

Научная статья

УДК 621.391

<https://doi.org/10.31854/1813-324X-2025-11-1-62-68>

EDN:RVENVC



Оптимизация использования ресурсов воздушных базовых станций на основе методов искусственного интеллекта

✉ Тунг Зыонг Чан, chan.tz@sut.ruАндрей Евгеньевич Кучерявый, akouch@sut.ru

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

Аннотация

В отдаленных областях и районах стихийных бедствий беспилотные летательные аппараты (БПЛА) могут выступать в качестве базовых станций, обеспечивающих беспроводную связь с наземными пользователями. Благодаря своей высокой мобильности, низкой стоимости, а также быстрому разворачиванию и поиску, БПЛА могут изменять свое местоположение в трехмерном пространстве, улучшая беспроводное соединение и повышая скорость передачи данных. В этой статье исследуются проблемы разворачивания воздушных базовых станций (ABS, аббр. от англ. Aerial Base Station) в трехмерном пространстве и распределения мощности в целях максимизации скорости передачи данных в системе. Для решения этих проблем предложено использовать алгоритм *Q-learning*, относящийся к методам обучения с подкреплением. Используя БПЛА в качестве агента, алгоритм позволяет ABS исследовать пространство состояний на основе политики *ε-greedy* (эпсилон жадный алгоритм) для определения местоположения в трехмерном пространстве и распределения мощности. Результаты моделирования показывают, что предложенный алгоритм превосходит известные методы размещения ABS в трехмерном пространстве и распределения мощности. **Целью** настоящей статьи является исследование эффективности применения современных методов искусственного интеллекта для оптимизации использования ресурсов воздушных базовых станций сетей связи общего пользования.

Сущность предлагаемого решения состоит в применении современных методов искусственного интеллекта, а именно: метода обучения *Q-learning* и эпсилон-жадного алгоритма *ε-greedy* для обеспечения совместной оптимизации размещения ABS и распределения мощности для максимизации скорости передачи данных. Система **имеет реализацию** в виде программы моделирования. **Эксперименты** при моделировании показали, что использование метода обучения с подкреплением *Q-learning* и эпсилон-жадного алгоритма *ε-greedy* для совместной оптимизации обеспечивает более высокую общую скорость передачи данных в системе по сравнению с оптимизацией только местоположения или распределения мощности.

Научная новизна предложенного решения состоит в том, что совместная оптимизация размещения ABS и распределения мощности позволила, в отличие от известных результатов, выявить, что высота полета БПЛА с установленной на нем базовой станцией при оптимизации только местоположения будет выше, чем высота полета БПЛА при совместной оптимизации местоположения и распределения мощности.

Практическая значимость заключается в возможности разработки методики планирования сетей связи общего пользования при использовании ABS для получения более высокой общей скорости передачи данных на соответствующем фрагменте сети.

Ключевые слова: воздушная базовая станция, трехмерное местоположение, распределение мощности, скорость передачи данных, *Q-learning*

Ссылка для цитирования: Чан Т.З., Кучерявый А.Е. Оптимизация использования ресурсов воздушных базовых станций на основе методов искусственного интеллекта // Труды учебных заведений связи. 2025. Т. 11. № 1. С. 62–68. DOI:10.31854/1813-324X-2025-11-1-62-68. EDN:RVENVC

Original research

<https://doi.org/10.31854/1813-324X-2025-11-1-62-68>

EDN:RVENV

Resource Optimization of Airborne Base Stations Using Artificial Intelligence Methods

 **Tung D. Tran** , chan.tz@sut.ru

 **Andrey E. Koucheryavy**, akouch@sut.ru

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

Annotation

In remote areas and disaster-stricken regions, unmanned aerial vehicles (UAVs) can serve as base stations, providing wireless communication to ground users. Due to their high mobility, low cost, and rapid deployment and retrieval capabilities, UAVs can continuously adjust their position in three-dimensional (3D) space, improving wireless connectivity and enhancing data transmission rates. In this paper, we investigate the problem of ABS (Aerial Base Station) deployment in 3D space and power allocation with the aim of maximizing the data transmission rate in the system. To address this non-convex problem, we propose Q-learning, a reinforcement learning algorithm. By using the ABS as an agent, the algorithm enables the ABS to explore the state space and take actions based on an ϵ -greedy policy (optimal epsilon value) to determine its 3D position and power allocation. Simulation results demonstrate that the proposed algorithm outperforms individual position optimization and power allocation optimization.

