Научная статья
УДК 621.391
https://doi.org/10.31854/1813-324X-2025-11-4-18-27
EDN:UOYTAY

Интеллектуальная бессерверная вычислительная система для услуг телеприсутствия

Захраа Ахмед Хуссейн Аль-Кереа, al-kerea.zah@sut.ru
Аммар Салех Али Мутханна ✉, muthanna.asa@sut.ru
Андрей Евгеньевич Кучерявый, akouch@sut.ru

Санкт-Петербургский государственный университет телекоммуникаций им. проф. М. А. Бонч-Бруевича, Санкт-Петербург, 193232, Российская Федерация

Аннотация

Актуальность. *В статье рассматривается проблема совместной оптимизации миграции сервисов и распределения ресурсов (SMRA) в среде периферийных мобильных вычислений с множественным доступом (MEC) для снижения задержки в системах телеприсутствия. MEC расширяет возможности облачных вычислений путем перемещения сервисов на границу сети максимально близко к пользователям, решая проблему задержки доступа. Однако высокая мобильность устройств и ограниченные ресурсы периферийных серверов усложняют поддержание качества обслуживания. Процесс миграции сервисов сам по себе приводит к дополнительной задержке, а различные серверы и пользовательские устройства имеют свои уникальные требования и политику распределения ресурсов, что требует сбалансированного подхода к решению данной задачи. Несмотря на достижения в области телеприсутствия, такие как высококачественные видеофайлы и пространственный звук, виртуальная и дополненная реальность, эффективное функционирование этих систем требует надежной инфраструктуры и минимальных задержек при взаимодействии.*
Постановка задачи. *В данной работе мы предлагаем совместный алгоритм SMRA+MEC, который учитывает специфику систем телеприсутствия, а также решает задачу оптимального распределения ресурсов и необходимости миграции сервисов.*
Цель работы. *Разработка и оценка эффективности совместного алгоритма SMRA+MEC, адаптированного для систем телеприсутствия.*
Методы исследования. *Для достижения поставленной цели в работе будут использованы математические модели для формализации задачи SMRA+MEC с учетом параметров систем телеприсутствия.*
Результаты *исследования показывают, что предложенный алгоритм обеспечивает значительное снижение задержки – на 50 %.*
Научная новизна. *Рассматривается новый способ расчета задержки, который позволяет минимизировать задержку и распределять ресурсы более оптимально. Показано, что комбинирование методов SMRA+MEC является наиболее эффективным подходом к минимизации задержек.*
Практическая значимость*. Разработанный алгоритм SMRA+MEC может быть использован операторами мобильной связи для оптимизации развертывания и управления MEC-инфраструктурой, обеспечивая высококачественное обслуживание для приложений телеприсутствия.*

Ключевые слова: *телеприсутствие, MEC, SMRA, задержка, нагрузка на сеть, распределение ресурсов*

Ссылка для цитирования: Аль-Кереа З.А.Х., Мутханна А.С.А., Кучерявый А.Е. Интеллектуальная бессерверная вычислительная система для услуг телеприсутствия // Труды учебных заведений связи. 2025. Т. 11. № 4. С. 18–27. DOI:10.31854/1813-324X-2025-11-4-18-27. EDN:UOYTAY

Original research
https://doi.org/10.31854/1813-324X-2025-11-4-18-27
EDN:UOYTAY

Intelligent Serverless Computing System for Telepresence Services

Zahraa A.H. Al-Kerea, al-kerea.zah@sut.ru
Ammar S.A. Muthanna ✉, muthanna.asa@sut.ru
Andrey E. Koucheryavy, akouch@sut.ru

The Bonch-Bruevich Saint Petersburg State University of Telecommunications,
St. Petersburg, 193232, Russian Federation

Annotation

Relevance*. This article addresses the problem of Joint Service Migration and Resource Allocation (SMRA) optimization in a Multi-access Edge Computing (MEC) environment, aiming to reduce latency in telepresence systems. MEC enhances cloud computing capabilities by relocating services to the network edge, as close as possible to users, thus resolving access latency issues. However, the high mobility of devices and the limited resources of edge servers complicate the maintenance of Quality of Service. The service migration process itself introduces additional latency, and different servers and user devices have their own unique requirements and resource allocation policies, necessitating a balanced approach to solving this problem. Despite advancements in the field of telepresence, such as high-quality video and spatial audio, virtual reality, and augmented reality, the effective operation of these systems requires a robust infrastructure and minimal interaction delays.*
Problem statement*. In this work, we propose a joint SMRA+MEC algorithm that considers the specific characteristics of telepresence systems and addresses the problem of optimal resource allocation and the necessity of service migration.*
Purpose of the work. *Development and evaluation of the effectiveness of a joint SMRA+MEC algorithm adapted for telepresence systems.*
Methods used. *To achieve the stated goal, mathematical models will be used in this work to formalize the SMRA+MEC problem, taking into account the parameters of telepresence systems.*
The results *show that the proposed algorithm achieves a significant latency reduction of 50 %.*
Scientific novelty. *A novel method for calculating latency is presented, which allows for minimizing latency and allocating resources more optimally. It is shown that combining SMRA+MEC methods is the most effective approach to latency minimization.*
Practical significance. *The developed SMRA+MEC algorithm can be used by mobile operators to optimize the deployment and management of MEC infrastructure, providing high-quality service for telepresence applications.*

Keywords: *telepresence, MEC, SMRA, latency, network load, resource allocation*

For citation: Al-Kerea Z.A.H., Muthanna A.S.A., Koucheryavy A.E. Intelligent Serverless Computing System for Telepresence Services. *Proceedings of Telecommunication Universities.* 2025;11(4):18–27. (in Russ.) DOI:10.31854/1813-324X-2025-11-4-18-27. EDN:UOYTAY

