисляется текущее расхождение $f(\mu_1) = q(\mu_1, A_i) - 0,7$.

2.4 Значению μ_0 присваивается текущее значение μ_1 и цикл повторяется.

2.5 При достижении заданной точности $|f(\mu_n)| < \delta$ алгоритм завершает работу.

3. Если ошибка между двумя последовательными операциями $\Delta = f(\mu_{n+1}, A_i) - f(\mu_n, A_i)$ увеличивается, значение коэффициента релаксации α уменьшается вдвое. Если ошибка между двумя последовательными операциями уменьшается, α присваивается значение 1.

4. Для поиска эффективных режимов работы при различных значениях A_0 и A_1 используется внешний цикл с итерационным перебором значений A_0 и A_1 от 0 до i и фиксированным шагом дискретизации.

Параметризация модели:

- начальное приближение $\mu_0 = 0,5$;
- диапазон варьирования A_0 : от 0,15 до 3 с шагом 0,15;
- диапазон варьирования A_1 : от 0,15 до 3 с шагом 0,15;
- $q_n = 0,7$;
- $\delta = 0,01$;

– параметры N и ν фиксированы на основе исходных данных, указанных в главе 3.

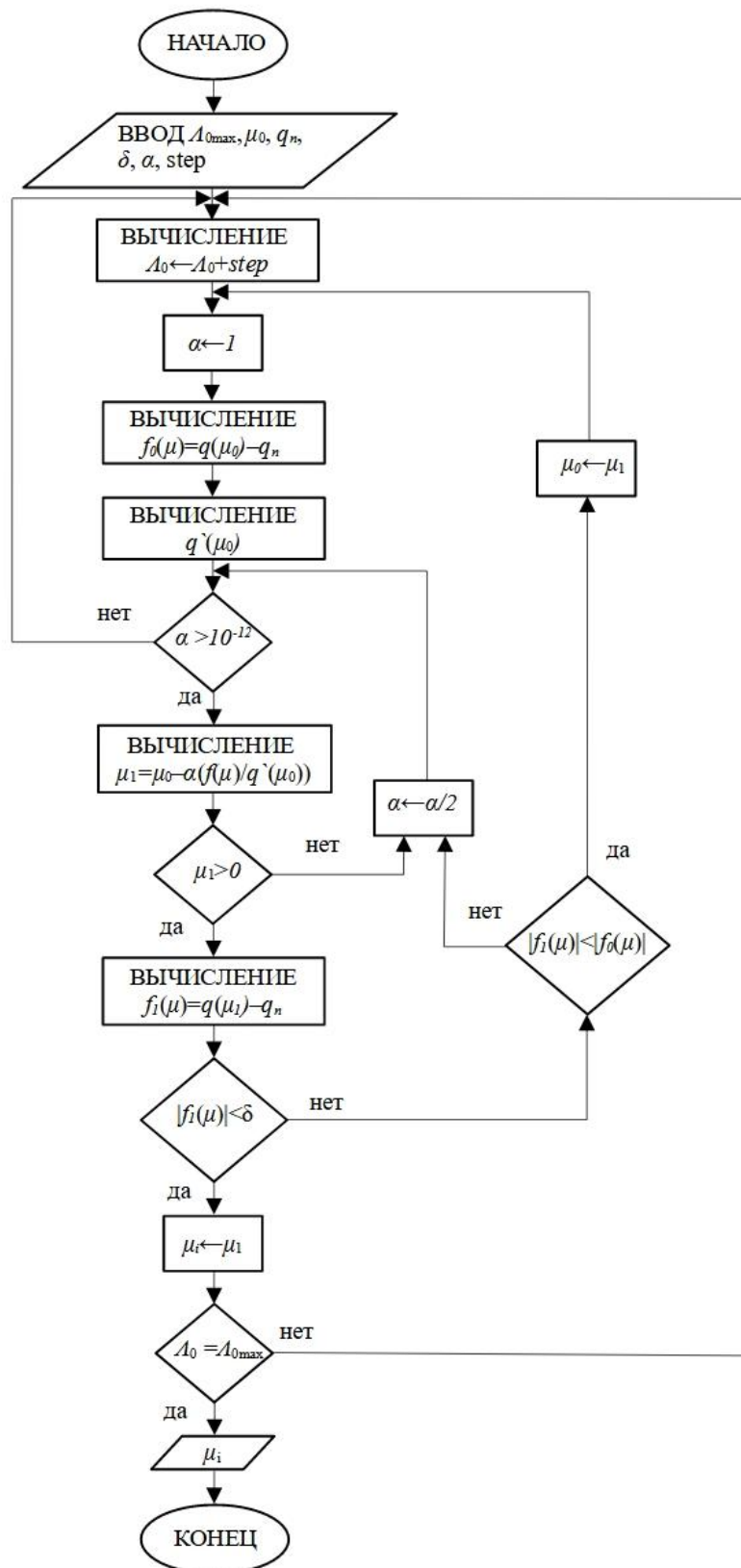


Рис. 5.5 – Алгоритм поиска значений μ и L_0 для $q=70\%$

```

...
for i_ ∈ 0..  $\left(\frac{\lambda\_t0\_max}{step}\right)$ 
    |  $\lambda\_t1 \leftarrow \lambda\_t1 + step$ 
    |  $\lambda\_t \leftarrow \lambda\_t + step$ 
    |  $\lambda\_t0 \leftarrow \lambda\_t0 + step$ 
    |  $f1 \leftarrow 5$ 
...
| Y

```

Рис. 5.6 – Программная реализация алгоритма поиска μ и λ_0 для $q=70\%$
(внешний цикл)

```

while |f1| >  $\delta$ 
    |  $\mu \leftarrow \mu_0$ 
    |  $f0 \leftarrow q(N, m, \lambda\_t0, \lambda\_t1, \lambda\_t, \mu, \nu) - qn$ 
    |  $q' \leftarrow \frac{d}{d\mu} q(N, m, \lambda\_t0, \lambda\_t1, \lambda\_t, \mu, \nu)$ 
    |  $\mu1 \leftarrow \mu_0 - \alpha \cdot \left(\frac{f0}{q'}\right)$ 
    | while  $\mu1 < 0 \vee \alpha > 10^{-12}$ 
    |     |  $\alpha \leftarrow \frac{\alpha}{2}$ 
    |     |  $\mu1 \leftarrow \mu_0 - \alpha \cdot \left(\frac{f0}{q'}\right)$ 
    |     | break if  $\mu1 > 0$ 
    | if  $\alpha \leq 10^{-12}$ 
    |     | Y ← "Method Error"
    |     | break
    |  $\mu \leftarrow \mu1$ 
    |  $f1 \leftarrow q(N, m, \lambda\_t0, \lambda\_t1, \lambda\_t, \mu, \nu) - qn$ 
    |  $\mu_0 \leftarrow \mu1$  if  $|f1| < |f0|$ 
    |  $\alpha \leftarrow 1$ 
    |  $\alpha \leftarrow \frac{\alpha}{2}$  if  $|f1| \geq |f0|$ 
    |  $Y_{0,0} \leftarrow \mu$ 
    |  $Y_{0,1} \leftarrow \lambda\_t0 \cdot 3$ 
    |  $Y_{0,2} \leftarrow f1 + qn$ 

```

Рис. 5.7 – Программная реализация алгоритма поиска μ и λ_0 для $q=70\%$
(внутренний цикл)

По аналогии с предыдущим примером рассмотрен случай отсутствия в потоке нулевой компоненты. Некоторые из результатов расчетов сведены в таблицы 5.3 и 5.4.

Таблица 5.3 – Матрица значений Λ_0 , μ и соответствующих им q

Λ_0	μ	q
0.15	0.227	0.692
0.3	0.438	0.695
0.45	0.693	0.694
0.6	0.831	0.692
0.75	1.028	0.694
0.9	1.208	0.691
1.05	1.398	0.692
1.2	1.589	0.693
1.35	1.777	0.694
1.5	1.965	0.695

Таблица 5.4 – Матрица значений Λ_1 , μ и соответствующих им q

Λ_1	μ	q
0.15	0.211	0.706
0.3	0.369	0.691
0.45	0.516	0.695
0.6	0.641	0.693
0.75	0.749	0.69
0.9	0.854	0.692
1.05	0.95	0.694
1.2	1.024	0.69
1.35	1.101	0.691
1.5	1.174	0.692

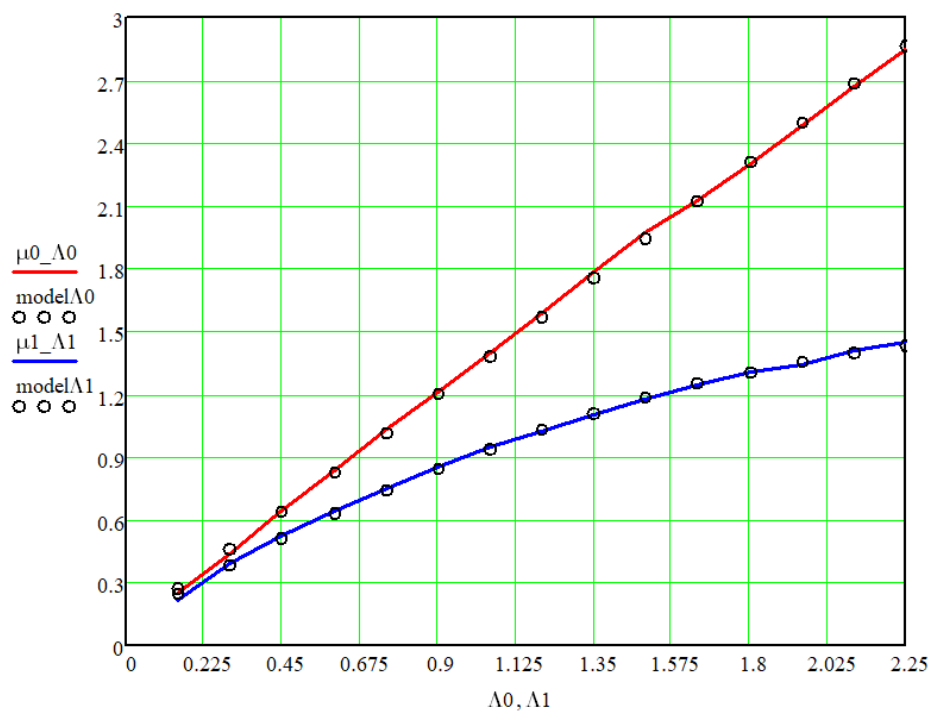


Рис. 5.8 – Изолинии $q(\mu, \Lambda_0)=70\%$ и $q(\mu, \Lambda_1)=70\%$

На рис. 5.8 представлены изолинии $q(\mu, \Lambda_0)=70\%$, $q(\mu, \Lambda_1)=70\%$, отображающие минимальные необходимые для выполнения условия $q=70\%$ комбинации параметров μ и Λ_0, Λ_1 . Рост параметра интенсивности входящего потока отражает необходимость компенсации перегрузки за счет увеличения скорости обслуживания:

- кривая $\mu_0_{\Lambda_0}$ соответствует изолинии $q(\mu, \Lambda_0)=70\%$ для трехкомпонентного потока (0-й, 1-й, 2-й типы заявок);
- кривая $\mu_1_{\Lambda_1}$ соответствует изолинии $q(\mu, \Lambda_1)=70\%$ для двухкомпонентного потока (1-й, 2-й типы заявок после исключения 0-го типа).

Для математического обоснования вида полученных зависимостей был проведен комплексный регрессионный анализ с использованием методов математической статистики. Анализ выполнен с использованием вычислительной платформы Wolfram Engine, что обеспечило высокую точность расчетов и строгость статистических оценок. При определении вида функциональных зависимостей рассматривались несколько математических

моделей, включая линейную, квадратичную, степенную, логарифмическую, экспоненциальную и обратную зависимости. Для оценки полученных функциональных зависимостей использовались:

- коэффициент детерминации R^2 для оценки качества аппроксимации экспериментальных данных;
- информационные критерии Акаике (AIC) и Байеса (BIC) для оценки качества модели и ее сложности
- среднеквадратическая ошибка RMSE для оценки точности прогнозирования;
- нормальность остатков (критерий Шапиро-Уилка «W») для проверки на нормальность распределения остатков.

Для двухкомпонентного потока анализ показал статистически значимое превосходство степенной модели:

$$\mu_1 = 0,8919 \cdot A_1^{0,6328},$$

которая продемонстрировала коэффициент детерминации $R^2=0,9666$ при минимальных значениях критериев $AIC=-62,98$ и $BIC=-61,56$, $RMSE=0,026$ и нормальность остатков $W=0,704$.

Результат расчета для трехкомпонентного потока показал адекватность применения линейной модели вида:

$$\mu_0 = 0,086638 + 1,23569 \cdot A_0.$$

Показатели модели: $R^2=0,996$, $AIC=-77,2$, $BIC=-75,78$, $RMSE=0,0162$, $W=0,754$.

Для двухкомпонентного потока при малых значениях A_1 ($<0,5$) интенсивность обслуживания μ растёт медленно, что связано с избытком свободных ресурсов системы. Эту область можно назвать областью малых нагрузок:

- при $A_1 = 0,25 \rightarrow \mu = 0,4$;
- при $A_1 = 0,5 \rightarrow \mu = 0,6$.

В диапазоне умеренных нагрузок $A_1 \in [0,5; 1,5]$ рост параметра μ становится более плавным, но остается близким к линейному. Это

обусловлено эффективным распределением заявок между обслуживающими устройствами.

- При $\Lambda_1 = 1,0 \rightarrow \mu = 1,0$;
- При $\Lambda_1 = 1,5 \rightarrow \mu = 1,205$.

В диапазоне высоких нагрузок $\Lambda_1 > 1,5$ замедление роста значения μ объясняется ограниченностью числа заявок в замкнутой СМО ($N=30$) и нелинейным накоплением заявок в системе, что отражено на рис. 5.9.

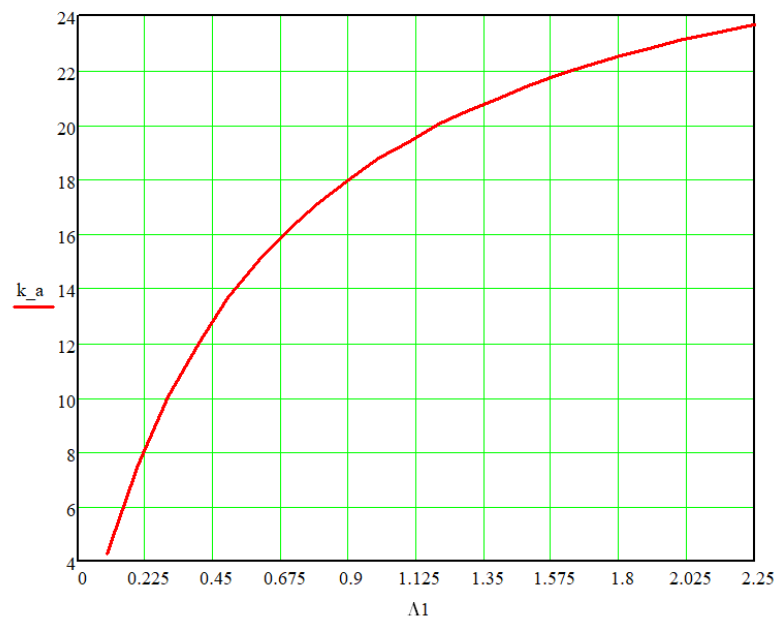


Рис. 5.9 – Зависимость \bar{k} от Λ_1

Таким образом, при $\Lambda_1 > 1,5$ источник заявок истощается, в результате чего снижаются требования к скорости обслуживания. В системе с 0-м типом заявок замедление роста μ отсутствует, что объясняется поддержанием баланса заявок, т.к. 0-й тип компенсирует потери, предотвращая истощение источника.

Для трёхкомпонентного потока (кривая $\mu_0_ \Lambda_0$) требуемые значения μ выше, чем для двухкомпонентного (кривая $\mu_1_ \Lambda_1$), а зависимость изменяется со степенной на линейную. Разница обусловлена наличием дополнительного потока заявок 0-го типа, создающего конкуренцию за ресурсы обслуживания. Исключение 0-го типа заявок снижает требуемую интенсивность обслуживания на 25–30% при одинаковой нагрузке ($\Lambda_0 = \Lambda_1$). Область между

кривыми соответствует сочетаниям параметров, при которых выполняется условие эффективности работы системы ($q=70\%$) как для трёхкомпонентного, так и для двухкомпонентного потоков. В этой области, очевидно, находятся значения $\lambda_0 \in [0; 0,25]$, удовлетворяющие заданному критерию эффективности. За пределами этой зоны (ниже $\mu_1 \Lambda_1$) условие $q=70\%$ нарушается. Таким образом, для трехкомпонентного потока при $\Lambda_0=1,5$ необходимое $\mu \approx 2,0$. Для двухкомпонентного потока при $\Lambda_1=1,5$ достаточно $\mu \approx 1,2$. Снижения значения параметра λ_0 расширяет зону эффективности работы системы.

На рис. 5.9 представлены изолинии $q(\mu, \Lambda_0)=70\%$ для разного количества обслуживающих устройств m . Кривые показывают, как увеличение числа устройств влияет на требования к скорости обслуживания и допустимую величину входного потока.

