

ФЕДЕРАЛЬНОЕ АГЕНТСТВО СВЯЗИ
ОРДЕНА ТРУДОВОГО КРАСНОГО ЗНАМЕНИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
МОСКОВСКИЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ СВЯЗИ И
ИНФОРМАТИКИ

На правах рукописи

СТЕПАНОВ МИХАИЛ СЕРГЕЕВИЧ

Разработка и анализ обобщённой
модели обслуживания вызовов в
перспективных контакт-центрах

Специальность 05.12.13 — Системы, сети и устройства
телекоммуникаций

Диссертация на соискание ученой степени
кандидата технических наук

Научный руководитель:
кандидат технических наук, профессор
Пшеничников Анатолий Павлович

Москва — 2016 г.

Аннотация

Диссертационная работа посвящена построению обобщенной модели обслуживания заявок в действующих и перспективных контакт-центрах и ее использованию для оценки показателей качества обслуживания клиентов. Для решения этой задачи процесс поступления и обслуживания заявок представлен в виде математической модели, описываемой многомерным марковским процессом.

При разработке модели учтены следующие особенности функционирования современных контакт-центров: наличие устройств IVR; использование операторов и консультантов; наличие мест ожидания начала обслуживания у оператора или консультанта; наличие ограничения на максимально возможное время пребывания на ожидании начала обслуживания; возможность повторения заблокированной заявки; возможность поступления заявки на обслуживание по каналам сети Интернет. Интервалы времени между осуществлением событий в модели имеют экспоненциальное распределение и не зависят друг от друга, как и вероятности переходов из состояния в состояние.

Даны определения основным показателям качества обслуживания заявок. Формальные выражения для характеристик получены через значения входных параметров и величины стационарных вероятностей модели. Построена система уравнений статистического равновесия, которая используется для оценки значений стационарных вероятностей модели. Точные величины характеристик получены путем решения системы уравнений равновесия итерационным методом Гаусса-Зейделя.

Разработаны алгоритмы приближенной оценки показателей качества обслуживания заявок, основанные на реализации принципа декомпозиции исследуемой системы на отдельные части, а также на применении асимптотических свойств модели при стремлении к бесконечности интервала времени между последовательными поступлениями повторных вызовов. Показано применение полученных результатов при оценке необходимого числа линий доступа, операторов, устройств IVR, а также величин характеристик, прямое измерение которых затруднено из-за необходимости разделения первичных вызовов и повторных. Построенная модель и разработанные на ее основе расчетные алгоритмы использованы в ПАО МГТС в виде методики оценки необходимого числа операторов справочно-информационной службы, а также использованы в учебном процессе на кафедре ССиСК МТУСИ, что подтверждено соответствующими актами.

Оглавление

Аннотация	1
Введение	5
Глава 1. Характеристика и анализ работы контакт-центров	9
1.1. Современное состояние и перспективы развития	9
1.1.1. Актуальность проблематики	9
1.1.2. Основные понятия и термины	11
1.1.3. Области использования контакт-центров	13
1.1.4. Бизнес-преимущества контакт-центров	13
1.1.5. Тенденции развития	14
1.1.6. Цели и задачи научных исследований	16
1.2. Основные структурные элементы и их взаимосвязь	17
1.2.1. Ресурс обслуживания заявок	17
1.2.2. Операторы и консультанты	17
1.2.3. Устройства интерактивного речевого взаимодействия (IVR)	20
1.3. Сценарии обработки запросов	23
1.4. Показатели поступления и обслуживания заявок	26
1.4.1. Типы и классификация характеристик	26
1.4.2. Операционные режимы	30
1.4.3. Измерение характеристик контакт-центра	30
1.5. Вопросы планирования работы контакт-центров	32
1.5.1. Методология планирования	32
1.5.2. Оценка числа операторов	33
1.6. Анализ выполненных исследований	38
1.7. Выводы и постановка задачи диссертационного исследования	45
Глава 2. Построение и анализ математической модели контакт-центра	47
2.1. Введение	47
2.2. Математическое описание модели	48
2.2.1. Схема функционирования	48

2.2.2.	Численные значения входных параметров	53
2.2.3.	Состояния модели и динамика их изменения	54
2.3.	Определение показателей качества обслуживания заявок	57
2.3.1.	Систематизация характеристик	57
2.3.2.	Средние значения функциональных компонент	57
2.3.3.	Интенсивности потоков заявок на этапах обслуживания	59
2.3.4.	Потери и длительности ожидания на этапах обслуживания	61
2.3.5.	Интегральные показатели качества обслуживания заявок	62
2.4.	Законы сохранения интенсивностей потоков заявок	63
2.4.1.	Вывод законов сохранения	63
2.4.2.	Косвенная оценка сходимости итерационного метода	66
2.4.3.	Оценка интенсивности поступления первичных заявок	67
2.4.4.	Оценка показателей обслуживания заявок	67
2.5.	Частные случаи построенной модели	68
2.5.1.	Условия выбора моделей	68
2.5.2.	Отсутствие консультантов	69
2.5.3.	Отсутствие операторского обслуживания	69
2.5.4.	Учет ограниченного ожидания	70
2.5.5.	Учёт ограниченного ожидания и возможности повторного запроса	75
2.5.6.	Отсутствие возможности повторного запроса	79
2.6.	Выводы по результатам второй главы	79
Глава 3.	Оценка показателей качества обслуживания заявок	81
3.1.	Введение	81
3.2.	Точные и приближённые методы расчёта модели	82
3.3.	Построение системы уравнений равновесия	85
3.3.1.	Перечень анализируемых событий	85
3.3.2.	Формирование левой части	86
3.3.3.	Формирование правой части	92
3.3.4.	Система уравнений равновесия	96
3.3.5.	Ограничение на число абонентов, повторяющих вызовы	98
3.4.	Формулировка рекурсивной схемы расчётов	99
3.5.	Численный анализ сходимости рекурсии	102
3.6.	Зависимость характеристик от параметров модели	103
3.7.	Выводы по результатам третьей главы	109
Глава 4.	Использование разработанной модели для решения задач плани-	
	рования и повышения эффективности работы контакт-центра	112
4.1.	Введение	112
4.2.	Асимптотические методы	113

4.2.1. Основные положения	113
4.2.2. Расчётные соотношения	114
4.2.3. Повышение эффективности расчётов	116
4.2.4. Общий случай	117
4.2.5. Оценка погрешности	118
4.3. Метод декомпозиции	120
4.3.1. Основные положения	120
4.3.2. Расчётные соотношения	122
4.4. Оценка влияния повторных вызовов на число операторов	125
4.5. Планирование числа операторов и линий доступа	127
4.6. Оценка влияния IVR на численность операторов	129
4.7. Выводы по результатам четвертой главы	134
Заключение	136
Список литературы	138
Приложение 1. Численный анализ характеристик контакт-центра	146
Приложение 2. Акты об использовании результатов диссертации	150

Введение

Актуальность темы исследования. Во всем мире расширяется использование контакт-центров для предоставления разного рода справочно-информационных услуг. Согласно статистике, более 60 % операционных затрат подобных служб составляет заработная плата операторов. По этой причине оптимизация численности операторов является одной из ключевых задач администрации контакт-центров. Ее решение невозможно без построения математической модели, адекватно описывающей работу контакт-центров. Использование классических моделей с потерями или ожиданием приводит к значительным погрешностям, поскольку в них не учитываются особенности построения действующих и перспективных контакт-центров. К ним относятся: наличие предварительного голосового обслуживания (IVR); использование операторов и консультантов; учет возможности повторения заявки из-за нехватки необходимого ресурса; возможность поступления заявок по каналам сети Интернет.

Таким образом, построение обобщенной модели контакт-центра, учитывающей перечисленные выше особенности их функционирования, проведение ее исследования и разработка методов расчета характеристик является актуальной задачей, имеющей большое значение для повышения эффективности использования инфраструктуры контакт-центров. Ее решение позволит находить оптимальные соотношения между параметрами потоков запросов на информационное обслуживание и характеристиками пропускной способности контакт-центра, обеспечивающие гарантированное качество обслуживания клиентов.

Степень разработанности темы. Поставленная задача решалась на базе моделей и методов теории телетрафика, а также возможностей, заложенных в механизмы управления процессом обслуживания заявок в современных контакт-центрах. Различным аспектам решения данной задачи посвящены работы российских и зарубежных авторов: Башарина Г.П., Гайдамаки Ю.В., Гольдштейна Б.С., Ефимушкина В.А., Кучерявого А.Е., Пшеничникова А.П., Рослякова А.В., Степанова С.Н., Самуйлова К.Е., Соколова Н.А., Borst S., Brandt A., Dawson K., Koole G., Mandelbaum A., Stollitz R., Reiman M., Zeltyn S. и др. Отдельные вопросы построения и исследования моделей call-и контакт-центров рассматривались в диссертационных работах: Белова М.Н., Бельской Н.М., Ваняшина С.В., Дибби В.Н., Зарубина А.А., Ухловской Л.Г., Чан Туан Миня и др. Анализ публикаций и выполненных диссертационных исследований показал, что в большинстве работ либо изучалось действие какого-то одного фактора на процесс функционирования справочного-информационного центра (например, квалификация операторов, наличие автоответчиков, возможность ожидания начала обслуживания, повторение заблокированной заявки, поствызывная работа оператора, наличие нескольких видов запросов и т.д.), либо предлагались эмпирические методики, направленные на повышение эффективности мониторинга и управления работой контакт-центра. При этом исполь-

зовались самые простые модели теории телетрафика, естественно, после соответствующей интерпретации входных параметров и характеристик. Задача анализа комплексного воздействия перечисленных факторов на работу контакт-центра в должном объеме не рассматривалась, что и определило направление исследований, выполненное в диссертации.

Цели и задачи работы. Построение и исследование обобщенной модели обслуживания заявок в контакт-центре с учётом совместного влияния основных факторов, определяющих их функционирование; разработка точных и приближенных алгоритмов оценки показателей качества обслуживания клиентов; разработка методик оценки необходимого числа линий доступа, операторов, устройств IVR, а также значений параметров, прямое измерение которых затруднено из-за необходимости разделения поступающих заявок на первичные и повторные.

Научная новизна.

1. Построена и исследована обобщенная модель обслуживания заявок в контакт-центре, которая позволила учесть совместное влияние основных значимых факторов, определяющих работу современных справочных служб. Среди них: дифференцированное обслуживание клиентов (IVR, операторы и консультанты); наличие ограничения на время пребывания заявки на ожидании; возможность повторения заблокированной заявки и возможность поступления заявок по каналам сети Интернет.
2. Построена система уравнений статистического равновесия, связывающая значения стационарных вероятностей модели и разработан алгоритм ее решения. Получены выражения для оценки характеристик качества обслуживания заявок через значения входных параметров и стационарных вероятностей обобщенной модели контакт-центра. Среди них: интенсивности и потери заявок, дифференцированные по этапам обслуживания; средние значения функциональных компонент состояния (среднее число абонентов, повторяющих вызов, среднее число занятых операторов и т.д.); вероятности и средние длительности пребывания заявок на ожидании начала обслуживания и др.
3. Получены соотношения между основными характеристиками работы контакт-центра. Показано их использование для косвенной оценки тех параметров и характеристик модели, которые затруднительно измерять прямыми методами, в частности, формулы для оценки величины интенсивности поступления первичных вызовов.
4. Разработаны две процедуры приближенного расчета построенной модели, основанные на применении асимптотических значений характеристик, полученных при стремлении к нулю интенсивности повторения запросов, и на реализации техники декомпозиции. В первом случае для вычисления оценок используются значения характеристик аналогичной модели контакт-центра, но без учета возможности повторения запроса и с увеличенной интенсивностью поступления первичных вызовов.

Во втором — исследуемая модель разбивается на отдельные части, которые затем рассчитываются независимо друг от друга. Влияние отдельных функциональных блоков друг на друга учитывается введением поправочных коэффициентов, которые получаются из результатов измерений или специальных соотношений. В исследуемой модели последовательность блоков выбирается из последовательности этапов обслуживания заявки: устройства IVR, операторы и консультанты.

5. Сформулирована двухэтапная процедура оценки числа операторов и линий доступа. Задача первого этапа — определение предварительного значения числа операторов при фиксированном количестве линий доступа. На втором этапе эта величина уточняется в результате последовательного изменения числа операторов и линий доступа и проверки достижения заданных значений показателей качества обслуживания поступающих заявок: доли потерянных заявок и средней величины задержки начала обслуживания.

Теоретическая и практическая значимость работы. Теоретическая значимость работы заключается в построении и исследовании обобщенной модели обслуживания заявок в контакт-центре, которая позволила учесть совместное влияние основных значимых факторов, определяющих функционирование современных справочных служб, а также в разработке точных и приближенных алгоритмов расчета подобных моделей. Получены программные реализации построенных в диссертации алгоритмов. Разработанный инструментарий рекомендуется использовать при решении задач планирования инфраструктуры контакт-центров и теоретическом обосновании действий администрации, направленных на повышение эффективности их работы. Результаты диссертации использованы в ПАО МГТС в виде методики оценки необходимого числа операторов справочно-информационной службы, а также использованы в учебном процессе на кафедре «Сети связи и системы коммутации» МТУСИ. Реализация результатов работы подтверждена соответствующими актами.

Методы исследования. Для решения поставленной задачи применялись методы теории телетрафика, теории вероятностей и вычислительной математики.

Основные положения, выносимые на защиту:

1. Анализ опыта работы контакт-центров, выполненный по результатам опубликованных исследований, показал эффективность дифференциации процесса обслуживания заявок. Она заключается в использовании устройств IVR и разделении операторов на отдельные группы по их компетенциям. Другой важной тенденцией является наличие существенной доли повторных запросов в поступающем потоке заявок, вызванных отказами на этапах их обслуживания.
2. Построенная обобщенная модель позволила учесть совместное влияние основных значимых факторов, определяющих качество работы современных контакт-центров. Среди них: наличие устройств IVR; использование операторов и консультантов;

наличие ограничения на время ожидания начала обслуживания; возможность повторения заявки; возможность поступления заявки по каналам сети Интернет.

3. Для оценки значений характеристик качества обслуживания заявок в рамках построенной модели контакт-центра рекомендуется использовать метод, основанный на решении системы уравнений равновесия итерационным алгоритмом Гаусса-Зейделя. Этот подход позволяет рассчитать характеристики для контакт-центров с числом операторов и консультантов, достигающим нескольких десятков.
4. В общем случае для оценки характеристик качества обслуживания заявок рекомендуется использовать приближенные методы, основанные на применении асимптотических значений характеристик, полученных при стремлении к нулю интенсивности повторения запросов, или на реализации техники декомпозиции. Численный анализ точности показал, что в большинстве практически интересных случаев относительная погрешность оценки характеристик лежит в пределах 5...10 %.
5. Разработанную модель и алгоритмы рекомендуется использовать для решения задач оценки числа операторов, консультантов и линий доступа, обеспечивающих заданные значения показателей качества обслуживания клиентов контакт-центра: доли потерянных заявок и средней величины задержки начала обслуживания, а также для исследования зависимости характеристик от особенностей поступления и обслуживания заявок, учитываемых в модели.

Степень достоверности и апробация результатов. Полученные теоретические результаты обоснованы доказательствами с использованием математических методов теории телетрафика, подтверждены численными экспериментами. Достоверность положений и выводов диссертации подтверждается апробацией работы, основные результаты которой обсуждались и докладывались на международной научно-технической конференции «Технологии информационного общества» (Москва, 2009–2014 гг.), на отраслевой научно-технической конференции «Телекоммуникационные и вычислительные системы» (Москва, 2009–2015 гг.), на Всероссийской конференции с международным участием «Информационно-телекоммуникационные технологии и математическое моделирование высокотехнологичных систем» (Москва, 2012, 2015 гг.), на кафедре «Сети связи и системы коммутации» МТУСИ. По материалам диссертации опубликованы 18 работ, в том числе 7 — в рецензируемых периодических изданиях, входящих в перечень ВАК при Минобрнауки России.

Основное содержание работы. Диссертация состоит из введения, четырех глав, заключения, списка литературы и приложений. Основная часть (без приложений) изложена на 146 страницах машинописного текста, содержит 40 рисунков и 8 таблиц; список литературы состоит из 106 наименований. Приложения изложены на 7 страницах машинописного текста.

Глава 1

Характеристика и анализ работы контакт-центров

1.1. Современное состояние и перспективы развития

1.1.1. Актуальность проблематики

Контакт-центр играет существенную роль в развитии бизнеса компаний, предоставляющих населению массовые услуги. Он является *единой точкой доступа* к бесплатной или платной информации о возможности получения и характеристиках продаваемых товаров и услуг. Наличие контакт-центра является решающим фактором в приобретении компанией лидирующих позиций на рынке. Конкуренцию чаще выигрывают не товары и услуги, а лучшие формы организации обслуживания существующих и потенциальных покупателей. По этой причине количество сотрудников, занятых в контакт-центрах, и перечень оказываемых ими услуг постоянно увеличивается.

По разным оценкам [70] в развитых странах число сотрудников контакт-центров составляет несколько процентов¹ от общего числа занятого трудового населения. По мнению экспертов ComNews², в России в 2015 г. объём рынка одних только аутсорсинговых справочных центров приблизился к 10,0 млрд рублей и будет иметь в ближайшие годы рост примерно в 10 %.

Контакт-центры находят самые разнообразные области использования. Они применяются в службах технической поддержки, в банковском и страховом деле, в организации маркетинговых исследований и продаж, служат основой при разработке центров личной и общественной безопасности (ситуационные центры). В зависимости от вида деятельности число операторов контакт-центра может меняться в пределах десяти человек, а может достигать и величины в несколько тысяч. В последнем случае операторы могут

¹В Великобритании — 1,5 %, в США — 4 %.

²См. выпуск от 18.02.2014 г.

разделяться по специализации или квалификации и обслуживать самые разнообразные запросы клиентов, а сами справочные службы могут быть территориально разнесены в пределах одной страны или даже нескольких стран. Учёт этих особенностей усложняет исследование контакт-центра, как единой точки доступа клиентов к справочно-информационному обслуживанию.

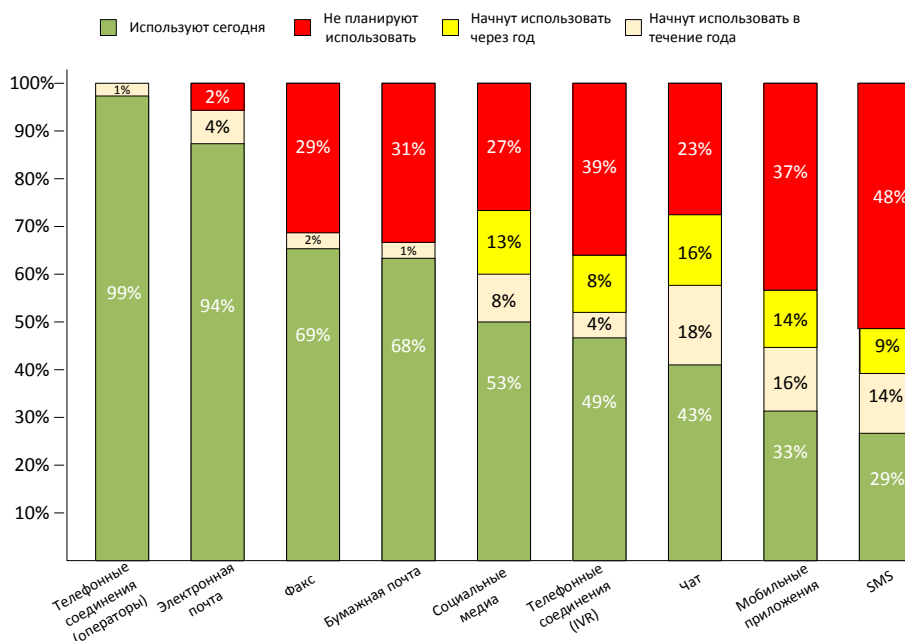


Рисунок 1.1. Каналы связи клиентов контакт-центров и перспективы их использования [80]

Для организации доступа клиентов к услугам контакт-центра применяются самые разные каналы связи. Их перечень определяется желанием администрации предоставить возможность приема заявки на информационное обслуживание от клиентов, находящихся *где угодно и в любой удобной и доступной для них форме*. В большинстве случаев — это обычное телефонное соединение, обеспечиваемое традиционными сетями фиксированной и подвижной связи. Заявка также может поступить с использованием возможностей и технологий сети Интернет: электронная почта, web-чат, социальные сети и т.п. Поиск требуемой информации часто осуществляется самим клиентом, а ответ может быть получен с применением автоответчиков. На рисунке 1.1, взятом из исследования компании Enghouse Interactive [80], показаны численные данные, иллюстрирующие текущее и будущее использование клиентами контакт-центров перечисленных выше коммуникационных каналов. Как видно из представленных данных, наибольшие ожидания операторы связывают с возможностями расширения каналов связи посредством социальных сетей, web-чатов и мобильных приложений.

Следует отметить, что не все клиенты готовы к полноценному общению со справочным центром с помощью неголосовых каналов связи. Об этом свидетельствуют данные, приведенные в таблице 1.1, подготовленной по материалам [106]. Из содержания таблицы

Таблица 1.1.

Предпочтения (в процентах) жителей разных стран по использованию различных видов доступа к услугам справочно-информационных центров [106]

Вид доступа	Страны									
	США	Канада	Мексика	Франц.	Герм.	Италия	Великобр.	Голл.	Австр.	Индия
Телефон	90	89	91	79	80	84	90	71	90	92
Видеоконф.	75	80	83	78	57	76	79	72	85	87
Электр. почта	67	65	71	71	73	74	72	60	68	81
Web-чат	47	38	72	54	36	54	38	30	33	73
СМС	22	17	39	22	17	48	26	11	22	65
Соц. сети	22	20	41	21	21	34	19	17	19	59
Автоотв.	20	13	36	17	14	27	10	6	10	49

видно, что клиенты по прежнему предпочитают общаться с оператором по привычным телефонным каналам связи. Однако количество запросов с использованием возможностей сети Интернет показывает устойчивый рост и в скором времени превысит голосовой трафик.

Приведенное краткое обсуждение показывает, что контакт-центр является достаточно сложным объектом. В нём взаимодействуют технические компоненты (разнообразные линии доступа и устройства коммутации, вычислительные комплексы), людские ресурсы (операторы, консультанты, супервизоры) и программные решения (управляющие системы, синтезаторы речи, базы данных). Эффективность работы контакт-центра во многом зависит от научно-обоснованной организации взаимодействия этих компонент. Различным аспектам решения этой задачи посвящено огромное число опубликованных статей, монографий и технических отчетов. Тем не менее, нерешенные проблемы остались и они станут предметом анализа в последующих разделах диссертационного исследования.

1.1.2. Основные понятия и термины

На русском языке нет общепринятого словосочетания для обозначения рассматриваемых справочно-информационных служб. Есть термин «центр обслуживания вызовов», зафиксированный в РД [65]. Однако в большинстве публикаций используются названия «call-центр» или «контакт-центр». Первое из них обозначает справочную службу с доступом клиентов только по телефонным линиям. Во втором случае также допускается приём заявок с использованием возможностей сети Интернет. Этот подход будет принят и в настоящем исследовании. Концептуальное определение контакт-центра сформулировано в РД [65]. Это: — *аппаратно-программный комплекс, предназначенный для приё-*

ма и обслуживания вызовов, поступающих на экстренные, справочно-информационные, консультационные, заказные, технические и другие службы, использующие для оказания услуг пользователям ресурсы телефонной сети связи общего пользования (ТфОП) и/или Интернет). Это довольно общее определение будет уточняться при рассмотрении конкретных справочно-информационных служб. Отметим, что в русскоязычной литературе по телетрафику [38,39] для обозначения контакт-центров также используется словосочетание «мультисервисный центр обслуживания вызовов».

Введём основные понятия и термины, которые применяются для описания моделей контакт-центров, выполним некоторую их классификацию и определимся с базовыми предположениями, которые будут использоваться в последующих разделах работы.

Типичный контакт-центр включает в себя «линии доступа» (англ. access lines), «операторов» (англ. agents) и «устройства интерактивного речевого взаимодействия» (англ. Interactive Voice Response — IVR). Перечисленные компоненты в совокупности составляют ресурс, который используется поступающими заявками для получения требуемого информационного обслуживания. При этом отдельные его виды также могут быть представлены в форме нескольких компонент. Так операторы могут подразделяться на группы в соответствии с их специализацией (англ. skilled agents) или квалификацией (англ. specialists). К ресурсу обслуживания заявок также относят каналы сети Интернет. Последовательность использования элементов ресурса при установлении соединения с контакт-центром и получении требуемой информации будет рассмотрена позднее.

Контакт-центры могут обслуживать только входящие заявки (англ. inbound contact center), только исходящие (англ. outbound contact center) или заявки обоих типов (англ. inbound/outbound or blended contact center). В дальнейшем будут рассматриваться справочные службы только первого типа. Ряд компаний не имеют достаточно средств, чтобы содержать собственный контакт-центр, и для исполнения его функций арендуют возможности других справочных служб. Такое решение называется «аутсорсингом». К нему прибегают небольшие компании, представленные в нишевых сегментах рынка, или компании, обладающие собственным контакт-центром, но с ограниченным спектром предоставляемых услуг. Аутсорсинг приносит значительную прибыль справочно-информационным службам и получил большое распространение. Будем далее предполагать, что заявка, переданная на обслуживание по процедуре аутсорсинга, не обладает приоритетом в занятии ресурса контакт-центра. В этом случае её появление не меняет процедуры обслуживания заявок и может быть отмечено при моделировании простым увеличением интенсивности их поступления.

Контакт-центр может иметь распределенную структуру. В этом случае он рассматривается как совокупность отдельных территориально разнесенных контакт-центров. Заявка, получившая отказ в одном из таких центров, пересылается в другой центр, имеющий в данный момент необходимый ресурс для ее обслуживания. Необходимость пересылки заявки может возникнуть и в случае, когда в центре первого выбора нет спе-

специалистов способных предоставить требуемое обслуживание. В модели эту особенность построения контакт-центра можно учесть разделением специалистов на операторов и консультантов.

1.1.3. Области использования контакт-центров

Контакт-центры широко применяются в разных сферах государственной деятельности и ведения бизнеса. Перечислим основные сегменты и приложения. К ним относятся:

- *Службы обеспечения государственной и личной безопасности (ситуационные центры).* Запрашиваемая и передаваемая информация связана с необходимостью поддержки требуемого уровня безопасности.
- *Транспортные службы.* Запрашиваемая и вносимая информация относится к получению сведений о расписании движения транспорта, резервировании билетов и их оплате.
- *Телекоммуникационные службы.* Запрашиваемая и вносимая информация связана с выбором тарифов, технической поддержкой, изменением условий подписки и т.д.
- *Розничные и оптовые продажи.* Характер запросов меняется от выяснения свойств товаров до их покупки. То обстоятельство, что в обсуждении свойств товара участвует оператор, повышает привлекательность этого способа продаж по сравнению с интернет-магазином.
- *Банки и страховые компании.* Запрашиваемая и вносимая информация относится к уточнению баланса лицевого счета, проведению транзакций, изменению условий договоров и т.п.
- *Государственные учреждения (больницы, учебные заведения и другие службы).* Запрашиваемая и вносимая информация связана получением справок, заполнением разного рода регистрационных форм и т.п.

Приведённый список подчеркивает важную роль, которую играют контакт-центры в жизни граждан, но не исчерпывает все возможности, перечень которых постоянно расширяется.

1.1.4. Бизнес-преимущества контакт-центров

Развитие телекоммуникационных технологий существенно расширили пути доступа клиентов к услугам современных контакт-центров. Если первые подобные службы по существу озвучивали телефонный справочник организации, то теперь и в ближайшей перспективе работа контакт-центр становится одной из важнейших составляющих

ведения бизнеса. Перечислим преимущества, которые обеспечивает компании наличие современного контакт-центра. По мнению экспертов [39] к ним относятся:

- Улучшение характеристик обслуживания потенциальных покупателей услуг или товаров компании, которое численно выражается в следующих изменениях их значений: уменьшается длительность промежутка времени ожидания ответа оператора; сокращаются затраты на обслуживание; увеличивается число запросов, получивших требуемую информацию с первого раза и т.д.
- Повышение лояльности потенциальных клиентов компании и увеличение их числа. Этот результат достигается за счёт расширения спектра предоставляемых услуг; привлечения сторонних консультантов; повышения благожелательности обслуживания со стороны операторов.
- Уменьшение затрат на обработку поступившей заявки за счет применения устройств интерактивного речевого взаимодействия, синтезаторов речи, и т.п.
- Возможность использования операторов для проведения маркетинговых акций и других подобных им мероприятий в целях продвижения товаров и услуг на рынке.
- Возможность предоставления наиболее востребованных услуг в круглосуточном режиме с помощью автоответчиков.
- Расширение числа оцениваемых характеристик качества обслуживания поступающих заявок и их дальнейшее использование при проведении мероприятий, направленных на повышение эффективности работы контакт-центра.
- Увеличение доходов за счет сдачи ресурсов контакт-центра компаниям, заинтересованным в развитии подобных услуг.

Наличие контакт-центра часто бывает решающим фактором для клиентов при выборе компании. Например в США 92 % клиентов формируют свое мнение о компании основываясь на опыте общения с её контакт-центром [82]. При этом 63 % клиентов отказываются от ее услуг, если работа контакт-центра признана неудовлетворительной. Эти факты ещё раз подтверждает необходимость использования научно обоснованных средств планирования элементов инфраструктуры анализируемых справочно-информационных служб.

1.1.5. Тенденции развития

В последнее десятилетие индустрия контакт-центров стремительно развивалась. По мнению экспертов, в ближайшие годы темпы роста несколько снизятся, но всё равно останутся выше, чем темпы роста телекоммуникационной отрасли в целом. Успехи в

развитии контакт-центров основаны на использовании новейших технических достижений, а так же на желании администрации компаний, представленных на массовом рынке, совершенствовать механизмы продаж товаров и услуг. В этом контексте контакт-центр превращается в стратегический актив компании, являющийся источником генерации новых доходов. Если говорить о тенденциях развития контакт-центров, то специалисты [98] отмечают следующие направления.

1. *Переход на модель аренды программных решений с использованием принципов SaaS (англ. Software as a Service — программное обеспечение как услуга).* Бизнес-модель SaaS — это часть глобальной концепции, известной как облачные решения (англ. cloud solution). Программные продукты, применяемые в работе контакт-центров, отличает высокая степень востребованности. Отсюда и следует экономическая целесообразность использования облачных подходов. В результате, небольшие контакт-центры получают возможность применения ранее недоступных программных продуктов. В свою очередь крупные справочные службы существенно экономят финансовые ресурсы.
2. *Размещение операторов в удалённом доступе, главным образом в домашних условиях.* Подобные решения устойчивы к разного рода бедствиям, включая преднамеренным. Другая позитивная черта — возможность гибкой организации работы операторов, снятие стресса, экономия в тратах операторов и контакт-центра и т.д. Техническая возможность реализации подобной модели ведения бизнеса обеспечивается применением облачных решений.
3. *Использование информационных возможностей социальных сетей.* Это могут быть рекламные акции или разного рода опросы.
4. *Учет индивидуальных особенностей клиента с целью предотвращения возможных проблем с его обслуживанием.* Клиенты различаются разной степенью переносимости ожидания начала обслуживания, общения с IVR или оператором. Эти проблемы должны решаться супервизорами контакт-центров для повышения лояльности клиентской базы.
5. *Использование единых коммуникационных решений.* Применение данного подхода направлено на повышение эффективности работы контакт-центров, достигаемое на базе единых коммуникационных платформ. Тем самым исключаются дополнительные издержки и возможные проблемы, возникающие в процессе интеграции решений разных производителей.
6. *Обратный звонок.* Эта возможность обычно используется для разгрузки очереди ожидающих абонентов. Наличие подобной услуги повышает лояльность клиентской базы и служит источником дополнительных доходов контакт-центра.

7. *Переход к коммутационным решениям на базе SIP-протокола.* Это решение базируется на общей тенденции отказа от использования услуг традиционной телефонной связи и переходу к возможностям, предоставляемым протоколом SIP.
8. *Расширение перечня и роли статистических данных, иллюстрирующих качество обслуживания клиентов.* Контакт-центры выгодно отличаются от прочих систем связи активным использованием результатов статистических измерений в процессе проведения мероприятий по повышению эффективности своей работы. Этот подход получит еще большее развитие за счет расширения перечня измеряемых характеристик и применения более совершенных методов обработки статистических данных.
9. *Расширение возможностей IVR.* Использование IVR удешевляет информационное обслуживание клиентов. Их возможности будут значительно расширены за счёт применения систем распознавания и синтеза устной речи.
10. *Увеличение доли web-чата в организации доступа клиентов.* Эта возможность получила особо широкое распространение среди молодежи, которая чаще предпочитает эту форму общения телефонному звонку. Её использование позволяет повысить эффективность работы контакт-центра в условиях перегрузки.

1.1.6. Цели и задачи научных исследований

Основной задачей научных исследований в области планирования и повышения эффективности работы контакт-центров является поиск функциональной зависимости между моментами поступления запросов, показателями качества их обслуживания и характеристиками пропускной способности контакт-центра. К последним относятся число линий доступа, количество устройств IVR, число операторов и консультантов. При построении функциональной зависимости необходимо учесть свойства поступающих информационных запросов; их зависимость от модели поведения клиента, получившего отказ в обслуживании; особенности использования устройств IVR и мест ожидания начала обслуживания; зависимость времени обработки запроса от его вида и квалификации оператора. Выбор зависимости ограничен требованием о возможности её дальнейшего использования при оценке стоимости потенциальных сценариев развития справочно-информационных служб. Решение перечисленных задач выполняется с помощью исследования свойств случайного процесса, описывающего динамику изменения состояний исследуемой системы связи. Для его определения необходимо ввести основные структурные компоненты контакт-центра и проанализировать их взаимодействие. Решение этих задач будет рассмотрено в последующих разделах главы.

1.2. Основные структурные элементы и их взаимосвязь

1.2.1. Ресурс обслуживания заявок

Оптимизация затрат при сохранении требуемых значений показателей качества обслуживания поступающих заявок является основной задачей администрации контакт-центров. Основными компонентами ресурса, используемого при обслуживании поступающих заявок, являются линии доступа, устройства IVR и операторы. При этом затраты на людские ресурсы (операторы, консультанты) существенно превосходят затраты на линии доступа и устройства IVR. Так, по сведениям [74], стоимость обслуживания одного запроса устройством IVR/CTI составляет всего $\frac{1}{16}$ от стоимости обслуживания аналогичного запроса оператором. Перечисленные компоненты ресурса ограничены по величине и поиск сбалансированного соотношения между отдельными его видами требует отдельного исследования средствами теории телетрафика и экономическими методами. Начнем анализ свойств ресурса обслуживания заявок с его основной компоненты — операторов.

1.2.2. Операторы и консультанты

Состояние оператора принято отображать с помощью специальной диаграммы. Её вид показан на рисунке 1.2. Здесь выделены пять состояний. Первое из них показывает

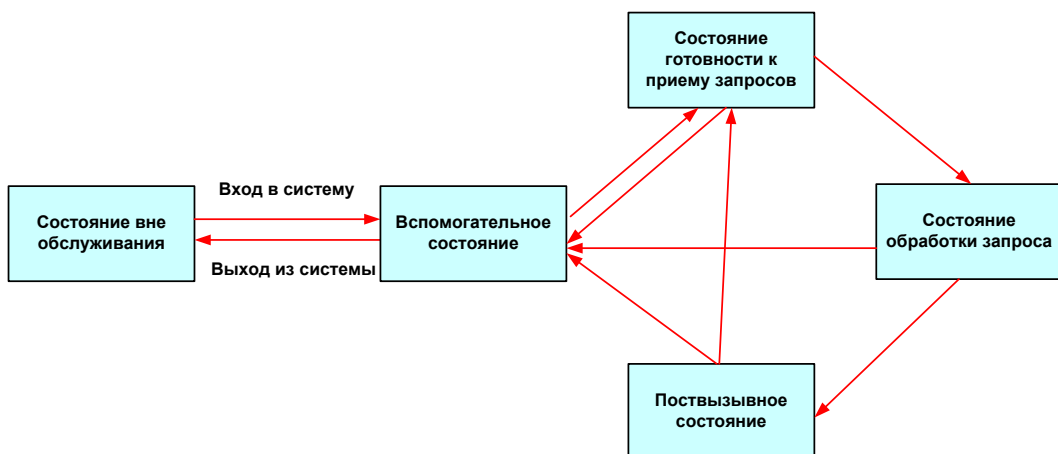


Рисунок 1.2. Состояния оператора и переходы между ними

пребывание оператора в *состоянии вне обслуживания*. Причины нахождения оператора в этом состоянии могут быть самыми разными, включая болезнь, отпуск и т.п. В начале рабочей смены оператор подключается к системе и переходит в *вспомогательное состояние*. Затем оператор переходит в *состояние готовности к приему запроса*. Если имеется запрос, ожидающий начала обслуживания, то он начинает выполняться и оператор переходит в *состояние обработки запроса*. После его завершения оператору требуется некоторое время на пребывание в *поствызовном состоянии*. За это время оператор может выполнить некоторые действия, связанные с окончанием обслуживания предыдущего запроса, например, внести информацию в базу данных. Далее оператор либо переходит в состояние готовности принять новый запрос, либо переходит во вспомогательное состояние, если того требуют обстоятельства: обеденный перерыв, рабочие встречи и т.п. Перечисленные состояния называются *первичными* в отличие от *вторичных*, пребывание в которых может указывать на дополнительную деятельность оператора, например, обработку запросов, поступающих по электронной почте.

В простейшем случае все поступающие заявки обрабатываются единой группой операторов с единой очередью ожидания начала обслуживания. Если поступающие запросы требуют дифференцированного обслуживания, то операторы могут быть представлены отдельными группами в соответствии с их специализацией. В этом случае они могут иметь отдельные очереди ожидания начала обслуживания. В рассматриваемой ситуации процесс обслуживания запросов в каждой группе происходит независимо. Чаще имеются группы операторов, которые способны обрабатывать запросы нескольких видов. Заявки, получившие отказ в группах операторов первого выбора, переходят на обслуживание в группы операторов второго выбора, обслуживающие избыточные потоки нескольких первичных групп и т.д. Группы 2-го и последующего выбора формируются из универсальных (опытных) операторов, прошедших дополнительную кроссфункциональную подготовку. Отмеченная особенность обслуживания информационных запросов показана на рисунке 1.3.

В более совершенных формах организации очередей и выбора ожидающих запросов не используется привязка операторов к конкретной очереди. Вместо этого для каждого оператора (или группы операторов) формируется перечень навыков (англ. *skills*), которые и служат основанием для направления к нему поступившего запроса. Наборы навыков готовятся на основе специализации контакт-центра и анализа опыта его работы. Маршрутизация запроса осуществляется в результате контакта клиента с устройством IVR. Таким образом поступившая заявка обрабатывается наиболее подготовленным для этих целей свободным оператором. Рассматриваемая модель формирования групп операторов показана на рисунке 1.4. Правильно организованная работа операторов на основе маршрутизации по навыкам повышает эффективность и качество обработки поступающих информационных запросов. Однако получить достаточно обоснованное решение сформулированной проблемы не так просто. Более того, возможные ошибки в ее решении

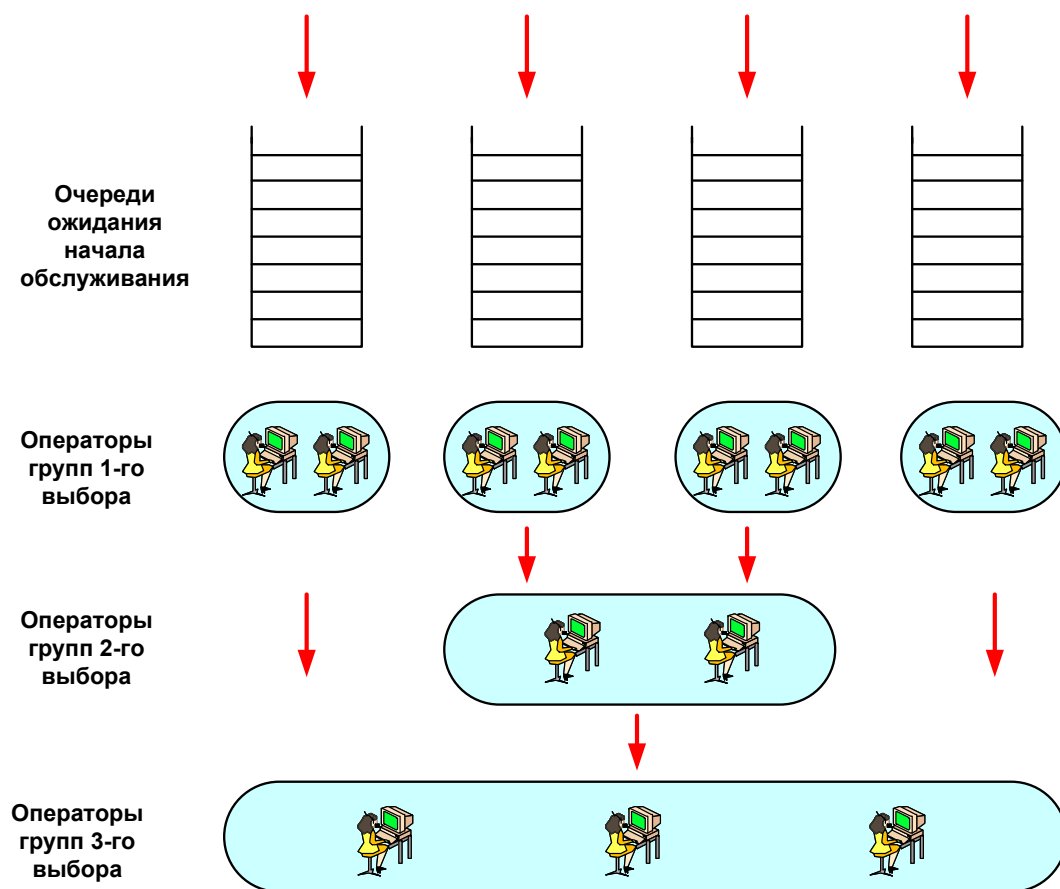


Рисунок 1.3. Обслуживание избыточных запросов группами универсальных операторов

могут существенно повлиять на качество обслуживания заявок [85].

В диссертационном исследовании рассмотрена другая форма учета специализации информационных запросов. Будем предполагать, что процесс обслуживания заявки состоит из трёх этапов. Первый — обработка запроса устройством IVR. По его завершении клиент может закончить обслуживание в контакт-центре или продолжить его, но уже у оператора. Задача оператора вступить в диалог с клиентом, дополнить недостающие сведения в ответ, полученный ранее от IVR, и в случае необходимости направить запрос клиента к консультанту для получения специализированной информации. Если такой необходимости не возникает, то пользователь услуг контакт-центра завершает обслуживание на уровне оператора и покидает анализируемую систему связи. Рассматриваемая модель обслуживания поступившего информационного запроса представлена на рисунке 1.5. Мотивация к использованию именно этой модели дифференциации операторов следует практике работы контакт-центров. Обучение и содержание большого штата специалистов компетентных во всех аспектах деятельности компании (продуктов, услуг и

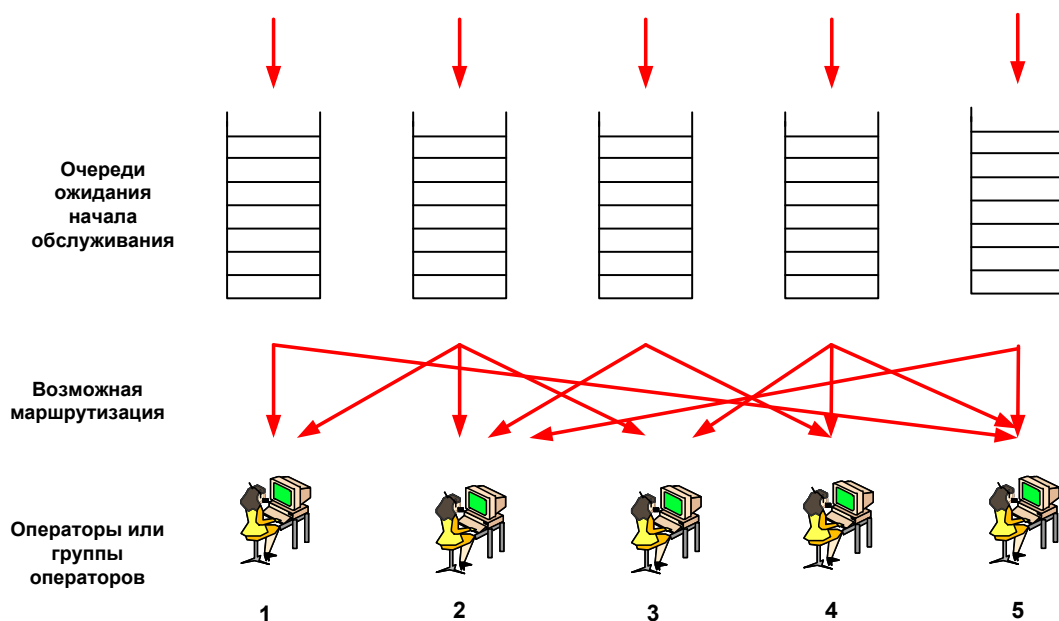


Рисунок 1.4. Обслуживание запросов на основе маршрутизации по навыкам операторов

т.д.) нецелесообразно с финансовой точки зрения. По этой причине, где это необходимо, операторы делятся на обычных и специализированных, обычно, в несколько вопросов. Таких сотрудников называют *консультантами, специалистами техподдержки, операторами второй линии*. В качестве примера можно привести контакт-центр магазина бытовой техники, где консультанты распределяются на группы, согласно перечню товаров, по которым они имеют экспертизу, например, группа по ТВ-оборудованию, видео/аудио технике, компьютерам или холодильникам. Если компания имеет много клиентов за границей, консультантов разделяют на группы по знанию иностранного языка и т.д. Если подобные группы в рассматриваемом контакт-центре не представлены, то в обобщенной модели их численность можно просто положить равной нулю. Это упростит анализ модели.

1.2.3. Устройства интерактивного речевого взаимодействия (IVR)

Устройство IVR можно рассматривать как роботизированного оператора контакт-центра. Клиент общается с IVR, используя кнопочный тональный набор или произнося необходимые словосочетания. Последний отвечает с помощью заранее записанных фраз или использует возможности синтезированного голоса. Современные IVR могут полу-

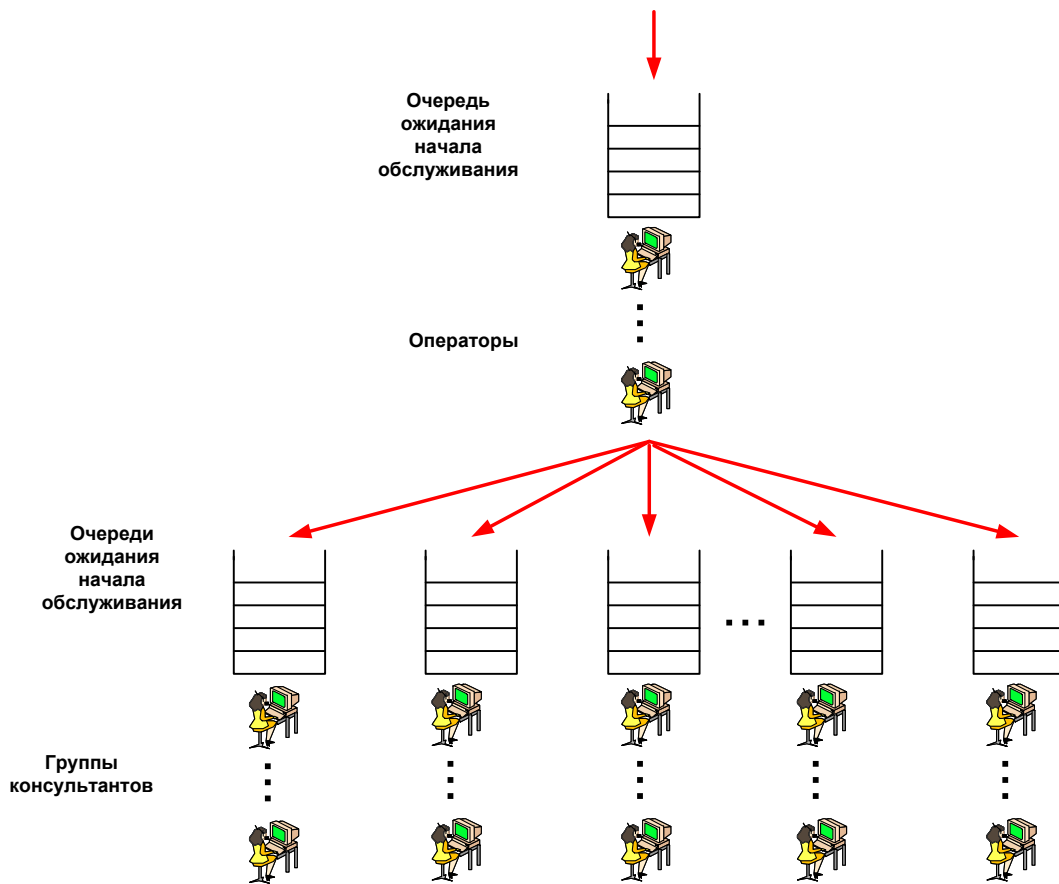


Рисунок 1.5. Обслуживание запросов на основе трёхуровневой модели: операторы и консультанты

чать затребованную информацию из баз данных, пересылать её на дисплей клиента, а также выполнять необходимую маршрутизацию вызова. Перечислим услуги, которые могут быть оказаны с помощью устройств IVR. Обычно к ним относятся следующие сервисы: предоставление разного рода справок, например, о состоянии счёта; обработка транзакций; выполнение действий (заполнение форм), относящихся к заказу чего-либо и т.п. Совокупность устройств IVR либо представляет из себя отдельный модуль, либо является частью коммутационной системы. Технические аспекты взаимодействия коммутатора и IVR зависят от производителя оборудования и здесь не рассматриваются. Обсудим влияние, которое оказывают подобные устройства на эффективность работы контакт-центра.

Использование элементов *самообслуживания* при обработке информационных запросов клиентов может существенно понизить затраты контакт-центра. Подобная возможность активно используется ими для обслуживания абонентов в неурочное время, например, в ночные часы, праздничные дни и т.п. Понятно, что автоматизированные системы не могут обеспечить 100 % ответов на поступившие запросы, но по данным [74],

60 % из них получают требуемую информацию и не обращаются за повторным обслуживанием. Достигнутая экономия на сокращении числа операторов позволяет окупить систему IVR уже за первый год её эксплуатации. Отношение клиентов к контактам с устройством IVR двойко. Большая часть из них, особенно молодежь, не видит в этом ничего отрицательного и часто предпочитают общение с IVR непосредственному контакту с оператором. Другие, наоборот, предпочитают получить обслуживание у оператора. Необходимо отметить, что в большинстве случаев источником проблем является плохая организация самообслуживания клиента, а не его непереносимость контактов с IVR.

В [90] перечислены основные причины, препятствующие позитивному восприятию подобных систем. Среди них:

- Избыточное число возможных сценариев развития событий, обеспечиваемое использованием всех двенадцати кнопок, представленных на телефонном аппарате.
- Большое количество уровней погружения в процессе выбора требуемого сервиса.
- Замкнутые циклы, которые случаются в ситуации, когда система IVR не обеспечена выходом на оператора. В результате клиент снова и снова возвращается в исходное состояние и повторяет цикл выбора услуги.
- Отказ в обслуживании в случае, когда абонент не успел осуществить выбор сервиса.
- Избыточное использование технического жаргона в предложениях IVR, затрудняющее понимание озвучиваемых опций.
- Частое изменение структуры меню и возможные ошибки в предлагаемых опциях.
- непрофессиональное изготовление звукового и письменного сопровождения в предложениях IVR.