1. Введение

Технология мобильных периферийных вычислений с множественным доступом (MEC, *аббр. от англ.* Multiaccess Edge Computing) обеспечивает приложения Интернета вещей (IoT, *аббр. от англ*. Internet of Things) и системы телеприсутствия услугами, критично важными с точки зрения ресурсоемкости и задержки за счет расширения возможностей облачных вычислений на уровень периферии сети в пределах досягаемости конечных устройств. Высокая мобильность IoT-устройств, например таких, как транспортные средства, вместе с ограниченными ресурсами периферийных серверов приводит к увеличению времени отклика и ухудшению стабильности обслуживания. В связи с этим, для поддержания требуемого уровня качества обслуживания (QoS, *аббр. от англ*. Quality of Service), воз-

никает необходимость в рациональном перераспределении ресурсов и выполнения миграции сервисов. Однако обслуживание процедуры миграции приводит к дополнительным задержкам. Кроме того, на необходимость миграции обслуживания влияют политики распределения ресурсов целевых периферийных серверов. Это связано с тем, что у различных пользователей систем телеприсутствия требования к сетевым и вычислительным ресурсам могут отличаться. Следовательно, вопрос оптимизации процессов миграции обслуживания и распределения ресурсов является важной задачей, требующей тщательной проработки. В этой статье рассматривается проблема совместной оптимизации миграции сервисов обслуживания и распределения ресурсов (SMRA, *аббр. от англ*. Service Migration And Resource Allocation) в условиях MEC для минимизации задержек в системах телеприсутствия. Предложенный в работе комбинированный метод представляет собой совместный алгоритм SMRA и MEC, который принимает в расчет специфические особенности систем телеприсутствия и определяет оптимальное решение по перемещению сервисов, вариантов их миграции и распределению ресурсов. Например, для задач чувствительных к временны́м задержкам время обработки не должно превышать 0,1 мс [1–3], что при использовании удаленного облака является сложно достижимой задачей.

С другой стороны, из-за ограниченных вычислительных ресурсов абонентских устройств обработка таких задач на пользовательских устройствах ведет к большому потреблению энергии и высоким задержкам. Для решения этой проблемы была предложена новая технология – упомянутые выше периферийные вычисления с множественным доступом [3]. Основной концепцией технологии MEC является разгрузка задач, которая подразумевает перенос задач с пользовательских устройств на сервера, размещенные ближе к «краю» сети, в непосредственной близости от пользователей или в удаленное облако. Данный подход позволяет существенно сократить задержки при обработке задач и уменьшить энергозатраты [1–2].

Технология телеприсутствия представляет собой совокупность информационных разработок для проведения сеансов видеоконференцсвязи, которая реализует «ощущение присутствия или воздействия» в месте, отдаленном от реального физического расположения пользователя. Суть этой технологии заключается в объединении технологий передачи видео высокой четкости (UHD, *аббр от англ.* Ultra High Definition), объемного звука и получения обратной связи от пользователя, что нужно для интерактивности, при этом обмен данными должен происходить в режиме реального времени. Кроме того, в эту систему могут входить технологии виртуальной (VR, *аббр. от англ.* Virtual Reality) и дополненной реальности (AR, *аббр. от англ.* Augmented Reality) – например, 3D-голограммы и устройства для имитации тактильных контактов. Концепция услуг телеприсутствия известна довольно давно, однако ее практическая реализация стала возможной только в последние годы благодаря развитию телекоммуникационных систем и информационных технологий, поддерживающих высокую пропускную способность передачи данных на гораздо более высоких частотах и низкую задержку. Однако внедрение и эксплуатация подобных систем (например, технологий 5G / 6G) являются весьма затратными, особенно для малого бизнеса и образовательных учреждений, что является одной из причин, тормозящих широкое распространение систем телеприсутствия [3].

Технология MEC стремительно развивается и предназначена для повышения эффективности передачи данных и обработки информации в современных сетях. В ходе ее реализации используется децентрализованный подход, который подразумевает размещение вычислительных ресурсов и хранилищ ближе к конечным пользователям, что в целом снижает время задержки и повышает QoS. Таким образом, в отличие от традиционных облачных вычислений, которые зависят от удаленных центров обработки данных, технология MEC распределяет ресурсы на периферии сети, обеспечивая быстрый доступ к данным. Это особенно актуально для приложений, требовательных к минимальным задержкам, таких, например, как беспилотные автомобили и IoT, а также системы телеприсутствия.

Технология MEC также помогает разгрузить ядро сети и центральные узлы, перераспределяя задачи обработки информации на периферийные устройства, что улучшает пропускную способность и управление трафиком. Тем не менее, MEC не всегда может обеспечить минимальную задержку при передаче данных, поскольку высока вероятность того, что периферийные узлы могут оказаться перегруженными. Следовательно, технология SMRA является важной и служит для оптимизации размещения / использования услуг и ресурсов. Основной целью SMRA является достижение эффективного использования коммуникационных и информационных систем и, как следствие, снижение издержек и затрат. Технология включает в себя ряд многоэтапных задач, причем эффективное распределение ресурсов является ключевым аспектом SMRA [3]. В зависимости от результатов анализа потребностей и фактического наличия ресурсов, разрабатываются стратегии их оптимального использования, перераспределения или увеличения.

Различия между технологиями SMRA и MEC заключаются в их основной направленности и функциональности. Ниже рассмотрим ключевые различия в принципах работы SMRA и MEC.

Во-первых, технология SMRA направлена на оптимизацию процессов миграции сервисов между различными узлами сети и эффективное распределение ресурсов. Главная задача SMRA заключается в том, чтобы обеспечить максимальную эффективность использования доступных ресурсов, снизить задержки и таким образом повысить QoS конечного пользователя. Как упоминалось выше, технология MEC представляет собой архитектуру вычислений «на краю» сети, где данные обрабатываются ближе к источнику их формирования (пользователям), что позволяет значительно сократить время задержки и разгрузить центральные серверы [3–4]. MEC предоставляет инфраструктуру и программные платформы для размещения разнообразных приложений и сервисов непосредственно на границе сети, что может быть использовано операторами связи для предоставления новых услуг и приложений с низкими задержками. Именно поэтому технология MEC широко используется в рамках концепций IoT, интеллектуальных транспортных систем (ITS, *аббр. от англ.* Intelligent Transportation System), AR / VR), беспилотных автомобилей и других приложений, требующих быстрой обработки данных и низких задержек.