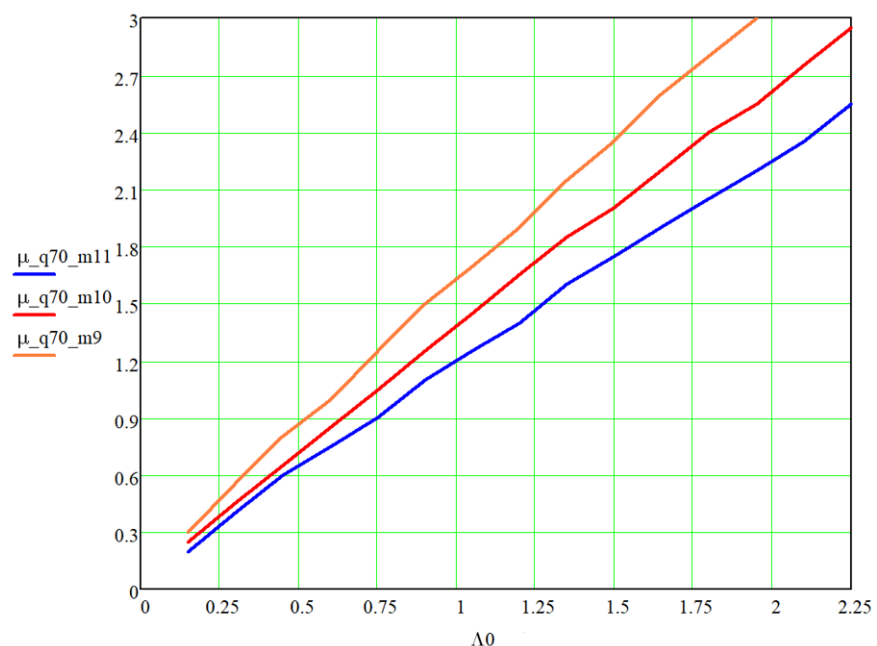


Рис. 5.10 – Изолинии $q(\mu, \Lambda_0)=70\%$ с учетом изменения m

Увеличение числа обслуживающих устройств позволяет распределить нагрузку более равномерно, что снижает требуемую интенсивность обслуживания, необходимую для сохранения эффективности работы системы. При интенсивности входящего потока $\Lambda_0 > 1,5$ влияние увеличения величины m становится особенно заметным. Например, для $\Lambda_0=1,75$:

- при $m=9 \rightarrow \mu \approx 2.7$;
- при $m=11 \rightarrow \mu \approx 2.0$

Таким образом, добавление двух обслуживающих устройств снижает требования к μ на 26 %. Это связано с тем, что каждое новое устройство берёт на себя часть нагрузки, предотвращая переполнение очереди. Это демонстрирует, что масштабирование позволяет системе адаптироваться к пиковым нагрузкам без критического роста затрат. Добавление даже одного устройства уменьшает μ на 15–20 %. На практике такая ситуация обоснована, т.к. зачастую экономически выгодней дополнить систему станцией обслуживания, нежели наращивать скорость обслуживания.

При фиксированной $\mu=2,7$ анализ изолиний показывает, что увеличение числа серверов позволяет значительно повысить допускаемую интенсивность входного потока:

- при $m=9 \rightarrow A_0 \approx 1,5$;
- при $m=11 \rightarrow A_0 \approx 2,1$.

Добавление двух обслуживающих устройств увеличивает допустимую нагрузку на 40 %, т.к. дополнительные серверы распределяют заявки более равномерно, снижая среднюю длину очереди и повышая относительную пропускную способность.

5.3 Анализ влияния интенсивности ухода заявок (ν) на фиксированный параметр относительной пропускной способности q

После определения эффективных режимов работы системы при варьировании ключевых параметров μ и A_0 , A_1 , перейдём к исследованию роли интенсивности ухода заявок из очереди (ν) – статистического фактора, определяющего поведение заявок 1-го типа в очереди. В отличие от μ и A , которые могут быть скорректированы в ходе проектирования или эксплуатации системы, параметр ν задается внешними условиями функционирования и определяется статистикой ухода заявок из очереди по завершении граничного времени ожидания.

На рис. 5.10 представлена зависимость q от ν для эффективного режима работы системы на основе параметров из предыдущих исследований: $\mu=0,6$, $\Lambda_0=0,4$.

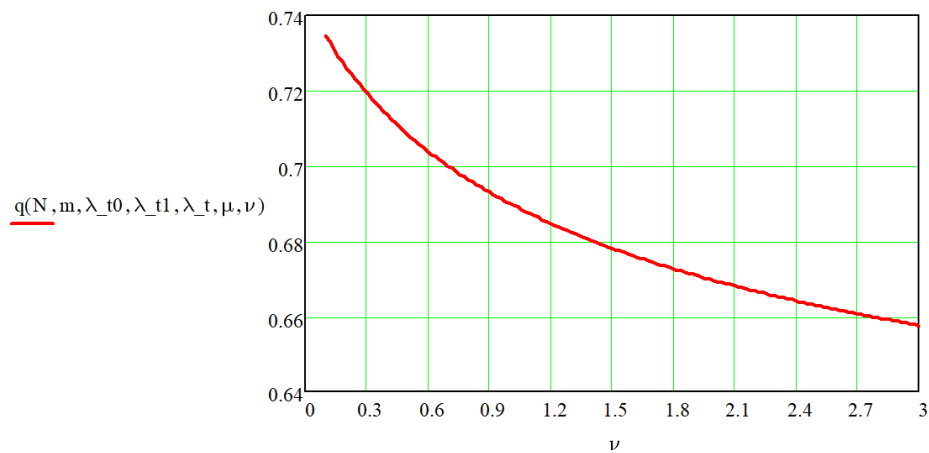


Рис. 5.11 – Зависимость $q(\nu)$ при $\mu=0,6$, $\Lambda_0=0,4$

Учитывая исходные данные, очевидно, что при $\nu \in [0,1; 0,6]$ получим диапазон значений параметра $q=70\%$. Дальнейший рост параметра ν снижает среднее время ожидания, но также и снижает относительную пропускную способность, т.к. потери в системе растут. Следует отметить, что при изменении параметра ν от 0,1 до 3,0 значение q колеблется в узком диапазоне (66 %–74 %).

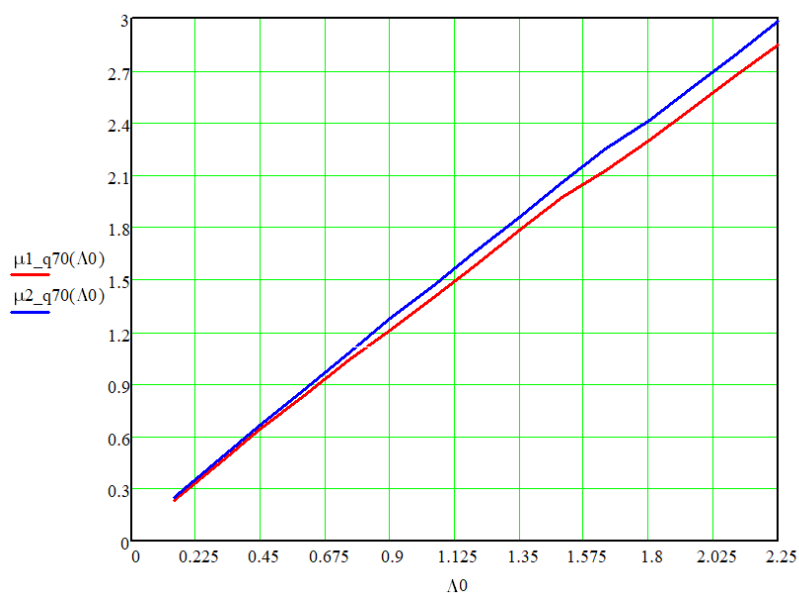


Рис. 5.12 – Зависимости $\mu(\Lambda_0)$ при $\nu=0,6$ и $\nu=1,2$:

– кривая $\mu_{0_q70}(\Lambda_0)$ – для трёхкомпонентного потока, $\nu=0,6$;

– кривая $\mu_1_q70(\Lambda_0)$ – для трёхкомпонентного потока, $\nu=1,2$.

Рост интенсивности ухода заявок ν с 0,6 до 1,2 приводит к незначительному, но систематическому росту требуемой интенсивности обслуживания μ (до 8 %). Этот эффект обусловлен необходимостью компенсации потерь заявок из-за ускоренного ухода из очереди и изменением баланса между временем ожидания и обслуживания.

Как уже отмечалось ранее, в трехкомпонентной системе заявки 0-го типа постоянно «подпитывают» входной поток. Таким образом, при наличии в потоке 0-й компоненты общее давление на систему сохраняется даже при учете роста числа вылетов, т.к. заявки 0-го типа все еще должны быть обслужены немедленно, следовательно размер очереди невелик. В результате вылетевшие из очереди заявки незначительно влияют на общую величину относительной пропускной способности q . Обратная ситуация представлена на графике, изображенном на рис. 5.13.

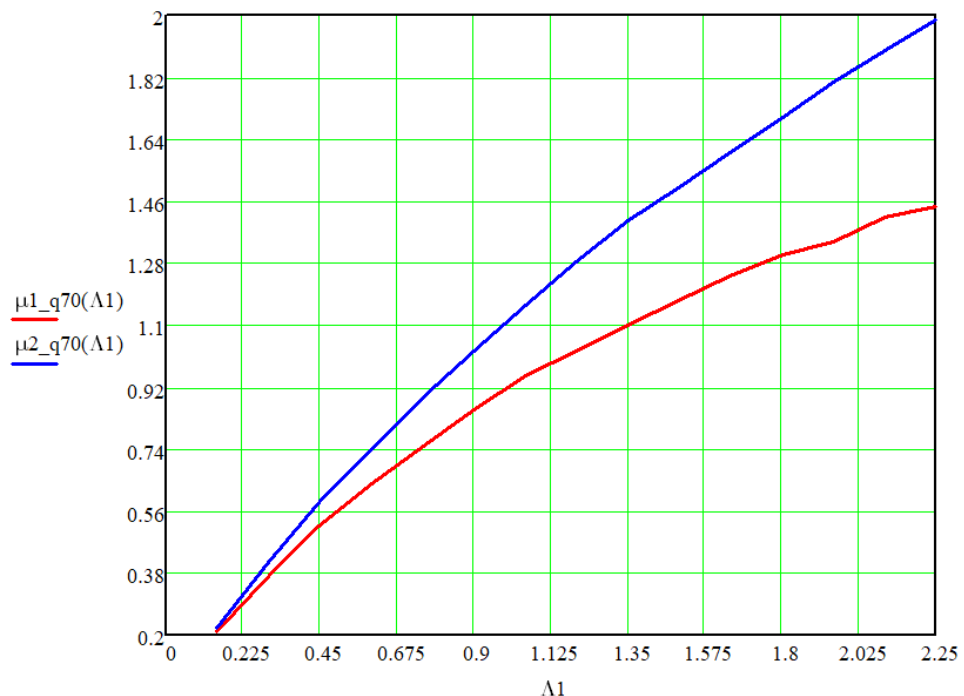


Рис. 5.13 – Зависимости $\mu(\Lambda_1)$ при $\nu=0,6$ и $\nu=1,2$:

- кривая $\mu_1_q70(\Lambda_1)$ – для двухкомпонентного потока, $\nu=0,6$;
- кривая $\mu_2_q70(\Lambda_1)$ – для двухкомпонентного потока, $\nu=1,2$.

Так как заявки 1-го и 2-го типов отказа в обслуживании не получают, при высоких значениях интенсивности λ_1 источник заявок истощается и все большее количество заявок находятся в очереди. Более подробно эта ситуация отображена на рис. 4.2, где при общем количестве заявок в источнике $N=30$ и $\lambda_1=2,25$ в системе может находиться 23 заявки, 13 из которых находится в очереди (при $m=10$). С ростом параметра ν все большее количество заявок покидает систему необслуженными, и т.к. 0-й компонент отсутствует, поступают на вход системы, значительно увеличивая нагрузку. Ввиду отсутствия отказов система вынуждена увеличивать скорость обслуживания для поддержания того же уровня пропускной способности. Увеличение интенсивности ухода заявок ν с 0,6 до 1,2 приводит к значительному росту требуемой интенсивности обслуживания μ (до 20 %). Полученные зависимости для $m=10$ подтверждают выводы главы 3, что наглядно продемонстрировано на графике, приведенном на рис. 5.14.

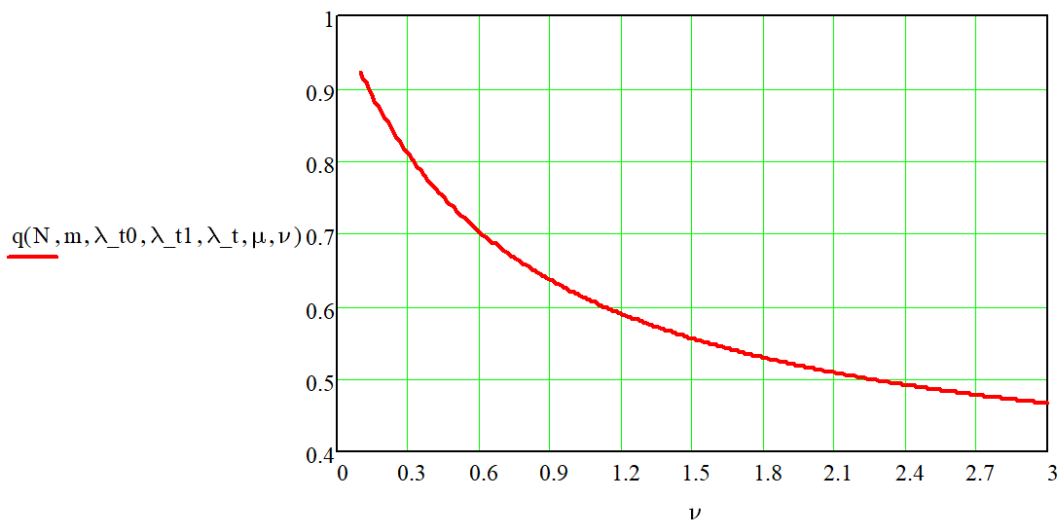


Рис. 5.14 – Зависимость $q(\nu)$ при $\mu=0,6$, $\lambda_1=0,5$

Таким образом, двухкомпонентная система более чувствительна к потерям. Ввиду «буфера» в виде заявок 0-го типа, чем выше параметр λ_0 в трехкомпонентном потоке, тем меньшее влияние ограничения времени ожидания оказывает на относительную пропускную способность,

Таким образом, рост параметра λ_0 в трехкомпонентной системе не только стабилизирует работу СМО, но и снижает зависимость относительной пропускной способности от ν , что делает ее более гибкой в условиях переменной нагрузки. При низких значениях λ_0 система, напротив, более чувствительна к потерям.

5.4 Исследование эффективных режимов СМО по параметру относительной пропускной способности при изменении параметров

μ и m

В данном подразделе рассматриваются минимальная интенсивность обслуживания (μ) и количество обслуживающих устройств (m) необходимые для обеспечения эффективной работы системы по критерию эффективности $q=70\%$ и реализующие условие:

$$q_{i,j}(m_i, \mu_j)=0,7.$$

Параметризация модели:

- диапазон варьирования m : от 5 до 20;
- начальное приближение $\mu_0=0,5$;
- $q_n=0,7$;
- $\delta=0,01$;
- параметры N , λ_0 , ν фиксированы на основе исходных данных, указанных в главе 3.

Блок-схема алгоритма нахождения минимального значения параметров, необходимых для обеспечения эффективной работы системы по критерию эффективности $q_n=70\%$ идентична с рассмотренной в разделе 4.2, за исключением замены параметра внешнего цикла с λ_0 на m . Алгоритм подразумевает циклический перебор значений m от минимального до максимального значения. Для каждого значения m выполняется внутренний цикл по μ . Для каждого m используется метода Ньютона с адаптивной релаксацией для поиска μ , также рассмотренный в разделе 4.2. Результатом

выполнения алгоритма являются сочетания минимальных значений μ и m , необходимых для обеспечения эффективного режима работы СМО с заданными параметрами.

Программная реализация в Mathcad выглядит следующим образом:

```

while |f1| > δ
    μ ← μ0
    f0 ← q(N, m, λ_t0, λ_t1, λ_t, μ, ν) - qn
    q' ←  $\frac{d}{d\mu} q(N, m, \lambda_{t0}, \lambda_{t1}, \lambda_t, \mu, \nu)$ 
    μ1 ← μ0 - α ·  $\left(\frac{f0}{q'}\right)$ 
    while μ1 < 0 ∨ α > 10-12
        α ←  $\frac{\alpha}{2}$ 
        μ1 ← μ0 - α ·  $\left(\frac{f0}{q'}\right)$ 
        break if μ1 > 0
    if α ≤ 10-12
        Y ← "Method Error"
        break
    μ ← μ1
    f1 ← q(N, m, λ_t0, λ_t1, λ_t, μ, ν) - qn
    μ0 ← μ1 if |f1| < |f0|
    α ← 1
    α ←  $\frac{\alpha}{2}$  if |f1| ≥ |f0|
Ym,0 ← μ

```

Рис. 5.15 – Программная реализация алгоритма поиска значений m и μ для $q=70\%$ (внутренний цикл)

На графике, изображенном на рис. 5.16 показана зависимость минимальной интенсивности обслуживания μ для трехкомпонентного потока требований, необходимая для достижения относительной пропускной способности $q=70\%$, при различных значениях количества обслуживающих устройств m .

