

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
федеральное государственное бюджетное образовательное учреждение высшего
образования
«Тольяттинский государственный университет»

Кафедра _____ «Прикладная математика и информатика»
(наименование)

01.03.02 Прикладная математика и информатика
(код и наименование направления подготовки / специальности)

Компьютерные технологии и математическое моделирование
(направленность (профиль) / специализация)

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА
(БАКАЛАВРСКАЯ РАБОТА)

на тему «Оптимизация обслуживания клиентов банка с помощью математического моделирования очередей»

Обучающийся	Н.А. Шурыгин (Инициалы Фамилия)	_____ (личная подпись)
Руководитель	М.А. Тренина (ученая степень (при наличии), ученое звание (при наличии), Инициалы Фамилия)	_____
Консультант	к.ф.н., М.В. Дайнеко (ученая степень (при наличии), ученое звание (при наличии), Инициалы Фамилия)	_____

Аннотация

Тема выпускной квалификационной работы: «Оптимизация обслуживания клиентов банка с помощью математического моделирования очередей».

Выпускная квалификационная работа посвящена оптимизации обслуживания клиентов банка, моделированию очередей и выявлению оптимального решения.

Структура работы включает в себя введение, 3 главы, заключение и список литературы.

В первой главе рассматриваются теоретические основы систем массового обслуживания, анализируются современные подходы к оптимизации обслуживания клиентов и рассматриваются соответствующие математические модели в контексте банковских операций.

Во второй главе представлен подробный анализ системы обслуживания клиентов в реальном банке, включая сбор данных, оценку эффективности и выявление узких мест в обслуживании.

В третьей главе описывается разработка программного средства для моделирования и оптимизации обслуживания клиентов банка с использованием теории массового обслуживания, а также тестирование и экспериментальный анализ.

В заключении представлены выводы по проделанной работе.

Выпускная квалификационная работа содержит пояснительную записку объемом 45 страниц, включая 17 иллюстраций, 3 таблицы и списка используемой литературы из 25 источников, включая иностранные.

Abstract

The title of the graduation work is Optimization of bank customer service using mathematical queue modeling.

The graduation work is devoted to optimizing the bank's customer service, queue modeling and identifying the optimal solution.

The structure of the graduation work includes an introduction, three parts, a conclusion and a list of references.

The first part examines the theoretical foundations of queuing systems, analyzes modern approaches to optimizing customer service, and examines relevant mathematical models in the context of banking operations.

The second part provides a detailed analysis of the customer service system in a real bank, including data collection, efficiency assessment, and identification of service bottlenecks.

The third part describes the development of a software tool for modeling and optimizing bank customer service using queuing theory, as well as testing and experimental analysis.

In conclusion, the conclusions of the work performed are presented.

The graduation work contains an explanatory note of 45 pages, including 17 figures, 3 tables, and a list of 25 references including foreign sources.

Оглавление

Введение.....	5
Глава 1 Теоретические основы математического моделирования очередей....	7
1.1 Обзор теории массового обслуживания	7
1.2 Текущее состояние проблемы	9
1.3 Предметная область и модель предметной области.....	10
1.4 Существующие методы и сравнительный анализ	15
Глава 2 Анализ текущего состояния обслуживания клиентов в банке	22
2.1 Описание исследуемого банка.....	22
2.2 Сбор и анализ данных.....	26
2.3 Выявление проблемных зон.....	31
Глава 3 Разработка и тестирование программного продукта.....	35
3.1 Программная реализация алгоритма.....	35
3.2 Тестирование программы и проведение экспериментов	39
Заключение	43
Список используемой литературы	44

Введение

В современной банковской среде обеспечение высокого качества обслуживания клиентов является важнейшим конкурентным преимуществом. В связи с увеличением клиентского потока и растущими ожиданиями в отношении скорости и удобства, банки вынуждены искать эффективные методы улучшения процессов обслуживания. Одним из наиболее эффективных подходов к анализу и оптимизации систем обслуживания является использование математического моделирования очередей.

Теория массового обслуживания позволяет проводить количественную оценку производительности системы, выявлять узкие места и моделировать различные варианты оптимизации. Эта работа направлена на применение моделей организации очередей в банковской сфере с целью сокращения времени ожидания клиентов, сбалансирования рабочей нагрузки сотрудников и повышения общей эффективности обслуживания.

Выпускная квалификационная работа посвящена оптимизации обслуживания клиентов банка, моделированию очередей и выявлению оптимального решения.

Актуальность данного исследования определяется растущей потребностью финансовых учреждений в повышении качества обслуживания клиентов и операционной эффективности. В современных банковских системах длинные очереди и чрезмерное время ожидания приводят к неудовлетворенности клиентов и операционным потерям. Поэтому применение математической теории массового обслуживания становится жизненно важным инструментом для оптимизации обслуживания.

Целью данной выпускной квалификационной работы является разработка и анализ математической модели системы обслуживания

клиентов банка, основанной на теории массового обслуживания, с целью определения оптимальных конфигураций обслуживания и сокращения времени ожидания.

Объектом исследования является система обслуживания клиентов банка.

Предметом исследования является применение методов математического моделирования для анализа и оптимизации динамики очередей в рамках данной системы.

Для достижения цели были поставлены следующие задачи:

- изучить существующие методы анализа систем массового обслуживания и их применение в банковской сфере;
- собрать и проанализировать статистические данные о потоках клиентов и параметрах обслуживания;
- построить имитационную модель процесса обслуживания в банке с использованием соответствующих математических инструментов;
- оценить производительность системы при различных сценариях и выявить узкие места;
- предложить стратегии оптимизации, направленные на повышение эффективности обслуживания и сокращение времени ожидания.

Структура работы включает в себя введение, 3 главы, заключение и список литературы.

Первая глава посвящена теории систем массового обслуживания. Во второй главе проводится анализ клиентского обслуживания в конкретном банке. Третья глава описывает разработку и тестирование программы для моделирования обслуживания клиентов.

Глава 1 Теоретические основы математического моделирования очередей

1.1 Обзор теории массового обслуживания

Эффективное управление процессами обслуживания клиентов является важнейшей задачей для современных банков, где задержки могут привести к неудовлетворенности клиентов и операционной неэффективности. Математическое моделирование очередей обеспечивает надежную основу для анализа и оптимизации систем обслуживания, учитывая динамику прихода клиентов, скорость обслуживания и пропускную способность системы.

Теория массового обслуживания – это раздел прикладной математики и исследования операций, который изучает поведение очередей, или очередей ожидания [4], [8]. Зародившись в работах А.К. Эрланга в начале 20-го века, эта теория развивалась для решения широкого круга проблем в области телекоммуникаций, транспорта, здравоохранения и финансовых услуг.

На протяжении многих лет такие исследователи, как Д.Г. Кендалл, Джон Литтл и Леонард Клейнрок, вносили свой вклад в развитие теории массового обслуживания, вводя такие ключевые понятия, как:

- система счисления Кендалла: система классификации моделей массового обслуживания (например, $M/M/1$, $M/M/c$);
- закон Литтла: фундаментальный принцип, который связывает среднее количество клиентов в системе со средним уровнем посещаемости и средним временем, проведенным в системе.

В банковской области теория массового обслуживания применяется для моделирования схем прибытия клиентов, времени обслуживания и численности персонала для повышения эффективности обслуживания [11], [19].

Системы массового обслуживания обычно характеризуются тремя параметрами [22].

Первый параметр характеризует процесс поступления клиентов, при этом моделируется пуассоновским процессом, в котором число поступающих клиентов за единицу времени распределено по закону Пуассона:

$$P(N=k)=\frac{(\lambda t)^k e^{-\lambda t}}{k!}, \quad (1)$$

где λ – средняя интенсивность потока;

t – время наблюдения.

Второй параметр – это характеристика очередности: определяет порядок обслуживания клиентов, например, по принципу «Первый пришел – первый вышел» (FIFO), «Последний пришел – первый обслужен» (LIFO) или по системе приоритетов.

Третий параметр – это механизм обслуживания: определяется числом серверов (кассиров, менеджеров) и законом распределения времени обслуживания, часто принимаемым экспоненциальным:

$$f(t)=\mu e^{-\mu t}, \quad (2)$$

где μ – интенсивность обслуживания.

На рисунке 1 представлен график, который показывает, что вероятность короткого времени обслуживания выше, чем длинного, что соответствует реальному процессу обслуживания клиентов в банке.