Во-вторых, MEC является более глобальной и занимается созданием архитектурных основ для обработки данных «на краю» сети.

В-третьих, SMRA стремится распределять задачи и ресурсы с точки зрения их текущей доступности и загруженности. MEC, в свою очередь, фокусируется на смещении вычислений ближе к точкам формирования данных [4–5].

В-четвертых, в SMRA акцент делается на управлении и распределении ресурсов между различными узлами сети, чтобы обеспечить оптимальную производительность, а в MEC ресурсы предоставляются на базе узлов сети (*от англ.* Edge Nodes), чтобы минимизировать время передачи данных и повысить скорость отклика.

Несмотря на эти различия, взаимодействие между SMRA и MEC возможно и, зачастую, необходимо для создания высокоэффективных и масштабируемых систем, способных удовлетворить потребности современных приложений и сервисов [6].

Рассмотрим сценарий, схематично представленный на рисунке 1, при котором пользователь $u1$ подключается к периферийному узлу $e1$ в момент времени t и запрашивает услугу У2 на периферийном узла $e1$. Мы предполагаем, что пользователь $u1$ переместится в окрестности периферийного узла $e3$ в момент времени $t + 1$. Если пользователь продолжит подключаться к исходному периферийному узлу $e1$, это приведет к длительной задержке связи между мобильным пользователем $u1$ и исходным периферийным узлом $e1$. В тоже время мы можем выполнить миграцию службы, чтобы перенести службу У2 с исходного периферийного узла $e1$ на целевой периферийный узел $e3$, чтобы сократить расстояние связи и уменьшить задержку связи.

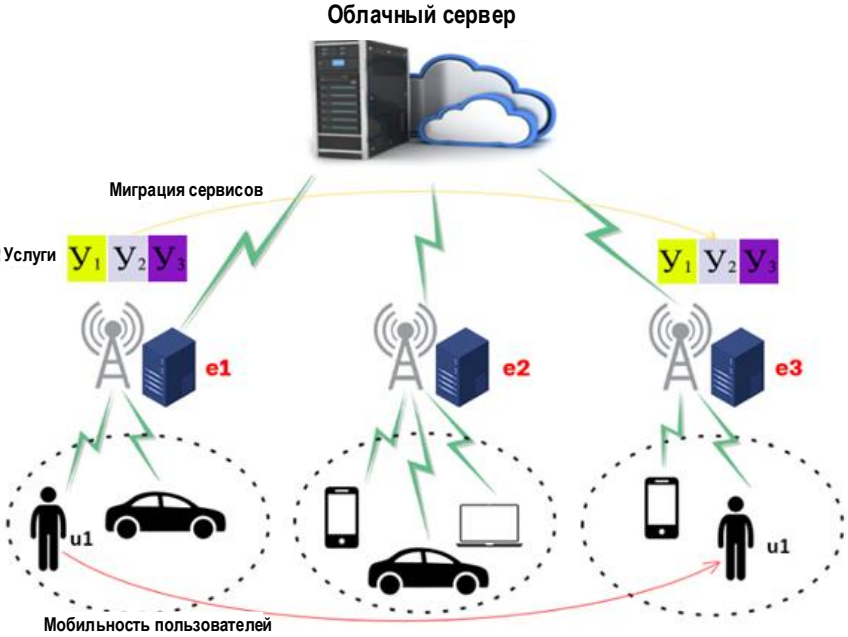

Рис. 1. Схема сети

Fig. 1. Network Diagram

Миграция сервисов является эффективным средством обеспечения непрерывности связи, но она дополнительно вызывает задержку в процессе миграции [7]. Кроме того, выполнение миграции сервисов ведет к изменению связей между пользователями систем телеприсутствия и пограничных

серверов (ES, *аббр от англ.* Edge Server) при этом нам необходимо учитывать неоднородность и ограниченность ресурсов целевых пограничных серверов. Перераспределение ресурсов на назначенных пограничных серверах после перемещения сервисов влияет на вычислительные задержки связи для мобильных пользователей. Это, в свою очередь, определяет, имеет ли смысл проведение данной миграции. Таким образом, важно одновременно рассчитать и учесть вероятную пользу от процесса миграции сервисов и перераспределения ресурсов и запускать процесс перемещения только в том случае, когда это способствует уменьшению времени отклика [2].

Для обеспечения непрерывного обслуживания и уменьшения задержек доступа будущие интеллектуальные системы телеприсутствия должны уметь распознавать мобильное поведение пользователей и заранее планировать схемы миграции услуг и распределения ресурсов. Однако предыдущие исследования в основном рассматривали миграцию услуг и распределение ресурсов как отдельные задачи, без учета их взаимного влияния. Результаты принятого решения о миграции услуг влияют на непрерывность обслуживания, тогда как стратегии распределения ресурсов оказывают воздействие на процессы миграции. К тому же разнообразие пользователей предъявляет различные требования к ресурсам и имеет уникальные шаблоны мобильности. Ресурсы в ES неоднородны и могут изменяться со временем, что добавляет сложности в совмещение оптимизации миграции услуг и распределения ресурсов [1–2]. Чтобы справиться с этими задачами, была исследована работа SMRA с учетом этих двух факторов. В ходе исследования были проанализированы определение необходимости миграции сервисов, выбор мест для их перемещения и оптимального распределения ресурсов. Авторы ставили своей целью обеспечить мобильность пользователей, оптимизировать использование ресурсов системы телеприсутствия, минимизировать задержки доступа, предотвращать сбои в обслуживании и улучшить общее QoS [1].

В этой статье, в отличие от имеющихся исследований, предлагается рассмотреть алгоритм совместного использования SMRA+MEC, который учитывает специфику систем телеприсутствия и решает задачу о целесообразности проведения миграции сервисов и оптимального распределения ресурсов. Предлагаемый способ расчета позволяет сократить задержки при передаче данных и вести распределение ресурсов более оптимально. Показано, что комбинация методов SMRA+MEC является наиболее эффективным подходом к минимизации задержек, поскольку итеративный подход к расчету задержки позволяет эффективно использовать имеющиеся вычислительные ресурсы. Стоит отметить, что числовые значения, используемые в статье, имеют относительный и условный характер и, в первую очередь, предназначены для иллюстрации предлагаемого подхода в минимизации задержки. Основная часть статьи организована следующим образом: в разделе 2 представлена системная модель и формулировка проблемы, раздел 3 включает в себя описание параметров и результаты моделирования. В заключении приводятся основные выводы [2–3].