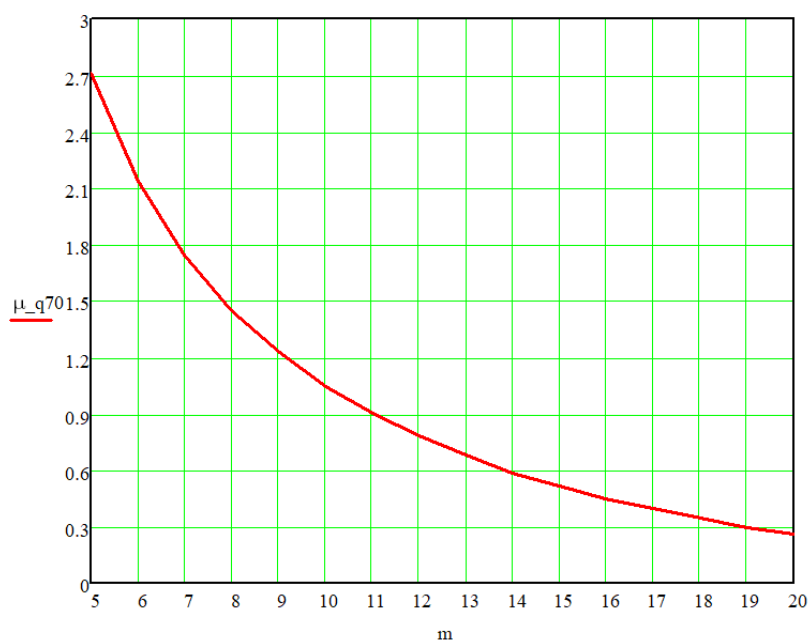


Рис. 5.16 – Зависимость $\mu(m)$ при $q=70\%$

По графику видно, что кривая имеет нелинейный характер, т.к. по мере роста количества обслуживающих устройств суммарная скорость обслуживания увеличивается.

При малых значениях m (<8) интенсивность обслуживания μ резко возрастает, что связано с тем, что система испытывает значительную нагрузку, что приводит к увеличению длины очереди и времени ожидания, что, в свою очередь, снижает относительную пропускную способность системы. При увеличении m с 5 до 6 требуемое значение μ снижается на 21 %.

По мере увеличения m кривая начинает плавно снижаться, что указывает на то, что с увеличением количества обслуживающих устройств нагрузка на каждое из них уменьшается, что позволяет системе более эффективно справляться с входящим потоком заявок. Например, при $m=9$ значение $\mu \approx 1,23$, а при $m=10$ значение μ снижается до $\mu \approx 1,05$ (на 15%). Это свидетельствует о том, что увеличение количества обслуживающих устройств существенно снижает требования к интенсивности обслуживания одним обслуживающим устройством для поддержания эффективного режима работы системы.

Когда m достигает значений выше 15, разница между соседними значениями μ составляет максимум $\pm 2\%$, что указывает на достижение предельной пропускной способности системы. Дальнейшее увеличение m уже не оказывает значительного влияния на значение μ , т.к. система становится достаточно ресурсоемкой для обработки входящего потока заявок даже при низкой интенсивности обслуживания. Следует отметить, что параметр μ является более гибким инструментом для тонкой настройки эффективности системы, тогда как m является инструментом грубой настройки при значительных изменениях нагрузки.

На следующем графике, изображенном на рис. 5.17, представлены зависимости минимальной интенсивности обслуживания μ и минимального количества обслуживающих устройств m , необходимых для обеспечения относительной пропускной способности $q=70\%$, при различных значениях интенсивности входящего потока заявок λ_0 .

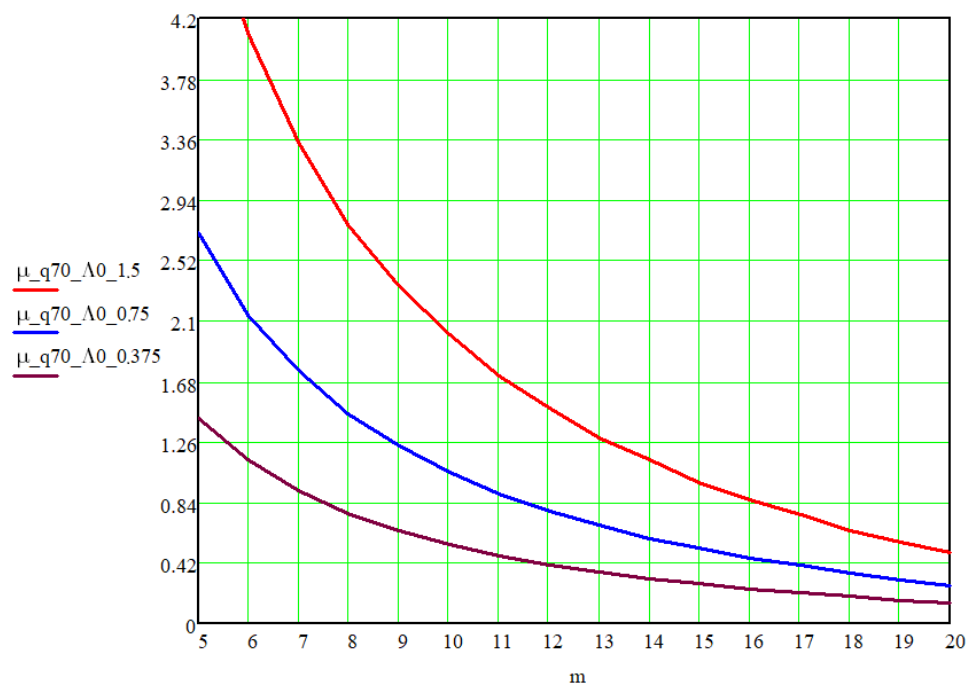


Рис. 5.17 – Зависимости $\mu(m)$ при $q=70\%$ для различных значений λ_0 :

- красная кривая ($\mu_q70_A0_1.5$) соответствует значению $\lambda_0 = 1.5$;
- синяя кривая ($\mu_q70_A0_0.75$) соответствует значению $\lambda_0 = 0.75$;
- фиолетовая кривая ($\mu_q70_A0_0.375$) соответствует значению $\lambda_0 = 0.375$.

Увеличение Λ_0 требует увеличения интенсивности обслуживания μ или числа обслуживаемых устройств для поддержания эффективного режима работы системы.

Например, для $m = 11$ и компенсации роста нагрузки за счет скорости обслуживания:

– при $\Lambda_0 = 0,75 \rightarrow \mu \approx 0,85$;

– при $\Lambda_0 = 1,5 \rightarrow \mu \approx 1,69$.

При компенсации за счет увеличения числа обслуживаемых устройств и $\mu=1,69$:

– при $\Lambda_0 = 0,75 \rightarrow m=7$;

– при $\Lambda_0 = 1,5 \rightarrow m=11$.

Таким образом, в данном случае для обеспечения $q_n=70\%$ при росте нагрузки в 2 раза требуется увеличить скорость обслуживания в 2 раза, либо увеличить число обслуживаемых устройств на 57 %.

5.5 Исследование эффективных режимов СМО по параметру относительной пропускной способности при изменении параметров

$$\Lambda_{0,1} \text{ и } m$$

В данном подразделе рассматриваются минимальное количество обслуживаемых устройств (m) и интенсивность входного потока Λ_0, Λ_1 , необходимые для обеспечения эффективной работы системы по критерию эффективности $q=70\%$ и реализующие условие:

$$q_{i,j}(m_i, \Lambda_j)=0,7.$$

Число обслуживаемых устройств m является дискретным параметром, что исключает возможность использования метода Ньютона для его поиска и реализацию внешнего цикла по Λ_0, Λ_1 аналогично тому, как это выполнено в разделе 4.2. В связи с этим, метода Ньютона реализован для параметров Λ_0, Λ_1 . Для сведения задачи внутреннего цикла к однопараметрической,

интенсивности всех составляющих потока λ_0 приняты равными между собой $\lambda_0 = \lambda_1 = \lambda$.

Параметризация модели:

- вспомогательная переменная $x = \lambda_0/3$;
- диапазон варьирования m : от 5 до 20;
- начальное приближение $x = 1$;
- $q_n = 0,7$;
- $\delta = 0,01$;
- параметры N, μ, ν фиксированы на основе исходных данных, указанных в главе 3.

Для вычисления численного значения производной в Mathcad использована симметричная разностная схема:

$$f'(x) \approx \frac{f(x+h) - f(x-h)}{2h},$$

$$f(x \pm h) = q \left(\frac{\lambda_0}{3} \pm h, \frac{\lambda_0}{3} \pm h, \frac{\lambda_0}{3} \pm h \right),$$

$$h = 10^{-5}.$$

Для двухкомпонентного потока λ_1 интенсивность входящего потока заявок 0-го типа распределена между 1-м и 2-м типом заявок:

$$f(x \pm h) = q \left(0, \frac{\lambda_1}{2} \pm h, \frac{\lambda_1}{2} \pm h \right).$$

Программная реализация алгоритма нахождения суммарной интенсивности входного потока λ_0 в Mathcad выглядит следующим образом:

```

h ← 10-5
λ_t0 ← λn + h
λ_t1 ← λn + h
λ_t ← λn + h
qp ← q(N, m, λ_t0, λ_t1, λ_t, μ, ν)
λ_t0 ← λn - h
λ_t1 ← λn - h
λ_t ← λn - h
qm ← q(N, m, λ_t0, λ_t1, λ_t, μ, ν)
dq ←  $\frac{qp - qm}{2 \cdot h}$ 

```

Рис. 5.18 – Программная реализация алгоритма поиска значений

m, A_0 для $q=70\%$ (численное значение производной)

```

while |f1| > δ
    λ_t0 ← λn
    λ_t1 ← λn
    λ_t ← λn
    f0 ← q(N, m, λ_t0, λ_t1, λ_t, μ, ν) - qn
    ...
    λn1 ← λn - α · (f0 / dq)
    while λn1 < 0 ∨ α > 10-12
        α ← α / 2
        λn1 ← λn - α · (f0 / dq)
        break if λn1 > 0
    if α ≤ 10-12
        Y ← "Method Error"
        break
    λ_t0 ← λn1
    λ_t1 ← λn1
    λ_t ← λn1
    f1 ← q(N, m, λ_t0, λ_t1, λ_t, μ, ν) - qn
    if |f1| < |f0|
        λn ← λn1
        α ← 1
    α ← α / 2 if |f1| ≥ |f0|
    Ym,0 ← λ_t · 3

```

Рис. 5.19 – Программная реализация алгоритма поиска значений

m, A_0 для $q=70\%$ (внутренний цикл)

Для нахождения требуемой минимальной интенсивности в случае двухкомпонентного потока A_1 параметру λ_0 присваивается нулевое значение. По найденным значениям параметров A_1 , A_0 и m построен график, представленный на рис. 5.20.

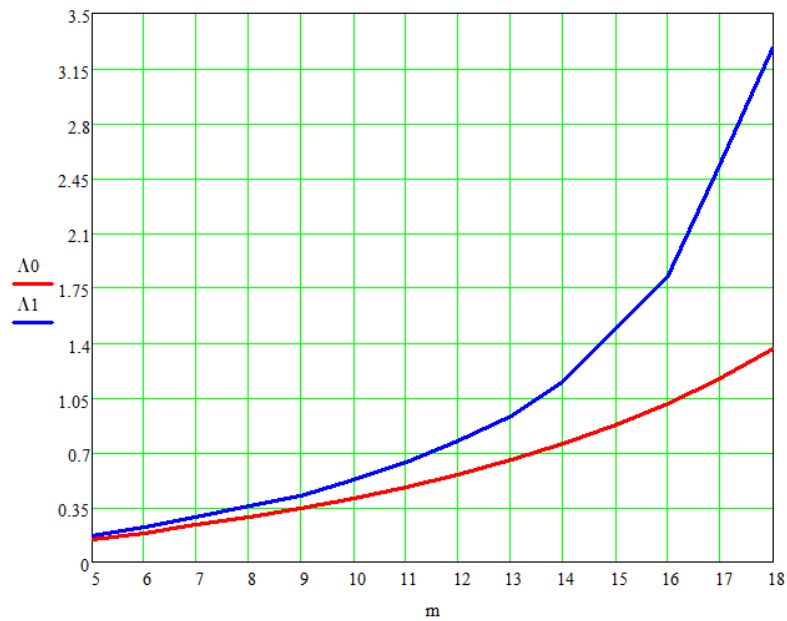


Рис. 5.20 – Зависимости $\Lambda_0(m)$ и $\Lambda_1(m)$ при $q=70\%$,
– красная кривая (Λ_0) соответствует двухкомпонентному потоку;
– синяя кривая (Λ_1) соответствует трехкомпонентному потоку.

По мере увеличения количества обслуживающих устройств m допускаемая интенсивность входящего потока заявок Λ_0 и Λ_1 также растет. Разница между кривыми становится более заметной по мере увеличения m . Это обусловлено с тем, что с ростом интенсивности входного потока заявок 0-го типа увеличивается и количество отказов в системе, накладывая ограничения на входящий поток для удовлетворения требования $q=70\%$. Таким образом, снижение интенсивности поступления «нетерпеливых» заявок позволяет существенно увеличить входящий поток заявок 1-го и 2-го типа при сохранении того же уровня относительной пропускной способности q_n .

На следующем графике, изображенном на рис. 5.21, представлены зависимости интенсивности входящего потока заявок Λ_0 от минимально необходимого количества обслуживающих устройств m при различных значениях интенсивности обслуживания μ для $q=70\%$.

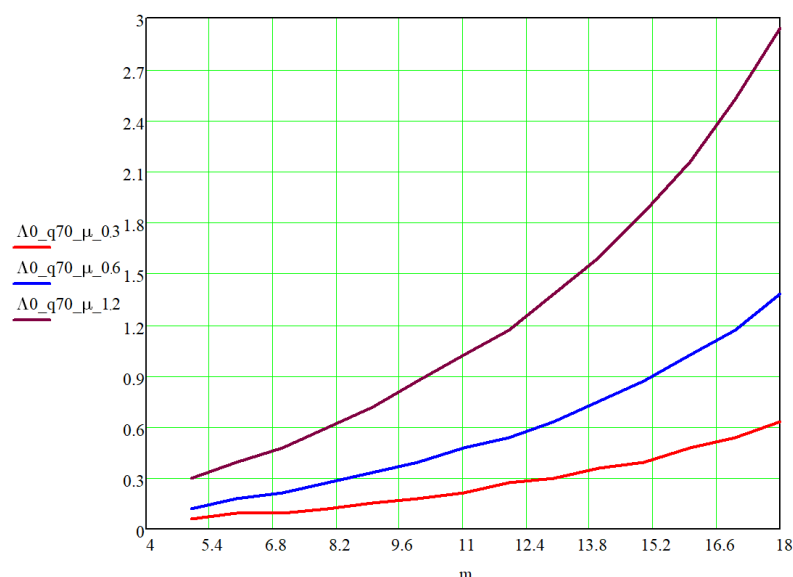


Рис. 5.21 – Зависимости $\lambda_0(m)$ при $q=70\%$ для различных значений μ :

- фиолетовая кривая ($\mu_{q70_A0_1.5}$) соответствует $\mu = 1,2$;
- синяя кривая ($\mu_{q70_A0_1.5}$) соответствует $\mu = 0,6$;
- красная кривая ($\mu_{q70_A0_1.5}$) соответствует $\mu = 0,3$.

По мере увеличения количества обслуживающих устройств m минимальная интенсивность входящего потока λ_0 , при которой система может обеспечить $q=70\%$, также возрастает. Большее количество каналов обслуживания позволяет снизить нагрузку на вход системы, что, в свою очередь, позволяет обрабатывать больший объем входящих заявок без снижения качества обслуживания. При фиксированном количестве обслуживающих устройств m увеличение интенсивности обслуживания μ приводит к росту допускаемой интенсивности входящего потока λ_0 .

Таким образом, пропорциональное повышение скорости обслуживания позволяет системе справиться с большим количеством входящих заявок, сохраняя при этом требуемый уровень пропускной способности.

На основании полученных результатов вычислительных экспериментов, построена область минимальных значений параметров СМО, в которой выполняется условие $q=70\%$. Каждая точка поверхности соответствует минимальному значению интенсивности обслуживания μ при которой достигается требуемый уровень пропускной способности для

заданных величин входного потока и числа обслуживающих устройств.
Цветовая шкала соответствует интенсивности обслуживания μ .