The purpose of this article is to study the efficiency of using modern artificial intelligence methods to optimize the use of resources of airborne base stations of public communication networks.

The essence of the proposed solution is to use modern artificial intelligence methods, namely: the Q-learning method and the epsilon-greedy ϵ -greedy algorithm to ensure joint optimization of the placement of airborne base stations and power distribution to maximize the data transfer rate. **The system has an implementation** in the form of a simulation program. **Simulation experiments** have shown that the use of the Q-learning reinforcement learning method and the epsilon-greedy ϵ -greedy algorithm for joint optimization provides a higher overall data transfer rate in the system compared to optimizing only the location or power distribution.

The scientific novelty of the proposed solution is that joint optimization of the placement of an airborne base station and power distribution made it possible, in contrast to known results, to establish that the flight altitude of a UAV with a base station installed on it when optimizing only the location will be higher than the flight altitude of a UAV when jointly optimizing the location and power distribution.

The practical significance is the possibility of developing a methodology for planning public communication networks using airborne base stations to obtain a higher overall data transfer rate on the corresponding network fragment.

Keywords: ABS, 3D positioning, power allocation, data transmission rate, Q-learning

For citation: Tran T.D., Koucheryavy A.E. Resource Optimization of Airborne Base Stations Using Artificial Intelligence Methods. *Proceedings of Telecommunication Universities*. 2025;11(1):62–68. (in Russ.) DOI:10.31854/1813-324X-2025-11-1-62-68. EDN:RVENV

1. Введение

Интернет вещей, являясь одной из определяющих концепций для сетей связи пятого поколения, позволяет осуществлять взаимодействие между устройствами Интернета вещей, такими как сенсоры, носимые устройства и смартфоны [1]. Однако появление большого числа услуг Интернета вещей создает серьезные проблемы для существующих наземных сетей. Если говорить более конкретно,

число устройств Интернета вещей может достичь 100 миллиардов к 2025 г. [2], тем самым создается огромное количество источников данных, которые еще больше усложняют работу текущей наземной сети. Из-за жесткой структуры наземной сети, в основе которой, как правило, используются стационарные базовые станции и которые могут с большой вероятностью быть повреждены в случае природных катастроф, системе Интернета вещей до-

статочно сложно обеспечить покрытие и сбор данных в реальном времени от устройств в зонах бедствий и отдаленных районах. Летающие базовые станции являются одним из перспективных решений для преодоления перегрузки сетевого трафика, обеспечивающих быстрый и экономически эффективный способ поддержки потребностей в беспроводных соединениях в таких условиях [3].

Беспилотные летательные аппараты (БПЛА) играют незаменимую роль в сетях связи пятого и последующих поколений. Благодаря простоте развертывания, беспроводные сети с поддержкой БПЛА могут быть достаточно легко построены, переконфигурированы и повторно использованы по назначению в зависимости от их предназначения. Применение БПЛА возможно во многих областях, например, в таких как мониторинг, наблюдение и доставка грузов [4]. По сравнению с традиционными наземными сетями БПЛА развертываются в воздухе, что обеспечивает более высокую вероятность соединений в зоне прямой видимости (LoS, от англ. Line of Sight) для канала «воздух – земля». Кроме того, мобильность позволяет БПЛА легко перемещаться с места на место во время сбора данных в удаленных районах, где нет наземной сетевой системы. Они также могут выступать в качестве трехмерных (3D) базовых станций для поддержки наземного пользовательского оборудования, а также для расширения зоны действия связи в труднодоступных районах или зонах стихийных бедствий.