2. Предлагаемый подход

В сфере современных систем телеприсутствия достижение минимальных значений задержки имеет решающее значение для эффективности работы и улучшения пользовательского опыта. В интересах этой цели одним из перспективных подходов является использование комбинации технологий SMRA и MEC. Этот раздел посвящен сценарию, предлагающему подходы к описанию сложных, но эффективных методов, применяемых для повышения оптимизации задержки [8–9]. Данный сценарий предназначен для моделирования и оптимизации задержки в системах телеприсутствия. Это осуществляется за счет стратегического распределения вычислительных задач между несколькими периферийными узлами, выбора оптимальных решений для миграции сервисов и управления ресурсами.

Предлагаемый сценарий предполагает балансировку вычислительных задач между различными периферийными узлами, динамическую миграцию сервисов и эффективное распределение ресурсов. Его цель – минимизировать общую задержку, которую испытывают пользователи. Далее рассмотрим ключевые аспекты модели, акцентируя внимание на сложных методах, используемых в ней. Имитационная модель строит сетевой график с помощью библиотеки NetworkX, чтобы смоделировать физическую и логическую структуру сети. График отображает взаимосвязи между разными периферийными узлами и центральным облачным сервером, определяет маршруты передачи данных и возможные задержки. Для воспроизведения реальности сценария модель симулирует миграцию сервисов и адаптацию распределения ресурсов среди периферийных узлов. Абстрактные функции моделируют эти процессы, сосредотачиваясь на принятии решений, направленных на выбор момента для миграции сервиса с учетом оптимального распределения полосы пропускания и вычислительной мощности. Центральная роль отводится алгоритму оптимизации, который минимизирует суммарную задержку в сети, последовательно подстраивая расположение сервисов и распределение ресурсов. В этом процессе может использоваться метод градиент-

ного спуска или другие техники нелинейного программирования. После завершения процесса оптимизации модель анализирует полученные результаты, выявляя конфигурации, обеспечивающие минимальную задержку. Эта обратная связь позволяет непрерывно улучшать и настраивать параметры в условиях изменяющейся сетевой среды в режиме реального времени [7].

В модели реализуется процесс назначения пользователей к узлам Edge-сети с целью минимизации общей задержки при выполнении задач. Основу составляют различные методы, включая MEC и SMRA (комбинированный метод, использующий итеративный процесс улучшения назначений). Суть этого метода состоит в оптимальном распределении пользовательских задач между граничными узлами для снижения задержки.

Предлагаемый комбинированный метод включает формулу для расчета задержки L:

$$L = \sum_{i=1}^{N} \left(D_{trans}(u_i, e_{a_i}) + \frac{R_{task}(u_i)}{R_{exec}(e_{a_i})} \right), \qquad (1)$$

где N – количество пользователей; u_i – позиция пользователя i; e_{a_i} – позиция назначенного Edge-узла для пользователя i; $D_{trans}(u_i, e_{a_i})$ – задержка передачи данных между пользователем u_i и назначенным узлом; $R_{task}(u_i)$ – скорость задачи пользователя u_i (количество задач в секунду); $R_{exec}(e_{a_i})$ – скорость выполнения задач на узле e_{a_i} (количество задач в секунду).

При этом задержка передачи данных $D_{trans}(u_i, e_{a_i})$ – время, затраченное на передачу данных от устройства пользователя к узлу, и может зависеть от расстояния, пропускной способности сети и текущей загруженности:

$$D_{trans}(u_i, e_{a_i}) = \frac{D_{data}(u_i)}{B_{net}(u_i, e_{a_i})}, \qquad (2)$$

где $D_{data}(u_i)$ – объем данных для передачи; $B_{net}(u_i, e_{a_i})$ – доступная пропускная способность между пользователем и узлом.

Кроме того, необходимо рассмотреть отдельный параметр: задержка выполнения $T_{exec}(u_i, e_{a_i})$ – это время, необходимое Edge-узлу для выполнения задачи пользователя. Данный параметр учитывает, как вычислительная мощность узла влияет на скорость выполнения задач. Задержка выполнения рассчитывается следующим образом [10–11]:

$$T_{exec}(u_i, e_{a_i}) = \frac{R_{task}(u_i)}{R_{exec}(e_{a_i})}. \qquad (3)$$

Следовательно, расчет задержки в предлагаемом методе можно представить в виде выражения:

$$L = \sum_{i=1}^{N} \left(D_{trans}(u_i, e_{a_i}) + T_{exec}(u_i, e_{a_i}) \right). \qquad (4)$$

Рисунок 2 наглядно показывает структуру алгоритма и представляет итеративный процесс улучшения назначения пользовательских задач Edge-узлам. Процесс начинается с произвольных назначений и использует методы случайного поиска для постепенного уменьшения задержки.