Тем не менее, система IVR, лишенная перечисленных выше недостатков, помогает администрации контакт-центра существенно повысить эффективность его работы. Перечислим основные моменты, лежащие в основе сформулированного утверждения [74]. Во-первых, повышается качество обслуживания клиентов. Этот результат достигается за счёт уменьшения числа запросов, обрабатываемых операторами. Существенная часть из них теперь обслуживается устройствами IVR. Абонент также позитивно оценивает возможность круглосуточного обслуживания и конфиденциальность обработки запросов, предлагаемую техническими системами. Последнее часто бывает желательным в запросах, касающихся состояния счета или содержания медицинских справок. Во-вторых, применение устройств IVR увеличивает доходность контакт-центра, поскольку, как уже было сказано, увеличивает часы работы справочной службы, привлекая потенциальных клиентов из других временных зон. В третьих, уменьшаются затраты контакт-центра на обеспечение информационного обслуживания. Этот фактор эксперты считают наиболее

существенным для принятия решения о внедрении систем речевого взаимодействия. Известно, что затраты на операторов составляют от 60 % до 70 % затрат контакт-центра на осуществление операционной деятельности [104]. Обслуживание телефонного запроса обходится компании в среднем в 7 \$. Тот же вызов, совершённый с использованием возможностей сети Интернет, обходится в 2,25 \$. Если запрос обслужен системой IVR, то он стоит уже 0,5 \$ [88]. Эффект от внедрения IVR, относящийся к уменьшению числа операторов, можно оценить численно с использованием модели, предусматривающей подобное обслуживание. Этот результат будет рассмотрен в главе 4.

1.3. Сценарии обработки запросов

Рассмотрим процесс обработки информационных запросов, поступающих в контакт-центр. Общая схема функционирования контакт-центра показана на рисунке 1.6.

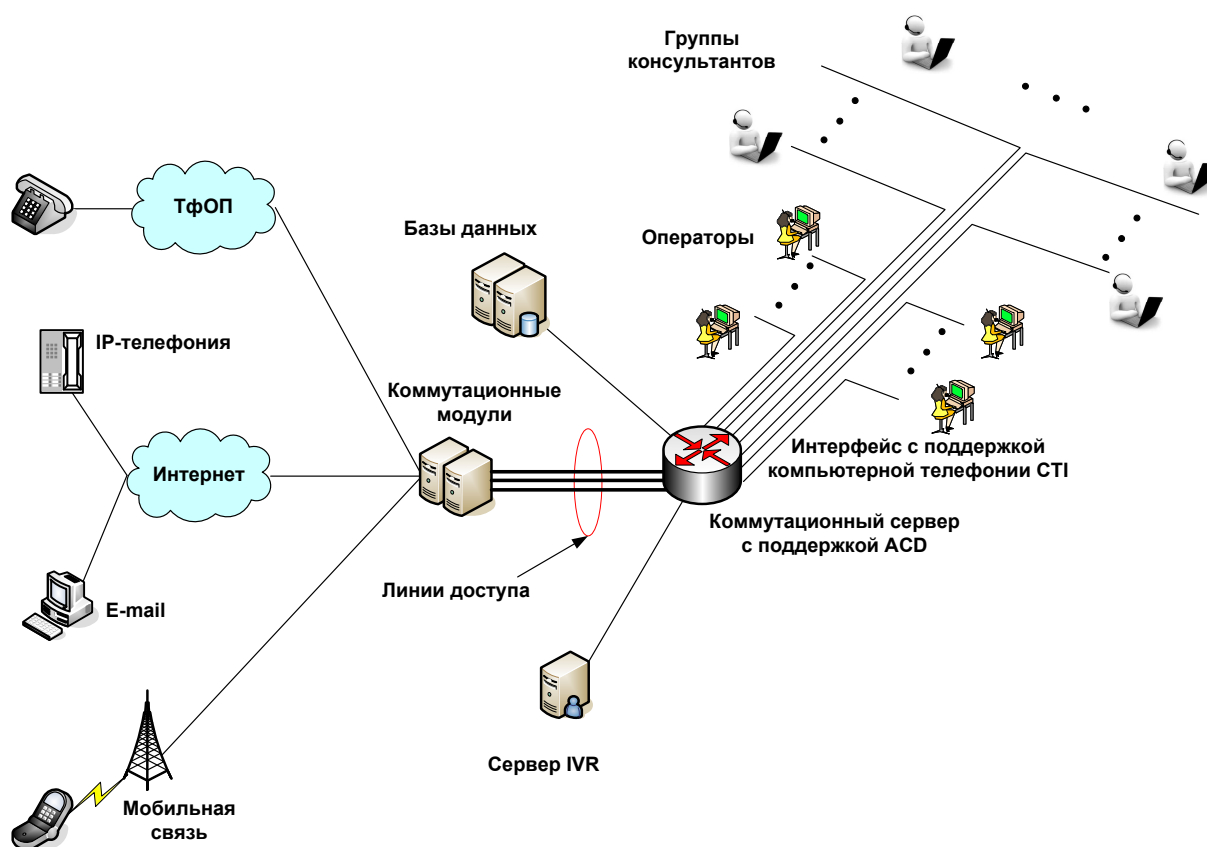


Рисунок 1.6. Взаимодействие основных функциональных компонентов современных контакт-центров

Перечислим основные этапы установления соединения клиента с контакт-центром посредством обычной телефонной связи и затем рассмотрим особенности, вносимые применением новых каналов передачи информационных сообщений. Клиент набирает номер контакт-центра и из своей сети (фиксированной или мобильной связи) с помощью соответствующего коммутационного модуля попадает на коммутационный сервер справочного узла, используя свободные линии доступа. На первом этапе установления соединения с контакт-центром пользователь вступает во взаимодействие с устройством IVR. Он реагирует на предложенные возможности, анализирует полученную информацию и довольно часто достигает своей цели уже на этом этапе обслуживания в справочной службе. Современные системы IVR достаточно развиты для удовлетворения рутинных запросов клиентов.

В большинстве случаев коммутационный сервер представляет из себя учрежденческую АТС с поддержкой функций автоматического распределения вызовов (англ. *automatic call distribution, ACD*). Их реализация базируется на инструкциях по маршрутизации поступивших запросов к операторам контакт-центра. Выбор инструкций в основном определяется в результате взаимодействия клиента с системой IVR. Также может использоваться дополнительная информация, затребованная из баз данных по идентификационному номеру клиента. Действующие аппаратно-программные комплексы ACD могут обрабатывать поступающие запросы на основе самых разнообразных критериев. Обычно они интегрированы в состав коммутационного сервера или представляют из себя отдельный модуль.

Оператор может косвенно участвовать в маршрутизации поступившей заявки. Происходит это следующим образом. После перехода оператора в состояние готовности к началу работы система ACD фиксирует изменившийся статус оператора, выясняет из баз данных его квалификационный уровень и направляет к нему подходящие запросы. Если все операторы, подходящие для обработки поступившего запроса заняты, то система ACD удерживает клиента на ожидании. В течение этого промежутка времени клиент прослушивает рекламу, дополнительную информацию о возможностях контакт-центра, музыку и т.п. Ожидающие абоненты могут не дожидаться начала обслуживания и покинуть систему. Часть абонентов повторит запрос, оставшиеся откажутся от него или отложат на некоторое время. В некоторых ситуациях ограничение на возможное время нахождения в очереди вводится администрацией контакт-центра в целях предотвращения его перегрузки.

После установления соединения с подходящим оператором начинается выполнение запроса клиента. Реализация этой функции основана на возможностях, обеспечиваемых технологией компьютерной телефонии (англ. *Computer Telephony Integration, CTI*), позволяющей обрабатывать поступающий звуковой сигнал (подавлять эхо, переводить звуковой сигнал в цифровой и обратно), осуществлять коммутацию запроса с использованием голосового меню или специальных таблиц маршрутизации. В частности, с использо-

ванием возможностей СТИ анализируется телефонный номер клиента, запрашивающего справочную информацию. По результатам анализа выполняется маршрутизация вызова и выясняется предыстория взаимодействия клиента с контакт-центром. Полученная информация высвечивается на мониторе оператора и используется в процессе выполнения поступившего запроса. Перечисленные возможности, обеспечиваемые технологией СТИ, уменьшают длительность промежутка времени обработки запроса, поскольку отпадает необходимость выполнения части стандартных функций, относящихся к идентификации клиента, например, выяснения номера его телефона, номера счета и т.п. В современных справочно-информационных службах возможности СТИ интегрированы в систему управления взаимодействием с клиентами (англ. Customer Relationship Management, CRM), позволяющую кастомизировать процесс их обслуживания устройством IVR или оператором на основе прошлых предпочтений.

Завершение работы с клиентом может развиваться по нескольким сценариям. Частично они показаны на рисунке 1.2. Оператору часто бывает необходимо внести в базу данных результаты обработки поступившего запроса или совершить действия, направленные на выполнение заказа клиента. Это состояние оператора названо на рисунке 1.2 поствызывным (англ. wrap-up). Выполнение указанной деятельности увеличивает длительность промежутка времени обслуживания вызова и должно быть учтено при построении математической модели работы контакт-центра. Приведенное описание определяет процесс функционирования типичной справочно-информационной службы. При построении модели конкретного контакт-центра необходимо учитывать особенности, вносимые используемым оборудованием и программным обеспечением, видом предоставляемой справочной информации, а также поведением пользователя услуг центра в процессе обслуживания. Как уже было сказано, в процессе эволюции справочных служб расширяется перечень каналов доступа к получению справочной информации. К линиям телефонной голосовой связи добавляются каналы доступа, организуемые посредством возможностей сети Интернет: электронная почта, интернет-телефония, web-чат и т.д. Для обработки подобных запросов обычно выделяются специальные группы операторов. В модели они могут быть учтены в виде отдельных групп консультантов. Они могут специализироваться только на выполнении перечисленных выше видов запросов или также обслуживать заявки, поступающие по телефонным каналам. В качестве примера приведем функциональную схему работу контакт-центра банка ВТБ. По данным компании Крок³, структурная схема системы имеет вид, представленный на рисунке 1.7.

Контакт-центр решает задачи обеспечения клиентов банка информационной поддержкой по вопросам кредитной политики, продажи розничных продуктов, обработки заявок, поступающих по электронной почте, а также проводит маркетинговые исследования путем обзвона клиентов. В банке работают 633 оператора для приема и обработки входящих информационных запросов; 100 операторов для организации опросов действи-

³www.croc.ru

тельных и потенциальных клиентов; имеются 682 устройства интерактивного голосового взаимодействия и 76 рабочих мест супервайзеров.

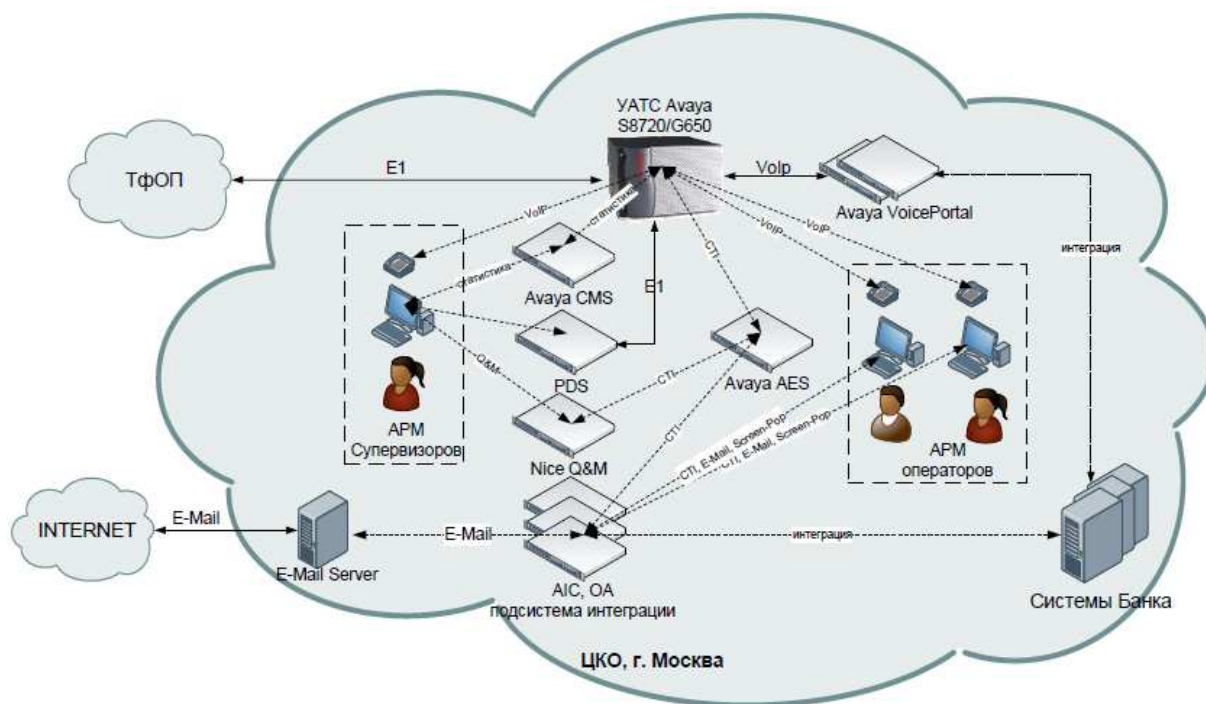


Рисунок 1.7. Структурная схема контакт-центра банка ВТБ

1.4. Показатели поступления и обслуживания заявок

1.4.1. Типы и классификация характеристик

Процесс управления работой контакт-центра и планирования количества необходимых элементов инфраструктуры опирается на значения показателей качества обслуживания поступающих информационных запросов. Большая часть из них измеряется в автоматизированном режиме и используется администрацией справочно-информационных узлов для принятия управляющих решений. Согласно [85] характеристики качества работы контакт-центра подразделяются на четыре категории. Первая и основная часть из них представлена показателями *операционной деятельности*, отражающими процесс обслуживания информационных запросов клиентов. Численные значения этих характеристик в основном собираются системами IVR и ACD. Они подразделяются на *операционные характеристики клиента* и *операционные характеристики оператора*. К первой группе относятся моменты времени появления запросов; начала и окончания их обслуживания; ухода из системы; идентификаторы операторов, к которым попали запросы и т.д.

Во вторую группу включены величины длительностей пребывания оператора в различных состояниях. Основные из них отображены на рисунке 1.2. Обычно они усредняются за интервалы длительности 15, 30 и 60 минут при принятии тактических решений и за промежутки времени длительностью одни сутки, одну неделю и один год при стратегическом анализе работы контакт-центра. Остальные три категории имеют вспомогательный характер и представлены показателями сервисного профиля клиента, производственного профиля оператора и качественными психологическими оценками восприятия предлагаемых информационных услуг.

Полный перечень операционных характеристик весьма обширен и может отличаться для контакт-центров разной специализации. Перечислим основные показатели, сформулируем их названия на английском языке и, руководствуясь отчетом [95], приведем для некоторых из них рекомендуемые численные значения, взятые из практики планирования и работы контакт-центров. Анализируемый перечень включает группу характеристик поступления и обслуживания заявок и группу характеристик работы операторов. В первое семейство входят следующие показатели [85]:

- *Число запросов, поступивших на каждый из этапов их обслуживания в контакт-центре, англ. Contact Volume.* Данная характеристика имеет большое значение при планировании работы справочного узла поскольку многие другие показатели являются производными от неё. Обычно её величина зависит от рассматриваемого этапа и получается путем усреднения по подходящим временным интервалам. Численное значение характеристики также зависит от специфики контакт-центра, его размеров и т.п. Результаты отдельных измерений представлены на рисунке 1.8.
- *Время обслуживания, англ. Call Duration.* Величина этой характеристики представляет собой среднее значение промежутка времени непосредственного общения оператора и клиента. Она зависит от типа запроса, квалификации оператора, способности клиента ясно изложить суть запроса и т.п. Оператору рекомендуется увеличить продолжительность общения с клиентом нежели спровоцировать его на повторный вызов. Среднее значение этого показателя составляет 4 мин.
- *Время поствызовной активности, англ. Call Wrap up Time.* Величина этой характеристики складывается из средних значений промежутков времени, относящихся к выполнению работы, связанной с завершением обслуживания информационного запроса. В этот перечень входят: заполнение форм, внесение изменений в базу данных, осуществление деятельности, связанной с выполнением заказа, сформулированного клиентом, и т.п. Величина характеристики определяется особенностями справочно-информационной службы и в среднем составляет 6 мин.
- *Длительность интервала времени обработки запроса, англ. Contact Handle Time.* Её значение складывается из величин *времени обслуживания* и *времени поствы-*

звой активности, введенных выше. Среднее значение характеристики составляет 10 мин.

- *Уровень обслуживания, англ. Service Level*, определяется через процент заявок, попавших на обслуживание к оператору за заданный промежуток времени. Чаще всего используется следующее соотношение: 80 % всех поступивших заявок попадут на обслуживание по истечении 20 с. Одна из наиболее часто используемых характеристик качества работы контакт-центра.
- *Среднее время ожидания ответа оператора, англ. Average Speed to Answer*. В соответствии с определением величина этой характеристики рассчитывается как среднее значение промежутка времени ожидания начала обслуживания. Для заявок, не получивших ответа в первые 20 с, значение этой характеристики составляет 28 с. Превышение этого значения является аргументом к изысканию возможностей увеличения числа операторов.
- *Доля запросов, обслуженных с первой попытки, англ. First Call Resolution Rate*. В соответствии с определением, значение данной характеристики показывает долю запросов, обслуженных в течение первой попытки соединения и именно тем оператором, к которому был адресован запрос. Таким образом, при оценке значения этого показателя не учитываются запросы переадресованные другим операторам и не рассматриваются ситуации, когда клиенту рекомендовано ждать обратного звонка от оператора. Величина характеристики лежит в пределах 70-75 %⁴.
- *Доля запросов, отказавшихся от попыток получить обслуживание, англ. Abandonment Rate*. Отказ в обслуживании может возникнуть из-за недостатка элементов ресурса (устройств IVR, операторов и консультантов), а также из-за ограничения на длительность максимально возможного времени пребывания на ожидании начала обслуживания. Это ограничение отражает нетерпеливость клиента или является результатом реализации технической политики администрации, направленной на борьбу с перегрузкой. Численное значение характеристики лежит в пределах 5–8 %. Следует отметить, что величина этой характеристики существенно зависит от особенностей поведения клиента, находящегося на ожидании.
- *Стоимость предоставления информационной услуги, англ. Cost Metrics*. Она может быть выражена в стоимости обработки одного запроса или в стоимости об-

⁴Эта весьма важная характеристика показывает потенциальное наличие повторных вызовов в общем потоке поступающих заявок. Так по данным Professional Association for Customer Engagement, опубликованным в 2012 г. [84], результаты измерения данного показателя привели к значению 70 %. Дополнительное исследование результатов статистических измерений показало, что оставшиеся 30 % запросов, не попавшие на обслуживание с первой попытки, явились причиной 55 % всех запросов, поступивших в контакт-центр, включая первичные и повторные.

служивания в единицу времени. Конкретный вид функционалов, определяющих величины этих характеристик, зависит от особенностей анализируемых контакт-центров.

- *Степень удовлетворения клиента, англ. Customer Satisfaction.* Выбор методик оценки данного показателя может быть весьма разнообразным. Величины этих характеристик определяются путем организации опросов клиентов, в частности, с использованием устройств IVR или с привлечением специализированных организаций. Они имеют большое значение для проведения мероприятий, направленных на повышение эффективности работы контакт-центров. Рекомендуемое значение характеристики составляет 90 %.

Во второе семейство включены такие показатели:

- *Степень выполнения планируемого расписания работы, англ. Adherence to Schedule.* Этот важный показатель деятельности работы оператора оценивается через отношение величины фактического времени работы оператора к планируемой величине. Нарушения графика связаны с заболеваниями, опозданиями, увеличением длительности отдыха и т.п. Рекомендуемое значение характеристики составляет 95 %. Особенно трудно выдерживать эту рекомендацию для контакт-центров, имеющих малое число операторов.
- *Занятость оператора, англ. Occupancy.* Оценивается значением доли времени пребывания оператора в состоянии ответа и в состоянии поствызывной обработки вызова. Рекомендуемая величина анализируемого показателя лежит в пределах 60–80 %.
- *Текущность кадров, англ. Turnover Rate.* Значение характеристики определяется через отношение числа операторов, оставивших работу в течение года, к общему числу задействованных операторов. Рекомендуемое значение характеристики составляет 15 %.

Перечисленные показатели характеризуют операционную деятельность контакт-центра и могут использоваться для обоснования мероприятий, направленных на повышения эффективности его работы. Большую часть из них после надлежащей формализации процессов поступления и обслуживания заявок можно рассчитать с использованием обобщенной модели контакт-центра. Определение характеристик и расчетные формулы будут приведены в разделе 2.3. Рассмотренные характеристики относятся к обслуживанию запросов, поступающих по обычным телефонным каналам. Можно привести аналогичный список показателей для запросов, поступающих по каналам, организованным на базе возможностей, обеспечиваемых технологиями сети Интернет.

1.4.2. Операционные режимы

Основной задачей администрации справочно-информационного узла является выполнение действий, направленных на обслуживание поступающих информационных запросов с надлежащим качеством. Для её решения используется ресурс контакт-центра в виде устройств IVR, операторов и консультантов. Операционные расходы на приобретение и поддержание необходимого количества ресурса составляют существенную часть расходного бюджета компании и поэтому должны быть надлежащим образом сбалансированы. Есть несколько подходов к решению этой проблемы. Обсудим основные результаты. *Операционная стратегия контакт-центра может быть направлена на максимально возможное использование имеющегося ресурса (англ. Efficiency Driven (ED) Regime)*. В рассматриваемой ситуации желание добиться максимальной загрузки ресурса превалирует над необходимостью поддержки требуемого качества обслуживания поступающих запросов. В анализируемом случае система спланирована с недостатком пропускной способности, почти все поступающие запросы должны ожидать начала обслуживания, степень загрузки ресурса близка к ста процентам. Используется и прямо противоположный подход. *Операционная стратегия контакт-центра может быть направлена на максимально достижимое качество обслуживания поступающих запросов (англ. Quality Driven (QD) Regime)*. В этой ситуации справочный узел спланирован с избытком пропускной способности, поступающие заявки практически не становятся в очередь, ресурс загружен слабо. Чаще всего применяется смешанный, так называемый *рациональный подход. Операционная стратегия контакт-центра направляется на достижение баланса между загрузкой ресурса и качеством обслуживания поступающих запросов (англ. Quality and Efficiency Driven (QED) Regime)*. В рассматриваемом случае тщательно оценивается соотношение между величиной поступающих запросов и количеством ресурса. Для этого используются модели и методы теории телетрафика. Поступающие заявки могут ожидать начала обслуживания, но доля таких заявок примерно равна 10 %. Загрузка ресурса приближается к 100 %. Это подход представляется оптимальным в большинстве ситуаций, если других решений не требует специфика работы анализируемого справочно-информационного центра⁵.

1.4.3. Измерение характеристик контакт-центра

Статистическая оценка и прогнозирование значений операционных характеристик контакт-центра являются важной составной частью его планирования и эффективного управления. Несмотря на отмеченную значимость проблематики и большое количество результатов статистических наблюдений, в литературе по теории телетрафика имеется

⁵Подобной спецификой обладают, например, справочные службы организаций, решающих задачи обеспечения условий безопасности граждан.

не так много работ, содержащих математический анализ полученных данных. Отмеченная ситуация стала меняться только в последнее время [100]. Для исследования данных используются стандартные подходы, применяемые в подобных ситуациях. В основном — статистические методы анализа временных рядов. Конечная цель исследования — построение математической модели и последующее её применение для прогноза будущих значений временного ряда. Модели могут быть получены *методом скользящего среднего* (англ. Moving-Average model — MA) или его обобщением, каковым является *интегрированная модель авторегрессии скользящего среднего* (англ. Autoregressive Integrated Moving-Average model — ARIMA). Могут использоваться и другие методы, учитывающие те или иные особенности анализируемого объекта. Не вдаваясь в подробности реализации этих процедур, известные из руководств по обработке данных статистических наблюдений [26], отметим, что лучшие результаты даёт применение более простых алгоритмов. Этот вывод обосновывается легкостью их реализации, а также разного рода интуитивными соображениями, обычно лежащими в основе их построения.

Рассмотрим вкратце применение одного из таких подходов, взяв в качестве примера анализ частоты поступления информационных запросов. Решение задачи основано на использовании алгоритма, который носит название *принцип сверху-вниз* (англ. top-down approach). Реализация данного подхода включает в себя суммирование значений частот появления запросов, измеренных на кратких интервалах времени (обычно это 30 мин), за месячный период наблюдения. Затем полученный временной ряд анализируется одним из методов, отмеченных выше, для оценки числа поступлений заявок за прогнозируемый месяц. Далее эта величина делится на отдельные недели, дни и часы. При этом могут использоваться значения пропорций между отдельными компонентами, если они выявлены по результатам измерений. Рассмотрим в качестве примера результаты анализа поступления запросов в небольшой контакт-центр банка [99]. В справочной службе работают 8 операторов, обслуживающих вызовы, поступающие по телефонным каналам и 5 операторов, обслуживающих запросы, поступившие по электронной почте. Величины обработанных запросов оценивались за день, неделю, месяц и год и представлены на рисунке 1.8 [99].

Результаты наблюдений показывают устойчивые изменения значений характеристик, которые проявляют себя на суточных, а также возможно и на недельных, месячных, сезонных и годовых интервалах времени. Отмеченные тренды выявляются в процессе исследования полученных временных рядов и служат основой для построения достаточно точных прогностических оценок значений характеристик. Обычно измеряются величины всех или части показателей, перечисленных в разделе 1.4.1. Отмечено [91], что значения частоты поступления запросов имеют наибольший разброс, по сравнению со значениями других характеристик.

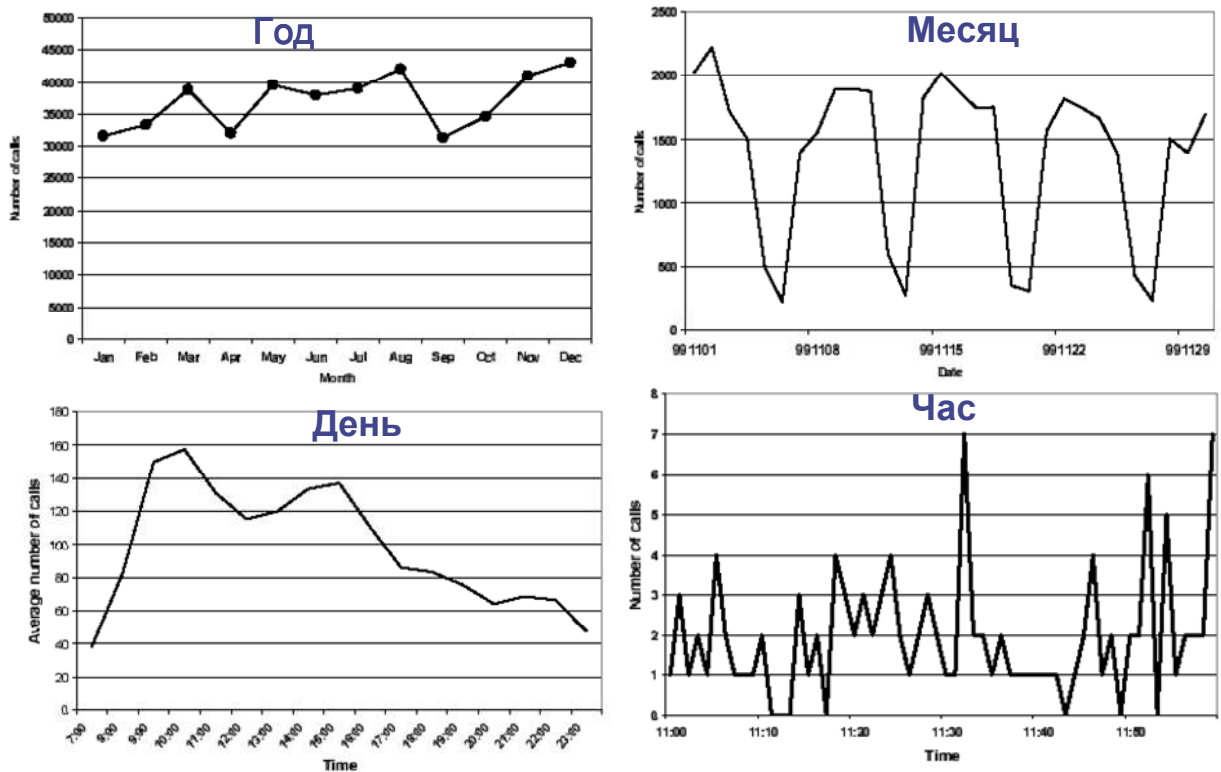


Рисунок 1.8. Результаты измерения числа информационных запросов в контакт-центре банка [99]

1.5. Вопросы планирования работы контакт-центров

1.5.1. Методология планирования

Планирование, построение и эффективная эксплуатация контакт-центра являются достаточно сложными техническими задачами. Трудности возникают из-за необходимости организации согласованной работы его технических компонент (линий связи, устройств коммутации, вычислительных комплексов), людских ресурсов (операторов, консультантов) и программных решений (управляющих систем, синтезаторов речи, баз данных). Появлению возможных проблем должна препятствовать правильно выбранная методология построения и эксплуатации контакт-центра. Её наличие особенно важно при создании нового контакт-центра, а также при существенном расширении функций существующих справочных центров. Этими вопросами обычно занимаются крупные консалтинговые фирмы и интеграторы. Ключевые моменты большинства подходов совпадают, поэтому достаточно рассмотреть один из них. Дадим краткую характеристику решения, предложенного известной консалтинговой компанией Prosci [97]. Она опреде-

ляет следующие пять компонент, взаимодействие которых является существенным для функционирования контакт-центра. Это: бизнес-стратегия; операционная деятельность; необходимые технологии; людские ресурсы и инфраструктура. Начальным этапом является разработка бизнес-стратегии. Она служит основой для последующего описания операционной деятельности и выбора поддерживающих её технологических решений. Далее следует подбор и управление людскими ресурсами и создание необходимой инфраструктуры. Каждая из отмеченных компонент детализируется, и все они в совокупности определяют так называемую *дорожную карту* (англ. *roadmap*) планирования контакт-центра. В качестве примера на рисунке 1.9 показан общий вид дорожной карты, составленной компанией Prosci [97].

Каждая из позиций, представленных на рисунке может быть подробно расписана. Рассмотрим, например, пункт «Цели и контроль» из блока «Стратегия». Качественная работа контакт-центра предполагает достижение следующих целей: операционную эффективность, оцениваемую в стоимости предоставления услуг; операционную продуктивность, выражающуюся в величине заработанных денежных средств; высокие уровни обслуживания клиентов и условия работы сотрудников. Факт достижения заявленных целей необходимо формализовать и выразить в численном виде. Для этого используются величины показателей качества обслуживания поступающих запросов, перечисленных в разделе 1.4. Так уровень обслуживания определяется степенью удовлетворения клиента, величиной доли запросов, обслуженных с первой попытки и другими формализованными характеристиками качества работы с клиентом. Важнейшей частью планирования контакт-центра является определение численности персонала. Она будет рассмотрена в следующем разделе.

1.5.2. Оценка числа операторов

Общие положения. Управление людскими ресурсами контакт-центра — это одновременно наука и искусство, совместное действие которых заключается в оценке необходимого числа операторов, размещении их в заданное время для обслуживания поступающего потока информационных запросов с требуемым качеством и за минимальную стоимость [81]. Это вполне очевидное утверждение определяет основную задачу администрации справочного узла. Уже отмечалось, что операционные расходы контакт-центра, отнесенные к персоналу, составляют около 70 % его общих затрат [85]. По этой причине оценка численности персонала должна производиться с особой тщательностью и с учетом всех значимых факторов.

Рассмотрим отдельные составляющие этой проблемы. Естественным образом она разделяется на две части [72]. Вначале ставится задача наёма или аренды необходимого числа операторов, затем определяются сценарии их занятости при обслуживании поступающих запросов. При этом наём персонала осуществляется на основании длительных

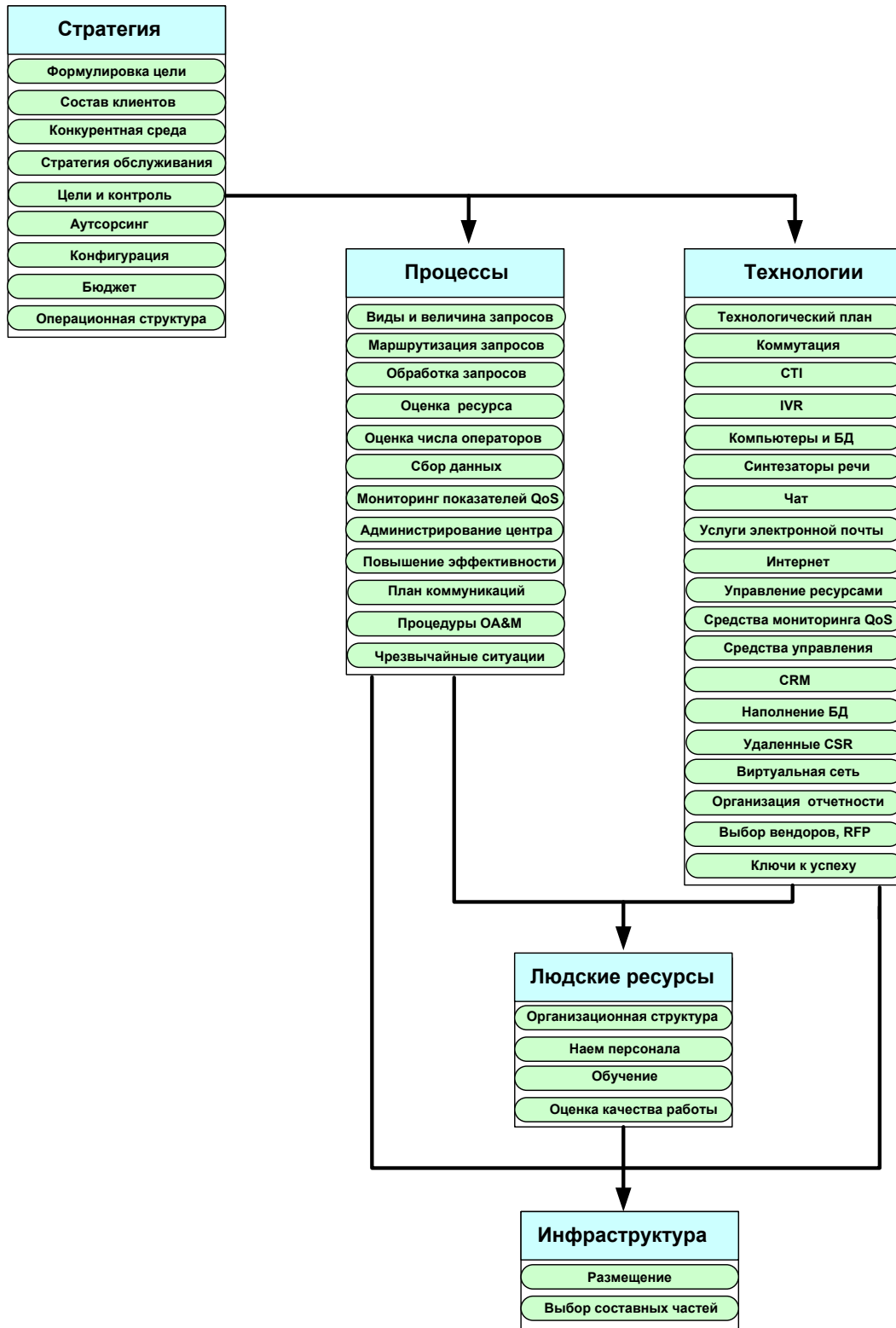


Рисунок 1.9. Дорожная карта планировании контакт-центра (по материалам компании Prosci [77])

прогнозов, как правило, исчисляемых на несколько месяцев, а расписание их работы составляются на недельной основе. Решение перечисленных задач выполняется в определенной последовательности, показанной на рисунке 1.10 [97]. Дадим характеристику основным её этапам.

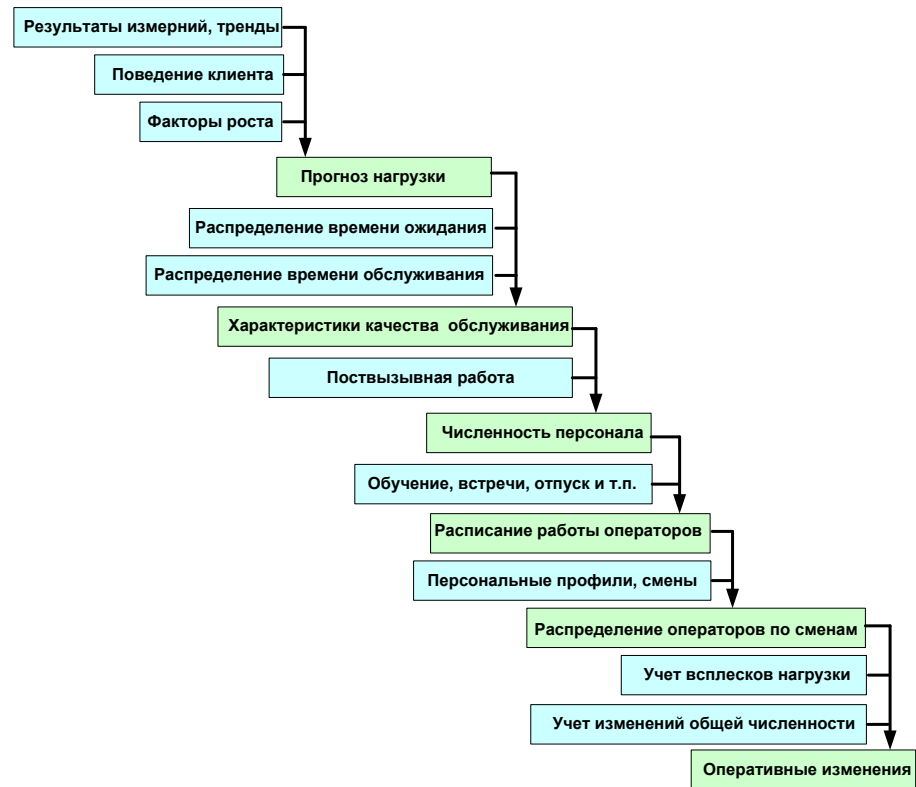


Рисунок 1.10. Основные факторы, влияющие на оценку численности персонала и составление расписания их работы [97]

Прогноз нагрузки. Данный этап является одним из главных, поскольку его результаты используются при проведении последующих вычислений. Основой расчетов являются данные измерения числа поступающих запросов. При обработке статистических результатов выявляются тренды, а также другие факторы, влияющие на значения показателей, например, особенности поведения клиента. Вид анализируемых характеристик показан на рисунке 1.8. Информация, пригодная для исследования, должна включать в себя результаты измерений как минимум за один год. Далее она рассматривается на интервалах времени длительности месяц, неделя, рабочий день, час, 30 мин и 15 мин. Выявляются и принимаются во внимание разного рода отклонения, связанные с праздниками и другими изменениями в поведении клиентов. После выполняется прогноз нагрузки. Для этого используются методы анализа временных рядов, отмеченные в разделе 1.4.3. Обычной в практике прогноза поступающей нагрузки является предпо-

ложение о равенстве константе интенсивности появления информационных запросов на интервалах времени 30 или 60 мин.

Характеристики качества обслуживания. Обычно клиент оценивает уровень обслуживания доступностью сервиса и временем, которое он тратит на получение запрошенной информации. Численная оценка перечисленных аспектов качества обслуживания выполняется с помощью характеристик, приведённых в разделе 1.4. К ним относятся: доля потерянных заявок; доля запросов, обслуженных с первой попытки; среднее время ожидания ответа оператора и ряд других показателей. При их вычислении требуется учитывать специфику анализируемой справочной службы, например, включать или нет в промежуток времени ожидания длительность обслуживания в IVR, аналогично, учитывать или нет во времени обслуживания длительность промежутка времени поствызовной обработки вызова.

Численность персонала. Необходимое число операторов в каждый из анализируемых интервалов времени определяется исходя из требования обслужить поступающие информационные запросы с заданным качеством. Для решения этой задачи необходимо воспользоваться математической моделью, которая устанавливает функциональное соотношение между нагрузкой, числом операторов и показателями качества ее обслуживания. Примером такой модели является модель Эрланга с ожиданием, в течение многих лет служащая основой создания разного рода калькуляторов численности персонала справочных узлов. Правда, она не учитывает многих факторов, определяющих работу современных справочных служб, например, наличие IVR и влияние поведения клиента в форме повторных запросов. Для устранения перечисленных проблем разработаны и введены в обиход специалистов новые более точные модели. Следует отметить, что найденное по нагрузке число операторов необходимо корректировать в сторону увеличения. Это действие направлено на отсутствие операторов на рабочем месте по уважительным причинам: отпуск, болезнь, деловые встречи, учеба и т.п. Также необходимо принимать во внимание текучесть кадров и занятость оператора на время поствызовной обработки заявки. Степень увеличения численности по разным оценкам составляет 20-30 %. Она зависит от специфики работы анализируемого справочного узла.

Составление смен и расписаний работы операторов. На данном этапе можно приступить к расчету численности персонала с учетом индивидуальных графиков его работы. Оценивается число смен и количество операторов в каждой смене. Принимаются во внимание следующие факторы: найденная ранее величина потребности в количестве операторов, присутствующих на рабочем месте в рассматриваемом интервале времени; длительность смены; начало и продолжительность перерывов в работе. Общее число операторов находится методом целочисленного линейного программирования путем минимизации стоимостного функционала, отражающего расходы, понесенные на содержание оцениваемого числа операторов. Для простоты в качестве функционала можно взять величину общего количества операторов. Далее следует распределить операторов

по сменам. Остановимся подробнее на решении этой задачи. На рисунке 1.11 приведено расписание выхода на работу операторов некоторого гипотетического контакт-центра. Промежуток времени работы справочного центра длится с 7.00 до 19.30, когда уходит по-

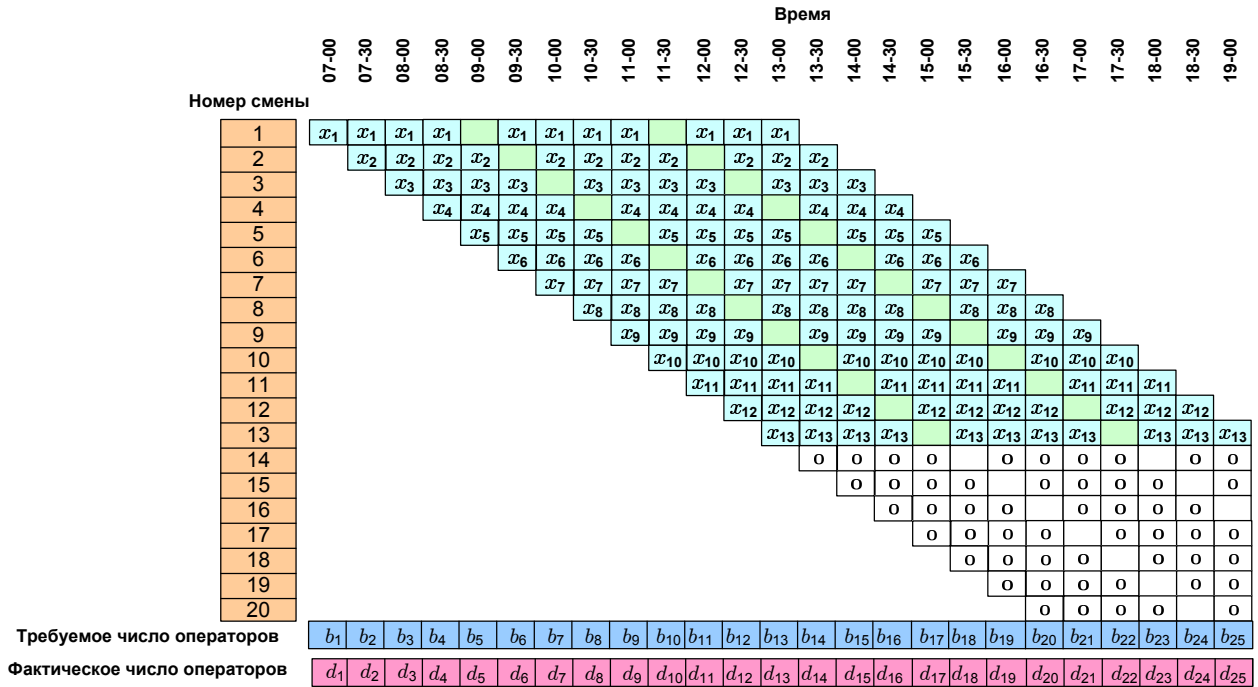


Рисунок 1.11. Оценка численности смен операторов

следняя смена. Первая смена приходит в 7.00, последняя — в 13.00. Продолжительность любой смены шесть с половиной часов и включает в себя два получасовых интервала для отдыха или приема пищи. На рисунке 1.11 через x_k обозначена численность k -ой смены, $k = 1, 2, \dots, 13$, через b_k обозначено необходимое по нагрузке число операторов, найденное с помощью соответствующей математической модели узла, через d_k обозначено фактическое число операторов, присутствующих на работе. Величины x_k минимизируют общее количество операторов, т.е. сумму

$$\min(x_1 + x_2 + \dots + x_{13})$$

при условии выполнения требуемых ограничений на число операторов для каждого из анализируемых интервалов. Вид этих соотношений прямо следует из содержания рисунка 1.11. Всего имеется 25 неравенств. Приведём несколько начальных соотношений. Они имеют вид:

$$x_1 \geq b_1;$$

$$x_1 + x_2 \geq b_2;$$

$$x_1 + x_2 + x_3 \geq b_3;$$

$$x_1 + x_2 + x_3 + x_4 \geq b_4;$$

$$x_2 + x_3 + x_4 + x_5 \geq b_5;$$

$$x_1 + x_3 + x_4 + x_5 + x_6 \geq b_6;$$

.....

Как уже было сказано, ответ ищется стандартными методами целочисленного линейного программирования и не вызывает затруднений.

Оперативные изменения. Они вносятся в принятый распорядок работы контакт-центра на основании анализа результатов текущих измерений характеристик качества обслуживания поступающих информационных запросов. Они призваны отразить возможные флуктуации в значениях и характере поступающих заявок, а также в числе присутствующих операторов.

1.6. Анализ выполненных исследований

Основные авторы. Число публикаций, посвященных различным аспектам анализа процессов функционирования действующих и перспективных контакт-центров, постоянно возрастает. Привлеченные актуальностью проблематики, в этой области теории телеграфика работают ученые самой разной профессиональной подготовки. Отметим малую часть зарубежных: S. Borst, A. Brandt, K. Dawson, G. Koole, A. Mandelbaum, R. Stollez, S. Zeltyn, M. Reiman и отечественных авторов: С.В. Ваняшин, Б.С. Гольдштейн, А.В. Росляков, А.Б. Самолюбова и дадим краткую характеристику выполненным исследованиям, сгруппировав их по используемым методам и полученным результатам.

Аналитические формулы и рекурсивные алгоритмы. Как уже было сказано, оценка необходимого по нагрузке числа операторов требует построения математической модели, отражающей процессы поступления и обслуживания информационных запросов клиентов. При разработке модели необходимо учесть последовательность использования ресурсов контакт-центра и особенности поведения пользователя его услуг при выполнении отдельных этапов сценария взаимодействия клиента и справочно-информационной службы. Для решения сформулированной задачи используются модели и методы теории телеграфика, представленные в основных руководствах, посвященных этой тематике [17,18,20,58,59,67,69,71,89]. На ранних этапах исследования контакт-центров для расчета числа операторов использовалась классическая модель Эрланга с ожиданием, имеющая в символике Кендалла обозначение $M/M/v$, а также некоторые её модификации, направленные на уточнение взаимодействия клиента и справочной службы. В общем виде схема функционирования модели показана на рисунке 1.12.

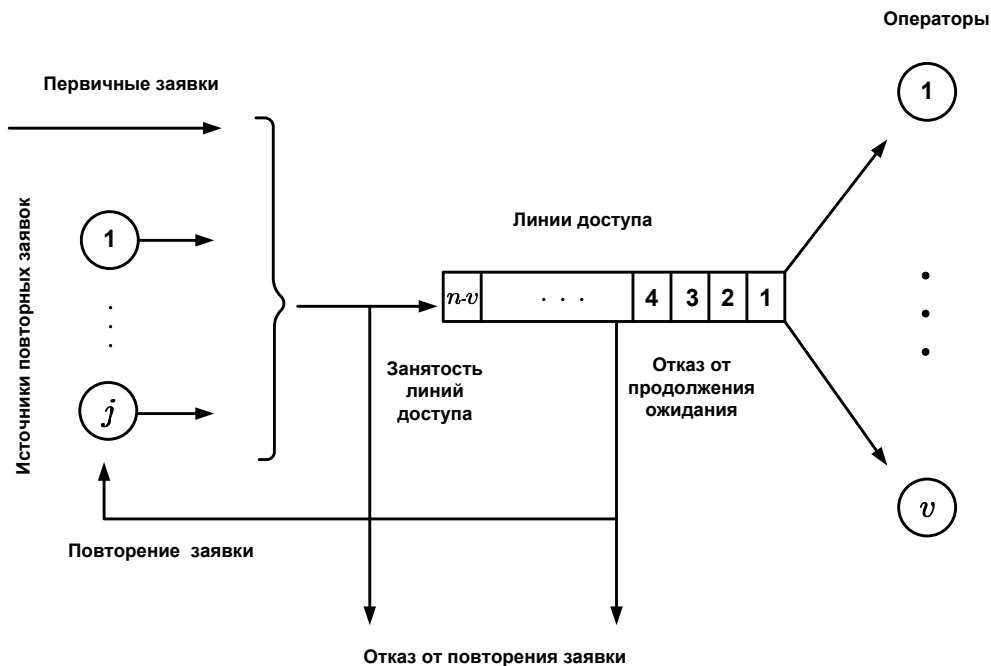


Рисунок 1.12. Схема поступления и обслуживания заявок в простейших моделях контакт-центра

Оценка показателей качества обслуживания поступающих заявок с помощью модели Эрланга с ожиданием не вызывает больших затруднений и основана на использовании значений формулы Эрланга, определяющей величину потерь заявок в аналогичной модели, но с потерями. Для вычисления значений формулы Эрланга разработаны специальные алгоритмы, широко представленные в литературе по теории телетрафика [20,69,89]. Обобщение анализируемой модели шло по нескольким направлениям. Отметим и обсудим основные. Это: ограничение очереди ожидающих заявок (англ. balking); ограничение длительности промежутка времени пребывания на ожидании (англ. reneging); возможность повторения заблокированной заявки (англ. retrials). Каждая из этих модификаций основной модели представлена на рисунке 1.12. В них меняется поведение абонента, получившего отказ. Эти изменения носят общее название *нетерпеливый абонент* (англ. impatient subscriber). Их действие может быть рассмотрено по отдельности или в совокупности. Перечислим основные полученные результаты. Начнем с моделей, в которых длительности интервалов времени между осуществлением событий, описываются случайными величинами, имеющими экспоненциальное распределение.

Ограничение очереди ожидающих заявок может исследоваться в более общем виде, если ввести понятие *функции внутренней блокировки*. Обозначим через f_i вероятность отказа в обслуживании для заявки, поступившей в то состояние модели, когда на обслуживании и ожидании находится i заявок. Выбрав значения f_i из соотношений $f_i = 0$,

$i = 0, 1, \dots, n - 1$ и $f_i = 1, i \geq n$, получаем модель с ограниченной очередью ожидающих абонентов длины $n - v > 0$. Нетрудно показать, что динамика изменения состояний моделей с возможностью внутренней блокировки и ограничением на время ожидания описывается марковским процессом рождения и гибели. В этой ситуации расчет характеристик качества обслуживания поступающих заявок не вызывает затруднений и выполняется с помощью хорошо изученных численных алгоритмов [20,69,87,89,105].

Учет возможности повторения заблокированной заявки резко усложняет оценку характеристик модели. Явные выражения получены только в случае одного обслуживающего прибора, когда абонент повторяет заблокированную попытку соединения с вероятностью равной единице. Для вычисления значений характеристик многолинейных систем без наличия мест ожидания заблокированных вызовов можно использовать рекурсивные численные методы решения систем уравнений равновесия. В этой ситуации решение всей системы уравнений разбивается на решение отдельных подсистем с размерностью на единицу большей числа обслуживающих приборов. Число подсистем на единицу больше, чем ограничение на максимальное число абонентов, одновременно повторяющих вызов. Эти подходы подробно исследованы в [14,15,19,58,103]. Если в модели предусмотрено ожидание заявки, получившей отказ, с возможным его прерыванием, но без повторения вызова, то для оценки значений характеристик также можно использовать рекурсивные методы решения систем уравнений равновесия, аналогичные представленным в [14,15,19,58,103]. Если же после неудачного завершения ожидания абонент имеет возможность для повторения вызова, то вычисление характеристик осуществляется с помощью решения системы уравнений равновесия каким-либо численным методом линейной алгебры. В анализируемом случае для расчета значений характеристик рекомендуется использовать итерационный метод Гаусса-Зейделя [57].

