

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

На правах рукописи



ШАЙДУЛЛИНА НАТАЛЬЯ КОНСТАНТИНОВНА

**МОДЕЛИРОВАНИЕ СИСТЕМ МАССОВОГО
ОБСЛУЖИВАНИЯ С ОГРАНИЧЕННЫМ ВРЕМЕНЕМ
ЖИЗНИ ЗАЯВОК**

1.2.2. Математическое моделирование,
численные методы и комплексы программ

ДИССЕРТАЦИЯ
на соискание ученой степени
кандидата технических наук

Научный руководитель:
кандидат технических наук,
доктор педагогических наук,
профессор Нуриев Наиль Кашапович

Казань – 2026

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ	4
1 СИСТЕМЫ МАССОВОГО ОБСЛУЖИВАНИЯ С ОГРАНИЧЕННЫМ ВРЕМЕНЕМ ЖИЗНИ ЗАЯВОК	13
1.1 Развитие теории систем массового обслуживания	13
1.2 Классификация систем массового обслуживания	18
1.3 Особенности систем массового обслуживания с ограниченным временем жизни заявок	23
1.4 Имитационное моделирование систем массового обслуживания	28
2 СИСТЕМЫ МАССОВОГО ОБСЛУЖИВАНИЯ СО СЛУЧАЙНЫМ ВРЕМЕНЕМ ЖИЗНИ ЗАЯВОК	34
2.1 Математическое моделирование одноканальной системы с очередью	35
2.2 Построение имитационной модели одноканальной системы с очередью и ее верификация	44
3 СИСТЕМЫ МАССОВОГО ОБСЛУЖИВАНИЯ С ФИКСИРОВАННЫМ ОГРАНИЧЕНИЕМ ВРЕМЕНИ ЖИЗНИ ЗАЯВОК	48
3.1 Математическое моделирование системы обслуживания ординарного потока заявок с отказами	48
3.2 Имитационное моделирование систем массового обслуживания ординарного потока заявок	54
3.2.1 Системы обслуживания ординарного входного потока заявок с отказами	55
3.2.2 Системы обслуживания ординарного входного потока заявок с очередью	67
3.3 Имитационное моделирование систем массового обслуживания группового входного потока заявок	74
3.3.1 Системы обслуживания группового входного потока заявок с отказами	74
3.3.2 Системы обслуживания группового входного потока заявок с очередью	78

3.4 Сравнительный анализ результатов имитационного моделирования различных СМО	83
4 ПОВЫШЕНИЕ ПОКАЗАТЕЛЯ ЭКОНОМИЧЕСКОЙ ЭФФЕКТИВНОСТИ СИСТЕМ С ФИКСИРОВАННЫМ ОГРАНИЧЕНИЕМ ВРЕМЕНИ ЖИЗНИ ЗАЯВОК	89
4.1 Математическое моделирование системы обслуживания группового входного потока с очередью	89
4.2 Постановка и решение задачи повышения показателя экономической эффективности систем для детерминированного потока обслуживания.....	99
4.3 Постановка и решение задачи повышения показателя экономической эффективности систем для случайного потока обслуживания	109
ЗАКЛЮЧЕНИЕ	121
СПИСОК ЛИТЕРАТУРЫ.....	124
ПРИЛОЖЕНИЯ.....	140

ВВЕДЕНИЕ

Актуальность темы исследования. В работе рассмотрены системы массового обслуживания, особенностью которых является наличие ограничения на время жизни заявок. В современном мире, где информация является наиважнейшим ресурсом, а ее ценность прямо пропорциональна актуальности, такие системы имеют критически важное значение. Ограниченное время жизни является неотъемлемым атрибутом требований не только в области обработки информации, но и в других важных сферах деятельности человека. В медицине срок жизни заявки может быть обусловлен ограниченным периодом времени для возможности экстренных манипуляций. В технических системах время жизни требования может ограничивать зона действия обслуживающего прибора. В торговых операциях естественным ограничением является срок годности товара.

В концепции федеральной программы «Развитие промышленной робототехники и автоматизации производства» разработка и внедрение программных и аппаратных средств управления роботизированными системами заложены как необходимые компоненты системного подхода к формированию промышленной роботизации. Системы массового обслуживания могут быть использованы для построения моделей управления потоками в роботизированных системах, а сформированный инструментарий теории массового обслуживания позволит выполнить предпроектный анализ эффективности, запланировать минимальные резервные мощности для бесперебойной работы, сбалансировать загрузку, обеспечить гибкую перенастройку системы без остановки производства.

Существующие на сегодняшний день модели систем массового обслуживания с ограничениями времени имеют ряд особенностей, которые не позволяют предложить универсальные решения для систем разного типа. Большинство аналитических моделей построены для пуассоновского входного потока заявок, экспоненциального характера обслуживания и ограниченного времени ожидания обслуживания. Они созданы для систем с накопителями и

подразумевают возможность потери требования только в процессе ожидания обслуживания. В отличие от них, в системах с ограниченным временем жизни заявок утрата заявки может произойти как в процессе ожидания, если имеется опция накопления очереди, так и в процессе обработки. Кроме того, классические модели СМО не пригодны для анализа функционирования систем с групповым поступлением требований.

Как правило, целью моделирования СМО является получение функциональных характеристик системы для анализа эффективности ее работы. Зачастую, однако, для принятия управленческих решений стандартных показателей эффективности системы, таких как относительная пропускная способность, среднее время ожидания заявки в очереди на обслуживание, вероятность отказа в обслуживании и др. оказывается недостаточно. Для объективного анализа поведения системы и прогнозирования результатов ее работы при изменении входных или управляющих параметров требуется не только знание показателей эффективности функционирования системы, но и оценка влияния изменения этих параметров на ее эксплуатационные характеристики. Именно это позволяет оценить целесообразность введения дополнительных мощностей или исключения действующих для оптимизации производительности системы.

В контексте сказанного, актуальным является создание инструментов для получения характеристик СМО с ограниченным временем жизни заявок и анализа их взаимозависимости, которые будут универсальными для разных систем: с ожиданием или без него, одноканальных или многоканальных, с ординарным или групповым поступлением требований.

Степень разработанности темы исследования. Принято считать, что научные исследования в области теории массового обслуживания начались с работ датского инженера А.К. Эрланга. Существенно способствовали развитию теории очередей на заре ее становления такие известные ученые, как А.А. Марков, А.Н. Колмогоров, Ф. Поллачек, А. Хинчин. Большой вклад в развитие математического аппарата внесли советские ученые: Е.С. Вентцель, Б.В.

Гнеденко, И.Н. Коваленко, А.А. Боровков; и зарубежные: К. Пальм, Д. Кендалл, Т. Саати, Л. Такач, Л. Клейнрок, Д. Литтл.

Усложнение аналитических моделей постепенно становилось препятствием к их практическому применению, что обусловило возрастание интереса к численным решениям. Появились методы статистического моделирования, на основе которых строились имитационные модели. Среди основоположников этого направления следует выделить А.Л. Лифшица, Р. Шеннона, Дж. Кляйнена, Н.П. Бусленко. Прогресс в области вычислительной техники способствовал появлению компьютерных сетей. В области применения теории очередей для изучения компьютерных сетей наиболее известными на начальном этапе были имена Х. Тагаки, Л. Клейнрока, С.Ф. Яшкова, В.М. Вишневого.

В настоящее время исследование систем, моделируемых как СМО, проводят К.Е. Самуйлов, Ю.В. Гайдамака, А.Н. Дудин, А.А. Назаров, С.П. Моисеева, А.В. Печинкин, А.Н. Моисеев, В.В. Рыков, С.В. Пауль, В.Н. Тарасов, Н.Ф. Бахарева, Б.Я. Лихтциндер, А.Ю. Привалов, Н.К. Нуриев, Е.А. Печеный, А.В. Зорин, А.Л. Казаков, И.Е. Тананко и др.

Изучением СМО с ограниченным временем занимались зарубежные ученые: Д. Баррер, Дж. Коэн, Дж. Дейли; и отечественные: Е. С. Вентцель, И.Н. Коваленко, Л. Г. Афанасьева, А.П. Кирпичников, Ю.П. Рыжиков. Описанные исследователями аналитические модели СМО с ограниченным временем действуют только в условиях, для которых они получены: простейший или детерминированный потоки, постоянное ограничение на время ожидания обслуживания, один канал обслуживания и др. Некоторые из этих моделей имеют ограниченную практическую полезность ввиду сложности доведения построенных конструкций до количественных результатов. Естественным путем преодоления этих затруднений является имитационное моделирование.

Объектом исследования являются системы массового обслуживания с ограниченным временем жизни заявок.

Предметом исследования являются математические и имитационные модели систем массового обслуживания с ограниченным временем жизни заявок.

Соответствие 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; 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.

Цель исследования. Целью диссертационной работы является разработка математических моделей, алгоритмов и программных комплексов для анализа характеристик систем массового обслуживания с ограниченным временем жизни заявок и возможностей повышения эффективности их функционирования.

Для достижения поставленной цели решены следующие задачи:

– построена математическая модель системы массового обслуживания пуассоновского потока заявок с ограниченным временем жизни без опции накопления очереди; построена математическая модель одноканальной системы массового обслуживания группового потока заявок с ограниченным временем жизни с опцией накопления очереди;

– построены и реализованы программно имитационные модели для различных систем массового обслуживания с ограниченным временем жизни заявок: с ожиданием и без него, одноканальных и многоканальных, с ординарным и групповым поступлением требований;

- подтверждена адекватность разработанных имитационных моделей посредством сравнения результатов их работы с аналитическими решениями;
- выполнен анализ влияния характеристик систем массового обслуживания с ограниченным временем жизни заявок на их эксплуатационные характеристики;
- проведено исследование возможностей повышения эффективности функционирования систем массового обслуживания с ограниченным временем жизни заявок;
- осуществлена апробация полученных результатов.

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

1. *Разработана новая математическая модель системы массового обслуживания пуассоновского потока заявок с фиксированным ограничением времени жизни без опции накопления очереди, отличающаяся нестандартным подходом к построению, позволяющая использовать ее в инженерных расчетах.*

2. *Построена новая математическая модель одноканальной системы массового обслуживания группового входного потока заявок с ограниченным временем жизни с опцией накопления очереди, отличающаяся наличием нестандартного показателя эффективности функционирования, позволяющая находить время и величину управляющего воздействия на систему с целью повышения этого показателя (п.8 паспорта специальности 1.2.2).*

3. *На основе концептуальной модели одноканальной системы с ожиданием и случайным временем жизни заявок, разработанной Б.В. Гнеденко и И.Н. Коваленко, получено аналитическое решение задачи нахождения основных характеристик СМО для простейшего входного потока, экспоненциального времени обслуживания и равномерно распределенного времени жизни*

заявок, *позволяющее* верифицировать имитационную модель (п.8 паспорта специальности 1.2.2).

4. *Разработаны имитационные модели* систем массового обслуживания с ограниченным временем жизни заявок, *реализованные* на языке программирования С#, *позволяющие* находить характеристики разных систем: с ожиданием и без него, одноканальных и многоканальных, с ординарным и групповым поступлением требований (п.6 паспорта специальности 1.2.2).

5. *Установлено* существование порогового значения времени жизни заявок в СМО, *позволяющее* определить зону возможного воздействия на систему для изменения ее эффективности функционирования. *Выявлено*, что для СМО, обрабатывающих поток заявок с ограниченным временем жизни, доля заявок, обслуженных в полном объеме, с увеличением числа действующих каналов стремится к некоторому пороговому значению, которое практически не зависит от интенсивности входного потока, но зависит от времени жизни заявки и среднего времени обслуживания. *Это позволяет* выполнить предпроектную оценку эффективности системы. (пп.8, 9 паспорта специальности 1.2.2).

6. Для одноканальной СМО с групповым потоком заявок с ограниченным временем жизни и случайным потоком обслуживания *найдены* время и величина управляющего воздействия на систему, которые *позволяют* обеспечить максимально возможную эффективность ее функционирования, определенную специальным экономическим критерием (пп.8, 9 паспорта специальности 1.2.2).

Теоретическая значимость диссертационной работы состоит в следующем: предложены новые математические модели систем массового обслуживания с ограниченным временем жизни заявок; получены пороговые значения относительной пропускной способности СМО и установлена зона влияния на него времени жизни заявок; предложен экономический критерий эффективности функционирования одноканальной системы массового обслуживания для группового поступления заявок с ограниченным временем жизни.

Практическая значимость диссертационной работы состоит в том, что полученные результаты могут быть применены в медицинских, информационных, телекоммуникационных, производственных и торговых областях в качестве теоретически обоснованных критериев для проектирования СМО, в которых заявки на обслуживание имеют ограниченный срок жизни и могут приходить группами. Они позволяют оценивать результативность работы системы при изменении параметров и находить время и величину управляющего воздействия с целью максимально возможного повышения эффективности их функционирования.

Методология и методы исследования включают математический аппарат теории массового обслуживания, анализа, имитационное моделирование, программная реализация которого выполнена на языке с#.

Положения, выносимые на защиту:

1. Математическая модель системы массового обслуживания пуассоновского потока заявок с фиксированным ограничением времени жизни без опции накопления очереди.
2. Математическая модель одноканальной системы массового обслуживания группового входного потока заявок с ограниченным временем жизни с опцией накопления очереди.
3. Аналитическое решение задачи нахождения основных характеристик СМО для простейшего входного потока, экспоненциального времени обслуживания и равномерно распределенного времени жизни заявок.
4. Имитационные модели систем массового обслуживания с фиксированным ограничением времени жизни заявок.
5. Результаты анализа характеристик систем массового обслуживания с фиксированным ограничением времени жизни заявок и исследования влияния параметров систем на их эксплуатационные характеристики.
6. Результаты исследований возможности повышения эффективности функционирования систем массового обслуживания с фиксированным ограничением времени жизни заявок.

Степень достоверности. Результаты работы являются достоверными благодаря обоснованному использованию математического аппарата теории массового обслуживания и анализа, подтвержденной адекватности имитационных моделей.

Личный вклад автора заключается в проведении основного объема теоретических исследований, в создании имитационных моделей и проведении экспериментов, в анализе и оформлении результатов в виде научных публикаций и докладов, в создании и регистрации программы «Эффективное обслуживание группы заявок с ограниченным временем жизни».

Апробация работы. Результаты диссертационной работы представлены докладами на XIII Международной научно-практической конференции «Вызовы современности и стратегии развития общества в условиях новой реальности» (Москва, 2022 г.), 37-й международной научной мультиконференции «Математические методы в технике и технологиях» (г. Казань, 2024 г.), XXXIX Международной научно-практической конференции «Актуальные проблемы науки и образования в условиях современных вызовов» (г. Москва, 2025 г.).

Разработанные методы повышения эффективности систем массового обслуживания используются в деятельности организации ООО «ТД ПОЛИХИМ-ТРЕЙД».

Публикации. Материалы диссертации опубликованы в 11 научных работах общим объемом 3,84 п. л. (в т.ч. лично автором – 3,02 п.л.), из них 8 статей в журналах из перечня рецензируемых научных изданий, рекомендованных высшей аттестационной комиссией Министерства науки и высшего образования Российской Федерации – «Современные наукоемкие технологии», «Вестник Технологического университета», «Научно-технический вестник Поволжья». Получено 1 свидетельство о государственной регистрации электронного ресурса с оценкой новизны («Эффективное обслуживание группы заявок с ограниченным временем жизни»).

Структура и объем диссертации Диссертационная исследование состоит из введения, 4 глав, заключения, списка литературы, двух приложений. Общий объем составляет 141 страница, включая 48 рисунков и 18 таблиц. Список литературы содержит 148 наименований.

1 СИСТЕМЫ МАССОВОГО ОБСЛУЖИВАНИЯ С ОГРАНИЧЕННЫМ ВРЕМЕНЕМ ЖИЗНИ ЗАЯВОК

1.1 Развитие теории систем массового обслуживания

Вне всякого сомнения, можно утверждать, что появлению теории массового обслуживания способствовали активные исследования в области теории вероятностей и теории случайных процессов. Они были вызваны, в том числе, и важными задачами прикладного характера, решение которых требовало применения соответствующего математического аппарата. Одной из таких задач в начале XX века стало определение допустимой нагрузки телефонной сети. В 1909 году датский инженер А.К. Эрланг опубликовал работу «Probability and telephone calls», в которой получил формулу вероятности отказа в обслуживании абонента при звонке на телефонную станцию [1]. В это же время выдающийся русский математик А.А. Марков вел активные исследования зависимых испытаний, фактически заложив основы теории случайных процессов [2]. В 1907 году он обнародовал первые результаты своих работ по этой теме в Известиях Императорской академии наук под названием «Исследование замечательного случая зависимых испытаний» [3]. Этот труд стал началом огромного цикла исследований в области теории вероятностей, которая в настоящее время носит название «Цепи Маркова». Существенный вклад в развитие марковский цепей внес крупнейший советский математик с мировым именем А.Н. Колмогоров. Его работа 1930 года «Об аналитических методах в теории вероятностей» дала возможность существенно расширить возможности математического анализа марковских процессов с непрерывным временем [4]. С именем Колмогорова неразрывно связана названная в его честь система дифференциальных уравнений для нахождения предельных (финальных) вероятностей состояний марковской системы с дискретными состояниями и непрерывным временем. 30-е годы XX века характеризовались также пристальным вниманием к теории очередей. Среди исследователей этого научного направления

нужно особо выделить австрийско-французского инженера Феликса Поллачека и известного советского математика Александра Хинчина, которые получили практически одновременно, но разными методами стационарное распределение вероятностей для величины очереди одноканальной системы массового обслуживания пуассоновского потока заявок с произвольным распределением времени обслуживания [5–8]. Позже метод Хинчина был назван «Методом вложенных цепей Маркова». В эти годы Александр Яковлевич как депутат Моссовета включился в решение проблем автоматических телефонных станций г. Москва, что и обусловило его интерес к задачам массового обслуживания. Именно ему приписывается появление в научном обиходе ныне общепринятого термина «Теория массового обслуживания» [8].

1940-е–1950-е годы, характеризующиеся развитием математического аппарата в исследованиях систем массового обслуживания, отмечены именами Конрада Пальма, Дэвида Кендалла, Александра Хинчина и др. Пальм рассмотрел стационарный, ординарный поток событий с ограниченным последствием, получивший впоследствии его имя и имевший большое значение в изучении систем с потерями [8]. Дэвид Кендалл в 1951 году ввел стандартную нотацию для классификации систем массового обслуживания $A/B/c$, где A – закон распределения интервалов времени между моментами поступления заявок в систему, B – закон распределения времени обслуживания, c – число каналов обслуживания [9]. Хинчин, понимая потребность в упорядоченном изложении применения методов теории вероятностей к вопросам массового обслуживания, структурирует и описывает их в своей монографии «Математические методы теории массового обслуживания». Этот труд явился в то время наиболее доступным систематизированным изложением общей теории, которого не имелось не только в Советском Союзе, но и за границей [10]. В этот же период интерес к теории массового обслуживания начинает проявлять советский математик, крупный специалист в области теории вероятностей Б.В. Гнеденко. Именно он справедливо считается основателем и лидером самой

известной и авторитетной в СССР школы теории массового обслуживания, подготовившей плеяду прекрасных ученых [11-14].

Следующий период, который можно назвать расцветом теории массового обслуживания, характеризуется активным формированием теоретической базы. Б.В. Гнеденко и И.Н. Коваленко выпускают монографию, в которой систематизируют марковские случайные процессы и их обобщения [11]. Под редакцией Гнеденко выходит книга «Работы по математической теории массового обслуживания», которая собрана из работ «выдающегося советского математика Хинчина и представляет собой прекрасно написанное введение в изучение теории массового обслуживания» [8]. А.А. Боровков публикует свою монографию, в которой рассматривает методы изучения свойств предельных распределений [15]. Иностранцы ученые в этот период тоже публикуют ряд фундаментальных трудов. Это работы Л. Такача [16], Т. Саати [17], Р.А. Ховарда [18], Д. Литтла [19], Л. Клейнрока [20], Д. Кокса [21] и других исследователей [22-26]. Тогда же теория массового обслуживания появляется и в учебной литературе для инженеров и студентов вузов как самостоятельный раздел теории вероятностей [2]. Для популяризации теории выполнена переработка монографии Хинчина 1955 года с добавлением конкретных задач из производственных и хозяйственных областей [27]. Проводятся теоретические исследования разных видов систем массового обслуживания. Известный советский и российский ученый Г. Башарин рассматривает системы связи и управления со сложными структурами, предлагает метод их моделирования [28]. Б. В. Гнеденко, Н. Джейсуол, О.И. Бронштейн изучают модели систем массового обслуживания с приоритетами [29-32]. Публикуются исследования поведения циклических систем [33, 34]. Появляются многочисленные работы по полумарковским системам, полноценный обзор которых приводится в работе [35]. Большой интерес наблюдается к задачам из области теории надежности [36-39].

Начавшееся в 60-х – 70-х гг. бурное развитие компьютерной техники формирует отдельное направление, связанное с практическим применением

теории массового обслуживания в вычислительных системах [40-46]. Приступают к разработке автоматизированных систем и моделей управления с целью повышения эффективности систем массового обслуживания [47-54]. Именно в это время сложность аналитических моделей становится препятствием к их практическому применению, и возникает необходимость в численных решениях, которые могут быть получены средствами имитационного моделирования [55-58].

Дальнейшее развитие теории массового обслуживания также тесно связано с непрерывным совершенствованием компьютерной техники, появлением и повсеместным распространением сетей беспроводной передачи данных, проводной передачей и обработкой больших объемов данных. Аналитические исследования продолжают и интерес к ним не ослабевает, но они большей частью направлены на решение узко специализированных задач. Также продолжается поиск эффективных способов управления производственными процессами с помощью моделирования их как систем массового обслуживания. Появляются автоматизированные системы для имитационного моделирования: Anylogic, Arena, Bizagi Modeler, Gpss, Rand Model Designer, Netlogo, Process simulator.

По результатам аналитических исследований опубликовано много монографий и книг, затрагивающих различные аспекты теории массового обслуживания. Одни из них представляют собой учебники и пособия разного уровня сложности [59-61], другие затрагивают специальные разделы теории и ориентированы на специалистов [62-66]. Направления теоретических исследований многочисленны и разнообразны: непуассоновские сети [67], замкнутые сети в условиях большой загрузки [68], сети многолинейных СМО [69], многофазные СМО [70], разомкнутые модели СМО [70], неоднородные марковские цепи [72], фрактальные СМО [73]. Отдельно можно выделить пул работ по оптимальному управлению системами и сетями массового обслуживания [74-81].

Несмотря на то, что с момента публикации первых работ в сфере телеграфика прошел почти век, интерес к этой тематике со стороны научного

сообщества не ослабевает. Это объясняется появлением ряда новых актуальных задач в различных отраслях человеческой активности, которые могут быть описаны в терминах теории массового обслуживания. В первую очередь следует отметить монографии и учебные пособия, в которых приведены математические модели телекоммуникационных систем и расчеты показателей эффективности качества обслуживания [82-85]. Среди конкретных направлений исследований в этой области можно выделить труды, посвященные моделированию и анализу процессов, протекающих в сетях передачи данных по Sip-протоколу [86-88].

Вообще, применением СМО для изучения компьютерных сетей начали активно заниматься в 80-х годах XX века. Список имен самых значимых ученых-основателей этой области нельзя составить без имен Х. Тагаки, Л. Клейнрока, С.Ф. Яшкова, В.М. Вишневого, В.А. Жожикашвили [89-92]. В настоящее время не прекращаются аналитические исследования в области моделирования телекоммуникационного трафика как потока заявок в системах массового обслуживания [93, 94]. Особое место в сфере сетевых коммуникаций занимает тема поллинговых систем. В сущности, к такому типу относится любая система опроса, которая содержит ряд очередей, обслуживаемых в циклическом порядке. Современные исследования поллинговых систем направлены в основном на изучение широкополосных беспроводных сетей и особенностей их использования в различных условиях [95-98].

Несмотря на активные теоретические изыскания в области СМО, многие задачи так и остались нерешенными; некоторые – в силу отсутствия или ограниченности математических моделей, некоторые – в силу сложности моделей, что признают даже их авторы [99]. Бурное развитие информационных технологий, являющееся неотъемлемым атрибутом как недавнего прошлого, так и настоящего, привело к необходимости искать и находить численные решения, пригодные для прикладного применения. Один из используемых для этого инструментов – имитационное моделирование. В связи с этим нельзя не упомянуть о вкладе казанских математиков И.М. Якимова и А.П. Кирпичникова,

которые параллельно с успешными аналитическими исследованиями активно использовали средства имитационного моделирования [100-102].

1.2 Классификация систем массового обслуживания

Любая система массового обслуживания представляет собой совокупность нескольких элементов, характеристики которых предполагаются известными:

- поток поступающих требований (один или несколько);
- накопитель, в котором требования ожидают обслуживания (их может быть несколько, а может не быть совсем);
- обслуживающие устройства;
- поток обслуженных требований;
- поток необслуженных требований.

Каждый из компонентов системы может обладать рядом специфических свойств, совокупность которых позволяет причислить систему к тому, или иному классу [103, 104].

По характеру входящего потока системы могут быть марковскими и немарковскими. Марковской называют систему, состояние которой в будущем зависит только от нынешнего состояния и не зависит от того, как система в этом состоянии оказалась. Кроме того, поток может поступать извне и не зависеть от процесса работы системы. Тогда говорят об открытой системе массового обслуживания. Если же источники заявок принадлежат системе, и заявки по завершении фазы обслуживания могут возвращаться в систему, то речь идет о системе замкнутой.

По типу накопителя СМО делятся на системы с ожиданием, с потерями и смешанные. В системе с ожиданием заявка, найдя в момент поступления все обрабатывающие устройства занятыми, направляется в очередь, где ожидает обслуживания. В системе с потерями если требование в момент поступления находит все каналы обслуживания занятыми, то оно тут же покидает систему.

В смешанных системах наличествует ограничение на процесс ожидания. Он может быть ограничен емкостью накопителя, а может временем, которое требование должно находиться в очереди, или в системе, включая и фазу ожидания и обслуживания. В системах с ограниченным временем ожидания требование будет потеряно вследствие ухода из очереди по истечении предельного времени ожидания. В системах же с ограниченным временем пребывания заявка может уйти как из очереди, так и в процессе обслуживания, то есть оказаться недообслуженной, а, значит, потерянной.

Очень важной характеристикой является так называемая дисциплина очереди, то есть порядок отбора заявки из накопителя. Очередь может формироваться в определенном порядке, и на обслуживание избирается первая по порядку заявка. В другом варианте выбор заявок из накопителя может быть случайным. Также заявки могут обладать различными признаками приоритета (инвалиды, ветераны, товары с ограниченным сроком годности и т. п.), что влияет на алгоритм их подключения к обслуживанию.

По числу обслуживающих устройств системы обычно разделяют на одноканальные и многоканальные. Кроме того, некоторые процессы требуют стадийной обработки заявки, то есть каждое требование должно пройти несколько фаз обслуживания. Такие системы называются многофазными.

При любом сочетании перечисленных свойств компонентов СМО неотъемлемой особенностью системы является случайность: случайны моменты времени поступления заявок; случайна длительность обслуживания требований; случайным может быть номер обслуживающего устройства, взявшего заявку на исполнение. Именно они вместе с остальными атрибутами определяют принадлежность системы к определенному классу. Как упоминалось выше, первую классификацию А/В/с ввел Д. Кендалл в 1951 году. И она определялась законом распределения первой случайной величины – интервалов времени между моментами поступления заявок в систему, законом распределения второй случайной величины – времени обслуживания требования и количеством каналов обслуживания. В свою очередь, закон распределения

интервалов времени между моментами поступления требований может обладать специфическими особенностями. Например, если все интервалы подчиняются одному закону распределения и средняя интенсивность поступления постоянна, то входной поток обладает свойством стационарности; если за бесконечно малый (для данной системы) промежуток времени может поступить только одна заявка – свойством ординарности; если заявки приходят независимо друг от друга – свойством, которое в терминах теории массового обслуживания, называется отсутствием последействия. Поток, который обладает всеми перечисленными выше свойствами, называется он простейшим. Своим названием он обязан тому факту, что именно для него математические модели имеют самый простой вид и позволяют получить результат, пригодный для практического применения. Исследование потоков, не обладающих этими особенностями, бывает затруднено и далеко не всегда удается довести модели до расчетных процедур приемлемой сложности [105].

По мере развития теории массового обслуживания классификация Кендалла претерпевала изменения. Самое минимальное ее расширение имеет вид: $A/B/c/d$. Здесь добавлен атрибут d – число мест в накопителе ($1 = 1, 2, \dots, \infty$). Еще шире классификация вида: $A/B/c/d/e$, где e – дисциплина обслуживания. Дисциплина FIFO (англ. first in, first out) означает, что из очереди на обслуживание выбирается та заявка, которая пришла первой. LIFO (англ. last in, first out) означает, что из очереди на обслуживание выбирается та заявка, которая пришла последней. И последний вариант RANDOM означает случайный выбор заявки из очереди. Автором самой широкой из признанных в настоящее время модификаций является наш соотечественник Гелий Башарин, а сама классификация носит название «Классификация Кендалла – Башарина» [106]. Выглядит она как набор 7-ми атрибутов: $A/B/S/K/N/f/z$, определение которых обусловлено сферой научных интересов автора – областью телекоммуникационных систем. Здесь, кроме упомянутых ранее законов распределений A и B , введены:

– структура коммутационной системы (S);

- максимальное состояние системы (K);
- число источников нагрузки (N);
- приоритетность обслуживания (f);
- число мест для ожидания (z).

Но даже такая полная классификация не всеобща, так как подразумевает независимость всех рассматриваемых случайных величин друг от друга [106].

Какова бы не была СМО, она всегда является математической моделью некоторого прикладного процесса. Он может относиться к самым разным областям человеческой деятельности: автоматические узлы связи, автозаправочные станции, вычислительные системы и сети, телекоммуникационные сети, производственные линии, приемные отделения больниц, магазины и т.п. Главной целью моделирования этих процессов является повышение эффективности их функционирования. Можно выделить три категории: показатели эффективности использования, качества обслуживания и экономической эффективности функционирования. Состав каждой группы представлено на рисунке 1.2.1.

Очевидно, что для разных систем массового обслуживания приоритетными являются разные критерии результативности. Если показатели эффективности использования могут служить оценками любой системы, то применимость показателей качества обслуживания обусловлена видом СМО.

Для систем с опцией накопления очереди ключевыми оценками результативности могут быть такие параметры как средняя длина очереди, вероятность того, что вновь прибывшая заявка будет немедленно взята на обслуживание, среднее время ожидания обслуживания, среднее время пребывания заявки в системе.

Для систем без накопителя и систем с накопителем и ограничениями (ограниченным временем пребывания требования в системе, ограниченным ожиданием обслуживания, ограниченной емкостью накопителя) основными характеристиками могут являться параметры законов распределения потоков потерянных и обслуженных требований.



Рисунок 1.2.1 – Основные показатели эффективности функционирования СМО (обобщено автором)

Причем, в СМО с ограниченным временем пребывания заявки в системе поток потерянных требований является двукомпонентным. Первую компоненту образуют потери заявок, не дождавшихся обслуживания в очереди, вторую – потери вследствие незавершенного обслуживания.

Стоит обратить внимание на то, что система, обладающая высокими оценками качества обслуживания, может оказаться нерациональной с экономической точки зрения. К примеру, для минимизации доли потерянных заявок можно увеличить количество каналов обслуживания, что несомненно приведет к понижению средней нагрузки на один канал, а, значит, ухудшит показатели загрузки. Совершенно очевидно, что повышение мощности системы требует финансовых вложений, как единовременных, так и расходов на содержание и эксплуатацию оборудования. Поэтому эффективность СМО – комплексный показатель, формирование которого определяется как функциональными характеристиками системы, так и конкретными задачами, поставленными в

ходе ее эксплуатации. В некоторых случаях предлагается применять, например, взвешенную сумму показателей эффективности разных категорий [107].

1.3 Особенности систем массового обслуживания с ограниченным временем жизни заявок

Системы массового обслуживания с ограниченным временем жизни заявок – это отдельный класс систем, который объединяет СМО с временными ограничениями, являющимися атрибутами требований. Обычно, в системах с временными ограничениями, касающимися заявок, речь идет либо о периоде ожидания обслуживания в очереди, либо об общем времени пребывания заявки в системе. Главной отличительной особенностью этих разновидностей является формирование потока потерянных заявок. В случае ограниченного времени ожидания поток формируется заявками, не дождавшимися обслуживания и ушедшими из очереди. В варианте с ограничением на общее время пребывания поток двухкомпонентен – к заявкам, ушедшими из очереди по причине окончания времени жизни, добавляется поток из «недообслуженных» заявок, то есть тех, которые поступили на обслуживание, но не были обработаны полностью по той же причине. Следует отметить, что при одинаковой интенсивности входных потоков требований и одинаковой интенсивности обслуживания эффективность систем этих двух типов будет различной. Это вызвано тем, что каналы, принимающие на обслуживание заявки с истекающим временем жизни затрачивают время на их обработку впустую, так как время жизни заявки истекает в процессе обслуживания, и требование остается необработанным, а следовательно, оказывается потерянным.

Исследования систем с ограниченным временем начались на заре развития теории массового обслуживания, поскольку лимитированное время ожидания тесно связано с темой оптимизации телефонии, которая и дала толчок развитию теории очередей. Первые аналитические результаты в этой области были получены в середине XX века. Д. Баррер исследовал СМО,

характеризующуюся простейшим входным потоком заявок, показательным законом распределения времени обслуживания, постоянным временем «терпения». Под временем «терпения» подразумевалось либо период ожидания в очереди, либо общее время пребывания заявки в системе. Кроме того, были изучены два варианта дисциплины очереди: выбор на обслуживание первого по порядку требования и случайный выбор [108]. Е. С. Вентцель была предложена модель СМО, в которой ограничение по времени является случайной величиной и подчиняется показательному закону распределения [109]. И.Н. Коваленко рассматривал постоянное ограничение на периоды ожидания или пребывания в многоканальных системах для простейшего входного потока заявок и показательного времени обслуживания [110, 111]. Л. Г. Афанасьева сначала доказала для простой очереди существование предельного распределения в СМО с ограниченным временем пребывания, вывела формулы для предельных распределений вероятностей состояний системы $D/M/1$ с постоянным ограничением на время пребывания заявки, систем $D/D/1$ и $M/M/n$ со случайным ограничением [112-114]. СМО вида $G1/G/1$ с постоянными временными ограничениями и определенными требованиями на распределения потока заявок и времени обслуживания исследовал Дж. Коэн [115]. Эту же разновидность систем рассматривал и Дж. Дейли [116]. Он получил интегральное уравнение для предельной функции распределения времени ожидания и сформулировал условия существования.

При анализе перечисленных выше аналитических результатов в области СМО с ограниченным временем жизни обнаруживается определенная ограниченность моделей в части условий, для которых они получены:

- простейший или детерминированный входной поток заявок;
- пуассоновский или детерминированный закон распределения времени обслуживания;
- одноканальная система;
- постоянное ограничение на время пребывания и (или) ожидания;

– искусственные ограничения (дробная рациональность преобразования Лапласа – Стилтеса распределений времени между поступлением требований, движущийся поток требований с постоянной скоростью, нахождение начала обслуживания в определенном временном интервале и др.).

Кроме того, эти результаты не дают ответа на актуальнейший вопрос в исследовании СМО – способ повышения эффективности ее функционирования.

Решение задачи повышения эффективности функционирования систем массового обслуживания неразрывно связано понятием управляющего воздействия и наличием возможности его формирования. Одним из первых результаты исследования управляемых СМО представил В.В. Рыков. Он формализует понятие управляемой СМО, выделяет критерии качества управления, предлагает возможный способ классификации, акцентирует внимание на обязательной возможности оперативного реагирования на изменения состояния системы в процессе ее функционирования [117]. Одной из задач управления системами массового обслуживания является повышение коэффициента использования ресурсов, низкое значение которого может быть обусловлено требованиями на допустимую долю потерянных заявок. Возможным вариантом решения этой задачи может стать изменение дисциплины обслуживания. В работе [118] рассматривается система с ограниченным временем пребывания, в которой дисциплина обслуживания основана на динамической сортировке заявок в очереди.

Современные тенденции в исследовании СМО с временными ограничениями обширны и разнообразны, большинство из них обусловлены необходимостью решения прикладных задач. В работе [119] работа колл-центра моделируется как многолинейная система массового обслуживания с конечным буфером и нетерпеливыми запросами, обслуживающая марковский поток запросов. Особенностью рассматриваемой СМО является использование вероятности, с которой требование остается в очереди на обслуживание или покидает систему. В работе [120] предпринята попытка повышения качества

функционирования системы с ожиданием путем минимизации затрат за определенный период эксплуатации. Другой вариант оптимизации управляемой СМО предложен в статье [121]. Здесь речь идет о подключении резервного канала обслуживания, когда время ожидания первой в очереди заявки достигает определенного значения. Очевидным направлением практического применения подобных моделей являются системы обработки информации, где роль каналов обслуживания выполняют серверы. Следующее исследование также использует «пороговое» время ожидания обслуживания в СМО типа $M/G1/n/\infty$, целью является нахождение оптимального количества каналов обслуживания [122]. Эту же цель преследуют и авторы другой работы [123], минимизируя вероятность отказа в обслуживании в марковских СМО с ожиданием. Инструментом для численного решения поставленной задачи минимизации выбрана система MATLAB.