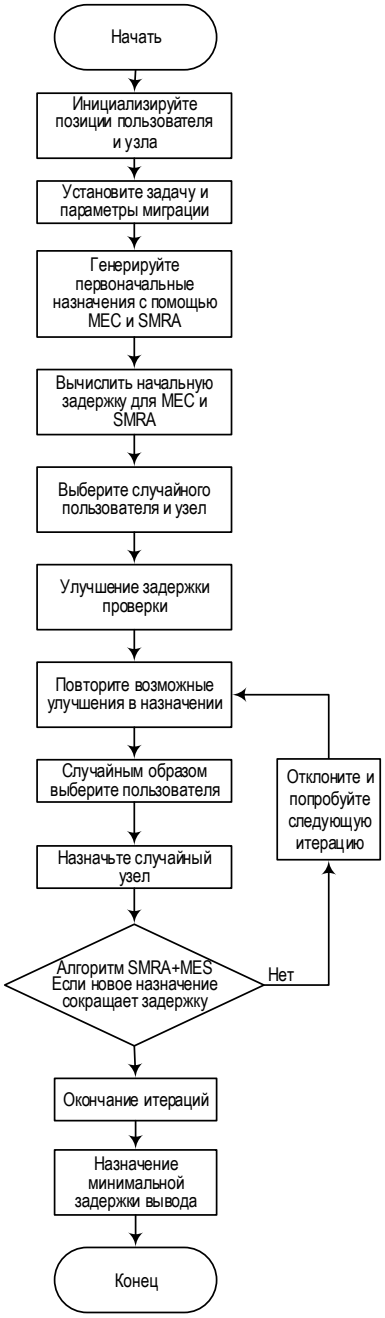

Рис. 2. Структура итеративной модели назначения граничных узлов алгоритма

Fig. 2. Structure of the Iterative Model for Assigning Edge Nodes of the Algorithm

3. Моделирование и анализ результатов

Ниже приводится краткое описание математической модели, которая включает в себя определение параметров, переменных и уравнений, описывающих представленную систему и процессы. Этот математический подход представляет собой набор параметрических задач оптимизации, связанных с распределением задач между узлами для минимизации общей задержки.

Описание модели и ее параметры

1) Позиции пользователей и граничных узлов

Пусть p_u – множество позиций пользователей:

$$p_u = \{i = 1, 2, \dots, N_u\},$$

где N_u – количество пользователей.

p_e – множество позиций граничных Edge-узлов:

$$p_e = \{j = 1,2, \dots, N_e\},$$

где N_e – количество граничных узлов.

2) Задачи и вычислительные ресурсы

Пусть T – вектор задач:

$$T = \{i = 1, 2, \dots, N_u\}.$$

R – вектор вычислительных мощностей Edge-узлов:

$$R = \{j = 1, 2, \dots, N_e\}.$$

3) Назначение пользователей на Edge-узлы

Пусть A – индекс узла, на который назначен пользователь i:

$$A = \{i = 1, 2, \dots, N_u\},$$

4) Задержка миграции

Пусть M – матрица стоимости миграции, где элемент $M_{j,k}$ обозначает стоимость миграции между узлами j и k.

Функции и метрики

1) Функция задержки передачи данных

Пусть $D_{trans}(u_i, e_j)$ – функция, которая вычисляет задержку передачи между пользователем i и узлом j:

$$D_{trans}(u_i, e_j) = \frac{1}{100}\sqrt{\left(x_u^{(i)} - x_e^{(j)}\right)^2 + \left(y_u^{(i)} - y_e^{(j)}\right)^2}. \quad (5)$$

2) Функция задержки выполнения задачи

Пусть $D_{exec}(t_i, r_j)$ – функция, которая вычисляет время выполнения задачи t_i на узле j:

$$D_{exec}(t_i, r_j) = \frac{t_i}{r_j}. \quad (6)$$

где t_i – вычислительная нагрузка пользователя i; r_j – вычислительная мощность узла j.

3) Общая задержка

Общая задержка для каждого пользователя i при назначении на узел j будет суммой задержек передачи данных и выполнения задачи:

$$D_{total(a_i)} = D_{trans(u_i,e_j)} + D_{exec(t_i,r_j)}. \quad (7)$$

4) Оптимизация назначения пользователей

Цель состоит в минимизации общей задержки для всех пользователей:

$$D^* = \sum_{i=1}^{N_u} D_{total(a_i)}. \quad (8)$$

Итерационная оптимизация

Пусть $A^{(t)}$ – текущее назначение на итерации t. Итерационный процесс обновления назначения осуществляется следующим образом:

1) выбор случайного пользователя i и случайный узел j;

2) расчет новой задержки $D_{total(a_i)}$ при изменении назначения;

3) обновление назначения $A^{(t+1)}$, если новая задержка меньше предыдущей.

Процесс продолжается до того момента, пока назначение не стабилизируется. Основная задача заключается в нахождении оптимального назначения пользователей на узлы с минимальной общей задержкой. Решение задачи происходит итерационно, что позволяет постепенно улучшать решение на основе случайных изменений назначения.

В рамках предлагаемой модели осуществляется настройка позиций пользователей и Edge-узлов случайным образом ограниченных территорией в 100×100 м (рисунок 3).

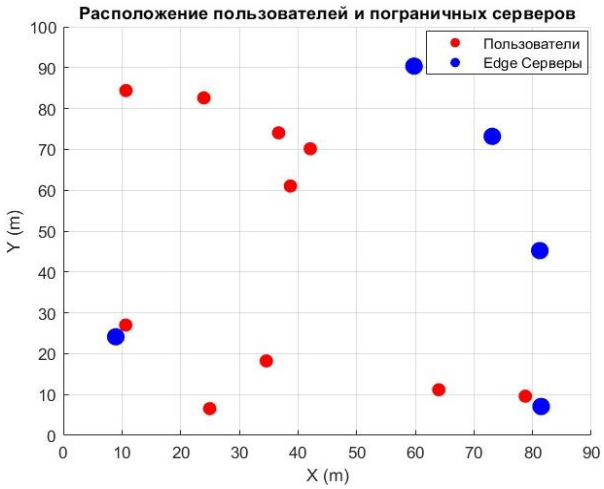

Рис. 3. Изначальное расположение пользователей и узлов

Fig. 3. Initial Placement of Users and Nodes

Задается число пользователей и Edge-узлов. В данном случае рассматривается 10 пользователей и 5 Edge-узлов, при этом можно проанализировать

любые значения. Именно эти значения были выбраны для наглядности и упрощения расчетов. Далее случайным образом определяются значения скорости выполнения задач для пользователей и – для узлов, а также затраты (стоимость) на миграцию между узлами.