Использование вышеупомянутых преимуществ сетей с поддержкой БПЛА требует решения многих технических проблем с точки зрения проектирования распределения ресурсов, а также – проблем, связанных с полезной нагрузкой, временем полета и зарядкой аккумулятора. В этой статье исследуется нисходящий канал связи пользователей, обслуживаемых воздушной базовой станцией (ABS, аббр от англ. Aerial Base Station). Цель исследования состоит в том, чтобы максимизировать общую пропускную способность системы за счет оптимизации местоположения ABS в 3D-пространстве, а также – управления мощностью передачи для каждого пользователя. Для решения этой задачи в статье предложено использование обучения с подкреплением на основе метода Q -learning. Поскольку оказание услуг абонентам сетей связи общего пользования осуществляется базовой станцией, расположенной на борту БПЛА во время его нахождения в стационарном состоянии, исследование характеристик собственно БПЛА в статье не требуется.

2. Модель системы и постановка задачи

Будем рассматривать нисходящую линию связи единой беспроводной сети с поддержкой ABS, состоящей из N наземных пользователей. Все назем-

ные пользователи являются стационарными, что обозначается множеством $TV = (1, 2, \dots, N)$. БПЛА развернут для зависания в определенном месте в этой области, чтобы предоставлять пользователям услуги беспроводной связи по нисходящей линии связи. Местоположение ABS в 3D-пространстве обозначается через (x_a, y_a, h_a) , где (x_a, y_a) – горизонтальная координата местоположения ABS, а h_a – высота ABS. Местоположение пользователя n обозначается через $(x_n, y_n, 0)$. Предполагается, что ABS обладает полной информацией о местоположении пользователя перед развертыванием. Кроме того, для предотвращения беспроводных помех применяется множественный доступ с частотным разделением, а полоса пропускания системы, обозначаемая как B , равномерно распределяется среди всех пользователей.

Канал связи ABS с землей можно описать как вероятностную модель, которая состоит из соединений в условиях LoS и непрямой видимости (NLoS, от англ. Non-Line of Sight).

Вероятность потери связи между ABS и пользователем k определяется следующим образом:

$$P_n^{LoS} = \frac{1}{1 + a \exp(-b[\frac{180}{\pi} \arcsin(\frac{h_a}{d_{a,n}}) - a])}, \quad (1)$$

где a и b – параметры среды распространения; $\arcsin(\frac{h_a}{d_{a,n}})$ – угол между ABS и пользователем n ; $d_{a,n}$ – 3D-расстояние от ABS до пользователя n .

Тогда вероятность соединения в условиях NLoS будет равна $P_n^{NLoS} = 1 - P_n^{LoS}$. Без потери общности, обозначим путь ABS между собой и пользователем n как $\eta_{a,n}^{LoS} = \beta_0 d_{a,n}^{-\hat{\alpha}}$ и $\eta_{a,n}^{NLoS} = \chi \beta_0 d_{a,n}^{-\hat{\alpha}}$ для LoS- и NLoS-подключения, соответственно, где β_0 – потери распространения на эталонном расстоянии 1 м для соединений в условиях LoS; $0 < \chi < 1$ – дополнительно необходимая средняя дальность распространения для соединений в условиях NLoS; $\hat{\alpha}$ – показатель потерь распространения.

Предположим, что коэффициент усиления канала между ABS и пользователем n составляет h_n , где h – случайная величина с двумя разными уровнями случайности, т. е. случайным возникновением связи в условиях LoS и NLoS и случайным замиранием.

Чтобы упростить задачу, среднее усиление мощности канала можно определить, как:

$$h_n = P_n^{LoS} \beta_0 d_{a,n}^{-\hat{\alpha}} + P_n^{NLoS} \chi \beta_0 d_{a,n}^{-\hat{\alpha}} = \overline{P_n^{LoS}} \beta_0 d_{a,n}^{-\hat{\alpha}}, \quad (2)$$

где $\overline{P_n^{LoS}} = P_n^{LoS} + (1 - P_n^{LoS})\chi$.

Будем предполагать, что общая мощность передачи ABS равна P_a . Пусть p_n представляет собой коэффициент распределения мощности для пользователя n , σ^2 – мощность аддитивного белого гауссовского шума.