Экспоненциальное распределение времени обслуживания

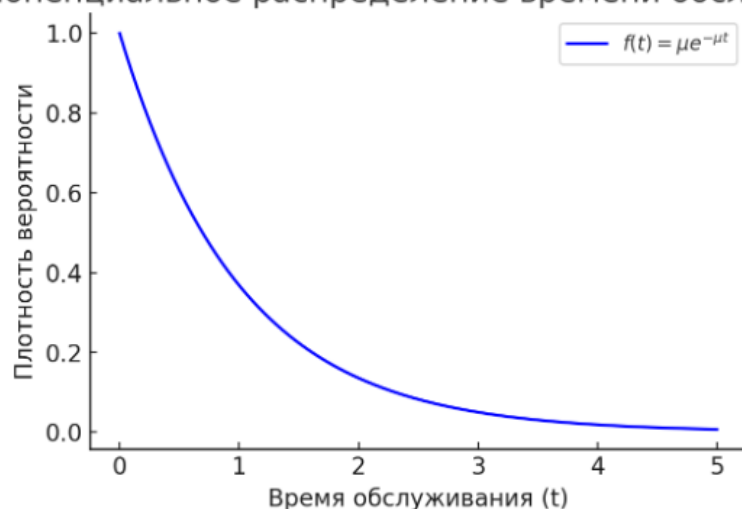


Рисунок 1 – График экспоненциального распределения времени обслуживания

1.2 Текущее состояние проблемы

Несмотря на достижения в области технологий и методов управления, многие банки по-прежнему сталкиваются с трудностями в обеспечении эффективного обслуживания клиентов [1]. Некоторые из ключевых проблем включают в себя: длительное ожидание – клиенты часто сталкиваются с долгим ожиданием, особенно в периоды пиковой нагрузки; неэффективное распределение ресурсов – персонал и другие ресурсы не всегда распределяются оптимально, что приводит к неэффективности; изменчивость потребительского спроса – приход клиентов сильно варьируется, что затрудняет прогнозирование и управление спросом. Цель состоит в том, чтобы сократить время ожидания, повысить эффективность обслуживания и повысить общую удовлетворенность клиентов [15].

Исследования в области систем массового обслуживания, применяемых в банковской сфере, были сосредоточены на оптимизации операций кассиров, услуг банкоматов и центров обработки вызовов [21]. Ключевые подходы включают: аналитические методы – использование математических моделей, таких как М/М/1 или М/М/с, для расчета показателей производительности,

включая среднее время ожидания, длину очереди и загрузку сервера. Методы моделирования – использование программных средств для моделирования сложных систем, которые дают возможность учитывать сложные факторы. Алгоритмы машинного обучения – прогнозирование поведения клиентов и динамическая настройка ресурсов для удовлетворения спроса [9].

Критический анализ литературы показывает, что недостаточно внимания уделяется использованию данных в режиме реального времени и адаптивному управлению очередями для персонализированного обслуживания клиентов. В большинстве исследований рассматривается общее поведение очередей и при этом мало внимания уделяется решениям, адаптированным для конкретных банковских подразделений.

1.3 Предметная область и модель предметной области

Теория массового обслуживания занимается созданием математических моделей, которые связывают параметры работы системы массового обслуживания (СМО), такие как количество каналов, их производительность, характер потока заявок и организация работы, с показателями её эффективности [2], [17].

Эти показатели эффективности можно разделить на три основные группы:

- 1) показатели использования СМО:
 - среднее число обслуживаемых заявок в единицу времени;
 - доля обслуженных заявок от поступивших;
 - средняя продолжительность периода занятости СМО;
 - время, в течение которого СМО занята обслуживанием;
 - время простоя СМО;
- 2) показатели качества обслуживания:
 - среднее время ожидания обслуживания;
 - среднее время пребывания заявки в СМО;

- вероятность отказа в обслуживании без ожидания;
- вероятность немедленного обслуживания поступившей заявки;
- среднее число заявок в СМО;
- среднее число заявок в очереди;
- вероятность превышения определённого числа заявок в очереди;

3) показатели экономической эффективности:

- средний доход, приносимый СМО в единицу времени;
- соотношение среднего дохода и затрат на содержание СМО;
- потери дохода из-за отказа в обслуживании части заявок.

Проблема оптимизации обслуживания клиентов в банках относится к области оптимизации системы обслуживания и стохастических процессов [10]. Моделирование таких систем позволяет анализировать и улучшать их работу, тем самым минимизируя время ожидания и повышая эффективность. Модель предметной области включает: модели прибытия клиентов – моделируются пуассоновские процессы с зависящими от времени темпами прибытия. Пуассоновский процесс предполагает, что события (прибытия клиентов) происходят независимо друг от друга с постоянной средней скоростью, что делает его подходящим для моделирования случайного времени прибытия. Скорость обслуживания – характеризуется экспоненциальным или не экспоненциальным распределением в зависимости от сложности транзакций. Например, экспоненциальное распределение используется для простых процессов, в то время как для более сложных транзакций может использоваться логарифмически нормальное распределение или распределение Вейбулла. Системные ограничения – ограниченная доступность сервера, физическое пространство для зон ожидания и часы работы. Эти ограничения являются неотъемлемой частью обеспечения соответствия модели реальным ограничениям [14].

Ключевыми метриками эффективности являются:

- среднее время ожидания W_q :

$$W_q = \frac{\rho}{\mu(1-\rho)}, \quad (3)$$

где $\rho = \lambda / (c\mu)$ – коэффициент загрузки системы;

– вероятность отказа в обслуживании P_b :

$$P_b = \frac{\left(\frac{\lambda}{\mu}\right)^c / c!}{\sum_{n=0}^c \left(\frac{\lambda}{\mu}\right)^n / n!}, \quad (4)$$

где c – число серверов (кассиров).

Требования к решению включают в себя: минимизация времени ожидания клиентов, максимальное использование сервера, обеспечение адаптивности к меняющемуся спросу [3].

На рисунке 2 показано, как средняя длина очереди (L_q) растет при увеличении коэффициента загрузки системы (ρ). Чем ближе ρ к 1, тем быстрее очередь увеличивается, что показывает необходимость эффективного управления очередями.

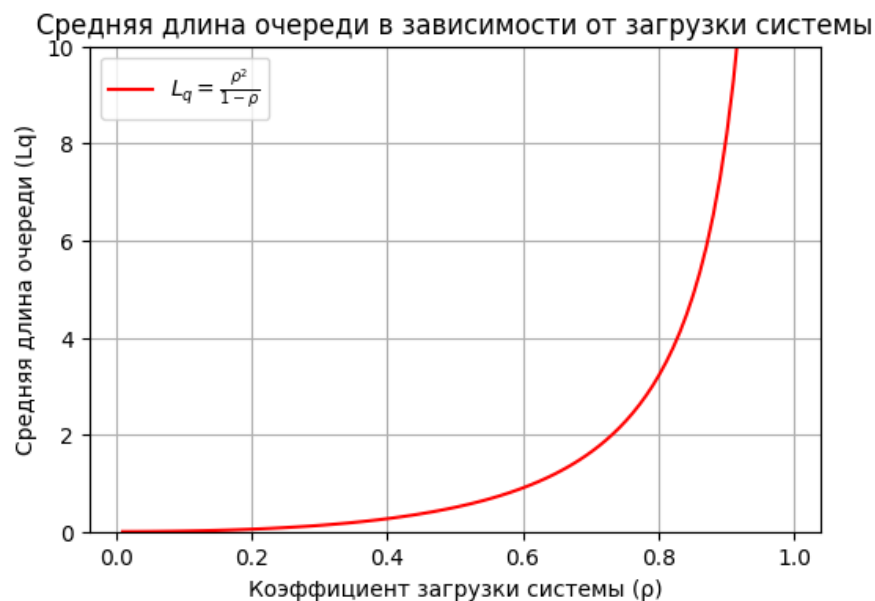


Рисунок 2 – График зависимости средней длины очереди от нагрузки системы

Основными характеристиками процесса обслуживания являются: конфигурация системы обслуживания, режим обслуживания.

Системы обслуживания различаются по числу каналов обслуживания. Обычно количество каналов можно определить, как число клиентов, обслуживание которых может быть начато одновременно. Другая характеристика – число фаз обслуживания или последовательных этапов обслуживания одного клиента [13]. Примером одноканальной системы обслуживания может служить банк, в котором открыто единственное окно для обслуживания клиентов. Если же в банке открыто несколько окон, клиент ожидает в общей очереди и подходит к первому освободившемуся окну, то мы имеем дело с многоканальной однофазовой системой обслуживания. Большинство банков, также как почтовые отделения и авиакассы, сейчас являются многоканальными системами обслуживания. В банковской сфере однофазовой системой можно назвать ситуацию, когда клиент обращается к одному специалисту (например, кассиру) для выполнения всех необходимых операций – снятия наличных, внесения платежа или обмена валюты, после чего покидает банк. Если же процесс обслуживания разделён на несколько этапов, например, клиент сначала получает консультацию у менеджера, затем оформляет документы в другом окне и только после этого проводит финансовую операцию у кассира, то такая система является многофазовой (трёхфазной) [23]. Разновидность систем представлены на рисунке 3.

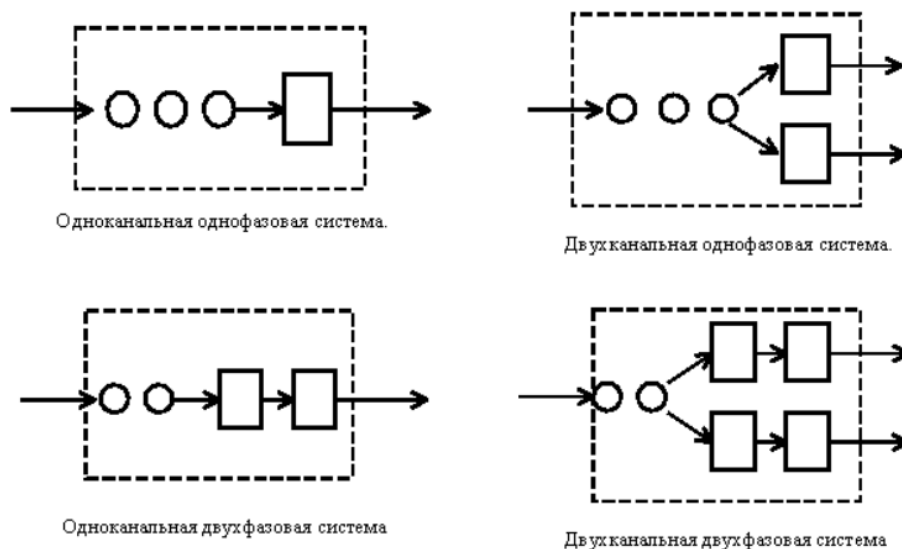


Рисунок 3 – Варианты каналов обслуживания

Режим обслуживания, так же как режим поступления заявок, может характеризоваться либо постоянным, либо случайным временем обслуживания. Во многих случаях можно предположить, что время обслуживания подчиняется экспоненциальному распределению со следующей функцией распределения:

$$F(\tau) = p(t < \tau) = 1 - e^{-\lambda \tau}, \quad (5)$$

где $p(t < \tau)$ – вероятность того, что фактическое время t обслуживания заявки не превысит заданной величины τ ;

λ – среднее число заявок, обслуживаемых в единицу времени.