Для повышения объективности модели, случайным образом запускается процесс генерации узлов и назначение их пользователям методов MEC и SMRA, после чего рассчитывается задержка для каждого из них. Комбинированный метод SMRA и MEC основан на итеративном поиске улучшений назначений для сокращения задержки: на каждом шаге итерации случайным образом выбирается пользователь, ему назначается новый узел; если такое изменение приводит к уменьшению задержки, оно фиксируется как текущее наилучшее решение. Таким образом, этот метод представляет собой процесс пошагового улучшения посредством случайного поиска, направленного на минимизацию задержки, начиная с подходов MEC и SMRA и постепенно совершенствуя назначения. Результаты работы метода представлены на рисунках 4–7.

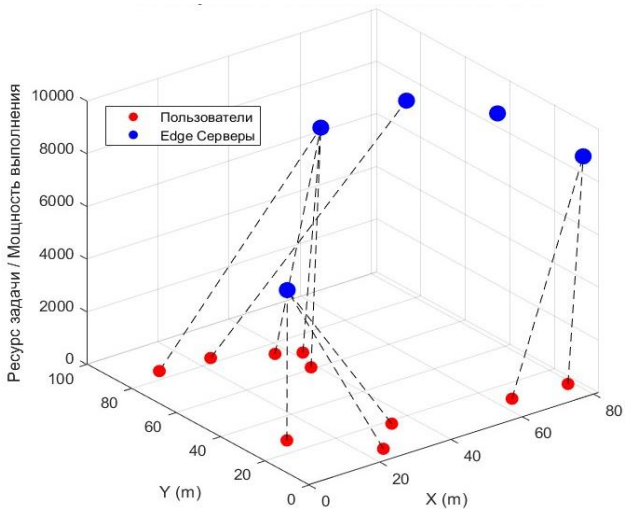

Рис. 4. 3D-визуализация назначения задач SMRA+MEC

Fig. 4. 3D Visualization of Task Assignment SMRA+MEC

Предлагаемая модель комбинирует случайные начальные назначения (используя методы MEC и SMRA) с последующим итеративным улучшением через случайный поиск. Такой подход позволяет снизить задержку выполнения задач благодаря динамическому перераспределению задач пользователей между Edge-узлами, что помогает находить самые эффективные маршруты передачи и обработки в сети. На рисунке 7 видно, что предложенный метод достигает минимальной задержки: задержка MEC составляет 60 мс, SMRA – 55 мс, а комбинированный метод SMRA+MEC дает результат всего в 31 мс. Рисунок 7 показывает, что задержка может быть снижена почти на 50 %, что говорит о значительном потенциале предложенной модели.

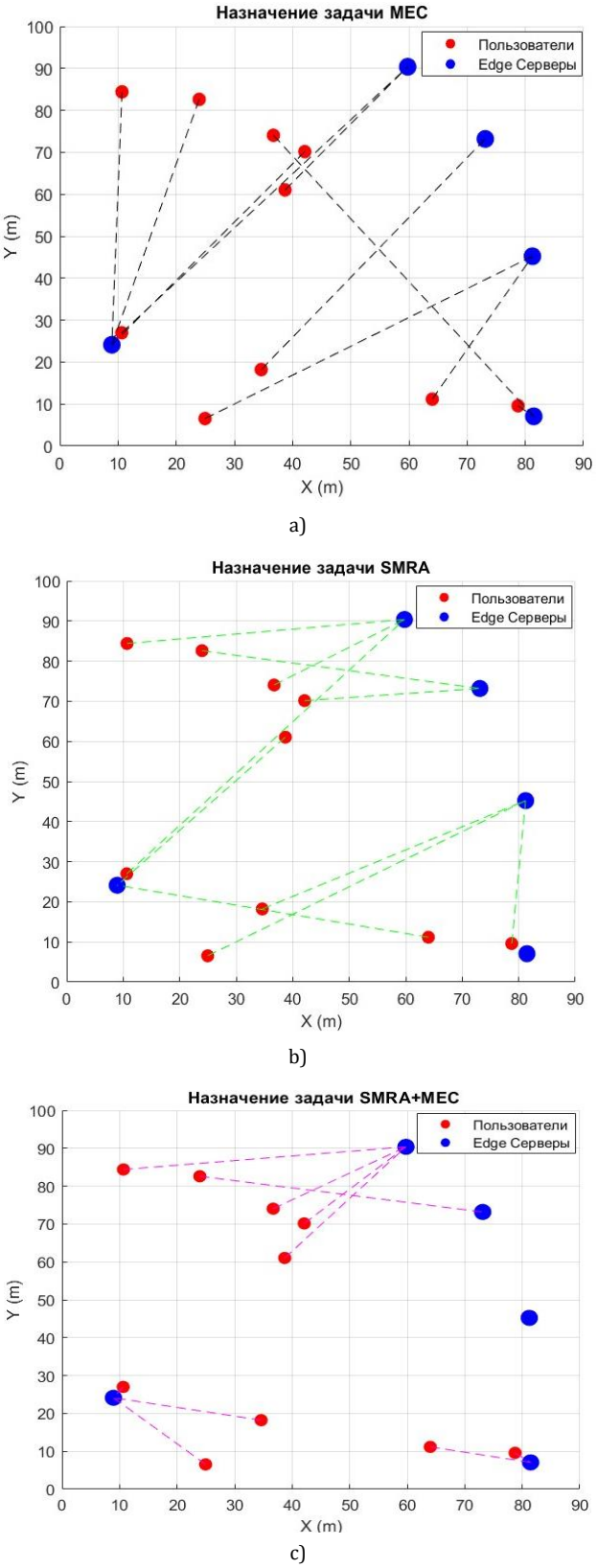

Рис. 5. Назначение задач методами MEC (a), SMRA (b) и комбинированным (c)

