

doi: 10.17586/2226-1494-2026-26-2-410-419

УДК 519.67

Генерация пространственно-временных рядов сетевой нагрузки в задачах граничных вычислений

Иван Викторович Филянин¹✉, Александр Александрович Капитонов²,
Алексей Петрович Мартынюк³

^{1,2,3} Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация

² Новый университет Узбекистана, Ташкент, 100000, Узбекистан

³ АО «ТБанк», Москва, 127287, Российская Федерация

¹ adeptvin1@gmail.com✉, <https://orcid.org/0000-0002-0536-2807>

² kap2fox@gmail.com, <https://orcid.org/0000-0001-5517-3038>

³ MartynyukAlexey05@gmail.com, <https://orcid.org/0009-0009-8697-7739>

Аннотация

Введение. Исследование систем принятия решений в задачах граничных вычислений с множественным доступом зачастую основывается на абстрактном представлении сети связи, не имеющей профилей сетевой нагрузки. Целью работы являлась разработка инструментов генерации пространственно-временных данных сетевой нагрузки в зависимости от архитектуры сети связи. **Метод.** Применены методы стохастической геометрии и статистические данные для формирования профиля возможной нагрузки. Для оценки работы методов стохастической геометрии разработан инструмент генерации и валидации пространственно-временных рядов с поиском паттернов из открытой базы вышек сотовой связи OpenCellID. **Основные результаты.** Выполнен анализ научных работ и открытых наборов данных расположения и нагрузки вышек сотовой связи. На основе анализа сделан вывод о низком качестве данных в задачах обучения систем принятия решений размещения вычислительных сервисов в геораспределенных узлах обработки данных. Разработан инструмент генерации и валидации пространственно-временных рядов с поиском паттернов из открытой базы вышек сотовой связи OpenCellID. Сравнительный анализ базового и калиброванного алгоритмов Hard-Core Point Process показал существенные различия в характеристиках генерируемых распределений. Для города Санкт-Петербурга калиброванная модель обеспечила увеличение плотности размещения станций в 99 раз, сокращение межстанционных расстояний в 52 раза при эффективной площади покрытия 0,04 км². В случае города Новосибирска наблюдались аналогичные тенденции с меньшей интенсивностью: увеличение плотности в 12,5 раз, сокращение расстояний в 21 раз при площади покрытия 0,32 км². **Обсуждение.** Использование пространственно-временных рядов, полученных с помощью разработанного инструмента генерации, позволит повысить точность алгоритмов размещения вычислительных сервисов и снизить задержки в системах граничных вычислений за счет предобучения на данных коррелирующих с реальным расположением вышек сотовой связи. С помощью предложенного инструмента генерации можно задать координаты местности предполагаемой сети связи, что окажет влияние на паттерны распределения вышек и позволит сгенерировать более точные пространственно-временные ряды.

Ключевые слова

граничные вычисления с множественным доступом, стохастическая геометрия, OpenCellID, пространственно-временные ряды, Hard-Core Poisson Process

Ссылка для цитирования: Филянин И.В., Капитонов А.А., Мартынюк А.П. Генерация пространственно-временных рядов сетевой нагрузки в задачах граничных вычислений // Научно-технический вестник информационных технологий, механики и оптики. 2026. Т. 26, № 2. С. 410–419. doi: 10.17586/2226-1494-2026-26-2-410-419

Generating spatiotemporal network load series in multi-access edge computing tasks using open data

Ivan V. Filianin¹, Aleksandr A. Kapitonov², Alexey P. Martynyuk³

^{1,2,3} ITMO University, Saint Petersburg, 197101, Russian Federation

² New University, Tashkent, 100000, Uzbekistan

³ “ТBank”, JSC, Moscow, 127287, Russian Federation

¹ adeptvin1@gmail.com, <https://orcid.org/0000-0002-0536-2807>

² kap2fox@gmail.com, <https://orcid.org/0000-0001-5517-3038>

³ martynyukalexey05@gmail.com, <https://orcid.org/0009-0009-8697-7739>

Abstract

Research into decision-making systems in multi-access edge computing systems is often based on an abstract representation of a communication network without network load profiles. The aim of this work was to develop tools for generating spatio-temporal network load data depending on the communication network architecture. In our work, we used stochastic geometry methods and statistical data to form a profile of possible load. To evaluate the performance of stochastic geometry methods, we developed a tool for generating and validating spatio-temporal series with pattern search from the OpenCellID open database of cell towers. During the work, an analysis of literature and public datasets on the location and load of cell towers was conducted. Based on the analysis, it was concluded that the data quality was low for the purposes of training decision-making systems for the placement of computing services in geographically distributed data processing nodes. A tool was also developed to generate and validate spatio-temporal series with pattern search from the OpenCellID open database of cell towers. A comparative analysis of the basic and calibrated Hard-Core Poisson Process algorithms showed significant differences in the characteristics of the generated distributions. For St. Petersburg, the calibrated model provided a 99-fold increase in station density and a 52-fold reduction in inter-station distances with an effective coverage area of 0.04 km². In the case of Novosibirsk, similar trends were observed with less intensity: a 12.5-fold increase in density and a 21-fold reduction in distances with a coverage area of 0.32 km². The use of spatio-temporal series obtained with the help of the developed generation tools will improve the quality of training decision-making systems for the placement of computing services through pre-training on data correlated with the actual location of cell towers. In addition, the generation tool allows you to specify the coordinates of the area of the proposed communication network which can also affect the distribution patterns of towers and which in turn will allow you to generate more accurate spatio-temporal series.

Keywords

multi-access edge computing, stochastic geometry, OpenCellID, spatiotemporal, Hard-Core Poisson Process

For citation: Filianin I.V., Kapitonov A.A., Martynyuk A.P. Generating spatiotemporal network load series in multi-access edge computing tasks using open data. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2026, vol. 26, no. 2, pp. 410–419 (in Russian). doi: 10.17586/2226-1494-2026-26-2-410-419

Введение

Исследование систем принятия решений для платформ граничных вычислений с множественным доступом сталкивается со сложностями моделирования распределения клиентского трафика между узлами сети связи во времени, что критически важно для предобучения систем и снижения эксплуатационных издержек.

