

**МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ  
РОССИЙСКОЙ ФЕДЕРАЦИИ  
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ  
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ  
«РОССИЙСКИЙ УНИВЕРСИТЕТ ДРУЖБЫ НАРОДОВ»**

Факультет физико-математических и естественных наук

Кафедра прикладной информатики и теории вероятностей

«Допустить к защите»

Заведующий кафедрой  
прикладной информатики  
и теории вероятностей

д.т.н., профессор

\_\_\_\_\_ К.Е. Самуйлов

« \_\_\_\_ » \_\_\_\_\_ 20\_\_ г.

**Выпускная квалификационная работа  
магистра**

Направление 02.04.02 «Фундаментальная информатика и информационные технологии»

Магистерская программа «Управление инфокоммуникациями и интеллектуальные системы»

ТЕМА «Разработка моделей с управляемым по сигналам перераспределением  
радиоресурсов между сегментами беспроводной сети 5G»

Выполнил студент Бурцева София Артемовна

(Фамилия, имя, отчество)

Группа НФИмд-02-20

Руководитель выпускной  
квалификационной работы

Студ. билет № 1032202160

Кочеткова Ирина Андреевна, к.ф.-м.н.,  
доцент кафедры прикладной  
информатики и теории вероятностей,  
(Ф.И.О., степень, звание, должность)

\_\_\_\_\_  
(Подпись)

Автор \_\_\_\_\_

(Подпись)

г. Москва

2022 г.

**Федеральное государственное автономное образовательное учреждение  
высшего образования  
«Российский университет дружбы народов»**

**АННОТАЦИЯ  
выпускной квалификационной работы**

\_\_\_\_\_ Бурцева София Артемовна \_\_\_\_\_

(фамилия, имя, отчество)

на тему: Разработка моделей с управляемым по сигналам перераспределением радиоресурсов между сегментами беспроводной сети 5G

Данная работа представлена следующими разделами: введение, основной части, включающей в себя три отдельных раздела, заключения, и источников используемых материалов.

Во введении описывается цель работы, актуальность выбранной темы, методы исследования и структура данной ВКР.

В первом разделе представлена краткая характеристика концепции рассматриваемой темы ВКР. Проведён обзор одной из моделей по удовлетворению потребностей к качеству нарезки ресурсов сети.

Во втором разделе представлена характеристика системной и математической моделей нарезки трафика для двух виртуальных операторов с различным поведением пользователей данных услуг. Приведен алгоритм нарезки ресурсов для двух виртуальных операторов, параметры оптимизации и их численный анализ.

В последнем разделе приведён численный анализ показателей эффективности модели из второго раздела. Представлены проверочные тестовые данные и определение реальных данных для численных результатов.

В заключении подводятся итог вышеописанным практическим задачам.

Работа содержит 64 страницы печатного текста, 10 таблиц, 13 рисунков. Для написания ВКР использовано 24 источника.

Автор ВКР \_\_\_\_\_

(Подпись)

\_\_\_\_\_ Бурцева. С.А \_\_\_\_\_

(ФИО)

## Оглавление

Список сокращений.....	4
Список обозначений.....	6
Введение .....	9
Глава 1. Нарезка радиоресурсов и задача исследования .....	12
1.1. Особенности нарезки радиоресурсов.....	12
1.2. Использование машинного обучения для нарезки ресурсов .....	17
1.3. Вероятностная модель динамической нарезки.....	29
1.4. Постановка задачи исследования.....	39
Глава 2. Модель управляемой контроллером нарезки радиоресурсов.....	40
2.1. Принципы нарезки .....	40
2.2. Системная модель.....	44
2.3. Система массового обслуживания с сигналами .....	46
2.4. Показатели эффективности перераспределения ресурса .....	49
Глава 3. Численный анализ показателей эффективности нарезки радиоресурсов	52
3.1. Пример численного сравнения с фиксированной нагрузкой .....	<b>Ошибка!</b>
<b>Закладка не определена.</b>	
3.2. Разработка сценария для просмотра веб-страниц и групповой передачи данных.....	54
Заключение.....	59
Литература.....	60
Приложение 1. Тезисы на конференцию ИТТММ-2022 .....	63
Приложение 2. Программа ЭВМ. Расчет показателей эффективности модели динамической нарезки между двумя сегментами радиоресурсов с управлением по внешнему событию .....	65
Приложение 3. Программа ЭВМ. Расчет показателей эффективности модели динамической нарезки между двумя сегментами радиоресурсов с управлением по внешнему событию .....	66

## Список сокращений

### Русскоязычные сокращения

ВКР	Выпускная квалификационная работа
ИИ	Искусственный интеллект
СУГБ	Система уравнений глобального баланса
СУЛБ	Система уравнений локального баланса

### Англоязычные сокращения

3GPP	3rd Generation Partnership Project
5G	5th Generation Network
AI	Artificial Intelligence
BBU	LTE baseband unit
CSC	Communication service client
DDPG	deep deterministic policy gradient
DRL	deep reinforcement learning
E2E	End-to-end
InP	Infrastructure Provider
KPI	Key Performance Indicator
MAE	Mean Absolute Error
MCS	Modulation and Coding Schemes
MNO	Mobile Network Operator
NF	Network Function
NFVI	Network Virtualization Infrastructure
NN	Neural Network
ONF	Open Network Foundation
OPEX	Operating Expenses
QoS	Quality of Service
RL	Reinforcement Learning
SDN	Software-defined Network
SLA	Service Level Agreements
SNR	Signal-to-Noise Ratio
vRAN	Virtualization Radio Access Network

WAN

Wide Area Network

## Список обозначений

### Первый раздел

$N$	Число требований в системе в момент времени $t$
$\lambda$	Интенсивность потока поступления требований ( $\lambda \sim Pois$ )
$b$	Минимальная гарантированная скорость
$\theta$	Среднее значение длины блока данных
$r$	Длина очереди
$\gamma$	Интенсивность потери нетерпеливых заявок
$C$	Ширина полосы пропускания (общий объем ресурса)
$M$	Фиксированное число источников
$\varepsilon$	Интенсивность потока поступления требований ( $\varepsilon \sim Pois$ )
$\alpha$	Среднее значение среди параметров соответствия SLA.
$\alpha_k$	Коэффициент, характеризующий соответствие SLA для $k$ -ого виртуального оператора – вычисляется следующим образом
$\beta$	Доля полезных вызовов нарезки.
$\overline{UTIL}$	Среднее значение по коэффициентам занятости ресурсов виртуальных операторов
$UTIL_k$	Коэффициент занятости ресурса $k$ -ого виртуального оператора
$K$	Количество виртуальных операторов, в данном случае 2.
$t_i$	Момент времени, в который происходит перераспределение ресурса.
$N_k(t_i)$	Максимальное число позиций на приборе, которые были отданы для обслуживания заявок $k$ -ого виртуального оператора в момент времени $t_i$ ,
$N_k^{SLA}$	Изначально выделенное количество мест на приборе заявки $k$ -ого виртуального оператора (в момент времени $t_0$ .)
$n_k$	Количество заявок на приборах $k$ -го виртуального оператора
$m_k$	Количество заявок в очереди $k$ -го виртуального оператора
$k$	Множество виртуальных операторов

### Второй раздел

$k$	Номер сервера, $k = \{1,2\}$
$N_1$	Количество виртуальных машин на первом сервере
$N_2$	Количество виртуальных машин на втором сервере

$\mu_1$	Интенсивность экспоненциального распределения времени вычисления блоков данных в виртуальных машинах первого сервера
$\mu_2$	Интенсивность экспоненциального распределения времени вычисления блоков данных в виртуальных машинах второго сервера
$j$	Тип поступившего запроса $j = \{1,2\}$
$\lambda_1$	Интенсивность входящего потока запросов первого типа
$s_1$	Средний объем трафика первого типа
$b$	Скорость передачи трафика
$N$	Максимально число обслуживаемых сессий
$R_1$	Число мест в очереди для обслуживания сессий первого типа
$\varepsilon_1$	Порог времени ожидания обслуживания запросов первого типа
$\delta$	Интенсивность управляющих сигналов
$\lambda_2$	Интенсивность входящего потока запросов второго типа
$s_2$	Средний объем трафика запросов второго типа
$N$	Максимально число обслуживаемых сессий
$R_2$	Число мест в очереди для обслуживания сессий второго типа
$\varepsilon_2$	Порог времени ожидания обслуживания запросов второго типа

### Третий раздел

$k$	Тип услуги, $k = \{1,2\}$
$V$	$V$
$\delta$	$\delta$
$\mu_1^{-1}$	Объем данных первой услуги
$T_1^1$	Нижняя граница диапазона времени задержки первой услуги
$T_2^1$	Верхняя граница диапазона времени задержки первой услуги
$b_1$	Скорость передачи данных первой услуги
$R_1$	Количество мест в очереди для запросов первой услуги
$\lambda_1$	Интенсивность поступления запросов первой услуги
$\mu_2^{-1}$	Объем данных второй услуги
$T_1^2$	Нижняя граница диапазона времени задержки второй услуги
$T_2^2$	Верхняя граница диапазона времени задержки второй услуги
$b_2$	Скорость передачи данных второй услуги
$R_2$	Количество мест в очереди для запросов второй услуги
$\lambda_2$	Интенсивность поступления запросов второй услуги

$T^2$	Верхний порог на допустимое время задержки в очереди
$T^1$	Нижний порог на допустимое время задержки в очереди
$\varepsilon$	Интенсивность ухода по нетерпеливости



## **Введение**

Если посмотреть статистические данные за 2021 года, то можно узнать, что количество пользователей интернета составляет 4,66 миллиарда человек по всему миру, что на 316 миллионов (7,3 %) больше, чем в прошлом году. Если судить об использовании мобильных устройств, то на сегодня смартфонами пользуются 5,22 миллиарда человек — 66,6 % мирового населения. С 2020 года на 1,8 % (93 миллиона) увеличилось количество уникальных пользователей, использующих мобильные устройства, в то время как общее количество мобильных подключений увеличилось на 72 миллиона (0,9 %) и достигло 8,02 миллиарда к 2021 году. Всё больше человек используют мобильную связь как основной способ взаимодействия с сетью. Уже в этом году имеется прирост на 124 миллиона (2,4%). На данный момент число пользователей составляет 5,19 миллиарда человек [1].

С увеличением числа пользователей у компаний, предоставляющих телекоммуникационные услуги, возникает ряд проблем масштабирования, организации и эксплуатации сетевых ресурсов. Поэтому в настоящий момент всё так же продолжается выпуск ряда новых спецификаций организациями ITU и 3GPP для сетей нового поколения [2,3]. Тема данной выпускной работы представляет собой описание одной из ключевых и перспективных функций, которые должны быть задействованы в сетях 5G [4,5,6]. Так же дано описание применимости технологии нейронных сетей для удовлетворения запросов к Network Slicing [14,15,16,17].

### **Цель работы**

Продолжение работы над моделями с разделением ресурсов, начатой в период курса бакалавриата, исследование двухсервисной модели из работы [23]. Описание основных параметров модели для большего удобства последующего моделирования. Проведение численного анализа описанной модели.

### **Методы исследования**

В работе применен комплексный анализ большого количества литературы, проведен обзор технологии нейронных сетей и их применимости к теме ВКР [18,19,20], построение моделей и структур, использующихся как наглядное пособие и следование структуре определённой модели [22]. Созданы иллюстрации отражающие построение моделей СМО для наглядного описания параметров приведённых моделей, а так же графики для наглядности представления численного

анализа. Используются некоторые существующие формулы и методы расчёта характеристик из [24]. Был проведён анализ существующих стандартов и спецификаций [2,5,6] для анализа результатов моделирования при реальных исходных данных.

### **Научная новизна**

Описания моделей, представленные в работе, писались самостоятельно с изучением дополнительных материалов и тщательным анализом представленной литературы [21].

Представленные в работе модели разрабатывались с учётом стандартов сетей следующего поколения 5G (5th Generation Network) [5, 6]. В настоящее время архитектура сетей следующего поколения находится в разработке, поэтому тема данной ВКР является актуальной [7, 8]. Развёртывание частных сетей 5G в некоторых странах так же показывает актуальность исследования по распределению ресурсов данных сетей.

### **Практическая ценность работы**

Представленные в настоящей работе модели могут быть применимы к существующим сетевым ресурсам, удовлетворяя требования, объявленные в стандартах 5G. Данные о производительности сетей 5G являются основополагающими для мониторинга, оценки, анализа, оптимизации и обеспечения качества сети. Моделирование нижеприведённых сетей может помочь в получении интересующей сетевых операторов информации.

### **Структура работы**

В первом разделе дана краткая характеристика представляемой концепции на основе изученной литературы [7,8,9,10] а также возможности использования ИИ в данной концепции[11,12,13]. Рассмотрены причины её возникновения и указаны принципы работы, приемлемые принятому стандарту. Во второй главе представлена модель реализации данной концепции. Описаны параметры сети и их подэлементов, а также обозначены оптимизируемые параметры.

### **Апробация работы**

По теме данной работы был подготовлен доклад на научный семинар ИТТММ-2022 (приложение 1). Была зарегистрирована программа ЭВМ в 2021 году (приложение 2). На регистрации находится программа ЭВМ «Расчет показателей

эффективности передачи данных с гарантированной минимальной скоростью в математической модели динамической нарезки радиоресурсов между двумя сегментами» (в приложении 3).

# Глава 1. Нарезка радиоресурсов и задача исследования

## 1.1. Особенности нарезки радиоресурсов

Нарезка радиоресурсов (Network slicing) в текущем времени является одной из главных разрабатываемых технологий беспроводных сетей 5G. Концепция позволяет определить несколько виртуальных сетей на одной физической сетевой инфраструктуре, благодаря чему операторы сотовой связи могут разворачивать изолированные друг от друга сети. Нарезка предоставляет возможность разделять одну физическую сеть на множество логически разделённых сегментов сети для удовлетворения требований различных сценариев использования. Одно из главных достоинств деления ресурсов на сегменты заключается в том, что такой подход требует меньших вложений в физическую инфраструктуру сети. Благодаря широкому покрытию данная технология позволяет расширить возможности эксплуатации сети, то есть станет важной составляющей сетей 5G, позволяющей принимать и обрабатывать данные в соответствии с конкретными требованиями клиентов.

В настоящее время поставщики технологий во всей отрасли разрабатывают решения, направленные на удовлетворение потребностей операторов, обусловленных быстрым ростом числа пользователей и устройств, подключенных к Интернету. По статистическим данным и на основании нескольких исследований в области телекоммуникаций, эта тенденция, которая будет оказывать все большее давление на системы и сети центров обработки данных, чтобы обеспечить соответствие спросу на предоставляемые сетями 5G услуги. Для каждой различных услуг требуются гарантированное качество обслуживания (QoS), статическое распределение ресурсов приводит к неэффективному использованию всех возможностей инфраструктуры оператора и приносит большие затраты. Вместо этого операторы заинтересованы в более экономных решениях по эксплуатации имеющихся у них ресурсов. Таким образом, только инновационная технология может решить указанные ранее задачи, и это технология Network Slicing или «Нарезка ресурсов сети».

Если описывать понятие Network Slicing более приближено к структуре глобальной сети – это метод создания нескольких уникальных логических и

виртуализированных сетей в общей многодоменной инфраструктуре. Используя виртуализацию сетевых функций (NFV), программно-определяемую сеть (SDN), оркестровку, аналитику и автоматизацию, операторы мобильной сети (MNO) могут быстро создавать сетевые сегменты, которые могут поддерживать конкретное приложение, услугу, группу пользователей или сеть. Нарезки сети могут охватывать несколько сетевых доменов, включая доступ, ядро и транспорт, и могут быть развернуты несколькими операторами[2].

Network Slicing относится к разделению физической сети на несколько виртуальных сетей; каждую сеть можно настроить и оптимизировать для конкретного типа приложения или абонента. Используя технологии облачных вычислений и виртуализации, общие физические сетевые ресурсы можно динамически и эффективно планировать для логических сегментов сети в зависимости от меняющихся требований пользователей. Сегмент сети 5G состоит из набора сетевых функций и настроек, которые объединены вместе для конкретного варианта использования или бизнес-модели. Сегмент сети радиодоступа, базовые сети, работающие в распределенной облачной инфраструктуре, и транспортные сети, поддерживающие гибкое размещение виртуальных функций. Основной принцип построения сегмента сети 5G заключается в предоставлении только настраиваемых функций, необходимых для обработки трафика конкретного варианта использования. Сегменты сети обладают не только настраиваемыми возможностями, необходимыми для соответствующих услуг, но также способностью адаптироваться к меняющимся требованиям. Предоставление виртуализированной сквозной среды, которая может быть открыта для третьих сторон, является одной из ключевых функций, которая отличает сегментирование сети от совместного использования сети.