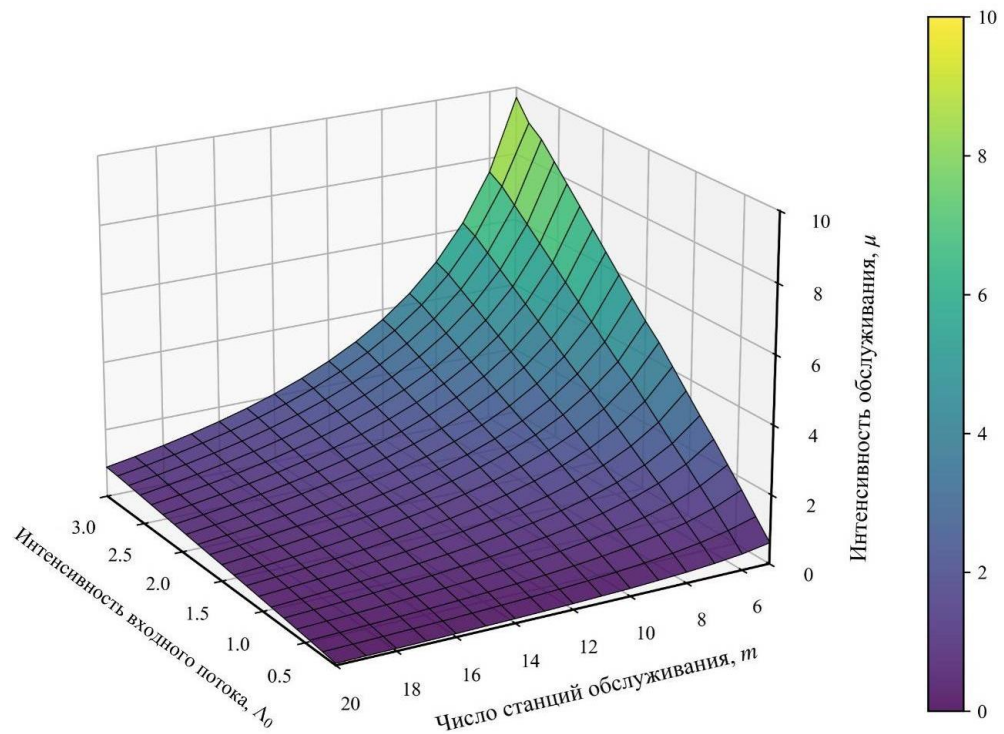


Рис. 5.22 – Область минимальных значений параметров СМО для трехкомпонентного потока, при которых $q=70\%$

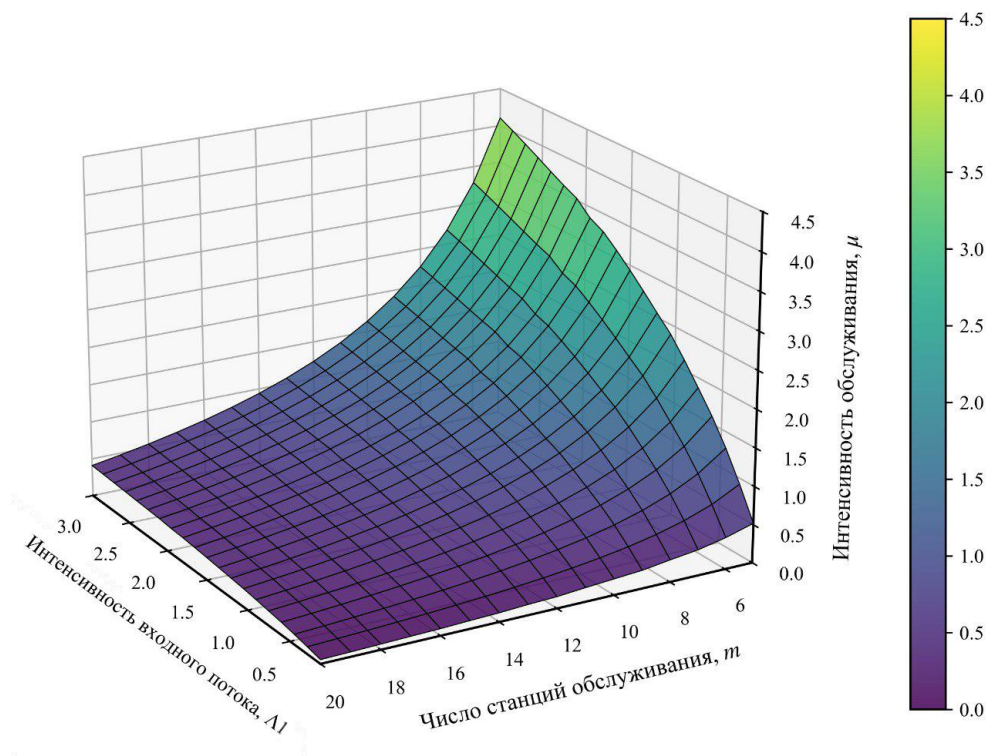


Рис. 5.23 – Область минимальных значений параметров СМО для двухкомпонентного потока, при которых $q=70\%$

5.6 Исследование эффективных режимов СМО по параметру среднего времени ожидания при изменении параметров m и λ_0, λ_1

Так как фактическое значение времени ожидания значительно варьируется в зависимости от исходных данных, в рамках проведения исследования задание некоего опорного значения среднего времени ожидания для поиска эффективных режимов работы нецелесообразно. Вместо этого в разделах 4.6-4.9 для поиска эффективных режимов работы использован корректирующий коэффициент среднего времени ожидания k , обеспечивающий заданный уровень корректировки времени ожидания $k \cdot \overline{t_{ож0}}$ для различных комбинаций параметров системы.

В данном подразделе исследуется зависимость минимального количества обслуживающих устройств m от интенсивности входящего потока λ_0 при котором выполняется условие:

$$\overline{t_{ож_{i,j}}}(m_i, \lambda_j) = k \cdot \overline{t_{ож0}}(m_0, \lambda_j),$$

где $\overline{t_{ож0}}(m_0, \lambda_j)$ – значения среднего времени ожидания, рассчитанные для $m=m_0$ и j -го значения $\lambda_{0,1}$.

Это позволяет определить минимально необходимое изменение количества обслуживающих устройств относительно значения m_0 для каждого значения параметра $\lambda_{0,1}$, при которых СМО работает с приемлемым средним уровнем задержек, равным $k \cdot \overline{t_{ож0}}$. Величина параметра k может изменяться в зависимости от поставленной задачи и требуемого времени ожидания.

Параметризация модели:

- диапазон значений λ_0 : от 0,15 до 3 с шагом 0,15;
- диапазон значений λ_1 : от 0,15 до 3 с шагом 0,15;
- диапазон варьирования m : от 5 до 20;
- начальное значение $m_0=5$;
- $k=0,5$;

– параметры N , μ , ν фиксированы на основе исходных данных, указанных в главе 3.

Программная реализация алгоритма нахождения минимального количества обслуживающих устройств $m(\lambda_0)$ в Mathcad с использованием метода вложенного перебора с равномерной дискретизацией:

```

for i_ ∈ 1..  $\frac{\lambda\_tmax}{step}$ 
     $\lambda\_t0 \leftarrow \lambda\_t0 + step$ 
     $\lambda\_t1 \leftarrow \lambda\_t1 + step$ 
     $\lambda\_t \leftarrow \lambda\_t + step$ 
     $m \leftarrow 5$ 
     $Z_{i\_} \leftarrow t\_w\_a(N, m, \lambda\_t0, \lambda\_t1, \lambda\_t, \mu, \nu)$ 
    for j_ ∈ 1.. m_max
         $T_{i\_} \leftarrow t\_w\_a(N, m, \lambda\_t0, \lambda\_t1, \lambda\_t, \mu, \nu)$ 
        if  $T_{i\_} \leq k \cdot Z_{i\_}$ 
             $Y_{i\_},1 \leftarrow m$ 
             $Y_{i\_},2 \leftarrow \lambda\_t0 \cdot 3$ 
             $Y_{i\_},3 \leftarrow Z_{i\_}$ 
             $Y_{i\_},4 \leftarrow T_{i\_}$ 
            break
         $m \leftarrow m + 1$ 

```

Рис. 5.24 – Программная реализация алгоритма поиска значений

$$m(\lambda_0) \text{ для } 0,5 \cdot \overline{t_{ож0}}$$

Полученные значения m для каждого $\lambda_{0,1}$ представляют собой минимальное число обслуживающих устройств, при котором среднее время ожидания не превышает половины от его значения при $m=5$. Полученные зависимости для двухкомпонентного и трехкомпонентного потоков приведены на рис. 5.25.

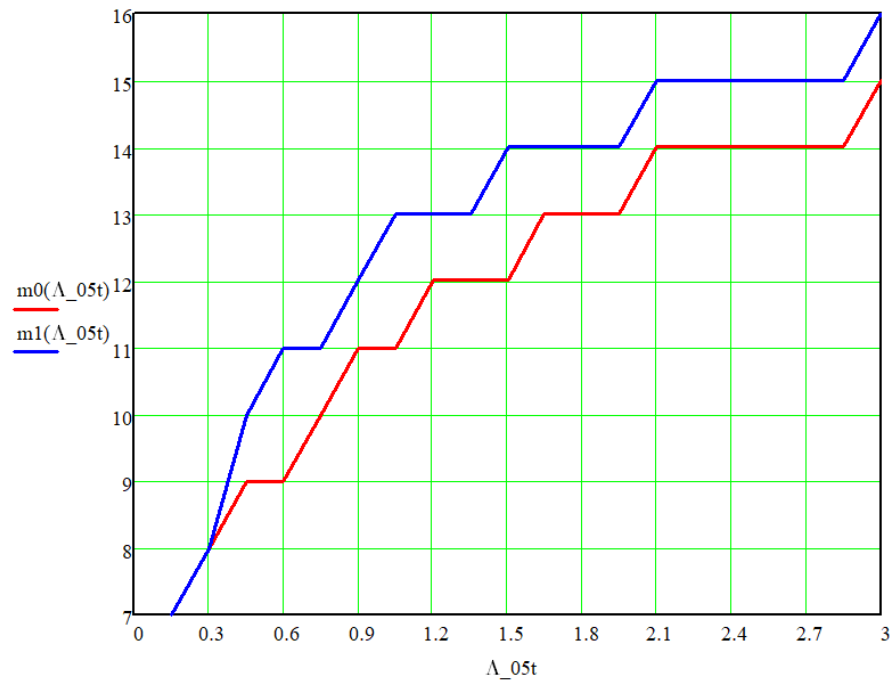


Рис. 5.25 – Зависимость $m(\Lambda_0)$ для $0,5 \cdot \overline{t_{ожс_0}}$

- красная кривая ($m0(\Lambda_0_{05t})$) соответствует трехкомпонентному потоку;
- синяя кривая ($m1(\Lambda_0_{05t})$) соответствует двухкомпонентному потоку.

На начальном участке кривой $\Lambda_0 \leq 0,3$ требуемое число обслуживающих устройств для потоков Λ_0 и Λ_1 составляет $m=8$. Это связано с тем, что при малой интенсивности входного потока влияние компоненты λ_0 несущественно ввиду малого количества отказов в обслуживании.

На участке $\Lambda_0 \in (0,3; 1,5]$ увеличение количества «нетерпеливых» заявок ведет к снижению нагрузки на очередь. Они, не занимая очередь ввиду наличия отказов, уменьшают среднее время ожидания заявок λ_1 и λ , позволяя достичь целевого значения с меньшими затратами ресурсов.

На участке $\Lambda_0 > 1,5$ рост входного потока частично нивелирует возможность заявок λ_0 разгружать очередь ввиду роста отказов, что соотносится с графиком на рис. 5.8.

Таким образом, при сокращении среднего времени ожидания, наличие во входном потоке заявок 0-го типа снижает ресурсоемкость СМО. Эффект максимален при умеренных нагрузках ($\Lambda_0 \approx 1$), где $\Delta m_0 = 2$ и сохраняется на уровне 6-7% при насыщении системы.

5.7 Исследование эффективных режимов СМО по параметру среднего времени ожидания при изменении параметров m и μ

В рамках исследования эффективных режимов работы замкнутой СМО с трехкомпонентным потоком заявок и ограничением по времени ожидания по аналогии с разделом 4.6 проведено численное моделирование для определения параметров системы, для которых обеспечивается условие:

$$\overline{t_{ож}}_{i,j}(m_i, \mu_j) = k \cdot \overline{t_{ож}}_0(m_0, \mu_j),$$

где $\overline{t_{ож}}_0(m_0, \mu_j)$ – значения среднего времени ожидания, рассчитанные для $m=m_0$ и j -го значения μ .

Параметризация модели:

- диапазон значений μ : от 0,01 до 2 с шагом 0,05;
- диапазон варьирования m : от 5 до 20;
- начальное значение $m_0=5$, $k=0,5$;
- параметры N , λ_0 , ν фиксированы на основе исходных данных, указанных в главе 3.

Программная реализация алгоритма нахождения минимального значения $m(\mu)$ для обеспечения требуемой величины среднего времени ожидания в Mathcad выглядит следующим образом:

```

μ ← μ + step
m ← 5
Zi ← t_w_a(N, m, λ_t0, λ_t1, λ_t, μ, ν)
for j_ ∈ 1..m_max
    Ti ← t_w_a(N, m, λ_t0, λ_t1, λ_t, μ, ν)
    if Ti ≤ k·Zi
        Yi,1 ← m
        Yi,2 ← μ
        Yi,3 ← Zi
        Yi,4 ← Ti
        break
    m ← m + 1

```

Рис. 5.26 – Программная реализация алгоритма поиска значений

$m(\mu)$ для $0,5 \cdot \overline{t_{ож}}_0$ (внутренний цикл)

На графике, приведенном на рис. 5.27, показана зависимость минимального количества обслуживающих устройств m от интенсивности обслуживания μ , необходимого для обеспечения среднего времени ожидания заявки в очереди, равного половине времени ожидания при $m_0=5$.

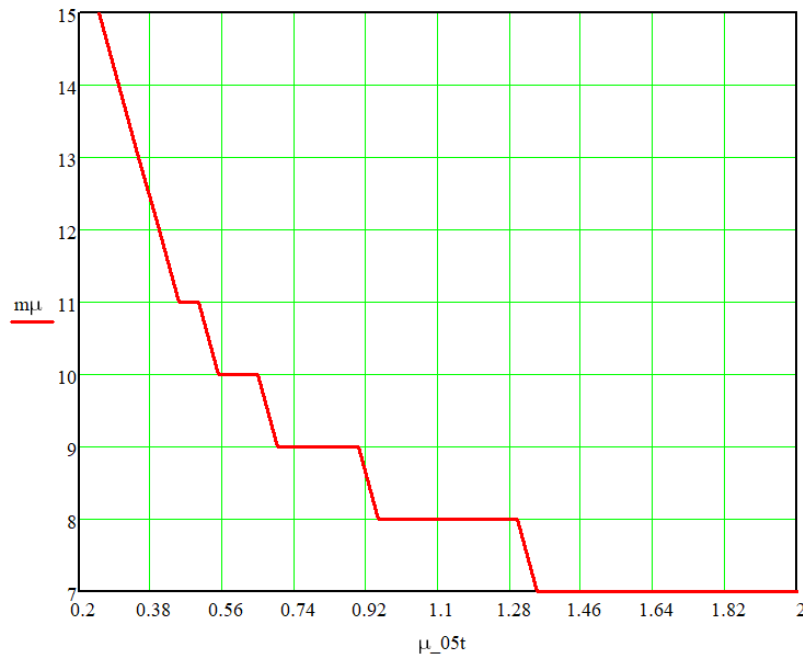


Рис. 5.27 – Зависимость $m(\mu)$ для $0,5 \cdot \overline{t_{ож0}}$

На начальном участке $\mu < 0,7$ система требует значительного увеличения числа приборов $\Delta m_0 \in [5; 10]$, так как низкая скорость обслуживания провоцирует рост очереди и времени ожидания.

В диапазоне $\mu \in [0,7; 1,35]$ скорость изменения Δm_0 становится менее заметной. Максимальное увеличение скорости обслуживания μ на данном участке снижает требуемое m на 11 %. Данный эффект вызван сокращением времени занятия обслуживающих устройств и уменьшением очереди. Точка перелома $\mu \approx 0,7$ разделяет зоны, в которых увеличение скорости обслуживания дает максимальный ($\mu < 0,7$) и минимальный ($\mu > 0,7$) эффект.

Дальнейшее увеличение скорости обслуживания $\mu > 1,35$ уже не дает пространства для столь же значительного снижения требуемого количества m . Так, для снижения времени обслуживания вдвое при $\mu \in (1,35; 3]$ потребуется увеличить количество обслуживающих устройств на 40 %.

Результаты вычислительных экспериментов, полученные в разделах 4.6, 4.7 позволяют построить область минимально допустимых значений $t(\mu, \Lambda_{0,1})$, в которой выполняется условие:

$$\overline{t_{ож}}_{i,j,k}(m_i, \mu_j, \Lambda_k) = 0,5 \cdot \overline{t_{ож0}}(m_0, \mu_j, \Lambda_k),$$

где $m_0=5$ и параметр $\overline{t_{ож0}}(m_0, \mu_j, \Lambda_k)$ рассчитан для j -го значения μ и k -го значения $\Lambda_{0,1}$.