Актуальность данного исследования обусловлена стремительным ростом числа устройств интернета вещей и приложений реального времени, которые требуют низкой сетевой задержки и высокой пропускной способности. По прогнозам аналитиков к 2030 году количество подключенных устройств превысит 25,4 млрд [1], что создаст беспрецедентную нагрузку на сетевую инфраструктуру. В условиях такого масштабирования традиционные подходы к планированию и управлению ресурсами систем Multi-Access Edge Computing (МЕС) становятся неэффективными.

Существующие методы генерации синтетических данных имеют ограничения: не учитывают пространственную корреляцию между базовыми станциями, упрощенно моделируют временные зависимости и не адаптируются к различным архитектурам сетей.

Практическая значимость работы заключается в возможности создания высококачественных обучаю-

щих данных для систем автоматического управления ресурсами МЕС без необходимости длительного сбора реальных метрик сети. Кроме того, разработанный инструмент генерации может быть использован для тестирования устойчивости систем принятия решений к различным сценариям нагрузки, включая аномальные и экстремальные ситуации.

Анализ открытых наборов данных и научных публикаций

Внедрение МЕС-платформ позволяет снизить задержки на 37,5 % по сравнению с облачными решениями [2], однако эффективность систем принятия решений критически зависит от качества обучающих данных, при этом большинство существующих работ [3, 4] фокусируются на алгоритмах обработки данных, оставляя без внимания методы создания репрезентативных обучающих выборок.

В работе [5] выполнено исследование проблем размещения вычислительных сервисов и рассмотрены связанные с этим решения. Также в [5] исследован вопрос качества данных при обучении систем принятия решений и влияние данных на конечный результат обучения.

В [6] проведен обзор научных публикаций, в которых использовали синтетические данные, в [7] —

данные из частных источников, в [8] — источников от операторов связи.

Рассмотрим наборы открытых наборов данных, как самый простой вариант, не требующий специфичных знаний и данных.

IBM имеет следующие поля: дата, идентификатор события, статус события, идентификатор пользователя, идентификатор сектора, технология передачи данных, модель устройства и IMEI. К сожалению, воспользоваться им не представляется возможным, так как данные представлены за 1 час 30 минут 2 февраля 2020 года. Данной выборки не хватит для формирования правильных корреляций дневного трафика.

Анализ набора данных Edge Computing Dataset¹ показал, что данные организованы в соответствии с принципом суточной сегментации, где каждый файл содержит полную информацию о сетевой активности за календарные сутки.

В наборе данных Edge Computing Dataset (табл. 1) исследованы характеристики городов Пекин, Гуанчжоу и Шанхай, что обеспечивает географическое разнообразие и различные паттерны пользовательской активности.

Набор данных города Гуанчжоу был выбран как наиболее репрезентативный по временному покрытию (75 %) и объему данных. Однако детальный анализ выявил хаотичное распределение сотовой связи с нереалистично низкой нагрузкой: для города с населением 18,8 млн человек отдельные вышки обслуживали 1–2 пользователей за несколько часов, что делает данные непригодными для обучения систем принятия решений.

Исходя из результатов проведенного анализа был сделан вывод, что использовать данные из приведенных наборов данных не представляется возможным по причине их ограниченности и неполноценности. А это значит, что необходимо рассмотреть механизмы генерации синтетических данных.

На основе анализа научных публикаций выделено три подхода к генерации пространственно-временных рядов: генерация синтетических данных [10], адаптация реальных данных [11] и генерация псевдослучайных данных [12]. Синтетические данные обеспечивают полный контроль параметров и воспроизводимость, но обладают недостаточной репрезентативностью реальных паттернов. Адаптация реальных данных сохраняет естественные корреляции, но ограничена доступностью

качественных открытых наборов данных и проблемами конфиденциальности. Псевдослучайная генерация проста в реализации, но не отражает реалистичных пространственно-временных корреляций, что может привести к завышенным результатам в тестовом окружении.

Для гибкого обучения систем принятия решений выбран подход генерации синтетических данных, приближенных к реальным. Реализация этого подхода требует предварительного моделирования сети связи и ее характеристик с использованием методов стохастической геометрии, методов точечного моделирования с отталкиванием и методов временного моделирования.

Рассмотрим методы стохастической геометрии и методы точечного моделирования с отталкиванием, позволяющие рассчитать в пространстве расположение усовершенствованных базовых станций сотовой связи (Evolved Node B, eNB) и пользовательское оборудование (User Equipment, UE). В исследованиях, изучающих использование методов стохастической геометрии в Multi-Access Edge Computing платформах в основном рассматривается применение алгоритмов в качестве инструмента расположения UE в пространстве, позволяющие рассчитать расстояние между UE и eNB, что дает возможность рассчитать соотношение сигнал шум (Signal Interference Noise Ratio) в статических моделях.

Например, в [13] рассматривалась модель сети MEC с поддержкой Graph-based Filtered Sparse Code Multiple Access (GF-SCMA). Алгоритм стохастической геометрии (Hard-Core Poisson Process, HCPP) использовался в качестве функции распределения eNB и UE в пространстве и их взаимосвязь. Применение данного метода позволило рассчитать преимущество GF-SCMA над ортогональным множественным доступом с частотным разделением (Orthogonal Multiple Access).

В работе [14] процесс Пуассона (Poisson Point Process, PPP) использовался для расчета распределения базовых станций (Base Station, BSs) и UEs/MEs (Mobile Users, MEs) в пространстве, на основе которого рассчитывалась модель беспроводной сети.

В результате возможно сделать вывод, что методы стохастической геометрии и точечного моделирования с отталкиванием можно использовать для решения поставленной задачи.

Методы временного моделирования позволяют создавать ряды сетевой нагрузки, отражающие временные закономерности (суточные, недельные, сезонные колебания). Среди распространенных подходов: аналитические модели с периодическими функциями, авторегрессионные процессы, марковские цепи и имитация на основе эмпирических данных. Использование

¹ Edge Computing Dataset [Электронный ресурс] // GitHub. URL: <https://github.com/BuptMecMigration/Edge-Computing-Dataset/tree/master> (дата обращения: 11.07.2025).

Таблица 1. Характеристики набора данных Edge Computing Dataset

Table 1. Characteristics of the Edge Computing Dataset

Город	Временное покрытие, ч	Покрытие, %	Объем, записей/день	Основные пропуски
Гуанчжоу	18/24	75	23 251	01:00-06:00
Шанхай	13/24	54	6048	01:00-07:00, 09:00-10:00, 16:00-17:00
Пекин	11/24	46	5555	01:00-09:00, 16:00-17:00

комбинации методов стохастической геометрии и временного моделирования позволяет создавать синтетические данные, близкие к реальным в пространственном и временном аспектах, формируя репрезентативные обучающие выборки для систем принятия решений в МЕС-платформах.