Перечисленные подходы основаны на численном решении системы уравнений равновесия и позволяют найти точные величины показателей качества обслуживания поступающих заявок. При анализе сложных моделей, к каковым безусловно относятся модели с учетом повторных вызовов, рекомендуется использовать приближенные алгоритмы. В теории телетрафика используются несколько способов решения сформулированной задачи. Отметим следующие два подхода, основанные на применении асимптотических разложений характеристик при стремлении входных параметров к своим предельным значениям. С точки зрения легкости реализации рекомендуется использовать асимптотические оценки характеристик, полученные при больших значениях входной нагрузки (стремлении интенсивности поступления первичных вызовов к бесконечности) [58,102] и больших значениях длительности интервала времени между последовательными повторными вызовами одного абонента (при стремлении интенсивности повторения заявок одного абонента к нулю) [58,61]. Численные исследования показали, что оба подхода отличаются хорошей точностью [58,61,102], которая в большинстве случаев лежит в пределах 10 %.

Назовем модель, представленную на рисунке 1.2, основной моделью контакт-центра. Дальнейшие усилия исследователей были направлены на снятие некоторых ограничений, использованных при ее построении. Часть работ посвящена уточнению процедуры моделирования промежутков времени между осуществлением событий, другая группа результатов связана с детализацией алгоритма обслуживания запросов. Остановимся на этих публикациях более подробно. Среди полученных результатов отметим модель, в которой ограничение на время пребывания в очереди имеет произвольное распределение [77]. Важным частным случаем этой модели является система с постоянным ограничением на время пребывания в очереди. Алгоритм расчета характеристик представлен в [7]. Дальнейшие продвижения в этом направлении приводят к очень сложным моделям, которые с большим трудом поддаются численному анализу и по этой причине не используются в практической деятельности инженеров связи.

Имеется ряд работ, где принимается во внимание наличие IVR и возможность разделения операторов по квалификации. Каждое из этих уточнений увеличивает число компонент в случайном процессе, описывающем функционирование математической модели, что естественным образом усложняет её описание и исследование. В [78] рассматривается модель контакт-центра, состоящая из двух групп приборов. Первая группа представляет из себя устройства IVR, вторая — операторов. Длительность времени обслуживания в каждой группе имеет экспоненциальное распределение. После завершения обслуживания в IVR запрос с некоторой вероятностью считается завершившим обслуживание, а с дополнительной вероятностью продолжает его у оператора. Если все операторы заняты, то предусмотрена очередь ожидания начала обслуживания. Введением дополнительной фиктивной группы приборов рассматриваемая открытая сеть трансформируется в замкнутую сеть, состоящую из трех узлов, для которой существует мультипликативное решение, используемое далее для оценки значений стационарных характеристик. В построенной модели не учитываются повторные вызовы, возможность прекращения ожидания обслуживания и наличие операторов разной квалификации.

Последние две из перечисленных проблем анализировались в [104], где построена и исследована модель контакт-центра с тремя группами операторов, различающихся своей квалификацией, и двумя конечными очередями ожидания начала обслуживания. Длительность промежутка времени пребывания в каждой из очередей ограничена случайной величиной, имеющей экспоненциальное распределение с параметром, зависящим от номера очереди. Такое же распределение имеют интервалы времени между последовательными поступлениями заявок и длительности их обслуживания. Построенная модель описывается четырехмерным марковским процессом. С помощью значений стационарных вероятностей модели получены соотношения для оценки основных показателей качества обслуживания поступающих заявок. Далее система уравнений равновесия преобразуется к виду удобному для проведения вычислений, выполняется расчет значений характеристик и анализируется их зависимость от величины входных параметров.

Отсутствие специальных свойств у матрицы системы уравнений равновесия ограничивает область использования стандартных численных методов при ее решении. К тому же в модели не учитываются влияние повторных вызовов и наличие IVR.

Имитационное моделирование. Перечисленные выше способы оценки значений характеристик качества обслуживания поступающих информационных запросов основаны на реализации аналитических методов анализа динамики изменения состояний контакт-центра. Альтернативой ему служат методы имитационного моделирования. При их использовании значения характеристик оцениваются путем подсчета значений счетчиков анализируемых событий. Для реализации имитационных алгоритмов созданы специальные программные комплексы. Имитационные модели позволяют более точно воссоздавать процессы поступления и обслуживания заявок в реальных справочно-информационных службах. Однако их использование требует значительных вычислительных ресурсов и не является оптимальным средством планирования инфраструктуры контакт-центров. В основном имитационные модели применяются для оценки погрешности приближенных аналитических методов и являются частью программно-аппаратных комплексов моделирования работы контакт-центров.

Функция распределения длительности ожидания начала обслуживания. В результате аналитического или имитационного моделирования работы контакт-центра обычно оцениваются величины стационарных характеристик качества обслуживания поступающих запросов из перечня, приведенного в разделе 1.4.1. К ним относятся: доля потерянных заявок; доля запросов, обслуженных с первой попытки; среднее время ожидания ответа оператора и ряд других показателей. Для некоторых вариантов основной модели контакт-центра могут быть получены и выражения для функции распределения времени пребывания в очереди на ожидании начала обслуживания. В первую очередь это относится к модели Эрланга с ожиданием. Обозначим через \mathcal{W} случайную длительность времени нахождения заявки на ожидании. Обозначим через $p\{\mathcal{W} > 0\}$ вероятность попадания поступившей заявки на ожидание. Справедливо соотношение [20,69,87,89,105]

$$p\{\mathcal{W} \leq t\} = 1 - p\{\mathcal{W} > 0\} \cdot e^{-v\mu t(1-\delta)} = 1 - D(v, a) \cdot e^{-v\mu t(1-\delta)}. \quad (1.1)$$

Здесь a — интенсивность предложенного трафика в эрлангах, v — число операторов, $D(v, a)$ — так называемая вторая формула Эрланга

$$\begin{aligned} D(v, a) &= \frac{\frac{a^v}{v!} \cdot \frac{v}{v-a}}{(1 - E(v, a)) \sum_{i=0}^v \frac{a^i}{i!} + \frac{a^v}{v!} \cdot \frac{v}{v-a}} = \\ &= \frac{v E(v, a)}{(v-a)(1 - E(v, a)) + v E(v, a)} = \frac{v E(v, a)}{v - a(1 - E(v, a))}. \end{aligned} \quad (1.2)$$

Полученное соотношение даёт возможность рассчитывать величину $D(v, a)$ с помощью эффективных алгоритмов, разработанных для вычисления значений формулы Эрланга

$E(v, a)$. Имеются обобщения соотношения (1.2) на случай аналогичной модели обслуживания заявок, но с ограниченной очередью, а также других дисциплин выбора из очереди ожидающих заявок (напомним, что в модели Эрланга используется дисциплина *первый пришел — первый на обслуживание*). Однако выглядят они громоздко [87, 96] и обычно при расчетах не используются.

Калькуляторы числа элементов инфраструктуры контакт-центров. Выражение (1.2) позволяет оценить значение уровня обслуживания поступающих заявок, определяемого через процент заявок, попавших на обслуживание к оператору за заданный промежуток времени. Как уже было сказано (см. раздел 1.4.1), чаще всего используется следующее соотношение: 80 % всех поступивших заявок попадут на обслуживание по истечении 20 с. Эта характеристика, а также величина среднего времени пребывания заявки на ожидании начала обслуживания являются основными рассчитываемыми показателями при использовании разного рода калькуляторов инфраструктуры контакт-центров, представленных в сети Интернет. Интерфейс одного из таких калькуляторов, разработанных компанией Westbay Engineering [79], показан на рисунке 1.13.

Hour	Calls	Avg. delay	Agents	Lines
Hour 1	100	13	10	12
Hour 2		-	-	-
Hour 3		-	-	-
Hour 4		-	-	-
Hour 5		-	-	-
Hour 6		-	-	-
Hour 7		-	-	-
Hour 8		-	-	-

Ср. продолжительность ответа оператора (180 с)

Ср. продолжительность поствызывная работа оператора (60 с)

Качество обслуживания у оператора:

80 % заявок будут обслужены в течение 20 с

Качество обслуживания на линиях доступа:

1 % заявок будут потеряны на доступе

Входные параметры:

Palls: среднее число вызовов, поступивших за час

Рассчитываемые характеристики:

Avg. delay - средняя длительность пребывания на ожидании для всех поступивших заявок (с)

Agents - требуемое число операторов

Lines - требуемое число линий доступа

Рисунок 1.13. Интерфейс калькулятора erlang@com для оценки числа элементов инфраструктуры контакт-центра

Для оценки значений показателей качества обслуживания поступающих информационных запросов используется модель Эрланга с ожиданием и ограниченной очередью (см. рисунке 1.12). Входными параметрами модели являются: средняя продолжительность обслуживания заявки, выраженная в секундах; средняя продолжительность поствызывной работы оператора, выраженная в секундах; среднее число заявок, поступающих в контакт-центр, за час наблюдений. Целевыми функционалами, регламентирующими качество работы контакт-центра, являются значения доли заявок, попавших на обслуживание к оператору за заданный промежуток времени (для определения числа операторов) и значение доли потерянных заявок (для определения числа линий доступа). Результаты вычисления для одного из частных случаев представлены на рисунке 1.13.

Использование эмпирических подходов. Эти подходы весьма популярны среди инженеров, занимающихся проектированием и эксплуатацией справочно-информационных служб. Одним из самых известных методов является так называемое *оценка численности персонала, основанная на принципе квадратного корня* (англ. *square root staffing principle*). Для оценки величины v числа операторов предлагается использовать выражение

$$v = a + k\sqrt{a}. \quad (1.3)$$

Здесь: a — интенсивность поступающей нагрузки, выраженная в эрлангах, k — эмпирический коэффициент, значение которого подбирается опытным путем из стоимостных соображений и величины характеристик качества обслуживания поступающих запросов. Соотношения вида (1.3) использовались еще со времен Эрланга, но только в последние годы они получили частичное теоретическое обоснование [75]. Отметим, что точность применения (1.3) повышается, если анализируются справочные службы с большой нагрузкой и большой численностью операторов.

Опыт диссертационных работ. Анализ выполненных диссертационных исследований [2,3,5,10,12,63,66] показал, что в большинстве работ либо изучается действие какого-то одного фактора на процесс функционирования справочного центра (например, самоподобные свойства потока заявок [12], поствызывная работа оператора [66], наличие нескольких видов запросов [5,10], повторные вызовы [63]), либо предлагаются методики повышающие эффективность мониторинга и управления работой контакт-центра [2,3]. При этом используются самые простые модели теории телетрафика, естественно, после соответствующей интерпретации входных параметров и характеристик. Актуальным направлением исследований представляется изучение комплексного воздействия перечисленных факторов на пропускную способность справочно-информационных служб. Однако исследование подобного рода требует построения и анализа достаточно сложных математических моделей.

Общий анализ опубликованных работ и направления будущих исследований. Сразу отметим разнообразие выполненных исследований. Диапазон опублико-

ванных результатов меняется от построения изоощренных математических моделей и методов до разработки простейших эмпирических правил определения числа элементов инфраструктуры контакт-центров, основанных на практическом опыте и инженерной интуиции. Эволюция интересов ученых шла нога в ногу с развитием справочных служб. Когда эти службы представляли из себя группу операторов, отвечающих на запросы абонентов телефонной сети о телефонных номерах других абонентов и организаций, то для оценки их численности было достаточно использовать простейшие модели Эрланга с потерями и ожиданием. В настоящий момент ситуация существенно изменилась. Расширился спектр деятельности справочных служб. Это уже давно не справочные службы телефонных операторов, а часто независимые компании, успешно ведущие свой бизнес на рынке услуг. В результате расширилось число и виды каналов доступа, более специфической стала запрашиваемая информация. Как следствие всего этого, обслуживание клиентов стало дифференцированным с привлечением элементов искусственного интеллекта. Большую роль в формировании информационных запросов стала играть модель поведения пользователя услуг контакт-центра. Все эти изменения должны анализироваться в комплексе и быть представлены в результатах теоретических исследований, посвященных данной тематике.

1.7. Выводы и постановка задачи диссертационного исследования

Анализ современного состояния и путей развития контакт-центров показал, что справочно-информационные службы будущего должны предоставить пользователю весь возможный спектр информационных услуг с использованием самых разнообразных линий доступа. Их перечень определяется желанием администрации обеспечить единую точку доступа к услугам для клиентов, находящихся где угодно и в любой доступной для них форме. В большинстве случаев это — обычное телефонное соединение, предлагаемое традиционными сетями фиксированной и подвижной связи. Заявка также может поступить с использованием возможностей и технологий сети Интернет: электронной почты, web-чата, социальных сетей и т.п. Расширение области применения контакт-центров усложнили характер запрашиваемой информации. Для повышения эффективности работы справочных служб стало необходимым использование дифференцированного обслуживания клиентов (операторы, консультанты), а также формирование ответа на запросы с помощью устройств IVR, синтезаторов речи и других элементов искусственного интеллекта. Усложнение модели обслуживания клиента, перенос на него части функций по коммутации запроса повысили зависимость процедуры формирования входного потока заявок от реакции абонента на неудачную попытку соединения и неудачное завершение времени ожидания. Для научно обоснованной оценки количества элементов

инфраструктуры контакт-центров, используемых для обслуживания поступающих информационных запросов (линии доступа, устройства IVR, операторы и консультанты) в изменившихся условиях построения современных справочных служб, требуется решить следующие задачи:

1. Разработать обобщенную математическую модель поступления и обслуживания информационных запросов, учитывающую особенности построения и работы действующих и перспективных контакт-центров. В перечень анализируемых особенностей включить следующие факторы: использование нескольких видов доступа; наличие предварительно голосового обслуживания IVR; использование операторов и консультантов; возможность повторения заявки из-за нехватки линий доступа, недостатка операторов или консультантов, а также присутствие ограничения на время ожидания начала обслуживания.
2. С помощью построенной модели получить выражения для вычисления значений основных показателей качества обслуживания поступающих информационных запросов. Установить соотношения между характеристиками работы контакт-центра, облегчающими их измерение или вычисление. Для обобщенной модели и ее частных случаев разработать точные и приближенные алгоритмы оценки введенных показателей, пригодные для практического использования во всей области изменения входных параметров. Оценить погрешность приближенных методов.
3. Использовать построенную модель и методы ее анализа и расчета для решения практических задач, возникающих в процессе проектирования и эксплуатации контакт-центров. В частности: оценить влияние входных параметров на значения показателей качества обслуживания поступающих заявок; получить выражение для косвенной оценки величины интенсивности поступления первичных вызовов и значений других характеристик, прямое измерение которых затруднено из-за необходимости отделения первичных вызовов от повторных; построить процедуру оценки числа линий доступа и операторов.

Глава 2

Построение и анализ математической модели контакт-центра

2.1. Введение

Аналитическое исследование процесса обслуживания заявок на предоставление различного рода услуг связи обычно проводится в рамках марковских моделей. Использование данного класса моделей существенно упрощает оценку характеристик, сохраняя при этом возможность учёта особенностей формирования входных потоков заявок. Схема функционирования обобщённой модели контакт-центра приводится в разделе 2.2. Там же анализируются основные этапы построения марковского процесса, описывающего динамику изменения числа пользователей услуг информационного центра, находящихся на различных этапах обслуживания, ожидания или повторения заявки. В разделе 2.3 сформулированы определения показателей качества обслуживания заявок через значения стационарных вероятностей пребывания модели в различных состояниях. В разделе 2.4 получены соотношения, связывающие значения введённых характеристик, облегчающие их вычисление или измерение. Подобные соотношения имеют характер законов сохранения интенсивностей потоков заявок, поступивших на определённый этап обслуживания и завершивших его. Построенная модель в общем виде описывает процесс поступления и обслуживания заявок в действующих и перспективных контакт-центрах. Её частные случаи составляют семейство упрощённых моделей, которые можно использовать для анализа отдельных функциональных компонент контакт-центра (места ожидания, операторы и т.д.), и разработки приближённых алгоритмов её расчёта. Эти модели и основные их свойства перечислены в разделе 2.5. В следующем разделе сформулированы выводы по 2-ой главе.

2.2. Математическое описание модели

2.2.1. Схема функционирования

Функциональная модель. Запросы абонентов на получение разного рода информационных услуг поступают в контакт-центр по линиям доступа и далее в зависимости от вида доступа обслуживаются устройствами IVR, операторами и, в случае необходимости, консультантами. Дадим краткую характеристику базовым функциональным составляющим математической модели контакт-центра. Структура и взаимодействие основных функциональных элементов действующих и перспективных контакт-центров исследованы в разделе 1.2.2 и показаны на рисунке 2.1. Вначале рассмотрим особенности поступления и обслуживания заявок по телефонным линиям доступа (т.е. с использованием ресурса ТфОП и сетей мобильной связи), а затем с использованием возможностей сети Интернет.

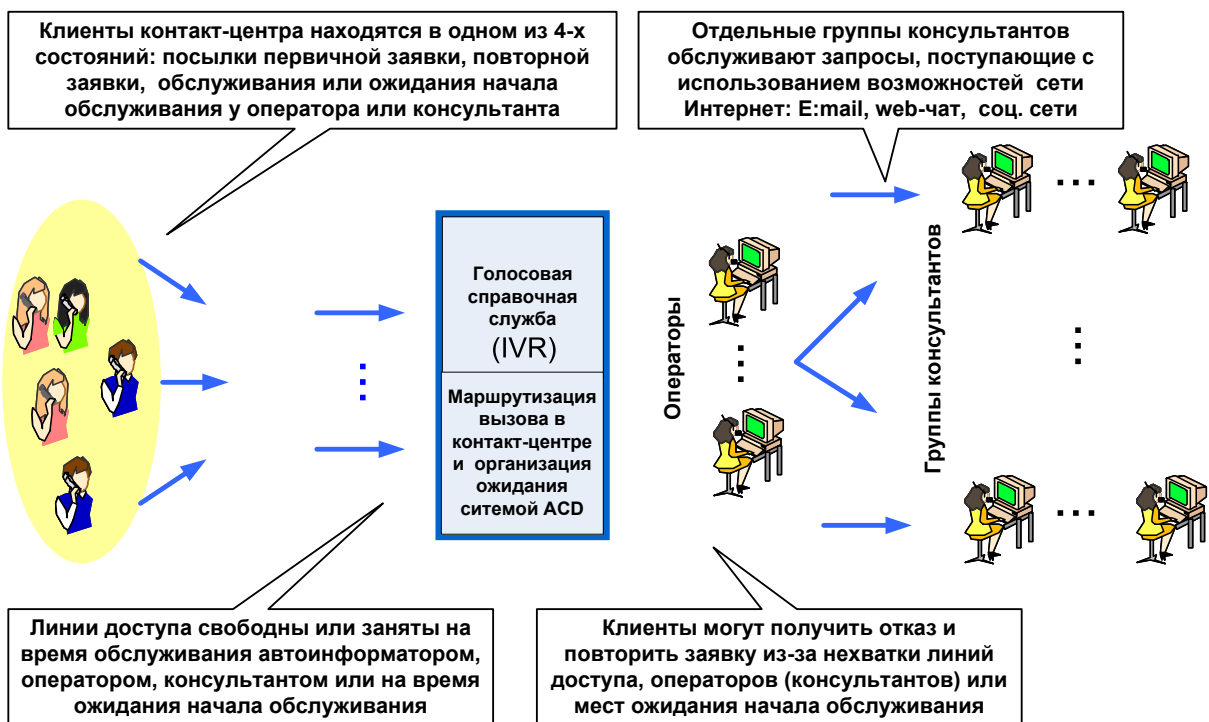


Рисунок 2.1. Функциональная модель контакт-центра

Модель поступления заявок. Будем полагать, что поток первичных заявок от пользователей услуг контакт-центра подчиняется закону Пуассона с интенсивностью λ . Используемая гипотеза является общепринятой в научной литературе, посвященной исследованию процессов формирования заявок в контакт-центрах [38-40,61,81,91,94,104]. Она имеет обоснование в известном математическом результате [1,4,7], из которого следует, что суперпозиция большого числа редких независимых потоков заявок с суммарной

интенсивностью, стремящейся к константе, приобретает свойства пуассоновского потока. Помимо первичных заявок от клиентов поступают потоки повторных заявок, образованные из-за отсутствия свободных линий доступа, нехватки операторов или консультантов, а также после неудачного завершения времени ожидания начала обслуживания у оператора или консультанта. Отметим, что в исследуемой модели пуассоновские свойства наличествуют только у потока первичных заявок. Суммарный поток поступающих первичных и повторных заявок не является пуассоновским, более того, интервалы времени между последовательными поступлениями заявок представляют собой зависимые случайные величины¹. При построении модели необходимо учитывать занятость линии доступа при обслуживании заявки оператором или консультантом. Она освобождается только после удачного завершения обслуживания или отказа клиента от попытки установить соединение. Для оценки величины λ необходимо в общем потоке поступающих заявок разделить первичные вызовы и повторные. Эта процедура трудоёмка и не всегда возможна. Решить сформулированную задачу можно косвенными методами после построения математической модели процесса поступления и обслуживания вызовов в контакт-центре. Одно из возможных решений рассмотрено в [51,53,60].

Последовательность обслуживания поступающих заявок. Анализ процедуры обработки заявок, поступающих в контакт-центр, выполненный в разделе 1.4, даёт основание использовать следующую математическую модель их обслуживания. Будем предполагать, что процесс обслуживания клиента контакт-центра состоит из трех фаз. Первая фаза заключается в получении информационного сообщения от устройства IVR. Пользователь услуг контакт-центра может получить требуемую информацию либо на уровне общения с IVR, либо перейти на обслуживание к оператору. Хотя длительности воспроизводства звуковых сообщений устройством IVR фиксированы, алгоритм их выбора и продолжительность выслушивания случайны и зависят от индивидуальных характеристик абонента. Это позволяет считать длительность обслуживания клиента системой IVR имеющим экспоненциальное распределение. Обозначим через α_1 параметр этого распределения, тогда величина $\frac{1}{\alpha_1}$ — среднее значение промежутка времени обслуживания клиента устройством IVR. Это соотношение можно использовать для оценки значения α_1 . Вторая и третья фазы обслуживания заключаются соответственно в возможности получения общей справочной информации от оператора или, в случае необходимости, запроса специализированной справочной информации от консультанта. Продолжительность пребывания клиента на каждой из этих фаз зависит от индивидуальных характеристик участников информационного обмена (вид запроса, понятливость абонента, квалификация сотрудника контакт-центра и т.п.) и не имеет каких-либо предсказуемых свойств. Это даёт основание считать длительности обслуживания клиента оператором и консультантом имеющими экспоненциальное распределение с параметра-

¹Момент поступления заявки может зависеть от состояния модели, предшествующего предыдущему моменту поступления заявки.

ми соответственно α_2 и β_k , где k — номер группы консультантов, $k = 1, 2, \dots, m$, а m — число групп консультантов. Исходя из свойств экспоненциального распределения выражения $\frac{1}{\alpha_1}$ и $\frac{1}{\beta_k}$ представляют собой среднюю длительность пребывания клиента на соответствующем этапе обслуживания и могут использоваться для оценки значений параметров α_2 и β_k .

Вероятности переходов между этапами обслуживания заявок. Переходы клиента с одной из отмеченных ранее фаз обслуживания на другую моделируется следующим образом. Будем предполагать, что с вероятностью q после завершения обслуживания устройством IVR абонент потребует продолжения обслуживания у оператора, а с дополнительной вероятностью $(1 - q)$ завершит обслуживание в контакт-центре, получив требуемую информацию. Также будем предполагать, что с вероятностью p после завершения обслуживания у оператора абоненту потребуется получить специализированную информацию у консультанта. Соответственно, предполагается, что с вероятностью $(1 - p)$ подобного желания не возникнет и абонент завершит обслуживание в контакт-центре. Выбор k -ой группы консультантов осуществляется с вероятностью p_k , $\sum_{k=1}^m p_k = 1$. Значения вероятностей q , p , p_k , $k = 1, 2, \dots, m$ определяются с использованием данных статистических наблюдений на основании анализа частот реализации каждой из перечисленных фаз обслуживания. Понятно, что величины этих вероятностей зависят от специфики информационных услуг, предоставляемых контакт-центром; от наличия устройств IVR и групп консультантов; от дня недели, сезона и конкретного времени проведения измерений.

Отказы в обслуживании поступающих заявок. В модели учитываются пять видов отказов в обслуживании заявки. Первый — из-за нехватки линий доступа, второй — из-за занятости всех операторов и мест ожидания их освобождения, третий — из-за превышения максимально возможного времени ожидания освобождения оператора, четвертый — из-за занятости консультантов выбранной группы и мест ожидания их освобождения, пятый — из-за превышения максимально возможного времени ожидания освобождения консультанта рассматриваемой группы. После отказа клиент с определенной вероятностью, определяемой принятой моделью поведения абонента после получения отказа в обслуживании, повторяет попытку соединения через случайный промежуток времени или отказывается от обслуживания. Величины вероятностей реализации всех перечисленных видов отказов в обслуживании могут быть найдены с помощью разработанной обобщенной математической модели контакт-центра и далее использоваться для решения задач планирования элементов инфраструктуры контакт-центров и проведения мероприятий по повышению эффективности его работы. Для этого необходимо составить и решить систему уравнений равновесия марковского процесса, описывающего функционирование построенной модели, или воспользоваться одним из предложенных приближенных методов оценки ее вероятностных характеристик. Часть из отмеченных вероятностей отказов могут найдены в результате измерений процесса поступления и об-

служивания вызовов в контакт-центре. Какие именно вероятности могут быть измерены зависит от возможностей установленной измерительной аппаратуры.

Ожидание начала обслуживания. В модели предполагается возможность ожидания начала обслуживания заявки, если в момент ее поступления заняты все операторы или все консультанты из выбранной группы. Время ожидания ограничено случайной величиной, имеющей экспоненциальное распределение с параметром σ , если речь идет об ожидании начала обслуживания у оператора, или σ_k , если рассматривается ожидание начала обслуживания у консультанта из k -ой группы. Средние значения отмеченных длительностей ожидания определяются соответственно из выражений $\frac{1}{\sigma}$ и $\frac{1}{\sigma_k}$, $k = 1, 2, \dots, m$. Эти величины либо заданы технической администрацией контакт-центра с целью ограничения длины очереди ожидающих абонентов, либо формируются на основе индивидуальных характеристик нетерпеливости абонента в процессе ожидания. В первом случае ограничение на длительность ожидания принимает фиксированное заранее известное значение, во втором — случайное. Выбор экспоненциального распределения длительностей рассматриваемых промежутков времени позволяет использовать марковский процесс при описании динамики изменения состояний модели и, как показали результаты численных экспериментов, не сильно сказывается на точности оценки характеристик модели.

Линии доступа, операторы и группы консультантов. Обозначим через n число линий доступа, через v — число операторов, а через v_k — число консультантов в k -ой группе, $k = 1, 2, \dots, m$. В соответствии с особенностями функционирования контакт-центров, линия доступа может быть занята на передачу следующих информационных потоков: на получение голосового сообщения от системы IVR; на время обслуживания или ожидания начала обслуживания у оператора; на время обслуживания или ожидания начала обслуживания у консультанта из k -ой группы, $k = 1, 2, \dots, m$. Понятно, что число линий доступа должно быть больше, чем суммарное количество операторов и консультантов, т.е. выполняется неравенство $n \geq v + \sum_{k=1}^m v_k$. Избыток линий доступа используется абонентами для ожидания начала обслуживания, если таковое необходимо из-за нехватки операторов или консультантов. Предполагается, что число линий доступа ограничено. Значения v и v_k , $k = 1, 2, \dots, m$ определяются исходя из принятых ограничений на показатели качества обслуживания поступающих заявок и технической политики, проводимой администрацией контакт-центра. Для этих целей используются модели и методы теории телетрафика.

Места ожидания. Обозначим через w число мест ожидания освобождения одного из v занятых операторов, а через w_k обозначим число мест ожидания освобождения одного из v_k занятых консультантов, $k = 1, 2, \dots, m$. Значения w и w_k , $k = 1, 2, \dots, m$ являются функцией v и v_k и определяются исходя из нормированных значений характеристик работы контакт-центра и технической политики проводимой её администрацией. Для этих целей используются модели и методы теории телетрафика. Из ограниченности

числа линий доступа следует ограниченность количество мест ожидания.

Модель поведения пользователя услуг контакт-центра после получения отказа в обслуживании. Клиент, получив отказ в обслуживании, с вероятностью H повторяет заявку через случайное время, имеющее экспоненциальное распределение с параметром, равным ν , а с дополнительной вероятностью $(1 - H)$ он отказывается от попыток получить обслуживание в контакт-центре. Величины H и ν не зависят от причины отказа в обслуживании. Повторная заявка обслуживается по тем же правилам, что и первичная. Длина интервала времени между последовательными повторными попытками одного абонента (в исследуемой модели её величина определяется из соотношения $\frac{1}{\nu}$) отсчитывается от момента получения отказа в обслуживании до момента поступления повторной заявки в контакт-центр. Таким образом, в рассматриваемый интервал включены: время принятия решения о посылке повторного вызова, время ожидания начала набора номера, время набора номера и время прохождения заявки по сети до контакт-центра. Результаты численного анализа, приведённые в [58], показывают слабую зависимость значений характеристик построенной модели от величины ν . Это свойство позволяет использовать при проведении вычислений достаточно грубые оценки значений ν . Оно также служит основой построения приближённых алгоритмов расчёта введённой модели контакт-центра.

Обратимся к анализу значений, принимаемых вероятностью повторения заявки H . Понятна трудоёмкость организации прямых измерений величины H , требующая организации наблюдения за полным циклом контактов клиента с анализируемой справочно-информационной системой. По этой причине результатов статистических измерений величины H довольно мало. Следует отметить исследование, проведённое Р.Эверсом на учрежденческих АТС [58], относящееся ко времени построения и анализа простейших моделей с учётом эффекта повторных вызовов. Для оценки значения H можно использовать косвенный подход. Обозначим через M_b среднее число повторных попыток, которое делает абонент, получившей в первичной попытке отказ в обслуживании. Будем полагать, что во всех последующих попытках абонент также получит отказ. Нетрудно в этих условиях рассчитать величину M_b . Она определяется из соотношения

$$\begin{aligned} M_b &= H(1 - H) \cdot 1 + H^2(1 - H) \cdot 2 + H^3(1 - H) \cdot 3 + \dots + = \\ &= H(1 - H) \cdot (1 + 2H + 3H^2 + \dots) = \frac{H}{1 - H}. \end{aligned} \quad (2.1)$$

Оценка величины M_b может быть известна из эмпирических соображений или задаваться документами, регламентирующими качество обслуживания абонентов контакт-центра. В этом случае выражение (2.1) можно использовать для косвенной оценки значения H . Расчётная формула следует из (2.1) и имеет вид

$$H = \frac{M_b}{1 + M_b}. \quad (2.2)$$

Для значений M_b , лежащих в интервале 2—3, получаем следующий интервал изменения H : $0,67 < H \leq 0,75$. Для величин M_b из интервалов 3—5 и 5—10, значение H , соответственно, удовлетворяет неравенствам: $0,75 < H \leq 0,83$ и $0,83 < H \leq 0,91$. Первый из рассмотренных интервалов изменения H соответствует случаю малой настойчивости абонента в установлении соединения, второй — средней настойчивости, а третий — сильной настойчивости.

Модель обслуживания запросов, поступающих с использованием возможностей сети Интернет. Будем полагать, что поток заявок от клиентов контакт-центра, использующих в качестве доступа возможности, предоставляемые сетью Интернет (далее этот вид запросов будем обозначать термином *интернет-запрос*), подчиняется закону Пуассона с интенсивностью λ_k . Номер k относится к характеристике доступа. Это может быть электронная почта, социальные сети, web-чат, мобильные приложения, СМС и т.д. (см. раздел 1.1.1). Заявки k -го потока попадают на обслуживание в k -ую группу консультантов. Предположим, что длительность времени обслуживания интернет-запроса в k -ой группе консультантов имеет экспоненциальное распределение с параметром β_k , где k — номер группы, $k = 1, 2, \dots, m$. Если все консультанты заняты, то запрос становится на ожидание. Промежуток времени ожидания ограничен случайной величиной, имеющей экспоненциальное распределение с параметром σ_k . Если заняты все консультанты и места ожидания начала обслуживания, то клиент с вероятностью H_k повторяет заявку через случайное время, имеющее экспоненциальное распределение с параметром равным ν_k , а с дополнительной вероятностью $(1 - H_k)$ он отказывается от попыток получить обслуживание в контакт-центре. Если при каком-то значении k поток интернет-запросов отличен от нуля т.е. $\lambda_k > 0$, то предполагается, что $p_k = 0$. Таким образом, k -ая группа консультантов может обслуживать либо телефонные запросы, либо интернет-запросы. Подобное разделение трафика связано со спецификой обработки интернет-запросов и практикуется в крупных контакт-центрах, например, в контакт-центре компании ОАО МТС. Анализ общего случая, когда k -ая группа консультантов может обрабатывать запросы всех видов доступа, не вызывает принципиальных затруднений, но здесь не рассматривается из-за громоздкости используемых вероятностных конструкций². Исследуемая модель обслуживания заявки в контакт-центре показана на рисунке 2.2.

2.2.2. Численные значения входных параметров

Величины, принимаемые входными параметрами модели, будут введены далее при проведении численных расчётов. Интенсивности поступления заявок, значения времени её пребывания на этапах обслуживания, длина интервала времени между последовательными повторными попытками одного пользователя услуг контакт-центра, длительности

²Отметим также, что наличие ожидающих интернет-запросов можно учесть, если увеличить длительность поствызывной работы консультанта и включить это время в общую продолжительность обслуживания запроса.

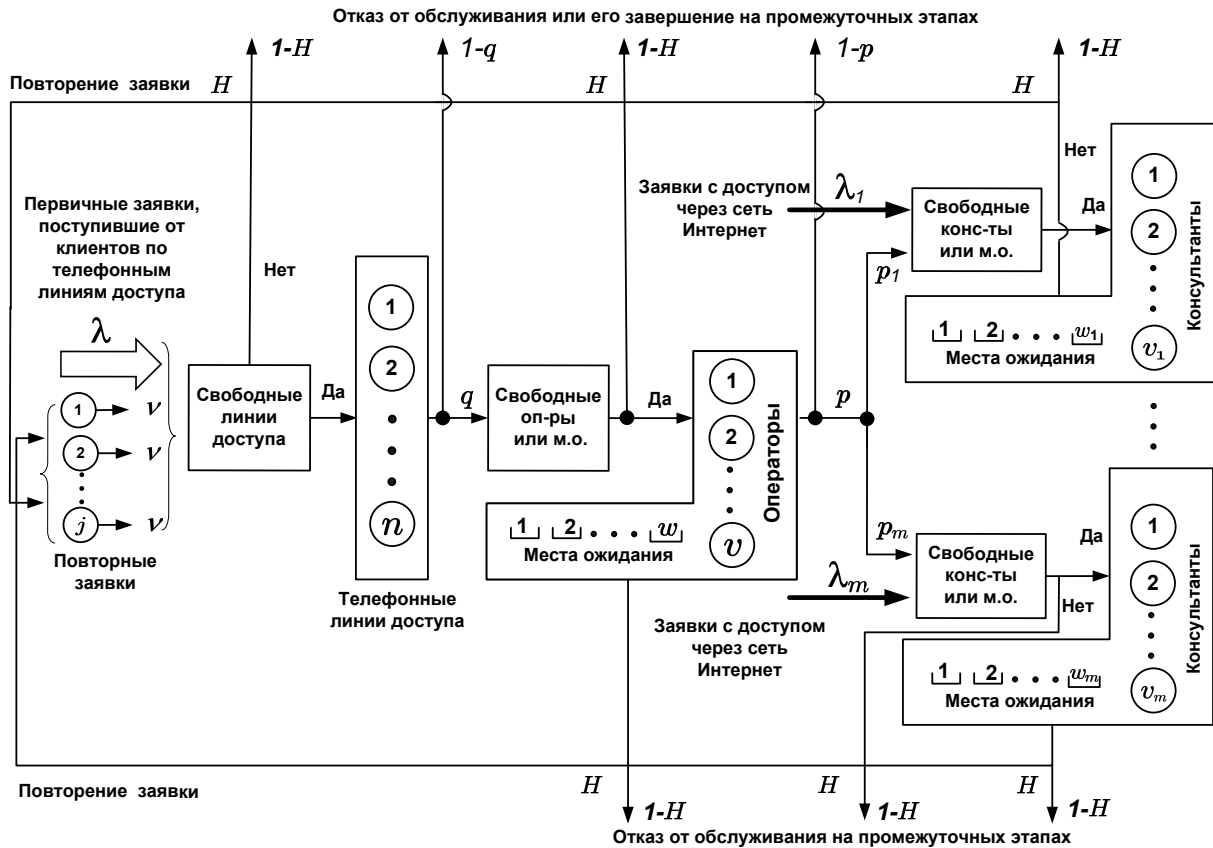


Рисунок 2.2. Схема функционирования математической модели контакт-центра

ограничений на ожидание начала обслуживания выражены в используемую в модели единицу времени. При проведении вычислений в качестве единицы времени обычно выбиралась средняя длительность обслуживания заявки устройством IVR или оператором. Соотношения между реальными значениями этих параметров обсуждались в разделе 1.4.1. Физическая интерпретация характеристик, которую можно использовать для оценки их значений рассмотрена в разделе 2.2.1. Отметим, что полученные далее расчетные выражения и соотношения между показателями качества обслуживания пользователей услуг контакт-центра в слабой степени ограничены конкретным выбором величин входных параметров и могут использоваться для любых их значений, представляющих практический интерес.

2.2.3. Состояния модели и динамика их изменения

Перейдем к математическому описанию процесса изменения состояний исследуемой модели контакт-центра. По построению процесс обслуживания интернет-запросов в каждой группе консультантов происходит независимо от других групп и от процес-

са поступления и обслуживания запросов, поступивших посредством телефонной сети. Динамика изменения числа интернет-запросов, находящихся в состоянии обслуживания, ожидания начала обслуживания или повторения запроса, описывается с помощью марковского процесса, построенного и исследованного в разделе 2.5.5. По этой причине они будут исключены из определения состояния модели контакт-центра. Для численной оценки значений показателей качества обслуживания интернет-запросов используются соотношения (2.29).

Рассмотрим процесс изменения в стационарном режиме числа запросов, поступивших по телефонным линиям доступа (т.е. с использованием ресурса ТфОП и сетей мобильной связи), в состоянии обслуживания, ожидания начала обслуживания и повторения запроса. Для существования стационарного режима достаточно потребовать выполнения соотношений: $H < 1$ и $\nu > 0$. Если параметры обобщённой модели контакт-центра удовлетворяют сформулированным условиям, то среднее число абонентов повторяющих вызов ограничено. Для доказательства этого утверждения будем полагать, что во всех попытках установить соединение абонент получает отказ. Обозначим через j число абонентов, находящихся в состоянии повторения вызова в рассматриваемой вспомогательной модели. Случайный процесс, описывающий изменение состояний (j) будет процессом рождения и гибели с интенсивностями перехода из состояния (j) в состояние ($j + 1$) и из состояния (j) в состояние ($j - 1$) равными, соответственно, λH и $j\nu(1 - H)$. Величина $p(j)$ стационарной вероятности состояния (j) определяется из соотношения

$$p(j) = \frac{1}{j!} \left(\frac{\lambda H}{\nu(1 - H)} \right)^j \times e^{-\frac{\lambda H}{\nu(1 - H)}}, \quad j = 0, 1, 2, \dots$$

Из приведённого выражения следует, что для обобщённой модели контакт-центра значение среднего числа абонентов, повторяющих вызов, ограничено величиной $\frac{\lambda H}{\nu(1 - H)}$.

Сформулированные выше ограничения на изменение H , ν являются необходимыми условиями существования стационарного режима. Строгое доказательство достаточности сформулированных условий требует привлечения достаточно сложного аппарата теории вероятностей и здесь не рассматривается.

Обозначим через j число клиентов, находящихся в состоянии повторного запроса заявки на информационное обслуживание; через i обозначим число занятых линий доступа; через ℓ обозначим суммарное число занятых операторов и мест ожидания их освобождения; через ℓ_k обозначим суммарное число занятых консультантов в k -ой группе и мест ожидания их освобождения, $k = 1, 2, \dots, m$. В соответствии с особенностями функционирования контакт-центров, линия доступа остается занятой при обслуживании клиента устройством IVR, оператором или консультантом из k -ой группы, $k = 1, 2, \dots, m$, а также при ожидании начала обслуживания у оператора или у консультанта из k -ой группы, $k = 1, 2, \dots, m$. Пусть: $(j, i, \ell, \ell_1, \dots, \ell_m)$ — вектор состояния исследуемой модели контакт-центра. Обозначим через S множество возможных состояний исследуемой

модели. В пространство S включены состояния $(j, i, \ell, \ell_1, \dots, \ell_m)$, компоненты которого удовлетворяют следующим условиям:

$$j = 0, 1, \dots; \quad i = 0, 1, \dots, n; \quad \ell = 0, 1, \dots, v + w; \quad (2.3)$$

$$\ell_k = 0, 1, \dots, v_k + w_k; \quad k = 1, 2, \dots, m.$$

Дополнительным требованием принадлежности состояния $(j, i, \ell, \ell_1, \dots, \ell_m)$ пространству S является выполнение неравенства, следующего из условия занятости линии доступа в процессе обслуживания или ожидания начала обслуживания у оператора или консультанта

$$i \geq \ell + \ell_1 + \ell_2 + \dots + \ell_m. \quad (2.4)$$

Введем компоненты случайного процесса $r(t)$, описывающего динамику изменения во времени состояний модели. Обозначим через $j(t)$ число клиентов, находящихся в момент времени t в состоянии повторения запроса заявки на информационное обслуживание, через $i(t)$ обозначим число занятых в момент времени t линий доступа, через $\ell(t)$ обозначим число занятых в момент времени t операторов и мест ожидания их освобождения, через $\ell_k(t)$ обозначим число занятых в момент времени t консультантов из k -ой группы и мест ожидания их освобождения, $k = 1, 2, \dots, m$. Таким образом, динамика изменения состояний модели в зависимости от момента времени t описывается случайным процессом

$$r(t) = \left(j(t), i(t), \ell(t), \ell_1(t), \dots, \ell_m(t) \right),$$

принимающим значения в пространстве состояний S . Построенный процесс будет марковским поскольку все случайные величины, определяющие длительности времени пребывания модели в различных состояниях, имеют экспоненциальное распределение и не зависят друг от друга так же как и вероятности перехода из одного состояния модели в другое состояние. Обозначим через $p(j, i, \ell, \ell_1, \dots, \ell_m)$ значения вероятностей стационарных состояний модели. Они имеют интерпретацию доли времени пребывания модели в состоянии $(j, i, \ell, \ell_1, \dots, \ell_m)$. Используя данную интерпретацию, приведём формальные выражения для основных вероятностных характеристик обобщённой модели контакт-центра, которые можно использовать для анализа показателей качества обслуживания поступающих заявок, введенных в разделе 1.4, и планирования необходимого количества инфраструктурных элементов контакт-центра. К последним относятся: число линий доступа; число операторов; число консультантов; число мест ожидания; величина ограничения на время пребывания на ожидании, и т.д. Алгоритм оценки численных значений стационарных вероятностей основан на составлении и последующем решении системы уравнений статистического равновесия. Он будет рассмотрен в разделах 3.3–3.4. Допустим значения $p(j, i, \ell, \ell_1, \dots, \ell_m)$ известны. В следующем разделе будут приведены формальные выражения для оценки основных показателей качества обслуживания поступающих заявок.

2.3. Определение показателей качества обслуживания заявок

2.3.1. Систематизация характеристик

Соотношения для оценки значений показателей качества обслуживания поступающих заявок следуют из их физического смысла (см. раздел 1.4) и получаются либо через отношение интенсивностей появления анализируемых событий, либо через усреднение стационарных вероятностей состояний контакт-центра по соответствующим образом подобранным подмножествам состояний пространства S . Приведём словесное определение характеристик и их формальные выражения через значения стационарных вероятностей $p(j, i, \ell, \ell_1, \dots, \ell_m)$. В перечень характеристик, представляющих интерес при теоретическом анализе обобщённой модели контакт-центра и их практическом использовании, входят следующие четыре группы характеристик:

1. Средние значения переменных функциональных компонент. К ним относятся: число абонентов, повторяющих заявку; число занятых линий доступа, число занятых операторов и т.д.
2. Интенсивности поступающих, потерянных и обслуженных заявок, дифференцированные по этапам обслуживания.
3. Потери заявок и длительности ожидания начала обслуживания, дифференцированные по этапам обслуживания.
4. Интегральные показатели обслуживания заявок, отражающие в целом качество работы контакт-центра.