Нарезка сети включает три уровня[3]:

- 1) Уровень экземпляра службы: представляет конечного пользователя услуги или бизнес-услуги. Каждая служба представлена экземпляром службы.
- 2) Уровень экземпляра сетевого фрагмента: включает экземпляры сетевого фрагмента, которые могут быть предоставлены. Экземпляр

сетевого сегмента предоставляет сетевые функции, которые требуются экземпляру службы.

- 3) Уровень ресурсов: предоставляет все виртуальные или физические ресурсы и сетевые функции, необходимые для создания экземпляра сетевого сегмента.

## **Ресурсы**

В общем смысле ресурс - это управляемая базовым оператором единица физической инфраструктуры, определяемая набором атрибутов или возможностей, которые могут быть использованы для предоставления услуги. Сетевой сегмент имеет в своей основе ресурс, виртуальные надстройки которого в соответствующем сочетании удовлетворяют требованиям обслуживания пользовательских запросов, поддерживаемых таким сегментом. В сетевом сегменте имеется два типа ресурсов[4].

**Сетевые функции:** Функциональные блоки, которые предоставляют определенные сетевые возможности для поддержки и реализации конкретных услуг, требуемых в каждом случае использования. Обычно реализуются как экземпляры программного обеспечения, работающие на инфраструктурных ресурсах. Сетевые функции могут быть физическими (комбинация аппаратного и программного обеспечения конкретного поставщика, определяющая традиционное целевое физическое устройство) и/или виртуализированными (программное обеспечение сетевых функций отделено от аппаратного обеспечения, на котором оно работает)[5].

**Инфраструктурные ресурсы:** Гетерогенные вычислительные системы и необходимое программное обеспечение для размещения и подключения **Сетевых функций**. Они включают в себя вычислительное оборудование, хранилище данных, сетевые ресурсы (например, каналы связи и устройства коммутации/маршрутизации, обеспечивающие сетевое подключение) и физические ресурсы для радиодоступа. Пригодные для использования в Network Slicing, вышеупомянутые ресурсы и их атрибуты должны иметь высокий уровень абстракции и иметь логическое разделение на основе использования механизмов виртуализации, определяющих виртуальные ресурсы, которые могут использоваться так же, как и физические.

## **Виртуализация**

Виртуализация является ключевым процессом в Network Slising, поскольку она позволяет эффективно распределять ресурсы между сегментами. Виртуализация - это абстрагирование ресурсов с использованием соответствующих методов. Абстракция ресурса - это настройка ресурса таким образом, что его характеристики соответствуют заранее определенным выбранным критериям, скрывая или игнорируя аспекты, которые не имеют отношения к обозначенным критериям, в попытке упростить использование и управление этим ресурсом. Ресурсы, подлежащие виртуализации, могут быть физическими или уже виртуализированными, поддерживая рекурсивный шаблон с различными уровнями абстракции[6].

Точно так же, как виртуализация серверов [7] делает виртуальные машины (ВМ) независимыми от базового физического оборудования, виртуализация сети [8] позволяет создавать несколько изолированных виртуальных сетей, которые полностью отделены от базовой физической сети и могут безопасно работать поверх нее.

Внедрение виртуализации в сетевую сферу позволяет создавать новые бизнес-модели с новыми участниками процесса предоставления услуг и различными бизнес-ролями. Стандартный набор участников имеет три типа ролей:

- Поставщик инфраструктуры (InP): владеет и управляет физической инфраструктурой сети и составляющими ее ресурсами. Такие ресурсы в виде глобальной компьютерной сети (WAN) и/или центров обработки данных виртуализируются и затем предлагаются через программные интерфейсы одному или нескольким арендаторам.

- Арендатор: арендует виртуальные ресурсы у одного или нескольких InP в форме виртуальной сети, где арендатор может реализовывать, управлять и предоставлять сетевые услуги своим пользователям.

- Конечный пользователь: потребитель (части) услуг, предоставляемых арендатором, не предоставляя их другим участникам бизнес-процесса.

Как было описано выше, виртуализация может иметь рекурсивный характер, где участник процесса может быть арендатором на одном уровне виртуализации, и действовать как InP на уровне выше. Упомянутая здесь рекурсия подразумевает, что арендатор может предоставлять сетевые услуги как конечному пользователю, так и

другому арендатору. В таком случае этот второй арендатор будет предоставлять более продвинутые сетевые услуги своим собственным пользователям.

### **Оркестровка**

Оркестровка также является ключевым процессом для разделения сети. В общем смысле оркестровку можно определить как искусство объединения и координации разрозненных функций в единое целое. В Network Slicing, где задействованные участники процесса настолько разнообразны, необходим оркестратор для координации, казалось бы, разрозненных сетевых процессов для создания, управления и предоставления услуг.

Согласно ONF [9], оркестровка определяется как непрерывный процесс выбора ресурсов для оптимального выполнения требований клиентского обслуживания. Идея оптимальности относится к политике оптимизации, которая управляет поведением оркестратора, которое, как ожидается, будет соответствовать всем конкретным политикам и соглашениям об уровне обслуживания (SLA), связанным с клиентами (например, арендаторами или конечными пользователями), которые запрашивают услуги. Так же ожидается, что доступные ресурсы, потребности в услугах и критерии оптимизации могут изменяться со временем. Интересно, что оркестровка также упоминается в [9] как определяющая характеристика контроллера Программно-определяемой сети (SDN).

ONF утверждает, что функции оркестратора включают в себя проверку потребности в услугах, зависящую от клиента, конфигурацию ресурсов и уведомление о событиях.

Однако при Network Slicing оркестровка не может выполняться одним централизованным объектом не только из-за сложности и широкого диапазона задач оркестровки, но и потому, что необходимо сохранить независимость управления и поддерживать возможность рекурсии. Структура, в которой каждый субъект виртуализации имеет объект, выполняющий функции оркестрации, кажется более подходящей для удовлетворения вышеуказанных требований. Сущности должны обмениваться информацией и делегировать функции между собой, чтобы гарантировать, что услуги, предоставляемые на определенном уровне абстракции, удовлетворяют требуемым уровням производительности при оптимальном использовании ресурсов.



## **Изоляция**

Сильная изоляция является основным требованием, которое должно быть выполнено для работы параллельных сегментов сети на общей многопользовательской основе. Изоляцию следует понимать с точки зрения:

**Производительности:** каждый сегмент создаётся для удовлетворения конкретных требований к услугам, обычно выражаемых в форме ключевых показателей эффективности (KPI). Изоляция производительности - это проблема E2E (End-to-End), и она должна гарантировать, что требования к производительности для конкретных услуг всегда выполняются на каждом сегменте, независимо от уровней перегрузки и производительности других сегментов.

**Безопасности и конфиденциальности:** атаки или сбои, происходящие в одном сегменте, не должны влиять на другие сегменты. Более того, каждый сегмент должен иметь независимые функции безопасности, которые не позволяют неавторизованным объектам иметь доступ на чтение или запись к информации о конфигурации / управлении / учете, зависящей от сегмента, и иметь возможность записывать любые из этих попыток, независимо от того, авторизованы они или нет.

**Управления:** каждый сегмент должен управляться независимо, как отдельная сеть.

Для достижения изоляции на каждом уровне виртуализации должен быть определен набор соответствующих согласованных политик и механизмов в соответствии с принципом рекурсии, представленным ранее. Политики (что должно быть сделано) содержат списки правил, которые описывают, как должны быть должным образом изолированы различные управляемые объекты, не вдаваясь в подробности того, как этого можно достичь. Механизмы (как это должно быть сделано) - это процессы, которые реализуются для обеспечения соблюдения определенных политик. С нашей точки зрения, чтобы полностью реализовать требуемый уровень изоляции, необходимо взаимодействие как виртуализации, так и оркестровки.

### **1.2. Использование машинного обучения для нарезки ресурсов**

Проблема управления инфраструктурой нарезки ресурсов сети может быть разделена на несколько фундаментальных задач, каждая из которых затем может

быть реализована с помощью подходящих инструментов искусственного интеллекта [10].

### ***Функции управления при нарезке ресурсов сети***

Совместное использование ресурсов между сегментами применяется на всех этапах управления жизненным циклом сетевого сегмента [11], который состоит из четырех основных этапов: подготовка, создание экземпляра, настройка и активация; плюс время работы и вывод из эксплуатации. Предлагаемая структура охватывает три ключевые функции, лежащие в основе первых трех этапов (последний этап не включает управленческих решений):

- Управление доступом - это важнейший механизм принятия решений во время подготовки сегмента, поскольку он определяет, могут ли предстоящие запросы к сетевому сегменту быть допущены к обработке в системе или нет, и выполняется так, чтобы гарантировать, что требования запросов будут удовлетворены.

- Реорганизация сетевых ресурсов является ключевой функцией, как для создания экземпляров сегментов, так и для операций во время выполнения, поскольку она распределяет доступные сетевые ресурсы для инициализированных сегментов наиболее эффективным из возможных способов, а затем динамически обновляет такое распределение во время выполнения, чтобы удовлетворять изменяющиеся во времени требования к каждому сегменту, избегая при этом перебоев в работе распределения ресурсов.

- Планирование радиоресурсов имеет первостепенное значение во время работы системы, поскольку оно управляет совместным использованием ресурсов радиодоступа между сегментами сети, обеспечивая выполнение потенциально строгих требований всех запросов (например, с точки зрения задержки и пропускной способности) через радиointерфейс.



Рис.1.1. Комплексная структура нарезки ресурсов сети.

На Рис.1.1 показаны вышеперечисленные функции, каждая из которых включает в себя разные временные рамки:

- Контроль допуска выполняется с частотой, которая соответствует частоте поступления запросов на новый сегмент сети, что может составлять несколько часов.
- Оркестровка ресурсов в программно-определяемых сетях происходит с частотой, которая зависит от времени, необходимого для изменения размера ресурсов виртуальных машин, обычно порядка минут.
- Планирование радиоресурсов применяется с более высокой степенью детализации, вплоть до миллисекундных интервалов в крайних случаях.

Отметим, что между ними могут взаимодействовать разные функции. В качестве примера информацию об использовании радиоресурсов можно использовать для управления доступом, чтобы понять, может ли прием новых секторов вызвать нехватку радиодоступа.

*Добавление ИИ в концепцию*

Все вышеперечисленные функции должны принимать решения для удовлетворения требований отдельных сегментов сети при максимизации общей производительности системы. Для этого им необходимо мониторить динамику трафика данных по сегментам и автоматически реагировать на их влияние на сетевую архитектуру для достижения соответствующих целей управления. Самонастраивающиеся конфигурации сетевых функций были введены более десяти лет назад; однако разработанные до сих пор решения обычно применяют функцию управления к ограниченному набору параметров, которые медленно меняются во времени. Кроме того, текущие подходы получают выходные данные, которые после требуют вмешательства человека для преобразования в модификации конфигурации сети (например, обновление транспортной сети с целью оптимизации передачи обслуживания в заданном регионе).

Эти характеристики несовместимы с новыми требованиями, предъявляемыми к нарезке ресурсов сети. Параметры, которые могут нуждаться в изменении конфигурации, гораздо более многочисленны, поскольку каждая функция виртуальной сети может отображать несколько из них программным способом. Временные рамки, в которые должны приниматься решения, резко сокращаются, так как в идеале нужно быть способным действовать на уровне радио- или даже проводной скорости. При принятии решений часто необходимо учитывать показатели, выходящие за рамки чистой производительности сети, такие как энергоэффективность или монетизация инфраструктуры, которые могут скрывать сложные перекрестные взаимосвязи.

Этот контекст создает благоприятную почву для использования ИИ в работе мобильных сетей. Для этой цели могут быть полезны все классы ИИ, включая контролируемые решения, требующие достоверных данных для обучения, неконтролируемые методы, которые работают при отсутствии достоверной информации, и подходы к обучению с подкреплением (RL), в которых возможны различные формы взаимодействия с управляемой системой [12]. Наиболее подходящие инструменты искусственного интеллекта следует выбирать в каждом конкретном случае, в зависимости от задействованных алгоритмических требований и сроков работы.

Например, RL особенно хорошо подходит, когда временная динамика проблемы может соответствовать кривой обучения, и цель состоит в том, чтобы определить последовательность действий, которые максимизируют определенное вознаграждение: это имеет место как в алгоритме управления доступом, так и в планировании радиоресурсов, как показано на практических реализациях, представленных ниже.

Обратите внимание, что методы RL необходимо обучать на очень большом количестве данных, что усложняет их применение в реальных средах, где исторических данных мало. Более того, подходы RL страдают от недостатка размерности, когда базовая модель становится слишком большой. Таким образом, в некоторых случаях, таких как описанный ниже, использование обучения без учителя в сочетании с RL может потребоваться для моделирования очень сложных отношений во входных данных. И наоборот, когда цель состоит в том, чтобы предоставить решения, которые не зависят от ранее принятых и качество которых можно оценить во время обучения системы, решения для обучения с учителем являются достаточно сильным вариантом: именно при этих настройках происходит оркестровка сетевых ресурсов, как показано на применяемом решении ниже. Здесь задача состоит в том, чтобы предоставить маркированные данные для алгоритма, что может быть невыполнимой задачей, если такие метки нельзя получить напрямую из сети.

Прежде чем продолжить, отметим, что представленные ниже примеры успешной интеграции ИИ в структуру, показанную на Рис. 1. Они не исчерпывают прикладное пространство ИИ для сетевых операций; скорее, они реализуют важные компоненты в комплексном дизайне самоорганизующихся сегментированных мобильных сетей.

### **3. Функции управления сегментами на основе ИИ**

Мы представляем жизнеспособные реализации функций управления сегментами в нашей структуре, которые полагаются на ИИ для управления доступом, оркестровкой и планирования радиоресурсов.

#### ***ИИ для контроля инициализации сегментов***

Ресурсы сетевой инфраструктуры ограничены, и требуются гарантии качества сетевых сегментов, что требует контроля доступа при запросах на инициализацию

новых сегментов. Согласно организации по стандартизации 3GPP по нарезке ресурсов сети, клиент службы связи (CSC) [10] (т. е. арендатор) будет запрашивать определенные услуги у поставщика услуг связи (CSP) (т. е. провайдера сети) из числа имеющихся в предложенном разделе услуг. Затем, он будет оплачивать услугу, сумма которой будет зависеть от таких показателей, как количество обслуживаемых пользователей, зона покрытия услуги и продолжительность работы экземпляра сегмента. Такие решения по контролю доступа имеют серьезные последствия для бизнеса: выбор количества сетевых сегментов для одновременной работы и способов совместного использования сетевой инфраструктуры между сегментами влияет на доходы сетевого провайдера.

На этапе контроля доступа необходимо найти компромисс между совместным использованием ресурсов и выполнением ключевых показателей эффективности (KPI). Если совместное использование ресурсов слишком агрессивно, требуемые ключевые показатели эффективности не могут быть достигнуты, то доходы падают, поскольку сегменты сети не предоставляют ожидаемых услуг; если вместо этого сетевые операторы будут придерживаться крайне консервативных подходов, они могут упустить существенные возможности для получения прибыли.

В конечном итоге тот факт, что более строгие гарантии ключевых показателей эффективности требуют изолированных, не разделяемых ресурсов, ограничивает допустимые конфигурации. В практических условиях этот и без того запутанный компромисс дополнительно осложняется многими технологическими и, возможно, изменяющимися во времени переменными, что делает поиск точки равновесия, которая максимизирует доходы на основе инициализированных сегментов, сложной задачей. Чтобы определить лучшее рабочее состояние, контроль инициализации сегментов должен изучать динамику поступления запросов и принимать решения по максимизации дохода на основе текущей занятости системы и ее ожидаемого долгосрочного развития.