При постоянном времени на обслуживание любого клиента затрачивается одинаковое время. Однако, часто встречаются ситуации, когда время обслуживания имеет случайное распределение. Например, открытие нового счета занимает больше времени, чем простой запрос баланса. В таких случаях для точного представления случайности времени обслуживания и анализа производительности системы в условиях неопределенности

используются модели, такие как $M/M/1$. Это позволяет проводить более реалистичное моделирование и принимать более эффективные решения по оптимизации ресурсов.

1.4 Существующие методы и сравнительный анализ

Обзор методологий позволяет выявить три основных подхода к решению проблем массового обслуживания [18]. Точные аналитические решения – подходят для простых систем, но ограничены в масштабируемости. Например, модель $M/M/1$: очередь с одним сервером, в которой поступления осуществляются в соответствии с пуассоновским процессом, время обслуживания распределяется экспоненциально, и длина очереди не ограничена.

Наиболее часто встречаются задачи массового обслуживания с единственным каналом. В этом случае клиенты формируют единственную очередь, которая обслуживается одним рабочим местом. Предположим, что для систем этого типа выполняются следующие условия:

- заявки обслуживаются по принципу: первым пришел – первым обслужен (FIFO), причем каждый клиент ожидает своей очереди до конца независимо от длины очереди;
- появления заявок являются независимыми событиями, однако, среднее число заявок, поступающих в единицу времени, неизменно;
- процесс поступления заявок описывается пуассоновским распределением, причем заявки поступают из неограниченного множества;
- время обслуживания описывается экспоненциальным распределением вероятностей;
- темп обслуживания выше темпа поступления заявок.

Если эти условия выполняются, то система массового обслуживания описывается уравнениями, в которых используются следующие параметры: λ

– число заявок в единицу времени; μ – число клиентов, обслуживаемых в единицу времени. Ниже приведены эти уравнения:

среднее число клиентов в системе

$$L_s = \frac{\lambda}{\mu - \lambda}, \quad (6)$$

среднее время обслуживания одного клиента в системе

$$W_s = \frac{1}{\mu - \lambda}, \quad (7)$$

среднее число клиентов в очереди

$$L_q = \frac{\lambda^2}{\mu(\mu - \lambda)}, \quad (8)$$

среднее время ожидания клиента в очереди

$$W_q = \frac{\lambda}{\mu(\mu - \lambda)}, \quad (9)$$

загруженность системы

$$\gamma = \frac{\lambda}{\mu}, \quad (10)$$

вероятность отсутствия заявок в системе

$$P_0 = 1 - \frac{\lambda}{\mu}. \quad (11)$$

Модель М/М/с: расширяет модель М/М/1 на несколько серверов (с-

серверов). Он вычисляет аналогичные показатели, принимая во внимание вероятность ожидания клиентов.

В многоканальной системе для обслуживания открыты два или более каналов. Предполагается, что клиенты ожидают в общей очереди и обращаются в первый освободившийся канал обслуживания [24]. Пример такой многоканальной однофазовой системы можно сейчас увидеть во многих банках. Из общей очереди клиенты обращаются в первое освободившееся окно для обслуживания.

В многоканальной системе поток заявок подчиняется пуассоновскому закону, а время обслуживания – экспоненциальному. Приходящий первым обслуживается первым, и все каналы обслуживания работают в одинаковом темпе [25].

Для многоканальной системы из n каналов должно выполняться условие $\frac{r}{n} < 1$, где r – параметр загрузки системы, n – минимальное количество каналов, при котором очередь не будет расти до бесконечности. В противоположном случае предельные вероятности существовать не могут.

Далее представлены формулы для описания системы:

вероятность того, что система свободна

$$P_0 = \left(\sum_{n=0}^{\infty} \frac{(\lambda/\mu)^n}{n!} \right)^{-1}, \quad (12)$$

вероятность того, что в системе находится n заявок

$$P_n = \frac{\gamma^n}{n!} \cdot P_0, \quad (13)$$

вероятность того, что заявка окажется в очереди

$$P_q = \frac{\gamma^{n+1}}{n!(n-\gamma)} \cdot P_0, \quad (14)$$

среднее число занятых каналов

$$\bar{k} = \frac{\lambda}{\mu} = \gamma, \quad (15)$$

среднее число заявок в очереди

$$L_q = \frac{\gamma^{n+1} \cdot P_0}{n \cdot n! \left(1 - \frac{\gamma}{n}\right)^2}, \quad (16)$$

среднее число заявок в системе

$$L_s = L_q + r, \quad (17)$$

время нахождения заявки в очереди

$$W_q = \frac{1}{\lambda} \cdot L_q, \quad (18)$$

время нахождения заявки в системе

$$W_s = \frac{1}{\lambda} \cdot L_s, \quad (19)$$

коэффициент загрузки персонала

$$U = \frac{\lambda}{c\mu}. \quad (20)$$

Одним из инструментов для анализа систем массового обслуживания являются цепи Маркова. В контексте модели М/М/с система может быть смоделирована как непрерывная цепь Маркова со следующими состояниями:

состояние 0: в системе нет клиентов;

состояние 1: в системе один клиент;

состояние 2: в системе два клиента;

...

состояние c : в системе c клиентов (все серверы заняты);

состояние $c+1$: в системе $c+1$ клиент (все серверы заняты, и один клиент ожидает).

Скорость перехода между состояниями может быть определена на основе скорости прибытия и скорости обслуживания.

Методы аппроксимации – полезны для больших систем, где точные решения требуют больших вычислительных затрат. Диффузионные аппроксимации, например, аппроксимация распределения длин очередей с использованием непрерывных функций плотности вероятности.

Имитационные модели – обеспечивают гибкость при моделировании сложностей реального мира, но требуют значительных вычислительных ресурсов [6]. Такие инструменты, как AnyLogic или Arena, моделируют динамические взаимодействия с клиентами и реакцию системы с течением времени.

Рассмотрим две схемы моделей массового обслуживания: $M/M/1$ и $M/M/c$. На рисунке 4 изображена диаграмма, которая представляет схему модели $M/M/1$, где клиенты поступают в очередь, а затем обслуживаются одной кассой.

Схема М/М/1



Рисунок 4 – Схема модели М/М/1

Рисунок 5 показывает диаграмму, которая иллюстрирует модель М/М/с, где клиенты сначала попадают в очередь, а затем распределяются между несколькими кассами.

Схема М/М/с

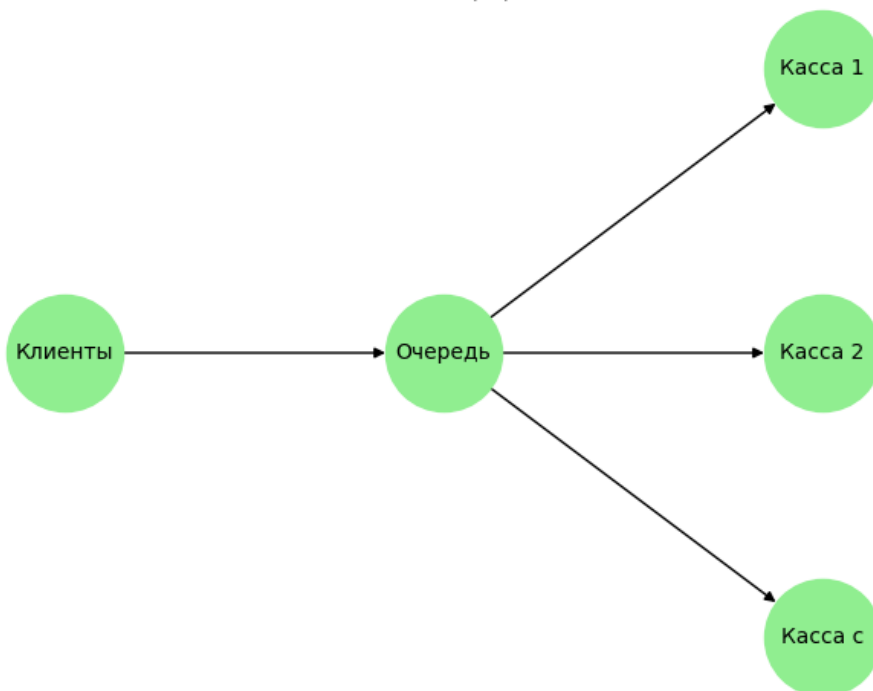


Рисунок 5 – Схема модели М/М/с

Модель $M/M/1$ проста и подходит для систем с одной точкой обслуживания, но требует долгого ожидания при большом потоке клиентов. Модель $M/M/c$ с несколькими серверами значительно сокращает время ожидания и повышает эффективность обслуживания, особенно в условиях высокой нагрузки. Однако сложность расчетов возрастает, что требует более совершенных методов оптимизации. Для банковских приложений, как правило, предпочтительнее использовать $M/M/c$, поскольку банки обрабатывают большие потоки клиентов, которым требуется несколько каналов обслуживания. Это обеспечивает более быстрое обслуживание, снижает загруженность и повышает удовлетворенность клиентов.