В качестве подхода временного моделирования была выбрана имитация на основе эмпирических данных. Результаты данного метода представлены в работе [15], где рассмотрено распределение нагрузки в Лондоне в течение 24 ч, которое использовано как часть системы генерации данных.

Для повышения реалистичности моделирования применены данные проекта OpenCellID¹ — открытой базы координат базовых станций. В отличие от рассмотренных наборов данных, OpenCellID предоставляет информацию о географическом расположении вышек связи, их идентификаторах, радиусе покрытия и числе наблюдений, что позволяет учесть реальную топологию сети при генерации синтетических данных.

При отборе данных основными критериями являлись соответствие региону по координатам и минимальное количество наблюдений для исключения ошибочных записей. Использование OpenCellID позволяет построить реалистичную карту базовых станций с учетом плотности размещения инфраструктуры и зон обслуживания в конкретных географических регионах.

Цель работы — сокращение средней сетевой задержки за счет размещения вычислительных сервисов в геораспределенных узлах обработки данных путем генерации пространственно-временных рядов для обучения системы принятия решений.

В качестве задач, определенных в рамках работы, можно выделить следующие: разработать математическую модель сети сотовой связи с использованием методов стохастической геометрии; разработать инструмент генерации пространственно-временных рядов сетевой нагрузки; провести эксперимент по генерации пространственно-временных рядов для нескольких вариаций сетей связи; выполнить сравнительный анализ и оценку сгенерированных данных и данных из открытых источников.

¹ OpenCellID — крупнейшая открытая база данных вышек сотовой связи [Электронный ресурс]. URL: <https://opencellid.org/> (дата обращения: 11.07.2025).

Методы генерации пространственно-временных рядов сетевой нагрузки

В настоящей работе не будем применять модель расчета параметров физического слоя, так как конкретная реализация передачи данных очень сильно зависит от использования конкретного сетевого оборудования. Предлагается также применять алгоритмы стохастической геометрии для физического размещения BSs и UEs, но в другой парадигме. Можно создать распределение UEs для каждой BS на определенной площади в двухмерном пространстве. Это позволит сформировать данные для проведения тренировки моделей машинного обучения приближенные к реальным.

Для моделирования распределения точек в пространстве могут использоваться различные методы стохастической геометрии, например PPP, который генерирует точки независимо и равномерно по площади. Однако он не учитывает ограничения на минимальное расстояние между точками.

Для учета отталкивания между точками применяются, например, детерминантный точечный процесс (Determinantal Point Process, DPP), где точки взаимно «отталкиваются» [16], или процесс Жинибра (Ginibre Point Process, GPP), основанный на случайных матрицах, также демонстрирующий сильное отталкивание [17]. Его обобщением является β -Ginibre Point Process (β -GPP), в котором степень отталкивания регулируется параметром β .

Если требуется жесткое минимальное расстояние между точками, используются процессы с жестким ядром, такие как процесс Матерна (Matérn Hard-Core Process, МНСП) и пуассоновский процесс с жестким ядром (НСРП) [18].

Эти модели позволяют гибко подстраивать распределение под реальные сценарии — от хаотичных (PPP) до регулярных с явно выраженным отталкиванием (DPP, GPP) или строгими минимальными расстояниями (МНСП, НСРП).

В табл. 2 приведен сравнительный анализ алгоритмов стохастической геометрии в задачах распределения BSs и UEs.

Основываясь на результатах сравнительного анализа (табл. 2), можно сделать вывод, что предпочтительным решением является использование алгоритма

Таблица 2. Сравнительный анализ алгоритмов стохастической геометрии
Table 2. Comparative analysis of stochastic geometry algorithms

Характеристика	GPP	β -GPP	МНСП	НСРП
Тип отталкивания	Мягкое, сильное отталкивание	Регулируемое отталкивание между PPP и GPP	Жесткое (hard-core), минимальное расстояние между точками	Жесткое, минимальное расстояние между точками
Математическая структура	Частный случай DPP с конкретным ядром	Разреженный и масштабированный GPP с параметром β	Основан на прореживании PPP с правилом удаления точек	Основан на прореживании PPP с правилом удаления точек
Сложность симуляции	Средняя, использует спектральные методы	Средняя, включает прореживание и масштабирование	Низкая, основана на простых правилах прореживания	Низкая, основана на простых правилах прореживания

НСРР. Такой подход позволяет учесть ключевое ограничение — минимальное расстояние между точками (например, между BS или UE), при этом оставаясь достаточно простым и интуитивно понятным. НСРР сохраняет хаотичность распределения, характерную для реальных сетей, но исключает нереалистично близкие точки, что важно для моделирования сценариев с физическими ограничениями на размещение оборудования.

Классический НСРР обеспечивает только жесткое ограничение на минимальное расстояние между базовыми станциями, что соответствует физическим требованиям размещения BS.

Таким образом, выбор НСРР представляет собой компромисс между реализмом, вычислительной сложностью и возможностью аналитического анализа. Рассмотрим НСРР с точки зрения математики [19].

Математически, конфигурация точек $X = \{x_1, x_2, \dots, x_n\}$ называется *hard-core* конфигурацией с параметром $r > 0$, если для любых двух различных точек $x_i, x_j \in X$ выполняется условие $\|x_i - x_j\| \geq r$, где r — минимальное допустимое расстояние между точками.

Плотность вероятности для конфигурации точек в НСРР определяется как произведение функций взаимодействия между точками:

$$p(x_n | x_1, \dots, x_{n-1}) = Z \prod_{i=1}^{n-1} h(x_i - x_n),$$

где $h(x_i - x_j)$ — функция взаимодействия между точками; Z — нормализующая константа.

Функция взаимодействия для *hard-core* процесса определяется как:

$$h(x_i - x_j) = \begin{cases} 0, & \|x_i - x_j\| < r \\ e^{-\beta \|x_i - x_j\|}, & \|x_i - x_j\| \geq r \end{cases},$$

где $\beta \geq 0$ — параметр, контролирующий степень отталкивания между точками на расстояниях больше r .