Точные методы требуют, чтобы все переменные были известны и не увеличивались вместе с количеством состояний в пространстве-состояний системы (которое экспоненциально растет с количеством сегментов). Вместо этого ИИ предоставляет подходящие инструменты для поиска золотой середины между максимальным доходом в единицу времени и пространством параметров, в котором

ключевые показатели эффективности больше не применяются, при этом активно учитывается весь набор изменяющихся во времени релевантных переменных [13]. В частности, подходы глубокого обучения с подкреплением (DRL) взаимодействуют с большим набором переменных, которые характеризуют систему, делают вывод о ее стохастическом поведении, а затем определяют лучшую стратегию инициализации в соответствии с заданной целью. Нейронные сети с прямой связью (NN) используются в компонентах настройки сети в качестве подходящей архитектуры при работе с аппроксимацией функций. Вкратце, решение обучения с подкреплением работает следующим образом:

- По прибытии нового запроса сегмента система выполняет действие (т. е. Принимает или отклоняет запрос), которое максимизирует долгосрочный доход; каждая сеть отвечает за прогнозирование дохода, связанного с одним действием.

- После выполнения действия алгоритм взаимодействует с системой и оценивает качество полученного дохода с помощью специальной функции потерь. Это значение затем используется для обучения соответствующей нейронной сети.

- Во время такой фазы обучения используется неправильное действие (то есть действие, связанное с более низким долгосрочным доходом) вместо того, что фактически выполняется с низкой вероятностью  $\epsilon$ , что позволяет исследовать систему и подстраиваться под ее изменения с течением времени.

- Когда наблюдается существенное изменение в поведении системы, может быть запущено полное переобучение системы, чтобы адаптировать алгоритм к новым условиям.

Этот подход на основе ИИ имеет ряд преимуществ, в том числе: гибкость в адаптации к изменениям настроек системы благодаря короткому времени конвергенции; эффективная работа в ситуациях, которые раньше не встречались, благодаря надежным оценкам, предоставляемым архитектурами NN с прямой связью в этих случаях; и масштабируемость до больших размеров сети.

### ***ИИ для оркестровки сетевых ресурсов***

После инициализации сегментам должны быть выделены достаточные ресурсы. Из-за преобладающего программного обеспечения мобильных сетей такие ресурсы становятся все более вычислительными по своей природе. Для обеспечения надежных гарантий часто требуется, чтобы вычислительные ресурсы выделялись

исключительно определенным сегментам и не могли использоваться другими сегментами в то же время [14]. Фактически, динамическое распределение сетевых ресурсов по разным инициализированным сегментам является главной задачей управления в области нарезки ресурсов сети.

В этом контексте оператору сети необходимо заранее решить, какой объем ресурсов следует выделить для различных сегментов, чтобы обеспечить наиболее эффективное использование доступной емкости и, таким образом, минимизировать операционные расходы (ОРЕХ). Ключевой компромисс между:

- Недостаточное выделение ресурсов - если оператор выделяет меньшую мощность, чем требуется для удовлетворения запроса, это приводит к нарушению Соглашения об уровне обслуживания (SLA), заключенного с арендатором.

- Избыточное выделение ресурсов - избыточные ресурсы, назначенные срезу, подразумевают затраты с точки зрения излишне выделенных ресурсов, которые остаются неиспользованными.

Для поиска правильной рабочей ситуации необходимо:

1. Спрогнозировать будущий спрос на услугу в каждом сегменте [15].
2. Решить, какое количество ресурсов необходимо для удовлетворения такого спроса.

Удовлетворить эти два требования - сложно, потому что для прогнозирования будущих потребностей на уровне обслуживания требуется разработка специальных и точных предсказателей; кроме того, распределение ресурсов таким образом, чтобы недостаточное или избыточное выделение ресурсов позволяло минимизировать операционные расходы оператора, требует оценки ожидаемой (отрицательной и положительной) ошибки прогноза. Более того, рассмотрение пунктов 1 и 2, указанных выше как отдельных проблем, может привести в значительной степени к неоптимальным решениям, поскольку унаследованные предикторы не предоставляют надежной информации об ожидаемой ошибке, которую можно фактически получить.

Хотя сложность полного решения при использовании традиционных методов может быть слишком высокой, ИИ можно использовать для одновременного решения обоих аспектов, решив проблему прогнозирования пропускной способности. Это может быть реализовано путем обучения типичной сверхточной



архитектуры нейронной сети (CNN) для прогнозирования временных рядов с помощью специальной функции потерь, которая вместо простой минимизации ошибки учитывает соответствующие затраты, связанные с нарушениями соглашения об уровне обслуживания (SLA) и избыточной пропускной способностью [16]. В отличие от рекуррентной модели нейронной сети, сверточная позволяет использовать внутренние пространственные корреляции в трафике, генерируемом в разных географических точках.

Типичный прогноз трафика, минимизирующий среднюю абсолютную ошибку (MAE), сравнивается с прогнозом пропускной способности, который учитывает фактические затраты на недостаточное выделение ресурсов и превышение размеров. Сеть, обученная с помощью MAE, пытается предвидеть точный спрос. Это влечет за собой дорогостоящее недофинансирование в течение значительной части времени. Вместо этого сеть, обученная с помощью DeepSeg, может научиться измерять емкость в следующем временном интервале, чтобы избежать нарушений Соглашения об уровне обслуживания, сохраняя при этом избыточное выделение ресурсов на минимальном уровне, как также показано на среднем графике. Такое правильно настроенное решение на основе искусственного интеллекта позволяет определить ресурсы, которые должны быть заранее выделены для каждого сегмента, чтобы удовлетворить их будущие потребности. В конечном счете, этот подход решает проблемы 1 и 2 одновременно с помощью ИИ, сводя к минимуму избыточное выделение ресурсов и избегая нарушений требований услуг, генерируемых устаревшим предсказателем трафика, с значительной денежной экономией для оператора.

### ***ИИ для планирования сегментов при радиодоступе***

При радиодоступе ключевой проблемой нарезки ресурсов сети является разработка механизма виртуализации (vRAN) сети радиодоступа (RAN), который совместно обеспечивает изоляцию между сегментами сети и адаптирует распределение объединенных физических ресурсов к потребностям каждой виртуальной сети радиодоступа. Как подчеркивалось выше, это должно происходить в гораздо более короткие сроки, чем рассмотренные ранее. Таким образом, разработка алгоритмов, реализующих адаптивные политики для мультиплексирования ресурсов, имеет первостепенное значение.

Однако оптимизация распределения ресурсов является особенно сложной задачей, поскольку существует сильная нелинейная связь между политиками вычислений и радиоуправления, что очень затрудняет определение количества вычислительных ресурсов, которые должны быть выделены для сегмента в зависимости от количества радиоресурсов, уже выделенных для этого. Традиционные подходы к моделированию, которые требуют предварительной калибровки для конкретных условий и платформ, не подходят для практического сечения RAN. Это связано с тем, что связь между схемами модуляции и кодирования (MCS) и временем ЦП далеко не тривиальна и зависит от условий SNR, нагрузки трафика и платформы, на которой размещен BBU.

В этом контексте перспективным решением является сочетание обучения без учителя и глубокого обучения с подкреплением [17]. Действительно, методы неконтролируемого обучения, такие как глубокие автокодировщики (и популярные варианты), особенно полезны для проецирования многомерных контекстных данных, таких как шаблоны загрузки данных или шаблоны SNR, в (разреженные) скрытые пространства, в то время как алгоритмы DRL разрабатывают политики, отображающие такие (закодированной) контекстной информации в оптимальные действия по управлению ресурсами, такие как решения по планированию ЦП или радио, которые решают вышеупомянутые проблемы. Основываясь на этих методах, алгоритм глубокого детерминированного градиента политики (DDPG), реализованный структурами субъект-критик нейронной сети, может иметь дело с большими и / или непрерывными пространствами состояний, которые являются общими для задач управления ресурсами [18]. Точнее, алгоритмы «субъект-критик» являются наиболее подходящим типом решения для класса проблем, которые мы решаем, поскольку они сочетают в себе преимущества как классических подходов RL (на основе политик), которые поддерживают большие пространства действий, так и RL на основе значений, что обеспечивает более высокую эффективность и стабильность образца.

### **Практические преимущества управления сегментами на основе искусственного интеллекта**

Частые споры по поводу интеграции ИИ в сетевые системы заключаются в том, действительно ли эти новые архитектуры приносят существенное

преимущество в производительности по сравнению с традиционными подходами, достаточное для оправдания дополнительной сложности и потери интерпретируемости. Нарезка ресурсов сети не исключение;

***Повышение монетизации инфраструктуры с помощью оптимальной политики доступа***

Многое можно сказать о преимуществах интеллектуального контроля доступа в сегментах сети с денежной точки зрения. Ниже приведены данные, полученные провайдером при принятии решения о том, принимать ли два вида сетевых сегментов: один (расширенная мобильная широкополосная связь, eMBB, менее дорогой), включая эластичный трафик, и второй (сверхнадежная связь с малой задержкой, URLLC, дороже), включая неэластичный трафик с высокими требованиями к QoS. Результаты рассмотрения запросов на инициализацию, поток которых является Пуассоновским для обоих видов сегментов из [10], приведены в Табл. 1.1.

**Табл. 1.1. Сводка преимуществ, достигнутых ИИ по сравнению с устаревшими решениями для управления сегментированной сетью.**

<b>Функция</b>	<b>Показатель производительности</b>	<b>% улучшения - пример использования</b>
Контроль инициализации сетевого сегмента	Улучшение доходов	- 0,23% - Оптимальное, коэффициент 1
		- 3,77% - Оптимальное, коэффициент 20
		33,3% - Случайная политика, коэффициент 15
Распределение облачных ресурсов	Снижение денежной стоимости	81,6% - Facebook, основной дата-центр
		59,2% - Snapchat, дата-центр MEC
		64,3% - YouTube, дата-центр C-RAN
Распределение ресурсов виртуальной сети радиодоступа	Экономия ЦП, задержка, пропускная способность	30% - экономия ЦП по сравнению с планировщиками без использования ЦП
		5% - QoS на основе задержки через планировщики, без использования ЦП

		25% - Пропускная способность при дефиците вычислительной мощности
--	--	---

Эксперименты проводятся для различных соотношений цен между URLLC и eMBB (где коэффициент  $x$  означает, что URLLC в  $x$  раз дороже, чем eMBB) и разных размеров сети (сравнение с оптимальной политикой выполняется в сценарии меньшего масштаба как оптимальная политика не может быть запущен для больших сценариев). Результаты показывают, что описанный ранее алгоритм на основе искусственного интеллекта принимает разумные решения о допуске и обеспечивает очень существенный выигрыш по сравнению с эвристическими подходами, в то же время вызывая почти незначительные потери по сравнению с оптимальным, неосуществимым на практике эталонным тестом.

### ***Прогнозирование пропускной способности в реальном мире и на основе данных***

Чтобы оценить производительность ИИ для оркестровки разделенных сетевых ресурсов, описанных выше, мы рассмотрим три типичных примера:

- Сегмент, предназначенный для трафика Facebook в центре обработки данных базовой сети, который контролирует все 470 4G eNodeB, развернутых в большом мегаполисе.
- Сегмент, выделенный для трафика Snapchat в центре обработки данных граничных вычислений с множественным доступом (MEC), который обрабатывает трафик примерно 70 узлов eNodeB.
- Сегмент, вмещающий трафик, генерируемый приложением YouTube в центре обработки данных облачной RAN (C-RAN), расположенном поблизости от радиодоступа, который выполняет обработку и планирование основной полосы частот для 11 узлов eNodeB.

Прогноз, возвращаемый этой моделью, завышен на 5 процентов - разумная цифра при наличии достаточно точного прогноза - чтобы попытаться избежать нарушений Соглашения об уровне обслуживания. Подход к оркестровке сетевых ресурсов на основе ИИ, описанный ранее, позволяет существенно снизить денежные

затраты, связанные с распределением ресурсов, с экономией более 50 процентов во всех случаях.

### ***Оптимизация ЦП на базе искусственного интеллекта на периферии***

Наконец, описанный ранее подход к совместному планированию радиоресурсов и ЦП оценивается посредством экспериментального подтверждения. Стенд включает в себя набор программно-определяемых радиостанций (SDR) USRP в качестве внешних интерфейсов радиосвязи, подключенных к вычислительному узлу, на котором размещен программно-управляемый BBU eNodeB LTE, а именно srsLTE eNB [19]. Вышеупомянутый алгоритм обучения с подкреплением DDPG и интерфейс с eNodeB реализуются через интерфейс прикладного программирования Docker (API) для управления распределением вычислительных ресурсов каждого стека (в частности, для управления планировщиком ЦП CFS, используемым Docker), а также через простой сокет, для установления границ приемлемой связи между схемами модуляции и кодирования со стороны UE, связанных с eNB.

Для оценки производительности рассматриваются два сценария. В первом сценарии вычислительная мощность не ограничена, и цель состоит в том, чтобы добиться хорошего баланса между стоимостью системы (использование ЦП) и производительностью QoS (прокси для системной задержки). Наблюдается, что даже в высокодинамичных сценариях подход на основе ИИ может обеспечить 25-процентное улучшение производительности QoS по сравнению со статическими политиками распределения ЦП, использующими в среднем такое же количество вычислительных ресурсов, и до 30 процентов средней экономии ЦП с минимальные потери QoS по сравнению с планировщиками, не использующими ЦП. При ограничении вычислительных ресурсов, доступные для двух конкурирующих eNB, ИИ обеспечивает на 25 процентов большую пропускную способность при почти нулевом уровне ошибок декодирования[20].

### ***1.3. Вероятностная модель динамической нарезки***

Для описания вероятностных моделей динамической нарезки возьмём следующий пример. Для представления системной модели представим, что у оператора есть базовая станция, передающая сигнал сотовой связи. Зона покрытия данного сигнала имеет форму в виде единичной соты (Рис.1.2). Передающая сигнал станция находится непосредственно в центре данной соты. В зоне действия сигнала

от указанной станции имеется некоторое количество мобильных устройств. В течение времени местоположение каждого устройства может быть изменено. Каждое из них имеет свой MSISDN (Mobile Subscriber Integrated Services Digital Number) – номер мобильного абонента цифровой сети с интеграцией служб.

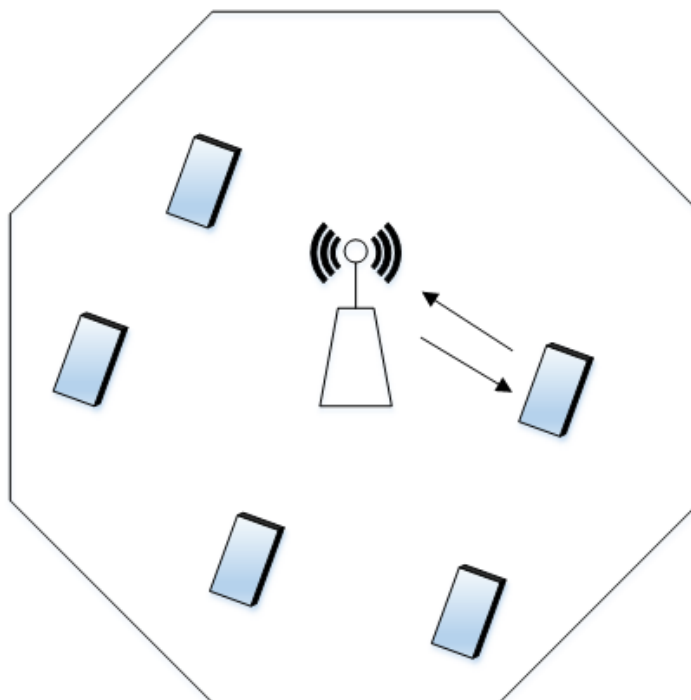


Рис.1.2. Схема модели

В зависимости от того, какое приложение на устройстве запрашивает услуги и инициирует обмен данными, т.е. учитывая систематизацию абонентов и их требования происходит определение типов сетевого трафика. Например, в услуге по скачиванию файлов трафик самого скачивания будет идти по восходящей линии, в то время как предоставление услуги от базовой станции, логично, по нисходящей. Если взять к примеру услугу по скачиванию файлов, то само скачивание ведётся по восходящей линии, а предоставление услуги, логично, по нисходящей. Поэтому, в рамках услуг трафик с разных линий, в дальнейшем, следует рассматривать отдельно [21].

Поведение пользователей, для описанной в предыдущем абзаце системы, может разниться, а, следовательно, существует множество вариантов сценариев. Для наибольшего охвата представления данных сценариев опишем две математические модели, и соответствующие им модели трафика. Первая модель имеет в своей основе следующий сценарий взаимодействия пользователя и системы:

1. Пользователь отправляет запрос на услугу со своего устройства.