Выводы по первой главе

В этой главе изложены основы теории массового обслуживания и ее значение для оптимизации обслуживания клиентов в банке. Критический анализ литературы и существующих методов показывает, что гибридный подход, объединяющий аналитику, моделирование и возможность адаптации в режиме реального времени, является наиболее эффективным решением.

Подробные математические модели, такие как $M/M/1$ и $M/M/c$, дают представление об основной динамике очередей, в то время как моделирование и оптимизация расширяют возможности их применения в сложных системах.

Используя принципы теории массового обслуживания и теории случайных процессов, банки могут разрабатывать сложные модели для оптимизации операций по обслуживанию клиентов. Внедрение таких моделей позволяет прогнозировать график прихода клиентов, эффективно распределять сервисные ресурсы и корректировать механизмы обслуживания в режиме реального времени, тем самым минимизируя время ожидания и повышая общее качество обслуживания.

Глава 2 Анализ текущего состояния обслуживания клиентов в банке

2.1 Описание исследуемого банка

Исследуемый банк является финансовым учреждением с хорошо развитой инфраструктурой и значительным объемом клиентских потоков. Банк предоставляет услуги частным лицам и корпоративным клиентам, включая депозиты, кредиты, денежные переводы и консультационные услуги. Эффективность обслуживания клиентов напрямую влияет на удовлетворенность клиентов и операционные показатели [5].

К основным характеристикам системы обслуживания банка относятся:

- количество отделений: банк имеет в городе несколько отделений с разной интенсивностью работы;
- количество обслуживаемых клиентов в день: в среднем каждое отделение обслуживает N клиентов в день;
- среднее время обслуживания одного клиента: среднее время обслуживания одного клиента составляет t_s минут;
- доступные банковские услуги: услуги включают кассовые операции, снятие наличных в банкоматах, консультации по кредитам и поддержку онлайн-банкинга.

Для примера сделаем график распределения потока клиентов в течение дня (рисунок 6). Это может проиллюстрировать часы пиковой нагрузки и помочь проанализировать изменения нагрузки в разное время.

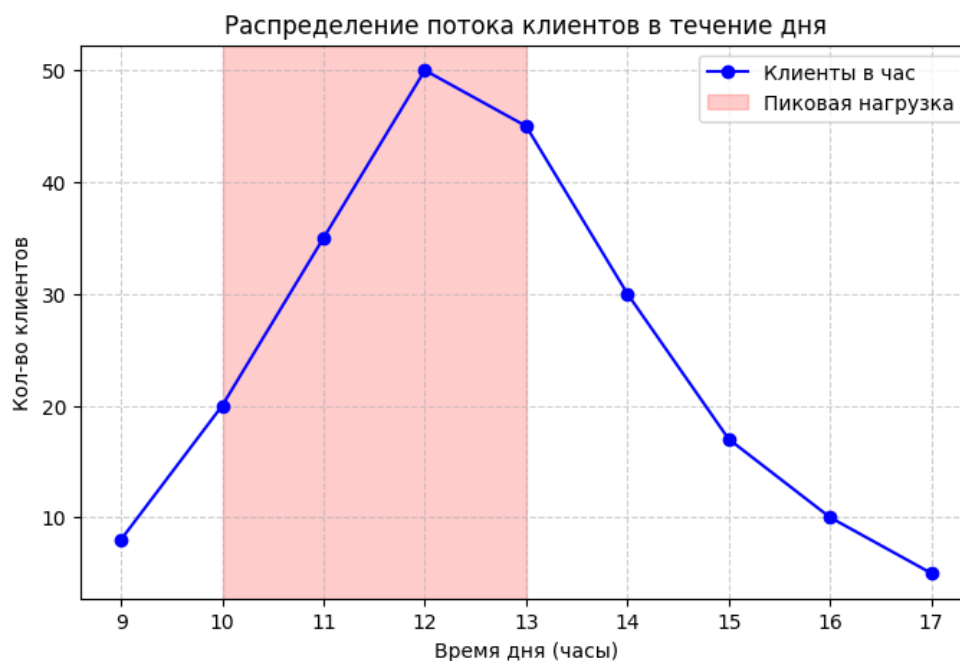


Рисунок 6 – График прихода клиентов за день

Потоки клиентов в банке распределены неравномерно в течение дня. Пиковая нагрузка приходится на утренние и обеденные часы, что приводит к увеличению длины очередей и времени ожидания. Рассмотрим значения в таблице 1.

Таблица 1 – Характеристика клиентопотока по времени суток

Время суток	Среднее кол-во клиентов в час (λ)	Среднее время обслуживания одного клиента ($1/\mu$), мин	Кол-во окон (с)
Утро (9:00 – 12:00)	29	5	3
День (12:00 – 15:00)	36	4	4
Вечер (15:00 – 18:00)	11	6	2

Для моделирования процесса обслуживания клиентов мы определяем: λ (коэффициент приема) – среднее количество клиентов, приходящих в единицу времени; μ (коэффициент обслуживания) – среднее количество клиентов,

обслуживаемых одним кассиром за единицу времени; c (количество серверов) – количество активных счетчиков обслуживания в любой момент времени.

Используя модели теории массового обслуживания, такие как M/M/1 и M/M/c, мы можем проанализировать производительность системы и выявить потенциальные узкие места [12]. Эти модели помогают рассчитать: предполагаемое время ожидания в очереди (W_q), ожидаемое количество клиентов в системе (L), вероятность перегрузки системы (P). Они также позволяют сравнивать различные конфигурации обслуживания и выбирать оптимальные параметры для повышения эффективности.

Пример расчета.

Рассмотрим отделение банка, где: средняя скорость прихода клиентов: $\lambda = 20$ клиентов в час; средняя скорость обслуживания одного кассира: $\mu = 10$ клиентов в час; количество кассиров: $c = 3$.

Для анализа производительности системы используем модель массового обслуживания M/M/c.

Шаг 1. Вычисляем загрузку системы.

Коэффициент загрузки определяется как: $\rho = \frac{\lambda}{c\mu} = \frac{20}{3 \cdot 10} = \frac{20}{30} = 0.67$.

Это означает, что система работает на 67% от своей общей мощности.

Шаг 2. Вероятность простоя системы.

Вероятность того, что в системе нет клиентов (вероятность простоя), определяется формулой Эрланга-B:

$$P_0 = \left(\sum_{n=0}^c \frac{(\lambda/\mu)^n}{n!} \right)^{-1},$$
$$P_0 = \left(\sum_{n=0}^3 \frac{(20/10)^n}{n!} \right)^{-1} = \left(1 + 2 + \frac{4}{2} + \frac{8}{6} \right)^{-1} = (1 + 2 + 2 + 1.33)^{-1},$$
$$P_0 = (6.33)^{-1} = 0.158.$$

Таким образом, вероятность того, что в системе нет клиентов, составляет 15,8%.

Шаг 3. Среднее количество клиентов в системе.

$$L = \frac{\lambda}{\mu} + \frac{p^c}{c!(1-p)} P_0.$$

Используя ранее найденные значения:

$$L = \frac{20}{10} + \frac{0.67^3}{3!(1-0.67)} \cdot 0.158,$$

$$L = 2 + \frac{0.3}{6 \cdot 0.33} \cdot 0.158,$$

$$L \approx 2 + 0.04 = 2.04.$$

Таким образом, в среднем в любой момент времени в системе находится 2,04 клиента.

Шаг 4. Среднее время ожидания в очереди.

Время ожидания в очереди (W_q) задается формулой:

$$W_q = \frac{L - \lambda/\mu}{\lambda} = \frac{2.04 - 2}{20},$$

$$W_q = \frac{0.04}{20} = 0.002 \text{ часа} = 7.2 \text{ секунды}.$$

Таким образом, среднее время ожидания в очереди составляет 7,2 секунды.

Шаг 5. Заключение:

- система работает эффективно, в среднем в любой момент времени в системе находится всего 2 клиента;
- вероятность того, что в банке не будет клиентов, составляет 15,8%;
- среднее время ожидания в очереди составляет всего 7,2 секунды, что свидетельствует о бесперебойном обслуживании;
- коэффициент загрузки в 67% показывает, что банк может обслужить большее количество клиентов до возникновения задержек.

Полученные результаты позволяют прогнозировать поведение системы при разном уровне количества потока клиентов. А также это решение демонстрирует, как модель М/М/с может применяться для анализа реальных банковских услуг.

2.2 Сбор и анализ данных

Для оценки эффективности системы обслуживания клиентов в банке используются различные методы сбора данных, такие как наблюдение, опросы, анализ статистики обращений и времени ожидания. Эти методы позволяют объективно оценить качество обслуживания и выявить факторы, влияющие на удовлетворенность клиентов [7], [16].