Следует подчеркнуть, что все вышеперечисленные работы, как и подавляющее большинство аналогичных исследований, рассматривают ограниченное время ожидания обслуживания. Имеется сравнительно мало публикаций, которые описывают системы массового обслуживания с ограниченным временем пребывания заявок в системе в целом. В первую очередь следует отметить работы Рыжикова Ю.И., которые посвящены изучению сетей массового обслуживания с ограниченным временем жизни заявок. В статье [124] «предложен алгоритм приближенного расчета разомкнутой сети массового обслуживания, в которой время прохождения маршрута (время жизни заявки) ограничено случайной величиной», входящие потоки заявок в СМО, коими являются узлы сети, являются простейшими. Для реализации алгоритма построены и верифицированы имитационные модели расчета моментов времени пребывания тех заявок в сети, которые были успешно обслужены. Средства имитационного моделирования для решения задачи поиска компромисса между ростом количества приборов и оплатой за их обслуживание в СМО с ограниченным временем пребывания используется и в работе [125]. Но в этой публикации предложен подход к описанию СМО в семимартингальных терминах, который

позволяет заложить основу имитационного процесса моделирования в виде итерационного алгоритма. Инструмент имитационного моделирования выбран и авторами работы [126], которые предлагают оптимизировать структуру системы для принятия управленческих решений.

Вообще говоря, идея применения инструментов имитационного моделирования для исследования поведения систем массового обслуживания лежит практически на поверхности. Это обусловлено, в частности, тем, что многие СМО, являющиеся моделями реальных задач, обладают гораздо более сложной структурой, чем системы массового обслуживания, для которых получены аналитические результаты. В таких случаях построение аналитической модели невозможно без ее упрощения, которое неизбежно снижает адекватность модели.

Математическое моделирование осуществимо для определенных узких классов систем массового обслуживания. В работе [127] объектом исследования является система массового обслуживания вида $GI/GI/1/r$ с групповым поступлением заявок, функциями распределения входящего потока и обслуживания фазового типа, ограниченной емкостью накопителя. Одним из результатов явился алгоритм расчета стационарного распределения числа заявок в системе. В статье других авторов исследуется нестационарный режим работы многоканальной СМО одиночных и парных заявок [128]. Особенностью является не только совместное поступление группы заявок в систему, но и совместное покидание ее: без обслуживания, если нет достаточного количества свободных каналов для принятия в работу всей группы заявок, после обслуживания – в противном случае. В результате построена система уравнений Колмогорова, которую предлагается решать численно. В работе [129] объектом исследования является бесконечно линейная система с экспоненциальным законом времени обслуживания и простейшими входящими потоками положительных и отрицательных заявок. Целью работы является установление асимптотической эквивалентности рассматриваемой системы с системами с потерей и ожиданием. Совсем другой объект исследования в работе [130] –

экспоненциальная сеть с ограниченным временем ожидания заявок и ненадежными системами обслуживания. Методом диффузной аппроксимации найдено среднее число заявок и исправных линий обслуживания в системах как в стационарном, так и в переходном режимах.

Проблематике систем массового обслуживания заявок с ограниченным временем пребывания в очереди посвятил часть своих трудов профессор Кирпичников. В работе [131] рассмотрена одноканальная система с однородным бесконечным простейшим потоком заявок, очередью неограниченной длины и временем пребывания заявки в очереди, ограниченным случайным значением с заданным средним. Получены аналитические формулы для нахождения основных числовых характеристик системы. Эти исследования продолжены в работе [132] для многоканальных СМО. Параллельно с теоретическими изысканиями автором ведутся и имитационные эксперименты, целью которых преимущественно является моделирование реальных прикладных задач во всевозможных средах с графическим вводом структур [133, 134].

1.4 Имитационное моделирование систем массового обслуживания

Имитационное моделирование – это метод изучения сложных систем, для которых средства аналитического моделирования или затруднительны для практического применения, или недостаточно разработаны ввиду особенностей структуры системы или процессов функционирования. Для процедуры симуляции строится компьютерная модель, воспроизводящая структуру и процессы функционирования реального прототипа, отражающая логику взаимодействия и закономерности поведения всех составляющих системы. Проводятся многочисленные вычислительные эксперименты, в результате которых находятся показатели эффективности функционирования. Неоспоримым преимуществом имитационного моделирования является возможность прогнозирования поведения системы под влиянием входных процессов, варьируя

параметры которых можно оптимизировать эксплуатационные характеристики системы.

Существуют разные методологии, применяемые при построении имитационных моделей, основные представлены на рисунке 1.4.1.

Агентное моделирование рассматривает систему как совокупность отдельных, самостоятельно функционирующих объектов, взаимодействующих друг с другом и окружающей средой. Определяются их параметры и поведение, устанавливаются связи между агентами. Итогом имитационного моделирования является получение закономерностей функционирования децентрализованной системы, поведение которой, в конечном счете, определяется действиями каждого объекта.



Рисунок 1.4.1 – Методологии имитационного моделирования
(обобщено автором)

Системная динамика моделирует системы целиком, абстрагируясь от индивидуальных свойств объектов. Основное предназначение этой концепции – стратегическое планирование, поскольку исследуются именно задачи, а не система целиком. В процессе моделирования с помощью накопителей и потоков между ними строятся глобальные взаимосвязи. В результате получаются

зависимости значений ключевых показателей или их статистических характеристик от выбора той или иной управленческой стратегии.

Дискретно-событийное моделирование имитирует поведение системы в хронологии происходящих событий. Случившееся событие влечет за собой изменение состояния всей системы. Если промежутки времени между событиями определены точно или аналитически, то такое моделирование называют детерминированным. Если время наступления событий случайно, то – стохастическим. Исторически первым, использующим случайность, а точнее, генератор псевдослучайных величин, стал метод «Монте-Карло». Значение случайной величины включалось в состав аналитической формулы для вычисления требуемого параметра. Повторялись многократные эксперименты с разными значениями случайной величины, проводился анализ статистического распределения результатов. В отличие от методологии «Монте-Карло» в стохастическом дискретно-событийном моделировании основополагающим является фактор времени, а именно то, что переходы системы из одного состояния в другое происходят в случайные моменты времени. Совершенно очевидно, что именно дискретно-событийный характер моделирования наиболее органичен системам массового обслуживания, стохастичность которых как раз и обусловлена случайным характером времени.

Независимо выбора методологии может быть предложена общая схема проведения имитационных исследований сложных систем, основу которой заложил Р. Шеннон [56, 135] (рис. 1.4.2).

1. Этап «Формулировка проблемы» подразумевает постановку целей и задач исследования. Согласно Эйнштейну, правильная постановка задачи даже более важна, чем ее решение. Это высказывание как нельзя лучше относится к имитационным исследованиям, поскольку проблема правильной постановки задачи зачастую является их обычным атрибутом. На этом же этапе формулируются показатели, которые будут использоваться для оценки эффективности модели.

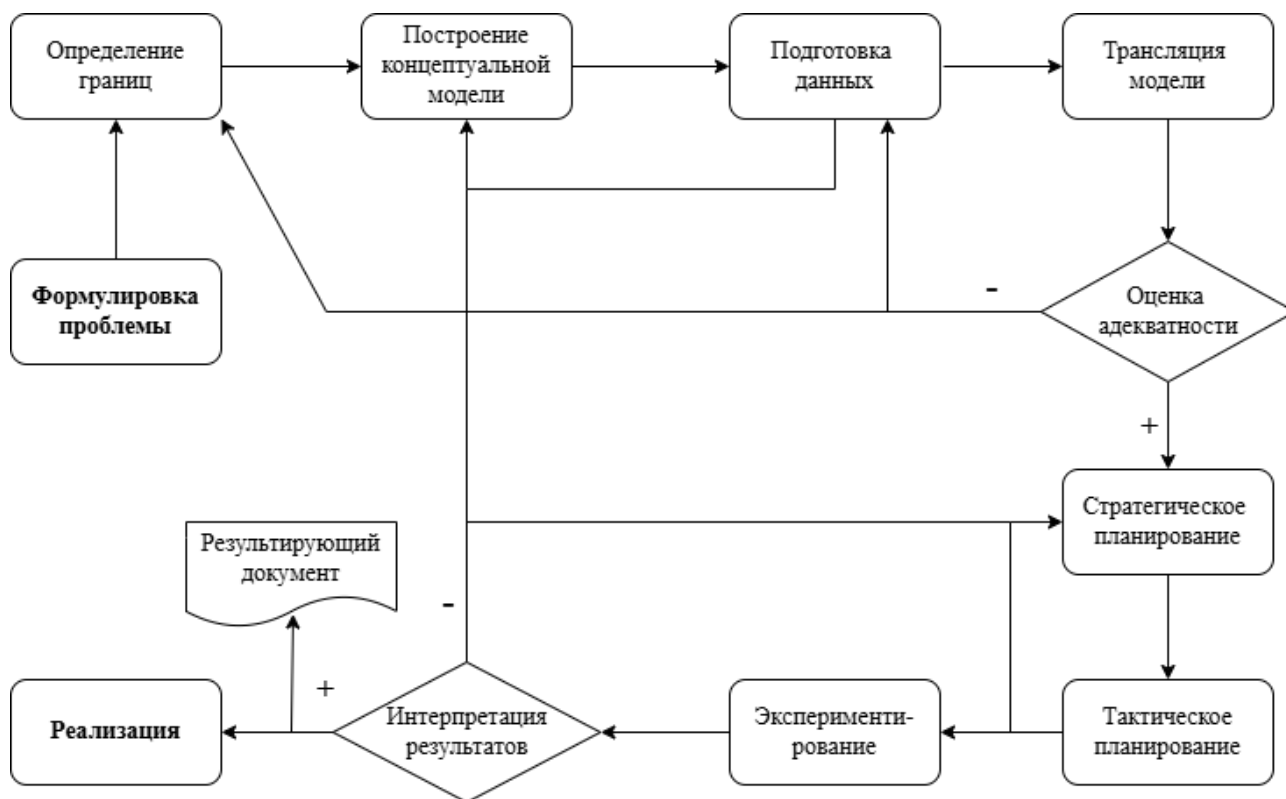


Рисунок 1.4.2 – Этапы процесса имитации (обобщено автором)

2. Этап «Определение границ» состоит в определении количества и свойств элементов и процессов системы, входных параметров, выходных характеристик. Происходит детализация структуры системы, то есть выделение основных компонентов, между которыми устанавливаются зависимости. Здесь нужно учитывать, что чрезмерная детализация чревата увеличением временных и финансовых затрат с одной стороны, и потерей ясности целеполагания – с другой.

3. Этап «Построение концептуальной модели» заключается в создании абстрактной содержательной модели в словесной форме. Если исследователь выбирает готовую систему моделирования, то его действия на данном этапе могут заключаться в работе с визуальными средствами построения. Этот этап носит характер сборки и отладки логического функционала без проверки результатов работы на конкретных исходных данных.

4. Этап «Подготовка данных» представляет собой сбор, анализ и верификацию количественных характеристик функционирования системы. При

анализе происходит статистическая обработка массивов случайных данных с целью нахождения функции распределения или ее аппроксимации известными законами.

5. Содержание этапа «Трансляция модели» зависит от выбора среды моделирования. Если выбрана современная система имитационного моделирования, то на данном этапе производится окончательная отладка модели с приведением форматов и размерностей посредством штучных экспериментов на конкретных исходных входных данных. Если же предполагается программная реализация без использования автоматизированных систем моделирования, то трансляция модели заключается в переводе ее на язык программирования и проведение отдельных экспериментов.

6. Этап «Оценка адекватности» является одновременно важным и сложным. Стандартами для оценки достоверности результатов работы имитационной модели приняты процедуры оценки валидации модели, верификации модели и валидации данных [135]. Валидация модели заключается в сравнении результатов работы имитационной модели с результатами деятельности реальной системы или с аналитическими результатами, если таковые имеются. Если нет, то валидность может определить экспертный совет в рассматриваемой предметной области. Верификация модели – это определение корректности перевода концептуальной модели в программную, осуществляется с помощью тестирования отдельных блоков и модулей, возможно, с помощью сторонних специалистов. Оценка валидности данных заключается в проверке достоверности, корректности и необходимой точности всех входных и выходных данных, параметров системы.

7. Стратегическое и тактическое планирование заключаются в построении глобальных и детальных планов вычислительных экспериментов в имитационной компьютерной модели с разными входными параметрами и ограничениями.

8. Этап «Экспериментирование» представляет собой серию имитационных экспериментов по тактическому плану, построенному на

предыдущем этапе. На этом этапе можно определить чувствительность модели, то есть выявить те параметры, которые оказывают наибольшее влияние на результат. Это является большим преимуществом имитационного моделирования, которое полностью подконтрольно разработчику в отличие от экспериментов с реальными системами.

9. На этапе «Интерпретация результатов» осуществляется обработка результатов, возможно, с привлечением сторонних программных средств, формируются рекомендации или выявляются ошибки и происходит возвращение на один из предыдущих этапов.

10. Этап «Реализация» является заключительным. На нем происходит оформление результатов, формирование рекомендаций по работе с моделью, возможным способам ее улучшения, расширения сферы применения.

Стоит обратить внимание на то, возврат на один из предыдущих этапов при обнаружении ошибок или неточностей может потребоваться не только на этапе «Интерпретация результатов», но практически на любой стадии разработки имитационной модели. Таким образом, говорить о простоте и быстроте симуляции как метода моделирования не приходится. Как и любая другая модель, имитационная обладает рядом недостатков и преимуществ. Любопытно, что ее особенности, заключающиеся в определенных требованиях к вычислительной технике, интерпретировались поначалу как недостаток, влекущий дорогостоящую и длительное время разработки в силу маломощности и небольшой распространённости компьютеров. В настоящее время быстродействие современных компьютеров и сетевые преимущества доступа к программному продукту определяют эти же особенности имитационного моделирования как преимущества.

2 СИСТЕМЫ МАССОВОГО ОБСЛУЖИВАНИЯ СО СЛУЧАЙНЫМ ВРЕМЕНЕМ ЖИЗНИ ЗАЯВОК

В настоящем разделе приведена математическая модель системы массового обслуживания, особенностью которой является наличие случайного ограничения на время жизни заявок. Полагается, что требование может быть потеряно по причине окончания срока жизни как в процессе обслуживания, так и в процессе ожидания в очереди.

Во избежание терминологической неопределенности следует отметить, системы такого типа нельзя смешивать с системами с «нетерпеливыми» заявками, где требования не попадают на обслуживание, либо покидая очередь по прошествии определенного времени, либо в случайные моменты, образуя еще один поток определенной интенсивности. То есть системы с «нетерпеливыми» заявками предполагают существование накопителя, тогда как в системах с ограниченным временем жизни заявок его наличие не является обязательным.

В подобных системах случайный характер времени жизни требования возникает из-за неопределенности внешней среды, которую балансируют с помощью выбора подходящего закона распределения. Примерами требований, имеющих случайное ограничение на время жизни могут являться запросы в вычислительных системах с микросервисной архитектурой, заявки на покупку или продажу акций, специальные предложения в маркетинге. В перечисленных ситуациях идет речь о случайном времени ожидания обслуживания, так как если работа над заявкой началась, то она должна завершиться полностью. В качестве примеров требований, имеющих случайное ограничение на общее время пребывания в системе, можно рассмотреть http-запрос с полным таймаутом или биологическую пробу для конвейерной обработки. Здесь требование может быть потеряно не только в процессе ожидания, но и в процессе самого обслуживания.

2.1 Математическое моделирование одноканальной системы с очередью

Системы обработки заявок с ограниченным временем жизни, которые подразумевают наличие накопителя, обладают той спецификой, что потеря заявок может произойти как в очереди в процессе ожидания обслуживания, так и в процессе самого обслуживания вследствие окончания времени жизни заявки. Поэтому, как показано в работах [136, 137], для создания аналитической модели недостаточно знать законы распределения потоков поступления заявок в систему и их обслуживания. Основываясь на результатах упомянутых работ Б.В. Гнеденко и И.Н. Коваленко, приведем возможную методику построения аналитической модели СМО с опцией накопления очереди и одним каналом обслуживания заявок с ограниченным случайным временем жизни.

Положим, что входной поток заявок в исследуемой СМО подчиняется закону распределения Пуассона с параметром λ . Этот постулат является очень важным фактором, значимо влияющим на процесс моделирования. Он гарантирует отсутствие последействия, стационарность и ординарность входного потока требований.

Обозначим через $H(t)$ функцию распределения случайной величины – времени обслуживания, через $v(t)$ – случайную величину – время ожидания начала обслуживания заявки, поступившей в момент времени t .

Введем вероятностные характеристики процесса обслуживания, необходимые для построения модели:

1. $F(t, \omega) = P(v(t) < \omega)$ – функция распределения фактического времени ожидания обслуживания.
2. $B(\omega)$ – функция распределения вероятностей того, что время пребывания заявки в накопителе равно некоторой величине ω .
3. $G_x(\omega)$ – функция распределения вероятностей того, что после нахождения заявки в накопителе на протяжении времени x , заявка попадет на обслуживание, где будет находиться в течение отрезка времени, не большего ω .

Дадим переменной t некоторое приращение Δt . Определим, при каких условиях время ожидания начала обслуживания в течение времени $t + \Delta t$ не превзойдет величины ω , значения которой могут различаться. То есть определим условия, при которых выполняется неравенство $v(t + \Delta t) \leq \omega$. Для этого рассмотрим все возможные варианты несовместных событий, которые могут произойти за рассматриваемый промежуток времени:

1. Не поступает заявка вовсе. Тогда имеем: $v(t) < \omega + \Delta t$.
2. За интервал времени $t + \varepsilon$ ($0 < \varepsilon < \Delta t$) поступает одна заявка и уходит из накопителя через время $v(t) = x < \omega + \Delta t$, не получив обслуживание по причине окончания времени жизни.
3. За интервал времени $t + \varepsilon$ ($0 < \varepsilon < \Delta t$) поступает одна заявка, находится в накопителе время x , получает обслуживание, но не заканчивает его по причине окончания времени жизни и покидает систему до момента времени $\omega - x + \Delta t$.
4. За интервал времени $t + \varepsilon$ ($0 < \varepsilon < \Delta t$) поступает одна заявка, находится в накопителе время x , получает обслуживание и завешает его, находясь на обслуживании время, большее чем $\omega - x + \Delta t$.

Поскольку рассмотренные события образуют полную группу, то вероятность выполнения неравенства $v(t + \Delta t) \leq \omega$ представима как сумма вероятностей этих событий.

$$\begin{aligned}
 F(\omega, t + \Delta t) &= (1 - \lambda \Delta t)F(\omega + \Delta t, t) + \\
 &+ \lambda \Delta t \int_0^{\omega + \Delta t} \{B(x) + [1 - B(x)]G_x(\omega - x + \Delta t) + \\
 &+ [1 - B(x)][1 - G_x(\omega - x + \Delta t)]H(\omega - x + \Delta t)\} d_x F(x, t)
 \end{aligned} \tag{2.1.1}$$

При построении формулы (2.1.1) использованы свойства пуассоновского входного потока требований и введенные выше вероятностные характеристики процесса обслуживания.

Первое слагаемое определяет вероятность того, что за время $t + \Delta t$ не поступит ни одна заявка, то есть $F(t + \Delta t) = 0$. В силу стационарности входного потока заявок для любого t верно равенство:

$$\lambda \Delta t = \sum_{i=0}^{\infty} i F_i(\Delta t), \quad (2.1.2)$$

а в силу ординарности верно равенство:

$$F_0(\Delta t) + F_1(\Delta t) + o(\Delta t) = 1 \text{ при } \Delta t \rightarrow 0, \quad (2.1.3)$$

то есть для малого промежутка времени вероятность поступления в систему более одной заявки ничтожно мала. При $\Delta t \rightarrow 0$ из формул (2.1.2) и (2.1.3) получаем равенство:

$$\begin{aligned} \lambda \Delta t &= 0 * F_0(\Delta t) + 1 * F_1(\Delta t) + o(\Delta t) \Rightarrow \\ &\Rightarrow \lambda \Delta t = 1 - F_0(\Delta t) \Rightarrow \\ &\Rightarrow F_0(\Delta t) = 1 - \lambda \Delta t. \end{aligned} \quad (2.1.4)$$

В силу же отсутствия последействия, которое гарантирует попарную независимость событий в потоке для любых t и Δt верно равенство:

$$F(t + \Delta t) = F(t) * F(\Delta t) \quad (2.1.5)$$

Таким образом, равенство (2.1.5) преобразуется к виду:

$$F(t + \Delta t) = (1 - \lambda \Delta t) F(\omega + \Delta t, t). \quad (2.1.6)$$

Три слагаемых под знаком интеграла в выражении (2.1.1) представляют собой вероятности наступления событий, описанных в пунктах 2, 3, 4.

Преобразуем уравнение (2.1.1) к дифференциальному виду, разделив его на Δt и устремив Δt к 0. Получим:

$$\begin{aligned} \frac{\partial F(\omega, t)}{\partial t} &= \frac{\partial F(\omega, t)}{\partial \omega} - \lambda F(\omega, t) + \lambda \int_0^{\omega} \{B(x) + [1 - B(x)]G_x(\omega - x) + \\ &+ [1 - B(x)][1 - G_x(\omega - x)]H(\omega - x)\} d_x P(x, t). \end{aligned} \quad (2.1.7)$$

Как показано в статье [111], у систем с ограниченным временем пребывания заявок существует стационарное состояние, значит, при $t \rightarrow \infty$ выполняется равенство:

$$\frac{\partial F(\omega, t)}{\partial t} = 0. \quad (2.1.8)$$

Учитывая это обстоятельство, уравнения (2.1.7) преобразуется к виду:

$$\frac{\partial F(\omega)}{\partial \omega} = \lambda \int_0^{\omega} [1 - B(x)][1 - G_x(\omega - x)] * [1 - H(\omega - x)] dF(x). \quad (2.1.9)$$

Построенная модель позволяет определить ключевые показатели работы системы, такие как вероятность ухода заявки из очереди α , вероятность потери заявки в процессе обслуживания β и другие.

$$\alpha = \int_0^{\infty} B(x) dF(x). \quad (2.1.10)$$

$$\beta = \int_0^{\infty} [1 - B(x)] \int_0^{\infty} G_x(y) dH(y) dF(x) \quad (2.1.11)$$

Как показано в работе [138], функция $F(x)$ терпит разрыв в точке 0 и абсолютно непрерывна при $x > 0$. Обозначим через $f(x)$ производную справа от функции $F(x)$. Равенство (2.1.9) принимает вид:

$$\begin{aligned} f(\omega) - \lambda \int_0^{\omega} (1 - B(x))(1 - G_{\omega}(\omega - x))(1 - H(\omega - x))f(x) dx = \\ = \lambda F(0)(1 - G_0(\omega))(1 - H(\omega)), \end{aligned} \quad (2.1.12)$$

$F(0)$ находится из нормирующего условия

$$F(0) + \int_0^T f(x) dx = 1, \quad (2.1.13)$$

где T – максимальное время жизни заявки.

Для систем с ограничением, накладываемым на общее время пребывания заявки в системе, и стандартной дисциплиной обслуживания FIFO, функция $B(x)$ определяет распределение времени пребывания заявок в системе, а функция $G_x(\omega)$ зависит только от времени пребывания заявки в накопителе и определяется как условная вероятность:

$$G_x(\omega) = (B(\omega + x) - B(x))/(1 - B(x)). \quad (2.1.14)$$

При этом, уравнение (12) преобразуется к виду:

$$\begin{aligned} f(\omega) - \lambda(1 - B(\omega)) \int_0^{\omega} (1 - H(\omega - x))f(x)dx = \\ = \lambda F(0)(1 - B(\omega))(1 - H(\omega)). \end{aligned} \quad (2.1.15)$$

Применение математической модели для решения задачи

Рассмотрим одноканальный процесс обслуживания заявок, в котором входной поток является простейшим, время обслуживания распределено экспоненциально, а время жизни заявки случайно и подчиняется равномерному закону распределения. Такие законы характерны для системы обслуживания сервером заявок торговых роботов на покупку или продажу акций; системы диспетчеризации скорой помощи; сервисов доставки еды; обработки видео-пакетов в реальном времени.

Пусть

$\lambda = 6$ – интенсивность простейшего входного потока заявок;

$H(x) = 1 - e^{-10x}$ – функция распределения случайной величины – времени обслуживания;

$B(x) = x/2$ – равномерное распределение на $[0, 2]$ – вероятностей того, что время ожидания заявки в очереди равно некоторой случайной величине x ;

$G_x(\omega) = \omega/(2 - x)$ – распределение вероятностей того, что после пребывания заявки в очереди в течение отрезка времени x , заявка попадет на обслуживание, где будет находиться в течение отрезка времени, не большего ω .

Тогда уравнение (2.1.15) принимает вид:

$$f(\omega) - 6 \left(1 - \frac{\omega}{2}\right) \int_0^{\omega} e^{-10(\omega-x)} f(x)dx = 6F(0) \left(1 - \frac{\omega}{2}\right) e^{-10\omega}. \quad (2.1.16)$$

Поскольку время пребывания заявки в системе подчинено равномерному закону распределения на интервале $[0, 2]$, возникает естественное ограничение $\omega \leq 2$.

Обозначим:

$$D(\omega) = \int_0^{\omega} e^{10x} f(x) dx. \quad (2.1.17)$$

Тогда:

$$f(\omega) = e^{-10\omega} D'(\omega). \quad (2.1.18)$$

При подстановке равенства (2.1.18) в уравнение (2.1.16) получаем:

$$e^{-10\omega} D'(\omega) - 6e^{-10\omega} \left(1 - \frac{\omega}{2}\right) D(\omega) = 6F(0)e^{-10\omega} \left(1 - \frac{\omega}{2}\right),$$

или в другом виде:

$$D'(\omega) - (6 - 3\omega)D(\omega) = F(0)(6 - 3\omega). \quad (2.1.19)$$

Это линейное дифференциальное уравнение первого порядка. Для решения введем интегрирующий множитель:

$$\mu(\omega) = e^{\int -(6-3\omega)d\omega} = e^{-6\omega+1,5\omega^2}.$$

Умножим уравнение (19) на интегрирующий множитель:

$$\begin{aligned} e^{-6\omega+1,5\omega^2} D'(\omega) - (6 - 3\omega)e^{-6\omega+1,5\omega^2} D(\omega) \\ = F(0)(6 - 3\omega)e^{-6\omega+1,5\omega^2}. \end{aligned} \quad (2.1.20)$$

Преобразуем полученное уравнение к дифференциальному виду:

$$\frac{d}{d\omega} (e^{-6\omega+1,5\omega^2} D(\omega)) = F(0)(6 - 3\omega)e^{-6\omega+1,5\omega^2}. \quad (2.1.21)$$

Интегрируем обе части уравнения (2.1.21) по ω :

$$\begin{aligned} e^{-6\omega+1,5\omega^2} D(\omega) &= F(0) \int (6 - 3\omega)e^{-6\omega+1,5\omega^2} d\omega + C = \\ &= F(0)(-e^{-6\omega+1,5\omega^2}) + C, \end{aligned} \quad (2.1.22)$$

откуда

$$D(\omega) = -F(0) + Ce^{6\omega-1,5\omega^2}. \quad (2.1.23)$$

Учитывая определение (2.1.17) функции $D(\omega)$, имеем, что $D(0) = 0$, поэтому

$$C = F(0).$$

Таким образом,

$$D(\omega) = F(0)(e^{6\omega-1,5\omega^2} - 1).$$

Отсюда следует:

$$D'(\omega) = F(0) \frac{d}{d\omega} (e^{6\omega-1,5\omega^2} - 1) = F(0)e^{6\omega-1,5\omega^2} (6 - 3\omega),$$

следовательно

$$f(\omega) = e^{-10\omega} F(0) e^{6\omega-1,5\omega^2} (6 - 3\omega) = F(0) e^{-4\omega-1,5\omega^2} (6 - 3\omega). \quad (2.1.24)$$

Найдем $F(0)$ из условия нормировки (2.1.13):

$$F(0) + F(0) \int_0^2 e^{-4x-1,5x^2} (6 - 3x) dx = 1. \quad (2.1.25)$$

Вычислим интеграл в выражении (2.1.25):

$$I = \int_0^2 e^{-4x-1,5x^2} (6 - 3x) dx = 10 \int_0^2 e^{-4x-1,5x^2} dx + \\ + \int_0^2 (-4 - 3x) e^{-4x-1,5x^2} dx = I_1 + I_2.$$

$$I_2 = \int_0^2 (-4 - 3x) e^{-4x-1,5x^2} dx = e^{-4x-1,5x^2} \Big|_0^2 = e^{-14} - 1 \approx -1,$$

$$I_1 = 10e^{\frac{8}{3}} \int_0^2 e^{-\frac{3}{2}(x+\frac{4}{3})^2} dx = \left\{ \begin{array}{l} u = \sqrt{\frac{3}{2}}(x + \frac{4}{3}) \\ dx = \sqrt{\frac{2}{3}} du \end{array} \right\} = \quad (2.1.26)$$

$$= 10e^{\frac{8}{3}} \sqrt{\frac{2}{3}} \int_{\frac{2\sqrt{6}}{3}}^{\frac{5\sqrt{6}}{3}} e^{-u^2} du.$$

Для нахождения интеграла в выражении (2.1.26) используем функцию ошибок:

$$\operatorname{erf}(u) = \frac{2}{\sqrt{\pi}} \int_a^b e^{-u^2} du.$$

Тогда

$$\begin{aligned} I_1 &= 10e^{\frac{8}{3}} \sqrt{\frac{2}{3}} \frac{\sqrt{\pi}}{2} \left(\operatorname{erf}\left(\frac{5\sqrt{6}}{3}\right) - \operatorname{erf}\left(\frac{2\sqrt{6}}{3}\right) \right) \approx \\ &\approx 10 * 14,39 * 0,724 * (1 - 0,979) \approx 2,19. \end{aligned}$$

Таким образом получаем, что $I \approx 1,19$, и из (2.1.25) следует:

$$F(0) = \frac{1}{1+I} \approx 0,456.$$

В результате искомая функция плотности вероятностей имеет вид:

$$f(x) \approx 0,456(6 - 3x)e^{-4x-1,5x^2}.$$

$F(0) \approx 0,456$ интерпретируется как доля заявок, принятых на обслуживание немедленно в момент поступления, без ожидания в очереди.

Доля заявок, получивших отказ в обслуживании, вычисляется по формуле (2.1.10):

$$\begin{aligned} \alpha &= 0,456 \int_0^2 \left(\frac{x}{2}\right) (6 - 3x)e^{-4x-1,5x^2} dx = \\ &= 0,228 \int_0^2 (6x - 3x^2)e^{-4x-1,5x^2} dx = \\ &= \frac{0,228}{18} \left(-43e^{\frac{8}{3}}\sqrt{6\pi}\operatorname{erf}\left(\frac{3x+4}{\sqrt{6}}\right) - e^{-\frac{x}{2}(3x+8)}(60 - 18x) \right) \Big|_0^2 \approx 0,048. \end{aligned}$$

Доля заявок, не получивших полное обслуживание, вычисляется по формуле (2.1.11) с применением равенства (2.1.14):

$$\begin{aligned}
\beta &= 10 \int_0^2 \left(\int_0^\infty \frac{z}{2-y} e^{-10z} dz \right) \left(1 - \frac{y}{2}\right) (6 - 3y) e^{-4y-1.5y^2} dy = \\
&= 10 \int_0^2 \left(\int_0^\infty z e^{-10z} dz \right) \frac{1}{2-y} \left(\frac{2-y}{2}\right) (6 - 3y) e^{-4y-1.5y^2} dy = \\
&= 5 \int_0^2 [(e^{-10x}(-0.1x - 0.01))|_0^\infty] (6 - 3y) e^{-4y-1.5y^2} dy = \\
&= 0.05 \int_0^2 (6 - 3y) e^{-4y-1.5y^2} dy = \\
&= \frac{0.05}{3} \left(5e^{\frac{8}{3}} \sqrt{6} \operatorname{erf} \left(\frac{3x+4}{\sqrt{6}} \right) - e^{-\frac{x}{2}(3x+8)} (60 - 18x) \right) \Big|_0^2 \approx 0,0589.
\end{aligned}$$

Доля заявок, получивших полное обслуживание, находится так:

$$Q = 1 - \alpha - \beta = 0,8927.$$

Следовательно, при заданных входных параметрах почти 90% заявок будет обслужено полностью.

Приведенный пример показывает, что решение подобных задач аналитически является долгим и крайне затруднительным процессом, причем, для каждой группы входных требуется индивидуальный подход. Кроме того, математическая модель строилась в предположении, что входной поток заявок простейший, в случае другого закона распределения она уже не работает, как и многие другие аналитические модели, вынужденные идеализировать реальность в угоду разрешимости. В отличие от них, имитационные модели позволяют моделировать любые законы распределения, работать с нестационарными потоками, мгновенно реагировать на изменение внешних параметров и логики функционирования систем.

2.2 Построение имитационной модели одноканальной системы с очередью и ее верификация

Для построения имитационной модели воспользуемся дискретно-событийным подходом. Он представляет собой методологию моделирования процессов во времени, где состояние системы изменяется мгновенно в те моменты модельного времени, когда в системе происходят события [138].

Базовый алгоритм работы имитационной модели СМО в самых общих представлен на рисунке 2.2.1.



Рисунок 2.2.1 – Блок-схема алгоритма работы имитационной модели СМО (разработано автором)

В качестве входных параметров для запуска процедуры симуляции используются:

1. Границы временного интервала моделирования. Они задаются на оси модельного времени, представляющую собой искусственную систему отсчета, в рамках которой функционирует модель. Она служит для выстраивания линии событий и определения последовательности изменений, происходящих в моделируемой системе, определяет шкалу измерения для всех параметров моделируемой системы.
2. Закон распределения интервалов времени между поступлением заявок и его параметры.
3. Закон распределения значений времени жизни заявки и его параметры.
4. Закон распределения значений времени обработки заявки и его параметры.

Для проектирования и реализации алгоритма, который имитирует работу одноканальной СМО, обрабатывающей пуассоновский поток заявок со случайным временем жизни и имеющей опцию накопления очереди, была разработана и использована блок-схема, изображенная на рисунке 2.2.2. Представленный алгоритм реализован на языке программирования C#.

Результаты работы программы для входных параметров задачи, приведенной в п. 2.1 (средней интенсивности входного потока заявок $\lambda = 6$, средней интенсивности обслуживания $\mu = 10$, равномерно распределенного на отрезке $[0, 2]$), таковы:

- доля заявок, получивших отказ в обслуживании, – 0,046;
- доля заявок, получивших неполное обслуживание, – 0,054;
- доля заявок, получивших полное обслуживание, – 0.9.

Полученные результаты с точностью до сотых долей совпадают с результатами аналитического решения, поэтому можно утверждать, что имитационная модель адекватна и применима для других функций распределения потоков заявок.