2. В ответ на запрос выделяется ресурс на обслуживание.
3. Получив ресурс, пользователь со своим устройством некоторое время занимают его, получая услугу.
4. Закончив использование услуги, устройство пропадает из системы. (данное событие может быть связано с уходом из зоны передачи сигнала или изменением типа услуги, а также с завершением обслуживания данного абонента в целом).

Вторая модель представляет сценарии, в которых количество пользователей в соте неизменно вне зависимости от момента времени, в который происходит фиксирование состояния системы. Данное количество будет обозначено как  $M$ . Вторая модель имеет в своей основе следующий сценарий взаимодействия пользователя и системы:

1. Пользователь отправляет запрос на услугу со своего устройства.
2. В ответ на запрос выделяется ресурс на обслуживание.
3. Получив ресурс, пользователь со своим устройством некоторое время занимают его, получая услугу.
4. После окончания обслуживания, пользователь некоторое время находится в ожидании.
5. Затем пользователь снова использует этот ресурс и деле сценарий будет повторен с шага 3.

Принимая во внимание описанные выше системные модели для определённых сценариев поведения пользователей, построим их математическое представление. Каждая из моделей представляется нам в виде систем массового обслуживания с эластичным трафиком и нетерпеливыми заявками. Для упрощения дальнейшего обозначения математического представления различных сценариев дадим им наименования Система I и Система II [22].

Чтобы имитировать переменное число пользователей представим входящий поток поступающих требований на услугу в виде пуассоновского потока первого рода. Для учёта иных параметров, необходимых для передачи данных, для входного потока следует определить подходящую интенсивность, минимальное требование к ресурсу и длину передаваемого блока данных (Табл.1.2). Для выведения формул и представления математической модели используются данные источника [24].

СМО для Системы I представлена следующей графической моделью (Рис.1.3):

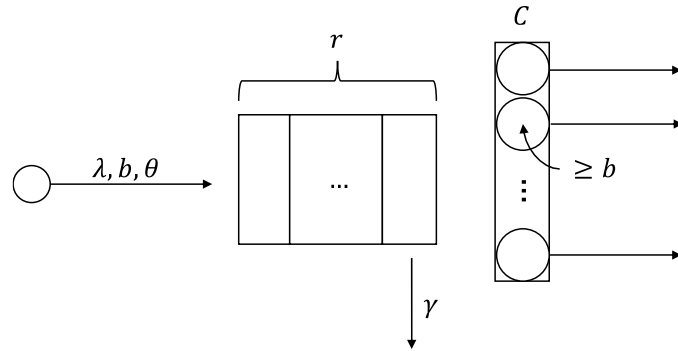


Рис.1.3. Схема модели с переменным числом пользователей (Система I)

Табл.1.2. Параметры модели

$N = N(t), t > 0$	число требований в системе в момент времени $t$
$\lambda$	интенсивность потока поступления требований ( $\lambda \sim Pois$ )
$b \in \{1, \dots, C\}$	минимальная гарантированная скорость
$\theta$	среднее значение длины блока данных
$r$	длина очереди
$\gamma$	интенсивность потери нетерпеливых заявок
$C$	ширина полосы пропускания (общий объем ресурса)
$\lceil \frac{C}{b} \rceil = N$	максимальное число требований, которые могут быть обслужены ресурсом одновременно

Представленные ниже формулы описывают распределение вероятностей для СМО (Система I):

$$p_n = \begin{cases} \left(\frac{\lambda\theta}{C}\right)^n p_0, & 0 < n < N + 1, \\ \left(\frac{\theta}{C}\right)^N \frac{\lambda^n}{\prod_{i=1}^{n-N} \left(\frac{C}{\theta} + i\gamma\right)} p_0, & N < n \leq N + r. \end{cases} \quad (1.1)$$

$$p_0 = \left( \sum_{n=0}^N \left(\frac{\lambda\theta}{C}\right)^n + \left(\frac{\theta}{C}\right)^N \sum_{n=N+1}^{N+r} \frac{\lambda^n}{\prod_{i=1}^{n-N} \left(\frac{C}{\theta} + i\gamma\right)} \right)^{-1}.$$

В рамках математического описания Система II будет отличаться от предыдущей наличием множественных источников поступления заявок. Количество данных источников будем обозначать  $K$ . Все источники могут инициировать только одну заявку, которая поступит в систему. Источник не может создать новую заявку,



пока предыдущая инициированная им находится в очереди или на обслуживании, занимая один из приборов (Табл. 1.3). Когда заявка по истечении некоторого времени завершит своё обслуживание и освободит прибор, а также если она уйдёт из системы по нетерпеливости, тогда источником будет инициирована новая заявка, которая поступит в систему.

СМО для Системы II представлена следующей графической моделью (Рис.1.4):

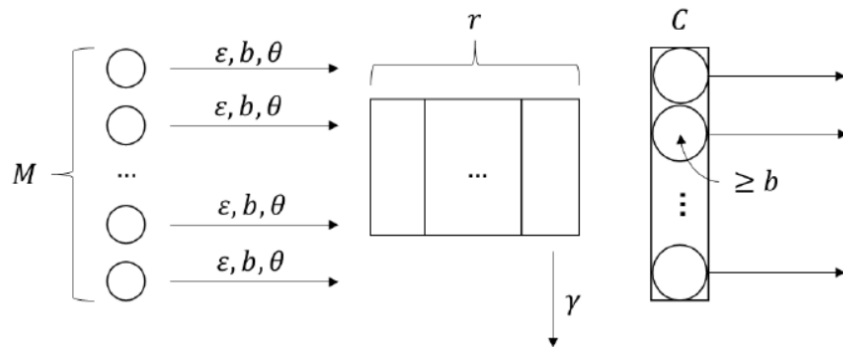


Рис.1.4. Схема модели с переменным числом пользователей (Система II)

Табл. 1.3. Параметры модели

$N = N(t), t > 0$	число требований в системе в момент времени $t$
$M$	фиксированное число источников
$\varepsilon$	интенсивность потока поступления требований ( $\varepsilon \sim Pois$ )
$b \in \{1, \dots, C\}$	минимальная гарантированная скорость
$\theta$	среднее значение длины блока данных
$r$	длина очереди
$\gamma$	интенсивность потери нетерпеливых заявок
$C$	ширина полосы пропускания (общий объем ресурса)
$\lceil \frac{C}{b} \rceil = N$	максимальное число требований, которые могут быть обслужены ресурсом одновременно

При описании Системы II должны быть рассмотрены три частных случая:

1.  $M \geq N + r$ , в текущем состоянии системы не все источники задействованы и число заявок, которые могут одновременно находиться в системе, меньше, чем может предоставить группа источников.

Принимая во внимание данное условие, запишем распределение вероятностей. Оно будет представлено в следующем виде:

$$p_n = \begin{cases} \left(\frac{\lambda\theta}{C}\right)^n p_0, & 0 < n < N + 1, \\ \left(\frac{\theta}{C}\right)^N \frac{\lambda^n}{\prod_{i=1}^{n-N} \left(\frac{C}{\theta} + i\gamma\right)} p_0, & N < n \leq N + r. \end{cases} \quad (1.2)$$

$$p_0 = \left( \sum_{n=0}^N \left(\frac{\lambda\theta}{C}\right)^n + \left(\frac{\theta}{C}\right)^N \sum_{n=N+1}^{N+r} \frac{\lambda^n}{\prod_{i=1}^{n-N} \left(\frac{C}{\theta} + i\gamma\right)} \right)^{-1}.$$

2.  $N \leq M < N + r$ , в текущем состоянии системы все источники задействованы и число заявок, которые могут одновременно находиться в системе, больше, чем может предоставить группа источников. При этом обслуживающих приборов, готовых принять заявки от источников – меньше чем сами источники.

Принимая во внимание второе условие, запишем распределение вероятностей для описанного случая. Оно будет представлено в следующем виде:

$$p_n = p_n = \begin{cases} \left(\frac{\varepsilon\theta}{C}\right)^n A_m^n p_0, & n = \overline{1, N}, \\ \left(\frac{\theta}{C}\right)^N \frac{\varepsilon^n}{\prod_{i=1}^{n-N} \left(\frac{C}{\theta} + i\gamma\right)} A_m^n p_0, & n = \overline{N + 1, N + r}. \end{cases} \quad (1.3)$$

$$p_0 = \left( \sum_{n=0}^N \left(\frac{\varepsilon\theta}{C}\right)^n A_m^n + \left(\frac{\theta}{C}\right)^N \sum_{n=N+1}^{N+r} \frac{\varepsilon^n}{\prod_{i=1}^{n-N} \left(\frac{C}{\theta} + i\gamma\right)} A_m^n \right)^{-1}.$$

3.  $M < N$ , в текущем состоянии системы все источники задействованы и число требований, которые могут одновременно обслуживаться на приборах, больше, чем может предоставить группа источников. Принимая во внимание третье условие, запишем распределение вероятностей для описанного случая. Оно будет представлено в следующем виде:

$$p_n = \frac{\left(\frac{\varepsilon\theta}{C}\right)^n M!}{(M - n)!} p_0, n = \overline{1, M} \quad (1.4)$$

$$p_0 = \left( \sum_{n=0}^N \left(\frac{\varepsilon\theta}{C}\right)^n A_m^n \right)^{-1}.$$

Представим теперь, что в данной СМО будет несколько систем, каждая из которых принадлежит своему виртуальному оператору (VNO). Обозначим множество таких операторов как  $\mathcal{K} = \{1, 2, \dots, K\}$ . Мощностью данного множества можно описать общее число виртуальных операторов, которое представлено как  $K = |\mathcal{K}|$ . Параметры всех виртуальных операторов идентичны, за исключением интенсивности поступления заявок.

Рассмотрим представленные параметры каждого виртуального оператора более подробно. Заметим, что для каждого  $k$ -того виртуального оператора, причём  $k \in \mathcal{K}$ , должен быть задан индивидуальный входящий поток. Определим для каждого виртуального оператора пуассоновский входящий поток первого рода с индивидуальной интенсивностью потока  $\lambda_k$ . Для каждой услуги имеется минимальная гарантированная скорость обслуживания заявки  $b$  (в дальнейшем для упрощения моделируемой системы будем считать её равной 1), которая задаётся вместе с входящим потоком, а также некоторый объём поступающих заявок, обозначенный как  $s$ . Обозначим как  $\gamma$  – интенсивность ухода заявок из очереди из-за свойства нетерпеливости, а  $r$  – максимальное количество заявок в очереди, т.е. её вместимость.

Принимая во внимание условие, указанное выше, минимальная гарантированная скорость обслуживания будет приравнена к единице. Тогда максимальное число заявок  $N$ , которое обслуживающий ресурс может принять на обслуживание, будет равняться количеству единиц самого ресурса, в противном случае, рассмотренном в разделе 2.3 настоящей работы, несколько единиц ресурса могут быть заняты одной или несколькими заявками. Если рассчитывать параметр максимального числа требований, которые могут быть обслужены ресурсом одновременно по следующей формуле  $N = \left\lceil \frac{c}{b} \right\rceil = \left\lceil \frac{c}{1} \right\rceil$ , то получается, что искомое количество одновременно обрабатываемых заявок будет равняться  $C$ . Примем обозначение  $N$  в последующем тексте работы, учитывая указанное в формуле равенство  $N$  и  $C$  для данного условия.

Обозначим первый момент времени, в который происходит вызов нарезки, за  $t_0 = 0$ , остальные же рассчитываются по формуле:

$$t_{i+1} = t_i + \Delta, i = 0, 1, \dots; \Delta = \text{const} \quad (1.5)$$

На (Рис. 2.1) представлена схема для двух виртуальных операторов нарезки ресурсов по времени. Каждый раз по истечении интервала времени  $\Delta$  будет задействован алгоритм перераспределения ресурсов, однако полностью выполняться он будет не при каждом вызове. Для работы алгоритма должно быть выполнено одно из ряда условий, описанный в разделе 2.1 настоящей работы. Количество вызовов алгоритма перенарезки ресурса в момент времени  $t$  будем записывать как  $L(t)$ .

Обозначим как  $N_k(t)$  число некоторых условных единиц ресурса для  $k$ -го виртуального оператора, которое было выделено ему в момент времени  $t$ . Данный параметр является изменяемым с течением времени, поскольку после каждой  $l$ -той перенарезки ресурса он будет заново определён. Опишем его в виде следующей системы:

$$N_k(t) \begin{cases} N_k^0, t \in [t_0, t_1), \\ N_k^1, t \in [t_1, t_2), \\ \vdots \\ N_k^l, t \in [t_l, t_{l-1}), \\ \vdots \end{cases}, k \in \mathcal{K}. \quad (1.6)$$

Для начального распределения ресурсов всем виртуальным операторам, обслуживающим пользователей в данной сети, по некоторому принципу выделяется некоторое количество всех доступных  $N$  ресурсов сети -  $N_k(t_0)$ . Причём так, чтобы выделенные операторам ресурсы в сумме не превышали ресурс имеющейся в сети у базового оператора  $N - \sum_{k=0}^K N_k(t_0) \leq N$ . Требуемое количество ресурсов для каждого виртуального оператора, которое распределено изначально, отражено в соглашении об уровне обслуживания SLA (Service Level Agreement). Запишем начальное распределение ресурсов в виде вектора -  $(N_1^{SLA}, N_2^{SLA}, \dots, N_K^{SLA}) = \overline{N^{SLA}}$ .

Для решения указанной выше задачи оптимальности динамической нарезки, параметры которой зависят от нагрузки на уже распределённые между виртуальными операторами ресурсы, необходимо разработать некоторый алгоритм, следуя которому система будет получать указания к следующим шагам. Для понимания, что заранее установленная последовательность действий систему будет работать верно, необходимо определить условия, удовлетворив которые данная последовательность будет оптимальна. Опишем далее указанные выше требования:

1. Стремление к первоначальной нарезке. Первоначальное распределение ресурсов, закреплённое в SLA, считается тем состоянием, к которому наша

система, и сами виртуальные операторы должны стремиться. Оно считается верным, поскольку оговорено между базовым и виртуальными операторами.

2. Минимальная загрузка системы служебными сообщениями. Каждое инициирование перераспределения ресурсов приносит с собой системное сообщение для администраторов сети. Однако, слишком объёмный поток таких сообщений загружает не только систему на отправку данных сообщений, но и устройство оператора или администратора. Следовательно, стоит стараться их минимизировать.
3. Необходимо исключить «простой» оборудования. Простой оборудования может случиться из-за излишне выданных ресурсов одному из виртуальных операторов. Предложение его услуги превышает ожидаемый спрос от пользователей. При этом, у другого VNO, наоборот, ощущается нехватка ресурсов для обслуживания. Спрос на услугу превышает возможное предложение. Рассмотрение нескольких VNO и системы в целом и даст возможность оптимизировать весь ресурс, находящийся у базового оператора, а не один сегмент сети.

Последние два условия можно считать взаимоисключающими, поскольку одно можно выразить через другое, но для описания мы постараемся учесть все возможные параметры оптимизации системы. При условии, что мы перестанем проводить перенарезку ресурсов, у нас будут минимизированы системные сообщения. Однако без проверки и алгоритма для действий системы нельзя с уверенностью утверждать, что весь ресурс базового оператора в руках различных VNO используется оптимально. Для лучшего понимания работы системы следует провести некоторые эксперименты. Проследим за работой системы в течение некоторого времени, чтобы математически описать и высчитать необходимые характеристики. (Рис. 1.5).

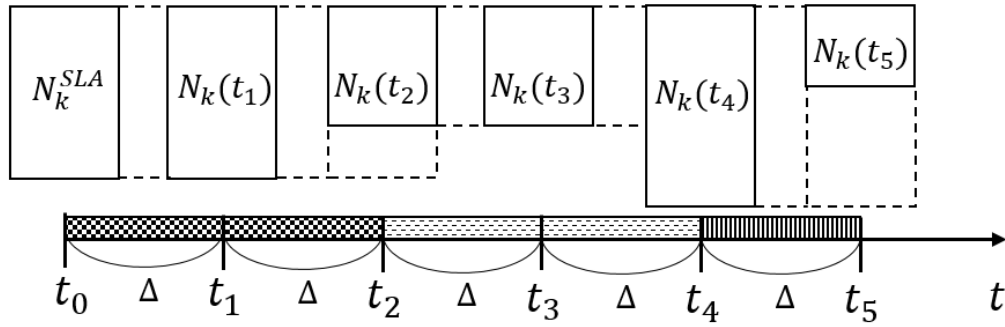


Рис. 1.5. Иллюстрирование работы нарезки ресурса для  $k$ -го оператора по времени

Отразим описанные ранее условия с помощью численных показателей и формул их нахождения:

- а)  $\alpha \in [0,1]$  – среднее значение среди параметров соответствия SLA.