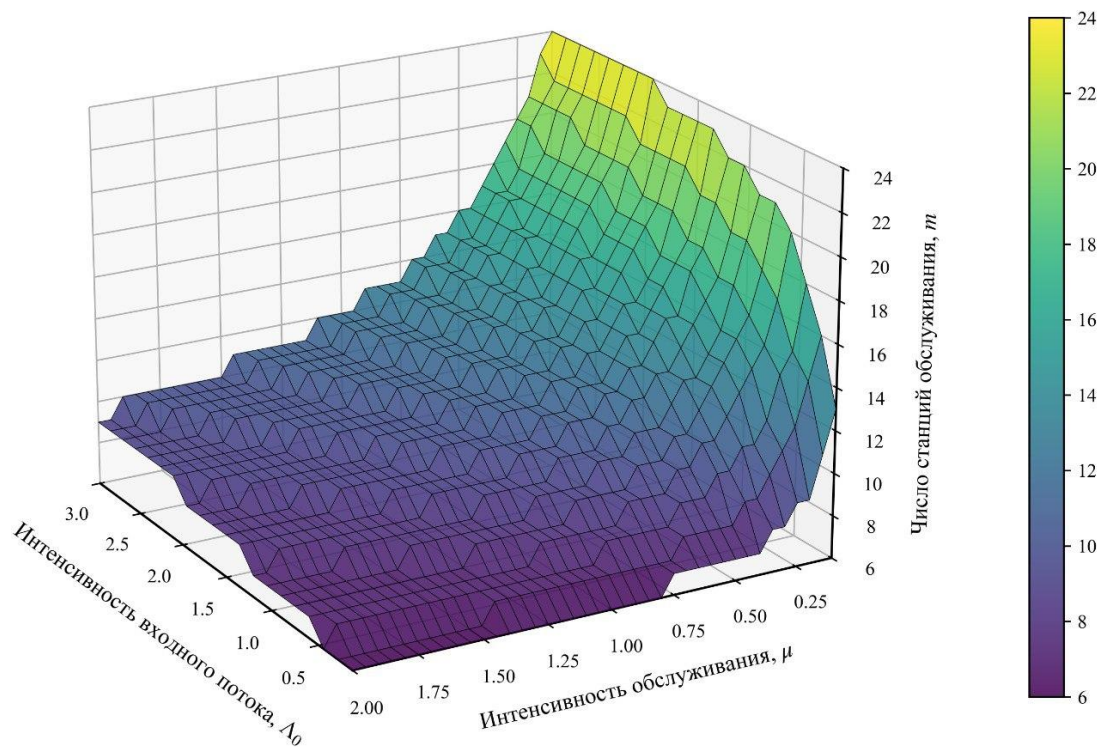


Рис. 5.28 – Область допустимых параметров СМО, обеспечивающих эффективную работу по параметру $\overline{t_{ож}}$ для трехкомпонентного потока заявок

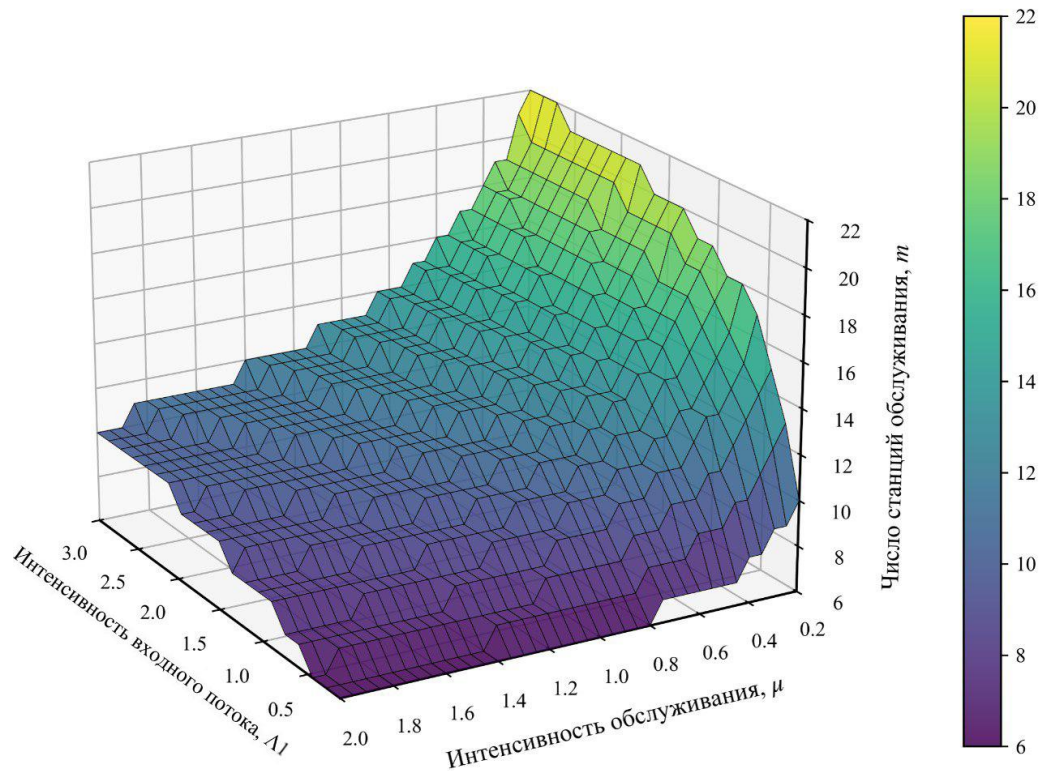


Рис. 5.29 – Область допустимых параметров СМО, обеспечивающих эффективную работу по параметру $\overline{t_{ож}}$ для двухкомпонентного потока заявок

5.8 Исследование эффективных режимов СМО по параметру среднего времени ожидания при изменении параметров m и ν

В рамках исследования эффективных режимов работы замкнутой СМО с трехкомпонентным потоком заявок и ограничением по времени ожидания по аналогии с 4.6 проведено численное моделирование для определения параметров системы, для которых обеспечивается условие:

$$\overline{t_{ож}}_{i,j}(m_i, \nu_j) = k \cdot \overline{t_{ож0}}(m_0, \nu_j),$$

где $\overline{t_{ож0}}(m_0, \nu_j)$ – значения среднего времени ожидания, рассчитанные для $m=m_0$ и j -го значения ν .

Параметризация модели:

- диапазон варьирования m : от 5 до 20;
- диапазон значений ν : от 0,2 до 3 с шагом 0,01;
- начальное значение $m_0=5$;

- $k=0,5$;
- параметры N , λ_0 , μ фиксированы на основе исходных данных, указанных в главе 3.

Программная реализация алгоритма нахождения минимального значения $m(v)$ для обеспечения требуемой величины среднего времени ожидания в Mathcad выглядит следующим образом:

```

m(v) :=
  νmax ← 3
  ν ← 0.2
  step ← 0.05
  m_max ← 20
  k ← 0.5
  for i_ ∈ 1..  $\frac{\nu_{\max}}{\text{step}}$ 
    ν ← ν + step
    m ← 5
    Zi_ ← t_w_a(N, m, λ_t0, λ_t1, λ_t, μ, ν)
    for j_ ∈ 1.. m_max
      Ti_ ← t_w_a(N, m, λ_t0, λ_t1, λ_t, μ, ν)
      if Ti_ ≤ k · Zi_
        Yi_,1 ← m
        Yi_,2 ← ν
        Yi_,3 ← Zi_
        Yi_,4 ← Ti_
        break
    m ← m + 1
  Y

```

Рис. 5.30 – Программная реализация алгоритма поиска значений

$$m(v) \text{ для } 0,5 \cdot \overline{t_{\text{ож}_0}}$$

На графике, изображенном рис. 5.31, показана зависимость минимального количества обслуживающих устройств m от интенсивности вылета заявок ν , необходимого для обеспечения среднего времени ожидания заявки в очереди, равного половине времени ожидания при $m_0=5$.

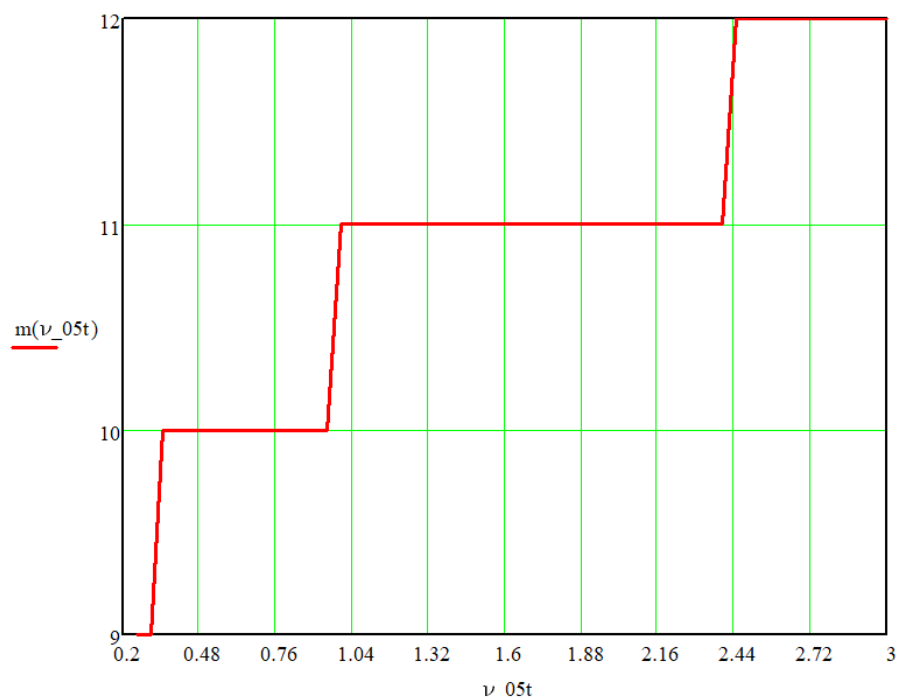


Рис. 5.31 – Зависимость $m(v)$ для $0,5 \cdot \overline{t_{ож0}}$

На начальном участке $v \leq 0,3$ количество заявок в очереди невелико и система способна сократить их среднее время ожидания вдвое путем увеличения m_0 на 80 %. Так, для $v=0,25$ и двухкратного сокращения времени ожидания необходимое изменение m_0 составит 4 станции обслуживания.

На участке $v \in (0,3; 0,95)$ количество уходов заявок 1-го типа увеличивается, сокращая среднее время ожидания, что в итоге приводит к необходимости добавления еще одного обслуживающего устройства ($\Delta m_0=5$) при $v \approx 0,95$ для достижения требуемого критерия эффективности.

С ростом параметра $v \geq 0,95$ система становится менее чувствительна к колебаниям параметра v для сокращения среднего времени ожидания заявок в очереди ввиду снижения в ней количества заявок 1-го типа. Данный вывод соотносится с графиками на рис. 4.25 и 4.30. Таким образом интервалы Δv , на концах которых происходит изменение требуемого количества обслуживающих устройств (Δm_0), становятся больше с ростом параметра v .

Сравнивая полученный график с графиком на рис. 5.30, можно сделать вывод о том, что для данной системы, в случае сокращения среднего времени

ожидания при малых значениях $\nu < 0,5$ требуются меньшие ресурсы (Δm_0), нежели при малых значениях $\mu < 0,5$. Обратная ситуация наблюдается при $\nu > 0,5$ и $\mu > 0,5$:

- при $\nu=0,3 \rightarrow \Delta m_0=9$;
- при $\mu=0,3 \rightarrow \Delta m_0=14$;
- при $\nu=1,5 \rightarrow \Delta m_0=6$;
- при $\mu=1,5 \rightarrow \Delta m_0=2$;

5.9 Исследование эффективных режимов СМО по параметру среднего времени ожидания при изменении параметров μ и ν

В рамках исследования эффективных режимов работы замкнутой СМО с трехкомпонентным потоком заявок и ограничением по времени ожидания проведено численное моделирование для определения параметров системы, для которых обеспечивается условие:

$$\overline{t_{ож\,i,j}}(\mu_i, \nu_j) = k \cdot \overline{t_{ож0}}(\mu_0, \nu_j),$$

где $\overline{t_{ож0}}(\mu_0, \nu_j)$ – значения среднего времени ожидания, рассчитанные для $\mu=\mu_0$ и j -го значения ν .

Параметризация модели:

- начальное приближение μ_0 : от 0,5
- диапазон значений ν : от 0,2 до 3 с шагом 0,05;
- $m=5$
- $k=0,5$;
- $\delta=0,001$;
- параметры N и L_0 , фиксированы на основе исходных данных, указанных в главе 3.

Программная реализация алгоритма нахождения минимального значения $\mu(\nu)$ для обеспечения требуемой величины среднего времени ожидания в Mathcad выглядит следующим образом:

```

while f1 > δ
    μ ← μ0
    f0 ← t_w_a(N, m, λ_t0, λ_t1, λ_t, μ, ν) - t_n
    q' ←  $\frac{d}{d\mu} t_w_a(N, m, \lambda_{t0}, \lambda_{t1}, \lambda_t, \mu, \nu)$ 
    μ1 ← μ0 - α ·  $\left(\frac{f0}{q'}$ 
    while μ1 < 0 ∨ α > 10-12
        α ←  $\frac{\alpha}{2}$ 
        μ1 ← μ0 - α ·  $\left(\frac{f0}{q'}$ 
        break if μ1 > 0
    Y ← "Method Error" if α ≤ 10-12
    break
    μ ← μ1
    f1 ← t_w_a(N, m, λ_t0, λ_t1, λ_t, μ, ν) - t_n
    μ0 ← μ1 if |f1| < |f0|
    α ← 1
    α ←  $\frac{\alpha}{2}$  otherwise
    Yi,0 ← μ

```

Рис. 5.32 – Программная реализация алгоритма поиска значений

$\mu(\nu)$ для $0,5 \cdot \overline{t_{ож0}}$ (внутренний цикл)

В отличие алгоритма на рис. 5.5, в данном случае необходимо фиксировать для каждой итерации по ν величину опорного значения среднего времени ожидания $\overline{t_{ож0}}$ при $\mu_0=0,6$ для последующего нахождения $\overline{t_{ож}}$.

На графике, представленном рис. 5.33, показана зависимость минимальной интенсивности обслуживания μ необходимой для обеспечения среднего времени ожидания заявки в очереди, равного половине времени ожидания при $\mu_0=0,6$ для различных значений ν .

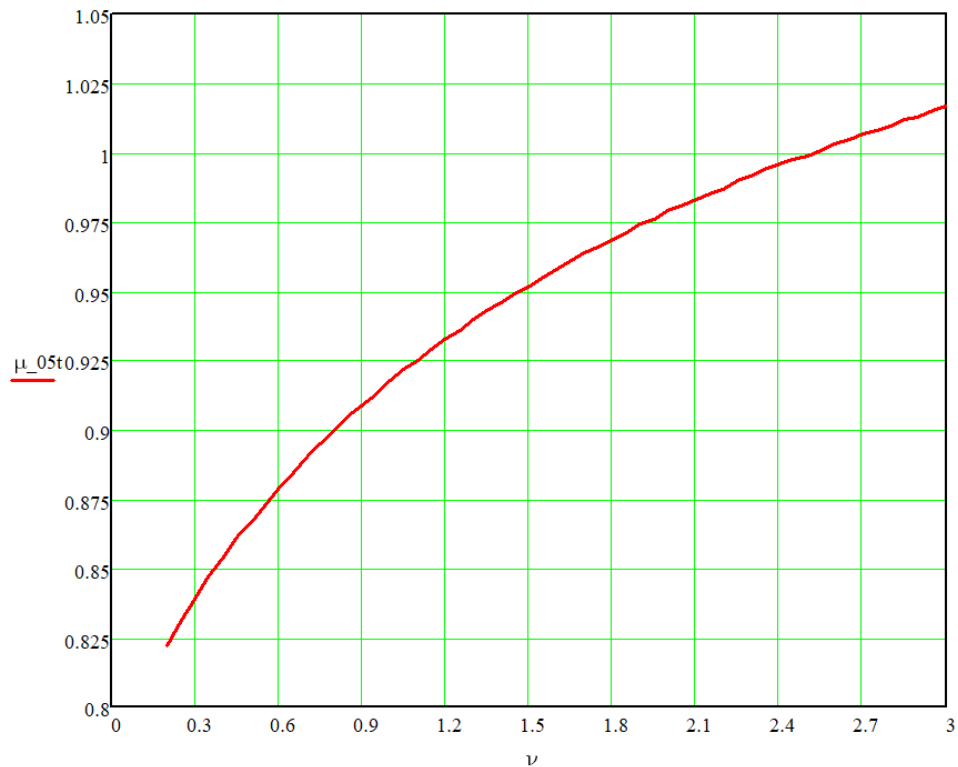


Рис. 5.33 – Зависимость $\mu(\nu)$ для $0,5 \cdot \overline{t_{ож0}}$

Полученный вид зависимости соотносится с представленным на рис. 5.31 ввиду схожей реакции системы на изменение параметров μ и m . Увеличение интенсивности обслуживания μ в ответ на увеличение параметра ν демонстрирует нелинейный характер.

На начальном этапе $\nu < 0,9$ система работает в режиме, где средняя длина очереди и конкуренция за ресурсы максимальны. Следствием увеличения параметра μ при наибольшей загруженности очереди является обратная связь, создающая дополнительную нагрузку на вход СМО ввиду ее замкнутости. В связи с этим на данном участке наблюдается более резкое увеличение μ с ростом ν :

- при $\nu=0,3 \rightarrow \sim 1,4\mu_0$;
- при $\nu=0,6 \rightarrow \sim 1,5\mu_0$.

При высоких значениях $\nu \in [0,9; 2,4]$ заявки 1-го типа чаще покидают систему до обслуживания, что уменьшает нагрузку на очередь и обслуживающие устройства. В этом случае рост ν уже не приводит к

пропорциональному увеличению $\overline{t_{ожс_0}}$, т.к. система частично разгружена через отказы и рост μ замедляется:

– при $\nu=1,2 \rightarrow \sim 1,56\mu_0$;

– при $\nu=1,5 \rightarrow \sim 1,59\mu_0$.

На конечном участке $\nu (>2.4)$ кривая растет медленней, что указывает на то, что количество заявок в очереди сокращается ввиду высокой интенсивности «вылетов». Поток уходов необслуженных заявок в таком случае близок к своему максимальному значению и параметр μ стабилизируется в диапазоне $1,69\mu_0-1,7\mu_0$.

Выводы:

1. В данной главе исследованы алгоритмы поиска эффективных режимов работы СМО при совместном варьировании несколькими исходными параметрами системы. При этом в качестве критериев эффективности использованы значения параметра относительной пропускной способности $q=70\%$ и времени ожидания, сокращенному вдвое.