Тогда отношение сигнал / шум у пользователя n определяется как:

$$\gamma_n = \frac{P_a p_n h_n}{\sigma^2}. \quad (3)$$

Предполагая, что B обозначает общую полосу пропускания канала, доступную для ABS, достижимая скорость нисходящей линии связи от ABS к пользователю n может быть рассчитана на основе формулы Шеннона [5]:

$$R_n = \frac{B}{N} \log_2(1 + \gamma_n) = \frac{B}{N} \log_2\left(1 + \frac{P_a p_n h_n}{\sigma^2}\right). \quad (4)$$

Как уже отмечалось выше, цель данной работы состоит в максимизации суммарной скорости передачи данных при ограничениях на местоположение ABS в 3D-пространстве и распределения мощности для каждого пользователя. Задача может быть сформулирована следующим образом:

$$\begin{aligned} (P) \quad & \max_{\{x_a, y_a, h_a, p_n\}} D = \sum_{n=1}^N R_n, \\ & C1: \sum_{n=1}^N p_n = 1, \\ & C2: p_n > 0, \forall n \in N, \\ & C3: \gamma_n \geq \gamma_0, \forall n \in N, \\ & C4: x_a, y_a, h_a \in S. \end{aligned} \quad (5)$$

Ограничение C1 означает, что общая сумма коэффициентов распределения мощности для всех пользователей в системе должна равняться 1; ограничение C2 представляет собой коэффициент распределения мощности для каждого пользователя; ограничение C3 указывает, что отношение сигнал / шум у пользователя n должно быть больше или равно пороговому значению γ_0 , тогда считается, что пользователь n подключен к ABS; ограничение C4 требует, чтобы ABS была размещена в заданном пространстве.

3. 3D-определение местоположения и распределения мощности ABS на основе Q-learning

В целом проблема оптимизации P является чрезвычайно сложной для решения из-за отсутствия выпуклости ограничений на переменные положения ABS и передающей мощности. Кроме того, число переменных увеличивается вместе с ростом количества пользователей. Поэтому проблема P является невыпуклой, найти оптимальное решение довольно затруднительно. Далее предложена методика решения проблемы P , основанная на методе искусственного интеллекта Q-learning. Будет рассмотрена связь Q-learning с обучением с подкреплением (RL, аббр. от англ. Reinforcement Learning).

В алгоритме Q-learning ABS выступает в роли агента, а сам алгоритм состоит из четырех частей: состояния, действий, вознаграждений и Q-значений. Основная цель Q-learning – получение политики, которая максимизирует наблюдаемый результат в процессе взаимодействия агентов. На каждом временном шаге итерации агенты наблюдают состояние из пространства состояний S . Соответственно, они выбирают действие из пространства действий A на основе политики J . Принцип политики на каждом временном шаге заключается в выборе действия, которое приводит модель к максимальному Q-значению. После выполнения действия состояние каждого агента переходит в новое состояние, и агент получает необходимый результат.

Состояния (S). Эти состояния представляют ситуацию среды системы ретрансляции на базе ABS, и принятие решения о действиях основывается на состоянии сети. В данной работе основными факторами, влияющими на состояние сетевой среды, являются положение ABS в 3D-пространстве и мощность передачи для разных пользователей в сети. Затем предоставляется обратная связь для корректировки основных факторов с целью улучшения производительности сети.

Обозначаем s_t как состояние ABS на временном слоте t , и оно может быть определено как:

$$s_t = \{x_a^t, y_a^t, h_a^t, p_1^t, p_2^t, \dots, p_n^t\}, \quad (6)$$

где $n \in N$, (x_a^t, y_a^t, h_a^t) – местоположение ABS на временном шаге t ; p_n^t – мощность, выделенная для пользователя n .

Действие (A). На каждом шаге обозначим a_t как действие ABS на временном слоте t . ABS принимает решения на основе состояния сети, что вызывает переход сети в новое состояние. Действие состоит из двух частей: одна часть – это перемещение в следующее состояние ABS, другая – распределение мощности. Перемещение ABS включает движение вперед, назад, вправо, влево, вверх и вниз. Распределение мощности для каждого пользователя может увеличиваться или уменьшаться.