Методы сбора данных:

- 1) мониторинг работы отделений банка:
 - наблюдение за потоками клиентов и временем обслуживания в реальных условиях,
 - регистрация часов работы в пиковые и нерабочие часы,
 - отслеживание загруженности персонала и скопления очередей;
- 2) анализ системных журналов:
 - извлечение времени транзакций из цифровых банковских систем,
 - анализ временных меток прихода и ухода клиентов,
 - оценка длительности очередей с течением времени;
- 3) опросы клиентов:
 - сбор отзывов о предполагаемом времени ожидания и удовлетворенности обслуживанием,
 - выявление проблемных точек в скорости и качестве обслуживания,
 - сравнение субъективных впечатлений с фактическими данными об очередях.

Далее рассмотрим ключевые показатели эффективности (КПЭ), которые позволяют объективно оценить работу системы обслуживания, выявить проблемы и обоснованно принимать решения по её оптимизации. На основе собранных данных анализируются следующие показатели.

- 1) Среднее время ожидания в очереди (W_q):

- представляет собой ожидаемое время, которое клиент тратит на ожидание обслуживания;
- рассчитывается с использованием закона Литтла:

$$W_q = \frac{L_q}{\lambda},$$

где L_q – среднее количество клиентов в очереди;

λ – количество приходящих клиентов.

На рисунке 7 показан график зависимости частоты прихода клиентов и среднего времени ожидания в очереди.

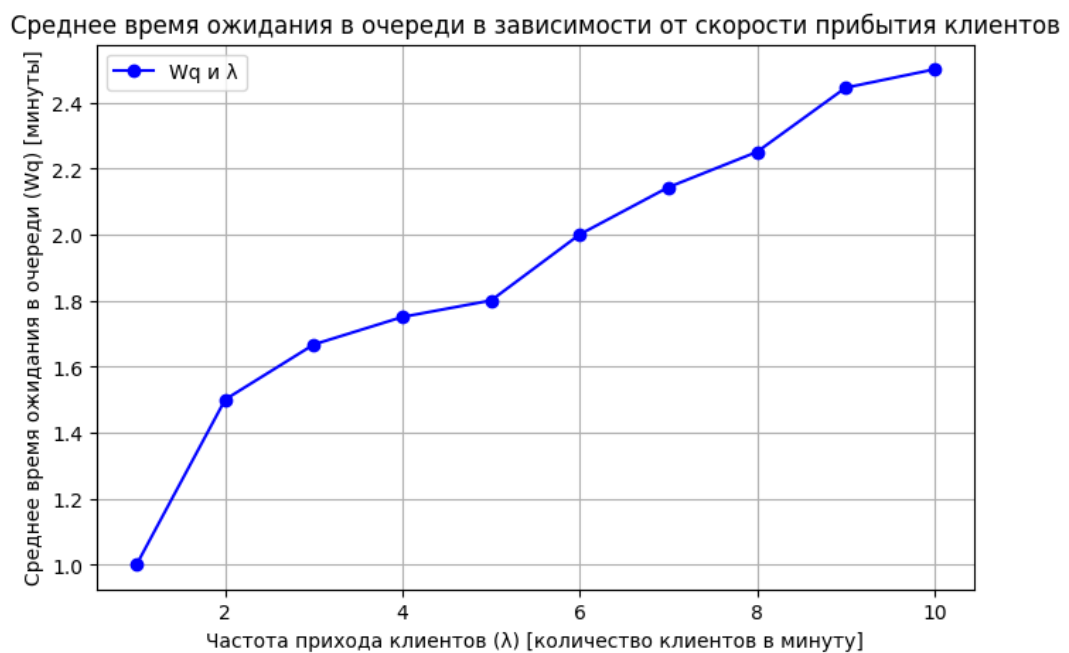


Рисунок 7 – График среднего времени ожидания в очереди

2) Длина очереди в разное время суток:

- измеряется путем отслеживания количества людей в очереди с течением времени;
- смоделировано с использованием системы М/М/с:

$$L_q = \frac{(p^c / c!)(1-p)P_0}{(1-p)^2},$$

где $p = \lambda / (c\mu)$ – нагрузка на систему;

c – количество серверов.

Далее на рисунке 8 увидим, как изменяется длина очереди в разное время суток.

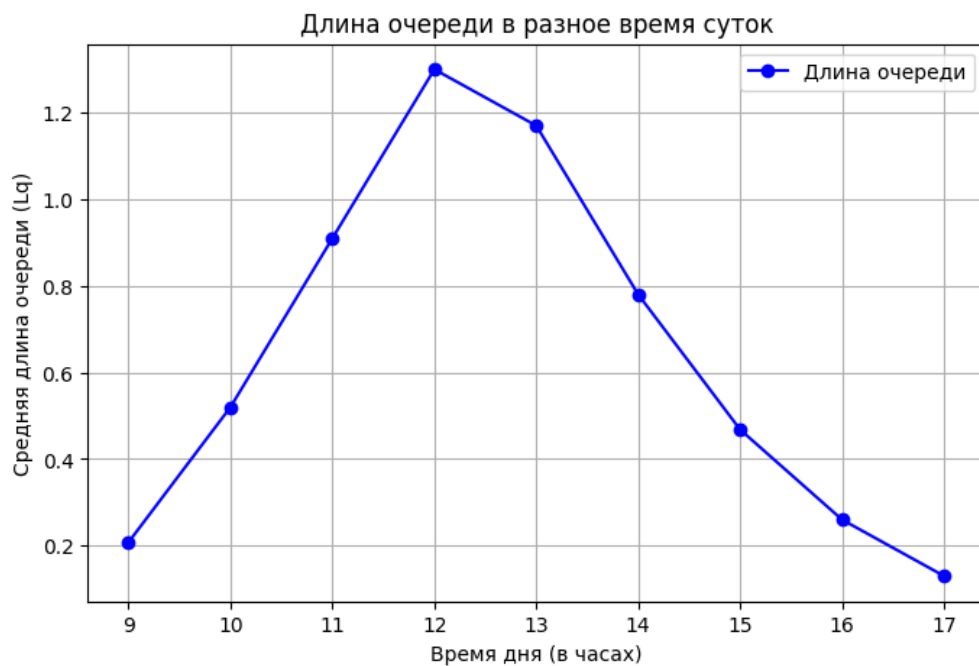


Рисунок 8 – График длины очереди в течение дня

3) Коэффициент загрузки персонала:

- показывает, насколько эффективно используются сотрудники;
- определяется как:

$$U = \frac{\lambda}{c\mu}.$$

Если $U \approx 1$, персонал перегружен работой.

Если $U \ll 1$, ресурсы используются недостаточно.

На основе формул коэффициента загрузки и средней длины очереди для системы М/М/с составим таблицу 2.

Таблица 2 – Расчет загрузки системы и средней длины очереди

Период времени	λ (чел/час)	μ (чел/час)	c	p	L_q (чел)
Утро	29	14	3	0.690	2.79
День	36	19	4	0.473	3.12
Вечер	11	5	2	1.1	Перегрузка

В вечернее время система перегружена ($p > 1$), требуется перераспределение ресурсов.

Пример расчета для отделения банка.

Предполагаем: частота прихода клиентов $\lambda = 15$ клиентов в час; скорость обслуживания на одного сотрудника $\mu = 6$ клиентов в час; количество кассиров $c = 3$.

Шаг 1. Загрузка системы.

$$p = \frac{\lambda}{c\mu} = \frac{15}{3 \cdot 6} = \frac{15}{18} = 0.83.$$

Банк загружен на 83%, что указывает на загруженность системы и означает, что сотрудники почти полностью заняты, но не перегружены.

Шаг 2. Средняя длина очереди.

$$L_q \approx 2.5 \text{ клиента.}$$

В среднем в очереди стоят 2-3 клиента.

Шаг 3. Среднее время ожидания в очереди.

$$W_q = \frac{L_q}{\lambda} = \frac{2.5}{15} = 0.167 \text{ часов} = 10 \text{ минут.}$$

Клиенты ждут около 10 минут, прежде чем их обслужат.

Шаг 4.: Вероятность того, что клиенту придется подождать.

$$P_w = \frac{\frac{(cp)^c}{c!(1-p)}}{\sum_{n=0}^c \frac{(cp)^n}{n!} + \frac{(cp)^c}{c!(1-p)}}.$$

Считаем числитель:

$$\frac{(3 \cdot 0.83)^3}{3!(1-0.83)} = \frac{(2.5)^3}{6 \cdot 0.17} = \frac{15.625}{1}.$$

Далее знаменатель:

$$\begin{aligned} \sum_{n=0}^3 \frac{(3 \cdot 0.83)^n}{n!} &= \frac{(2.5)^0}{0!} + \frac{(2.5)^1}{1!} + \frac{(2.5)^2}{2!} + \frac{(2.5)^3}{3!} = \\ &= 1 + 2.5 + \frac{6.25}{2} + \frac{15.625}{6} = 1 + 2.5 + 3.125 + 2.604 = 9.229, \\ 9.229 + 15.625 &= 24.854, \\ P_w &= \frac{15.625}{24.854} \approx 0.628. \end{aligned}$$

Таким образом, вероятность того, что клиенту придется стоять в очереди, составляет 62,8%.

Вывод:

- в часы пика в среднем собирается очередь из 2-3 клиентов;
- среднее время ожидания клиентов составляет 10 минут, что может повлиять на удовлетворенность;
- сотрудники работают с загрузкой 83%, что говорит о том, что система эффективна, но близка к предельной;
- дальнейшая оптимизация (например, динамическое распределение кассиров) может повысить скорость обслуживания.