2. Проведена количественная оценка требуемых минимальных мощностей системы для выхода на эффективный режим работы в случае двухкомпонентного и трехкомпонентного потока заявок. Проведенный анализ эффективных режимов работы по параметру относительной пропускной способности системы показал, что при добавлении «нетерпеливых» заявок в двухкомпонентный поток зависимость интенсивности обслуживания от суммарной интенсивности входного потока меняется со степенной на линейную.

3. Система с двухкомпонентным потоком заявок λ_1 более чувствительна к потерям, по отношению к системе с «нетерпеливыми» заявками λ_0 . Чем выше параметр λ_0 в трехкомпонентном потоке λ_0 , тем меньшее влияние ограничение времени ожидания ν оказывает на относительную пропускную способность q . При низких значениях λ_0 система, напротив, более чувствительна к потерям.

4. С ростом количества обслуживающих устройств m и скорости обслуживания μ система становится менее чувствительна к изменению параметра интенсивности ухода заявок из очереди ν .

5. Следует учитывать, что найденные эффективные режимы работы системы, основаны на математическом ожидании параметров и в условиях реальных стохастических процессов с экспоненциальным распределением интенсивностей будут подвержены колебаниям вокруг математического ожидания, что требует резервирования устойчивости системы к возможным отклонениям параметров при практической реализации.

6 ПРАКТИЧЕСКИЕ РЕКОМЕНДАЦИИ ПО ВНЕДРЕНИЮ РЕЗУЛЬТАТОВ ДИССЕРТАЦИОННОЙ РАБОТЫ

6.1 Общие положения

Результаты диссертационной работы, направленные на разработку и анализ математической модели замкнутой СМО с трехкомпонентным потоком заявок и ограничением по времени ожидания, могут быть использованы для поиска эффективных режимов работы как для уже существующих, так и для проектируемых объектов, работающих по принципу замкнутых СМО. Основными областями применения являются логистика, телекоммуникации, управление производственными процессами, call-центры, облачные вычисления и другие сферы, где возникает необходимость эффективного управления ресурсами и минимизации задержек в обслуживании.

В рамках данной главы предлагаются практические рекомендации по внедрению полученной математической модели в соответствии с принятыми допущениями:

- 1) **Вид случайного процесса** – марковский;
- 2) **Входящий поток заявок** – пуассоновский, двухкомпонентный или трехкомпонентный;
- 3) **Время обслуживания** – экспоненциально распределенное;
- 4) **Ограничение на время ожидания** – экспоненциально распределенное;
- 5) **Дисциплина очереди** – FIFO;
- 6) **Число заявок в источнике** – фиксировано.

Эти характеристики позволяют адаптировать модель под широкий спектр задач, где входящий поток заявок является случайным, а время обслуживания зависит от интенсивности работы станций обслуживания.

6.2 Проверка соответствия входного потока заявок пуассоновскому потоку

Рассмотрим алгоритм проверки соответствия входного потока заявок пуассоновскому потоку.

1. Пуассоновский поток является одной из ключевых моделей, используемых для описания случайных событий в СМО и характеризуется следующими свойствами, выполнение которых обязательно для рассматриваемой СМО:

– стационарность: средняя интенсивность поступления заявок постоянна во времени;

– ординарность: вероятность появления более одной заявки за бесконечно малый промежуток времени стремится к нулю;

– отсутствие последействия: события происходят независимо друг от друга, то есть вероятность поступления новой заявки не зависит от истории предыдущих событий.

2. Необходимо проанализировать состав входящего потока заявок в проектируемую или анализируемую систему на наличие заявок 0-го, 1-го и 2-го типа.

3. Зафиксировать временные метки поступления заявок каждого типа в систему за отчетный период R .

4. Для анализа стационарного режима работы системы период R должен быть достаточно большим, чтобы обеспечить статистическую устойчивость данных. Также, в случае, например, сезонных или циклических изменений, период должен охватывать полный цикл изменений. В замкнутых СМО количество заявок в источнике (N) фиксировано, поэтому при малом количестве заявок в источнике, отчетный период R может быть выбран исходя из времени полного оборота заявок в системе.

5. Отчетный период R разбивается на равные временные сегменты t (например, минуты, часы, дни). Чем больше временных сегментов t содержит в себе отчетный период R , тем точнее результат вычислений. Оценивается

интенсивность входящего потока L_i на каждом временном сегменте t (количество поступивших заявок за время t).

6. Для оценки репрезентативности интенсивности потоков заявок, для нескольких отчетных периодов $R_1, R_2, R_3, \dots R_j$, для начала, рассчитывается среднее значение интенсивности входного потока Q_j для каждого из доступных периодов R_j :

$$Q_j = \frac{\sum_{i=1}^n L_i}{n_i}, \quad (6.1)$$

Где L_i – интенсивности входящего потока заявок в каждом временном сегменте t периода R_j ;

n_i – количество временных сегментов, на которые разделен период R_j .

Размерность величины Q_j зависит от размерности временного интервала t . Например, если $t=60$ сек., то Q_j выражается в размерности [ед. заявок в мин.]. Необходимо учитывать, что статистические данные по заявкам 0-го, 1-го и 2-го типа собираются отдельно.

7. Далее определяется среднее значение интенсивности входящего потока \bar{Q} между несколькими отчетными периодами:

$$\bar{Q} = \frac{\sum_{j=1}^k Q_j}{k}, \quad (6.2)$$

Где k – количество отчетных периодов R_j .

8. Рассчитывается относительное стандартное отклонение σ , %, характеризующее степень рассеяния данных Q_j по выборкам относительно среднего значения интенсивности входящего потока \bar{Q} :

$$\sigma = \left(\frac{\sum_{j=1}^k (Q_j - \bar{Q})^2}{k-1} \right)^{0,5} \cdot \bar{Q}^{-1} \cdot 100\%. \quad (6.3)$$

По результатам проведенного расчета для двух и более периодов, в зависимости от стабильности потока заявок, принимается решение о необходимости увеличения или фиксации длительности и количества отчетных периодов R_j .

9. Формирование репрезентативной выборки. Репрезентативная выборка может быть сформирована на основании расчета среднего значения интенсивностей \bar{Q} между выборками R_j на каждом временном сегменте t . Либо репрезентативной признается одна из выборок R_j , среднее значение интенсивности входящего потока Q которой ближе всего к среднему значению входящего потока между выборками \bar{Q} . Выдвигается нулевая гипотеза о соответствии входящих потоков заявок распределению Пуассона.

10. На основании полученной репрезентативной выборки строится эмпирическая функция распределения. Для этого:

- данные сортируются по возрастанию количества поступивших в систему заявок n за временные сегменты t ;
- при равных значениях n на нескольких временных сегментах t данные группируются в категории k ;
- для каждой группы k определяется частота Y появления величины n в выборке.

11. Вычисляется среднее значение интенсивности потока входящих требований в репрезентативной выборке. Полученное значение используется в качестве параметра λ для вычисления теоретического значения функции вероятности:

$$p(k) = \frac{\lambda^k}{k!} e^{-\lambda}. \quad (6.4)$$

12. Для вычисления теоретических частот $Y_m(k)$, каждое из полученных значений вероятности $p(k)$ умножается на величину выборки.

13. В случае, если соседние категории k имеют малые теоретические или эмпирические частоты, группы допускается объединять.

14. При необходимости графического отображения строится гистограмма частот. Результаты сравниваются с теоретической экспоненциальной кривой.

15. Для оценки соответствия входящего потока распределению Пуассона в рамках данной работы предлагается использовать критерий хи-

квадрат Пирсона [84], позволяющий определить, насколько полученная выборка близка к эталонной:

$$\chi^2 = \sum_k \frac{(Y_o(k) - Y_m(k))^2}{Y_m(k)}. \quad (6.5)$$

16. Определим количество степеней свободы df :

$$df = N - \hat{k} - 1, \quad (6.6)$$

где N – число категорий после объединения;

\hat{k} – число оцениваемых параметров (т.к. оценивается только λ , $\hat{k}=1$).

17. Для полученного значения степеней свободы определяется критическое значения $\chi^2_{крит}$ из справочных таблиц хи-квадрат [85, 86], либо значение $\chi^2_{крит}$ рассчитывается в ПО Microsoft Excel для выбранного значения уровня значимости α .

18. В случае $\chi^2 < \chi^2_{крит}$, нулевая гипотеза о соответствии входного потока распределению Пуассона не отвергается и данные могут быть использованы для моделирования работы рассматриваемой СМО. Дополнительно рекомендуется провести расчет параметра СКО, т.к. в распределении Пуассона он численно равен математическому ожиданию. В случае экспериментальной выборки параметр СКО не должен значительно отклоняться от математического ожидания.

19. В случае значительного расхождения между теоретическим и эмпирическим распределением входного потока рекомендуется разбить данные на подвыборки и выполнить тест хи-квадрат Пирсона для каждой из них.

20. Определяется используемое в математической модели математическое ожидание интенсивности одной заявки:

$$\lambda_i = \frac{Q}{N}, \quad (6.7)$$

где Q – средняя интенсивность входящего потока выборки, признанной репрезентативной;

N – общее число заявок в источнике.

21. Рассмотренный алгоритм выполняется отдельно для потоков заявок 0-го, 1-го и 2-го типов.

22. Определяются суммарные интенсивности потоков заявок для ввода в математическую модель:

$$A_0 = \lambda_0 + \lambda_1 + \lambda,$$

$$A_1 = \lambda_1 + \lambda,$$

где λ_0 , λ_1 , и λ – определены либо на основании расчета среднего значения интенсивностей между выборками, либо по единственной выборке, среднее значение интенсивности входящего потока которой ближе всего к среднему по выборкам.

Далее, необходимо определить соответствует ли время обслуживания экспоненциальному распределению.

Основные свойства:

- беспамятность: вероятность наступления события в будущем не зависит от прошедшего времени наблюдения;
- вероятность наступления события за единицу времени не меняется со временем.

6.3 Проверка соответствия времени обслуживания экспоненциальному распределению

Рассмотрим алгоритм проверки соответствия времени обслуживания экспоненциальному распределению.

1. Необходимо зафиксировать временные метки между началом и окончанием обслуживания. При наличии очереди временная метка окончания обслуживания заявки должна совпадать с началом обслуживания следующей заявки.

2. Период R , на котором определяются временные метки, должен быть достаточно большим, чтобы обеспечить статистическую устойчивость

данных. Также, в случае, например, сезонных или циклических изменений, период должен охватывать полный цикл изменений.

3. Для определения времени обслуживания вычисляется расстояние между временными метками и определяется среднее значение выборки \bar{x} . В случае многокомпонентной системы допускается рассчитывать усредненное значение \bar{x} между несколькими обслуживающими устройствами.

4. Параметр μ для экспоненциального распределения равен обратному значению среднего значения выборки:

$$\mu = \frac{1}{\bar{x}}. \quad (6.8)$$

5. Оценка расхождений между выборками $R_1, R_2, R_3, \dots R_j$ определяется по параметру μ аналогично рассмотренному алгоритму для распределения Пуассона (п. 6-8).

6. Совокупность полученных статистических данных о времени обслуживания разбивается на равные интервалы Δt по времени обслуживания и определяются наблюдаемые частоты событий $Y_o(T)$ на интервалах $(t, t+\Delta t)$.

7. Определяются наблюдаемые частоты (количество данных, попадающих в каждый из диапазонов). Величина интервала Δt должна выбираться так, чтобы количество попадающих в него частот $Y_o(T) \geq 1$.

8. При необходимости графического отображения строится график наблюдаемых частот $Y_o(T)$.

9. Теоретическую частоту $Y_m(T)$ можно определить как:

$$Y_m(T) = P(a \leq t < b) \cdot n,$$

где $P(a \leq t < b)$ – теоретическая вероятность нахождения времени обслуживания в диапазоне $[a; b)$;

n – число наблюдений в выборке.

Теоретическая функция вероятности для экспоненциального распределения имеет вид:

$$F(t) = 1 - e^{-\mu t},$$

тогда

$$P(a \leq t < b) = F(b) - F(a),$$

и значит

$$Y_m(T) = n \cdot (e^{-\mu a} - e^{-\mu b}). \quad (6.9)$$

10. При необходимости строится кривая теоретических частот $Y_m(t)$ для сравнения с эмпирической функцией $Y_e(T)$.

11. Вычисление статистики хи-квадрат проводят аналогично с рассмотренным в п.15-19 алгоритмом для распределения Пуассона. Дополнительно рекомендуется провести расчет параметра СКО. В случае экспериментальной выборки он не должен значительно отклоняться от математического ожидания, что в свою очередь является свойством экспоненциального распределения.

6.4 Проверка соответствия ограничения на время ожидания экспоненциальному распределению

Информация о **граничном значении времени ожидания** собирается на основании статистических данных ухода из очереди заявок 1-го типа. Алгоритм проверки соответствия потока экспоненциальному распределению идентичен приведенному выше, за исключением фиксации временных меток. В данном случае необходимо фиксировать временные метки между двумя последовательными уходами заявок из очереди.

Следует учитывать, что в отличие от модели, представленной в работе [86], где ограничение на время ожидания может быть использовано как регулятор длины очереди, в рамках данной работы ограничение на время ожидания является дескриптивной величиной, определяемой на основе статистических данных. На практике это означает, что ключевая задача данного параметра – корректно отразить реальное поведение заявок типа 1, основанное на статистике. Таким образом, эффективное применение модели требует строгого соблюдения статистической обоснованности граничного времени ожидания и разделения заявок на типы на этапе анализа данных.

6.5 Математическое моделирование работы СМО в стационарном режиме

Для расчета параметров проектируемого или анализируемого объекта в качестве исходных данных используются:

- средняя интенсивность входящего потока заявок λ_0 , λ_1 , и λ , определенная на основании статистических данных и имеющая пуассоновский характер распределения;
- средняя интенсивность обслуживания заявки одним обслуживающим устройством μ , определенная на основании статистических данных и распределенная по экспоненциальному закону;
- средняя интенсивность ухода из системы заявок 1-го типа ввиду ограничения на время ожидания ν , определенная на основании статистических данных и распределенная по экспоненциальному закону;
- общее число заявок в источнике N ;
- число обслуживающих устройств m .

На основании вышеперечисленных исходных данных, в соответствии с разделом 2 рассчитываются характеристики производительности СМО. В случае, если система не достигает требуемого уровня эффективности, в соответствии с разделами 3 и 4 система перенастраивается под требуемый уровень эффективности путем варьирования численными значениями исходных данных.

Пример оценки соответствия входного потока распределению Пуассона и времени обслуживания экспоненциальному распределению приведен в приложении А.

Выводы:

1. В данной главе приведены примеры алгоритмов оценки применимости разработанной математической модели для описания исследуемого объекта. Приведены основные допущения, используемые при разработке математической модели, а также способы проверки соответствия реальных статистических данных этим допущениям. Даны общие

рекомендации по выбору отчетных периодов отбора статистических данных. Рассмотрен порядок определения расхождений между выборками.

2. Использование приведенных алгоритмов позволяет проверить нулевую гипотезу о соответствии статистических данных теоретическим моделям распределения Пуассона и экспоненциального распределения на основании критерия хи-квадрат Пирсона.

ЗАКЛЮЧЕНИЕ

Основные результаты, полученные в диссертационной работе, опубликованы в работах [88-96] и заключаются в следующем.

1. Разработана математическая модель замкнутой СМО с учетом трех типов поступающих в систему заявок и ограничения на время ожидания. В общем виде получены формулы для расчета характеристик модели в стационарном режиме.

2. С использованием метода Монте-Карло и языка программирования Python разработана имитационная модель замкнутой СМО, учитывающая три типа поступающих в систему заявок и ограничение на время ожидания. Предложенная модель обеспечивает эффективный расчет вероятностных характеристик СМО в нестационарном режиме функционирования. Получены результаты расчета времени выхода СМО на квазистационарный режим.

3. Для частного случая, при котором вероятность полной загрузки системы пренебрежимо мала, во входном потоке отсутствуют «нетерпеливые» заявки и интенсивность обслуживания близка к интенсивности уходов «нетерпеливых» заявок, методом производящих функций получено решение системы дифференциальных уравнений Колмогорова. Сравнение аналитических результатов для частного случая с данными численного моделирования методом Монте-Карло показало их высокую схожесть, что подтвердило адекватность предложенной модели для описания динамики функционирования систем с аналогичной структурой в нестационарном режиме.

4. Разработаны численные алгоритмы исследования СМО на предмет определения параметров модели, обеспечивающих требуемый уровень эффективности функционирования системы. Проведена количественная оценка требуемых минимальных мощностей системы для выхода на эффективный режим работы в случае двухкомпонентного и

трехкомпонентного потока заявок. Проведенный анализ эффективных режимов работы по параметру относительной пропускной способности системы показал, что при добавлении «нетерпеливых» заявок в двухкомпонентный поток зависимость интенсивности обслуживания от суммарной интенсивности входного потока меняется со степенной на линейную.

5. Разработаны рекомендации для применения и адаптации полученной модели к реальным объектам.

6. Разработан комплекс специализированных программных средств для расчета ключевых характеристик СМО и оценки эффективности ее функционирования в стационарном и нестационарном режимах.

Рекомендуется полученные формулы для расчета характеристик системы брать за основу для дальнейших исследований сложных стохастических процессов, включая анализ квазистационарных режимов и нахождение параметров, обеспечивающих эффективные режимы работы систем.