Вознаграждение (R). Целью Q-learning является максимизация накопленного ожидаемого результата. Чтобы достичь цели, поставленной в уравнении (5), необходимо включить общую системную скорость передачи данных в функцию вознаграждения. На каждом временном шаге t общая скорость передачи данных для пользователя D_t может быть рассчитана на основе местоположения ABS. ABS будет регулировать свое местоположение и распределение мощности таким образом, чтобы поддерживать оптимальное состояние.

Таким образом, можно выразить функцию вознаграждения во временном интервале t как:

$$r_t = D_t - \lambda(N - N_u), \quad (7)$$

где N – общее число пользователей в рассматриваемой области; N_u – число пользователей, обслуживаемых ABS, статус пользователя может быть основан на формуле (3); λ – штрафной коэффициент.

Политика (P). Политика – это механизм выбора действий, используемый для определения тех действий, которые агент будет выполнять во время обучения. Ее главная цель – установить баланс между исследованием и эксплуатацией таким образом, чтобы агент мог подкрепить те суждения, которые он уже знает как хорошие, но также исследовать возможные новые действия [6]. В данной работе далее будем рассматривать наиболее известную политику ϵ -greedy (эпсилон-жадный алгоритм), в которой агент выбирает случайное действие с вероятностью ϵ и лучшее действие, соответствующее наилучшему Q -значению в данный момент, с вероятностью $1-\epsilon$.

Таким образом, вероятность выбора действия a_t в состоянии s_t на временном слоте t задается следующей формулой:

$$\pi_m(s_m, \theta_m) = \begin{cases} 1 - \epsilon, & \text{если } Q_m(\theta) \text{ наибольшее} \\ \epsilon, & \text{в противном случае} \end{cases}, \quad (8)$$

где $\epsilon \in (0,1)$.

Чтобы обеспечить сходимость Q -learning, скорость обучения α_t является функцией временного шага [7], который задается формулой:

$$\alpha_t = \frac{1}{(t + c_\alpha)^\phi}, \quad (9)$$

где $c_\alpha > 0$, $\phi \in (\frac{1}{2}, 1]$.

Во время обучения функция значения состояния-действия для агента может итеративно обновляться следующим образом:

$$Q_{t+1}(s_t, a_t) \leftarrow (1 - \alpha_t) \cdot Q_t(s_t, a_t) + \alpha_t \cdot [r_t + \beta \cdot \max_{a'} Q_t(s_{t+1}, a')], \quad (10)$$

где β – коэффициент дисконтирования.

В Q -learning параметр ϵ со временем уменьшается, чтобы обеспечить баланс между исследованием и эксплуатацией. В начале обучения высокое значение ϵ помогает агенту выбирать случайные действия для исследования окружающей среды и построения точной таблицы значений Q -learning. Однако, если ϵ останется высоким, агент всегда будет выбирать случайные действия, не используя полученные знания, что приведет к отсутствию сходимости процесса обучения. Напротив, если ϵ постепенно уменьшается, агент начнет отдавать предпочтение действиям с наивысшими значениями, согласно таблице значений Q -learning, тем самым оптимизируя политику действий. Это помогает агенту освоить лучшую стратегию после достаточного изучения пространства состояний, обеспечивая высокую производительность в долгосрочной перспективе.

Таким образом, в данной работе ϵ регулируется после каждого эпизода, в соответствии с формулой:

$$\epsilon_{p+1} = \epsilon_p \epsilon_0, \quad (11)$$

где ϵ_0 – постоянное значение, на которое ϵ уменьшается после каждого эпизода; ϵ_p и ϵ_{p+1} – значение ϵ в эпизоде p и $(p+1)$.

4. Результаты моделирования

Рассмотрим случай для $N = 5$ пользователей, равномерно распределенных на площади $300 \text{ м} \times 300 \text{ м}$. Горизонтальное положение ABS ограничено в пределах рассматриваемой области, а высота ABS фиксирована и составляет $h_a = 50 \text{ м}$. Другие параметры моделирования приведены в таблице 1.