Для расчёта используем следующую формулу:

$$\alpha = \frac{1}{K} \sum_{k \in K} a_k \quad (1.7)$$

где,  $\alpha_k \in [0,1]$  – коэффициент, характеризующий соответствие SLA для  $k$ -ого виртуального оператора – вычисляется следующим образом:

$$a_k = \lim_{T \rightarrow \infty} \frac{\sum_{l=1}^{L(T)} ((t_l - t_{l-1}) \cdot \min(N_k(t_l), N_k^{SLA}))}{T \cdot N_k^{SLA}} \quad (1.8)$$

- б)  $\beta \in [0,1]$  - доля полезных вызовов нарезки. Для учёта данного параметра потребуется формула полной вероятности, где количество благоприятных исходов делится на количество всех исходов. Для числителя требуется посчитать сумму из таких моментов, что вызванная нарезка меняла значение ресурсов  $N_k(t)$  для какого-нибудь VNO из множества  $K$ . И для получения необходимого значения разделим данное число на количество всех вызовов нарезки.

$$\beta = \lim_{T \rightarrow \infty} \frac{\sum_{l=1}^{L(T)} \max_{k \in K} \{N_k(t_l) \neq N_k(t_{l-1})\}}{L(T) + 1} \quad (1.9)$$

- с)  $\overline{UTIL} \in [0,1]$  - среднее значение по коэффициентам занятости ресурсов виртуальных операторов.

$$\overline{UTIL} = \frac{1}{K} \sum_{k \in K} UTIL_k \quad (1.10)$$

где,  $UTIL_k \in [0,1]$  - коэффициент занятости ресурса k-ого виртуального оператора – определяется формулой

$$UTIL_k = \lim_{T \rightarrow \infty} \frac{\sum_{i=1}^{I(T)} (t_i - \tau_i)}{T} \quad (1.11)$$

где  $\tau_i$  и  $t_i$  – начало и конец интервала занятости ресурса.

#### 1.4. Постановка задачи исследования

Задачей настоящей выпускной работы является изучение существующих моделей и принципов Network Slising. Исследование характеристик оптимизации и различных подходов её осуществления. Описание основных параметров моделей для большего удобства последующего моделирования. Проведение численного анализа для поиска оптимального временного интервала нарезки ресурсов с последующей иллюстрацией его результатов. Анализ системы с управлением перераспределением ресурсов между двумя операторами с учетом временного интервала при разном соотношении нагрузки первого и второго класса заявок. Анализ влияния интервала нарезки на показатели эффективности базового и виртуальных операторов для модели из двух сегментов с эластичным трафиком

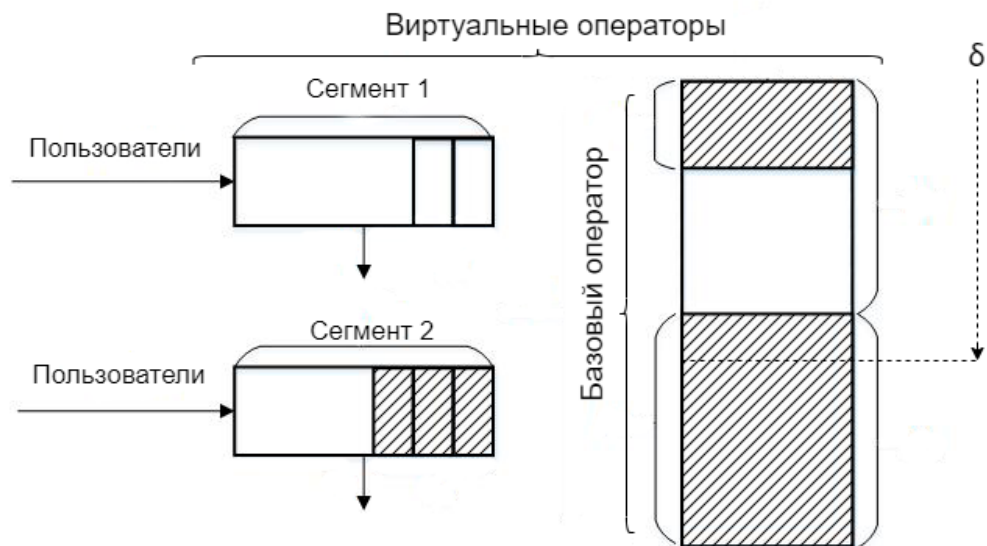


Рис. 1.6. Графическое представление системной модели

## Глава 2. Модель управляемой контроллером нарезки радиоресурсов

### 2.1. Принципы нарезки

Представим, что СМО описанная в разделе 2.1.3. выше, имеет частный случай с двумя виртуальными операторами.

Следующей схема (Рис.2.1) даёт нам наглядное представление системы:

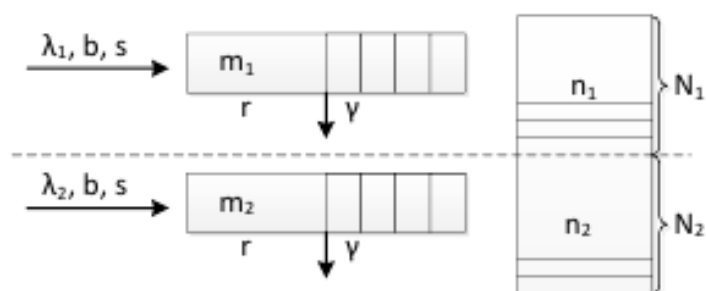


Рис.2.1. Схема системы с 2 виртуальными операторами в момент времени  $t$

Табл. 2.1. Характеризующие систему в момент перераспределения ресурсов параметры

Параметр	Описание
$K$	Количество виртуальных операторов, в данном случае 2.
$N$	Максимальное число заявок, обрабатываемых прибором одновременно
$t_i$	Момент времени, в который происходит перераспределение ресурса.
$N_k(t_i),$ $\sum_{k=0}^K N_k(t_i) = N$	Максимальное число позиций на приборе, которые были отданы для обслуживания заявок $k$ -ого виртуального оператора в момент времени $t_i$ ,
$N_k^{SLA}$	Изначально выделенное количество мест на приборе заявки $k$ -ого виртуального оператора (в момент времени $t_0$ .)
$n_k$	Количество заявок на приборах $k$ -го виртуального оператора
$m_k$	Количество заявок в очереди $k$ -го виртуального оператора
$k \in \{1,2\}$	Множество виртуальных операторов



Опишем, как в системе используются параметры, указанные в таблице выше (Табл. 2.1).

Представим что потоки заявок на услугу, а так же ресурсные блоки для каждого VNO распределены таким образом, что  $\lambda_1 < \lambda_2 < \dots < \lambda_k$ .

Для обозначения количества заявок, которое находится на обслуживании ресурса у каждого  $k$ -того оператора в определённый момент времени  $t$ , введём вектор с обозначением описанных выше параметров  $(n_1(t), n_2(t), \dots, n_K(t)) = \vec{n}(t)$ . Для описания информации о том, сколько заявок на услугу находится в очереди у каждого виртуального оператора в момент времени  $t$  сформируем вектор со следующими обозначениями  $(m_1(t), m_2(t), \dots, m_K(t)) = \vec{m}(t)$ . Каждый из этих векторов необходим, для формирования состояния системы в каждый отдельно взятый момент времени  $t$ .

При группировке описанных выше векторов  $(\vec{n}(t), \vec{m}(t))$  можно описать полученные в процессе работы системы значения каждого виртуального оператора, которые будут использованы в алгоритме:

$$\left( (n_1(t), m_1(t)), (n_2(t), m_2(t)), \dots, (n_K(t), m_K(t)) \right).$$

Представим пространство состояний в системе в следующем виде:

$$X_l := \prod_{k \in \mathcal{K}} \left\{ \begin{array}{l} (n_k(t), m_k(t)) : n_k(t) \geq 0, m_k \geq 0, \\ n_k(t) = 0, 1, \dots, N_k(t) - 1, m_k(t) = 0, \\ n_k(t) = N_k(t), m_k(t) = 0, 1, \dots, r \end{array} \right\}, t \in [t_l, t_{l+1}), l = 0, 1, \dots \quad (2.1)$$

Далее, рассмотрим возможные действия по изменению состояния системы при поступлении сигналов от контроллера о перенарезке ресурсов.

Рассмотрим возможные сценарии работы системы при иницировании данного события:

1. При имеющемся свободном ресурсе у одного виртуального оператора, и недостатке данного у другого (ресурс обслуживания переполнен, очередь частично или полностью занята), незанятый ресурс первого VNO будет перераспределён нуждающемуся в ресурсах второму VNO.
2. При активной занятости ресурсов у обоих виртуальных операторов (ресурс обслуживания переполнен, очередь частично или полностью занята),

перенарезка, а следовательно и изменение состояния системы происходить не будет.

Зададим SLA для рассматриваемого частного случая системы, где множество виртуальных операторов представлено как  $K = \{1,2\}$ , в момент времени  $t_0$ :

$$N_1(t_0 = 0) = N_2(t_0 = 0) = N_1^{SLA} = N_2^{SLA} = \begin{bmatrix} C \\ 2 \end{bmatrix} = \begin{bmatrix} N \\ 2 \end{bmatrix} \quad (2.2)$$

Каждый раз после вызова  $l$ -ой нарезки сетевых ресурсов базового оператора в момент времени  $t_l$  для каждого состояния системы в установленный момент времени будет образовываться вектор описания ресурсов, имеющих в распоряжении каждого VNO:  $\vec{N}(t_l) = (N_1(t_l), N_2(t_l))$ .

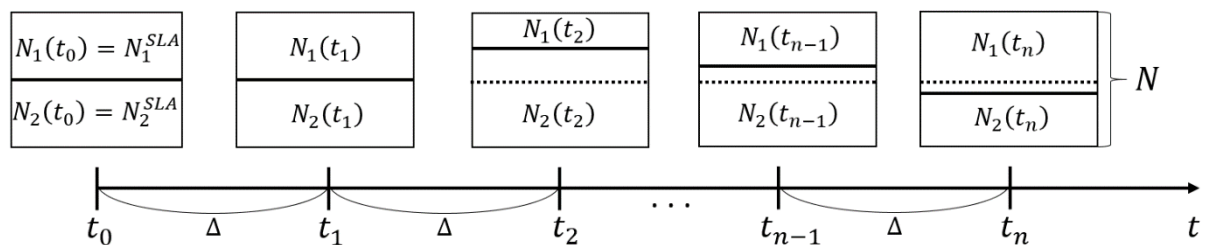


Рис. 2.1. Схема нарезки ресурсов во времени с интервалом нарезки  $\Delta$

В данном конкретном случае для двух виртуальных операторов состояние системы будет состоять из двух пар показателей работы каждого виртуального оператора -  $((n_1, m_1), (n_2, m_2))$ , на основе которых алгоритмом будет приниматься решение об изменении выделенных каждому виртуальному оператору ресурсов. Путём аналитических рассуждений было установлено, что существует всего три условных различных состояния, которые может застать алгоритм при вызове перенарезки ресурсов в момент времени  $t_l$ :

1.  $m_1 = 0, m_2 = 0$ , т.е. очереди всех операторов пусты.

Если алгоритмом будет зафиксировано такое состояние, распределение ресурсов обоих виртуальных операторов остаётся без изменений:  $\vec{N}(t_l) = \vec{N}(t_{l-1})$

2. Очередь из заявок на услугу хотя бы одного VNO не пуста.

Рассмотрим различные вариации данного условия, комбинируя разные значения занятости каждой из очередей виртуальных операторов:

- a.  $m_1 > 0, m_2 = 0$  – первая очередь не пуста, а, следовательно, все предоставленные первому виртуальному оператору ресурсы заняты.

б.  $m_1 = 0, m_2 > 0$  – вторая очередь не пуста, а, следовательно, все предоставленные второму виртуальному оператору ресурсы заняты.

Действия алгоритма для каждого из указанных случаев будут рассмотрены ниже.

3.  $m_1 > 0, m_2 > 0$  – в обеих очередях виртуальных операторов находятся заявки, распределённый ресурс используется наиболее полно.

В данном случае считаем, что при отсутствии простаивающего ресурса у каждого виртуального оператора перенарезка ресурсов инициирована не будет:  $\vec{N}(t_l) = \vec{N}(t_{l-1})$ .

Далее, опишем получившиеся состояния системы в условиях 2а и 2б более подробно. В каждом конкретном условии возможны различные варианты изменения состояния системы, в зависимости от параметров, задающих предыдущее её состояние для анализа алгоритмом. При более подробном рассмотрении случая из пункта 2а, оказывается, что при различных показателях занятости ресурсов каждого из операторов, а также количества заявок в первой очереди возможны две ортогональные ситуации. Для того чтобы более понятно описать каждую из них представим на схеме (Рис. 2.2) состояния системы уже после перераспределения ресурсов в момент времени  $t_l$ , и предыдущее состояние в момент времени  $t_{l-1}$ :

А.  $m_1 \leq N_2(t_{l-1}) - n_2$ , данное условие описывает, что количество свободных единиц ресурса второго виртуального оператора имеется в достаточном количестве, чтобы принять все ожидающие заявки в очереди заявок первого VNO.

Б.  $m_1 > N_2(t_{l-1}) - n_2$ , данное условие описывает, что количество свободных единиц ресурса второго виртуального оператора недостаточно, чтобы принять все ожидающие заявки в очереди заявок первого VNO.

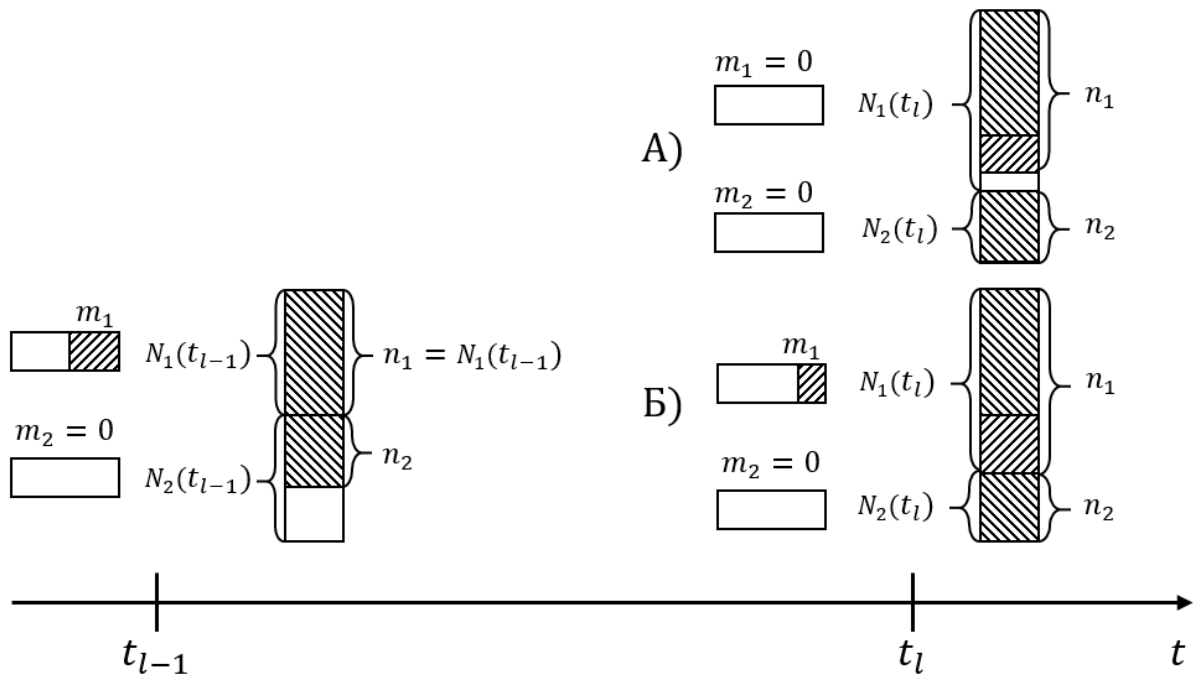


Рис. 2.2. Схема изменения нарезки для состояния 2а

Аналогичные условия запишем и для состояния системы из пункта 2б. Как (2.3) представлены функции вычисления нового распределения ресурса для указанных в условии состояний.

$$N_1(t_l) = \begin{cases} N_1(t_{l-1}) + m_1, & m_1 \leq N_2(t_{l-1}) - n_2 \\ N_1(t_{l-1}) + N_2(t_{l-1}) - n_2, & m_1 > N_2(t_{l-1}) - n_2 \end{cases} \quad (2.3)$$

$$N_2(t_l) = \begin{cases} N_2(t_{l-1}) - m_1, & m_1 \leq N_2(t_{l-1}) - n_2 \\ n_2, & m_1 > N_2(t_{l-1}) - n_2 \end{cases}$$

## 2.2. Системная модель

Описанные выше принципы перераспределения ресурсов применимы и к данной системе, за исключением различия пространства состояний и учёта имеющейся у виртуальных операторов очередей. Рассмотрим систему управлением перераспределением ресурсов между различными услугами с учетом временного интервала, в продолжении работы [23]. В системе имеется базовый оператор, имеющий в своём распоряжении физическую структуру сети и ресурс, который имеет определённую скорость обслуживания пользователей. А также имеются два виртуальных оператора сотовой связи, требующие доступа к использованию ресурсов базового оператора. Виртуальные операторы предоставляют пользователям два типа услуг с минимальной и стабильной передачей данных. Задержка во время

предоставления услуг для пользователей является динамической и может изменяться во время процесса предоставления услуг благодаря процессу перераспределения радиоресурсов между двумя виртуальными сегментами сети. Схематическое изображение системы представлено на Рис. 1.6.