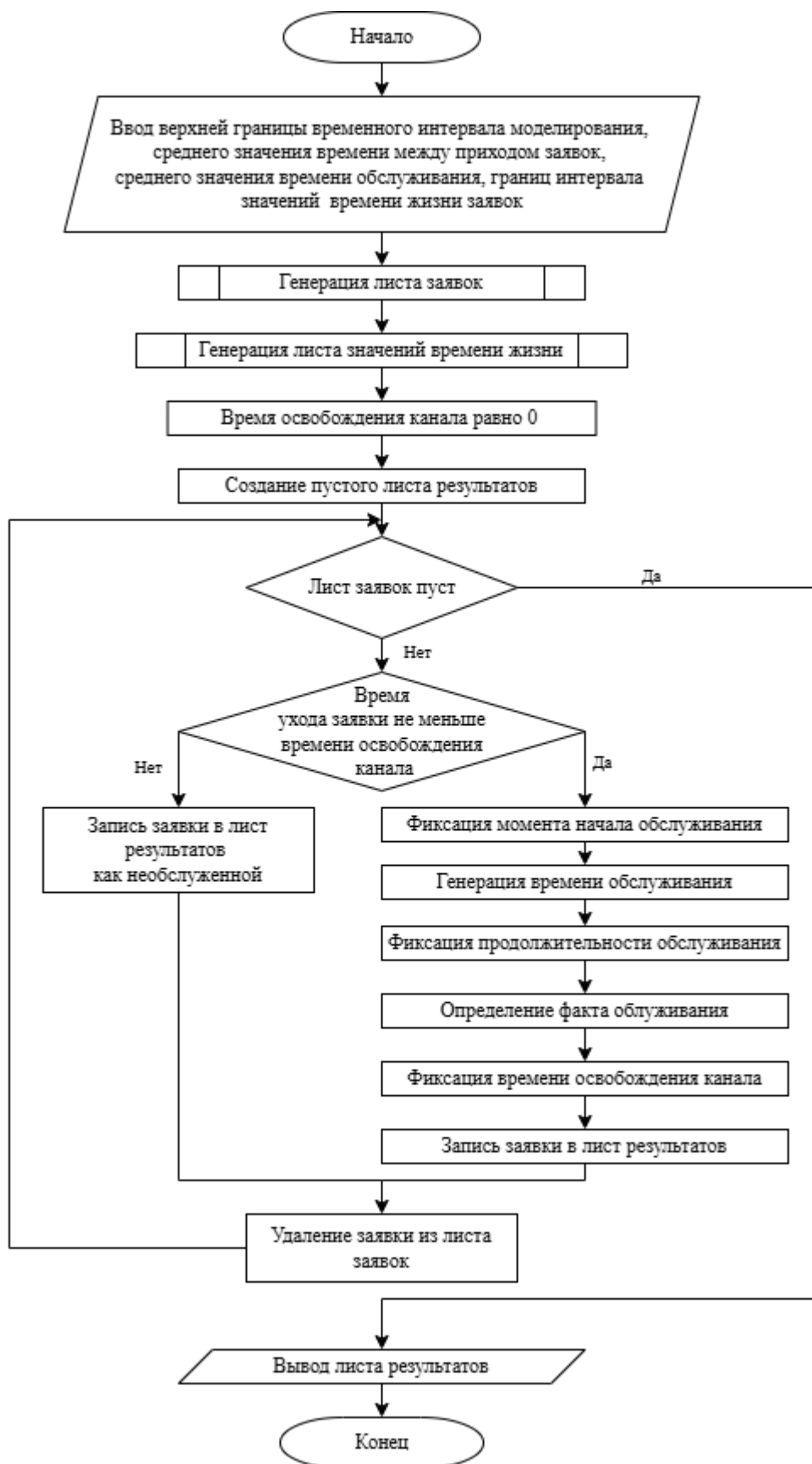


Рисунок 2.2.2 – Блок-схема алгоритма работы имитационной модели одноканальной СМО с очередью и случайным временем жизни заявок

(разработано автором)

Таким образом, для одноканальной СМО со случайным временем жизни заявок и опцией накопления очереди изложена методика построения математической модели, основанная на результатах работ [136, 137]. Показано, что модель работает для СМО простейшего потока заявок, и для ограниченного времени пребывания задается законами распределения входного потока, потока обслуживания и времени пребывания заявок в системе. На основе описанной концептуальной модели для равномерно распределенного времени жизни заявок, выполнена постановка задачи и получено аналитическое решение. Сделан вывод о том, что процесс решения каждой конкретной задачи требует больших временных затрат. Кроме того, для решения большинства подобных задач требуется применение численных методов, которые необходимо адаптировать под специфику задачи. Более эффективным представляется подход имитационного моделирования, который позволяет решать широкий спектр задач без существенной переработки алгоритмов.

Изложенные результаты представлены в работе [139, 140].

3 СИСТЕМЫ МАССОВОГО ОБСЛУЖИВАНИЯ С ФИКСИРОВАННЫМ ОГРАНИЧЕНИЕМ ВРЕМЕНИ ЖИЗНИ ЗАЯВОК

3.1 Математическое моделирование системы обслуживания ординарного потока заявок с отказами

Рассмотрим систему массового обслуживания потока заявок, которые обслуживает единственный канал, и опция накопления очереди отсутствует. Положим, что входной поток требований является пуассоновским со средней частотой появления заявок λ . Также будем считать пуассоновским поток обслуживания, обозначив через μ среднее число заявок, которое может обслужить канал в единицу времени. Граф состояний такой системы представлен на рисунке 3.1.1.

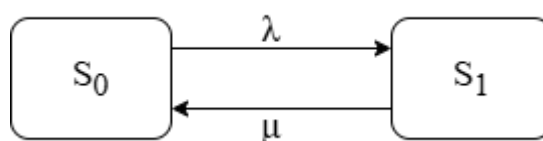


Рисунок 3.1.1 – Граф состояний одноканальной СМО нетерпеливых
заявок [103]

Здесь S_0 – состояние, в котором система ожидает заявку» (единственный канал свободен), S_1 – состояние системы, в котором система обслуживает заявку (единственный канал занят обслуживанием). Такие системы хорошо известны, и их поведение досконально изучено.

Вероятности нахождения системы в своих состояниях P_0 и P_1 зависят от времени, поскольку их описание задается системой дифференциальных уравнений Колмогорова. Но при $t \rightarrow \infty$ они достигают своих предельных значений и фактически теряют эту зависимость. Этот режим называют стационарным, а вероятности – финальными или предельными. Доказано, что для системы с

конечным числом состояний финальные вероятности существуют и достигаются в темпе экспоненты.

Так как S_0 – состояние ожидания заявки, то финальная вероятность P_0^* нахождения системы в этом состоянии есть вероятность того, что поступившая заявка найдет систему свободной и сразу попадет на обслуживание. То есть P_0^* имеет смысл доли обслуженных заявок. В состоянии S_1 система занята обслуживанием, значит, финальная вероятность P_1^* нахождения системы в этом состоянии есть вероятность того, что поступившая заявка найдет систему занятой и моментально покинет ее. То есть P_1^* имеет смысл доли потерянных заявок.

Финальные вероятности P_0^* и P_1^* равны соответственно

$$P_0^* = \frac{\mu}{\lambda + \mu}, \quad P_1^* = \frac{\lambda}{\lambda + \mu} \quad (3.1.1)$$

Наличие ограничений на время жизни заявок требует внесения некоторых дополнений, которые прокомментированы ниже. Дело в том, что в этом случае для адекватного описания функционирования СМО оказывается недостаточно средней интенсивности обслуживания λ . Связано это прежде всего с важной эксплуатационной характеристикой – загруженностью канала, которое, при отсутствии ограничений на время жизни заявок, равно обратной величине интенсивности μ пуассоновского входного потока. Если же срок жизни заявок ограничен некоторой величиной t_{end} , то те заявки, которые должны были бы обслуживаться в течение более продолжительного времени, покинут систему в момент времени t_{end} . Следствием это станет уменьшение среднего времени пребывания заявок в системе, которое окажется тем более значительным, чем меньше срок жизни заявок. Именно это обстоятельство обуславливает необходимость введения еще одной характеристики – интенсивности выхода заявок из системы ($\hat{\mu}$).

Поскольку при неограниченном времени жизни заявок интенсивность выхода заявок из системы совпадает с интенсивностью обслуживания – μ , справедливо неравенство

$$\hat{\mu} > \mu. \quad (3.1.2)$$

Обозначив через \widehat{P}_0 и \widehat{P}_1 –вероятности стационарных состояний СМО с ограниченным временем жизни, получим следующие формулы для их нахождения:

$$\widehat{P}_0 = \frac{\hat{\mu}}{\lambda + \hat{\mu}}, \quad \widehat{P}_1 = \frac{\lambda}{\lambda + \hat{\mu}}. \quad (3.1.3)$$

Покажем, как можно найти величину $\hat{\mu}$.

Заметим, что показатель средней интенсивности обслуживания μ остаётся постоянным вне зависимости от внешних факторов, так как определяется техническими характеристиками обслуживающего устройства.

Поскольку входной поток обслуживания полагается простейшим, плотность распределения интервалов времени обслуживания имеет вид:

$$f(t) = \mu e^{-\mu t}. \quad (3.1.4)$$

Установленное ограничение на время жизни заявок позволяет определить долю полностью обработанных заявок, принятых каналом на обслуживание, следующим образом:

$$q = \int_0^{t_{end}} f(t) dt. \quad (3.1.5)$$

Среднее время пребывания в системе полностью обслуженных заявок определяется как средневзвешенная величина, где в качестве весовой характеристики выступают доли обслуженных заявок. То есть фактически является медианным значением t_M , которое находится из условия:

$$\int_0^{t_M} f(t) dt = \int_{t_M}^{t_{end}} f(t) dt = \frac{q}{2}. \quad (3.1.6)$$

На рисунке 3.1.2 представлена графическая иллюстрация этого соотношения.

И, поскольку время пребывания в системе всех не полностью обслуженных требований одинаково и равно t_{end} , среднее время пребывания \bar{t} всех заявок, попавших на обслуживание, определяется как

$$\bar{t} = qt_M + (1 - q)t_{end}, \quad (3.1.7)$$

а среднее значение интенсивности ухода заявок будет:

$$\hat{\mu} = \frac{1}{\bar{t}}. \quad (3.1.8)$$

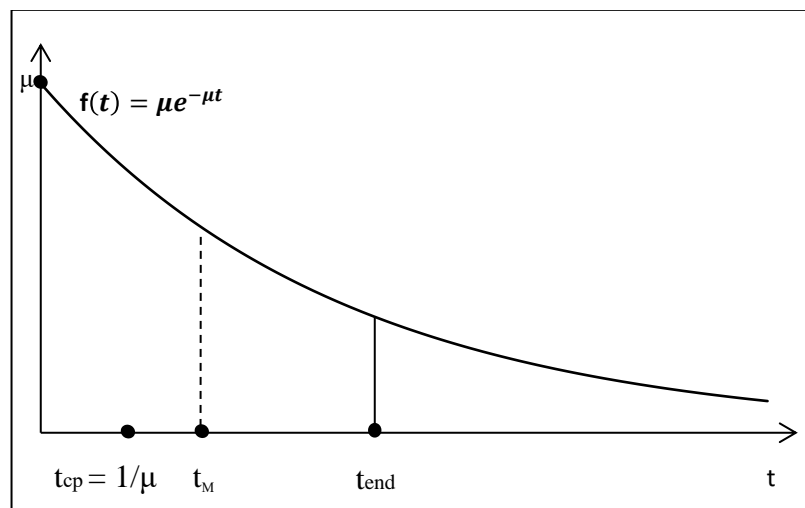


Рисунок 3.1.2 – Доля заявок, не получивших обслуживание в полном объеме (разработано автором)

Чтобы получить долю заявок, обслуженных полностью (Q), следует применить формулу:

$$Q = q\hat{P}_0. \quad (3.1.9)$$

Получить долю заявок, попавших на обслуживание, но потерянных по причине окончания срока жизни (R), можно так:

$$R = (1 - q)\hat{P}_0. \quad (3.1.10)$$

Доля заявок, не попавших на обслуживание вовсе, то есть тех, на момент прихода которых обслуживающее устройство было занято, равна \hat{P}_1 .

Очевидно, что справедливо равенство: $Q + R + \hat{P}_1 = 1$.

При рассмотрении СМО с несколькими каналами одинаковой производительности и отсутствием очереди будем использовать тот же алгоритм построения модели. Это означает, что значение интенсивности ухода заявок из системы можно получить с помощью соотношений (3.1.6) – (3.1.8), и использовать их в формулах Эрланга для нахождения финальных вероятностей состояний системы:

$$\widehat{P}_0 = \left[\sum_{i=0}^n \frac{\widehat{\rho}^i}{i!} \right]^{-1}, \widehat{P}_i = \frac{\widehat{\rho}^i}{i!} \widehat{P}_0, \widehat{\rho} = \frac{\lambda}{\widehat{\mu}}, i = \overline{1, n}, \quad (3.1.11)$$

где n – число каналов обслуживания, $\widehat{\rho}$ – интенсивность нагрузки на канал.

На основании вышеизложенного вероятность совсем не попасть на обслуживание для заявки вычисляется по формуле:

$$\widehat{P}_{\text{отк}} = \frac{\widehat{\rho}^n}{n!} \widehat{P}_0, \quad (3.1.12)$$

доля полностью обслуженных заявок находится так:

$$Q = q \sum_{j=0}^{n-1} \widehat{P}_j, \quad (3.1.13)$$

доля заявок, попавших на обслуживание, но потерянных по причине окончания срока жизни, определяется следующим образом:

$$R = (1 - q) \sum_{j=0}^{n-1} \widehat{P}_j. \quad (3.1.14)$$

В качестве примера рассмотрим одноканальную систему без опции накопления с ограниченным временем жизни заявок с параметрами $\lambda = 10$, $\mu = 6$, $t_{\text{end}} = 0,5$. Применим формулы (3.1.6) – (3.1.10):

$$\int_0^{t_M} 6e^{-6t} dt = \int_{t_M}^{0,5} 6e^{-6t} dt,$$

откуда $t_M \approx 0,1074$.

Доля полностью обработанных заявок, принятых каналом на обслуживание, составляет

$$q = 2 \int_0^{0,1074} 6e^{-6t} dt = 0,95.$$

Среднее время пребывания всех заявок, попавших на обслуживание:

$$\bar{t} = 0,95 * 0,1074 + 0,05 * 0,5 = 0,127.$$

Среднее значение интенсивности ухода заявок: $\widehat{\mu} = 7,87$.

Вероятности нахождения системы в состояниях ожидания и обслуживания соответственно равны:

$$\widehat{P}_0 = \frac{7,87}{10 + 7,87} = 0,44, \quad \widehat{P}_1 = \frac{10}{10 + 7,87} = 0,56.$$

Доля заявок, обслуженных полностью: $Q = 0,95 * 0,44 = 0,418$.

Доля заявок, попавших на обслуживание, но потерянных по причине окончания срока жизни: $R = 0,05 * 0,44 = 0,022$.

Доля заявок, не попавших на обслуживание вовсе: $\widehat{P}_1 = 0,56$.

Для сравнения рассчитаем доли обработанных и потерянных требований для одноканальной системы без очереди и без ограничения на время жизни заявки с теми же средними значениями времени:

$$P_0^* = \frac{6}{10 + 6} = 0,375, \quad P_1^* = \frac{10}{10 + 6} = 0,625.$$

Результаты вычислений показывают, что в случае наличия ограничения на время жизни заявок интенсивность ухода выше, чем интенсивность обслуживания, и поэтому, доля заявок, совсем не попавших на обслуживание выше в системе без ограничений.

В качестве другого примера рассмотрим трехканальную систему без опции накопления с ограниченным временем жизни заявок с теми же входными параметрами $\lambda = 10$, $\mu = 6$, $t_{\text{end}} = 0,5$.

Интенсивности нагрузки на канал составляет

$$\widehat{\rho} = \frac{\lambda}{\widehat{\mu}} = \frac{10}{7,87} = 1,27.$$

По формулам Эрланга найдём финальные вероятности системы.

Вероятность того, что в момент поступления заявки все каналы свободны вычисляется так:

$$\widehat{P}_0 = \left[\sum_{i=0}^n \frac{\widehat{\rho}^i}{i!} \right]^{-1} = \left[1 + 1,27 + \frac{1,27^2}{2!} + \frac{1,27^3}{3!} \right]^{-1} = 0,292.$$

Вероятности того, что в момент поступления заявки заняты один, два или все три канала равны соответственно:

$$\widehat{P}_1 = 1,27 * 0,292 = 0,372; \widehat{P}_2 = \frac{1,27^2}{2!} 0,292 = 0,236;$$

$$\widehat{P}_3 = \frac{1,27^3}{3!} 0,292 = 0,1.$$

В концепции предложенной модели доля полностью обслуженных заявок будет вычислена так:

$$(\widehat{P}_0 + \widehat{P}_1 + \widehat{P}_2) * 0,95 = 0,855,$$

доля заявок, попавших на обслуживание, но не завершивших его:

$$(\widehat{P}_0 + \widehat{P}_1 + \widehat{P}_2) * 0,05 = 0,045,$$

доля заявок, не попавших на обслуживание совсем по причине занятости всех трех каналов:

$$P_{\text{отк}} = \widehat{P}_3 = 0,1.$$

Итак, для одноканальных и многоканальных систем обслуживания пуассоновского потока заявок с ограниченным временем жизни без накопителя построена математическая модель, обладающая очевидной прикладной полезностью. Используя стандартные входные параметры СМО – средние интенсивности потока заявок и потока обслуживания, и постоянное время жизни заявки в качестве ограничения, модель позволяет отыскать не только интенсивность выхода заявок из системы и вероятности стационарных состояний системы, но и основные функциональные характеристики системы, такие, как средняя доля требований, потерянных в момент поступления, средняя доля требований, потерянных в процессе обслуживания по причине окончания срока жизни, и средняя доля требований, обслуженных полностью.

Изложенные результаты представлены в работе [139].

3.2 Имитационное моделирование систем массового обслуживания ординарного потока заявок

Рассмотрим систему массового обслуживания, которая обрабатывает заявки, приходящие группами и имеющие фиксированный, одинаковый для всех

срок жизни. Система может иметь один или несколько каналов обслуживания, может обладать или не обладать опцией накопления очереди. Группы заявок могут иметь случайный объем и приходить через случайные промежутки времени. Как и полагалось выше, ограниченный срок жизни заявок для систем без накопителя означает возможность потери заявки в процессе обслуживания, для систем с накопителем – возможность потери заявки как в процессе обслуживания, так и в процессе ожидания обслуживания в очереди. Самым простым вариантом такой системы является одноканальная СМО с простейшим входным потоком заявок, время жизни которых ограничено постоянной величиной, и экспоненциальным характером обслуживания. Аналитические модели этой системы существуют, в настоящей работе представлена одна из них. Для систем с другим распределением входного потока заявок или их групповым поступлением, весьма проблематично осуществить построение математической модели при наличии ограничения на время жизни заявок. Существование накопителя в СМО также создает дополнительные сложности для процесса аналитического моделирования. Как показано в настоящем исследовании, даже при успешной реализации намерения создать математическую модель возникают значительные затруднения в области ее практического применения из-за сложности полученных конструкций. Кроме того, аналитические модели не могут обеспечить высокую скорость принятия решений, которая имеет важное значение для управления системами в концепте повышения эффективности их функционирования. Эти обстоятельства обуславливают целесообразность численного моделирования работы системы для анализа ее функционирования, и применения готовых программных продуктов – для управления ею.

3.2.1 Системы обслуживания ординарного входного потока заявок с отказами

Одноканальная СМО с отказами

Рассмотрим одноканальную систему массового обслуживания заявок с фиксированным временем жизни, образующих пуассоновский входной поток,

без опции накопления очереди. Будем считать, что значения времени обслуживания заявки – случайные величины, распределенные по экспоненциальному закону. Для проектирования и реализации алгоритма, который имитирует работу рассматриваемой в этом подразделе СМО, разработана блок-схема, изображенная на рисунке 3.2.1.

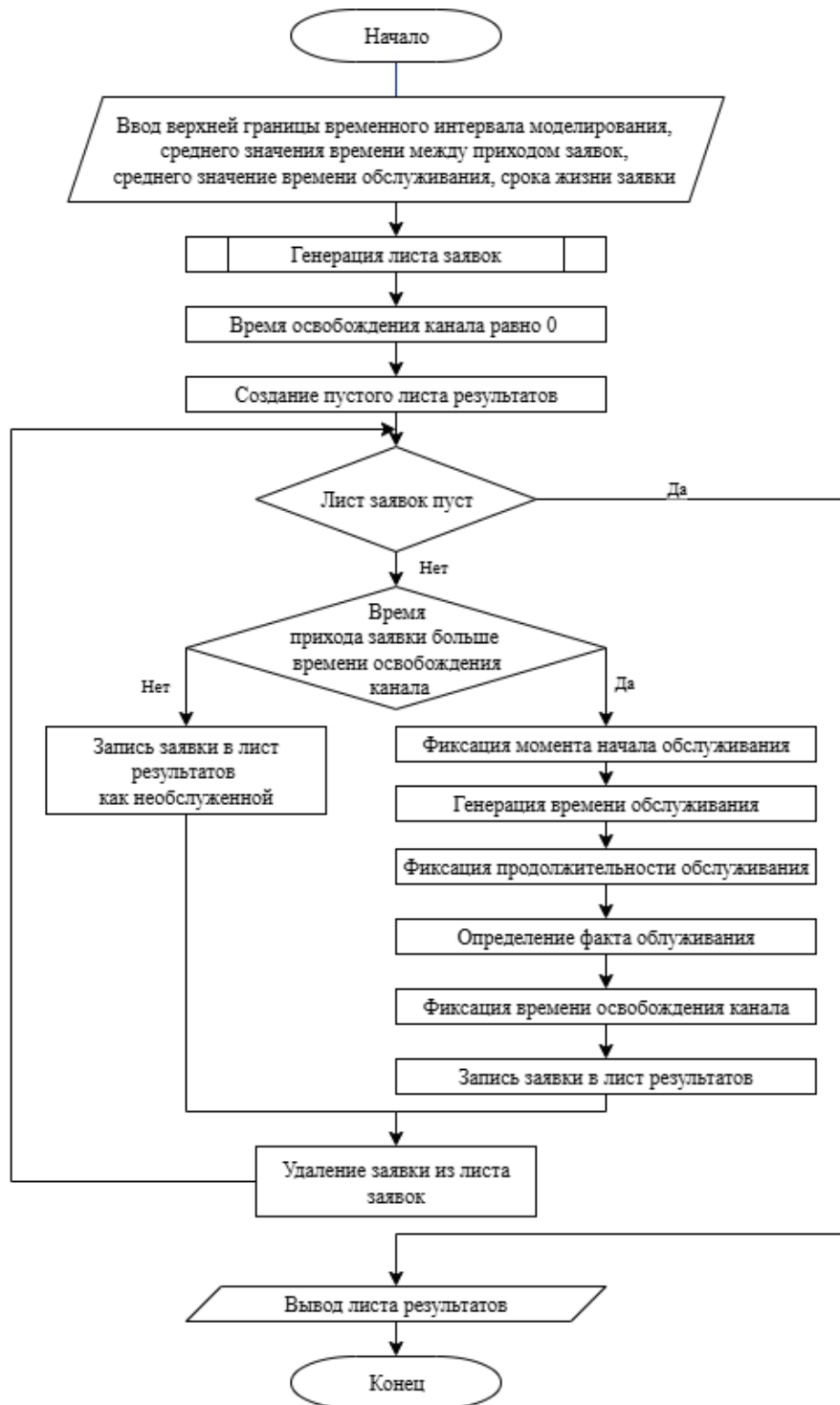


Рисунок 3.2.1 – Блок-схема алгоритма работы имитационной модели одноканальной СМО без очереди (разработано автором)

Приведенный алгоритм реализован на языке программирования C#.

Входными параметрами для старта моделирования являются:

1. Границы временного интервала моделирования.
2. Среднее значение времени между приходом заявок.
3. Среднее значение времени обработки заявок.
4. Срок жизни заявок.

Дисциплина обслуживания требований является стандартной, в порядке их поступления (FIFO).

Результатом работы имитационного эксперимента является лист, содержащий множество записей. Каждая запись включает следующие сведения об одной заявке:

- момент модельного времени прихода заявки;
- сгенерированная в соответствии с заданным законом распределения продолжительность обработки заявки;
- фактическая длительность обслуживания, которая зависит от длительности обслуживания и ограничения времени жизни заявки;
- момент модельного времени выхода заявки из системы;
- показатель факта полного обслуживания.

Выходные данные модуля моделирования передаются в модуль анализа для статистической обработки.

В таблице 3.2.1 представлена часть массива результатов работы, описанного выше алгоритма для следующих входных параметров:

- 1) верхняя граница интервала моделирования 1000 врем. ед. Нижняя – 0;
- 2) экспоненциальный закон распределения интервалов времени между поступлением заявок со средним $\lambda = 10$;
- 3) экспоненциальный закон распределения значений времени обработки заявок со средним $\mu = 6$;
- 4) срок жизни заявок с момента поступления в систему 0,5 врем. ед.

Таблица 3.2.1 – Часть массива результатов эксперимента по имитации одноканальной СМО без очереди с простейшим входным потоком заявок (рассчитано автором)

№ п/п	Момент времени прихода заявки	Расчетное время обработки заявки	Длительность обработки заявки	Момент времени выхода заявки из системы	Индикатор полного обслуживания
1	0,191269	0,007905	0,007905	0,199174	1
2	0,214675	0,078231	0,078231	0,292906	1
3	0,41163	0,41969	0,41969	0,83132	1
4	0,460752	0	0	0,460752	0
5	0,624517	0	0	0,624517	0
6	0,663678	0	0	0,663678	0
7	0,737211	0	0	0,737211	0
8	0,769391	0	0	0,769391	0
9	0,83763	0,025448	0,025448	0,863078	1
10	0,926585	0,063987	0,063987	0,990572	1
11	0,940888	0	0	0,940888	0
12	1,464652	0,363099	0,363099	1,827751	1
13	1,570829	0	0	1,570829	0
14	1,864172	0,387997	0,387997	2,252169	1
15	1,902801	0	0	1,902801	0
16	2,693411	0,004623	0,004623	2,698034	1
17	2,726247	0,078653	0,078653	2,8049	1
18	2,838919	0,249986	0,249986	3,088905	1
19	2,8598	0	0	2,8598	0
...
99696	9999,552	0	0	9999,552	0
99697	9999,685	0	0	9999,685	0

В результате многократно проведенных имитационных экспериментов с одинаковым набором входных параметров были получены следующие характеристики рассматриваемой СМО:

- 1) средняя доля заявок, получивших полное обслуживание, равна 0,37;
- 2) средняя доля заявок, начавших обслуживание, но не закончивших его по причине окончания срока жизни, равна 0,02;
- 3) средняя доля заявок, не попавших на обслуживание совсем, равна 0,61.

Результаты работы имитационной модели отличаются от аналитических расчетов, приведенных в разделе 3.1 настоящего исследования, не более, чем на 0,05, что укладывается в пределы допустимой погрешностью вычислений.

Приняв время жизни заявки равным 2, мы получим несколько иную картину. Средняя доля заявок, получивших полное обслуживание, увеличится до 0,375, средняя доля заявок, не попавших на обслуживание совсем, увеличится до 0,625. Примечательно, что совсем не будет заявок, начавших обслуживание и не завершивших его. Представляется важным заметить, что средние доли в этом случае равны финальным вероятностям состояний СМО без ограничения на время жизни заявки с теми же входными параметрами. То есть при данном сочетании входных параметров ограничение на время жизни заявок не влияет на функционирование системы.

На рисунке 3.2.2 представлены графики зависимостей средних долей от времени жизни заявок для фиксированных $\lambda = 10$ и $\mu = 6$.

Анализ представленных результатов позволяет сделать вывод о том, что для данного набора режимных параметров существует пороговое значение времени жизни заявок, при достижении которого его влияние на функционирование системы практически исчезает. При указанном сочетании входных параметров это происходит при значении времени жизни заявки 0,3.

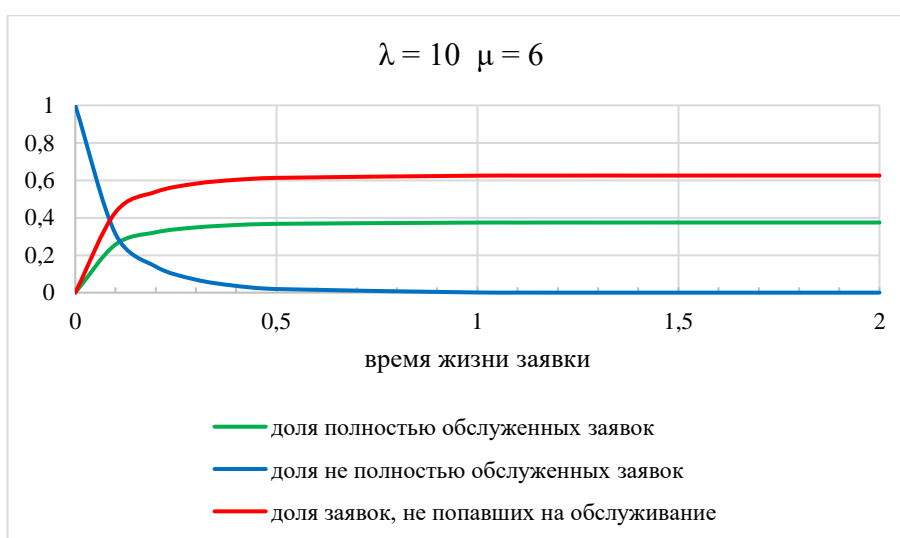


Рисунок 3.2.2 – Зависимости долей заявок от времени их жизни в одноканальной СМО без накопителя

На рисунке 3.2.3 представлены графики зависимостей средних долей полностью обслуженных заявок для разных сочетаний параметров λ и μ .

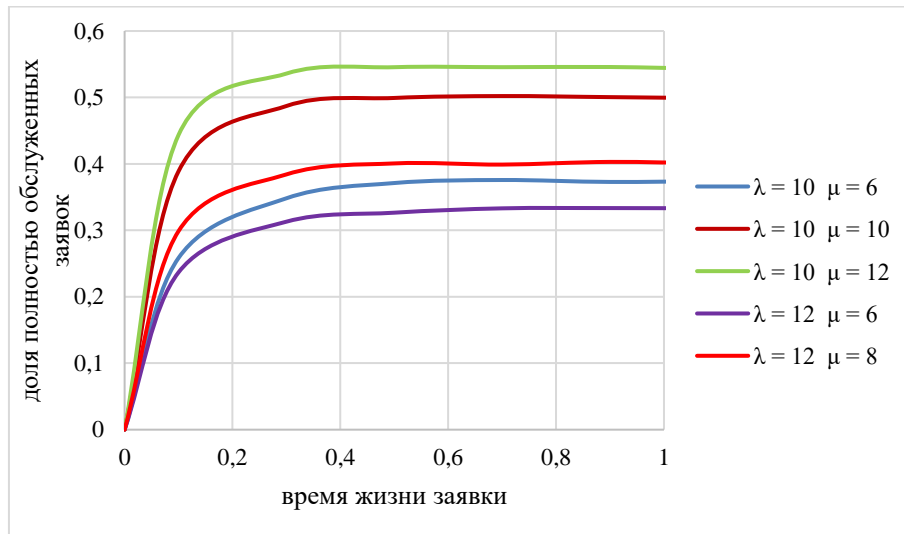


Рисунок 3.2.3 – Зависимости долей полностью обслуженных заявок от времени их жизни в одноканальных СМО без накопителя

При малых интенсивностях процессов поступления и обслуживания заявок тенденция сохраняется, меняется лишь значение времени жизни заявки, при котором прекращается ее влияние на поведение системы (рис. 3.2.4).

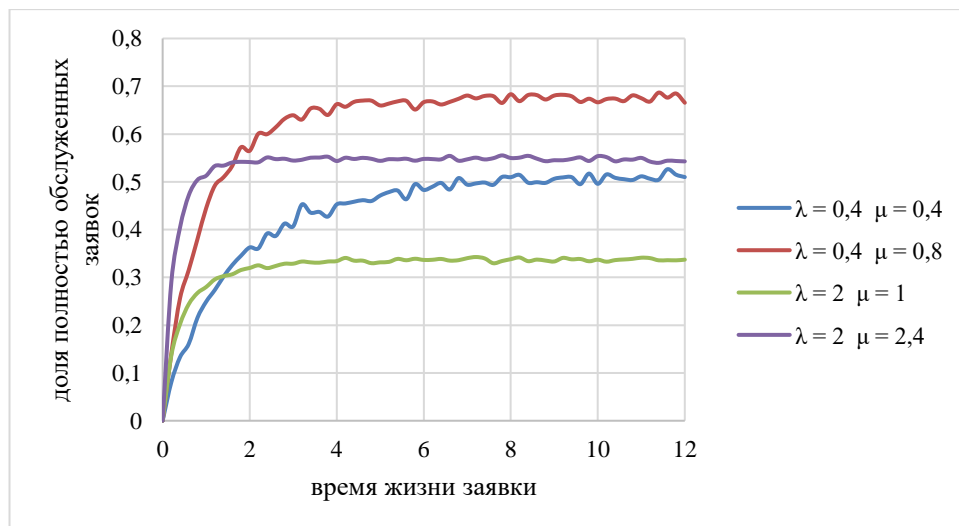


Рисунок 3.2.4 – Зависимости долей полностью обслуженных заявок от времени их жизни в одноканальных СМО без накопителя при малых интенсивностях

В заключении можно сказать, что, что в одноканальной СМО без опции накопления очереди время жизни заявок имеет влияние на функционирование системы в достаточно узком диапазоне его значений, границы которого зависят от интенсивностей входного потока заявок и процесса обслуживания. На рисунке 3.2.5 изображены пороговые значения времени жизни заявок, при которых система становится нечувствительной к этому ограничению.

		λ									
		1	2	3	4	5	6	7	8	9	10
μ	1	1,8	1,1	0,8	0,6	0,5	0,4	0,3	0,3	0,3	0,3
	2	1,2	0,9	0,7	0,6	0,5	0,4	0,3	0,3	0,3	0,3
	3	0,9	0,7	0,6	0,5	0,5	0,4	0,3	0,3	0,3	0,3
	4	0,7	0,6	0,5	0,5	0,4	0,4	0,3	0,3	0,3	0,3
	5	0,6	0,5	0,5	0,4	0,4	0,4	0,3	0,3	0,3	0,3
	6	0,5	0,5	0,4	0,4	0,4	0,3	0,3	0,3	0,3	0,3
	7	0,4	0,4	0,4	0,4	0,3	0,3	0,3	0,3	0,3	0,2
	8	0,4	0,4	0,4	0,3	0,3	0,3	0,3	0,3	0,2	0,2
	9	0,4	0,3	0,3	0,3	0,3	0,3	0,3	0,3	0,2	0,2
	10	0,3	0,3	0,3	0,3	0,3	0,3	0,3	0,2	0,2	0,2

Рисунок 3.2.5 – Пороговые значения времени жизни заявок для одноканальных СМО без накопителя с высокими интенсивностями

Аналогичные пороговые значения времени жизни заявок были получены и для СМО с низкими интенсивностями входного потока заявок и обслуживания (рис. 3.2.6).

		λ									
		0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1
μ	0,1	17	11	7	6	5	4	3	3	3	3
	0,2	11	9	7	6	5	4	3	3	3	3
	0,3	8	7	6	5	4	4	3	3	3	3
	0,4	7	6	6	5	4	4	3	3	3	3
	0,5	6	5	5	4	4	4	3	3	3	3
	0,6	5	4	4	4	4	3	3	3	3	3
	0,7	5	4	4	4	3	3	3	3	3	2
	0,8	4	4	4	3	3	3	3	3	2	2
	0,9	4	3	3	3	3	3	3	2	2	2
	1	3	3	3	3	3	3	2	2	2	2

Рисунок 3.2.6 – Пороговые значения времени жизни заявок для одноканальных СМО без накопителя с низкими интенсивностями

Проиллюстрированные данные демонстрируют наличие изохронных областей значений порогового времени жизни заявок для разных средних интенсивностей входного потока и обслуживания.

Многоканальная СМО с отказами

Рассмотрим многоканальную систему массового обслуживания заявок с ограниченным временем жизни, образующих пуассоновский входной поток, без опции накопления очереди. Каналы обслуживания полагаются эквивалентными с одинаковым средним значением времени обслуживания, распределенным по экспоненциальному закону. Каналы образуют очередь на обслуживание по мере освобождения, то есть при поступлении требования в систему, его примет на обслуживание тот канал, который до этого освободился раньше всех.

Блок-схема алгоритма работы имитационной модели, реализованной на языке программирования C#, представлена на рисунке 3.2.7.

Здесь появляется новый входной параметр – количество каналов обслуживания, и новый выходной параметр – номер канала, взявшего заявку в работу.

Имитационный эксперимент был проведен для следующих входных параметров:

1. Верхняя граница интервала моделирования 1000 врем. ед. Нижняя – 0.
2. Экспоненциальный закон распределения интервалов времени между поступлением заявок со средним $\lambda = 10$.
3. Экспоненциальный закон распределения значений времени обработки заявок со средним $\mu = 6$.
4. Срок жизни заявок с момента поступления 0,5 врем. ед.
5. Количество каналов 3.

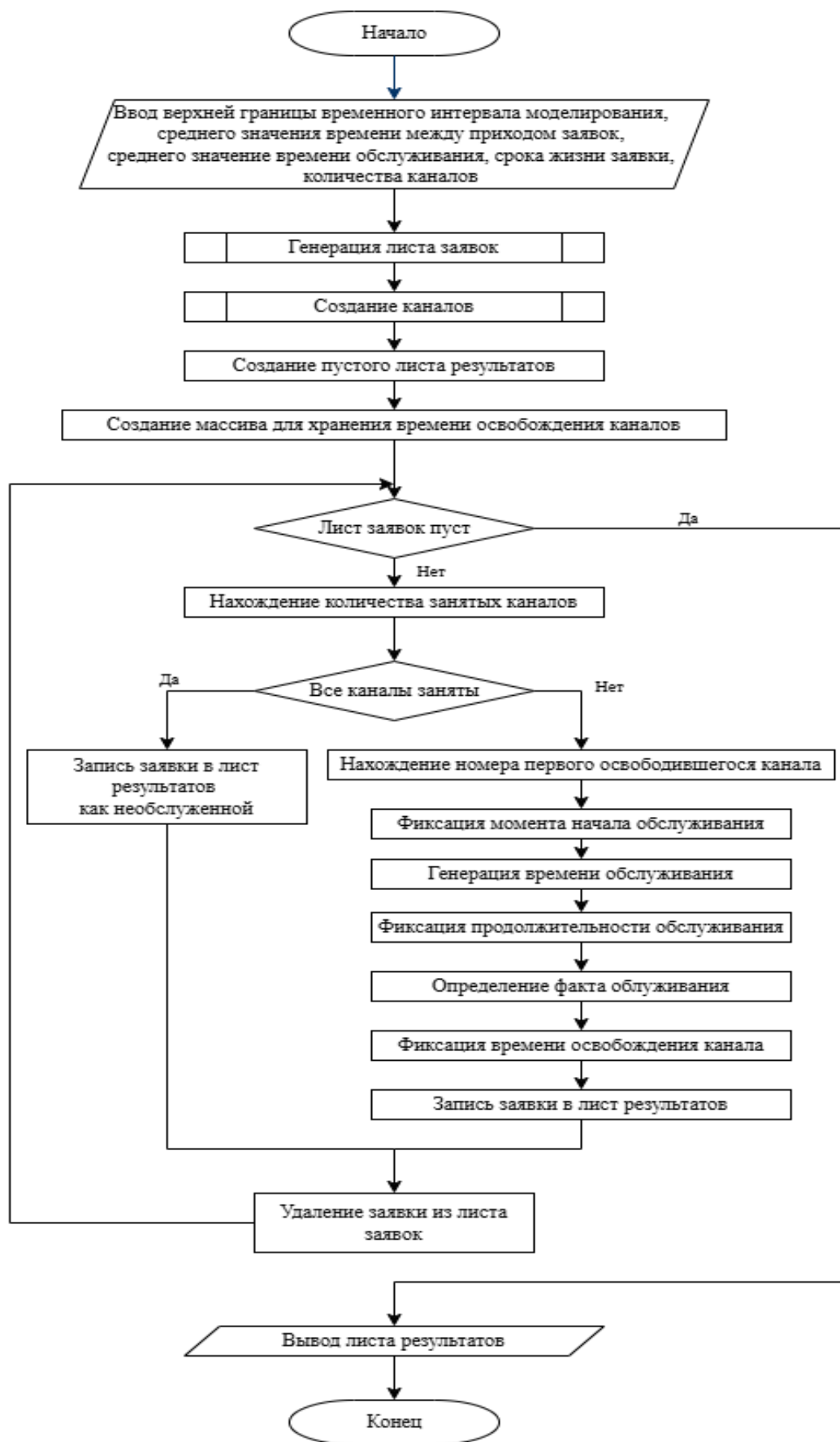


Рисунок 3.2.7 – Блок-схема алгоритма работы имитационной модели многоканальной СМО без очереди (разработана автором)

В таблице 3.2.2 представлена часть массива результатов работы описанного выше алгоритма для выбранных параметров.