ТАБЛИЦА 1. Параметры моделирования

TABLE 1. Simulation Parameters

Параметр	Значение	Параметр	Значение
P_a	100 мВт	B	5 МГц
a	9,16	b	0,16
σ^2	-110 дБм	β_0	-40 дБ
γ_0	5 дБ	$\hat{\alpha}$	-2,3
χ	0,1	–	–

Параметры для предложенного Q -learning следующие: $\beta = 0,99$; $c_\alpha = 0,5$; $\phi_\alpha = 0,7$; $\epsilon = 1$; $\epsilon_0 = 0,999$; штрафной коэффициент $\lambda = 20$. Число эпизодов обучения составляет 250 с шагом по времени $T = 500$.

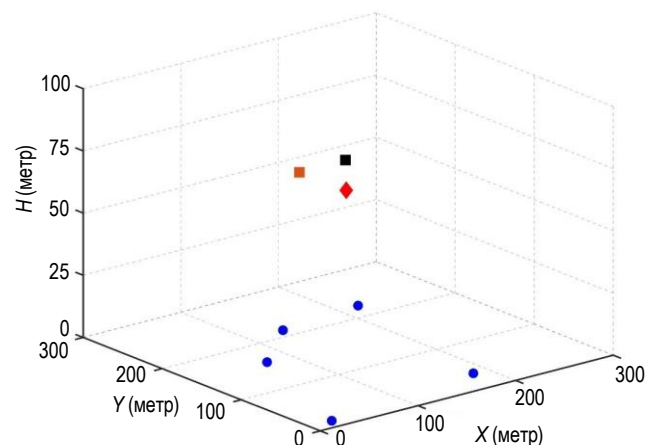


Рис. 1. 3D-положение ABS и пользователей

Fig. 1. 3D-Location of Aerial Base Station (ABS) and Users

На рисунке 1 показано 3D-расположение ABS в различных случаях. Красный ромб отображает позицию ABS при оптимизации местоположения и распределения мощности. Эта точка имеет координаты (140, 140,5, 65). Оранжевый квадрат представляет положение ABS при использовании оптимальной позиции, равномерно распределяющей мощность между пользователями. Эта точка имеет координаты (149, 151, 74). Черные квадраты представляют любое положение ABS и отображают распре-

деление мощности между пользователями. Как видим, высота ABS при оптимизации только местоположения будет выше, чем высота полета ABS при оптимизации местоположения и распределения мощности. Это связано с тем, что при одинаковом уровне распределения мощности среди пользователей ABS будет находиться выше, чтобы уменьшить связь в условиях NLoS в канале «воздух – земля» для удаленных пользователей. Между тем, в случае, когда энергия может быть распределена между пользователями, при наличии удаленных пользователей ABS могут увеличить распределение энергии по сравнению с близлежащими пользователями.

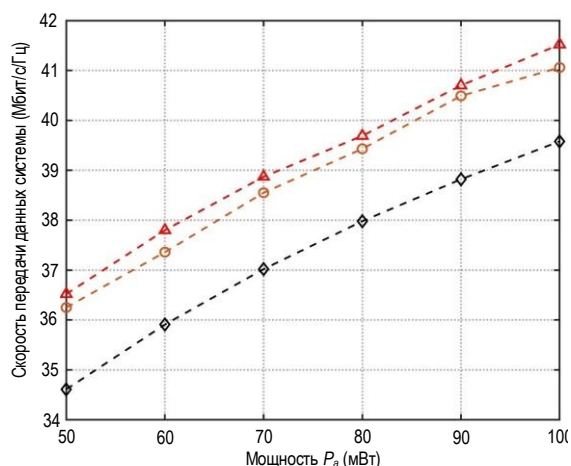


Рис. 2. Скорости передачи данных системы с различной мощностью передачи ABS

Fig. 2. Data Transmission Rate of the System with Different ABS Transmission Power

Скорость передачи данных системы при различной мощности передачи ABS показана на рисунке 2 в зависимости от положения ABS (см. рисунок 1). Как видим, что при оптимизации положения ABS производительность будет выше, чем при случайном формировании положения ABS. При этом скорость передачи данных системы при оптимизации положения ABS имеет примерно такое же значение, как скорость передачи данных системы при оптимизации положения и распределения энергии. Из этого следует, что местоположение играет важную роль в увеличении показателя скорости.