В начальный момент времени, согласно соглашению об уровне обслуживания (SLA), имеется равное распределение ресурсов между двумя сегментами сети. В зависимости от загрузки системы во время фиксированного момента времени состояние системы будет проверяться и определять потребность перераспределения ресурсов между сегментами. В результате основная идея состоит в том, чтобы оптимально распределить потребности пользователей между двумя сегментами сети методом принятия решений по распределению ресурсов в определенные интервалы.

Решение данной задачи позволяет обеспечить эффективное предоставление сетевых услуг. Изменение интервала нарезки также приводит к изменениям в системе, т.к. уменьшение значения позволяет увеличить степень загрузки ресурсов. Тем не менее, перераспределение фактических радиоресурсов осуществляется только в том случае, если он позволяет увеличить пропускную способность сети, и это не может произойти во всех моментах времени проверки системы. Следовательно, система может быть перегружена сигнальными сообщениями, что приводит к затратам и услугам времени доставки для пользователей.

**Задача оптимизация** решается в отношении следующих параметров:

Необходимо дать рекомендации по поводу значения частоты поступления сервисных сообщений от контроллера. При этом, предоставленное значение должно удовлетворять целям работы контроллера в системе.

Основное требования к базовому оператору - оптимальное «справедливое» деление ресурса между виртуальными операторами.

1. Имеющиеся у базового оператора ресурсы не должны простаивать.
2. Справедливое деление ресурсов между виртуальными операторами, для удовлетворения их требований.
3. Сообщения контроллера должны оптимизировать систему. Если после прихода сообщения в системе не произошло перераспределения ресурсов

между виртуальными операторами, сервисное сообщение считается непродуктивным.

Для Виртуального оператора важны следующие показатели системы:

1. Вероятность блокировки запроса
2. Среднее число обслуживаемых пользователей
3. Среднее число пользователей в очереди

### 2.3. Система массового обслуживания с сигналами

Опишем параметры системы более подробно (Табл. 2.2). Пусть входящие потоки запросов на передачу каждого класса услуг эластичного трафика пуассоновские с интенсивностями  $\lambda$ , а объем трафика распределен по экспоненциальному закону с параметрами  $s_k$ , с учетом порога  $b$  для скорости передачи трафика максимальное число обслуживаемых сессий составляет  $N$ , а число мест в очереди для ожидающих начала обслуживания сессий  $R_k$ . Пусть пороги для времени ожидания начала обслуживания сессий распределены по экспоненциальному закону с параметрами  $\epsilon_k$ , где  $k = 1, 2$ . Поток сигналов, управляющих перераспределением ресурса, является пуассоновским с интенсивностью  $\delta$ . Схема модели представлена на Рис. 2.2.

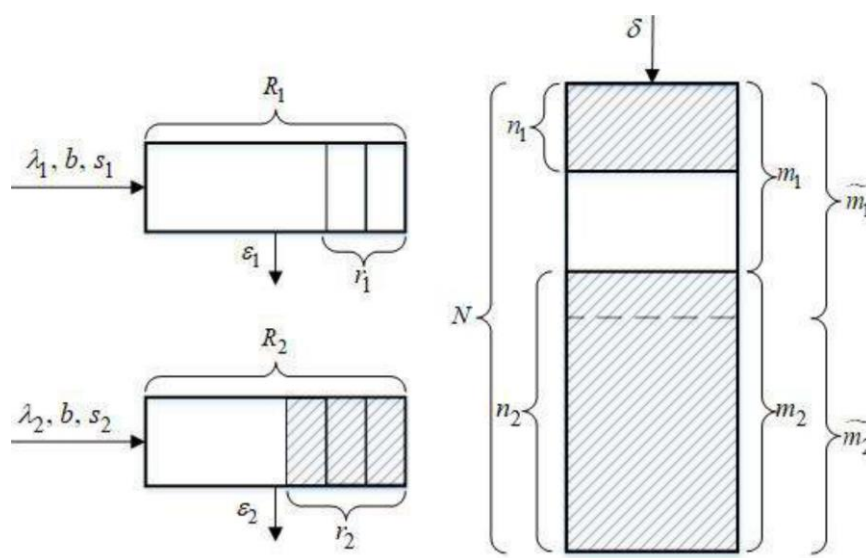


Рис. 2.2. Графическое представление математической модели

Табл. 2.2. Описание параметров математической модели

Обозначение	Описание
$\lambda_k$	Интенсивность входящего потока запросов
$s_k$	Средний объем трафика
$b$	Скорость передачи трафика
$N$	Максимально число обслуживаемых сессий
$R_k$	Число мест в очереди для обслуживания сессий
$\varepsilon_k$	Порог времени ожидания обслуживания
$\delta$	Интенсивность управляющих сигналов

Функционирование системы описывает случайный процесс  $X(t)$  с состояниями вида  $x = (m_1, m_2, n_1, n_2, r_1, r_2)$ , где  $m_k$  - порог на максимальное число обслуживаемых  $k$ -сессий,  $n_k$  число обслуживаемых  $k$ -сессий,  $r_k$  число ожидающих начала обслуживания  $k$ -сессий. Таким образом, пространство состояний  $X(t)$  имеет вид (2.4, 2.5).

$$\begin{aligned}
 X = & \{(m_1, N - m_1, n_1, n_2, 0, 0): 0 \leq m_1 \leq N, 0 \leq n_1 \leq m_1, 0 \leq n_2 \leq m_2\} \cup \\
 & \cup \{(N - m_2, m_2, N - m_2, n_2, r_1, 0): 0 \leq m_2 \leq N, 0 \leq n_2 \leq m_2, 0 < r_1 \leq R_1\} \cup \\
 & \cup \{(m_1, N - m_1, n_1, N - m_1, 0, r_2): 0 \leq m_1 \leq N, 0 \leq n_1 \leq m_1, 0 < r_2 \leq R_2\} \cup \\
 & \cup \{(m_1, N - m_1, m_1, N - m_1, r_1, r_2): 0 \leq m_1 \leq N, 0 < r_1 \leq R_1, 0 < r_2 \leq R_2\}
 \end{aligned} \quad (2.4)$$

$$\begin{aligned}
 x' = (m'_1, m'_2, n'_1, n'_2, r'_1, r'_2) \succ (m_1, m_2, n_1, n_2, r_1, r_2) = x \Leftrightarrow \\
 n'_1 + n'_2 + r'_1 + r'_2 > n_1 + n_2 + r_1 + r_2 \\
 n'_1 + n'_2 + r'_1 + r'_2 = n_1 + n_2 + r_1 + r_2, \quad m'_1 > m_1 \\
 n'_1 + n'_2 + r'_1 + r'_2 = n_1 + n_2 + r_1 + r_2, \quad m'_1 = m_1, \quad n'_1 + r'_1 > n_1 + r_1
 \end{aligned} \quad (2.5)$$

В Табл. 2.3 перечислены возможные интенсивности переходов между состоянием  $x$  и другими состояниями системы. Политика управления перераспределением ресурсов определена в строках 4. Исходя из описанных правил записывается матрица интенсивностей переходов и могут быть найдены стационарные вероятности  $\pi(x)$ ,  $x \in X$ .

Табл. 2.3. Интенсивности переходов

№ п/п	Интенсивность события	Условие на $x$	Состояние $x'$
1а-1	$\lambda_1$	$n_1 + 1 \leq m_1$	$(m_1, N - m_1, n_1 + 1, n_2, 0, r_2)$
1б-1	$\lambda_1$	$n_1 + 1 > m_1, r_1 + 1 \leq R_1$	$(m_1, N - m_1, m_1, n_2, r_1 + 1, r_2)$
1а-2	$\lambda_2$	$n_2 + 1 \leq m_2$	$(N - m_2, m_2, n_1, n_2 + 1, r_1, 0)$
1б-2	$\lambda_2$	$n_2 + 1 > m_2, r_2 + 1 \leq R_2$	$(N - m_2, m_2, n_1, m_2, r_1, r_2 + 1)$
2а-1	$\frac{m_1}{N} V \mu_1$	$r_1 > 0$	$(m_1, N - m_1, m_1, n_2, r_1 - 1, r_2)$
2б-1	$\frac{m_1}{N} V \mu_1$	$r_1 = 0, n_1 > 0$	$(m_1, N - m_1, n_1 - 1, n_2, 0, r_2)$
2а-2	$\frac{m_2}{N} V \mu_2$	$r_2 > 0$	$(N - m_2, m_2, n_1, m_2, r_1, r_2 - 1)$
2б-2	$\frac{m_2}{N} V \mu_2$	$r_2 = 0, n_2 > 0$	$(N - m_2, m_2, n_1, n_2 - 1, r_1, 0)$
3а-1	$r_1 \varepsilon_1$	$r_1 > 0$	$(m_1, N - m_1, m_1, n_2, r_1 - 1, r_2)$
3а-2	$r_2 \varepsilon_2$	$r_2 > 0$	$(N - m_2, m_2, n_1, m_2, r_1, r_2 - 1)$
4а-1	$\delta$	$n_1 = m_1, r_1 > 0,$ $n_2 < m_2, r_2 = 0,$ $r_1 \leq m_2 - n_2$	$(m_1 + r_1, m_2 - r_1, m_1 + r_1, n_2, 0, 0)$
4б-1	$\delta$	$n_1 = m_1, r_1 > 0,$ $n_2 < m_2, r_2 = 0,$ $r_1 > m_2 - n_2$	$\left( \begin{array}{l} m_1 + m_2 - n_2, n_2, \\ m_1 + m_2 - n_2, n_2, \\ r_1 - m_2 + n_2, 0 \end{array} \right)$
4а-2	$\delta$	$n_2 = m_2, r_2 > 0,$ $n_1 < m_1, r_1 = 0,$ $r_2 \leq m_1 - n_1$	$(m_1 - r_2, m_2 + r_2, n_1, m_2 + r_2, 0, 0)$



46-2	$\delta$	$n_2 = m_2, r_2 > 0,$ $n_1 < m_1, r_1 = 0,$ $r_2 > m_1 - n_1$	$\left( \begin{array}{l} n_1, m_2 + m_1 - n_1, \\ n_1, m_2 + m_1 - n_1, \\ 0, r_2 - m_1 + n_1 \end{array} \right)$
------	----------	---	--

## 2.4. Показатели эффективности перераспределения ресурса

Отметим, что параметрами эффективности перераспределения ресурса для базового оператора, соответствующими критериям оценки эффективности перераспределения ресурса между классами трафика, будут  $\beta$  — коэффициент успеха перераспределения ресурса (2.6), и  $\gamma$  — коэффициент использования ресурса (2.7).

$$\beta = \sum_{x \in B} \pi(m_1, m_2, n_1, n_2, r_1, r_2) \quad (2.6)$$

$$B = \{x \in X: n_1 = m_1, n_2 < m_2, r_1 > 0, r_2 = 0\} \cup \\ \cup \{x \in X: n_2 = m_2, n_1 < m_1, r_2 > 0, r_1 = 0\}$$

$$\gamma = \frac{\gamma_1 + \gamma_2}{2}, \gamma_k = \sum_{x \in X: m_k > 0} \frac{n_k}{m_k} \pi(m_1, m_2, n_1, n_2, r_1, r_2) \quad (2.7)$$

Определим показатели эффективности для виртуального оператора.

Вероятность блокировки запроса

$$B_1 = \{(m_1, m_2, n_1, n_2, r_1, r_2) \in X: n_1 = m_1, r_1 = R_1\} \\ B_2 = \{(m_1, m_2, n_1, n_2, r_1, r_2) \in X: n_2 = m_2, r_2 = R_2\}$$

$$B_k = \sum_{x \in B_k} \pi(m_1, m_2, n_1, n_2, r_1, r_2) \quad (2.8)$$

Вероятность ухода пользователя по нетерпеливости — это означает что пользователь находится в очереди, и уйдёт из системы, так и не дождавшись начала обслуживания. При этом, не сам пользователь уходит из-за недостаточной скорости обслуживания. Специальный обработчик, по прошествии определённого времени инициирует события вывода заявки из очереди.

$$I_1 = \{(m_1, m_2, n_1, n_2, r_1, r_2) \in X: r_1 > 0\}$$

$$I_2 = \{(m_1, m_2, n_1, n_2, r_1, r_2) \in X: r_2 > 0\}$$

$$P(A) = \sum_i P(A|H_i)P(H_i)$$

$$P(A) = \sum_i P(A|H_i)P(H_i) \quad (2.9)$$

$P(H_i)$  - вероятность того что система будет находится в том состоянии, в котором есть очередь. Т.е. пользователь находится в системе, которая находится в состоянии  $\vec{x}$ . Пользователь ожидает в очереди.

$$P(A) = \sum_{\vec{x} \in X} \pi(\vec{x})P(A|H_i) = \sum_{\vec{x} \in I_k} \pi(\vec{x})P(A|H_i) \quad (2.10)$$

$P(A|H_i)$  – это условная вероятность того, что пользователь находится в очереди и не дождётся начала обслуживания =  $P(A, B, C|H_i)$ . При этом должно быть инициировано выбивание заявки из очереди с интенсивностью  $\varepsilon_k$  ( $B$ ) и именно заявка пользователя должна быть выбрана из остальных заявок, находящихся в очереди ( $C$ ).

$P(A, B|H_i) =$   
 $\frac{r_k \varepsilon_k - \text{событие вывода будет иницировано с интенсивностью } \varepsilon_k \text{ для одной из } r_k}{\lambda_1 1\{n_1+1 \leq m_1\} + \lambda_1 1\{n_1+1 > m_1, r_1+1 \leq R_1\} + \dots}$  -  
 вероятность того, что будет инициировано событие выведение заявки.

$P(C) = \frac{1}{r_1}$  – вероятность выбора пользователя из очереди.

$$P(A) = \sum_{\vec{x} \in I_k} \pi(\vec{x}) \frac{r_k \varepsilon_k}{\text{все интенсивности } r_k} \frac{1}{r_k} \quad (2.11)$$

Среднее число обслуживаемых пользователей

$$N_k = \sum_{x \in X} n_k \pi \quad (m_1, m_2, n_1, n_2, r_1, r_2) \quad (2.12)$$

Среднее число пользователей в очереди

$$N_k^r = \sum_{x \in X} r_k \pi \quad (m_1, m_2, n_1, n_2, r_1, r_2) \quad (2.13)$$

Таким образом, необходимо оптимально распределять запросы пользователей между двумя виртуальными операторами путем принятия решений по распределению ресурсов с определенным интервалом времени при разном соотношении нагрузки первого и второго классов заявок.