Таблица 3.2.2 – Часть результатов эксперимента по имитации трехканальной СМО без очереди с простейшим входным потоком заявок

№ п/п	Момент времени прихода заявки	Расчетное время обработки заявки	Длительность обработки заявки	Момент времени выхода заявки из системы	Номер канала	Индикатор полного обслуживания
1	0,04949	0,087687	0,087687	0,137177	1	1
2	0,055992	0,174521	0,174521	0,230513	2	1
3	0,09607	0,002494	0,002494	0,098564	3	1
4	0,112714	0,300642	0,300642	0,413356	3	1
5	0,277754	0,0501	0,0501	0,327854	1	1
6	0,303097	0,075467	0,075467	0,378564	2	1
7	0,404826	0,086299	0,086299	0,491125	1	1
8	0,557317	0,095766	0,095766	0,653083	2	1
9	0,89285	0,195263	0,195263	1,088113	3	1
10	1,009103	0,04039	0,04039	1,049493	1	1
11	1,164765	0,564076	0,5	1,664765	2	0
12	1,177982	0,014555	0,014555	1,192537	1	1
13	1,731771	0,617116	0,5	2,231771	3	0
14	2,112367	0,134921	0,134921	2,247288	1	1
15	2,202635	0,00391	0,00391	2,206545	2	1
16	2,470698	0,019783	0,019783	2,490481	2	1
17	2,525831	0,717632	0,5	3,025831	3	0
18	2,653784	0,071284	0,071284	2,725068	1	1
19	2,673485	0,070466	0,070466	2,743951	2	1
...
100028	9999,602	0,0152	0,0152	9999,617	1	1
100029	9999,609	0	0	9999,609	0	0

В результате ряда имитационных экспериментов с указанным набором параметров были получены следующие характеристики рассматриваемой СМО:

- 1) средняя доля заявок, получивших полное обслуживание, равна 0,852;

2) средняя доля заявок, начавших обслуживание, но не закончивших его по причине окончания срока жизни, равна 0,048;

3) средняя доля заявок, не попавших на обслуживание по причине занятости всех каналов в момент прихода требований, равна 0,1.

Расчеты, выполненные с помощью аналитической модели, полученной в разделе 3.1, для аналогичной системы дали следующие результаты: доля заявок, получивших полное обслуживание, равна 0,855, доля заявок, получивших неполное обслуживание, равна 0,045, доля заявок, не попавших на обслуживание, равна 0,1. Это свидетельствует о высокой степени валидности имитационной модели.

На рисунке 3.2.8 приведены кривые зависимостей доли полностью обслуженных заявок от времени жизни заявки для разных сочетаний параметров λ и μ .

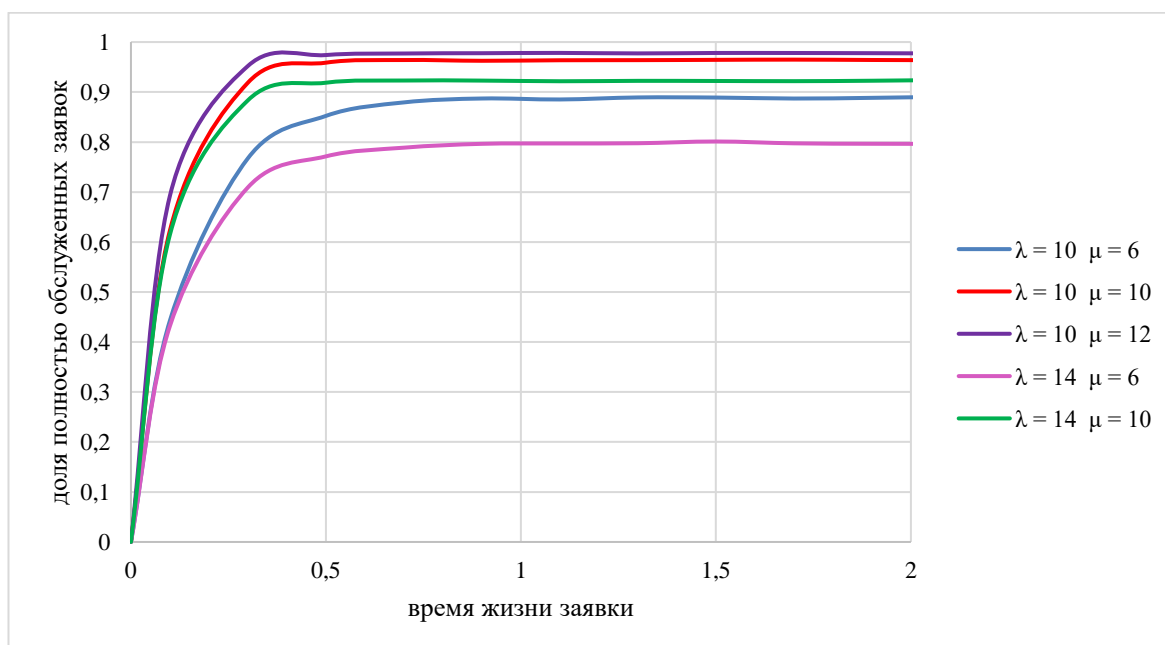


Рисунок 3.2.8 – Зависимости долей полностью обслуженных заявок от времени их жизни в трехканальной СМО без очереди

Графики на этом рисунке демонстрируют ту же тенденцию к «выходу на плато» при увеличении времени жизни заявки, что и в одноканальной СМО.

На рисунке 3.2.9 представлены зависимости долей полностью обслуженных требований от времени жизни заявки для одних и тех же входных параметров и разного числа каналов.

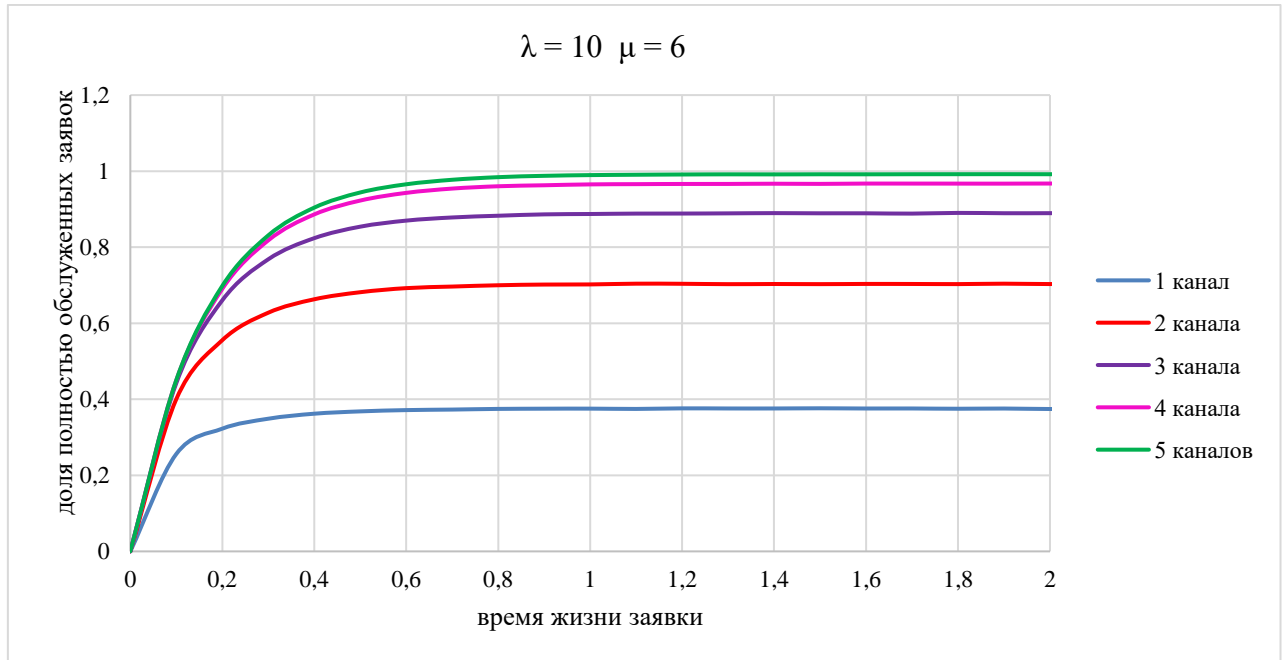


Рисунок 3.2.9 – Зависимости долей полностью обслуженных заявок от времени их жизни в многоканальной СМО без очереди

Графики показывают, что при указанных средних интенсивностях поступления и обслуживания заявок при времени жизни заявок, большем 1 при наличии 5 каналов, весь поток требований будет обработан. По мере уменьшения числа задействованных каналов доля заявок, получивших полное обслуживание, существенно снижается.

3.2.2 Системы обслуживания ординарного входного потока заявок с очередью

Одноканальная СМО с очередью

Теперь рассмотрим одноканальную систему массового обслуживания заявок с ограниченным временем жизни, образующих пуассоновский входной поток и имеющих возможность ожидать обслуживания в очереди в пределах своего срока жизни. Примем такую дисциплину обслуживания, при которой первой обслуживается заявка, пришедшая раньше всех (FIFO). Будем считать закон распределения значений времени обслуживания заявок экспоненциальным. Алгоритм работы модели, имитирующей работу такой системы, отличается от представленного на рисунке 3.2.2 только условным блоком, в котором со временем освобождения канала сравнивается не время прихода заявки, а ее срок жизни.

При проведении большой серии экспериментов с указанными входными параметрами найдена средняя доля заявок, получивших полное обслуживание. Она составляет приблизительно 56,4%. Средняя доля заявок, принятых в работу, но обслуженных полностью составляет примерно 43,6%. Средняя доля заявок, не попавших на обслуживание совсем, равна 0, то есть абсолютно все заявки, пришедшие в систему, успели дождаться начала обслуживания, но не все его завершили.

В таблице 3.2.3 продемонстрирована часть массива результатов работы означенного алгоритма для тех же входных данных, что и в примере выше.

Таблица 3.2.3 – Часть массива результатов эксперимента по имитации одноканальной СМО с очередью и простейшим входным потоком заявок

№ п/п	Момент времени прихода заявки	Момент времени окончания жизни	Момент начала обслуживания	Расчетное время обработки заявки	Длительность обработки заявки	Момент времени выхода заявки из системы	Индикатор полного обслуживания
1	0,041657	0,541657	0,041657	0,466007	0,466007	0,507664	1
2	0,042923	0,542923	0,507664	0,25916	0,035259	0,542923	0
3	0,164112	0,664112	0,542923	0,458341	0,121189	0,664112	0
4	0,210255	0,710255	0,664112	0,226876	0,046143	0,710255	0
5	0,479284	0,979284	0,710255	0,193976	0,193976	0,904231	1
6	0,528382	1,028382	0,904231	0,119126	0,119126	1,023357	1
7	0,599066	1,099066	1,023357	0,022414	0,022414	1,045771	1
8	0,667438	1,167438	1,045771	0,036758	0,036758	1,082529	1
9	0,898878	1,398878	1,082529	0,058052	0,058052	1,140581	1
10	1,173344	1,673344	1,173344	0,247213	0,247213	1,420557	1
11	1,178174	1,678174	1,420557	0,10637	0,10637	1,526927	1
12	1,234354	1,734354	1,526927	0,081933	0,081933	1,60886	1
13	1,300316	1,800316	1,60886	0,305248	0,191456	1,800316	0
14	1,365115	1,865115	1,800316	0,121015	0,064799	1,865115	0
15	1,654397	2,154397	1,865115	0,021014	0,021014	1,886129	1
16	1,768189	2,268189	1,886129	0,017637	0,017637	1,903766	1
17	1,950999	2,450999	1,950999	0,018897	0,018897	1,969896	1
...
99867	9999,924	10000,42	10000,33	0,187983	0,095335	10000,42	0
99868	9999,933	10000,43	10000,42	0,085273	0,008998	10000,43	0

На рисунке 3.2.10 представлены графики зависимостей средних долей полностью обработанных, частично обработанных и необработанных заявок от времени их жизни для фиксированных $\lambda = 10$ и $\mu = 6$, полученных в результате многократно проведенных симуляционных опытов.

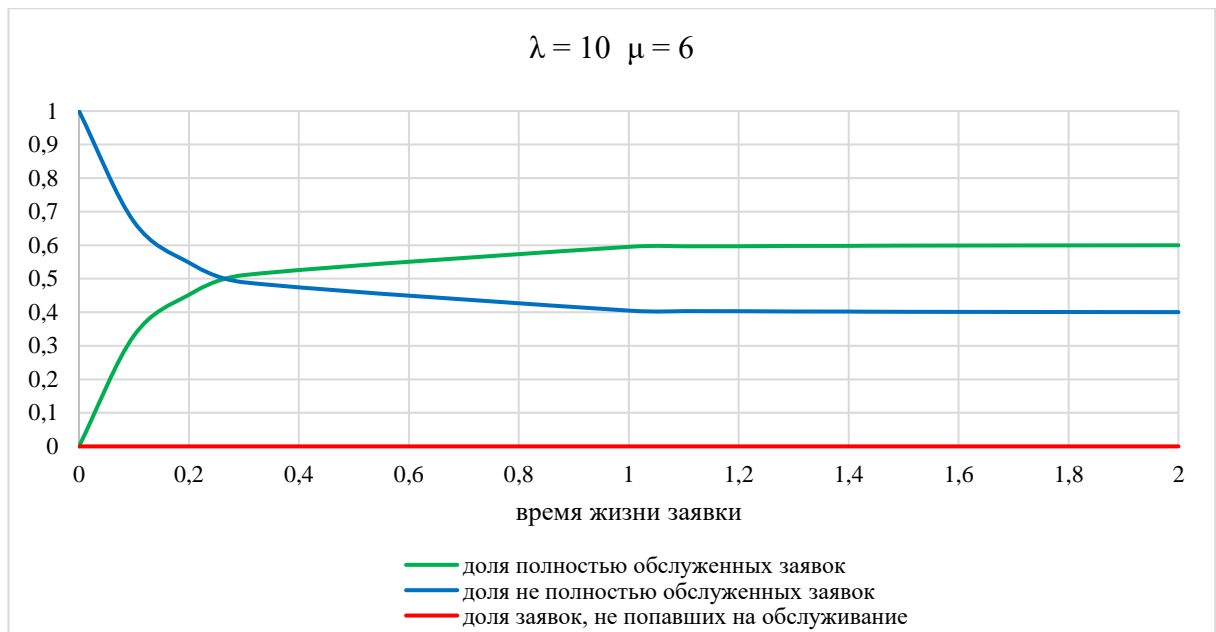


Рисунок 3.2.10 – Зависимости долей заявок от времени их жизни в одноканальной СМО с накопителем

Как и в варианте системы с отсутствием опции накопления очереди, здесь есть тенденция к «выходу на плато» кривых зависимостей долей от времени жизни заявки. Причем, если интенсивность обслуживания не меньше интенсивности входного потока, то доля полностью обслуженных требований стремится к относительной пропускной способности аналогичной системы без ограничения времени жизни заявки, которая равна 1. Если интенсивность обслуживания меньше интенсивности входного потока, то относительная пропускная способность системы с ограничением времени жизни заявки стремится к значению, меньшему 1. Оно зависит от интенсивностей входного потока заявок и процесса обслуживания. Это иллюстрируют графики зависимостей средних долей полностью обслуженных заявок для разных сочетаний параметров λ и μ , полученных в результате большой серии экспериментов, представленные на рисунке 3.2.11.

Как и в случае с одноканальной СМО без опции накопления очереди, здесь тоже можно определить границы диапазона значений, в рамках которого время жизни заявок имеет влияние на функционирование системы.

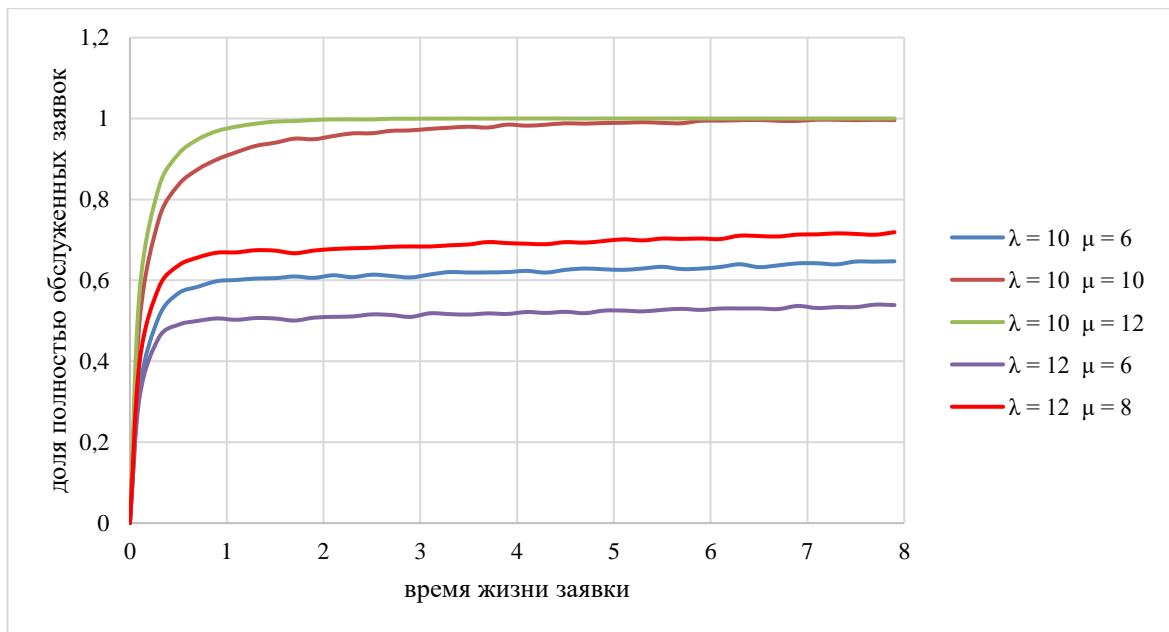


Рисунок 3.2.11 – Зависимости долей полностью обслуженных заявок от времени их жизни в одноканальной СМО с очередью

Для различных интенсивностей входного потока и обслуживания были проведены многочисленные опыты, позволившие найти значения времени жизни заявок, при которых система стремится к пороговым значениям относительной пропускной способности (рис. 3.2.12).

		λ									
		1	2	3	4	5	6	7	8	9	10
μ	1	1,2	0,8	0,6	0,5	0,4	0,4	0,4	0,3	0,3	0,3
	2	1,1	1,1	0,9	0,7	0,6	0,5	0,5	0,4	0,4	0,4
	3	1	1,1	0,9	0,8	0,7	0,6	0,5	0,5	0,5	0,4
	4	0,9	0,9	0,9	0,8	0,8	0,7	0,6	0,6	0,5	0,5
	5	0,8	0,8	0,8	0,9	0,9	0,8	0,7	0,6	0,6	0,5
	6	0,7	0,8	0,8	0,8	0,8	0,8	0,8	0,7	0,6	0,6
	7	0,7	0,7	0,7	0,8	0,8	0,8	0,7	0,7	0,7	0,6
	8	0,6	0,7	0,7	0,8	0,8	0,7	0,7	0,8	0,7	0,6
	9	0,6	0,6	0,6	0,7	0,7	0,7	0,7	0,8	0,7	0,6
	10	0,5	0,6	0,6	0,6	0,6	0,7	0,7	0,7	0,7	0,7

Рисунок 3.2.12 – Пороговые значения времени жизни заявок для одноканальных СМО с накопителем при высоких интенсивностях
Многоканальная СМО с очередью

Теперь рассмотрим многоканальную СМО, обслуживающую простейший поток заявок с ограниченным временем жизни и имеющую опцию накопления очереди. При моделировании работы этой СМО были полностью сохранены как порядок, так и условия имитационных экспериментов, описанных выше. Часть массива результатов симуляции функционирования СМО приведена в таблице 3.2.4.

Таблица 3.2.4 – Часть результатов эксперимента по имитации трехканальной СМО с очередью и простейшим входным потоком заявок

№ п/п	Момент времени прихода заявки	Момент времени окончания жизни	Момент начала обслуживания	Расчетное время обработки заявки	Длительность обработки заявки	Момент времени выхода заявки из системы	Номер канала	Индикатор обслуживания
1	0,094304	0,594304	0,094304	0,044146	0,044146	0,13845	1	1
2	0,126653	0,626653	0,126653	0,087358	0,087358	0,214011	2	1
3	0,279651	0,779651	0,279651	0,056977	0,056977	0,336628	3	1
4	0,315781	0,815781	0,315781	0,254377	0,254377	0,570158	1	1
5	0,321915	0,821915	0,321915	0,039624	0,039624	0,361539	2	1
6	0,335807	0,835807	0,336628	0,53203	0,499179	0,835807	3	0
7	0,547291	1,047291	0,547291	0,116176	0,116176	0,663467	2	1
8	0,739177	1,239177	0,739177	0,138056	0,138056	0,877233	1	1
9	0,947771	1,447771	0,947771	0,64686	0,5	1,447771	2	0
10	1,020254	1,520254	1,020254	0,018869	0,018869	1,039123	3	1
11	1,132662	1,632662	1,132662	0,284311	0,284311	1,416973	1	1
12	1,264228	1,764228	1,264228	0,037079	0,037079	1,301307	3	1
13	1,371137	1,871137	1,371137	0,268834	0,268834	1,639971	3	1
14	1,389134	1,889134	1,416973	0,124436	0,124436	1,541409	1	1
15	1,486735	1,986735	1,486735	0,049073	0,049073	1,535808	2	1
16	1,521352	2,021352	1,535808	0,029257	0,029257	1,565065	2	1
17	1,638829	2,138829	1,638829	0,505339	0,5	2,138829	1	0
...
99780	9999,751	10000,25	9999,751	0,117789	0,117789	9999,868	3	1
99781	9999,77	10000,27	9999,77	0,226757	0,226757	9999,997	2	1

Усредненные значения функциональных характеристик этой СМО, вычисленные по достаточно представительному множеству имитационных экспериментов, таковы:

- 1) средняя доля заявок, получивших полное обслуживание, равна 0,937;
- 2) средняя доля заявок, начавших обслуживание, но не закончивших его по причине окончания срока жизни, равна 0,063;
- 3) средняя доля заявок, не попавших на обслуживание, равна 0.

На рисунке 3.2.13 показаны кривые зависимостей доли полностью обслуженных заявок от времени жизни заявки для разных входных параметров λ и μ .

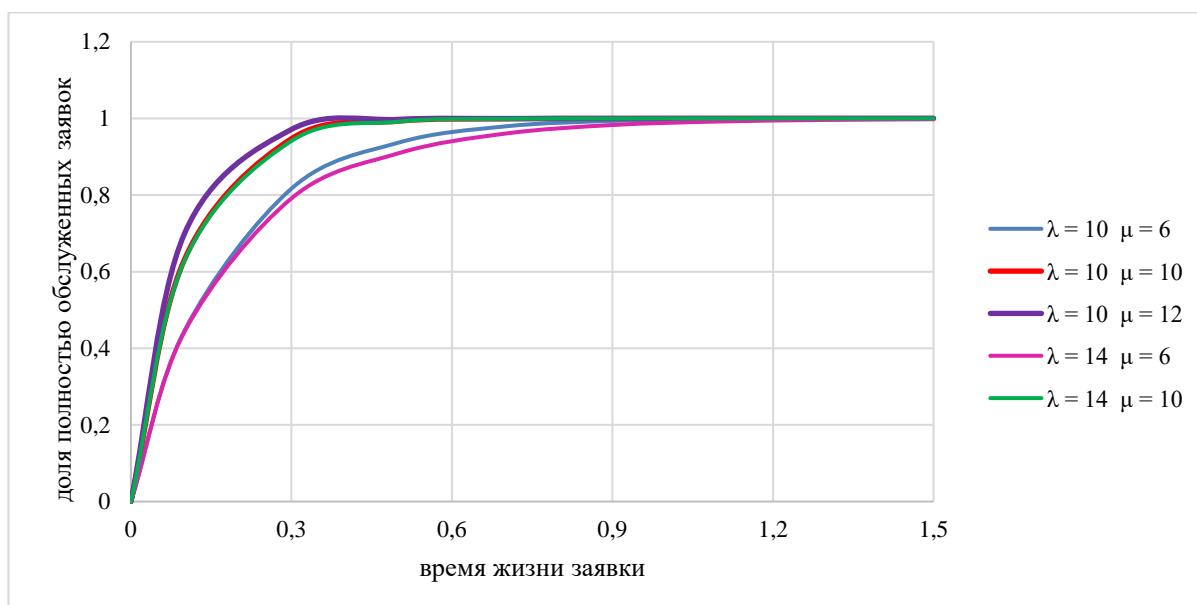


Рисунок 3.2.13 – Зависимости долей полностью обслуженных заявок от времени их жизни в трехканальной СМО с очередью

Графики на этом рисунке демонстрируют ту же тенденцию к «выходу на плато», что и в СМО с отказами, но, в отличие от нее, при увеличении времени жизни заявок относительная пропускная способность системы стремится к 1.

На рисунке 3.2.14 представлены зависимости долей полностью обслуженных требований от времени жизни заявок для одних и тех же входных параметров и разного числа каналов.

Графики на рисунке показывают, что при указанных средних интенсивностях поступления и обслуживания заявок при времени жизни заявок, большем 1, достаточно трех каналов для обслуживания всех заявок. При наличии двух каналов такой же результат достижим, если время жизни заявки превышает 2 врем. ед. А при наличии 1 канала система сможет обслужить не более 65% требований при любом времени жизни заявки для заданных средних интенсивностей поступления и обслуживания заявок.

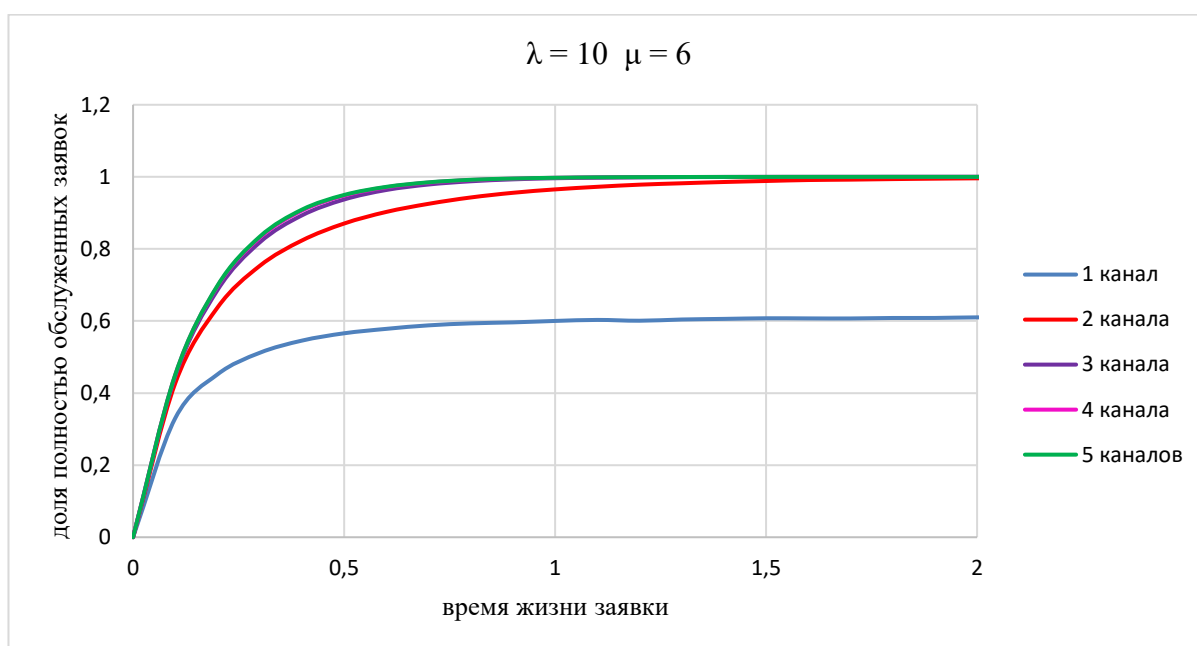


Рисунок 3.2.14 – Зависимости долей полностью обслуженных заявок от времени их жизни в многоканальной СМО с очередью

Результаты, приведенные на рисунках 3.2.3, 3.2.8, 3.2.11, 3.2.13 свидетельствуют о заметном отличии в поведении систем с опцией и без опции накопления очереди, которые выражаются не только в различии их относительных пропускных способностей, но и в различии пороговых значений времени жизни заявок, при которых чувствительность к этому ограничению теряется.

Таким образом, для систем, обрабатывающих простейший поток заявок с ограниченным временем жизни, независимо от наличия опции накопления

очереди, можно указать значение времени жизни, при котором теряется его влияние на поведение системы, а относительная пропускная способность системы стремится к определенному пороговому значению.

3.3 Имитационное моделирование систем массового обслуживания группового входного потока заявок

3.3.1 Системы обслуживания группового входного потока заявок с отказами

Одноканальная СМО с отказами

Рассмотрим одноканальную систему массового обслуживания заявок с ограниченным временем жизни без опции накопления очереди. Положим, что характер поступления заявок групповой, то есть одновременно в систему может прийти как одно, так и несколько требований. Примем, что количество заявок, попадающих в систему в составе группы, подчиняется равномерному закону распределения на интервале $[1, k]$, где k – максимально возможное количество требований в группе. Так же, как и выше, установим экспоненциальный закон распределения времени обслуживания.

Имитационный алгоритм отличается от представленного на рисунке 3.2.1 лишь тем, что на начальном этапе задается значение параметра k и при создании листа заявок происходит не только генерация интервалов времени, но и количества заявок в группе.

Для примера рассмотрим процесс обслуживания группового потока заявок со следующими входными параметрами:

- 1) верхняя граница интервала моделирования 1000 врем. ед. Нижняя – 0;
- 2) экспоненциальный закон распределения интервалов времени между поступлением групп заявок со средним $\lambda = 10$;
- 3) экспоненциальный закон распределения значений времени обработки заявок со средним $\mu = 6$;
- 4) срок жизни заявок с момента поступления 0,5 врем. ед.;

5) максимально возможное количество заявок в группе $k = 4$.

В таблице 3.3.1 представлена часть результатов работы описанного выше алгоритма для данных входных параметров.

Таблица 3.3.1 – Часть результатов эксперимента по имитации одноканальной СМО без очереди с групповым входным потоком заявок

№ п/п	Момент времени прихода заявки	Расчетное время обработки заявки	Длительность обработки заявки	Момент времени выхода заявки из системы	Индикатор полного обслуживания
250328	9990,857	0,01831	0,01831	9990,875	1
250329	9990,857	0	0	9990,857	0
250330	9991,187	0,779344	0,5	9991,687	0
250331	9991,187	0	0	9991,187	0
250332	9991,471	0	0	9991,471	0
250333	9991,471	0	0	9991,471	0
250334	9991,471	0	0	9991,471	0
250335	9991,626	0	0	9991,626	0
250336	9991,626	0	0	9991,626	0
250337	9991,626	0	0	9991,626	0
250338	9991,626	0	0	9991,626	0
250339	9991,777	0,221847	0,221847	9991,998	1
250340	9991,777	0	0	9991,777	0

Материалы этой таблицы демонстрируют, что одноканальная система без очереди для обработки группового потока заявок малоэффективна в силу невозможности одновременного обслуживания нескольких требований. Для приведенных выше входных параметров доля полностью обслуженных заявок составила всего 15%.

Многоканальная СМО с отказами

Теперь в том же порядке, что был использован ранее, и с помощью тех же инструментов исследуем поведение системы, имеющей в своем составе несколько эквивалентных каналов обслуживания. Время обслуживания, как и

выше, полагается случайной величиной, распределенной по экспоненциальному закону. Пришедшую в систему заявку принимает на обслуживание первый освободившийся ранее канал. После завершения работы над заявкой канал возвращается в очередь на прием новых требований.

Для того же входного набора параметров ($\lambda = 10$, $\mu = 6$, время жизни заявок 0,5 врем. ед., максимально возможное количество заявок в группе $k = 4$, число каналов 3) выполним симуляцию работы системы, которая не имеет опции накопления очереди. Фрагмент совокупности результатов работы имитационного алгоритма ее функционирования представлена в таблице 3.3.2.

Таблица 3.3.2 – Часть массива результатов эксперимента по имитации трехканальной СМО без очереди с групповым входным потоком заявок

№ п/п	Момент времени прихода заявки	Расчетное время обработки заявки	Длительность обработки заявки	Момент времени выхода заявки из системы	Номер канала	Индикатор полного обслуживания
292	11,69379	0,174054	0,174054	11,86785	3	1
293	11,72926	0,283102	0,283102	12,01236	2	1
294	11,66078	0	0	11,66078	0	0
295	11,94868	0,360618	0,360618	12,3093	1	1
296	12,02939	0,140072	0,140072	12,16946	3	1
297	12,02939	0,163132	0,163132	12,19252	2	1
298	12,05496	0	0	12,05496	0	0
299	12,07543	0	0	12,07543	0	0
300	12,07543	0	0	12,07543	0	0
301	12,07543	0	0	12,07543	0	0
302	12,31707	0,694782	0,5	12,81707	3	0
303	12,31707	0,06579	0,06579	12,38286	2	1
304	12,31707	0,019725	0,019725	12,3368	1	1
305	12,31707	0	0	12,31707	0	0
306	12,79382	0,08502	0,08502	12,87884	1	1
307	12,81915	0,208087	0,208087	13,02724	2	1
308	13,01454	0,192831	0,192831	13,20737	3	1
309	13,06601	0,092501	0,092501	13,15851	1	1
310	13,06601	0,151945	0,151945	13,21795	2	1
311	13,15851	0,083182	0,083182	13,24169	1	1

Усреднение по множеству, насчитывающему 100000 имитационных экспериментов с одинаковым набором входных параметров, позволило найти следующие характеристики рассматриваемой СМО:

- 1) средняя доля заявок, получивших полное обслуживание, равна 0,481.
- 2) средняя доля заявок, начавших обслуживание, но не закончивших его по причине окончания срока жизни, равна 0,031.
- 3) средняя доля заявок, не попавших на обслуживание совсем, равна 0,488.

На рисунке 3.3.1 приведены кривые зависимостей доли полностью обслуженных заявок от времени жизни заявок для различных сочетаний параметров λ и μ .