На рисунке 3 показана средняя суммарная скорость передачи данных системы при различном количестве и распределении пользователей. В целом, с ростом их количества уменьшается средняя суммарная скорость передачи данных в системе. Это можно объяснить тем, что по мере увеличения количества пользователей потери при передаче данных будут возрастать вследствие большего количества подключений пользователей к ABS.

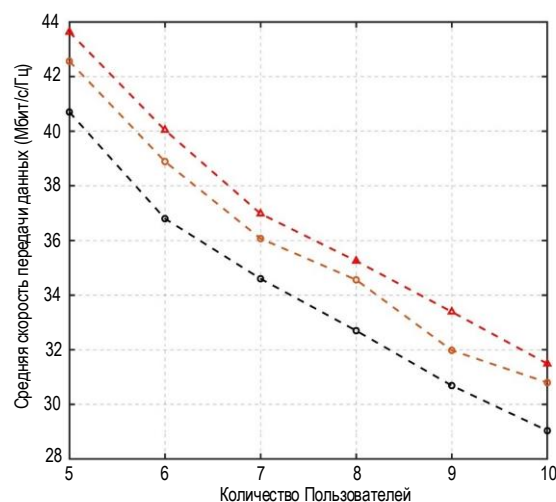


Рис. 3. Средняя суммарная скорость передачи данных системы при различном числе и местоположении наземных пользователей

Fig. 3. Average Data Transmission Rate of the System with Different Number and Locations of Ground Users

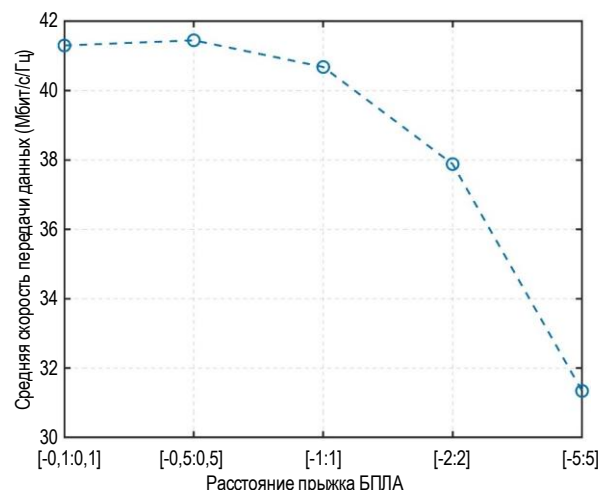


Рис. 4. Скорость передачи данных системы при различных расстояниях перелета ABS

Fig. 4. Data Transfer Rate of the System at Different ABS Flight Distances

Средняя скорость передачи данных внешней системы зависит от расстояния после каждого временного шага ABS. Более конкретно [-1:1] означает, что БПЛА может изменять положение максимум на 1 м в каждом направлении в 3D-пространстве. Как показано на рисунке 4, при увеличении расстояния перехода средняя скорость передачи данных системы вначале остается неизменной, а затем резко снижается. Это связано с тем, что при большом расстоянии перелета ABS может легко пропустить глобальную оптимальную точку и вернуться обратно будет очень сложно.

5. Выводы

Во-первых, в статье исследуются проблемы совместной оптимизации размещения ABS и распределения мощности для максимизации скорости передачи данных в системе. В интересах решения

этой проблемы предложено использовать метод обучения с подкреплением *Q-learning* и эпсилон-жадный алгоритм *ε-greedy*.

Во-вторых, совместная оптимизация размещения БПЛА и распределения мощности позволила, в отличие от известных результатов, установить, что высота ABS при оптимизации только местоположения будет выше, чем высота ABS при сов-

местной оптимизации местоположения и распределения мощности.