В этом разделе теоретические модели объединены с реальным анализом, что обеспечивает основанный на данных подход к пониманию эффективности банковских услуг. Применение имитационного моделирования позволило проверить поведение системы в различных условиях. Это помогло найти способы сделать работу системы быстрее и удобнее для клиентов.

2.3 Выявление проблемных зон

На основе собранных данных и расчетов были выявлены ключевые узкие места в системе массового обслуживания банка [20].

1) Длительное время ожидания в час пик.

Анализ данных о времени ожидания показывает, что в час пик средняя длина очереди значительно увеличивается, что приводит к длительным задержкам обслуживания. Вероятность ожидания в очереди рассчитывается по формуле Эрланга-С ($P_w \approx 62.8\%$) и подтверждает, что более половины клиентов сталкиваются с задержками.

Используя закон Литтла $L_q = \lambda W_q$, мы можем оценить среднюю длину очереди в час пик. Если среднее время ожидания в очереди (W_q) составляет 10 минут (0,167 часа), а скорость прибытия (λ) составляет 15 клиентов в час, тогда: $L_q = 15 * 0.167 = 2.505$. Это означает, что в час пик в среднем по меньшей мере 2-3 клиента всегда стоят в очереди, а при резком росте спроса длина очереди значительно увеличивается.

2) Неравномерное распределение нагрузки между окнами обслуживания.

Из наблюдений за длиной очереди было установлено, что некоторые кассиры перегружены, в то время как другие остаются недостаточно загруженными. Это связано с различием во времени обработки транзакций; клиенты отдают предпочтение определенным кассирам из-за предполагаемой скорости обслуживания; отсутствие динамического распределения ресурсов. Более сбалансированное распределение клиентов по доступным точкам обслуживания могло бы повысить эффективность. Алгоритмы динамического распределения, такие как организация очередей на основе приоритетов или модели прогнозирования на основе машинного обучения, могут помочь распределить нагрузку более равномерно.

3) Высокая частота отказов в обслуживании из-за длинных очередей.

Увеличение длины очереди приводит к увеличению вероятности отказа

в обслуживании, когда клиенты уходят, не будучи обслуженными. Это можно проанализировать с помощью формулы Эрланга-В, которая оценивает вероятность отказа в обслуживании при отсутствии зоны ожидания:

$$P_B = \frac{(cp)^c / c!}{\sum_{n=0}^c (cp)^n / n!}.$$

С $c=3$, $p=0,83$, и, используя факторные расчеты, вероятность отказа в обслуживании составляет приблизительно 7-10%. Это означает, что почти 1 из 10 клиентов может покинуть банк из-за чрезмерного времени ожидания, что приведет к потере дохода и неудовлетворенности клиентов.

Составим таблицу 3 для наглядности как количество кассиров (окон) влияет на загруженность.

Таблица 3 – Оптимизация числа окон

Кол-во окон (с)	Среднее время ожидания W_q (мин)	Вероятность отказа P_b (%)	Загрузка (%)
2	18.6	22.8	94
3	8.2	5.5	81
4	3.2	1.2	66

К основным факторам, влияющим на качество обслуживания и способствующим неэффективности системы массового обслуживания, относятся:

- недостаточное количество операционных окон: численность персонала банка может быть недостаточной для того, чтобы справляться с пиковыми нагрузками, что приводит к перегрузкам;
- неэффективное распределение клиентов: текущая политика создания очередей не учитывает различные потребности клиентов, такие как быстрые транзакции и сложные операции, требующие больше времени;

- сложность операций: некоторые транзакции занимают значительно больше времени, чем другие (например, заявки на получение кредита или снятие наличных), что влияет на среднее время обслуживания.

Возможные решения:

- 1) увеличение штата сотрудников в час пик – наем дополнительных кассиров или внедрение гибких смен;
- 2) внедрение интеллектуального распределения очередей – использование аналитики данных для направления клиентов к наименее загруженным окнам обслуживания;
- 3) поощрение цифровых банковских решений – сокращение количества физических посещений отделений за счет продвижения онлайн-банкинга для проведения простых транзакций.

Устраняя эти проблемные зоны, банк может сократить время ожидания, повысить эффективность обслуживания и повысить удовлетворенность клиентов.

Выводы по второй главе

В результате проведенного анализа текущего состояния обслуживания клиентов в банке были получены объективные оценки эффективности работы системы. Исследование показало, что наибольшая загруженность наблюдается в утренние и дневные часы, когда длина очередей и время ожидания возрастают. Математическое моделирование с применением моделей М/М/1 и М/М/с позволило оценить ключевые параметры: коэффициент загрузки, среднюю длину очереди, вероятность ожидания и отказа в обслуживании.

Выявлены основные проблемные зоны: перегруженность в вечернее время, неравномерное распределение клиентов по кассам, высокая вероятность отказа в обслуживании из-за длинных очередей. Эти слабые стороны системы связаны как с недостаточным количеством рабочих окон в пиковые часы, так и с отсутствием динамического механизма распределения

потока клиентов.

Проведенные расчеты подтвердили необходимость оптимизации существующей системы обслуживания. Предлагаются следующие направления улучшения: перераспределение персонала по сменам, внедрение интеллектуальных систем управления очередями и развитие удаленных сервисов. Полученные результаты создают основу для разработки более точных моделей, адаптированных под конкретные условия банковского отделения. Таким образом, предложенные меры могут существенно повысить качество обслуживания, сократить время ожидания и снизить нагрузку на персонал. Результаты анализа показывают, что для эффективной организации обслуживания в банке нужно использовать математические методы и инструменты моделирования.

Глава 3 Разработка и тестирование программного продукта

3.1 Программная реализация алгоритма

Для реализации модели очереди и визуализации работы системы массового обслуживания используем язык программирования Python. Этот язык был выбран по следующим причинам: простота синтаксиса – Python легко читается и пишется, что упрощает реализацию и отладку моделей; наличие специализированных библиотек – используемые в проекте библиотеки `simpy` и `matplotlib` позволяют удобно реализовывать моделирование по событиям, имитировать работу системы и строить графики; поддержка научных расчетов – Python широко используется в научных и исследовательских задачах благодаря таким модулям, как `numpy`, `scipy`, `simpy`; распространенность – Python работает на всех операционных системах и легко интегрируется в любые среды.

Код реализует имитационную модель обслуживания клиентов в банке, где посетители поступают с заданной статистической интенсивностью, обслуживаются и попадают в очередь.

Для начала установим библиотеку `simpy` командой *`pip install simpy`*, которая используется для создания моделей и моделирования процессов (расчет времени ожидания, длина очереди, загрузка ресурсов и др.).

Далее подключаем библиотеки (рисунок 9):

- `simpy`: библиотека для дискретно-событийного моделирования. Позволяет моделировать процессы, очереди и ресурсы;
- `random`: стандартная библиотека Python для генерации случайных чисел (в том числе из экспоненциального распределения);
- `matplotlib.pyplot`: библиотека для построения графиков (используется для визуализации времени ожидания клиентов).

```
import simpy
import random
import matplotlib.pyplot as plt
```

Рисунок 9 – Импорт библиотек

Следом идет ввод параметров пользователем, таких как количество окон, интенсивность прихода клиентов в минуту и интенсивность обслуживания клиентов в минуту (рисунок 10). Также используется try-эксерпт блок для обработки ошибок (если ввод неверный, программа завершится).

```
try:
    NUM_SERVERS = int(input("Введите количество окон (с): "))
    ARRIVAL_RATE = float(input("Введите интенсивность прихода клиентов в минуту ( $\lambda$ ): "))
    SERVICE_RATE = float(input("Введите интенсивность обслуживания клиентов в минуту ( $\mu$ ): "))
except ValueError:
    print("Ошибка: введите корректные числовые значения.")
    exit()
```

Рисунок 10 – Ввод параметров

Затем на рисунке 11 показано, как мы инициализируем параметры симуляции. Устанавливаем время моделирования на 10 мин и фиксируем генерацию случайных чисел для воспроизводимости результатов.

```
SIM_TIME = 10
RANDOM_SEED = 42
random.seed(RANDOM_SEED)
```

Рисунок 11 – Параметры симуляции

Переходим к основному блоку программы. Здесь мы моделируем главную функцию системы массового обслуживания и задаем параметры, которые ввели вручную. Следующим шагом моделируем одного клиента, то

есть клиент приходит в систему, фиксируется его время прихода, происходит ожидание окна (время ожидания до начала обслуживания добавляется в список) и затем клиент обслуживается `service_time`, сгенерированным из экспоненциального распределения. Далее генерируем следующих клиентов. Каждые $\exp(\lambda)$ минут создается новый клиент и запускается процесс его обработки. После чего происходит запуск симуляции и работы модели в течении заданного времени. Следующим этапом создается среда моделирования, в которой моделируется работа банка и получается список времен ожидания. В конце считается среднее время ожидания путем деления суммы всех времен ожидания каждого клиента в очереди до начала обслуживания на общее количество клиентов, а также по формуле Литтла считается средняя длина очереди (средняя интенсивность потока клиентов умножается на среднее время ожидания). Данный блок продемонстрирован на рисунке 12.

```

# --- Основная функция моделирования ---
def bank(env, num_servers, arrival_rate, service_rate):
    server = simpy.Resource(env, num_servers)
    wait_times = []

    def customer(name, env, server):
        arrival_time = env.now
        with server.request() as request:
            yield request
        wait = env.now - arrival_time
        wait_times.append(wait)
        service_time = random.expovariate(service_rate)
        yield env.timeout(service_time)

    def generate_customers(env):
        i = 0
        while True:
            yield env.timeout(random.expovariate(arrival_rate))
            env.process(customer(f"Клиент {i}", env, server))
            i += 1

    env.process(generate_customers(env))
    env.run(until=SIM_TIME)
    return wait_times

# --- Моделирование ---
env = simpy.Environment()
wait_times = bank(env, NUM_SERVERS, ARRIVAL_RATE, SERVICE_RATE)
avg_wait = sum(wait_times) / len(wait_times)
queue_len = ARRIVAL_RATE * avg_wait