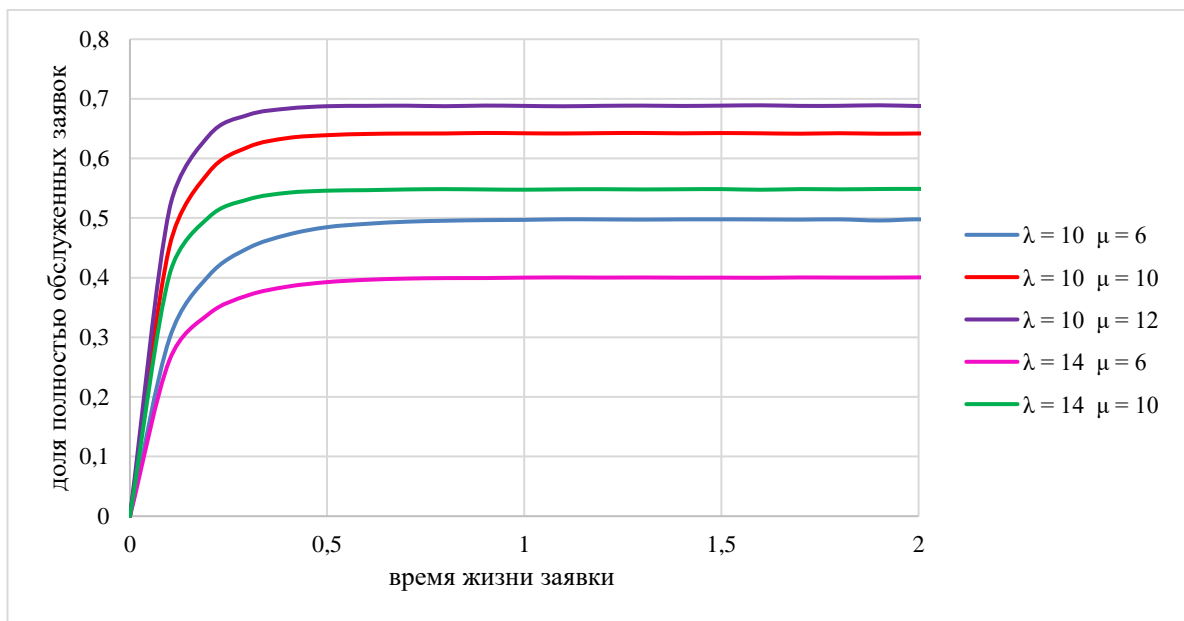


Рисунок 3.3.1 – Зависимости долей полностью обслуженных заявок группового потока от времени их жизни в трехканальной СМО без очереди

Полученные графики позволяют оценить пороговые значения относительных пропускных способностей СМО с групповым поступлением требований, которые значительно ниже, чем у аналогичной системы с ординарным входным потоком, кривые долей полностью обслуженных заявок которой представлены на рисунке 3.2.9. Разница оказывается существенной и превышает 30%.

На рисунке 3.3.2 представлены зависимости долей полностью обслуженных требований от времени жизни заявки для одних и тех же входных параметров и разного числа каналов.

Графики на рисунке убедительно доказывают, что для входного группового потока, где существует возможность одновременного поступления от 1 до 4 требований, среднего времени между приходом групп $1/10$ и среднего времени обработки $1/6$ наличие 5-ти каналов обслуживания позволяет достичь величины относительной пропускной способности системы не более 75%.

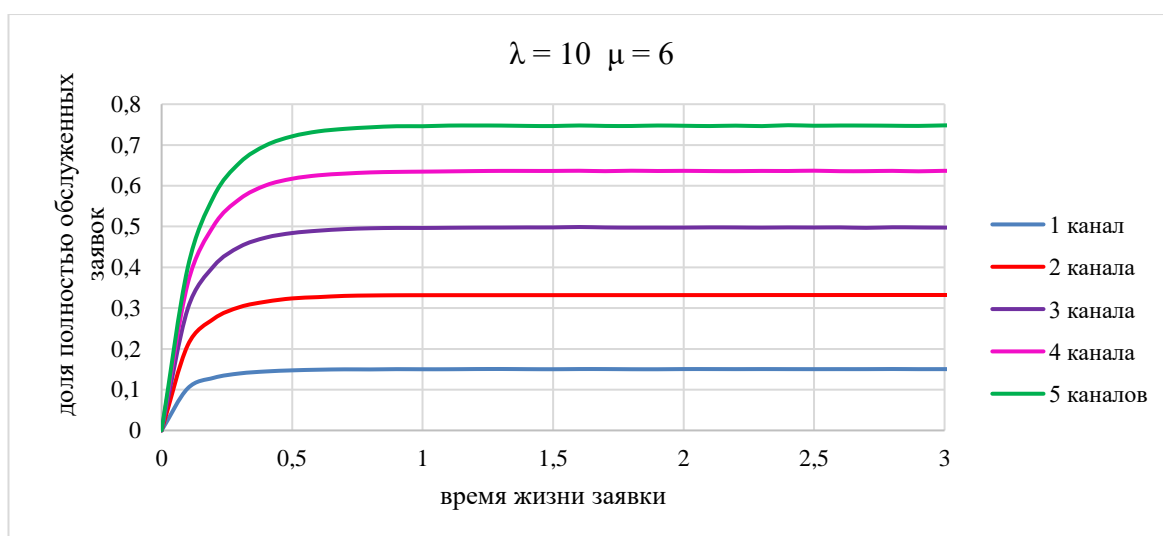


Рисунок 3.3.2 – Зависимости долей полностью обслуженных заявок группового потока от времени их жизни в многоканальной СМО без очереди

3.3.2 Системы обслуживания группового входного потока заявок с очередью

Одноканальная СМО с очередью

Для тех же входных параметров одноканальная система с опцией накопления очереди имеет несколько лучшие эксплуатационные характеристики: относительная пропускная способность системы составила около 24%; 34% заявок попали на обслуживание, но не завершили его; 42% всех пришедших требований были утеряны в процессе ожидания обслуживания в очереди.

В таблице 3.3.3 приведен фрагмент совокупности результатов работы имитационного алгоритма работы такой одноканальной системы.

Для анализа зависимости относительной пропускной способности системы от срока жизни заявок были проведены многократные имитационные эксперименты с разными комбинациями входных параметров.

Таблица 3.3.3 – Часть результатов эксперимента по имитации одноканальной СМО с очередью и групповым входным потоком заявок

№ п/п	Момент времени прихода заявки	Момент времени окончания жизни	Момент начала обслуживания	Расчетное время обработки заявки	Длительность обработки заявки	Момент времени выхода заявки из системы	Индикатор полного обслуживания
2173	84,5359	85,0359	84,76716	0,084934	0,084934	84,8521	1
2174	84,5359	85,0359	84,8521	0,022986	0,022986	84,87508	1
2175	84,8236	85,3236	84,87508	0,058237	0,058237	84,93332	1
2176	84,8236	85,3236	84,93332	0,073275	0,073275	85,00659	1
2177	85,09749	85,59749	85,09749	0,845345	0,5	85,59749	0
2178	85,09749	85,59749	85,09749	0	0	85,09749	0
2179	85,09749	85,59749	85,09749	0	0	85,09749	0
2180	85,10178	85,60178	85,59749	0,127596	0,004289	85,60178	0
2181	85,10178	85,60178	85,10178	0	0	85,10178	0
2182	85,13197	85,63197	85,60178	0,318154	0,030192	85,63197	0
2183	85,13197	85,63197	85,13197	0	0	85,13197	0
2184	85,13197	85,63197	85,13197	0	0	85,13197	0
2185	85,16894	85,66894	85,63197	0,34371	0,03697	85,66894	0
2186	85,16894	85,66894	85,16894	0	0	85,16894	0
2187	85,16894	85,66894	85,16894	0	0	85,16894	0
2188	85,16894	85,66894	85,16894	0	0	85,16894	0

Графики зависимостей средних долей полностью обслуженных заявок от времени их жизни для разных сочетаний параметров λ и μ представлены на рисунке 3.3.3. Они показывают, что этот вид СМО тоже обнаруживает стремление к определенному пороговому значению относительной пропускной способности при увеличении времени жизни заявок. Причем, величина доли полностью обработанных заявок крайне невелика для того, чтобы можно было

признать одноканальную СМО с очередью эффективным инструментом для обслуживания группового потока заявок с фиксированным временем жизни.

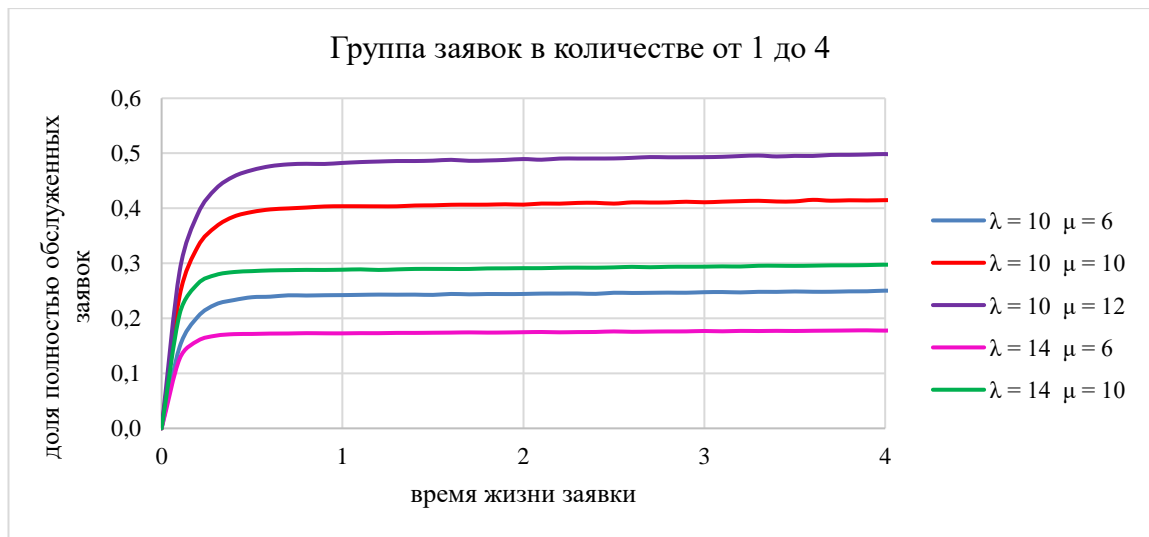


Рисунок 3.3.3 – Зависимости долей полностью обслуженных заявок группового входного потока от времени их жизни в одноканальной СМО с очередью

Многоканальная СМО с очередью

Теперь рассмотрим многоканальную СМО, обслуживающую групповой поток заявок с ограниченным временем жизни и имеющую опцию накопления очереди, куда заявки помещаются при отсутствии свободных каналов на момент прибытия. Будем считать, что обслуживание заявок осуществляется по принципу «первый пришел – первый обработан». Как только канал освобождается, он немедленно принимает на обслуживание первую в очереди заявку.

Для тех же входных параметров, что и у рассмотренной выше системы без накопителя, был проведен ряд имитационных экспериментов. Фрагмент совокупности результатов имитационного эксперимента функционирования СМО приведена в таблице 3.3.4.

Усредненные значения эксплуатационных характеристик системы, вычисленные по всему множеству проведенных имитационных экспериментов, таковы:

- 1) средняя доля заявок, получивших полное обслуживание, равна 0,66;

2) средняя доля заявок, начавших обслуживание, но не закончивших его по причине окончания срока жизни, равна 0,332;

3) средняя доля заявок, не попавших на обслуживание по причине окончания их времени жизни во время пребывания в очереди, равна 0,008.

Таблица 3.3.4 – Часть массива результатов эксперимента по имитации трехканальной СМО с очередью и групповым входным потоком заявок

№ п/п	Момент времени прихода заявки	Момент времени окончания жизни	Момент начала обслуживания	Расчетное время обработки заявки	Длительность обработки заявки	Момент времени выхода заявки из системы	Номер канала	Индикатор полного обслуживания
3142	131,747	132,247	131,747	0,030968	0,030968	131,778	1	1
3143	131,747	132,247	131,747	0,761409	0,5	132,247	2	0
3144	131,7611	132,2611	131,778	0,061591	0,061591	131,8396	1	1
3145	132,0908	132,5908	132,0908	0,013667	0,013667	132,1045	1	1
3146	132,0908	132,5908	132,0908	0,087049	0,087049	132,1779	3	1
3147	132,3454	132,8454	132,3454	0,55925	0,5	132,8454	1	0
3148	132,3454	132,8454	132,3454	0,018796	0,018796	132,3642	3	1
3149	132,3454	132,8454	132,3454	0,041361	0,041361	132,3867	2	1
3150	132,3454	132,8454	132,3642	0,41717	0,41717	132,7813	3	1
3151	132,4016	132,9016	132,4016	0,42296	0,42296	132,8246	2	1
3152	132,4016	132,9016	132,7813	0,105887	0,105887	132,8872	3	1
3153	132,5069	133,0069	132,8246	0,47799	0,182286	133,0069	2	0
3154	132,5069	133,0069	132,8454	0,22688	0,1615	133,0069	1	0
3155	132,5069	133,0069	132,8872	0,033847	0,033847	132,9211	3	1

На рисунке 3.3.4 приведены кривые зависимостей доли полностью обслуженных заявок от времени жизни заявки для разных сочетаний параметров λ и μ . Кривые, изображенные на рисунке, демонстрируют заметно большую эффективность системы с опцией накопления очереди по сравнению с системой без нее. Причем, для системы с четырьмя каналами обслуживания при

времени жизни заявок, большем 2, разница между значениями относительной пропускной способности составляет более 30%.

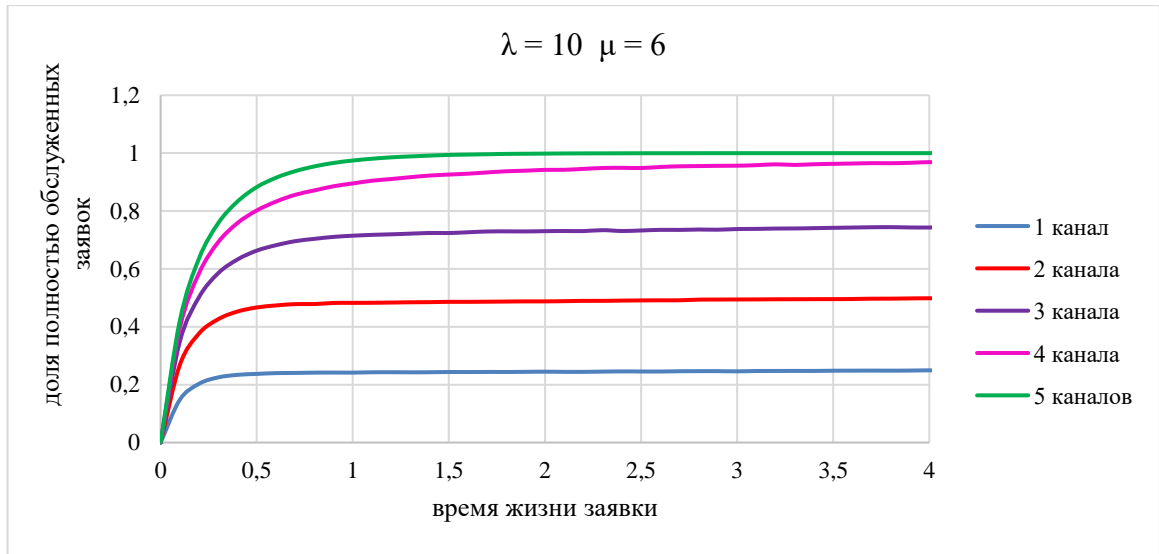


Рисунок 3.3.4 – Зависимости долей полностью обслуженных заявок группового потока от времени их жизни в многоканальной СМО с очередью

Результаты предельной относительной пропускной способности для некоторых значений времени жизни заявок для системы с очередью и без нее, обслуживающей групповой поток заявок с максимальным размером группы, равным 4, представлены в таблице 3.3.5.

Таблица 3.3.5 – Значения относительной пропускной способности систем с очередью и без нее

$\lambda = 10, \mu = 6, \text{ размер группы } 1-4$										
время жизни заявки	0,3		0,5		1		2		3	
наличие накопителя	-	+	-	+	-	+	-	+	-	+
1 канал	0,14	0,23	0,15	0,23	0,15	0,24	0,15	0,25	0,15	0,25
2 канала	0,30	0,43	0,32	0,47	0,33	0,48	0,33	0,49	0,33	0,49
3 канала	0,45	0,59	0,48	0,66	0,50	0,71	0,50	0,73	0,50	0,74
4 канала	0,57	0,69	0,62	0,80	0,63	0,90	0,63	0,94	0,63	0,96
5 каналов	0,66	0,76	0,72	0,88	0,75	0,97	0,75	0,99	0,75	1,00

3.4 Сравнительный анализ результатов имитационного моделирования различных СМО

Для сравнения производительности разных СМО возьмем те же входные параметры, что и раньше. А именно:

- 1) верхняя граница интервала моделирования 1000 врем. ед., нижняя – 0;
- 2) экспоненциальный закон распределения интервалов времени между поступлением заявок (групп заявок) со средним $\lambda = 10$;
- 3) экспоненциальный закон распределения значений времени обработки заявок со средним $\mu = 6$;
- 4) срок жизни заявок, отсчитываемый с момента поступления, 0,5 врем. ед.;
- 5) количество заявок в группе равно 1 для ординарного потока или распределено равномерно от 1 до 4 и от 1 до 10.

Проведем имитационные эксперименты для разного количества каналов. Добавим к сравнению значения относительной пропускной способности системы без опции накопления очереди, обслуживающей простейший поток заявок с неограниченным временем жизни, вычисленные по известным формулам. (3.4.1, 3.4.2).

Относительная пропускная способность одноканальной СМО с отказами и неограниченным временем жизни заявок:

$$Q = (1 + \rho)^{-1}, \quad (3.4.1)$$

многоканальной СМО с отказами:

$$Q = 1 - \frac{\rho^n}{n!} \left(\sum_{k=0}^n \frac{\rho^k}{k!} \right)^{-1}, \quad (3.4.2)$$

где ρ – средняя нагрузка на канал, n – количество каналов.

При наличии очереди для $\rho > 1$ одноканальная система попадает в поглощающее состояние и оказывается неработоспособной. Условие

поглощающего состояния при большем количестве каналов выглядит так:

$$\frac{\lambda}{\sum_{k=1}^n \mu_k} > 1. \text{ Здесь } \lambda - \text{ интенсивность потока заявок, } \mu_k - \text{ интенсивность } k\text{-го}$$

канала обслуживания, n – количество каналов.

В таблице 3.4.1 представлены значения относительной пропускной способности СМО с ограниченным временем жизни заявок для разного числа каналов обслуживания.

Данные этой таблицы свидетельствуют о том, что относительная пропускная способность различных видов систем с ограниченным временем жизни при увеличении числа каналов обслуживания стремится к значению доли полностью обслуженных заявок ординарного потока.

Таблица 3.4.1 – Значения относительной пропускной способности разных СМО с одинаковыми средними интенсивностями и одинаковым временем жизни заявок

Кол-во каналов	Без очереди				С очередью		
	простейший (аналит.)	ординарный	групповой 1-4	групповой 1-10	ординарный	групповой 1-4	групповой 1-10
1	0,375	0,368	0,147	0,067	0,566	0,237	0,109
2	0,658	0,682	0,324	0,150	0,871	0,467	0,217
3	0,840	0,853	0,484	0,233	0,937	0,663	0,322
4	0,938	0,922	0,617	0,312	0,948	0,802	0,425
5	0,980	0,922	0,721	0,389	0,950	0,883	0,523
6	0,994	0,944	0,800	0,460	0,950	0,920	0,615
7	1	0,950	0,856	0,528	0,950	0,936	0,694
8		0,950	0,892	0,589	0,950	0,943	0,762
9		0,950	0,918	0,644	0,950	0,947	0,815
10		0,950	0,931	0,694	0,950	0,949	0,855
...	
100		0,950	0,950	0,950	0,950	0,950	0,950

Графическое представление значений относительной пропускной способности систем наглядно иллюстрирует эту тенденцию (рис. 3.4.1). Еще одной характерной особенностью полученных результатов, заслуживающей упоминания, является стремление всех типов систем при одних и тех же средних

интенсивностях потоков и одинаковых значениях времени жизни заявок к одному финальному значению.

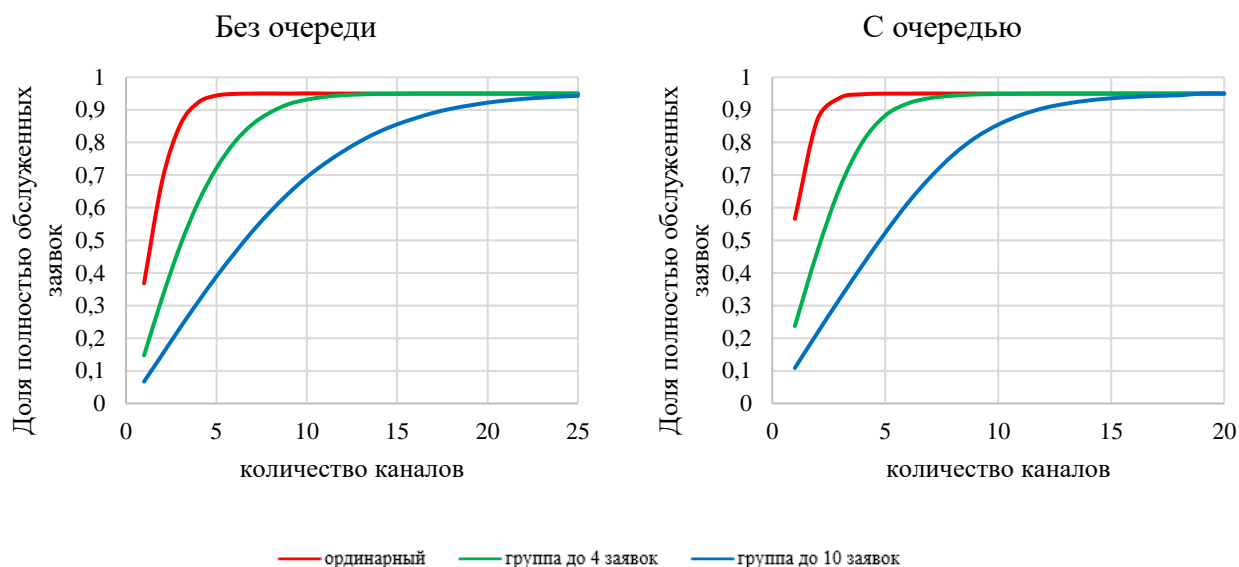


Рисунок 3.4.1. – Зависимости относительной пропускной способности системы от количества каналов обслуживания

По материалам масштабного имитационного эксперимента, выполненного для различных типов СМО с ограниченным временем жизни заявок, было установлено, что пороговая величина, к которой стремится доля полностью обслуженных заявок не зависит ни от интенсивности входного потока заявок, ни от его характера (группового или ординарного), ни от наличия или отсутствия очереди, а только от интенсивности обслуживающих устройств и времени жизни заявок. Сводные данные приведены в таблице 3.4.2.

Поскольку предельная относительная пропускная способность является ключевым показателем эффективности функционирования СМО, то целью исследования является не только ее нахождение, но и поиск значения, при котором система будет надежно работать с очень низкой вероятностью сбоев. Показатель в 95% можно считать приемлемым уровнем, на котором она стабильно справляется с нагрузкой. В этом аспекте представляет интерес нахождение минимального количества каналов, которое потребуется для обслуживания 95% предельно возможного количества заявок.

Таблица 3.4.2 – Значения пороговой относительной пропускной способности СМО с ограниченным временем жизни заявок

		время жизни заявок									
		0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1
μ	1	0,1	0,18	0,26	0,33	0,39	0,45	0,50	0,55	0,59	0,63
	2	0,18	0,33	0,45	0,55	0,63	0,70	0,75	0,80	0,84	0,86
	3	0,26	0,45	0,59	0,70	0,78	0,83	0,88	0,91	0,93	0,95
	4	0,33	0,55	0,70	0,80	0,87	0,91	0,94	0,96	0,97	0,98
	5	0,40	0,63	0,78	0,87	0,92	0,95	0,97	0,98	0,99	0,99
	6	0,45	0,70	0,83	0,91	0,95	0,97	0,98	0,99	1	1
	7	0,50	0,75	0,88	0,94	0,97	0,99	0,99	1	1	1
	8	0,55	0,80	0,91	0,96	0,98	0,99	1	1	1	1
	9	0,59	0,84	0,93	0,97	0,99	1	1	1	1	1
	10	0,63	0,86	0,95	0,98	0,99	1	1	1	1	1

Для решения этой задачи были проведены имитационные эксперименты с различными комбинациями входных параметров. Для системы без опции накопления очереди часть множества результатов опытов, проведенных для средней интенсивности $\lambda = 6$ прихода групп заявок размером до 10 штук представлена на рисунке 3.4.2.

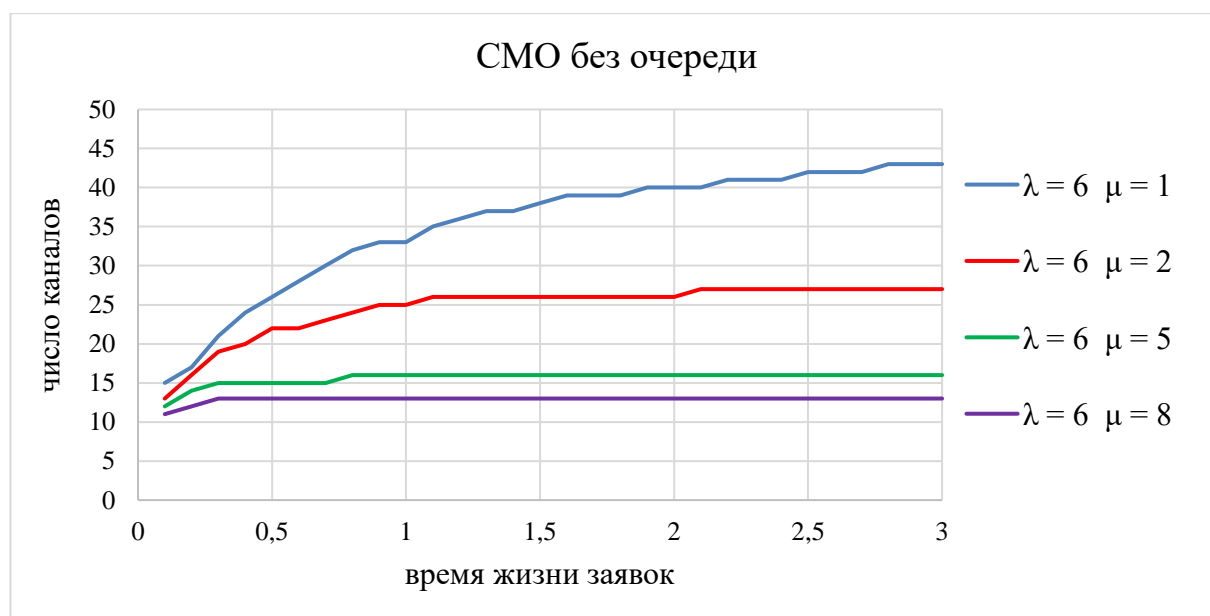


Рисунок 3.4.2 – Минимальное число каналов для выхода системы без накопителя на 95%-ю мощность

Эти графики демонстрируют очевидную тенденцию к остановке процесса роста числа каналов при увеличении времени жизни заявок. Анализ поведения фиолетовой кривой, соответствующей $\lambda = 6$ и $\mu = 8$, говорит о том, что для системы с такими параметрами нужно 13 каналов, чтобы обеспечить относительную пропускную способность, близкую к максимальной, независимо от времени жизни заявок. Такая система при значении времени жизни заявок, равном 0,1, обеспечит долю полностью обслуженных заявок в размере 0,52, а при значении 2 – 0,96. Синяя кривая, соответствующая $\lambda = 6$ и $\mu = 1$, демонстрирует более разнообразные результаты, и стабилизируется только при времени жизни заявок, равном 2,8. Для системы с такими параметрами при времени жизни заявки, равном 0,5, нужно 26 каналов, чтобы обеспечить относительную пропускную способность 0,37, а при времени жизни заявки, равном 2, нужно 40 каналов, чтобы обеспечить относительную пропускную способность 0,82.

Для системы с опцией накопления очереди часть массива результатов опытов, проведенных для средней интенсивности $\lambda = 6$ прихода групп заявок размером до 10 штук представлена на рисунке 3.4.3.

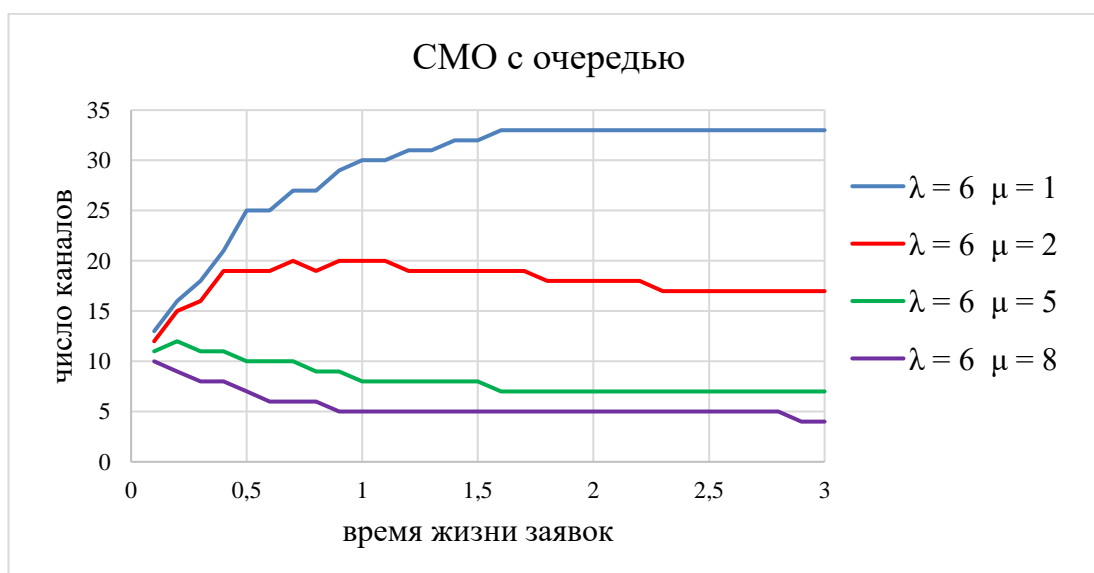


Рисунок 3.4.3 – Минимальное число каналов для выхода системы с накопителем на 95%-ю мощность

Здесь при увеличении времени жизни заявки также просматривается тенденция к стабилизации значения количества каналов обслуживания, необходимого для выхода на 95%-ю мощность. Для параметров $\lambda = 6$ и $\mu = 8$ при времени жизни заявок, равном 0,1, 10 каналов могут обеспечить относительную пропускную способность, близкую к максимальной и составляющую приблизительно 0,53. А при времени жизни заявок, равном 1, достаточно 5 каналов для обеспечения доли полностью обслуженных заявок в размере 0,96.

Таким образом, можно заключить, что время жизни заявок оказывает влияние на работу системы в определенном диапазоне его значений, границы которого зависят от интенсивностей входного потока заявок и процесса обслуживания. При выходе за эти границы система теряет чувствительность к ограничению и выходит на стабильные показатели эффективности.

Обнаружено, что в СМО с накопителем и групповым поступлением требований при наличии ограничения на время жизни заявок и интенсивности обслуживания, меньшей интенсивности входного потока, относительная пропускная способность системы имеет предельное значение, меньшее 1. Выявлено, что относительная пропускная способность системы с ограниченным временем жизни при увеличении числа каналов обслуживания стремится к доли полностью обслуженных заявок ординарного потока. Установлено, что пороговая величина, к которой стремится доля полностью обслуженных заявок не зависит ни от интенсивности входного потока заявок, ни от его характера (группового или ординарного), ни от наличия очереди, а зависит только от интенсивности обслуживающих устройств и времени жизни заявок. Для систем с различными параметрами расчетным путем получено минимальное число каналов обслуживания, необходимое для достижения 95%-го значения установленной максимальной относительной пропускной способности, считающимся достаточным для стабильной работы системы.

Часть изложенных результаты представлена в работе [141, 142].

4 ПОВЫШЕНИЕ ПОКАЗАТЕЛЯ ЭКОНОМИЧЕСКОЙ ЭФФЕКТИВНОСТИ СИСТЕМ С ФИКСИРОВАННЫМ ОГРАНИЧЕНИЕМ ВРЕМЕНИ ЖИЗНИ ЗАЯВОК

4.1 Математическое моделирование системы обслуживания группового входного потока с очередью

Рассмотрим одноканальную систему массового обслуживания, предназначенную для обработки заявок, которые имеют ограниченное время жизни, одинаковое для всех. Предполагается, что заявки поступают группами фиксированного объема и помещаются в накопитель, откуда отбираются на обслуживание. В качестве примеров таких систем можно привести розничную реализацию товаров с ограниченным сроком годности, некоторые многостадийные производства, где строго ограничено время между последовательными операциями и т.п.

Как и в описанном выше случае, поток потерь заявок двухкомпонентен, то есть заявки могут выбывать как из очереди, так и в процессе обслуживания по истечении срока жизни. Но в отличие от системы, описанной в разделе 3.3 настоящей работы, рассматриваемая СМО обладает важными особенностями, определяемыми групповым характером входного потока заявок. А именно, по окончании срока жизни группы система теряет в общем случае множество заявок одновременно. Причем, только одна заявка теряется, будучи в процессе обслуживания, все остальные выбывают из накопителя, находясь в процессе ожидания. Заметим также, что в отличие от описанных выше, в рассмотренной в этом разделе СМО, предполагается, что срок жизни заявки значительно превышает время ее обработки. Следовательно, концепцию построения модели одноканальной СМО с очередью, в основу которой положены свойства простейшего потока заявок, применить в рамках предлагаемой постановки не представляется возможным.

Особенности рассматриваемой здесь СМО обосновывают целесообразность сосредоточения основного внимания на массив заявок, которые уходят из системы необслуженными. Это обуславливает возникновение задачи отыскания возможности минимизации доли потерянных требований и экономических потерь, которые при этом неизбежны. Какие меры следует предпринять для интенсификации обслуживания без введения в эксплуатацию дополнительных устройств обработки: очевидно, что для решения подобных задач критерии оценки качества работы СМО, такие как средняя длина очереди, среднее время ожидания и т.п., традиционно используемые в теории массового обслуживания, оказываются малоэффективными. Экономическое содержание задачи требует использования иных критериев эффективности и введения дополнительного параметра, обеспечивающего возможность управления состоянием системы. Содержательный смысл такого параметра можно продемонстрировать на примере сферы реализации товаров с ограниченным сроком годности, приняв за него стоимость продажи товара. Очевидно, что, уменьшая цену реализации, можно ускорить темп процесса продажи. Тогда цель работы СМО можно будет сформулировать как достижение максимальной выручки от продажи того объема товаров, который будет реализован до истечения их срока годности. В сферах оказания услуг и производства цена тоже может служить дополнительной характеристикой эффективности обслуживания. В этом случае ее понижение может быть осуществлено за счет уменьшения издержек на обработку требований, например, путем исключения дублирующих операций или второстепенных этапов обслуживания.

Для полноты картины следует дать некоторые разъяснения по причинам возникновения ситуаций, при которых СМО не сможет обслужить всю группу заявок до окончания их срока жизни. Это может произойти из-за ошибок прогнозирования поведения системы или некорректно определенных первоначально параметрах. Например, при возникновении ситуаций, когда объем группы заявок по какой-то причине оказывается больше обычного. Тогда администратору системы следует принять решение об интенсификации процесса

обслуживания. Задача состоит в том, чтобы сделать это в такой момент и выбрать такую величину управляющего воздействия, чтобы получить максимально возможный эффект.

Итак, положим, что обработка каждой заявки имеет установленную стоимость, одинаковую для всех. В качестве меры эффективности рассматриваемой системы примем величину суммарной выручки, полученной за обслуживание той части заявок группы, которую система успеет обработать за установленный период жизни требований. В данной постановке предполагается, что темп обслуживания заявок задан некоторой функцией времени, которая заведомо известна или может быть найдена экспериментально. В тот момент времени, когда становится очевидным, что обычная интенсивность обслуживания не позволит обработать всю группу заявок, и это приведет к неприемлемым потерям, следует интенсифицировать процесс путем удешевления стоимости обработки каждой заявки. Суть задачи состоит в отыскании момента времени приложения управляющего воздействия и его величины для получения максимально возможного эффекта.

Для построения модели примем следующие обозначения:

Q – объем группы заявок;

t_{end} – время жизни группы заявок;

R_1 – установленная первоначально стоимость обработки одной заявки;

$v_1(t)$ – функция интенсивности при первоначально установленной стоимости R_1 ;

t – момент времени принятия управляющего воздействия для интенсификации процесса обслуживания, $0 < t < t_{\text{end}}$;

R_2 – новая стоимость обработки одной заявки, установленная в момент принятия управляющего воздействия;

$v_2(t)$ – функция интенсивности при новой стоимости R_2 .

Для нахождения количества заявок группы, которую система успеет обработать за установленный период их жизни, можно применить формулу:

$$Q_1 = \int_0^{t_{\text{end}}} v_1(t) dt. \quad (4.1.1)$$

Тогда выручка за обслуживание заявок в количестве Q_1 вычисляется следующим образом:

$$M_1 = R_1 * Q_1. \quad (4.1.2)$$

При применении в некоторый момент времени t управляющего воздействия общая выручка может быть найдена по формуле:

$$M_2 = R_1 \int_0^t v_1(t) dt + R_2 \int_t^{t_{\text{end}}} v_2(t) dt, \quad (4.1.3)$$

при соблюдении условия:

$$Q_2 = \int_0^t v_1(t) dt + \int_t^{t_{\text{end}}} v_2(t) dt < Q. \quad (4.1.4)$$

На рисунке 4.1.1 штриховкой выделены подграфики функций $v_1(t)$ и $v_2(t)$, площади которых соответствуют объемам заявок, обслуженных системой до окончания их срока жизни при применении управляющего воздействия в момент времени t .