Fig. 5. Assignment of Tasks by MEC (a), SMRA (b) and Combined (c) Methods

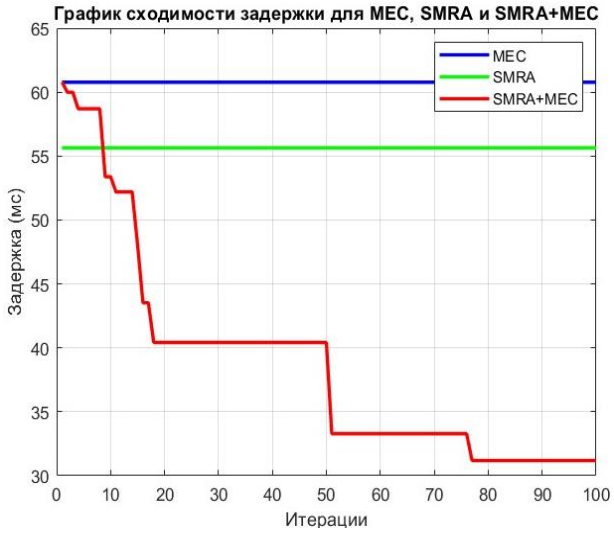

Рис. 6. График сходимости задержки для MEC, SMRA и предлагаемый алгоритм SMRA+MEC

Fig. 6. Graph of Latency Convergence for MEC, SMRA, and the Proposed SMRA+MEC Algorithm

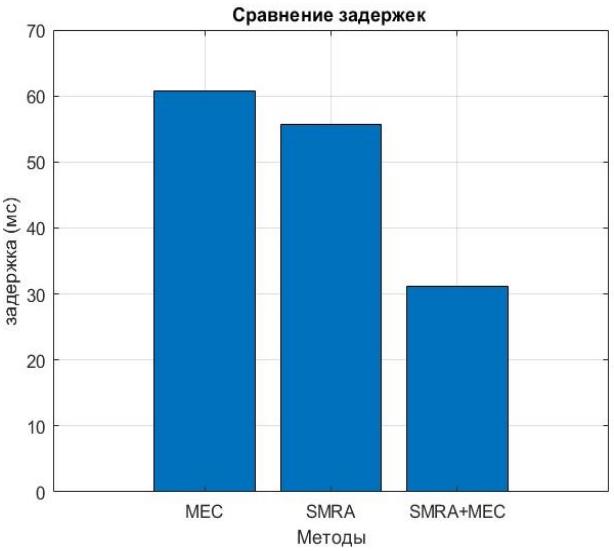

Рис. 7. Сравнение задержек

Fig. 7. Comparison of Latency

Однако стоит отметить, что представленные в ходе исследования значения являются ориентировочными и могут потребовать уточнения для конкретных сетей или систем. Полученное итоговое значение и процент уменьшения дают представление об уровне повышения эффективности, достигаемого в ходе оптимизации.

Заключение

Описанный комбинированный метод успешно объединяет затраты на миграцию и задержку обработки в периферийных узлах для оптимизации размещения сервисов и минимизации задержек в устройствах телеприсутствия в среде мобильных периферийных вычислений. Формула учитывает основные факторы, влияющие на задержку, а итеративный процесс оптимизации обеспечивает наилучшее распределение, что подтверждается сравнением расчетных, визуальных распределений и средних значений задержек. Этот сценарий демонстрирует сложную взаимосвязь между миграцией услуг, распределением ресурсов и мобильными периферийными вычислениями для повышения эффективности систем телеприсутствия.

Благодаря рациональному распределению вычислительных задач и динамической оценке задержек достигается минимальная задержка для устройств телеприсутствия. Задержка может быть снижена примерно на 50 %, а итоговые значения и процент снижения отражают повышение эффективности в результате оптимизации. Этот сценарий является моделью для разработки более эффективных систем связи в реальном времени и расширения границ технологий телеприсутствия. В условиях быстро меняющегося цифрового ландшафта оптимизация задержек с использованием инновационных подходов остается ключевым фактором. Этот сценарий показывает, как теоретические концепции в области вычислений и сетевой оптимизации могут улучшить повседневные технологические приложения на практике.

Список источников

1. Ateya A.A., Abd El-Latif A.A., Muthanna A., Volkov A., Koucheryavy A. Enabling Metaverse and Telepresence Services in 6G Networks. New York: CRC Press, 2025. DOI:10.1201/9788770046749

2. Thang D.V., Volkov A., Muthanna A., Koucheryavy A., Ateya A.A., Jayakody D.N.K. Future of Telepresence Services in the Evolving Fog Computing Environment: A Survey on Research and Use Cases // Sensors. 2025. Vol. 25. Iss. 11. P. 3488. DOI:10.3390/s25113488

3. Van Thang D., Volkov A., Muthanna A., Elgendy I.A., Alkanhel R., Jayakody D.N.K., Koucheryavy A. A Framework Integrating Federated Learning and Fog Computing Based on Client Sampling and Dynamic Thresholding Techniques // IEEE Access. 2025. Vol. 13. PP. 95019–95033. DOI:10.1109/ACCESS.2025.3571979

4. Taleb T., Samdanis K., Mada B., Flinck H., Dutta S., Sabella D. On Multi-Access Edge Computing: A Survey of the Emerging 5G Network Edge Cloud Architecture and Orchestration // IEEE Communications Surveys & Tutorials. 2017. Vol. 19. Iss. 3. PP. 1657–1681. DOI:10.1109/COMST.2017.2705720

5. Чистова Н.А., Парамонов А.И., Кучерявый А.Е. Метод формирования цифровых кластеров сетей связи пятого и последующих поколений на основе качества предоставления услуг // Электросвязь. 2020. № 7. С. 22–28. DOI:10.34832/ELSV.2020.8.7.003. EDN:QDEUQG

6. Yu H., Ming Z., Wang C., Taleb T. Network Slice Mobility for 6G Networks by Exploiting User and Network Prediction // Proceedings of the International Conference on Communications (ICC, Rome, Italy, 28 May – 01 June 2023). IEEE, 2023. DOI:10.1109/ICC45041.2023.10279739