```

Рисунок 12 – Основной блок программы

Далее выводятся ключевые результаты и строится линейный график, который показывает время ожидания каждого клиента (рисунок 13).

```

print("\nРезультаты моделирования:")
print(f"Среднее время ожидания (Wq): {round(avg_wait, 2)} мин")
print(f"Средняя длина очереди (Lq): {round(queue_len, 2)} клиентов")

plt.figure(figsize=(6, 4))
plt.plot(wait_times, label="Время ожидания клиентов", color='blue')
plt.axhline(avg_wait, color='red', linestyle='--', label=f"Среднее ожидание: {round(avg_wait, 2)} мин")
plt.title("Время ожидания клиентов в очереди")
plt.xlabel("Номер клиента")
plt.ylabel("Время ожидания (мин)")
plt.legend()
plt.tight_layout()
plt.show()

```

Рисунок 13 – Вывод результатов

3.2 Тестирование программы и проведение экспериментов

Протестируем разработанное программное обеспечение. Далее будет серия экспериментов с программой, моделирующей работу банка как системы массового обслуживания. Мы будем варьировать параметры: количество окон, интенсивность прихода клиентов (λ) и интенсивность обслуживания (μ). После каждого запуска – вывод и анализ результатов.

Эксперимент 1: перегрузка системы. Были введены следующие данные: количество окон – 2, интенсивность прихода клиентов в минуту – 8; интенсивность обслуживания клиентов в минуту – 3. Результаты вывода программы представлены на рисунке 14.

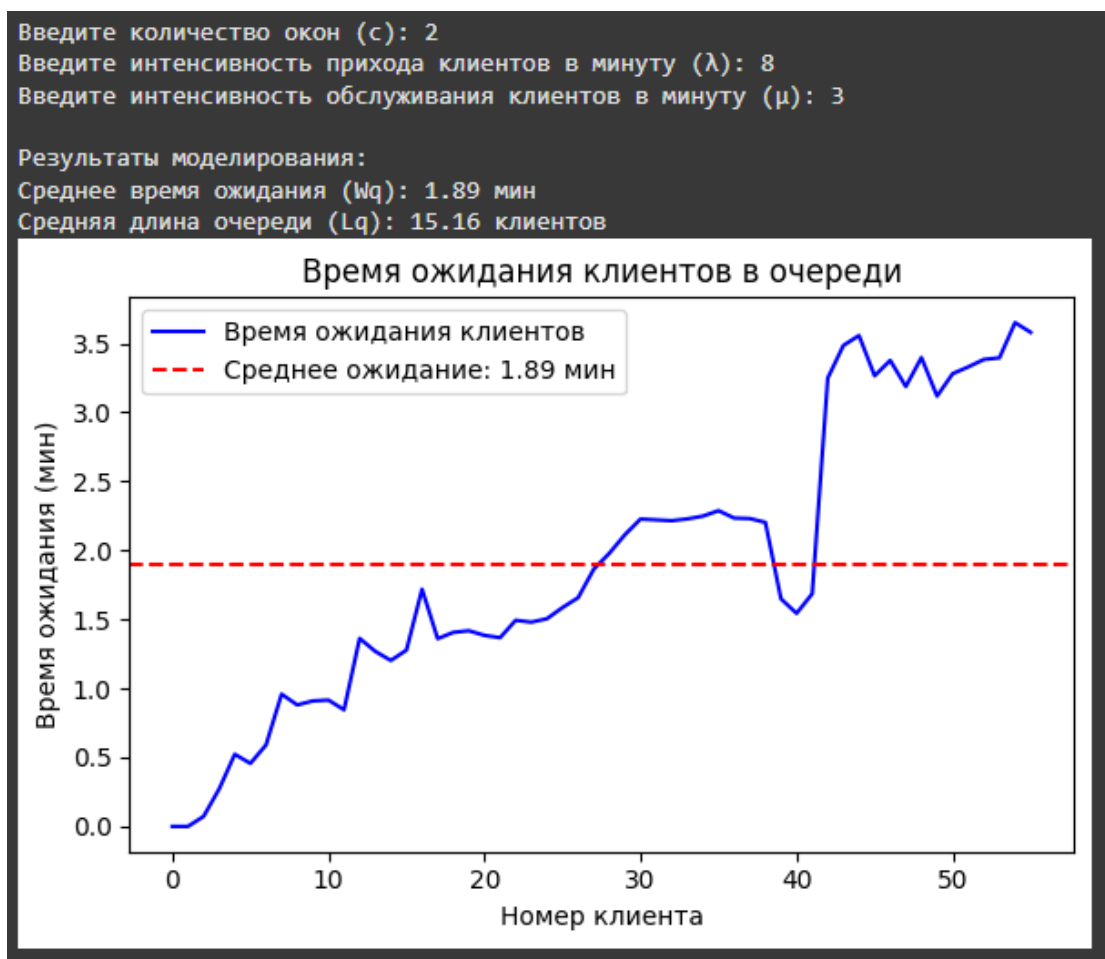


Рисунок 14 – Эксперимент 1

Суммарная пропускная способность $2 * 3 = 6$ клиентов/мин, а приходит – 8. Система перегружена: поступает больше клиентов, чем обслуживается. Очередь и время ожидания нарастают, высокая нагрузка на операторов.

Эксперимент 2: баланс на грани. Параметры: окна – 3; приход – 9; обслуживание – 3. Результат на рисунке 15.

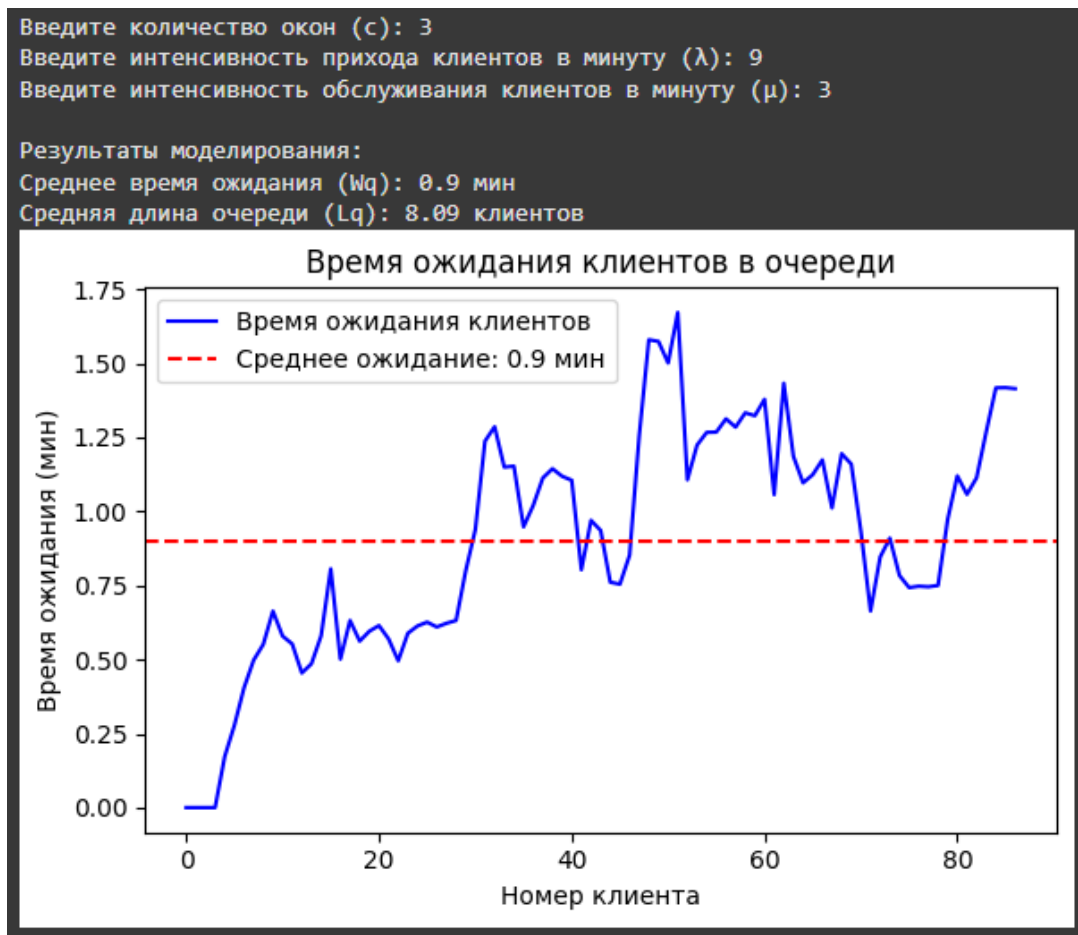


Рисунок 15 – Эксперимент 2

Пропускная способность $3 * 3 = 9$ клиентов/мин, полностью совпадает с потоком. Очередь не исчезает полностью, но стабилизирована. Это граничное состояние устойчивости, могут возникать кратковременные задержки.