Для генерации выборок из НСРР используется алгоритм последовательной генерации. Алгоритм начинает с пустой конфигурации и на каждой итерации предлагает новую точку, которая принимается или отвергается на основе вероятности принятия. Ключевым условием является проверка *hard-core* ограничения — если предлагаемая точка находится слишком близко к существующим точкам, она автоматически отвергается.

После генерации координат базовых станций строится диаграмма Вороного для определения зон покрытия каждой станции. Зона покрытия станции s_i определяется как множество точек, для которых данная станция является ближайшей.

Временная динамика сетевой нагрузки моделируется на основе типичных дневных паттернов активности пользователей. Базовый профиль нагрузки определяется как вектор коэффициентов для каждого часа суток. Для создания более реалистичных временных рядов в базовый профиль вносится случайная девиация:

$$N(t) = \lfloor I_{t/3600}(1 + \varepsilon(t)N_{\max}) \rfloor,$$

где $I_{t/3600}$ — коэффициент нагрузки для текущего часа суток из базового профиля нагрузки (вектор значений от

0 до 1, каждый элемент которого соответствует доле активных пользователей в данный час); $N(t)$ — количество активных пользователей в момент времени t ; N_{\max} — максимальное количество пользователей в системе; $\varepsilon(t)$ — случайная девиация, равномерно распределенная в интервале $[-\delta, \delta]$, δ — максимально допустимое отклонение нагрузки от базового профиля (в данной работе $\delta = 0,1$, что соответствует девиации $\pm 10\%$).

Временные ряды генерируются с фиксированным шагом дискретизации $\Delta t = 10$ с, что обеспечивает детальное временное разрешение для анализа кратковременных флуктуаций нагрузки.

Интеграция пространственного и временного моделирования осуществляется следующим образом. Для каждого момента времени t_k генерируется пространственное распределение $N(t_k)$ пользователей с использованием одного из трех методов: равномерное распределение по всей области, нормальное распределение относительно центра области, или кластерное распределение вокруг базовых станций. Каждый пользователь назначается ближайшей базовой станции, что формирует распределение нагрузки по станциям.

Разработка инструмента генерации пространственно-временных рядов на основе паттернов с реальных данных

Практическая реализация выбранного метода НСРР требует предварительного анализа статистических характеристик реальных сетевых топологий. Для этого необходимо извлечь ключевые паттерны распределения вышек сотовой связи на участке местности из базы данных OpenCellID. Можно выделить следующие основные паттерны: плотность размещения станций [20], расстояние между соседними станциями [21], и степень кластеризации [22].

Расчет каждого из паттернов осуществляется следующим образом. Плотность определяется как $\lambda = \frac{N}{A}$, где N — количество станций; A — площадь области покрытия. Расстояние между соседями можно представить как $s = \frac{1}{N} \sum_{i=1}^N \min \{\|x_i - x_j\| : j \neq i\}$. Кластеризация определяется как $N_\varepsilon(p) = \{q \in D : \text{dist}(p, q) \leq \varepsilon\}$, где ε — радиус окрестности.

На основе описанных методов был разработан инструмент генерации пространственно-временных рядов, блок-схема которого изображена на рис. 1.

На первом этапе для получения репрезентативных данных был проведен анализ базовых станций LTE (Long-Term Evolution) в нескольких географических регионах с различной плотностью застройки из данных OpenCellID. Выбор LTE обусловлен его широким распространением и релевантностью для современных МЕС-платформ.

Процедура анализа включала несколько последовательных этапов. Сначала осуществлялась загрузка пространственных данных для заданного региона через OpenCellID в формате JSON с последующей фильтрацией по типу сети (LTE) и достоверности координат. Затем производился расчет пространственных харак-



Рис. 1. Блок-схема инструмента генерации пространственно-временных рядов

Fig. 1. Block diagram of the operation of the tool for generating space-time series

теристик согласно описанным методикам в разделе «Анализ открытых наборов данных и научных публикаций» и выполнялся статистический анализ полученных паттернов.

Следующим этапом являлась генерация базового и калиброванного НСРР и оценка распределения, используя методы визуального анализа и статистическую оценку. На основе данных распределения BSs и UEs калиброванного НСРР было сформировано пространственно-временное распределение пользователей в течение дня на основе статистики в работе [15], рассмотренной в разделе «Анализ открытых наборов данных и научных публикаций», и сгенерирован график с 00:00 до 23:59 с шагом в 10 с.

Для каждой базовой станции и каждого временного интервала формируется профиль нагрузки, включающий количество одновременных пользователей, интервал между запросами и длительность профиля. Эти профили сохраняются в формате JSON для последующего использования в процессе обучения систем принятия решений.

Разработанная система¹ поддерживает различные временные интервалы (6, 12, 24 ч) и генерирует выходные данные в формате JSON, совместимом с системой тестирования LWMECPS. Качество генерируемых данных оценивается по критериям пространственной корректности, временной консистентности и статистической репрезентативности.

Данный подход позволяет создавать синтетические наборы данных сетевой нагрузки, которые сохраняют ключевые пространственно-временные корреляции реальных сетей МЕС и могут быть использованы для предварительного обучения систем принятия решений.

Генерация пространственно-временных рядов на основе паттернов с реальными данными

Разработанный инструмент генерации пространственно-временных рядов на основе паттернов реаль-

ных данных требует экспериментальной валидации для оценки его эффективности и адаптивности к различным условиям размещения базовых станций. Для экспериментальной оценки работы разработанного инструмента генерации пространственно-временных рядов был проведен сравнительный анализ работы алгоритма НСРР на данных двух городов с различными характеристиками плотности застройки и распределения базовых станций.

В качестве тестовых регионов были выбраны город Санкт-Петербург, координаты: $59^{\circ}47'60''$ – $60^{\circ}06'00''$ с. ш., $30^{\circ}00'00''$ – $30^{\circ}36'00''$ в. д. как крупный мегаполис с высокой плотностью базовых станций и город Новосибирск с координатами $54^{\circ}57'18''$ – $55^{\circ}12'00''$ с. ш., $82^{\circ}47'60''$ – $83^{\circ}05'60''$ в. д. в качестве представителя регионального центра с умеренной плотностью покрытия.

Выбор данных городов обусловлен их различными характеристиками урбанизации, что позволяет оценить адаптивность алгоритма к различным пространственным паттернам размещения вышек сотовой связи.