В-третьих, результаты моделирования показали также, что использование метода обучения с подкреплением *Q-learning* и эпсилон-жадного алгоритма *ε-greedy* для совместной оптимизации обеспечивает более высокую общую скорость передачи данных в системе по сравнению с оптимизацией только местоположения или распределения мощности.

Список источников

1. Ding G., Wu Q., Zhang L., Lin Y., Tsiftsis T.A., Yao Y.D. An amateur drone surveillance system based on the cognitive Internet of Things // *IEEE Communications Magazine*. 2018. Vol. 56. Iss. 1. PP. 29–35. DOI:10.1109/MCOM.2017.1700452. EDN:YBEOMH
2. Rose K., Eldridge S., Chapin L. The internet of things: An overview // *The internet society (ISOC)*. 2015. Vol. 80(15). PP. 1–53.
3. Zhao N., Lu W., Sheng M., Chen Y., Tang J., Yu F.R., et al. UAV-Assisted Emergency Networks in Disasters // *IEEE Wireless Communications*. 2019. Vol. 26. Iss. 1. PP. 45–51. DOI:10.1109/MWC.2018.1800160
4. Li B., Fei Z., Zhang Y. UAV communications for 5G and beyond: Recent advances and future trends // *IEEE Internet of Things Journal*. 2018. Vol. 6. Iss. 2. PP. 2241–2263. DOI:10.1109/JIOT.2018.2887086. EDN:UPZMGQ
5. Shannon C.E. A mathematical theory of communication // *The Bell System Technical Journal*. 1948. Vol. 27. Iss. 3. PP. 379–423. DOI:10.1002/j.1538-7305.1948.tb01338.x
6. Sutton R.S. *Reinforcement learning: An introduction*. The MIT Press, 2018.
7. Jaakkola T., Jordan M., Singh S. Convergence of Stochastic Iterative Dynamic Programming Algorithms // *Advances in Neural Information Processing Systems 6 (NIPS 1993)*. 1993.

References

1. Ding G., Wu Q., Zhang L., Lin Y., Tsiftsis T.A., Yao Y.D. An amateur drone surveillance system based on the cognitive Internet of Things. *IEEE Communications Magazine*. 2018;56(1):29–35. DOI:10.1109/MCOM.2017.1700452. EDN:YBEOMH
2. Rose K., Eldridge S., Chapin L. The internet of things: An overview. *The internet society (ISOC)*. 2015;80(15):1–53.
3. Zhao N., Lu W., Sheng M., Chen Y., Tang J., Yu F.R., et al. UAV-Assisted Emergency Networks in Disasters. *IEEE Wireless Communications*. 2019;26(1):45–51. DOI:10.1109/MWC.2018.1800160
4. Li B., Fei Z., Zhang Y. UAV communications for 5G and beyond: Recent advances and future trends. *IEEE Internet of Things Journal*. 2018;6(2):2241–2263. DOI:10.1109/JIOT.2018.2887086. EDN:UPZMGQ
5. Shannon C.E. A mathematical theory of communication. *The Bell System Technical Journal*. 1948;27(3):379–423. DOI:10.1002/j.1538-7305.1948.tb01338.x
6. Sutton R.S. *Reinforcement learning: An introduction*. The MIT Press; 2018.
7. Jaakkola T., Jordan M., Singh S. Convergence of Stochastic Iterative Dynamic Programming Algorithms. *Advances in Neural Information Processing Systems 6 (NIPS 1993)*. 1993.

Статья поступила в редакцию 13.01.2025; одобрена после рецензирования 15.02.2025; принята к публикации 21.02.2025.

The article was submitted 13.01.2025; approved after reviewing 15.02.2025; accepted for publication 21.02.2025.

Информация об авторах:

**ЧАН
Тунг Зыонг**

аспирант кафедры сетей связи и передачи данных Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича
 <https://orcid.org/0009-0006-0080-9477>

**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.

Koucheryavy A.E. has been a member of the journal "Proceedings of Telecommunication Universities" Editorial Council since 2016, but has nothing to do with the decision to publish this article. The article has passed the review procedure accepted in the journal. The authors have not declared any other conflicts of interest.