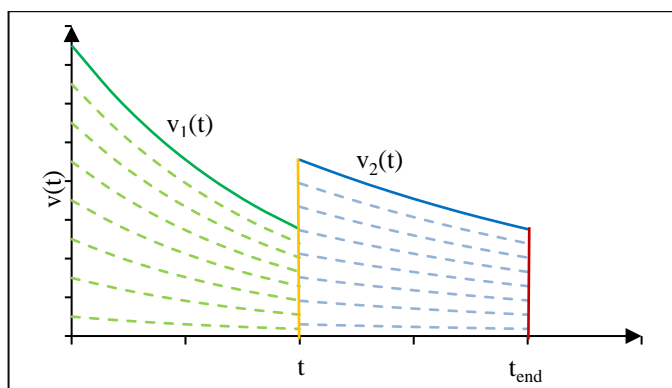


Рисунок 4.1.1 – Объем обработанных заявок при интенсификации

Требуется найти максимальное значение меры эффективности системы при выполнении управленческих действий, т.е. решить задачу:

$$M_2 \rightarrow \max. \quad (4.1.5)$$

Задача (4.1.3) эквивалентна нахождению максимума разности M_2 и M_1 , то есть преобразуется к виду:

$$M(t) = \int_t^{t_{\text{end}}} (R_2 v_2(t) - R_1 v_1(t)) dt \rightarrow \max. \quad (4.1.6)$$

Для поиска стационарной точки функции $M(t)$ используем необходимое условие существования экстремума функции одной переменной, то есть равенство нулю ее производной. Поскольку функция $M(t)$ фактически представляет собой определенный интеграл с переменным нижним пределом от непрерывной функции, то воспользовавшись соответствующей теоремой математического анализа [133], получим равенство:

$$M'(t) = -R_2 v_2(t) + R_1 v_1(t) = 0. \quad (4.1.7)$$

Решив уравнение 47, найдем точку t^* , подозрительную на экстремум. Для доказательства того, что она является точкой максимума функции $M(t)$ останется проверить достаточное условие существования точки экстремума или исследовать поведение функции в окрестности точки t^* .

В качестве примера рассмотрим возможный вид функции интенсивности обработки заявок, не противоречащий экономической парадигме, $y(t) = ke^{-\lambda t}$ ($\lambda > 0$). Примем, что форма функциональной зависимости скорости обслуживания от времени после применения управляющего воздействия не меняется, модифицируются лишь ее коэффициенты.

Допустим, первоначальный темп обслуживания заявок имеет вид:

$$v_1(t) = k_1 e^{-\lambda_1 t} (\lambda_1 > 0),$$

а темп обслуживания заявок после применения управляющего воздействия

$$v_2(t) = k_2 e^{-\lambda_2 t} (\lambda_2 > 0).$$

Тогда уравнение (33) преобразуется к виду:

$$M'(t) = -R_2 k_2 e^{-\lambda_2 t} + R_1 k_1 e^{-\lambda_1 t} = 0. \quad (4.1.8)$$

Решение уравнения (34) выглядит так:

$$t^* = \frac{\ln \frac{R_1 k_1}{R_2 k_2}}{\lambda_1 - \lambda_2}. \quad (4.1.9)$$

Поскольку время – величина неотрицательная, то числитель и знаменатель дроби в формуле (4.1.7) должны быть одноименного знака.

На рисунке 4.1.2 наглядно продемонстрировано влияние параметров λ и k на поведение кривой темпа обслуживания.

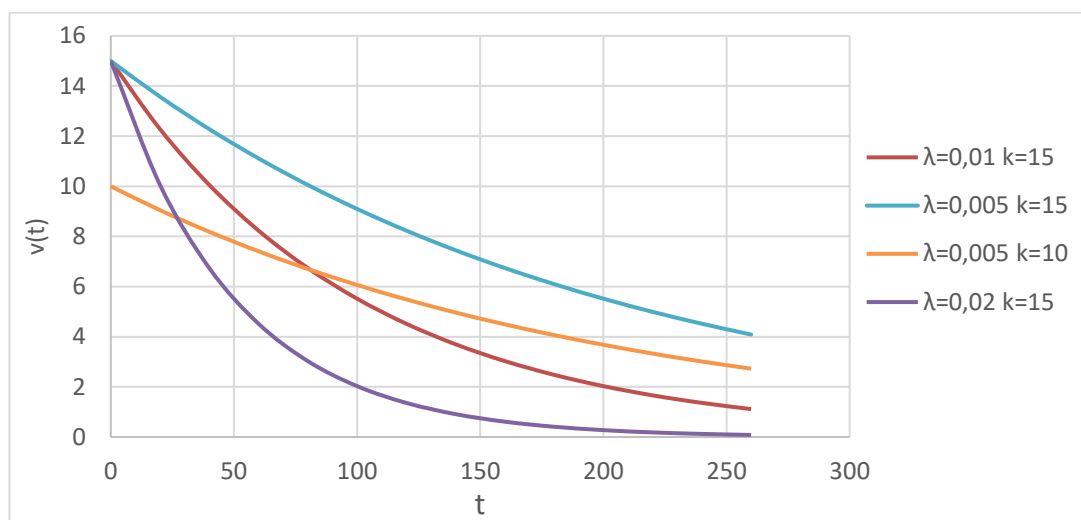


Рисунок 4.1.2 – Функции скорости обслуживания

Кривая, изображающая более высокую интенсивность обслуживания, лежит выше кривой с более низкой интенсивностью хотя бы на некотором временном интервале. Такое положение обеспечивается коэффициентом λ . Таким образом, верно неравенство:

$$\lambda_2 > \lambda_1, \quad (4.1.10)$$

а следовательно, должно выполняться и неравенство:

$$\frac{R_1 k_1}{R_2 k_2} > 1. \quad (4.1.11)$$

Для проверки достаточного условия существования точки экстремума найдем вторую производную функции $M(t)$ в точке t^* :

$$M''(t^*) = R_2 k_2 \lambda_2 e^{-\lambda_2 \left(\frac{\ln \frac{R_1 k_1}{R_2 k_2}}{\lambda_1 - \lambda_2} \right)} - R_1 k_1 \lambda_1 e^{-\lambda_1 \left(\frac{\ln \frac{R_1 k_1}{R_2 k_2}}{\lambda_1 - \lambda_2} \right)} =$$

$$= R_2 k_2 \left(\frac{R_1 k_1}{R_2 k_2} \right)^{\frac{-\lambda_2}{\lambda_1 - \lambda_2}} (\lambda_2 - \lambda_1)$$

При выполнении условий (4.1.10) и (4.1.11) $M''(t^*) < 0$, значит, t^* – точка максимума функции $M(t)$.

В качестве первого примера рассмотрим процесс со следующими параметрами:

$$Q = 1500;$$

$$t_{\text{end}} = 200;$$

$$R_1 = 100;$$

$$v_1(t) = 15e^{-0,01t},$$

$$R_2 = 90;$$

$$v_2(t) = 10e^{-0,005t}.$$

Без применения управляющего воздействия будут обработаны заявки в количестве:

$$Q_1 = \int_0^{200} 15e^{-0,01t} dt \approx 1296.$$

Мера эффективности системы составит 129600.

По формуле (49) найдем время принятия решения об удешевлении обслуживания:

$$t^* = \frac{1,67}{0,005} = 102,56 \approx 103.$$

Проверим выполнения условия (44):

$$Q_2 = \int_0^{103} 15e^{-0,01t} dt + \int_{103}^{200} 10e^{-0,005t} dt \approx 1423 < Q.$$

Найдем значение функции $M(t^*)$:

$$M(t^*) = \int_{103}^{200} (900e^{-0,005t} - 1500e^{-0,01t}) dt \approx 8081.$$

Рассмотрим другую величину управляющего воздействия: $R_2 = 80$ и соответствующую ему новую интенсивность обслуживания $v_2(t) = 8e^{-0,004t}$. Для данных характеристик процесса имеем:

$$t^* = \frac{2,34}{0,006} = 141,95 \approx 142.$$

Проверим выполнения условия (44):

$$Q_2 = \int_0^{142} 15e^{-0,01t} dt + \int_{142}^{200} 8e^{-0,004t} dt \approx 1372 < Q.$$

Найдем значение функции $M(t^*)$:

$$M(t^*) = \int_{142}^{200} (640e^{-0,004t} - 1500e^{-0,01t}) dt \approx 2816.$$

Приведенный пример показывает, что оптимальное время принятия управляющего воздействия зависит и от новой стоимости обслуживания заявки, и от параметров функции интенсивности обслуживания. Правильный выбор этих характеристик находится в компетенции лица, принимающего решение.

Изменим вид функции интенсивности на линейную и рассмотрим процесс обслуживания со следующими характеристиками:

Допустим, первоначальный темп обслуживания заявок имеет вид:

$$v_1(t) = a_1 t + b_1,$$

а темп обслуживания заявок после применения управляющего воздействия

$$v_2(t) = a_2 t + b_2.$$

Постановка задачи подразумевает увеличение интенсивности обслуживания при уменьшении коэффициента результативности заявок и падение темпа обслуживания со временем. Поэтому справедливы следующие неравенства:

$$R_2 < R_1, a_1, a_2 < 0, |a_1| > |a_2|. \quad (4.1.12)$$

На рисунке 4.1.3 продемонстрировано поведение линейных функций темпа обслуживания с разными параметрами.

Для заданных видов функций интенсивности обслуживания уравнение (4.1.7) преобразуется к виду:

$$M'(t) = -R_2(a_2t + b_2) + R_1(a_1t + b_1) = 0. \quad (4.1.13)$$

Решение уравнения (55) выглядит так:

$$t^* = \frac{R_1b_1 - R_2b_2}{R_2a_2 - R_1a_1}. \quad (4.1.14)$$

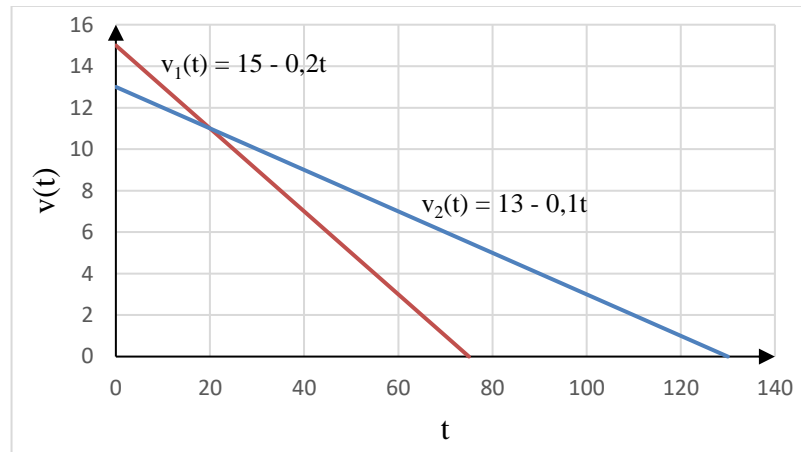


Рисунок 4.1.3 – Линейные функции скорости обслуживания

Поскольку время – величина неотрицательная, то числитель и знаменатель дроби в формуле (56) должны быть одноименного знака. Этот факт накладывает достаточно жесткие условия на характеристики процесса.

Для проверки достаточного условия существования точки экстремума найдем вторую производную функции \$M(t)\$ в точке \$t^*\$:

$$M''(t) = -R_2a_2 + R_1a_1 = R_1 \left(a_1 - \frac{R_2}{R_1} a_2 \right).$$

Очевидно, что при выполнении условий (52) \$M''(t) < 0\$, а значит, \$t^*\$ – точка максимума функции \$M(t)\$.

В качестве примера рассмотрим процесс со следующими параметрами:

$$Q = 700;$$

$$t_{\text{end}} = 50;$$

$$R_1 = 100;$$

$$v_1(t) = 15 - 0,2t,$$

$$R_2 = 90;$$

$$v_2(t) = 13 - 0,1t.$$

Найдем точку максимума по формуле (55): $t^* = 30$.

Проверим выполнения условия (46):

$$Q_2 = \int_0^{30} (15 - 0,2t)dt + \int_{30}^{50} (13 - 0,1t)dt = 540 < Q.$$

Теперь вычислим прирост выручки при применении управляющего воздействия в момент времени t^* :

$$M(t^*) = \int_{30}^{50} (90(13 - 0,1t) - 100(15 - 0,2t))dt = 2200.$$

Обратим внимание на то, что, как показано в примерах, для повышения эффективности работы рассматриваемой СМО нужно владеть полным спектром знаний о поведении системы. Этот факт обуславливает зависимость результата от знаний и навыков лица, осуществляющего управленческое воздействие.

Таким образом, построена математическая модели для одноканальной СМО, которая обрабатывает групповой входной поток требований, имеющих ограниченное время жизни. Существенными отличиями системы от классических являются два фактора. Первый – в момент окончания срока жизни система может потерять множество заявок одновременно. Второй – срок жизни заявки значительно превышает время ее обработки. Они обуславливают введение нестандартной меры эффективности функционирования системы – выручки за обработку той части требований, которые будут обслужены за установленный период их жизни. Темп обслуживания задан функцией времени, аналитическое выражение которой полагается известным, причем, при изменении интенсивности меняются только параметры функции, а ее вид остается неизменным. Показано, что при выполнении определенных условий можно найти оптимальное время приложения управляющего воздействия для интенсификации процесса обслуживания и увеличения выручки. Отмечено, что

решение задачи зависит от величины управляющего воздействия и вида функции, задающей интенсивность обслуживания. Чтобы успешно адаптировать модель для более широкого класса задач, необходимо применение методов имитационного моделирования.

Изложенные результаты представлены в работах [143,144].

4.2 Постановка и решение задачи повышения показателя экономической эффективности систем для детерминированного потока обслуживания

Математическая модель, предложенная в разделе 4.1 а построена для темпа обслуживания, заданного функцией времени с известным аналитическим выражением, вид которого не меняется при изменении интенсивности обработки. Предполагается, что известен не только вид зависимости темпа от времени, но и его параметры при определенных интенсивностях. Трансформируем этот постулат. Будем считать, что известна $v_1(t)$ – функция интенсивности при первоначально установленной стоимости R_1 и $v_2(R_1/R_2)$ – зависимость между темпом продажи и ценой товара, здесь R_2 – новая сниженная цена товара. Нужно, как и прежде, найти время и величину снижения стоимости товара для интенсификации процесса реализации и получения максимально возможной суммарной выручки.

Ввиду того, что математическая модель не предусматривает возможность решения задачи в такой постановке, была построена и программно реализована имитационная модель.

Эксперименты проводились с различными зависимостями темпа от времени: постоянной, линейной, экспоненциальной и квадратичной:

$$v_{11}(t) = 8;$$

$$v_{12}(t) = -0,06 t + 10,9;$$

$$v_{13}(t) = e^{2,52 - 0,01t};$$

$$v_{14}(t) = -0,00025 t^2 + 8,768.$$

Параметры этих зависимостей были подобраны так, чтобы объем реализованных товаров за выбранный период жизни заявок был примерно одинаков.

Определим входные параметры:

- $Q = 1000$ шт. – объем партии товара;
- $t_{\text{end}} = 96$ врем. ед. – срок годности партии товара;
- $R_1 = 100$ д. ед. – установленная первоначально цена товара.

Если не предпринимать никаких управляющих воздействий, то для данных входных параметров получим значения количества проданных товаров и выручки за него, представленные в таблице 4.2.1.

Таблица 4.2.1 – Значения выручки без применения управляющего воздействия

$v_1(t)$	Количество проданных товаров за 96 часов	Выручка
8	768	76800
$-0,06t + 10,9$	769	76900
$e^{2,52 - 0,01t}$	766	76600
$-0,00025t^2 + 8,7684$	768	76800

При определении вида функции $v_2(R_1/R_2)$ мы руководствовались практическими соображениями. На их основе были определены следующие варианты зависимостей между темпом продажи и соотношением цен:

$$v_{21}(t) = v_1(t) \left(\frac{R_1}{R_2}\right)^2,$$

$$v_{22}(t) = v_1(t) e^{\left(\frac{R_1}{R_2} - 1\right)},$$

$$v_{23}(t) = v_1(t) \sqrt{\left(\frac{R_1}{R_2}\right)},$$

$$v_{24}(t) = v_1(t) \ln\left(\frac{R_1}{R_2} + 1\right).$$

Для разных комбинаций функций v_{1i} и v_{2j} ($i, j = 1, 2, 3, 4$) были проведены численные эксперименты, заключающиеся в том, чтобы в некоторый момент времени снизить стоимость товара на определенную величину, увеличив тем самым темп обслуживания. Результаты экспериментов для функции v_{11} и соотношения v_{21} представлены на рисунке 4.2.1.

Графики демонстрируют интуитивно понятную логику – чем ближе точка принятия решения к окончанию срока годности товара, тем больше должна быть скидка на товар для получения большей выручки. Для рассматриваемых зависимостей оптимальным решением является снижение цены на 20 денежных единиц в момент времени 44. При принятии такого решения выручка составит 87040.

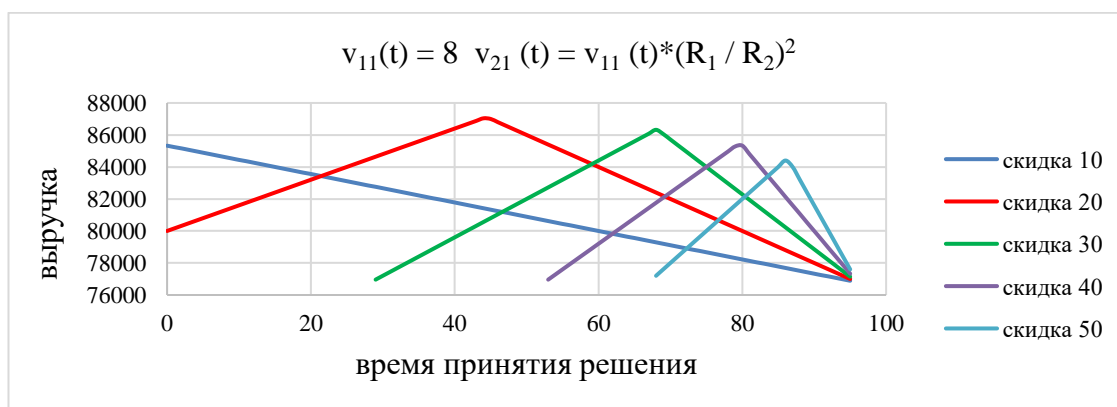


Рисунок 4.2.1 – Значения выручки для зависимостей v_{11} и v_{21} при принятии решения о снижении цены

Результаты экспериментов для функции v_{11} и соотношения v_{22} представлены на рисунке 4.2.2.

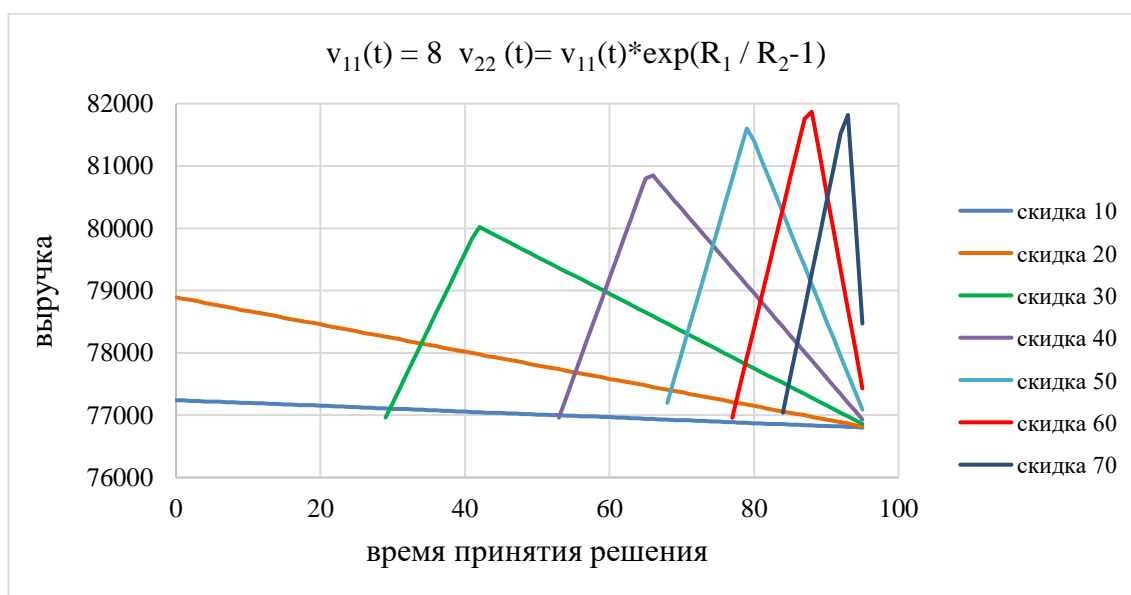


Рисунок 4.2.2 – Значения выручки для зависимостей v_{11} и v_{22} при принятии решения о снижении цены

В этом случае наилучший момент времени для принятия управляющего воздействия – 88 врем. ед. с начала продажи товара, а величина скидки – 60 д. ед. Выручка в таком случае увеличится на 5070 д. ед. по сравнению с первоначальным планом продажи без изменения цены. Если же снизить цену на 70 д. ед. в момент 93 врем. ед., то прирост выручки составит 5020 д. ед.

Результаты экспериментов для функции v_{13} и соотношения v_{21} представлены на рисунке 4.2.3.

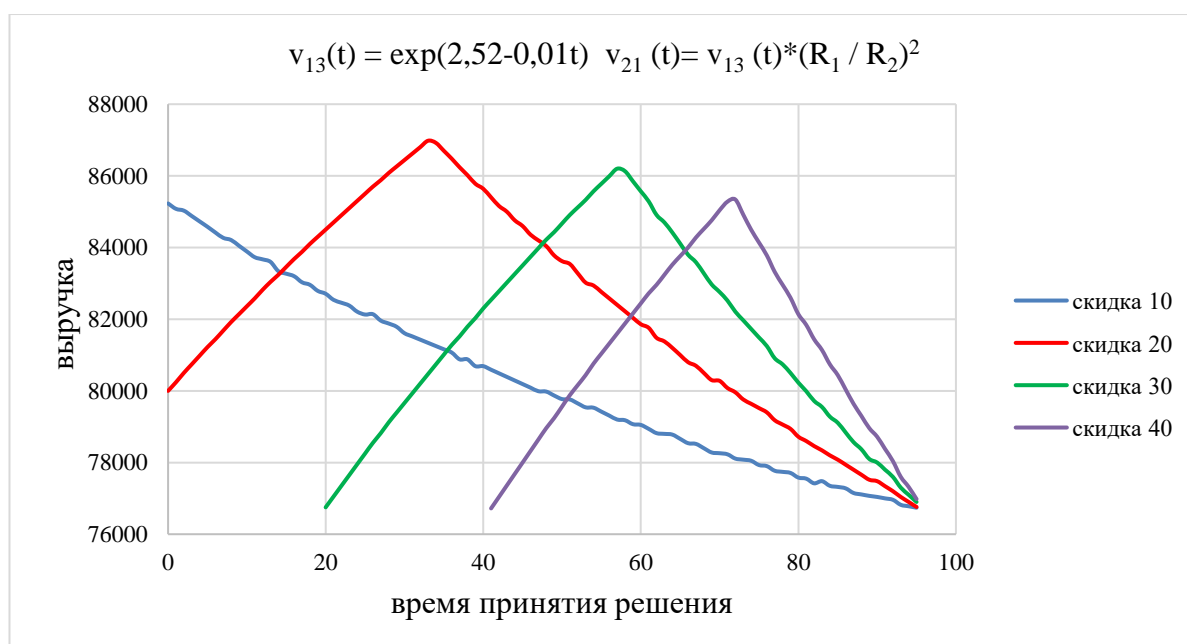


Рисунок 4.2.3 – Значения выручки для зависимостей v_{31} и v_{21} при принятии решения о снижении цены

В этом случае наилучший момент времени для принятия управляющего воздействия – 33 врем. ед. с начала продажи товара, а величина скидки – 20 д. ед. Выручка в таком случае увеличится на 10280 д. ед. по сравнению с первоначальным планом продажи без изменения цены, причем будет реализована вся партия товара.

Результаты экспериментов для функции v_{13} и соотношения v_{22} представлены на рисунке 4.2.4.

Для функции v_{31} и соотношения v_{22} наилучший момент времени для снижения стоимости товара – 83 врем. ед. с начала продажи товара, а величина скидки – 60 д. ед. Выручка в таком случае составит 81930 д. ед., что на 5230 д. ед. больше по сравнению с первоначальным планом продажи. Причем, если при первоначальном темпе реализации могло быть продано всего 767 единиц продукции, то при снижении цены в наилучший момент времени – 997, то есть почти вся партия целиком.

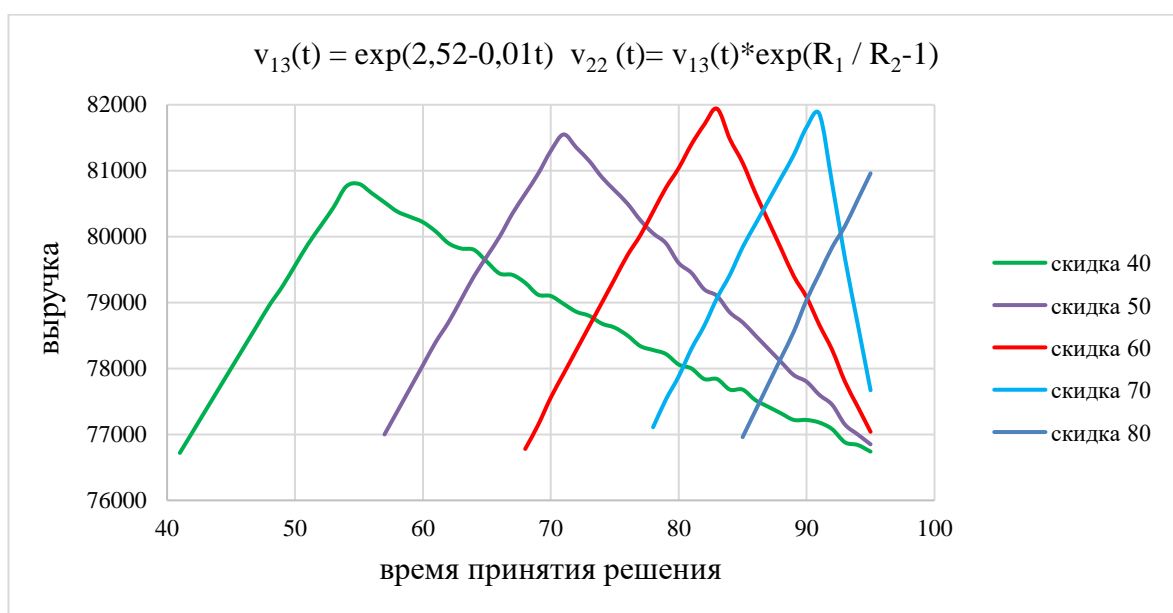


Рисунок 4.2.4 – Значения выручки для зависимостей v_{31} и v_{22} при принятии решения о снижении цены

В ходе экспериментов было получено, что для соотношения $v_2 = v_1 (R_1 / R_2)^2$ независимо от вида функции v_1 самым выгодным решением будет снижение цены товара на 20 д. ед. Графики зависимости значения выручки от времени принятия решения о снижении цены на 20 д. ед. представлены на рисунке 4.2.5.

Графики на этом рисунке показывают, что для всех видов функций темпа продажи время принятия решения об удешевлении товара для получения максимально возможной выручки лежит в окрестности 40 врем. ед.

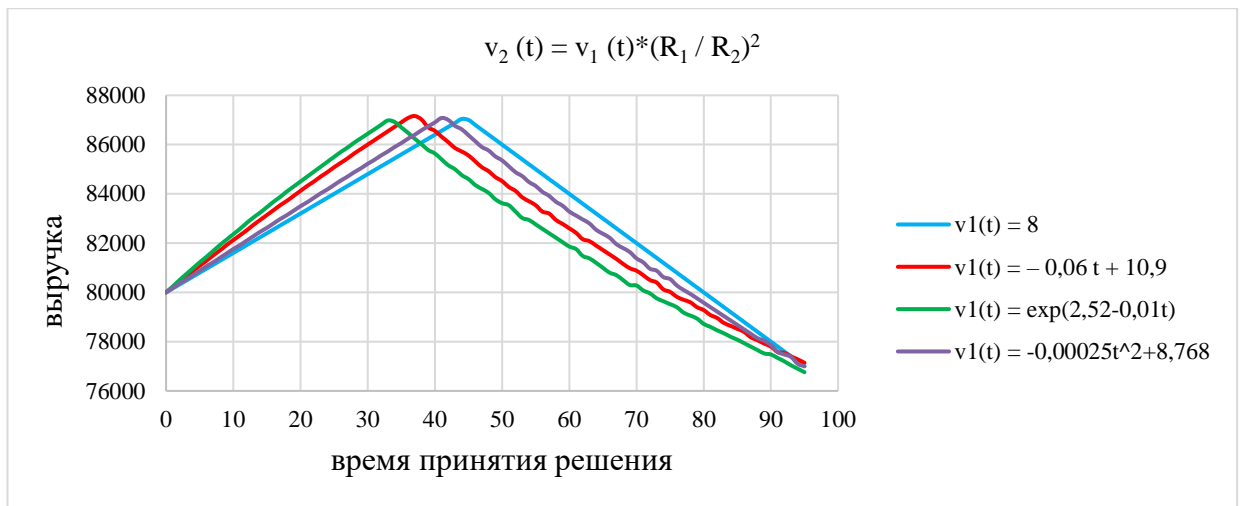


Рисунок 4.2.5 – Значения выручки для соотношения v_{21} при принятии решения о снижении цены на 20 д. ед.

Для другого соотношения $v_2 = v_1(t) \cdot \exp(R_1 / R_2 - 1)$ также нашлось одно для всех видов функции v_1 самое выгодное снижение цены товара – на 60 д. ед. Графики зависимости значения выручки от времени принятия решения о снижении цены на 60 д. ед. представлены на рисунке 4.2.6.

Анализ кривых на этом рисунке показывает, что для всех видов функций темпа продажи моменты времени принятия решения об удешевлении товара лежат в диапазоне [83, 88] врем. ед. При этом, середину этого интервала – 85 врем. ед. можно считать единым решением для всех видов функций темпа продажи, так как при снижении цены на 60 д. ед. в этот момент времени получаем величину выручки, близкую к максимальной.

При проведении опытов с соотношениями v_{23} и v_{24} оказалось, что для всех видов функций темпа продажи принятие решения о снижении цены в любой момент времени на любую величину дает суммарную выручку ниже той, которую получим без применения управляющего воздействия.

Отличием функций v_{23} и v_{24} от v_{21} и v_{22} является выпуклость последних, если рассматривать их как зависимости от соотношения цен (R_1 / R_2) . Поэтому в принципе можно утверждать, что выпуклая форма является необходимым

условием для существования точки принятия решения по интенсификации обслуживания для получения большего экономического эффекта.

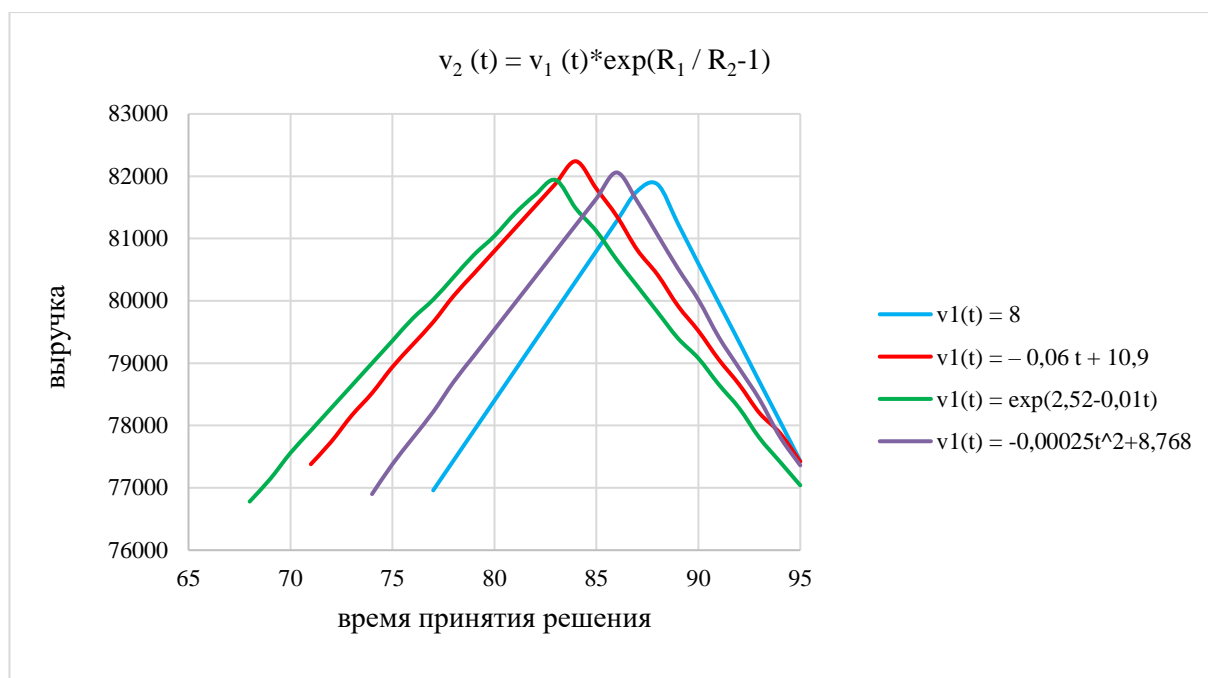


Рисунок 4.2.6 – Значения выручки для соотношения v_{22} при принятии решения о снижении цены на 60 д. ед.

Полученные результаты могут быть интерпретированы немного иначе.

При анализе графиков, представленных на рисунках 4.2.1 – 4.2.4, для функций v_{21} и v_{22} было выявлено, что различная величина скидки, осуществляемой в разные моменты времени, может давать сопоставимые значения суммарной выручки. Объединим такие моменты во множества и поставим в соответствие каждому наименьшее значение выручки. Назовем его уровнем выручки, который будет гарантированно достигнут при принятии решения об уценке в некотором временном диапазоне. Зададим пороговое значение для формирования множеств как 0,5% от первоначального значения суммарной выручки, которое получается без применения управляющего воздействия. При размере первоначальной выручки 76800 во множество самого низкого уровня будут объединены моменты времени, которые дают значение выручки в промежутке $[76800, 76800 \cdot 1,005)$, в следующий – $[76800 \cdot 1,005, 76800 \cdot 1,01)$ и т.д.

На рисунке 4.2.7 графически представлены множества самых высоких по значению уровней выручки для функции темпа продажи v_{11} и соотношения v_{21} .

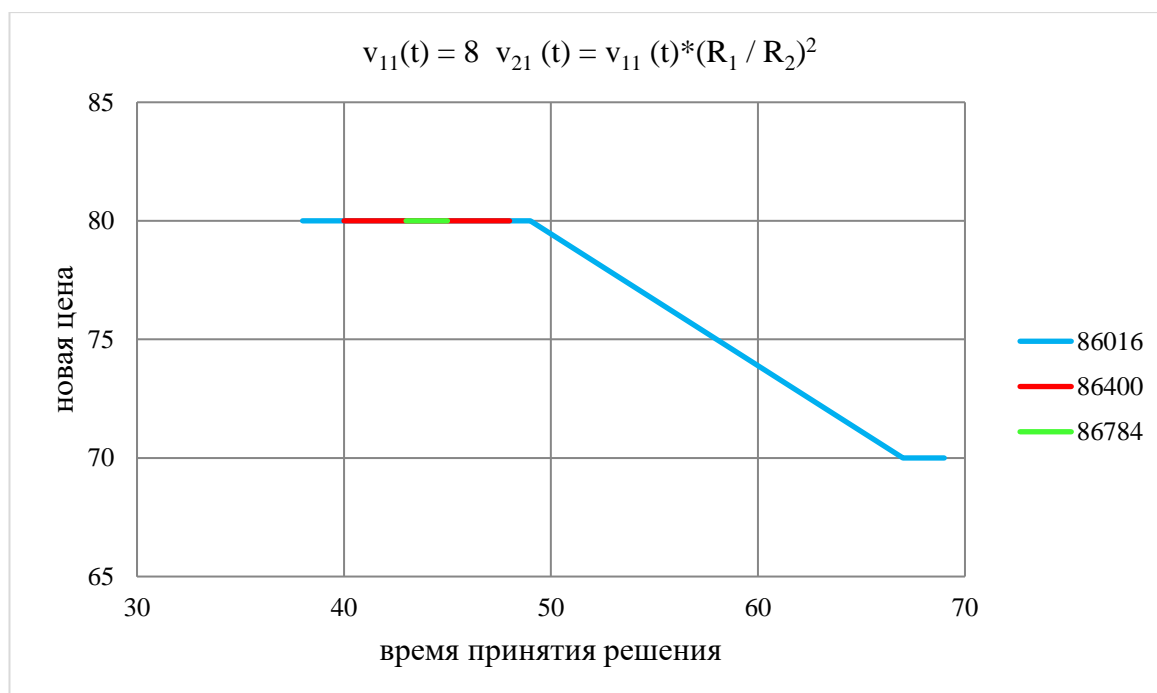


Рисунок 4.2.7 – Множества высоких уровней выручки для зависимостей v_{11} и v_{21}

Линии на этом рисунке наглядно демонстрируют наилучший временной диапазон принятия решения об уценке – [43 врем. ед., 46 врем. ед.] и новую стоимость товаров – 80 д. ед. Если наилучший момент был упущен, и решения принимается по прошествии 48 временных единиц с начала процесса продажи, то максимально достижимым становится уровень со значением выручки 86016, а значение новой цены зависит от момента времени управляющего воздействия.