Экспериментальное исследование проводилось по следующей методологии. На первом этапе осуществлялось извлечение координат базовых станций из базы данных OpenCellID для выбранных регионов с последующим расчетом статистических паттернов, включающих плотность размещения станций, межстанционные расстояния и параметры кластеризации. Далее выполнялась калибровка параметров алгоритма НСРР на основе извлеченных паттернов и генерация синтетических данных с использованием как калиброванного, так и базового вариантов алгоритма. Результаты визуализировались посредством диаграмм Вороного с последующим анализом распределений, что позволило провести сравнительную оценку качества генерации синтетических данных.

Базовый НСРР представляет собой НСРР с расширенными взаимодействиями, который генерирует точки в пространстве с учетом жесткого ограничения на минимальное расстояние между точками, используя фиксированные параметры из конфигурации для универсального применения.

Калиброванный НСРР является адаптированной версией базового НСРР, параметры которой (минималь-

¹ lwmeeps-testapp [Электронный ресурс] // GitHub. URL: <https://github.com/adeptvin1/lwmeeps-testapp> (дата обращения: 11.07.2025).

ное расстояние, интенсивность и коэффициент взаимодействия) определяются на основе анализа реальных данных OpenCellID для конкретного региона, обеспечивая более точное соответствие реальным паттернам размещения базовых станций в городской среде.

На основе описанной методологии были получены экспериментальные данные (табл. 3).

Основываясь на данных из табл. 3, можно сделать вывод, что высокая интенсивность вышек сотовой связи в Санкт-Петербурге влияет на эффективную площадь покрытия в калиброванном НСРР. Эффективная площадь покрытия, калиброванного НСРР меньше в 100 раз по сравнению с базовым НСРР. При этом плотность вышек в Новосибирске в 8 раз меньше, чем в Санкт-Петербурге, и это также влияет на эффективную площадь покрытия в калиброванном НСРР. Эффективная площадь покрытия, калиброванного НСРР меньше в 12,5 раз по сравнению с базовым НСРР.

Для проведения визуального анализа покрытия вышек сотовой связи было сгенерировано распределение BSs и UEs в пространстве для базового и калиброванного НСРР для города Санкт-Петербург.

На рис. 2 изображены распределения BS с использованием диаграмм Вороного и НСРР на участке в 2×2 км.

На рис. 2 заметна разница в площади покрытия, отображенная в табл. 3, а также изменение контуров зоны покрытия сети связи и пользователей. Калиброванная версия обеспечивает более реалистичное распределение пользователей относительно BS, избегая чрезмерной концентрации нагрузки на отдельных узлах, характерной для базового НСРР. Это, в свою очередь, будет влиять на пространственно-временной ряд.

Для генерации временной части было использовано следующее распределение, в зависимости от времени, полученное в работе [15], где каждое значение это процент от наибольшего числа пользователей. В результате было

Таблица 3. Сравнительный анализ алгоритмов стохастической геометрии
Table 3. Comparative analysis of stochastic geometry algorithms

Параметр	Базовый НСРР	Калиброванный НСРР	
		Санкт-Петербург	Новосибирск
Количество базовых станций	4	4	4
Плотность λ (число станций на км ²)	1	99,21	12,31
Эффективная площадь покрытия, км ²	4	0,04	0,32
Минимальное расстояние, км	1	0,019	0,047

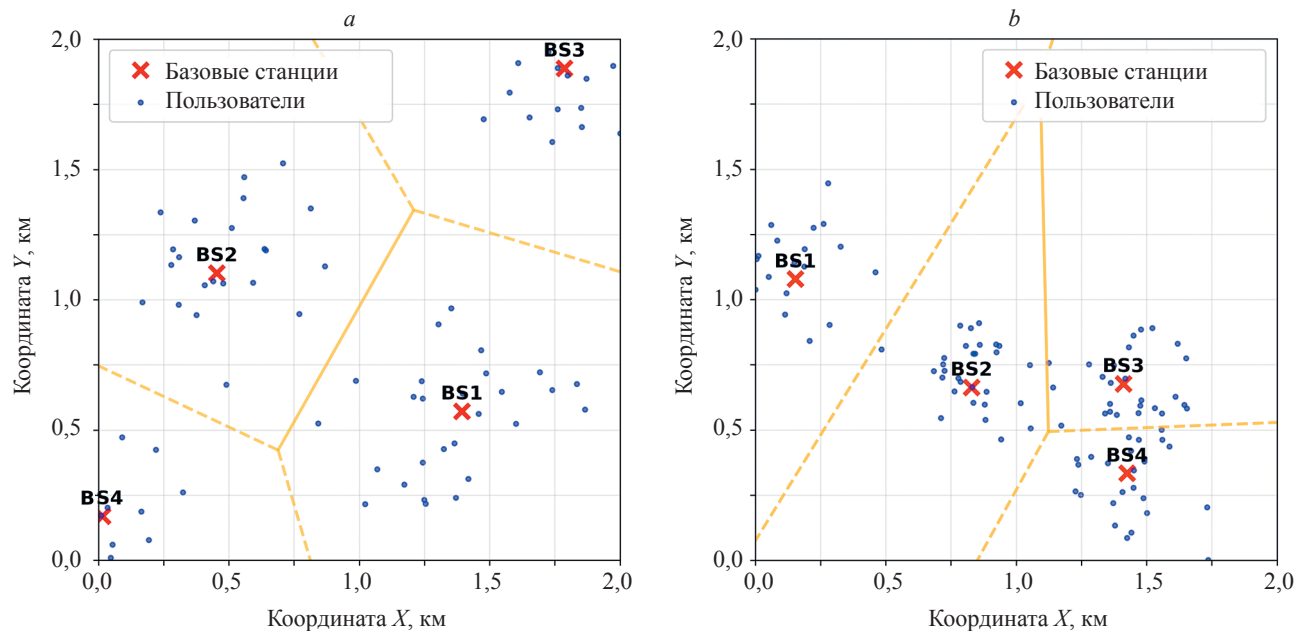


Рис. 2. Распределение базовых станций (BS) с использованием диаграмм Вороного, а также базового (a) и калиброванного (b) Hard-Core Poisson Process.

Красные крестики — базовые станции, синие точки — пользователи, желтые сплошные линии — границы зон покрытия (Вороного) внутри исследуемой области, желтые пунктирные линии — продолжение границ зон покрытия за пределами исследуемой области

Fig. 2. Distribution of base stations (BS) using Voronoi diagrams as well as the basic (a) and calibrated (b) Hard-Core Poisson Process.