Перспективы развития темы заключаются в применении полученных результатов в качестве универсального инструментария для оценки эффективности функционирования систем с разнородными типами заявок, в которых имеют место уходы заявок из очереди и ограниченность источника, что позволит производить точную настройку в системах с циклическим характером функционирования, обеспечивая баланс между качеством обслуживания и ресурсозатратами.

СПИСОК ЛИТЕРАТУРЫ

1. Li, H., Ge, Y., Duan, Y. et al. Collaborative optimization of signals and ecological driving speed guidance for buses without dedicated bus lanes in the connected environment // *Scientific Reports*. – 2024. – Vol. 14. – P. 32.
2. Гнеденко Б. В., Коваленко И. Н. Введение в теорию массового обслуживания. – 2-е изд., перераб. и доп. – М.: Наука, 1987. – 336 с.
3. Кирпичников, А. П. Методы прикладной теории массового обслуживания. – Казань: Издательство Казанского университета, 2011. – С. 177–188.
4. Кирпичников А. П. Прикладная теория массового обслуживания. – Казань: Издательство КГУ, 2008. – 112 с.
5. Саати Т. Л. Элементы теории массового обслуживания и её приложения / Т. Л. Саати ; пер. с англ. под ред. И. Н. Коваленко. – 2-е изд. – М.: Советское радио, 1971. – 520 с.
6. Клейнрок Л. Теория массового обслуживания / Л. Клейнрок. – М.: Мир, 1979. – 432 с.
7. Хинчин А. Я. Математические методы теории массового обслуживания / А. Я. Хинчин // Труды Математического института имени В. А. Стеклова АН СССР. – М.: Издательство АН СССР, 1955. – Т. 49. – 400 с.
8. Хинчин А. Я. Потоки случайных требований без последствия // Теория вероятностей и её применения. – 1956. – Т. 1, вып. 1. – С. 3–18.
9. Хинчин А. Я. О пуассоновских потоках случайных событий // Теория вероятностей и её применения. – 1956. – Т. 1, вып. 3. – С. 320–327.
10. Монсик В. Б. Система массового обслуживания неделимых групповых заявок с очередью неограниченной длины / В. Б. Монсик, А. А. Скрынников, А. Ю. Федотов // Научный вестник Московского государственного технического университета гражданской авиации. – 2012. – № 184. – С. 108–112.

11. Тарасов В. Н. Система массового обслуживания E2/M/1 с обычными и сдвинутыми входными распределениями / В. Н. Тарасов, Н. Ф. Бахарева, Э. Г. Ахметшина // Инфокоммуникационные технологии. – 2018. – Т. 16, № 4. – С. 387–393.
12. Тарасов В. Н. Модель задержки на основе сдвинутых гиперэкспоненциального и эрланговского распределений / В. Н. Тарасов, Н. Ф. Бахарева // Физика волновых процессов и радиотехнические системы. – 2022. – Т. 25, № 1. – С. 21–26.
13. Konovalov M. G., Razumchik R. V. Numerical analysis of improved access restriction algorithms in a GI/G/1/N system // Journal of Communications Technology and Electronics. – 2018. – Vol. 63. – P. 616–625.
14. Коновалов М. Г., Разумчик Р. В. Comparison of two active queue management schemes through the M/D/1/N queue // Информатика и её применения. – 2018. – Т. 12, № 4. – С. 9–15.
15. Гильмутдинов Р. Ф., Кирпичников А. П. Математическая модель замкнутой одноканальной системы массового обслуживания // Вестник Казанского государственного технологического университета – Казань: Издательство Казан. гос. технол. ун-та, 2012. – № 6 – С.189–194.
16. Гильмутдинов, Р. Ф. Модель многоканальной замкнутой системы массового обслуживания / Р. Ф. Гильмутдинов, А. П. Кирпичников // Современные инструментальные системы, информационные технологии и инновации : Сборник научных трудов XII-ой Международной научно-практической конференции, Курск, 19–20 марта 2015 года / Том 1 – 2015.
17. Гильмутдинов Р. Ф. Замкнутые системы массового обслуживания с различным количеством обслуживающих устройств / Р. Ф. Гильмутдинов, А. П. Кирпичников // За нами будущее: взгляд молодых ученых на инновационное развитие общества : сборник научных статей Всероссийской молодежной научной конференции: в 4 томах, Курск, 05 июня 2020 года. Том 2. – Курск: Юго-Западный государственный университет, 2020. – С. 334–337.

18. Гильмутдинов Р. Ф., Кирпичников А. П. Замкнутые модели систем массового обслуживания с ограничениями // Вестник Казанского государственного технологического университета - Казань: Издательство Казан. гос. технол. ун-та, 2006. – № 4 – С. 220–224.

19. Kendall D. G. Stochastic Processes Occurring in the Theory of Queues and Their Analysis by the Method of the Imbedded Markov Chain // Annals of Mathematical Statistics. – 1953. – Vol. 24. – P. 338–354.

20. Erlang A.K. The Theory of Probabilities and Telephone Conversations // Nyt Tidsskrift for Matematik Afd. B. – 1909. – Vol. 20. – P. 33–39.

21. Erlang A.K. Solution of Some Problems in the Theory of Probabilities of Significance in Automatic Telephone Exchanges // Elektroteknikeren. – 1917. Vol.10 – P. 189–197.

22. Erlang A.K. Telephone waiting times // Nyt Tidsskrift for Matematik Afd. B. – 1920. – Vol. 31. – P. 25–42.

23. Cohen J. W., Harkema P. A study of the delay encountered in telegraph timerelay switching // Commun. News – 1954. – Vol. 15. – P. 47.

24. Cohen J. W. Berechnung der verkehrsgrossen in wartezeitsystem aus den verkehrsgrossen eines verlustsystems // NTZ – Nachrtech. Z. – 1955. – Vol. 8. – P. 139.

25. Cohen J. W. Das warteproblem fur das vollkommene bundle mit einer end lichen quellenzahl // NTZ – Nachrtech. Z. – 1955. – Vol. 8. – P. 641–645.

26. Cohen J. W. Some examples in the use of implication in switching algebra // Commun. News. – 1955. – Vol. 16. – № 1. – P. 1–10.

27. Cohen J. W. On the queueing process of lanes // Philips Tech. Rept. – 1956.

28. Cohen J. W. Certain delay problems for a full availability trunk group load- ed by two sources // Communication News. – 1956. – Vol. 16. – № 3. – P. 105–113.

29. Cohen J. W. Basic problems of telephone traffic theory and the influence of repeated calls // Philips Telecommun. Rev. – 1957. – Vol. 18. – № 2. – P. 49–100.

30. Cohen J. W. The full availability group of trunks with an arbitrary distribution of the interarrival times and a negative exponential holding time distribution // Natuurk. Tydschr. (Ghent) – 1957. – Vol. 26. – № 4. – P. 169–181.

31. Cohen J. W. A survey of queueing problems occurring in telephone and telegraph traffic theory // In Proc. 1st Int. Conf. Operat. Res., Oxford: English Universities Press, 1957. – P. 138–146.

32. Cohen J. W. The generalized Engset formulae // Philips Telecommun. Rev. – 1957. – Vol. 18. – № 4. – P. 158–170.

33. Cohen J. W. On the fundamental problem of telephone traffic theory and the influence of repeated calls // Philips Telecommun. Rev. – 1957. – Vol. 18. – P. 49–100.

34. Cohen J. W., Beukelman B.J. Call congestion of transposed multiples, Philips Telecommun. Rev. – 1957. – Vol. 17. – № 4. – P. 145–154.

35. Нгуен Тхань Банг. Системы массового обслуживания с конечным объемом накопителя и ограниченным средним временем нахождения требований в очереди: дис. ... канд. физ.-мат. наук. – Казань, 2019. – 129 с.

36. Валеев И. Н. Характеристики многоканальных систем селективного массового обслуживания с поликомпонентным входным потоком заявок : автореф. дис. канд. техн. наук. – Казань, 2011. – 24 с.

37. Титовцев А. С. Системы массового обслуживания поликомпонентных потоков с произвольным количеством источников и ограничений по длине очереди : автореф. дис. докт. техн. наук. – Казань, 2020. – 24 с.

38. Смирнова Н. А. Имитационная модель оптимизации управления производством строительных материалов : дис. канд. техн. наук. – СПб., 2002. – 219 с.

39. Чаплыгин В. В. Математические методы и алгоритмы расчета некоторых немарковских моделей массового обслуживания : автореф. дис. канд. техн. наук. – 2005. – 24 с.

40. Алиев Т. И. Трехмоментная аппроксимация вероятностных распределений в моделях массового обслуживания / Т. И. Алиев // Научно-технический вестник информационных технологий, механики и оптики. – 2014. – № 2(90). – С. 107–110.

41. Алиев Т. И. Погрешности моделирования высоконагруженных систем в GPSS WORLD / Т. И. Алиев // Научно-технический вестник информационных технологий, механики и оптики. – 2013. – № 1(83). – С. 70–75.

42. Алиев Т. И. Приоритетные стратегии управления трафиком в мультисервисных компьютерных сетях / Т. И. Алиев, Л. А. Муравьева-Витковская // Известия высших учебных заведений. Приборостроение. – 2011. – Т. 54, № 6. – С. 44–48.

43. Алиев Т. И. Распределение приоритетов в системах с комбинированными ограничениями / Т. И. Алиев // Вестник компьютерных и информационных технологий. – 2017. – № 4(154). – С. 3–7.

44. Алиев Т. И. Проектирование систем с приоритетами / Т. И. Алиев // Известия высших учебных заведений. Приборостроение. – 2014. – Т. 57, № 4. – С. 30–35.

45. Антонова П. В. Характеристики нестационарного режима систем массового обслуживания смешанного типа с ограниченным временем пребывания заявки в очереди / П. В. Антонова, А. С. Титовцев // Научно-технический вестник Поволжья. – 2021. – № 12. – С. 171–173.

46. Титовцев А. С. Численное исследование стабильных режимов работы систем массового обслуживания с ограниченным средним временем пребывания заявки в системе и временем обслуживания, распределенным по закону Эрланга / А. С. Титовцев, П. В. Антонова // Научно-технический вестник Поволжья. – 2021. – № 8. – С. 83–85.

47. Задорожный В. Н. Исследование эффективности приоритетных дисциплин на основе метамодели многопоточковых систем массового обслуживания / В. Н. Задорожный, Д. А. Тулубаев // Омский научный вестник. – 2012. – № 2(110). – С. 255–258.

48. Задорожный В. Н. Методы моделирования систем массового обслуживания с тяжелыми хвостами распределений / В. Н. Задорожный, Т. Р. Захаренкова // Информационные технологии и математическое моделирование (ИТММ-2016) : Материалы XV Международной конференции имени А. Ф. Терпугова, Алтайский край, Алтайский р-он, пос. Катунь, 12–16 сентября 2016 года. Т. 1 Часть 1. – Алтайский край, Алтайский р-он, пос. Катунь: Национальный исследовательский Томский государственный университет, 2016. – С. 56–61.

49. Задорожный В. Н. Оптимизация распределения ресурсов и переходных вероятностей в немарковских сетях массового обслуживания / В. Н. Задорожный // Информационные технологии и математическое моделирование (ИТММ-2016) : Материалы XV Международной конференции имени А. Ф. Терпугова, Алтайский край, Алтайский р-он, пос. Катунь, 12–16 сентября 2016 года. Т. 1 Часть 1. – Алтайский край, Алтайский р-он, пос. Катунь: Национальный исследовательский Томский государственный университет, 2016. – С. 61–66.

50. Моисеева С. П. Исследование бесконечнолинейной системы массового обслуживания с разнотипным обслуживанием и входящим потоком марковского восстановления / С. П. Моисеева, Е. В. Панкратова, Е. Г. Убонова // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. – 2016. – № 2(35). – С. 46–53.

51. Моисеева С. П. Распределение вероятностей двумерного потока обращений в бесконечнолинейной системе массового обслуживания с повторным обращением / С. П. Моисеева, А. С. Морозова, А. А. Назаров //

Вестник Томского государственного университета. – 2006. – № S16.
– С. 125–128.

52. Моисеев А. Н. Исследование математических моделей систем и сетей массового обслуживания с высокоинтенсивными непуассоновскими входящими потоками : специальность 05.13.18 "Математическое моделирование, численные методы и комплексы программ" : диссертация на соискание ученой степени доктора физико-математических наук / Моисеев Александр Николаевич, 2016. – 333 с.

53. Моисеев А. Н. Асимптотический анализ многофазной системы массового обслуживания с высокоинтенсивным рекуррентным входящим потоком / А. Н. Моисеев, А. А. Назаров // Автометрия. – 2014. – Т. 50, № 2. – С. 67–76.

54. Назаров А. А. Асимптотический анализ систем массового обслуживания с неограниченным числом приборов и полумарковским входящим потоком / А. А. Назаров, И. А. Семенова // Известия Томского политехнического университета. – 2012. – Т. 320, № 5. – С. 12–17.

55. Романенко В. А. Оптимизация управления технологическими процессами узлового аэропорта как системы массового обслуживания с нестационарными потоками и частичной взаимопомощью каналов / В. А. Романенко // Управление большими системами: сборник трудов. – 2012. – № 36. – С. 209–247.

56. Рыков В. В. Основы теории массового обслуживания (Основной курс: марковские модели, методы марковизации) : Учебное пособие / В. В. Рыков, Д. В. Козырев. – Москва: ООО «Научно-издательский центр ИНФРА-М», 2016. – 223 с.

57. Самуйлов К. Е. К анализу стационарных характеристик системы массового обслуживания со случайными требованиями / К. Е. Самуйлов, Э. С. Сопин, О. Г. Вихрова // Вероятностные методы в дискретной математике. IX Международная Петрозаводская конференция : расширенные тезисы, Петрозаводск, 30 мая–03 июня 2016 года / Институт прикладных

математических исследований Карельского научного центра Российской академии наук ; науч. ред. В. В. Мазалов ; отв. ред. Е. Н. Спектор. – Петрозаводск: Петрозаводский государственный университет, 2016. – С. 87–89.

58. Сущенко С. П. Моделирование разделяемой памяти двухпроцессорной вычислительной системы / С. П. Сущенко, М. С. Сущенко, Д. В. Биматов // Вестник Томского государственного университета. – 2003. – № 280. – С. 319–323.

59. Фархадов М. П. Принципы построения автоматизированных систем массового обслуживания нового поколения (на примере компьютерных систем резервирования) : специальность 05.13.13 : диссертация на соискание ученой степени кандидата технических наук / Фархадов Маис Паша оглы. – Москва: 1999. – 187 с.

60. Валеев И. Н. Многоканальная система массового обслуживания с отказами / И. Н. Валеев, А. П. Кирпичников // Вестник Казанского технологического университета. – 2006. – № 4. – С. 75–79.

61. Кирпичников А. П., Титовцев А. С. Системы обслуживания с неоднородным входным потоком требований, отказами и очередью // Вестник Казанского государственного технологического университета – Казань: Издательство Казан. гос. технол. ун-та, 2011. – Т. 14. – № 5. – С. 154–161.

62. Смирнов А. В. Векторные односвязные цепи Маркова / А. В. Смирнов // Математические методы в технике и технологиях. – 2003. – Т. 1. – № 3. – С. 102–105.

63. Писарева Н. Д. Марковские цепи в исследовании суточного слоя осадков // Математика и ее приложения в современной науке и практике : Сборник научных статей VIII Международной научно-практической конференции, Курск, 27–28 апреля 2018 года / Ответственный редактор Е. А. Бойцова. – Курск: Юго-Западный государственный университет, 2018. – С. 92–98.

64. Лабскер Л. Г., Бабешко Л. О. Теория массового обслуживания в экономической сфере: учебное пособие для вузов. – М.: Банки и биржи, ЮНИТИ, 1998. – 319 с.

65. Шведов А. С. О методах Монте-Карло с цепями Маркова // Экономический журнал ВШЭ. – 2010. – Т. 14. – № 2. – С. 243–277.

66. Данилова Н. В. Расчёт интервала справедливых цен для бинарной модели (b,s)-рынка с волатильностью, являющейся марковской цепью / Н. В. Данилова // Известия высших учебных заведений. Северо-Кавказский регион. Серия: Естественные науки. – 2016. – № 4(192). – С. 17–20.

67. Домбровский В. В. Управление с прогнозированием взаимосвязанными гибридными системами с марковскими скачками при ограничениях / В. В. Домбровский, Т. Ю. Обьедко // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. – 2012. – № 3(20). – С. 5–12.

68. Цимбал В. А. Основы теории чувствительности конечных марковских цепей / В. А. Цимбал // Проблемы обеспечения эффективности и устойчивости функционирования сложных технических систем, Серпухов, 18–19 июня 2000 года. Часть 3. – Серпухов: Серпуховской военный институт ракетных войск, 2000. – С. 44–50.

69. Подлегаев А. В. Нахождение характеристик оперативности взаимодействия оператора в человеко-машинной системе на основе конечных марковских цепей / А. В. Подлегаев, С. И. Рысьев // Информационные технологии в проектировании и производстве. – 2008. – № 1. – С. 66–68.

70. Магазев А. А. Оценка среднего времени до отказа безопасности на основе марковских цепей с непрерывным временем / А. А. Магазев, А. С. Мельникова, В. Ф. Цырульник // Математические структуры и моделирование. – 2020. – № 4(56). – С. 112–125.