2.3.2. Средние значения функциональных компонент

Характеристики данной группы имеют самостоятельное значение или используются как вспомогательные компоненты в выражениях для других характеристик. В число анализируемых функциональных компонент входят все случайные переменные, участвующие в определении состояния $(j, i, \ell, \ell_1, \dots, \ell_m)$ исследуемой марковской модели контакт-центра. Обозначим через c число линий доступа, занятых абонентами, находящимися на обслуживании у оператора или консультанта или ожидающими начала соответствующего обслуживания. Понятно, что $c = \ell + \ell_1 + \dots + \ell_m$. В перечень характеристик первой группы входят:

- среднее число абонентов, повторяющих заявку по одной из следующих пяти причин отказа в обслуживании, различаемых в модели: из-за нехватки линий доступа; из-за занятости всех операторов и мест ожидания их освобождения; из-за превышения

времени ожидания освобождения оператора; из-за занятости всех консультантов выбранной группы и мест ожидания их освобождения; из-за превышения времени ожидания освобождения консультанта выбранной группы

$$M_r = \sum_{\{(j,i,\ell,\ell_1,\dots,\ell_m) \in S \mid j > 0\}} p(j, i, \ell, \ell_1, \dots, \ell_m) j;$$

- среднее число занятых линий доступа

$$M_a = \sum_{\{(j,i,\ell,\ell_1,\dots,\ell_m) \in S \mid i > 0\}} p(j, i, \ell, \ell_1, \dots, \ell_m) i;$$

- среднее число абонентов, обслуживаемых устройствами IVR,

$$M_i = \sum_{\{(j,i,\ell,\ell_1,\dots,\ell_m) \in S \mid i > c\}} p(j, i, \ell, \ell_1, \dots, \ell_m) (i - c);$$

- среднее число абонентов, находящихся на обслуживании у операторов,

$$M_\ell = \sum_{\{(j,i,\ell,\ell_1,\dots,\ell_m) \in S \mid \ell \leq v\}} p(j, i, \ell, \ell_1, \dots, \ell_m) \ell + \\ + \sum_{\{(j,i,\ell,\ell_1,\dots,\ell_m) \in S \mid \ell > v\}} p(j, i, \ell, \ell_1, \dots, \ell_m) v;$$

- среднее число абонентов, ожидающих освобождения оператора,

$$M_w = \sum_{\{(j,i,\ell,\ell_1,\dots,\ell_m) \in S \mid \ell > v\}} p(j, i, \ell, \ell_1, \dots, \ell_m) (\ell - v);$$

- среднее число абонентов, находящихся на обслуживании у консультантов k -ой группы, $k = 1, 2, \dots, m$,

$$M_{\ell,k} = \sum_{\{(j,i,\ell,\ell_1,\dots,\ell_m) \in S \mid 0 < \ell_k \leq v_k\}} p(j, i, \ell, \ell_1, \dots, \ell_m) \ell_k + \\ + \sum_{\{(j,i,\ell,\ell_1,\dots,\ell_m) \in S \mid \ell_k > v_k\}} p(j, i, \ell, \ell_1, \dots, \ell_m) v_k;$$

- среднее число абонентов, ожидающих освобождения консультантов k -ой группы, $k = 1, 2, \dots, m$,

$$M_{w,k} = \sum_{\{(j,i,\ell,\ell_1,\dots,\ell_m) \in S \mid \ell_k > v_k\}} p(j, i, \ell, \ell_1, \dots, \ell_m) (\ell_k - v_k).$$

2.3.3. Интенсивности потоков заявок на этапах обслуживания

Характеристики данной группы представляют собой усредненные значения интенсивностей потоков заявок, дифференцированные по результатам прохождения этапа обслуживания (получившие отказ или успешно завершившие обслуживание). В дальнейшем они будут использоваться для определения величины доли заявок, потерянных на каждом этапе обслуживания и для формирования законов сохранения интенсивностей потоков заявок, поступивших на определённый этап обслуживания и завершивших его. В перечень характеристик второй группы входят следующие показатели:

- суммарная интенсивность потоков первичных и повторных заявок, поступивших на линии доступа,

$$\Lambda_a = \sum_{\{j,i,\ell,\ell_1,\dots,\ell_m\} \in S} p(j,i,\ell,\ell_1,\dots,\ell_m)(\lambda + j\nu);$$

- суммарная интенсивность потоков первичных и повторных заявок, получивших отказ из-за занятости линий доступа,

$$\Lambda_{a,bl} = \sum_{\{j,i,\ell,\ell_1,\dots,\ell_m\} \in S | i=n} p(j,i,\ell,\ell_1,\dots,\ell_m)(\lambda + j\nu);$$

- интенсивность потока заявок, завершивших обслуживание в системе IVR и решивших продолжить обслуживание у оператора,

$$\Lambda_o = \sum_{\{j,i,\ell,\ell_1,\dots,\ell_m\} \in S | i>c} p(j,i,\ell,\ell_1,\dots,\ell_m)(i - c)\alpha_1q;$$

- интенсивность потока заявок, завершивших обслуживание в системе IVR и допущенных к обслуживанию у оператора,

$$\Lambda'_o = \sum_{\{j,i,\ell,\ell_1,\dots,\ell_m\} \in S | i>c, \ell < v+w} p(j,i,\ell,\ell_1,\dots,\ell_m)(i - c)\alpha_1q;$$

- суммарная интенсивность потока заявок, потерянных на доступе к оператору, которая складывается из интенсивности потока заявок, завершивших обслуживание в системе IVR, решивших продолжить обслуживание у оператора и получивших отказ из-за занятости всех операторов и мест ожидания их освобождения, и интенсивности потока заявок, получивших отказ в обслуживании из-за превышения допустимого времени ожидания освобождения оператора,

$$\Lambda_{o,bl} = \sum_{\{j,i,\ell,\ell_1,\dots,\ell_m\} \in S | i>c, \ell=v+w} p(j,i,\ell,\ell_1,\dots,\ell_m)(i - c)\alpha_1q +$$

$$+ \sum_{\{(j,i,\ell,\ell_1,\dots,\ell_m) \in S \mid \ell > v\}} p(j, i, \ell, \ell_1, \dots, \ell_m)(\ell - v)\sigma;$$

- интенсивность потока заявок, завершивших обслуживание у оператора и решивших продолжить обслуживание у консультанта k -ой группы, $k = 1, 2, \dots, m$,

$$\begin{aligned} \Lambda_{c,k} = & \sum_{\{(j,i,\ell,\ell_1,\dots,\ell_m) \in S \mid 0 < \ell \leq v\}} p(j, i, \ell, \ell_1, \dots, \ell_m) \ell \alpha_2 p p_k + \\ & + \sum_{\{(j,i,\ell,\ell_1,\dots,\ell_m) \in S \mid \ell > v\}} p(j, i, \ell, \ell_1, \dots, \ell_m) v \alpha_2 p p_k; \end{aligned}$$

- интенсивность потока заявок, завершивших обслуживание у оператора и допущенных к обслуживанию у консультанта k -ой группы,

$$\begin{aligned} \Lambda'_{c,k} = & \sum_{\{(j,i,\ell,\ell_1,\dots,\ell_m) \in S \mid 0 < \ell \leq v, \ell_k < v_k + w_k\}} p(j, i, \ell, \ell_1, \dots, \ell_m) \ell \alpha_2 p p_k + \\ & + \sum_{\{(j,i,\ell,\ell_1,\dots,\ell_m) \in S \mid \ell > v, \ell_k < v_k + w_k\}} p(j, i, \ell, \ell_1, \dots, \ell_m) v \alpha_2 p p_k; \end{aligned}$$

- суммарная интенсивность потока заявок, потерянных на доступе к консультантам k -ой группы, которая складывается из интенсивности потока заявок, завершивших обслуживание у оператора, решивших продолжить обслуживание у консультанта k -ой группы и получивших отказ из-за занятости консультантов k -ой группы и мест ожидания их освобождения, и интенсивности потока заявок, получивших отказ в обслуживании из-за превышения допустимого времени ожидания освобождения консультанта k -ой группы, $k = 1, 2, \dots, m$,

$$\begin{aligned} \Lambda_{c,k,bl} = & \sum_{\{(j,i,\ell,\ell_1,\dots,\ell_m) \in S \mid 0 < \ell \leq v, \ell_k = v_k + w_k\}} p(j, i, \ell, \ell_1, \dots, \ell_m) \ell \alpha_2 p p_k + \\ & + \sum_{\{(j,i,\ell,\ell_1,\dots,\ell_m) \in S \mid \ell > v, \ell_k = v_k + w_k\}} p(j, i, \ell, \ell_1, \dots, \ell_m) v \alpha_2 p p_k + \\ & + \sum_{\{(j,i,\ell,\ell_1,\dots,\ell_m) \in S \mid \ell_k > v_k\}} p(j, i, \ell, \ell_1, \dots, \ell_m) (\ell_k - v_k) \sigma_k. \end{aligned}$$

2.3.4. Потери и длительности ожидания на этапах обслуживания

Характеристики данной группы определяют величины доли заявок, поступивших на один из анализируемых этапов обслуживания и потерянных либо из-за занятости обслуживающих устройств и мест ожидания их освобождения, представленных на рассматриваемом этапе, либо в результате выполнения ограничения на допустимое время ожидания освобождения этих устройств. В перечень характеристик третьей группы входят следующие показатели:

- доля времени, в течение которого заняты все линии доступа (в следствии пуассоновского характера поступления первичных заявок эта характеристика совпадает с долей потерянных первичных заявок),

$$\pi_{t,a} = \sum_{\{(j,i,\ell,\ell_1,\dots,\ell_m) \in S \mid i=n\}} p(j,i,\ell,\ell_1,\dots,\ell_m);$$

- доля первичных и повторных заявок, потерянных из-за занятости линий доступа,

$$\pi_a = \frac{\Lambda_{a,bl}}{\Lambda_a};$$

- доля времени, в течение которого заняты все операторы и места ожидания их освобождения,

$$\pi_{t,o} = \sum_{\{(j,i,\ell,\ell_1,\dots,\ell_m) \in S \mid \ell=v+w\}} p(j,i,\ell,\ell_1,\dots,\ell_m);$$

- доля заявок, потерянных либо из-за занятости всех операторов и мест ожидания их освобождения, либо в результате выполнения ограничения на допустимое время ожидания освобождения операторов,

$$\pi_o = \frac{\Lambda_{o,bl}}{\Lambda_o};$$

- доля заявок, вставших на ожидание освобождения оператора,

$$p\{\mathcal{W}_o > 0\} = \frac{1}{\Lambda_o} \times \sum_{\{(j,i,\ell,\ell_1,\dots,\ell_m) \in S \mid i>c, v \leq \ell < v+w\}} p(j,i,\ell,\ell_1,\dots,\ell_m)(i-c)\alpha_1 q;$$

- среднее время нахождения заявки на ожидании начала обслуживания у оператора, найденное с использованием формулы Литтла, по результатам обслуживания всех заявок, допущенных к данной услуге,

$$W_o = \frac{M_w}{\Lambda'_o};$$

- доля времени, в течение которого заняты все консультанты k -ой группы и места ожидания их освобождения, $k = 1, 2, \dots, m$,

$$\pi_{t,k} = \sum_{\{(j,i,\ell,\ell_1,\dots,\ell_m) \in S \mid \ell_k = v_k + w_k\}} p(j, i, \ell, \ell_1, \dots, \ell_m);$$

- доля заявок, потерянных либо из-за занятости всех консультантов k -ой группы и мест ожидания их освобождения, либо в результате выполнения ограничения на допустимое время ожидания освобождения консультантов k -ой группы, $k = 1, 2, \dots, m$,

$$\pi_{c,k} = \frac{\Lambda_{c,k,bl}}{\Lambda_{c,k}};$$

- доля заявок, вставших на ожидание освобождения консультанта k -ой группы,

$$p\{\mathcal{W}_k > 0\} = \frac{\alpha_2 p p_k}{\Lambda_{c,k}} \times \left\{ \sum_{\{(j,i,\ell,\ell_1,\dots,\ell_m) \in S \mid 0 < \ell \leq v, v_k \leq \ell_k < v_k + w_k\}} p(j, i, \ell, \ell_1, \dots, \ell_m) \ell + \right. \\ \left. + \sum_{\{(j,i,\ell,\ell_1,\dots,\ell_m) \in S \mid \ell > v, v_k \leq \ell_k < v_k + w_k\}} p(j, i, \ell, \ell_1, \dots, \ell_m) v \right\};$$

- среднее время нахождения заявки на ожидании начала обслуживания у консультанта k -ой группы, найденное с использованием формулы Литтла, по результатам обслуживания всех заявок, допущенных к данной услуге,

$$W_k = \frac{M_{w,k}}{\Lambda'_{c,k}}.$$

2.3.5. Интегральные показатели качества обслуживания заявок

Характеристики данной группы с общих позиций описывают процесс поступления и обслуживания заявок в контакт-центре. Как правило, значения этих показателей используются при проведении мероприятий, направленных на повышение эффективности работы контакт-центра, и решении задач планирования (см. раздел 1.5) величины отдельных компонент его инфраструктуры. В перечень характеристик четвёртой группы входят:

- доля заявок, потерянных по разным причинам, от общего числа поступивших первичных и повторных заявок,

$$\pi_g = \frac{\Lambda_{a,bl} + \Lambda_{o,bl} + \sum_{k=1}^m \Lambda_{c,k,bl}}{\Lambda_a};$$

- доля заявок, отказавшихся от установления соединения,

$$\pi_r = \frac{(\Lambda_{a,bl} + \Lambda_{o,bl} + \sum_{k=1}^m \Lambda_{c,k,bl}) \cdot (1 - H)}{\lambda};$$

- доля повторных заявок в общем потоке заявок, поступивших на линии доступа,

$$\delta_r = \frac{M_r \nu}{\Lambda_a};$$

- среднее число повторных заявок на одну первичную,

$$M = \frac{M_r \nu}{\lambda};$$

- среднее число отказов на одно завершённое обслуживание.

$$Q = \frac{\Lambda_{a,bl} + \Lambda_{o,bl} + \sum_{k=1}^m \Lambda_{c,k,bl}}{M_i \alpha_1 (1 - q) + M_\ell \alpha_2 (1 - p) + \sum_{k=1}^m M_{\ell,k} \beta_k}.$$

2.4. Законы сохранения интенсивностей потоков заявок

2.4.1. Вывод законов сохранения

Важную роль в исследовании моделей систем связи с учётом возможности повторения заблокированной заявки играют законы сохранения, связывающие интенсивности поступающих и обслуженных системой потоков вызовов. Для их доказательства воспользуемся общей схемой, изложенной в [56, 58]. Умножим систему уравнений статистического равновесия, связывающую значения $p(j, i, \ell, \ell_1, \dots, \ell_m)$ (она построена и исследована в разделе 3.3), последовательно на значения $j, i, \ell, \ell_1, \dots, \ell_m$. После суммирования полученных соотношений и приведения подобных членов нетрудно установить справедливость следующих равенств:

$$M_r \nu = \left(\Lambda_{a,bl} + \Lambda_{o,bl} + \sum_{k=1}^m \Lambda_{c,k,bl} \right) \cdot H; \quad (2.5)$$

$$\Lambda_a = \Lambda_{a,bl} + M_i \alpha_1; \quad (2.6)$$

$$\Lambda_o = \Lambda_{o,bl} + M_\ell \alpha_2; \quad (2.7)$$

$$\Lambda_{c,k} = \Lambda_{c,k,bl} + M_{\ell,k} \beta_k, \quad k = 1, 2, \dots, m. \quad (2.8)$$

Полученные соотношения имеют характер законов сохранения интенсивностей потоков заявок, поступивших и обработанных в исследуемой модели контакт-центра на определённом этапе обслуживания.

- Соотношение (2.5) показывает, что интенсивность потока заявок, повторно требующих обслуживания в контакт-центре (выражение в левой части (2.5)), равна интенсивности потока отказов, приводящих к появлению повторных заявок (слагаемые из правой части (2.5)). Соотношение между интенсивностями потоков заявок, сформулированные в законе сохранения (2.5), поясняются на рисунке 2.3.

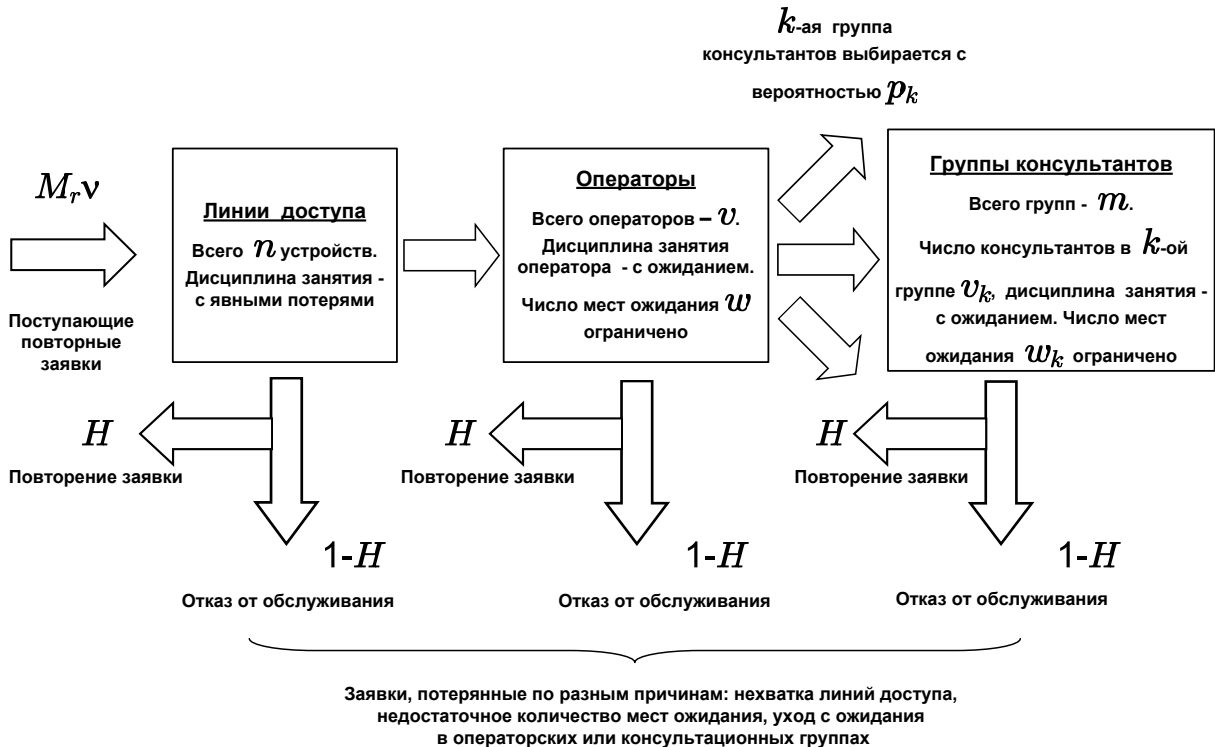


Рисунок 2.3. Соотношение между интенсивностями повторных вызовов и заблокированных заявок, инициирующих повторные запросы

- Соотношение (2.6) показывает, что интенсивность потока первичных и повторных заявок, поступивших на линии доступа (выражение в левой части (2.6)) равна сумме интенсивностей потока первичных и повторных заявок заблокированных на линиях доступа (первое слагаемое из правой части (2.6)) и интенсивности потока заявок, обслуженных устройствами IVR (второе слагаемое из правой части (2.6)).
- Соотношение (2.7) показывает, что интенсивность потока заявок, завершивших обслуживание на устройствах IVR и решивших продолжить обслуживание у оператора (выражение в левой части (2.7)) равна сумме интенсивностей потока заявок, заблокированных из-за занятости операторов и мест ожидания их освобождения или ушедших из системы из-за превышения допустимого времени ожидания начала обслуживания у оператора (первое слагаемое из правой части (2.7)) и интенсивно-

сти потока заявок, обслуженных операторами (второе слагаемое из правой части (2.7)).

- Соотношение (2.8) показывает, что интенсивность потока заявок, завершивших обслуживание у операторов и решивших продолжить обслуживание в k -ой группе консультантов (выражение в левой части (2.8)) равна сумме интенсивностей потока заявок, заблокированных из-за занятости k -ой группы консультантов и мест ожидания их освобождения или ушедших из системы из-за превышения допустимого времени ожидания начала обслуживания у консультантов (первое слагаемое из правой части (2.8)) и интенсивности потока заявок, обслуженных в k -ой группы консультантов, $k = 1, 2, \dots, m$ (второе слагаемое из правой части (2.8)).

Поскольку только q -ая часть потока заявок, завершивших обслуживание на IVR, продолжает обслуживание у операторов, то получаем соотношение

$$M_i \alpha_1 q = \Lambda_o. \quad (2.9)$$

Аналогично, поскольку только $p \cdot p_k$ -ая часть потока заявок, завершивших обслуживание у оператора продолжит обслуживание в k -ой группе консультантов, то получаем соотношение

$$M_\ell \alpha_2 p p_k = \Lambda_{c,k}. \quad (2.10)$$

Сложим соотношения (2.6)–(2.10). После приведения подобных членов нетрудно установить справедливость равенства

$$\Lambda_a = \Lambda_{a,bl} + \Lambda_{o,bl} + \sum_{k=1}^m \Lambda_{c,k,bl} + M_i \alpha_1 (1 - q) + M_\ell \alpha_2 (1 - p) + \sum_{k=1}^m M_{\ell,k} \beta_k. \quad (2.11)$$

Полученное равенство показывает, что интенсивность потока заявок, поступающих на линии доступа, равна сумме интенсивности потока заявок, заблокированных по разным причинам, и интенсивности потока заявок, завершивших обслуживание в контакт-центре на одном из этапов. На рисунке 2.4 показаны соотношения между интенсивностями поступления и обслуживания заявок, дифференцированные по этапам обслуживания. Эти соотношения формально представлены в законах сохранения (2.6)–(2.8) и соотношении (2.11).

Далее выражения (2.5)–(2.8) будут применяться для поиска соотношений между введёнными показателями обслуживания заявок (см. разделы 2.4.3 и 2.4.4) и при разработке точных и приближенных алгоритмов их оценки (см. разделы 4.2 и 4.3). В следующем разделе будет рассмотрена возможность использования законов сохранения для косвенной оценки сходимости итерационных методов решения систем уравнений равновесия.

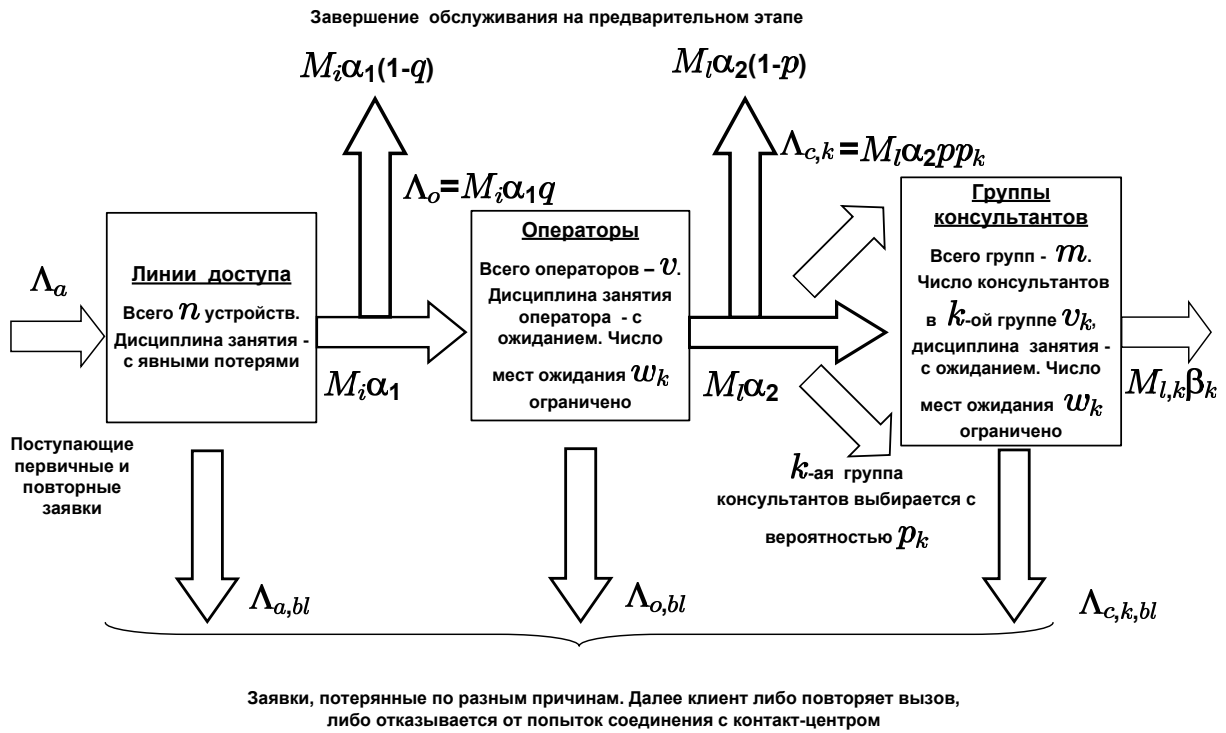


Рисунок 2.4. Соотношение между интенсивностями поступающих и заблокированных потоков заявок, использованные при записи законов сохранения (2.6)–(2.8) и соотношения (2.11)

2.4.2. Косвенная оценка сходимости итерационного метода

Основным методом оценки введённых характеристик обслуживания абонентов является численное решение системы уравнений статистического равновесия с использованием итерационных методов. Реализация этого подхода будет подробно рассмотрена в следующей главе. Используя итерационный метод Гаусса-Зейделя, можно решать системы уравнений равновесия из нескольких миллионов неизвестных. Из результатов теоретических исследований известно, что данный метод не всегда сходится, если речь идет о решении системы уравнений равновесия марковских моделей, используемых в теории телетрафика для анализа характеристик работы систем связи [58,69]. Тем не менее, в большинстве случаев сходимость имеет место. Её необходимо проверять косвенными методами. Для этого рекомендуется контролировать выполнение каких-либо известных теоретических соотношений. К ним, в частности, относятся законы сохранения (2.5)–(2.8). Таким образом, после достижения сходимости итерационного метода, выявленного с использованием традиционного критерия, основанного на степени близости нормированной разницы между последовательными значениями оцениваемой характеристики,

дополнительно проверяется выполнение равенств (2.5)–(2.8). Если они выполняются в пределах заданной точности представления численных значений характеристик и параметров модели, то этот результат может служить еще одним косвенным подтверждением сходимости итерационного метода. Строго говоря, сформулированный результат является необходимым условием сходимости, но не достаточным. Однако опыт проведения вычислительных работ показал, что выполнение указанных соотношений практически гарантирует сходимость итерационного метода с учётом условий, сформулированных выше.

2.4.3. Оценка интенсивности поступления первичных заявок

Законы сохранения интенсивностей поступающих и обслуженных заявок, сформулированные в виде соотношений (2.5)–(2.8), можно использовать для оценки величины λ интенсивности поступления первичных вызовов. Отметим, что прямая оценка значения λ затруднена из-за необходимости разделять в общем потоке поступающих заявок первичные и повторные вызовы. Предположим, что из результатов измерений известны значение Λ_a общей интенсивности поступающих заявок и величина π_g доли заявок потерянных по разным причинам. Значения обеих характеристик могут быть найдены с использованием стандартных измерительных средств. При проведении измерений этих характеристик не требуется отделять первичные вызовы от повторных. Из определения π_g получаем

$$\Lambda_{a,bl} + \Lambda_{o,bl} + \sum_{k=1}^m \Lambda_{c,k,bl} = \pi_g \cdot \Lambda_a. \quad (2.12)$$

Прибавим к левой и правой частям соотношения (2.5) величину λ . Получаем равенство

$$\lambda + \left(\Lambda_{a,bl} + \Lambda_{o,bl} + \sum_{k=1}^m \Lambda_{c,k,bl} \right) \cdot H = \Lambda_a. \quad (2.13)$$

Из (2.12), (2.13) следует выражение для оценки λ

$$\lambda = \Lambda_a \cdot (1 - \pi_g H). \quad (2.14)$$

2.4.4. Оценка показателей обслуживания заявок

С помощью законов сохранения (2.5)–(2.8) можно установить ряд зависимостей между введёнными характеристиками, которые упрощают их расчёт или измерение. В частности, значения некоторых интегральных показателей обслуживания заявок, рассмотренных в разделе 2.3.5, можно выразить через величину π_g и значение H . Приведем

соответствующие выражения. Для оценки значения π_r доли заявок, отказавшихся от установления соединения, получаем такую формулу

$$\pi_r = \frac{(\Lambda_{a,bl} + \Lambda_{o,bl} + \sum_{k=1}^m \Lambda_{c,k,bl}) \cdot (1 - H)}{\lambda} = \frac{\pi_g \Lambda_a (1 - H)}{\Lambda_a (1 - \pi_g H)} = \frac{\pi_g (1 - H)}{1 - \pi_g H}. \quad (2.15)$$

Для определения величины δ_r доли повторных заявок в общем потоке заявок, поступающих на линии доступа, имеем следующее выражение

$$\delta_r = \frac{M_r \nu}{\Lambda_a} = \frac{(\Lambda_{a,bl} + \Lambda_{o,bl} + \sum_{k=1}^m \Lambda_{c,k,bl}) \cdot H}{\Lambda_a} = \frac{\pi_g \Lambda_a H}{\Lambda_a} = \pi_g H. \quad (2.16)$$

Для вычисления значения M среднего числа повторных заявок на одну первичную, получаем такое соотношение

$$M = \frac{M_r \nu}{\lambda} = \frac{(\Lambda_{a,bl} + \Lambda_{o,bl} + \sum_{k=1}^m \Lambda_{c,k,bl}) \cdot H}{\lambda} = \frac{\pi_g H}{1 - \pi_g H}. \quad (2.17)$$

Воспользовавшись (2.11), нетрудно показать, что для оценки величины Q среднего числа отказов на одно завершённое обслуживание можно использовать следующее выражение

$$Q = \frac{\Lambda_{a,bl} + \Lambda_{o,bl} + \sum_{k=1}^m \Lambda_{c,k,bl}}{M_i \alpha_1 (1 - q) + M_\ell \alpha_2 (1 - p) + \sum_{k=1}^m M_{\ell,k} \beta_k} = \frac{\pi_g \Lambda_a}{\Lambda_a - \pi_g \Lambda_a} = \frac{\pi_g}{1 - \pi_g}. \quad (2.18)$$

Если дополнительно известно значение Λ_a общей интенсивности поступающих заявок, то можно оценить величину M_r среднего числа абонентов, повторяющих заявку. Действительно, из определения M_r и законов сохранения (2.5)–(2.8) получаем такое равенство

$$M_r = \frac{\Lambda_a \pi_g H}{\nu}. \quad (2.19)$$

Отметим, что формулы для косвенной оценки показателей обслуживания заявок и параметров их поступления справедливы только в рамках рассматриваемой модели. Важно также учитывать точность измеряемых характеристик.

2.5. Частные случаи построенной модели

2.5.1. Условия выбора моделей

Построенная модель в общем виде описывает процесс поступления и обслуживания заявок в действующих и перспективных контакт-центрах. Её частные случаи составляют семейство упрощённых моделей, которые можно использовать для анализа отдельных функциональных компонент контакт-центра. Будучи частными случаями эти модели обладают более простыми алгоритмами оценки показателей обслуживания заявок. Оценка числа линий доступа (устройств IVR) ведётся в рамках моделей с потерями и обычно не составляет большой проблемы, поскольку эти модели хорошо изучены и результаты

исследования широко представлены в многочисленных руководствах по теории теле-трафика [1,20,69,87,89,105]. Для их использования необходимо оценить интенсивность поступления первичных вызовов. Решение этой проблемы рассмотрено в разделе 2.4.3.

Для оценки влияния числа устройств IVR на численность операторов и консультантов следует исследовать действие этих функциональных компонент в совокупности. Анализ сформулированной задачи выполнен в разделе 4.6. Далее будут рассмотрены пять частных случаев построенной модели, которые можно использовать для оценки числа мест ожидания, операторов и групп консультантов в небольших контакт-центрах, имеющих более простую процедуру обслуживания поступающих заявок, или в больших контакт-центрах, но рассматриваемых в условиях, когда влиянием отдельных факторов можно пренебречь.

2.5.2. Отсутствие консультантов

Вначале упростим анализ квалификации обслуживающего персонала. Рассмотрим модель контакт-центра, в которой отсутствуют консультанты. Для этого в исходной модели достаточно положить $p = 0$. Пусть: j — число клиентов, находящихся в состоянии повторения заявки, i — число занятых линий доступа, ℓ — число занятых операторов и мест ожидания. Определение и последующая оценка показателей обслуживания заявок выполняется с использованием значений стационарных вероятностей упрощенной модели $P(j, i, \ell)$. Таким образом, расчет характеристик сводится к решению системы уравнений равновесия. Поскольку она не обладает какими-либо специальными свойствами, то ее проще всего решить методом Гаусса-Зейделя. В рассматриваемом случае число неизвестных существенно меньше, чем в исходной модели, поэтому оценку показателей можно выполнить для большинства практически интересных случаев. Схема функционирования введённой модели контакт-центра показана на рисунке 2.5.

2.5.3. Отсутствие операторского обслуживания

Другим интересным для практики частным случаем является модель контакт-центра, в которой отсутствует предварительное операторское обслуживание. Абонент, завершив обслуживание на IVR, с заданной вероятностью продолжает его в выбранной группе консультантов. Данная модель следует из исходной, если положить $\alpha_1 = \infty$, $q = 1$. Схема функционирования введённой модели контакт-центра показана на рисунке 2.6. Определение показателей обслуживания заявок и алгоритм их оценки аналогичны исходной модели и здесь не приводятся.

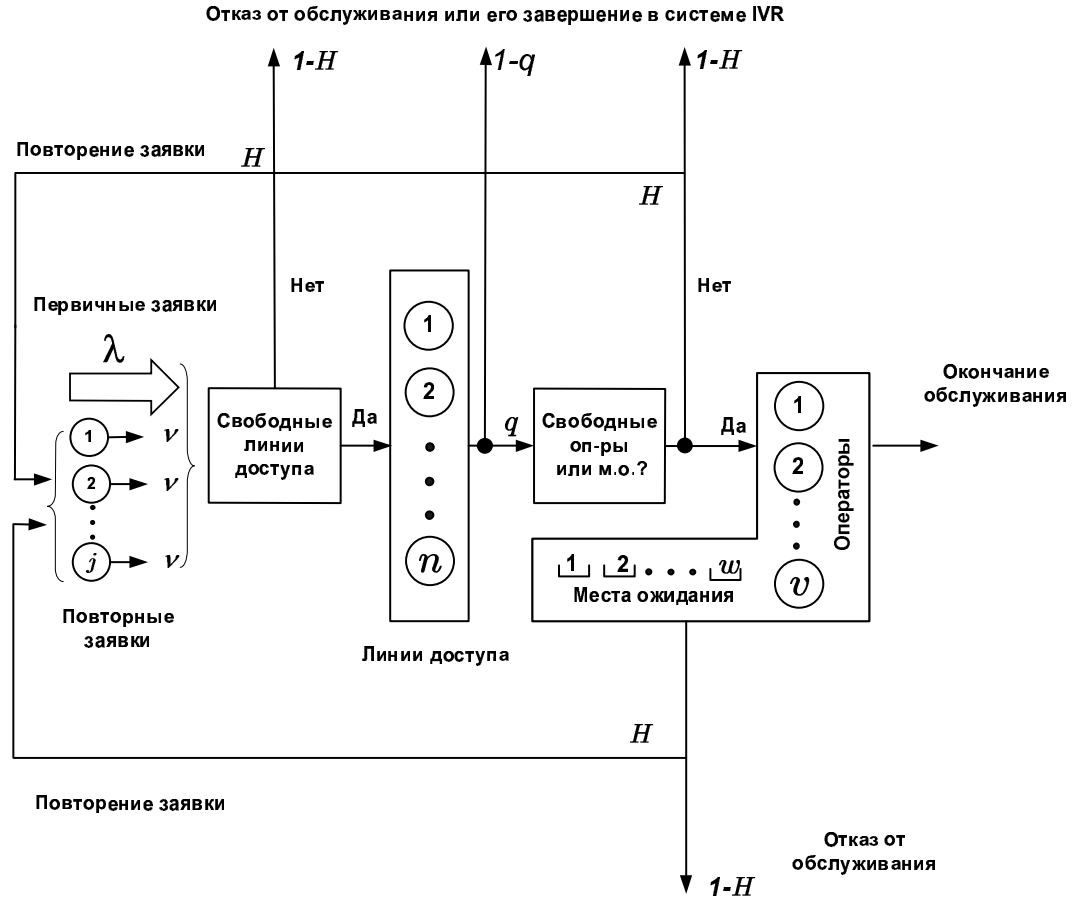


Рисунок 2.5. Модель контакт-центра без дифференциации операторов по квалификации

2.5.4. Учет ограниченного ожидания

Перейдём к анализу частных случаев, которые можно использовать как отдельные расчётные модули при разработке приближенных методов анализа обобщённой модели контакт-центра, основанных на асимптотических свойствах её характеристик, или на реализации техники декомпозиции. Эти результаты будут получены, соответственно, в разделах 4.2 и 4.3. Начнем с наиболее известной модели, в которой рассматривается процесс обслуживания поступающих заявок группой операторов. Для заявок, получивших отказ предусматривается возможность ожидания освобождения оператора. Время пребывания на ожидании ограничено. Данную модель можно построить специальным выбором значений входных параметров. Для этого достаточно положить $\alpha_1 = \infty$ (при проведении вычислений можно взять $\alpha_1 = 10^{10}$), $q = 1$, $p = 0$, $H = 0$. Первое и второе из приведённых соотношений устраняет из рассмотрения линии доступа и устройства IVR, третье — возможность продолжения обслуживания у консультантов, четвертое — возможность повторения заблокированной заявки. Построим модель контакт-центра в анализируемом частном случае и приведём необходимые расчётные выражения, воспользовавшись результатами [1,20,69,87,89,105].

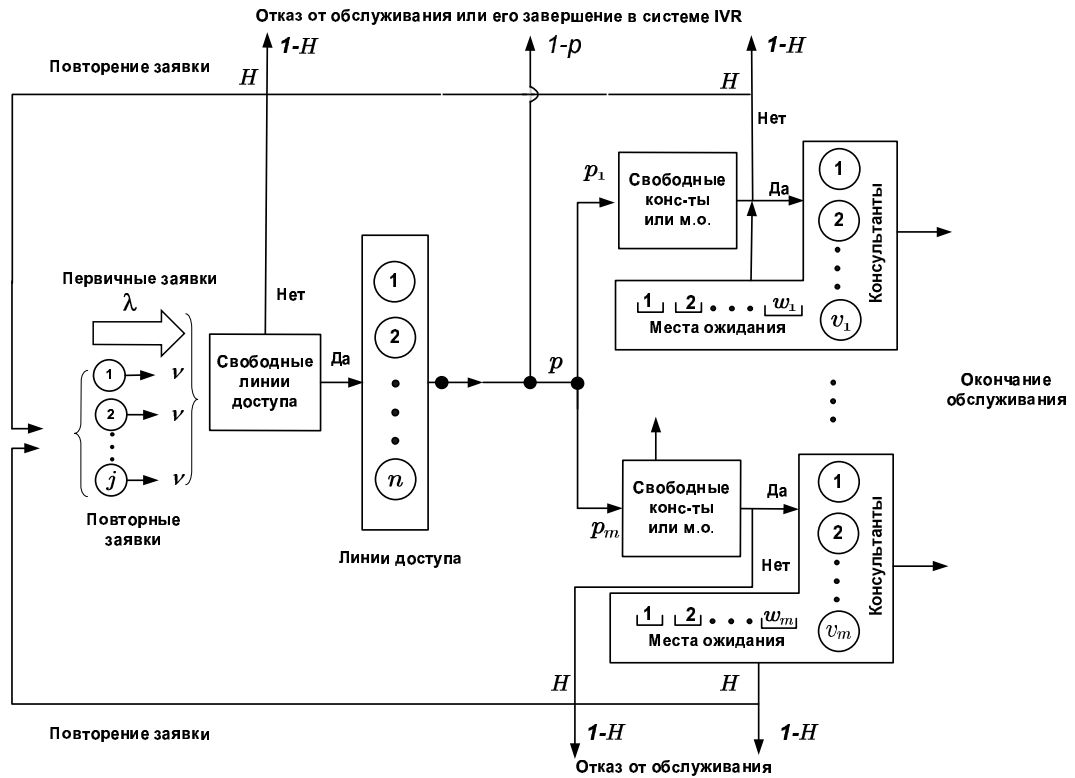


Рисунок 2.6. Модель контакт-центра без предварительного операторского обслуживания

Рассмотрим систему, состоящую из v операторов, на которую поступает пуассоновский поток заявок с интенсивностью λ . Время обслуживания заявки имеет экспоненциальное распределение с параметром α . Если в момент поступления заявки свободных операторов нет, то она становится на ожидание (занимает линию доступа). Число мест ожидания обозначим через w . Время ожидания ограничено случайной величиной, имеющей экспоненциальное распределение, с параметром σ . Если по завершении этого времени заявка не попала на обслуживание, то она покидает систему. Будем считать, что заявки выбираются на обслуживание в порядке постановки в очередь (первый пришел, первый на обслуживание). Схема функционирования введённой модели показана на рисунке 2.7.

Построенная модель описывается процессом рождения и гибели, поэтому расчёт её характеристик не вызывает больших затруднений. Воспользовавшись упрощенной моделью, выпишем оценки для основных характеристик обслуживания поступающих заявок. Обозначим идентичные по смыслу характеристики упрощённой и обобщённой моделей контакт-центра одинаковыми символами. Пусть $p(i)$ — стационарная вероятность занятости i операторов и мест ожидания в исследуемой упрощённой модели. Основные показатели качества обслуживания заявок из перечня, введённого в разделе 2.3, находятся

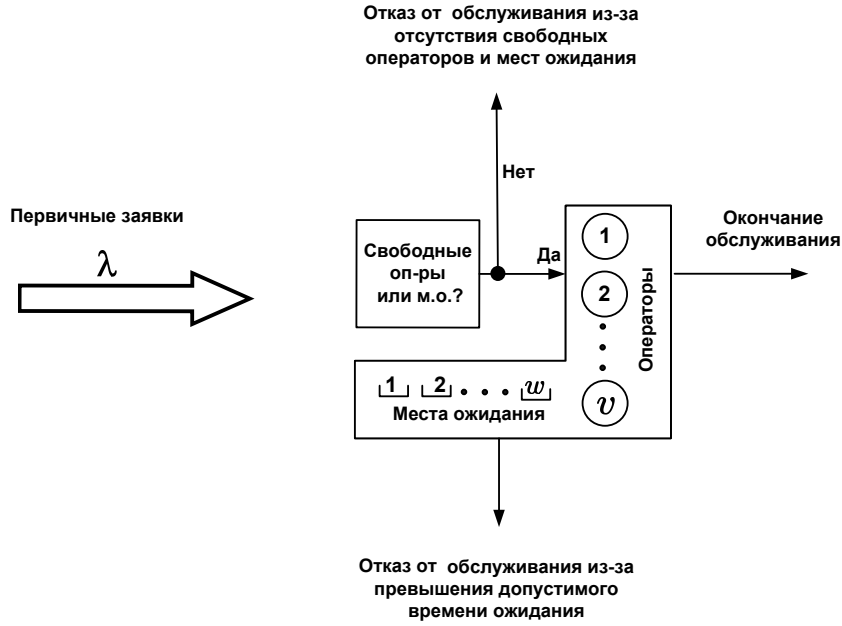


Рисунок 2.7. Модель контакт-центра с ограниченными числом мест и временем ожидания

из следующих соотношений:

$$M_\ell = \sum_{i=1}^v p(i)i + v \sum_{i=v+1}^{v+w} p(i); \quad M_w = \sum_{i=v+1}^{v+w} p(i)(i-v); \quad (2.20)$$

$$\Lambda_o = \lambda; \quad \Lambda'_o = \lambda(1 - p(v+w)); \quad \Lambda_{o,bl} = \lambda p(v+w) + \sigma \sum_{i=v+1}^{v+w} p(i)(i-v);$$

$$\pi_{t,o} = p(v+w); \quad \pi_g = \pi_o = p(v+w) + \frac{\sigma}{\lambda} \times \sum_{i=v+1}^{v+w} p(i)(i-v);$$

$$p\{\mathcal{W} > 0\} = \sum_{i=v}^{v+w-1} p(i); \quad W_o = \frac{M_w}{\Lambda'_o}.$$

Чтобы воспользоваться введёнными определениями, необходимо найти значения $p(i)$. Обозначим здесь и далее ненормированные значения стационарных вероятностей прописной буквой $P(\cdot)$. Соответственно, строчными буквами $p(\cdot)$ обозначим нормированные величины стационарных вероятностей. Значения $P(i)$ связаны системой уравнений статистического равновесия, имеющей вид

$$P(0)\lambda = P(1)\alpha; \quad (2.21)$$

$$P(1)(\lambda + \alpha) = P(0)\lambda + P(2)2\alpha;$$

$$P(v-1)(\lambda + (v-1)\alpha) = P(v-2)\lambda + P(v)v\alpha;$$

$$P(v)(\lambda + v\alpha) = P(v-1)\lambda + P(v+1)(v\alpha + \sigma);$$

$$P(v+1)(\lambda + v\alpha + \sigma) = P(v)\lambda + P(v+2)(v\alpha + 2\sigma);$$

$$P(v+w-1)(\lambda + v\alpha + (w-1)\sigma) = P(v+w-2)\lambda + P(v+w)(v\alpha + w\sigma);$$

$$P(v+w)(v\alpha + w\sigma) = P(v+w-1)\lambda.$$

Для $P(i)$ выполнено условие нормировки

$$\sum_{i=0}^{v+w} P(i) = 1.$$

Рассмотрев (2.21) последовательно при $i = 0, 1, \dots, v+w-1$, получаем рекуррентные формулы для вычисления $P(i)$

$$P(i)\lambda = P(i+1)(i+1)\alpha, \quad i = 0, 1, \dots, v-1; \quad (2.22)$$

$$P(i)\lambda = P(i+1)(v\alpha + (i+1-v)\sigma), \quad i = v, v+1, \dots, v+w-1.$$

Используя (2.22), величины $P(i)$, $i = 1, 2, \dots, v+w$ можно выразить через $P(0)$. Получаем такие выражения

$$P(i) = \begin{cases} P(0) \left(\frac{\lambda}{\alpha}\right)^i \frac{1}{i!}, & i = 0, 1, \dots, v; \\ P(0) \frac{\lambda^i}{\alpha^v v! (v\alpha + \sigma) \dots (v\alpha + (i-v)\sigma)}, & i = v+1, \dots, v+w. \end{cases} \quad (2.23)$$

Соотношения (2.23) дают возможность найти явные выражения для основных характеристики модели через величины входных параметров. Однако они не всегда удобны для проведения вычислений, особенно в области больших значений v . Построим алгоритм вычисления $p(i)$, $i = 0, 1, \dots, v$, свободный от указанного недостатка. Последовательность расчётов выглядит следующим образом:

1. Положим $P(0) = 1$.
2. Используя (2.22) находим рекурсивные выражения для оценки значений $P(i)$ через величину $P(0)$

$$P(i+1) = P(i) \cdot \frac{\lambda}{\alpha(i+1)}, \quad i = 0, 1, \dots, v-1, \quad (2.24)$$

$$P(i+1) = P(i) \cdot \frac{\lambda}{v\alpha + (i+1-v)\sigma}, \quad i = v, v+1, \dots, v+w-1.$$

3. Находим величину нормировочной константы $N = \sum_{i=0}^{v+w} P(i)$ и нормированные значения вероятностей всех состояний модели

$$p(i) = \frac{P(i)}{N}, \quad i = 0, 1, \dots, v + w.$$

Построенный алгоритм даёт возможность рассчитать нормированные значения $p(i)$, а с ними величины характеристик обслуживания заявок в исследуемой модели контакт-центра. Необходимые расчётные выражения задаются равенствами (2.20).

Несложная структура исследуемой модели позволяет получить явные выражения для характеристик модели через значения входных параметров (см., например, [1, 20, 69, 87, 89, 105]). Приведём искомые соотношения для части характеристик, перечисленных в (2.20). Величины оставшихся характеристик находятся исходя из их определений и приведённых ниже формул. Для простоты, в последующих выражениях предполагается выполненным соотношение $\lambda \neq v\alpha$ и используется вспомогательная переменная $a = \frac{\lambda}{\alpha}$. Расчётные выражения имеют следующий вид:

$$p\{\mathcal{W} > 0\} = \frac{v \left(1 - \left(\frac{a}{v}\right)^w\right)}{\frac{v-a}{E(v,a)} + a \left(1 - \left(\frac{a}{v}\right)^w\right)}; \quad p(v+w) = \frac{(v-a) \left(\frac{a}{v}\right)^w}{\frac{v-a}{E(v,a)} + a \left(1 - \left(\frac{a}{v}\right)^w\right)}; \quad (2.25)$$

$$M_\ell = a \left(1 - p(v+w)\right);$$

$$M_w = \lambda p\{\mathcal{W} > 0\} \times \left(\frac{1}{v-a} - \frac{w \left(\frac{a}{v}\right)^w}{v \left(1 - \left(\frac{a}{v}\right)^w\right)} \right) \times \frac{1}{\alpha};$$

$$W_o = \frac{p\{\mathcal{W} > 0\}}{1 - p(v+w)} \times \left(\frac{1}{v-a} - \frac{w \left(\frac{a}{v}\right)^w}{v \left(1 - \left(\frac{a}{v}\right)^w\right)} \right) \times \frac{1}{\alpha}.$$

$$\pi_g = \pi_o = p(v+w) + \frac{\sigma}{\lambda} \times M_w;$$

В соотношениях (2.25) символ $E(v, a)$ обозначает величину формулы Эрланга для неподступной системы из v каналов с интенсивностью предложенного трафика a .

Простая структура исследуемой модели позволяет в явном виде найти функцию распределения времени пребывания заявки в очереди и тем самым рассчитать долю заявок, которым придётся ожидать начала обслуживания сверх заданного промежутка времени. Эта характеристика имеет большое значение при планировании систем связи с ожиданием, в частности, для оценки требуемого числа операторов и линий доступа в контакт-центрах. Приведём вид функции в случае, когда число мест ожидания неограничено и

в модели не учитывается ограничение на максимально возможное время пребывания на ожидании. Искомая функция записывается следующим образом:

$$p\{\mathcal{W} \leq t\} = 1 - p\{\mathcal{W} > 0\} \cdot e^{-v\alpha t(1-\frac{a}{v})} = 1 - D(v, a) \cdot e^{-v\alpha t(1-\frac{a}{v})}. \quad (2.26)$$

В соотношениях (2.26) символ $D(v, a)$ обозначает величину второй формулы Эрланга для полнодоступной системы из v каналов с интенсивностью предложенного трафика a и ожиданием начала обслуживания в случае блокировки. Для расчёта значения $D(v, a)$ используется следующее соотношение

$$D(v, a) = \frac{v E(v, a)}{v - a(1 - E(v, a))}. \quad (2.27)$$

В общем случае определение явного вида функции распределения времени ожидания представляет собой очень сложную задачу. Её решение получено для нескольких несложных обобщений модели Эрланга с ожиданием. Отметим модель с ограниченным числом мест ожидания [87]. В более сложных ситуациях используется имитационное моделирование или адаптация входных параметров модели для реализации возможности применения явных выражений (2.26).

Таким образом, для модели с ожиданием оценка показателей обслуживания заявок не вызывает больших затруднений и сводится к использованию простых рекуррентных зависимостей (2.22). Назовем построенную схему обслуживания заявок 1-ой вспомогательной моделью контакт-центра. Она будет активно использоваться при построении приближенных процедур расчёта обобщённой модели, основанных на пуассоновской замене потока повторных вызовов и реализации техники декомпозиции. Эти результаты будут получены, соответственно, в разделах 4.2 и 4.3. Теперь учтём в рассмотренной модели возможность повторения заявки в ситуации отказа в обслуживании из-за занятости операторов и мест ожидания или превышения допустимого времени ожидания начала обслуживания. Эта модель также является частным случаем исследуемой обобщённой модели контакт-центра. Её можно получить специальным выбором значений входных параметров. Для этого достаточно положить $\alpha_1 = \infty$ (при проведении вычислений можно взять $\alpha_1 = 10^{10}$), $q = 1$, $p = 0$. Первое и второе из приведенных соотношений устраняет из рассмотрения линии доступа и устройства IVR, третье — возможность продолжения обслуживания у консультантов. Построим исследуемую модель контакт-центра и, воспользовавшись результатами [57,58], приведём необходимые расчётные формулы.

2.5.5. Учёт ограниченного ожидания и возможности повторного запроса

Рассмотрим систему, состоящую из v операторов, на которую поступает пуассоновский поток заявок интенсивности λ . Время обслуживания заявки имеет экспоненциальное распределение с параметром α . Если в момент поступления заявки свободных

операторов нет, то она становится на ожидание (занимает линию доступа). Число мест ожидания обозначим через w . Время ожидания ограничено случайной величиной, имеющей экспоненциальное распределение с параметром σ . Если заняты все операторы и места ожидания или неудачно завершилось время ожидания начала обслуживания, то абонент с вероятностью H повторяет заявку через случайное время, имеющее экспоненциальное распределение с параметром ν , а с дополнительной вероятностью $1 - H$ покидает систему отказавшись от попыток установить соединение. Схема функционирования введённой модели контакт-центра показана на рисунке 2.8.

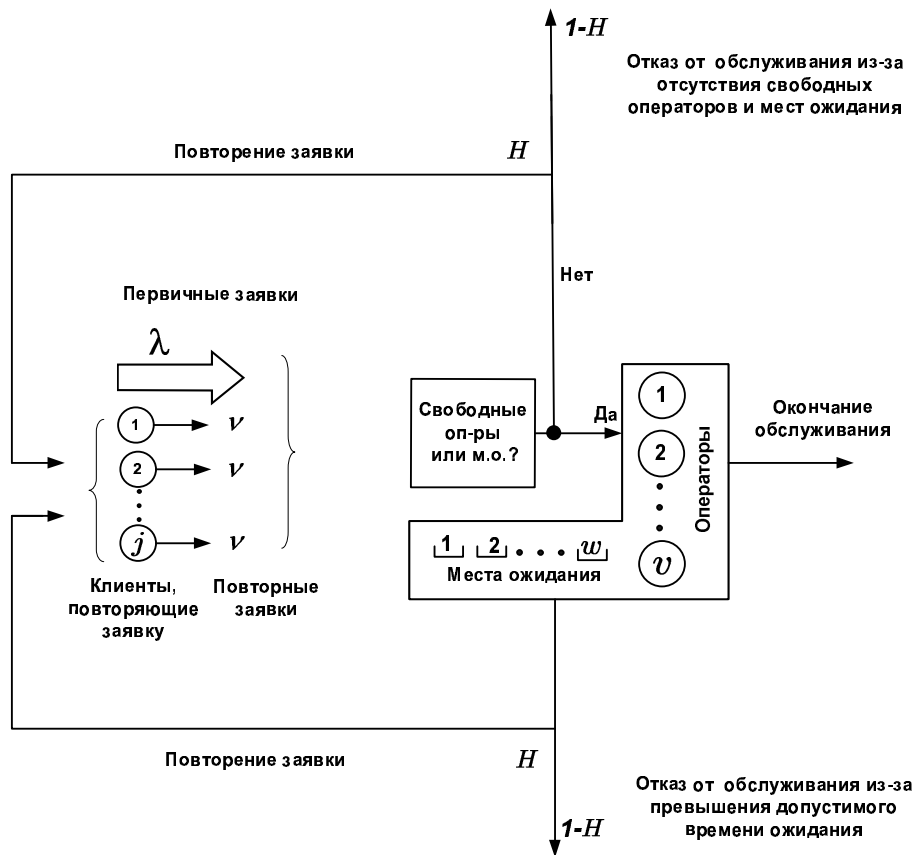


Рисунок 2.8. Модель контакт-центра с учётом возможности ожидания и повторения заблокированной заявки

Построенная модель взаимодействия абонента с контакт-центром описывается марковским процессом $r(t) = (j(t), i(t))$, где $j(t)$ — число абонентов, повторяющих в момент времени t вызов по причине занятости всех операторов и мест ожидания, а также из-за неудачного завершения времени ожидания начала обслуживания, а $i(t)$ — суммарное число занятых в момент времени t мест ожидания и операторов. Обозначим через (j, i) значения принимаемые процессом $r(t)$. Здесь: j — число абонентов, повторяющих вызов, а i — суммарное число занятых мест ожидания и операторов. Через S обозначим множество всех значений, принимаемых процессом $r(t)$. Состояние $(j, i) \in S$ для вели-

чин j, i , удовлетворяющих условиям: $j = 0, 1, 2, \dots$ и $i = 0, 1, \dots, v + w$. Будем считать выполненными достаточные условия существования стационарного режима у процесса $r(t)$. Они формулируются из тех же соображений, что были использованы при исследовании обобщённой модели контакт-центра (см. раздел 2.2.3) и накладывают следующие ограничения на значения входных параметров модели: $H < 1$ и $\nu > 0$. Обозначим через $p(j, i)$ величину стационарной вероятности состояния (j, i) . Она имеет интерпретацию доли времени пребывания модели в состоянии (j, i) . Этот результат имеет большое значение для практических приложений, поскольку позволяет рассчитать значения основных стационарных показателей качества обслуживания поступающих заявок.

Вероятности $p(j, i)$ определяются с помощью решения системы уравнений статистического равновесия процесса $r(t)$, которая является частным случаем системы уравнений равновесия исследуемой модели, рассмотренной в следующей главе. Выглядит она существенно проще, поэтому получим и приведем её вид. Приравняв интенсивности входа и выхода в каждое из состояний $(j, i) \in S$, получаем следующую бесконечную систему уравнений, связывающую ненормированные значения стационарных вероятностей $P(j, i)$

$$\begin{aligned}
& \left(\lambda I(i < v + w) + \lambda H I(i = v + w) + \right. & (2.28) \\
& + j \nu I(i < v + w) + j \nu (1 - H) I(i = v + w) + \\
& \left. + (i - v) \sigma I(i > v) + i \alpha I(i \leq v) + v \alpha I(i > v) \right) P(j, i) = \\
& = \lambda I(i > 0) P(j, i - 1) + \\
& + \lambda H I(j > 0, i = v + w) P(j - 1, i) + \\
& + (j + 1) \nu I(i > 0) P(j + 1, i - 1) + \\
& + (j + 1) \nu (1 - H) I(i = v + w) P(j + 1, i) + \\
& + (i + 1 - v) \sigma H I(j > 0, v \leq i < v + w) P(j - 1, i + 1) + \\
& + \left((i + 1 - v) \sigma (1 - H) I(v \leq i < v + w) + \right. \\
& \left. + (i + 1) \alpha I(i < v) + v \alpha I(v \leq i < v + w) \right) P(j, i + 1),
\end{aligned}$$

где

$$j = 0, 1, \dots; \quad i = 0, 1, \dots, v + w.$$

Для $P(j, i)$ выполнено условие нормировки

$$\sum_{j=0}^{\infty} \sum_{i=0}^{v+w} P(j, i) = 1.$$

Приведём формальные определения основных показателей обслуживания заявок. Они следуют из соотношений, сформулированных в разделе 2.3 и имеют следующий вид:

$$M_r = \sum_{j=0}^{\infty} \sum_{i=0}^{v+w} P(j, i)j; \quad (2.29)$$

$$M_\ell = \sum_{j=0}^{\infty} \left[\sum_{i=0}^v P(j, i)i + v \sum_{i=v+1}^{v+w} P(j, i) \right];$$

$$M_w = \sum_{j=0}^{\infty} \sum_{i=v+1}^{v+w} P(j, i)(i - v);$$

$$\Lambda_o = \lambda + M_r\nu; \quad \Lambda'_o = \sum_{j=0}^{\infty} \sum_{i=0}^{v+w-1} P(j, i)(\lambda + j\nu);$$

$$\Lambda_{o,bl} = \sum_{j=0}^{\infty} P(j, v+w)(\lambda + j\nu) + \sum_{i=v+1}^{v+w} P(j, i)(i - v)\sigma;$$

$$\pi_{t,o} = \sum_{j=0}^{\infty} P(j, v+w); \quad \pi_g = \pi_o = \frac{\Lambda_{o,bl}}{\Lambda_o};$$

$$p\{\mathcal{W}_o > 0\} = \frac{1}{\Lambda_o} \times \sum_{j=0}^{\infty} \sum_{i=v}^{v+w-1} P(j, i)(\lambda + j\nu);$$

$$W_o = \frac{M_w}{\Lambda'_o}; \quad \pi_r = \frac{\Lambda_{o,bl}(1 - H)}{\Lambda_o}; \quad \delta_r = \frac{M_r\nu}{\Lambda_o};$$

$$M = \frac{M_r\nu}{\lambda}; \quad Q = \frac{\Lambda_{o,bl}}{M_\ell\alpha}.$$

Законы сохранения (2.5)–(2.8), полученные для исходной модели, в данном частном случае имеют вид:

$$M_r\nu = \Lambda_{o,bl}H, \quad (2.30)$$

$$\lambda + M_r\nu = \Lambda_{o,bl} + M_\ell\alpha. \quad (2.31)$$

Таким образом, для модели с ожиданием и возможностью повторения заблокированной заявки, точная оценка показателей качества их обслуживания сводится к решению системы уравнений равновесия. В общем виде она не обладает какими-либо специальными свойствами. По этой причине искомое решение проще всего получить с помощью итерационной процедуры Гаусса-Зейделя. Особенности её реализации будут рассмотрены

в следующей главе. Для данной модели можно построить эффективные приближённые алгоритмы оценки характеристик, основанные на замене потока повторных заявок на пуассоновский с неизвестной интенсивностью, вычисляемой из решения неявного уравнения, вытекающего из соотношения (2.30). В этом случае, значения (2.29) оцениваются с использованием модели, построенной в разделе 2.5.4. Если в модели отсутствует возможность повторения заблокированной заявки после отказа в обслуживании, вызванного превышением допустимого времени ожидания, то решение всей системы уравнений разбивается на решение отдельных подсистем с размерностью на единицу большей числа операторов. Число подсистем на единицу больше, чем ограничение на максимальное число абонентов, одновременно повторяющих вызов. Эти подходы подробно исследованы в [14,15,19,58,103]. Назовем построенную процедуру обслуживания заявок 2-ой вспомогательной моделью контакт-центра.