Red crosses — base stations, blue dots — users (UE), yellow solid lines — coverage zone boundaries (Voronoi) within the study area, yellow dashed lines — extension of coverage zone boundaries outside the study area

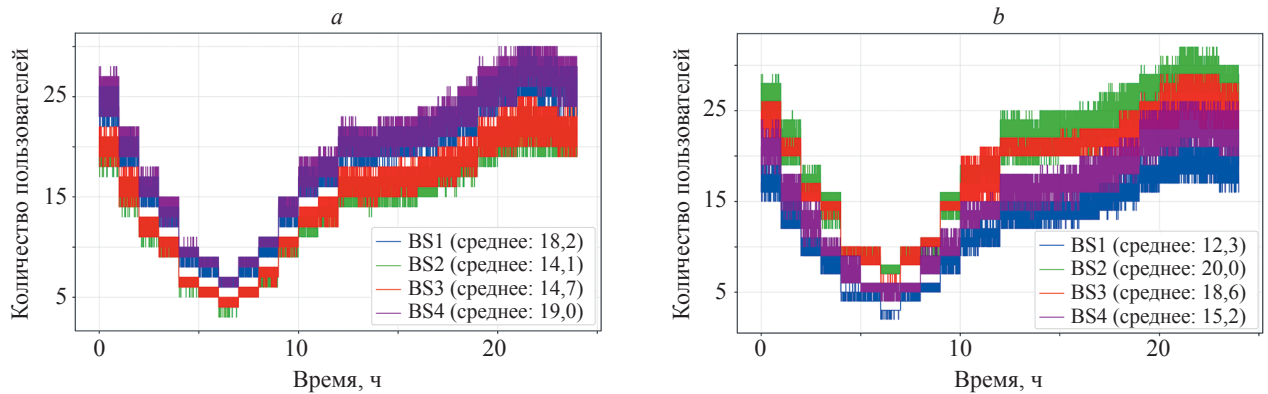


Рис. 3. Диаграммы распределения UEs и BS в течение одних суток для базового (a) и калиброванного (b) Hard-Core Poisson Process

Fig. 3. Distribution diagrams of UEs and BS over one day for the basic (a) and optimized (b) Hard-Core Poisson Process

выбрано 100 пользователей, для более точного распределения добавлена девиация в размере 10 % для каждого временного участка и шагом в 10 с между значениями.

На рис. 3 показаны диаграммы распределения UEs для каждой BS в течение одних суток.

Визуальный анализ рис. 3 также показал более равномерное распределение пользователей по вышкам связи для калиброванного HCSP. Используемые методы и инструмент генерации позволяют сделать вывод, что калиброванный HCSP лучше подходит для генерации пространственно-временных рядов. Отметим, что сравнение полученных данных с реальными не было выполнено из-за качества связи реальными данными. Но при этом можно сделать предположение, что на выборках больших, чем четыре вышки и при фиксированной площади покрытия рассматриваемые методы покажут хорошие результаты.

Предложенное решение может использоваться в качестве исходных данных для обучения систем принятия решений в задачах граничных вычислений и повысить качество принимаемых решений за счет использования реалистичных пространственных конфигураций BS, учитывающих реальные паттерны размещения в городской среде.

Обсуждение

Полученные результаты демонстрируют потенциал применения калиброванных методов стохастической геометрии для генерации пространственно-временных рядов сетевой нагрузки в системах граничных вычислений с множественным доступом.

Основная новизна метода заключается в интеграции калибровки параметров HCSP на основе реальных данных OpenCellID, адаптации к специфике различных типов урбанизации и учета пространственно-временных корреляций при генерации синтетических данных.

В отличие от существующих подходов с фиксированными параметрами, предложенное решение обеспечивает автоматическую адаптацию к реальным паттернам размещения BS в конкретном географическом регионе, что позволяет воспроизводить не только

статистические характеристики плотности размещения вышек, но и их пространственные корреляции.

Улучшение реалистичности обучающих данных непосредственно влияет на ключевые показатели производительности граничных вычислений: точность предсказания оптимальных местоположений сервисов, минимизацию задержек обслуживания, время отклика приложений и эффективность распределения нагрузки.

Основные ограничения исследования связаны с использованием единого временного паттерна нагрузки и необходимостью валидации на больших сетевых топологиях. Перспективными направлениями являются интеграция методов машинного обучения для автоматического выявления временных паттернов и разработка адаптивных алгоритмов калибровки.

Практическое применение разработанного инструмента может снизить затраты на разработку систем управления ресурсами МЕС за счет улучшения качества предварительного обучения систем принятия решений.

В перспективе представляет интерес рассмотрение вопросов развития методов оценки обеспечения надежности и отказоустойчивости распределенных систем применительно к системам граничных вычислений [23, 24].

Заключение

Выполненный анализ научных работ и открытых наборов данных расположения и нагрузки вышек сотовой связи показал низкое качество данных в задачах обучения систем принятия решений размещения вычислительных сервисов в геораспределенных узлах обработки данных. В результате сравнения методов генерации пространственно-временных рядов был выбран алгоритм Hard-Core Poisson Process в качестве метода генерации.

Сравнительный анализ базового и калиброванного алгоритмов Hard-Core Point Process показал существенные различия в характеристиках генерируемых распределений. Для Санкт-Петербурга калиброванная модель обеспечила увеличение плотности размещения станций в 99 раз, сокращение межстанционных расстояний в 52 раза при эффективной площади покрытия 0,04 км².

В случае Новосибирска наблюдались аналогичные тенденции с меньшей интенсивностью: увеличение плотности в 12,5 раз, сокращение расстояний в 21 раз при площади покрытия 0,32 км². Использование данного инструментария позволяет генерировать пространственно-временные ряды, указывая географические координаты интересующего города без необходимости проведения дополнительных работ.

Полученные результаты демонстрируют способность калиброванного Hard-Core Point Process адаптироваться к специфике городской инфраструктуры различных типов — от высокоплотной застройки мегаполиса до умеренной плотности регионального центра. Различия в коэффициентах изменения параметров отражают особенности урбанизации исследуемых регионов и подтверждают необходимость калибровки модели по реальным данным.