71. Фаворов А. В. Поиск участков специфического связывания белков-регуляторов транскрипции с ДНК методом Монте-Карло Марковскими цепями : специальность 03.00.02, автореферат диссертации на соискание

ученой степени кандидата физико-математических наук / А. В. Фаворов. – Москва: Московский государственный университет им. М.В. Ломоносова, 2005. – 23 с.

72. Юрчук А. А. Прогнозирование респираторных вирусных заболеваний с использованием скрытых марковских цепей / А. А. Юрчук // Международный научный журнал. – 2016. – № 7. – С. 37–39.

73. Зуев В. С. Стохастическая и детерминированная модели распространения эпидемий типа SIRS / В. С. Зуев, Н. С. Шишмакова, И. С. Петрова // Дневник науки. – 2020. – № 1(37). – С. 24.

74. Марасанова Т. П. Оптимизация бизнес-процессов предприятия путем анализа СМО с приоритетной очередью / Т. П. Марасанова // Формирование конкурентной среды, конкурентоспособность и стратегическое управление предприятиями, организациями и регионами: Сборник статей V Международной научно-практической конференции, Пенза, 11–12 мая 2020 года. – Пенза: Пензенский государственный аграрный университет, 2020. – С. 116–119.

75. Павлюченко К. Г. Оптимизация работы торгового предприятия на примере магазина шаговой доступности. Многоканальная СМО с ограниченной очередью / К. Г. Павлюченко, В. П. Пивоваров, Т. В. Богачев // Экономика. Социология. Право: материалы международной научно-практической конференции, Саратов, 21 октября 2016 года / Институт управления и социально-экономического развития, Саратовский государственный технический университет. – Саратов: ООО "Центр профессионального менеджмента "Академия Бизнеса", 2016. – С. 3–6.

76. Княжева Ю. В. Повышение эффективности системы массового обслуживания торгового предприятия посредством численного статистического моделирования / Ю. В. Княжева // Вестник Новосибирского государственного университета. Серия: Социально-экономические науки. – 2014. – Т. 14, № 2. – С. 83–100.

77. Калашников В. В., Рачев С. Т. Математические методы построения стохастических моделей обслуживания. – М.: Наука, 1988. – 310 с.

78. Бочарникова В. С. Моделирование систем массового обслуживания на основе метода монте-карло / В. С. Бочарникова, Л. Ф. Фатхуллин, Э. Н. Мифтахов // Математическое моделирование процессов и систем : Материалы VII Международной молодежной научно-практической конференции, Уфа, 07–09 декабря 2017 года / Ответственный редактор С. А. Мустафина. Часть I. – Уфа: Стерлитамакский филиал ФГБОУ ВО "Башкирский государственный университет", 2017. – С. 168–170.

79. Бородина А. В. Имитационное моделирование неоднородного процесса деградации в системе с постепенными и внезапными отказами / А. В. Бородина, В. А. Тищенко // Труды Карельского научного центра Российской академии наук. – 2018. – № 7. – С. 3–13.

80. Килюшева Е. Д. Массовое обслуживание в маршрутизирующих устройствах / Е. Д. Килюшева, П. Д. Лебедев, В. М. Шек // Научный вестник Московского государственного горного университета. – 2011. – № 6. – С. 32–43.

81. Назаров А. А., Терпугов А. Ф. Теория массового обслуживания: учебное пособие. – 2-е изд., испр. – Томск: Издательство НТЛ, 2010. – 228 с.

82. Титовцев А. С. Численно-аналитическое моделирование систем массового обслуживания поликомпонентных потоков / А. С. Титовцев, А. П. Кирпичников // Математические методы в технике и технологиях – ММТТ. – 2019. – Т. 2. – С. 80–84.

83. Плескунов М. А. Теория массового обслуживания : учебное пособие / М. А. Плескунов ; Министерство науки и высшего образования РФ, Уральский федеральный университет. – Екатеринбург: Издательство Урал. ун-та, 2022. – 264 с.

84. Крохалев В. Я. Статистика : учебное пособие / В. Я. Крохалев, С. А. Скопинов, В. А. Телешев; ФГБОУ ВО УГМУ Минздрава России. – Екатеринбург: Издательство УГМУ, 2018. – 114 с.

85. Гмурман В. Е. Теория вероятностей и математическая статистика: учебник для вузов / В. Е. Гмурман. – М.: Юрайт, 2020. – 479 с.

86. Богданова Р. А. Оценка данных на соответствие закону нормального распределения критерием Хи-квадрат в программе STATISTICA / Р. А. Богданова, Н. М. Пак // Информация и образование: границы коммуникаций. – 2022. – № 14(22). – С. 166–170.

87. Кирпичников А. П., Банг Нгуен Тхань, Куи Чан Куанг. Суммарное число требований, находящихся в системе массового обслуживания с ограниченным средним временем пребывания заявки в очереди // Вестник Казанского Технологического университета. – 2017. – Т. 20, № 9. – С. 104–107.

88. Хасанов Н. А. Имитационное моделирование нестационарного режима в замкнутой системе массового обслуживания с поликомпонентным потоком заявок // Научно-технический вестник Поволжья. – 2025. – № 10. – С. 64–67.

89. Хасанов Н. А., Титовцев А. С., Петров Т. Э. Трехкомпонентный поток заявок в замкнутых системах массового обслуживания с бесконечной емкостью накопителя и ограничением по времени ожидания // Инженерный вестник Дона. – 2024. – № 10(118). – С. 156–163.

90. Хасанов Н. А., Титовцев А. С., Петров Т. Э. Математическая модель замкнутых систем обслуживания с поликомпонентным потоком заявок, ожиданием и потерями // Научно-технический вестник Поволжья. – 2024. – № 10. – С. 37–40.

91. Петров Т. Э., Титовцев А. С., Хасанов Н. А. Вероятностные характеристики замкнутых систем обслуживания с ожиданием, отказами и ограничением по длине очереди // Научно-технический вестник Поволжья. – 2023. – № 12. – С. 101–102.

92. Петров Т. Э., Хасанов Н. А., Титовцев А. С. Вероятностные характеристики замкнутых систем массового обслуживания с ожиданием, отказами и ограничением по длине очереди // Школа теории массового обслуживания (ШТМО-2025): тезисы I международной конференции (21–26 апреля 2025 г.). – Томск: Издательство Томского государственного университета, 2025. – С. 42–43.

93. Хасанов Н. А., Петров Т. Э., Титовцев А. С. Вероятностные характеристики замкнутых систем массового обслуживания с трехкомпонентным потоком заявок и ограничением на время ожидания заявки в очереди // Школа теории массового обслуживания (ШТМО-2025): тезисы I международной конференции (21–26 апреля 2025 г.). – Томск: Издательство Томского государственного университета, 2025. – С. 59–60.

94. Хасанов Н. А. Программа для расчета характеристик замкнутых СМО // Навигатор в мире науки и образования. – 2025. – № 1(66). – С. 13–18.

95. Хасанов Н. А. Свидетельство о регистрации электронного ресурса № 25470 «Программа для расчета характеристик замкнутых СМО». – М.: ОФЭРНиО, 2025.

96. Хасанов Н. А. Замкнутые системы массового обслуживания с трехкомпонентным потоком заявок и ограничением на время ожидания заявки в очереди // Математические методы в технологиях и технике. – 2025. – № 10. – С. 95–100.

ПРИЛОЖЕНИЯ

**Пример оценки соответствия входного потока распределению Пуассона и
времени обслуживания экспоненциальному распределению**

1. Рассмотрим реализацию алгоритма проверки соответствия входного потока заявок 1-го типа распределению Пуассона. Отчетный период составляет 3 дня. Пример собранных статистических данных приведен в таблице 1.

Таблица 1 – Временные метки поступления заявок 1-го типа в систему

Порядковый номер поступившей заявки	Выборка № 1	Выборка № 2	Выборка № 3
	Время поступления		
1	12:20	8:52	12:12
2	13:14	10:58	19:18
3	14:23	15:34	19:44
4	18:30	16:27	19:48
5	11:30	17:38	11:50
6	14:38	8:24	12:22
7	18:41	12:50	13:16
8	10:06	15:17	14:30
9	15:02	12:05	14:32
10	15:54	15:28	15:03
11	17:40	18:02	18:17
12	19:44	8:12	18:21
13	-	8:21	18:53
14	-	11:54	9:19
15	-	15:15	9:26
16	-	16:24	11:22
17	-	16:55	13:45
18	-	17:34	18:22

Порядковый номер поступившей заявки	Выборка № 1	Выборка № 2	Выборка № 3
	Время поступления		
19	-	19:12	19:38
20	-	19:48	-

Разделим полученный временной интервал на сегменты по 2 часа и определим количество входящих заявок в каждом временном сегменте.

Таблица 2 – Интенсивность входящего потока 1-го типа за отчетный период на интервалах $t = 2$ часа

№ п/п	Временной интервал		Выборка № 1	Выборка № 2	Выборка № 3
1	8:00	10:00	0	1	0
2	10:00	12:00	0	1	0
3	12:00	14:00	2	0	1
4	14:00	16:00	1	1	0
5	16:00	18:00	0	2	0
6	18:00	20:00	1	0	3
7	8:00	10:00	0	1	0
8	10:00	12:00	1	0	1
9	12:00	14:00	0	2	2
10	14:00	16:00	1	2	3
11	16:00	18:00	0	0	0
12	18:00	20:00	1	1	3
13	8:00	10:00	0	2	2
14	10:00	12:00	1	1	1
15	12:00	14:00	0	0	1
16	14:00	16:00	2	1	0
17	16:00	18:00	1	3	0
18	18:00	20:00	1	2	2

Результаты расчета относительного стандартного отклонения между выборками σ и интенсивности входящего потока одной заявки λ_1 для общего количества заявок в источнике ($N=11$) и $t = 1$ ч, приведены в таблице 3.

Таблица 3

Характеристика	Расчетная формула	Значение		
		Выборка № 1	Выборка № 2	Выборка № 3
Q_j	(6.1)	0,666	1,111	1,055
λ_1	(6.7)	0,035	0,051	0,048
\bar{Q}	(6.2)	0,944		
$\sigma, \%$	(6.3)	25,6		

Таким образом, относительное стандартное отклонение в количестве заявок между выборками составило 25,6 %. Так как результат интенсивности Q_j в выборке № 3 ближе всего к среднему значению, она используется в дальнейших расчетах как референтная.

Отсортируем выборку № 3 по возрастанию количества поступающих в систему заявок и получим следующий числовой ряд: 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 2, 2, 2, 3, 3, 3.

Определим эмпирическую частоту $Y_e(k)$ появления каждого из элементов ряда k в выборке № 3.

$k=0$, частота появления 8;

$k=1$, частота появления 4;

$k=2$, частота появления 3;

$k=3$, частота появления 3.

Определим теоретическое значение функции вероятности для полученного количества событий k и рассчитанного значения $Q_j=1,055$ по (6.4):

$$p(0) = \frac{1,055^0}{0!} e^{-1,055} = 0,3482 \%,$$

$$p(1) = \frac{1,055^1}{1!} e^{-1,055} = 0,3673 \%,$$

$$p(2) = \frac{1,055^2}{2!} e^{-1,055} = 0,1938 \%,$$

$$p(3) = \frac{1,055^3}{3!} e^{-1,055} = 0,0681 \%.$$

Определим теоретическую частоту $Y_m(k)$ появления каждого из элементов ряда k в выборке № 3:

$$Y_m(0) = 0,3482 \cdot 19 = 6,6157,$$

$$Y_m(1) = 0,3673 \cdot 19 = 6,9795,$$

$$Y_m(2) = 0,1938 \cdot 19 = 3,6817,$$

$$Y_m(3) = 0,0681 \cdot 19 = 1,2927.$$

Проведем оценку значимости различий между полученными результатами используя критерий хи-квадрат Пирсона. Для этого рассчитаем параметр χ^2 по (6.5):

$$\chi^2 = \sum_k \frac{(Y_o(k) - Y_m(k))^2}{Y_m(k)} = 3,9338.$$

Количество степеней свободы для теста хи-квадрат рассчитывается по (6.6):

$$df = 4 - 1 - 1 = 2.$$

Для уровня значимости $\alpha = 0.05$ и количества степеней свободы 2, критическое значение $\chi_{крит}^2$ из таблицы хи-квадрат составляет 5,991. Так как полученное значение χ^2 меньше $\chi_{крит}^2$, полученная выборка может использоваться для моделирования работы СМО.

2. Рассмотрим пример проверки соответствия времени обслуживания экспоненциальному распределению. Алгоритм оценки погрешности между выборками в данном случае идентичен рассмотренному выше. В связи с чем, для времени обслуживания, будем считать, что референтная выборка уже

определена и отчетный период составляет 3 дня. Пример собранных статистических данных приведен в таблице 4.

Таблица 4 – Временные метки начала и конца обслуживания

№ п/п	Время начала обслуживания	Время окончания обслуживания	Временной интервал, x , мин
1	8:40	17:00	500
2	17:00	10:10	310
3	10:10	10:50	40
4	10:50	14:00	190
5	14:00	18:20	260
6	18:20	15:00	520
7	15:00	15:30	30
8	15:30	16:10	40
9	16:10	16:55	45

В соответствии с таблицей 4 среднее время обслуживания \bar{x} составляет 215 мин. Средняя интенсивность по выборке, рассчитанная в соответствии с (6.8) составляет 0,0047.

Разделим совокупность полученных статистических данных на равные группы по $t=60$ мин и найдем количество попавших в полученные интервалы значений времени обслуживания x , мин.

Таблица 5 – Группировка данных по временным интервалам

№ п/п	Временные интервалы t , мин.	$Y_s(t)$
1	[30;90)	4
2	[90;150)	0
3	[150;210)	1
4	[210;270)	1
5	[270;330)	1
6	[330;390)	0
7	[390;450)	0
8	[450;510)	1
9	[510;570)	1

В соответствии с (6.9) определим теоретические частоты для экспоненциального распределения:

$$\begin{aligned} Y_m(1) &= 9 \cdot (e^{-0,0047 \cdot 30} - e^{-0,0047 \cdot 90}) = 1,92, \\ Y_m(2) &= 9 \cdot (e^{-0,0047 \cdot 90} - e^{-0,0047 \cdot 150}) = 1,45, \\ Y_m(3) &= 9 \cdot (e^{-0,0047 \cdot 150} - e^{-0,0047 \cdot 210}) = 1,09, \\ Y_m(4) &= 9 \cdot (e^{-0,0047 \cdot 210} - e^{-0,0047 \cdot 270}) = 0,82, \\ Y_m(5) &= 9 \cdot (e^{-0,0047 \cdot 270} - e^{-0,0047 \cdot 330}) = 0,62, \\ Y_m(6) &= 9 \cdot (e^{-0,0047 \cdot 330} - e^{-0,0047 \cdot 390}) = 0,47, \\ Y_m(7) &= 9 \cdot (e^{-0,0047 \cdot 390} - e^{-0,0047 \cdot 450}) = 0,35, \\ Y_m(8) &= 9 \cdot (e^{-0,0047 \cdot 450} - e^{-0,0047 \cdot 510}) = 0,27, \\ Y_m(9) &= 9 \cdot (e^{-0,0047 \cdot 510} - e^{-0,0047 \cdot 570}) = 0,20, \end{aligned}$$

Таким образом, в соответствии с (6.5):

$$\chi^2 = \sum_k \frac{(Y_o(k) - Y_m(k))^2}{Y_m(k)} = 9,984.$$

Число степеней свободы:

$$df = 9 - 1 - 1 = 7.$$

Критическое значение $\chi_{крит}^2$, полученное из таблицы хи-квадрат для уровня значимости $\alpha=0,05$, составляет 14,068. Так как $\chi^2 < \chi_{крит}^2$, нулевая гипотеза о соответствии данных экспоненциальному распределению не отвергается и результаты выборки могут быть использованы в качестве исходных данных для расчета.

3. Приведем размерность полученных интенсивностей к одному временному промежутку. Так, $\lambda_1=0,048$ [заявок в час], приводя параметр интенсивности обслуживания к той же размерности, получим: $\mu=0,00465$ [заявок в 1 минуту] или $\mu = 0,28$ [заявок в час].



ООО НПФ «ГЕОНИК»

422594, Республика Татарстан, р-н Верхнеуслонский,

с. Набережные Мокваши, ул. Покровская, зд. 5

ИНН 1658053275 ОГРН 1041626802128

«__» _____ 2025 г.

АКТ О ВНЕДРЕНИИ

Настоящим актом подтверждается, что результаты диссертационной работы Хасанова Н.А. на соискание ученой степени кандидата технических наук «Замкнутые системы массового обслуживания с трехкомпонентным потоком заявок и ограничением по времени ожидания заявки в очереди» использовались в ООО НПФ «ГЕОНИК» для минимизации задержек в обслуживании производственного оборудования.

Представленная в диссертационной работе математическая модель системы массового обслуживания с наличием «нетерпеливых» заявок и учетом временных ограничений позволяет достаточно достоверно описать процессы технического обслуживания и ремонта производственного оборудования. Применимость модели обоснована ее способностью отражать ключевые особенности процессов: разнородность заявок и фиксированное число требований. На основании проведенного статистического анализа исторических данных о поломках и заявках на обслуживание оборудования допущение о пуассоновском характере входного потока оказалось корректным. Внедрение математической модели позволило получить достоверную оценку загрузки сервисных ресурсов и выявить узкие места в процессах обслуживания оборудования, что создало основу для повышения стабильности производственных операций и снижения рисков аварийных остановок.

Генеральный директор
ООО НПФ «ГЕОНИК»



Ясаков А.Н.