2.5.6. Отсутствие возможности повторного запроса

Последним из рассмотренных частных случаев является модель контакт-центра, которая отличается от обобщённой модели отсутствием возможности повторения вызова. Для её формального построения достаточно положить в обобщённой модели вероятность повторения вызова H равной нулю. Назовем эту модель обслуживания заявок 3-ой вспомогательной или базовой моделью контакт-центра. Определение показателей обслуживания заявок и алгоритм их оценки аналогичен исходной модели и здесь не рассматривается.

2.6. Выводы по результатам второй главы

1. Выполненное исследование процесса функционирования современных контакт-центров показало целесообразность использования устройств IVR, операторов и консультантов для обслуживания клиентов. Построена новая обобщенная математическая модель контакт-центра, в которой учтена эта особенность работы справочных служб, а также возможность повторения заявки из-за отказов на различных этапах обслуживания, наличие мест ожидания начала обслуживания у оператора или консультанта, наличие ограничения на время пребывания на ожидании и возможность поступления заявки на обслуживание по каналам сети Интернет.
2. С использованием разработанной модели получены выражения для основных показателей качества обслуживания поступающих заявок. Формальные выражения для характеристик получены через значения входных параметров и величины стационарных вероятностей марковского процесса, описывающего функционирование модели. Среди них: интенсивности и потери заявок, дифференцированные по этапам обслуживания; средние значения функциональных компонент состояния (среднее

число абонентов, повторяющих заявку, среднее число занятых операторов и т.д.); вероятности и средние длительности пребывания заявок на ожидании начала обслуживания и т.д.

3. В рамках построенной модели получены соотношения между основными характеристиками работы контакт-центра, которые имеют характер законов сохранения интенсивностей поступающих и обслуженных системой потоков заявок. Показано использование найденных соотношений для косвенной оценки сходимости итерационного метода решения системы уравнений равновесия и для разработки приближенных методов оценки показателей качества обслуживания заявок.
4. Законы сохранения позволили выполнить косвенную оценку значений тех входных параметров и характеристик модели, прямое измерение которых стандартной аппаратурой вызывает затруднение из-за сложностей в разделении первичных и повторных вызовов. В частности, получены выражения для оценки значений интенсивности поступления первичных вызовов, среднего числа повторных вызовов на один первичный, вероятности ухода абонента без обслуживания, среднего числа отказов на одно установленное соединение через величины общей доли потерянных заявок и общей интенсивности их поступления.
5. Фиксируя значения отдельных параметров, получены несколько частных случаев обобщенной модели контакт-центра, которые имеют важное самостоятельное значение. В их числе: модель без дифференциации операторов по квалификации; модель без предварительного операторского обслуживания; модель с ограниченным числом мест и временем ожидания; модель с ожиданием и повторением заблокированной заявки; модель без учёта повторных запросов. Показано, что в каждом из рассмотренных случаев оценка показателей обслуживания заявок требует существенно меньше вычислительных усилий, чем это необходимо для обобщенной модели контакт-центра. Предложенные модели будут использованы для расчёта характеристик несложных по структуре моделей контакт-центров, для анализа его отдельных функциональных компонент (числа мест ожидания, числа операторов и т.д.) и для разработки приближённых алгоритмов расчёта обобщенной модели.

Глава 3

Оценка показателей качества обслуживания заявок

3.1. Введение

Математическая модель контакт-центра, построенная в предыдущей главе, имеет достаточно громоздкий вид, который следует из необходимости учёта всех особенностей формирования и обслуживания поступающих заявок в современных справочно-информационных службах. Помимо имитационного моделирования, единственным реальным способом оценки показателей обслуживания заявок и других характеристик работы контакт-центров является применение алгоритмов, основанных на составлении и последующем решении системы линейных уравнений статистического равновесия методами вычислительной математики. Сформулированную задачу можно решать либо для всей модели, рассматривая её как единое целое, и находить точные значения показателей качества обслуживания заявок, либо для какой-то её части и затем использовать результаты вычислений в качестве основы приближённых алгоритмов, построенных на принципах декомпозиции исходной модели на отдельные сегменты. Обе из перечисленных возможностей будут исследованы в этой и последующей главах диссертации.

В разделе 3.2 выполнен анализ основных методов решения систем уравнений равновесия и сформулированы принципы построения приближённых процедур оценки характеристик моделей систем связи. В разделе 3.3 кратко перечислены положения итерационного метода Гаусса-Зейделя. В следующем разделе получено алгебраическое представление системы уравнений равновесия построенной модели в виде удобном для последующего решения алгоритмом Гаусса-Зейделя. В разделе 3.4 исследована зависимость введённых показателей качества обслуживания заявок от изменения входных параметров построенной модели контакт-центра. Выявлены свойства, которые могут быть использованы для построения приближённых алгоритмов оценки её характеристик. В заключительном разделе сформулированы выводы по главе.

3.2. Точные и приближённые методы расчёта модели

Эффективность расчётных методов, используемых в теории телетрафика, во многом определяется лёгкостью представления вычислительных алгоритмов на языках программирования и универсальным характером рассматриваемых подходов. Отмеченным свойством обладают алгоритмы оценки характеристик, основанные на составлении и последующем решении системы линейных уравнений статистического равновесия. Данный подход позволяет рассчитывать модели с числом состояний в несколько миллионов. Время расчёта обычно составляет несколько минут и постоянно уменьшается с ростом быстродействия компьютеров. Поскольку система уравнений равновесия представляет собой систему линейных уравнений, то теоретической базой расчётных алгоритмов являются положения и методы линейной алгебры. В вычислительной математике получено достаточно много решений сформулированной задачи. Выбор конкретной расчётной схемы существенно зависит от свойств матрицы системы уравнений равновесия, которая в матричной форме имеет вид:

$$AP = 0. \tag{3.1}$$

В (3.1) символ A — квадратная матрица размерности $s \times s$, величина s — число неизвестных в системе уравнений равновесия (число состояний в используемом пространстве состояний S), $P = (P_1, P_2, \dots, P_s)$ — вектор размерности s , P_k — стационарная вероятность состояния с номером k , где $k = 1, 2, \dots, s$.

Матрица A обычно имеет большое или даже бесконечное число элементов. Это свойство может усложнить и даже сделать невозможным применение стандартных процедур решения системы уравнений равновесия. Существуют несложные приёмы, позволяющие избавиться от указанного недостатка. Они основаны на урезании исходного пространства состояний. Процедура урезания заключается в удалении состояний, имеющих пренебрежимо малую вероятность существования. Для этого необходимо ограничить количество мест ожидания начала обслуживания и число пользователей, повторяющих заявку каким-либо достаточно большим целым числом. Подбор ограничения выполняется опытным путем. Матрица A также обладает свойством разреженности. Оно означает большое количество нулевых элементов в каждой строке. Элементы матрицы A связаны простыми рекуррентными соотношениями, позволяющими находить их как функцию целочисленных компонент, определяющих состояние. Для исследуемой модели эти зависимости будут далее приведены. Таким образом, для любого состояния модели можно без особого труда выписать левую и правую части соответствующего уравнения системы уравнений равновесия. При этом слагаемые, участвующие в записи этого уравнения и коэффициенты при них не берутся из памяти компьютера, а вычисляются в момент формирования анализируемого уравнения.

Существуют два основных способа решения систем уравнений равновесия: прямой и итерационный. При использовании прямых методов система уравнений равновесия преобразуется к виду, удобному для реализации стандартных алгоритмов линейной алгебры. Обычно эти преобразования являются встроенными процедурами в пакетах прикладных программ типа MathCAD или MathLab. Если прямые методы не основаны на каких-либо специальных свойствах матрицы системы уравнений равновесия, то область их применения существенно зависит от размеров анализируемого пространства состояний. Обычно $s \leq 200$ —300. В ряде частных случаев значительного продвижения в увеличении скорости решения системы уравнений равновесия и экономии памяти компьютера удастся достигнуть, если воспользоваться блочной структурой матрицы A . Речь идет о тех случаях, когда решение всей системы удастся разбить на решение отдельных подсистем, имеющих существенно меньшую размерность, чем исходная система. Подобные методы применяются при анализе простейших моделей контакт-центров с учётом влияния повторных заявок.

При использовании итерационных методов строится сходящаяся последовательность приближений к решению. Перечисленные выше свойства матрицы A позволяют рекомендовать для решения (3.1) итерационные методы, основанные на реализации принципа последовательных подстановок. К таким методам относится итерационный алгоритм Гаусса-Зейделя. Сформулируем его основные положения [58,59,69]. Обозначим вектор r -го приближения к искомому вектору неизвестных вероятностей $P = (P_1, P_2, \dots, P_s)$ через $P^{(r)} = (P_1^{(r)}, P_2^{(r)}, \dots, P_s^{(r)})$. При реализации итерационного метода Гаусса-Зейделя $(r+1)$ -ое приближение получается из r -го с использованием следующих рекуррентных соотношений

$$P_\ell^{(r+1)} = -\frac{1}{a_{\ell,\ell}} \left(\sum_{j=1}^{\ell-1} a_{\ell,j} P_j^{(r+1)} + \sum_{j=\ell+1}^s a_{\ell,j} P_j^{(r)} \right), \quad \ell = 1, 2, \dots, s. \quad (3.2)$$

В процессе определения значения $P_\ell^{(r+1)}$, являющегося ℓ -ым компонентом вектора $(r+1)$ -го приближения, необходимо использовать уже найденные компоненты $(r+1)$ -го приближения $P_j^{(r+1)}$, $j = 1, 2, \dots, \ell-1$ и имеющиеся компоненты r -го приближения $P_j^{(r)}$, $j = \ell+1, \ell+2, \dots, s$.

В качестве начального приближения можно взять любое известное приближение с положительными компонентами

$$P^{(0)} = (P_1^{(0)}, P_2^{(0)}, \dots, P_s^{(0)}).$$

Сходимость рассмотренной вычислительной процедуры оценивается исходя из достижения малости нормированной разницы между двумя последовательными приближениями к вектору неизвестных вероятностей, т.е. проверяется справедливость соотношения

$$\frac{|P_1^{(r+1)} - P_1^{(r)}| + |P_2^{(r+1)} - P_2^{(r)}| + \dots + |P_s^{(r+1)} - P_s^{(r)}|}{P_1^{(r+1)} + P_2^{(r+1)} + \dots + P_s^{(r+1)}} \leq \varepsilon.$$

Как правило, значение ε берётся из интервала $10^{-8} \dots 10^{-10}$. Как только сформулированное неравенство выполнено, итерационный цикл считается оконченным. В качестве значений вероятностей p_1, p_2, \dots, p_s берутся нормированные значения компонент вектора, являющегося последним из найденных приближений к искомому решению

$$p_\ell = \frac{P_\ell^{(r+1)}}{P_1^{(r+1)} + P_2^{(r+1)} + \dots + P_s^{(r+1)}}, \quad \ell = 1, 2, \dots, s.$$

Далее находятся значения показателей качества обслуживания заявок.

Число итераций зависит от значений входных параметров и величины ε , задающей относительную разницу между двумя последовательными приближениями. Обычно оно меняется от нескольких десятков до нескольких сотен. Максимальное число ненулевых элементов в строке матрицы системы уравнений равновесия исследуемой модели определяется максимальным числом слагаемых в соотношении (3.4) и не зависит от ограничений на число абонентов повторяющих вызов, а также от числа операторов, консультантов и мест ожидания. Каждый ненулевой элемент находится из рекуррентных формул, которые определяются компонентами состояния $(j, i, \ell, \ell_1, \dots, \ell_m)$. При реализации итерационной процедуры в памяти компьютера достаточно хранить один вектор размерности, равной числу состояний в исследуемом марковском процессе. Поскольку это число равно нескольким миллионам, то отсюда получаем диапазон изменения структурных параметров модели, для которого возможен расчёт характеристик рассматриваемым способом. Значения $n, v + w, v_k + w_k, k = 1, 2, \dots, m$ и ограничение на максимально возможное число абонентов, повторяющих вызов, не превосходят нескольких десятков, а число групп консультантов ограничено значениями 2 – 3. Для частных случаев построенной модели эти ограничения могут быть существенно увеличены.

Выполненное обсуждение даёт основание рекомендовать итерационный алгоритм Гаусса-Зейделя в качестве основного метода решения системы уравнений равновесия обобщённой модели контакт-центра. Особенности его реализации будут рассмотрены в следующих двух разделах. В частных случаях для решения системы уравнений равновесия могут быть использованы прямые методы, основанные на свойствах блочной структуры матрицы A , позволяющей свести решение всей системы к решению отдельных подсистем. Перечисленные выше подходы применяются к моделям, изменение состояний которых описывается марковским процессом. В более общих случаях оценка точных значений показателей качества обслуживания поступающих заявок находится средствами имитационного моделирования.

Основная область использования точных алгоритмов — анализ зависимости характеристик от изменения входных параметров и оценка погрешности инженерных методик расчета контакт-центров. Для практических приложений необходимо построить приближенные процедуры, имеющие приемлемую точность и несложные в реализации. Они основаны на использовании частных случаев модели, применении асимптотических разложений характеристик при стремлении интенсивности повторения к нулю и реализации

техники декомпозиции анализируемой модели на отдельные фрагменты со специальным образом подобранными параметрами. Эти результаты изложены, соответственно, в разделах 2.5, 4.1, 4.2.

3.3. Построение системы уравнений равновесия

3.3.1. Перечень анализируемых событий

Вероятности $p(j, i, \ell, \ell_1, \dots, \ell_m)$ определяются из решения системы уравнений статистического равновесия процесса $r(t)$. Она формируется из условия равенства интенсивности выхода из произвольного состояния $(j, i, \ell, \ell_1, \dots, \ell_m) \in S$ (левая часть системы уравнений равновесия) и интенсивности попадания в рассматриваемое состояние $(j, i, \ell, \ell_1, \dots, \ell_m)$ (правая часть системы уравнений равновесия). Формальная запись этого утверждения выглядит следующим образом:

$$\begin{aligned} & \{\text{интенсивность выхода } r(t) \text{ из } (j, i, \ell, \ell_1, \dots, \ell_m)\} = \\ & = \{\text{суммарная интенсивность входа } r(t) \text{ в } (j, i, \ell, \ell_1, \dots, \ell_m)\}. \end{aligned}$$

Построим процедуру формирования системы уравнений равновесия, которую будет удобно использовать для последующего её решения итерационным методом. Как уже было сказано, в анализируемой модели существует $4 + 3 \cdot m$ типов событий, меняющих её состояние. Их перечень выглядит следующим образом:

- поступление первичной заявки;
- поступление повторной заявки;
- завершение обслуживания в системе IVR с возможным его продолжением у оператора;
- превышение допустимого времени ожидания освобождения оператора;
- завершение обслуживания у оператора с возможным его продолжением в k -ой группе консультантов, $k = 1, 2, \dots, m$;
- превышение допустимого времени ожидания у консультанта k -ой группы; $k = 1, 2, \dots, m$;
- освобождение консультанта из k -ой группы; $k = 1, 2, \dots, m$.

Выпишем левые и правые части системы уравнений равновесия.

3.3.2. Формирование левой части

Для записи левой части системы уравнений равновесия необходимо найти условия осуществления всех событий, изменяющих произвольное состояние $(j, i, \ell, \ell_1, \dots, \ell_m)$ исследуемого процесса, и выписать интенсивность выхода из этого состояния. Под условиями осуществления событий будем понимать область в пространстве состояний, где указанное событие происходит, а также вероятности реализации событий, задающих рассматриваемый переход. Для выделения этой области будем использовать индикаторную функцию $I(\cdot)$, определяемую следующим соотношением

$$I(\cdot) = \begin{cases} 1, & \text{если выполнено условие, сформулированное} \\ & \text{в скобках,} \\ 0, & \text{если это условие не выполнено.} \end{cases} \quad (3.3)$$

В приведённом выражении в скобках указаны границы изменения компонент состояния $(j, i, \ell, \ell_1, \dots, \ell_m)$, где может осуществиться анализируемое событие.

Перечислим эти события и укажем частоту (интенсивность) их осуществления. Здесь и далее для произвольного состояния $(j, i, \ell, \ell_1, \dots, \ell_m) \in S$ обозначим через $c = \ell + \ell_1 + \dots + \ell_m$ число линий доступа, занятых на время обслуживания или ожидания начала обслуживания у оператора или консультанта из k -ой группы, $k = 1, 2, \dots, m$. Поступление первичных заявок, происходящее с интенсивностью λ для любого состояния $(j, i, \ell, \ell_1, \dots, \ell_m)$, изменяет состояние $(j, i, \ell, \ell_1, \dots, \ell_m)$ в следующих случаях:

1. С вероятностью, равной единице, если есть свободные линии доступа, т.е. для всех состояний $(j, i, \ell, \ell_1, \dots, \ell_m)$, удовлетворяющих условию $i < n$. В этой ситуации с интенсивностью $\lambda P(j, i, \ell, \ell_1, \dots, \ell_m)$ осуществляется переход из состояния $(j, i, \ell, \ell_1, \dots, \ell_m)$ в состояние $(j, i+1, \ell, \ell_1, \dots, \ell_m)$. В левой части системы уравнений равновесия получаем слагаемое: $\lambda I(i < n) P(j, i, \ell, \ell_1, \dots, \ell_m)$.
2. С вероятностью, равной H , если все линии доступа заняты, т.е. для всех состояний $(j, i, \ell, \ell_1, \dots, \ell_m)$, удовлетворяющих условию $i = n$. В этой ситуации с интенсивностью $\lambda H P(j, i, \ell, \ell_1, \dots, \ell_m)$ осуществляется переход из состояния $(j, i, \ell, \ell_1, \dots, \ell_m)$ в состояние $(j+1, i, \ell, \ell_1, \dots, \ell_m)$. В левой части системы уравнений равновесия получаем слагаемое: $\lambda H I(i = n) P(j, i, \ell, \ell_1, \dots, \ell_m)$.

Поступление повторных заявок, происходящее в состоянии $(j, i, \ell, \ell_1, \dots, \ell_m)$ с интенсивностью $j\nu$, изменяет состояние $(j, i, \ell, \ell_1, \dots, \ell_m)$ в следующих случаях:

1. С вероятностью, равной единице, если есть свободные линии доступа, т.е. для всех состояний $(j, i, \ell, \ell_1, \dots, \ell_m)$, удовлетворяющих условию $i < n$. В этой ситуации с интенсивностью $j\nu P(j, i, \ell, \ell_1, \dots, \ell_m)$ осуществляется переход из состояния $(j, i, \ell, \ell_1, \dots, \ell_m)$ в состояние $(j-1, i+1, \ell, \ell_1, \dots, \ell_m)$. В левой части системы уравнений равновесия получаем слагаемое: $j\nu I(i < n) P(j, i, \ell, \ell_1, \dots, \ell_m)$.

2. С вероятностью, равной $1 - H$, если заняты все линии доступа, т.е. для всех состояний $(j, i, l, l_1, \dots, l_m)$, удовлетворяющих условию $i = n$. В этой ситуации с интенсивностью $j\nu(1 - H)P(j, i, l, l_1, \dots, l_m)$ осуществляется переход из состояния $(j, i, l, l_1, \dots, l_m)$ в состояние $(j-1, i, l, l_1, \dots, l_m)$. В левой части системы уравнений равновесия получаем слагаемое: $j\nu(1 - H)I(i = n)P(j, i, l, l_1, \dots, l_m)$.

Диаграмма переходов марковского процесса $r(t)$ в результате поступления первичных и повторных заявок показана на рисунке 3.1.

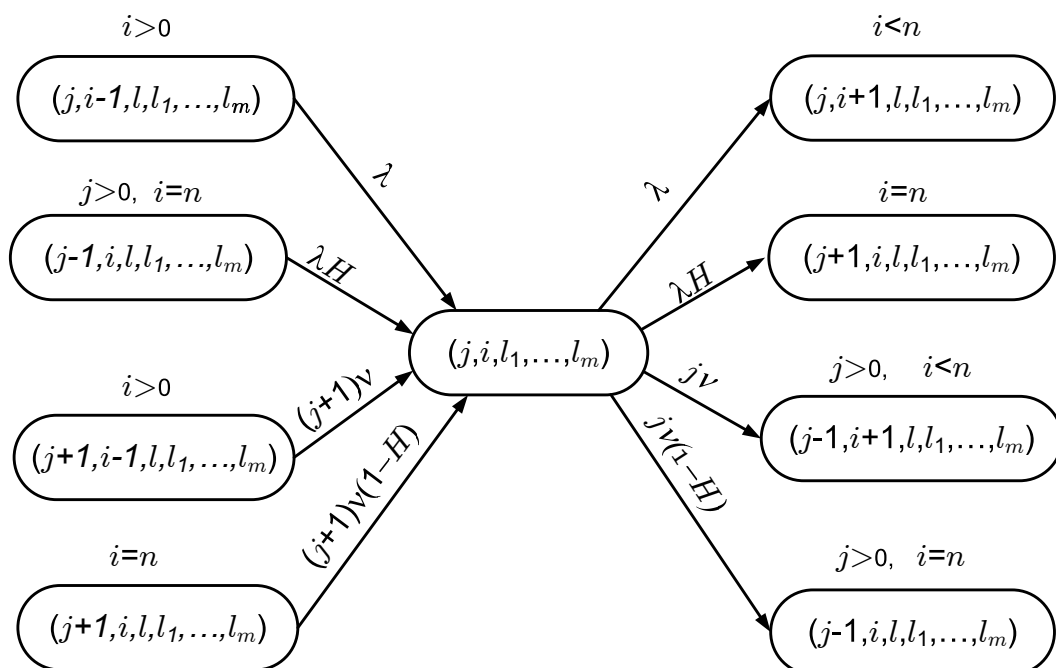


Рисунок 3.1. Направления и интенсивности переходов $r(t)$ в результате поступления первичных и повторных заявок

Завершение обслуживания в IVR, происходящее в состоянии $(j, i, l, l_1, \dots, l_m)$ с интенсивностью $(i - c)\alpha_1$ изменяет состояние $(j, i, l, l_1, \dots, l_m)$ в следующих случаях:

1. С вероятностью q , если абонент решит продолжить обслуживание и есть свободные операторы или места ожидания их освобождения, т.е. для всех состояний $(j, i, l, l_1, \dots, l_m)$, удовлетворяющих условию $i - c > 0, l < v + w$. В этой ситуации с интенсивностью $(i - c)\alpha_1 q P(j, i, l, l_1, \dots, l_m)$ осуществляется переход из состояния $(j, i, l, l_1, \dots, l_m)$ в состояние $(j, i, l+1, l_1, \dots, l_m)$. В левой части системы уравнений равновесия получаем слагаемое: $(i - c)\alpha_1 q I(i - c > 0, l < v + w) P(j, i, l, l_1, \dots, l_m)$.

2. С вероятностью q , если абонент решит продолжить обслуживание у оператора, но все операторы и места ожидания их освобождения заняты, т.е. для всех состояний $(j, i, \ell, \ell_1, \dots, \ell_m)$, удовлетворяющих условию $i - c > 0, \ell = v + w$. В этой ситуации с интенсивностью $(i - c)\alpha_1 q P(j, i, \ell, \ell_1, \dots, \ell_m)$ осуществляется переход из состояния $(j, i, \ell, \ell_1, \dots, \ell_m)$ в состояние $(j, i - 1, \ell, \ell_1, \dots, \ell_m)$. В левой части системы уравнений равновесия получаем слагаемое: $(i - c)\alpha_1 q I(i - c > 0, \ell = v + w) P(j, i, \ell, \ell_1, \dots, \ell_m)$.
3. С вероятностью, равной $1 - q$, в случае, если клиент получил необходимую информацию от системы IVR и отказывается от услуг оператора, т.е. для всех состояний $(j, i, \ell, \ell_1, \dots, \ell_m)$, удовлетворяющих условию $i - c > 0$. В этой ситуации с интенсивностью $(i - c)(1 - q)\alpha_1 P(j, i, \ell, \ell_1, \dots, \ell_m)$ происходит переход из состояния $(j, i, \ell, \ell_1, \dots, \ell_m)$ в состояние $(j, i - 1, \ell, \ell_1, \dots, \ell_m)$. В левой части системы уравнений равновесия получаем слагаемое: $(i - c)\alpha_1 (1 - q) I(i - c > 0) P(j, i, \ell, \ell_1, \dots, \ell_m)$.

Превышение допустимого времени ожидания обслуживания у оператора, происходящее в состоянии $(j, i, \ell, \ell_1, \dots, \ell_m)$ с интенсивностью $(\ell - v)\sigma$ (для осуществления данного события необходимо потребовать выполнения соотношения $\ell > v$), изменяет состояние $(j, i, \ell, \ell_1, \dots, \ell_m)$ в следующих случаях:

1. С вероятностью, равной H , при освобождении места для ожидания и принятии клиентом решения о повторении заявки, т.е. для всех состояний $(j, i, \ell, \ell_1, \dots, \ell_m)$, удовлетворяющих условию $v < \ell \leq v + w$. В рассматриваемой ситуации с интенсивностью $(\ell - v)\sigma H P(j, i, \ell, \ell_1, \dots, \ell_m)$ совершается переход из состояния $(j, i, \ell, \ell_1, \dots, \ell_m)$ в состояние $(j + 1, i - 1, \ell - 1, \ell_1, \dots, \ell_m)$. В левой части формируемой системы уравнений равновесия получаем слагаемое: $(\ell - v)\sigma H I(v < \ell \leq v + w) P(j, i, \ell, \ell_1, \dots, \ell_m)$.
2. С вероятностью, равной $1 - H$, при освобождении занятого места для ожидания и отказе от повторной попытки, т.е. для всех состояний $(j, i, \ell, \ell_1, \dots, \ell_m)$, удовлетворяющих условию $v < \ell \leq v + w$. В этой ситуации совершается переход с интенсивностью $(\ell - v)\sigma(1 - H) P(j, i, \ell, \ell_1, \dots, \ell_m)$ из состояния $(j, i, \ell, \ell_1, \dots, \ell_m)$ в состояние $(j, i - 1, \ell - 1, \ell_1, \dots, \ell_m)$. В левой части системы уравнений равновесия получаем слагаемое: $(\ell - v)\sigma(1 - H) I(v < \ell \leq v + w) P(j, i, \ell, \ell_1, \dots, \ell_m)$.

Завершение обслуживания у оператора, происходящее в состоянии $(j, i, \ell, \ell_1, \dots, \ell_m)$ с интенсивностью $\lambda\alpha_2$, если $\ell \leq v$ и $v\alpha_2$, если $\ell > v$, изменяет состояние в следующих случаях:

1. С вероятностью p , если клиент продолжает обслуживание у консультанта и выполняется неравенство $0 < \ell \leq v$. Выбор k -ой группы консультантов осуществляется с вероятностью p_k , $k = 1, 2, \dots, m$. Изменение состояния происходит, если есть свободный консультант k -ой группы или место ожидания, т.е. для всех состояний

$(j, i, \ell, \ell_1, \dots, \ell_m)$, удовлетворяющих условию $0 < \ell \leq v, \ell_k < v_k + w_k$. В этой ситуации с интенсивностью $\ell\alpha_2pp_kP(j, i, \ell, \ell_1, \dots, \ell_m)$ происходит переход из состояния $(j, i, \ell, \ell_1, \dots, \ell_m)$ в состояние $(j, i, \ell - 1, \ell_1, \dots, \ell_k + 1, \dots, \ell_m)$. В левой части формируемой системы уравнений равновесия получаем слагаемое: $\ell\alpha_2pp_kI(0 < \ell \leq v, \ell_k < v_k + w_k)P(j, i, \ell, \ell_1, \dots, \ell_m)$.

2. С вероятностью p , если клиенту требуется продолжение обслуживания у консультанта k -ой группы, и выполняется неравенство $0 < \ell \leq v$. Изменение состояния происходит, если заняты все консультанты k -ой группы и места ожидания их освобождения, т.е. для всех состояний $(j, i, \ell, \ell_1, \dots, \ell_m)$, удовлетворяющих условию $0 < \ell \leq v, \ell_k = v_k + w_k$. В этой ситуации с интенсивностью $\ell\alpha_2pp_kHP(j, i, \ell, \ell_1, \dots, \ell_m)$ происходит переход из состояния $(j, i, \ell, \ell_1, \dots, \ell_m)$ в состояние $(j + 1, i - 1, \ell - 1, \ell_1, \dots, \ell_m)$. В левой части формируемой системы уравнений равновесия исследуемой модели получаем слагаемое: $\ell\alpha_2pp_kHI(0 < \ell \leq v, \ell_k = v_k + w_k)P(j, i, \ell, \ell_1, \dots, \ell_m)$.
3. С вероятностью p , если клиенту требуется продолжение обслуживания у консультанта k -ой группы, и выполняется неравенство $0 < \ell \leq v$. Изменение состояния происходит, если заняты все консультанты k -ой группы и места ожидания их освобождения, т.е. для всех состояний $(j, i, \ell, \ell_1, \dots, \ell_m)$, удовлетворяющих условию $0 < \ell \leq v, \ell_k = v_k + w_k$. В этой ситуации с интенсивностью $\ell\alpha_2pp_k(1 - H)P(j, i, \ell, \ell_1, \dots, \ell_m)$ происходит переход из состояния $(j, i, \ell, \ell_1, \dots, \ell_m)$ в состояние $(j, i - 1, \ell - 1, \ell_1, \dots, \ell_m)$. В левой части формируемой системы уравнений равновесия исследуемой модели получаем слагаемое: $\ell\alpha_2pp_k(1 - H)I(0 < \ell \leq v, \ell_k = v_k + w_k)P(j, i, \ell, \ell_1, \dots, \ell_m)$.
4. С вероятностью $1 - p$, если клиент получил всю необходимую информацию от оператора и отказывается от услуг консультанта и выполняются неравенства $0 < \ell \leq v$. В этой ситуации с интенсивностью $\ell\alpha_2(1 - p)P(j, i, \ell, \ell_1, \dots, \ell_m)$ происходит переход из состояния $(j, i, \ell, \ell_1, \dots, \ell_m)$ в состояние $(j, i - 1, \ell - 1, \ell_1, \dots, \ell_m)$. В левой части системы уравнений равновесия получаем слагаемое: $\ell\alpha_2(1 - p)I(0 < \ell \leq v)P(j, i, \ell, \ell_1, \dots, \ell_m)$.
5. С вероятностью p , если клиент продолжает обслуживание у консультанта и выполняются неравенства $v < \ell \leq v + w$. Выбор k -ой группы консультантов осуществляется с вероятностью p_k , если есть свободный консультант k -ой группы или места ожидания освобождения консультантов, т.е. для всех состояний $(j, i, \ell, \ell_1, \dots, \ell_m)$, удовлетворяющих условию $v < \ell \leq v + w, \ell_k < v_k + w_k$. В этих условиях с интенсивностью $v\alpha_2pp_kP(j, i, \ell, \ell_1, \dots, \ell_m)$ происходит переход из состояния $(j, i, \ell, \ell_1, \dots, \ell_m)$ в состояние $(j, i, \ell - 1, \ell_1, \dots, \ell_k + 1, \dots, \ell_m)$. В левой части формируемой системы уравнений равновесия получаем слагаемое: $v\alpha_2pp_kI(v < \ell \leq v + w, \ell_k < v_k + w_k)P(j, i, \ell, \ell_1, \dots, \ell_m)$.

6. С вероятностью p , если клиенту требуется продолжение обслуживания у консультанта, и выполняется неравенство $v < \ell \leq v + w$. Изменение состояния происходит, если заняты все консультанты k -ой группы и места ожидания их освобождения, т.е. для всех состояний $(j, i, \ell, \ell_1, \dots, \ell_m)$, удовлетворяющих условию $v < \ell \leq v + w$, $\ell_k = v_k + w_k$. В этой ситуации с интенсивностью $v\alpha_2 pp_k HP(j, i, \ell, \ell_1, \dots, \ell_m)$ происходит переход из состояния $(j, i, \ell, \ell_1, \dots, \ell_m)$ в состояние $(j+1, i-1, \ell-1, \ell_1, \dots, \ell_m)$. В левой части формируемой системы уравнений равновесия получаем слагаемое: $v\alpha_2 pp_k HI(v < \ell \leq v + w, \ell_k = v_k + w_k)P(j, i, \ell, \ell_1, \dots, \ell_m)$.
7. С вероятностью p , если клиенту требуется продолжение обслуживания у консультанта, и выполняется неравенство $v < \ell \leq v + w$. Изменение состояния происходит, если заняты все консультанты k -ой группы и места ожидания их освобождения, т.е. для всех состояний $(j, i, \ell, \ell_1, \dots, \ell_m)$, удовлетворяющих условию $v < \ell \leq v + w$, $\ell_k = v_k + w_k$. В этой ситуации с интенсивностью $v\alpha_2 pp_k (1-H)P(j, i, \ell, \ell_1, \dots, \ell_m)$ происходит переход из состояния $(j, i, \ell, \ell_1, \dots, \ell_m)$ в состояние $(j, i-1, \ell-1, \ell_1, \dots, \ell_m)$. В левой части формируемой системы уравнений равновесия получаем слагаемое: $v\alpha_2 pp_k (1-H)I(v < \ell \leq v + w, \ell_k = v_k + w_k)P(j, i, \ell, \ell_1, \dots, \ell_m)$.
8. С вероятностью $1-p$, если клиент, получив необходимую информацию от оператора, отказывается от услуг консультанта, и выполняются неравенства $v < \ell \leq v + w$. В этой ситуации с интенсивностью $v\alpha_2(1-p)P(j, i, \ell, \ell_1, \dots, \ell_m)$ происходит переход из состояния $(j, i, \ell, \ell_1, \dots, \ell_m)$ в состояние $(j, i-1, \ell-1, \ell_1, \dots, \ell_m)$. В левой части системы уравнений равновесия получаем слагаемое: $v\alpha_2(1-p)I(v < \ell \leq v + w)P(j, i, \ell, \ell_1, \dots, \ell_m)$.

Диаграмма переходов марковского процесса $r(t)$ в результате поступления заявки на обслуживание к оператору показана на рисунке 3.2.

Превышение допустимого времени ожидания начала обслуживания у консультанта из k -ой группы, $k = 1, 2, \dots, m$, происходящее в состоянии $(j, i, \ell, \ell_1, \dots, \ell_m)$ с интенсивностью $(\ell_k - v_k)\sigma_k$ (для осуществления события необходимо потребовать, чтобы $\ell_k > v_k$), изменяет состояние:

1. С вероятностью, равной H , при освобождении занятого места ожидания в k -ой группе консультантов, $k = 1, 2, \dots, m$ и принятии клиентом решения о повторении заявки, т.е. для всех состояний $(j, i, \ell, \ell_1, \dots, \ell_m)$, удовлетворяющих условию $v_k < \ell_k \leq v_k + w_k$. В этих условиях с интенсивностью $(\ell_k - v_k)\sigma_k HP(j, i, \ell, \ell_1, \dots, \ell_m)$ совершается переход из состояния $(j, i, \ell, \ell_1, \dots, \ell_k, \dots, \ell_m)$ в состояние $(j+1, i-1, \ell, \ell_1, \dots, \ell_k-1, \dots, \ell_m)$. В левой части формируемой системы уравнений равновесия получаем слагаемое: $(\ell_k - v_k)\sigma_k HI(v_k < \ell_k \leq v_k + w_k)P(j, i, \ell, \ell_1, \dots, \ell_m)$.
2. С вероятностью, равной $1-H$, при освобождении занятого места ожидания в k -ой группе консультантов, $k = 1, 2, \dots, m$ и отказе от повторной попытки, т.е. для всех

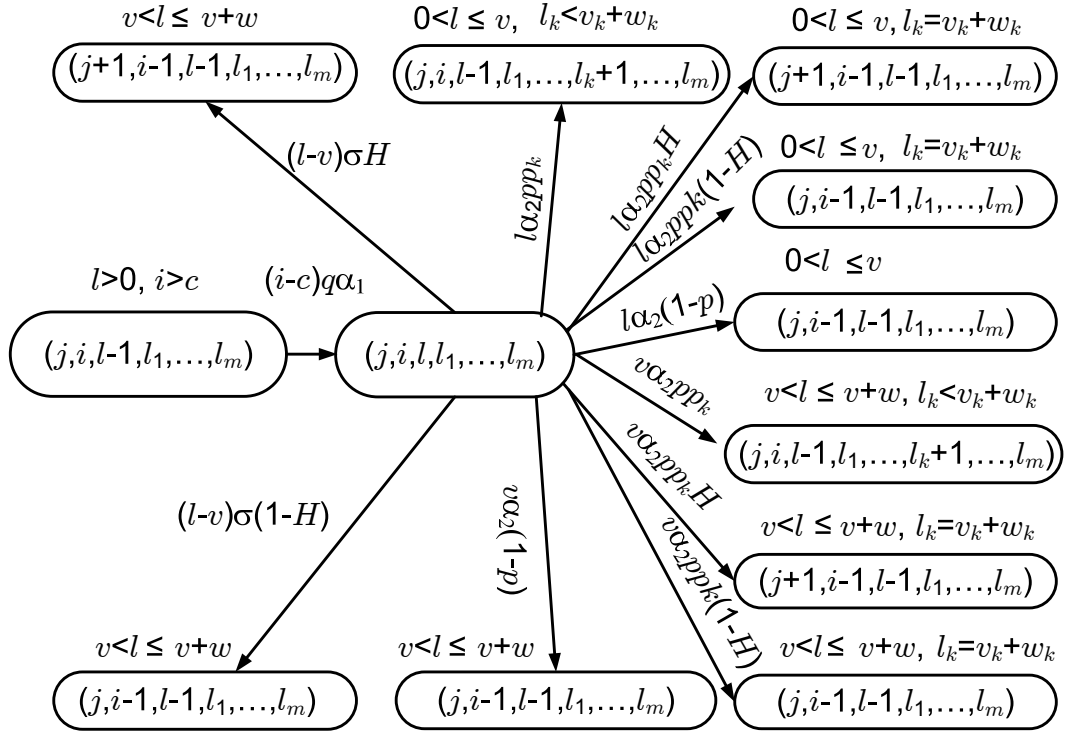


Рисунок 3.2. Направления и интенсивности переходов $r(t)$ после поступления заявки на обслуживание к оператору, последующего завершения обслуживания (или отказа от него) и возможных результатов установления соединения с консультантом

состояний $(j, i, l, l_1, \dots, l_m)$, удовлетворяющих условию $v_k < l_k \leq v_k + w_k$. В этой ситуации с интенсивностью $(l_k - v_k)\sigma_k(1-H)P(j, i, l, l_1, \dots, l_m)$ совершается переход из состояния $(j, i, l, l_1, \dots, l_k, \dots, l_m)$ в состояние $(j, i-1, l, l_1, \dots, l_k-1, \dots, l_m)$. В левой части системы уравнений равновесия получаем слагаемое: $(l_k - v_k)\sigma_k(1-H)I(v_k < l_k \leq v_k + w_k)P(j, i, l, l_1, \dots, l_m)$.

Освобождение консультанта из k -ой группы, $k = 1, 2, \dots, m$, осуществляемое в состоянии $(j, i, l, l_1, \dots, l_m)$, с интенсивностью $l_k\beta_k$, если $0 < l_k \leq v_k$, и $v_k\beta_k$, если $v_k < l_k \leq v_k + w_k$, изменяет состояние $(j, i, l, l_1, \dots, l_m)$ в следующих случаях:

1. С вероятностью, равной единице, если есть занятые консультанты в k -ой группе, $k = 1, 2, \dots, m$, и число занятых мест ожидания в k -ой группе равно 0, т.е., для всех состояний $(j, i, l, l_1, \dots, l_m)$, удовлетворяющих условию $0 < l_k \leq v_k$. В этой ситуации с интенсивностью $l_k\beta_kP(j, i, l, l_1, \dots, l_k, \dots, l_m)$ совершается переход из состояния $(j, i, l, l_1, \dots, l_k, \dots, l_m)$ в состояние модели $(j, i-1, l, l_1, \dots, l_k-1, \dots, l_m)$. В левой части системы уравнений получаем слагаемое: $l_k\beta_kI(0 < l_k \leq v_k)P(j, i, l, l_1, \dots, l_m)$.
2. С вероятностью, равной единице, если есть занятые консультанты k -ой группы,

$k = 1, 2, \dots, m$ и число занятых мест ожидания в k -ой группе больше 0, т.е., для всех состояний $(j, i, l, l_1, \dots, l_m)$, удовлетворяющих условию $v_k < l_k \leq v_k + w_k$. В этой ситуации с интенсивностью $v_k \beta_k P(j, i, l, l_1, \dots, l_k, \dots, l_m)$ совершается переход из состояния $(j, i, l, l_1, \dots, l_k, \dots, l_m)$ в состояние $(j, i - 1, l, l_1, \dots, l_k - 1, \dots, l_m)$. В левой части системы уравнений получаем слагаемое: $v_k \beta_k I(v_k < l_k \leq v_k + w_k) P(j, i, l, l_1, \dots, l_m)$.

Диаграмма переходов марковского процесса $r(t)$ в результате поступления на обслуживание к консультанту k -ой группы, $k = 1, 2, \dots, m$ показана на рисунке 3.3.

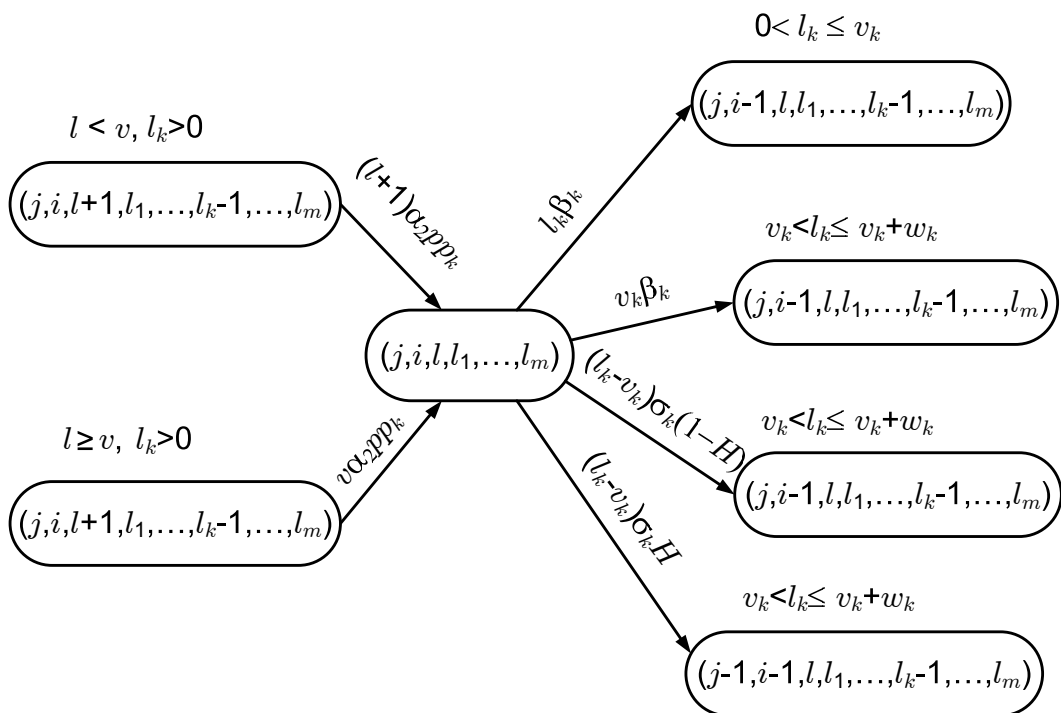


Рисунок 3.3. Направления и интенсивности переходов $r(t)$ в результате поступления заявки на обслуживание к консультанту k -ой группы, $k = 1, 2, \dots, m$, последующего завершения обслуживания или отказа от попытки установить соединение с консультантом

3.3.3. Формирование правой части

Для определения правой части системы уравнений равновесия в той же последовательности перебираются основные события, изменяющие состояние исследуемого процесса, и находятся состояния и интенсивности перехода $r(t)$ в произвольное заданное состояние $(j, i, l, l_1, \dots, l_m)$.

Поступление первичной заявки приводит к переходу $r(t)$ в состояние $(j, i, \ell, \ell_1, \dots, \ell_m)$ в следующих случаях:

1. В результате занятия линии доступа. Данное событие происходит с вероятностью, равной единице, в состоянии $(j, i-1, \ell, \ell_1, \dots, \ell_m)$, удовлетворяющем условию $i-1 \geq c$. Это даёт в правой части системы уравнений равновесия слагаемое, имеющее вид: $\lambda I(i-1 \geq c)P(j, i-1, \ell, \ell_1, \dots, \ell_m)$.
2. В результате увеличения числа клиентов, повторяющих заявку из-за отсутствия свободных линий доступа. Данное событие происходит с вероятностью, равной H , в состоянии $(j-1, i, \ell, \ell_1, \dots, \ell_m)$, удовлетворяющем условию $j > 0, i = n$. Это даёт слагаемое в правой части исследуемой системы уравнений равновесия, имеющее вид: $\lambda HI(j > 0, i = n)P(j-1, i, \ell, \ell_1, \dots, \ell_m)$.

Поступление повторной заявки осуществляет переход $r(t)$ в состояние $(j, i, \ell, \ell_1, \dots, \ell_m)$ в следующих случаях:

1. В результате занятия линии доступа. Данное событие происходит с вероятностью, равной единице, в состоянии $(j+1, i-1, \ell, \ell_1, \dots, \ell_m)$, удовлетворяющем условию $i-1 \geq c$. Это даёт слагаемое в правой части системы уравнений равновесия, имеющее вид: $(j+1)\nu I(i-1 \geq c)P(j+1, i-1, \ell, \ell_1, \dots, \ell_m)$.
2. В результате уменьшения числа клиентов, повторяющих заявку. Данное событие происходит с вероятностью, равной $1-H$, в состоянии $(j+1, i, \ell, \ell_1, \dots, \ell_m)$, удовлетворяющем условию $i = n$. Это даёт слагаемое в правой части, имеющее вид: $(j+1)\nu(1-H)I(i = n)P(j+1, i, \ell, \ell_1, \dots, \ell_m)$.

Завершение обслуживания заявки в системе IVR осуществляет переход $r(t)$ в состояние $(j, i, \ell, \ell_1, \dots, \ell_m)$ в следующих случаях:

1. В результате перевода заявки на обслуживание к оператору. Данное событие происходит с вероятностью равной q в состоянии $(j, i, \ell-1, \ell_1, \dots, \ell_m)$, удовлетворяющем условию $\ell > 0$. Это даёт слагаемое в правой части, имеющее вид: $(i-c+1)\alpha_1 q I(\ell > 0)P(j, i, \ell-1, \ell_1, \dots, \ell_m)$.
2. В результате получения от IVR необходимого информационной услуги и последующего отказа клиента от продолжения обслуживания у оператора. Это событие происходит с вероятностью, равной $1-q$, в состоянии $(j, i+1, \ell, \ell_1, \dots, \ell_m)$, удовлетворяющем условию $i+1 \leq n$. Это даёт слагаемое в правой части, имеющее вид: $(i+1-c)\alpha_1(1-q)I(i+1 \leq n)P(j, i+1, \ell, \ell_1, \dots, \ell_m)$.
3. В результате решения клиента о повторении вызова. Это событие происходит с вероятностью qH , в состоянии $(j-1, i+1, \ell, \ell_1, \dots, \ell_m)$, удовлетворяющем условию $j > 0, i+1 \leq n$ и $\ell = v+w$. Это даёт слагаемое в правой части, имеющее вид: $(i+1-c)\alpha_1 q HI(j > 0, i+1 \leq n, \ell = v+w)P(j-1, i+1, \ell, \ell_1, \dots, \ell_m)$.

4. В результате отказа клиента от возможности повторить вызов и ухода из системы. Это событие происходит с вероятностью $q(1 - H)$, в состоянии $(j, i + 1, \ell, \ell_1, \dots, \ell_m)$, удовлетворяющем условию $i + 1 \leq n$ и $\ell = v + w$. Это даёт слагаемое в правой части, имеющее вид: $(i + 1 - c)\alpha_1 q(1 - H)I(i + 1 \leq n, \ell = v + w)P(j, i + 1, \ell, \ell_1, \dots, \ell_m)$.

Превышение допустимого времени ожидания освобождения оператора осуществляет переход $r(t)$ в состояние $(j, i, \ell, \ell_1, \dots, \ell_m)$ в следующих случаях:

1. В результате увеличения числа клиентов, повторяющих заявку, после превышения допустимого времени ожидания начала обслуживания у оператора. Данное событие происходит с вероятностью, равной H , в состоянии $(j - 1, i + 1, \ell + 1, \ell_1, \dots, \ell_m)$, удовлетворяющем условию $j > 0, i + 1 \leq n, v < \ell + 1 \leq v + w$. Это даёт следующее слагаемое в правой части системы уравнений равновесия исследуемой модели: $(\ell + 1 - v)\sigma HI(j > 0, i + 1 \leq n, v < \ell + 1 \leq v + w)P(j - 1, i + 1, \ell + 1, \ell_1, \dots, \ell_m)$.
2. В результате уменьшения числа клиентов, ожидающих освобождения операторов, после превышения допустимого времени ожидания и отказа от продолжения обслуживания. Данное событие происходит с вероятностью, равной $1 - H$, в состоянии $(j, i + 1, \ell + 1, \ell_1, \dots, \ell_m)$, удовлетворяющем условию $i + 1 \leq n, v < \ell + 1 \leq v + w$. Это даёт слагаемое в правой части системы уравнений, имеющее вид: $(\ell + 1 - v)\sigma(1 - H)I(i + 1 \leq n, v < \ell + 1 \leq v + w)P(j, i + 1, \ell + 1, \ell_1, \dots, \ell_m)$.