## Глава 3. Численный анализ показателей эффективности нарезки радиоресурсов

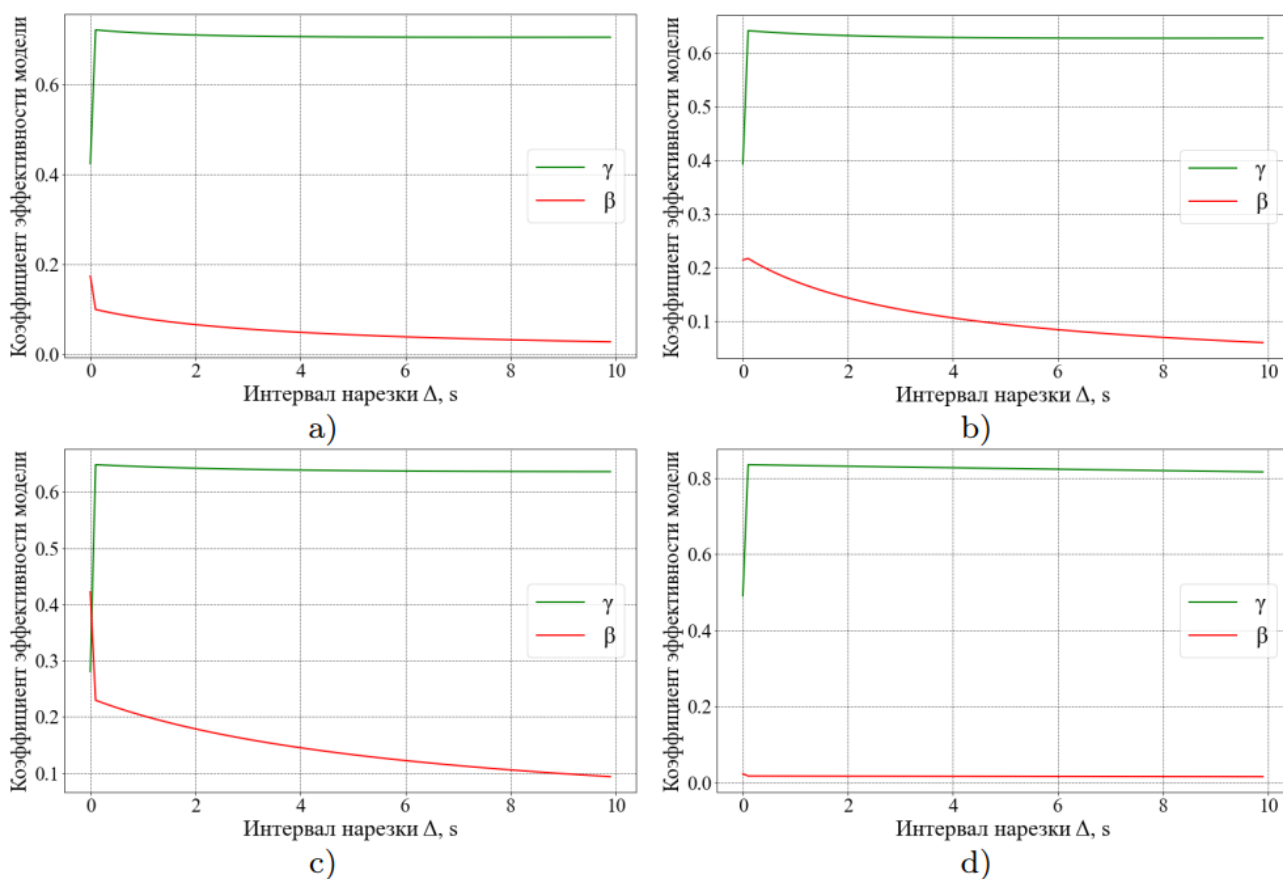
### 3.1. Численный пример

Перейдем к численному анализу построенной системы. Для понятия работы системы и того, правильно ли была получена политика оптимизации, был проведён следующий эксперимент. Для работы системы были задействованы следующие исходные данные. Более подробные иллюстрации приведены в виде графиков показателей эффективности по временным интервалам сервисных сообщений контроллера (Рис. 3.1).

Для проведения численного анализа была написана программа на языке Python. Используя (2.5), получим пространство состояний системы. Матрица переходов построена исходя из описанных в статье [23] правил. При этом отдельно рассчитываются коэффициент успеха перераспределения ресурса  $\beta$  (2.6) и коэффициент использования ресурса  $\gamma$  (2.7). Исходные данные приведены в Табл. 3.1. Интервал нарезки  $\delta$  от 0 до 10 с шагом 0.1.

Табл. 3.1. Тестовые данные

Сценарий	C	b	$\lambda_1$	$\lambda_2$	$\mu_1$	$\mu_2$	$\epsilon_1$	$\epsilon_2$	$R_1$	$R_2$
a	2	1	4	3	1	1	1	1	1	1
B	2	1	3	4	3	2	1	1	1	1
C	2	1	7	5	4	4	2	2	1	1
d	2	1	100	150	83	111	1	1	1	1



**Рис. 3.1. Результаты моделирования тестовых исходных данных для модели управляемой контроллером перенарезки ресурсов**

Проиллюстрируем зависимость коэффициентов эффективности перераспределения ресурса от среднего значения интервала между поступлениями сигналов. На рис.2 показана зависимость коэффициентов эффективности перераспределения ресурса от среднего значения  $\delta$  интервала между поступлениями сигналов. Чем чаще вызывается перераспределение ресурсов, тем более полно используются предоставленные для операторов ресурсы, на что указывает коэффициент  $\gamma$ . Тем не менее, при этом менее полезным для работы системы становится каждый его вызов. Если после инициирования перераспределения в системе не происходит никаких изменений, поскольку состояние её уже оптимально, то данный вызов становится непродуктивным, что и демонстрирует нам график коэффициента  $\beta$ .

### 3.2. Сценарий просмотра просмотра веб-страниц и групповой передачи данных

Для проведения численного анализа на реальных данных было принято взять информацию по услугам просмотра веб-страниц и групповой передачи данных. Данные были получены из стандартов, предоставленных ITU и 3GPP [2,3].

Данные по услугам представлены в Табл. 3.2:

Табл. 3.2. Данные по услугам

Услуга	Степень симметрии (degree of symmetry)	Стандартное значение объёма данных	Порог на время задержки	Потеря данных
Групповая передача данных	Primarily one-way	10 Кбайт - 10 Мбайт	Рекомендуемое – 15 с Допустимое – 60 с	Zero
Просмотр веб-страниц	Primarily one-way	10 Кбайт	Рекомендуемое – 2 с Допустимое – 4 с	Zero

Опишем характеристики участников процесса эксплуатации сети:

1. Характеристики ресурса со стороны базового оператора
  - a. Пропускная способность (МГц)
  - b. Метод кодирования MIMO
  - c. Схема модуляции

Из данных параметров получаем минимальную скорость передачи данных в бит/с. У базового оператора имеется контролер для перераспределения ресурсов. Характеристики ресурса со стороны контроллера:

- d. Частота поступления сервисных сообщений со стороны коннектора в систему ( $\delta$ )
2. Характеристики ресурса со стороны Виртуального оператора
    - a. Предоставляемая оператором услуга характеризуется следующими параметрами:
      - i. Средний объём передаваемых данных по услуге
      - ii. Порог на реальное время обработки запроса услуги.

Для каждой услуги есть время максимальной задержки пакетов на передачу данных. Обозначим для одной услуги данное время как  $T_1^1$ . Тогда, скорость передачи трафика будет определяться как отношение объёма передаваемых данных к времени задержки  $b_1 = \frac{\mu_1^{-1}}{T_1^1}$ .

в. Количество мест в системе для постановки заявки в ожидание.

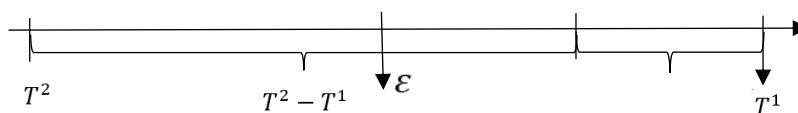
### 3. Характеристики ресурса со стороны пользователя

а. Частота обращения пользователя к услуге ( $\lambda_1$ )

в. Порог на допустимое время задержки

Будем считать, что пользователи готовы ожидать некоторый период времени между приходом в момент времени  $T^1$  и окончанием обслуживания  $T^2$ . Пользователи, после отправки запроса, ждут обработки своего запроса, но при этом, если время ожидания превышает их лимит, после истечения данного времени, они отзывают запрос.

Следовательно, время ожидания определяет период, находящийся между данными моментами  $\varepsilon^{-1} = T^2 - T^1$



Численные показатели описанных выше данных составлены на основе услуг, заявленных в стандартах ITU и 3GPP [2,3] и приведены в Табл. 3.3. Результаты проиллюстрированы на Рис. 3.2. и Рис. 3.3.

Табл. 3.3. Данные численного анализа для услуг просмотра веб-страниц и групповой передачи данных

Название	Значение услуги Bulk data transfer/retrieval	Значение услуги Web-browsing
<b>Базовый оператор</b>		
V	3 МГц, QPSK, Without MIMO(SISO), <b>2.344 Мбит/с</b>	
$\delta$	0.00001	
<b>Виртуальный оператор</b>		
$\mu_k^{-1}$ – объём данных	1 Мбайт = 8 Мбит	136,58 Кбайт = 1,067 Мбит

	2 Мбайт = 16 Мбит	273,02 Кбайт = 2,133 Мбит
$T_1^k, T_2^k$ - Порог на время задержки	Рекомендуемое – 15 с	Рекомендуемое – 2 с
$b_k = \frac{\mu_k^{-1}}{T_k^1}$ – скорость передачи данных	0,533 Мбит/с	
	1,067 Мбит/с	
$R_k$ – количество мест в очереди	30	15
<b>Пользователь</b>		
$\lambda_k$ – интенсивность поступления запросов	0,03	0,6
$(T^2 - T^1)$ - порог на допустимое время задержки в очереди	Допустимое – 60 с 60-15=45с	Допустимое – 4 с
$\varepsilon = \frac{1}{T^2 - T^1}$ – интенсивность ухода по нетерпеливости	0,01с <sup>-1</sup>	0,25 с <sup>-1</sup>

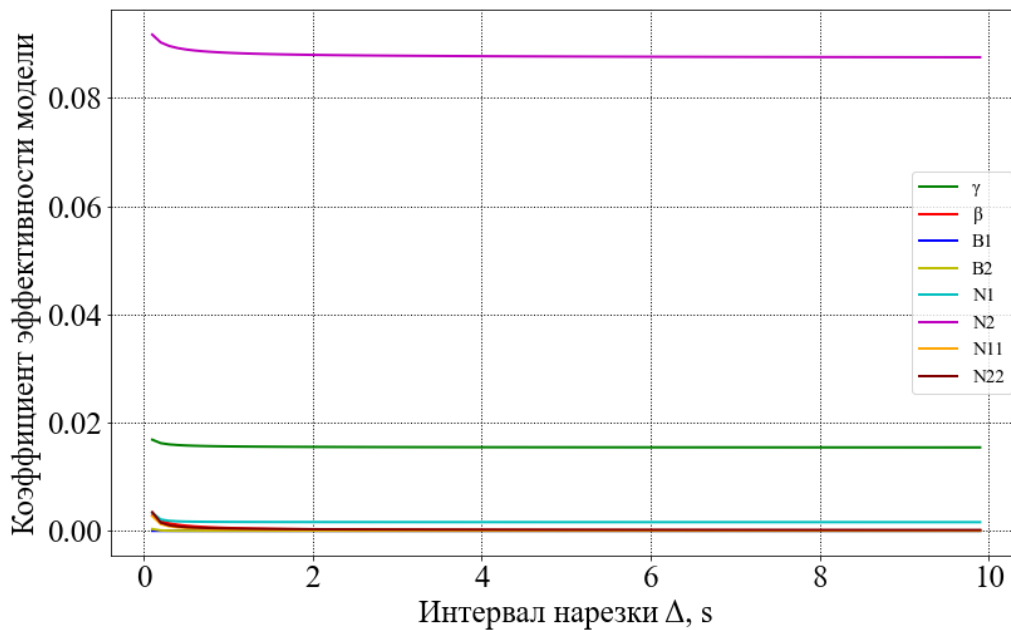


Рис. 3.2. Показатели эффективности системы для скорости передачи данных 0,533 Мбит/с



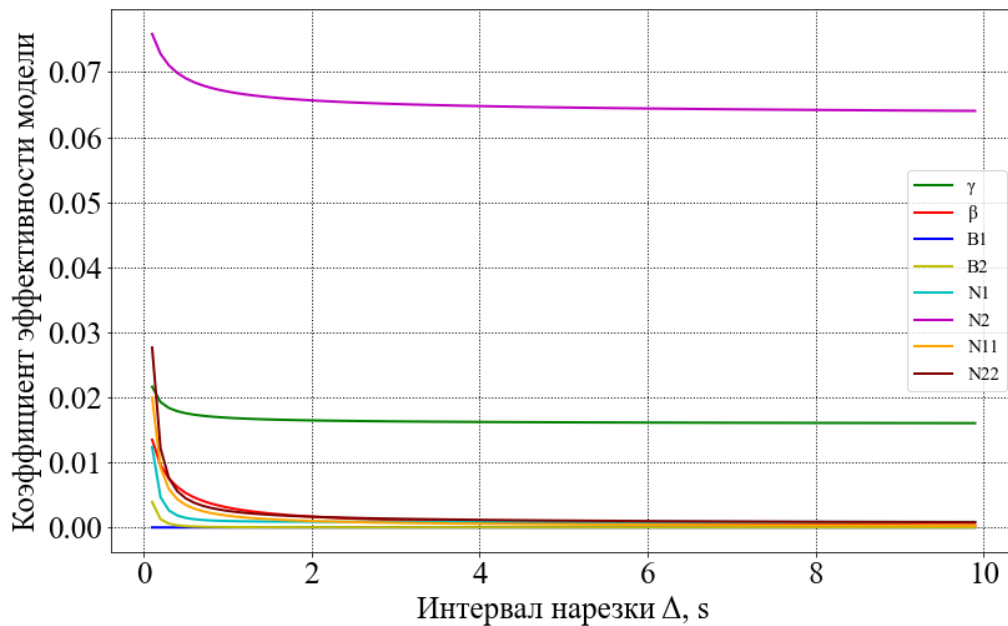


Рис. 3.3. Показатели эффективности системы для скорости передачи данных 1,067 Мбит/с

Заметим, что текущие данные основаны на предположении интенсивностей потока заявок на услугу, а также определённом установленном числу мест в очереди. Изменим данные (Табл. 3.4) для того, чтобы подобрать оптимальные интенсивности поступления заявок. Дополним представленные выше зависимости от интервала нарезки графиками зависимости от интенсивности поступления заявок (Рис. 3.4, Рис. 3.5).