7. Addad R.A., Dutra D.L.C., Taleb T., Flinck H. Toward Using Reinforcement Learning for Trigger Selection in Network Slice Mobility // IEEE Journal on Selected Areas in Communications. 2021. Vol. 39. Iss. 7. PP. 2241–2253. DOI:10.1109/jsac.2021.3078501. EDN:UGJVKC
8. Волков А.Н., Мутханна А.С.А., Кучерявый А.Е., Бородин А.С., Парамонов А.И., Владимиров С.С. и др. Перспективные исследования сетей и услуг 2030 в лаборатории 6G MEGANETLAB СПБГУТ // Электросвязь. 2023. № 6. С. 5–14. DOI:10.34832/ELSV.2023.43.6.001. EDN:CJSYLS
9. Hu L., Tian Y., Yang J., Taleb T., Xiang L., Hao Y. Ready Player One: UAV-Clustering-Based Multi-Task Offloading for Vehicular VR/AR Gaming // IEEE Network. 2019. Vol. 33. Iss. 3. PP. 42–48. DOI:10.1109/MNET.2019.1800357
10. Chen Y., Sun Y., Wang C., Taleb T. Dynamic Task Allocation And Service Migration in Edge-Cloud IoT System Based on Deep Reinforcement Learning // IEEE Internet of Things Journal. 2022. Vol. 9. Iss. 18. PP. 16742–16757. DOI:10.1109/JIOT.2022.3164441. EDN:XSAUUL
11. Ming Z., Li X., Sun C., Fan Q., Wang X., Leung V.C.M. Dependency-Aware Hybrid Task Offloading in Mobile Edge Computing Networks // Proceedings of the International Conference on Parallel and Distributed Systems (ICPADS, Beijing, China, 14–16 December 2021). IEEE, 2021. PP. 225–232. DOI:10.1109/ICPADS53394.2021.00034

References

1. Ateya A.A., Abd El-Latif A.A., Muthanna A., Volkov A., Koucheryavy A. *Enabling Metaverse and Telepresence Services in 6G Networks*. New York: CRC Press; 2025. DOI:10.1201/9788770046749
2. Thang D.V., Volkov A., Muthanna A., Koucheryavy A., Ateya A.A., Jayakody D.N.K. Future of Telepresence Services in the Evolving Fog Computing Environment: A Survey on Research and Use Cases. *Sensors*. 2025;25(11):3488. DOI:10.3390/s25113488
3. Van Thang D., Volkov A., Muthanna A., Elgendy I.A., Alkanhel R., Jayakody D.N.K., Koucheryavy A. A Framework Integrating Federated Learning and Fog Computing Based on Client Sampling and Dynamic Thresholding Techniques. *IEEE Access*. 2025;13: 95019–95033. DOI:10.1109/ACCESS.2025.3571979
4. Taleb T., Samdanis K., Mada B., Flinck H., Dutta S., Sabella D. On Multi-Access Edge Computing: A Survey of the Emerging 5G Network Edge Cloud Architecture and Orchestration. *IEEE Communications Surveys & Tutorials*. 2017;19(3):1657–1681. DOI:10.1109/COMST.2017.2705720
5. Chistova N.A., Paramonov A.I., Koucheryavy A.E. The method of forming the digital clusters for communication networks of fifth and subsequent generations based on QoS. *Electrosvyaz*. 2020;7:22–28. (in Russ.) DOI:10.34832/ELSV.2020.8.7.003. EDN:QDEUQG
6. Yu H., Ming Z., Wang C., Taleb T. Network Slice Mobility for 6G Networks by Exploiting User and Network Prediction. *Proceedings of the International Conference on Communications, ICC, 28 May – 01 June 2023, Rome, Italy*. IEEE; 2023. DOI:10.1109/ICC45041.2023.10279739
7. Addad R.A., Dutra D.L.C., Taleb T., Flinck H. Toward Using Reinforcement Learning for Trigger Selection in Network Slice Mobility. *IEEE Journal on Selected Areas in Communications*. 2021;39(7):2241–2253. DOI:10.1109/jsac.2021.3078501. EDN:UGJVKC
8. Volkov A.N., Muthanna A.S.A., Kucheryavy A.E., Borodin A.S., Paramonov A.I., Vladimirov S.S. et al. Perspective research of networks and services 2030 in the laboratory 6G MEGANETLAB SPBSUT. *Electrosvyaz*. 2023;6:5–14. (in Russ.) DOI:10.34832/ELSV.2023.43.6.001. EDN:CJSYLS
9. Hu L., Tian Y., Yang J., Taleb T., Xiang L., Hao Y. Ready Player One: UAV-Clustering-Based Multi-Task Offloading for Vehicular VR/AR Gaming. *IEEE Network*. 2019;33(3):42–48. DOI:10.1109/MNET.2019.1800357
10. Chen Y., Sun Y., Wang C., Taleb T. Dynamic Task Allocation And Service Migration in Edge-Cloud IoT System Based on Deep Reinforcement Learning. *IEEE Internet of Things Journal*. 2022;9(18):16742–16757. DOI:10.1109/JIOT.2022.3164441. EDN:XSAUUL
11. Ming Z., Li X., Sun C., Fan Q., Wang X., Leung V.C.M. Dependency-Aware Hybrid Task Offloading in Mobile Edge Computing Networks. *Proceedings of the International Conference on Parallel and Distributed Systems, ICPADS, 14–16 December 2021, Beijing, China*. IEEE; 2021. p.225–232. DOI:10.1109/ICPADS53394.2021.00034

Статья поступила в редакцию 04.08.2025; одобрена после рецензирования 29.08.2025; принята к публикации 01.09.2025.
The article was submitted 04.08.2025; approved after reviewing 29.08.2025; accepted for publication 01.09.2025.

Информация об авторах:

АЛЬ-КЕРЕА Захраа Ахмед Хуссейн — аспирант кафедры сетей связи и передачи данных Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича
https://orcid.org/0009-0002-1659-8169

МУТХАННА Аммар Салех Али — доктор технических наук, доцент, профессор кафедры сетей связи и передачи данных Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича
https://orcid.org/0000-0003-0213-8145

КУЧЕРЯВЫЙ Андрей Евгеньевич — доктор технических наук, профессор, заведующий кафедрой сетей связи и передачи данных Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича
https://orcid.org/0000-0003-4479-2479

Кучерявый А.Е. является членом редакционного совета журнала «Труды учебных заведений связи» с 2016 г., но не имеет никакого отношения к решению опубликовать эту статью. Статья прошла принятую в журнале процедуру рецензирования. Об иных конфликтах интересов авторы не заявляли.