Завершение обслуживания заявки оператором осуществляет переход $r(t)$ в состояние $(j, i, \ell, \ell_1, \dots, \ell_m)$ в следующих случаях:

1. С вероятностью pp_k в результате перевода заявки на обслуживание в k -ую группу консультантов, $k = 1, 2, \dots, m$, при отсутствии занятых мест ожидания обслуживания у оператора. Данное событие происходит с вероятностью pp_k в состоянии $(j, i, \ell + 1, \ell_1, \dots, \ell_k - 1, \dots, \ell_m)$, удовлетворяющем условию $\ell + 1 \leq v, \ell_k > 0$. Это даёт слагаемое в правой части, имеющее вид: $(\ell + 1)\alpha_2 pp_k I(\ell + 1 \leq v, \ell_k > 0)P(j, i, \ell + 1, \ell_1, \dots, \ell_k - 1, \dots, \ell_m)$.
2. С вероятностью pp_k в результате отказа клиента от дальнейшего обслуживания из-за отсутствия свободных консультантов и мест ожидания в k -ой группе, $k = 1, 2, \dots, m$ и при отсутствии занятых мест ожидания обслуживания у оператора. Данное событие происходит с вероятностью $pp_k(1 - H)$ в состоянии $(j, i + 1, \ell + 1, \ell_1, \dots, \ell_m)$, удовлетворяющем условию $i + 1 \leq n, \ell + 1 \leq v, \ell_k = v_k + w_k$. Это даёт слагаемое в правой части, имеющее вид: $(\ell + 1)\alpha_2 pp_k(1 - H)I(i + 1 \leq n, \ell + 1 \leq v, \ell_k = v_k + w_k)P(j, i + 1, \ell + 1, \ell_1, \dots, \ell_m)$.
3. С вероятностью pp_k в результате повторения заявки клиентом из-за отсутствия свободных консультантов и мест ожидания в k -ой группе, $k = 1, 2, \dots, m$ и при отсутствии занятых мест ожидания обслуживания у оператора. Данное событие

происходит с вероятностью $pp_k H$ в состоянии $(j - 1, i + 1, \ell + 1, \ell_1, \dots, \ell_m)$, удовлетворяющем условию $j > 0, i + 1 \leq n, \ell + 1 \leq v, \ell_k = v_k + w_k$. Это даёт слагаемое в правой части, имеющее вид: $(\ell + 1)\alpha_2 pp_k HI(j > 0, i + 1 \leq n, \ell + 1 \leq v, \ell_k = v_k + w_k)P(j - 1, i + 1, \ell + 1, \ell_1, \dots, \ell_m)$.

4. С вероятностью $1 - p$ в результате получения необходимой информации у оператора и отказа абонента от дальнейшего обслуживания у консультанта, при отсутствии занятых мест ожидания обслуживания у оператора. Данное событие происходит в состоянии $(j, i + 1, \ell + 1, \ell_1, \dots, \ell_m)$, удовлетворяющем условию $i + 1 \leq n, \ell + 1 \leq v$. Это даёт слагаемое в правой части, имеющее вид: $(\ell + 1)\alpha_2(1 - p)I(i + 1 \leq n, \ell + 1 \leq v)P(j, i + 1, \ell + 1, \ell_1, \dots, \ell_m)$.
5. С вероятностью pp_k в результате перевода заявки на обслуживание в k -ую группу консультантов, $k = 1, 2, \dots, m$, при наличии занятых мест ожидания обслуживания у оператора. Данное событие происходит с вероятностью pp_k в состоянии $(j, i, \ell + 1, \ell_1, \dots, \ell_k - 1, \dots, \ell_m)$, удовлетворяющем условию $v < \ell + 1 \leq v + w, \ell_k > 0$. Это даёт слагаемое в правой части, имеющее вид: $v\alpha_2 pp_k I(v < \ell + 1 \leq v + w, \ell_k > 0)P(j, i, \ell + 1, \ell_1, \dots, \ell_k - 1, \dots, \ell_m)$.
6. С вероятностью pp_k в результате отказа клиента от дальнейшего обслуживания из-за отсутствия свободных консультантов и мест ожидания в k -ой группе, $k = 1, 2, \dots, m$ и при наличии занятых мест ожидания обслуживания у операторов. Данное событие происходит с вероятностью $pp_k(1 - H)$ в состоянии $(j, i + 1, \ell + 1, \ell_1, \dots, \ell_m)$, удовлетворяющем условию $i + 1 \leq n, v < \ell + 1 \leq v + w, \ell_k = v_k + w_k$. Это даёт слагаемое в правой части, имеющее вид: $v\alpha_2 pp_k(1 - H)I(i + 1 \leq n, v < \ell + 1 \leq v + w, \ell_k = v_k + w_k)P(j, i + 1, \ell + 1, \ell_1, \dots, \ell_m)$.
7. С вероятностью pp_k в результате повторения заявки клиентом из-за отсутствия свободных консультантов и мест ожидания в k -ой группе, $k = 1, 2, \dots, m$. Данное событие происходит с вероятностью $pp_k H$ в состоянии $(j - 1, i + 1, \ell + 1, \ell_1, \dots, \ell_m)$, удовлетворяющем условию $j > 0, i + 1 \leq n, v \leq \ell + 1 < v + w, \ell_k = v_k + w_k$. Это даёт слагаемое в правой части, имеющее вид: $v\alpha_2 pp_k HI(j > 0, i + 1 \leq n, v \leq \ell + 1 < v + w, \ell_k = v_k + w_k)P(j - 1, i + 1, \ell + 1, \ell_1, \dots, \ell_m)$.
8. С вероятностью $1 - p$ в результате отказа абонента от дальнейшего обслуживания у консультантов, $k = 1, 2, \dots, m$, при наличии занятых мест ожидания обслуживания у оператора. Данное событие происходит в состоянии $(j, i + 1, \ell + 1, \ell_1, \dots, \ell_m)$, удовлетворяющем условию $i + 1 \leq n, v < \ell + 1 \leq v + w$. Это даёт слагаемое в правой части, имеющее вид: $v\alpha_2(1 - p)I(i + 1 \leq n, v < \ell + 1 \leq v + w)P(j, i + 1, \ell + 1, \ell_1, \dots, \ell_m)$.

Превышение допустимого времени ожидания обслуживания у консультанта из k -ой группы, $k = 1, 2, \dots, m$ осуществляет переход $r(t)$ в состояние $(j, i, \ell, \ell_1, \dots, \ell_m)$ в следующих случаях:

1. В результате увеличения числа клиентов, повторяющих заявку, после превышения допустимого времени ожидания обслуживания у консультанта k -ой группы, $k = 1, 2, \dots, m$. Данное событие происходит с вероятностью, равной H , в состоянии $(j - 1, i + 1, \ell, \ell_1, \dots, \ell_k + 1, \dots, \ell_m)$, удовлетворяющем условию $j > 0, i + 1 \leq n, v_k < \ell_k + 1 \leq v_k + w_k$. Это даёт следующее слагаемое в правой части системы уравнений равновесия: $(\ell_k - v_k + 1)\sigma_k H I(j > 0, i + 1 \leq n, v_k < \ell_k + 1 \leq v_k + w_k) P(j - 1, i + 1, \ell, \ell_1, \dots, \ell_k + 1, \dots, \ell_m)$.
2. В результате уменьшения числа клиентов, находящихся на ожидании, после превышения допустимого времени ожидания обслуживания у консультанта k -ой группы, $k = 1, 2, \dots, m$ и отказа от повторной заявки. Данное событие происходит с вероятностью, равной $1 - H$, в состоянии $(j, i + 1, \ell, \ell_1, \dots, \ell_k + 1, \dots, \ell_m)$, удовлетворяющем условию $i + 1 \leq n, v_k < \ell_k + 1 \leq v_k + w_k$. Это даёт слагаемое в правой части системы уравнений, имеющее вид: $(\ell_k - v_k + 1)\sigma_k(1 - H)I(i + 1 \leq n, v_k < \ell_k + 1 \leq v_k + w_k)P(j, i + 1, \ell, \ell_1, \dots, \ell_k + 1, \dots, \ell_m)$.

Завершение обслуживания заявки консультантом из k -ой группы консультантов, $k = 1, 2, \dots, m$ осуществляет переход $r(t)$ в состояние $(j, i, \ell, \ell_1, \dots, \ell_m)$ в следующих случаях:

1. В результате освобождения консультанта k -ой группы, $k = 1, 2, \dots, m$ при отсутствии занятых мест ожидания их освобождения. Данное событие происходит с вероятностью, равной единице, в состоянии $(j, i + 1, \ell, \ell_1, \dots, \ell_k + 1, \dots, \ell_m)$ удовлетворяющем условию $i + 1 \leq n, \ell_k + 1 \leq v_k$. Это даёт слагаемое в правой части системы уравнений, имеющее вид: $(\ell_k + 1)\beta_k I(i + 1 \leq n, \ell_k + 1 \leq v_k)P(j, i + 1, \ell, \ell_1, \dots, \ell_k + 1, \dots, \ell_m)$.
2. В результате освобождения консультанта k -ой группы, $k = 1, 2, \dots, m$ при наличии занятых мест ожидания их освобождения. Данное событие происходит с вероятностью, равной единице, в состоянии $(j, i + 1, \ell, \ell_1, \dots, \ell_k + 1, \dots, \ell_m)$, удовлетворяющем условию $i + 1 \leq n, v_k < \ell_k + 1 \leq v_k + w_k$. Это даёт в правой части системы уравнений слагаемое, имеющее вид: $v_k \beta_k I(i + 1 \leq n, v_k < \ell_k + 1 \leq v_k + w_k)P(j, i + 1, \ell, \ell_1, \dots, \ell_k + 1, \dots, \ell_m)$.

3.3.4. Система уравнений равновесия

Приравняв интенсивности выхода и входа для всех состояний $(j, i, \ell, \ell_1, \dots, \ell_m)$ исследуемого марковского процесса $r(t)$ и приведя подобные члены, получаем следующую бесконечную систему уравнений статистического равновесия, связывающую стационарные вероятности $P(j, i, \ell, \ell_1, \dots, \ell_m)$

$$P(j, i, \ell, \ell_1, \dots, \ell_m) \times \left\{ I(i < n)(\lambda + j\nu) + I(i = n)(\lambda H + j\nu(1 - H)) + \right. \tag{3.4}$$

$$\begin{aligned}
& +I(i > c)(i - c)\alpha_1 + I(\ell > v)\left((\ell - v)\sigma + v\alpha_2\right) + I(0 < \ell \leq v)\ell\alpha_2 + \\
& + \sum_{k=1}^m \left(I(\ell_k > v_k)\left((\ell_k - v_k)\sigma_k + v_k\beta_k\right) + I(0 < \ell_k \leq v_k)\ell_k\beta_k \right) \Big\} = \\
& = P(j, i - 1, \ell, \ell_1, \dots, \ell_m)I(i > 0, i - 1 \geq c)\lambda + \\
& + P(j - 1, i, \ell, \ell_1, \dots, \ell_m)I(j > 0, i = n)\lambda H + \\
& + P(j + 1, i - 1, \ell, \ell_1, \dots, \ell_m)I(i - 1 \geq c)(j + 1)\nu + \\
& + P(j + 1, i, \ell, \ell_1, \dots, \ell_m)I(i = n)(j + 1)\nu(1 - H) + \\
& + P(j, i, \ell - 1, \ell_1, \dots, \ell_m)I(\ell > 0)(i - c + 1)\alpha_1 q + \\
& + P(j, i + 1, \ell, \ell_1, \dots, \ell_m)I(i + 1 \leq n)(i + 1 - c)\alpha_1 \times \\
& \quad \times \left((1 - q) + q(1 - H)I(\ell = v + w) \right) + \\
& + P(j - 1, i + 1, \ell, \ell_1, \dots, \ell_m)I(i + 1 \leq n, j > 0, \ell = v + w)(i + 1 - c)\alpha_1 q H + \\
& + P(j - 1, i + 1, \ell + 1, \ell_1, \dots, \ell_m)I(j > 0, i + 1 \leq n)H \times \\
& \times \left\{ I(v < \ell + 1 \leq v + w)(\ell + 1 - v)\sigma + I(\ell_k = v_k + w_k)\alpha_2 p p_k \times \right. \\
& \times \left. \left(I(\ell + 1 \leq v)(\ell + 1) + I(v \leq \ell + 1 < v + w)v \right) \right\} + \\
& + P(j, i + 1, \ell + 1, \ell_1, \dots, \ell_m)I(i + 1 \leq n) \times \\
& \times \left\{ (\ell + 1 - v)\sigma(1 - H)I(v < \ell + 1 \leq v + w) + \right. \\
& + I(\ell + 1 \leq v)(\ell + 1)\alpha_2(1 - p) + \\
& + I(v < \ell + 1 \leq v + w)v\alpha_2(1 - p) + \\
& + \sum_{k=1}^m I(\ell_k = v_k + w_k)\alpha_2 p p_k(1 - H) \times \\
& \times \left. \left(I(\ell + 1 \leq v)(\ell + 1) + I(v < \ell + 1 \leq v + w)v \right) \right\} + \\
& + \sum_{k=1}^m P(j, i, \ell + 1, \ell_1, \dots, \ell_k - 1, \dots, \ell_m)I(\ell_k > 0)\alpha_2 p p_k \times
\end{aligned}$$

$$\begin{aligned}
& \times \left(I(\ell + 1 \leq v)(\ell + 1) + I(v < \ell + 1 \leq v + w)v \right) + \\
& + \sum_{k=1}^m P(j - 1, i + 1, \ell, \ell_1, \dots, \ell_k + 1, \dots, \ell_m) \times \\
& \times I(j > 0, v_k < \ell_k + 1 \leq v_k + w_k, i + 1 \leq n)(\ell_k - v_k + 1)\sigma_k H + \\
& + \sum_{k=1}^m P(j, i + 1, \ell, \ell_1, \dots, \ell_k + 1, \dots, \ell_m) I(i + 1 \leq n) \times \\
& \times \left\{ I(v_k < \ell_k + 1 \leq v_k + w_k) \left((\ell_k - v_k + 1)\sigma_k(1 - H) + v_k\beta_k \right) + \right. \\
& \left. + I(\ell_k + 1 \leq v_k)(\ell_k + 1)\beta_k \right\}.
\end{aligned}$$

Для $P(j, i, \ell, \ell_1, \dots, \ell_m)$ выполнено условие нормировки

$$\sum_{(j, i, \ell, \ell_1, \dots, \ell_m) \in S} P(j, i, \ell, \ell_1, \dots, \ell_m) = 1.$$

3.3.5. Ограничение на число абонентов, повторяющих вызовы

Для применения численных методов необходимо ограничить число неизвестных в системе уравнений (3.4). С этой целью будем предполагать, что максимально возможное число абонентов, которые могут одновременно повторять попытку соединения не превосходит некоторого целого числа j_m . Из условий существования стационарного режима следует, что можно подобрать величину j_m так, чтобы значения характеристик «урезанной модели» с любой заданной наперед точностью совпадали со значениями аналогичных характеристик исходной модели без ограничения на максимально возможное число абонентов, повторяющих вызовы. Теоретические исследования простых моделей с повторными вызовами показали [53], что погрешность оценки пропорциональна величине вероятности нахождения модели в состоянии, когда число абонентов, повторяющих вызов равно j_m . Интуитивно понятно, что это свойство должно присутствовать и в исследуемой обобщённой модели контакт-центра. В дальнейших вычислениях величина j_m подбиралась опытным путем из условия выполнения сформулированной гипотезы, справедливости законов сохранения (2.5)–(2.8) и сходимости оценок характеристик при последовательном увеличении значения j_m . Величина j_m будет указываться в перечне входных параметров модели во всех последующих численных примерах. Наличие ограничения на изменение числа абонентов, повторяющих вызовы, меняет вид системы уравнений (3.4). Рассматриваемая модификация (3.4) не вызывает больших затруднений, однако её окончательный результат выглядит достаточно громоздко и здесь не приводится.

Для простоты, последующие расчётные выражения и алгоритмы будут сформулированы для случая $j_m = \infty$.

Как показало исследование, проведённое в разделах 3.2-3.3, построенную систему уравнений следует решать итерационным методом Гаусса-Зейделя. Используя результаты раздела 3.3, сформулируем основные расчетные выражения.

3.4. Формулировка рекурсивной схемы расчётов

В разделе 3.3 приведено алгебраическое представление системы уравнений равновесия в виде удобном для его последующего применения при реализации метода Гаусса-Зейделя. Покажем, как это общее выражение можно использовать для решения системы уравнений равновесия исследуемой модели контакт-центра. В соответствии с результатами раздела 3.3 и основными положениями итерационного метода Гаусса-Зейделя последовательные приближения к значениям стационарных вероятностей вычисляются по следующим рекуррентным формулам. Обозначим через $P^{(r)}(j, i, \ell, \ell_1, \dots, \ell_m)$ величину r -го найденного приближения к ненормированному значению стационарной вероятности $P(j, i, \ell, \ell_1, \dots, \ell_m)$, $r = 0, 1, \dots$. Реализация итерационной схемы Гаусса-Зейделя состоит в выполнении следующих шагов:

1. Выбирается начальное приближение из условия

$$P^{(0)}(j, i, \ell, \ell_1, \dots, \ell_m) = 1, \quad (j, i, \ell, \ell_1, \dots, \ell_m) \in S.$$

2. Для каждого $(j, i, \ell, \ell_1, \dots, \ell_m) \in S$ и значений r равных $0, 1, \dots$ вычисляются величины последовательных приближений к значениям стационарных вероятностей модели $P^{(r+1)}(j, i, \ell, \ell_1, \dots, \ell_m)$, используя соотношение

$$P^{(r+1)}(j, i, \ell, \ell_1, \dots, \ell_m) = \frac{R^{(r+1)}(j, i, \ell, \ell_1, \dots, \ell_m)}{L(j, i, \ell, \ell_1, \dots, \ell_m)}. \quad (3.5)$$

В (3.5) функция $L(j, i, \ell, \ell_1, \dots, \ell_m)$ представляет собой коэффициент при оцениваемой вероятности $P(j, i, \ell, \ell_1, \dots, \ell_m)$ в левой части (3.4). Значение рассматриваемой функции не зависит от величины r и в зависимости от значений целочисленных компонент состояния $(j, i, \ell, \ell_1, \dots, \ell_m)$ определяется из соотношения

$$\begin{aligned} L(j, i, \ell, \ell_1, \dots, \ell_m) = & \\ = & I(i < n)(\lambda + j\nu) + I(i = n)(\lambda H + j\nu(1 - H)) + \\ & + I(i > c)(i - c)\alpha_1 + I(\ell > v)((\ell - v)\sigma + v\alpha_2) + I(0 < \ell \leq v)\ell\alpha_2 + \end{aligned}$$

$$+ \sum_{k=1}^m \left(I(\ell_k > v_k) \left((\ell_k - v_k)\sigma_k + v_k\beta_k \right) + I(0 < \ell_k \leq v_k) \ell_k \beta_k \right).$$

Функция $R^{(r+1)}(j, i, \ell, \ell_1, \dots, \ell_m)$ представляет собой правую часть соотношения (3.4). Она зависит от значений целочисленных компонент состояния $(j, i, \ell, \ell_1, \dots, \ell_m)$ и значений r -го или $(r+1)$ -го приближений для стационарных вероятностей. Величина рассматриваемой функции определяется из следующего соотношения:

$$\begin{aligned} R^{(r+1)}(j, i, \ell, \ell_1, \dots, \ell_m) &= \\ &= P^{(r+1,r)}(j, i-1, \ell, \ell_1, \dots, \ell_m) I(i > 0, i-1 \geq c) \lambda + \\ &+ P^{(r+1,r)}(j-1, i, \ell, \ell_1, \dots, \ell_m) I(j > 0, i = n) \lambda H + \\ &+ P^{(r+1,r)}(j+1, i-1, \ell, \ell_1, \dots, \ell_m) I(i-1 \geq c) (j+1) \nu + \\ &+ P^{(r+1,r)}(j+1, i, \ell, \ell_1, \dots, \ell_m) I(i = n) (j+1) \nu (1-H) + \\ &+ P^{(r+1,r)}(j, i, \ell-1, \ell_1, \dots, \ell_m) I(\ell > 0) (i-c+1) \alpha_1 q + \\ &+ P^{(r+1,r)}(j, i+1, \ell, \ell_1, \dots, \ell_m) I(i+1 \leq n) (i+1-c) \alpha_1 \times \\ &\quad \times \left((1-q) + q(1-H) I(\ell = v+w) \right) + \\ &+ P^{(r+1,r)}(j-1, i+1, \ell, \ell_1, \dots, \ell_m) I(i+1 \leq n, j > 0, \ell = v+w) (i+1-c) \alpha_1 q H + \\ &+ P^{(r+1,r)}(j-1, i+1, \ell+1, \ell_1, \dots, \ell_m) I(j > 0, i+1 \leq n) H \times \\ &\quad \times \left\{ I(v < \ell+1 \leq v+w) (\ell+1-v) \sigma + I(\ell_k = v_k + w_k) \alpha_2 p p_k \times \right. \\ &\quad \left. \times \left(I(\ell+1 \leq v) (\ell+1) + I(v \leq \ell+1 < v+w) v \right) \right\} + \\ &+ P^{(r+1,r)}(j, i+1, \ell+1, \ell_1, \dots, \ell_m) I(i+1 \leq n) \times \\ &\quad \times \left\{ (\ell+1-v) \sigma (1-H) I(v < \ell+1 \leq v+w) + \right. \\ &\quad \left. + I(\ell+1 \leq v) (\ell+1) \alpha_2 (1-p) + \right. \end{aligned}$$

$$\begin{aligned}
& +I(v < \ell + 1 \leq v + w)v\alpha_2(1 - p)+ \\
& + \sum_{k=1}^m I(\ell_k = v_k + w_k)\alpha_2pp_k(1 - H) \times \\
& \times \left(I(\ell + 1 \leq v)(\ell + 1) + I(v < \ell + 1 \leq v + w)v \right) \Big\} + \\
& + \sum_{k=1}^m P^{(r+1,r)}(j, i, \ell + 1, \ell_1, \dots, \ell_k - 1, \dots, \ell_m) I(\ell_k > 0) \alpha_2 pp_k \times \\
& \times \left(I(\ell + 1 \leq v)(\ell + 1) + I(v < \ell + 1 \leq v + w)v \right) + \\
& + \sum_{k=1}^m P^{(r+1,r)}(j - 1, i + 1, \ell, \ell_1, \dots, \ell_k + 1, \dots, \ell_m) \times \\
& \times I(j > 0, v_k < \ell_k + 1 \leq v_k + w_k, i + 1 \leq n) (\ell_k - v_k + 1) \sigma_k H + \\
& + \sum_{k=1}^m P^{(r+1,r)}(j, i + 1, \ell, \ell_1, \dots, \ell_k + 1, \dots, \ell_m) I(i + 1 \leq n) \times \\
& \times \left\{ I(v_k < \ell_k + 1 \leq v_k + w_k) \left((\ell_k - v_k + 1) \sigma_k (1 - H) + v_k \beta_k \right) + \right. \\
& \left. + I(\ell_k + 1 \leq v_k) (\ell_k + 1) \beta_k \right\}.
\end{aligned}$$

В приведённом соотношении запись вида $P^{(r+1,r)}(\cdot)$ означает использование при проведении вычислений последнего из найденных приближений к стационарной вероятности $P(\cdot)$, т.е., если уже известно $(r + 1)$ -ое приближение, то применяется оно, если оно ещё не найдено, то используется r -ое приближение.

3. При каждом фиксированном r вычисляется нормированная разница между последовательными значениями нормировочной константы

$$\frac{\sum_{(j,i,\ell,\ell_1,\dots,\ell_m) \in S} |P^{(r+1)}(j, i, \ell, \ell_1, \dots, \ell_m) - P^{(r)}(j, i, \ell, \ell_1, \dots, \ell_m)|}{\sum_{(j,i,\ell,\ell_1,\dots,\ell_m) \in S} |P^{(r+1)}(j, i, \ell, \ell_1, \dots, \ell_m)|}$$

и сравнивается с величиной ε , которая берётся из интервала $10^{-8} \dots 10^{-12}$. Если нормированная разница меньше чем ε , то дополнительно проверяется справедливость законов сохранения (2.5)–(2.8). Если они выполняются в пределах выбранной точности вычислений, то итерационный цикл заканчивается.

4. Вычисляется нормировочная константа

$$N = \sum_{(j,i,\ell,\ell_1,\dots,\ell_m) \in S} P^{(r+1)}(j, i, \ell, \ell_1, \dots, \ell_m)$$

и нормированные значения стационарных вероятностей

$$P(j, i, \ell, \ell_1, \dots, \ell_m) = \frac{1}{N} \times P^{(r+1)}(j, i, \ell, \ell_1, \dots, \ell_m), \quad (j, i, \ell, \ell_1, \dots, \ell_m) \in S.$$

5. В соответствии с определениями, введенными в разделе 2.3, находятся значения показателей обслуживания заявок и других характеристик работы контакт-центра.

Данная рекурсия может быть реализована для 2–3 групп консультантов при общем числе неизвестных в системе уравнений равновесия до нескольких миллионов. Наличие удобных вычислительных схем даёт возможность провести численное исследование зависимости характеристик работы контакт-центра от значений параметров входных потоков, поведения клиента и структурных характеристик центра. Результаты соответствующего анализа будут приведены в разделе 3.6 и в главе 4.

3.5. Численный анализ сходимости рекурсии

Выполним численное исследование скорости сходимости итерационного метода Гаусса-Зейделя в зависимости от значений входных параметров модели. Для иллюстрации полученных результатов рассмотрим модель контакт-центра со следующими величинами параметров: $v = 8$; $n = 23$; $m = 3$, $j_m = 20$; $H = 0,9$; $\nu = 10$; $v_1 = 2$; $v_2 = 3$; $v_3 = 2$; $w = 4$; $w_1 = 1$; $w_2 = 2$; $w_3 = 1$; $\alpha_1 = 1$; $\alpha_2 = 1$; $\beta_1 = 0,5$; $\beta_2 = 0,5$; $\beta_3 = 0,5$; $q = 0,7$; $p = 0,7$; $p_1 = 0,5$; $p_2 = 0,3$; $p_3 = 0,2$; $\sigma = 2$; $\sigma_1 = 3$; $\sigma_2 = 4$; $\sigma_3 = 5$ и найдём значения основных показателей обслуживания заявок в зависимости от изменения интенсивности первичных заявок λ . Результаты вычисления доли потерянных заявок для каждого этапа обслуживания приведены в таблице 3.1. Расчёты выполнялись с помощью итерационного метода Гаусса-Зейделя. Итерационный процесс прекращался, когда величина Δ относительной разности последовательных значений нормировочной константы достигала уровня $\varepsilon = 10^{-12}$. Общее число неизвестных в (3.4) при данном выборе входных параметров принимало значение 327600. Зависимость величины Δ от номера итерации и значения λ показана на рисунке 3.4.

Приведённые данные показывают хорошую скорость сходимости итерационной процедуры. Число итераций не превышало 1000. Сходимость увеличивается при уменьшении или возрастании интенсивности поступления первичных вызовов. На начальных шагах итерационного процесса возможны некоторые флуктуации в уменьшении относительной разницы значений нормировочной константы. Затем сходимость устанавливается и

Таблица 3.1.

Показатели качества обслуживания заявок при различных значениях интенсивности поступления первичных заявок

λ (Эрл)	π_a	π_o	$\pi_{c,1}$	$\pi_{c,2}$	$\pi_{c,3}$	π_g
4	0,002438	0,004281	0,400050	0,091349	0,156163	0,131321
6	0,140692	0,033008	0,537972	0,186368	0,257830	0,313828
8	0,495021	0,070095	0,597741	0,242141	0,311918	0,619636
10	0,704680	0,090706	0,617405	0,262851	0,331334	0,783146

далее относительная разница значений нормировочной константы монотонно убывает. Полученные результаты позволяют рекомендовать итерационный метод Гаусса-Зейделя для проведения численного анализа марковских моделей контакт-центров.

3.6. Зависимость характеристик от параметров модели

Алгоритм, построенный в предыдущих разделах главы, даёт возможность выполнить численное исследование зависимости основных показателей качества обслуживания заявок от значений входных параметров модели. Результаты анализа могут быть использованы для изучения особенностей построения и работы действующих и перспективных контакт-центров. К ним относятся: наличие предварительно голосового обслуживания; разделение операторов по классификации; учет возможности повторения заявки из-за нехватки линий доступа, недостатка операторов или консультантов, наличия ограничения на длительность ожидания начала обслуживания.

Рассмотрим использование модели для анализа условий возникновения лавинообразного роста входного потока заявок, инициированного реакцией абонента на отказ в установлении соединения с оператором контакт-центра. Подобная ситуация имеет место при увеличении интенсивности поступления первичных вызовов и возрастании настойчивости абонента в попытках получить обслуживание. В обоих случаях резкий рост интенсивности входного потока обусловлен увеличением доли повторных заявок в общем потоке поступающих вызовов.

Приведём численные примеры, иллюстрирующие анализируемое явление. Для выполнения расчётов воспользуемся обобщённой моделью контакт-центра со следующими фиксированными значениями входных параметров: $\alpha_1 = 1$; $\alpha_2 = 0,5$; $\beta_1 = 0,2$; $q = 0,7$; $p = 0,3$; $m = 1$; $p_1 = 1,0$; $\sigma = 1$; $\sigma_1 = 1$; $H = 0,9$; $\nu = 10$; $n = 50$; $v = 15$; $v_1 = 2$; $w = 7$; $w_1 = 2$; $j_m = 50$. Чтобы не усложнять запись значений параметров, выберем для определенности в качестве единицы времени длительность обслуживания заявки в

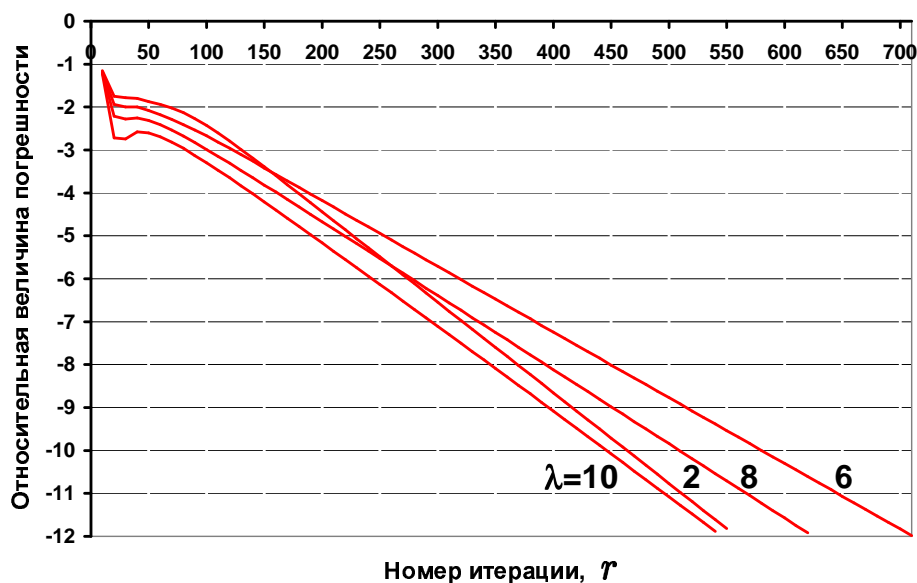


Рисунок 3.4. Зависимость относительной разности последовательных значений нормировочной константы от номера итерации и λ

IVR. Значения оставшихся параметров соответствуют результатам измерений или предположений, которые обсуждались в разделе 2.2.2. На рисунке 3.5 показано изменение величины δ_r доли повторных заявок в общем потоке поступающих вызовов от увеличения λ . На рисунке 3.6 для тех же фиксированных значений входных параметров показано изменение величины δ_r от увеличения H .

Полученные численные данные показывают темпы роста доли повторных вызовов в анализируемых асимптотических условиях и дают возможность численно оценить отрицательные последствия этого явления. В наибольшей степени лавинообразный рост трафика наблюдается с увеличением настойчивости абонента в установлении соединения. Изменение доли потерянных заявок в этих же условиях представлены на рисунке 3.10. В исследуемых обстоятельствах рост интенсивности поступления заявок не вызывает существенного увеличения оплачиваемого трафика, поскольку большая часть заявок — повторные. Не принося дохода и являясь по сути паразитным трафиком, порождённым плохой работой контакт-центра, повторные заявки наряду с первичными обрабатываются системой распределения вызовов контакт-центра, быстро вводя её в состояние перегрузки со всеми вытекающими отсюда отрицательными последствиями.

Теперь рассмотрим как меняется структура общего потока поступающих заявок в зависимости от изменения перечисленных выше параметров и этапа установления соединения. Рассмотрим следующие характеристики поступающего потока заявок: интенсивность поступления первичных заявок — λ ; добавим к этой величине интенсивность поступления повторных заявок, обусловленных отказами из-за занятости линий досту-

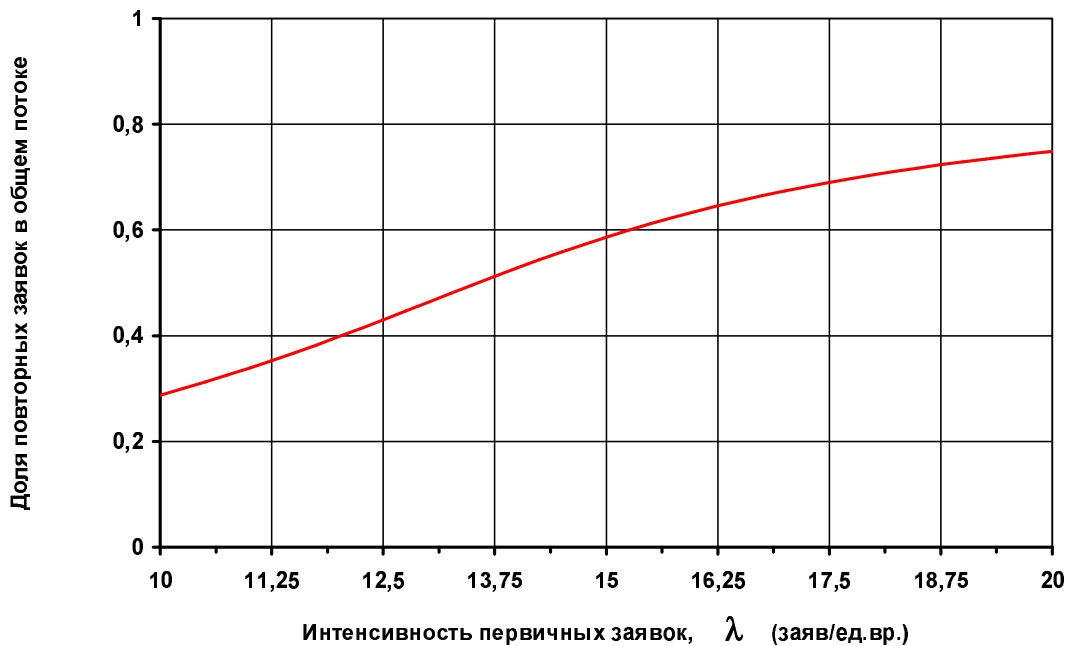


Рисунок 3.5. Темпы роста доли повторных заявок в общем потоке поступающих вызовов с увеличением интенсивности поступления первичных вызовов

па, получим значение $\lambda + \Lambda_{a,bl}H$; добавим к этой величине интенсивность повторных заявок, обусловленных отказами из-за занятости операторов и мест ожидания, а также неудачного завершения времени ожидания освобождения операторов, получим значение $\lambda + \Lambda_{a,bl}H + \Lambda_{o,bl}H$; добавим к этой величине интенсивность поступления повторных заявок, обусловленных отказами из-за занятости консультантов и мест ожидания, а также неудачного завершения времени ожидания освобождения консультантов, получим $\lambda + \Lambda_{a,bl}H + \Lambda_{o,bl}H + \Lambda_{c,bl}H$. Перечисленные характеристики показывают увеличение интенсивности входного потока заявок в зависимости от учета дополнительных потоков повторных вызовов, обусловленных потерями заявок на промежуточных этапах установления соединения. Все введенные характеристики показаны на рисунках 3.7 и 3.8, где анализируется их зависимость от изменения λ и H . Вторая и третья из перечисленных характеристик обозначены, соответственно, цифрами 1 и 2.

Сравнение различных причин образования повторных заявок показывает, что наибольший вклад дают потери на доступе, следующее по величине слагаемое дают потери из-за недостаточного количества операторов и мест ожидания освобождения операторов, и наименьший вклад дают потери из-за недостаточного количества консультантов и мест ожидания освобождения консультантов.

Помимо анализа особенностей функционирования современных справочно-информационных систем возможность точной оценки основных показателей обслуживания поступающих заявок может быть использована для обоснования методик построения прибли-

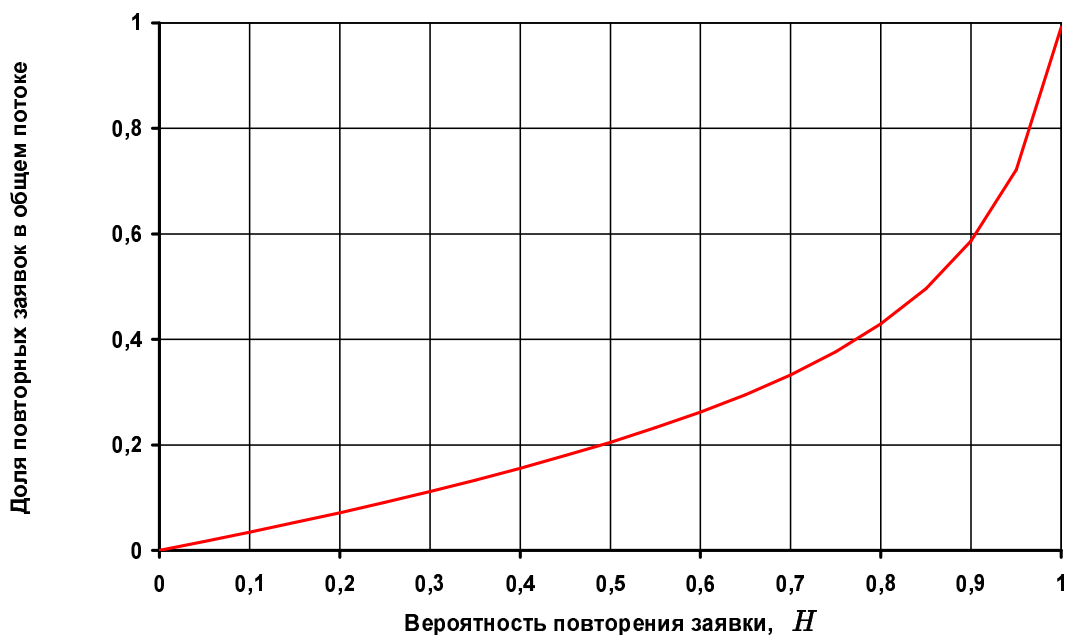


Рисунок 3.6. Темпы роста доли повторных заявок в общем потоке поступающих вызовов с увеличением вероятности повторной попытки

женных алгоритмов расчета контакт-центра. Проведём необходимые численные исследования. Обратимся к модели с параметрами, использованными при построении зависимостей, показанных на рисунках 3.5 и 3.6.

Для проведения исследований выберем зависимость доли заявок, потерянных по разным причинам, от значений интенсивности поступления первичных вызовов λ , вероятности повторения заявки после получения отказа в обслуживании H и интенсивности повторения ν . Напомним, что в модели рассматривается четыре типа характеристик, оценивающих прохождение заявкой определенного этапа обслуживания. Это: π_a — доля первичных и повторных заявок, потерянных из-за занятости линий доступа; π_o — доля заявок, потерянных либо из-за занятости всех операторов и мест ожидания их освобождения, либо в результате выполнения ограничения на максимально возможное время ожидания освобождения операторов; $\pi_{c,k}$ — доля заявок, потерянных либо из-за занятости всех консультантов k -ой группы (в исследуемом случае рассматривается одна группа консультантов, поэтому индекс номера группы опустим) и мест ожидания их освобождения, либо в результате выполнения ограничения на максимально возможное время ожидания освобождения консультантов; π_g — доля всех потерянных по разным причинам заявок от общего числа поступивших.

На рисунке 3.9 показана зависимость перечисленных характеристик от изменения λ , на рисунке 3.10 — от изменения H и на рисунке 3.11 — от изменения ν . На рисунках 3.9–3.11 кривая с номером 1 соответствует изменению значений π_a , кривая с номером 2 —

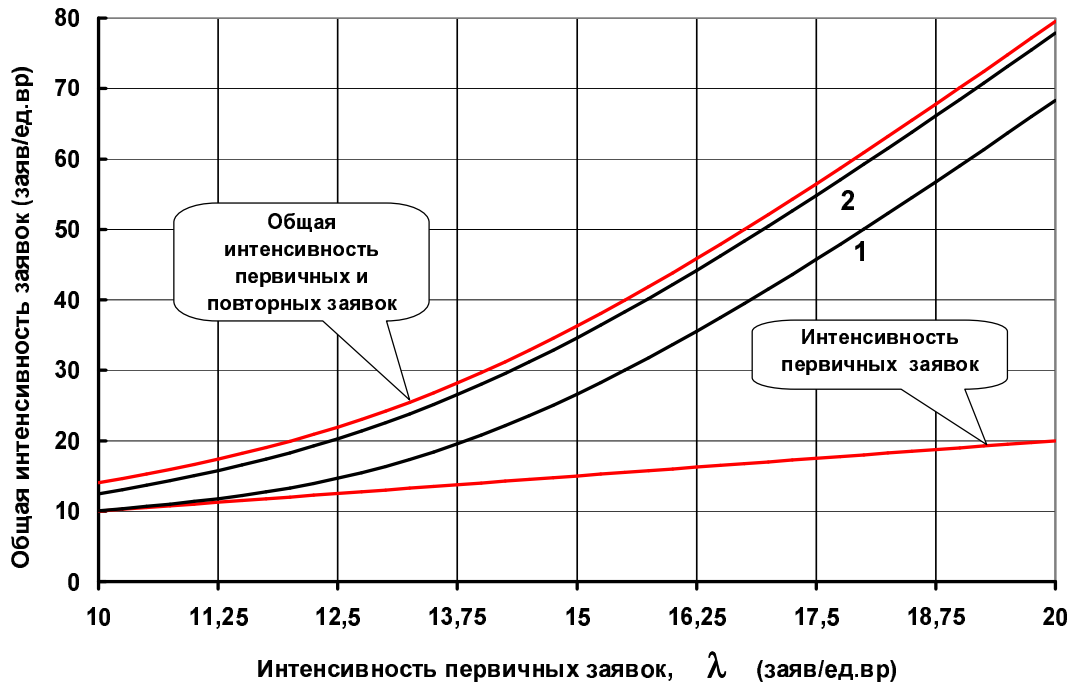


Рисунок 3.7. Интенсивности первичных и повторных вызовов, дифференцированные по этапам обслуживания в зависимости от λ : кривая 1 — значение $\lambda + \Lambda_{a,bl}H$, кривая 2 — значение $\lambda + \Lambda_{a,bl}H + \Lambda_{o,bl}H$

изменению значений π_o , кривая с номером 3 — изменению значений π_c и кривая с номером 4 — изменению значений π_g . Анализ представленных численных результатов позволяет сделать следующие выводы:

- Наиболее сильно величина доли потерянных заявок зависит от увеличения интенсивности первичных вызовов и настойчивости абонента к установлению соединения и существенно в меньшей степени от изменения значения $\frac{1}{\nu}$ длительности интервала времени между последовательными повторными попытками одного абонента. Отметим, что зависимость характеристик от вероятности повторения резко возрастает при стремлении H к единице. В этой ситуации начинается лавинообразный рост интенсивности повторных попыток, приводящей к резкому ухудшению показателей обслуживания заявок.
- Устойчивость доли потерянных заявок к изменениям ν , позволяет рекомендовать для приближенной оценки характеристик работы контакт-центров метод, основанный на пуассоновской замене потока повторных вызовов на пуассоновский с интенсивностью, вычисляемой либо решения системы неявных уравнений, следующей из законов сохранения (2.5)–(2.8), либо из выражений, получаемых из соотношения

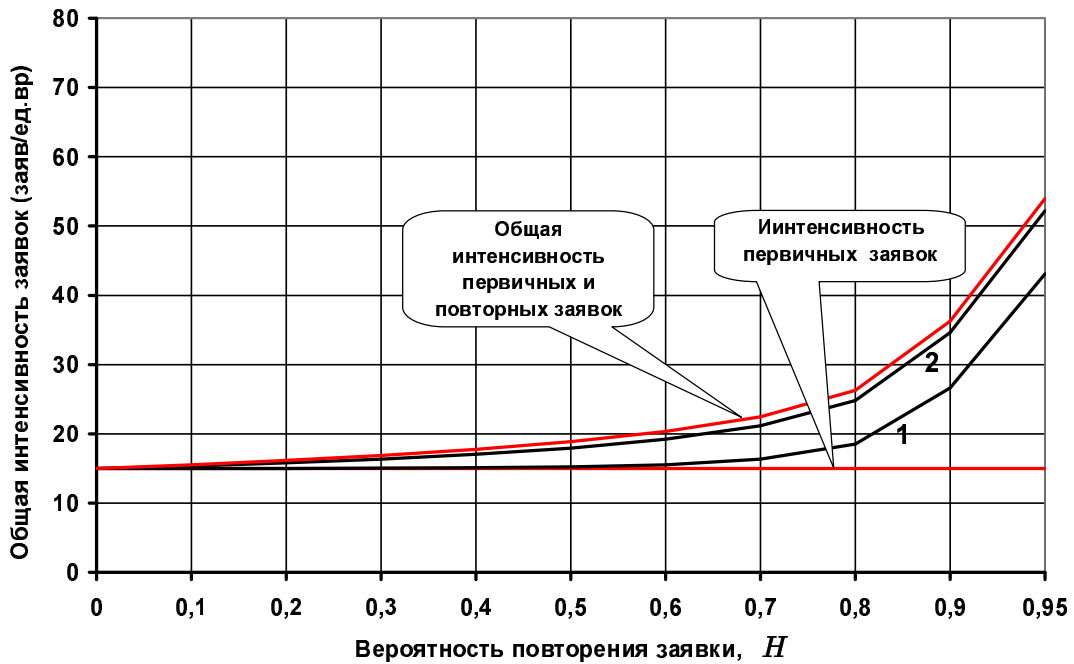


Рисунок 3.8. Интенсивности первичных и повторных вызовов, дифференцированные по этапам обслуживания в зависимости от H : кривая 1 — значение $\lambda + \Lambda_{a,bl}H$, кривая 2 — значение $\lambda + \Lambda_{a,bl}H + \Lambda_{o,bl}H$

(2.15)–(2.19) и результатов измерения величины π_g . Алгоритм расчёта характеристик работы контакт-центра, основанный на этой идее, будет построен в разделе 4.2.

- Из анализируемых характеристик наиболее сильно от значений входных параметров зависят доля заявок, потерянных на доступе, в меньшей степени доля заявок, потерянных на обслуживании у операторов и в еще меньшей степени доля заявок, потерянных на обслуживании у консультантов. Сглаживание потока заявок, при прохождении последовательных этапов обслуживания в контакт-центре, позволяет рекомендовать для приближенной оценки характеристик метод, основанный на реализации процедуры декомпозиции, когда характеристики обслуживания заявок на отдельных этапах рассчитываются независимо друг от друга, а взаимное влияние этапов учитывается через использование соответствующим образом подобранных поправочных коэффициентов. Алгоритм расчёта характеристик работы контакт-центра, основанный на этой идее, будет построен в разделе 4.3.

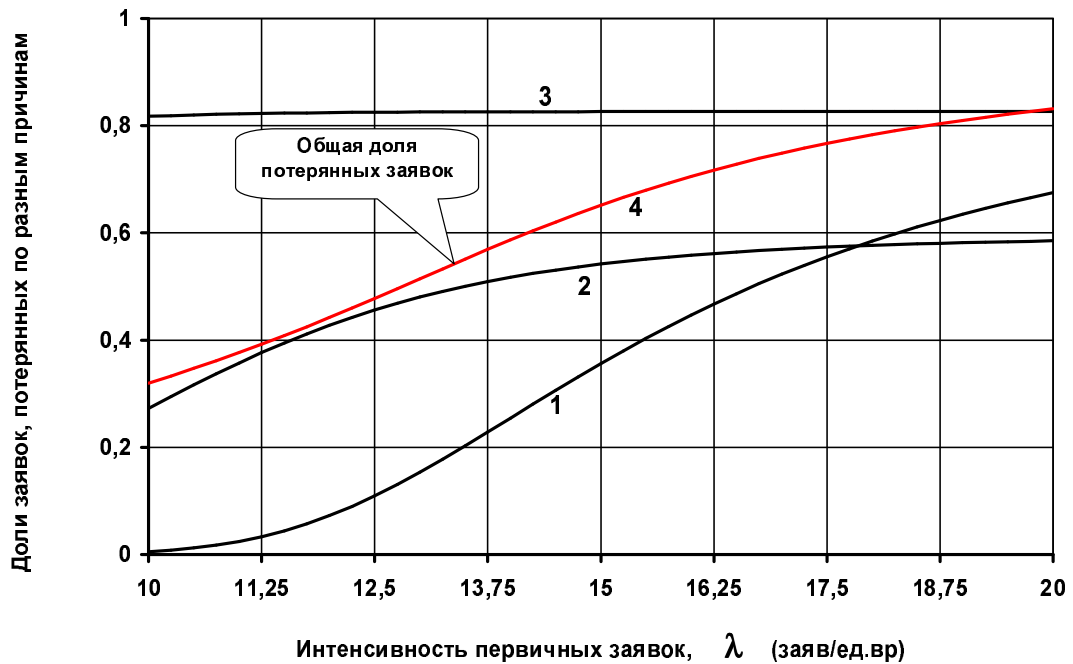


Рисунок 3.9. Зависимость доли заявок, потерянных по разным причинам, от λ : кривая 1 — значение π_a , кривая 2 — значение π_o , кривая 3 — значение π_c и кривая 4 — значение π_g

3.7. Выводы по результатам третьей главы

1. Анализ выполненных исследований показал, что наиболее эффективным способом оценки показателей обслуживания заявок и других характеристик качества работы контакт-центров являются алгоритмы, основанные на составлении и последующем решении системы линейных уравнений статистического равновесия с использованием методов вычислительной математики. Изучение свойств матрицы системы уравнений равновесия, среди которых важнейшими являются большое количество элементов, разреженность, наличие рекуррентных зависимостей для оценки коэффициентов, дало основания рекомендовать для её решения итерационный метод Гаусса-Зейделя. Сформулированы основные этапы реализации алгоритма для обобщенной модели контакт-центра (выбор начальных значений итерационного процесса, формула для рекурсивного вычисления последовательных приближений, критерий останова итерационного процесса).
2. Для исследуемой обобщенной модели контакт-центра получено алгебраическое представление системы уравнений равновесия в виде удобном для последующей реализации метода Гаусса-Зейделя. Полученное выражение даёт возможность записать все уравнения системы уравнений равновесия в виде одного соотношения с ко-

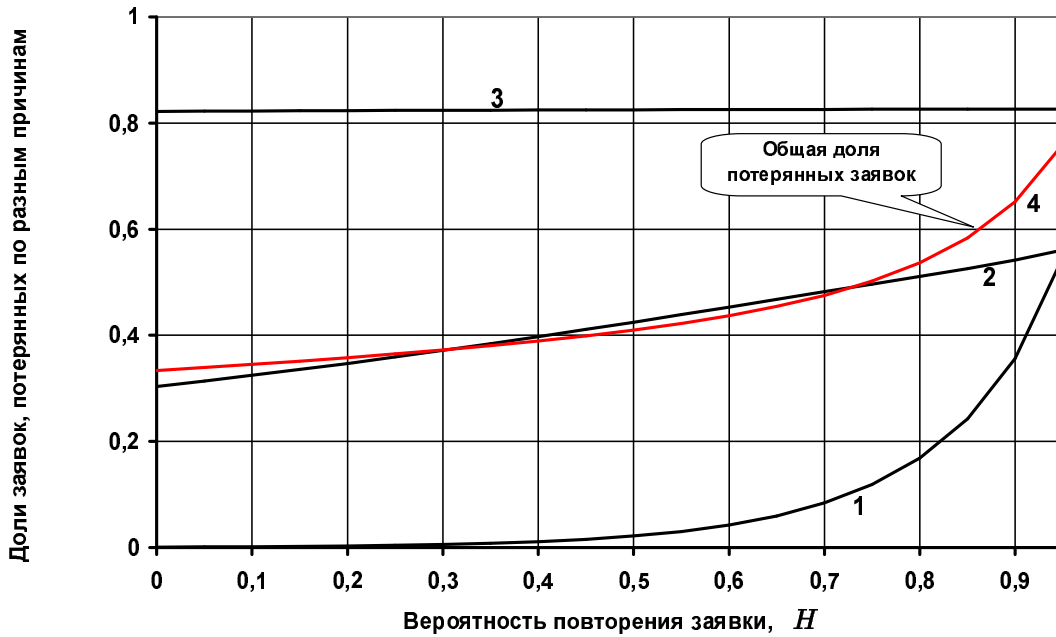


Рисунок 3.10. Зависимость доли заявок, потерянных по разным причинам, от H : кривая 1 — значение π_a , кривая 2 — значение π_o , кривая 3 — значение π_c и кривая 4 — значение π_g

эффицентами, вычисляемыми с помощью рекуррентных формул, зависящих от компонент состояния модели. Это позволяет вычислять элементы матрицы в процессе реализации итерационного алгоритма, а не хранить их в памяти компьютера. Всё это значительно упрощает реализацию метода и даёт возможность увеличить число состояний в исследуемой модели до нескольких миллионов.