Табл. 3.4. Изменённые данные численного анализа для услуг просмотра веб-страниц и групповой передачи данных

Название	Значение услуги Bulk data transfer/retrieval	Значение услуги Web-browsing
<b>Базовый оператор</b>		
V	3 МГц, QPSK, Without MIMO(SISO), <b>2.344 Мбит/с</b>	
$\delta$	0.000001	
<b>Виртуальный оператор</b>		
$\mu_k^{-1}$ – объём данных	3 Мбайт = 24 Мбит	409,6 Кбайт = 3,2 Мбит
$T_1^k, T_2^k$ - Порог на время задержки	Рекомендуемое – 15 с	Рекомендуемое – 2 с
$b_k = \frac{\mu_k^{-1}}{T_k^1}$ – скорость передачи данных	1,6 Мбит/с	
$R_k$ – количество мест в очереди	50	2

Пользователь		
$\lambda_k$ – интенсивность поступления запросов	0,35	0,6
$(T^2 - T^1)$ - порог на допустимое время задержки в очереди	Допустимое – 60 с 60-15=45с	Допустимое – 4 с
$\varepsilon = \frac{1}{T^2 - T^1}$ – интенсивность ухода по нетерпеливости	0,01с <sup>-1</sup>	0,25 с <sup>-1</sup>

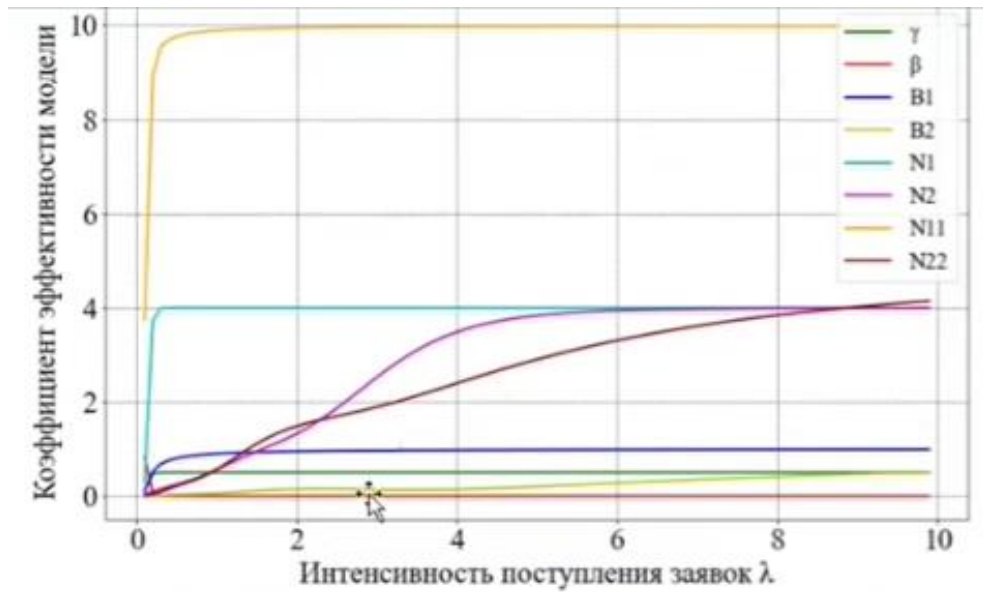


Рис. 3.4. Значения показателей эффективности в зависимости от интенсивности

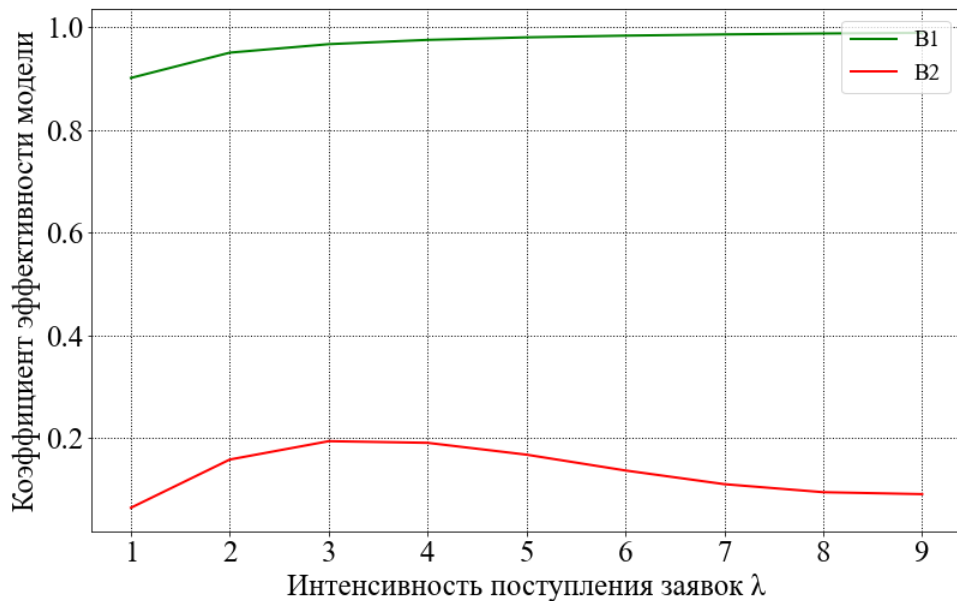


Рис. 3.5. Значения блокировок поступления заявок в зависимости от интенсивности

## **Заключение**

1. В ходе исследования было выявлено несколько концепций построения сетевой сегментов.
2. В ходе исследования были определены функции задания входящего потока, связанные с управлением и согласованием срезов сети.
3. Проанализирована двухсервисная сетевая модель, выявлены качественные характеристики.
4. Проанализирован метод динамической нарезки ресурса между виртуальными операторами.
5. Выведены показатели эффективности нарезки.
6. Описана имитационная модель для системы.

В рассмотренных в настоящей работе методах оптимизация разделения ресурсов сетей идёт по разным характеристикам. Поэтому, применение данных методов зависит от целей сетевого оператора по оптимизации. Более подробный разбор моделей и добавления к ним других показателей, а также вариация исходных данных может стать задачей дальнейших исследований.

## Литература

1. *Вся статистика интернета на 2021 год — цифры и тренды в мире и в России [Электронный ресурс]. – Режим доступа: <https://www.web-canape.ru/business/vsya-statistika-interneta-i-socsetej-na-2021-god-cifry-i-trendy-v-mire-i-v-rossii/> – Заглавие с экрана. – (Дата обращения: 04.05.2022).*
2. *Rec. 3GPP TS 23.203 V17.2.0 (2021-12). 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Policy and charging control architecture (Release 17). France, 2021*
3. *Rec. ITU-T Rec. G.1010 (06/2018). ITU-T SERIES E: OVERALL NETWORK OPERATION, TELEPHONE SERVICE, SERVICE OPERATION AND HUMAN FACTORS. Quality of telecommunication services: concepts, models, objectives and dependability planning – Terms and definitions related to the quality of telecommunication services.*
4. *Ordonez-Lucena, J., Ameigeiras, P., Lopez, D., Ramos-Munoz, J.J., Lorca, J., Folgueira, J. Network Slicing for 5G with SDN/NFV: Concepts, Architectures, and Challenges (2017) IEEE Communications Magazine, 55 (5), art. no. 7926921, pp. 80-87.*
5. *ITU-T Y.3112, «TELECOMMUNICATION STANDARDIZATION SECTOR OF ITU (12/2018), SERIES Y: GLOBAL INFORMATION INFRASTRUCTURE, INTERNET PROTOCOL ASPECTS, NEXT-GENERATION NETWORKS, INTERNET OF THINGS AND SMART CITIES»*
6. *3GPP «TR 28.801 V15.1.0 (2018-01)» - 3GPP, 2018*
7. *M. Pearce, S. Zeadally, R. Hunt, «Virtualization: Issues, Security Threats, and Solutions» - ACM Computer Surveys, vol. 45, no. 2, Feb. 2013, pp. 1–38.*
8. *N. M. M. K. Chowdhury, R. Boutaba, «A Survey of Network Virtualization» - Computer Networks, vol. 54, no. 5, Apr. 2010, 862–76*
9. *ONF TR-521, “SDN Architecture,” Feb. 2016.*
10. *Bega D., Gramaglia M., Garcia-Saavedra A., Fiore M., Banchs A., Costa-Perez X (2020). Network Slicing Meets Artificial Intelligence: An AI-Based Framework for Slice Management. IEEE Communications Magazine, 58(6), 32-38. <https://doi.org/10.1109/MCOM.001.1900653>*

11. 3GPP, “Telecommunication Management; Study on Management and Orchestration of Network Slicing for Next Generation Network,” Tech. Rep. (TR) 28.801, 2018; <http://www.3gpp.org/DynaReport/28801.htm>.
12. I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning, MIT Press, 2016; <http://www.deeplearningbook.org>
13. J. X. Salvat et al., “Overbooking Network Slices Through Yield-driven End-to-end Orchestration,” Proc. ACM CoNEXT, Dec. 2018, pp. 353–65.
14. J. Mendes et al., “Cellular Access Multi-Tenancy Through Small-Cell Virtualization and Common Rf Front-End Sharing,” Computer Commun., vol. 133, Jan. 2019, pp. 59–66.
15. C. Zhang, P. Patras, and H. Haddadi, “Deep Learning in Mobile and Wireless Networking: A Survey,” IEEE Commun. Surveys & Tutorials, vol. 21, no. 3, Mar. 2019, pp. 2224–87.
16. D. Bega et al., “DeepCog: Cognitive Network Management in Sliced 5G Networks with Deep Learning,” Proc. IEEE INFOCOM, May 2019, pp. 280–88.
17. J. A. Ayala-Romero et al., “vrAIn: A Deep Learning Approach Tailoring Computing and Radio Resources in Virtualized RANs,” Proc. ACM MobiCom, Oct. 2019, pp. 1–16.
18. T. P. Lillicrap et al., “Continuous Control With Deep Reinforcement Learning,” arXiv preprint arXiv:1509.02971, 2015.
19. I. Gomez-Miguel et al., “srsLTE: An Open-source Platform for LTE Evolution and Experimentation,” Proc. ACM WiNTECH, Oct. 2016, pp. 25–32.
20. Vincenzo Sciancalepore, Lanfranco Zanzi, Xavier Costa-Perez, Antonio Capone «ONETS: Online Network Slice Broker From Theory to Practice» – IEEE, 2018 г.
21. Ana Margarida Pina Simões Temporal Modelling of Mobile Data Traffic Applications for Network Optimisation // Thesis to obtain the Master of Science Degree in Electrical and Computer Engineering – May – 2017.
22. Makeeva E.D., Харин П.А., Поляков Н.А., Маркова Е.В., Гудкова И.А., Галинина О. С. Анализ установления соединения пар передатчик-приемник на mmWave // Всероссийская конференция с международным участием «Информационно-телекоммуникационные технологии и математическое

моделирование высокотехнологичных систем» ИТТММ-2018 (16–20 апреля 2018 г., Москва): Материалы. – М.: РУДН. – 2018. – С.134-136.

23. *Kochetkova I., Vlaskina A., N. N. Vva and V. S. Shorgin. Queuing System with Signals for Dynamic Resource Allocation for Analyzing Network Slicing in 5G Networks // Inform. Primen., 2021, Volume 15, Issue 3, Pages 91–97. doi:10.14357/19922264210312.*
24. *Башарин Г.П. Лекции по математической теории телетрафика. – М.: РУДН, 2004.*

**Information and  
Telecommunication  
Technologies and  
Mathematical Modeling of  
High-Tech Systems 2022  
(ITTMM 2022)**

Monday, April 18, 2022 - Friday, April 22, 2022

3 Ordzhonikidze st.

**Book of Abstracts**

## Анализ показателей эффективности нарезки радиоресурсов в сети 5G в виде системы массового обслуживания с сигналами

**Authors:** Anastasiia Vlaskina<sup>Nonm</sup>; Feroiz Khusainova<sup>Nonm</sup>; Sofiya Burtseva<sup>1</sup>; Veronika Filippova<sup>2</sup>

<sup>1</sup> RUDN university

<sup>2</sup> RUDN University

**Corresponding Author:** sofiyaburtseva@gmail.com

Нарезка радиоресурсов (Network slicing) является одной из главных технологий беспроводных сетей 5G. Концепция позволяет определить несколько виртуальных сетей на одной физической сетевой инфраструктуре, благодаря чему операторы сотовой связи могут разворачивать изолированные друг от друга сети. Нарезка предоставляет возможность разделять одну физическую сеть на множество логически разделённых сегментов сети для удовлетворения требований различных сценариев использования. Одно из главных достоинств разделения ресурсов на сегменты заключается в том, что такой подход требует меньших вложений в физическую инфраструктуру сети. Благодаря широкому покрытию данная технология позволяет расширить возможности эксплуатации сети, то есть станет важной составляющей сетей 5G, позволяющей принимать и обрабатывать данные в соответствии с конкретными требованиями клиентов. В данной работе мы рассмотрим математическую модель для двух классов нетерпеливого эластичного трафика с минимальным порогом скорости в виде системы массового обслуживания с очередями и сигналами, управляющими перераспределением ресурса. Разберем несколько сценариев поведения системы при разных входных данных, а также проведем численный анализ системы. Задача оптимизации численно решается относительно параметра интервала времени  $\delta^{-1}$  при разном соотношении нагрузки первого и второго класса заявок для каждого из приведенных сценариев.

MTT + NGN: Joint section / 357

## Поиск луча в системах связи ТГц диапазона частот в условиях микроподвижности абонентских терминалов

**Author:** Svetlana Dugaeva<sup>1</sup>

**Co-authors:** Дмитрий Молчанов<sup>2</sup>; Вячеслав Бегушев

<sup>1</sup> Peoples' Friendship University of Russia (RUDN University)

<sup>2</sup> РУДН

**Corresponding Author:** 1032182508@rudn.ru

В работе проведено математическое моделирование для оценки эффективности применения датчиков для определения положения беспроводных устройств в пространстве в условиях микроподвижности в сетях терагерцового диапазона частот. В связи отсутствием полных данных по корректности измерений датчиками устройств, исследование было проведено для различных коэффициентов точности. В работе использовались методы математического моделирования, аппарата теории вероятностей, а именно случайное блуждание с непрерывным временем. Основным результатом исследования является получение характеристик среднего времени нахождения беспроводного устройства в состоянии устойчивой связи.

Poster session. Applied information systems / 358

## Методы автоматической коррекции направления движения робототехнических комплексов



# Приложение 2. Программа ЭВМ. Расчет показателей эффективности модели динамической нарезки между двумя сегментами радиоресурсов с управлением по внешнему событию

РОССИЙСКАЯ ФЕДЕРАЦИЯ



RU

**2021661716**

ФЕДЕРАЛЬНАЯ СЛУЖБА  
ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ  
(12) ГОСУДАРСТВЕННАЯ РЕГИСТРАЦИЯ ПРОГРАММЫ ДЛЯ ЭВМ

Номер регистрации (свидетельства): <a href="#">2021661716</a>	Авторы: Ву Ньят Нам (VN), Власкина Анастасия Сергеевна (RU), Кочеткова Ирина Андреевна (RU), Бурцева София Артемовна (RU)
Дата регистрации: 14.07.2021	
Номер и дата поступления заявки: 2021660847 08.07.2021	
Дата публикации: <a href="#">14.07.2021</a>	Правообладатель: федеральное государственное автономное образовательное учреждение высшего образования «Российский университет дружбы народов» (РУДН) (RU)
Контактные реквизиты: titova-ekbr@rudn.ru	

Название программы для ЭВМ:

**Расчет показателей эффективности модели динамической нарезки между двумя сегментами радиоресурсов с управлением по внешнему событию**

## Реферат:

Программа предназначена для расчета показателей эффективности управления радиоресурсами в модели двух сегментов с эластичным трафиком и перераспределением радиоресурсов с некоторым интервалом нарезки. Программа включает модуль ввода данных, модуль вывода данных, модуль, реализующий расчет состояний системы, и модуль, реализующий расчет показателей эффективности динамической нарезки. Программа предусматривает вывод коэффициента использования ресурсов, коэффициента соответствия начальной нарезки и долю успешных нарезок.

**Язык программирования:** Python 3.7

**Объем программы для ЭВМ:** 111 КБ

# Приложение 3. Программа ЭВМ. Расчет показателей эффективности модели управления нарезкой радиоресурсов беспроводной сети между двумя виртуальными операторами по сигналам контроллера

## УВЕДОМЛЕНИЕ-ЗАЯВКА на регистрацию ПРЭВМ (программы для ЭВМ)

Авторы: Кочеткова И. А., Власкина А. С., Бурцева С. А., Филиппова В. С., Хусайнова Ф. Д.

Структурное подразделение: кафедра прикладной информатики и теории вероятностей, факультета физико-математических и естественных наук, РУДН

Контакты: +7 (937) 7088890

Основания для создания ПРЭВМ: ПРЭВМ создана по заданию на разработку сотрудником РУДН

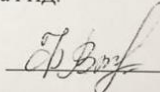
Номер темы НИР (НИОКР): Работа выполнена вне темы НИР

Наименование ПРЭВМ: Расчет показателей эффективности модели управления нарезкой радиоресурсов беспроводной сети между двумя виртуальными операторами по сигналам контроллера


Индивидуализирующие признаки ПРЭВМ: ПРЭВМ обеспечивает распределение вычислений по локальной одноранговой сети.

Приложение:

1. заявление на регистрацию ПРЭВМ
2. согласие автора на указание сведений об авторе;
3. согласие на обработку персональных данных;
4. служебное задание;
5. отчет о выполнении служебного задания;
6. реферат;
7. машиночитаемый носитель;
8. регистрационная карта РИД.

 / Филиппова В. С./

« » 2022 г.

Принято: 

« 16 » 05 2022 г.