Для функции темпа продажи v_{13} и соотношения v_{21} . множества самых высоких по значению уровней выручки изображены на рисунке 4.2.8.

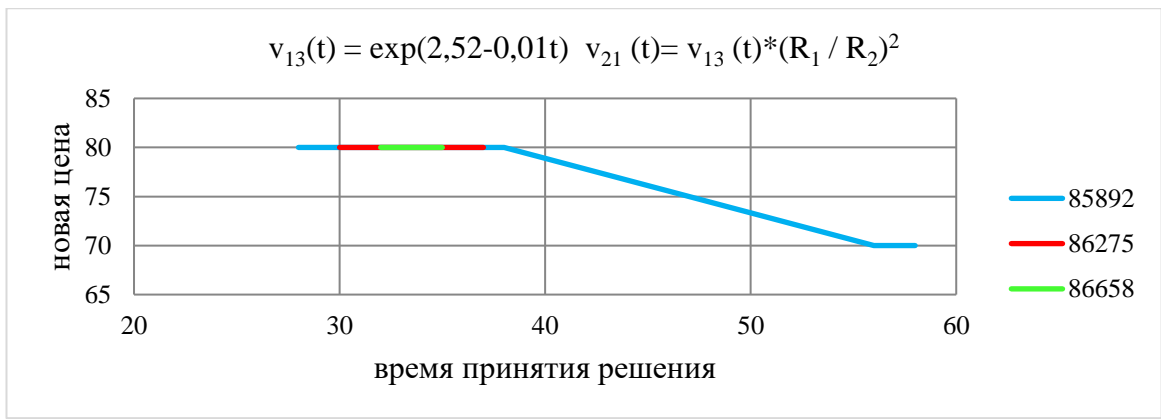


Рисунок 4.2.8 – Множества высоких уровней выручки для зависимостей v_{13} и v_{21}

Графики на этом рисунке демонстрируют ту же тенденцию, что и на рис. 4.2.7, значения уровней отличаются незначительно, разнятся только диапазоны принятия решений. Это может говорить о том, что вид функции темпа продажи не влияет на величину уровня. Данный результат прямо коррелируется с полученным выше: основное влияние на результат оказывает не вид функции темпа продажи, а вид зависимости между скоростью продажи и ценой товара. Это подтверждают и эксперименты, проведенные с другим соотношением между темпом продажи и стоимостью товара.

На рисунке 4.2.9 графически представлены множества самых высоких по значению уровней выручки для функции темпа продажи v_{11} и соотношения v_{22} .

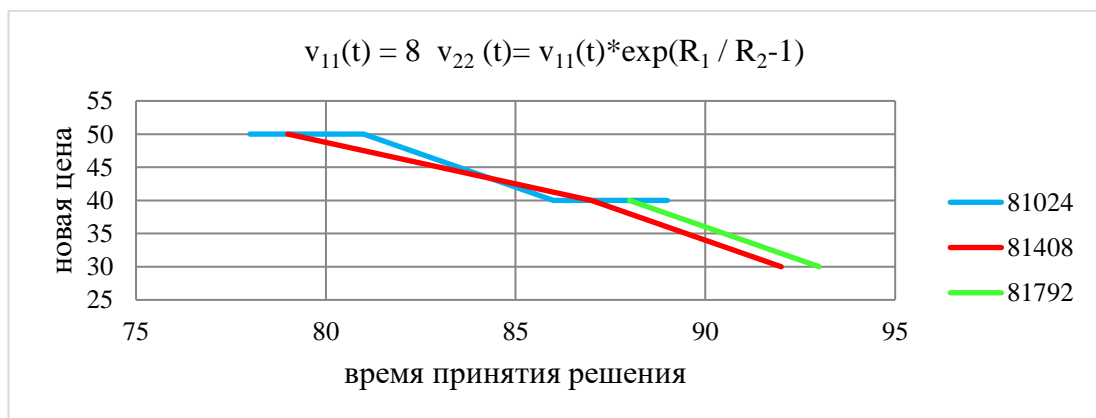


Рисунок 4.2.9 – Множества высоких уровней выручки для зависимостей v_{11} и v_{22}

На рисунке 4.2.10 графически представлены множества самых высоких по значению уровней выручки для функции темпа продажи v_{13} и соотношения v_{22} .

Таким образом, для известных функций интенсивности обслуживания и заданных зависимостей между темпом обслуживания и соотношением цен, можно найти момент приложения управляющего воздействия и его величину для получения максимально возможного экономического эффекта.

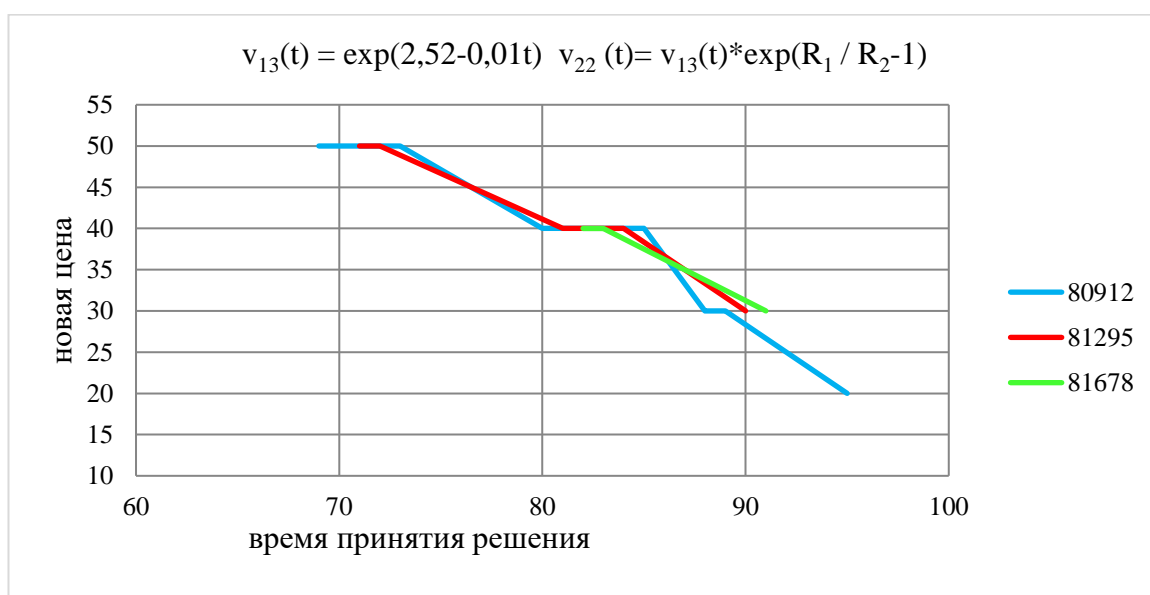


Рисунок 4.2.10 – Множества высоких уровней выручки для зависимостей v_{13} и v_{22}

При необходимости, например, когда этот момент упущен, предлагается расширение временного диапазона управления с помощью объединения моментов времени во множества с определенным значением выручки. Принадлежность момента времени конкретному множеству гарантирует значение выручки не ниже определенного уровня.

Изложенные результаты представлены в работах [143, 145].

4.3 Постановка и решение задачи повышения показателя экономической эффективности систем для случайного потока обслуживания

Рассмотрим ту же задачу повышения экономической эффективности одноканальной СМО с групповым поступлением требований, имеющих ограниченное время жизни, что и в предыдущем разделе, но со случайным потоком обслуживания. Пусть время обработки заявки подчинено нормальному закону распределения $N(\tau, \sigma^2)$. Как и раньше, будем полагать, что существует соотношение между стоимостью и темпом обработки требований, то есть при применении управляющего воздействия среднее время обслуживания заявки по измененной цене R_2 есть некоторая функция $\tau_2 = f(\tau_1, R_1, R_2)$, где τ_1 – среднее время обслуживания по цене R_1 . В качестве этих соотношений будем рассматривать те зависимости, которые дали положительный результат для детерминированного потока обслуживания:

$$\tau_2 = \tau_1 / \left(\frac{R_1}{R_2}\right)^2, \quad \tau_2 = \tau_1 / e^{\left(\frac{R_1}{R_2} - 1\right)}.$$

Первая серия имитационных экспериментов была проведена для следующие входных параметров:

- $Q = 1000$ шт. – объем партии товара;
- $t_{\text{end}} = 96$ врем. ед. – срок годности партии товара;
- $R_1 = 100$ д. ед. – установленная первоначально цена товара;
- $\tau_1 = 7,5$ минут;
- $\sigma = 1$ (2, 3);
- $\tau_2 = \tau_1 / \left(\frac{R_1}{R_2}\right)^2$.

Поиск момента принятия решения об изменении цены осуществляется в часах.

В таблице 4.3.1 представлены лучшие результаты одного опыта для различных значений среднеквадратичного отклонения в порядке убывания суммарной выручки.

Таблица 4.3.1 – Лучшие значения общей выручки одного эксперимента для

$$\tau_2 = \tau_1 / (R_1 / R_2)^2$$

σ = 1			σ = 2			σ = 3		
t*	R ₂	выручка	t*	R ₂	выручка	t*	R ₂	выручка
0	100	76300	0	100	76600	0	100	75300
44	80	87020	46	80	87280	47	80	87460
43	80	86860	45	80	87200	46	80	87380
45	80	86760	47	80	87200	45	80	87220
42	80	86700	44	80	87060	49	80	87080
41	80	86540	48	80	87060	44	80	87050
46	80	86520	43	80	86900	48	80	87040
47	80	86460	49	80	86760	43	80	86880
40	80	86380	42	80	86740	51	80	86820
39	80	86220	41	80	86580	52	80	86820
48	80	86220	40	80	86420	50	80	86750

В результате ряда опытов были получен средний процент прироста выручки для описанных входных параметров. В таблице 4.3.2 представлены лучшие средние результаты по 1000 опытам для различных значений среднеквадратичного отклонения в порядке убывания процента прироста общей выручки.

Таблица 4.3.2 – Лучшие средние значения прироста выручки по 1000 экспери-

$$\tau_2 = \tau_1 / (R_1 / R_2)^2$$

σ = 1				σ = 2				σ = 3			
t*	R ₂	частота	сред- ний %	t*	R ₂	частота	сред- ний %	t*	R ₂	частота	сред- ний %
42	80	0,017	13,60	42	80	0,076	14,04	42	80	0,054	14,14
43	80	0,091	13,94	43	80	0,137	13,85	43	80	0,102	14,02
44	80	0,322	13,89	44	80	0,192	13,87	44	80	0,123	13,87
45	80	0,393	13,86	45	80	0,204	13,84	45	80	0,124	14,08
46	80	0,160	13,84	46	80	0,158	13,88	46	80	0,109	14,10
47	80	0,016	13,95	47	80	0,071	14,10	47	80	0,087	14,10
				48	80	0,067	13,90	48	80	0,09	13,79
								49	80	0,056	13,85

На основании полученных результатов можно заключить, что однозначно определена оптимальная величина снижения цены обработки заявки, она составляет 20 д. ед. Средний процент прироста выручки для всех решений варьируется в малом диапазоне значений – (13,6; 14,14), что позволяет определить порядок величины процента прироста выручки с высокой точностью.

Представим частоты появления точек принятия решения в 1000 экспериментах для разных значений среднеквадратичного отклонения в графическом виде (рис. 4.3.1).

Анализ результатов, изображенных на этом рисунке, позволяет утверждать, что для каждого среднеквадратичного отклонения определен диапазон значений времени для осуществления управляющего воздействия, причем, область принятия оптимального решения расширяется по мере увеличения дисперсии. Кроме того, важно заметить, что результаты опытов для $\sigma = 1$ хорошо

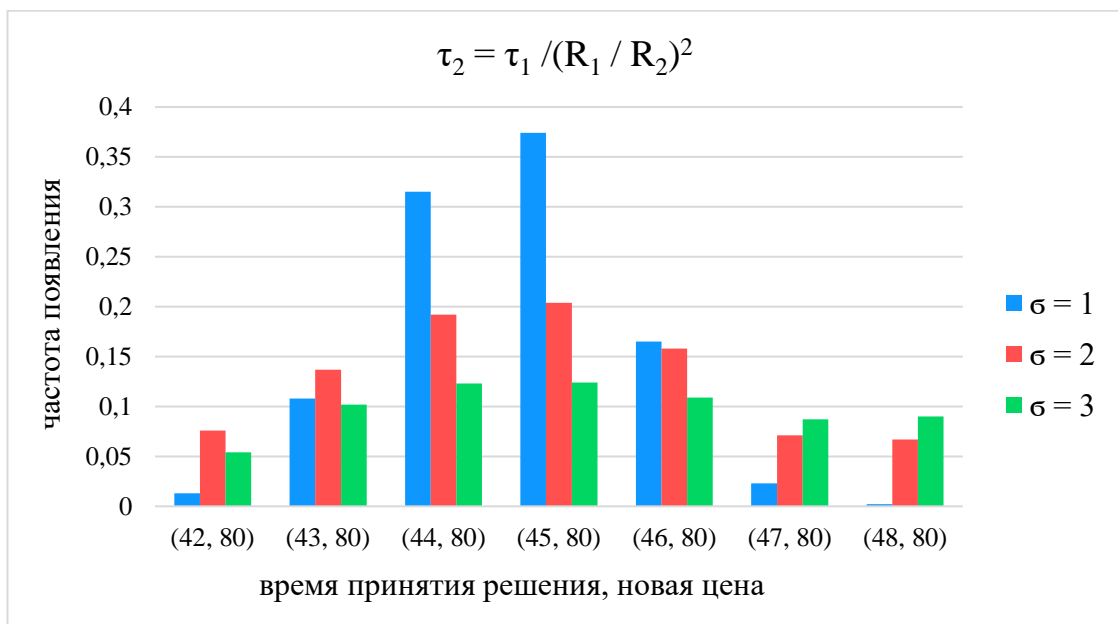


Рисунок 4.3.1 – Частоты появления лучших точек принятия решения для

$$\tau_2 = \tau_1 (R_1 / R_2)^2$$

согласуются с итогами экспериментов, описанных в разделе 4.2.2, проведенных для постоянного темпа обслуживания. Причем, обнаруживается совпадение и в найденном интервале значений оптимального времени принятия решения, и в величине прироста выручки.

Следующие эксперименты были осуществлены для описанных выше входных параметров и соотношения $\tau_2 = \tau_1 / e^{\left(\frac{R_1}{R_2}-1\right)}$. В таблице 4.3.3 представлены лучшие средние результаты по 1000 опытам для различных значений среднеквадратичного отклонения в порядке убывания процента прироста общей выручки.

Таблица 4.3.3 – Лучшие средние значения прироста выручки по 1000 экспериментам для $\tau_2 = \tau_1 / \exp(R_1 / R_2 - 1)$

$\sigma = 1$				$\sigma = 2$				$\sigma = 3$			
t^*	R_2	частота	сред- ний%	t^*	R_2	частота	сред- ний%	t^*	R_2	частота	сред- ний%
88	40	0,854	7,15	88	40	0,680	7,16	88	40	0,575	7,14
87	40	0,139	7,16	87	40	0,246	7,41	87	40	0,217	7,55
79	50	0,005	7,10	80	50	0,039	7,23	80	50	0,069	7,52
80	50	0,002	6,90	79	50	0,026	7,51	79	50	0,038	7,72
								89	40	0,029	6,51

Представим частоты появления точек принятия решения в 1000 экспериментах для разных значений среднеквадратичного отклонения в графическом виде (рис. 4.3.2).

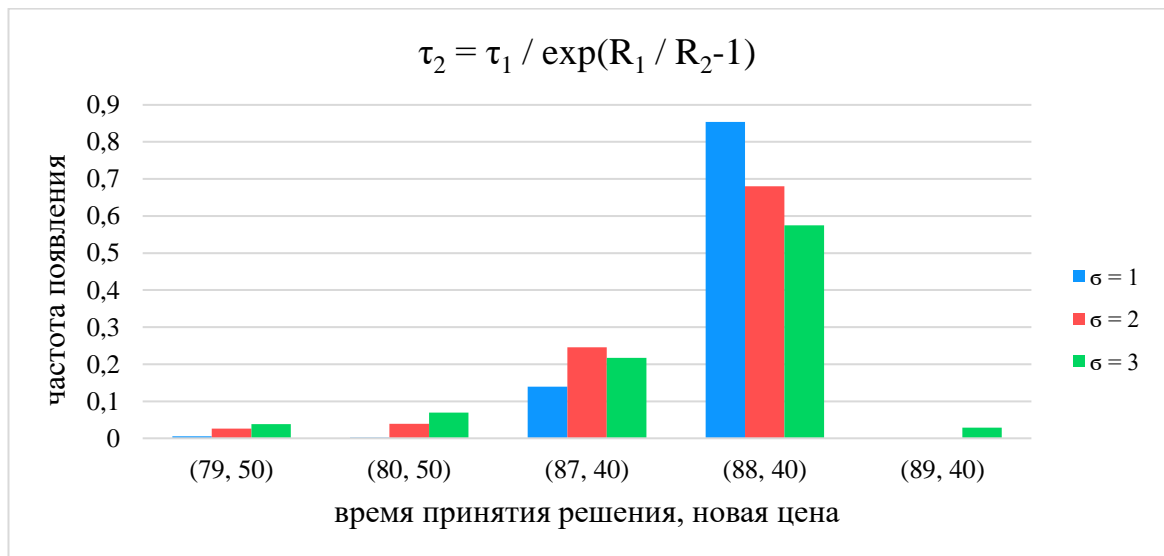


Рисунок 4.3.2 – Частоты появления лучших точек принятия решения для $\tau_2 = \tau_1 / \exp(R_1 / R_2 - 1)$

Учитывая существенное преобладание в результатах пары значений оптимальной величины снижения цены обработки заявки, которая равна 60 д. ед., и момента времени осуществления управляющего воздействия – 88 врем. ед., можно утверждать, что оптимальное решение найдено. Здесь также обнаруживается согласие результатов с итогами экспериментов, осуществленных для постоянного темпа обслуживания, и в найденном значении оптимального времени принятия решения, и в величине прироста выручки.

Вторая серия имитационных экспериментов была проведена для следующие входных параметров:

- $Q = 1000$ шт. – объем партии товара;
- $t_{\text{end}} = 365$ врем. ед. – срок годности партии товара;
- $\tau_1 = 5/12$ врем. ед.;
- $R_1 = 100000$ д. ед. – установленная первоначально цена товара;
- $\sigma = 5/144$ ($5/72, 5/36$);
- $\tau_2 = \tau_1 / \left(\frac{R_1}{R_2}\right)^2, \tau_2 = \tau_1 / e^{\left(\frac{R_1}{R_2} - 1\right)}$.

Существенным отличием этой постановки от приведенной выше является порядок входных параметров. Они характеризуют продажу дорогостоящих товаров с длинным сроком годности. Проведя серию имитационных экспериментов для $\tau_2 = \tau_1 / (R_1 / R_2)^2$, удалось установить, что оптимальная скидка на товар 20000 д. ед., средний прирост выручки равен примерно 9%. Частоты появления значений моментов времени снижения цены представлены на рисунке 4.3.3.

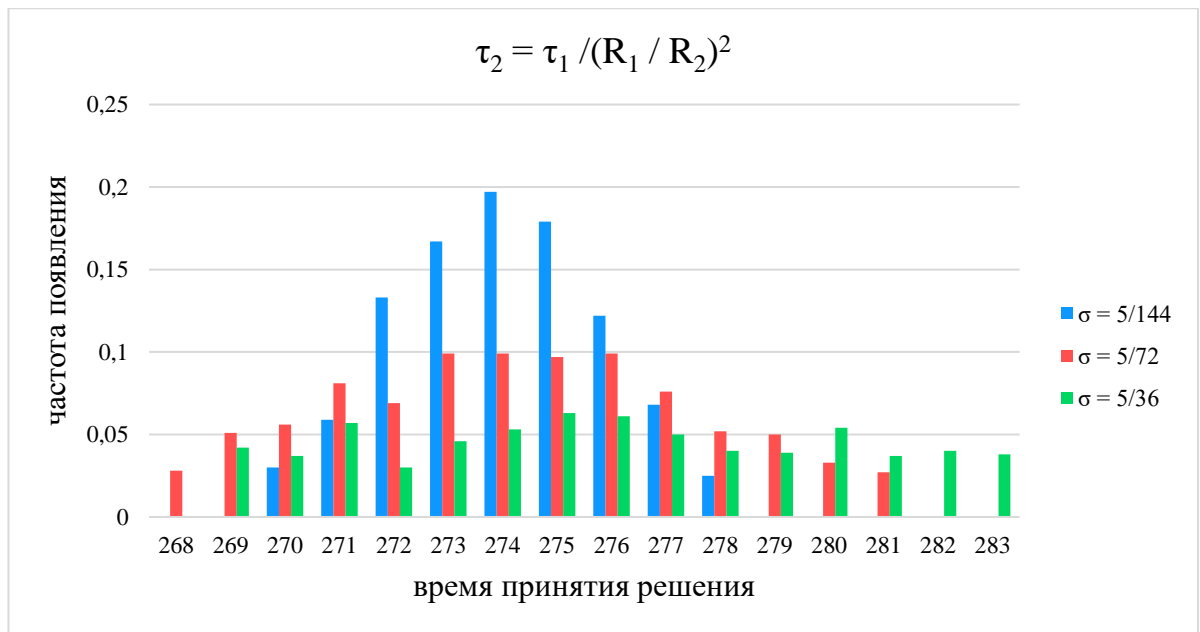


Рисунок 4.3.3 – Частоты появления лучших точек принятия решения для $\tau_2 = \tau_1 / (R_1 / R_2)^2$ в случае большого срока годности

Проведя серию имитационных экспериментов для $\tau_2 = \tau_1 / \exp(R_1 / R_2 - 1)$, удалось установить, что оптимальная скидка на товар 60000 д. ед., средний прирост выручки составляет приблизительно 5,5%. Частоты появления значе- ний моментов времени снижения цены представлены на рисунке 4.3.4.

Здесь, как и в предыдущих опытах, решение зависит от выбора вида за- висимости между темпом обслуживания и соотношением цен. Экспоненци- альная зависимость определяет более узкий диапазон принятия решения о сни- жении стоимости, чем квадратичная. Но и в том, и другом случае величина снижения цены определяется однозначно.

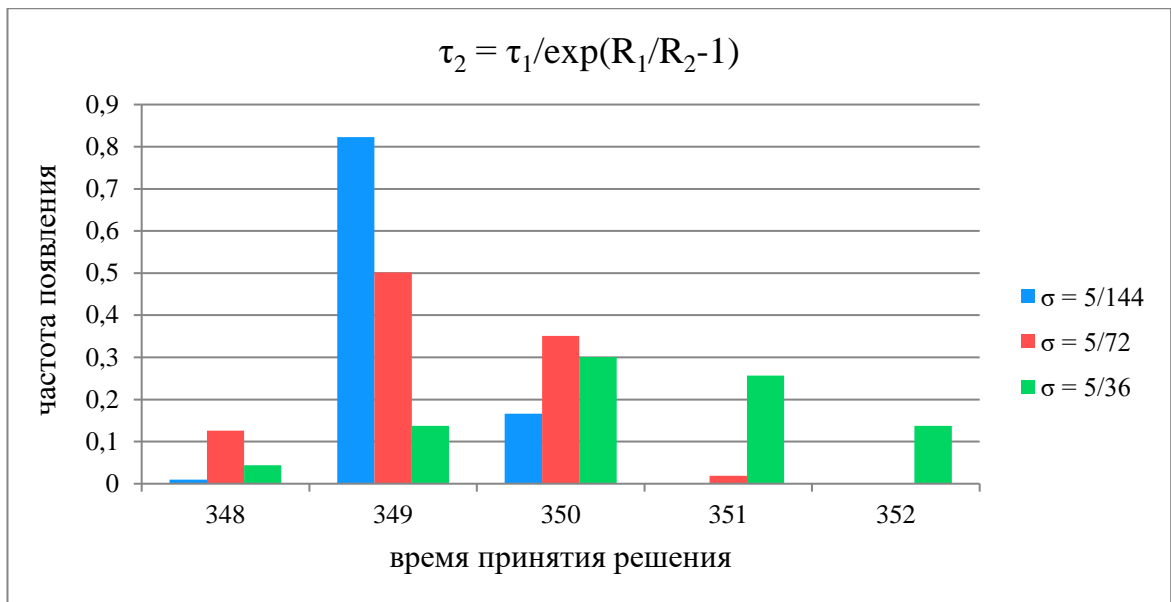


Рисунок 4.3.4 – Частоты появления лучших точек принятия решения для $\tau_2 = \tau_1/\exp(R_1/R_2-1)$ в случае большого срока годности

Таким образом, для нормального закона распределения времени обслуживания и заданных зависимостей между темпом обслуживания и соотношением цен, можно определить диапазон времени для осуществления управляющего воздействия и его величину для получения максимально возможного экономического эффекта. Положение этого диапазона на оси модельного времени и величина управляющего воздействия обусловлены видом зависимости между темпом обслуживания и соотношением цен, и набором входных параметров. Поэтому, каждая задача требует индивидуального подхода с обязательным учетом специфики предметной области.

Изложенные результаты представлены в работе [146].

Трансформируем задачу следующим образом.

Допустим, что главной целью процесса обслуживания является не максимизация выручки, а полная обработка всей группы заявок до истечения их срока жизни. Способом интенсификации обслуживания в данном контексте может стать введение дополнительных устройств, которое сопряжено с увеличением стоимости обработки заявок. Тогда целесообразно искать время и

величину управляющего воздействия, такие, которые обеспечат достижение цели с минимальными дополнительными вложениями.

Положим:

$Q = 1000$ шт. – объем группы заявок;

$t_{\text{end}} = 64$ часа – время жизни заявок с момента поступления в систему;

[6 мин., 8 мин.] – интервал равномерно распределенного времени обслуживания заявки τ_1 ;

$R_1 = 100$ д. ед. – установленная первоначально стоимость обработки каждой заявки;

$\tau_2 = \tau_1 \left(\frac{R_1}{R_2}\right)^2$, $\tau_2 = \tau_1 e^{\left(\frac{R_1}{R_2}-1\right)}$ – зависимости между интенсивностью обслуживания и соотношением цен.

Положим, что решение об интенсификации процесса обслуживания принимается в минутах. В таблице 4.3.4 представлены лучшие результаты одного опыта в порядке увеличения процента дополнительных затрат.

Проведение ряда экспериментов принесло понимание того, что минутная шкала для времени принятия решения не является удачной, поскольку сколько-нибудь устойчивого решения получить не удастся. Поэтому была выбрана другая единица измерения шкалы времени – час.

Таблица 4.3.4 – Лучшие значения дополнительных затрат одного эксперимента

$\tau_2 = \tau_1 \left(\frac{R_1}{R_2}\right)^2$			$\tau_2 = \tau_1 e^{\left(\frac{R_1}{R_2}-1\right)}$		
t*(мин)	R ₁	Итоговая цена в%	t*(мин)	R ₁	Итоговая цена в%
602	140	149,5795	150	260	371,4706
601	140	149,5795	149	260	371,4706
600	140	149,5795	146	260	371,7647
599	140	149,5795	145	260	371,7647
598	140	149,6527	144	260	371,7647
597	140	149,6527	143	260	371,7647
596	140	149,6527	142	260	371,7647
595	140	149,6527	137	260	372,0588
594	140	149,6527	135	260	372,0588

В таблице 4.3.5 представлены частоты появления значений времени принятия решения об интенсификации процесса обслуживания и новой стоимости обслуживания с минимальными дополнительными затратами в 1000 опытах.

Таблица 4.3.5– Частоты появления лучших результатов в 1000 экспериментах

$\tau_2 = \tau_1 \left(\frac{R_1}{R_2}\right)^2$			$\tau_2 = \tau_1 e^{\left(\frac{R_1}{R_2}-1\right)}$		
t*(час)	R₁	частота	t*(час)	R₁	частота
9	140	850	2	260	620
8	140	150	0	250	170
			1	260	130
			1	250	50
			3	260	30

На основании полученных результатов можно утверждать, что в случае квадратичной зависимости между соотношением цен обработки заявок и ее временем однозначно определены оптимальное значение новой стоимости обработки заявки – 140 д. ед. и время принятия решения – интервал между 8 и 9 часами с начала обслуживания. В случае экспоненциальной зависимости имеем более размытый интервал принятия решения и неоднозначные, но близкие по значению, новые стоимости обслуживания заявок.

Изложенные результаты представлены в работе [147].

Завершающим этапом работы стало создание инструмента для управления СМО группового входного потока заявок с ограниченным временем жизни с целью повышения меры ее экономической эффективности. Поскольку в общем случае закон распределения времени обслуживания заявок неизвестен, с учетом ограниченного времени жизни требований было решено следовать общим для таких случаев рекомендациям и установить треугольный закон распределения. Следующим ключевым моментом явилось определение зависимости между временем обработки заявок и соотношением цен. Так как нахождение функциональной зависимости выходит за рамки компетенции администратора, требовалось трансформировать ее в приемлемый для ЛПР

вариант. Им стала непрерывная кусочно-линейная функция, которая определяет зависимость времени обслуживания заявок от соотношения цен (R_0/R_i), $i = \overline{1, n}$. Здесь R_0 – установленная первоначально цена товара, R_i – цена, соответствующая i -й скорости обслуживания, $(n+1)$ – количество пар значений, которые задает администратор. Это решение хорошо согласуется с профессиональными компетенциями лица, который определяет необходимость интенсификации процесса обслуживания.

Установим следующие входные параметры:

$Q = 1000$ шт. – объем партии товара;

$t_{\text{end}} = 96$ часов – срок годности партии товара;

$R_1 = 100$ д. ед.– установленная первоначально цена товара;

$\tau_1 = 7,5$ минут – среднее время обслуживания по цене R_1 ;

$\sigma = 1$ минута – разброс среднего времени обслуживания;

τ_2 – непрерывная кусочно-линейная функция, построенная по точкам: (100 д. ед., 8 шт./час), (90 д. ед., 10 шт./час), (70 д. ед., 20 шт./час), (50 д. ед., 30 шт./час).

На рисунке 4.3.5 изображены кривые зависимостей времени обслуживания заявок от соотношения цен: линейной кусочно-непрерывной функции, построенной по 4-м точкам и квадратичной функции, которая использовалась ранее – $\tau_2 = \tau_1 / (R_1 / R_2)^2$.

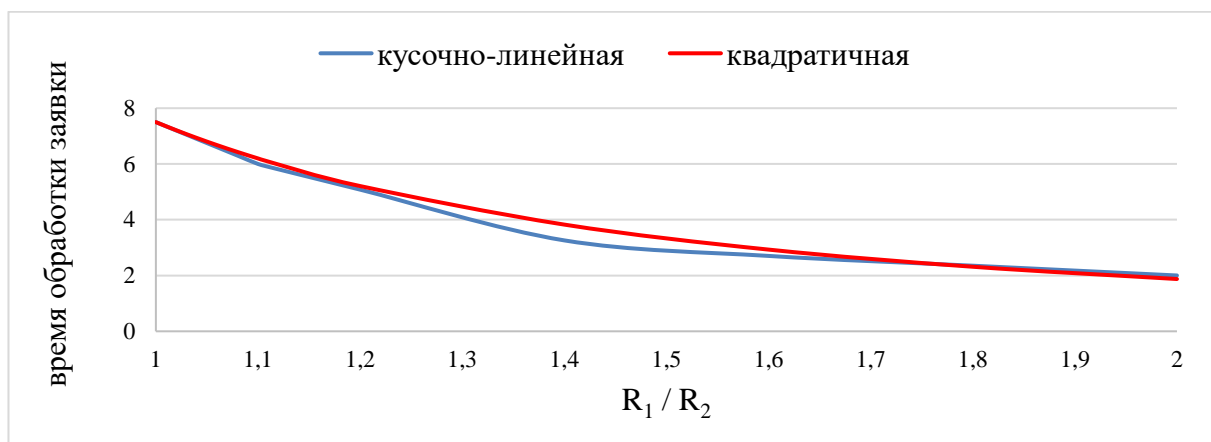


Рисунок 4.3.5 – Кривые зависимостей времени обслуживания заявок от соотношения цен

Для означенных входных параметров была проведена серия экспериментов. В результате локализованы две точки принятия решения об интенсификации продаж. Первый вариант – снизить цену товара до 70 д. ед. в момент времени 76 часов с момента начала продажи. Такое решение позволит увеличить выручку примерно на 15%. Второй вариант – снизить цену товара до 80 д. ед. в момент времени 47 часов с момента начала продажи. Это решение позволит увеличить выручку примерно на 14%. Следует отметить, что второе решение близко к результатам исследований, проведенных для квадратичной зависимости времени обслуживания от соотношения цен.

Построим функцию τ_2 по другому набору точек: (100 д. ед., 8 шт./час), (70 д. ед., 20 шт./час), (50 д. ед., 30 шт./час), (30 д. ед., 60 шт./час) (рис. 4.3.6).

Серия имитационных экспериментов, проведенных для новой функции τ_2 , позволила локализовать интервал времени принятия решения – [75 час., 77 час.] и найти оптимальную величину снижения цены – 30 д. ед. Такое решение позволит увеличить выручку более, чем на 15%.

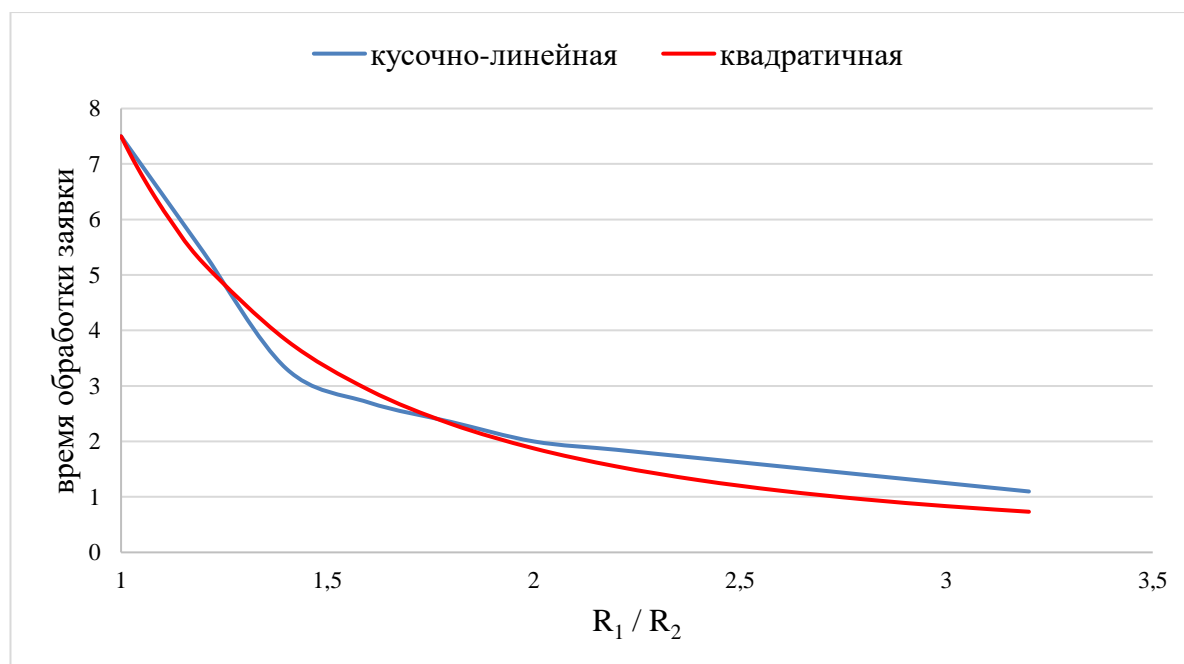


Рисунок 4.3.6 – Кривые зависимостей времени обслуживания заявок от соотношения цен (2)

Таким образом, был сформирован и реализован простой и понятный механизм для ситуационного управления групповым входным потоком заявок с

фиксированным временем жизни с целью повышения меры экономической эффективности системы обслуживания.

Программный продукт зарегистрирован, о чем получено «Свидетельство о регистрации электронного ресурса с оценкой новизны» (см. Приложение 2).

ЗАКЛЮЧЕНИЕ

В диссертационном исследовании представлены результаты разработки математических моделей, алгоритмов и программных комплексов для анализа характеристик систем массового обслуживания с ограниченным временем жизни заявок и возможностей повышения эффективности их функционирования.

В процессе исследования были получены следующие основные результаты.

1. Построена математическая модель системы массового обслуживания пуассоновского потока заявок с ограниченным временем жизни без опции накопления очереди, которая в отличие от большого количества аналитических моделей обладает тем неоспоримым преимуществом, что имеет высокую практическую применимость. Построена математическая модель одноканальной системы массового обслуживания группового потока заявок с ограниченным временем жизни с опцией накопления очереди. Ее особенностью является то, что для оценки эффективности функционирования системы выбран экономический критерий, и с помощью введения управляющего параметра найден способ повышения экономического эффекта.