3. Численное исследование сходимости итерационной процедуры показало зависимость скорости сходимости от значений входных параметров. При количестве неизвестных порядка одного миллиона число итераций не превышало 1000 при относительной погрешности вычисления нормировочной константы 10^{-12} . Сходимость увеличивается при уменьшении или возрастании интенсивности поступления первичных вызовов. На начальных шагах итерационного процесса возможны некоторые флуктуации в уменьшении относительной разницы значений нормировочной константы. Затем сходимость устанавливается и далее относительная разница значений нормировочной константы монотонно убывает.
4. Численное исследование зависимости характеристик от значений входных параметров показало устойчивость доли потерянных заявок к изменениям ν . Это дает основание рекомендовать для приближенной оценки характеристик работы контакт-центров метод, основанный на замене потока повторных вызовов на пуассоновский

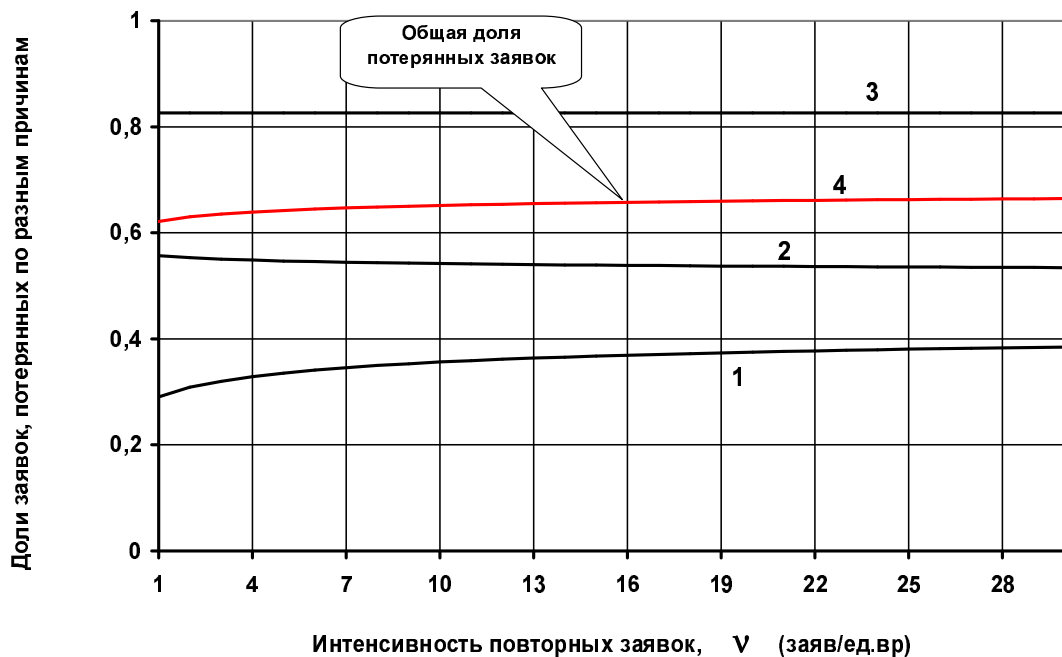


Рисунок 3.11. Зависимость доли заявок, потерянных по разным причинам, от ν : кривая 1 — значение π_a , кривая 2 — значение π_o , кривая 3 — значение π_c и кривая 4 — значение π_g

с интенсивностью, вычисляемой либо путем решения системы неявных уравнений, следующей из законов сохранения (2.5)–(2.8), либо из выражений, получаемых из соотношения (2.15)–(2.19) и результатов измерения величины π_g . Алгоритм расчёта характеристик работы контакт-центра, основанный на этой идее, будет построен в разделе 4.2.

- Из анализируемых характеристик наиболее сильно от значений входных параметров зависят доля заявок, потерянных на доступе, в меньшей степени доля заявок, потерянных на обслуживании у операторов и в еще меньшей степени доля заявок, потерянных на обслуживании у консультантов. Сглаживание потока заявок, при прохождении последовательных этапов обслуживания в контакт-центре, позволяет рекомендовать для приближенной оценки характеристик метод, основанный на реализации процедуры декомпозиции, когда характеристики обслуживания заявок на отдельных этапах рассчитываются независимо друг от друга, а взаимное влияние этапов учитывается через использование соответствующим образом подобранных поправочных коэффициентов. Алгоритм расчёта характеристик работы контакт-центра, основанный на этой идее, будет построен в разделе 4.3.

Глава 4

Использование разработанной модели для решения задач планирования и повышения эффективности работы контакт-центра

4.1. Введение

Построенную в предыдущих главах модель можно использовать для решения ряда задач, относящихся к практике построения и эксплуатации современных контакт-центров. Решение практических проблем требует разработки инженерных методик, которые применимы для любых реальных значений входных параметров. В основе подобных решений лежит использование приближённых алгоритмов оценки показателей обслуживания заявок. В разделе 4.2 разработана процедура расчёта модели, основанная на применении асимптотических значений характеристик, полученных при стремлении к бесконечности длины интервала времени между последовательными повторными попытками одного абонента. В разделе 4.3 излагаются особенности оценки характеристик обобщённой модели контакт-центра на основе реализации принципов декомпозиции исследуемой модели на отдельные части. В следующем разделе исследована погрешность, вносимая в оценку числа операторов отсутствием средств дифференцированного анализа числа поступающих заявок в зависимости от его типа: первичная или повторная. В разделе 4.5 предложена процедура определения необходимого по нагрузке числа линий доступа и операторов. В следующем разделе анализируется влияние использования устройств IVR на уменьшение загрузки операторов и консультантов. В разделе 4.7 сформулированы выводы по 4-ой главе.

4.2. Асимптотические методы

4.2.1. Основные положения

Большинство приближённых методов, используемых в теории телетрафика для расчёта моделей систем связи с учётом повторных вызовов, основаны на асимптотических свойствах исследуемых моделей, когда в качестве оценок вероятностных характеристик берутся их предельные значения, полученные при стремлении одного или нескольких входных параметров к границам своих интервалов изменения. Выбор параметров, используемых для построения асимптотических методов, диктуется возможностями расчёта получаемой упрощённой модели и значением погрешности оценки вероятностных характеристик. Решение этих проблем во многом опирается на качественные свойства исследуемой системы связи. Покажем это на примере анализируемой обобщённой модели контакт-центра.

Устремим величину ν к нулю. Обратное значение данного параметра определяет среднюю длительность интервала времени между последовательными повторными заявками, поступающими от одного абонента. Когда длина этого интервала стремится к бесконечности, то зависимость между моментами последовательных поступлений повторных требований уменьшается. В анализируемой ситуации поток повторных вызовов по свойствам приближается к пуассоновскому с некой неизвестной интенсивностью, которая зависит от загрузки контакт-центра и параметров поведения абонента, получившего отказ в обслуживании. Положим эту информацию в основу приближенного метода, который будем использовать для оценки характеристик исходной модели. Понятно, что точность метода зависит от величины ν . Чем меньше ν , тем она выше. Хорошую точность метода можно ожидать и для достаточно больших значений ν , если на интервале времени длины $\frac{1}{\nu}$ система часто меняет свое состояние. Последнее свойство выполняется для моделей со сложной структурой, когда в ней имеется большое число этапов обслуживания заявок, обслуживающих устройств и входных потоков.

Будем предполагать, что поступление повторных вызовов подчиняется закону Пуассона с некой неизвестной интенсивностью. Чтобы не усложнять запись формул, будем использовать для обозначения оценок те же символы, которые были введены при определении соответствующих вероятностных характеристик исходной модели. Из перечисленных допущений следует, что для оценки показателей обслуживания заявок обобщённой модели контакт-центра будет использоваться 3-ая вспомогательная модель, введённая в разделе 2.5.6, но с изменённой интенсивностью поступления первичных вызовов. Для оценки её величины необходимо составить и решить неявное уравнение, вытекающее из соотношения, определяющего величину интенсивности поступления повторных вызовов.

4.2.2. Расчётные соотношения

Поскольку исследуемая модель контакт-центра имеет достаточно громоздкий вид, выполним подробное изложение приближённого алгоритма на примере её частного случая, когда в ней отсутствуют устройства IVR и консультанты, и только затем рассмотрим этапы реализации алгоритма для обобщённой модели. Упрощённая подобным образом модель носит название 2-ая вспомогательная модель контакт-центра. Она была введена и исследована в разделе 2.5.5. В соответствии с положениями анализируемого расчётного метода для приближённого вычисления её характеристик предлагается использовать 1-ую вспомогательную модель с увеличенным значением интенсивности поступления заявок [58,61,63]. Предполагается, что входной поток заявок состоит из двух независимых компонент: пуассоновского потока первичных вызовов интенсивности λ и пуассоновского потока повторных вызовов интенсивности $x - \lambda$. Суммарный поток поступающих заявок подчиняется пуассоновскому закону и имеет интенсивность x .

Чтобы упростить запись промежуточных выкладок и окончательных формул в качестве единицы времени, примем среднюю продолжительность выдачи справочной информации $\frac{1}{\alpha}$. Таким образом, длительность промежутка времени занятия оператора имеет экспоненциальное распределение с параметром, равным единице. Обозначим через $p(i)$ стационарную вероятность занятости i операторов и мест ожидания. Используя значения характеристик 1-ой вспомогательной модели, нетрудно получить выражения для оценки соответствующих характеристик 2-ой вспомогательной модели, перечисленных в (2.29). Приведём часть этих выражений:

$$\pi_{t,o} = p(v+w); \quad M_\ell = \sum_{i=1}^v p(i)i + v \sum_{i=v+1}^{v+w} p(i); \quad (4.1)$$

$$M_w = \sum_{i=v+1}^{v+w} p(i)(i-v); \quad \pi_g = \pi_{t,o} + \frac{\sigma}{x} \cdot M_w.$$

Значения оставшихся оценок характеристик получаются с использованием определений и соотношений, введённых в разделах 2.3, 2.4, 2.5.5.

Для учёта зависимости характеристик 1-ой вспомогательной модели $\pi_{t,o}$, M_ℓ и M_w от значений неизвестной интенсивности входного потока x будем использовать обозначения $\pi_{t,o}(x)$, $M_\ell(x)$ и $M_w(x)$. Установим характер зависимости функций $\pi_{t,o}(x)$, $M_\ell(x)$ и $M_w(x)$ от x . Для этого достаточно обратиться к полученному далее соотношению (4.8), откуда следует, что $\pi_{t,o}(x)$ монотонно возрастает с увеличением интенсивности поступающей нагрузки x при фиксированных остальных входных параметрах. Используя преобразования, приведшие к соотношению (4.8), или качественные методы анализа систем с повторными вызовами [58], можно показать, что указанное свойство выполняется и для

выражений следующего вида

$$\sum_{i=k}^{v+w} P(i), \quad k = 1, 2, \dots, v + w. \quad (4.2)$$

Просуммировав (4.2) по $k = 1, 2, \dots, v$ и по $k = v + 1, v + 2, \dots, v + w$, получаем, что свойство монотонного возрастания с увеличением интенсивности поступающей нагрузки x выполняется, соответственно, для $M_\ell(x)$ и $M_w(x)$. Этот результат будет использован при анализе алгоритма вычисления x .

Для того, чтобы воспользоваться соотношениями (4.1) при вычислении искомых оценок характеристик, необходимо найти значение x . Величину x определим, потребовав для оценок $\pi_{t,o}(x)$, $M_\ell(x)$ и $M_w(x)$ выполнения соотношений, по форме совпадающих с законами сохранения (2.5)–(2.8). Выполнение этого требования даёт следующие равенства, связывающие характеристики и параметры 1-ой вспомогательной модели,

$$x = x\pi_{t,o}(x) + \sigma M_w(x) + M_\ell(x), \quad (4.3)$$

$$x = xH\pi_{t,o}(x) + \sigma HM_w(x) + \lambda.$$

Первое из соотношений является тождеством и выполняется для любых значений x . Выполнив несложные преобразования второго из приведенных выражений, получаем неявное уравнение для определения x

$$x = \frac{\lambda + \sigma HM_w(x)}{1 - H\pi_{t,o}(x)}. \quad (4.4)$$

Рассмотрим процедуру решения (4.4) более подробно. Вначале убедимся в существовании решения. Для этого достаточно ввести функцию

$$f(x) = x - \frac{\lambda + \sigma HM_w(x)}{1 - H\pi_{t,o}(x)},$$

представляющую собой разницу между левой и правой частями уравнения, и установить наличие значений x , при которых $f(x) = 0$. Рассмотрим поведение $f(x)$ в зависимости от изменения x . При $x = 0$ значение $f(0) < 0$. Поскольку правая часть неявного уравнения для любого x ограничена величиной $\frac{\lambda + \sigma Hw}{1 - H}$ (напомним сформулированное ранее предположение о выполнении неравенства $H < 1$ для исследуемой модели справочно—информационного центра), то при $x = \frac{\lambda + \sigma Hw}{1 - H}$ величина $f(x) > 0$. Таким образом, на концах интервала $(0, \frac{\lambda + \sigma Hw}{1 - H})$ функция $f(x)$ принимает разные знаки. Поскольку $f(x)$ непрерывная функция, то внутри интервала $(0, \frac{\lambda + \sigma Hw}{1 - H})$ существует по крайней мере одна точка x^* такая, что $f(x^*) = 0$. Существование решения доказано. Теперь покажем, что оно единственно.

Сделаем это с помощью соотношения, которое следует после вычитания из первого соотношения (4.3) второго. После несложных преобразований получаем следующее равенство

$$\lambda = x(1 - H)\pi_{t,o}(x) + (1 - H)\sigma M_w(x) + M_\ell(x). \quad (4.5)$$

Ранее было показано, что функции $\pi_{t,o}(x)$, $M_w(x)$, $M_\ell(x)$ монотонно возрастают с увеличением значения x . Допустим, что не существуют два решения уравнения x_1 , x_2 . Поскольку $x_1 \neq x_2$, то между ними должно существовать соотношение типа неравенства. Пусть $x_1 > x_2$. Однако это утверждение противоречит (4.5), поскольку левая часть (4.5) от x не зависит, а правая увеличится после подстановки увеличенного значения x . Таким образом, решение неявного уравнения единственно. Для его определения можно использовать метод деления пополам или метод секущих. Для оценки вероятностных характеристик 2-ой вспомогательной модели соответствующими характеристиками 1-ой вспомогательной модели необходимо рассчитать величины

$$\frac{\lambda + \sigma H M_w(x^{(r)})}{1 - H \pi_{t,o}(x^{(r)})}$$

для некоторой последовательности значений $x^{(r)}$, сходящейся к x^* . С вычислительной точки зрения этот приближенный алгоритм расчёта исходной модели сводится к многократному расчёту характеристик $\pi_{t,o}(x)$, $M_w(x)$ для 1-ой вспомогательной модели. Это свойство алгоритма требует выполнения работы по повышению эффективности оценки характеристик 1-ой вспомогательной модели. Решение сформулированной задачи будет рассмотрено в следующем разделе.

4.2.3. Повышение эффективности расчётов

Значения $p(i)$ находятся после реализации рекурсивных зависимостей (2.22)

$$p(i)\lambda = p(i+1)(i+1)\alpha, \quad i = 0, 1, \dots, v-1; \quad (4.6)$$

$$p(i)\lambda = p(i+1)(v\alpha + (i+1-v)\sigma), \quad i = v, v+1, \dots, v+w-1,$$

следующих из системы уравнений статистического равновесия, и последующего использования условия нормировки. Выразим значения $p(i)$ не через $p(0)$, как это обычно делается, а через $p(v+w)$. Получаем соотношение

$$p(v+w-i) = p(v+w)r_i, \quad i = 0, 1, \dots, v+w, \quad (4.7)$$

где $r_0 = 1$, а

$$r_i = \begin{cases} \frac{(v+w\sigma) \cdots (v+(w+1-i)\sigma)}{\lambda^i}, & i = 1, 2, \dots, w, \\ \frac{(v+w\sigma) \cdots (v+\sigma)v(v-1) \cdots (v+w+1-i)}{\lambda^i}, & i = w+1, \dots, w+v. \end{cases}$$

Для оценки $\pi_{t,o} = p(v+w)$ воспользуемся соотношением

$$\frac{1}{\pi_{t,o}} = 1 + \sum_{i=1}^{v+w} r_i. \quad (4.8)$$

Коэффициенты r_i с ростом значения i быстро уменьшаются по величине. Выполнение этого свойства дает возможность использовать для вычисления $\pi_{t,o}$ частную сумму приведённого выше ряда. Рассмотрим выражение

$$\frac{1}{\pi_{t,o}^{(s)}} = 1 + \sum_{i=1}^s r_i.$$

и найдем оценку сверху для относительной ошибки вычисления $\pi_{t,o}$ с помощью $\pi_{t,o}^{(s)}$. Она определяется из следующего соотношения

$$\begin{aligned} \frac{1}{\pi_{t,o}} - \frac{1}{\pi_{t,o}^{(s)}} &= \sum_{i=s+1}^{v+w} r_i \leq r_{s+1} \times \\ &\times \left(1 + \frac{v+w\sigma}{\lambda} + \dots + \frac{(v+w\sigma) \cdots (v+\sigma)v(v-1) \cdots 1}{\lambda^{v+w}} \right) = \frac{r_{s+1}}{\pi}. \end{aligned} \quad (4.9)$$

Относительная погрешность $\delta\pi$ вычисления $\pi_{t,o}$ с помощью $\pi_{t,o}^{(s)}$ оценивается сверху величиной

$$\delta\pi_{t,o} = \frac{\pi_{t,o}^{(s)} - \pi_{t,o}}{\pi_{t,o}^{(s)}} \leq r_{s+1},$$

т.е. не превосходит следующего рассчитываемого слагаемого в разложении $\frac{1}{\pi_{t,o}}$. Таким образом, последовательно вычисляя значения r_i , $i = 1, 2, \dots, s$ и накапливая сумму $1 + \sum_{i=1}^s r_i$, можно по величине следующего рассчитываемого слагаемого r_{s+1} , определить достигнутую точность оценки $\pi_{t,o}$ и остановить процесс вычисления $\pi_{t,o}$, если требуемая точность уже получена. Действуя подобным образом, можно построить процедуру эффективной оценки величины M_ℓ и M_w . Полученный результат позволяет существенно снизить объём вычислительной работы, поскольку при инженерных расчетах и графических построениях часто бывает достаточно иметь относительную погрешность оценки характеристик в несколько процентов.

4.2.4. Общий случай

Построение расчётного алгоритма для обобщённой модели контакт-центра происходит по тем же правилам, что были использованы для 2-ой вспомогательной модели. Для оценки характеристик обобщённой модели используются значения соответствующих характеристик 3-ей вспомогательной модели с интенсивностью поступления первичных и повторных вызовов, равной x . Напомним, что эта модель отличается от обобщённой модели отсутствием возможности повторения вызова. Обозначим величины её вероятностных характеристик теми же символами, что были использованы для исходной обобщённой модели. Формально она следует из исходной модели, если в ней положить $H = 0$.

Из соотношения (2.5) и определений характеристик получаем неявное уравнение для оценки величины x . Оно имеет следующий вид:

$$x = \frac{\lambda + (\Lambda_{o,bl}(x) + \sum_{k=1}^m \Lambda_{c,k,bl}(x))H}{1 - H\pi_{t,a}(x)}. \quad (4.10)$$

Обозначим через x^* решение (4.10). Для вычисления значения x^* можно использовать итерационный метод. Существование решения, его единственность и сходимость метода оцениваются из анализа численных данных. Теоретический анализ перечисленных задач, рассмотренный в предыдущих разделах для частного случая исследуемой модели, требует привлечения достаточно сложного математического аппарата и здесь не приводится.

Для оценки вероятностных характеристик обобщённой модели соответствующими характеристиками 3-ей вспомогательной модели необходимо рассчитать значения

$$\frac{\lambda + (\Lambda_{o,bl}(x^{(r)}) + \sum_{k=1}^m \Lambda_{c,k,bl}(x^{(r)}))H}{1 - H\pi_{t,a}(x^{(r)})}$$

для некоторой последовательности значений $x^{(r)}$, сходящейся к x^* . С вычислительной точки зрения этот приближённый алгоритм расчёта исходной модели сводится к многократному вычислению значений характеристик $\Lambda_{o,bl}(x)$, $\sum_{k=1}^m \Lambda_{c,k,bl}(x)$, $\pi_{t,a}(x)$ для 3-ей вспомогательной модели.

4.2.5. Оценка погрешности

Проведём численное исследование точности построенного приближённого метода. Для выполнения расчётов воспользуемся обобщённой моделью контакт-центра со следующими фиксированными значениями входных параметров: $\alpha_1 = 1$; $\alpha_2 = 0,5$; $\beta_1 = 0,2$; $q = 0,8$; $p = 0,15$; $m = 1$; $p_1 = 1,0$; $\sigma = 1$; $\sigma_1 = 1$; $H = 0,8$; $n = 50$; $v = 15$; $v_1 = 7$; $w = 7$; $w_1 = 4$; $j_m = 30$. В качестве единицы времени выбрана средняя продолжительность обслуживания поступающей заявки устройством *IVR*. Результаты вычисления основных показателей обслуживания заявок, введённых в разделе 2.3, представлены в таблицах 4.1–4.3 в зависимости от изменения величины ν интенсивности повторения заявки. Значение ν менялось от 1 до 10. Результаты расчёта характеристик при $\nu = 0$ соответствуют их приближённым значениям, полученным с помощью асимптотических методов, предложенных в предыдущем разделе. Они выделены жирным шрифтом. Приведённые в таблицах точные и приближённые значения характеристик найдены в области средних потерь, когда общая доля потерянных заявок составляет величину порядка 20–30 процентов. В Приложении 1 к диссертации приведены данные вычисления этих же характеристик в области больших и относительно малых потерь. В первом случае значение общих потерь составляет примерно 60 процентов, во втором случае — 10 процентов.

Анализ численных данных позволяет сделать следующие выводы:

Таблица 4.1.

Доля потерянных заявок в зависимости от этапа обслуживания и величины интенсивности повторения заявки ν

Значение ν	Показатели доступности ресурса			
	π_a	π_o	$\pi_{c,1}$	π_g
10	0,004972	0,275666	0,102575	0,233288
9	0,004820	0,275565	0,102605	0,233094
8	0,004650	0,275435	0,102640	0,232864
7	0,004454	0,275265	0,102685	0,232584
6	0,004227	0,275036	0,102742	0,232234
5	0,003958	0,274715	0,102818	0,231782
4	0,003630	0,274240	0,102925	0,231165
3	0,003213	0,273480	0,103088	0,230260
2	0,002645	0,272096	0,103367	0,228759
1	0,001735	0,268738	0,103962	0,225467
0	0,000484	0,257478	0,105890	0,215805

- Значения приведённых характеристик (за исключением π_a) являются медленно меняющимися функциями параметра ν . Десятикратное изменение величины ν меняет значение характеристик менее, чем на 10 процентов. Эти результаты получены для близких к реальным соотношениям между величинами анализируемых входных параметров модели H и ν , определяющих поведение абонента после отказа в обслуживании.
- Отмеченное свойство лежит в основе асимптотического метода оценки характеристик обобщённой модели контакт-центра. Относительная погрешность вычисления характеристик составляет величину порядка 5...10 процентов.
- Число операций, которые необходимо потратить на оценку характеристик, определяется умножением числа операций, используемых при вычислении характеристик аналогичной модели контакт-центра только без учета возможности повторения заявки, на число итераций рекурсивного решения неявного уравнения типа (4.4) и (4.10), составленного для определения значения интенсивности повторных вызовов. В зависимости от ожидаемой погрешности оценки характеристик число итераций составляет величину порядка 10–30.
- Погрешность расчётов уменьшается в области больших потерь и увеличивается в области малых потерь. Это свойство особенно заметно при оценке величины π_a . В

Таблица 4.2.

Характеристики ожидания начала обслуживания в зависимости от его этапа и величины интенсивности повторения заявки ν

Значение ν	Показатели ожидания обслуживания			
	$p\{\mathcal{W}_o > 0\}$	W_o	$p\{\mathcal{W}_1 > 0\}$	W_1
10	0,685177	0,231387	0,140637	0,101878
9	0,685860	0,231425	0,140669	0,101942
8	0,685487	0,231406	0,140651	0,101907
7	0,686322	0,231443	0,140692	0,101986
6	0,686906	0,231457	0,140724	0,102042
5	0,687673	0,231464	0,140769	0,102118
4	0,688725	0,231452	0,140838	0,102224
3	0,690271	0,231396	0,140953	0,102385
2	0,692792	0,231216	0,141173	0,102661
1	0,697781	0,230501	0,141737	0,103249
0	0,711741	0,227142	0,144092	0,105159

области сверхмалых¹ потерь (см. таблицы 4.1–4.3 и Приложение 1) погрешность приближённого вычисления величины π_a становится высокой.

- Проведённое выше обсуждение определяет область использования разработанной процедуры расчёта модели контакт-центра с величиной потерь, превышающей уровень в 10 процентов. Именно в этой области наиболее существенно проявляется влияние повторных вызовов.

4.3. Метод декомпозиции

4.3.1. Основные положения

Построенная модель относится к классу структурно-сложных моделей, поскольку одновременно рассматривает три этапа обслуживания заявок в контакт-центре и к тому же учитывает для заблокированной заявки возможности повторения запроса и ожидания начала обслуживания. Детальность анализа усложняет модель и сказывается на увеличении времени счёта особенно в случаях больших значений структурных параметров: числа линий доступа, числа операторов и консультантов. В этой ситуации предлагается

¹Менее одного процента.

Таблица 4.3.

Интегральные характеристики обслуживания заявок в зависимости от величины интенсивности повторения заявки ν

Значение ν	Интегральные показатели			
	π_r	δ_r	M	Q
10	0,05738	0,186630	0,2295	0,304270
9	0,05730	0,186475	0,2292	0,303941
8	0,05723	0,186291	0,2289	0,303549
7	0,05715	0,186067	0,2286	0,303074
6	0,05705	0,185787	0,2282	0,302481
5	0,05690	0,185425	0,2276	0,301713
4	0,05673	0,184932	0,2269	0,300669
3	0,05645	0,184208	0,2258	0,299140
2	0,05600	0,183003	0,2240	0,296611
1	0,05498	0,180260	0,2199	0,291100
0	0,05218	0,172644	0,2087	0,275193

использовать приближённые методы расчёта характеристик, а для обоснования используемых подходов и оценки точности вычислений можно использовать точные алгоритмы, разработанные для исследуемой модели в 3-ей главе.

Одним из самых известных способов построения приближённых алгоритмов является использование техники декомпозиции. Суть метода состоит в разбиении исследуемой системы, а следовательно и модели, на отдельные части, которые затем рассчитываются независимо друг от друга. Влияние отдельных функциональных блоков друг на друга учитывается введением поправочных коэффициентов, которые находятся из результатов измерений или специальных соотношений. В исследуемой модели последовательность блоков естественно выбрать из последовательности этапов обслуживания заявки. В модели рассматривается три подобных этапа: обслуживание на линии доступа, обслуживание у оператора и обслуживание у консультанта. Процедура оценки числа линий доступа будет рассмотрена в разделе 4.5. Ограничимся анализом процесса обслуживания поступающих заявок у операторов и консультантов. С этой целью будем предполагать, что задача подбора требуемого количества линий доступа решена и потерями заявок из-за их недостатка можно пренебречь.

В соответствии с сформулированным выше принципом декомпозиции, в качестве приближенной оценки доли заявок, потерянных в попытке получить обслуживание у оператора или консультанта, возьмем величину потерь рассчитанную с использованием

1-ой вспомогательной модели, из v операторов (или v_k консультантов, если рассматривается k -ая группа консультантов, $k = 1, 2, \dots, m$), на которую поступает пуассоновский поток заявок интенсивности Λ_o , если рассматриваются операторы контакт-центра, или $\Lambda_{c,k}$, если рассматривается k -ая группа консультантов. Время обслуживания заявки имеет экспоненциальное распределение с параметром α_2 , если рассматриваются операторы контакт-центра, или β_k , если рассматривается k -ая группа консультантов. Если в момент поступления заявки свободных операторов нет, то она становится на ожидание. Операторы имеют w мест ожидания начала обслуживания, а консультанты из k -ой группы — w_k мест. Время ожидания начала обслуживания ограничено случайной величиной, имеющей экспоненциальное распределение, с параметрами σ для операторов и σ_k для k -ой группы консультантов, $k = 1, 2, \dots, m$. Величины Λ_o и $\Lambda_{c,k}$ представляют собой значения суммарной интенсивности потока заявок, поступающих для обслуживания у операторов или в k -ой группе консультантов, $k = 1, 2, \dots, m$ без разделения на первичные и повторные. Значения Λ_o и $\Lambda_{c,k}$ могут быть получены в результате измерения.

4.3.2. Расчётные соотношения

Проведём численное исследование точности оценки характеристик обслуживания заявок у операторов или консультантов с использованием сформулированного подхода. Для простоты рассмотрим модель контакт-центра с одной группой консультантов $m = 1$. Будем также предполагать, что в исследуемой ситуации решена задача подбора числа линий доступа, обеспечивающего нормированные значения потерь поступающих заявок, т.е. потери заявок на уровне доступа в контакт-центр пренебрежимо малы. Фиксированные параметры модели принимают значения: $\alpha_1 = 1$; $\alpha_2 = 0,5$; $\beta_1 = 0,2$; $q = 0,7$; $p = 0,25$; $\sigma = 1$; $\sigma_1 = 1$; $H = 0,9$; $\nu = 2$; $v = 25$; $v_1 = 10$; $w_1 = 5$; $w = 10$. Значение λ меняется в пределах, обеспечивающих весь интервал изменения величины доли потерянных заявок: от малых значений до больших. Величины интенсивностей выражены через значение длительности времени обслуживания в системе IVR, принятой за единицу. При проведении вычислений величина j_m максимально возможного числа абонентов, одновременно повторяющих вызов, принимала значение 40. Число линий доступа n принято равным 100. Для выбранных значений входных параметров величина доли заявок, потерянных на доступе, менялась от 0,0000001 при $\lambda = 10$ до 0,000983 при $\lambda = 20$. В таблице 4.4 приведены точные значения доли потерянных заявок на этапе обслуживания у операторов π_o и в единственной группе консультантов $\pi_{c,1}$, полученные после решения системы уравнений равновесия, и их приближенные величины (обозначены верхним индексом *), найденные с помощью метода декомпозиции.

Результаты расчетов говорят о хорошей точности вычислений, особенно в области больших потерь. Погрешность расчётов остаётся приемлемой не только при оценке величины доли потерянных заявок, но также и при вычислении других характеристик

Таблица 4.4.

Доля потерянных заявок в зависимости от этапа обслуживания и величины интенсивности первичных вызовов λ

Значение λ	Показатели использования ресурса					
	Λ_o	π_o	π_o^*	$\Lambda_{c,1}$	$\pi_{c,1}$	$\pi_{c,1}^*$
10,0	7,1465	0,000861	0,000766	1,7849	0,126724	0,124515
12,5	9,1476	0,012515	0,010686	2,2580	0,228662	0,227152
15,0	11,5281	0,068334	0,060062	2,6848	0,314329	0,314252
17,5	14,7230	0,192806	0,183168	2,9708	0,365727	0,366191
20,0	18,8259	0,343573	0,340164	3,0891	0,385814	0,386056
22,5	23,3955	0,466777	0,466179	3,1184	0,390773	0,390832

модели, в частности, среднего числа занятых операторов M_ℓ , среднего числа занятых мест освобождения операторов M_w , а также среднего числа занятых консультантов $M_{\ell,1}$ и среднего числа занятых мест освобождения операторов $M_{w,1}$. Точные значения этих характеристик и их оценка (она обозначена верхним индексом *), полученная методом декомпозиции, приведены в таблице 4.5.

Таблица 4.5.

Точные и приближенные величины характеристик в зависимости от этапа обслуживания заявок в контакт-центре и значения λ

Значение λ	Показатели использования ресурса							
	M_ℓ	M_ℓ^*	M_w	M_w^*	$M_{\ell,1}$	$M_{\ell,1}^*$	$M_{w,1}$	$M_{w,1}^*$
10,0	14,2807	14,2821	0,0062	0,0055	7,7936	7,8133	0,2242	0,2204
12,5	18,0661	18,0997	0,1137	0,0974	8,7086	8,7255	0,5057	0,5027
15,0	21,4807	21,6714	0,6802	0,7596	9,2045	9,2055	0,8138	0,8129
17,5	23,7687	24,0524	2,5110	2,4937	9,4215	9,4146	1,0333	1,0326
20,0	24,7157	24,8440	4,8783	5,0620	9,4866	9,4827	1,1248	1,1241
22,5	24,9501	24,9780	6,7248	6,9052	9,4992	9,4981	1,1470	1,1468

Для вычисления методом декомпозиции значений характеристик, зависящих от параметров поведения абонента, получившего отказ в обслуживании, используются их формальные определения (см., раздел 2.3), законы сохранения (2.5)–(2.8) и альтернативные соотношения для их вычисления, перечисленные в разделе 2.4.4. Приведём необходимые утверждения на примере оценки значений: π_g , π_r , δ_r и M . Учитывая, что потерь

заявок на уровне линий доступа нет, получаем следующие расчётные соотношения:

$$\Lambda_{o,bl} = \pi_o \cdot \Lambda_o; \quad \Lambda_{c,1,bl} = \pi_{c,1} \cdot \Lambda_{c,1}; \quad \Lambda_a = \frac{1}{q} \cdot \Lambda_o.$$

Отсюда находим выражение для оценки π_g . Оно имеет такой вид:

$$\pi_g = q \left(\pi_o + \pi_{c,1} \cdot \frac{\Lambda_{c,1}}{\Lambda_o} \right).$$

Соотношения для оценки π_r , δ_r и M приведены в разделе 2.4.4. Они выглядят следующим образом:

$$\pi_r = \frac{\pi_g(1-H)}{(1-\pi_g H)}; \quad \delta_r = \pi_g H; \quad M = \frac{\pi_g H}{1-\pi_g H}.$$

В таблице 4.6 для модели с теми же значениями входных параметров, что были использованы при расчёте таблиц 4.4–4.5, приведены точные значения π_g , π_r , δ_r и M и их оценки, найденные с помощью перечисленных выше соотношений и результатов реализации процедуры декомпозиции. Погрешность оценки примерно та же, что была достигнута при расчёте характеристик, представленных в таблицах 4.4–4.5.

Таблица 4.6.

Интегральные показатели обслуживания заявок в зависимости от величины интенсивности первичных вызовов λ

λ	Показатели использования ресурса							
	π_g	π_g^*	π_r	π_r^*	δ_r	δ_r^*	M	M^*
10,0	0,022776	0,022051	0,002325	0,002250	0,020498	0,019846	0,020927	0,020248
12,5	0,048289	0,046729	0,005048	0,004878	0,043460	0,042056	0,045434	0,043902
15,0	0,099094	0,093274	0,010880	0,010182	0,089184	0,083947	0,097917	0,091640
17,5	0,186635	0,179875	0,022431	0,021462	0,167972	0,161888	0,201882	0,193157
20,0	0,284844	0,282430	0,038304	0,037869	0,256360	0,254187	0,344736	0,340819
22,5	0,363841	0,362786	0,054099	0,053866	0,327457	0,326507	0,486893	0,484798

Низкая погрешность оценки величин характеристик контакт-центра с использованием метода декомпозиции можно объяснить хорошей точностью процедур расчёта показателей обслуживания заявок в контакт-центре, основанных на замене потока повторных вызовов на пуассоновский с последующим применением моделей с потерями и ожиданием заблокированных заявок. Особенности реализации этого подхода для исследуемой модели контакт-центра были рассмотрены в разделе 4.2. Оба подхода не требуют разделения первичных и повторных вызовов при оценке перечисленных выше интегральных характеристик обслуживания поступающих заявок. Следует отметить, что при оценке

числа операторов и консультантов, требуемых для обеспечения заданного уровня потерь, дифференциация заявок по их типу необходима, иначе итоговый результат будет сильно завышен. Иллюстрация сформулированного положения будет рассмотрена в следующем разделе.

4.4. Оценка влияния повторных вызовов на число операторов

Первичные и повторные заявки, поступающие на обслуживание в контакт-центр, не различаются между собой. Отметим, что дифференциация заявок по истории их взаимоотношения с контакт-центром в принципе возможна, но требует значительных вычислительных усилий на анализ и хранение информации и обычно не проводится. Между тем, численное исследование, выполненное в разделе 3.6, показало, что даже для умеренных значений загрузки контакт-центра повторные заявки могут составить существенную часть от общего количества поступивших заявок с дальнейшим увеличением этой пропорции с ростом трафика и настойчивости абонента в установлении соединения. Понятно, что в данных условиях оценка канального ресурса на базе традиционных методик, основанных на моделях с потерями, где все поступившие заявки рассматриваются как первичные, может привести к значительной погрешности.

Покажем это на примере анализируемой модели. Чтобы упростить процедуру выбора значений входных параметров предположим, что в исследуемой модели контакт-центра отсутствуют консультанты. Схема функционирования модели рассмотрена в разделе 2.5.2. Традиционное решение поставленной задачи заключается в последовательном увеличении числа операторов и оценке π_g для каждого значения v при фиксированном значении Λ_a интенсивности общего потока заявок, поступающих на линии доступа. Минимальное значение числа операторов, при котором выполняется неравенство $\pi_g < \pi_{norm}$, где π_{norm} — регламентированное значение величины доли всех потерянных заявок, и будет искомым решением. При реализации традиционного метода предполагается, что значение Λ_a не зависит от числа операторов v , при котором происходит расчёт величины π_g . Такая постановка задачи значительно завышает итоговый результат. Этот факт имеет следующее объяснение. В реальности величина Λ_a состоит из двух компонент: один соответствует потоку первичных заявок (интенсивность этого потока действительно не зависит от имеющегося числа операторов и определяется возникающим потоком первичных заявок), другой — потоку повторных заявок (значение данной интенсивности определяется достаточностью используемого числа операторов с его увеличением будет уменьшаться). Учёт подобного расслоения поступающего потока заявок возможен только в рамках моделей с возможностью повторения заблокированной заявки. Приведём численный пример, иллюстрирующий решение задачи оценки числа операторов

традиционным методом и алгоритмом, основанным на учёте повторных заявок.

Рассмотрим модель контакт-центра с параметрами: $\lambda = 30$; $\alpha_1 = 1$; $\alpha_2 = 0,5$; $\beta_1 = 0,2$; $q = 0,75$; $\sigma = 1$; $H = 0,9$; $\nu = 5$; $n = 150$; $v = 35$; $w = 10$; $j_m = 40$. Поскольку в анализируемой модели контакт-центра отсутствуют консультанты, то источниками повторных вызовов могут быть лишь недостаток линий доступа, нехватка операторов и мест ожидания освобождения операторов и нетерпеливость клиентов контакт-центра, ограничивающая время ожидания. Для простоты ограничимся учетом влияния второй и третьей из упомянутых причин образования повторных вызовов. С этой целью положим число линий доступа равным 150. Будем предполагать, что значение Λ_a известно и определяется из соотношения $\Lambda_a = 50,8$ заяв/ед.вр. При использовании традиционного метода все поступающие заявки рассматриваются как первичные, т.е. вероятность повторения равна нулю. Воспользовавшись 3-ей вспомогательной моделью, находим необходимое число операторов из условия, что доля потерянных заявок не превосходит 5 %. На рисунке 4.1 показаны промежуточные результаты вычисления π_g при использовании традиционного способа оценки необходимого числа операторов (в этой ситуации минимально необходимое значение числа операторов $v=75$), и результаты решения сформулированной задачи, полученное с помощью модели с учётом эффекта повторных вызовов (минимально необходимое значение канального ресурса $v=49$). Разница составляет 26 операторов. Ошибка весьма существенная. Понятно, что погрешность в оценке необходимого количества числа операторов будет тем больше, чем больше значение вероятности повторения.

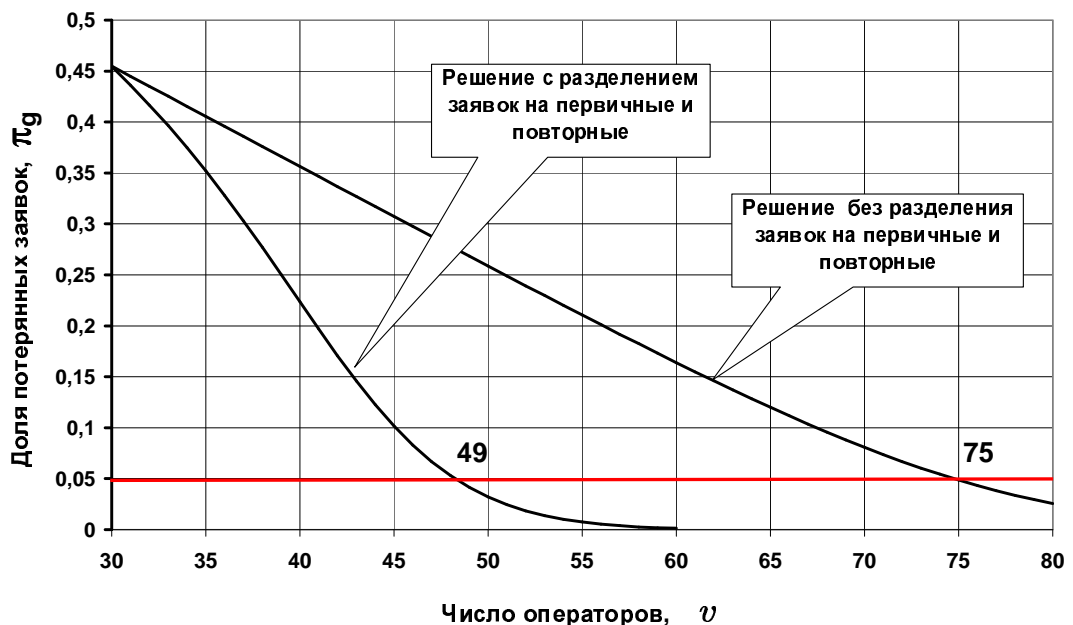


Рисунок 4.1. Результаты оценки числа операторов традиционным методом и с использованием модели с учётом возможности повторения заблокированной заявки

4.5. Планирование числа операторов и линий доступа

Решение сформулированной задачи зависит от стоимостных соотношений между основными инфраструктурными компонентами контакт-центра. Для простоты изложения последующего материала будем полагать, что в системе отсутствуют IVR и консультанты. Решение задачи в общем случае выполняется аналогичными действиями. Естественно предположить, что значения показателей CAPEX и OPEX, отнесенные на одного оператора, выше, чем значение аналогичного показателя, рассчитанного на одну линию доступа². Следовательно, экономически целесообразно увеличивать число линий доступа с одновременным уменьшением числа операторов. Выполнение указанных действий ограничено требованием о соблюдении заданных норм обслуживания абонентов: вероятности допуска заявки и ограничения на время ожидания начала обслуживания. Процедура оценки числа операторов и линий доступа разбивается на два этапа. На первом этапе число операторов принимается равным числу линий доступа и ожидание не предусмотрено. Необходимое число операторов определяется с использованием классических моделей с явными потерями заблокированных заявок или с их повторением. Полученное значение будет оценкой сверху искомой величины операторов. На втором этапе найденное значение числа операторов уменьшается с одновременным увеличением числа линий доступа и постановкой заблокированных заявок на ожидание. Рассмотрим решение сформулированных задач с большей детальностью. Начнём с исходных данных. Предполагается, что из маркетинговых исследований или измерений известны значения интенсивности поступления первичных заявок λ и параметры поведения абонента после отказа в обслуживании H и ν . В качестве единицы времени выберем среднюю продолжительность обслуживания заявки оператором. Будем считать фиксированным и известным величину $\frac{1}{\sigma}$ ограничения на длительность максимально возможного времени пребывания заявки на ожидании.

На первом этапе решается задача определения минимального числа операторов, при котором величина π_g доли потерянных заявок не превосходит заданного значения π_{norm} , являющегося нормативным показателем при решении задач планирования инфраструктуры контакт-центров. Для этого число операторов последовательно увеличивается на единицу, начиная с некоторого начального значения³ v_0 и проверяется условие $\pi_g < \pi_{norm}$. На данном этапе предполагается, что число линий доступа равно числу операторов. Для оценки π_g можно воспользоваться точными методами, основанными

²Для обоснования сформулированного утверждения можно воспользоваться результатами [104]. Согласно данным, приведенным в монографии, для типичного call-центра с числом сотрудников от нескольких десятков до нескольких сотен зарплата операторов составляет 63 % операционных расходов; обучение персонала — 6 %; аренда помещений — 5 %; аренда линий доступа — 5 %; IT-услуги и оборудование — 10 %; прочее — 1 %.

³В качестве v_0 можно взять единицу или целую часть от потенциального числа соединений λ .

на решении системы уравнений равновесия⁴, или приближенными методами, изложенными в разделах 4.2 или 4.3. Обозначим через v число операторов, найденное после выполнения указанных выше действий. На начальном шаге второго этапа значение v уменьшается на единицу и одновременно на единицу увеличивается число линий доступа. Величина π_g может оказаться меньше π_{norm} , тогда число операторов опять уменьшается на единицу с одновременным увеличением числа линий доступа. В противном случае по одной добавляются линии доступа. В результате, уменьшается значение π_g и увеличивается продолжительность времени W_o пребывания заявки на ожидании начала обслуживания у оператора. Величина W_o при фиксированных значениях числа операторов и линий доступа находится из формулы Литтла как отношение среднего числа заявок, находящихся на ожидании начала обслуживания у оператора, к интенсивности потока заявок попадающих на это ожидание. Формальное определение W_o приведено в разделе 2.3.4. Увеличение числа линий доступа происходит до тех пор, пока значение π_g не станет меньше π_{norm} при одновременном выполнении условия $W_o < W_{norm}$, где W_{norm} — регламентированное значение величины задержки начала обслуживания. Далее опять значение v уменьшается на единицу и продолжается увеличение n пока не выполнится соотношение $\pi_g < \pi_{norm}$, $W_o < W_{norm}$ и т.д.

Приведём численный пример, иллюстрирующий особенности использования сформулированного алгоритма. На рисунке 4.2 показаны результаты реализации 1-го этапа алгоритма для модели со значениями входных параметров: $\lambda = 30$; $\nu = 5$; $\alpha_2 = 1$; $\pi_{norm} = 0,03$ и $H = 0,8$. Требуемое число операторов для $H = 0,8$ определяется из соотношения $v = 40$. Для сравнения на этом же рисунке приведены результаты оценки v для других значений настойчивости абонента к установлению соединения. Рассмотрены следующие величины: $H = 0,5$; $H = 0,9$ и $H = 0,95$. Представленные численные данные показывают наличие зависимости решения от значения H особенно при больших потерях. Когда величина π_g становится близка к π_{norm} эта зависимость сильно ослабевает. Отсюда следует важный вывод. При решении задач планирования числа операторов контакт-центра нет необходимости использовать модели, учитывающие влияние повторных вызовов, поскольку эти задачи решаются в области малых потерь заявок. Это утверждение справедливо только в случае, если известно значение интенсивности поступления первичных вызовов. Часто бывают ситуации, когда подобная информация отсутствует. Это происходит в условиях перегрузки, когда потоки повторных вызовов искажают информацию об интенсивности поступающего потока первичных заявок. В этом случае модели с повторными вызовами позволяют оценить долю первичных и повторных заявок в общем потоке поступающих заявок (см. соотношение (2.14)).

Рассмотрим теперь результаты реализации 2-го этапа решения сформулированной задачи. Расчётные данные показаны на рисунке 4.3. Напомним входные параметры моде-

⁴В общем виде это система (3.4). В частных случаях — это одна из систем уравнений равновесия, приведенных в разделе 2.5.

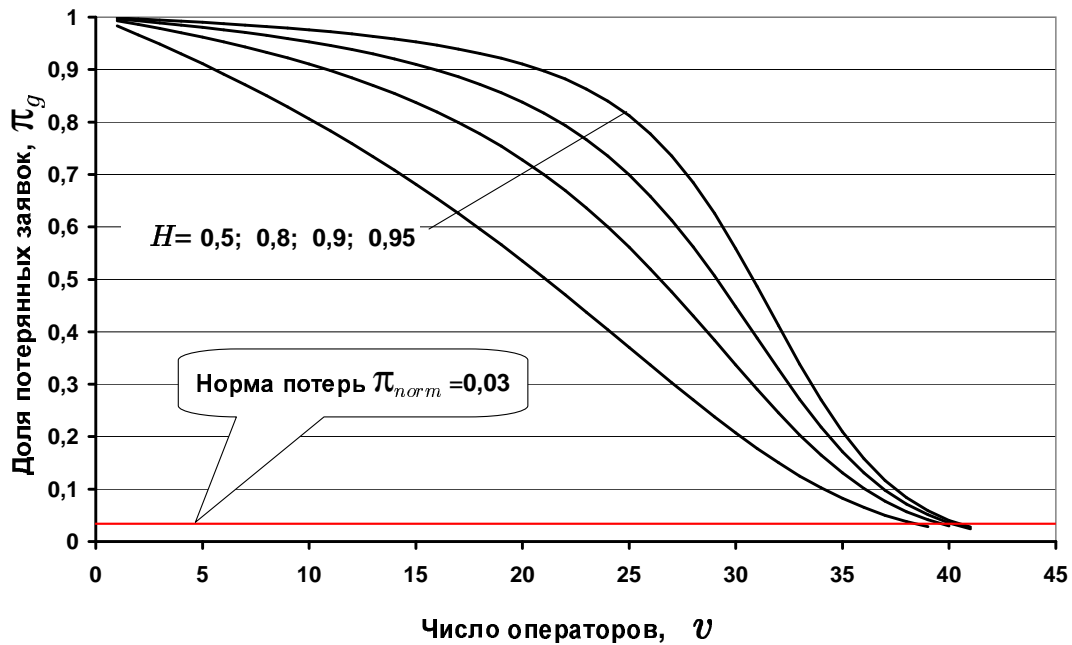


Рисунок 4.2. Результаты оценки числа операторов для $\pi_{norm} = 0,03$, $\lambda = 30$ и разных H

ли: $\lambda = 30$; $\nu = 5$; $\alpha_2 = 1$; $\pi_{norm} = 0,03$; $H = 0,8$; $\nu = 40$; $\frac{1}{\sigma} = 100$. Величину W_{norm} зададим равенством $W_{norm} = 0,5$. Из результатов вычислений видно, что с возрастанием количества уточнений числа операторов сильно увеличивается количество добавочных линий доступа. При этом возрастает и время задержки заявки. Понятно, что в этих условиях требуется раньше остановить процесс уточнения ν , воспользовавшись ограничениями на задержку заявки в системе и максимально возможное число линий доступа⁵. Выбор этих ограничений определяется экономическими соображениями, и здесь не рассматривается. Следуя данным, приведенным на рисунке 4.3, одним из возможных ответов сформулированной задачи является выбор числа линий доступа и операторов из соотношений: $n = 78$ ($w = 47$), $\nu = 31$. Для этих значений входных параметров величины характеристик, определяющих качество обслуживания заявок, находятся из следующих равенств: $\pi_g = 0,029736$; $W_o = 0,4624$.

4.6. Оценка влияния IVR на численность операторов

Использование интерактивной системы голосового ответа (IVR) дает возможность сократить число операторов контакт-центра необходимых для обслуживания поступающих заявок с заданным качеством. Особенности использования этого решения в прак-

⁵Необходимо отметить, что процесс уточнения числа операторов также ограничен требованием о существовании стационарного режима у случайного процесса, описывающего функционирование исследуемой модели. В системах с ожиданием потенциальное число соединений не должно превосходить максимальную пропускную способность системы.