Научная новизна исследования заключается в следующих аспектах:

- разработана методология калибровки алгоритмов стохастической геометрии на основе реальных данных OpenCellID, позволяющая адаптировать па-

раметры Hard-Core Poisson Process к конкретным географическим регионам с различной плотностью урбанизации;

- предложен гибридный подход к генерации пространственно-временных рядов, сочетающий калиброванные методы стохастической геометрии для пространственного моделирования с эмпирическими временными паттернами нагрузки, что обеспечивает сохранение реалистичных корреляций;
- создан инструмент автоматической генерации синтетических данных сетевой нагрузки, способный учитывать специфику различных типов городской инфраструктуры без необходимости длительного сбора реальных данных эксплуатации.

Из этого следует, что использование пространственно-временных рядов, полученных с помощью калиброванного Hard-Core Point Process, позволит повысить точность прогнозирования размещения вычислительных сервисов и снизить среднюю задержку обработки запросов в системах граничных вычислений за счет предобучения на данных коррелирующих с реальным расположением вышек сотовой связи.

Литература

1. Al-Bahri M., Alkishri W., Ahmed F.Y.H., Alshar'e M., Maskari S.A. Enhancing IoT network security through digital object architecture-based approaches // *Qubahan Academic Journal*. 2024. V. 4. N 1. P. 224–239. <https://doi.org/10.48161/qaj.v4n1a413>
2. Li N., Hao W., Zhou F., Chu Z., Yang S., Muta O., Gacanin H. Min-max latency optimization for IRS-aided cell-free mobile edge computing systems // *IEEE Internet of Things Journal*. 2024. V. 11. N 5. P. 8757–8770. <https://doi.org/10.1109/jiot.2023.3322751>
3. Hua H., Li Y., Wang T., Dong N., Li W., Cao J. Edge computing with artificial intelligence: a machine learning perspective // *ACM Computing Surveys*. 2023. V. 55. N 9. P. 1–35. <https://doi.org/10.1145/3555802>
4. Pandey C., Tiwari V., Rathore R.S., Jhaveri R.H., Roy D.S., Selvarajan S. Resource-efficient synthetic data generation for performance evaluation in mobile edge computing over 5G networks // *IEEE Open Journal of the Communications Society*. 2023. V. 4. P. 1866–1878. <https://doi.org/10.1109/OJCOMS.2023.3306039>
5. Ismail A.A., Khalifa N.E., El-Khoribi R.A. A survey on resource scheduling approaches in multi-access edge computing environment: a deep reinforcement learning study // *Cluster Computing*. 2025. V. 28. N 3. P. 184. <https://doi.org/10.1007/s10586-024-04893-7>
6. Talpur A., Gurusamy M. DRLD-SP: A Deep-reinforcement-learning-based dynamic service placement in edge-enabled Internet of vehicles // *IEEE Internet of Things Journal*. 2022. V. 9. N 8. P. 6239–6251. <https://doi.org/10.1109/jiot.2021.3110913>
7. Kabeer M., Nordin R., Behjati M., Shaharuddin F.Y.B.M. An urban multi-operator QoE-aware dataset for cellular networks in dense environments // *arXiv*. 2025. arXiv:2506.22484. <https://doi.org/10.48550/arXiv.2506.22484>
8. Luo F., Zheng S., Ding W., Fuentes J., Li Y. An edge server placement method based on reinforcement learning // *Entropy*. 2022. V. 24. N 3. P. 317. <https://doi.org/10.3390/e24030317>
9. Pandey C., Tiwari V., Rodrigues J.J.P.C., Roy D.S. 5GT-GAN-NET: Internet traffic data forecasting with supervised loss based synthetic data over 5G // *IEEE Transactions on Mobile Computing*. 2024. V. 23. N 11. P. 10694–10705. <https://doi.org/10.1109/TMC.2024.3364655>
10. Xiang B., Elias J., Martignon F., Di Nitto E. A dataset for mobile edge computing network topologies // *Data in Brief*. 2021. V. 39. P. 107557. <https://doi.org/10.1016/j.dib.2021.107557>
11. Pasqualini L., Parton M. Pseudo random number generation: a reinforcement learning approach // *Procedia Computer Science*. 2020. V. 170. P. 1122–1127. <https://doi.org/10.1016/j.procs.2020.03.057>
12. Liu P., Lei J., Cao H., Garg S., Kaur K., Kaddoum G. A Stochastic geometry model and analysis scheme for SCMA aided mobile edge

References

1. Al-Bahri M., Alkishri W., Ahmed F.Y.H., Alshar'e M., Maskari S.A. Enhancing IoT network security through digital object architecture-based approaches. *Qubahan Academic Journal*, 2024, vol. 4, no. 1, pp. 224–239. <https://doi.org/10.48161/qaj.v4n1a413>
2. Li N., Hao W., Zhou F., Chu Z., Yang S., Muta O., Gacanin H. Min-max latency optimization for IRS-aided cell-free mobile edge computing systems. *IEEE Internet of Things Journal*, 2024, vol. 11, no. 5, pp. 8757–8770. <https://doi.org/10.1109/jiot.2023.3322751>
3. Hua H., Li Y., Wang T., Dong N., Li W., Cao J. Edge computing with artificial intelligence: a machine learning perspective. *ACM Computing Surveys*, 2023, vol. 55, no. 9, pp. 1–35. <https://doi.org/10.1145/3555802>
4. Pandey C., Tiwari V., Rathore R.S., Jhaveri R.H., Roy D.S., Selvarajan S. Resource-efficient synthetic data generation for performance evaluation in mobile edge computing over 5G networks. *IEEE Open Journal of the Communications Society*, 2023, vol. 4, pp. 1866–1878. <https://doi.org/10.1109/OJCOMS.2023.3306039>
5. Ismail A.A., Khalifa N.E., El-Khoribi R.A. A survey on resource scheduling approaches in multi-access edge computing environment: a deep reinforcement learning study. *Cluster Computing*, 2025, vol. 28, no. 3, pp. 184. <https://doi.org/10.1007/s10586-024-04893-7>
6. Talpur A., Gurusamy M. DRLD-SP: A Deep-reinforcement-learning-based dynamic service placement in edge-enabled Internet of vehicles. *IEEE Internet of Things Journal*, 2022, vol. 9, no. 8, pp. 6239–6251. <https://doi.org/10.1109/jiot.2021.3110913>
7. Kabeer M., Nordin R., Behjati M., Shaharuddin F.Y.B.M. An urban multi-operator QoE-aware dataset for cellular networks in dense environments. *arXiv*, 2025. arXiv:2506.22484. <https://doi.org/10.48550/arXiv.2506.22484>
8. Luo F., Zheng S., Ding W., Fuentes J., Li Y. An edge server placement method based on reinforcement learning. *Entropy*, 2022, vol. 24, no. 3, pp. 317. <https://doi.org/10.3390/e24030317>
9. Pandey C., Tiwari V., Rodrigues J.J.P.C., Roy D.S. 5GT-GAN-NET: Internet traffic data forecasting with supervised loss based synthetic data over 5G. *IEEE Transactions on Mobile Computing*, 2024, vol. 23, no. 11, pp. 10694–10705. <https://doi.org/10.1109/TMC.2024.3364655>
10. Xiang B., Elias J., Martignon F., Di Nitto E. A dataset for mobile edge computing network topologies. *Data in Brief*, 2021, vol. 39, pp. 107557. <https://doi.org/10.1016/j.dib.2021.107557>
11. Pasqualini L., Parton M. Pseudo random number generation: a reinforcement learning approach. *Procedia Computer Science*, 2020, vol. 170, pp. 1122–1127. <https://doi.org/10.1016/j.procs.2020.03.057>
12. Liu P., Lei J., Cao H., Garg S., Kaur K., Kaddoum G. A Stochastic geometry model and analysis scheme for SCMA aided mobile edge