Эксперимент 3: оптимальный режим. Количество окон – 3; скорость прихода клиентов – 6; скорость обслуживания – 3. Результаты на рисунке 16.

Введите количество окон (c): 3
Введите интенсивность прихода клиентов в минуту (λ): 6
Введите интенсивность обслуживания клиентов в минуту (μ): 3

Результаты моделирования:

Среднее время ожидания (W_q): 0.13 мин

Средняя длина очереди (L_q): 0.76 клиентов



Рисунок 16 – Эксперимент 3

Система устойчива: клиентов меньше, чем могут обработать. Средняя длина очереди минимальна, время обслуживания почти отсутствует. Такой режим подходит для качественного обслуживания без задержек.

Эксперимент 4: недоиспользование ресурсов. Входные данные: окна – 5; приход клиентов – 5; обслуживание – 2. Результаты на рисунке 17.

Введите количество окон (с): 5
Введите интенсивность прихода клиентов в минуту (λ): 5
Введите интенсивность обслуживания клиентов в минуту (μ): 2

Результаты моделирования:

Среднее время ожидания (W_q): 0.01 мин

Средняя длина очереди (L_q): 0.07 клиентов

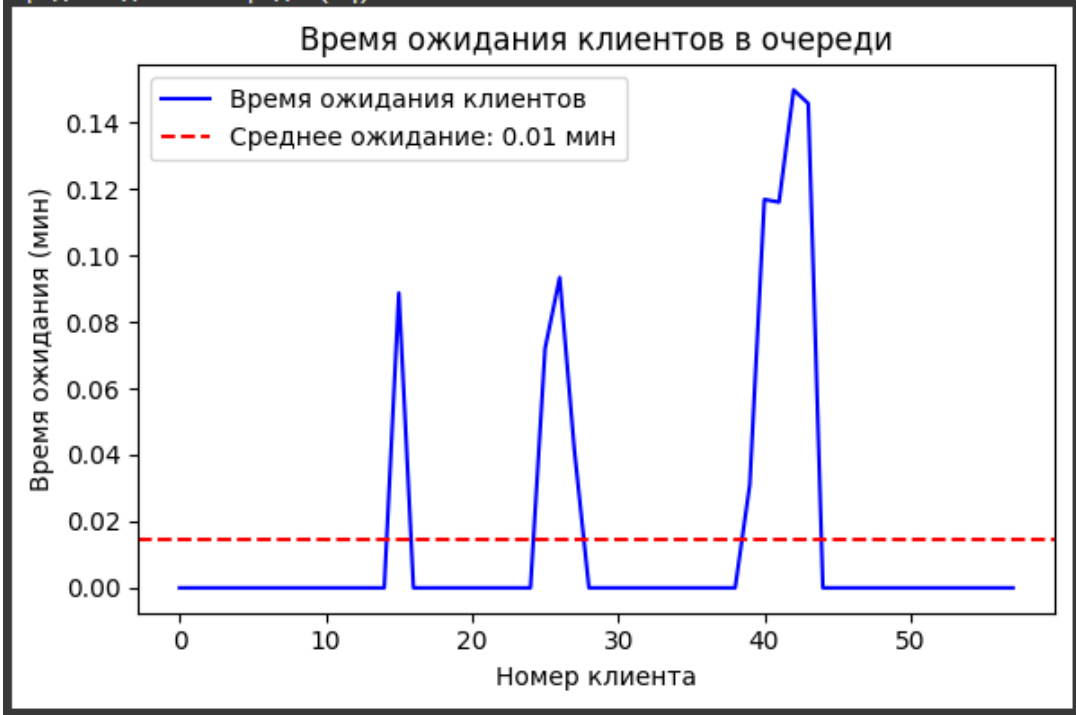


Рисунок 17 – Эксперимент 4

Пропускная способность $5 * 2 = 10$ клиентов/мин, нагрузка в 2 раза ниже. Очереди почти нет, но происходит перерасход ресурсов – окна простаивают. Нежелательно в плане оптимизации затрат.

Выводы по третьей главе

Нагрузка должна быть строго меньше совокупности интенсивности обслуживания клиентов и количества окон, иначе очередь растет бесконечно. Увеличение числа окон резко снижает ожидание, но не всегда целесообразно с точки зрения затрат. Оптимизация равна балансу между качеством обслуживания (малое W_q) и эффективностью загрузки ресурсов.

Заключение

В ходе выполнения данной выпускной работы было проведено исследование, которое показало, что применение математического моделирования очередей эффективно для оптимизации процессов обслуживания клиентов в банке. На основе анализа методов теории массового обслуживания, данных о клиентском потоке и созданной имитационной модели удалось оценить, как разные варианты числа окон влияют на показатели эффективности, такие как время ожидания и длину очереди.

Проведенные эксперименты показали, что правильная настройка количества окон обслуживания и правильное распределение потока клиентов позволяют значительно сократить очереди, повысить скорость обслуживания и удовлетворенность клиентов. Разработанное программное решение позволяет удобно проверять разные сценарии и помогает принимать обоснованные решения для повышения эффективности.

Результаты, полученные в этой работе, могут послужить основой для дальнейших исследований в области оптимизации обслуживания, включая разработку систем, которые автоматически подстраиваются под поток клиентов, а также использование методов машинного обучения для более точного прогнозирования и принятия решений. В общем и целом, математическое моделирование помогает банкам работать стабильнее и оставаться конкурентоспособными.

Таким образом, использование моделей массового обслуживания оказывается эффективным и полезным инструментом для современных банков, стремящихся улучшить качество обслуживания и удовлетворенность клиентов.

Список используемой литературы

1. Банки и банковское дело: в 2 ч. Ч. 2: учебник и практикум для вузов / В.А. Боровкова [и др.]; под ред. В.А. Боровковой. – 5-е изд., перераб. и доп. – М.: Юрайт, 2021. – 189 с.
2. Банковское дело: учебник / под ред. Г.Г. Коробовой. – 2-е изд., перераб. и доп. – М.: Магистр: ИНФРА-М, 2019. – 592 с.
3. Бережная Е.В. Математические методы моделирования экономических систем. – М.: Инфра-М, 2005. – 320 с.
4. Вентцель Е.С. Исследование операций: задачи, принципы, методология. – М.: Дрофа, 2006. – 206 с.
5. Вентцель Е.С. Теория вероятностей. – 3-е изд., перераб. – М.: Инфра-М, 2004. – 384 с.
6. Данилов А.М., Гарькина И.А., Домке Э.Р. Математическое и компьютерное моделирование сложных систем: учеб. пособие. – Пенза: ПГУАС, 2011. – 296 с.
7. Журавлёв Ю.И. Имитационное моделирование систем. – М.: Наука, 2012. – 256 с.
8. Захарова Е.А. Оценка результативности деятельности предприятия / Е.А. Захарова // Международный студенческий научный вестник. – 2017. – № 5. – С. 71.
9. Каштанов В.А. Теория массового обслуживания. – М.: ЮНИТИ, 2008. – 221 с.
10. Котляров П.А. Теория вероятностей и математическая статистика. – М.: КНОРУС, 2017. – 320 с.
11. Кутузов А.А. Моделирование бизнес-процессов: учеб. пособие. – СПб.: Питер, 2018. – 240 с.
12. Малыгина Л.А. Основы математического моделирования: учеб. пособие. – М.: Юрайт, 2020. – 198 с.
13. Осипов Г.С. Теория систем и системный анализ. – М.:

Академический проект, 2020. – 290 с.

14. Плескунов М.А. Теория массового обслуживания: учеб. пособие. – Екатеринбург: УрФУ, 2022. – 264 с.

15. Полищук В.И. Моделирование и оптимизация систем массового обслуживания. – Новосибирск: НГУ, 2015. – 168 с.

16. Радюкова Я.Ю., Чернышова О.Н., Федорова А.Ю. Банковский менеджмент: учебник. – М.: ИНФРА-М, 2020. – 379 с.

17. Сериденко В.В., Бялецкая Е.М. Алгоритмы повышения эффективности банковской системы // Вестник науки. – 2024. – № 6 (75). – С. 347–352.

18. Теория массового обслуживания: учеб. пособие / Т.В. Черушева, Н.В. Зверовщикова. – Пенза: ПГУ, 2021. – 224 с.

19. Шапиро Л.И. Теория массового обслуживания: задачи и решения. – М.: Физматлит, 2016. – 224 с.

20. Ястремская О.А. Сервисные технологии в банковской сфере. – М.: Финансы и статистика, 2019. – 210 с.

21. Banks J., Carson J.S., Nelson B.L., Nicol D.M. Discrete-Event System Simulation. – 5th ed. – Upper Saddle River: Pearson, 2010. – 640 p.

22. Gross D., Shortle J.F., Thompson J.M., Harris C.M. Fundamentals of Queueing Theory. – 5th ed. – Hoboken: Wiley, 2018. – 560 p.

23. Hillier F.S., Lieberman G.J. Introduction to Operations Research. – 10th ed. – New York: McGraw-Hill, 2014. – 1136 p.

24. Law A.M., Kelton W.D. Simulation Modeling and Analysis. – 5th ed. – New York: McGraw-Hill, 2014. – 768 p.

25. Teknomo K. Queueing Theory Tutorial [Электронный ресурс]. – 2019. – URL: <https://people.revoledu.com/kardi/tutorial/Queueing/index.html>.