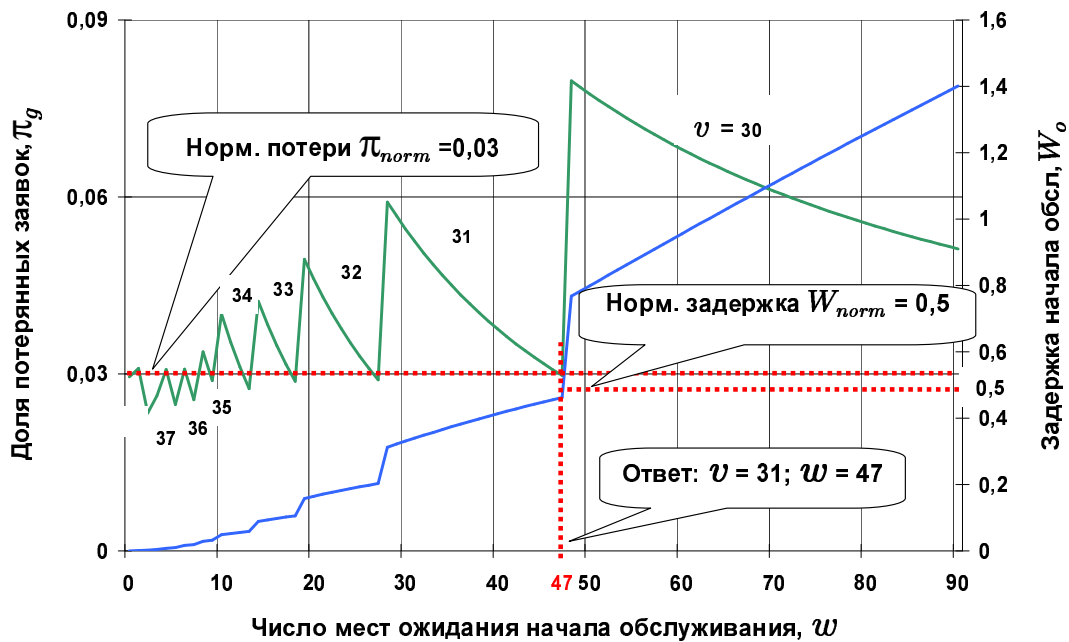


Рисунок 4.3. Результаты уточнения числа операторов v и значения задержки начала обслуживания в зависимости числа мест ожидания. Число линий доступа определяется из соотношения $n = v + w$. Единица времени — среднее время обслуживания заявки у оператора

тике работы контакт-центра рассмотрены в разделе 1.2.3. Наличие модели позволяет точно оценить число операторов, которые освобождаются после внедрения IVR и исследовать зависимость достигаемого выигрыша от различных условий, в которых происходит работа контакт-центра: модели поведения пользователя услуг контакт-центра и интенсивности входного потока заявок. Получаемую выгоду оценим разницей между числом операторов, требуемых для обслуживания поступающих заявок без использования устройств IVR, и при наличии таковых. Для простоты предположим, что в исследуемой модели отсутствуют консультанты, и речь идет об оценке числа операторов, обеспечивающих заданную величину общих потерь π_g . Схема функционирования модели рассмотрена в разделе 2.5.2. Условие достаточности количества операторов формулируется в виде необходимости выполнения неравенства $\pi_g \leq 0,05$. Фиксированные входные параметры принимают значения: $\alpha_1 = 1$; $\alpha_2 = 0,5$; $\sigma = 1$; $\nu = 5$; $n = 100$; $j_m = 40$. За единицу времени принята средняя длительность времени обслуживания устройством IVR. При фиксированной величине числа операторов значение числа мест ожидания w находится как целая часть выражения $\frac{v}{2}$.

На рисунках 4.4–4.6 показана зависимость числа операторов, необходимых для обслуживания фиксированного потока поступающих первичных заявок от доли заявок, получающих обслуживание на IVR для различных условий поступления заявок.

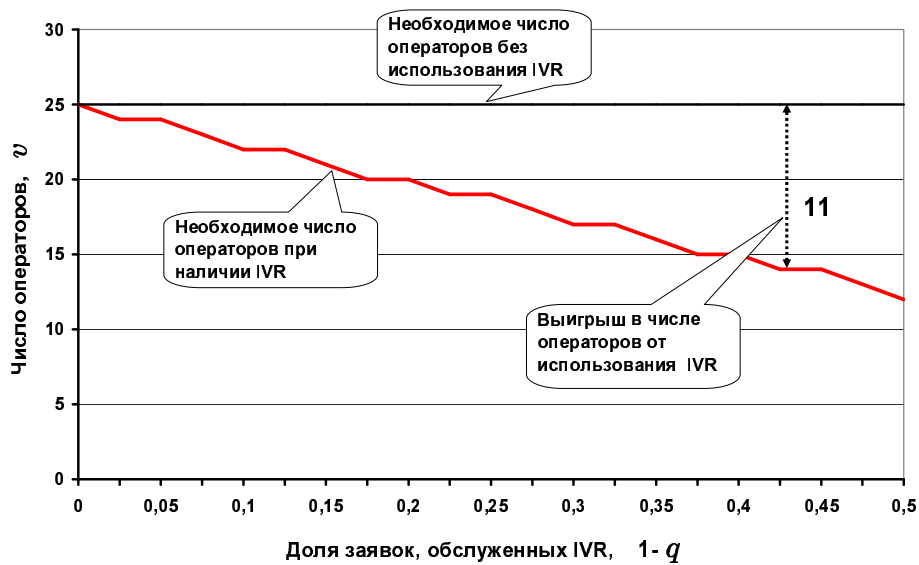


Рисунок 4.4. Зависимость числа операторов от доли заявок, завершивших обслуживание на IVR, $H = 0,9$, $\lambda = 10$

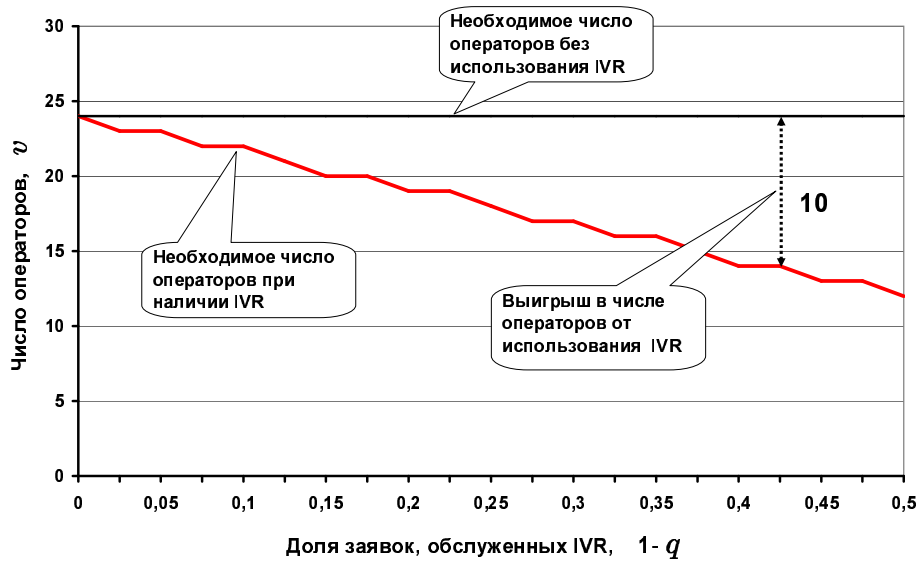


Рисунок 4.5. Зависимость числа операторов от доли заявок, завершивших обслуживание на IVR, $H = 0,5$, $\lambda = 10$

Для данных, показанных на рисунке 4.4 предполагалось, что $H = 0,9$; $\lambda = 10$. При выполнении расчётов, результаты которых приведены на рисунке 4.5, полагалось, что $H = 0,5$; $\lambda = 10$ и для данных, показанных на рисунке 4.6, $H = 0,9$; $\lambda = 20$. Анализ представленных численных результатов позволяет сделать следующие выводы:

- Выигрыш в числе операторов от внедрения устройств IVR пропорционален доле

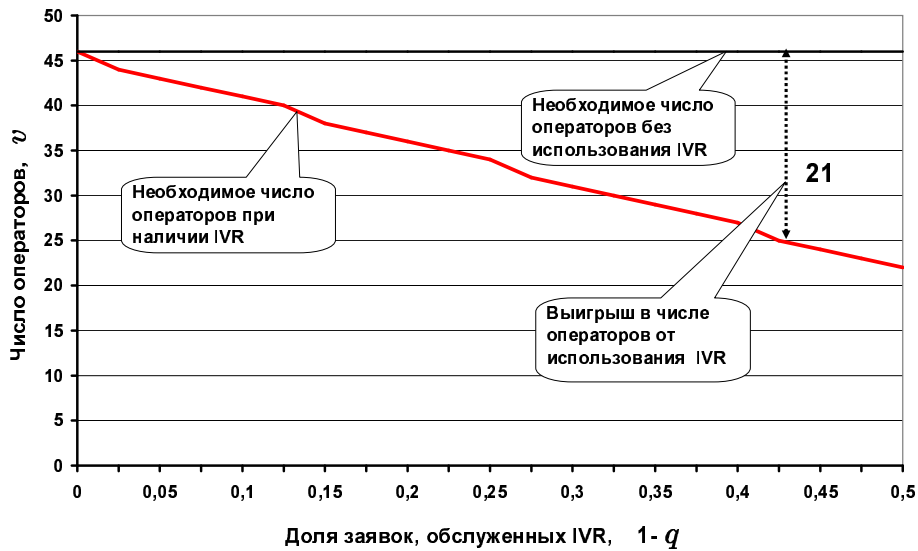


Рисунок 4.6. Зависимость числа операторов от доли заявок, завершивших обслуживание на IVR, $H = 0,9$, $\lambda = 20$

заявок, получивших завершивших обслуживание на IVR и покинувших систему и может достигать нескольких десятков процентов при величине $1 - q$ в несколько десятков процентов.

- Величина выигрыша в слабой степени зависит от вероятности повторения заблокированной заявки H и в большей — от значения вероятности завершения обслуживания в системе IVR $1 - q$ и величины интенсивности поступления первичных вызовов λ .
- Наличие модели позволяет численно оценить число операторов, которые освобождаются после внедрения IVR, и использовать полученные результаты при решении задач планирования инфраструктуры контакт-центра.

Позитивная роль IVR может выражаться не только в уменьшении числа операторов, необходимых для обслуживания поступающих потоков заявок с заданным качеством, но также и в уменьшении доли повторных заявок в общем числе заявок, поступающих в контакт-центр. Приведём соответствующие примеры. На рисунках 4.7 и 4.8 показаны значения π_g доли потерянных заявок (рисунок 4.7) и Λ_a общей интенсивности поступления заявок (рисунок 4.8) в зависимости от величины $1 - q$ доли заявок, завершивших обслуживание в системе IVR.

Фиксированные входные параметры принимают значения: $\alpha_1 = 1$; $\alpha_2 = 0,5$; $H = 0,9$; $\lambda = 12$ $\sigma = 1$; $H = 0,9$; $\nu = 5$; $n = 100$; $r_m = 40$. За единицу времени принята средняя длительность времени обслуживания устройством IVR. При $1 - q = 0$ устройство IVR не используется, и контакт-центр испытывает проблемы от большого числа повторных

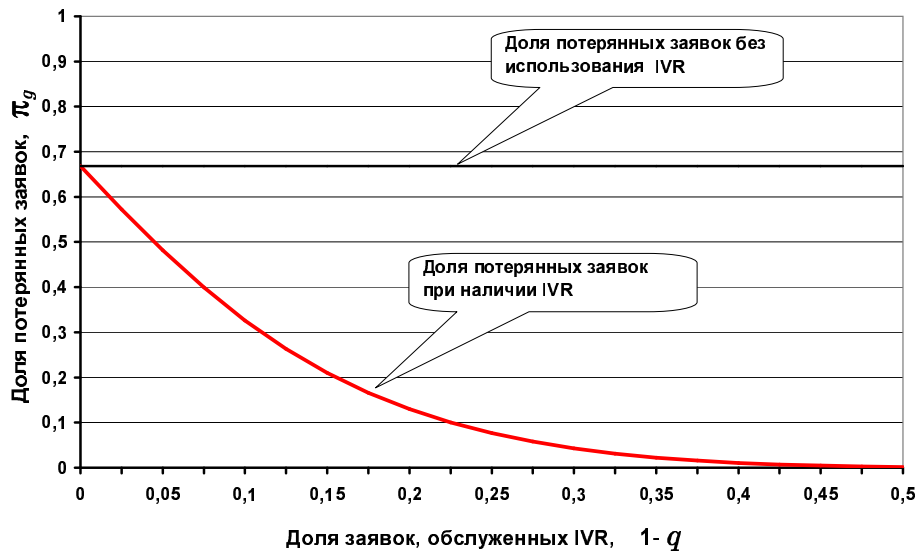


Рисунок 4.7. Зависимость доли потерянных заявок от доли заявок, завершивших обслуживание в системе IVR

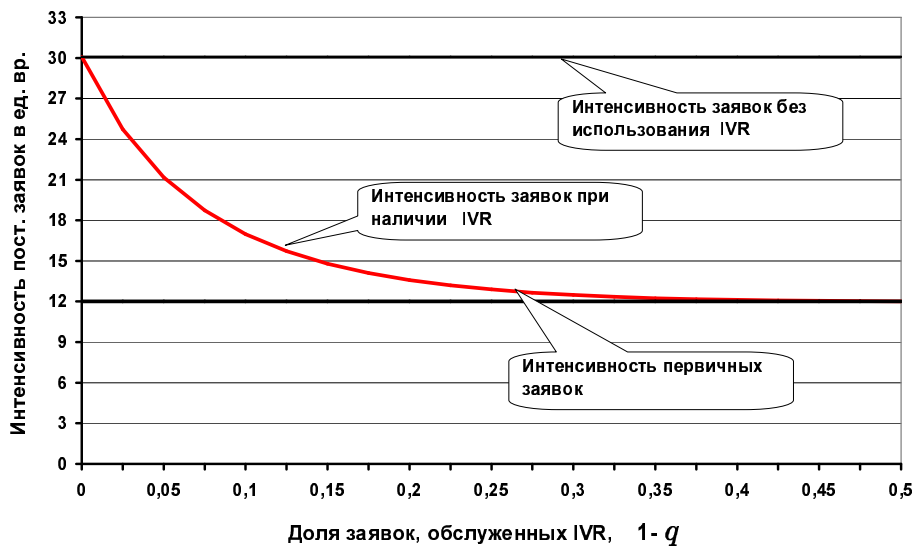


Рисунок 4.8. Зависимость интенсивности поступающих заявок от доли заявок, завершивших обслуживание в системе IVR

заявок. Далее, при подключении IVR число заявок, завершивших свое обслуживание в системе IVR увеличивается, соответственно, меньшее число заявок попадает и, соответственно, теряется в операторской группе. В результате доля потерянных заявок стремится к нулю, а интенсивность поступающих заявок стремится к интенсивности поступающих первичных вызовов.

4.7. Выводы по результатам четвертой главы

1. Разработана процедура расчёта обобщённой модели контакт-центра основанная на применении асимптотических значений характеристик, полученных при стремлении к бесконечности интервала времени между последовательными повторными попытками одного абонента. Для вычисления оценок используются характеристики аналогичной модели контакт-центра, но без учета возможности повторения вызова и с увеличенной интенсивностью поступления первичных вызовов. Результаты численного исследования точности оценки характеристик контакт-центра показали, что относительная погрешность составляет величину порядка 3–5 процентов. Погрешность расчётов уменьшается в области больших потерь и увеличивается в области малых потерь.
2. Сформулирована процедура оценки доли заявок, потерянных на доступе к операторам и консультантам, основанная на использовании техники декомпозиции. Суть метода состоит в разбиении исследуемой системы, а следовательно и модели, на отдельные части, которые затем рассчитываются независимо друг от друга. Взаимное влияние отдельных функциональных блоков учитывается введением поправочных коэффициентов, которые получаются из результатов измерений или специальных соотношений. В исследуемой модели последовательность блоков естественно выбрать из последовательности этапов обслуживания заявки. Этих этапов в модели рассматривается три: обслуживание на линии доступа, обслуживание у оператора и обслуживание у консультанта. Показано, что предложенный метод имеет высокую точность. Относительная погрешность оценки характеристик составляет несколько процентов. Точность возрастает с ростом потерь. Доля потерянных заявок при обслуживании у консультантов оценивается с большей точностью, чем доля потерянных заявок при обслуживании у операторов.
3. Численный анализ показал, что даже для умеренных значений загрузки контакт-центра повторные заявки составляют существенную часть от общего количества поступивших. С ростом трафика и настойчивости абонента в установлении соединения имеет место дальнейшее увеличение этой пропорции. В рассматриваемых условиях оценка канального ресурса на базе традиционных методик, основанных на моделях с потерями, где все поступившие заявки рассматриваются как первичные, может привести к значительной погрешности, которая может составит десятки процентов от общего объема канального ресурса.
4. Сформулирована двухэтапная процедура планирования числа операторов и линий доступа. Задача первого этапа — определение предварительного значения числа операторов при фиксированном количестве линий доступа. На втором этапе эта

величина уточняется изменением числа линий доступа. Показано, что подобным образом можно уменьшить число требуемых операторов на величину до 10%.

5. Показано, что выигрыш в числе операторов от внедрения IVR пропорционален доле заявок, получивших обслуживание на IVR и покинувших систему. Он может достигать нескольких десятков процентов при величине $1 - q$ в несколько десятков процентов. Величина выигрыша в слабой степени зависит от вероятности H повторения заблокированной заявки и в большей — от значения $1 - q$ вероятности завершения обслуживания на IVR и величины интенсивности поступления первичных вызовов λ . Стоит отметить, что использование IVR уменьшает долю повторных заявок.

Заключение

Основные результаты работы состоят в следующем.

1. Проведен анализ работы современных контакт-центров. Он показал эффективность использования устройств IVR, операторов и консультантов для дифференцированного обслуживания клиентов, а также необходимость учета повторных вызовов при формировании входного потока заявок. Исследование, выполненное по материалам опубликованных работ, показывает отсутствие методов оценки ресурса контакт-центров, учитывающих совокупное влияние перечисленных факторов. Подобные методы необходимы для научно обоснованного решения инженерных задач, возникающих при проектировании и эксплуатации справочных служб.
2. Построена и исследована новая обобщенная математическая модель контакт-центра, в которой для обслуживания клиентов используются устройства IVR, операторы и консультанты, а также учитывается возможность повторения заявки из-за отказов на различных этапах обслуживания, наличие ограничения на время пребывания на ожидании и возможность поступления заявки на обслуживание по каналам сети Интернет.
3. С использованием модели даны определения основным показателям качества обслуживания заявок, поступающих в контакт-центр. Формальные выражения для характеристик получены через значения входных параметров и величины стационарных вероятностей модели. Среди них: интенсивности и потери заявок, дифференцированные по этапам обслуживания; средние значения функциональных компонент состояния (среднее число абонентов, повторяющих заявку, среднее число занятых операторов и т.д.); вероятности и средние длительности пребывания заявок на ожидании начала обслуживания и т.д.
4. Получены соотношения между характеристиками работы контакт-центра, которые имеют характер законов сохранения интенсивностей поступающих и обслуженных системой потоков заявок. Найденные соотношения можно использовать для косвенной оценки тех входных параметров и характеристик модели, прямое измерение которых стандартной аппаратурой затруднено из-за сложностей в разделении первичных и повторных вызовов, а также для разработки приближенных методов оценки показателей качества обслуживания заявок.

5. Получено алгебраическое представление системы уравнений равновесия исследуемой модели контакт-центра в виде, удобном для последующей реализации метода Гаусса-Зейделя. Найденное выражение дает возможность записать все уравнения системы в виде одного соотношения с коэффициентами, вычисляемыми с помощью рекуррентных формул, зависящих от компонент состояния модели. Это значительно упрощает реализацию метода и дает возможность увеличить число состояний в исследуемой модели до нескольких миллионов.
6. Разработаны две процедуры приближенного расчета построенной модели контакт-центра, основанные на применении асимптотических значений ее характеристик, полученных при стремлении к нулю интенсивности повторения запросов, и на реализации техники декомпозиции. В первом случае для вычисления оценок используются значения характеристик аналогичной модели контакт-центра, но без учета возможности повторения вызова и с увеличенной интенсивностью поступления первичных вызовов. Во втором — исследуемая модель разбивается на отдельные части, которые затем рассчитываются независимо друг от друга. Влияние отдельных функциональных блоков друг на друга учитывается введением поправочных коэффициентов, которые получаются из результатов измерений или специальных соотношений. В исследуемой модели последовательность блоков выбирается из последовательности этапов обслуживания заявки (IVR, операторы, консультанты).
7. Сформулирована двухэтапная процедура планирования числа операторов и линий доступа. Задача первого этапа — определение предварительного значения числа операторов при фиксированном количестве линий доступа. На втором этапе эта величина уточняется изменением числа линий доступа. Показано, что подобным образом можно уменьшить число требуемых операторов на величину до 10%.
8. Показано, что выигрыш в числе операторов от внедрения IVR пропорционален доле заявок, получивших обслуживание на IVR и покинувших систему. Он может достигать нескольких десятков процентов при величине $1 - q$ в несколько десятков процентов.
9. Результаты диссертации использованы в ПАО МГТС в виде методики оценки необходимого числа операторов справочно-информационной службы, а также использованы в учебном процессе на кафедре «Сети связи и системы коммутации» МТУСИ. Реализация результатов диссертации подтверждается соответствующими актами.

Список литературы

1. *Башарин Г.П., Гайдамака Ю.В., Самуйлов К.Е.* Математическая теория телетрафика и ее приложения к анализу мультисервисных сетей связи следующих поколений // Автоматика и вычислительная техника. — Латвия, Рига. — 2013. — №2. — С.11- 21.
2. *Белов М.Н.* Исследование и разработка метода организации информационных центров обслуживания вызовов: Дис. ... канд. техн. наук: 05.12.13. — Москва, 2002. — 99 с.
3. *Бельская Н. М.* Исследование и разработка алгоритмов мониторинга и анализа качества работы операторов контакт-центра: Дис. ... канд. техн. наук: 05.12.13. — Самара, 2012. — 144 с.
4. *Бузюкова И.Л., Бязров А.С., Гайдамака Ю.В.* Имитационная модель узла управления услугами интеллектуальной сети // Т-Сomm: Телекоммуникации и Транспорт. — 2010. — №7. — С. 18-22.
5. *Ваняшин С.В.* Исследование и разработка моделей мультисервисного центра обслуживания вызовов: Дис. ... канд. техн. наук: 05.12.13. — Самара, 2006. — 157 с.
6. *Глушак Е.В., Росляков А.В.* Анализ гомогенной модели распределенного центра обслуживания вызовов // Инфокоммуникационные технологии. — 2013. — Т.12. — №3. — С. 23-26.
7. *Глушак Е.В., Росляков А.В.* Экспериментальное исследование модели РЦОВ на основе имитационного моделирования // Т-Сomm: Телекоммуникации и транспорт. — 2014. — Том 8. — №11. — С. 42-46.
8. *Гольдштейн Б.С., Фрейнкман В.А.* Call-центры и компьютерная телефония. — СПб.: БХВ-Санкт-Петербург, 2002. — 372 с.
9. *Гольдштейн Б. С., Соколов Н. А., Яновский Г. Г.* Сети связи. Учебник для вузов. — СПб.: БХВ-Санкт-Петербург, 2010. — 400 с.
10. *Дибс В. Н.* Модели и методы расчета мультисервисных контакт-центров: Дис. ... канд. техн. наук: 05.12.13. — Санкт-Петербург, 2011. — 125 с.
11. *Ершов В. А, Кузнецов Н. А.* Мультисервисные телекоммуникационные сети. — М.: Изд-во МГТУ им. Н. Э. Баумана, 2003. — 432 с.
12. *Зарубин А.А.* Исследование контакт-центров NGN: Дис. ... канд. техн. наук: 05.12.13. — Санкт-Петербург, 2004. — 141 с.

13. *Зарубин А., Мараев В.* От центра обслуживания вызовов к видеоконтакт-центру // *Connect! Мир связи.* — 2009. — № 4. — С. 54-57.
14. *Зелинский А. М., Корнышев Ю. Н.* Эквивалентные модели системы с повторными вызовами // *Тр. уч. ин-тов связи.* — 1976. — Вып. 80. — С. 37-42.
15. *Ионин Г. Л., Седол Я. Я.* Таблицы вероятностных характеристик полнодоступного пучка при повторных вызовах. — М.: Наука, 1970. — 154 с.
16. *Исследование рынка контакт-центров, проведенное компанией Dimension Data.* [Электронный ресурс]. Режим доступа: [http : //www.callcenterguru.ru/news/1906](http://www.callcenterguru.ru/news/1906) (дата обращения 01.03.2016).
17. *Клейнрок Л.* Теория массового обслуживания / Пер. с англ. под ред. В. И. Неймана. — М.: Машиностроение, 1979. — 452 с.
18. *Клейнрок Л.* Вычислительные системы с очередями / Пер. с англ. под ред. Б. С. Цыбакова. — М.: Мир, 1979. — 600 с.
19. *Корнышев Ю. Н.* Расчёт полнодоступной коммутационной системы с повторными вызовами // *Электросвязь.* — 1969. — №11. — С. 65-72.
20. *Корнышев Ю. Н., Пшеничников А. П., Харкевич А. Д.* Теория телетрафика. Учебник для вузов. — М.: Радио и связь, 1996. — 272 с.
21. *Крылов В. В., Самохвалова С. С.* Теория телетрафика и её приложения. — СПб.: БХВ-Санкт-Петербург, 2005. — 288 с.
22. *Кучерявый Е. А.* Управление трафиком и качество обслуживания в сети Интернет. — СПб.: Наука и техника, 2004. — 336 с.
23. *Лагутин В. С., Степанов С. Н.* Телетрафик мультисервисных сетей связи. — М.: Радио и связь, 2000. — 320 с.
24. *Леснова Л.* Инфраструктурные решения для контакт-центров // *Connect! Мир связи.* — 2010. — № 1-2. — С. 68-71.
25. *Леснова Л.* Тенденции в бизнесе call-центров // *Connect! Мир связи.* — 2010. — № 4. — С.18-19.
26. *Мишулина О.А.* Статистический анализ и обработка временных рядов. — М.: МИ-ФИ, 2004. — С. 180.
27. *Наумов В. А., Самуйлов К. Е., Яркина Н. В.* Теория телетрафика мультисервисных сетей. — М.: Изд-во РУДН, 2007. — 191 с.
28. *Орлов С.* Оптимизация ЦОВ // *Журнал сетевых решений LAN.* — 2009. — № 10. — С.32-38.
29. *Павлов А.* Особенности планирования и осуществления обзвонov в рамках современных call/contact-центров // *Connect! Мир связи.* — 2007. — № 4. — С. 54-55.

30. *Пинчук А.В., Фрейнкман В.А.* Контакт-центры в мультисервисных сетях // Технологии и средства связи. — 2002. — № 1. — С. 72-76.
31. *Приказ Министерства связи и массовых коммуникаций РФ от 12.12.2011 №340 "Об утверждении Правил применения оборудования центров обслуживания вызовов информационно-справочного обслуживания"* [электронный ресурс] // Министерство связи и массовых коммуникаций РФ [Офиц. сайт].
URL : <http://www.minsvyaz.ru/ru/documents/4065/> (дата обращения 01.03.2016).
32. *Пшеничников А.П., Васькин Ю.А., Степанов М.С.* Распределение канального ресурса при обслуживании мультисервисного трафика // Т-Comm: Телекоммуникации и транспорт.— 2009. — Том 3. — №4. — С.46–50.
33. *Пшеничников А. П., Степанов М.С.* Асимптотические свойства модели с учетом повторения заблокированных заявок // Труды конференции “Телекоммуникационные и вычислительные системы”. — 2009. — С.44–46.
34. *Пшеничников А.П., Степанов М.С.* Асимптотические модели оценки качества обслуживания вызовов call-центром методов// Т-Comm: Телекоммуникации и транспорт. —2010. — Том 4. — №7. — С.93–97.
35. *Пшеничников А.П., Степанов М.С.* Обобщенная модель call-центра // Т-Comm: Телекоммуникации и транспорт. — 2011. — Том 5. — №7. — С.125–129.
36. *Пшеничников А. П., Степанов М.С.* Особенности функционирования контакт-центров // Труды конференции “Телекоммуникационные и вычислительные системы”. — 2012. — С.32–33.
37. *Росляков А.В.* Мультисервисные платформы сетей следующего поколения NGN. Т. 1. Отечественные системы: — Самара: Ас Гард, 2012. — 312 с.
38. *Росляков А.В., Самсонов М.Ю., Шибалева И.В.* Центры обслуживания вызовов (Call centre). — М.: Эко-Трендз, 2002. — 272 с.
39. *Росляков А. В., Ваняшин С. В.* Математические модели центров обслуживания вызовов. — М.: Ириас, 2006. — 336 с.
40. *Росляков А.В., Ваняшин С.В.* Call-центры. Теоретические основы. — Berlin: LAP LAMPERT Academic Publishing, 2011. — 382 с.
41. *Росляков А. В., Ваняшин С. В., Решодько А. А.* Сравнительный анализ математических моделей центров обслуживания вызовов // Электросвязь. — 2004. — № 9. — С. 32-34.
42. *Самолюбова А.Б.* CALL-CENTER на 100%. Практическое руководство по организации обслуживания вызовов. М.: Альпина Бизнес Букс, 2004. — 309 с.
43. *Степанов М.С.* Использование асимптотических свойств модели с повторениями заявок для оценки показателей качества их обслуживания // Труды конференции “Телекоммуникационные и вычислительные системы”. — 2009. — С.46–47.

44. *Степанов М.С.* Обобщенная модель call-центра с учетом IVR // Труды конференции “Телекоммуникационные и вычислительные системы”. — 2010. — С.46–47.
45. *Степанов М.С.* Особенности построения перспективных контакт-центров // Труды конференции “Телекоммуникационные и вычислительные системы”. — 2011. — С.27.
46. *Степанов М.С.* Оценка влияния IVR на численность операторов в современных контакт-центрах // Труды конференции “Телекоммуникационные и вычислительные системы”. — 2012. — С.50–51.
47. *Степанов М.С.* Оценка характеристик работы контакт-центра с использованием итерационных методов // Т-Сomm: Телекоммуникации и транспорт. — 2012. — Том 6. — №7. — С.188–192.
48. *Степанов М.С.* Использование модели контакт-центра для анализа процесса поступления и обслуживания заявок // Тезисы докладов Всероссийской конференции с международным участием «Информационно-телекоммуникационные технологии и математическое моделирование высокотехнологичных систем». — 2012. — С.55–57.
49. *Степанов М.С.* Планирование числа операторов и линий доступа в современных контакт-центрах // Т-Сomm: Телекоммуникации и транспорт. — 2014. — Том 8. — №7. — С.89–91.
50. *Степанов М.С.* Принципы организации работы контакт-центре нового поколения // Труды конференции “Телекоммуникационные и вычислительные системы”. — 2014. — С.52.
51. *Степанов М.С.* Использование законов сохранения для оценки интенсивности поступления первичных заявок в современных контакт-центрах // Труды международной конференции, посвященной дню Радио (выпуск LXVIII). РНТОРЭС и. А.С. Попова. — 2013. — С.92–94.
52. *Степанов М.С.* Обобщенная модель контакт-центра и частные случаи ее использования // Т-Сomm: Телекоммуникации и транспорт. — 2013. — Том 7. — №7. — С.126–130.
53. *Степанов М.С.* Оценка погрешности планирования числа операторов контакт-центров, внесенная наличием повторных заявок // Труды конференции “Телекоммуникационные и вычислительные системы”. — 2013. — С.42.
54. *Степанов М.С.* Определение и свойства входных параметров обобщенной модели контакт-центра // Т-Сomm: Телекоммуникации и транспорт. — 2015. — Том 9. — №7. — С.25–30.
55. *Степанов М.С.* Косвенная оценка параметров и характеристик обобщенной модели контакт-центра // Тезисы докладов Всероссийской конференции с международным участием «Информационно-телекоммуникационные технологии и математическое моделирование высокотехнологичных систем». — 2015. — С.56–58.

56. *Степанов С. Н.* Интегральные соотношения равновесия для недоступной системы с повторными попытками и их применение // Проблемы передачи информации. — 1980. — Т. 16. — Вып. 4. — С. 88-93.
57. *Степанов С. Н.* Расчёт пучка линий с повторными вызовами и ожиданием // Электросвязь. — 1983. — № 6. — С. 9-13.
58. *Степанов С. Н.* Численные методы расчета систем с повторными вызовами. — М.: Наука, 1983. — 230 с.
59. *Степанов С. Н.* Основы телетрафика мультисервисных сетей. — М.: Эко-Трендз, 2010. — 392 с.
60. *Степанов С.Н., Степанов М.С.* Построение и анализ обобщенной модели контакт-центра // Автоматика и телемеханика. — 2014. — №11. — С.55-69.
61. *Степанов С.Н., Ухловская Л.Г.* Моделирование потоков вызовов, поступающих и обслуженных справочно-информационными службами ГТС // Электросвязь. — 1995. — № 4. — С. 26-29.
62. *Степанова И. В.* Развитие концепции построения центров обслуживания вызовов // Электросвязь. — 2006. — № 2. — С. 16-17.
63. *Ухловская Л.Г.* Исследование принципов построения и разработка методов повышения пропускной способности справочно-информационных служб ГТС большой емкости: Дис. ... канд. техн. наук: 05.13.01 — Москва, 1993. — 150 с.
64. *Холли А.* Контакт-центры как инструмент бизнеса // Connect! Мир связи. — 2007. — № 4. — С.50-52.
65. *Центры обслуживания вызовов. Общие технические требования:* РД.45.191-2001/ Министерство связи РФ. — 2001. — 40 с.
66. *Чан Туан Минь.* Исследование и разработка метода обслуживания вызовов в контакт-центрах: Дис. ... канд. техн. наук: 05.12.13. — Москва, 2012. — 166 с.
67. *Шварц М.* Сети связи: протоколы, моделирование и анализ / Пер. с англ. В. И. Неймана Ч.1: — М.: Наука, 1992. — 336 с.; Ч. 2: — М.: Наука, 1992. — 272 с.
68. *Шерстнева О., Шерстнева А.* Call-центр: алгоритм функционирования с учетом повторных вызовов // Мобильные телекоммуникации. — 2010. — № 2-3. — С. 15-19.
69. *Шнепс-Шнеппе М. А.* Системы распределения информации. Методы расчёта. Справочное пособие. — М.: Связь, 1979. — 344 с.
70. *Шнепс-Шнеппе М. А.* Численные методы теории телетрафика. — М.: Связь, 1974. — 232 с.
71. *Akimaru H., Kawashima K.* Teletraffic Theory and Applications. — London, Berlin, New-York: Springer-Verlag, 1993. — 225 p.

72. *Aksin O.Z., Armony M., Mehrotra V.* The modern call-center: A multi-disciplinary perspective on operations management research // *Production and Operations Management*. — 2007. — V. 16. — N 6. — P. 665-688.
73. *Artalejo J. R., Gyzmez-Corral A.* Retrial Queueing Systems: A Computational Approach. — Springer, 2008. — 318 p.
74. *Aspect Communications Corporation.* Why Your Customers Hate Your IVR Systems. White paper, 2003. [Электронный ресурс]. Режим доступа: <http://www.aspect.com/mm/pdf/products/interactive> (дата обращения 01.03.2016).
75. *Borst S., Mandelbaum A., Reiman M.* Dimensioning large call centers // *Operation Research*. — 2004. — V. 52. — N 1. — P. 17-34.
76. *Boxma O.J., de Waal P.R.* Multiserver queues with impatient customers // *ITC*. — 1994. — N 14. — P. 743-756.
77. *Brandt A., Brandt M.* On the M(n)/M(m)/s Queue with impatient calls // *Performance Evaluation*. — 1999. — V. 35. — N. 1-2. — P. 1-18.
78. *COPC-2000® CSP Standard Release 4.3.* January 2009 [Electronic resource] // Customer Operations Performance Center Inc.[Web portal]. URL : <http://www.proseed.co.jp/download/upload,mages/COPC%20CSP%20Standard%20Release%204.3%20Version%201.1.pdf> (дата обращения 01.03.2016).
79. *Call Center Calculator (Ansapoint).* [Электронный ресурс]. Режим доступа: <http://erlang.com/calculator/call/> (дата обращения 01.03.2016).
80. *ContactBabel Decision Makers Guide 2013.* [Электронный ресурс]. Режим доступа: <http://blog.enghouseinteractive.com/multimedia-management-and-the-universal-queue/>(дата обращения 01.03.2016).
81. *Dawson K.* The Call Center Handbook. CMP Books — San Francisco, 2004.
82. *Delorey E.* Correlating IVR Performance and Customer Satisfaction. May 2003. [Электронный ресурс]. Режим доступа: <http://www.easyivr.com> (дата обращения 01.03.2016)
83. *Eilon S.* A simpler proof of $L = \lambda W$ // *Operations Research*. — 1969. — V. 17. — P. 915-917.
84. *First Contact Resolution.* [Электронный ресурс]. Режим доступа: <http://www.nice.com/first-contact-resolution> (дата обращения 01.03.2016).
85. *Gans N., Koole G., Mandelbaum A.* A telephone call centers: Tutorial, review, and research prospects // *Manufacturing and Service Operation Management*. — 2003. — N 5. — P. 79-141.
86. *Garnett O., Mandelbaum A., Reiman M.I.* Designing a call center with impatient customers // *Manufacturing and Service Operation Management*. — 2002. — N 4. — P. 208-227.
87. *Gross D., Harris C. M.* Fundamentals of Queueing Theory. — Wiley, 1985. — 588 p.

88. *Horovitz B.* Whatever happened to customer service?, USA TODAY, posted 25.09.2003. [Электронный ресурс]. Режим доступа: http://usatoday30.usatoday.com/money/economy/services/2003-09-25-services-frontcover_x.htm (дата обращения 01.03.2016).
89. *Iversen V. B.* Teletraffic Engineering and Network Planning. — Technical University of Denmark, May 2010. — 370 p.
90. *Klenke M.* Endless Loops: Are Your Customers Getting Stuck?. EasyIVR Tech Library. [Электронный ресурс]. Режим доступа: <http://www.easyivr.com> (дата обращения 01.03.2016).
91. *Koole G.* *Call Center Mathematics — A Scientific Method for Understanding and Improving Contact Centers. Vision of Jan. 2007.* [Электронный ресурс]. Режим доступа: <http://www.freetechbooks.com/call-center-mathematics-a-scientific-method-for-understanding-and-improving-contact-centers-t475.html> (дата обращения 01.03.2016).
92. *Leland W. E., Taqqu M. S., Willinger W., Wilson D. V.* On the Self-Similar Nature of Ethernet Traffic (Extended Version) // IEEE/ACM Transactions of Networking. — 1994. — V. 2. — N. 1. — P. 1–15.
93. *Little J. D. C.* A proof of the queueing formula $L = \lambda W$ // Operations Research. — 1961. — V. 9. — P. 383–387.
94. *Mandelbaum A., Zeltyn S.* The Impact of customers patience on delay and abandonment: some empirically-driven experiments with the M/M/N+G queue // OR Spectrum. Special Issue on Call Center Management. — 2004.— № 26.— P. 377-411.
95. *Measuring Call Center Performance. Global Best Practice.* [Электронный ресурс]. Режим доступа: <http://www.ifc.org/wps/wcm/connect/75ce96004cf85d4f8752c7f81ee631cc/Tool+9.4.+Measuring+Call+Center+Performance.pdf?MOD=AJPERES>. (дата обращения 01.03.2016).
96. *Medhi J.* Stochastic Models in Queueing Theory. — Academic Press, 2002. — 482 p.
97. *Prosci.* Call Center Planning and Design Toolkit — A Blueprint for Building a Successful Contact Center. — Prosci Business Performance Series. Learning Center Publications. 2004.
98. *Puent J.* Top Ten Trends For Contact Centers in 2013. [Электронный ресурс]. Режим доступа: <https://www.google.ru/webhp?sourceid=chrome-instant&ion=1&espv=2&ie=UTF-8#q=Top+Ten+Trends+For+Contact+Centers+in+2013> (дата обращения 01.03.2016).
99. *Sakov A.* *Empirical Analysis of a Telephone Call Center.* [Электронный ресурс]. Режим доступа: https://www.google.ru/url?sa=t&rct=j&q=&esrc=s&source=web&cd=2&ved=0ahUKewj-svEp97LAhXh_nIKHQBXAYMQFggmMAE&url=

- [http%3A%2F%2Fie.technion.ac.il%2Fserveng%2FReferences%2FSakovTelAviv_talk.ppt&usg=AFQjCNEgHC_6t5sHOPVh3_sJ1sO-GETrzQ&sig2=ioTRBPupYyxB6Rp6qRLt8A&bvum=bv.117868183,d.bGQ](http://3A%2F%2Fie.technion.ac.il%2Fserveng%2FReferences%2FSakovTelAviv_talk.ppt&usg=AFQjCNEgHC_6t5sHOPVh3_sJ1sO-GETrzQ&sig2=ioTRBPupYyxB6Rp6qRLt8A&bvum=bv.117868183,d.bGQ) (дата обращения 01.03.2016).
100. *Shen H., Huang J.Z.* Analysis of call center arrival data using singular value decomposition // *Applied Stochastic Models in Business an Industry*. — 2005. — V. 21. — N 3. — P. 251-263.
 101. *Shrinivasan R., Talim J., Wang J.* Performance analysis of a call center with interactive voice response units // *Sociedad de Estadistica e Investigacion Operativa Top*. — 2004. — V. 12. — N. 1. P. 91-110.
 102. *Stepanov S. N.* Generalized model with retrials in case of extreme load // *Queueing Systems*. — 1998. — V. 27. — N. 1-2. — P. 131-151.
 103. *Stepanov S. N.* Markov Models with Retrials: The Calculation of Stationary Performance Measures Based on the Concept of Truncation // *Mathematical and Computer Modelling*. — 1999. — V. 30. — N. 3-4. — P. 207-228.
 104. *Stolletz R.* Performance Analysis and Optimization of Inbound Call Centers. — Berlin, Heidelberg: Springer-Verlag, 2003.
 105. *Surenda M. Gupta* Queueing Model with State Dependent Balking and Reneging: Its Complementary and Equivalence // *ACM SIGMETRICS Performance Evaluation Review*. — 1995. — V. 22. — N. 2-4. — P. 63-72.
 106. *The Future of Contact Centers. Cognizant 20-20 insights, may 2014.* [Электронный ресурс]. Режим доступа: [http : //www.cognizant.com/InsightsWhitepapers/The - Future - of - Contact - Centers.pdf](http://www.cognizant.com/InsightsWhitepapers/The-Future-of-Contact-Centers.pdf) (дата обращения 01.03.2016).

Приложение 1.

Численный анализ характеристик контакт-центра

В данном разделе приведены результаты точного и приближенного вычисления характеристик обобщённой модели контакт-центра асимптотическим методом, построенным в разделе 4.2. Фиксированные параметры модели принимают значения: $\alpha_1 = 1$; $\alpha_2 = 0,5$; $\beta_1 = 0,2$; $q = 0,8$; $p = 0,15$; $m = 1$; $p_1 = 1,0$; $\sigma = 1$; $\sigma_1 = 1$; $H = 0,8$; $n = 50$; $v = 15$; $v_1 = 7$; $w = 7$; $w_1 = 4$. Рассмотрены случаи больших потерь ($\lambda = 15$, $j_m = 30$) и малых потерь ($\lambda = 8$, $j_m = 15$). Величина ν меняется от 1 до 10. Выводы по результатам использования приближенного метода сформулированы в разделе 4.2.4.

Таблица П.1.1.

Доля потерянных заявок в зависимости от этапа обслуживания и величины интенсивности повторения заявки ν (случай больших потерь)

Значение ν	Показатели доступности ресурса			
	π_a	π_o	$\pi_{c,1}$	π_g
10	0,270251	0,554331	0,120507	0,598575
9	0,268631	0,554932	0,120520	0,598030
8	0,266791	0,555611	0,120535	0,597410
7	0,264671	0,556389	0,120552	0,596696
6	0,262187	0,557293	0,120573	0,595859
5	0,259216	0,558365	0,120599	0,594856
4	0,255557	0,559673	0,120631	0,593621
3	0,250869	0,561327	0,120673	0,592037
2	0,244483	0,563541	0,120731	0,589876
1	0,231356	0,566492	0,120811	0,584535
0	0,220335	0,571674	0,120905	0,581755

Таблица П.1.2.

Характеристики ожидания начала обслуживания в зависимости от его этапа и величины интенсивности повторения заявки ν (случай больших потерь)

Значение ν	Показатели ожидания обслуживания			
	$p\{\mathcal{W}_o > 0\}$	W_o	$p\{\mathcal{W}_1 > 0\}$	W_1
10	0,731556	0,400819	0,117510	0,119478
9	0,731173	0,401157	0,117443	0,119490
8	0,730750	0,401545	0,117366	0,119505
7	0,730276	0,401993	0,117276	0,119521
6	0,729741	0,402522	0,117169	0,119542
5	0,729124	0,403159	0,117039	0,119567
4	0,728398	0,403948	0,116877	0,119598
3	0,727510	0,404962	0,116668	0,119639
2	0,726354	0,406341	0,116384	0,119695
1	0,724972	0,408270	0,115989	0,119772
0	0,721510	0,411262	0,115478	0,119863

Таблица П.1.3.

Интегральные характеристики обслуживания заявок в зависимости от величины интенсивности повторения заявки ν (случай больших потерь)

Значение ν	Интегральные показатели			
	π_r	δ_r	M	Q
10	0,2297	0,478860	0,9189	1,491125
9	0,2293	0,478424	0,9173	1,487748
8	0,2289	0,477928	0,9154	1,483919
7	0,2284	0,477357	0,9134	1,479520
6	0,2277	0,476687	0,9109	1,474383
5	0,2270	0,475885	0,9080	1,468260
4	0,2261	0,474897	0,9044	1,460757
3	0,2250	0,473630	0,8998	1,451204
2	0,2234	0,471897	0,8936	1,438286
1	0,2186	0,466460	0,8743	1,406942
0	0,2177	0,465404	0,8706	1,390942

Таблица П.1.4.

Доля потерянных заявок в зависимости от этапа обслуживания и величины интенсивности повторения заявки ν (случай малых потерь)

Значение ν	Показатели доступности ресурса			
	π_a	π_o	$\pi_{c,1}$	π_g
10	0,000045	0,098217	0,068671	0,086056
9	0,000043	0,098096	0,068680	0,085959
8	0,000040	0,097946	0,068692	0,085840
7	0,000038	0,097759	0,068706	0,085690
6	0,000035	0,097516	0,068724	0,085497
5	0,000032	0,097189	0,068748	0,085238
4	0,000028	0,096726	0,068781	0,084871
3	0,000023	0,096016	0,068831	0,084310
2	0,000017	0,094784	0,068911	0,083338
1	0,000008	0,092087	0,069060	0,081211
0	0,000001	0,083247	0,069103	0,074210

Таблица П.1.5.

Характеристики ожидания начала обслуживания в зависимости от его этапа и величины интенсивности повторения заявки ν (случай малых потерь)

Значение ν	Показатели ожидания обслуживания			
	$p\{\mathcal{W}_o > 0\}$	W_o	$p\{\mathcal{W}_1 > 0\}$	W_1
10	0,353509	0,091397	0,121075	0,068358
9	0,353418	0,091320	0,121101	0,068368
8	0,353304	0,091225	0,121132	0,068379
7	0,353155	0,091105	0,121172	0,068393
6	0,352955	0,090948	0,121223	0,068412
5	0,352675	0,090735	0,121293	0,068436
4	0,352254	0,090428	0,121393	0,068469
3	0,351563	0,089949	0,121548	0,068519
2	0,350241	0,089096	0,121822	0,068599
1	0,346876	0,087149	0,122434	0,068749
0	0,332040	0,080168	0,124345	0,068803

Таблица П.1.6.

Интегральные характеристики обслуживания заявок в зависимости от величины интенсивности повторения заявки ν (случай малых потерь)

Значение ν	Интегральные показатели			
	π_r	δ_r	M	Q
10	0,01848	0,068845	0,0739	0,094159
9	0,01845	0,068767	0,0738	0,094042
8	0,01842	0,068672	0,0737	0,093900
7	0,01840	0,068552	0,0736	0,093721
6	0,01835	0,068398	0,0734	0,093491
5	0,01830	0,068191	0,0732	0,093181
4	0,01820	0,067897	0,0728	0,092743
3	0,01808	0,067448	0,0723	0,092073
2	0,01785	0,066670	0,0714	0,090915
1	0,01738	0,064969	0,0695	0,088389
0	0,01578	0,059368	0,0631	0,080159

Приложение 2.

Акты об использовании результатов диссертации

1. Акт об использовании результатов диссертационной работы Степанова М.С. «Разработка и анализ обобщенной модели обслуживания вызовов в перспективных контакт-центрах» в ОАО МГТС.
2. Акт об использовании результатов диссертационной работы М.С. Степанова на тему: «Разработка и анализ обобщенной модели обслуживания вызовов в перспективных контакт-центрах» в учебном процессе кафедры Сети Связи и Системы Коммутации МТУСИ.



Открытое акционерное общество
«Московская городская телефонная сеть»
ул. Большая Ордынка, дом 25, стр. 1
Москва, Россия, 119991

Т +7 (495) 950 00 00
Ф +7 (495) 950 06 18
www.mgts.ru

№ _____
от _____
на № _____
от _____

ОТВЕРЖДАЮ
начальник Центра обучения и развития
персонала ОАО МГТС
Вяткин А.Б.
_____ 2015 г.



АКТ

об использовании результатов диссертационной работы Степанова М.С. «Разработка и анализ обобщенной модели обслуживания вызовов в перспективных контакт-центрах»
в ОАО МГТС.

Комиссия в составе: председателя Начальника Контактного Центра ДОК Михайловой Л.М. и членов комиссии Начальника Отдела справочных услуг Кравцова Е.Н. и Главного специалиста ГОК ОСУ КЦ ДОК Закутаевой Л.Р. составила настоящий акт в том, что:

- При анализе влияния повторных вызовов на процесс формирования входного потока заявок в справочно-информационную службу были использованы следующие результаты и методики, полученные в диссертационной работе Степанова М.С. «Разработка и анализ обобщенной модели обслуживания вызовов в перспективных контакт-центрах», представленной на соискание ученой степени кандидата технических наук:
 - математическая модель поступления и обслуживания заявок в контакт-центре, учитывающая возможность повторения заявки из-за занятости операторов и линий доступа, а также из-за неудачного завершения времени ожидания.
 - методика оценки доли потерянных заявок, а также других показателей качества обслуживания поступающих заявок, основанная на составлении и последующем решении системы уравнений статистического равновесия при заданных значениях входных параметров.
- Внедрение разработанных Степановым М.С. методов и алгоритмов позволило:
 - оценить интенсивность поступления первичных вызовов, а также ряд других характеристик, прямое измерение которых затруднено из-за невозможности разделения первичных и повторных вызовов, и тем самым уточнить процедуры планирования необходимого по нагрузке числа операторов.

Председатель комиссии:

Начальник Контактного Центра ДОК _____ Михайлова Л.М.

Члены:

Начальник Отдела справочных услуг КЦ ДОК _____ Кравцов Е.Н.

Главный специалист ГОК ОСУ КЦ ДОК _____ Закутаева Л.Р.

УТВЕРЖДАЮ

Проректор по учебной работе ФГБОУ ВО
«Московский технический университет
связи и информатики»,



АКТ

**об использовании результатов диссертационной работы М.С. Степанова на
тему: «Разработка и анализ обобщенной модели обслуживания вызовов в
перспективных контакт-центрах» в учебном процессе кафедры Сетей
связи и систем коммутации ФГБОУ ВО МТУСИ**

Комиссия в составе:

- директора Департамента организации и управления учебным процессом МТУСИ Карпушиной Н.Д.;
- заведующей Центром планирования и сопровождения учебного процесса МТУСИ Патенченковой Е.К.;

- заведующего кафедрой СС и СК Пшеничникова А.П. удостоверяет, что в учебном процессе кафедры СС и СК при чтении курса лекций по дисциплине «Сети связи» используются результаты диссертационного исследования Степанова М.С., а именно: проведенный диссертантом анализ используемых математических моделей контакт-центров, а также алгоритмы оценки характеристик современных и перспективных справочно-информационных служб, полученные в диссертационном исследовании.

Директор департамента организации и
управления учебным процессом МТУСИ  Н.Д. Карпушина

Зав. центром планирования и сопровождения
учебного процесса МТУСИ  Е.К. Патенченкова

Заведующий кафедрой СС и СК  А.П. Пшеничников