- computing // Proc. of the IEEE Wireless Communications and Networking Conference (WCNC). 2024. P. 1–6. <https://doi.org/10.1109/WCNC57260.2024.10570726>
13. Gu Y., Yao Y., Li C., Xia B., Xu D., Zhang C. Modeling and analysis of stochastic mobile-edge computing wireless networks // IEEE Internet of Things Journal. 2021. V. 8. N 18. P. 14051–14065. <https://doi.org/10.1109/JIOT.2021.3068382>
 14. Facchini C., Holland O., Granelli F., da Fonseca N.L.S., Aghvami H. Dynamic green self-configuration of 3G base stations using fuzzy cognitive maps // Computer Networks. 2013. V. 57. N 7. P. 1597–1610. <https://doi.org/10.1016/j.comnet.2013.02.011>
 15. Bufetov A.I. Rigidity of determinantal point processes with the Airy, the Bessel and the Gamma kernel // Bulletin of Mathematical Sciences. 2016. V. 6. N 1. P. 163–172. <https://doi.org/10.1007/s13373-015-0080-z>
 16. Deng N., Zhou W., Haenggi M. The Ginibre point process as a model for wireless networks with repulsion // IEEE Transactions on Wireless Communications. 2015. V. 14. N 1. P. 107–121. <https://doi.org/10.1109/twc.2014.2332335>
 17. Teichmann J., Ballani F., van den Boogaart K.G. Generalizations of Matérn’s hard-core point processes // Spatial Statistics. 2013. V. 3. P. 33–53. <https://doi.org/10.1016/j.spasta.2013.02.001>
 18. Illian J., Penttinen A., Stoyan H., Stoyan D. Statistical Analysis and Modelling of Spatial Point Patterns. Wiley-Interscience, 2008. 560 p.
 19. Stoyan D., Kendall W.S., Mecke J. Stochastic Geometry and its Applications. John Wiley & Sons, 2013. 584 p.
 20. ElSawy H., Hossain E., Haenggi M. Stochastic geometry for modeling, analysis, and design of multi-tier and cognitive cellular wireless networks: a survey // IEEE Communications Surveys and Tutorials. 2013. V. 15. N 3. P. 996–1019. <https://doi.org/10.1109/surv.2013.052213.00000>
 21. Ester M., Krieger H.P., Sander J., Xu X. A density-based algorithm for discovering clusters // Proc. of the 2nd International Conference on Knowledge Discovery and Data Mining. 1996. P. 226–231.
 22. Богатырев В.А., Богатырев С.В., Богатырев А.В. Граничная оценка надежности кластерных систем на основе декомпозиции марковской модели при ограниченном восстановлении узлов с накоплением отказов // Научно-технический вестник информационных технологий, механики и оптики. 2025. Т. 25. № 3. С. 574–583. <https://doi.org/10.17586/2226-1494-2025-25-3-574-583>
 23. Богатырев В.А., Богатырев С.В., Богатырев А.В. Оценка готовности компьютерной системы к своевременному обслуживанию запросов при его совмещении с информационным восстановлением памяти после отказов // Научно-технический вестник информационных технологий, механики и оптики. 2023. Т. 23. № 3. С. 608–617. <https://doi.org/10.17586/2226-1494-2023-23-3-608-617>

Авторы

Филянин Иван Викторович — инженер, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 58665582700](https://orcid.org/0000-0002-0536-2807), <https://orcid.org/0000-0002-0536-2807>, adeptvin1@gmail.com

Капитонов Александр Александрович — кандидат технических наук, доцент, заместитель декана, Новый университет Узбекистана, Ташкент, 100000, Узбекистан; доцент, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 57202255729](https://orcid.org/0000-0001-5517-3038), <https://orcid.org/0000-0001-5517-3038>, kap2fox@gmail.com

Мартынюк Алексей Петрович — инженер, АО «ТБанк», Москва, 127287, Российская Федерация; студент, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, <https://orcid.org/0009-0009-8697-7739>, martynyukAlexey05@gmail.com

Authors

Ivan V. Filianin — Engineer, ITMO University, Saint Petersburg, 197101, Russian Federation, [sc 58665582700](https://orcid.org/0000-0002-0536-2807), <https://orcid.org/0000-0002-0536-2807>, adeptvin1@gmail.com

Aleksandr A. Kapitonov — Uzbekistan PhD, Associate Professor, Associate Dean, New University, Tashkent, 100000, Uzbekistan; Associate Professor, ITMO University, Saint Petersburg, 197101, Russian Federation, [sc 57202255729](https://orcid.org/0000-0001-5517-3038), <https://orcid.org/0000-0001-5517-3038>, kap2fox@gmail.com

Alexey P. Martynyuk — Engineer, “ТБанк”, JSC, Moscow, 127287, Russian Federation; Student, ITMO University, Saint Petersburg, 197101, Russian Federation, <https://orcid.org/0009-0009-8697-7739>, martynyukAlexey05@gmail.com

Статья поступила в редакцию 21.07.2025
Одобрена после рецензирования 13.01.2026
Принята к печати 15.03.2026

Received 21.07.2025
Approved after reviewing 13.01.2026
Accepted 15.03.2026



Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»