2. Исследована модель одноканальной системы с ожиданием, концепция которой приведена в работах Б.В. Гнеденко и И.Н. Коваленко. На ее основе получено аналитическое решение задачи нахождения основных характеристик СМО для простейшего входного потока и равномерно распределенного времени жизни заявок.

3. Построены имитационные алгоритмы функционирования СМО, учитывающие специфику систем, изучаемых в настоящей работе: ограниченное время жизни заявок и групповое поступление требований. Они созданы для разных систем массового обслуживания: с ожиданием и без него, одноканальных и многоканальных, с ординарным и групповым поступлением требований. Осуществлена их программная реализация на языке С#. Использование библиотеки параллельных вычислений позволило значительно ускорить

работу программных модулей. Проведена верификация имитационных моделей путем сравнения результатов их работы с аналитическими вычислениями.

4. Проанализированы характеристики систем массового обслуживания с ограниченным временем жизни заявок. Обнаружено, для многоканальных систем, обслуживающих групповые потоки заявок с ограниченным временем жизни, при увеличении числа каналов обслуживания относительная пропускная способность стремится к доли полностью обслуженных заявок ординарного потока. Кроме того, существует предельное значение относительной пропускной способности системы, обслуживающей поток заявок с фиксированным временем жизни, которое зависит от интенсивности обслуживающих устройств и времени жизни заявок и не зависит ни от количества каналов обслуживания, ни от размера группы, ни от наличия очереди, ни от интенсивности входного потока заявок. Найдено пороговое значение времени жизни заявок – значение, при котором время жизни заявок теряет свое влияние на поведение системы и делает бесполезным управление ею.

5. Имитационным путем получено решение задачи повышения экономической эффективности функционирования одноканальной системы массового обслуживания группового потока заявок с ограниченным временем жизни. Для детерминированных и случайных потоков обслуживания найдены момент времени и величина управляющего воздействия на систему для максимально возможного повышения значения критерия эффективности. Опытным путем доказано, что величина управляющего воздействия зависит от вида зависимости между темпом обслуживания и соотношением управляющих параметров.

Полученные результаты могут быть применены для любых процессов, моделями которых являются системы массового обслуживания заявок с ограниченным временем жизни, как обоснованные критерии на этапе предпроектного анализа. На этапе эксплуатации их можно использовать для повышения эффективности функционирования систем.

По причине актуальности федеральной программы по роботизации производств, основными компонентами которой являются внедрение программных и аппаратных средств управления, результаты исследования становятся практически значимыми, поскольку СМО с рассматриваемой спецификой могут служить цифровыми двойниками потоков обслуживания, протекающих в роботизированных системах.

Рекомендуется на этапах проектирования и предпроектной проработки систем, моделируемых средствами теории массового обслуживания заявок с ограниченным временем жизни, принимать во внимание выводы, изложенные в данной работе.

Перспективы дальнейшей разработки темы должны охватывать построение аналитических выражений для оценки эксплуатационных характеристик, формирование правил принятия решений для повышения эффективности, исследование влияния законов распределения случайных величин и дисциплин обслуживания на показатели эффективности функционирования систем массового обслуживания с ограниченным временем жизни заявок.

СПИСОК ЛИТЕРАТУРЫ

1. Erlang, A. K. Probability and telephone calls / A.K. Erlang // *Nyt Tidsskr. Mat. Ser. B.* – 1909. – Vol. 20. – P. 33-39.
2. Вентцель, Е.С. Теория вероятностей / Е.С. Вентцель – М.: Наука, 1969. – 576 с.
3. Марков, А.А. Исследование замечательного случая зависимых испытаний // *Известия Императорской академии наук. Серия 6.* – Санкт-Петербург, 1907. – Том 1, вып. 3. – С. 61-80. – Отд. оттиск.
4. Колмогоров, А.Н. Über die Grenzwertsätze der Wahrscheinlichkeitsrechnung / von A. Kolmogorov (A. Kolmogoroff) // *Известия Академии наук СССР. Серия 7. Отделение математических и естественных наук.* – М.; Л., 1933. – Вып. 3. – С. 363–372. – Отд. оттиск.
5. Кендалл, Д. Стохастические процессы, встречающиеся в теории очередей, и их анализ методом вложенных цепей Маркова / Д. Кендал // *Математика.* – 1959. – Т. 3. – № 6. – С. 97-111.
6. Хинчин, А. Я. Математическая теория стационарной очереди / А.Я. Хинчин // *Матем. сб.* – 1932. – Вып. 39. – С. 73-84.
7. Pollaczek, F. Über eine Aufgabe der Wahrscheinlichkeitstheorie / F. Pollaczek // *Math. Zeit.* – 1930(32). – С. 64-110; С. 729-750.
8. Хинчин, А. Я. Работы по математической теории массового обслуживания / А. Я. Хинчин. – М.: ФИЗМАТЛИТ, 1963. – 236 с.
9. Kendall D.G. Some Problems in the theory of queues / D.G. Kendall // *Journal of Royal Statistical Society.* – 1951. – Vol. 13. – № 2. – P. 151-173.
10. Хинчин, А.Я. Математические методы теории массового обслуживания / А.Я Хинчин // *Тр. МИАН СССР*, 49, М.: Изд-во АН СССР, 1955. – 122 с.
11. Гнеденко, Б.В. Введение в теорию массового обслуживания / Б.В. Гнеденко, И.Н. Коваленко. – М.: Наука, 1966. – 301 с.

12. Золотарев, В.М. Об одной гипотезе Б. В. Гнеденко / В.М. Золотарев, В.С. Королюк. // Теория вероятн. и ее примен. – 1961. – Т.6. – Вып. 4. – С. 469–474.
13. Скороход, А. В. Случайные процессы с независимыми приращениями / А.В. Скороход. — М.: Наука, 1964. – 280 с.
14. Б. Григелионис, Б. О сходимости сумм ступенчатых случайных процессов к пуассоновскому / Б. Григелионис. Теория вероятн. и ее примен., 1963. –Т.8. – Вып.2.–С. 189–194.
15. Боровков, А.А. Вероятностные процессы в теории массового обслуживания / А.А. Боровков. – М.: Наука, 1972. – 368 с.
16. Такач, Л. Комбинированные методы в теории случайных процессов / Л. Такач. – М.: Мир, 1971. – 266 с.
17. Саати, Т. Л. Элементы теории массового обслуживания и ее приложения / Т. Л. Саати. – М.: Издательство Советское радио, 1971. – 520 с.
18. Ховард Р.А. Динамическое программирование и марковские процессы / Р.А. Ховард – М.: Советское радио, 1964. – 200 с.
19. Little, John D. C. A Proof for the Queuing Formula: $L = \lambda W$ / John D. C. Little// Operations Research. – 1961. – Vol. 9, No. 3 – P. 383-387.
20. Клейнрок, Л. Коммуникационные сети: Стохастические потоки и задержки сообщений / Пер. с англ. В. В. Дубницкого и М. И. Ланина. Под ред. А. А. Первозванского. – М.: Наука, 1970. — 255 с.: ил.
21. Кокс, Д. Теория очередей / Д. Кокс, В. Смит — М.: Мир, 1966. – 218 с.
22. Cohen, J. The Single Server Queue / J. Cohen. - Amsterdam: North Holland, 1969. – 657 pp.
23. Кофман, А. Массовое обслуживание. Теория и приложения / А. Кофман, Р. Крюон. – М.: Мир, 1965. – 302 с.
24. Риордан, Дж. Вероятностные системы обслуживания / Дж. Риордан. – М.: Связь, 1966. – 190 с.

25. Benes, V.E. General Stochastic Processes in the Theory of Queues / Vaclav E. Benes/ – Published by Addison-Wesley, 1963. – 88с.
26. Siski, R., Introduction to congestion theory in telephone systems / R. Siski, N. H. G. Morris. – Oliver and Boyd, Edinburgh and London, 1960. – 742 с.
27. Розенберг, В.Я. Что такое теория массового обслуживания / Розенберг В.Я., Прохоров А.И. – М.: Радио и связь, 1963.– 256 с.
28. Башарин, Г.П. Массовое обслуживание в телефонии / Башарин Г.П., Харкевич А.Д., Шнепс М.А. – М.: Наука, 1968. — 247 с.
29. Гнеденко, Б. В. Приоритетные системы обслуживания / Б. В. Гнеденко, Э. А. Даниэлян, Б.Н. Димитров и др. – М.: Изд-во МГУ, 1973. – 448 с.
30. Джейсуол, Н. Очереди с приоритетами / Н. Джейсуол – М.: Мир, 1973. – 279 с.
31. Бронштейн, О.И. Модели приоритетного обслуживания в информационно-вычислительных системах / О.И. Бронштейн, И.М. Духовный – М.: Наука, 1976. – 220 с.
32. Ушаков В. Г. Система обслуживания с эрланговским входящим потоком и относительным приоритетом / В.Г. Ушаков // Теория вероятности и ее применение. -1977. – Т.22. – В.4. – С. 860-866.
33. Жданов, В.С. Условия существования установившихся режимов в циклических системах массового обслуживания / В.С. Жданов, Е.А. Саксонов // Автоматика и телемеханика. –1979. – № 2. – С. 29-38.
34. Müntz, R. Wating time distribution for round-robin queueing systems / R. Müntz // Proceedings of Symposium on Computer Communications Networks, and Teletraffic. – New York. – 1972.–P. 429-439.
35. Королюк, В.С. Полумарковские процессы и их приложения / В.С. Королюк, А.Ф. Турбин – Киев: Наукова думка, 1976. – 184 с.
36. Коваленко, И.Н. Исследования по анализу надежности сложных систем / И.Н. Коваленко – Киев: Наукова думка, 1976. – 211 с.

37. Гнеденко, Б.В. Математические методы в теории надежности: Основные характеристики надёжности и их стат. анализ / Б. В. Гнеденко, Ю. К. Беляев, А. Д. Соловьёв. – Москва: Наука, 1965. – 524 с.
38. Ивницкий, В.А. О времени первого переполнения очереди при ненадежной обслуживающей системе / В.А. Ивницкий, В.А. Капырин // Изв. АН СССР. Техническая кибернетика. – 1975. – № 5.
39. Соловьёв, А.Д. Резервирование с быстрым восстановлением / А.Д. Соловьёв // Изв. АН СССР. Техн. кибернетика. – 1970. – № 1 – С. 56 - 71.
40. Липаев, В.В. Эффективность методов организации вычислительного процесса в АСУ / В.В. Липаев, С.Ф. Яшков. – М.: Статистика, 1975. – 255 с.
41. Яшков, С.Ф. О дисциплине преимущественного разделения процессора требованиями с наименьшей обслуженной длиной/ С.Ф. Яшков // Техника средств связи. Сер. АСУ. – 1978. – Вып. 2. – С. 51-61.
42. Липаев, В.В. Распределение ресурсов в вычислительных системах. / Липаев В.В. – М.: Статистика, 1979. – 247 с.
43. Голованов, О.В. Моделирование сложных дискретных систем на ЭВМ третьего поколения / О.В. Голованов, С.Г. Дуванов, В.Н. Смирнов. – М.: Энергия, 1978. – 160 с.
44. Артамонов, Г.Т Аналитические вероятностные модели функционирования ЭВМ / Г.Т Артамонов, О.М. Брехов. – М.: Энергия, 1978. – 368 с.
45. Kleinrock, L. Resource allocation in computer systems and computer communication networks / L. Kleinrock // Proceedings of IFIP Cong. Proc., 1974. – pp. 1118.
46. Базилевич, Е.В. Системы коммутации сообщений на базе техники ЭВМ / Е.В. Базилевич, Г.Ф. Прамнэк. – М.: Связь, 1971. – 180 с.
47. Дьяченко, В.Ф. Управление на сетях связи / В.Ф. Дьяченко, В.Г. Лазарев, Г.Г. Саввин. – М.: Связь, 1967. – 223 с.

48. Файнберг, М.А. Управление в системах массового обслуживания / М.А. Файнберг, Е.А. Файнберг // Зарубежная радиоэлектроника, 1975. – № 3. – С. 3 - 34.
49. Перельман, А.Е. Построение моделей автоматизированных систем оперативного управления производством / А.Е. Перельман. – М.: Статистика, 1973. – 265 с.
50. Лифшиц, А.Л. Статистическое моделирование систем массового обслуживания / А.Л. Лифшиц, Э.А. Мальц. – М.: Советское радио, 1978. – 216 с.
51. Дынкин, Е.Б. Управляемые марковские процессы и их приложения / Е.Б. Дынкин, А.А. Юшкевич. – М: Наука, 1975. – 154 с.
52. Рыков, В.В. Управляемые СМО / В.В. Рыков // Итоги науки и техники. Сер. Теория вероятностей. Математическая статистика. Теоретическая кибернетика. – 1974. – Т. 12. – С. 43-153.
53. Crabill, T.V., Gross D., Magazine M. J. A classified bibliography of research on optimal design and control of queues / T.V. Crabill, D. Gross, M. J. Magazine // Operation Research. – 1977. – V. 25. – С. 219-232.
54. Андронов, А.М. Теория массового обслуживания и научная организация труда в гражданской авиации / Андронов А.М. – М.: РИО МГА СССР, 1969. – 118 с.
55. Бусленко, Н.П. Автоматизация имитационного моделирования сложных систем / Н.П. Бусленко. – М.: Наука, 1979. – 239 с.
56. Шеннон, Р. Имитационное моделирование систем – искусство и наука / Р. Шеннон. – М.: Мир, 1978. – 420 с.
57. Моисеев, Н.Н. Имитационные системы / Н.Н. Моисеев, Ю.Г. Евтушенко, П.С. Краснощеков, Ю.Н. Павловский // Экономика и организация промышленного производства. Новосибирск: Наука СО. – 1973. – №6. – С. 39-46.
58. Клейнен, Дж. Статистические методы в имитационном моделировании / Дж. Клейнен. – М.: Статистика, 1978. – 290 с.

59. Ивченко, Г.И. Теория массового обслуживания / Г.И. Ивченко, В.А. Каштанов, И.Н. Коваленко. – М.: Высшая школа, 1982. – 256 с.
60. Матвеев, В.Ф. Системы массового обслуживания / В.Ф. Матвеев, В.Г. Ушаков. – М.: Изд-во МГУ, 1984. – 242 с.
61. Бочаров, П. П. Теория массового обслуживания / П. П. Бочаров, А.В. Печинкин. – М.: Изд-во РУДН, 1995. – 529 с.
62. Назаров, А.А., Моисеева С.П. Метод асимптотического анализа в теории массового обслуживания / А.А. Назаров, С.П. Моисеева. – Томск: Изд-во НТЛ, 2006. – 112 с.
63. Зейфман, А.И. Марковские цепи и модели с непрерывным временем / А.И. Зейфман, В.Е. Бенинг, И.А. Соколов. – М.: ЭЛЕКС-КМ, 2008. – 168 с.
64. Наумов, В. А. Мультипликативные решения конечных цепей Маркова / В.А. Наумов, К.Е. Самуйлов, Ю.В. Гайдамака – М.: РУДН, 2015. – 159 с.
65. Моисеев, А. Н. Бесконечнолинейные системы и сети массового обслуживания / А. Н. Моисеев, А. А. Назаров. – Томск : Изд-во НТЛ, 2015. – 240 с.
66. Бахарева, Н. Ф. Аппроксимативные методы и модели массового обслуживания. Исследование компьютерных сетей : монография / Н. Ф. Бахарева, В. Н. Тарасов. – Самара: СНЦ РАН, 2017. – 328 с.
67. Морозов, Е.В. Критерий стационарности одного класса непуассоновских сетей обслуживания / Е.В. Морозов // Изв. АН СССР. Техническая кибернетика. – 1988. – № 1. – С. 129-133.
68. Коган, Я.Ф. Об асимптотике стационарного распределения длин очередей в замкнутых сетях в условиях большой загрузки / Я.Ф. Коган // Автоматика и телемеханика. – 1991. – № 10. – С. 112-120.
69. Ивницкий, В.А. Об условии инвариантности стационарных вероятностей состояний для сетей многолинейных систем массового

обслуживания с абсолютным приоритетом поступающего требования. / В.А. Ивницкий // Автоматика и вычислительная техника. – 1997. – № 4. С. 59-68.

70. Наумов, В.А. Приближенный анализ вероятностно-временных характеристик многофазных систем массового обслуживания / В.А. Наумов, П.О. Абаев // Вестник РУДН. Серия «Прикладная и компьютерная математика». – 2007. - № 3. – Вып. 4. – С. 64-77.

71. Абрамов, П.Б. Существование и устойчивость решения систем дифференциальных уравнений для разомкнутых моделей систем массового обслуживания / П.Б. Абрамов, А.В. Леньшин // Вестник Воронежского ин-та МВД России. – 2013. - № 4. – С. 182-189.

72. Зейфман, А.И. Оценки погрешности аппроксимаций неоднородных марковских цепей с непрерывным временем / А.И. Зейфман, А.В. Коротышева, В.Ю. Королев, Я.А. Сатин // Теория вероятн. и ее примен. – 2016. – Т. 61. – Вып. 3. – С. 563-569.

73. Кутузов, О. И. К оцениванию и сопоставлению классических и фрактальных систем массового обслуживания / О. И. Кутузов, Т. М. Татарникова // Информационно-управляющие системы. – 2016. – № 2(81). – С. 48-55.

74. Назаров, А.А. Управляемые системы массового обслуживания и их оптимизация / А.А. Назаров. – Томск: Изд-во ТГУ, 1984. - 234 с.

75. Дудин, А.Н. Оптимальное управление ненадежной двухскоростной системой массового обслуживания / А. Н. Дудин // Автоматика и телемеханика. – 1985. – № 9. – С. 56-62.

76. Герасимов А.И., Оптимизация замкнутых сетей массового обслуживания с несколькими классами требований / А.И. Герасимов // Проблемы передачи информации. – 1994. – Т. 30. – Вып. 1. – С. 85-96.

77. Boel, R. K., Performance analysis and optimal threshold policies for queueing systems with several heterogeneous servers and Markovmodulated arrivals / R.K. Boel., N.K. Talat // Technical Report, Katholik University, Belgium. –1997. – 22 С.

78. Рыков, В. В. Об условиях монотонности оптимальных политик управления системами массового обслуживания / В.В. Рыков // Автоматика и Телемеханика. – 1999. – Т. 9. – С. 92-106.
79. Altman, E., Jimenez T., Koole G. On optimal call admission control in a resource-sharing system / E. Altman, T. Jimenez, G. Koole // Technical Report, Vrije University, Amsterdam. – 2000. – 18 С.
80. Задорожный, В. Н. Оптимизация однородных немарковских сетей массового обслуживания / В. Н. Задорожный // Проблемы управления. – 2009. – № 6. – С. 68-75.
81. Тарасов, В.Н. Оптимизация расчета характеристик системы $M_2/M/1$ / В.Н. Тарасов, Н.Ф. Бахарева, И.В. Карташевский, Л.В. Липилина // Инфокоммуникационные технологии. – 2017. – №4. – С.360-364.
82. Ионин, Г.Л. Статистическое моделирование систем телетрафика / Г.Л. Ионин, Я.Я. Седол. – М.: Радио и связь, 1982. – 184 с.
83. Корнышев, Ю.Н. Теория телетрафика. Учебник для вузов / Ю.Н. Корнышев, А.П. Пшеничников, А.Д. Харкевич. – М.: Радио и связь, 1996. – 272 с.
84. Башарин, Г.П. Лекции по математической теории телетрафика: Учеб. пособие. Изд. 3-е, исп. и доп. / Г.П. Башарин. – М.: Изд-во РУДН, 2009. – 342 с.
85. Гайдамака, Ю.В. Модели обслуживания вызовов в сети сотовой подвижной связи: Учебно-метод. Пособие / Ю.В. Гайдамака, Э.Р. Зарипова, К.Е. Самуйлов. – М.: Изд-во РУДН, 2008. – 72 с.
86. Абаев, П.О. Моделирование работы SIP-сервера с помощью системы массового обслуживания с гистерезисом и прогулками в дискретном времени / П.О. Абаев, Р.В. Разумчик // Т-Comm - Телекоммуникации и транспорт. - 2012. - № 7. - С. 5-8.
87. Гайдамака, Ю.В. Анализ некоторых характеристик СМО $M |G|1|r$ с гистерезисным управлением для исследования перегрузок SIP-сервера /

Ю.В. Гайдамака, Р.И. Закирова // Вестник РУДН. Серия «Математика. Информатика. Физика». – М.: Изд-во РУДН. – 2014. – № 2. – С. 43-50.

88. Гайдамака, Ю.В. К анализу эффектов группового поступления сообщений на время ожидания в очереди обработки SIP-сервера / Ю.В. Гайдамака, Р.И. Закирова, Ю.Н. Орлов // Информационные технологии и вычислительные системы. – 2015. – № 3. – С. 74-80.

89. Жожикашвили, В.А. Сети массового обслуживания. Теория и применение к сетям ЭВМ / В.А. Жожикашвили, В.М. Вишневский – М.: Радио и связь, 1988. – 192 с.

90. Яшков, С.Ф. Математические вопросы теории систем обслуживания с разделением процессора / С.Ф. Яшков // Итоги науки и техники. Серия «Теория вероятностей. Математическая статистика. Теоретическая кибернетика». – 1990. – Т. 29. – С. 3-82.

91. Вишневский, В. М. Теоретические основы проектирования компьютерных сетей: монография / В. М. Вишневский. – Москва : Техносфера, 2003. – 512 с.

92. Takagi, H. A tutorial on the analysis of polling systems / H. Takagi, L. Kleinrock // Computer Science Department, University of California, Los Angeles. – Report No. CSD-850005. – 1985. – 172 p.

93. Lichtzinder, B. Ya. Generalization of Formulas for Queue Length Moments under Nonordinary Poissonian Arrivals for Batch Queues in Telecommunication Systems / B. Ya. Lichtzinder, A. Yu. Privalov // Problems of Information Transmission. – 2023. – Vol. 59. – No. 4. – P. 243-248.

94. Лихтциндер, Б. Я. Два алгоритма формирования группового квазипуассоновского потока / Б. Я. Лихтциндер, А. Ю. Привалов, Р. И. Гильмуллин // Инфокоммуникационные технологии. – 2024. – Т. 22, № 2(86). – С. 13-19.

95. Вишневский, В.М. Модель системы поллинга для исследования широкополосных беспроводных сетей / В.М. Вишневский, О.В. Семенова, С.А. Шпилев // Автоматика И Телемеханика. – 2006. – № 12. – С. 123-135.

96. Маликова, Е.Е. Стратегия группового поллинга в широкополосных беспроводных сетях мониторинга / Е.Е. Маликова, И.И. Цитович // *Обозрение прикладной и промышленной математики*. – 2010. – Т. 17. – №2. – С. 284-285.
97. Вишневский, В. М. Исследование системы поллинга с адаптивным циклическим опросом и ее применение для проектирования широкополосных беспроводных сетей / В. М. Вишневский, О. В. Семенова, З. Т. Буй // *Пробл. управл.* – 2020. – № 5. – С. 50–55.
98. Рыков, В.В. К анализу поллинг-систем / В. В. Рыков// *Автоматика и телемеханика*. – 2009. – № 6. – С. 90–114.
99. Рыжиков Ю.И. Теория очередей и управление запасами / Ю.И. Рыжиков. – СПб.: Питер, 2001. – 384 с.
100. Якимов, И.М. Оценка достоверности результатов имитационного моделирования по результатам аналитического моделирования / И.М. Якимов, А.П. Кирпичников, Г.Р. Зайнуллина, З.Т. Яхина // *Вестник технологического университета*. – 2015. – Т. 18. – № 6. – С. 173-178.
101. Якимов, И.М. Сравнение результатов имитационного моделирования вероятностных объектов в системах: Anylogic, Arena, Bizagi Modeler, Grps W / И.М. Якимов, А.П. Кирпичников, Ю.Г. Исаева // *Вестник технологического университета*. – 2015. – Т. 18. – № 16. – С. 260-264.
102. Кирпичников, А.П. Методы прикладной теории массового обслуживания / А.П. Кирпичников – Казань: Изд-во Казанск. гос. ун-та, 2011. – 200 с.
103. Плескунов, М.А. Теория массового обслуживания: учебное пособие / М. А. Плескунов. – М-во науки и высшего образования РФ, Урал. федер. ун-т. — Екатеринбург : Изд-во Урал. ун-та, 2022. — 264 с.
104. Рыков В.В. Основы теории массового обслуживания (Основной курс: марковские модели, методы марковизации): учебное пособие / В.В. Рыков, Д.В. Козырев. – Москва: ИНФРА-М, 2024. – 223с.

105. Алиев, Т.И. Основы моделирования дискретных систем / Т.И. Алиев. – СПб: СПбГУ ИТМО. – 2009. – 363 с.
106. Лихтциндер, Б. Я. О недостатках классификации систем массового обслуживания / Б. Я. Лихтциндер // Вестник Самарского государственного технического университета. Серия: Технические науки. – 2020. – Т. 28, № 3(67). – С. 155-157. – DOI 10.14498/tech.2020.3.10. – EDN HJQRDU.
107. Рыжиков Ю.И. Теория очередей и управление запасами / Ю.И. Рыжиков. – СПб.: Питер, 2001. – 384 с.
108. Коваленко, И.Н Теория массового обслуживания / И.Н. Коваленко // Итоги науки. Сер. Теор. вероятн. – 1963, 1965. – С. 73– 125.
109. Вентцель, Е.С, Теория вероятностей. (Учебник для вузов). Изд. 2-е переработ, и доп. / Е.С. Вентцель – М., Физматгиз, 1962. – 564 с.
110. Коваленко, И. Н. Исследование многолинейной системы обслуживания с очередью и ограниченным временем пребывания в системе / И. Н. Коваленко // Укр. матем. ж.. – 12, 4 (1960). – С. 471—476.
111. Коваленко, И. Н. Некоторые задачи массового обслуживания с ограничением / И. Н. Коваленко // Теория вероятностей и ее применение. – 1961. – Т.6, в. 2. – С. 222—228.
112. Афанасьева, Л. Г. О существовании предельного распределения в системах массового обслуживания с ограниченным временем пребывания / Л. Г. Афанасьева // Теория вероятностей и ее применения. – 1965. – Т.10, № 3. – С. 570— 578.
113. Афанасьева, Л. Г. Аналитический вид распределения времени ожидания в некоторых системах обслуживания с ограниченным временем пребывания / Л. Г. Афанасьева // Научи. тр. Краснодарск. политехи, ин-т. – 1968. – Вып. 23, № 10. – С. 109—117.
114. Афанасьева, Л. Г. О некоторых задачах массового обслуживания с ограниченным временем ожидания / Л. Г. Афанасьева // Изв. АН ССОР. Техн. Кибернетика. – 1964. – № 6. С. 27—37.

115. Cohen, J. W. On two integral equations of queueing theory / J. W. Cohen // J. Appl. Probab.. – 1967. – 4, № 2. – С. 343-355 (РЖ/Мат, 1968, 4В51).
116. Daley, D. J. General customer impatience in the queue $G1/G/1$ / D. J. Daley – Appl. Probab. – 1965. – 2, № 1, 186-205 (РЖМат, 1966, 1В61).
117. Рыков, В. В. Управляемые системы массового обслуживания / В.В. Рыков // В кн.: Итоги науки и техники. Сер. Теория вероятностей. Математическая статистика. Теоретическая кибернетика. – М.: ВИНТИ. – 1975. – Т. 12. – С. 43-153.
118. Бурлаков, М. В. О динамической сортировке заявок в системах с ограниченным временем пребывания / М. В. Бурлаков, Е. Е. Подоляк, Е. И. Шаповал // Автоматика и телемеханика. – 1985. – Выпуск 10. – С.39-43.
119. Дудин С.А. Модель функционирования колл-центра как система $MAR/PH/N/R-N$ с нетерпеливыми запросами / С. А. Дудин, О. С. Дудина // Пробл. передачи информ. – 47:4 (2011). С. 68–83; Problems Inform. Transmission, 47:4 (2011), 364–377.
120. Полетаев, В. П. Оптимизация управления пропускной способностью системы массового обслуживания с ожиданием / В. П. Полетаев, А. Н. Ершов // Труды международного симпозиума "Надежность и качество". – 2007. – Т. 1. – С. 239-240.
121. Самочернова, Л. И. Оптимизация системы массового обслуживания с резервным прибором с управлением, зависящим от времени ожидания / Л. И. Самочернова // Известия Томского политехнического университета. – 2010. – Т. 316, № 5. – С. 94-97.
122. Сеченова С. А. Исследование процесса времени ожидания в системе массового обслуживания типа $M/GI/N/\infty$ / С. А. Сеченова, Е. Ю. Лисовская // Научное творчество молодежи. Математика. Информатика [Электронный ресурс]: материалы XIX Всероссийской научно-практической конференции, 15-16 мая 2015 г. Анжеро-Судженск, 2015. С. 75-79. URL: <http://vital.lib.tsu.ru/vital/access/manager/Repository/vtls:000616189>

123. Афонин, В. В. Оптимизация марковских систем массового обслуживания с ожиданием в системе MATLAB / В. В. Афонин, В. В. Никулин // Вестник Астраханского государственного технического университета. Серия: Управление, вычислительная техника и информатика. – 2017. – № 2. – С. 39-47.
124. Рыжиков, Ю.И. Расчет сети обслуживания с ограничением времени жизни заявок / Ю.И. Рыжиков, А.В. Уланов // XII Всероссийское совещание по проблемам управления. – М., 16-19 июня 2014. – С.8620-8624.
125. Савинов Ю. Г. Оптимизация в СМО с нетерпеливыми заявками / Ю. Г. Савинов, Е. Д. Табакова, И. Д. Сафиуллов // Ученые записки УлГУ. Серия «Математика и информационные технологии». – 2019. – Выпуск 1. – С.92–98.
126. Бобков, С. П. Имитационное моделирование системы массового обслуживания с целью анализа ее работы / С. П. Бобков, И. А. Астраханцева, В. С. Волков // Современные наукоемкие технологии. Региональное приложение. – 2021. – № 3(67). – С. 58-62.
127. Бочаров, П. П. Стационарное распределение очереди в системе обслуживания конечной емкости с групповым потоком и временем обслуживания фазового типа / П. П. Бочаров, С. В. Якутина // Автомат. и телемех. – 1994. – Выпуск 9. – С. 106–119.
128. Монсик, В. Б. Показатели эффективности функционирования системы массового обслуживания с неординарным входным потоком заявок в нестационарном режиме работы / В. Б. Монсик, А. А. Скрынников, А. Ю. Федотов // Научный вестник Московского государственного технического университета гражданской авиации. – 2009. – № 145. – С. 113-118.
129. Назаров, А. А. Исследование бесконечнолинейной системы обслуживания с отрицательными заявками и их ожиданием в пустой системе / А. А. Назаров, Н. М. Феропонтова // Известия вузов. Физика. – 2015. – Т. 58, № 11-2. – С. 204-210.

130. Монько, В. Д. Асимптотический анализ экспоненциальной сети с ограниченным временем ожидания заявок и ненадежными системами обслуживания / В. Д. Монько, Д. Я. Копать, С. Э. Статкевич // Вестник Гродненского государственного университета имени Янки Купалы. Серия 2. Математика. Физика. Информатика, вычислительная техника и управление. – 2015. – № 1(186). – С. 138-147.
131. Бусарев, М. И. Одноканальная система массового обслуживания с ограниченным средним временем пребывания заявки в системе в целом / М. И. Бусарев, А. П. Кирпичников, Д. Б. Флакс // Вестник Казанского технологического университета. – 2011. – № 22. – С. 155-161.
132. Кирпичников, А. П. Системы массового обслуживания с ограниченным временем пребывания заявки в системе / А. П. Кирпичников, Д. Б. Флакс, Л. Р. Валеева // Актуальные проблемы гуманитарных и естественных наук. – 2015. – № 6-1. – С. 68-73.
133. Анализ динамической модели «Оптимизация дорожного движения на участке дороги» в среде STATISTICA и Anylogic / В. В. Мокшин, А. П. Кирпичников, Р. Р. Шириазданов, Ш. А. Шайхутдинов // Вестник Технологического университета. – 2019. – Т. 22, № 5. – С. 119-127.
134. Моделирование задачи производства изделий с помощью AnyLogic / Л.М. Шарнин, А.П. Кирпичников, Р.А. Нитшаев [и др.] // Вестник Технологического университета. – 2019. – Т. 22, № 4. – С. 153-157.
135. Девятков, В.В. Методология и технология имитационных исследований сложных систем: современное состояние и перспективы развития: монография / В.В. Девятков. – Москва: Вузовский учебник: ИНФРА-М, 2021. – 445 с.
136. Гнеденко, Б.В., Über einige Aspekte der Entwicklung der Theorie der Warteschlangen / Б.В. Гнеденко // Mathematik Technik Wirtschaft. – 1960. –Н. 3. – S.160-166.
137. Коваленко, И.Н. Теория вероятностей и ее применение / И.Н. Коваленко. – 6, 2, 222-228 (1961).

138. Лоу, А.М. Имитационное моделирование. Классика CS. 3-е изд. / А.М. Лоу, В.Д. Кельтон – СПб.: Питер; Киев: Издательская группа BHV, 2004. – 847 с.
139. Нуриев, Н.К. Математическое моделирование систем массового обслуживания заявок с ограниченным временем жизни / Н.К. Нуриев, Н.К. Шайдуллина, Е.А. Печеный // Вестник технологического университета. – 2024. – Т.27, №12. – С.165-169.
140. Шайдуллина, Н.К. Математическое моделирование и решение задачи нахождения характеристик системы массового обслуживания заявок со случайным временем жизни / Н.К. Шайдуллина, Н.К. Нуриев, Е.А. Печеный // Современные наукоемкие технологии. – 2026. – №6. – С. 204-209.
141. Шайдуллина, Н.К. Численный анализ поведения смешанного потока заявок с ограниченным временем жизни / Н.К. Шайдуллина, Е.А. Печеный, Н.К. Нуриев // Современные наукоемкие технологии. – 2024. – №10. – С. 94-99.
142. Шайдуллина Н.К. Численное решение задачи эффективного обслуживания группы заявок с ограниченным временем жизни / Н.К. Шайдуллина // Сборник материалов XXXIX Международной научно-практической конференции «Актуальные проблемы науки и образования в условиях современных вызовов». Сб. материалов XXXIX Международной научно-практической конференции. – Москва: ООО «Издательство Академическая среда». – 2025. – С.408-413.
143. Шайдуллина, Н.К. Проблема эффективного администрирования системы массового обслуживания с ограниченным временем жизни заявок / Н.К. Шайдуллина, Е.А. Печеный, Н.К. Нуриев // Современные наукоемкие технологии. – 2022. – №7. – С. 69-73.
144. Шайдуллина, Н.К. Модель управления системой массового обслуживания с ограниченным временем жизни заявок / Н.К. Шайдуллина, Е.А. Печеный, Н.К. Нуриев // Научно-технический вестник Поволжья. – 2023. – №7. – С.102-104.

145. Шайдуллина, Н.К. Имитационное моделирование процесса администрирования продажи товаров с ограниченным сроком годности / Вестник технологического университета. – 2022. – Т.25, №12. – С.158-161.

146. Шайдуллина, Н.К. Моделирование процесса администрирования системы массового обслуживания с ограниченным временем жизни заявок / Н.К. Шайдуллина, Е.А. Печеный, Н.К. Нуриев // Современные наукоемкие технологии. – 2023. – №11. – С.81-86.

147. Н.К. Шайдуллина, Задача управления системой массового обслуживания с ограниченным временем жизни заявок / Научно-технический вестник Поволжья. – 2023. – №12. – С.157-159.

148. Шайдуллина Н.К. Свидетельство о регистрации электронного ресурса №25476 «Эффективное обслуживания группы заявок с ограниченным временем жизни». – М.: ОФЭРНиО, 2025.

ПРИЛОЖЕНИЯ





ООО «ТД «ПОЛИХИМТРЕЙД»
420111, Казань, ул. Большая красная 13, оф 1-6
Тел. +7 987 225 77 71/+7 987 006 20 80
e-mail: Dmitriy800.16@gmail.com
ОГРН 1171690012086 / ИНН 1655380789
Расчетный счёт: 40702810762000024240
доп. офис номер 8610/0077 ПАО Сбербанк

20.01.2026

г. Казань

В диссертационный совет 24.2.312.08,
функционирующий в ФГБОУ ВО
«Казанский национальный
исследовательский
технологический университет»

Справка

о внедрении результатов диссертационного исследования Шайдуллиной Натальи Константиновны «Моделирование систем массового обслуживания ограниченным временем жизни заявок» на соискание учёной степени кандидата технических наук по специальности 1.2.2. Математическое моделирование, численные методы и комплексы программ.

Настоящий документ подтверждает, что результаты диссертационного исследования Шайдуллиной Н.К. нашли применение в деятельности предприятия ООО «ТД ПОЛИХИМТРЕЙД». В настоящий момент они используются для оценки эффективности производственной деятельности и сбыта выпускаемой предприятием продукции. Планируется применение результатов исследований Шайдуллиной Н.К. на стадии предпроектных работ по запуску новых линий производства.

Генеральный директор
ООО "ТД "ПОЛИХИМТРЕЙД"

Зайцев Д. В.